Finite Fields and Their Applications 68 (2020) 101733

Contents lists available at ScienceDirect

Finite Fields and Their Applications

www.elsevier.com/locate/ffa

Classification of quadratic APN functions with coefficients in $\mathbb{F}_2$ for dimensions up to 9

Yuyin Yu [a], Nikolay Kaleyski [b,*], Lilya Budaghyan [b], Yongqiang Li [c]

[a] *College of Mathematics and Information Science, Guangzhou University, Guangzhou, China*
[b] *Department of informatics, University of Bergen, Norway*
[c] *State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, China*

ARTICLE INFO

Article history:
Received 28 December 2019
Received in revised form 15 May 2020
Accepted 28 July 2020
Available online 28 August 2020
Communicated by Pascale Charpin

MSC:
94A60
06E30

Keywords:
Boolean functions
Almost perfect nonlinear
Almost bent
Quadratic functions

ABSTRACT

Almost perfect nonlinear (APN) and almost bent (AB) functions are integral components of modern block ciphers and play a fundamental role in symmetric cryptography. In this paper, we describe a procedure for searching for quadratic APN functions with coefficients in $\mathbb{F}_2$ over the finite field $\mathbb{F}_{2^n}$ and apply this procedure to classify all such functions over $\mathbb{F}_{2^n}$ with $n \leq 9$. We discover two new APN functions (which are also AB) over $\mathbb{F}_{2^9}$ that are CCZ-inequivalent to any known APN function over this field. We also verify that there are no quadratic APN functions with coefficients in $\mathbb{F}_2$ over $\mathbb{F}_{2^n}$ with $6 \leq n \leq 8$ other than the currently known ones.

* Corresponding author.
E-mail addresses: yuyuyin@163.com (Y. Yu), Nikolay.Kaleyski@uib.no (N. Kaleyski), Lilya.Budaghyan@uib.no (L. Budaghyan), yongq.lee@gmail.com (Y. Li).

https://doi.org/10.1016/j.ffa.2020.101733

1. Introduction

A vectorial Boolean (n, m)-function is a function between the vector spaces $\mathbb{F}_2^n$ and $\mathbb{F}_2^m$ over the finite field $\mathbb{F}_2 = \{0, 1\}$ for some two positive integers m, n. Vectorial Boolean functions play a crucial role in the design of modern block ciphers (where they are referred to as "S-boxes" or "substitution boxes"), in which they typically represent the only nonlinear part of the encryption. For this reason, the resistance of a block cipher to cryptanalytic attacks directly depends on the properties of its substitution boxes. Vectorial Boolean (n, n)-functions are of particular importance in cryptography since one typically wishes to substitute a sequence of bits for another sequence of the same length. In this case, the vector space $\mathbb{F}_2^n$ is usually identified with the finite field $\mathbb{F}_{2^n}$, and (n, n)-functions are expressed as polynomials over $\mathbb{F}_{2^n}$.

Among the most powerful cryptanalytic attacks known to date are the so-called "differential cryptanalysis" introduced by Biham and Shamir [1], and the "linear cryptanalysis" introduced by Matsui [25]. Almost perfect nonlinear (APN) functions were introduced by Nyberg [26] as the class of (n, n)-functions offering optimal resistance to differential cryptanalysis, while almost bent (AB) functions are the ones that are optimal against linear cryptanalysis [21]. Finding new examples and constructions of APN and AB functions is very important not only for the purpose of constructing new block ciphers in cryptography, but also for other areas of computer science and discrete mathematics (such as combinatorics, sequence design, coding theory, design theory) in which some APN functions correspond to optimal objects. Furthermore, finding new APN and AB functions is a difficult task, especially for large dimensions n: indeed, to date only six infinite monomial APN families and twelve infinite polynomial APN families have been discovered,[1] despite ongoing research on the topic since the early 90's. Among these, there are four infinite families of AB monomials and eight infinite families of AB polynomials.

The case of quadratic APN functions is more tractable than the general one, which is evinced by the fact that all the infinite polynomial families constructed so far are quadratic, and only one known sporadic example of a non-quadratic (up to CCZ-equivalence) APN function (which is defined over $\mathbb{F}_{2^6}$) is known [23]. Nevertheless, quadratic APN functions are an important ongoing direction of research: in 2010, Dillon et al. discovered an APN permutation in dimension $n = 6$, thereby disproving the conjecture that APN functions over fields of even dimension could never be bijective [5]. Despite Dillon's permutation not being a quadratic APN function per se, it was constructed by traversing the CCZ-equivalence class of a quadratic function. The question of the existence of other APN permutations for even n remains open, and investigating new instances of quadratic APN functions is a promising way to approach it.

A lot of research has been done on the topic of APN functions in recent years. An infinite construction of APN binomials inequivalent to power functions is given

[1] Tables of the known infinite monomial and polynomial families can be found at https://boolean.h.uib.no/mediawiki/.

in [13], disproving the long-standing conjecture that all infinite APN families must be monomials. Further infinite constructions of APN and AB functions are proposed in [2,8–15,29,32]. Previously, a classification of all APN functions over $\mathbb{F}_{2^n}$ for n up to 5 was given in [3], with classification for dimensions n higher than 5 remaining incomplete at the time of writing. In the case of $n = 6$, classification is complete for the particular cases of quadratic and cubic functions: in [4], 13 CCZ-inequivalent quadratic functions over $\mathbb{F}_{2^6}$ are listed, and it is shown that these encompass all quadratic CCZ-classes over $\mathbb{F}_{2^6}$ in [22]; as for the case of cubic APN functions, their classification is given in [24]. Furthermore, a study of the EA-equivalence classes corresponding to all known APN functions over $\mathbb{F}_{2^6}$ is presented in [16,17]. More background on APN functions and their construction can be found e.g. in [7] or [19].

Using a matrix construction, a large number of CCZ-inequivalent APN functions were found over $\mathbb{F}_{2^7}$ and $\mathbb{F}_{2^8}$ [31], bringing the total number of known APN functions over these fields to 490 and 8180, respectively. To the best of our knowledge, no systematic search of this kind has been performed over $\mathbb{F}_{2^n}$ for any dimension $n \geq 9$. The main reason for this is that the complexity of a computer search (which increases exponentially with the dimension n) becomes too demanding over dimensions of this magnitude.

Results similar to those in [31] have been independently obtained in [30], wherein 285 and 10 previously unknown quadratic APN functions are obtained over $\mathbb{F}_{2^7}$ and $\mathbb{F}_{2^8}$, respectively. Another similar approach based on the concept of antidifferentiation is developed in and [27] and [28].

In this paper, we focus on the particular case of quadratic APN functions over $\mathbb{F}_{2^n}$ with $n \leq 9$ and with coefficients in $\mathbb{F}_2$. We employ a specialization of the matrix method presented in [31] to conduct our search, and obtain a complete classification (up to CCZ-equivalence) of these functions over $\mathbb{F}_{2^9}$. In particular, we discover two instances of APN functions over $\mathbb{F}_{2^9}$ that are inequivalent to any known APN function over this field. For dimensions n with $6 \leq n \leq 8$, we show that there are no quadratic APN functions with coefficients in $\mathbb{F}_2$ other than the already known ones.

In our classification, we list a shortest possible representative from each discovered CCZ-equivalence class. In dimensions n up to 6, these shortest representatives are all monomials. In dimensions $n \in \{7, 8\}$, the longest representative has 6 terms, while in dimension $n = 9$, the longest representative has 9 terms. This raises the question of whether any quadratic APN function over $\mathbb{F}_{2^n}$ represented by a polynomial with coefficients in $\mathbb{F}_2$ is CCZ-equivalent to a function that can be represented by a polynomial with coefficients in $\mathbb{F}_2$ with at most n terms.

Furthermore, although all of the functions that we find over $\mathbb{F}_{2^8}$ are equivalent to representatives from [23], we find shorter representatives for two of these functions, viz. $x^3 + x^6 + x^{72}$ for $x^3 + \mathrm{Tr}(x^9)$ and $x^3 + x^6 + x^{144}$ for $x^9 + \mathrm{Tr}(x^3)$. Thus, to the best of our knowledge, our classification lists the shortest known representatives for these CCZ-equivalence classes.

2. Preliminaries

Let n be a positive integer. We denote by $\mathbb{F}_{2^n}$ the finite field with 2^n elements, by $\mathbb{F}_{2^n}^\star$ its multiplicative group, and by $\mathbb{F}_{2^n}[x]$ the univariate polynomial ring over $\mathbb{F}_{2^n}$ in indeterminate x. The trace function $\mathrm{Tr} : \mathbb{F}_{2^n} \to \mathbb{F}_2$ is defined by $\mathrm{Tr}(x) = \sum_{i=0}^{n-1} x^{2^i}$ for $x \in \mathbb{F}_{2^n}$. By $\mathbb{F}_{2^n}^{m\times k}$, we denote the set of m-by-k matrices with entries in $\mathbb{F}_{2^n}$, and if $M \in \mathbb{F}_{2^n}^{m\times k}$, we denote by $M[i,j]$ the entry in the i-th row and j-th column of M, for $0 \le i \le m-1$, $0 \le j \le k-1$. By $\mathrm{Submatrix}(M,i,j,p,q)$, we will denote the $p \times q$ submatrix of M rooted at (i,j), for $0 \le i \le m-1$, $0 \le j \le k-1$, $1 \le p \le m-i$, $1 \le q \le k-1$. Note that we index matrix rows and columns from zero.

We will use the following conventions and notation throughout the paper:

(i) When working over $\mathbb{F}_{2^n}$, integers indexing i.a. basis elements and matrix rows and columns will be considered modulo n. For instance, a normal basis $\{\alpha_0, \alpha_1, \ldots, \alpha_{n-1}\}$ satisfies $\alpha_{i+1} = \alpha_i^2$ for $0 \le i \le n-1$; this means that $\alpha_{i+1} = \alpha_i^2$ for $0 \le i \le n-2$, and $\alpha_0 = \alpha_{n-1}^2$.

(ii) Suppose $\{\alpha_0, \alpha_1, \ldots, \alpha_{n-1}\}$ is a normal basis of $\mathbb{F}_{2^n}$ over $\mathbb{F}_2$, so that $\alpha_{i+1} = \alpha_i^2$ for $0 \le i \le n-1$, and suppose $\{\theta_0, \theta_1, \ldots, \theta_{n-1}\}$ is its dual basis, i.e. $\mathrm{Tr}(\alpha_i\theta_j) = 0$ for $i \ne j$ and $\mathrm{Tr}(\alpha_i\theta_i) = 1$ for $0 \le i,j \le n-1$. Note that $\{\theta_0, \theta_1, \ldots, \theta_{n-1}\}$ is also a normal basis, so that without loss of generality, we can assume $\theta_{i+1} = \theta_i^2$ for $0 \le i \le n-1$.
Let $M_\alpha \in \mathbb{F}_{2^n}^{n\times n}$ and $M_\theta \in \mathbb{F}_{2^n}^{n\times n}$ be such that

$$M_\alpha[i,u] = \alpha_u^{2^i} \text{ and } M_\theta[i,u] = \theta_u^{2^i} \tag{1}$$

for $0 \le u,i \le n-1$. Then $M_\alpha^t M_\theta[u,j] = \mathrm{Tr}(\alpha_u\theta_j)$ for $0 \le u,j \le n-1$, so that $M_\alpha^t M_\theta = I_n$, where I_n is the identity matrix of order n. Thus $M_\theta^{-1} = M_\alpha^t$, where M_α^t is the transpose of M_α.

(iii) Let $B \in \mathbb{F}_{2^n}^m$ be a vector $B = (\eta_0, \eta_1, \cdots \eta_{m-1})$ where $\eta_i \in \mathbb{F}_{2^n}$ for $0 \le i \le m-1$. Then $\mathrm{Span}(B) = \mathrm{Span}(\eta_0, \eta_1, \cdots, \eta_{m-1})$ is the subspace spanned by $\{\eta_0, \eta_1, \cdots, \eta_{m-1}\}$ over $\mathbb{F}_2$. The dimension of this subspace is denoted by $\mathrm{Rank}(B) = \mathrm{Rank}(\eta_0, \eta_1, \cdots, \eta_{m-1})$, and is referred to as the rank of B over $\mathbb{F}_2$.
If $\eta_i = \sum_{j=0}^{n-1} \lambda_{i,j}\alpha_j$ for $0 \le j \le m-1$, with $\lambda_{i,j} \in \mathbb{F}_2$ for $0 \le i,j \le n-1$, and we define an m-by-n matrix $\Lambda \in \mathbb{F}_2^{m\times n}$ by $\Lambda[i,j] = \lambda_{i,j}$, then the rank of B is equal to the rank of Λ.

An (n,n)-function, or vectorial Boolean function, is any mapping $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ from the field with 2^n elements to itself. Any (n,n)-function can be represented as a polynomial $F(x) = \sum_{i=0}^{2^n-1} a_i x^i$ over $\mathbb{F}_{2^n}$ with $a_i \in \mathbb{F}_{2^n}$; this representation is referred to as the univariate representation of F, and is unique. The binary weight $wt_2(i)$ of a positive integer i is the number of ones in its binary notation; equivalently, if we write i as a sum of powers of two, so that $i = \sum_{j=0}^{k} b_j 2^j$ for $b_j \in \{0,1\}$, then its binary weight

is $wt_2(x) = \sum_{i=0}^{k} b_j$, with the sum taken over the integers. The largest binary weight of an exponent i with non-zero coefficient a_i in the univariate representation of an (n, n)-function F is called the algebraic degree of F and is denoted by $\deg(F)$. A function of algebraic degree 1, resp. 2, resp. 3 is called affine, resp. quadratic, resp. cubic. An affine F satisfying $F(0) = 0$ is called linear.

In the following, we concentrate on the case of homogeneous quadratic functions, which can be written as

$$F(x) = \sum_{0 \le i < j \le n-1} a_{i,j} x^{2^i + 2^j}$$

for $a_{i,j} \in \mathbb{F}_{2^n}$, i.e. quadratic functions with no linear terms in their univariate representation.

Definition 1. A mapping $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is called **differentially $\delta(F)$-uniform** if

$$\delta(F) = \max_{a \in \mathbb{F}_{2^n}^{\star}, b \in \mathbb{F}_{2^n}} \#\Delta_F(a, b),$$

where $\Delta_F(a, b) = \{x \in \mathbb{F}_{2^n} : F(x + a) + F(x) = b\}$, and $\#\Delta_F(a, b)$ is the cardinality of $\Delta_F(a, b)$. If $\delta(F) = 2$, F is called **almost perfect nonlinear (APN)**.

Definition 2. Let F and F' be two functions from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^n}$. We say that F and F' are **EA-equivalent** (Extended affine equivalent) if we can write F' as

$$F'(x) = A_1(F(A_2(x))) + A_3(x),$$

where A_1 and A_2 are affine permutations of $\mathbb{F}_{2^n}$, and A_3 is an affine function on $\mathbb{F}_{2^n}$.

We say that F and F' are **CCZ-equivalent** (Carlet-Charpin-Zinoviev equivalent) [20], if there exists an affine permutation which maps G_F onto $G_{F'}$, where $G_F = \{(x, F(x)) : x \in \mathbb{F}_{2^n}\}$ is the graph of F, and $G_{F'}$ is the graph of F'.

EA-equivalence is a special case of CCZ-equivalence, and the latter, which also includes taking inverses of permutations as a particular case, is known to be strictly more general that the combination of both of the aforementioned transformations [6,8,18]. An important property of CCZ-equivalence is that it leaves the differential uniformity $\delta(F)$ invariant, i.e. if two (n, n)-functions F and F' are CCZ-equivalent, then $\delta(F) = \delta(F')$. For this reason, APN functions are typically classified up to CCZ-equivalence, and this makes the classification process somewhat easier despite the large amount of (n, n)-functions.

Computationally testing whether two (n, n)-functions are CCZ-equivalent is typically done by associating a linear code to each function and then testing whether the resulting two codes are isomorphic [4]. To the best of our knowledge, this test is reliable for finite fields $\mathbb{F}_{2^n}$ with $n \le 9$, but sometimes fails for higher values of n due to a lack of computational resources. The Γ-rank, Δ-rank, and the order of the multiplier group

are CCZ-invariants introduced in [23]. In our search, we use the code isomorphism test to partition the APN functions that we find into CCZ-equivalence classes, and use the Γ-ranks of the two new functions that we find as proof that they lie outside the bounds of all previously known APN functions over $\mathbb{F}_{2^9}$.

We recall a couple of useful notions from [31].

Definition 3. Let $H \in \mathbb{F}_{2^n}^{m\times k}$ $(m, k \leq n)$. We say that H is **proper** if every nonzero linear combination over $\mathbb{F}_2$ of the m rows of H has rank at least $k-1$.

Definition 4. Let H be an $n\times n$ matrix defined on $\mathbb{F}_{2^n}$. Then H is called a **QAM** (quadratic APN matrix) if:
i) H is symmetric and the elements in its main diagonal are all zeros;
ii) H is proper, i.e. every nonzero linear combination of the n rows (or, equivalently, columns, due to H being symmetric) of H has rank $n-1$.

3. Construction of quadratic APN functions

3.1. Correspondence between quadratic functions with coefficients in $\mathbb{F}_2$ and a class of matrices

As shown in [31], there is a one-to-one correspondence between quadratic APN functions and QAM's. The precise statement is given in Theorem 1 below.

Theorem 1. *[31] Let $F(x) = \sum_{0\leq t<i\leq n-1} c_{i,t}x^{2^i+2^t} \in \mathbb{F}_{2^n}[x]$ be a homogeneous quadratic (n,n)-function and let $C_F \in \mathbb{F}_{2^n}^{n\times n}$ be defined by $C_F[i,t] = C_F[t,i] = c_{i,t}$, $C_F[i,i] = 0$ for $0 \leq i < t \leq n-1$. Let $H = M_\alpha^t H M_\alpha$ where M_α is as defined in (1). Then $\delta(F) = 2^k$ if and only if any non-zero linear combination over $\mathbb{F}_2$ of the n rows of H has rank at least $n-k$. In particular, F is APN if and only if H is a QAM.*

The following theorem addresses the specific case when all coefficients of the function are in $\mathbb{F}_2$.

Theorem 2. *Let $F(x) = \sum_{0\leq t<i\leq n-1} c_{i,t}x^{2^i+2^t}$ be a quadratic homogeneous (n,n)-function. Define an $n\times n$ matrix C_F by $C_F[t,i] = C_F[i,t] = c_{i,t}$ for $0 \leq t < i \leq n-1$ and $C_F[i,i] = 0$ for $0 \leq i \leq n-1$. Finally, take*

$$H = M_\alpha^t C_F M_\alpha.$$

Then

$$H[u+1, v+1] = H[u,v]^2 \tag{2}$$

(with the indices taken modulo n) for $0 \leq v, u \leq n-1$ *if and only if* $c_{i,t} \in \mathbb{F}_2$ *for* $0 \leq t < i \leq n-1$.

Proof. ($\Leftarrow$) Suppose $c_{i,t} \in \mathbb{F}_2$ for $0 \leq t < i \leq n-1$. From $H = M_\alpha^t C_F M_\alpha$ we have, for all $0 \leq v, u \leq n-1$,

$$H[u,v] = \sum_{0 \leq t < i \leq n-1} c_{it}(\alpha_u^{2^i}\alpha_v^{2^t} + \alpha_u^{2^t}\alpha_v^{2^i}).$$

It is easy to see that $H[u+1, v+1] = H[u,v]^2$ for $0 \leq v, u \leq n-1$, since $\{\alpha_0, \alpha_1, \ldots, \alpha_{n-1}\}$ is a normal basis such that $\alpha_{i+1} = \alpha_i^2$ for $0 \leq i \leq n-1$. Note that in the case $i = n-1$, this means that $\alpha_{n-1}^2 = \alpha_0$.

($\Rightarrow$) Suppose now that H satisfies (2). From $H = M_\alpha^t C_F M_\alpha$, we have $C_F = (M_\alpha^t)^{-1} H M_\alpha^{-1} = M_\theta H M_\theta^t$, which means that, for all $0 \leq v, u \leq n-1$,

$$c_{i,t} = C_F[i,t] = \sum_{0 \leq u,v \leq n-1} (\theta_u^{2^i}\theta_v^{2^t}) H[u,v].$$

Since $\theta_{i+1} = \theta_i^2$ for $0 \leq i \leq n-1$, if $H[u+1, v+1] = H[u,v]^2$ for $0 \leq v, u \leq n-1$, we have

$$c_{i,t} = \sum_{0 \leq k \leq n-1} Tr(\theta_0^{2^i}\theta_{0+k}^{2^t} H[0, 0+k]),$$

which clearly belongs to $\mathbb{F}_2$. □

By Theorem 2, any matrix H representing a quadratic APN function with coefficients in $\mathbb{F}_2$ satisfies (2); this significantly reduces the search space, and allows an exhaustive search to be performed in practice for higher dimensions.

3.2. Conditions on QAM's

In the following subsection, we describe how we conduct an exhaustive search over all $n \times n$ QAM's corresponding to (n,n)-functions represented by univariate polynomials with coefficients in $\mathbb{F}_2$. The condition $H[u+1, v+1] = H[u,v]^2$ greatly reduces the search space, and, in fact, implies that the values of only $\lfloor n/2 \rfloor$ entries of the matrix have to be guessed before the values of the remaining entries can be uniquely reconstructed. Depending on the parity of n, the situation is slightly different, and so, in the following we look at two concrete examples, one for $n = 5$, and one for $n = 6$.

Example 1. In the case of $n = 5$, suppose that H is a symmetric 5×5 matrix with zero diagonal and such that $H[u+1, v+1] = H[u,v]^2$ for all $0 \leq u, v \leq 4$. If we denote the entries of this matrix at $H[0,1]$ and $H[0,2]$ by a and b, respectively, we can readily see that H must take the form

$$H = \begin{pmatrix} 0 & a & b & b^8 & a^{16} \\ a & 0 & a^2 & b^2 & b^{16} \\ b & a^2 & 0 & a^4 & b^4 \\ b^8 & b^2 & a^4 & 0 & a^8 \\ a^{16} & b^{16} & b^4 & a^8 & 0 \end{pmatrix}.$$

Thus, knowing the values of only two entries of the matrix completely determines the rest. For comparison, without the condition $H[u+1, v+1] = H[u, v]^2$, we would have to guess $1 + 2 + 3 + 4 = 10$ entries of the matrix.

In the case of $n = 6$, we once again label the entries of a 6×6 matrix H at $H[0, 1]$, $H[0, 2]$, and $H[0, 3]$ by a, b, and c, respectively. The matrix then takes the form

$$H = \begin{pmatrix} 0 & a & b & c & b^{16} & a^{32} \\ a & 0 & a^2 & b^2 & c^2 & b^{32} \\ b & a^2 & 0 & a^4 & b^4 & c^4 \\ c^8 & b^2 & a^4 & 0 & a^8 & b^8 \\ b^{16} & c^{16} & b^4 & a^8 & 0 & a^{16} \\ a^{32} & b^{32} & c^{32} & b^8 & a^{16} & 0 \end{pmatrix}.$$

Since H must be symmetric, from $H[0, 3] = c$ and $H[3, 0] = c^8$ we get an additional condition on the value of c, namely $c^8 = c$, i.e. $c \in \mathbb{F}_{2^3}$. In this case, only 3 entries of H need to be guessed before the entire matrix can be reconstructed. For comparison, omitting the condition $H[u+1, v+1] = H[u, v]^2$ would require us to guess $1+2+3+4+5 = 15$ entries of the matrix.

The above principles can be generalized as follows.

Proposition 1. *Let n be a positive integer and H be a symmetric $n \times n$ matrix over $\mathbb{F}_{2^n}$ with zeros on its main diagonal such that $H[u+1, v+1] = H[u, v]^2$ for all $0 \le u, v \le n-1$, with the indices being taken modulo n. Then:*

1. *$H[i, j] = H[0, j-i]^{2^i}$ for any $0 \le i, j \le n-1$;*
2. *$H[0, j] = H[0, -j]^{2^j}$ for any $0 \le j \le n-1$;*
3. *if n is even, then $H[0, n/2] \in \mathbb{F}_{2^{n/2}}$.*

Consequently, the entries of H at $H[0, j]$ for $1 \le j \le \lfloor n/2 \rfloor$ uniquely determine the values of all entries of H.

Proof. The first point follows from (2) by induction on i. For the second point, we have

$$H[0, n-j+1] = H[n-j+1, 0] = H[0, j-n-1]^{2^{n-j+1}} = H[0, j-1]^{2^{n-j+1}}$$

using the symmetry of H and the first point. The third point then follows from the second one by taking $j = n/2$. □

In general (that is, without the condition from Theorem 2), a symmetric $n \times n$ matrix with zeros on the main diagonal is determined by $1 + 2 + \cdots + (n-1) = n(n-1)/2$ entries. By restricting ourselves to matrices satisfying (2), the number of entries drops to $\lfloor n/2 \rfloor$ as pointed out in Proposition 1, which decreases the number of guesses from quadratic to linear in the dimension n.

The following proposition allows us to further reduce the search complexity by discarding QAM's which a priori correspond to equivalent functions. Proposition 2 follows from Theorem 3 of [31], which asserts that if $H \in \mathbb{F}_{2^n}^{n \times n}$ is a symmetric matrix, and $H' \in \mathbb{F}_{2^n}^{n \times n}$ is defined by applying a linear permutation $L : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ to all elements of H, then the quadratic functions defined by H and H' are EA-equivalent. As the mapping $x \mapsto x^2$ is a linear permutation on account of $\gcd(2, 2^n - 1) = 1$, the proposition is an immediate consequence of this theorem. The restriction to linear permutations of the form $x \mapsto x^{2^k}$ comes from the fact that the property (2) remains invariant under such permutations.

Proposition 2. *Suppose $F_1 \in \mathbb{F}_{2^n}[x]$ is a homogeneous quadratic APN function with coefficients in $\mathbb{F}_2$, and H is its corresponding QAM. Let H' be the matrix defined by $H'[i,j] = H[i,j]^2$ for $0 \leq i, j < n$. Then H' is also a QAM, and its corresponding function $F_2 \in \mathbb{F}_2[x]$ is EA-equivalent to F_1.*

Following the statement of Proposition 2, recall that we will two elements $a, b \in \mathbb{F}_{2^n}$ *conjugates* if there is a non-negative integer k such that $a = b^{2^k}$. The relation "conjugate to" is an equivalence relation which induces a partition of $\mathbb{F}_{2^n}$ into conjugacy classes.

To summarize, by Proposition 2, only a single representative from each conjugacy class has to be considered for the first entry that we guess, which further reduces the number of possibilities that have to be considered. Furthermore, in the case of even n, the set of possible values for the last entry that we guess can be restricted to the subfield $\mathbb{F}_{2^{n/2}}$.

The results from Theorems 1, 2 and Proposition 2 are combined into an efficient procedure for searching for quadratic APN functions over $\mathbb{F}_{2^n}$ with coefficients in $\mathbb{F}_{2^n}$ in Algorithm 1.

3.3. The algorithm

The algorithm is essentially an exhaustive search which traverses all possible $n \times n$ QAM's by starting with the $n \times n$ zero matrix and iteratively assigning concrete values to its entries. Condition (2) greatly reduces the search space.

The *Search(j, H)* procedure implements the basic logic of the exhaustive search. An invocation of *Search(j, H)* attempts to assign a value to the entry of the matrix in the first row and j-th column, i.e. $H[0, j]$. In order to achieve this, it first invokes the *GetPossibleValues(j, H)* function which returns a list W of all possible values that $H[0, j]$

Input: An integer $n = 2m + 1$
Output: A list of APN functions over $\mathbb{F}_{2^n}$ represented by univariate polynomials with coefficients in $\mathbb{F}_2$

```
1  procedure Search(n);
2  H ← an n × n zero matrix;
3  Search(1, H);
4  end procedure;
5
6  procedure Search(j, H);
7  W ← GetPossibleValues(j, H);
8  for w ∈ W do
9  |  Assign(j, H, w);
10 |  if j = m then
11 |  |  if H is a QAM then
12 |  |  |  output the polynomial corresponding to H;
13 |  |  end
14 |  else
15 |  |  Search(j + 1, H);
16 |  end
17 end
18 end procedure;
19
20 procedure Assign(j, H, w);
21 H[0, j] ← w;
22 H[j, 0] ← w;
23 for t ∈ 1, · · · , n − 1 do
24 |  //Note that all indices are modulo n
25 |  H[t, j + t] ← H[t − 1, j + t − 1]^2;
26 |  H[j + t, t] ← H[t, j + t];
27 end
28 end procedure;
29
30 function GetPossibleValues(j, H);
31 if j = 1 then
32 |  return GetConjugacyClassRepresentatives(n);
33 else
34 |  S ← Span({H[0, i], H[0, n − i] : i ∈ 1, 2, · · · , j − 1});
35 |  if #S < 2^{2j−2} then
36 |  |  return ∅;
37 |  end
38 |  E ← F*_{2^n} \ S;
39 |  for e ∈ E do
40 |  |  H[0, j] ← e;
41 |  |  A ← Submatrix(H, 0, 0, j, j + 1);
42 |  |  if A is not proper then
43 |  |  |  E ← E \ {e};
44 |  |  end
45 |  end
46 |  return E;
47 end
48 end function;
```

Algorithm 1: A procedure for searching for QAM's corresponding to APN functions with coefficients in $\mathbb{F}_2$.

can take; using Proposition 2, a number of impossible values are filtered out by *GetPossibleValues*, which further reduces the complexity of the search. For all possible values $w \in W$, the *Search* procedure attempts to assign w to $H[0, j]$. This is performed by calling the *Assign*(j, H, w) procedure, which assigns w to $H[0, j]$ and derives the values

of all other entries of the matrix that follow from $H[0,j]$ by symmetry and by (2). If $j = \lfloor n/2 \rfloor$ and a value w is assigned to $H[0,j]$, then all entries in the matrix are already known, and it remains to check whether the obtained matrix is a QAM. If $j < m$, then *Search*$(j+1, H)$ is called recursively to assign a value to the next variable.

The *GetPossibleValues*(j, H) function distinguishes between two cases. Since a QAM must contain zeros on the main diagonal, $H[0,1]$ is the first variable to be assigned a value. By Proposition 2, it suffices to consider a single representative from every conjugacy class in $\mathbb{F}_{2^n}$; this is precisely what the function *GetConjugacyClassRepresentatives(n)* returns.

When $j > 2$, we can no longer restrict ourselves to a single representative from each conjugacy class, but can reduce the range of possible values for $H[0,j]$ in other ways. Recall that by the definition of a QAM, every nonzero linear combination of rows must have rank $n-1$. Since every row contains a zero element on the main diagonal, this is equivalent to saying that the elements of each row that do not lie on the main diagonal must be linearly independent. For this reason, the subspace S spanned by the entries in the first row that have already been assigned is removed from the list E of possible values. After S is computed, its size is used to test whether the known elements on the first row are linearly independent; note that while the element at $H[0,j]$ is always selected so that it is linearly independent on the previously assigned elements, the same is not necessarily true for the value of $H[0,-j]$ derived by Proposition 1, and this necessitates the test for linear independence. If the test fails, *GetPossibleValues* returns an empty set for the possible values of $H[0,j]$, which immediately forces the search procedure to backtrack to $H[0,j-1]$. By Corollary 2 of [31], every submatrix of a QAM must be proper. This condition is also exploited by *GetPossibleValues* in order to reduce the set E of possible values; once all entries $H[0,j]$ for $1 \leq j \leq j-1$ are known, the submatrix of H consisting of the first j rows and $j+1$ columns is fully determined. All values of $H[0,j]$ for which this submatrix is not proper are removed from E.

The entire search procedure begins by initializing H to an $n \times n$ zero matrix and invoking *Search*$(1, H)$ to assign a value to the first variable.

As observed in Subsection 3.2, the cases for an even and for an odd dimension n are slightly different. The only major difference is that the values of one of the entires of the matrix can be restricted to the subfield $\mathbb{F}_{2^{n/2}}$ when n is even. When implementing the search in practice, the distinction between the odd and even case manifests in the indexing of the variables. Algorithm 1 provides an explicit description of the search procedure in the case of odd n; this is motivated by the fact that our experiments for $n = 9$ constitute the main point of interest in our experimental output, as, to the best of our knowledge, no search of this type has been performed for dimensions greater than 8. The algorithm in the case of an even n is principally the same, keeping in mind that the value of $H[0,n/2]$ can be restricted to $\mathbb{F}_{2^{n/2}}$.

Table 1
List of representatives from all CCZ-equivalence classes of quadratic APN functions over $\mathbb{F}_{2^n}$ represented by polynomials with coefficients in $\mathbb{F}_2$, for $4 \leq n \leq 9$.

n	ID	Functions	Γ-rank	Δ-rank	$\lvert\mathcal{M}(G_F)\rvert$
4	4.1	x^3	100	20	5760
5	5.1	x^3	330	42	4960
	5.2	x^5	330	42	4960
6	6.1	x^3	1102	94	24192
7	7.1	x^3	3610	198	113792
	7.2	x^5	3708	198	113792
	7.3	x^9	3610	198	113792
	7.4	$x^3 + x^5 + x^6 + x^{12} + x^{33} + x^{34}$	4050	210	896
	7.5	$x^3 + x^5 + x^{10} + x^{33} + x^{34}$	4040	212	896
	7.6	$x^3 + x^6 + x^{20}$	4038	212	896
	7.7	$x^3 + x^6 + x^{34} + x^{40} + x^{72}$	4048	212	896
	7.8	$x^3 + x^9 + x^{10} + x^{66} + x^{80}$	4026	212	896
	7.9	$x^3 + x^9 + x^{18} + x^{66}$	4044	212	896
	7.10	$x^3 + x^{12} + x^{17} + x^{33}$	4048	210	896
	7.11	$x^3 + x^{12} + x^{40} + x^{72}$	4048	210	896
	7.12	$x^3 + x^{17} + x^{20} + x^{34} + x^{66}$	4040	210	896
	7.13	$x^3 + x^{17} + x^{33} + x^{34}$	4040	212	896
	7.14	$x^3 + x^{20} + x^{34} + x^{66}$	4048	210	896
	7.15	$x^5 + x^{18} + x^{34}$	4034	210	896
8	8.1	x^3	11818	420	522240
	8.2	x^9	12370	420	522240
	8.3	$x^3 + x^5 + x^{18} + x^{40} + x^{66}$	14044	446	2048
	8.4	$x^3 + x^6 + x^{72}$	13800	432	6144
	8.5	$x^3 + x^6 + x^{68} + x^{80} + x^{132} + x^{160}$	14040	454	2048
	8.6	$x^3 + x^6 + x^{144}$	13804	434	6144
	8.7	$x^3 + x^{12} + x^{40} + x^{66} + x^{130}$	14046	438	2048
9	9.1	x^3	38470	872	2354688
	9.2	x^5	41494	872	2354688
	9.3	x^{17}	38470	872	2354688
	9.4	$x^{136} + x^{132} + x^{96} + x^{80} + x^{36} + x^{34} + x^{18} + x^{17} + x^{12}$	48856	940	4608
	9.5	$x^{144} + x^{130} + x^{72} + x^{65} + x^{18} + x^9 + x^3$	48428	930	4608
	9.6	$x^{257} + x^{144} + x^{130} + x^{72} + x^{65} + x^{18} + x^9$	48460	944	4608
	9.7	$x^{264} + x^{160} + x^{144} + x^{132} + x^{80} + x^{72} + x^{66} + x^{40} + x^{17}$	47890	920	4608
	9.8	$x^{288} + x^{272} + x^{264} + x^{160} + x^{144} + x^{130} + x^{48} + x^{34}$	48858	940	4608

3.4. Summary of experimental results

Running the search for $n = 9$ on a server operating with an Intel Xeon E5 CPU at 3.5G GHz took approximately 33 days and produced a list of 21504 functions. Partitioning them into CCZ-equivalence classes by the code isomorphism test was performed by running several parallel processes on a server with an Intel Xeon E5 CPU at 2.60 GHz, and around 15-16 months. As a result, we obtain the 8 CCZ-inequivalent representatives given in Table 1. Computing the Γ-rank of one representative on the same server takes around an hour, while computing the Δ-rank takes approximately 3 days.

The running times for lower dimensions are negligible, and the computations were performed on a personal computer running an Intel m5-6Y54 CPU at 1.5 GHz. For $n = 8$, performing the exhaustive search took around 3 hours and produced 7616 functions,

which were partitioned into CCZ-classes in 8 hours. For $n = 7$, 4410 functions were found in 2 minutes, and partitioned into CCZ-classes within 12 hours. For $4 \leq n \leq 6$, both performing the search and partitioning the resulting functions into CCZ-equivalence classes takes less than a second; the number of functions found was 4 for $n = 4$, 72 for $n = 5$, and 32 for $n = 6$.

Table 1 lists representatives from all CCZ-equivalence classes found by our method. Note that the search is complete, i.e. the CCZ-equivalence classes containing these representatives cover all possible homogeneous quadratic APN functions with coefficients in $\mathbb{F}_2$ over $\mathbb{F}_{2^n}$ with $4 \leq n \leq 9$. For each representative, we have also computed its Γ-rank, Δ-rank, and the order $|\mathcal{M}(G_F)|$ of its multiplier group [23].

In dimensions $n \leq 6$, we only find power functions as expected. In dimension $n = 7$, besides three power functions, we find 12 polynomials, among which are two trinomials, five quadrinomials, four pentanomials, and one hexanomial. In dimension $n = 8$, we find two power functions and 5 polynomials, which consist of two trinomials, two pentanomials, and one hexanomial. In dimension $n = 9$, we find three power functions, along with 5 polynomials: two of them have 7 terms, one has 8 terms, and two have 9 terms. All the representatives given in the tables are in shortest possible presentation.

In the case of dimension $n \leq 8$, all of the representatives that we have discovered are identical or equivalent to switching class representatives from [23]. Despite this, in dimension $n = 8$, we discover very "short" and previously undocumented representatives (namely, trinomials) for two of the switching classes from [23]: $x^3 + x^6 + x^{72}$ is CCZ-equivalent to $x^3 + \mathrm{Tr}(x^9)$, and $x^3 + x^6 + x^{144}$ is CCZ-equivalent to $x^9 + \mathrm{Tr}(x^3)$. Both of these trinomials consist of monomials from the cyclotomic cosets of x^3 and x^9, and despite their nearly identical structure, they belong to distinct CCZ-equivalence classes. Note that the $x^3 + \mathrm{Tr}(x^9)$ belongs to the infinite family of APN functions from [14], while the second has not be generalized into any infinite family so far.

Furthermore, in dimension $n = 9$, we discover two representatives, viz.

$$s_1(x) = x^{136} + x^{132} + x^{96} + x^{80} + x^{36} + x^{34} + x^{18} + x^{17} + x^{12}$$

and

$$s_2(x) = x^{288} + x^{272} + x^{264} + x^{160} + x^{144} + x^{130} + x^{48} + x^{34}$$

which are CCZ-inequivalent to any currently known APN function over $\mathbb{F}_{2^9}$. We have verified this inequivalence in two ways: by means of the code isomorphism test, and, in addition, by computing their Γ-ranks, which turn out to be 48856 AND 48858, respectively.

We have computationally checked that these newly found functions are not CCZ-equivalent to a permutation, which took us about 40 hours computation. Thus, no quadratic APN function with coefficients in $\mathbb{F}_2$ can be CCZ-equivalent to a permutation over $\mathbb{F}_{2^n}$ with $n \leq 9$, except for the Gold APN monomials in the case of odd n.

Based on the computational results for dimensions $n \leq 9$, we can observe that any quadratic APN function F_1 with coefficients in $\mathbb{F}_2$ appears to be CCZ-equivalent to a quadratic APN function F_2 with at most n non-zero coefficients in $\mathbb{F}_{2^n}$. It would be interesting to establish whether this is true in general; if so, it would indicate the existence of a simple polynomial form for functions of this type, which would significantly simplify the complexity of searching for them.

This is closely related to the problem of finding the "simplest" possible polynomial representation for a given (n, n)-function F. A simple representation not only results in a polynomial representation that can be evaluated more efficiently in practice, but facilitates the mathematical analysis of the function in question and its properties.

Problem 1. Given an (n, n)-function F, find a function G, such that G is CCZ-equivalent to F and its univariate representation has the least possible number of non-zero coefficients.

4. Conclusion

We have described a procedure for searching for quadratic APN functions with coefficients in $\mathbb{F}_2$ over $\mathbb{F}_{2^n}$ by constructing matrices of a particular type, and have used this procedure to classify all such functions over the finite fields $\mathbb{F}_{2^n}$ with $n \leq 9$. We have discovered two previously unknown APN functions over $\mathbb{F}_{2^9}$, and a representation of two of the switching class representatives over $\mathbb{F}_{2^8}$ in the form of trinomials, which is simpler than their currently known representations. In the case of $6 \leq n \leq 8$, we have experimentally verified that there are no quadratic APN functions with coefficients in $\mathbb{F}_2$ other than the previously known ones.

CRediT authorship contribution statement

Yuyin Yu: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Validation, Writing - original draft. **Nikolay Kaleyski:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Writing - original draft, Writing - review & editing. **Lilya Budaghyan:** Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation. **Yongqiang Li:** Formal analysis, Funding acquisition, Investigation, Methodology, Resources.

Acknowledgments

The research of the second and the third authors is supported by the "Construction of Optimal Boolean Functions" project of the Trond Mohn Foundation. Yuyin Yu is supported by the NSF of China (Grant No. 61502113), and the Guangdong Provincial

NSF (Grant No. 2015A030310174). Yongqiang Li is supported by the NSF of China (Grant No. 61772517).

References

[1] E. Biham, A. Shamir, Differential cryptanalysis of DES-like cryptosystems, J. Cryptol. 4 (1) (1991) 3–72.
[2] C. Bracken, E. Byrne, N. Markin, G. McGuire, A few more quadratic APN functions, Cryptogr. Commun. 3 (3) (2011) 43–53.
[3] M. Brinkmann, G. Leander, On the classification of APN functions up to dimension five, Des. Codes Cryptogr. 49 (1–3) (2008) 273–288.
[4] K. Browning, J.F. Dillon, M. McQuistan, APN polynomials and related codes, in: Honoring the 75-th Birthday of Prof. D.K. Ray-Chaudhuri, J. Comb. Inf. Syst. Sci. 34 (1-4) (2009) 135–159, Special volume.
[5] K. Browning, J.F. Dillon, M. McQuistan, A.J. Wolfe, An APN permutation in dimension six, Contemp. Math. 58 (2010) 33–42.
[6] L. Budaghyan, The simplest method for constructing APN polynomials EA-inequivalent to power functions, in: Proceedings of the International Workshop on the Arithmetic of Finite Fields, WAIFI 2007, Madrid, Spain, in: Lecture Notes in Computer Science, vol. 4547, June 2007, pp. 177–188.
[7] L. Budaghyan, Construction and Analysis of Cryptographic Functions, Springer Verlag, 2014.
[8] L. Budaghyan, C. Carlet, A. Pott, New classes of almost bent and almost perfect nonlinear polynomials, IEEE Trans. Inf. Theory 52 (3) (2006) 1141–1152.
[9] L. Budaghyan, C. Carlet, Classes of quadratic APN trinomials and hexanomials and related structures, IEEE Trans. Inf. Theory 54 (5) (2008) 2354–2357.
[10] L. Budaghyan, C. Calderini, C. Carlet, R. Coulter, I. Villa, Constructing APN functions through isotopic shifts, https://eprint.iacr.org/2018/769.
[11] L. Budaghyan, C. Carlet, P. Felke, G. Leander, An infinite class of quadratic APN functions which are not equivalent to power mappings, in: IEEE International Symposium on Information Theory, 2006, pp. 2637–2641.
[12] L. Budaghyan, T. Helleseth, N. Kaleyski, A new family of APN quadrinomials, Cryptology ePrint Archive, Report 2019/994, 2019.
[13] L. Budaghyan, C. Carlet, G. Leander, Two classes of quadratic APN binomials inequivalent to power functions, IEEE Trans. Inf. Theory 54 (9) (2008) 4218–4229.
[14] L. Budaghyan, C. Carlet, G. Leander, Constructing new APN functions from known ones, Finite Fields Appl. 15 (2) (2009) 150–159.
[15] L. Budaghyan, C. Carlet, G. Leander, On a construction of quadratic APN functions, in: 2009 IEEE Information Theory Workshop, 2009.
[16] L. Budaghyan, M. Calderini, I. Villa, On equivalence between known families of quadratic APN functions, https://eprint.iacr.org/2019/793.
[17] M. Calderini, On the EA-classes of known APN functions in small dimensions, https://eprint.iacr.org/2019/369.
[18] A. Canteaut, L. Perrin, On CCZ-equivalence, extended-affine equivalence, and function twisting, Finite Fields Appl. 56 (2019) 209–246.
[19] C. Carlet, Vectorial Boolean functions for cryptography, in: Boolean Models and Methods in Mathematics, Computer Science, and Engineering, in: Encyclopedia of Mathematics and Its Applications, vol. 134, 2010, pp. 398–469.
[20] C. Carlet, P. Charpin, V. Zinoviev, Codes, bent functions and permutations suitable for DES-like cryptosystems, Des. Codes Cryptogr. 15 (2) (1998) 125–156.
[21] F. Chabaud, S. Vaudenay, Links between differential and linear cryptanalysis, in: Workshop on the Theory and Application of Cryptographic Techniques, Springer, Berlin, Heidelberg, 1994 May 9, pp. 356–365.
[22] Y. Edel, Quadratic APN functions as subspaces of alternating bilinear forms, in: Proceedings of the Contact Forum Coding Theory and Cryptography III, Belgium, vol. 2009, 2011, pp. 11–24.
[23] Yves Edel, Alexander Pott, A new almost perfect nonlinear function which is not quadratic, Adv. Math. Commun. 3 (1) (2009) 59–81.
[24] P. Langevin, Classification of APN cubics in dimension 6 over GF(2), http://langevin.univ-tln.fr/project/apn-6/apn-6.html.

[25] M. Matsui, Linear cryptanalysis method for DES cipher, in: Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings, Springer, Berlin, Heidelberg, 1993.
[26] K. Nyberg, Differentially uniform mappings for cryptography, in: Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings, Springer, Berlin, Heidelberg, 1993.
[27] A. Sălăgean, Discrete antiderivatives for functions over $\mathbb{F}_p^n$, Des. Codes Cryptogr. 88 (3) (2020) 471–486.
[28] V. Suder, Antiderivative functions over $\mathbb{F}_{2^n}$, Des. Codes Cryptogr. 82 (1–2) (2017) 435–447.
[29] H. Taniguchi, On some quadratic APN functions, Des. Codes Cryptogr. 87 (9) (2019) 1973–1983.
[30] G. Weng, T. Yin, G. Guang, On quadratic almost perfect nonlinear functions and their related algebraic object, in: Workshop on Coding and Cryptography, WCC, 2013.
[31] Y. Yu, M. Wang, Y. Li, A matrix approach for constructing quadratic APN functions, Des. Codes Cryptogr. 73 (2) (2014) 587–600.
[32] Y. Zhou, A. Pott, A new family of semifields with 2 parameters, Adv. Math. 234 (2013) 43–60.