On two fundamental problems on APN power functions

Lilya Budaghyan, Marco Calderini, Claude Carlet,
Diana Davidova and Nikolay Kaleyski
Department of Informatics, University of Bergen Bergen, Norway
{lilya.budaghyan,marco.calderini,diana.davidova,nikolay.kaleyski}@uib.no,
claude.carlet@gmail.com

Abstract

The six infinite families of power APN functions are among the oldest known instances of APN functions, and it has been conjectured in 2000 that they exhaust all possible power APN functions. Another long-standing open problem is that of the Walsh spectrum of the Dobbertin power family, which is the only one among the six families for which it remains unknown. In this paper, we derive alternative representations for functions from the infinite APN monomial families, with the hope that this will pave the way for further progress in this area. More concretely, we show how the Niho, Welch, and Dobbertin functions can be represented as the composition $x^i \circ x^{1/j}$ of two power functions, and prove that our representations are the simplest possible in the sense that no two power functions of lesser algebraic degree can produce the same composition. We also investigate compositions of the form $x^i \circ L \circ x^{1/j}$ for a linear polynomial L, and computationally determine all APN functions of this form for $n \leq 9$ and for L with coefficients in $\mathbb{F}_2$ in order to confirm that our theoretical constructions exhaust all possible cases. We present some observations and computational data on power functions with exponent of the form $\sum_{i=1}^{k-1} 2^{2ni} - 1$, which can be seen as generalizations of both the inverse and the Dobbertin APN families. Finally, we present our computational data on the Walsh coefficients of the Dobbertin function over $\mathbb{F}_{2^n}$ for $n \leq 35$, and conjecture the exact form of its Walsh spectrum.

I. Introduction

Let n and m be positive integers, and let $\mathbb{F}_{2^n}$ denote the finite field with 2^n elements. The multiplicative group of $\mathbb{F}_{2^n}$ will be denoted by $\mathbb{F}_{2^n}^*$, and Tr_n will denote the *absolute trace function* from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ given by

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

if the dimension n is clear from context, we will simply write Tr instead of Tr_n.

An (n, m)-function, or vectorial Boolean function, is any mapping F from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$. When $n = m$, any (n, n)-function can be uniquely represented as a univariate polynomial of the form $F(x) = \sum_{i=0}^{2^n-1} a_i x^i$, for $a_i \in \mathbb{F}_{2^n}$. We say that an (n, n)-function F is a *power, or monomial, function*, if its univariate representation is of the form $F(x) = x^d$ for some positive integer d. An n-dimensional *Boolean function* is simply an $(n, 1)$-function for some n.

Given a positive integer i, its *binary weight* (also called 2-weight) is the number of ones in its binary notation. More precisely, if $i = \sum_{j=0}^{K} c_i 2^i$ for some positive integer K and for $c_i \in \{0, 1\}$ for $0 \leq j \leq K$, then the binary weight of i is $\mathrm{w}(i) = \sum_{j=0}^{K} c_i$. The largest binary weight of any exponent i in the univariate representation of an (n, n)-function F with $a_i \neq 0$ is the *algebraic degree* of F. A function of algebraic degree 1, resp, 2, resp. 3 is said to be *affine*, resp. *quadratic*, resp. *cubic*. A *linear* function is an affine function F with $F(0) = 0$.

Vectorial Boolean functions are widely applied to the design of block ciphers in cryptography, where they are used in the design of so-called substitution boxes, or S-boxes, whose input and output are then both sequences of bits. This is possible since $\mathbb{F}_{2^n}$ can be seen as an n-dimensional vector space over the prime field $\mathbb{F}_2$, thanks to which $\mathbb{F}_{2^n}$ can be identified with $\mathbb{F}_2^n$. This then implies that any element of $\mathbb{F}_{2^n}$ can be interpreted as an n-dimensional binary vector, i.e. a vector consisting of zeros and ones. A prominent example is the AES, or Rijndael, block cipher, which contains an $(8, 8)$-function at its core [14].

It is clearly important to analyze the resistance of any given vectorial Boolean function against various kinds of cryptanalytic attacks when it is used as an S-box. One of the most powerful attacks against block ciphers is differential cryptanalysis [2], which exploits statistical dependencies between the difference $a = x - y$ (or, equivalently, $a = x + y$ since addition and subtraction coincide in characteristic 2) of two inputs and the difference $b = F(x) - F(y)$ (or $b = F(x) + F(y)$) of their corresponding outputs under $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$; if, for some input difference $a \in \mathbb{F}_{2^n}$, the probability of obtaining some output difference $b \in \mathbb{F}_{2^n}$ is greater than uniform, this correlation can be used to mount an attack on the corresponding block cipher. Furthermore, the efficiency of the attack is directly related to the largest probability among all pairs $(a, b) \in \mathbb{F}_{2^n}^2$ of input and output differences. The relationship between the difference of two inputs and their corresponding outputs under an (n, n)-function F is expressed by the so-called *derivative* $D_a F$ of F in the direction $a \in \mathbb{F}_{2^n}$, which is defined as the (n, n)-function $D_a F$ given by $D_a F(x) = F(a + x) + F(a)$.

The notion of the differential uniformity of a function is introduced in [26] as a measurement of contribution of the function to the resistance of the block cipher against differential cryptanalysis. More precisely, the *differential uniformity* Δ_F of an

(n,n)-function F is defined as the largest number of solutions $x \in \mathbb{F}_{2^n}$ to any equation of the form $D_aF(x) = b$, i.e. $F(x) + F(a+x) = b$ for $a, b \in \mathbb{F}_{2^n}$ with $a \neq 0$, i.e.

$$\Delta_F = \max_{a \in \mathbb{F}_{2^n}^*, b \in \mathbb{F}_{2^n}} \#\{x \in \mathbb{F}_{2^n} : F(a+x) + F(x) = b\}.$$

Since $a + x$ is a solution to $F(x) + F(a + x) = b$ whenever x is, Δ_F must be even for any F, and hence can be no lower than 2. The (n, n)-functions attaining this lower bound with equality are called *almost perfect nonlinear (APN)* and provide the best possible resistance to differential cryptanalysis.

Another powerful attack employed against block ciphers is linear cryptanalysis [25], which is successful if a nonzero linear combination of the bits of the output of a function F (viewed as valued in $\mathbb{F}_2^n$ thanks to the identification of $\mathbb{F}_{2^n}$ with $\mathbb{F}_2^n$ described above) can be approximated by a linear combination of the bits of its input. The nonlinearity of F is introduced as a measurement of the contribution of the function to the resistance against this kind of cryptanalysis. In contrast to the case of differential uniformity, higher values of the nonlinearity correspond to stronger resistance to linear attacks. It is also worth remarking that taking a function at random usually has very bad, i.e. high, differential uniformity, while its nonlinearity is typically close to acceptable. To make this notion more precise, we first introduce the concept of component functions. If F is an (n, m)-function for some n and m, then a *component function* of F is any $(n, 1)$-function of the form $F_c(x) = \mathrm{Tr}(cF(x))$ for $c \in \mathbb{F}_{2^m}^*$. The *nonlinearity* $\mathcal{NL}(F)$ of an (n, m)-function F is defined as the minimum Hamming distance between any component of F and any affine $(m, 1)$-function; symbolically,

$$\mathcal{NL}(F) = \min_{\substack{c \in \mathbb{F}_{2^m}, l:\mathbb{F}_{2^n} \to \mathbb{F}_2 \\ \deg(l)=1}} d_H(F_c, l),$$

where the *Hamming distance* between two (n, m)-functions F and G is defined as $d_H(F, G) = \#\{x \in \mathbb{F}_{2^n} : F(x) \neq G(x)\}$, i.e. as the number of inputs on which the values of F and G disagree.

A useful tool that is frequently used in the analysis of vectorial Boolean function is the so-called *Walsh transform* $W_F : \mathbb{F}_{2^n} \times \mathbb{F}_{2^m} \to \mathbb{Z}$, which, for an (n, m)-function F, is defined as

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

with the sum computed over $\mathbb{Z}$. The function F can be uniquely reconstructed from W_F, and in this sense the Walsh transform constitutes another possible representation of (n, m)-functions. The multiset $\{W_F(a, b) : a \in \mathbb{F}_{2^n}, b \in \mathbb{F}_{2^m}\}$ of all values of W_F is called the *Walsh spectrum* of F; the multiset $\{|W_F(a, b)| : a \in \mathbb{F}_{2^n}, b \in \mathbb{F}_{2^m}\}$ of the absolute values of W_F is referred to as the *extended Walsh spectrum*, and is noteworthy because (unlike the Walsh spectrum) it remains invariant under certain equivalence relations to be discussed below.

The nonlinearity of an (n, m)-function F can be expressed as

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

Furthermore, the nonlinearity of any (n, n)-function is bounded from above by $2^{n-1} - 2^{(n-1)/2}$ [13]. This bound is tight, and functions attaining it with equality are called *almost bent (AB)*; consequently, AB functions provide the best possible resistance to linear cryptanalysis. Clearly, AB functions can exist only for odd values of n; in the case of even n, functions with nonlinearity $2^{n-1} - 2^{n/2}$ are known, and it is conjectured that this is the highest possible value of the nonlinearity in the even case. Any AB function is necessarily APN [13], and thus AB functions provide the best possible resistance to both differential and linear cryptanalysis. We note that AB functions do have certain drawbacks, however: as shown in [11], their composition with vectorial functions does not provide a function with a large enough algebraic degree, and this is problematic since the algebraic degree of the vectorial function equal to the output of the r-th round of a block cipher should reach the optimum for r as small as possible.

APN and AB functions are typically classified with respect to CCZ-equivalence, which is currently the most general known equivalence relation that preserves the differential uniformity and nonlinearity [12]. Two (n, n)-functions F and G are said to be *Carlet-Charpin-Zinoviev-equivalent, or CCZ-equivalent*, if their graphs $\Gamma_F = \{(x, F(x)) : x \in \mathbb{F}_{2^n}\}$ and $\Gamma_G = \{(x, G(x)) : x \in \mathbb{F}_{2^n}\}$ are affine equivalent, i.e. if there is an affine permutation $\mathcal{A} : \mathbb{F}_{2^n}^2 \to \mathbb{F}_{2^n}^2$ such that $\mathcal{A}(\Gamma_F) = \Gamma_G$. For instance, a permutation and its inverse are always CCZ-equivalent. Another equivalence relation preserving differential uniformity is the so-called extended affine equivalence, or EA-equivalence. Two functions F and G are said to be *EA-equivalent* if there exist affine permutations A_1, A_2 of $\mathbb{F}_{2^n}$ and an affine function $A : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ such that $A_1 \circ F \circ A_2 + A = G$. EA-equivalence is a particular case of CCZ-equivalence, with the latter being strictly more general than EA-equivalence and taking inverses of permutations [8].

In the case of power functions, CCZ-equivalence (as well as EA-equivalence) coincides with cyclotomic equivalence. Two power functions $F(x) = x^d$ and $G(x) = x^e$ over $\mathbb{F}_{2^n}$, where d, e, n are positive integers, are said to be *cyclotomic equivalent* if $d \equiv 2^k e \mod (2^n - 1)$ for some positive integer k, or if $d^{-1} \equiv 2^k e \mod (2^n - 1)$ for some positive integer k in the

TABLE I
KNOWN INFINITE FAMILIES OF APN POWER FUNCTIONS OVER $\mathbb{F}_{2^n}$

Family	Exponent	Conditions	Algebraic degree	Source
Gold	2^i+1	$\gcd(i,n)=1$	2	[20], [26]
Kasami	$2^{2i}-2^i+1$	$\gcd(i,n)=1$	$i+1$	[23], [27]
Welch	2^t+3	$n=2t+1$	3	[15]
Niho	$2^t+2^{t/2}-1$, t even $2^t+2^{(3t+1)/2}-1$, t odd	$n=2t+1$	$(t+2)/2$ $t+1$	[16]
Inverse	$2^{2t}-1$	$n=2t+1$	$n-1$	[1], [26]
Dobbertin	$2^{4i}+2^{3i}+2^{2i}+2^i-1$	$n=5i$	$i+3$	[17]

case that $\gcd(d,2^n-1)=1$, with d^{-1} being the multiplicative inverse of d modulo 2^n-1. Cyclotomic equivalence has the advantage of being significantly simpler to test than both EA- and CCZ-equivalence.

APN functions have been studied since the 90's, and only around 16 infinite families of such functions are known to date. In particular, this illustrates that it is quite challenging to construct such functions. Among the known APN functions, the power APN functions play a particularly prominent role. For one, they have contributed the earliest known examples of APN functions and of infinite families of APN functions. For another, all known APN functions (including both instances of infinite families and unclassified sporadic examples) are CCZ-equivalent to power functions or quadratic functions (that is, functions of algebraic degree 2, which is considered too small for cryptographic purposes), with only one known exception in $\mathbb{F}_{2^6}$ [19].

The six known infinite families of APN monomials are given in Table I. It is conjectured by Dobbertin that this classification is complete up to CCZ-equivalence [17], i.e. any APN power function is CCZ-equivalent to an instance from one of the families in Table I. The conjecture is verified computationally for $n \leq 24$ by Anne Canteaut according to [17] and later by Yves Edel for $n \leq 34$ and $n = 36, 38, 40, 42$ (unpublished).

To date, the Walsh spectrum and even the nonlinearity of the Dobbertin family of power functions remain unknown. This is remarkable, as the exact Walsh spectra of the remaining five power families in Table I have been determined. The problem of determining the Walsh spectrum of the Dobbertin family has already been open for 20 years, and without any progress since the seminal work of Canteaut, Charpin and Dobbertin from 2000, in which they proved that all Walsh coefficients of the Dobbertin function over $\mathbb{F}_{2^{5m}}$ are divisible by 2^{2m} [10]. As hinted above, the Walsh spectrum of a function contains a lot of information about its properties, and so its computation is an important result per se. Moreover, there is a correspondence between the Walsh coefficients of a power function and the weight distribution of an associated linear code, as shown in e.g. [12]; thus, knowing the Walsh spectrum of the APN power functions has fundamental significance for the theory of linear codes. Furthermore, it is known that the extended Walsh spectrum is invariant under CCZ-equivalence, and knowing it can potentially allow to justify the inequivalence of functions belonging to distinct CCZ-equivalence classes.

Motivated by the above, in this paper we investigate alternative representations of the infinite power families from Table I, and for the Dobbertin functions in particular, in the hope that they can facilitate the computation of its Walsh spectrum. For n odd, the Kasami exponent $2^{2i}-2^i+1$, where $\gcd(i,n)=1$, which is itself of algebraic degree $i+1$, can be represented as the fraction $(2^{3i}+1)/(2^i+1)$ of two quadratic exponents, i.e. the corresponding Kasami function equals the compositions of a quadratic power function $x^{2^{3i}+1}$ and the inverse $x^{1/(2^i+1)}$ of another quadratic power function [18]. As shown in [18], this representation leads to a simpler proof of the AB-ness of the Kasami function for odd dimensions, and facilitates the derivation of its Walsh spectrum. We thus look for similar representations of the known APN power functions (and the Dobbertin function in particular). In Section II, we show how the Welch, Niho and Kasami functions can be expressed as fractions of low-degree exponents, and argue that these representations are optimal in the sense that fractions of exponents of lower algebraic degree cannot possibly provide a representation of these functions. In Section III, we examine a construction in which linear functions are composed on the left and on the right with power functions, allowing us to obtain one class of APN power functions from another. In Section IV, we present some observations and computational data on the differential spectrum of power functions x^d with exponent of the form $d=\sum_{i=1}^{k-1} 2^{ni}-1$ over $\mathbb{F}_{2^{nk}}$; since the exponents of both the inverse and the Dobbertin family are special cases of this form, investigating these exponents is a potential direction for approaching the problem of the existence of APN power functions inequivalent to the ones in Table I. Finally, in Section V, we formulate a conjecture based on our experimental data which completely describes the Walsh spectrum of the Dobbertin function. We hope that the simplified representation for the Dobbertin function developed in Section II will allow this conjecture to be approached more easily.

Some of the contents from Sections II, III and V has previously been presented at *Sequences and their Application (SETA)* 2020 [6], [7]. All the results presented in Section II with the exception of those on the Dobbertin power functions are completely new, as well as all the results of Section IV.

II. SIMPLIFYING THE EXPONENTS OF THE INFINITE APN MONOMIAL FAMILIES

It is known that the exponent of the Kasami function in the case of odd n can be represented as $2^{2i}-2^i+1=\frac{2^{3i}+1}{2^i+1}$, that is, the function can be expressed as the composition of a quadratic power function with the inverse of another quadratic

power function. As shown in [12], this property gives a simple explanation of the AB-ness of the Kasami function for n odd. In this section we study whether a similar property can be derived for other APN power functions, and in particular, for the Dobbertin function.

We show that the Niho APN power functions (in the case of both even and odd dimension) can be represented as the composition of x^3 and the inverse of a cubic function; and the Dobbertin function is cyclotomic-equivalent to a composition of a cubic function and the inverse of a quadratic function. Moreover, we prove that the derived representations are optimal, in the sense that the exponents of the Niho and Dobbertin functions cannot be represented as a fraction of two quadratic exponents. In this sense, the Welch exponent is optimal as well.

Consider first the Welch function: x^{2^t+3}, $n = 2t+1$. Clearly, the binary weight of its exponent is 3 for $t > 1$. Lemma 1 shows that the exponent 2^t+3 cannot be represented as a fraction of two numbers of binary weight 2, in general. Thus, the canonical representation of the Welch function is optimal in this sense.

Proposition 1. *Let t be a positive integer strictly greater than 3. Then, for any positive integers j, l, r such that $1 \leq j, l, r < 2t+1$, we have*

$$\left(2^t + 2^1 + 1\right)(2^j + 1) \not\equiv 2^l + 2^r \mod (2^{2t+1} - 1). \tag{1}$$

Proof. We shall show that for any $1 \leq j < 2t+1$ and $t > 3$, the binary weight of the left-hand side of (1) is always strictly greater than 2. Consider the following possible cases:

1) $1 \leq j \leq t$:
   a) $j \notin \{1, t-1, t\}$,
   b) $j \in \{1, t-1, t\}$;
2) $j = t + j', 1 \leq j' < t$:
   a) $j \notin \{1, 2\}$,
   b) $j \in \{1, 2\}$;
3) $j = 2t$.

In all the cases the binary weight of $(2^t + 2^1 + 1)(2^j + 1)$ is equal to an integer between 4 and 6. Indeed, for $1 \leq j \leq t$, we get

$$(2^t + 2^1 + 1)(2^j + 1) = 2^{t+j} + 2^t + 2^{j+1} + 2^j + 2^1 + 1.$$

(a) If $j \notin \{1, t-1, t\}$, then, obviously, $wt\Big((2^t + 2^1 + 1)(2^j + 1)\Big) = 6$.
(b) If $j \in \{1, t-1, t\}$, then, for instance, when $j = t$, we obtain

$$(2^t + 2^1 + 1)(2^t + 1) = 2^{2t} + 2^{t+2} + 2 + 1,$$

therefore

$$wt\Big((2^t + 2^1 + 1)(2^t + 1)\Big) = 4.$$

In the same way, we show that for $j = 1$, $wt\Big((2^t + 2^1 + 1)(2^j + 1)\Big)$ is 4 and for $j = t-1$ that is 5.

Similarly, for $j = t + j', 1 \leq j' \leq t-1$, we can write

$$(2^t + 2^1 + 1)(2^{t+j'} + 1) = 2^{2t+j'} + 2^{t+j'+1} + 2^{t+j'} + 2^t + 2^1 + 1 \equiv$$
$$2^{t+j'+1} + 2^{t+j'} + 2^t + 2^{j'-1} + 2 + 1 \mod (2^{2t+1} - 1).$$

(a) If $j' \notin \{1, 2\}$, then $wt\Big((2^t + 2^1 + 1)(2^{t+j'} + 1)\Big) = 6$.
(b) If $j' \in \{1, 2\}$, then $wt\Big((2^t + 2^1 + 1)(2^{t+j'} + 1)\Big)$ is either 4 or 5. For instance, taking j' equals to 2, we get

$$(2^t + 2^1 + 1)(2^{t+2} + 1) \equiv 2^{t+3} + 2^{t+2} + 2^t + 2^1 + 2 + 1 \equiv$$
$$2^{t+3} + 2^{t+2} + 2^t + 2^2 + 2 \mod (2^{2t+1} - 1).$$

Since $t > 3$, we then have

$$wt\Big((2^t + 2^1 + 1)(2^{t+2} + 1)\Big) = 5.$$

In the exact same way, we obtain that for $j' = 1$, the binary weight of the left-hand side of (1) is 4 in the case $j = 2t$ as well. □

Remark 2. It is easy to verify that for $t = 2$, the binary weight of $(2^t + 2^1 + 1)(2^j + 1)$ is either 1 or 3; and for $t = 3$, $wt\Big((2^t + 2^1 + 1)(2^j + 1)\Big)$ is either 2, 4 or 5 (it equals to 2 only for $j = 2t = 6$).

We now consider the Niho functions. Bellow we prove that the Niho power APN functions in the even case can be represented as the composition of x^3 and the inverse of a cubic power function.

Lemma 3. *For any positive even integer t the following congruences hold:*

$$\begin{aligned}(2^t+2^{\frac{t}{2}}-1)\equiv 2^{\frac{t}{2}-1}\frac{3}{2^{\frac{3t}{2}}+2^t+1} \quad &\equiv \\ 2^t\frac{3}{2^{\frac{3t}{2}+1}+2^{\frac{t}{2}+1}+1} \quad &\equiv \\ 2^{\frac{3t}{2}}\frac{3}{2^{t+1}+2^{\frac{t}{2}}+1} \quad &\mod (2^{2t+1}-1).\end{aligned}$$

Proof. We first prove that $2^{\frac{3t}{2}}+2^t+1$, $2^{\frac{3t}{2}+1}+2^{\frac{t}{2}+1}+1$ and $2^{t+1}+2^{\frac{t}{2}}+1$ are invertible modulo $2^{2t+1}-1$, i.e. that

$$\begin{aligned}\gcd\left(2^{\frac{3t}{2}}+2^t+1,2^{2t+1}-1\right)=\gcd\left(2^{\frac{3t}{2}+1}+2^{\frac{t}{2}+1}+1,2^{2t+1}-1\right)= \\ \gcd\left(2^{t+1}+2^{\frac{t}{2}}+1,2^{2t+1}-1\right)=1.\end{aligned}$$

For simplicity, we denote $2^{\frac{t}{2}}$ by y. By the Euclidean algorithm, we easily get:

$$\gcd\left(2y^4-1,y^3+y^2+1\right)=\gcd\left(y^3+y^2+1,2y^2-2y+1\right)=\gcd\left(2y^2-2y+1,y\right)=1,$$

$$\gcd\left(2y^4-1,2y^3+2y+1\right)=\gcd\left(2y^3+2y+1,2y^2+y+1\right)=\gcd\left(2y^2+y+1,y+1\right)=$$

$$\gcd\left(y+1,2\right)=1,$$

and

$$\gcd\left(2y^4-1,2y^2+y+1\right)=\gcd\left(2y^2+y+1,y-1\right)=\gcd(y-1,4)=1.$$

To prove the first congruence it remains to check that $(y^2+y-1)(y^3+y^2+1)\equiv\frac{3}{2}y \mod (2y^4-1)$ which is straightforward. Indeed, computing the left-hand side of this congruence, we get:

$$(y^2+y-1)(y^3+y^2+1)=y^5+2y^4+y-1\equiv\frac{3}{2}y \mod (2y^4-1).$$

This proves the first statement of the lemma. The other two congruences are proven in the same way. □

The following lemma justifies the optimality of this representation.

Lemma 4. *Let t be a positive even integer. Then, for any positive integers j,l,r such that $1\le j,l,r<2t+1$, the following incongruence holds:*

$$(2^t+2^{\frac{t}{2}}-1)(2^j+1)\not\equiv 2^l+2^r \mod (2^{2t+1}-1). \tag{2}$$

Proof. Following the same approach as in the proof of Lemma 1, we will show that for any $1\le j<2t+1$, the binary weight of the left-hand side of (2) is always strictly greater than 2.

We consider the following cases:

1) $1\le j\le\frac{t}{2}$:
   a) $j\notin\{\frac{t}{2},\frac{t}{2}-1\}$,
   b) $j\in\{\frac{t}{2},\frac{t}{2}-1\}$;
2) $j=\frac{t}{2}+j',1\le j'\le\frac{t}{2}$:
   a) $j'\notin\{1,\frac{t}{2}-1,\frac{t}{2}\}$,
   b) $j'\notin\{1,\frac{t}{2}-1,\frac{t}{2}\}$;
3) $j=t+j',1\le j'\le\frac{t}{2}$:
   a) $j'\notin\{1,2,\frac{t}{2}\}$,
   b) $j'\in\{1,2,\frac{t}{2}\}$[1];
4) $j=\frac{3t}{2}+j',1\le j'\le\frac{t}{2}$:
   a) $j'\notin\{1,2,3,\frac{t}{2}\}$,
   b) $j'\in\{1,2,3,\frac{t}{2}\}$[2].

We now show that the binary weight of $(2^t+2^{\frac{t}{2}}-1)(2^j+1),1\le j\le 2t$ takes every integer value from the interval $[\frac{t}{2}+2;t+2]$ depending on j.

[1] $j'=2$ is meaningful only for $t>2$.
[2] $j'=2$ is meaningful only for $t>2$, $j'=3$ is meaningful only for $t>4$.

Case 1. For $1 \leq j \leq \frac{t}{2}$, we get:

$$(2^t + 2^{\frac{t}{2}} - 1)(2^j + 1) = 2^{t+j} + 2^{\frac{t}{2}+j} + 2^t + 2^{\frac{t}{2}} - 2^j - 1.$$

(a) If $j \notin \{\frac{t}{2}, \frac{t}{2} - 1\}$, then we immediately get that

$$wt\Bigg(\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^j + 1\Big)\Bigg) = wt\Big(2^{t+j} + 2^{\frac{t}{2}+j} + 2^t\Big) + wt\Big(2^{\frac{t}{2}} - 2^j - 1\Big) = 3 + \Big(\frac{t}{2} - 1\Big) = \frac{t}{2} + 2.$$

(b) If $j \in \{\frac{t}{2}, \frac{t}{2} - 1\}$, then, for instance, when $j = \frac{t}{2} - 1$, we get:

$$\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^{\frac{t}{2}-1} + 1\Big) = 2^{\frac{3t}{2}-1} + 2^{t-1} + 2^t + 2^{\frac{t}{2}} - 2^{\frac{t}{2}-1} - 1 = 2^{\frac{3t}{2}-1} + 2^t + 2^{t-1} + 2^{\frac{t}{2}-1} - 1,$$

hence

$$wt\Bigg(\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^{\frac{t}{2}-1} + 1\Big)\Bigg) = wt\Big(2^{\frac{3t}{2}-1} + 2^t + 2^{t-1}\Big) + \Big(2^{\frac{t}{2}-1} - 1\Big) = 3 + \Big(\frac{t}{2} - 1\Big) = \frac{t}{2} + 2;$$

In the same way, for $j = \frac{t}{2}$ we get

$$\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^{\frac{t}{2}} + 1\Big) = 2^{\frac{3t}{2}} + 2^t + 2^t + 2^{\frac{t}{2}} - 2^{\frac{t}{2}} - 1 = 2^{\frac{3t}{2}} + 2^{t+1} - 1,$$

hence

$$wt\Bigg(\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^{\frac{t}{2}-1} + 1\Big)\Bigg) = wt\Big(\ 2^{\frac{3t}{2}}\ \Big) + wt\Big(2^{t+1} - 1\Big) = 1 + (t+1) = t + 2.$$

Case 2. For $j = \frac{t}{2} + j', 1 \leq j' \leq \frac{t}{2}$.

$$\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^{\frac{t}{2}+j'} + 1\Big) = 2^{\frac{3t}{2}+j'} + 2^{t+j'} + 2^t - 2^{\frac{t}{2}+j'} + 2^{\frac{t}{2}} - 1.$$

(a) If $j' \notin \{1, \frac{t}{2} - 1, \frac{t}{2}\}$, then

$$wt\Bigg(\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^j + 1\Big)\Bigg) = wt\Big(2^{\frac{3t}{2}+j'} + 2^{t+j'}\Big) + wt\Big(2^t - 2^{\frac{t}{2}+j'}\Big) + wt\Big(2^{\frac{t}{2}} - 1\Big) =$$

$$2 + \Big(t - \Big(\frac{t}{2} + j'\Big)\Big) + \frac{t}{2} = (t+2) - j'.$$

Thus, $\frac{t}{2} + 3 < wt\Bigg(\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^j + 1\Big)\Bigg) < t + 1$, for $1 < j' < \frac{t}{2} - 1$.

(b) If $j' \in \{1, \frac{t}{2} - 1, \frac{t}{2}\}$, then:

- for $j' = 1$:

$$\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^{\frac{t}{2}+1} + 1\Big) = 2^{\frac{3t}{2}+1} + 2^{t+1} - 2^{\frac{t}{2}+1} + 2^t + 2^{\frac{t}{2}} - 1 = 2^{\frac{3t}{2}+1} + 2^{t+1} + 2^t - 2^{\frac{t}{2}} - 1,$$

hence

$$wt\Bigg(\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^{\frac{t}{2}+1} + 1\Big)\Bigg) = wt\Big(2^{\frac{3t}{2}+1} + 2^{t+1}\Big) + \Big(2^t - 2^{\frac{t}{2}} - 1\Big) = 2 + (t-1) = t + 1.$$

- for $j' = \frac{t}{2} - 1$:

$$\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^{t-1} + 1\Big) = 2^{2t-1} + 2^{\frac{3t}{2}-1} - 2^{t-1} + 2^t + 2^{\frac{t}{2}} - 1 = 2^{2t-1} + 2^{\frac{3t}{2}-1} + 2^{t-1} + 2^{\frac{t}{2}} - 1,$$

thus

$$wt\Bigg(\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^{t-1} + 1\Big)\Bigg) = wt\Big(2^{2t-1} + 2^{\frac{3t}{2}-1} + 2^{t-1}\Big) + wt\Big(2^{\frac{t}{2}} - 1\Big) = \frac{t}{2} + 3.$$

- for $j' = \frac{t}{2}$:

$$\Big(2^t + 2^{\frac{t}{2}} - 1\Big)\Big(2^t + 1\Big) = 2^{2t} + 2^{\frac{3t}{2}} - 2^t + 2^t + 2^{\frac{t}{2}} - 1 = 2^{2t} + 2^{\frac{3t}{2}} + 2^{\frac{t}{2}} - 1,$$

hence

$$wt\left(\left(2^t+2^{\frac{t}{2}}-1\right)\left(2^t+1\right)\right)=wt\left(2^{2t}+2^{\frac{3t}{2}}\right)+wt\left(2^{\frac{t}{2}}-1\right)=\frac{t}{2}+2.$$

Thus, in Case 2 we can conclude that $wt\left(\left(2^t+2^{\frac{t}{2}}-1\right)\left(2^{\frac{t}{2}+j'}+1\right)\right)$, $1\le j'\le\frac{t}{2}$ takes every integer value from the interval $[\frac{t}{2}+2;t+1]$.

In the exact same way, we can obtain the following results for the two remaining cases; we omit the proofs for the sake of brevity.

Case 3. $wt\left(\left(2^t+2^{\frac{t}{2}}-1\right)\left(2^{t+j'}+1\right)\right)$, $1\le j'\le\frac{t}{2}-1$ takes every integer value from the interval $[\frac{t}{2}+2;t-1]$.

Case 4. $wt\left(\left(2^t+2^{\frac{t}{2}}-1\right)\left(2^{\frac{3}{2}t+j'}+1\right)\right)$, $1\le j'<\frac{t}{2}$ equals to either $\frac{t}{2}+2$ or $t+2$.

Thus, the weight of the right-hand side of (2) takes every integer value from the interval $[\frac{t}{2}+2;t+2]$ depending on j. □

The following corollary follows immediately from Lemmas 3 and 4. Recall that, in the case of power functions, cyclotomic equivalence, EA-equivalence, and CCZ-equivalence coincide. Throughout the rest of this section, if we refer to two power functions being equivalent or inequivalent, we are having cyclotomic equivalence in mind.

Corollary 5. *Let x^d be a power function defined over the field $\mathbb{F}_{2^{2t+1}}$ with $d=2^t+2^{\frac{t}{2}}-1$, with t. Then x^d is cyclotomic equivalent to the power functions with exponents $\frac{3}{2^{\frac{3t}{2}}+2^t+1}$, $\frac{3}{2^{\frac{3t}{2}+1}+2^{\frac{t}{2}+1}+1}$ and $\frac{3}{2^{2^{t+1}+\frac{t}{2}}+1}$. Furthermore, these representations are optimal, in the sense that x^d is cyclotomic inequivalent to any power function whose exponent is a fraction of two quadratic exponents.*

The next two lemmas address the exponent of the Niho power APN function in the odd case.

Lemma 6. *For any positive odd integer t, we have*

$$\begin{aligned}(2^{\frac{3t+1}{2}}+2^t-1)\equiv 2^t\frac{3}{2^{\frac{3(t+1)}{2}}+2^{\frac{t+1}{2}}+1}\equiv\\ 2^{t-1}\frac{3}{2^{\frac{3t+1}{2}}+2^{t+1}+1}\equiv\\ 2^{\frac{3t-1}{2}}\frac{3}{2^t+2^{\frac{t-1}{2}}+1}\mod(2^{2t+1}-1)\end{aligned}$$

Proof. Following the proof of Lemma 3, we first show that $2^{\frac{3(t+1)}{2}}+2^{\frac{t+1}{2}}+1, 2^{\frac{3t+1}{2}}+2^{t+1}+1$ and $2^t+2^{\frac{t-1}{2}}+1$ are invertible modulo $2^{2t+1}-1$, then we prove the corresponding congruences.

For simplicity, we denote $2^{\frac{t+1}{2}}$ by y and apply the Euclidean algorithm to obtain

$$\gcd\left(\frac{1}{2}y^4-1,y^3+y+1\right)=\gcd\left(y^3+y+1,\frac{1}{2}y^2+\frac{1}{2}y+1\right)=\gcd\left(\frac{1}{2}y^2+\frac{1}{2}y+1,3\right)=1,$$

$$\gcd\left(\frac{1}{2}y^4-1,\frac{1}{2}y^3+y^2+1\right)=\gcd\left(\frac{1}{2}y^3+y^2+1,2y^2-y+1\right)=$$

$$\gcd\left(2y^2-y+1,y-1\right)=\gcd(y-1,2)=1$$

and

$$\gcd\left(\frac{1}{2}y^4-1,\frac{1}{2}y^2+\frac{1}{2}y+1\right)=\gcd\left(\frac{1}{2}y^2+\frac{1}{2}y+1,y\right)=1.$$

To prove the first statement of this lemma it remains to verify the congruence: $2(y^3+y^2-1)(y^3+y+1)\equiv 3y^2 \mod (y^4-2)$ which is straightforward. Indeed, computing the left-hand side of it, we get

$$(y^3+y^2-2)(y^3+y+1)=y^6+y^5+y^4+y^2-2y-1\equiv 3y^2 \mod (y^4-2).$$

The remaining two congruences are proven in a similar way. □

Lemma 7. *Let t be a positive odd integer. Then, for any positive integers j, l, r such that $1 \leq j, l, r < 2t+1$, the following incongruence holds:*

$$(2^{\frac{3t+1}{2}} + 2^t - 1)(2^j + 1) \not\equiv 2^l + 2^r \mod (2^{2t+1} - 1). \tag{3}$$

Proof. Similar to Lemma 4, we shall show that for any $1 \leq j < 2t+1$, the binary weight of $\left(2^{\frac{3t+1}{2}} + 2^t - 1\right)\left(2^j + 1\right)$ is always strictly greater than 2. Consider the following possible cases:

1) $1 \leq j < \frac{t+1}{2}$;
2) $j = \frac{t-1}{2} + j', 1 \leq j' \leq \frac{t+1}{2}$:
   a) $j' \notin \{1, 2, \frac{t-1}{2}, \frac{t+1}{2}\}$,
   b) $j' \in \{1, 2, \frac{t-1}{2}, \frac{t+1}{2}\}$[3];
3) $j = t + j', 1 \leq j' \leq \frac{t+1}{2}$:
   a) $j' \notin \{1, 2, \frac{t-1}{2}, \frac{t+1}{2}\}$,
   b) $j' \in \{1, 2, \frac{t-1}{2}, \frac{t+1}{2}\}$[4];
4) $j = \frac{3t+1}{2} + j', 1 \leq j' \leq \frac{t-1}{2}$:
   a) $j' \notin \{1, \frac{t-1}{2}\}$,
   b) $j' \in \{1, \frac{t-1}{2}\}$.

In all cases, the binary weight the left-hand side of (3) equals either to $t+2$, $\frac{t+3}{2}$ or $\frac{t+1}{2}$, depending on a choice of j.

Case 1: $1 \leq j < \frac{t+1}{2}$.

$$\left(2^{\frac{3t+1}{2}} + 2^t - 1\right)\left(2^j + 1\right) = 2^{\frac{3t+1}{2}+j} + 2^{\frac{3t+1}{2}} + 2^{t+j} + 2^t - 2^j - 1,$$

hence

$$wt\left(\left(2^{\frac{3t+1}{2}} + 2^t - 1\right)\left(2^j + 1\right)\right) = wt\left(2^{\frac{3t+1}{2}+j} + 2^{\frac{3t+1}{2}} + 2^{t+j}\right) + wt\left(2^t - 2^j - 1\right) = 3 + (t-1) = t+2.$$

Case 2. $j = \frac{t-1}{2} + j'$, $1 \leq j' \leq \frac{t+1}{2}$:

$$\left(2^{\frac{3t+1}{2}} + 2^t - 1\right)\left(2^{\frac{t-1}{2}+j'} + 1\right) = 2^{2t+j'} + 2^{\frac{3t-1}{2}+j'} - 2^{\frac{t-1}{2}+j'} + 2^{\frac{3t+1}{2}} + 2^t - 1 \equiv$$

$$2^{\frac{3t-1}{2}+j'} + 2^{\frac{3t+1}{2}} + 2^t - 2^{\frac{t-1}{2}+j'} + 2^{j'-1} - 1 \mod (2^{2t+1} - 1).$$

(a) If $j' \notin \{1, 2, \frac{t-1}{2}, \frac{t+1}{2}\}$, then

$$wt\left(\left(2^{\frac{3t+1}{2}} + 2^t - 1\right)\left(2^{\frac{t+1}{2}+j'} + 1\right)\right) = wt\left(2^{\frac{3t-1}{2}+j'} + 2^{\frac{3t+1}{2}}\right) + wt\left(2^t - 2^{\frac{t-1}{2}+j'}\right) + wt\left(2^{j'-1} - 1\right) =$$

$$2 + \left(t - \left(\frac{t-1}{2} + j'\right)\right) + (j' - 1) = \frac{t+3}{2}.$$

(b) If $j' \in \{1, 2, \frac{t-1}{2}, \frac{t+1}{2}\}$, then, for instance, take $j' = 1$:

$$\left(2^{\frac{3t+1}{2}} + 2^t - 1\right)\left(2^{\frac{t-1}{2}+1} + 1\right) \equiv 2^{\frac{3t-1}{2}+1} + 2^{\frac{3t+1}{2}} + 2^t - 2^{\frac{t-1}{2}+1} \equiv 2^{\frac{3t+1}{2}+1} + 2^t - 2^{\frac{t+1}{2}} \mod (2^{2t+1} - 1),$$

hence

$$wt\left(\left(2^{\frac{3t+1}{2}} + 2^t - 1\right)\left(2^{\frac{t-1}{2}+1} + 1\right)\right) = wt\left(2^{\frac{3t+1}{2}+1}\right) + wt\left(2^t - 2^{\frac{t+1}{2}}\right) = 1 + \left(t - \frac{t+1}{2}\right) = \frac{t+1}{2}.$$

The remaining cases are proven in the exact same way. □

The next statement follows from Lemma 6 and Lemma 7.

Corollary 8. *Let x^d be a power function defined over the field $\mathbb{F}_{2^{2t+1}}$ with $d = 2^{\frac{3t+1}{2}} + 2^t - 1$, with t odd. Then x^d is cyclotomic equivalent to the power functions with exponents $\frac{3}{2^{\frac{3(t+1)}{2}} + 2^{\frac{t+1}{2}} + 1}$, $\frac{3}{2^{\frac{3t+1}{2}} + 2^{t+1} + 1}$ and $\frac{3}{2^t + 2^{\frac{t-1}{2}} + 1}$. Furthermore, these representations are optimal, in the sense that x^d is cyclotomic inequivalent to any function whose exponent is a fraction of two quadratic exponents.*

[3] $j' = 2$ is meaningful only for $t > 3$.
[4] $j' = 2$ is meaningful only for $t > 3$.

Let us focus now on the exponent of the Dobbertin power function $d = 2^{4m} + 2^{3m} + 2^{2m} + 2^m - 1$. In the next two lemmas we prove that x^d is equivalent to a power function composed from a cubic power function and the inverse of a quadratic power function.

Lemma 9. *For any positive integer m the following equivalences are true:*

$$\sum_{i=1}^{4} 2^{im} - 1 \equiv 2^{2m+1}\frac{2^{2m}+2^m+1}{2^m+1} \equiv$$
$$2^{m+1}\frac{2^{3m}+2^{2m}+1}{2^{2m}+1} \equiv$$
$$2^{m+1}\frac{2^{3m}+2^m+1}{2^{3m}+1} \equiv$$
$$2^{m+1}\frac{2^{2m}+2^m+1}{2^{4m}+1} \mod (2^{5m}-1).$$

Proof. Consider the first congruence. We first prove that 2^m+1 is invertible modulo $2^{5m}-1$, i.e. that $\gcd(2^m+1, 2^{5m}-1) = 1$. This follows from

$$\gcd(2^k+1, 2^l-1) = \begin{cases} 1, & \text{if } l/\gcd(l,k) \text{ is odd;} \\ 2^{\gcd(l,k)}+1, & \text{if } l/\gcd(l,k) \text{ is even.} \end{cases} \tag{4}$$

Indeed, since $\gcd(m, 5m) = m$, we have $\gcd(2^m+1, 2^{5m}-1) = 1$.
For simplicity, we denote 2^m by y. It remains to check the equivalence $(y+1)(y^4+y^3+y^2+y-1) \equiv 2y^2(y^2+y+1) \mod (y^5-1)$ which is straightforward. Indeed, computing the left-hand side of this equivalence, we get
$(y+1)(y^4+y^3+y^2+y-1) = y^5+2y^4+2y^3+2y^2-1 \equiv 2y^4+2y^3+2y^2 \mod (y^5-1)$. This proves the first statement of the lemma.

The other three equivalences are proven in the same way. A justification that $2^{2m}+1$, $2^{3m}+1$ and $2^{4m}+1$ are invertible modulo $2^{5m}-1$ easily follows from (4). The corresponding congruences are then straightforward to check. □

Lemma 10. *Let m be a positive integer. Then, for any positive integers j, l, r such that $1 \le j, l, r < 5m$, the following inequivalence holds:*

$$\left(\sum_{i=1}^{4} 2^{im} - 1\right)(2^j+1) \not\equiv 2^l + 2^r \mod (2^{5m}-1). \tag{5}$$

Proof. We shall show that for any $1 \le j < 5m$, the binary weight of the left-hand side of (5) is always strictly greater than 2. The cases $j \in \{m, 2m, 3m, 4m\}$ are covered in Lemma 1 when the binary weight of $\left(\sum_{i=1}^{4} 2^{im} - 1\right)(2^j+1)$ equals 3. We thus consider the remaining 5 cases:

1) $1 \le j < m$,
2) $j = m + j', 1 \le j' < m$,
3) $j = 2m + j', 1 \le j' < m$,
4) $j = 3m + j', 1 \le j' < m$,
5) $j = 4m + j', 1 \le j' < m$.

In all of these cases, the binary weight of $\left(\sum_{i=1}^{4} 2^{im} - 1\right)(2^j+1)$ is equal to $m+6$. Indeed, for $1 \le j < m$ we get

$$\left(\sum_{i=1}^{4} 2^{im} - 1\right)(2^j+1) \equiv \left(\sum_{i=1}^{4} 2^{im+j} - 2^j + \sum_{i=1}^{4} 2^{im} - 1\right) \mod (2^{5m}-1).$$

Hence,

$$wt\left(\left(\sum_{i=1}^{4} 2^{im} - 1\right)(2^j+1)\right) = wt\Big(\sum_{i=1}^{4} 2^{im+j} + \sum_{i=2}^{4} 2^{im}\Big) + wt(2^m - 2^j - 1) = 7 + (m-1) = m+6.$$

Similarly, for $j = m + j', 1 \le j' < m$:

$$\left(\sum_{i=1}^{4} 2^{im} - 1\right)(2^j+1) = \left(\sum_{i=1}^{4} 2^{im} - 1\right)(2^{m+j'}+1) = \sum_{i=2}^{5} 2^{im+j'} - 2^{m+j'} + \sum_{i=1}^{4} 2^{im} - 1$$
$$\equiv \Big(\sum_{i=2}^{4} 2^{im+j'} - 2^{m+j'} + \sum_{i=1}^{4} 2^{im} + 2^{j'} - 1\Big) \mod (2^{5m}-1).$$

Therefore,

$$\begin{aligned} wt\left(\left(\sum_{i=1}^{4} 2^{im}-1\right)(2^j+1)\right) &= wt\left(\sum_{i=2}^{4} 2^{im+j'}-2^{m+j'}+\sum_{i=1}^{4} 2^{im}+2^{j'}-1\right) \\ &= wt(\sum_{i=2}^{4} 2^{im+j'}+2^{4m}+2^{3m}+2^m)+wt(2^{2m}-2^{m+j'}) \\ &\quad +wt(2^{j'}-1)=6+(m-j')+j'=6+m. \end{aligned}$$

The remaining cases are proven in the exact same way. □

The following corollary is a straightforward consequence of Lemma 1 and Lemma 2.

Corollary 11. *Let x^d be the power function defined over the field $\mathbb{F}_{2^{5m}}$ with $d = 2^{4m}+2^{3m}+2^{2m}+2^m-1$. Then x^d is cyclotomic equivalent to the power functions with the exponents $\frac{2^{2m}+2^m+1}{2^m+1}$, $\frac{2^{3m}+2^{2m}+1}{2^{2m}+1}$, $\frac{2^{3m}+2^m+1}{2^{3m}+1}$ and $\frac{2^{2m}+2^m+1}{2^{4m}+1}$. Furthermore, these representations are optimal, in the sense that x^d is cyclotomic inequivalent to any power function whose exponent is a fraction of two quadratic exponents.*

III. Composition of power functions with linear functions

In this section, we consider a different way of representing some of the known APN power functions, in which a linear polynomial is "inserted" between two power functions, so that the composition has the form $x^i \circ L \circ x^j$. To facilitate the discussion, we introduce the following shorthand notation for the various APN power functions:

1) $P_i(x) = x^i$ for any positive integer i;
2) $G_i(x) = x^{2^i+1}$ is the Gold function with parameter i;
3) $K_i(x) = x^{2^{2i}-2^i+1}$ is the Kasami function with parameter i;
4) $W(x) = x^{2^t+3}$ is the Welch function, where $n = 2t+1$;
5) $N(x) = x^{2^t+2^{t/2}-1}$ and $N(x) = x^{2^t+2^{(3t+1)/2}-1}$ is the Niho function for t even and for t odd, respectively, where $n = 2t+1$;
6) $I(x) = x^{2^n-2}$ is the inverse function;
7) $D(x) = x^{2^{4i}+2^{3i}+2^{2i}+2^i-1}$ is the Dobbertin function, where $n = 5i$.

A. The case of odd dimension

This direction of study is motivated by an initial observation that, over any finite field $\mathbb{F}_{2^n}$ with n odd, composing the Gold function $G_i(x) = x^{2^i+1}$ with its inverse $G_i^{-1}(x)$ (where i is any positive integer with $\gcd(i,n) = 1$) and the linear polynomial $L(x) = x^{2^{2i}} + x$ in between gives a function EA-equivalent to the Kasami function $K_i(x) = x^{2^{2i}-2^i+1}$ with the same parameter i. More precisely, we observe that

$$G_i \circ L \circ G_i^{-1}(x) = K_i(x) + x^{2^{2i}} + x^{2^i} + x.$$

More generally, taking $L_\mu(x) = x^{2^{2i}} + \mu x$, we have

$$G_i \circ L_\mu \circ G_i^{-1}(x) = \mu K_i(x) + x^{2^{2i}} + \mu^{2^i} x^{2^i} + \mu^{2^i+1} x$$

for any $\mu \in \mathbb{F}_{2^n}^*$.

We thus see that in certain cases, a function CCZ-equivalent to a Kasami function can be obtained by combining a Gold function and the inverse of a Gold function with a linear polynomial. A formal treatment of this observation is provided in the following proposition. This suggests that functions CCZ-inequivalent to P_i and P_j can be expressed as $P_i \circ L \circ P_j$. We contrast this with EA-equivalence, in which an (n,n)-function F is composed with two linear permutations L_1, L_2 in the form $L_1 \circ F \circ L_2$. We note that all linear polynomials L that we compose with in Propositions 12 and 13 are 2-to-1 over $\mathbb{F}_{2^n}$, while the linear functions L_1 and L_2 in the definition of EA-equivalence are necessarily bijective. In particular, this shows that while the Kasami functions (and their inverses) are always 1-to-1 functions for odd dimensions, the addition of certain linear functions can make them 2-to-1 functions.

Proposition 12. *Let $n = 2m+1$, and denote $L_i^\mu(x) = \mu x^{2^i} + x$. Then, for any $1 \le i \le n-1$, we have*

$$G_i \circ L_{2i}^\mu \circ G_i^{-1}(x) = A_i^\mu(x) + \mu^{2^i} K_i(x), \tag{6}$$

where $A_i^\mu(x) = \mu^{2^i+1} x^{2^{2i}} + \mu x^{2^i} + x$.

Similarly, for any $1 \le i \le n-1$, we have

$$G_i \circ L_{n-2i}^\mu \circ G_i^{-1}(x) = \mu K_i(x^{2^{-2i}}) + C_i^\mu(x^{2^{-2i}}), \tag{7}$$

where $C_i^\mu(x) = \mu^{2^i+1}x + \mu^{2^i}x^{2^i} + x^{2^{2i}}$.

Proof. Denoting $x = y^{2^i+1}$, we obtain

$$\begin{aligned} G_i \circ L_{2i}^\mu \circ G_i^{-1}(x) &= \left(\mu y^{2^{2i}} + y\right)^{2^i+1} \\ &= \mu^{2^i+1}y^{2^{2i}(2^i+1)} + \mu^{2^i}y^{2^{3i}+1} + \mu y^{2^i(2^i+1)} + y^{2^i+1} \\ &= \mu^{2^i+1}x^{2^{2i}} + \mu^{2^i}x^{(2^{3i}+1)/(2^i+1)} + \mu x^{2^i} + x \\ &= A_i^\mu(x) + \mu^{2^i}K_i(x) \end{aligned}$$

due to $K_i(x) = x^{2^{2i}-2^i+1} = x^{(2^{3i}+1)/(2^i+1)}$. The proof in the case of L_{n-2i}^μ is similar. □

A natural question is whether APN functions other than the Kasami function can be obtained in the same manner. The following two propositions demonstrate two ways in which we can reach the EA-equivalence class of the inverse of the Kasami function by composing a Gold function and the inverse of a Gold function (with different parameters) with a linear polynomial in between. We note that the polynomial expression of the inverse of the Kasami APN function in odd dimension (that is, the expression of its exponent as a power function) can be quite complex [29]. The expression of K_i^{-1} in Proposition 13 is therefore rather interesting in this sense. We note that explicit formulas for the inverses of the Dobbertin and Welch exponents have previously been studied in [28].

Proposition 13. *Let* $n = 3s \pm r$, $3s \geq r$ *and* $\gcd(3s, r) = 1$, n *odd, and let* $L_i^\mu(x) = \mu x^{2^i} + x$. *Then*

$$G_s \circ L_{2s}^\mu \circ G_r^{-1}(x) = \begin{cases} A^\mu \circ K_s^{-1}(x^{2^{3s}}) + \mu^{2^s}x^{2^{3s}} & n = 3s + r \\ A^\mu \circ K_s^{-1}(x) + \mu^{2^s}x^{2^s} & n = 3s - r, \end{cases} \tag{8}$$

where $A^\mu(x) = \mu^{2^s+1}x^{2^{2s}} + \mu x^{2^s} + x$ *is a linear permutation.*

Similarly, we have

$$G_s \circ L_{n-2s}^\mu \circ G_r^{-1}(x) = \begin{cases} B_s^\mu \circ K_s^{-1}(x) + \mu x^{2^{-2s}} & n = 3s - r \\ B_s^\mu \circ K_s^{-1}(x^{2^{3s}}) + \mu x^{2^s} & n = 3s + r, \end{cases} \tag{9}$$

where $B_s^\mu(x) = x + \mu^{2^s}x^{2^{n-s}} + \mu^{2^s+1}x^{2^{n-2s}}$ *is a linear permutation.*

Proof. Denoting by $y = x^{1/(2^r+1)}$ the inverse of $G_r(x)$, we obtain by straightforward manipulation

$$\begin{aligned} G_s \circ L_{2s}^\mu \circ G_r^{-1}(x) = G_s \circ L_{2s}^\mu(y) &= \left(\mu y^{2^{2s}} + y\right)^{2^s+1} \\ &= \mu^{2^s+1}y^{2^{2s}(2^s+1)} + \mu^{2^s}y^{2^{3s}+1} + \mu y^{2^s(2^s+1)} + y^{2^s+1} \\ &= A^\mu\left(y^{2^s+1}\right) + \mu^{2^s}y^{(2^{3s}+1)}. \end{aligned}$$

Suppose now that $n = 3s + r$. Then

$$\frac{1}{2^r+1} \equiv \frac{2^n}{2^r+2^n} \equiv \frac{2^{3s+r}}{2^r(2^{3s}+1)} \equiv \frac{2^{3s}}{2^{3s}+1} \mod (2^n - 1),$$

so that $y^{2^s+1} = x^{(2^s+1)/(2^r+1)} = x^{2^{3s}(2^s+1)/(2^{3s}+1)}$, which is precisely $K_s^{-1}(x^{2^{3s}})$ since $K_s(x) = x^{2^{2s}-2^s+1}$; equivalently, $K_s(x) = x^{(2^{3s}+1)/(2^s+1)}$, whence $K_s^{-1}(x) = x^{(2^s+1)/(2^{3s}+1)}$. Similarly, $\mu y^{2^{3s}+1} = \mu x^{(2^{3s}+1)/(2^r+1)} = \mu x^{2^{3s}}$, which concludes the proof in the case of $n = 3s + r$.

When $n = 3s - r$, we have

$$\frac{1}{2^r+1} \equiv \frac{1}{2^{n+r}+1} \equiv \frac{1}{2^{3s}+1} \mod (2^n - 1),$$

so that $y^{2^s+1} = x^{(2^s+1)/(2^{3s}+1)} = K_s^{-1}(x)$, and $\mu y^{2^{3s}+1} = \mu x^{2^{3s}+1}2^r + 1 = x$, concluding the proof for L_{2s}^μ.

Let j be a positive integer. We will prove that $\mu^{2^j+1}x^{2^{2j}} + \mu x^{2^j} + x$ permutes $\mathbb{F}_{2^n}$ whenever $3 \nmid n$ by showing that it has a trivial kernel. Suppose that $\mu^{2^j+1}x^{2^{2j}} + \mu x^{2^j} + x = 0$. Raising both sides to the power 2^j and multiplying by μ, we obtain $\mu^{2^{2j}+2^j+1}x^{2^{3j}} + \mu^{2^j+1}x^{2^{2j}} + \mu x^{2^j} = 0$. Summing both of these identities, we have $x = \mu^{2^{2j}+2^j+1}x^{2^{3j}}$, and hence, assuming $x \neq 0$, $x^{2^{3j}-1} = (1/\mu)^{2^{2j}+2^j+1}$. Since $2^{3j} - 1 = (2^{2j} + 2^j + 1)(2^j - 1)$, and $\gcd(2^{2j} + 2^j + 1, 2^n - 1) = 1$ for $3 \nmid n$, this implies $x^{2^j-1} = 1/\mu$, whence $x^{2^j} = x/\mu$ and $x^{2^{2j}} = x/\mu^{2^j+1}$. Substituting this into $\mu^{2^j+1}x^{2^{2j}} + \mu x^{2^j} + x = 0$, we obtain $x/\mu = 0$, implying $x = 0$ and contradicting our assumption that $x \neq 0$.

The proof for B_s^μ follows the same logic. Denoting once again $y = x^{1/(2^r+1)}$, we obtain

$$\begin{aligned} G_s \circ L_{n-2s}^\mu \circ G_r^{-1}(x) &= \left(y + \mu y^{(2^{n-2s})}\right)^{2^s+1} \\ &= y^{2^s+1} + \mu y^{2^{n-2s}+2^s} + \mu^{2^s} y^{2^{n-s}+1} + \mu^{2^s+1} y^{2^{n-2s}+2^{n-s}} \\ &= B_s^\mu(y^{2^s+1}) + \mu y^{2^{n-2s}+2^s}. \end{aligned}$$

We have already seen that y^{2^s+1} becomes $K_s^{-1}(x^{2^{3s}})$, resp. $K_s^{-1}(x)$ when $n = 3s+r$, resp. $n = 3s-r$. When $n = 3s+r$, the term $\mu y^{2^{n-2s}+2^s}$ becomes

$$\mu y^{2^{s+r}+2^s} = \mu x^{2^s(2^r+1)/(2^r+1)} = \mu x^{2^s};$$

when $n = 3s-r$, we have

$$\mu y^{2^{n-2s}+2^s} = \mu y^{2^{s-r}+2^s} = \mu x^{2^{s-r}(2^r+1)/(2^r+1)} = \mu x^{2^{s-r}} = \mu x^{2^{-2s}}.$$

Finally, showing that $B_s^\mu(x)$ is a permutation is done in the same way as for $\mu^{2^j+1}x^{2^{2j}} + \mu x^{2^j} + x$. □

While Proposition 13 explicitly describes only compositions of the form $G_s \circ L \circ G_r^{-1}$ over $\mathbb{F}_{2^n}$, where $n = 3s \pm r$, we can observe that G_s and G_{n-s} yield EA-equivalent functions, and so the parameters s and r can be freely replaced with $n-s$ and $n-r$, respectively, thereby allowing for a wider range of compositions. Furthermore, if $s \equiv s' \mod n$, then G_s and $G_{s'}$ correspond to the same function, and so arbitrary multiples of the dimension n can be added or subtracted, allowing us even more freedom. We thus have the following general principle.

Remark 14. Assuming the notation of Proposition 13, the following compositions are all equivalent for any linear function L:

$$\begin{aligned} G_i \circ L \circ G_j^{-1}, \\ G_{n-i} \circ L \circ G_j^{-1}, \\ G_i \circ L \circ G_{n-j}^{-1}, \\ G_{n-i} \circ L \circ G_{n-j}^{-1}. \end{aligned}$$

For instance, the composition $G_1 \circ L \circ G_3^{-1}$ over $\mathbb{F}_{2^7}$ cannot be directly expressed using Proposition 13; but taking $s = n-1 = 6$, and $r = 11 \equiv 4 \mod n$ so that $n-3 = 4$, we have $n = 3 \cdot s - r$, and we obtain the case $G_1 \circ L \circ G_3^{-1}$.

Corollary 15. *Let $n = 2m+1$ be odd with $3 \nmid n$, and let i be a positive integer in the range $1 \le i \le n-1$ such that $\gcd(i,n) = 1$. Let $\mu \in \mathbb{F}_{2^n}^*$ be arbitrary, and denote $L_i^\mu(x) = \mu x^{2^i} + x$ as before. Then the functions*

$$G_i \circ L_{2i}^\mu \circ G_{3i}^{-1}$$

and

$$G_i \circ L_{n-2i}^\mu \circ G_{3i}^{-1}$$

are APN, and EA-equivalent to the inverse K_i^{-1} of the Kasami function with parameter i.

Proof. Take $s = i+n$ and $r = 3s-n$. We have $3s-r = n$. Furthermore, $s \equiv i \mod n$, and $r \equiv 3i \mod n$. Thus, we only have to show that the pair (s,r) satisfies the hypothesis of Proposition 13 in order to finish the proof. We want to show that $|r| \le 3s$, i.e. $-3s \le 3s-n \le 3s$, which gives the inequalities $n \ge 0$ and $n \le 6s \le 6i+6n$. Both of these are clearly always satisfied. Finally, we need to show that $\gcd(3s,r) = 1$. Clearly, $3 \nmid r$ since $3 \nmid n$ by the hypothesis; thus, we only need to show that $\gcd(s,r) = 1$. Suppose d is a non-trivial common divisor of s and $r = 3s-n$; then d is a non-trivial common divisor of $s = i+n$ and n, and hence of i and n. But since $1 \le i \le n-1$ by assumption, we reach a contradiction, and thus $\gcd(s,r) = \gcd(3s,r) = 1$ as claimed. Now, all conditions on (s,r) from the hypothesis of Proposition 13 are satisfied, and an application of the latter concludes the proof.

□

Remark 16. We note that while Propositions 12 and 13 describe cases in which a composition of the form $P_i \circ L \circ P_j$ is EA-equivalent to a Kasami K_i function (or its inverse), in some cases we obtain K_1 (or its inverse), which is actually the Gold function G_1 (or its inverse). In particular, this happens in Proposition 12 for $i = 1$, and in Proposition 13 for $s = 1$.

In our experimental results, we also observe combinations of the form $G_t^{-1} \circ L \circ G_t$, which are EA-equivalent to G_t^{-1}, and combinations of the form $I \circ L \circ I$, which gives a function EA-equivalent to the inverse function I.

Observation 17. *Let $n = 2t+1$. Then the compositional inverse of $G_t(x) = x^{2^t+1}$ is $x^{2^{t+1}(2^{t+1}-1)}$. Consequently, the composition $G_t^{-1} \circ L \circ G_t$ becomes*

$$G_t^{-1} \circ L \circ G_t(x) = \left(x^{2^t+1} + x^{2^{2t}+2^t}\right)^{2^{t+1}\cdot(2^{t+1}-1)} \tag{10}$$

for $L = x^{2^t} + x$, and

$$G_t^{-1} \circ L \circ G_t(x) = \left(x^{2^t+1} + x^{2^{2t+1}+2^{t+1}}\right)^{2^{t+1}\cdot(2^{t+1}-1)} \tag{11}$$

for $L = x^{2^{t+1}} + x$. Similarly, we get

$$I \circ L \circ I(x) = \left(x^{2^{2t}-1} + x^{2^{2t+1}-2}\right)^{2^{2t}-1} \tag{12}$$

for $L = x^2 + x$, and

$$I \circ L \circ I(x) = \left(x^{2^{2t}-1} + x^{2^{4t}-2^{2t}}\right)^{2^{2t}-1} \tag{13}$$

for $L = x^{2^{2t}} + x$.

The functions in (10) *and* (11)*, and* (12) *and* (13) *are EA-equivalent to G_t^{-1} and I, respectively. Furthermore, for $n \in \{3, 5, 7, 9\}$, the combinations described in* (10)*,* (11)*,* (12)*, and* (13)*, and Propositions 12 and 13 exhaust all APN functions over $\mathbb{F}_{2^n}$ that can be obtained as $P_i \circ L \circ P_j$ for any affine function L with coefficients in $\mathbb{F}_2$.*

Proof. We show that the functions from (10) and (12) are EA-equivalent to the Gold and inverse functions, respectively.

In the Gold case, we have $n = 2t + 1$, and $G_t^{-1} \circ L \circ G_t = (x^{2^{t+1}+1} + x^{2^t+1})^{2^{t+1}-1}$. Since $2^{t+1} - 1 = 2^t + 2^{t-1} + \cdots + 1$, we have that this is equal to

$$\prod_{j=0}^{t}(x^{2^{t+1}+1} + x^{2^t+1})^{2^j} = \prod_{j=0}^{t} x^{2^j(2^t+1)} \prod_{j=0}^{t}(x^{2^t}+1)^{2^j} = x^{2^t} \prod_{j=0}^{t}(x^{2^t}+1)^{2^j} = x^{2^t} \sum_{j=0}^{2^{t+1}-1}(x^{2^t})^j.$$

The latter function is EA-equivalent to

$$\sum_{j=1}^{2^{t+1}} x^j = \frac{(x^{2^{t+1}+1}+1)}{x+1} + 1.$$

Using the transformation $x \mapsto x + 1$ (and adding 1), we get the function

$$\frac{(x^{2^{t+1}+1} + x^{2^{t+1}} + x)}{x} = x^{2^{t+1}} + x^{2^{t+1}-1} + 1,$$

which is EA-equivalent to G_t^{-1}.

As for the inverse case, the function from (12) can be written as $1/(x+1) + x + 1$. Indeed, $I \circ L \circ I = (\frac{1}{x^2} + \frac{1}{x})^{-1} = (\frac{1+x}{x^2})^{-1} = \frac{x^2}{1+x} = \frac{1}{x+1} + x + 1$. □

B. The case of even dimension

Our experimental results indicate that the case for even values of n is somewhat less interesting. For $n = 6$, no APN functions can be obtained as $P_i \circ L \circ P_j$ for L with coefficients in $\mathbb{F}_2$, while for $n \in \{4, 8\}$, only APN functions from the EA-equivalence class of P_i can be obtained in this manner, as described in the following proposition.

Proposition 18. *Let $n = 2m$, $l_n = \frac{2^{n-1}+1}{3}$, $L(x) = \sum_{j=1}^{t} x^{2^{2i_j}}$ be a permutation for some positive integer t and for some non-negative integers i_j for $1 \le j \le t$, and let $1 \le i \le 2^n - 2$ be arbitrary with $3 \mid i$. Then*

$$P_i \circ L \circ P_{l_n}(x) = P_i \circ M, \tag{14}$$

and

$$P_i \circ L \circ P_{2l_n+1} = P_{2i} \circ M' \circ x^2,$$

where $M(x) = \sum_{j=1}^{t} x^{2^{2i_j-1}}$ and $M'(x) = \sum_{j=1}^{t} x^{2^{2i_j+1}}$. In particular, both $P_i \circ L \circ P_{l_n}$ and $P_i \circ L \circ P_{2l_n+1}$ are linear equivalent to P_i.

Proof. Let us denote $y = x^{l_n}$. We will prove that

$$L(y)^3 = \left(\sum_{j=1}^{t} y^{2^{2i_j}}\right)^3 = \left(\sum_{j=1}^{t} x^{2^{i_j-1}}\right)^3 = M(x)^3;$$

this then implies the case for general i due to $3 \mid i$.

In the following, we use the fact that

$$\frac{2^n+2}{3} 3j \equiv 3j \mod (2^n - 1)$$

for any integer j, and, in particular

$$\frac{2^n+2}{3}(2^{2i_j}-1) \equiv 2^{n-k}+1 \mod (2^n-1) \tag{15}$$

for any integer i_j.

Clearly, $(x^{(2^n-1)/3}f(x))^3 = f(x)^3$ for any polynomial $f(x)$ over $\mathbb{F}_{2^n}$ with $f(0)=0$. We apply this to $L(y)^3 = L(x^{l_n})^3$. The exponent of x in $y^{2^{2i_j}} = x^{2^{2i_j}l_n} = x^{2^{2i_j}(2^{n-1}+1)/3}$ becomes

$$\begin{aligned} &2^{2i_j}\frac{2^{n-1}+1}{3} + \frac{2^n-1}{3} = \frac{2^{n+2i_j-1}+2^{2i_j}+2^n-1}{3} \\ &= \frac{2^n+2}{3}(2^{2i_j-1}+1) - 1 \equiv 2^{2i_j-1} \mod (2^n-1) \end{aligned}$$

for any non-negative integer i_j. Thus, $L(y)^3 = M(x)^3$ as claimed.

The case for $2l_n+1$ follows in the same way, but we multiply the expression by $(x^{(2^n-1)/3})^2$. Denoting $z = x^{2l_n+1}$, the exponent of x in $z^{2^{2i_j}}$ becomes

$$\begin{aligned} &2^{2i_j}\left(\frac{2^n+2}{3}+1\right) + \frac{2^{n+1}-2}{3} \\ &= (2^{2i_j}-1)\left(\frac{2^n+2}{3}\right) + 2^{2i_j} + \frac{2^n+2}{3} + \frac{2^{n+1}-2}{3} \\ &= \frac{2^n+2}{3} + 2^{2i_j} - 1 + 2^{2i_j} + \frac{2^{n+1}-2}{3} \\ &= 2^n - 1 + 2^{2i_j+1} \equiv 2^{2i_j+1} \mod (2^n-1). \end{aligned}$$

The rest follows in the same way as in the previous case. □

We then immediately have the following generalization.

Corollary 19. *Let $n=2m$ be even, $l_n = \frac{2^{n-1}+1}{3}$, $L(x) = \sum_{j=1}^{t} x^{2^{2i_j}}$ be a permutation for some positive integer t and for non-negative integers i_j for $1 \le j \le t$, and let $F(x) = G(x^3)$ for some (n,n)-function G. Then*

$$F \circ L \circ P_{l_n}(x) = F \circ M,$$

$$F \circ L_j \circ P_{2l_n+1}(x) = F \circ P_2 \circ L,$$

where $M(x) = \sum_{j=1}^{t} x^{2^{n-k_j-1}} + x^{2^{n-1}}$. In particular, $F \circ L \circ P_{l_n}$, and $F \circ L \circ P_{2l_n+1}$ are linear equivalent to F.

We note that all APN functions that we computationally obtain as $P_i \circ L \circ P_j$ for L linear with coefficients in $\mathbb{F}_2$ over $\mathbb{F}_{2^n}$ with $n \in \{4,6,8\}$ are described by Proposition 18.

C. Experimental results

For $\mathbb{F}_{2^n}$ with $4 \le n \le 9$, we consider the function $F = P_i \circ L \circ P_j$ for all possible linear L over $\mathbb{F}_{2^n}$ with coefficients in $\mathbb{F}_2$ and for a single i and j from each cyclotomic coset, and record the instances in which F is APN. We confirm that all such cases correspond to one of the cases treated in Sections III-A and III-B.

IV. On power functions over $\mathbb{F}_{2^{nk}}$ with exponents of the form $\sum_{i=1}^{k-1} 2^{in} - 1$

In this section, we summarize some experimental results on power functions of the form x^d with $d = \sum_{i=1}^{k-1} 2^{in} - 1$ over $\mathbb{F}_{2^{nk}}$ for some n and $k \ge 2$. We note that both the Dobbertin and the inverse power function can be expressed in this form. Indeed, for $n=1$, we have $d = \sum_{i=1}^{k-1} 2^i - 1 = 2^k - 2$ over $\mathbb{F}_{2^k}$, which is precisely the inverse; and for $k=5$, we get $d = \sum_{i=1}^{4} 2^i - 1 = 2^{4i} + 2^{3i} + 2^{2i} + 2^i - 1$, which is precisely the Dobbertin exponent over $\mathbb{F}_{2^{5n}}$. In this way, exponents of this form can be seen as generalizations of both the inverse and the Dobbertin families, and investigating them is one potential direction of approaching the problem of the existence (or non-existence) of APN power functions cyclotomic inequivalent to the six known infinite families from Table I.

Power functions of this form have already been studied, and it has been shown in [5, Proposition 12] that this type of functions cannot be APN over $\mathbb{F}_{2^{kn}}$ whenever $k = 2^l + 2$ for some positive integer l, or when $k=2$ and $n > 2$. In [5, Corollary 6], it has also been proved that these power functions cannot be AB. The case of $k=2$ has also been characterized in [3, Theorem 1], where the authors study power functions of the form x^{2^t-1} over a field $\mathbb{F}_{2^m}$, for some $m > 2$.

We first recall the following theorem, which we can use for studying the behavior of the exponent $d = \sum_{i=1}^{k-1} 2^{in} - 1$ for $k=2$.

TABLE II
DIFFERENTIAL SPECTRA OF x^d FOR $d = \sum_{i=1}^{k-1} 2^{ni}$ OVER $\mathbb{F}_{2^{nk}}$

$n \setminus k$	2	3	4
2	2^8	4^{16}	$2^{96}, 12^4, 16^1$
3	$2^{27}, 4^1, 6^1$	8^{64}	$2^{1792}, 56^8, 64^1$
4	$2^{121}, 14^1$	16^{256}	$2^{30720}, 240^{16}, 256^1$
5	$2^{495}, 4^1, 30^1$	32^{1024}	$2^{507904}, 992^{32}, 1024^1$
6	$2^{2017}, 62^1$	64^{4096}	$2^{8257536}, 4032^{64}, 4096^1$
7	$2^{8127}, 4^1, 126^1$	128^{16384}	$2^{133169152}, 16256^{18}, 16384^1$
8	$2^{32641}, 254^1$	256^{65536}	
9	$2^{130815}, 4^1, 510^1$	512^{262144}	
10	$2^{523777}, 1022^1$	$1024^{133169152}$	

$n \setminus k$	6
2	$2^{1434}, 4^{262}, 6^{30}$
3	$2^{82812}, 4^{19422}, 6^{2558}, 8^{378}, 10^{46}$
4	$2^{5147136}, 4^{1261920}, 6^{200688}, 8^{24576}, 10^{2784}, 12^{168}, 14^{24}, 16^{136}, 18^{120}$

$n \setminus k$	7
2	$2^{4753}, 4^{1079}, 6^{329}, 8^{56}, 10^{14}$
3	$2^{655526}, 4^{155316}, 6^{24101}, 8^{2261}, 10^{189}, 12^{21}$
4	$2^{81089988}, 4^{20333107}, 6^{3464027}, 8^{447412}, 10^{49287}, 12^{4970}, 14^{413}, 16^{56}, 18^{14}, 22^7$

$n \setminus k$	8
2	$2^{21920}, 4^{4632}, 6^{464}, 8^{48}$
3	$2^{5204996}, 4^{1259380}, 6^{190068}, 8^{21736}, 10^{1512}, 12^{24}$

$n \setminus k$	9
2	$2^{87087}, 4^{19135}, 6^{1773}, 8^{99}$
3	$2^{41588551}, 4^{10087560}, 6^{1535275}, 8^{165807}, 10^{14121}, 12^{891}, 14^{27}$

$n \setminus k$	10
2	$2^{289405}, 4^{70910}, 6^{22935}, 8^{3860}, 10^{1350}, 12^{240}, 14^{45}, 16^{11}, 18^{25}$
3	$2^{332825256}, 4^{80618460}, 6^{12297172}, 8^{1331025}, 10^{109576}, 12^{6910}, 14^{540}$

Theorem 20 ([3])**.** *Let $n > 1$. Then the power function x^{2^n-1} is 2^n-2-uniform over $\mathbb{F}_{2^{2n}}$. Let $\delta(b)$ be the number of solutions of $x^{2^n-1} + (x+1)^{2^n-1} = b$, we have $\delta(0) = 2^n - 2$,*

$$\delta(1) = \begin{cases} 2 & \text{if } n \text{ is even} \\ 4 & \text{if } n \text{ is odd} \end{cases}$$

and $\delta(b) \in \{0, 2\}$ if $b \notin \mathbb{F}_2$.

Now, if $k = 3$, we can see that the function $x^{2^{2n}+2^n-1}$ over $\mathbb{F}_{2^{3n}}$ is equivalent to $x^{2^{2n}-2^n+1}$, which is a Kasami function. The differential uniformity of $x^{2^{2k}-2^k+1}$ over $\mathbb{F}_{2^n}$ in the case of $s = \gcd(k, n) \neq 1$ is determined in [21], where it is shown that all its derivatives are 2^s-to-1 mappings. Applying this, we then immediately obtain the following result.

Proposition 21. *Let $n > 1$. Then the power function $x^{2^{2n}+2^n-1}$ is 2^n-uniform over $\mathbb{F}_{2^{3n}}$; moreover, all of its derivatives are 2^n-to-1.*

In the following Table II we report our computational results on the differential spectrum of the monomial x^d with $d = \sum_{i=1}^{k-1} 2^{in} - 1$ for values of k and n with $kn \leq 30$. For each combination of n and k, we list the number of solutions x to $x^d + (x+1)^d = b$ for all $b \in \mathbb{F}_{2^{nk}}$; more precisely, we write i^j if there are j elements $b \in \mathbb{F}_{2^{nk}}$ having i solutions x to the aforementioned equation. The number of elements b having no solutions at all is omitted for the sake of brevity.

In addition to the cases of $n = 1$ and $= 2, 3, 5$, the case of $k = 4$ is also noteworthy. Based on the computational data, we formulate the following conjecture.

Conjecture 22. Let $d = 2^{3n} + 2^{2n} + 2^n - 1$ and consider the power function x^d over $\mathbb{F}_{2^{4n}}$. Then the equation $x^d + (x+1)^d = b$ has 2^{2n} solutions for one value of b; it has $2^{2n} - 2^n$ solutions for 2^n values of b; and has at most 2 solutions for all remaining points b.

V. A CONJECTURE ABOUT THE WALSH SPECTRUM OF THE DOBBERTIN FUNCTION

As remarked previously, the problem of computing the Walsh spectrum of the Dobbertin APN functions has been open for quite some time, with virtually no progress since 2000, when it was shown that all values of its Walsh transform are divisible by 2^{2m}, where $n = 5m$ [10].

For n odd, the Gold, Niho, Welch and Niho APN functions from Table 1 are also AB (for the proofs of the AB property, see [9], [10], [20], [22], [26], [27]). In the case of n even, the Gold and the Niho functions have the same Walsh spectrum: $\{0, \pm 2^{\frac{n}{2}}, \pm 2^{\frac{n+2}{2}}\}$. The Walsh transform of the inverse function takes any value divisible by 4 in the interval $[1-2^{\frac{n}{2}+1}, \ldots, 1+2^{\frac{n}{2}+1}]$ [24].

In order get some insight about the form of the Walsh spectrum of the Dobbertin function, we experimentally computed the Walsh coefficients of its instances over the fields $\mathbb{F}_{2^{5m}}$ for $m \leq 7$. Below, we present our computational data in two tables: for n odd, and for n even.

Based on Tables 2 and 3, we make the following conjecture.

Conjecture 23. The Walsh spectrum of the Dobbertin function x^d, where $d = 2^{4m} + 2^{3m} + 2^{2m} + 2^m - 1$ over $\mathbb{F}_{2^{5m}}$ has the following possible forms depending on the parity of m:

- $\{0, 2^{2m}(2^m+1), \pm 2^{5k-2}, \pm a \cdot 2^{2m} \mid 1 \leq a \leq k \cdot (k+1), a \text{ odd}\}$ for $m = 2k-1, k \in \mathbb{N}$;
- $\{0, -2^{2m}(2^m+1), \pm 2^{5k}, \pm 2^{5k+1}, \pm a \cdot 2^{2m} \mid 1 \leq a \leq k \cdot (k+2), a \text{ odd}\}$ for $m = 2k, k \in \mathbb{N}$.

Moreover, $W_F(u,v)$ takes the maximum absolute value $2^{2m}(2^m+1)$ for $u = v = 1$: for even m, we have $\min W_F(u,v) = -2^{2m}(2^m+1)$, and for odd m, we have $\max W_F(u,v) = 2^{2m}(2^m+1)$. Hence, the nonlinearity of the Dobbertin function is

$$2^{5m-1} - 2^{2m-1}(2^m+1).$$

TABLE III
WALSH COEFFICIENTS OF THE DOBBERTIN FUNCTION OVER $\mathbb{F}_{2^{5m}}$ WITH $m = 2k-1, 1 \leq k \leq 4$

$n=5, m=1, k=1$	$n=15, m=3, k=2$	$n=25, m=5, k=3$	$n=35, m=7, k=4$
0	0	0	0
$12 = 2^2(2^1+1)$	$576 = 2^6(2^3+1)$	$33792 = 2^{10}(2^5+1)$	$2113536 = 2^{14}(2^7+1)$
$\pm 8 = \pm 2^3$	$\pm 64 = \pm 2^6$	$\pm 1024 = \pm 2^{10}$	$\pm 16384 = \pm 2^{14}$
$\pm 4 = \pm 2^2$	$\pm 256 = \pm 2^8$	$\pm 8192 = \pm 2^{13}$	$\pm 262144 = \pm 2^{18}$
	$\pm 192 = \pm 3 \cdot 2^6$	$\pm 3072 = \pm 3 \cdot 2^{10}$	$\pm 49152 = \pm 3 \cdot 2^{14}$
	$\pm 320 = \pm 5 \cdot 2^6$	$\pm 5120 = \pm 5 \cdot 2^{10}$	$\pm 81920 = \pm 5 \cdot 2^{14}$
		$\pm 7168 = \pm 7 \cdot 2^{10}$	$\pm 114688 = \pm 7 \cdot 2^{14}$
		$\pm 9216 = \pm 9 \cdot 2^{10}$	$\pm 147456 = \pm 9 \cdot 2^{14}$
		$\pm 11264 = \pm 11 \cdot 2^{10}$	$\pm 180224 = \pm 11 \cdot 2^{14}$
			$\pm 212992 = \pm 13 \cdot 2^{14}$
			$\pm 245760 = \pm 15 \cdot 2^{14}$
			$\pm 278528 = \pm 17 \cdot 2^{14}$
			$\pm 311296 = \pm 19 \cdot 2^{14}$

TABLE IV
WALSH COEFFICIENTS OF THE DOBBERTIN FUNCTION OVER $\mathbb{F}_{2^{5m}}$ WITH $m = 2k, 1 \leq k \leq 3$

$n=10, m=2, k=1$	$n=20, m=4, k=2$	$n=30, m=6, k=3$
0	0	0
$-80 = -2^4(2^2+1)$	$-4352 = -2^8(2^4+1)$	$-266240 = -2^{12}(2^6+1)$
$\pm 16 = \pm 2^4$	$\pm 256 = \pm 2^8$	$\pm 4096 = \pm 2^{12}$
$\pm 32 = \pm 2^5$	$\pm 1024 = \pm 2^{10}$	$\pm 32768 = 2^{15}$
$\pm 64 = \pm 2^6$	$\pm 2048 = \pm 2^{11}$	$\pm 65536 = \pm 2^{16}$
$\pm 48 = \pm 3 \cdot 2^4$	$\pm 768 = \pm 3 \cdot 2^8$	$\pm 12288 = \pm 3 \cdot 2^{12}$
	$\pm 1280 = \pm 5 \cdot 2^8$	$\pm 20480 = \pm 5 \cdot 2^{12}$
	$\pm 1792 = \pm 7 \cdot 2^8$	$\pm 28672 = \pm 7 \cdot 2^{12}$
		$\pm 36864 = \pm 9 \cdot 2^{12}$
		$\pm 45056 = \pm 11 \cdot 2^{12}$
		$\pm 53248 = \pm 13 \cdot 2^{12}$
		$\pm 61440 = \pm 15 \cdot 2^{12}$

Conclusion

We have investigated two different approaches for obtaining alternative representations for functions from the known infinite APN families, and have justified the optimality of our representation in some cases. We have described APN functions that have the form $x^i \circ L \circ x^j$ for L linear with coefficients in $\mathbb{F}_2$, and have computationally verified that our constructions exhaust all possible cases over $\mathbb{F}_{2^n}$ with $4 \leq n \leq 9$. We have also documented our experimental data for the Walsh spectrum of the Dobbertin power functions and, based on it, we have formulated a conjecture describing the exact form of this Walsh spectrum. In addition, we have reported experimental data on the differential spectrum of functions x^d with $d = \sum_{i=1}^{k-1} 2^{ni} - 1$ over $\mathbb{F}_{2^{nk}}$.

Acknowledgements

The research of this paper was supported by the Trond Mohn Foundation (TMS).

REFERENCES

[1] Thomas Beth and Cunsheng Ding. "On almost perfect nonlinear permutations." Workshop on the Theory and Application of of Cryptographic Techniques. Springer, Berlin, Heidelberg, 1993.
[2] Eli Biham and Adi Shamir. "Differential cryptanalysis of DES-like cryptosystems." Journal of CRYPTOLOGY 4.1 (1991): 3-72.
[3] Céline Blondeau, Anne Canteaut, and Pascale Charpin. "Differential Properties of $x \mapsto x^{2^t-1}$." IEEE Transactions on Information Theory 57.12 (2011): 8127-8137.
[4] Marcus Brinkmann and Gregor Leander. "On the classification of APN functions up to dimension five." Designs, Codes and Cryptography 49.1-3 (2008): 273-288.
[5] Lilya Budaghyan. "The equivalence of Almost Bent and Almost Perfect nonlinear functions and their generalization." PhD Dissertation, Otto-von-Guericke-University, Magdeburg, Germany, 2005.
[6] Lilya Budaghyan, Marco Calderini, Claude Carlet, Diana Davidova and Nikolay Kaleyski. "A note on the Walsh spectrum of Dobbertin APN functions." Sequences and Their Applications 2020, Russia, Saint-Petersburg.
[7] Lilya Budaghyan, Marco Calderini, Claude Carlet, Diana Davidova and Nikolay Kaleyski. "On a Relationship between Gold and Kasami Functions and other Power APN Functions." Sequences and Their Applications 2020, Russia, Saint-Petersburg.
[8] Lilya Budaghyan, Claude Carlet, and Alexander Pott. "New classes of almost bent and almost perfect nonlinear polynomials." IEEE Transactions on Information Theory 52.3 (2006): 1141-1152.
[9] Anne Canteaut, Pascale Charpin, and Hans Dobbertin. "Binary m-sequences with three-valued crosscorrelation: a proof of Welch's conjecture." IEEE Transactions on Information Theory 46.1 (2000): 4-8.
[10] Anne, Canteaut, Pascale Charpin, and Hans Dobbertin. "Weight divisibility of cyclic codes, highly nonlinear functions on F2m, and crosscorrelation of maximum-length sequences." SIAM Journal on Discrete Mathematics 13.1 (2000): 105-138.
[11] Anne Canteaut and Marion Videau. "Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis." International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Berlin, Heidelberg, 2002.
[12] Claude Carlet, Pascale Charpin, and Victor Zinoviev. "Codes, bent functions and permutations suitable for DES-like cryptosystems." Designs, Codes and Cryptography 15.2 (1998): 125-156.
[13] Florent Chabaud and Serge Vaudenay. "Links between differential and linear cryptanalysis." Workshop on the Theory and Application of of Cryptographic Techniques. Springer, Berlin, Heidelberg, 1994.
[14] Joan Daemen and Vincent Rijmen. "The design of Rijndael: AES-the advanced encryption standard." Springer Science & Business Media, 2013.
[15] Hans Dobbertin. "Almost perfect nonlinear power functions on $GF(2^n)$) the Welch case." IEEE Transactions on Information Theory 45.4 (1999): 1271-1275.
[16] Hans Dobbertin. "Almost perfect nonlinear power functions on GF (2n): the Niho case." Information and Computation 151.1-2 (1999): 57-72.
[17] Hans Dobbertin. "Almost perfect nonlinear power functions on GF (2 n): a new case for n divisible by 5." Finite Fields and Applications. Springer, Berlin, Heidelberg, 2001. 113-121.
[18] Hans Dobbertin. "Another Proof of Kasami's Theorem". Designs, Codes and Cryptography 17, 177–180 (1999).
[19] Yves Edel and Alexander Pott. "A new almost perfect nonlinear function which is not quadratic." Adv. in Math. of Comm. 3.1 (2009): 59-81.
[20] Robert Gold. "Maximal recursive sequences with 3-valued recursive cross-correlation functions (Corresp.)." IEEE transactions on Information Theory 14.1 (1968): 154-156.
[21] Doreen Hertel and Alexander Pott. "Two results on maximum nonlinear functions." Designs, Codes and Cryptography 47.1-3 (2008): 225-235.
[22] Henk DL Hollmann and Qing Xiang. "A proof of the Welch and Niho conjectures on cross-correlations of binary m-sequences." Finite Fields and Their Applications 7.2 (2001): 253-286.
[23] Heeralal Janwa and Richard M. Wilson. "Hyperplane sections of Fermat varieties in P 3 in char. 2 and some applications to cyclic codes." International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes. Springer, Berlin, Heidelberg, 1993.
[24] Gilles Lachaud, and Jacques Wolfmann. "The weights of the orthogonals of the extended quadratic binary Goppa codes." IEEE transactions on information theory 36.3 (1990): 686-692.
[25] Mitsuru Matsui. "Linear cryptanalysis methods for DES cipher", Advances in Cryptology, Eurocrypt'93, Lecture Notes in Comput.Sci., vol. 65, 1993, pp. 386–397.
[26] Kaisa Nyberg. "Differentially uniform mappings for cryptography." Workshop on the Theory and Application of of Cryptographic Techniques. Springer, Berlin, Heidelberg, 1993.
[27] Tadao Kasami. "The weight enumerators for several classes of subcodes of the 2nd order binary Reed-Muller codes." Information and Control 18.4 (1971): 369-394.
[28] Gohar M. Kyureghyan and Valentin Suder. "On inversion in Z_{2n-1}." Finite Fields and Their Applications 25 (2014): 234-254.
[29] Lukas Kölsch. "On the inverses of Kasami and Bracken–Leander exponents". Designs, Codes and Cryptography, 2020.