Classification of all DO planar polynomials with prime field coefficients over $\mathbb{F}_{3^n}$ for $n \leq 7$*

Diana Davidova and Nikolay S. Kaleyski

Abstract

We describe how any function over a finite field $\mathbb{F}_{p^n}$ can be represented in terms of the values of its derivatives. In particular, we observe that a function of algebraic degree d can be represented uniquely through the values of its derivatives of order $(d-1)$ up to the addition of terms of algebraic degree strictly less than d. We identify a set of elements of the finite field, which we call the degree d extension of the basis, which has the property that for any choice of values for the elements in this set, there exists a function of algebraic degree d whose values match the given ones. We discuss how to reconstruct a function from the values of its derivatives, and discuss the complexity involved in transitioning between the truth table of the function and its derivative representation.

We then specialize to the case of quadratic functions, and show how to directly convert between the univariate and derivative representation without going through the truth table. We thus generalize the matrix representation of qaudratic vectorial Boolean functions due to Yu et al. to the case of arbitrary characteristic. We also show how to characterize quadratic planar functions using the derivative representation. Based on this, we adapt the method of Yu et al. for searching for quadratic APN functions with prime field coefficients to the case of planar DO functions. We use this method to find all such functions (up to CCZ-equivalence) over $\mathbb{F}_{3^n}$ for $n \leq 7$. We conclude that the currently known planar DO polynomials cover all possible cases for $n \leq 7$. We find representatives simpler than the known ones for the Zhou-Pott, Dickson, and Lunardon-Marino-Polverino-Trombetti-Bierbrauer families for $n = 6$. Finally, we discuss the computational resources that would be needed to push this search to higher dimensions.

1 Introduction

Functions over finite fields, typically referred to as discrete functions or vectorial functions, play an important in many areas of mathematics, computer science,

*Parts of this work were previously presented as a conference paper at WCC 2022. The contents in Sections 3, 4, and 5 is completely new. The proof of Proposition 3 was omitted from the conference version due to space constraints. Theorem 1 and Proposition 4 are generalizations of Propositions 1 and 4 from the conference version, respectively, whose proof was also omitted due to space constraints.

and engineering. For instance, the non-linear components of modern cryptographic block ciphers are typically modeled as functions between two finite fields of characteristic 2; since linearity by itself does not provide any cryptographic strength, the security of the entire encryption hinges on the properties of these functions (usually called "substitution boxes", or "S-boxes", in this context). Boolean functions, i.e. functions from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ for some natural number n, have been widely studied and applied in areas as varied as cryptography, artificial intelligence and combinatorics. In the case of odd characteristic, perhaps one of the most remarkable applications of discrete functions is the correspondence between quadratic planar functions and commutative semifields explored in [7], which has led to recent breakthroughs and advances in an area that has historically been studied since the early twentieth century. In general, discrete functions can be used to encode almost any kind of data, and this has led to their widespread use in many areas and contexts.

Discrete functions can be represented in many different ways. Arguably one of the most natural such representations is the truth table, or look-up table representation, which is simply a list of the values $F(x) \in \mathbb{F}_{p^m}$ of a function $F : \mathbb{F}_{p^n} \to \mathbb{F}_{p^m}$ on all possible inputs $x \in \mathbb{F}_{p^n}$. This representation has the advantage of being simple to implement and fast to use in practice: indeed, evaluating F at some given $x \in \mathbb{F}_{p^n}$ amounts to nothing more than indexing an array in the computer's memory containing the truth table, and is certainly the fastest way of working with a discrete function in practice. This natural simplicity is, however, overshadowed by multiple drawbacks, such as the high memory consumption (since every single value of F needs to be explicitly stored in memory), and the fact that the values in the truth table reveal very little about the structure and properties of the function. For instance, it is very hard to say anything about the cryptographic properties, the algebraic degree, the polynomial form, etc. of a function from its truth table without doing extensive computations.

For this reason, alternative representations of discrete functions are studied and used in practice. There is no overall "best" representation, and different representations can be optimal with respect to different goals or parameters. For instance, many cryptographically optimal functions have a very compact representation as a univariate polynomial over $\mathbb{F}_{p^n}$, while the algebraic normal form (ANF) makes it very easy to compute the algebraic degree of a function (or to construct functions with a prescribed algebraic degree). Another motivation for exploring different representations of discrete functions is related to computational searches: finding instances of cryptographically optimal functions such as planar functions and APN functions is very hard, and due to the exponentially growing search space, it is not possible to examine all possible functions. For this reason, computational searches for e.g. APN functions have to be restricted to some small set of functions, such as those having a simple form under some given representation; then it is possible that the functions having a simple form under one representation, e.g. the ANF, will be completely different from the functions having a simple representation under a different representation, such as a univariate polynomial. Another consideration is that the representation

might make it easier (or harder) to determine whether a function possesses certain properties (such as being APN), or to decide whether a pair of functions is equivalent or inequivalent under a given notion of equivalence (we note that e.g. APN functions are classified under relations such as CCZ-equivalences, and so computational searches are only done up to CCZ-equivalence).

We recall that the first-order derivatives of a function F are functions of the form $D_aF(x) = F(a+x) - F(x)$ that express the difference between a pair of values of F whose pre-images have a prescribed difference; and higher-order derivatives are basically "derivatives of derivatives" and express the relation between more than two values of F. Many important cryptographic properties such as the differential uniformity are defined in terms of derivatives, and so the derivatives of a function are significant and natural objects to study.

In this paper, we show how any discrete function can be represented in terms of the values of its derivatives. More precisely, we show that if F is a discrete function of degree d, then knowing its derivatives of order $(d-1)$ allows us to reconstruct the truth table of F uniquely up to the addition of a function of algebraic degree strictly less than d. Since this "lower-degree" part of F can be itself described in terms of derivatives of degree $(d-2)$, $(d-3)$, etc., knowing the values of all derivatives of degree up to d allows F to be reconstructed uniquely. In this way, the set of values of the derivatives of F constitutes an alternative representation of F. While the memory needed for storing this representation can be the same as that of the truth table in the worst case, memory can be saved in comparison if the function is of low algebraic degree or if we ignore some of the lower-degree terms of the function (for example, if we ignore the quadratic and affine terms of a cubic function).

More importantly, our representation allows us to analyze the behavior of the derivatives of the function; this is particularly useful in the case of quadratic functions, where the representation consists of values of the first order derivatives, and these are used in the definition of cryptographically optimal classes of functions such as planar functions and APN functions. Indeed, specializations of this representation to quadratic functions in characteristic 2 have previously been used in works such as [18] and [17] to search for APN functions, although the connection of this representation with derivatives does not seem to have been made in those papers. This has made important computational results possible: the approach in [18] was used to find thousands of new examples of CCZ-inequvalent APN functions over $\mathbb{F}_{2^8}$ (whereas only several such functions were known prior to that work), while the computations in [17] resulted in a complete classification of all quadratic APN functions with coefficients in $\mathbb{F}_2$ over F_{2^n} for $n \leq 9$ (we remark that such classifications are very difficult to obtain, even for specific subclasses of functions; APN functions have been completely classified over $\mathbb{F}_{2^n}$ only for $n \leq 5$ [2]; for cubic APN functions up to $n \leq 6$ [13]; and for quadratic APN functions up to $n \leq 7$ [13], [12]). Since this approach was previously developed only for functions over finite fields of characteristic 2, it was not possible to perform similar computations for e.g. planar functions or other important classes of functions over finite fields of odd characteristic. In the present paper, we show how quadratic functions of any differential uniformity

over a field of any characteristic can be characterized in terms of the derivative representation; we then apply this characterization to the case of characteristic 3, and computationally classify all quadratic planar functions with coefficients in $\mathbb{F}_3$ over $\mathbb{F}_{3^n}$ for $n \leq 7$. We confirm that there are no planar functions other than the currently known ones (up to equivalence), but find representatives of some of the known equivalence classes that are significantly simpler than the previously known ones. We also discuss the computational challenges for taking this approach farther, and propose several directions for future work.

We note that the exposition in the paper focuses on functions from $\mathbb{F}_{p^n}$ to $\mathbb{F}_{p^m}$ with $n = m$ for the sake of simplicity and since this is the most practically relevant case (for instance, the cryptographically optimal classes of planar and APN functions cannot be defined otherwise). Nonetheless, all of the principles should be applicable to the general case of $n \neq m$ (with the exception of the conversion to and from the univariate polynomial form, which may not be defined when $n \neq m$).

2 Preliminaries

Let n, m, p be natural numbers, with p prime. We denote by $\mathbb{F}_{p^n}$ the finite field with p^n elements. An (n, m, p)**-function** F is any function F from $\mathbb{F}_{p^n}$ to $\mathbb{F}_{p^m}$; when the values of n, m, p are understood from the context or are not important, we will also refer to such functions as **discrete functions** or **vectorial functions**. When $p = 2$, these are also called **vectorial Boolean functions**; if, in addition, $m = 1$, we talk about **Boolean functions**.

The finite field $\mathbb{F}_{p^n}$ can identified with the vector space $\mathbb{F}_p^n$ of dimension n over $\mathbb{F}_p$; in particular, every element $x \in \mathbb{F}_{p^n}$ can be represented as a coordinate vector in $\mathbb{F}_p^n$. Let $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ be a basis of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$.

We denote by $\mathrm{wt}_{\mathcal{B}}(x)$ the Hamming weight of the coordinate vector of x with respect to $\mathcal{B}$; in other words, if $x = \sum_{i=1}^{n} a_i b_i$ for $a_i \in \mathbb{F}_p$, then $\mathrm{wt}(x) = \sum_{i=1}^{n} a_i$, with the sum taken over the integers. We note that the weight does not depend on the concrete choice of basis, and so we will write wt instead of $\mathrm{wt}_{\mathcal{B}}$.

The **algebraic normal form (ANF)** of an (n, m, p)-function F is the polynomial

$$F(x_1, x_2, \ldots, x_n) = \sum_{u \in \mathbb{F}_{p^n}} a_u \prod_{i=1}^{n} x_i^{u_i},$$

where $x = (x_1, x_2, \ldots, x_n)$ and $u = (u_1, u_2, \ldots, u_n)$ are the coordinate vectors of x and u, respectively. The ANF always exists and is uniquely defined. The **algebraic degree** $\deg(F)$ of F is the largest Hamming weight of an exponent with a non-zero coefficient, i.e.

$$\deg(F) = \max\{\mathrm{wt}(u) : u \in \mathbb{F}_{p^n} \mid a_u \neq 0\}.$$

A function F of algebraic degree at most 1 are called **affine**, and has the property that $F(x) + F(y) + F(z) = F(x + y + z)$ for any $x, y, z \in \mathbb{F}_{p^n}$. If, in addition,

$F(0) = 0$, we say that F is **linear**. Functions of algebraic degree 2 are called **quadratic**, functions of algebraic degree 3 are called **cubic**, and so forth.

If F is an (n, m, p)-function with $m \mid n$, so that $\mathbb{F}_{p^m}$ is a subfield of $\mathbb{F}_{p^n}$, the function F can be represented as a polynomial over $\mathbb{F}_{p^n}$ of the form

$$F(x) = \sum_{i=0}^{p^n-1} a_i x^i.$$

This is called the **univariate representation** of F, and always exists; it can be obtained, for instance, by Lagrange interpolation on the list of values of the function. When, in addition, $n = m$, this representation is unique; in the cases when m is a proper divisor of n, some additional restrictions have to be imposed on the polynomial in order to ensure uniqueness. Since this is not relevant to our work, we do not go into details here.

In fact, in the following, we will assume that we are working with (n, n, p)-functions, i.e. that the domain and co-domain are the same. This is arguably the most widely studied case in practice, and is motivated by the focus on the cryptographically optimal classes of planar and almost perfect nonlinear functions in the latter half of the paper. While the assumption that $n = m$ makes it possible to define a unique univariate polynomial representation, we stress that virtually all of the principles and procedures discussed in the sequal can immediately be applied to the general case of (n, m, p)-functions with $n \neq m$ as well (except statements such as Proposition 3 that concern the univariate representation).

For any (n, n, p)-function F, its (first-order) **derivative** in **direction** $a \in \mathbb{F}_{p^n}$ is the (n, n, p)-function

$$\Delta_a F(x) = F(a + x) - F(x) - F(a) + F(0).$$

We note that the derivative is sometimes defined as $D_a F(x) = F(a+x) - F(x)$, but in the context of studying properties such as differential uniformity, the difference between the two representations is $\Delta_a F(x) - D_a F(x)$ is a constant, and thus the two definitions can be used interchangeably. In our work, we prefer the form $\Delta_a F(x)$ since it is symmetric in a and x; for this reason, we will also denote it by $\Delta_F(a, x)$ or, equivalently, $\Delta_F(x, a)$.

The **differential uniformity** δ_F of an (n, n, p)-function F is the maximum number of solutions x to any equation of the form $F(a + x) - F(x) = b$ for any choice of $a, b \in \mathbb{F}_{p^n}$ with $a \neq 0$. Clearly, this is the same as the maximum number of solutions $x \in \mathbb{F}_{p^n}$ to $\Delta_F(a, x) = b$ for any $a, b \in \mathbb{F}_{p^n}$ with $a \neq 0$. Symbolically:

$$\delta_F = \max\{\#\{x \in \mathbb{F}_{p^n} : \Delta_F(a, x) = b\} : a, b \in \mathbb{F}_{p^n} \mid a \neq 0\}.$$

The differential uniformity is an important cryptographic parameter, and it is desirable for it to be low in order for a function to provide reliable resistance against differential cryptanalysis when employed as a component in a block cipher. If $\delta_F = 1$, we say that F is a **perfect nonlinear (PN)** or **planar**

function. Since $\Delta_F(a,x) = \Delta_F(a+x,x)$ when $p = 2$ (due to addition and subtraction being the same operation), planar functions only exist when the characteristic of the field is odd. In the case of even characteristic, the optimal value of the differential uniformity is 2. If $\delta_F = 2$, we say that F is **almost perfect nonlinear (APN)**.

Planar and APN functions are important objects of study due to their cryptographic significance, but they are also of interest thanks to various correspondences with optimal objects in other areas of mathematics. For instance, the correspondence between quadratic planar functions over fields of odd characteristic and commutative semifields studied in [7] is an important breakthrough, which has led to the construction of multiple infinite families of semifields which were not previously known despite the study of semifields since their introduction at the beginning of the twentieth century. We refer the reader to [6] for more background on APN functions, and to [14] for a survey on constructions and properties of APN and planar functions.

Finding new examples of APN and planar functions is thus clearly a matter of significant theoretical and practical significance. A potential issue is that the number of such functions is very large; in fact, the number of functions even over a finite field of relatively small order such as $\mathbb{F}_{2^4}$ is so large that even just recording all of these functions becomes a significant challenge. For this reason, the classes of both APN and planar functions are only considered up to an appropriate equivalence relations; that is, some notion of equivalence is introduced (which must necessarily preserve the property of the functions being APN or planar, respectively, so that such a classification would make sense), and then any two functions belonging to the same equivalence class are considered to be "the same"; in particular, only a single representative from each equivalence class needs to be considered.

The most general known relation on (n,n,p)-functions used in practice that preserves the differential uniformity is called CCZ-equivalence. We say that $F, G : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ are **CCZ-equivalent** if there exists an affine permutation A of $\mathbb{F}_{p^{2n}}$ mapping the graph $\{(x, F(x)) : x \in \mathbb{F}_{p^n}\}$ of F to the graph $\{(x, G(x)) : x \in \mathbb{F}_{p^n}\}$ of G (we note that $\mathbb{F}_{p^n} \times \mathbb{F}_{p^n}$ can be identified with $\mathbb{F}_{p^{2n}}$).

Less general notions of equivalence are frequently used in the literature as well, since they can be easier to work with or reason about, and since they coincide with CCZ-equivalence for some important classes of functions. We say that F and G are **linear equivalent** if there exist linear permutations L_1, L_2 of $\mathbb{F}_{p^n}$ such that $L_1 \circ F \circ L_2 = G$; similarly, we say that F and G are **affine equivalent** if there exist affine permutations A_1, A_2 of $\mathbb{F}_{p^n}$ such that $A_1 \circ F \circ A_2 = G$. Finally, we say that F and G are **extended affine equivalent (EA-equivalent)** if there exist affine permutations A_1, A_2 of $\mathbb{F}_{p^n}$ and an affine (n,n,p)-function A such that $A_1 \circ F \circ A_2 + A = G$. Clearly, linear equivalence is a special case of affine equivalence, which in turn is a special case of EA-equivalence; it can also be shown that EA-equivalence is a special case of CCZ-equivalence, and so these notions of equivalence form a hierarchy of increasing generality.

We know that two quadratic APN functions are CCZ-equivalent if and only

if they are EA-equivalent [16]; that two planar functions are CCZ-equivalent if and only if they are affine equivalent [4]; and that two quadratic planar functions are CCZ-equivalent if and only if they are linear equivalent [4]. Since the vast majority of the currenty known APN and planar functions are quadratic, this means that in practice it is often enough to consider and reason about some of these less general equivalence relations.

This can be advantageous since deciding whether two given functions are CCZ-equivalent is one of the largest bottlenecks in finding new examples of i.a. APN and planar functions. At the time of writing, the only efficient way of deciding CCZ-equivalence is through linear codes. More precisely, a linear code C_F can be associated with any (n, n, p)-function F; and then two functions F and G are CCZ-equivalent if and only if C_F and C_G are isomorphic [9]. Algorithms for testing code isomorphism are known and implementations are available for instance in the *Magma* algebra system [1]. Unfortunately, this approach requires a lot of memory, and is only applicable to relatively small finite fields (for instance, $\mathbb{F}_{2^{11}}$ and $\mathbb{F}_{3^8}$ are already too large for the algorithm to run on our department server).

In the case of quadratic APN functions, EA-equivalence between a pair of functions can be decided by computing their orthoderivatives [5]. We omit the details here, but note that this approach can be used to very quickly confirm that two quadratic APN functions are inequivalent. If the functions are equivalent (so that they cannot be differentiated through the orthoderivatives), it is possible to verify this with algorithms such as the ones described in [5] or [11], although these approaches are significantly slower and do not work for some classes of functions.

The orthoderivative approach, unfortunately, does not work for planar functions. The only way to test CCZ-equivalence in general is through code isomorphism, although this takes significant time. In the case of even planar functions, i.e. planar functions $F(x)$ such that $F(x) = F(-x)$ for all $x \in \mathbb{F}_{P^n}$, a recent algorithm allows equivalence to be decided without going through linear codes, and significantly reduces the computation time and memory requirements [10]. We apply this approach to partition the functions that we find from our computational searches into equivalence classes since all functions produced by our computational search are even.

3 Higher-order derivatives and their properties

In this section, we formally define the notion of higher-order derivatives and justify some of their fundamental properties. While many of these properties may seem intuitively clear, we consider it necessary to provide formal proofs since we are not aware of a similar treatment available in the literature, and since the correctness of the derivative hypermatrix representation discussed in the rest of the paper crucially depends on some of these properties.

For a function F over $\mathbb{F}_{p^n}$, we define its $(k-1)$-st order derivative Δ_F^k as the function $\Delta_F^k : \mathbb{F}_{p^n}^k \to \mathbb{F}_{p^n}$ given by

$$\Delta_F^k(x_1, x_2, \ldots, x_k) = \sum_{I \subseteq \{1,2,\ldots,k\}} (-1)^{k-\#I} F\left(\sum_{i \in I} x_i\right). \tag{1}$$

For example, the first order derivative takes the form

$$\Delta_F^2(x, y) = F(x + y) - F(x) - F(y) + F(0),$$

and the second order derivative is of the form

$$\Delta_F^3(x, y, z) = F(x+y+z) - F(x+y) - F(x+z) - F(y+z) + F(x) + F(y) + F(z) - F(0).$$

We can observe that a k-th order derivative can be expressed as the first order derivative of a $(k-1)$-st order derivative. This is formalized in Proposition 1 below.

In the following proof (as well as further on in the paper), we will identify sets of elements with values of F; for instance, if $\{x_1, x_2, x_3\}$ is a set of elements, then the corresponding value of F will be $F(x_1 + x_2 + x_3)$. This is mostly useful when the elements x_i are multiples of elements from a basis, since any element x can be written as a sum of such elements.

Proposition 1. *For any (n, n, p)-function F, we have*

$$\begin{aligned} \Delta_F^{d+1}(x_1, x_2, \ldots, x_{n-1}, x_n, x_{n+1}) = &\Delta_F^d(x_1, x_2, \ldots, x_{n-1}, x_n + x_{n+1}) - \\ &\Delta_F^d(x_1, x_2, \ldots, x_{n-1}, x_n) - \\ &\Delta_F^d(x_1, x_2, \ldots, x_{n-1}, x_{n+1}). \end{aligned} \tag{2}$$

Proof. From the definition of Δ_F^d in (1), we can see that the value of the derivative of F is the sum of the values of F on all subsets of the inputs to the derivative, with the signs in front of the values of F being determined by the parity of the corresponding subset. Let us denote $A = \Delta_F^d(x_1, x_2, \ldots, x_{n-1}, x_n + x_{n+1})$, $B = \Delta_F^d(x_1, x_2, \ldots, x_{n-1}, x_n)$ and $C = \Delta_F^d(x_1, x_2, \ldots, x_{n-1}, x_{n+1})$. We can see that subsets containing both x_n and x_{n+1} are contained in A, subsets containing only x_n are contained in B, and subsets containing only x_{n+1} are contained in C; furthermore, we can immediately see that these subsets will have the correct signs. Subsets not containing any of x_{n-1}, x_n, x_{n+1} occur in all of A, B, and C, and they occur with the same sign in all three. However, since in (2) we have $A - B - C$ on the right-hand side, the values of F corresponding to such subsets coming from e.g. A and B will cancel out, and only one value will be left; again, it can be easily verified that the sign of this value is the same as the one coming from the derivative on the left-hand side of (2). □

We can also observe that, as one might expect, the k-th order derivative of a function F of algebraic degree d has algebraic degree at most $d - k$. This is most easily seen by considering the ANF of F.

Proposition 2. *Let F be an (n, n, p)-function with $\deg(F) = d$. Let D be the function*

$$D : x \mapsto \Delta_F^{k+1}(a_1, a_2, \ldots, a_k, x)$$

for some natural number k and some $a_1, a_2, \ldots, a_k \in \mathbb{F}_{p^n}$. Then the degree of $D(x)$ is at most $d - k$.

Proof. We prove the statement by induction on k. If $k = 1$, then the first-order derivative of F takes the form

$$\Delta_F^2(a, x) = F(a + x) - F(a) - F(x).$$

Let $x = (x_1, x_2, \ldots, x_n)$ and $a = (a_1, a_2, \ldots, a_n)$ for $x_i, a_i \in \mathbb{F}_p$. Since $\deg(F) = d$, then the ANF of F does not contain any term of degree greater than d. Consider the term $cx_1x_2 \cdots x_d$ of degree d for some $c \in \mathbb{F}_{p^n}$. In the first-order derivative

$$F(a_1 + x_1, a_2 + x_2, \ldots, a_n + x_n) - F(a_1, a_2, \ldots, a_n) - F(x_1, x_2, \ldots, x_n)$$

this becomes $c((a_1 + x_1)(a_2 + x_2) \cdots (a_n + x_n) - a_1a_2 \cdots a_n - x_1x_2 \cdots x_n)$, and the terms of degree d cancel out; the resulting expression is therefore of degree at most $d - 1$. Since this argument applies to any term of degree d, and since F does not contain terms of higher degree, we can conclude that $\deg(D) = 1$ as claimed.

For $k > 1$, the proof follows by induction using the decomposition in Proposition 1; for instance, for $k = 2$, we obtain

$$D(x) = \Delta_F^3(a, b, x) = \Delta_F^2(a, b + x) - \Delta_F^2(a, x) - \Delta_F^2(a, b) = \Delta_g^2(b, x)$$

for $g(x) = \Delta_F^2(a, x)$. We then have $\deg(g) \leq d - 1$ and $\deg(D) \leq (d - 1) - 1 = d - 2$ by the base case $k = 1$. □

Combining the above two propositions, we can see that the derivatives as defined in (1) have the same linearity properties that one would intuitively expect when applied to functions of low enough algebraic degree. More precisely, we see that the $(d - 1)$-st order derivative of a function of algebraic degee d is linear.

Corollary 1. *Let F be an (n, n, p)-function of algebraic degree d. Then $G(x) = \Delta_F^d(a_1, a_2, \ldots, a_{d-1}, x)$ is linear in x for any choice of $a_1, a_2, \ldots, a_{d-1} \in \mathbb{F}_{p^n}$.*

Proof. By Proposition 2, $G(x)$ has algebraic degree at most 1. From the decomposition in Proposition 1, we can see that $G(0) = 0$, and therefore $G(x)$ is linear as claimed. □

From Proposition 1 and Corollary 1, we can make the simple but fundamental observation that for any function F of degree d, Δ_F^k is the zero function for any $k > d$.

Corollary 2. *Let F be an (n, n, p)-function of algebraic degree d, and let $k > d$ be a natural number. Then, for any $x_1, x_2, \ldots, x_k \in \mathbb{F}_{p^n}$, we have $\Delta_F^k(x_1, x_2, \ldots, x_k) = 0$.*

Proof. By Proposition 1, it suffices to prove the statement for $k = d+1$. Again by Proposition 1, we can decompose

$$\Delta_F^{d+1}(x_1, x_2, \ldots, x_d, x_{d+1}) = \Delta_F^d(x_1, x_2, \ldots, x_{d-1}, x_d + x_{d+1}) - \Delta_F^d(x_1, x_2, \ldots, x_{d-1}, x_d) - \Delta_F^d(x_1, x_2, \ldots, x_{d-1}, x_{d+1}).$$

For the sake of brevity, let us denote $x_1, x_2, \ldots, x_{d-1}$ by A. From Corollary 1, $\Delta_F^d(A, x)$ is linear in x, and so

$$\Delta_F^{d+1}(A, x_d, x_{d+1}) = \Delta_F^d(A, x_d + x_{d+1} - x_d - x_{d+1}) = \Delta_F^d(A, 0),$$

which clearly evaluates to 0. This completes the proof. □

4 The derivative hypermatrix

We now introduce the main object of our study, the degree d derivative hypermatrix, which is simply an indexed set containing the values of all degree d derivatives of a function F on the elements of a given linear basis $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$. Knowing the degree d derivative hypermatrix of a degree d function allows the function to be uniquely reconstructed up to the addition of terms of degree strictly less than d. In particular, knowing the degree k derivative hypermatrices of a degree d function F for $k = 1, 2, \ldots, d$ allows d to be uniquely reconstructed, and is therefore a representation of F.

Definition 1. *Let $F : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ for some prime p and some positive integer n. Let $\mathcal{B} = \{b_1, b_2, \cdots, b_n\}$ be a basis of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$. The degree k* ***derivative hypermatrix*** *of F is the hypermatrix $H_F \in \mathbb{F}_{p^n}^{n^{k+1}}$ whose entries are given by*

$$(H_F)_{i_1, i_2, \ldots, i_k} = \Delta_F(b_{i_1}, b_{i_2}, \ldots, b_{i_k}),$$

where $1 \leq i_1, i_2, \ldots, i_k \leq n$.

We will now show that a degree d function F over $\mathbb{F}_{p^n}$ can be uniquely reconstructed from its degree d derivative hypermatrix up to the addition of terms of algebraic degree strictly less than d. In particular, for $d = 2$ we will obtain that we can reconstruct quadratic functions up to the addition of affine terms, i.e. up to EA-equivalence.

We will assume that if the function F being reconstructed is of degree d, then $F(x) = 0$ for all $x \in \mathbb{F}_{p^n}$ with $\mathrm{wt}(x) < d$. This can be assumed without losing generality (up to the addition of terms of degree less than d), since if F does not have this property, i.e. if there exist elements $x \in \mathbb{F}_{p^n}$ with $F(x) \neq 0$ and $\mathrm{wt}(x) < d$, then we can always find a function f of degree $\deg(f) < d$ such that $f(x) = F(x)$ for all $x \in \mathbb{F}_{p^n}$ with $\mathrm{wt}(x) < d$. Then $H_F = H_{F-f}$, and the reconstruction procedure can be applied to $F - f$ instead.

The fact that a function f as described above always exists can, in fact, be seen as a corollary of the fact that functions can be reconstructed from their derivative hypermatrix. For this reason, we will first show how a function F of degree d with $F(x) = 0$ for all x with $\mathrm{wt}(x) < d$ can be reconstructed from H_F; and the general statement that any function F of degree d can be reconstructed (up to the addition of lower-degree terms) will follow as a corollary.

We begin by examining the case for cubic functions for the sake of simplicity. The statement and proof for arbitrary algebraic degree is given in Theorem 1 below.

Suppose F is a cubic function and H_F is its derivative hypermatrix of degree 3. In the following, we will write (x) as shorthand for $F(x)$. Knowing H_F gives us knowledge of $\Delta_F(a, b, c)$ for any $a, b, c \in \mathcal{B}$, where $\mathcal{B}$ is a fixed basis of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$. Let $x \in \mathbb{F}_{p^n}$. If $\mathrm{wt}(x) < 3$, then $F(x) = 0$ by assumption. If $wt(x) = 3$, then $x = a + b + c$ for some $a, b, c \in \mathcal{B}$, and then we have

$$\Delta_F(a, b, c) = (a+b+c)-(a+b)-(a+c)-(b+c)+(a)+(b)+(c)-(0) = (a+b+c)$$

since all elements of Hamming weight less than 3 evaluate to 0; thus, $F(a+b+c) = \Delta_F(a, b, c)$.

Suppose now that $x = a + b + c + d$ for $a, b, c, d \in \mathcal{B}$. Then

$$\Delta_F(a + b, c, d) = (x) - (a + b + c) - (a + b + d).$$

We know that $(a+b+c) = \Delta_F(a, b, c)$ and $(a+b+d) = \Delta_F(a, b, d)$. Substituting this above and re-arranging the terms yields

$$(a + b + c + d) = \Delta_F(a + b, c, d) + \Delta_F(a, b, c) + \Delta_F(a, b, d).$$

Using the linearity of Δ_F (Corollary 1), the above becomes

$$\begin{aligned} F(a + b + c + d) =& \Delta_F(a, c, d) + \Delta_F(b, c, d)+ \\ & \Delta_F(a, b, c) + \Delta_F(a, b, d). \end{aligned}$$

As a final example, if $x = a + b + c + d + e$ for $a, b, c, d, e \in \mathcal{B}$, then

$$\Delta_F(a + b + c, d, e) = (x) - (a + b + c + d) - (a + b + c + e) + (a + b + c).$$

Using the expressions for values of Hamming weight 3 and 4 from above, this becomes

$$\begin{aligned} (x) =& \Delta_F(a + b + c, d, e) + \Delta_F(a, b, c) + \Delta_F(a, c, d) + \Delta_F(b, c, d)+ \\ & \Delta_F(a, b, c) + \Delta_F(a, b, e) + \Delta_F(b, c, e) - \Delta_F(a, b, c). \end{aligned}$$

The term $\Delta_F(a, b, c)$ occurs three times here, once with a negative sign and twice with a positive sign; after cancellation, and another application of Corollary 1, this leaves us with the sum of the values of Δ_F on all three-element subsets of $\{a, b, c, d, e\}$.

The above discussion shows how to derive the value of x if all of the non-zero entries in its coordinate vector are equal to 1. If this is not the case, e.g. if

$x = a + 2b + 2c + d$, then the decomposition of $F(x)$ into a sum of third-order derivatives works in the same exact way, except that instead of taking the sum of derivatives over three-element subsets of $\{a, b, c, d\}$, we take their sum over three-element subsets of $\{a, 2b, 2c, d\}$; in other words, in this case we would get

$$F(a + 2b + 2c + d) = \Delta_F(a, 2b, 2c) + \Delta_F(a, 2b, d) + \Delta_F(a, 2c, d) + \Delta_F(2b, 2c, d).$$

Furthermore, since the order d derivatives of a degree d function are linear, we have e.g. $\Delta_F(2b, 2c, d) = 2 \cdot 2 \cdot \Delta_F(b, c, d)$, and so the values of these derivatives can be immediately recovered from the derivative hypermatrix.

As the above illustration shows, there are three cases to consider depending on the Hamming weight $w = \mathrm{wt}(x)$: if $w < d$, if $w \geq d$ but $d - w < d$, and if $w \geq 2d$. In the following, we formalize the same argument for functions of arbitrary algebraic degree.

Theorem 1. *Let F be an (n, n, p)-function of degree d such that $F(x) = 0$ for all $x \in \mathbb{F}_{p^n}$ with $\mathrm{wt}(x) < d$. Let $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ be a basis of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$. Then, for any $x \in \mathbb{F}_{p^n}$ with $\mathrm{wt}(x) = w$, we have*

$$F(x) = \sum_{I \subseteq \{1,2,\ldots,w\}, \#I = d} \Delta_F(I), \tag{3}$$

where $j_1, j_2, \ldots, j_w$ are indices in $\{1, 2, \ldots, n\}$ such that $x = c_1 b_{j_1} + c_2 b_{j_2} + \cdots c_w b_{j_w}$, and $\Delta_F(I)$ is shorthand for $\Delta_F(c_{k_1} b_{j_{k_1}}, c_{k_2} b_{j_{k_2}}, \ldots, c_{k_m} b_{j_{k_m}})$ with $I = \{k_1, k_2, \ldots, k_m\}$ (note that the value of Δ_F does not depend on the order of its arguemnts since the derivatives are symmetric).

Proof. Without loss of generality, we can assume that the coefficients $c_1, c_2, \ldots, c_w$ are all equal to 1; otherwise we can multiply the elements of $\mathcal{B}$ by the appropriate coefficients, and obtain a different basis of $\mathbb{F}_{p^n}$ under which the coordinate vector of x only contains 0 and 1 as coefficients. The proof proceeds by induction on w.

If $w < d$, then $F(x) = 0$ by assumption, and the sum in (3) is empty, and so the statement is true.

If $w = d$, then $\{j_1, j_2, \ldots, j_d\}$ is the only subset in the sum in (3), and so the latter becomes

$$F(x) = \Delta_F(b_{j_1}, b_{j_2}, \ldots, b_{j_d}).$$

On the other hand, the derivative on the right-hand side equals $\Delta_F(b_{j_1}, \ldots, b_{j_d}) = F(b_{j_1} + b_{j_2} + \cdots + b_{j_d}) = F(x)$ since all the remaining terms in the definition of Δ_F are values of F on elements of Hamming weight less than d, which are all equal to 0 by assumption.

If $w > d$ but $w - d < d$, then the derivative

$$\Delta_F(b_{j_1}, b_{j_2}, \ldots, b_{j_{d-1}}, b_{j_d} + b_{j_{d+1}} + \cdots + b_{j_w})$$

equals the sum of the values of F

$$\sum_{\emptyset \neq S \subseteq \{j_1, j_2, \ldots, j_{d-1}\}} (-1)^{\#S} F\left(b_{j_d} + b_{j_{d+1}} + \cdots + b_{j_w} + \sum_{s \in S} b_s\right). \tag{4}$$

By the induction hypothesis, we can express the values of F corresponding to all subsets S except $S = \{j_1, j_2, \ldots, j_{d-1}\}$ as the sum of the values of Δ_F^d on all d-element subsets of basis elements indexed by $S \cup \{j_d, j_{d+1}, \ldots, j_{d-1}\}$. Let D be a d-element subset of $\{j_1, j_2, \ldots, j_w\}$. Let $L = \{j_1, j_2, \ldots, j_{d-1}\}$ and $R = \{j_d, j_{d+1}, \ldots, j_w\}$, and denote $D_l = D \cap L$ and $D_r = D \cap R$. Note that for all the values $F(x)$ in (4), x contains all the elements of $\mathcal{B}$ indexed by R in its coordinate vector, and so the elements indexed by D will occur in the coordinate vectors corresponding to precisely $2^{\#L - \#D_l}$ values of F in (4). Suppose $\#L > \#D_l$; then there exists some $l \in L \setminus D_l$, and the map $S \mapsto S \oplus \{l\}$, where $\oplus$ denotes symmetric difference, is a bijection between the subsets of L containing D_l but not l, and the subsets of L containing $D_l \cup \{l\}$. The number of sets of both types is thus equal, but one of them is the set corresponding to x (the sum of all possible elements in $L \cup R$); thus, all but one of these values will cancel out. The only remaining set (besides the full $L \cup R$) is therefore $(L \cup R) \setminus \{l\}$; $L \cup R$ contains precisely one element more, and so we can deduce the corresponding values of F appear on both sides of the equation with the same sign.

For those D for which $L = D_l$, the subsets D only occur in the set of indices corresponding to x itself. We thus obtain that the derivative above is equal to

$$\Delta_F = F(x) - \sum_{S \subseteq \{j_1, j_2, \ldots, j_w\}, \#S = d, L \subsetneq S} F\left(\sum_{s \in S} b_j\right).$$

Replacing the values of F with Hamming weight d on the right-hand side of the above equation with the corresponding order d derivatives, and transferring them to the left-hand side, along with using the linearity from Corollary 1, we obtain the sum of the values of F on all elements of Hamming weight d with coordinates among $\{j_1, j_2, \ldots, j_w\}$ on the left-hand side, and $F(x)$ on the right-hand side as desired.

In the case when $w - d > d$, the proof proceeds in the same manner, except that $F(\sum_{s \in R} b_s)$ itself needs to be considered as a value since its Hamming weight is d or more. We skip this part of the proof for the sake of brevity. □

Theorem 1 shows that given a function F of degree d with $F(x) = 0$ on x with $\mathrm{wt}(x) < 0$, the truth table of F can be uniquely reconstructed from H_F. On the other hand, suppose that F' is another function with the same property that $F'(x) = 0$ whenever $\mathrm{wt}(x) < d$. If $F \neq F'$, then there exists some x with $\mathrm{wt}(x) \geq d$ such that $F(x) \neq F'(x)$, and therefore $H_F \neq H_{F'}$. In this sense, the mapping from functions to derivative hypermatrices is injective. Furthermore, the number of derivative hypermatrices of degree d is precisely $(p^n)^{\binom{n}{d}}$, while the number of functions of algebraic degree at most d is precisely

$$(p^n)^{\binom{n}{d} + \binom{n}{d-1} + \cdots + \binom{n}{1} + 1}.$$

By Corollary 2, adding lower-degree terms to F will not change the hypermatrix H_F, and since there are

$$(p^n)^{\binom{n}{d-1} + \binom{n}{d-2} + \cdots + \binom{n}{1} + 1}$$

functions of algebraic degree no greater than $d-1$, we can see that there are $(p^n)^{\binom{n}{d}}$ classes of degree d functions (up to addition of lower-degree terms). Since this is the same as the number of distinct degree d hypermatrices, we can conclude that for any degree d derivative hypermatrix H_F, there exists some function F of algebraic degree d having H_F as its derivative hypermatrix. We thus obtain the following corollary.

Corollary 3. *Let H_F be a degree d derivative hypermatrix over $\mathbb{F}_{p^n}$. Then there exists an (n,n,p)-function F of algebraic degree d having H_F as a degree d derivative hypermatrix.*

Noting that the entries in the degree d derivative hypermatrix are the values of F on all elements $x \in \mathbb{F}_{p^n}$ of weight $wt(x) = d$ with only 0 and 1 in their coordinate vector motivates the following definition.

Definition 2. *Let $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ be a basis of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$ for some natural numbers p, n with p prime. Let d be a natural number with $d \leq (p-1)n$. We call the set*

$$\mathcal{B}_d = \left\{ \sum_{i=1}^{n} a_i b_i : (a_1, a_2, \ldots, a_n) \in \{0,1\}^n \mid \mathrm{wt}(a_1, a_2, \ldots a_n) = d \right\}$$

*of all possible sums of d elements from the basis $\mathcal{B}$ the **degree** d **extension** of $\mathcal{B}$.*

Using this terminology, we can now reformulate Corollary 3 as an "interpolation lemma" as follows.

Corollary 4. *Let $\mathcal{B}_d$ be a degree d extension of a basis $\mathcal{B}$ of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$ for some appropriate d, p, n. Let the elements of $\mathcal{B}_d$ be $e_1, e_2, \ldots, e_m$ for $m = \binom{n}{d}$. Then, for any choice of values $v_1, v_2, \ldots, v_m \in \mathbb{F}_{p^n}$, there exists an (n,n,p)-function F of algebraic degree d such that $F(e_i) = v_i$ for $i = 1, 2, \ldots, m$. Furthermore, this function is uniquely defined up to the addition of terms of algebraic degree strictly less than d.*

We can contrast the above with the classical Lagrange interpolation, in which knowledge of the values of a function on k elements of the field would allow us to reconstruct a degree k polynomial matching these points. The statement of the above corollary is similar in principle, although the ordinary notion of the degree of a polynomial is replaced with that of the algebraic degree; and the algebraic degree of the interpolated polynomial is determined by the Hamming weight of the interpolated points rather than the number of point-value pairs that we consider.

We note that it is still possible to have a function F' with $\deg(F') > d$ and having the same degree d derivative hypermatrix as F of $\deg(F) = d$. A trivial example would be to take a function F of algebraic degree $\deg(F) > 1$, and consider its values on a basis $\mathcal{B}$ of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$. The latter define a linear function L whose values agree with those of F on all elements of $\mathcal{B}$. Finally, we

note that $\mathcal{B}$ itself is essentially the degree 1 derivative hypermatrix of both F and L.

Nonetheless, we know that a function F of degree d always exists; and furthermore, we can show that the functions obtained via the procedure described in Theorem 1 will always be of algebraic degree d. This is because in the proof of the theorem, we assume that the derivatives Δ_F^d are linear, and reconstruct the values of the function F based on this assumption; as a consequence, Δ_F^k will be the zero function for any $k > d$ (cf. Proposition 1 and the proof of Corollary 2). It remains to formalize the intuitive fact that if all derivatives of a function F of degree greater than k vanish, then the algebraic degree of F cannot be greater than k. Note that we already know that the converse holds by Corollary 2.

We can prove this constructively using Corollary 3 as follows. Suppose F is such that $\Delta_F^k(x_1, x_2, \ldots, x_k) = 0$ for all $x_1, x_2, \ldots, x_k$ and all $k > d$ for some natural number d. Assume without loss of generality that there are arguments $x_1, x_2, \ldots x_d$ for which $\Delta_F^d(x_1, x_2, \ldots x_d)$ does not vanish.

First, we construct the degree 1 derivative matrix of F; by Corollary 3, there exits a function F_1 of degree 1 such that $F - F_1$ vanishes on all points of weight 1. We then construct the degree 2 matrix of $F - F_1$, and obtain by Corollary 3 a function F_2 of degree 2 such that $F - F_1 - F_2$ vanishes on all points of weight 2 or less. We continue in this manner until we have constructed $F' = F - F_1 - F_2 - \cdots - F_d$. Since $\deg(F_i) \leq i$, we have $\deg(F' - F) \leq d$, and so F' still satisfies $\Delta_{F'}^k = 0$ for $k > d$. Thus e.g. the degree $d + 1$ derivative hypermatrix of F' is the zero matrix. Applying Theorem 1, we can see that $F'(x) = 0$ for all x with $\mathrm{wt}(x) > d$. Since the same is true for x with $\mathrm{wt}(x) \leq d$ by the construction of F', we get that F' is the zero function, and thus $\deg(F) \leq d$.

5 On some computational aspects of the derivative representation

The degree d derivative hypermatrix contains $\binom{n}{d}$ elements from $\mathbb{F}_{p^n}$, which is significantly less than the $(p^n)^{p^n}$ needed to represent the truth table. Of course, this only allows us to represent the function up to the addition of terms of lower-algebraic degree. Nonetheless, even in the case when we want to know the exact function (in which case we need to store the derivative hypermatrices of all degrees up to d), we need to store

$$\sum_{j=0}^{d} \binom{n}{j}$$

elements. While this expression does not seem to have an obvious closed form, we can see that it is substantially more compact than the truth table for functions of low algebraic degree. Even in the worst case (for $d = n$), we need to store 2^n elements, while the truth table requires the storage of p^n elements.

As we have seen in the previous section, reconstructing a single value $F(x)$ of a degree d function F involves summing the values of its degree d derivatives on all d-element subsets of its degree d derivative hypermatrix. The values of the derivatives may also need to be multiplied by coefficients from the prime field if some of the coordinates of x are distinct from 0 and 1. Thus, reconstructing the value of x with $\mathrm{wt}(x) = w$ involves summing up and $\binom{w}{d}$ values, and at most $\binom{w}{d}$ multiplications. Clearly, the number of multiplications cannot be greater than the number of additions (plus one), and so evaluating the complexity of the procedure reduces to counting the number of additions.

If we want to recover the entire truth table of the function, we will need

$$\sum_{w=0}^{n} (p-1)^w \binom{w}{d}$$

additions, since for every possible Hamming weight w, there are $(p-1)^w$ elements x with $\mathrm{wt}(x) = w$.

In the case of $p = 2$, the above sum evaluates to $\binom{n+1}{d+1}$ using the identity $\sum_{j=k}^{n} \binom{j}{k} = \binom{n+1}{k+1}$. In the case of $p > 2$, computing the exact form of the sum appears to be less trivial, but we can immediately give the bounds

$$\binom{n+1}{d+1} = \sum_{w=0}^{n} \binom{w}{d} \leq \sum_{w=0}^{n} (p-1)^w \binom{w}{d} \leq (p-1)^n \sum_{w=0}^{n} \binom{w}{d} = (p-1)^n \binom{n+1}{d+1}.$$

On the other hand, if we need to recover all values of the truth table, we can begin by reconstructing all values of Hamming weight d, followed by all values of Hamming weight $d+1$, etc. In this way, a lot of redundant summation can be avoided. From the identity

$$\Delta_F(x_1+x_2+\cdots+x_{d-1}, x_d, \ldots, x_w) = F(x) - \sum_{S \subseteq \{d,d+1,\ldots,w\}} F\left(\sum_{s \in S \cup \{1,2,\ldots,d-1\}} x_s \right),$$

we can see that if we already have the values of all elements of weight less than d computed, and if we know the value of the derivative on the left-hand side in the expression above, we only need to do as many additions as the size of the power set of $\{d, d+1, \ldots, w\}$, i.e. 2^{w-d+1}. The value of the derivative can be recovered with $w - d + 1$ additions using linearity. In total, we would need

$$\sum_{w=d+1}^{n} (p-1)^w (2^{d-1} - 1 + w - d + 1) = \sum_{w=d+1}^{n} w(p-1)^w + (2^{d-1} - d) \sum_{w=d+1}^{n} w$$

additions to reconstruct all elements.

The second sum on the right-hand side is simply

$$\sum_{w=1}^{n} w - \sum_{w=1}^{d} w = \frac{n(n+1)}{2} - \frac{d(d+1)}{2}.$$

n\d	1	2	3	4	5	6	7	8	9	10	11	12	13	14
5	1.01	1.79	1.68	0.98										
6	1.00	2.22	2.68	1.93	0.92									
7	1.00	2.68	4.03	3.71	2.18	0.89								
8	1.00	3.15	5.72	6.55	4.90	2.44	0.87							
9	1.00	3.63	7.74	10.70	9.91	6.25	2.70	0.86						
10	1.00	4.11	10.10	16.37	18.26	14.25	7.78	2.97	0.85					
11	1.00	4.60	12.79	23.83	31.17	29.22	19.70	9.49	3.25	0.84				
12	1.00	5.09	15.81	33.30	50.03	54.94	44.44	26.39	11.37	3.53	0.84			
13	1.00	5.58	19.16	45.03	76.44	96.34	91.21	64.92	34.45	13.44	3.81	0.83		
14	1.00	6.077	22.84	59.27	112.19	159.65	173.38	144.33	91.73	44.00	15.67	4.10	0.83	
15	1.00	6.57	26.85	76.26	159.27	252.61	309.61	295.48	219.53	126.05	55.18	18.09	4.38	0.83

Table 1: Proportion of number of operations when truth table is reconstructed element by element to number of operations when truth table is reconstructed iteratively by Hamming weight

On the other hand, the first sum on the right-hand side can be upper-bounded by $\sum_{w=d+1}^{n}(p-1)^n w$, which evaluates to

$$(p-1)^n \left(\frac{n(n+1)}{2} - \frac{d(d+1)}{2} \right).$$

Furthermore, it is possible to precompute the values of the order d derivatives everywhere, so that their computation from the derivative hypermatrix can then be avoided. This will then dispose of the term $w - d + 1$ in the sum above at the cost of memory.

While the above formulas can be used to obtain the exact number of additions needed to reconstruct the truth table for given values of p, n, d, it is in general difficult to get a sense of how their values related to each other for a given set of parameters. To give an idea of how much more efficient the reconstruction is when the values of F are reconstructed iteratively by Hamming weight, we have computed for $p = 3$ and n in the range $5 \leq n \leq 15$, the proportion of the number of additions needed to compute all values of F one by one to the number of additions needed using the iterative approach. These proportions are displayed in Table 1. All proportions are rounded up or down to two decimal digits. As we can see from the table, the efficiency of the iterative approach is especially pronounced when the degree d of the derivative hypermatrix is approximately $n/2$, and grows with the dimension.

We note that the calculations above are based on the assumption that we only want to reconstruct the function F with $F(x) = 0$ for $\mathrm{wt}(x) < \deg(F)$. If the values of F need to be reconstructed exactly, then the same procedure as above needs to be applied to the derivative hypermatrices of lower degree as well.

In any case, it is evident that while it possible to naturally convert between the truth table and derivative hypermatrix representation, the latter does not offer any particular advantage in terms of memory consumption over the former. The real advantage in storing a function as a matrix of derivatives is when analyzing its behavior on d-dimensional additive subsets; for instance, in the case of APN and planar functions. Furthermore, when the functions considered are of degree $d = 2$, the derivative representation is unique up to EA-equivalence, and this representation is thus advantageous for the study of e.g. quadratic

APN and planar functions that are studied up to EA-equivalence to begin with. In the following section, we restrict to the case of quadratic functions, and discuss how the derivative representation can be applied to the mathematical and computational study of functions with optimal differential uniformity.

Before we proceed, we would like to highlight one advantage that the derivative representation offers from the point of view of computational searches as opposed to the truth table representation. A typical way of searching for new e.g. APN or planar functions is to consider the representation of some known function, and then to search for functions whose representation is "similar" in some sense to that of the original function. If the representation in question is e.g. the univariate one, the search might involve adding a small number of new terms to the polynomial representation of the original function; if the representation is the truth table, the search might involve changing the values of the function at a small number of inputs.

We can easily see that changing a single value of the degree d derivative hypermatrix of F will change the values of $F(x)$ at all x containing the coordinates of the value that was changed. For example, if $d = 3$ and $\Delta_F(b_1, b_2, b_3)$ was changed, then any x with coordinate vector $(x_1, x_2, x_3, \ldots, x_n)$ with $x_1x_2x_3 \neq 0$ will be affected by the change. In total, there are $(p-1)^3p^{n-3}$ such points in the case of $d = 3$, or $(p-1)^dp^{n-d}$, in general; and thus, functions that have very similar derivative representations will have vastly different truth table representations. This means that a computational search based on the derivative hypermatrix will go through a different set of functions than one based on the truth table.

While it is not quite as straightforward to compute how many values will change if a single term in the univariate representation is modified, we can observe by example that this can cause the entire derivative hypermatrix to change. For instance, the function x^2 over $\mathbb{F}_{3^4}$ has a degree 2 derivative matrix

$$\begin{pmatrix} \alpha^{42} & \alpha^{44} & \alpha^{50} & \alpha^{68} \\ \alpha^{44} & \alpha^{46} & \alpha^{52} & \alpha^{70} \\ \alpha^{50} & \alpha^{52} & \alpha^{58} & \alpha^{76} \\ \alpha^{68} & \alpha^{70} & \alpha^{76} & \alpha^{14} \end{pmatrix},$$

while the one for $x^2 + x^6$ is

$$\begin{pmatrix} a^{33} & a^{73} & 2 & a^{51} \\ a^{73} & a^{19} & a^{59} & 2 \\ 2 & a^{59} & a^{57} & a^{17} \\ a^{51} & 2 & a^{17} & a^{11} \end{pmatrix};$$

in both cases, α is the default primitive element chosen by *Magma*, and the basis used is the normal basis $\{\alpha, \alpha^3, \alpha^9, \alpha^{27}\}$.

6 Application to the case of quadratic planar and APN functions

As we have indicated in the discussion above, the derivative hypermatrix representation is particularly useful in the case of quadratic and in the context of planar and APN functions for two reasons. First, the two classes of planar and APN functions are characterized in terms of their first-order derivatives, whose values form the entries of the degree 2 derivative hypermatrix. Second: the functions can be reconstructed up to the addition of affine terms, i.e. up to EA-equivalence, under which planar and affine functions are typically classified anyway. Since in the quadratic case H_F is an ordinary matrix (as opposed to a higher-dimensional hypermatrix), it is also somewhat more intuitive to work with and, as we shall see in the sequel, allows for some characterizations and simplifications that are not quite as straightforward in the multi-dimensional case.

In the quadratic case, the degree 2 derivative hypermatrix (or simply, the derivative matrix of F) takes the form

$$H_F = \begin{bmatrix} \Delta_F(b_1, b_1) & \Delta_F(b_1, b_2) & \dots & \Delta_F(b_1, b_n) \\ \Delta_F(b_2, b_1) & \Delta_F(b_2, b_2) & \dots & \Delta_F(b_2, b_n) \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_F(b_n, b_1) & \Delta_F(b_n, b_2) & \dots & \Delta_F(b_n, b_n) \end{bmatrix}.$$

In the following, we will describe how the differential uniformity of a quadratic function can be characterized directly in terms of its derivative matrix. One of the immediate applications of such a characterization would be a computational search for e.g. APN or planar functions; it is then imperative to be able to quickly convert such functions to univariate representation so that they can be conveniently compared and classified up to equivalence against known representatives. The general procedure described in Theorem 1 allows us to reconstruct the truth table of a function F representing the CCZ-class given by by H_F; the univariate representation can then be obtained by Lagrange interpolation from the truth table.

In the quadratic case, however, it is possible to obtain the univariate representation directly from H_F and vice-versa. This involves only matrix multiplication, and is thus both simple to implement and highly efficient. This is formalized in the following proposition.

Proposition 3. *Let $\{b_1, \dots, b_n\}$ be a basis of $\mathbb{F}_p^n$. Let $F : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ be given by $F(x) = \sum_{1 \le i,j \le n-1, i \le j} a_{ij} x^{p^{i-1}+p^{j-1}}$. Then*

$$H_F = B^T A B, \tag{5}$$

where $B = \begin{pmatrix} b_1^{p^0} & b_2^{p^0} & \dots & b_n^{p^0} \\ b_1^{p^1} & b_2^{p^1} & \dots & b_n^{p^1} \\ \vdots & \vdots & \ddots & \vdots \\ b_1^{p^{n-1}} & b_2^{p^{n-1}} & \dots & b_n^{p^{n-1}} \end{pmatrix}$, $A = \begin{pmatrix} 2a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & 2a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & 2a_{nn} \end{pmatrix}$,
and B^T *is the transpose of* B.

Proof. We shall show that $\sum_{l=1}^{n}(B^TA)_{il}B_j = (H_F)ij$, for any $1 \le i, j \le n$. Indeed, for any $1 \le i \le n$:

$(B^TA)_{i1} = 2a_{11}b_i^{p^0} + a_{12}b_i^p + a_{13}b_i^{p^2} + \ldots + a_{1n}b_i^{p^{n-1}}$;
$(B^TA)_{i2} = a_{12}b_i^{p^0} + 2a_{22}b_i^p + a_{23}b_i^{p^2} + \ldots + a_{2n}b_i^{p^{n-1}}$;
...
$(B^TA)_{in} = a_{1n}b_i^{p^0} + a_{2n}b_i^p + a_{3n}b_i^{p^2} + \ldots + a_{nn}b_i^{p^{n-1}}$.

Then, for any $1 \le i, j \le n$, we obtain

$$(B^TAB)_{ij} = \sum_{l=1}^{n}(B^TA)_{il}B_j =$$
$$(2a_{11}b_i^{p^0} + a_{12}b_i^p + a_{13}b_i^{p^2} + \ldots + a_{1n}b_i^{p^{n-1}})b_j^{p^0} +$$
$$(a_{12}b_i^{p^0} + 2a_{22}b_i^p + a_{23}b_i^{p^2} + \ldots + a_{2n}b_i^{p^{n-1}})b_j^{p^1} +$$
$$\ldots$$
$$(a_{1n}b_i^{p^0} + a_{2n}b_i^p + a_{3n}b_i^{p^2} + \ldots + a_{nn}b_i^{p^{n-1}})b_j^{p^{n-1}} =$$
$$\sum_{1\le l,k\le n, l\le k} a_{lk}(b_i^{p^{l-1}}b_j^{p^{k-1}}) + \sum_{1\le l,k\le n, l\le k} a_{lk}b_j^{p^{l-1}}b_i^{p^{k-1}} =$$
$$\sum_{1\le l,k\le n, l\le k} a_{lk}(b_i+b_j)^{p^{k-1}+p^{l-1}} - \sum_{1\le l,k\le n, l\le k} a_{lk}b_i^{p^{k-1}+p^{l-1}} -$$
$$\sum_{1\le l,k\le n, l\le k} a_{lk}b_j^{p^{k-1}+p^{l-1}} = F(b_i+b_j) - F(b_i) - F(b_j) = \Delta_{b_i}F(b_j).$$

This completes the proof. □

Following [18], we define the rank of a vector $v \in \mathbb{F}_{p^n}^n$ to be the dimension of the subspace spanned by its elements. In other words, if $v = (v_1, v_2, \ldots, v_n)$ with $v_i \in \mathbb{F}_{p^n}$, the **rank** of v is $r(v) = \log_p \#\{a_1v_1 + a_2v_2 + \cdots + a_nv_n : a_1, a_2, \ldots, a_n \in \mathbb{F}_p\}$. We now have the following characterization.

The following proposition provides an "algebraic" characterization of the differential uniformity of a quadratic function in terms of its derivative matrix, and is the basis for our computational search described in the next section. We note that special cases of this characterization have been previously used in computational searches such as [18] and [17]. Both of these approaches are remarkable in their own right: the work in [18] allowed the authors of that paper to find thousands of previously unknown examples of CCZ-inequivalent

APN functions over $\mathbb{F}_{2^8}$, while prior to this work, only a small number of such functions (less than 50) were known. In [17], a classification of all quadratic APN functions with prime field coefficients over $\mathbb{F}_{2^n}$ for $n \leq 9$ was obtained; for comparison, complete classifications of APN functions over $\mathbb{F}_{2^n}$ were only known for $n \leq 5$ in general, for $n \leq 6$ for cubic functions, and for $n \leq 7$ for quadratic functions. These were significant computational advances in the study of APN functions, and promise that the adaptation of these methods to the case of e.g. planar functions might prove to be similarly useful.

Proposition 4. *Let F be an (n, n)-function and H_F be its derivative matrix. Then F has differential uniformity $\delta = p^k$ if and only if any non-zero linear combination of the rows of H_F has rank $n - (k + 1)$.*

Proof. First, note that each row of the derivative matrix represents the values of a derivative $\Delta_F b_i(x) = \Delta_F^2(b_i, x)$ on $x \in \mathcal{B}$ for some $b_i \in \mathcal{B}$, where $\mathcal{B}$ is a basis of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$. Then a linear combination of the rows is a vector giving the values of a derivative $\Delta_F^2(a, x)$ on $x \in \mathcal{B}$ for some $a \in \mathbb{F}_{p^n}$. The rank of this linear combination is the linear subspace spanned by its elements, and since the first-order derivatives $x \mapsto \Delta_F^2(a, x)$ are linear in x, this subspace consists of precisely all values taken by $x \mapsto \Delta_F^2(a, x)$.

Suppose that the rank of some linear combination of the rows of H_F is at most $n - k$ for some $k > 1$, and let $\Delta_F^2(a, x)$ be the corresponding derivative. Then the corresponding subspace must contain p^k zeros, i.e. $\Delta_F^2(a, x)$ must take value 0 for p^k distinct $x \in \mathbb{F}_{p^n}$. This means that the equation $\Delta_F^2(a, x) = 0$ has at least p^n solutions $x \in \mathbb{F}_{p^n}$, and so the differential uniformity of F must be at least p^k.

Conversely, suppose $\delta_F \geq p^k$; then there are at least p^k elements $x \in \mathbb{F}_{p^n}$ solving $\Delta_F^2(a, x) = b$ for some $0 \neq a \in \mathbb{F}_{p^n}$ and some $b \in \mathbb{F}_{p^n}$, and since $\Delta_F^2(a, x)$ is linear in x, this means that its kernel is of size at least p^k. Consequently, the dimensions of the subspace $\{\Delta_F^2(a, x) : x \in \mathbb{F}_{p^n}\}$ is at most p^{n-k}, i.e. the rank of the vector $(\Delta_F^2(a, b_1), \ldots, \Delta_F^2(a, b_n))$ is at most $n-k$. This vector can clearly be expressed as a linear combination of the rows of the derivative matrix, which concludes the proof. □

Since our practical interest is predominantly focused on the case of planar and APN functions, we explicitly state the implications of Proposition 4 for these two special cases as follows.

Corollary 5. *Let F be an (n, n, p)-function and H_F be its derivative matrix. Then:*

- *F is planar if and only if any non-zero linear combination of the rows of H_F has rank n;*
- *If $p = 2$, then H_F only has zeros on its main diagonal, and F is APN if and only any linear combination of the rows of H_F has rank $n - 1$.*

This characterization can be compared with Theorem 1 and Definition 5 of [18]. The condition on the rank of the linear combinations of rows of H_F is the same in both cases; the advantage of our approach is that the matrix H_F has a clear intuitive meaning (containing the values of the first-order derivatives of F on the basis elements), and is consequently easier to analyze and to construct from F in practice. Note that in the case of odd characteristic, the main diagonal of H_F is not necessarily zero since $\Delta_x F(x)$ is not equal to 0 in general.

In particular, from the interpretation of H_F in terms of the derivatives of F, we see that applying a linear permutation $L : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ to all elements of H_F gives the derivative matrix of $L(F)$. Compare this with Theorem 3 of [18] which requires a non-trivial proof. We state this as an observation; in practice, we use it to restrict the number of matrices that we consider in our search.

Observation 1. *Let H_F be the derivative matrix of $F : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$, and let $L : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ be a linear function. Then the matrix H'_F defined by $(H'_F)_{i,j} = L((H_F)_{i,j})$ for all i, j is the derivative matrix of $L \circ F$. In particular, if L is a permutation, then H_F and H'_F correspond to EA-equivalent functions.*

7 Functions with prime field coefficients

As in [17], we now consider the case of quadratic functions $F : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ with prime field coefficients, i.e. with coefficients in the subfield $\mathbb{F}_p$. Since the Frobenius automorphism $x \mapsto x^p$ fixes $\mathbb{F}_p$, we have $F(x^p) = F(x)^p$ (and, more generally, $F(x^{p^k}) = F(x)^{p^k}$ for any non-negative integer k) for any such function. Consequently, we have $\Delta_{a^{p^k}} F(x^{p^k}) = (\Delta_a F(x))^{p^k}$ for any non-negative integer k. If we construct the matrix H_F corresponding to F with respect to a normal basis, i.e. with respect to a basis $\mathcal{B} = \{b, b^p, b^{p^2}, \ldots, b^{p^{n-1}}\}$ for some suitable $b \in \mathbb{F}^*_{p^n}$, then H_F will be such that $(H_F)_{i+1,j+1} = (H_F)^p_{i,j}$ for any $0 \leq i, j \leq n-1$; here we index the rows and columns from 0 to $n-1$, since $(H_F)_{i+1,j+1} = (H_F)^p_{i,j}$ is true if the indices i, j are considered modulo n; in other words, we have e.g. $(H_F)_{0,1} = (H_F)^p_{n-1,0}$.

This severely restricts the number of elements that we have to guess in order to completely determine H_F. For instance, the matrices H^6_F and H^7_F

corresponding to a $(6,6)$- and $(7,7)$-function become

$$H_F^6 = \begin{bmatrix} A & B & C & D & C^{p^4} & B^{p^5} \\ B & A^p & B^p & C^p & D^p & C^{p^5} \\ C & B^p & A^{p^2} & B^{p^2} & C^{p^2} & D^{p^2} \\ D^{p^3} & C^p & B^{p^2} & A^{p^3} & B^{p^3} & C^{p^3} \\ C^{p^4} & D^{p^4} & C^{p^2} & B^{p^3} & A^{p^4} & B^{p^4} \\ B^{p^5} & C^{p^5} & D^{p^5} & C^{p^3} & B^{p^4} & A^{p^5} \end{bmatrix},$$

$$H_F^7 = \begin{bmatrix} A & B & C & D & D^{p^5} & C^{p^4} & B^{p^6} \\ B & A^p & B^p & C^p & D^p & D^{p^5} & C^{p^6} \\ C & B^p & A^{p^2} & B^{p^2} & C^{p^2} & D^{p^2} & D^{p^6} \\ D & C^p & B^{p^2} & A^{p^3} & B^{p^3} & C^{p^3} & D^{p^3} \\ D^{p^4} & D^p & C^{p^2} & B^{p^3} & A^{p^4} & B^{p^4} & C^{p^4} \\ C^{p^5} & D^{p^5} & D^{p^2} & C^{p^3} & B^{p^4} & A^{p^5} & B^{p^5} \\ B^{p^6} & C^{p^6} & D^{p^6} & D^{p^3} & C^{p^4} & B^{p^5} & A^{p^6} \end{bmatrix},$$

respectively, with $A, B, C, D \in \mathbb{F}_{p^6}$ for H_F^6, and $A, B, C, D \in \mathbb{F}_{p^7}$ for H_F^7. It is easy to see that in the case of even n, we have to guess $n/2+1$ values in order to specify H_F, while for odd n, we have to guess $(n+1)/2$ values. When n is even, we can restrict one of the values to the subfield $\mathbb{F}_{p^{n/2}}$: for instance, in H_F^6, we have $D = D^{p^3}$ due to the fact that H_F^6 is symmetric (since $\Delta_a F(x) = \Delta_x F(a)$), and so we must have $D \in \mathbb{F}_{p^3}$. This naturally generalizes to an arbitrary even dimension n.

Some further necessary conditions can be obtained by observing that the linear combinations of the rows of any submatrix of H_F must also have full rank. Following [18], we say that a matrix $S \in \mathbb{F}_{p^n}^{m \times k}$ is **proper** if any non-zero linear combination of the rows of S has rank k. Thus, H_F is proper if and only if it represents a planar function; and, clearly, if H_F is proper, then the same is true for any submatrix of H_F (since if some linear combination of the rows of a submatrix $S \in \mathbb{F}_p^{m \times k}$ of H_F spans a subspace of dimension less than k, then the same linear combination of the rows of the entire matrix H_F will have rank less than n since appending $n-k$ elements can increase the rank by at most $n-k$).

This submatrix condition is particularly valuable for submatrices that only depend on a subset of the variables needed to specify the matrix. For instance, the submatrix of H_F^6 on the rows with indices $\{0,1\}$ and the columns with indices $\{0,1,2,5\}$ depends on A, B, C, but not on D. Similarly, the submatrix with rows and columns with indices $\{0,1\}$ depends only on A and B. After guessing the value of e.g. A and B, we can check whether all submatrices that depend only on A and B are proper; if not, we can backtrack immediately, thus saving significant computation time.

In this paragraph, we will denote the matrix corresponding to the rows with indices R and columns with indices C by (R, C). For H_F^6, we use the submatrices corresponding to $(\{0,5\},\{0,5\})$, $(\{0,1\},\{0,1\})$ that depend only on A and B; and those corresponding to $\{(0,1,5\},\{0,1,5\})$, $(\{0,2,4\},\{0,2,4\})$,

$(\{0,1\},\{0,1,2,5\})$, $(\{0,2\},\{0,1,2,4\})$, $(\{0,5\},\{0,1,4,5\})$, $(\{0,1,4\},\{0,2,5\})$ and $(\{0,1,2\},\{0,1,2\})$ depending only on A, B, C.

In the case of H_F^7, we use $(\{0,6\},\{0,6\})$ and $(\{0,1\},\{0,1\})$ that only depend on A, B, and $(\{0,1,6\},\{0,1,6\})$, $(\{0,5,6\},\{0,5,6\})$, $(\{0,1\},\{0,1,2,6\})$, $(\{0,6\},\{0,1,5,6\})$, $(\{0,1,2\},\{0,1,2\})$ that depend on A, B, C.

We note that the above lists do not exhaust all submatrices that only depend on a subset of values, but according to our empirical observations, verifying whether other submatrices are proper does not detect any contradictions beyond the ones obtained from the submatrices listed above. For dimensions n less than 6, the computation time is so short that we do not have to consider submatrix conditions of this form. As an example of how this improves the efficiency of the search, we note that for $n = 6$, conducting the search for one fixed value of A without the submatrix conditions takes around 7000 seconds as opposed to around 5500 seconds with the submatrix conditions.

8 Computational results

We run our searches on a server with 56 3.2 GHz cores and 500 GB of memory. We perform an exhaustive search over all possible matrices H_F corresponding to quadratic functions with prime field coefficients over $\mathbb{F}_{3^n}$ for $n \leq 7$. In order to facilitate the search, we use Observation 1 to restrict the value of one of the entries of H_F. However, while there is a linear permutation L such that $L(c) = c'$ for any two non-zero $c, c' \in \mathbb{F}_{p^n}$, the composition of such a permutation with a function having prime field coefficients is not necessarily going to have prime field coefficients, and so we cannot simply fix the value of the first variable in H_F to 1. However, we consider all linear permutations with prime field coefficients over $\mathbb{F}_{3^n}$, and use them to restrict the choice of the first variable, A, in H_F. More precisely, we define an equivalence relation $\sim$ on $\mathbb{F}_{p^n}^*$ with $a \sim b$ if there exists a linear permutation $L : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ with prime field coefficients such that $L(a) = b$. The number of such linear permutations is sufficiently small for us to partition $\mathbb{F}_{3^n}^*$ into equivalence classes according to $\sim$ for all the dimensions that we consider. Since composing two functions with prime field coefficients gives a function that also has prime field coefficients, it then suffices to consider only one element from each class as the value of the first variable in H_F. For $n = 5$ and $n = 7$, we get 3 equivalence classes; for $n = 4$, we get 7; and for $n = 6$, we get 15.

In the case of $n = 4$, the search takes less than 2 seconds, and yields 24 functions. For $n = 5$, it takes about 17 seconds and yields 616 functions. For $n = 6$, we run 15 parallel processes, one for each equivalence class of the relation $\sim$ described in the previous paragraph; each process (with the submatrix conditions) takes around 5500 seconds (as pointed out above); in total, we get 2928 functions. Finally, in the case of $n = 7$, we conduct the search by running 22 processes in parallel on the server (each process handling all three equivalence classes under $\sim$), with each process handling 100 (out of $3^7 - 1$) possible values for B. Each of the 22 processes takes around 150 000 seconds to

finish. Ultimately, we obtain 5093 functions.

The real bottleneck is classifying the functions under CCZ-equivalence. The code isomorphism test can take up to around 5 seconds for $n = 5$, around a minute for $n = 6$, and around an hour for $n = 7$. In the conference paper [8], we use an ad-hoc method to speed up the classification for $n = 7$, which involves composing the functions obtained from the search with linear permutations with coefficients in the prime field at random in order to hopefully eliminate equivalent pairs of functions without having to go through the computationally expensive code isomorphism test. The entire computation takes about three months; we refer the reader to [8] for details.

We also classify the functions using the recently developed algorithm for testing linear equivalence of even planar functions [10]. This turns out to be significantly faster, and finishes within three days of computation; the results match those of the code isomorphism test.

We omit a list of the known CCZ-classes of planar functions since we see that all functions that we find are CCZ-equivalent to one of the known instances. We refer the reader to [14] for an excellent survey on planar functions which includes all known families and sporadic instances. The families referenced in the last column of Table 2 refer to the names used in [14].

We only find functions that are CCZ-equivalent to known ones. However, in the case of $n = 6$, we find representatives for the Zhou-Pott, the Dickson, and the Lunardon-Marino-Polverino-Trombetti-Bierbrauer (LMPTB) functions that are simpler than the known ones. Comparing with the representatives given in [3] and [15], we see that in the case of Zhou-Pott, our representative

$$x^{162} + 2x^{108} + 2x^{90} + x^{82} + 2x^{10} + x^4 + x^2$$

has 7 terms with prime field coefficients, while the one in [3], viz.

$$\alpha^{140}x^{324} + \alpha^{504}x^{246} + \alpha^{284}x^{108} + \alpha^{504}x^{90} + \alpha^{674}x^{82} + \alpha^{506}x^{54} + \alpha^{726}x^{30} + \alpha^{225}x^{28} + \alpha^{140}x^{12} + \alpha^{388}x^4 + \alpha^{532}x^2$$

has 11 terms with various coefficients; for the Dickson case, our representative

$$2x^{270} + 2x^{244} + x^{54} + x^{36} + x^{10} + x^2$$

has 5 terms, as opposed to the 6 terms in [15], viz.

$$x^{162} + x^{84} + \alpha^{58}x^{54} + \alpha^{58}x^{28} + x^6 + \alpha^{531}x^2$$

and the 7 terms in [3], which we omit here; finally, for LMPTB, our representative

$$2x^{486} + x^{270} + 2x^{162} + x^{90} + x^2$$

has 5 terms while the one in [15], viz.

$$2x^{270} + x^{246} + 2x^{90} + x^{82} + x^{54} + 2x^{30} + x^{10} + x^2$$

has 7 terms; a different representative is given in [3], and also has 7 terms.

Table 2: CCZ-representatives from all quadratic planar functions with prime field coefficients over $\mathbb{F}_{3^n}$ with $4 \leq n \leq 7$

n	F	Family
4	x^2	Finite field
	$x^{36} + 2x^{10} + 2x^4$	Dickson
5	x^2	Finite field
	x^4	Albert
	x^{10}	Albert
	$x^{10} + x^6 + 2x^2$	Coulter-Matthews-Ding-Yuan
	$x^{10} + 2x^6 + 2x^2$	Coulter-Matthews-Ding-Yuan
	$x^{90} + x^2$	sporadic
	$x^{162} + x^{108} - x^{84} + x^2$	sporadic
6	x^2	Finite field
	x^{10}	Albert
	$x^{162} + 2x^{108} + 2x^{90} + x^{82} + 2x^{10} + x^4 + x^2$	Zhou-Pott (*)
	$2x^{270} + 2x^{244} + x^{54} + x^{36} + x^{10} + x^2$	Dickson (*)
	$2x^{486} + x^{270} + 2x^{162} + x^{90} + x^2$	LMPTB (*)
7	x^2	Finite field
	x^4	Albert
	x^{10}	Albert
	x^{28}	Albert
	$x^{10} + x^6 + 2x^2$	Coulter-Matthews-Ding-Yuan
	$x^{10} + 2x^6 + 2x^2$	Coulter-Matthews-Ding-Yuan

More importantly, we obtain a complete classification of all quadratic planar functions with prime field coefficients over $\mathbb{F}_{3^n}$ up to $n = 7$. A complete overview is given in Table 2.

We also try to extend our search to the case of dimension $n = 8$. In this case, we end up with 31 possible choices for the value of the first element A of the derivative matrix, namely

$$\{1, \alpha, \alpha^2, \alpha^4, \alpha^7, \alpha^{10}, \alpha^{11}, \alpha^{13}, \alpha^{16}, \alpha^{19}, \alpha^{35}, \alpha^{37}, \alpha^{41}, \alpha^{43}, \alpha^{55}, \alpha^{65}, \alpha^{71}, \alpha^{82}, \alpha^{164}, \alpha^{187}, \alpha^{236}, \alpha^{319}, \alpha^{410}, \alpha^{413}, \alpha^{436}, \alpha^{484}, \alpha^{533}, \alpha^{820}, \alpha^{1066}, \alpha^{1640}, \alpha^{2173}\}$$

where α is the default primitive element of $\mathbb{F}_{3^8}$ selected by *Magma*. Unfortunately, even with the consideration of sub-matrices in order to prune branches of the search tree as before, the computational load seems to be too much: after approximately three weeks of computation for one of the choices of A, only 9 (out of $3^8 - 1 = 6560$) possibilities for the second element B have been processed. We thus conclude that pushing the same approach farther would be difficult, and exploring dimensions beyond 7 would require significant optimizations to the approach, or large-scale parallel computations.

9 Conclusion and directions for future work

We have described how an (n, n, p)-function of algebraic degree d can be represented (up to addition of terms of lower algebraic degree) using a so-called derivative hypermatrix containing the values of its derivatives of order $(d-1)$ on the elements of a basis of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$, and how to transition between this derivative representation and the truth table of the function. We have identified a set $\mathcal{B}_d = \{b_1, b_2, \ldots, b_m\}$ (called a degree d extension) corresponding to any given basis $\mathcal{B}$ of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$ such that for any choice of values $v_1, v_2, \ldots, v_m$ for the elements in $\mathcal{B}_d$ it is possible to find an (n, n, p)-function F with $F(b_i) = v_i$ for all i, and with $\deg(F) = d$. We have discussed some of the advantages provided by this derivative representation from the point of view of computational searches, and considered the complexity of converting between it and the truth table representation.

In the case of quadratic functions, we have described how to directly obtain the univariate from the derivative representation, and vice-versa; and we have shown how the differential uniformity of a quadratic function can be characterized directly in terms of the derivative matrix. In this way, we have generalized the matrix representation and characterization of quadratic APN functions from [18] to the case of characteristics other than 2. We have used this generalized characterization to perform a computational search for quadratic planar functions over $\mathbb{F}_{3^n}$ with coefficients in $\mathbb{F}_3$, and have provided a complete classification of such functions up to CCZ-equivalence. Our search has also produced simpler representatives for some of the known classes of planar functions. Furthermore, we have confirmed that the previously known representatives exhaust all planar functions of this type up to CCZ-equivalence. Finally, we have motivated why continuing this search in higher dimensions would require significant computational resources, and so a dedicated computational effort would be required to obtain further results in this direction.

There are multiple directions left for future work. As outlined above, a potential direction would be to conduct similar searches for dimensions greater than 8, or for characteristics other than 3. This would require a non-trivial amount of computational resources, and it would likely be necessary to parallelize the computation and run it on highly efficient hardware in order to conduct the search within a reasonable amount of time.

At the moment, we have focused our practical applications to the quadratic case, since the main motivation for our study were the classes of planar and APN functions which are defined in terms of their first-order derivatives. It would be interesting to see what classes of functions can be defined in characterized using their higher-order derivatives, which would then allow these classes to be advantageously represented and studied using the higher-order derivative hypermatrices. In this respect, finding an analogue to Proposition 3 for higher-order functions would be useful since it would allow the polynomial form of such functions to be extracted directly, i.e. without going through the truth table.

References

[1] Wieb Bosma, John Cannon, and Catherine Playoust. The Magma algebra system I: The user language. *Journal of Symbolic Computation*, 24(3-4):235–265, 1997.

[2] Marcus Brinkmann and Gregor Leander. On the classification of APN functions up to dimension five. *Designs, codes and cryptography*, 49(1):273–288, 2008.

[3] Lilya Budaghyan, Marco Calderini, Claude Carlet, Robert Coulter, and Irene Villa. On isotopic shift construction for planar functions. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 2962–2966. IEEE, 2019.

[4] Lilya Budaghyan and Tor Helleseth. New commutative semifields defined by new PN multinomials. *Cryptography and communications*, 3(1):1–16, 2011.

[5] Anne Canteaut, Alain Couvreur, and Léo Perrin. Recovering or testing extended-affine equivalence. *IEEE Transactions on Information Theory*, 2022.

[6] Claude Carlet. *Boolean functions for cryptography and coding theory*. Cambridge University Press, 2021.

[7] Robert S Coulter and Marie Henderson. Commutative presemifields and semifields. *Advances in Mathematics*, 217(1):282–304, 2008.

[8] Diana Davidova and Nikolay S Kaleyski. Classification of all DO planar polynomials with prime field coefficients over $\mathbb{F}_{3^n}$ for $n \leq 7$.

[9] Yves Edel and Alexander Pott. On the equivalence of nonlinear functions. In *Enhancing cryptographic primitives with techniques from error correcting codes*, pages 87–103. IOS Press, 2009.

[10] Ivana Ivkovic and Nikolay Kaleyski. Deciding and reconstructing linear equivalence of uniformly distributed functions. Cryptology ePrint Archive, Paper 2022/666, 2022. https://eprint.iacr.org/2022/666.

[11] Nikolay Kaleyski. Deciding EA-equivalence via invariants. *Cryptography and Communications*, 14(2):271–290, 2022.

[12] Konstantin Kalgin and Valeriya Idrisova. The classification of quadratic APN functions in 7 variables and combinatorial approaches to search for APN functions. *Cryptography and Communications*, pages 1–18, 2022.

[13] Philippe Langevin, Z Saygi, and E Saygi. Classification of APN cubics in dimension 6 over gf (2), 2020.

[14] Alexander Pott. Almost perfect and planar functions. *Designs, Codes and Cryptography*, 78(1):141–195, 2016.

[15] Alexander Pott and Yue Zhou. Switching construction of planar functions on finite fields. In *International Workshop on the Arithmetic of Finite Fields*, pages 135–150. Springer, 2010.

[16] Satoshi Yoshiara. Equivalences of quadratic APN functions. *Journal of Algebraic Combinatorics*, 35(3):461–475, 2012.

[17] Yuyin Yu, Nikolay Kaleyski, Lilya Budaghyan, and Yongqiang Li. Classification of quadratic APN functions with coefficients in $\mathbb{F}_2$ for dimensions up to 9. *Finite Fields and Their Applications*, 68:101733, 2020.

[18] Yuyin Yu, Mingsheng Wang, and Yongqiang Li. A matrix approach for constructing quadratic APN functions. *Designs, codes and cryptography*, 73(2):587–600, 2014.