Normality of Boolean bent functions in eight variables, revisited

Alexandr Polujan[1], Luca Mariot[2], and Stjepan Picek[3]

[1]Otto von Guericke University Magdeburg
Universitätsplatz 2, 39106, Magdeburg, Germany
alexandr.polujan@gmail.com

[2]Semantics, Cybersecurity and Services Group, University of Twente,
Drienerlolaan 5, 7511GG Enschede, The Netherlands
l.mariot@utwente.nl

[3]Digital Security Group, Radboud University
Postbus 9010, 6500 GL Nijmegen, The Netherlands
stjepan.picek@ru.nl

Abstract

There are approximately 2^{106} bent functions in 8 variables, and the known constructions cover only a tiny part of all these functions [9]. However, finding *"rare"* bent functions, i.e., those which do not arise from generic classes of functions or those of which examples are only a few known, is still a non-trivial problem. In this paper, we give for the first time an example of a non-normal partial spread bent function in 8 variables by analyzing the list of all partial spread bent functions [8], thus solving two open problems by Charpin [4, Open problem 5] and Leander [10, p.17], respectively. Additionally, we show that all partial spread bent functions in $n = 8$ variables are either normal or weakly normal. Finally, using evolutionary algorithms, we show that it is possible to construct bent functions which do not belong, up to equivalence, to the Maiorana-McFarland class.

Keywords: Boolean bent function, partial spread class, normality, evolutionary computation.

1 Preliminaries

A mapping $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ is called a *Boolean function*. For $a \in \mathbb{F}_2^n$, the *Walsh transform* $\hat{\chi}_F\colon \mathbb{F}_2^n \to \mathbb{Z}$ is defined by $\hat{\chi}_f(a) = \sum_{x\in\mathbb{F}_2^n}(-1)^{f(x)+a\cdot x}$, where $a \cdot x = a_1x_1 + \cdots + a_nx_n$. A Boolean function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ is called *bent* if its Walsh transform satisfies $\hat{\chi}_f(a) = \pm 2^{n/2}$ for all $a \in \mathbb{F}_2^n$.

Definition 1.1. A Boolean function f on $\mathbb{F}_2^n$ is said to be *normal* if it is constant on some affine subspace U of $\mathbb{F}_2^n$ of dimension $\lceil n/2 \rceil$. In this case, f is said to be normal with respect to the affine space U. If no such an affine space exists, f is said to be *non-normal*.

To prove theoretically that a given bent function f on $\mathbb{F}_2^n$ is non-normal is a very challenging task. Nevertheless, for small values of n (i.e., $n \leq 8$), one can check the normality of a given bent function with the help of Algorithm 1.1. With a recursive algorithm suggested in [2, Algorithm 1], several examples of non-normal bent functions in $n = 10, 12, 14$ variables were obtained. For example, the restriction of the Kasami–Welch function $x \in \mathbb{F}_{2^{11}} \mapsto Tr(x^{241})$ to the trace 0 (and trace 1) elements is a non-normal bent function in $n = 10$ variables [11, Fact 14]. Note that $n = 10$ is the smallest number of variables for which such a bent function is known. Using the direct sum construction, one can construct new non-normal bent functions in an arbitrary number of variables from the known in the following way.

Result 1.2. *[10, p. 24] Let f be a Boolean bent function on $\mathbb{F}_2^n$ and g be a quadratic Boolean bent function on $\mathbb{F}_2^m$. Then $h(x, y) = f(x) + g(y)$ is normal on $\mathbb{F}_2^n \times \mathbb{F}_2^m$ iff f is normal on $\mathbb{F}_2^n$.*

Despite the progress on the normality of bent functions in $n \geq 10$ variables, the following two questions (the first is due Charpin [4, Open problem 5] and the second due Leander [10, p.17]) still remain not answered:

Algorithm 1.1. Checking normality (according to [4, Theorem 1]).

Require: Bent function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$.

1: **for all** subspaces V of dimension $n/2$ **do**

2: **Check** the following condition: f is constant on $b+V$ if and only if

$$(-1)^{b\cdot v}\hat{\chi}_f(v) = \varepsilon 2^k, \text{ for all } v \in V^\perp = \{u \in \mathbb{F}_2^n \colon u \cdot v = 0 \text{ for all } v \in V\},$$

where ε is constant, equal either to $+1$ or -1.

3: **Output** affine subspaces $b+V$, on which f is constant.

4: **end for**

1. Do non-normal bent functions of 8 variables and degree 4 exist?
2. Do non-normal bent functions in the $\mathcal{PS}^- \setminus \mathcal{PS}_{ap}$ class exist?

In the following section, we positively answer both of the mentioned questions by finding among all $\mathcal{PS}$ bent functions in $n = 8$ variables [8] a non-normal bent function in $\mathcal{PS}^- \setminus \mathcal{PS}_{ap}$.

2 A non-normal partial spread bent function in eight variables

First, we give a definition of a partial spread and define its canonical representation.

Definition 2.1. A partial spread of order s in $\mathbb{F}_2^n$ with $n = 2k$ is a set of s vector subspaces $U_1, \ldots, U_s$ of $\mathbb{F}_2^n$ of dimension k each, such that $U_i \cap U_j = \{0\}$ for all $i \neq j$. The partial spread of order $s = 2^k + 1$ in $\mathbb{F}_2^n$ with $n = 2k$ is called a spread.

Following the notation in [8], for two matrices $A, B \in \mathbb{F}_2^{(k,k)}$ s.t. $\operatorname{rank}[A\ B] = k$, we denote by $[A : B]$ the linear span of the rows of $[A\ B]$. Let 0_k and I_k denote the all-zero and all-one matrix of order k, respectively. Any partial spread of order s is equivalent to one of the form

$$\mathcal{S} = \{\underbrace{[0_k : I_k]}_{U_1}, \underbrace{[I_k : 0_k]}_{U_2}, \underbrace{[I_k : I_k]}_{U_3}, \underbrace{[I_k : A_4]}_{U_4}, \ldots, \underbrace{[I_k : A_s]}_{U_s}\}, \tag{2.1}$$

where $A_2(= 0_k), A_3(= I_k), A_4, \ldots, A_s$ have the property that $A_i - A_j$ is invertible for all $2 \leq i < j \leq s$. In the following, we denote by $\mathbb{1}_U : \mathbb{F}_2^n \to \mathbb{F}_2$ the *indicator function* of $U \subseteq \mathbb{F}_2^n$, i.e., $\mathbb{1}_U(x) = 1$ if $x \in U$, and 0 otherwise. Let the vector spaces $U_1, \ldots, U_{2^{k-1}+1}$ of $\mathbb{F}_2^n$ form a partial spread in $\mathbb{F}_2^n$. The *partial spread class* $\mathcal{PS}$ of bent functions on $\mathbb{F}_2^n$ is the union of the following two classes [5]: the $\mathcal{PS}^+$ *class* is the set of Boolean bent functions of the form $f(x) = \sum_{i=1}^{2^{k-1}+1} \mathbb{1}_{U_i}(x)$; the $\mathcal{PS}^-$ *class* is the set of Boolean bent functions of the form $f(x) = \sum_{i=1}^{2^{k-1}} \mathbb{1}_{U_i^*}(x)$, where $U_i^* := U_i \setminus \{0\}$. The *Desarguesian partial spread* class $\mathcal{PS}_{ap} \subset \mathcal{PS}^-$ is the set of Boolean bent functions f on $\mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$ of the form $f\colon (x, y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k} \mapsto h\,(x/y)$, where $\frac{x}{0} = 0$, for all $x \in \mathbb{F}_{2^k}$ and $h\colon \mathbb{F}_{2^k} \to \mathbb{F}_2$ is a balanced Boolean function with $h(0) = 0$.

Clearly, every $\mathcal{PS}^+$ bent function f on $\mathbb{F}_2^n$ is normal, since $f|_{U_i} = 1$ for every spread line U_i. Moreover, all functions in $\mathcal{PS}_{ap}$ class are normal, since they vanish on the k-dimensional subspace $\{0\} \times \mathbb{F}_{2^k}$. However, the question about the normality of bent functions in $\mathcal{PS}^- \setminus \mathcal{PS}_{ap}$, becomes non-trivial since, in this case, one deals with the sets U_i^*, which are not vector subspaces anymore.

Partial spreads on $\mathbb{F}_2^8$ were completely classified in [8]; the representatives of the corresponding bent functions are available (at the moment of submission of this article) at [7]. Remarkably, there exist 9,316 partial spreads of order 8 on $\mathbb{F}_2^8$, and each of them gives rise to a partial spread bent function in the $\mathcal{PS}^-$ class. Now, we give an example of such a bent function, which is non-normal.

Example 2.2. Let $n = 2k = 8$. Let us define invertible $k \times k$-matrices $A_4, \dots, A_8$, which, in turn, define the partial spread $\mathcal{S}$ of order $s = 8$, given by its canonical representation (2.1):

$$A_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, A_5 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}, A_6 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, A_7 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix}, A_8 = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

The corresponding bent function $f(x) = \sum_{i=1}^{2^{k-1}} \mathbb{1}_{U_i^*}(x)$ is in the $\mathcal{PS}^- \setminus \mathcal{PS}_{ap}$ class (it is the function psf=970 in [7, psf-8.txt]). The ANF of this function is given by:

$$\begin{aligned} f(x) = {} & x_1 + x_2 + x_1x_2 + x_3 + x_1x_3 + x_2x_3 + x_1x_2x_3 + x_4 + x_1x_4 + x_2x_4 + x_1x_2x_4 + x_3x_4 \\ & + x_1x_3x_4 + x_2x_3x_4 + x_1x_2x_3x_4 + x_5 + x_1x_5 + x_1x_2x_5 + x_1x_3x_5 + x_2x_3x_5 + x_4x_5 + x_1x_4x_5 \\ & + x_2x_4x_5 + x_1x_2x_4x_5 + x_2x_3x_4x_5 + x_6 + x_1x_6 + x_2x_6 + x_3x_6 + x_1x_3x_6 + x_2x_3x_6 \\ & + x_1x_2x_3x_6 + x_1x_4x_6 + x_1x_2x_4x_6 + x_3x_4x_6 + x_1x_3x_4x_6 + x_5x_6 + x_2x_5x_6 + x_3x_5x_6 \\ & + x_2x_3x_5x_6 + x_4x_5x_6 + x_7 + x_2x_7 + x_1x_2x_7 + x_3x_7 + x_2x_3x_7 + x_2x_4x_7 + x_1x_2x_4x_7 \\ & + x_1x_3x_4x_7 + x_2x_3x_4x_7 + x_5x_7 + x_2x_5x_7 + x_1x_2x_5x_7 + x_3x_5x_7 + x_1x_3x_5x_7 + x_4x_5x_7 \\ & + x_1x_4x_5x_7 + x_2x_4x_5x_7 + x_6x_7 + x_1x_6x_7 + x_2x_6x_7 + x_3x_6x_7 + x_2x_3x_6x_7 + x_1x_4x_6x_7 \\ & + x_5x_6x_7 + x_1x_5x_6x_7 + x_2x_5x_6x_7 + x_4x_5x_6x_7 + x_8 + x_1x_8 + x_1x_2x_8 + x_4x_8 + x_1x_4x_8 \\ & + x_2x_4x_8 + x_3x_4x_8 + x_1x_3x_4x_8 + x_2x_3x_4x_8 + x_5x_8 + x_1x_2x_5x_8 + x_4x_5x_8 + x_2x_4x_5x_8 \\ & + x_6x_8 + x_1x_6x_8 + x_2x_6x_8 + x_1x_3x_6x_8 + x_4x_6x_8 + x_5x_6x_8 + x_1x_5x_6x_8 + x_4x_5x_6x_8 \\ & + x_7x_8 + x_1x_7x_8 + x_2x_7x_8 + x_1x_2x_7x_8 + x_3x_7x_8 + x_2x_3x_7x_8 + x_4x_7x_8 + x_5x_7x_8 \\ & + x_1x_5x_7x_8 + x_3x_5x_7x_8 + x_6x_7x_8 + x_1x_6x_7x_8 + x_3x_6x_7x_8 + x_5x_6x_7x_8. \end{aligned}$$

Using Algorithm 1.1, one can check that this function is non-normal. With this example, we give positive answers to both mentioned questions and also make the following conclusion (we give a short proof for completeness).

Corollary 2.3. *Let f be a non-normal bent function on $\mathbb{F}_2^n$. Then, $n \geq 8$.*

Proof. Since f is bent on $\mathbb{F}_2^n$ and $n \leq 6$, we have that f is either quadratic or cubic (the latter is only possible for $n = 6$). Every quadratic bent function f on $\mathbb{F}_2^n$ is normal, see [4, Theorem A.1]. For $n = 6$, every cubic bent function is equivalent, up to a nonsingular affine transformation on the variables, to the function $g(x, y) = g(x_1, x_2, x_3, y_1, y_2, y_3) = x \cdot \pi(y) + x_1x_2x_3$, where π is a permutation of $\mathbb{F}_2^3$, see [3, Proposition 4]. Clearly, $g|_V = 0$ for a vector space $V = \{0\} \times \mathbb{F}_2^3$, and hence $f|_{b+V'} = 0$ for some affine space $b + V'$. With Example 2.2 and Result 1.2, we conclude that non-normal bent functions exist on $\mathbb{F}_2^n$ for all even $n \geq 8$. □

Surprisingly, the function $f(x)$ in Example 2.2 is the only non-normal bent function from the list of all partial spread bent functions [7]. This function is, however, *weakly normal*, i.e., $f + l$ is normal for a non-zero linear function l on $\mathbb{F}_2^8$. Indeed, it is possible to verify that for a linear function $l(x) = x_8$ and an affine subspace $V = (0,1,0,1,0,0,0,0) + \langle (0,0,1,1,0,0,0,1), (0,0,0,0,1,0,0,1), (0,0,0,0,0,1,0,0), (0,0,0,0,0,0,1,1) \rangle$ the following holds $(f + l)|_V = 1$. With this observation, we make the following conclusion.

Result 2.4. *All $\mathcal{PS}$ bent functions in $n = 8$ variables are either normal or weakly normal.*

3 Computational construction methods of bent functions

Aimed to generate more non-normal bent functions and to find the first examples of non-weakly normal bent functions in 8 variables, we use two computational approaches for the generation of large sets of bent functions, based on the *cellular automata (CA)* and the *genetic programming (GP)*. In the following, we briefly discuss the used approaches and bent functions obtained with their help.

3.1 Generating bent functions with CA and GP

Cellular Automata (CA) can be seen as a particular kind of discrete dynamical system equipped with a shift-invariant update function that acts over a regular lattice of cells. When the state set of the cells is a finite field, and the local rule is linear, a cellular automaton can be interpreted as a linear recurring sequence (LRS). The authors of [6] studied families of LRS of order d, whose feedback polynomials are pairwise coprime. In this way, it is possible to define a partial spread by considering the projection of the LRS onto their first $2d$ coordinates. Such families exist only when the degree of the feedback polynomials is either 1 or 2. The former case corresponds to the Desarguesian spread. For degree 2, the authors of [6] found 273 $\mathcal{PS}^-$ functions of 8 variables, most of which are inequivalent to Maiorana-McFarland and Desarguesian spread-based functions. Therefore, they seem to be good candidates to test for non-normality.

Genetic Programming (GP) is an optimization algorithm loosely inspired by the principles of biological evolution. The underlying idea is to encode a Boolean function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ as a syntactic tree where the leaves represent the input variables, the internal nodes are Boolean operators (such as AND, XOR, NOT, etc.) acting on the inputs received from their children, and the root node gives the output of the function. Therefore, one can define the truth table of the function by evaluating the circuit encoded by the tree over all 2^n input combinations. The GP algorithm randomly initializes a population of trees encoding n-variables Boolean functions, then evaluates their *fitness*, which measures the optimization criterion to optimize. In our case, the fitness function is defined as the nonlinearity of the functions to be maximized. Then, the GP algorithm iteratively evolves the population by applying mutation and crossover operators, which give a new population to be evaluated against the fitness function. The fittest individuals are then carried over to the next iteration. For our problem, we employed the GP algorithm proposed in [12], adopting the same experimental settings and parameters. In particular, the GP algorithm performed 10 000 optimization runs, where in each run, a population of 50 trees encoding Boolean functions of 8 variables is evolved for 500 000 iterations.

3.2 Analysis of generated bent functions

CA. Aimed to analyze whether it is possible to generate non-normal partial spread bent functions using CA, we revised all 273 $\mathcal{PS}^-$ bent functions generated with this approach in [6]. It turned out that all partial spread bent functions constructed with this approach are normal.

Genetic Programming. With this approach, we were able to generate 7,478 different bent functions. Among them, there are 4 690 quadratic, 2 367 cubic, and 421 of degree 4. Since all quadratic bent functions and all cubic functions in 8 variables are normal, it is enough to analyze only bent functions of degree 4. We note that all the generated bent functions of degree 4 turned out to be normal as well, which was reasonable to expect since most of them have only a few monomials of degree 4. For this reason, and due to the fact that the majority of generated bent functions are quadratic and cubic (and hence are equivalent to the Maiorana-McFarland class), it was essential to check whether these bent functions of degree 4 are equivalent to the Maiorana-McFarland class. Among 421 functions of degree 4, we identified a function inequivalent to a member of the Maiorana-McFarland class (this fact was checked with the corresponding algorithms described in [1, 13]). The ANF of this function is given by

$$\begin{aligned} g(x) = {} & 1 + x_2 + x_5 + x_6 + x_8 + x_1x_5 + x_1x_7 + x_1x_8 + x_2x_6 + x_2x_7 + x_3x_8 + x_4x_7 \\ & + x_2x_5x_8 + x_1x_3x_6x_7 + x_2x_5x_7x_8. \end{aligned} \tag{3.1}$$

Again, with Algorithm 1.1, one can check that the function g given in (3.1) is normal, since $g|_V = 0$ for the affine subspace $V = (1,1,0,0,0,0,0,0) + \langle (0,0,1,0,0,0,0,0), (0,0,0,1,0,0,0,0), (0,0,0,0,1,0,0,0), (0,0,0,0,0,1,0,0)\rangle$ of dimension 4.

4 Conclusion and open problems

In this paper, we completely analyzed all partial spread bent functions in $n = 8$ variables with respect to normality, thus providing the first example of a non-normal bent function in $n = 8$ variables. The next essential step is to find (if possible) the examples of non-weakly normal bent functions on $\mathbb{F}_2^8$, as well as non-weakly normal bent functions in the $\mathcal{PS}^- \setminus \mathcal{PS}_{ap}$ class; the latter question was essentially asked by Leander in [10, p.17].

Aimed to generate more non-normal bent functions and even to find non-weakly normal ones, we used evolutionary algorithms to construct such functions. Being unable to find such examples (mostly due to the reason that we evolved only non-linearity), we still, however, were able to find bent functions, which, up to equivalence, do not belong to the Maiorana-McFarland class. This finding indicates, that using suitably chosen evolutionary algorithms (e.g., by additionally minimizing the number of flats on which a bent function is affine), it might be possible to construct "rare" bent functions.

Finally, we want to underline that future research on generating new bent functions should be focused on the construction of algorithms 1) generating bent functions outside the known classes with a high probability, 2) generating non-normal bent functions, and 3) generating non-weakly normal bent functions. We believe that based on the analysis of big sets of bent functions not coming from the known analytic constructions, it should be possible to develop generic theoretical construction methods of new families of bent functions.

References

[1] Bapić, A., Pasalic, E., Polujan, A., Pott, A.: Vectorial Boolean functions with the maximum number of bent components beyond the Nyberg's bound. Designs, Codes and Cryptography (2023). p. 4.

[2] Canteaut, A., Daum, M., Dobbertin, H., Leander, G.: Finding nonnormal bent functions. Discrete Applied Mathematics **154**(2), 202–218 (2006). p. 1.

[3] Carlet, C.: Two new classes of bent functions. In: T. Helleseth (ed.) Advances in Cryptology — EUROCRYPT '93, pp. 77–101. Springer Berlin Heidelberg, Berlin, Heidelberg (1994). p. 3.

[4] Charpin, P.: Normal Boolean functions. J. Complexity **20**(2-3), 245–265 (2004). pp. 1, 2, and 3.

[5] Dillon, J.F.: Elementary Hadamard difference sets. Ph.D. thesis, University of Maryland (1974). p. 2.

[6] Gadouleau, M., Mariot, L., Picek, S.: Bent functions in the partial spread class generated by linear recurring sequences. Designs, Codes and Cryptography **91**(1), 63–82 (2023). p. 4.

[7] Langevin, P.: Classification of partial spread functions in eight variables. Philippe Langevin's numerical project page (2010). pp. 2 and 3.

[8] Langevin, P., Hou, X.D.: Counting partial spread functions in eight variables. IEEE Transactions on Information Theory **57**, 2263–2269 (2011). pp. 1 and 2.

[9] Langevin, P., Leander, G.: Counting all bent functions in dimension eight 99270589265934370305785861242880. Designs, Codes and Cryptography **59**(1), 193–205 (2011). p. 1.

[10] Leander, G.: Normality of bent functions. Monomial- and binomial-bent functions. Ph.D thesis, Ruhr-Universität Bochum, Universitätsbibliothek (2005) pp. 1 and 5.

[11] Leander, G., McGuire, G.: Construction of bent functions from near-bent functions. Journal of Combinatorial Theory, Series A **116**(4), 960–970 (2009) p. 1.

[12] Picek, S., Jakobovic, D., Miller, J.F., Batina, L. Cupic, M.: Cryptographic Boolean functions: One output, many design criteria. Appl. Soft Comput. 40: 635-653 (2016) p. 4.

[13] Polujan, A., Pott, A.: Cubic bent functions outside the completed Maiorana-McFarland class. Designs, Codes and Cryptography **88**(9), 1701–1722 (2020). p. 4.