THE DISTANCE FROM A VECTORIAL BOOLEAN FUNCTION TO THE AFFINE FUNCTIONS: CERTIFIED VALUES FOR $n \leq 4$, AND MAXIMIZERS THAT ARE NOT APN

KOREA SUPERINTELLIGENCE LABS

Abstract. For an (n, m)-function $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ let $d_A(F)$ be the minimum Hamming distance from F to an affine function, and let $M(n) = \max\{d_A(F)\}$ over (n, n)-functions. Problem 11 of the Eighth International Olympiad in Cryptography (NSUCRYPTO'2021) asks for functions in a fixed number of variables whose distance to the affine functions is as large as possible, and the organisers' status table still records the problem as unsolved. For $n = 4$ it is not. Carlet's interpolation bound $d_A(F) \leq 2^n - n - 1$ is strict for (n, n)-functions with $n \geq 4$ — proved by Carlet, and proved again by Biryukov, Tureček and Udovenko at EUROCRYPT 2026, whose introduction names the olympiad problem — while Ryabov computes $d_A(f) = 10$ for *every* APN function in four variables and states that this value is the maximum over all $(4, 4)$-functions. So $M(4) = 10$ is in print. We state that attribution first, and then contribute two things. First, a complete machine-checked account of the range $n \leq 4$: the interpolation bound for all n and m with no hypothesis on F, the exact values $M(2) = 1$ and $M(3) = 4$ over *all* functions rather than the APN ones, an explicit $(4, 4)$-function at distance at least 10 from every one of the 2^{20} affine maps and exactly 10 from one of them, and a third proof of the matching upper bound at $n = 4$ — a finite search of 52 six-point certificates over 1 374 496 assignments, confined to the even-weight subspace, using neither character sums nor a counting argument. Second, one statement the sources do not make: at $n = 4$ the maximum is attained by functions that are *not* APN. An explicit F_w of differential uniformity 4 and image size 10 has $d_A(F_w) = 10$, while x^3 over $\mathbb{F}_{16}$ is APN, has image size 6, and reaches the same 10; with Ryabov's theorem the set of maximizers at $n = 4$ therefore contains the APN functions strictly. Every assertion below, apart from the computations explicitly labelled as lying outside the formal development, is machine-checked in Lean 4.

1. Introduction

The quantity. An (n, m)-*function* is a map $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$; it is *affine* when $F(x) = L(x) + c$ with L linear and $c \in \mathbb{F}_2^m$ constant. The Hamming distance between two (n, m)-functions is the number of inputs at which they differ,

$$d_H(F, G) = \#\{x \in \mathbb{F}_2^n : F(x) \neq G(x)\},$$

and the distance from F to the affine functions is

$$d_A(F) = \min\{d_H(F, A) : A \text{ affine}\}.$$

The object of this paper is the maximum of that minimum over (n, n)-functions,

$$M(n) = \max\{d_A(F) : F\colon \mathbb{F}_2^n \to \mathbb{F}_2^n\}.$$

The cryptographic literature works with the complementary quantity, the *vectorial linearity*

$$\mathrm{VecLin}_F = \max\{|X| : X \subseteq \mathbb{F}_2^n,\ F|_X = A|_X \text{ for some affine } A\} = 2^n - d_A(F),$$

introduced under that name in [1], so that $M(n) = 2^n - \min_F \mathrm{VecLin}_F$. For $m = 1$ the analogous maximum is the covering radius of the first-order Reed–Muller code $RM(1, n)$, a classical quantity, equal to $2^{n-1} - 2^{n/2-1}$ for even n by Rothaus' theorem on bent functions and open for odd $n > 7$; Carlet [2, §7] poses the vectorial question as the analogue still to be settled, not as the same problem.

The source, and its status. The question comes from the Eighth International Olympiad in Cryptography, NSUCRYPTO'2021, unsolved research problem 11, *"Distance to affine functions"* [3]. The problem statement recalls the definition of Hamming distance and then says that the minimum distance to the affine functions "is known to be strictly smaller than $2^n - n - 1$ if $n \geq 4$", and its third question asks

Problem 1.1 ([3, Problem 11, Q3])**.** Find constructions (possibly with a computer; then a representation of these functions will be needed) of functions F in fixed numbers of variables having a distance to affine functions as large as possible.

The solution section of [3] reports no value of $M(n)$ for any n, and the organisers' live table of unsolved problems still carried the row "2021 | Distance to affine functions | Unsolved" when we last consulted it, on 3 September 2026 [10].

What is settled, and by whom. We record the attribution before anything else, because the live status flag is misleading.

(1) *The upper bound.* Carlet [2, Prop. 6] proves $d_A(F) \leq 2^n - n - 1$ for every n and every (n, m)-function F, by interpolating F at a point and a basis translated by it. In the paragraph after his Proposition 7 he sharpens it: "Hence, for every $n \geq 4$ and every m, the inequality in Proposition 6 is in fact strict", so $d_A(F) \leq 2^n - n - 2$ for $n \geq 4$; the argument there is a sixth-moment character-sum identity. The displayed computation there is carried out for $m \leq n$, and Ryabov, reading the same paper, records the hypothesis as $m \leq 2n - 5$ or $m = n = 4$ [7, (6)]; all three readings agree on $m = n$, which is the only case used below. Biryukov, Tureček and Udovenko [1, Prop. 6] prove the same inequality again, in the form $\mathrm{VecLin}_S \geq n + 2$ for $n \geq 4$ and $m \leq 2n - 3$, by a pigeonhole count on 3-subsets producing a six-element zero-sum. Their introduction states that this improvement "was posed as an open problem at the NSUCRYPTO 2021 international cryptographic olympiad".

(2) *The value at $n \leq 5$.* Ryabov [7, §3] computes the distance to the affine functions of all APN functions in $n \leq 5$ variables, class by class of extended-affine equivalence, and finds one value in each dimension: $0, 1, 4, 10, 25$ for $n = 1, \ldots, 5$. At $n = 4$ there are two EA-classes, both of distance 10, and neither contains a permutation. His conclusion draws the consequence in so many words: "all APN functions in dimension up to 5 have the maximum possible nonlinearity among all mappings in the corresponding dimension". Nagy [6, Thm. 17] restates the table as a theorem attributed to [7, §3], and [1, §6.4] report verifying the $n \leq 5$ values with their own algorithms.

Since $2^4 - 4 - 2 = 10$, the two together give $M(4) = 10$: no $(4, 4)$-function exceeds 10 by (1), and every APN one attains it by (2). Biryukov, Tureček and Udovenko put both halves in a single sentence [1, §6.4]: "the vectorial linearity of APN functions is forced to be equal to $4, 6, 7$ respectively for $n = 3, 4, 5$ simply by the theoretical bounds." So Problem 1.1 is answered at $n = 4$ in the published record, whatever the olympiad's status table says, and nothing in this paper claims otherwise.

What this paper adds. Two things, of different kinds.

A certified account of $2 \leq n \leq 4$. Sections 3–5 give a complete formal development of the range: Carlet's interpolation bound for every n and every m with no hypothesis on F (Theorem 3.1); the exact values $M(2) = 1$ and $M(3) = 4$ over all functions, not only the APN ones (Theorem 4.1); an explicit $(4, 4)$-function at distance at least 10 from every one of the 2^{20} affine maps, with equality for one of them (Theorem 5.1); and the matching upper bound at $n = 4$ (Theorem 5.2). The mathematics of all four is known. What is new is that the $n = 4$ upper bound is obtained here by a third route, independent of both published ones: after normalising F by its own interpolant, a depth-first search over six values — six of the eight points of the even-weight subspace — closes with 52 six-point certificates, no character sums and no counting argument. The search is small enough to be checked inside a proof assistant, which is what Section 8 reports.

One new observation. At $n = 4$ the maximum is attained by functions that are not APN (Theorem 6.1). Carlet asks, in the section that poses the problem, "what are the functions which reach it (which would be the equivalent of bent functions and of almost bent functions for this notion of nonlinearity)" [2, §7]; every published answer we located describes the extremal functions as APN functions — Ryabov's theorem, in the form Nagy states it, is about "all APN functions in n variables", [1, §6.4] apply their algorithms to "several known APN extended affine (EA) equivalence classes", and [6] is framed around APN and bent functions throughout. The one place where the literature computes the parameter for non-APN functions at $n = 4$ is Ryabov's §4, and it goes the other way: all differentially 4-uniform *permutations* of $\mathbb{F}_2^4$ have distance 9, one short of the maximum. Nor do the available inequalities forbid a non-APN maximizer: the bound $\mathrm{VecLin}_S \le \sqrt{\delta_S}\, 2^{n/2} + \frac{1}{2}$ of [1, Prop. 3], attributed there to Nagy and to Ryabov independently, requires only $\delta \ge 2$ at $n = 4$. We are explicit about the risk in §6: once the question is asked, the computation that answers it takes a second, so the fact may well be folklore; what we did not find in the sources is the statement.

Larger n. Nothing here touches $n \ge 5$. For the record, the published data bracket the next values. At $n = 5$, $M(5) = 25 = 2^5 - 5 - 2$ by the same two ingredients as at $n = 4$. At $n = 6$ the bound gives $M(6) \le 56$, while Dillon's APN permutation has $d_A = 55$ [6, Thm. 17] and [1, §6.4] confirm vectorial linearity 9, i.e. $d_A = 55$, for all 534 EA-classes of cubic APN functions in six variables (Langevin's classification, cited there); so $M(6) \in \{55, 56\}$. At $n = 7$ the bound gives $M(7) \le 119$ and [6, Thm. 18(ii)] computes $d_A = 117$ for x^3 and x^9, so $117 \le M(7) \le 119$; [1, §6.4] report the same VecLin $= 11$ for most of the $n = 7$ APN functions they checked. At $n = 8$ the bound gives $M(8) \le 246$, and the largest distance we found in print is 241: an APN function of vectorial linearity 15, "the smallest known for $n = 8$ bits" [1, Appendix C]. Their discussion [1, §6.5] adds: "We recall that the proven lower bound for $n = 8$ is 10. It is thus a challenging open problem to find 8-bit S-boxes with vectorial linearity lower than 14, or to prove that they do not exist." All of these lower bounds come from APN functions. Theorem 6.1 says that at $n = 4$ the extremal set is larger than the APN functions, so a search for $M(n)$ at $n \ge 6$ has no reason to be confined to them; we make no claim about whether that helps.

2. Notation and the formal encoding

Throughout, $e_0, \dots, e_{n-1}$ is the canonical basis of $\mathbb{F}_2^n$, indexed from zero. A linear (n, m)-function L is determined by the images $b_i = L(e_i)$, and every affine (n, m)-function is $A(x) = L(x) + c$ for a unique tuple $(b_0, \dots, b_{n-1}; c)$; there are $2^{m(n+1)}$ of them, in particular 2^{20} for $n = m = 4$. We write $\mathrm{agr}(F, A)$ for the number of points at which F and A agree, so $\mathrm{agr}(F, A) + d_H(F, A) = 2^n$ and $d_A(F) = 2^n - \mathrm{VecLin}_F$.

The *differential uniformity* of F is $\delta_F = \max_{a \ne 0} \max_b \#\{x : F(x) + F(x + a) = b\}$, and F is *APN* when $\delta_F = 2$; the *image size* of F is $|F(\mathbb{F}_2^n)|$.

For the formal development we use a flat encoding, which is also how the tables below are printed. A point of $\mathbb{F}_2^n$ is a natural number $x < 2^n$, bit i of x being its i-th coordinate x_i, so that $e_i = 2^i$; a value in $\mathbb{F}_2^m$ is a natural number $< 2^m$; and addition in $\mathbb{F}_2^m$ is bitwise XOR, written $\oplus$. A linear map is a function $b\colon i \mapsto b_i$ with $\mathrm{lin}(b, x) = \bigoplus_{i<n,\, x_i=1} b_i$, and $\mathrm{aff}(b, c, x) = \mathrm{lin}(b, x) \oplus c$. A $(4, 4)$-function is displayed as the list $[F(0), F(1), \dots, F(15)]$ of 4-bit integers, and a linear part as the tuple (b_0, b_1, b_2, b_3). We identify $\mathbb{F}_{16}$ with $\mathbb{F}_2[x]/(x^4 + x + 1)$ and encode a field element by the integer whose bits are its coordinates in the basis $1, x, x^2, x^3$.

3. The interpolation bound

The first theorem is Carlet's Proposition 6, formalised for all parameters at once. Its proof is his, in one paragraph; we reproduce it because everything below rests on it and because the formal statement carries no hypothesis whatsoever: no bound on m, and no condition on F of any kind.

Theorem 3.1 (Carlet [2, Prop. 6])**.** *Let $n, m \geq 1$ and let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ be any function. Let A_F be the affine function with constant term $F(0)$ and linear part $e_i \mapsto F(e_i) + F(0)$. Then A_F agrees with F at the $n+1$ points $0, e_0, \ldots, e_{n-1}$, so*

$$d_A(F) \leq d_H(F, A_F) \leq 2^n - n - 1.$$

Proof. $A_F(0) = F(0)$ because a linear map vanishes at 0, and $A_F(e_i) = (F(e_i)+F(0))+F(0) = F(e_i)$. The $n+1$ points $0, e_0, \ldots, e_{n-1}$ are distinct, so F and A_F differ at no more than the remaining 2^n-n-1 points. □

Nothing in Theorem 3.1 is new, and [1, Prop. 2] restate it as $\mathrm{VecLin}_S \geq n+1$ with a two-line proof. We isolate it because it is the only statement in this paper that holds for every n, and because it is what makes the finite searches below finite: the interpolant A_F is also the normalisation used in §5.2.

4. The exact values at $n = 2$ and $n = 3$

Carlet's bound is attained below $n = 4$, which is the non-vacuity control for its strictness from $n = 4$ on.

Theorem 4.1. *$M(2) = 1$ and $M(3) = 4$; that is, $\max_F d_A(F) = 2^n - n - 1$ for $n = 2, 3$, the maximum being over* all *(n, n)-functions. Witnesses are*

$$\Phi_2 = [0, 1, 0, 0], \qquad \Phi_3 = [0, 7, 4, 2, 1, 0, 0, 0],$$

which satisfy $d_A(\Phi_2) = 1$ and $d_A(\Phi_3) = 4$.

Proof sketch. The upper halves are Theorem 3.1 with $n = 2$ and $n = 3$, and need no computation. The lower halves are exhaustive: $d_H(\Phi_2, A) \geq 1$ is checked against all $4^3 = 64$ affine $(2, 2)$-functions and $d_H(\Phi_3, A) \geq 4$ against all $8^4 = 4096$ affine $(3, 3)$-functions, in each case by enumerating the linear part and the constant term and evaluating the agreement count at all 2^n points. The two witnesses themselves were found outside the formal development, by exhausting every F with $F(0) = 0$ — a normalisation that costs nothing, since adding a constant to F permutes the affine functions — which is 4^3 functions at $n = 2$ and 8^7 at $n = 3$. □

5. The value at $n = 4$

5.1. A function at distance exactly ten.

Theorem 5.1. *Let*

$$F_w = [0, 12, 11, 0, 10, 2, 0, 0, 3, 4, 5, 7, 1, 2, 4, 2].$$

Then $d_H(F_w, A) \geq 10$ for every one of the 2^{20} affine functions $A\colon \mathbb{F}_2^4 \to \mathbb{F}_2^4$, with equality for the A of linear part $(3, 5, 1, 0)$ and constant term 0. Hence $d_A(F_w) = 10 = 2^4 - 4 - 2$ and $M(4) \geq 10$.

Proof sketch. Exhaustive. The $2^{20} = 1\,048\,576$ affine maps are indexed by a single integer whose four low nibbles are b_0, b_1, b_2, b_3 and whose top nibble is c; for each the agreement count over the 16 points is computed and compared with 6. The attaining map is exhibited separately, as a single evaluation. No structure of F_w is used. □

The mathematical content of Theorem 5.1 — that the sharpened bound is attained at $n = 4$ — is published: by Ryabov's computation [7, §3] every APN $(4, 4)$-function is at distance 10 from the affine functions. What the theorem adds is a witness in explicit form, as Problem 1.1 asks for, checked against every affine map rather than through the Sidon-set machinery the sources use — and, as §6 shows, a witness that is not APN and so is not an instance of the published family. The table was found by a restarted hill climb and verified twice in C and once in Python before any formal statement was written.

5.2. **The upper bound, by a third route.**

Theorem 5.2. *Every $(4,4)$-function agrees with some affine function at 6 or more of the 16 points; equivalently, $d_A(F) \le 10 = 2^4 - 4 - 2$ for every $F\colon \mathbb{F}_2^4 \to \mathbb{F}_2^4$.*

This is Carlet's strict inequality at $n = 4$, and equally [1, Prop. 6] at $n = m = 4$. The proof below is neither of theirs. We describe it in full, because the interest is how small the finite search is and because the verification section has to say exactly what was searched.

Normalisation. For a fixed affine A_0 the map $A \mapsto A + A_0$ is a bijection of the affine functions and $d_H(F, A) = d_H(F + A_0, A + A_0)$, so F and $G := F + A_0$ have the same agreement spectrum. Take $A_0 = A_F$, the interpolant of Theorem 3.1; then G vanishes at the five points $0, e_0, e_1, e_2, e_3$, i.e. at $0, 1, 2, 4, 8$ in the flat encoding. Suppose, for contradiction, that every affine map agrees with F — equivalently with G — at most 5 times.

The search space. Eleven values of G remain unknown. The search assigns only six of them, namely

$$G(3),\ G(5),\ G(6),\ G(9),\ G(10),\ G(12),$$

each ranging over the 16 elements of $\mathbb{F}_2^4$. Together with 0 these are seven of the eight points of the even-weight subspace $E = \{0, 3, 5, 6, 9, 10, 12, 15\}$: the four odd-weight points $7, 11, 13, 14$ are never assigned, and neither is $G(15)$.

The certificates. If some six-element set $T \subseteq \mathbb{F}_2^4$ carries an affine interpolant of G — an affine A with $A|_T = G|_T$ — then $\mathrm{agr}(G, A) \ge 6$, contradicting the hypothesis. A *certificate* is such a candidate set T together with (i) one left-null-space vector of the 6×5 matrix $[1 \mid x]_{x \in T}$ over $\mathbb{F}_2$, recorded as the set of value slots it involves, and (ii) the five coefficients $c, b_0, \ldots, b_3$ of the interpolant, each written as an XOR of the assigned values. A certificate *fires* on an assignment when the null-space relation is satisfied *and* the affine map reconstructed from (ii) is verified, point by point, to agree with G at all six points of T. The relation is only a filter; the six explicit agreement checks are the whole evidence, so an erroneous coefficient table can make the search prune less but can never make it unsound.

Proof sketch of Theorem 5.2. The certificates were produced outside the formal development by enumerating all $\binom{12}{6} = 924$ six-element subsets of the twelve points $\{0, 1, 2, 4, 8\} \cup E$, computing the null-space relation and the interpolation masks of each, and retaining those that fire at least once during a full search; 52 survive, and they involve only the eleven points $\{0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 12\}$. They are distributed by the depth at which they become checkable: $1, 2, 4, 7, 14, 24$ certificates at depths $0, \ldots, 5$. The formal search is a depth-first traversal of the six unknowns in the order above, testing at each level exactly the certificates of that depth and abandoning a prefix as soon as one fires. It visits 1 374 496 candidate assignments — the surviving prefixes per level are $1, 15, 210, 2520, 22680, 60480, 0$ — and no assignment survives. The bridge from that emptiness to the theorem needs two facts about the table, both decided by kernel computation: every certificate's point set is a six-element duplicate-free subset of the eleven points, and every point of a depth-d certificate either is one of the five base points, where G vanishes, or has a slot $\le d$, so that the depth-d prefix already determines its G-value. Given those, a firing certificate yields an affine map agreeing with G at six distinct points, contradicting the assumption; hence no assignment can be consistent with it, and the search's emptiness is exactly the statement that no F is farther than 10 from every affine map. □

The same search was run in C, in about a second on one core, performing 12 670 621 certificate tests, once with the relation filter alone and once with the relation filter and the full agreement check, giving identical node and work counts; and the tables were replayed in Python under exactly the semantics of the formal statement before any of them was written down.

Corollary 5.3. $M(4) = 10$.

Proof. Theorem 5.2 gives $d_A(F) \le 10$ for every $(4,4)$-function and Theorem 5.1 exhibits an F with $d_A(F) = 10$. □

As explained in §1, Corollary 5.3 restates a published value; the two theorems it combines are a formal certification of it, with a proof of the upper half that is independent of the two in the literature.

6. The maximizers at $n = 4$ are not only the APN functions

Theorem 6.1. *Let F_w be as in Theorem 5.1 and let $C\colon \mathbb{F}_{16} \to \mathbb{F}_{16}$ be $C(x) = x^3$, with $\mathbb{F}_{16} = \mathbb{F}_2[x]/(x^4 + x + 1)$, whose table in the encoding of §2 is $[0, 1, 8, 15, 12, 10, 1, 1, 10, 15, 15, 12, 8, 10, 8, 12]$. Then*

1. *$\delta_{F_w} = 4$, so F_w is not APN, its image has 10 of the 16 values, and $d_A(F_w) = 10$;*
2. *$\delta_C = 2$, so C is APN, its image has 6 of the 16 values, and $d_A(C) = 10$, the value being attained at the affine map of linear part $(4, 6, 0, 0)$ and constant term 14.*

In particular the set of $(4, 4)$-functions attaining $M(4) = 10$ contains functions of differential uniformity 2 and functions of differential uniformity 4.

Proof sketch. Each item is an exhaustive evaluation. The differential uniformities are read off the 16×16 difference distribution table; the image sizes are counted directly; and $d_A(C) = 10$ is checked, as in Theorem 5.1, against all 2^{20} affine maps, with the attaining map exhibited by a single evaluation. Part (1)'s distance claim is Theorem 5.1. □

Corollary 6.2. *The set of $(4, 4)$-functions at maximum distance from the affine functions contains the APN functions strictly.*

Proof. Every APN $(4, 4)$-function is at distance 10 from the affine functions, by [7, §3]; that inclusion is not ours. Theorem 6.1(1) exhibits a function at distance 10 that is not APN. □

Why we state this. Carlet poses the general question — "what are the functions which reach it" [2, §7] — and every published answer we located describes the extremal functions at $n \leq 5$ exclusively as APN functions. Ryabov's theorem, in the form Nagy states it, is about "all APN functions in n variables"; [1, §6.4] compute vectorial linearity for "several known APN extended affine (EA) equivalence classes"; [6] is organised around APN and bent functions throughout. The nearest computation for non-APN functions at $n = 4$ is [7, §4], which finds that the thirteen EA-classes of differentially 4-uniform mappings of $\mathbb{F}_2^4$ that contain permutations all have distance 9, and concludes that no permutation of $\mathbb{F}_2^4$ exceeds 9 — so that computation, far from exhibiting a non-APN maximizer, rules out the bijective ones. What we found in no source is the statement that some non-APN $(4, 4)$-function reaches 10. The question is not settled by the available inequalities either: [1, Prop. 3] gives $\mathrm{VecLin}_S \leq \sqrt{\delta_S} \cdot 2^{n/2} + \frac{1}{2}$, which at $n = 4$ and $\mathrm{VecLin} = 6$ requires only $\delta_S \geq 2$ and therefore excludes nothing.

Remark 6.3 (stated as a risk). Theorem 6.1 is one exhaustive computation on two explicit tables. Once the question is asked, answering it takes a second of computer time, and we cannot rule out that the fact is folklore among specialists even though it is absent from the sources we searched. What we claim as new is the observation, not its difficulty.

Remark 6.4 (outside the formal development). Non-APN maximizers are not rare. Sampling 3 000 functions normalised by $F(0) = 0$ found one with maximum affine agreement at most 6, so the density of maximizers among normalised $(4, 4)$-functions is of order 3×10^{-4} and their number of order 6×10^9 (about 6×10^{15} functions in all, the affine translate being free). A depth-first enumeration run for 600 CPU-seconds visited 70 146 386 nodes and emitted 2 562 661 maximizers before its cost cap; every one of them had differential uniformity 4. That region is a lexicographic prefix and not a random sample, so the last observation is not evidence about the global distribution of δ among maximizers, and we draw no conclusion from it. A complete census is out of reach at this size. None of the numbers in this remark is part of the formal development.

7. Controls, and one reproduction

Negative controls. A finite search that proves an emptiness is only as good as the evidence that it could have found something. Each of the following is computed by the same pipeline as the result it guards.

Proposition 7.1. (1) The distance is not trivially large. *The affine function with linear part* $(3, 9, 6, 12)$ *and constant term* 5*, whose table is* $[5, 6, 12, 15, 3, 0, 10, 9, 9, 10, 0, 3, 15, 12, 6, 5]$*, is at distance* 0 *from itself.*

(2) The claim is not too large. *It is false that* $d_H(F_w, A) \ge 11$ *for every affine* A*: the map of Theorem 5.1 attains exactly* 10*. Here* $11 = 2^4 - 4 - 1$ *is Carlet's unsharpened bound, so the witness does not reach it.*

(3) A wrong witness fails. *Changing the single value* $F_w(7)$ *from* 0 *to* 1 *produces a function at distance* 9 *from the affine map of linear part* $(6, 6, 1, 3)$ *and constant term* 0*; so Theorem 5.1 is a property of this table and not of tables of this shape.*

(4) The certificates fire. *On the normalised witness*

$$G = F_w + A_{F_w} = [0, 0, 0, 7, 0, 4, 1, 13, 0, 11, 13, 3, 8, 7, 6, 12],$$

whose six search values $\big(G(3), G(5), G(6), G(9), G(10), G(12)\big)$ *are* $(7, 4, 1, 11, 13, 8)$*, a certificate fires — as it must, since the maximum agreement of* F_w *really is* 6*. A certificate also fires on the zero function at the first level. But the test is not constant: on the assignment* $G(3) = 1$ *the first level passes.*

An S-box, and non-invariance under inversion.

Proposition 7.2 (reproducing [1, Rem. 1])**.** *Let* $S = [3, 7, 4, 8, 6, 11, 0, 5, 13, 10, 9, 1, 12, 2, 14, 15]$*, a permutation of* $\mathbb{F}_2^4$*, with inverse* $S^{-1} = [6, 11, 13, 0, 2, 7, 4, 1, 3, 10, 9, 5, 12, 8, 14, 15]$*. Then*

$$d_A(S) = 9, \qquad d_A(S^{-1}) = 8,$$

that is $\mathrm{VecLin}_S = 7$ *and* $\mathrm{VecLin}_{S^{-1}} = 8$*. The values are attained at the affine maps of linear part* $(8, 3, 0, 9)$ *with constant term* 3*, and of linear part* $(13, 11, 6, 5)$ *with constant term* 6*, respectively.*

Proof sketch. Two exhaustive sweeps over all 2^{20} affine maps, one for S and one for S^{-1}, plus the two attaining evaluations; that S^{-1} is the inverse of S is checked pointwise. □

The two numbers are those of [1, Rem. 1], where they are given to show that vectorial linearity is not preserved by inversion — a fact the authors report as resolving the question, posed in [6] and numbered Problem 2 in its journal version, of whether d_A is CCZ-invariant. The paper states the values without attaining maps; the maps above are ours, and are what makes the reproduction independently checkable. Note also that this permutation has $d_A = 9$, one *less* than the maximum $M(4) = 10$.

Remark 7.3 (not ours; the published answer). That gap is not an accident of this example. No permutation of $\mathbb{F}_2^4$ attains $d_A = 10$: Ryabov [7, §4] computes $d_A = 9$ for canonical representatives of all thirteen EA-classes of differentially 4-uniform mappings of $\mathbb{F}_2^4$ that contain permutations, and argues from the same classification that every EA-class containing a permutation of differential uniformity at least 6 has $d_A \le 9$; since $\mathbb{F}_2^4$ admits no APN permutation, 9 is the maximum of d_A over $\mathbb{F}_2^4$-permutations. Vectorial linearity is an EA-invariant [1, Prop. 5], which is what makes a class-by-class computation decide the question. We record this because we did not find it before pricing a direct attack: outside the formal development, 400 000 uniformly random 4-bit permutations gave vectorial linearity at least 7 every time — against a rate of about one in 3 000 for unrestricted $(4, 4)$-functions — and an exhaustive depth-first search, normalised by $S(0) = 0$ and $S(1) = 1$ (an output translation followed by an output linear bijection, both of which preserve bijectivity and every agreement count) and pruned incrementally over all 65 536 linear parts, covered 3 532 644 nodes in 42 CPU-minutes and entered only 7 of the several thousand depth-6 branches, which prices the full search near 700 CPU-hours. Both are consistent with Ryabov's answer, and neither was needed for it. We did recompute the distances and the differential uniformities in his Tables 7 and 8 — again outside the formal development — and every one came out as printed.

8. Verification

Every theorem and proposition stated above, and Corollary 5.3, is formally verified in Lean 4 [4] (toolchain `v4.32.0`); the development contains no `sorry` and no hand-written axiom. The two exceptions are the remarks that say so themselves — Remarks 6.4 and 7.3, whose numbers come from computations outside the development — and Corollary 6.2, which combines Theorem 6.1 with a published theorem that is not formalised here. It consists of seven modules: the encoding and its elementary lemmas; Theorem 3.1; the witness of Theorem 5.1; the certificate search of §5.2; the bridge and Corollary 5.3; Theorem 4.1; and the controls of §7. Four of the seven are written against Lean's built-in prelude alone, so that the statements they carry are equations between natural numbers, lists and booleans; the three that need finite-set reasoning — Theorem 3.1, Theorem 4.1 and the bridge — are stated over *Mathlib* [5]. The reasoning layer is kernel-clean: Theorem 3.1 for all n and m, the two structural facts about the certificate table used in the proof of Theorem 5.2, and the whole of the $n = 2$ case of Theorem 4.1 — including its 64-map sweep — are checked by the kernel and depend on no evaluation axiom. Everything that touches a 2^{20}-map sweep, the 4096-map sweep at $n = 3$, or the certificate search is delegated to compiled evaluation, and each such statement carries the standard axioms of the ambient library together with one evaluation axiom for each compiled sweep it invokes — one where a single sweep settles the statement, two where a sweep is combined with the exhibition of an attaining map, three for Corollary 5.3, which combines the search with both halves of Theorem 5.1. This covers Theorems 5.1 and 6.1, the emptiness of the search in Theorem 5.2, Propositions 7.1 and 7.2, and Corollary 5.3. Re-checking the development module by module costs about 500 CPU-seconds in total — the controls of §7, which contain three separate 2^{20}-map sweeps, account for 275 of them and the certificate search for 96 — at a peak resident set of about 7 GB, reached by two of the three *Mathlib*-importing modules. Every number reported anywhere in this paper was produced first by an independent implementation in C, and replayed in Python under the semantics of the formal statements, before being stated formally; the formal run therefore re-checks a known answer rather than searching for one. The development lives in a repository that is private at the time of writing; repository reference to be added at submission.

9. What is open

Problem 1.1 at $n \geq 6$. The first undecided value is $M(6) \in \{55, 56\}$: equivalently, is there a $(6, 6)$-function of vectorial linearity 8? We priced a direct attack and parked it. Deciding VecLin ≤ 8 for a single function by the method used here means interpolating on every affine frame of $\mathbb{F}_2^6$, of which there are $64 \cdot 63 \cdot 62 \cdot 60 \cdot 56 \cdot 48 \cdot 32/7! = 2.56 \times 10^8$, each at 64 points: about 1.6×10^{10} operations, or 30–60 CPU-minutes *per function*. Enumerating over linear parts instead is worse, 2^{36} of them. Any search that evaluates thousands of candidates is therefore two to three orders of magnitude beyond what we were willing to spend, and a route that changes this would have to bound vectorial linearity from a cheaper invariant — the gerbera configurations of [6], or the reduction algorithms of [1] — rather than from frames.

Formalising the general bound. Theorem 5.2 is a search at $n = 4$ only. Formalising [1, Prop. 6] for general n would replace it by a theorem for every $n \geq 4$; its proof is a pigeonhole count on 3-subsets together with the extension of an affinely independent set to an affine basis, and it is the second half that would be the work.

The extremal set. Corollary 6.2 says the maximizers at $n = 4$ are not only the APN functions. Which differential uniformities occur among them? Remark 6.4 reports $\delta = 4$ throughout a large but lexicographically biased sample, and $\delta = 2$ for x^3; whether $\delta = 6$ or larger occurs is a one-shot witness hunt we did not run, and the complete census is out of reach. The same question at $n = 5$, where $M(5) = 25$ is known, is untouched.

Not open: 4-bit permutations. Whether some permutation of $\mathbb{F}_2^4$ reaches $d_A = 10$ is the natural companion question, and it is settled in print: none does, by [7, §4]. Remark 7.3 records the answer and our own late arrival at it.

The best general upper bound. Ryabov's olympiad follow-up [8] collects the upper bounds into a single "universal" one, $d_A(F) \le UB(n,m)$ with $UB(n,m) = \min\{(2^m-1)2^{n-m} - 2^{n/2-m},\ 2^n - n - 1\}$, the second entry replaced by $2^n - n - 2$ when $m \le 2n-5$ or $m = n = 4$. At $m = n$ the first entry is $2^n - 1 - 2^{-n/2}$, which never binds, and the bound reduces to the one used here; we know of no published upper bound on $M(n)$ better than $2^n - n - 2$. The conjecture of Liu, Mesnager and Chen, $d_A(F) \le (1 - 2^{-m})(2^n - 2^{n/2})$, would give $M(6) \le 55$ and hence decide $M(6)$, but [6, Thm. 18] disproves it at $n = 8$ and $n = 9$.

References

[1] A. Biryukov, P. Tureček and A. Udovenko, *Algorithmic toolkit for linearization of S-boxes*, in: Advances in Cryptology — EUROCRYPT 2026, Part VI, Lecture Notes in Computer Science 16546, Springer, 2026, 243–272, doi:10.1007/978-3-032-25333-0_9. We read and cite the full version, IACR Cryptology ePrint Archive, Report 2026/913: all numbering used above — Definition 5, Propositions 2, 3, 5 and 6, Remark 1, §§6.4–6.5 and Appendix C — is that of the ePrint, and we did not check it against the proceedings version.

[2] C. Carlet, *Bounds on the nonlinearity of differentially uniform functions by means of their image set size, and on their distance to affine functions*, IEEE Trans. Inform. Theory **67** (2021), no. 12, 8325–8334, doi:10.1109/TIT.2021.3114958; preprint IACR Cryptology ePrint Archive, Report 2020/1529. We cite the numbering of the ePrint, whose Proposition 6 and the paragraph after Proposition 7 are the two statements used here.

[3] A. A. Gorodilova, N. N. Tokareva, S. V. Agievich et al., *An overview of the Eight International Olympiad in Cryptography "Non-Stop University CRYPTO"*, Sib. Èlektron. Mat. Izv. **19** (2022), no. 1, A.9–A.37, doi:10.33048/semi.2022.19.023; preprint arXiv:2204.11502v1. Unsolved research problem 11, "Distance to affine functions". We quote the arXiv version, from the LaTeX source file distributed with it.

[4] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction — CADE 28, Lecture Notes in Computer Science 12699, Springer, 2021, 625–635, doi:10.1007/978-3-030-79876-5_37.

[5] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, New York, 2020, 367–381.

[6] G. P. Nagy, *On the minimum Hamming distance between vectorial Boolean and affine functions*, Cryptogr. Commun. **17** (2025), no. 6, 1703–1720, doi:10.1007/s12095-025-00808-4; preprint arXiv:2503.03905v1. We read the preprint and cite its numbering, in which theorems, lemmas and definitions share one counter: Theorem 17 is the table attributed to [7, §3], Theorem 18 the computation at $6 \le n \le 9$. The preprint's two "Problem" environments are unnumbered; the one cited as Problem 2 in [1] is the question whether d_A is CCZ-invariant.

[7] V. G. Ryabov, *Nonlinearity of APN functions: comparative analysis and estimates*, Prikl. Diskretn. Mat. 2023, no. 61, 15–27, doi:10.17223/20710410/61/2. Read here in the original (in English), from the open-access full text at `mathnet.ru/eng/pdm810`.

[8] V. G. Ryabov, *Distance between vectorial Boolean functions and affine analogues (following the Eighth International Olympiad in Cryptography)* (in Russian), Mat. Vopr. Kriptogr. **15** (2024), no. 1, 127–142, doi:10.4213/mvk465. Read here in the original, from the open-access full text at `mathnet.ru/eng/mvk466`.

[9] O. S. Rothaus, *On "bent" functions*, J. Combin. Theory Ser. A **20** (1976), no. 3, 300–305, doi:10.1016/0097-3165(76)90024-8.

[10] Non-Stop University CRYPTO, *Unsolved problems*, `nsucrypto.nsu.ru/unsolved-problems`, consulted 3 September 2026: the row "2021 – Distance to affine functions – Unsolved".