A $(6, 4)$ Vectorial Boolean Function With Nonlinearity 26 and the Maximum Nonlinearity for Six-Bit Permutations

Yufei Yuan[1,2(✉)], Lei Zhang[1,3], and Wenling Wu[1,4(✉)]

[1]Trusted Computing and Information Assurance Laboratory,
Institute of Software, Chinese Academy of Sciences, Beijing 100190, China

[2]University of Chinese Academy of Sciences, Beijing 100049, China

[3]State Key Laboratory of Cryptology, P.O. Box 5159, Beijing 100878, China

[4]Zhongguancun Laboratory, Beijing 100194, China

[(✉)]Corresponding authors:
yufei2021@iscas.ac.cn; wenling@iscas.ac.cn

Abstract

When the output dimension of a vectorial Boolean function exceeds half its input dimension, not all nonzero components can be bent. The best attainable componentwise nonlinearity in this range, however, is generally unknown. We ask whether the conjectured bound for even-dimensional square mappings extends to this high-output regime, and show that it does not. Specifically, we construct a six-input, four-output function with nonlinearity 26, thereby improving the previous lower bound of 24. Its seven bent components form the nonzero part of a three-dimensional component subspace, whereas the remaining eight components all have maximum absolute Walsh coefficient 12. Accordingly, the associated binary linear code has length 64, dimension 11, and minimum distance 26. We then address the distinct problem of six-bit permutations and determine its exact maximum. A computer-assisted evaluation of the complete classification of Boolean functions in six variables bounds the autocorrelation energy of every balanced component whose Walsh coefficients have magnitude at most 12. Combined with a vectorial fourth-moment identity, this bound forces every six-bit permutation to have a component with maximum absolute Walsh coefficient at least 16, and hence nonlinearity at most 24. Inversion

over the field with 64 elements attains this value. Finally, the same argument gives necessary coding conditions for any non-bijective six-input, six-output function whose nonlinearity exceeds 24.

Index Terms

Bent functions, Boolean functions, covering radius, nonlinearity, Reed–Muller codes, S-boxes, vectorial Boolean functions, Walsh transform.

I. Introduction

For a Boolean function, nonlinearity measures the distance from the affine functions. The vectorial notion takes the minimum of this distance over all nonzero components. Consequently, the associated extremal problem asks for a component space in which every nonzero function remains far from every affine Boolean function. In coding terms, this is equivalent to finding a binary linear code of large minimum distance that contains the first-order Reed–Muller code [1]–[3].

For $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$, put

$$\mathrm{NL}(F) = \min_{0 \neq b \in \mathbb{F}_2^m} \mathrm{NL}(b \cdot F), \qquad N(n,m) = \max_{F} \mathrm{NL}(F). \tag{1}$$

When $n = 2k$ and $m \leq k$, vectorial bent functions attain the Boolean covering-radius bound. Once $m > k$, however, Nyberg's theorem excludes such functions [4]. The resulting high-output regime is

$$k < m < 2k, \tag{2}$$

where individual components may be bent even though the entire component space cannot be. Parseval gives a baseline Walsh peak of at least 2^k for each Boolean function. In the square case $m = n = 2k$, the stronger conjecture of Wadayama *et al.* states that $N(2k, 2k) = 2^{2k-1} - 2^k$ [5]; equivalently, some nonzero component must have Walsh peak at least 2^{k+1}. This spectral formulation suggests asking whether the same factor-of-two jump persists throughout (2).

Write

$$\Lambda(F) = \max_{\substack{a \in \mathbb{F}_2^{2k} \\ 0 \neq b \in \mathbb{F}_2^m}} |W_F(a,b)|. \tag{3}$$

With this notation, the proposed extension asks whether every function in (2) satisfies

$$\Lambda(F) \geq 2^{k+1}, \qquad \mathrm{NL}(F) \leq 2^{2k-1} - 2^k. \tag{4}$$

Although (4) holds for two broad structured classes, it is false in general. For quadratic functions, the bound follows from the rank of the associated polar forms. For functions with the maximum number of bent components, it instead follows from a divisibility theorem controlled by the span of the non-bent masks. We disprove the unrestricted claim with an explicit $(6, 4)$ function.

At $(n, m) = (6, 4)$, Nyberg's construction gives the lower bound 24 [5], [6], whereas the linear-code analysis of Beelen and Leander gives the upper bound 27 and the exact adjacent value $N(4, 3) = 4$ [2]. Hence $(6, 4)$ is the first unresolved boundary parameter after $(4, 3)$, with published interval

$$24 \leq N(6, 4) \leq 27 \tag{5}$$

By extending the representative F_7^3 in the EA classification of $(6, 3)$-bent functions [7], our construction raises the lower endpoint to 26 and yields

$$\mathrm{RM}(1, 6) \subset C_G = [64, 10, 28] \subset C_F = [64, 11, 26]. \tag{6}$$

Because the added word is at distance 26 from C_G, the covering radius of this maximal bent code is either 26 or 27.

The failure of (4) at output dimension four does not settle the permutation problem, because every nonzero component of a six-bit permutation is balanced. Although several families of even-dimensional permutations attain $2^{n-1} - 2^{n/2}$, the literature reviewed in Section II records this only as the best known value; we found no corresponding upper-bound proof for six-bit permutations. We prove that 24 is the exact maximum. To do so, we replace the codimension and divisibility argument with a fourth-moment identity that couples all components through the derivative distribution and the weight-4 words of an associated dual code.

The main results are as follows.

1) A projection argument reduces any bound that depends only on n and m, in the range $m > k$, to the boundary case $m = k + 1$.
2) A codimension-to-divisibility theorem gives a hierarchy of bounds in terms of the span of the non-bent component masks. When combined with a separate rank argument for quadratic components, it proves (4) for two broad structured classes.
3) An explicit $(6, 4)$ construction has seven bent components and an affine coset of eight even-weight non-bent components with a common four-valued Walsh distribution. Its vectorial nonlinearity is 26.

4) Shifted Walsh identities and derivative-image constraints characterize part of the joint structure required by any function satisfying $\Lambda(F) < 2^{k+1}$.
5) Finally, a scalar classification bound and a vectorial autocorrelation identity prove the exact permutation maximum. The same identity also restricts any non-bijective 6-to-6 map that could exceed 24.

The construction and the permutation theorem both use the correspondence between component Walsh spectra and Reed–Muller supercodes, but neither follows formally from the other. The former concerns a coset of minimum weight 26 for a specified maximal bent code, whereas the latter depends on balanced components and their coupled fourth moments.

The remainder of the paper is organized as follows. Section II places both problems in the literature, and Section III gives their Walsh and coding formulations. Section IV derives bounds from the span of the non-bent masks; Sections V and VI then apply these bounds to the new construction. After Section VII develops the scalar and vectorial fourth-moment identities, Sections VIII and IX use them to study six-output maps.

II. Relation to Previous Work

The two problems treated in this paper have related spectral formulations but different prior bounds. We therefore discuss first the high-output extension of the square-case conjecture, then the associated covering-radius question, and finally the permutation problem. Table I summarizes the comparison and states the scope of each result.

A. Bounds beyond Nyberg's threshold

The proposed extension (4) originates from the conjecture $N(2k, 2k) = 2^{2k-1} - 2^k$ for even-dimensional square mappings [5]. Bracken and Leander use the equivalent spectral formulation in their discussion of highly nonlinear even-dimensional permutations [8]. Whereas that conjecture concerns $m = 2k$, (4) asks the analogous question for $k < m < 2k$; it is not a previously established bound in this wider range.

Boolean Parseval gives the covering-radius bound

$$\mathrm{NL}(F) \le 2^{n-1} - 2^{n/2-1} \tag{7}$$

for even n, but Nyberg's theorem makes equality impossible once $m > n/2$. To refine this general estimate, Carlet and Ding combined imbalance and coding bounds [1]. Later estimates incorporate

component structure, image-set size, distance to affine mappings, or differential uniformity [9], [10]. Although these results improve (7) when additional information about F is available, none forces the spectral jump in (4).

Known constructions also clarify which parameter ranges remain relevant. The five-valued-spectrum families of Zhang, Song, and Pasalic lie at $m = n/2 + 1$ and have component Walsh values in $\{0, \pm 2^{n/2}, \pm 2^{n/2+1}\}$ [11], so they attain the scale $2^{n/2+1}$. Earlier balanced and resilient constructions can exceed the corresponding nonlinearity threshold only when the output dimension is below $n/2$ [12], [13]. They therefore do not address (2), in which a common bent component space is impossible.

A separate literature measures distance to affine vectorial maps and calls that quantity vectorial nonlinearity [14]–[17]. In contrast, throughout this paper nonlinearity means the minimum Boolean nonlinearity of the nonzero components, as in (1). Because the two metrics define different extremal problems, results for the former do not directly bound the latter.

Upper bounds can also be derived from the number of bent components. Pott *et al.* treated the square-output case, while Zheng *et al.* proved that a $(2k, m)$-function with $m \geq k$ has at most

$$2^m - 2^{m-k} \tag{8}$$

bent components. Equality defines the maximum-number-of-bent-components (MNBC) class [18], [19]. Within this class, equivalence properties, six-variable classifications, and nonlinearity bounds are known for several plateaued and non-plateaued subclasses [20]–[23]. These results do not cover our $(6, 4)$ function, which has seven bent components rather than the maximum of fourteen.

B. Extensions of vectorial bent codes and covering radii

For a vectorial bent mapping $G : \mathbb{F}_2^{2k} \to \mathbb{F}_2^k$, the code

$$C_G = \mathrm{RM}(1, 2k) + \langle G_1, \ldots, G_k \rangle \tag{9}$$

has parameters

$$[2^{2k}, 3k + 1, 2^{2k-1} - 2^{k-1}]. \tag{10}$$

Adding a Boolean coordinate to G is equivalent to adjoining a coset of C_G, so extendability and covering radius are two formulations of the same problem. Meidl, Polujan, and Pott accordingly asked how the covering radius depends on the maximal bent mapping [24, Question 5.18]. If (4)

TABLE I

COMPARISON WITH PREVIOUSLY ESTABLISHED BOUNDS

Problem	Previous result	This paper	Method	Scope
High-output mappings, $n = 2k$, $k < m < 2k$	Square-case conjecture at $m = 2k$; general bounds do not imply (4) [1], [5]	The extension is false in general; valid for quadratic mappings and the stated codimension classes	Projection, rank, Walsh divisibility; counterexample	All mappings (negative); stated subclasses (positive)
$(6, 4)$ mappings	$24 \leq N(6, 4) \leq 27$ [2], [5], [6]	$26 \leq N(6, 4) \leq 27$	Walsh spectra and the associated binary code	All six-to-four maps
Six-bit permutations	Nonlinearity 24 attained; no upper-bound proof identified in the reviewed literature	Exact maximum 24	Complete scalar classification; vectorial fourth-moment identity	Six-bit permutations
Unrestricted $(6, 6)$ mappings	$24 \leq N(6, 6) \leq 26$ [2], [5]	Interval unchanged; necessary nested-code conditions above 24	Fourth moments; dual weight-4 words	All six-to-six maps; the non-bijective case remains open

were valid, it would imply $\rho(C_G) \leq 2^{2k-1} - 2^k$. Our construction disproves this implication for the code associated with F_7^3 and places its radius in $\{26, 27\}$. Although unrestricted $[64, 11, 26]$ codes were already known [25], the prescribed chain (6) is new: it identifies a distance-26 coset of a specific maximal bent code. Thus the result concerns a covering radius rather than merely the existence of a linear code with these parameters.

C. Even-dimensional permutations

Several power and butterfly permutations in even dimension attain $2^{n-1} - 2^{n/2}$. Carlet's survey records this as the best known nonlinearity and lists families that attain it [26, Sec. 3.1.7]; later construction papers retain the same status [8], [27]–[30]. For unrestricted $(6, 6)$ maps, however, general coding bounds leave the interval $24 \leq N(6, 6) \leq 26$ [2], [5]. Moreover, in the literature reviewed for this submission, we found no published proof that 24 upper-bounds all six-bit permutations. It was therefore an attained value, not a proved permutation optimum. By coupling all 63 balanced components, our argument resolves the permutation problem without determining $N(6, 6)$ for unrestricted maps.

III. PRELIMINARIES

We first fix the spectral and coding notation used throughout the paper. Put $V_n = \mathbb{F}_2^n$ and use the standard inner product. For $F : V_n \to V_m$, define its vectorial Walsh transform by

$$W_F(a,b) = \sum_{x \in V_n} (-1)^{a\cdot x + b\cdot F(x)}. \tag{11}$$

For a Boolean function $f : V_n \to \mathbb{F}_2$, the corresponding definitions are

$$W_f(a) = \sum_{x \in V_n} (-1)^{f(x)+a\cdot x} \tag{12}$$

and

$$\mathrm{NL}(f) = 2^{n-1} - \frac{1}{2}\max_{a\in V_n} |W_f(a)|. \tag{13}$$

Consequently, (1) is equivalent to

$$\mathrm{NL}(F) = 2^{n-1} - \frac{1}{2}\Lambda(F). \tag{14}$$

We shall also use the Parseval identity

$$\sum_{a\in V_n} W_f(a)^2 = 2^{2n}. \tag{15}$$

In addition to its second moment, the Walsh spectrum retains the parity of the truth-table weight. Indeed, for $n \geq 2$, $\mathrm{wt}(f + a\cdot x) \equiv \mathrm{wt}(f) \pmod 2$, and therefore

$$W_f(a) \equiv \begin{cases} 0 \pmod 4, & \mathrm{wt}(f) \text{ even}, \\ 2 \pmod 4, & \mathrm{wt}(f) \text{ odd}. \end{cases} \tag{16}$$

For a vectorial function $F : V_n \to V_m$, these component parities form the linear functional

$$\eta_F(b) = \mathrm{wt}(b\cdot F) \pmod 2 = b \cdot \sum_{x\in V_n} F(x). \tag{17}$$

When n is even, a Boolean function f is bent if $|W_f(a)| = 2^{n/2}$ for every $a \in V_n$. Accordingly, a vectorial function is bent if each of its nonzero components is bent [4], [31].

The same component spectra have a coding interpretation. Associate with F the code

$$\begin{aligned} C_F = \{&(a\cdot x + b\cdot F(x) + \epsilon)_{x\in V_n} : \\ &(a,b,\epsilon) \in V_n \times V_m \times \mathbb{F}_2\}. \end{aligned} \tag{18}$$

Provided that F has no nonzero affine component, this code satisfies

$$\dim C_F = n + m + 1, \qquad d(C_F) = \mathrm{NL}(F),$$

and $\mathrm{RM}(1,n) \subset C_F$ [1], [3]. For later use, let $B_4(C)$ denote the number of weight-4 words in a binary linear code C. Also put $d(v,C) = \min_{c\in C} d_{\mathrm{H}}(v,c)$ for $v \in \mathbb{F}_2^{2^n}$; then the covering radius of C is

$$\rho(C) = \max_{v\in\mathbb{F}_2^{2^n}} d(v,C).$$

A. Reduction to the boundary output dimension

Before deriving structural bounds, we reduce the high-output regime to its smallest output dimension.

Proposition 1 (Projection reduction): Let $F : V_{2k} \to V_m$ with $m > k+1$, and let $\pi : V_m \to V_{k+1}$ be surjective. Then

$$\Lambda(F) \geq \Lambda(\pi \circ F), \qquad \mathrm{NL}(F) \leq \mathrm{NL}(\pi \circ F). \tag{19}$$

Consequently, an upper bound that depends only on n and m for all $(2k, k+1)$-functions extends to every output dimension $m > k$.

Proof: Every nonzero component of $\pi \circ F$ is a component of F indexed by a nonzero mask in $\mathrm{im}(\pi^*)$. Therefore, the maximum defining $\Lambda(\pi \circ F)$ is taken over a subset of the masks used for $\Lambda(F)$. ■

It therefore remains to understand the boundary case $m = k+1$. A function without nonzero affine components then gives a binary $[2^{2k}, 3k+2]$ supercode of $\mathrm{RM}(1,2k)$, and (4) is equivalent to

$$d(C_F) \leq 2^{2k-1} - 2^k. \tag{20}$$

If an affine component is present, then $\mathrm{NL}(F) = 0$, so the bound is immediate. Otherwise, the problem reduces to a Reed–Muller supercode with a prescribed Hadamard subcode rather than an unrestricted coding problem.

IV. Structural Bounds from Component Masks

We now derive the structural bounds that explain why the conjectured jump holds for several important classes. The first argument uses quadratic rank; the second tracks how the non-bent masks span the output space.

A. Quadratic functions

Proposition 2 (Quadratic case): Let $F : V_{2k} \to V_m$ be quadratic and $m > k$. Then

$$\Lambda(F) \geq 2^{k+1}, \qquad \mathrm{NL}(F) \leq 2^{2k-1} - 2^k. \tag{21}$$

Proof: Nyberg's dimension bound supplies a non-bent component $f = b \cdot F$. Let s denote the dimension of the radical of the alternating polar form of f. Every nonzero Walsh coefficient of a quadratic Boolean function on V_{2k} then has magnitude $2^{k+s/2}$, while bentness is equivalent to $s = 0$. Since an alternating form has even rank, non-bentness implies $s \geq 2$ and therefore produces a coefficient of magnitude at least 2^{k+1}. ■

B. Codimension and Walsh divisibility

Define

$$\begin{aligned} \mathcal{S}_F &= \{0\} \cup \{b \in V_m : b \cdot F \text{ is not bent}\}, \\ H_F &= \mathrm{span}(\mathcal{S}_F), \end{aligned} \tag{22}$$

and put $r = m - \dim H_F$.

The codimension r measures how far the non-bent masks are from spanning the whole output space. It is this codimension that controls the divisibility of the remaining Walsh coefficients.

Theorem 1 (Codimension-to-divisibility): Let $F : V_{2k} \to V_m$. For every $a \in V_{2k}$ and $b \in H_F$,

$$2^{r+1} \mid W_F(a, b). \tag{23}$$

Proof: After an invertible output transformation, write

$$V_m = V_{m-r} \oplus V_r, \qquad F = (G, Z), \qquad H_F = V_{m-r} \oplus \{0\}. \tag{24}$$

For $d \neq 0$, every component $(c, d) \cdot F$ is bent, so

$$W_F(a, (c, d)) = \varepsilon_{a,c,d} 2^k, \qquad \varepsilon_{a,c,d} \in \{\pm 1\}. \tag{25}$$

For $z \in V_r$, define the signed fiber sum

$$T_{a,c}(z) = \sum_{Z(x)=z} (-1)^{a \cdot x + c \cdot G(x)}. \tag{26}$$

Fourier inversion in the last output coordinate consequently gives

$$\begin{aligned} 2^r T_{a,c}(z) = {} & W_F(a, (c, 0)) \\ & + 2^k \sum_{d \neq 0} \varepsilon_{a,c,d} (-1)^{d \cdot z}. \end{aligned} \tag{27}$$

The map Z is vectorial bent, so $r \le k$. To determine the parity of the fibers, let $N_z = |Z^{-1}(z)|$. A second Fourier inversion yields

$$2^r N_z = 2^{2k} + 2^k Q_z, \tag{28}$$

where Q_z is a sum of $2^r - 1$ signs. For $r \ge 1$, Q_z is odd; hence N_z is even when $r < k$ and odd when $r = k$. Since $T_{a,c}(z) \equiv N_z \pmod 2$, the signed fiber sum has the same parity.

We now distinguish the two possible values of r. If $r < k$, the left side of (27) and its final term are both divisible by 2^{r+1}. If $r = k$, division of (27) by 2^k gives

$$\frac{W_F(a,(c,0))}{2^k} = T_{a,c}(z) - S_{a,c}(z),$$

where $S_{a,c}(z)$ denotes the sign sum in (27). Both terms on the right are odd, so their difference is even. Thus in either case $W_F(a,(c,0))$ is divisible by 2^{r+1}. When $r = 0$, the assertion reduces to the usual evenness of Boolean Walsh coefficients. ■

Corollary 1: If $m > k$, then

$$\Lambda(F) \ge 2^k + \min\{2^{r+1}, 2^k\}, \tag{29}$$

$$\mathrm{NL}(F) \le 2^{2k-1} - 2^{k-1} - \min\{2^r, 2^{k-1}\}. \tag{30}$$

In particular, (4) holds whenever $r \ge k - 1$.

Proof: At least one nonzero component is non-bent, so Parseval gives a Walsh coefficient with magnitude greater than 2^k. Theorem 1 requires this magnitude to be a multiple of 2^{r+1}. Because $r \le k$, the smallest permitted magnitude above 2^k is the first expression in the corollary. Substitution into (14) gives the second. ■

For the maximum-number-of-bent-components class, Zheng *et al.* proved that $\mathcal{S}_F$ is an $(m-k)$-dimensional subspace [19]. Consequently, $r = k$ in this class.

Corollary 2 (Maximum number of bent components): If $m > k$ and $F : V_{2k} \to V_m$ has $2^m - 2^{m-k}$ bent components, then

$$\mathrm{NL}(F) \le 2^{2k-1} - 2^k. \tag{31}$$

C. Affine component cosets at the boundary

The preceding codimension bound is strongest when the non-bent masks leave a nontrivial quotient. At the boundary $m = k+1$, a complementary description in terms of an affine component coset is more useful.

Assume $k \geq 2$ and $m = k + 1$, and suppose that a k-dimensional mask subspace $\mathcal{E}$ consists of zero and bent components. After an output transformation, write

$$F = (G, h), \qquad G : V_{2k} \to V_k \text{ vectorial bent.} \tag{32}$$

The complementary mask hyperplane then indexes the affine coset of component functions

$$h + \langle G_1, \ldots, G_k \rangle = \{h + c \cdot G : c \in V_k\}. \tag{33}$$

Bent functions on V_{2k} have even weight for $k \geq 2$. Because η_F is linear, all functions in (33) therefore have the same weight parity. If $\eta_F = 0$, they are even and all their Walsh coefficients are divisible by 4. If $\eta_F \neq 0$ and $\mathcal{E} = \ker \eta_F$, they are odd and every Walsh coefficient is congruent to 2 modulo 4.

If every function in this affine coset is non-bent, then

$$\mathcal{S}_F = \{0\} \cup (e_{k+1} + \mathcal{E}), \qquad \operatorname{span}(\mathcal{S}_F) = V_{k+1}. \tag{34}$$

Thus $r = 0$ in Theorem 1, regardless of the common weight parity.

Proposition 3 (One-coordinate extension formula): Let $G : V_6 \to V_3$ be vectorial bent, let $h : V_6 \to \mathbb{F}_2$, and put $F = (G, h)$. For

$$C_G = \mathrm{RM}(1, 6) + \langle G_1, G_2, G_3 \rangle,$$

one has

$$\mathrm{NL}(F) = \min\{28, d(h, C_G)\}, \tag{35}$$

where

$$d(h, C_G) = 32 - \frac{1}{2} \max_{a \in V_6,\ c \in V_3} |W_{h + c \cdot G}(a)|. \tag{36}$$

Proof: The seven nonzero components of G have nonlinearity 28, whereas the other eight components are $h + c \cdot G$. The minimum distance of the latter family from the affine functions is exactly the distance of h from C_G. Taking the minimum over the two families proves (35). ■

V. A $(6, 4)$ Function of Nonlinearity 26

A. Construction rationale

Proposition 3 reduces the $(6, 4)$ problem to a one-coordinate extension of a maximal vectorial bent function. Once G is fixed, the $[64, 10, 28]$ code C_G is fixed as well, and the task is to find a coset of large minimum weight. If the coset functions have even weight, their Walsh coefficients

TABLE II

WALSH DISTRIBUTIONS OF THE NONZERO COMPONENTS. PARENTHESES GIVE MULTIPLICITIES.

Masks	Walsh distribution	$\max_a \lvert W_{F_b}(a)\rvert$	Component weights
$1 \le b \le 7$	-8 (28), 8 (36)	8	$28, 28, 28, 28, 36, 28, 28$
$8 \le b \le 15$	-12 (10), -4 (18), 4 (22), 12 (14)	12	$30, 26, 26, 26, 26, 26, 34, 34$

are divisible by 4. A non-bent component must nevertheless have a coefficient larger than 8, so the first possible maximum is 12; an even coset attaining this value has distance 26 from C_G.

For $x = (x_1, \ldots, x_6) \in V_6$, index its truth-table position by

$$j(x) = x_1 + 2x_2 + 4x_3 + 8x_4 + 16x_5 + 32x_6.$$

Let $G = (g_1, g_2, g_3) : V_6 \to V_3$ be given by

$$g_1 = x_1x_4 + x_2x_5 + x_3x_6 + x_1x_2x_3, \tag{37}$$

$$g_2 = x_1x_3 + x_1x_5 + x_2x_3 + x_4x_6 + x_1x_2x_4, \tag{38}$$

$$\begin{aligned} g_3 = {} & x_1x_2 + x_1x_3 + x_2x_4 + x_2x_5 + x_3x_5 + x_4x_5 \\ & + x_5x_6 + x_1x_2x_5 + x_3x_4x_5. \end{aligned} \tag{39}$$

All sums in this section are over $\mathbb{F}_2$. The resulting G is the representative F_7^3 in the EA classification of $(6, 3)$-bent functions [7]; in particular, every nonzero component of G is bent.

Define $p : V_6 \to \mathbb{F}_2$ by

$$\begin{aligned} p = {} & x_1 + x_2 + x_2x_3 + x_1x_4 + x_1x_2x_4 + x_3x_4 + x_1x_5 \\ & + x_2x_3x_5 + x_1x_4x_5 + x_2x_4x_5 + x_3x_4x_5 \\ & + x_1x_3x_4x_5 + x_1x_6 + x_2x_6 + x_3x_6. \end{aligned} \tag{40}$$

We identify a mask $b = (b_1, b_2, b_3, b_4)$ with $b_1 + 2b_2 + 4b_3 + 8b_4$ and put

$$F = (g_1, g_2, g_3, p) : V_6 \to V_4. \tag{41}$$

Theorem 2: The function in (41) satisfies

$$\Lambda(F) = 12, \qquad \mathrm{NL}(F) = 26. \tag{42}$$

Its associated code has parameters $[64, 11, 26]$ and weight enumerator

$$W_{C_F}(z) = 1 + 192z^{26} + 448z^{28} + 320z^{30} + 126z^{32} + 320z^{34} + 448z^{36} + 192z^{38} + z^{64}. \tag{43}$$

Proof: Direct evaluation of the Walsh transforms of the ANFs (37)–(40) gives Table II. The first row therefore shows that the seven components indexed by $1 \le b \le 7$ are bent, while the remaining eight have the distribution in the second row. It follows that $\Lambda(F) = 12$ and, by (14), $\mathrm{NL}(F) = 26$. Since no nonzero component is affine, $\dim C_F = 11$. Finally, taking every affine shift of each component spectrum gives the weight enumerator. The supplementary material contains the exact Walsh and code enumerations. ■

The multiplicities in the second row sum to 64, and their squared moment is

$$24 \cdot 12^2 + 40 \cdot 4^2 = 4096.$$

Thus the common distribution satisfies Parseval. The listed component weights also show that $\eta_F = 0$, equivalently $\sum_x F(x) = 0$.

B. Mask geometry and joint spectra

Let $\mathcal{E} = V_3 \times \{0\} \le V_4$. The masks in $\mathcal{E} \setminus \{0\}$ index the seven bent components, whereas $e_4 + \mathcal{E}$ indexes the eight non-bent components. Consequently,

$$\mathcal{S}_F = \{0\} \cup (e_4 + \mathcal{E}), \qquad \mathrm{span}(\mathcal{S}_F) = V_4. \tag{44}$$

Theorem 1 therefore has $r = 0$ and supplies no enhanced divisibility. Weight parity still gives divisibility by 4, but the common maximum 12 remains below $16 = 2^{k+1}$. Thus (4) is false for this function. The example is nonquadratic because p has degree 4; it is also outside the MNBC and plateaued classes, since it is not an MNBC function and its non-bent components are not plateaued.

The eight spectra share additional shifted-correlation structure, as the next identity shows.

Proposition 4 (Shifted Walsh identity): For Boolean functions f, g on V_n and every $u \in V_n$,

$$\sum_{a \in V_n} W_f(a) W_g(a + u) = 2^n W_{f+g}(u). \tag{45}$$

Proof: Expand both Walsh transforms and sum first over a. Character orthogonality forces the two input variables to be equal, leaving the Walsh transform of $f + g$ at u. ■

Put $f_c = p + c \cdot G$. If $c \neq d$, then $f_c + f_d = (c+d) \cdot G$ is bent. The shifted identity therefore gives

$$\left| \sum_a W_{f_c}(a) W_{f_d}(a+u) \right| = 512 \tag{46}$$

for every u. When $c = d$, the shifted autocorrelation is 4096 at zero and zero elsewhere. Equations (45) and (46) consequently prescribe every shifted cross-correlation among the eight spectra in Table II.

This correlation structure also disproves the claim that every maximal vectorial bent $G : V_{2k} \to V_k$ and every Boolean h must satisfy

$$\max_{a,c} |W_{h+c\cdot G}(a)| \geq 2^{k+1}. \tag{47}$$

For the class F_7^3, the left side can instead equal 12.

C. Code parameters and covering radius

Proposition 5: Let

$$C_G = \mathrm{RM}(1,6) + \langle g_1, g_2, g_3 \rangle.$$

Then

$$C_G = [64, 10, 28], \qquad d(p, C_G) = 26, \tag{48}$$

$C_F = C_G + \langle p \rangle$, and

$$26 \leq \rho(C_G) \leq 27. \tag{49}$$

Proof: Because every nonzero component of G is bent, C_G has the stated parameters. Proposition 3, together with Table II, gives

$$d(p, C_G) = \min_{c \in V_3} \mathrm{NL}(p + c \cdot G) = 26.$$

It follows that $\rho(C_G) \geq 26$. If $\rho(C_G) \geq 28$, a word at distance at least 28 could be adjoined to C_G. The resulting $[64, 11, 28]$ supercode of $\mathrm{RM}(1,6)$ would define a vectorial bent map $V_6 \to V_4$, contrary to Nyberg's bound. Hence $\rho(C_G) \leq 27$. ■

Combining the construction with the established upper bound gives

$$26 \leq N(6,4) \leq 27, \tag{50}$$

where the upper bound is the established linear-code bound [2]. The prescribed structure in (6), rather than the unrestricted parameters $[64, 11, 26]$, is the distinguishing feature of the code.

Corollary 3 (Maps with even-weight components): Every $F : V_6 \to V_4$ satisfying $\sum_x F(x) = 0$ has $\mathrm{NL}(F) \leq 26$. The function in (41) attains equality.

Proof: All nonzero components have even weight. If $\mathrm{NL}(F) \geq 27$, their Walsh coefficients have magnitude at most 10; divisibility by 4 therefore reduces the maximum magnitude to at most 8. Parseval then forces every nonzero component to be bent, which would give a vectorial bent map $V_6 \to V_4$. Nyberg's theorem rules this out. ■

Proposition 6 (A characterization of $N(6,4) = 27$): The equality $N(6,4) = 27$ holds if and only if there are a vectorial bent map $G : V_6 \to V_3$ and an odd-weight Boolean function h such that

$$d(h, C_G) = 27. \tag{51}$$

Proof: Suppose $F : V_6 \to V_4$ has nonlinearity 27. Corollary 3 then gives $\eta_F \neq 0$, so its kernel has dimension 3. Every nonzero component in the kernel has even weight and Walsh magnitude at most 10. Divisibility by 4 reduces the maximum to 8, and Parseval consequently makes all seven components bent. After an output transformation, $F = (G, h)$ with G vectorial bent and h odd. Proposition 3 gives $d(h, C_G) \geq 27$. The distance cannot be 28, because that would produce a vectorial bent map with four outputs; hence it is 27.

Conversely, (51) and Proposition 3 give $\mathrm{NL}(G, h) = 27$. ■

VI. Further Structural Consequences

A. General upper bounds

The preceding construction shows that the factor-of-two jump can fail, but elementary parity still gives a strict bound for every high-output map.

Proposition 7: Let $F : V_{2k} \to V_m$ with $m > k$. Then

$$\mathrm{NL}(F) \leq 2^{2k-1} - 2^{k-1} - 1. \tag{52}$$

If $k \geq 2$ and $\dim \ker \eta_F > k$, then

$$\mathrm{NL}(F) \leq 2^{2k-1} - 2^{k-1} - 2. \tag{53}$$

The second condition holds when $\eta_F = 0$ or when $m \geq k + 2$.

Proof: Some component is non-bent, and Parseval therefore forces a Walsh coefficient with magnitude greater than 2^k. Since Boolean Walsh coefficients are even, that magnitude is at least $2^k + 2$, which gives (52).

If $\dim\ker\eta_F > k$, that kernel cannot consist entirely of zero and bent components, since it would define a vectorial bent map with more than k outputs. It contains a non-bent even-weight component. Its Walsh coefficients are divisible by 4, so one of them has magnitude at least 2^k+4. This proves (53). Finally, $\dim\ker\eta_F = m$ when $\eta_F = 0$, whereas it is $m-1$ otherwise. ■

B. Linear structures of the derivatives

For a Boolean function f on V_{2k}, let

$$L(f) = \{u \in V_{2k} : D_u f \text{ is constant}\}. \tag{54}$$

Lemma 1: If $\ell = \dim L(f)$, then

$$\max_a |W_f(a)| \geq 2^{k+\ell/2}. \tag{55}$$

Proof: The Walsh support lies in an affine coset of $L(f)^\perp$, which has size $2^{2k-\ell}$. Parseval gives the stated lower bound. ■

For $u \neq 0$, define

$$K_u = \{b \in V_m : D_u(b\cdot F) \text{ is constant}\}. \tag{56}$$

If s_u is the affine dimension of $\mathrm{im}(D_uF)$, then $\dim K_u = m - s_u$. To see this, write the affine hull as $y_u + U_u$; the definition then gives $K_u = U_u^\perp$.

Proposition 8 (Derivative partial-spread condition): If $\Lambda(F) < 2^{k+1}$, then

$$K_u \cap K_v = \{0\} \qquad (u \neq v,\ u, v \neq 0), \tag{57}$$

and

$$\sum_{u\neq 0} \left(2^{m-s_u} - 1\right) \leq 2^m - 1. \tag{58}$$

Proof: If $0 \neq b \in K_u \cap K_v$, then $b\cdot F$ has the two independent linear structures u and v. Lemma 1 gives a Walsh coefficient of magnitude at least 2^{k+1}, contradicting the hypothesis. Thus the spaces K_u have pairwise disjoint nonzero parts, and counting those vectors gives (58). ■

The function in (41) satisfies the hypothesis because $\Lambda(F) = 12 < 16$. Its derivative-image annihilators therefore obey (57) and (58), even though (4) fails.

VII. Fourth-Moment Identities for Six-Variable Maps

The divisibility and derivative-image bounds constrain individual spectra and spaces of linear structures, but they do not couple the fourth moments of all components. In six variables, the derivative distribution provides the missing coupling through the weight-4 words of the associated dual code.

We now make this coupling explicit. For a Boolean function $f : V_6 \to \mathbb{F}_2$, define

$$\mathrm{AC}_f(u) = \sum_{x \in V_6} (-1)^{f(x)+f(x+u)} \tag{59}$$

and

$$e_8(f) = \sum_{u \neq 0} \left(\frac{\mathrm{AC}_f(u)}{8} \right)^2. \tag{60}$$

The Walsh and autocorrelation transforms give

$$e_8(f) = \frac{1}{4096} \sum_{a \in V_6} W_f(a)^4 - 64. \tag{61}$$

A. *The affine quotient and the orbit action*

Let $\mathcal{B}_6$ be the $\mathbb{F}_2$-space of Boolean functions on V_6 and let $\mathcal{A}_6 = \mathrm{RM}(1, 6)$. We use the quotient

$$\mathcal{Q}_6 = \mathcal{B}_6/\mathcal{A}_6, \qquad [f] = f + \mathcal{A}_6. \tag{62}$$

Because $\dim \mathcal{B}_6 = 64$ and $\dim \mathcal{A}_6 = 7$, the quotient has 2^{57} elements before the $\mathrm{AGL}(6, 2)$ action is taken. For $A \in \mathrm{AGL}(6, 2)$, composition $f \mapsto f \circ A$ preserves $\mathcal{A}_6$ and hence induces an action on $\mathcal{Q}_6$. Thus

$$[f] \sim [g] \quad \Longleftrightarrow \quad g = f \circ A + \ell, \qquad A \in \mathrm{AGL}(6, 2), \quad \ell \in \mathcal{A}_6. \tag{63}$$

Here $\sim$ denotes membership in the same $\mathrm{AGL}(6, 2)$-orbit. For scalar Boolean functions this is the usual extended-affine relation: the affine output permutation contributes only a constant, which is already contained in ℓ.

The quantities used below depend only on the quotient class. More precisely, if $A(x) = Lx + t$ and $\ell(x) = v \cdot x + c$, a change of variables gives

$$W_{f \circ A + \ell}(u) = (-1)^{c + (u+v) \cdot L^{-1} t}, W_f\big(L^{-\mathsf{T}}(u + v)\big). \tag{64}$$

Consequently, the multiset of absolute Walsh coefficients, the coset minimum weight

$$\mu([f]) = \min_{\ell \in \mathcal{A}_6} \mathrm{wt}(f + \ell), \tag{65}$$

and $e_8(f)$ are invariant under (63). In particular, $\mu([f]) = \mathrm{NL}(f)$. The classification required for Lemma 2 is therefore a classification of $\mathrm{AGL}(6,2)$-orbits on $\mathcal{Q}_6$, rather than a classification of individual truth tables.

B. Completeness of the finite orbit census

The completeness of the census is needed in the proof of Lemma 2, so we state the orbit construction explicitly. Let $\mathcal{B}_{\le d}$ denote the Boolean functions of algebraic degree at most d, let $\mathcal{H}_d$ be the span of the degree-d monomials, and set

$$R_d = \mathcal{B}_6/\mathcal{B}_{\le d-1} \quad (2 \le d \le 6), \qquad R_7 = \{0\}. \tag{66}$$

Every class in R_d has a unique representative whose ANF contains only terms of degree at least d. For such a representative p of a parent class in R_{d+1}, put

$$\Gamma_p = \{A \in \mathrm{AGL}(6,2) : p \circ A - p \in \mathcal{B}_{\le d}\}. \tag{67}$$

Thus Γ_p is the stabilizer of the parent class $[p]$ in R_{d+1} and, in particular, is a subgroup of $\mathrm{AGL}(6,2)$. If $[q]_d$ denotes the homogeneous degree-d part of an ANF, then the action of Γ_p on the possible degree-d extensions is

$$\Phi_{p,A}(h) = \big[(p+h) \circ A - p\big]_d, \qquad h \in \mathcal{H}_d. \tag{68}$$

This map is an affine permutation of $\mathcal{H}_d$. Its linear part is $h \mapsto [h \circ A]_d$, which is invertible, and its translation part is $[p \circ A - p]_d$. Thus (68) is the precise meaning of the "boundary action" used in the enumeration.

Proposition 9 (Complete orbit recursion): Suppose that p represents a parent orbit in R_{d+1}. The orbits of Γ_p on $\mathcal{H}_d$ are in one-to-one correspondence with the orbits of extensions $p+h$ in R_d. For an extension representative h, its stabilizer is

$$\Gamma_{p+h} = \{A \in \Gamma_p : \Phi_{p,A}(h) = h\}. \tag{69}$$

Starting with $p = 0$ and $\Gamma_0 = \mathrm{AGL}(6,2)$ at $d = 6$, and recursively retaining one representative from every orbit in $\mathcal{H}_d$, produces exactly one representative of every orbit on $\mathcal{Q}_6$.

Proof: Let $h_1, h_2 \in \mathcal{H}_d$. If an affine permutation A maps the class of $p+h_1$ to that of $p+h_2$ in R_d, projection to R_{d+1} first gives $p \circ A - p \in \mathcal{B}_{\le d}$, so $A \in \Gamma_p$. Taking the degree-d parts of $(p+h_1) \circ A - (p+h_2) \in \mathcal{B}_{\le d-1}$ then gives $\Phi_{p,A}(h_1) = h_2$. Conversely, this equality implies that

the same difference has degree at most $d-1$, so the two extensions represent the same class in R_d. This proves the first assertion and (69).

Assume inductively that the retained parents are complete and pairwise inequivalent in R_{d+1}. The first assertion partitions all extensions of each parent into disjoint Γ_p-orbits. Extensions of inequivalent parents cannot become equivalent, since their projections to R_{d+1} would then be equivalent. Thus the children are complete and pairwise inequivalent. At $d=2$, $\mathcal{B}_{\le 1}=\mathcal{A}_6$, so $R_2=\mathcal{Q}_6$; the induction therefore proves the claim. ∎

This degree-by-degree refinement is the descending orbit method used in the classification literature. The argument above states the filtration and stabilizer steps needed here explicitly [32].

To certify the subgroup computations in the finite recursion, let $e_1,\dots,e_6$ be the standard basis. The three affine generators used in the enumeration are

$$\begin{aligned} &\sigma(e_i)=e_{\,1+(i \bmod 6)} && (1\le i\le 6),\\ &\tau(e_2)=e_2+e_1, \quad \tau(e_i)=e_i && (i\neq 2),\\ &\nu(x)=x+e_1. \end{aligned} \tag{70}$$

Write T_{ij} for the elementary transvection that sends e_j to e_j+e_i and fixes the other basis vectors. The conjugates of τ by powers of σ give the six transvections on the directed cycle. The matrix identity

$$[T_{ij},T_{jk}]=T_{ik} \qquad (i,j,k \text{ distinct}) \tag{71}$$

shows inductively that a directed path in the cycle gives the transvection between its endpoints. Every ordered pair of distinct coordinates is joined by such a path, so all T_{ij} occur. Gaussian elimination uses these transvections and, over $\mathbb{F}_2$, a row interchange is a product of three of them. Hence they generate $\mathrm{GL}(6,2)$. The translation ν has vector e_1, and $\mathrm{GL}(6,2)$ is transitive on the nonzero vectors; its conjugates therefore give every translation $x\mapsto x+v$. The generated group is consequently $\mathrm{AGL}(6,2)$, whose order is

$$|\,\mathrm{AGL}(6,2)| = 2^6\prod_{i=0}^{5}(2^6-2^i) = 1\,290\,157\,424\,640. \tag{72}$$

At each recursion step, a breadth-first search applies the generators and their inverses to an unmarked element of $\mathcal{H}_d$. It marks every element reached, so the selected representatives partition $\mathcal{H}_d$. We use the right action convention $h^A=\Phi_{p,A}(h)$ and composition $(AB)(h)=(h^A)^B$. To

TABLE III

ORBIT COUNTS AFTER ADDING THE INDICATED ANF LAYERS.

Layer degree d	$\dim \mathcal{H}_d$	Number of retained classes
6	1	2
5	6	3
4	15	11
3	20	205
2	15	150357

compute the exact stabilizer, choose a transporter t_h from a fixed base element to each h in the orbit, with $t_h = 1$ at the base element. The Schreier elements

$$t_h A t^{-1}_{\Phi_{p,A}(h)} \qquad (A \text{ a generator of } \Gamma_p) \tag{73}$$

fix the base element and generate its full stabilizer. Indeed, inserting $t_h t_h^{-1}$ between successive letters of a word $A_1 \cdots A_s$ that returns the base element to itself, and writing $h_0 = h$ and $h_i = \Phi_{p,A_i}(h_{i-1})$, gives the explicit factorization

$$A_1 \cdots A_s = \prod_{i=1}^{s} \big(t_{h_{i-1}} A_i t_{h_i}^{-1}\big), \qquad h_s = h. \tag{74}$$

Thus every element of the stabilizer is generated by (73). The orbit-stabilizer identity gives

$$|\Gamma_{p+h}| = \frac{|\Gamma_p|}{|\operatorname{Orb}_{\Gamma_p}(h)|}. \tag{75}$$

All layer actions and stabilizer orders are evaluated over finite integer tables. The implementation retains Schreier generators until the recursively computed subgroup has the order on the right of (75). Because the generated subgroup is contained in the stabilizer, equality of orders proves equality of subgroups. This supplies both the no-omission and no-duplication checks required by Proposition 9.

The numbers of retained classes after adding the layers $6, 5, \ldots, d$ are listed in Table III. The last number agrees with Maiorana's complete classification [33]. This agreement is an independent check: completeness and uniqueness follow from the filtration and stabilizer arguments above, rather than from matching the final integer alone.

C. A scalar classification bound

Lemma 2 (Six-variable classification bound): Let $f : V_6 \to \mathbb{F}_2$.

1) If $\mathrm{wt}(f)$ is even and $\max_a |W_f(a)| \le 12$, then $e_8(f) \le 60$.
2) If $\mathrm{wt}(f)$ is odd and $\max_a |W_f(a)| \le 14$, then $e_8(f) \le 327/4$.

Proof: Proposition 9 gives exactly one representative for each of the 150357 $\mathrm{AGL}(6,2)$-orbits on $\mathcal{Q}_6$, which is the orbit formulation of Maiorana's classification [33]. For a representative f, the Walsh coefficients can be obtained without selecting affine representatives one at a time. For each $a \in V_6$,

$$W_f(a) = 64 - 2\,\mathrm{wt}(f + a \cdot x), \tag{76}$$

and the constant affine term changes $\mathrm{wt}(f + a \cdot x)$ to its complement. Thus the 64 linear masks, together with complementation, account for all 128 elements of the coset $[f]$. Substituting these 64 integers into (61) gives $e_8([f])$ exactly. Equation (64) shows that this value is independent of the chosen representative.

For even f, all Walsh coefficients are divisible by 4, so

$$\max_a |W_f(a)| \le 12 \quad \Longleftrightarrow \quad \mu([f]) \ge 26. \tag{77}$$

For odd f, all Walsh coefficients are congruent to 2 modulo 4, and similarly

$$\max_a |W_f(a)| \le 14 \quad \Longleftrightarrow \quad \mu([f]) \ge 25. \tag{78}$$

The complete minimum-weight distribution in the orbit census has the following upper tail:

$\mu([f])$	25	26	27	28
# $\mathrm{AGL}(6,2)$-orbits	983	84	4	4

No orbit has $\mu([f]) > 28$. (79)

Because the parity of every word in a coset is fixed, the even cases in (77) have minimum weight 26 or 28, whereas the odd cases in (78) have minimum weight 25 or 27. Evaluating (61) for every representative in these four rows gives Table IV. Its row multiplicities sum to $84 + 4$ in the even case and $983 + 4$ in the odd case, so the table contains every case covered by either hypothesis. The largest entry proves the two claims. The supplementary implementation performs the orbit recursion, the Schreier order checks, and the integer fourth-moment evaluations. ■

TABLE IV

VALUES OF e_8 IN THE RELEVANT $\mathrm{AGL}(6,2)$-ORBITS ON $\mathcal{Q}_6$ (EQUIVALENTLY, SCALAR EA CLASSES). THE NUMBER OF CLASSES IS SHOWN IN PARENTHESES.

Weight parity	Coset minimum weight	e_8 (number of EA classes)
Even	26	30 (8), 42 (13), 48 (7), 54 (1), 60 (55)
Even	28	0 (4)
Odd	25	159/4 (11), 183/4 (10), 207/4 (48), 231/4 (153), 255/4 (237), 279/4 (347), 303/4 (135), 327/4 (42)
Odd	27	63/4 (4)

D. Derivative collisions and the associated code

Let $H : V_6 \to V_t$, where $t \geq 3$. For $u \neq 0$, partition V_6 into the 32 unordered pairs $\{x, x+u\}$. Let $m_u(y)$ count the pairs for which

$$H(x+u) + H(x) = y,$$

and define

$$c_u = \sum_{y \in V_t} \binom{m_u(y)}{2}.$$

Lemma 3: For every $u \neq 0$,

$$\sum_{b \neq 0} \left(\frac{\mathrm{AC}_{H_b}(u)}{8} \right)^2 = 2^{t+1} - 64 + 2^{t-3} c_u.$$

Proof: Since the derivative is constant on each unordered pair,

$$\mathrm{AC}_{H_b}(u) = 2 \sum_{y \in V_t} m_u(y)(-1)^{b \cdot y}.$$

Character orthogonality on V_t gives

$$\begin{aligned} \sum_{b \neq 0} \left(\frac{\mathrm{AC}_{H_b}(u)}{8} \right)^2 &= 2^{t-4} \sum_{y} m_u(y)^2 - 64 \\ &= 2^{t+1} - 64 + 2^{t-3} \sum_{y} \binom{m_u(y)}{2}, \end{aligned}$$

where the last equality uses $\sum_y m_u(y) = 32$. ∎

Lemma 4: Let

$$C_H = \mathrm{RM}(1,6) + \langle H_1, \ldots, H_t \rangle.$$

Then

$$\sum_{u\neq 0} c_u = 3B_4(C_H^{\perp}).$$

Proof: Two distinct pairs counted by $\binom{m_u(y)}{2}$ form the four points of an affine two-flat. Their derivative values agree exactly when the sum of the four values of H is zero; the sum of the four input points is zero as well. Thus the corresponding support is a weight-4 codeword in $C_H^{\perp}$. Conversely, every weight-4 word in $C_H^{\perp}$ is supported on four distinct points. Orthogonality to $\mathrm{RM}(1,6)$ forces their xor to be zero, so they form an affine two-flat. Orthogonality to each coordinate of H then forces the four values of H to sum to zero. Such a flat has three nonzero directions, and its two opposite pairs have the same derivative value in each direction. It therefore contributes once for each direction. ■

Proposition 10: For every map $H : V_6 \to V_t$, where $t \geq 3$,

$$\sum_{b\neq 0} e_8(H_b) = 63(2^{t+1} - 64) + 3 \cdot 2^{t-3} B_4(C_H^{\perp}). \tag{80}$$

In particular,

$$\sum_{b\in V_5\setminus\{0\}} e_8(H_b) = 12B_4(C_H^{\perp}), \qquad t = 5, \tag{81}$$

$$\sum_{b\in V_6\setminus\{0\}} e_8(H_b) = 4032 + 24B_4(C_H^{\perp}), \qquad t = 6. \tag{82}$$

Proof: Summing Lemma 3 over the 63 nonzero directions and applying Lemma 4 gives the identity. ■

E. The fourth moment of the constructed function

The identity also records the joint structure of the function in Theorem 2. Although it is not needed to prove the function's nonlinearity, it provides a direct numerical link between the construction and the six-output bound.

Proposition 11: For the function F in (41),

$$e_8(F_b) = \begin{cases} 0, & b \in \mathcal{E} \setminus \{0\}, \\ 60, & b \in e_4 + \mathcal{E}, \end{cases} \tag{83}$$

and

$$\sum_{b\neq 0} e_8(F_b) = 480, \qquad B_4(C_F^{\perp}) = 416. \tag{84}$$

Consequently, (80) at $t = 4$ reads

$$480 = -2016 + 6 \cdot 416. \tag{85}$$

Proof: For a bent component, (61) gives $64 \cdot 8^4/4096 - 64 = 0$. Each non-bent component has 24 Walsh coefficients of magnitude 12 and 40 of magnitude 4, so

$$e_8(F_b) = \frac{24 \cdot 12^4 + 40 \cdot 4^4}{4096} - 64 = 60.$$

There are seven components of the first type and eight of the second; this gives the first equality in (84). Applying the MacWilliams identity to (43) gives

$$B_4(C_F^{\perp}) = [z^4]\frac{(1+z)^{64}}{2^{11}} W_{C_F}\left(\frac{1-z}{1+z}\right) = 416.$$

Substitution into (80) proves (85). ∎

The eight non-bent components therefore attain the scalar bound in Lemma 2 simultaneously. At output dimension four, the negative constant term in (80) is offset by the 416 weight-4 dual words. At output dimension six, by contrast, the constant term alone is 4032. Because this exceeds the total scalar allowance $63 \cdot 60 = 3780$ for 63 even components, the scalar bounds and the vectorial identity are incompatible when all components have even weight.

VIII. The Exact Bound for Six-Bit Permutations

We first state this incompatibility for maps whose output sum vanishes; the permutation upper bound will then follow immediately.

Proposition 12 (Even-component obstruction): If $H : V_6 \to V_6$ satisfies

$$\sum_{x \in V_6} H(x) = 0, \tag{86}$$

then $\mathrm{NL}(H) \leq 24$.

Proof: Assume, for a contradiction, that $\mathrm{NL}(H) \geq 25$. Every nonzero component then has Walsh peak at most 14. Because (86) makes all component weights even,

$$W_f(a) = 64 - 2\,\mathrm{wt}(f + a \cdot x)$$

is divisible by 4 for each such component. Consequently, every nonzero component of H has Walsh peak at most 12, and Lemma 2 gives

$$\sum_{b \neq 0} e_8(H_b) \leq 63 \cdot 60 = 3780. \tag{87}$$

On the other hand, (82) gives

$$\sum_{b\neq 0} e_8(H_b) = 4032 + 24B_4(C_H^{\perp}) \geq 4032, \tag{88}$$

a contradiction. ∎

Corollary 4 (Permutation upper bound): Every permutation $S : V_6 \to V_6$ satisfies $\mathrm{NL}(S) \leq 24$.

Proof: Since S is bijective,

$$\sum_{x\in V_6} S(x) = \sum_{y\in V_6} y = 0.$$

Therefore Proposition 12 applies. ∎

A. Attainment

Theorem 3: Let $S : \mathbb{F}_{64} \to \mathbb{F}_{64}$ be defined by

$$S(0) = 0, \qquad S(x) = x^{-1} \quad (x \neq 0).$$

Then S is a permutation and

$$\mathrm{NL}(S) = 24.$$

Proof: With the convention $0^{-1} = 0$, inversion is a bijection of $\mathbb{F}_{64}$. The trace pairing is nondegenerate, so every nonzero linear functional on $\mathbb{F}_{64}$ has the form $y \mapsto \mathrm{Tr}_{64/2}(\beta y)$. Thus the nonzero components of S are

$$f_\beta(x) = \mathrm{Tr}_{64/2}(\beta x^{-1}), \qquad \beta \in \mathbb{F}_{64}^*.$$

The identity $f_\beta(\beta x) = f_1(x)$ makes them linearly equivalent. Consequently, it suffices to evaluate the Walsh transform of f_1. A direct evaluation over $\mathbb{F}_{64} = \mathbb{F}_2[X]/(X^6 + X + 1)$ gives the distribution

t	-12	-8	-4	0	4	8	12	16
multiplicity	6	7	12	13	6	9	8	3.

The maximum absolute Walsh coefficient is 16, so each nonzero component has nonlinearity $32 - 16/2 = 24$. This proves the lower bound, while Corollary 4 supplies the matching upper bound. The finite-field calculation is included in the supplementary material. ∎

Corollary 5:

$$\max_{\substack{S:V_6\to V_6 \\ S \text{ permutation}}} \mathrm{NL}(S) = 24.$$

IX. Non-Bijective Six-to-Six Maps

The preceding obstruction settles the zero-sum case. Any six-to-six map with nonlinearity above 24 must therefore have the complementary parity pattern.

Corollary 6 (Component parity split): Every map $F : V_6 \to V_6$ with $\mathrm{NL}(F) \geq 25$ has exactly 31 nonzero even-weight components and 32 odd-weight components.

Proof: By Proposition 12, $\sum_x F(x) \neq 0$. Hence the parity functional η_F in (17) is nonzero. Its kernel has dimension 5 and contains the zero mask and the 31 nonzero even masks; the remaining 32 masks index odd-weight components. ∎

This parity split leads to two nested Reed–Muller supercodes. The next proposition records the parameters and the fourth-moment restrictions that any such map must satisfy.

Proposition 13: Let $F : V_6 \to V_6$ satisfy $\mathrm{NL}(F) \geq 25$. After an invertible output transformation, write $F = (G, h)$, where $G : V_6 \to V_5$, all components of G have even weight, and all functions $h + u \cdot G$ have odd weight. Define

$$E = \mathrm{RM}(1, 6) + \langle G_1, \ldots, G_5 \rangle, \qquad C = E + \langle h \rangle,$$

and put

$$q_E = B_4(E^\perp), \qquad q_C = B_4(C^\perp).$$

Then

$$E = [64, 12, 26], \qquad C = [64, 13, 25], \tag{89}$$

and

$$118 \leq q_E \leq 155, \qquad q_E - 2q_C \geq 118, \qquad 0 \leq q_C \leq 18. \tag{90}$$

Proof: Corollary 6 gives $\eta_F \neq 0$. Choose the first five output coordinates as a basis of $\ker \eta_F$, and choose h outside that kernel. Since the assumed nonlinearity excludes nonzero affine components, $\dim E = 12$ and $\dim C = 13$.

Every nonzero component of G has even weight, and its maximum absolute Walsh coefficient is at most 12. Therefore $d(E) \geq 26$. Because $5 > 6/2$, Nyberg's bound implies that at least one component of G is not bent. Its Walsh coefficients are divisible by 4, and Parseval then forces the maximum magnitude to be 12. Thus $d(E) = 26$.

Equations (81) and (82) give

$$\sum_{u\neq 0} e_8(u\cdot G) = 12q_E, \tag{91}$$

$$\sum_{u\in V_5} e_8(h+u\cdot G) = 4032 + 24q_C - 12q_E. \tag{92}$$

For the even components, Lemma 2 yields $12q_E \leq 31\cdot 60$, and hence $q_E \leq 155$. For the odd components, the same lemma yields

$$4032 + 24q_C - 12q_E \leq 32\cdot\frac{327}{4},$$

which is equivalent to $q_E - 2q_C \geq 118$. Nonnegativity of q_E and q_C then gives $q_E \geq 118$ and $q_C \leq 18$ as well.

It remains to determine $d(C)$. Suppose that every function $h + u\cdot G$ had maximum absolute Walsh coefficient at most 10. Because each such function has odd weight, its Walsh magnitudes would lie in $\{2, 6, 10\}$. Let N_j denote the number of coefficients of magnitude j. Parseval gives

$$N_6 + 3N_{10} = 120.$$

Substituting this relation into (61) gives

$$e_8(h+u\cdot G) = \frac{6N_{10}-105}{4} \leq \frac{135}{4}.$$

Summing over the 32 odd components would make the left-hand side of (92) at most 1080. In contrast, that equation, together with $q_C \geq 0$ and $q_E \leq 155$, gives $4032 - 12q_E \geq 2172$, which is impossible. Hence some odd component has maximum absolute Walsh coefficient 14; its distance from $\mathrm{RM}(1,6)$ is 25, and therefore $d(C) = 25$. ■

Thus, if $N(6,6) > 24$, the associated map must induce the nested codes in (89), and their dual codes must satisfy the weight-4 multiplicity constraints in (90). These necessary conditions do not establish that such a non-bijective map exists.

X. Conclusion

The results first establish a componentwise framework for the high-output problem. The projection proposition reduces bounds that depend only on n and m to the boundary case $m = k + 1$. Within that case, the quadratic rank argument and the codimension-to-divisibility theorem identify two settings in which $\Lambda(F) \geq 2^{k+1}$. In particular, the divisibility theorem

recovers the MNBC bound from the span of the non-bent masks. When this inequality does not hold, the derivative partial-spread condition supplies an independent necessary condition.

The function in (41) exhibits the boundary behavior. Its zero component and seven bent components form a three-dimensional subspace whose seven nonzero elements are bent. The remaining eight components form an even-weight affine coset with Walsh peak 12. Because these masks span the full output space, the codimension parameter is zero. The associated code chain

$$\mathrm{RM}(1,6) \subset C_G = [64, 10, 28] \subset C_F = [64, 11, 26]$$

has $d(p, C_G) = 26$ and $26 \leq \rho(C_G) \leq 27$. Consequently, $26 \leq N(6,4) \leq 27$. The shifted Walsh identity describes the correlations within the affine coset, and the fourth-moment identity gives $B_4(C_F^{\perp}) = 416$.

For six-output maps, the same fourth-moment identity leads to a sharper obstruction. Its constant term exceeds the sum of the scalar bounds for the 63 even-weight components. Therefore every map $H : V_6 \to V_6$ satisfying $\sum_x H(x) = 0$ has $\mathrm{NL}(H) \leq 24$. This gives the permutation upper bound, and inversion on $\mathbb{F}_{64}$ attains it. A non-bijective map above 24 must instead induce the nested codes in Proposition 13, whose dual weight-4 multiplicities obey the stated bounds.

The results leave several questions open. A maximal bent code may or may not admit an odd-weight coset of minimum weight 27; this is equivalent to $N(6,4) = 27$. The covering radius of the specific code C_G also remains to be determined, namely whether it is 26 or 27. Finally, it is unknown whether a non-bijective six-to-six map can exceed nonlinearity 24. Classifying maximal vectorial bent mappings that admit even affine cosets with Walsh peak 12 could clarify whether the construction extends to higher even input dimensions.

REFERENCES

[1] C. Carlet and C. Ding, "Nonlinearities of S-boxes," *Finite Fields and Their Applications*, vol. 13, no. 1, pp. 121–135, 2007.

[2] P. Beelen and G. Leander, "A new construction of highly nonlinear S-boxes," *Cryptography and Communications*, vol. 4, no. 1, pp. 65–77, 2012.

[3] D. Tang, C. Carlet, and Z. Zhou, "Binary linear codes from vectorial boolean functions and their weight distribution," *Discrete Mathematics*, vol. 340, no. 12, pp. 3055–3072, 2017.

[4] K. Nyberg, "Perfect nonlinear S-boxes," in *Advances in Cryptology—EUROCRYPT '91*, ser. Lecture Notes in Computer Science, vol. 547. Springer, 1991, pp. 378–386.

[5] T. Wadayama, T. Hada, K. Wakasugi, and M. Kasahara, "Upper and lower bounds on maximum nonlinearity of n-input m-output boolean function," *Designs, Codes and Cryptography*, vol. 23, no. 1, pp. 23–34, 2001.

[6] K. Nyberg, "S-boxes and round functions with controllable linearity and differential uniformity," in *Fast Software Encryption*, ser. Lecture Notes in Computer Science, vol. 1008. Springer, 1995, pp. 111–130.

[7] A. A. Polujan and A. Pott, "On design-theoretic aspects of boolean and vectorial bent function," *IEEE Transactions on Information Theory*, vol. 67, no. 2, pp. 1027–1037, 2021.

[8] C. Bracken and G. Leander, "A highly nonlinear differentially 4-uniform power mapping that permutes fields of even degree," *Finite Fields and Their Applications*, vol. 16, no. 4, pp. 231–242, 2010.

[9] Y. Xu, C. Carlet, S. Mesnager, and C. Wu, "Classification of bent monomials, constructions of bent multinomials and upper bounds on the nonlinearity of vectorial functions," *IEEE Transactions on Information Theory*, vol. 64, no. 1, pp. 367–383, 2018.

[10] 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 Transactions on Information Theory*, vol. 67, no. 12, pp. 8325–8334, 2021.

[11] W. Zhang, C. Song, and E. Pasalic, "A design of five-valued spectra (vectorial) boolean functions and their use in constructing bent functions outside $\mathcal{M}^{\#}$," *IEEE Transactions on Information Theory*, vol. 72, no. 7, pp. 5217–5236, 2026.

[12] W.-G. Zhang and E. Pasalic, "Constructions of resilient S-boxes with strictly almost optimal nonlinearity through disjoint linear codes," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1638–1651, 2014.

[13] ——, "Highly nonlinear balanced S-boxes with good differential properties," *IEEE Transactions on Information Theory*, vol. 60, no. 12, pp. 7970–7979, 2014.

[14] J. Liu, S. Mesnager, and L. Chen, "On the nonlinearity of S-boxes and linear codes," *Cryptography and Communications*, vol. 9, no. 3, pp. 345–361, 2017.

[15] G. P. Nagy, "Sidon sets, thin sets, and the nonlinearity of vectorial boolean functions," *Journal of Combinatorial Theory, Series A*, vol. 212, p. 106001, 2025.

[16] ——, "On the minimum hamming distance between vectorial boolean and affine functions," *Cryptography and Communications*, vol. 17, no. 6, pp. 1703–1720, 2025.

[17] V. G. Ryabov, "Nonlinearity of vectorial functions over finite fields," *Discrete Mathematics and Applications*, vol. 36, no. 1, pp. 43–58, 2026.

[18] A. Pott, E. Pasalic, A. Muratović-Ribić, and S. Bajrić, "On the maximum number of bent components of vectorial functions," *IEEE Transactions on Information Theory*, vol. 64, no. 1, pp. 403–411, 2018.

[19] L. Zheng, J. Peng, H. Kan, Y. Li, and J. Luo, "On constructions and properties of (n, m)-functions with maximal number of bent components," *Designs, Codes and Cryptography*, vol. 88, no. 10, pp. 2171–2186, 2020.

[20] S. Mesnager, F. Zhang, C. Tang, and Y. Zhou, "Further study on the maximum number of bent components of vectorial functions," *Designs, Codes and Cryptography*, vol. 87, no. 11, pp. 2597–2610, 2019.

[21] A. Bapić, E. Pasalic, A. Polujan, and A. Pott, "Vectorial boolean functions with the maximum number of bent components beyond the nyberg's bound," *Designs, Codes and Cryptography*, vol. 92, no. 3, pp. 531–552, 2024.

[22] N. Anbar, T. Kalaycı, W. Meidl, and L. Mérai, "On a class of functions with the maximal number of bent components," *IEEE Transactions on Information Theory*, vol. 68, no. 9, pp. 6174–6186, 2022.

[23] X. Xie, Y. Ouyang, and H. Hu, "On vectorial functions with maximal number of bent components," *Designs, Codes and Cryptography*, vol. 93, no. 6, pp. 1889–1910, 2025.

[24] W. Meidl, A. A. Polujan, and A. Pott, "Linear codes and incidence structures of bent functions and their generalizations," *Discrete Mathematics*, vol. 346, no. 1, p. 113157, 2023.

[25] M. Grassl, "Bounds on the minimum distance of linear codes," Online database: https://www.codetables.de/BKLC/BKLC.php?q=2&n=64&k=11, accessed Aug. 19, 2026.

[26] C. Carlet, "Vectorial Boolean functions for cryptography," in *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, ser. Encyclopedia of Mathematics and Its Applications. Cambridge University Press, 2010, vol. 134, pp. 398–470.

[27] L. Perrin, A. Udovenko, and A. Biryukov, "Cryptanalysis of a theorem: Decomposing the only known solution to the big APN problem," in *Advances in Cryptology—CRYPTO 2016*, ser. Lecture Notes in Computer Science, vol. 9814. Springer, 2016, pp. 93–122.

[28] A. Canteaut, S. Duval, and L. Perrin, "A generalisation of Dillon's APN permutation with the best known differential and nonlinear properties for all fields of size 2^{4k+2}," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7575–7591, 2017.

[29] S. Fu, X. Feng, and B. Wu, "Differentially 4-uniform permutations with the best known nonlinearity from butterflies," *IACR Transactions on Symmetric Cryptology*, vol. 2017, no. 2, pp. 228–249, 2017.

[30] K. H. Kim, S. Mesnager, J. H. Choe, D. N. Lee, S. Lee, and M. C. Jo, "On permutation quadrinomials with boomerang uniformity 4 and the best-known nonlinearity," *Designs, Codes and Cryptography*, vol. 90, no. 6, pp. 1437–1461, 2022.

[31] O. S. Rothaus, "On "bent" functions," *Journal of Combinatorial Theory, Series A*, vol. 20, no. 3, pp. 300–305, 1976.

[32] V. Gillot and P. Langevin, "Classification of some cosets of the Reed–Muller code," *Cryptography and Communications*, vol. 15, no. 6, pp. 1129–1137, 2023.

[33] J. A. Maiorana, "A classification of the cosets of the Reed–Muller code R(1,6)," *Mathematics of Computation*, vol. 57, no. 195, pp. 403–414, 1991.