(Balanced) Boolean functions on 375 variables with nonlinearity exceeding the Patterson-Wiedemann direct-sum bound

Deng Tang and Subhamoy Maitra
Shanghai Jiao Tong University, Shanghai 200240, China
Indian Statistical Institute, Kolkata 700108, India,
{dengtang@sjtu.edu.cn, subho@isical.ac.in, maitra.subhamoy@gmail.com}

Abstract

Patterson and Wiedemann [IEEE-IT 1983] constructed Boolean functions on 15 variables with nonlinearity 16276, which is strictly greater than the bent concatenation bound. Taking their direct sum with a bent function gives nonlinearity $2^{n-1}-27\cdot 2^{(n-11)/2}$ for every odd $n \geq 15$. Much later, Schmidt [JCT-A 2019] proved that the asymptotic maximum nonlinearity is $2^{n-1}-(1+o(1))2^{n/2-1}$ as $n \to \infty$. In this work, we construct Boolean functions f on 375 variables such that the nonlinearity is greater than $2^{374}-442367\cdot 2^{168} = 2^{374}-2^{187}+81921\cdot 2^{168}$, which exceeds the direct sum of a 15-variable Patterson-Wiedemann function and a bent function on 360 variables. To be specific, the direct sum of a 15-variable Patterson-Wiedemann function and a 360-variable bent function has nonlinearity $2^{374}-2^{187}+20\cdot 2^{180} = 2^{374}-2^{187}+81920\cdot 2^{168}$. Thus our construction improves on this direct-sum value by more than 2^{168}. Our construction idea is as follows. Let H be the multiplicative subgroup of order $(2^{375}-1)/751$ in $\mathbb{F}_{2^{375}}^*$. The support of f is the disjoint union of 375 prescribed cosets and a subset $B \subseteq H$. The evaluation of Gauss sums reduces the contribution of the prescribed cosets to a computation of 751 integers. The conditions on B form a 0-1 linear feasibility problem. Using a potential criterion for simultaneous linear constraints, we prove the existence of a single subset $B \subseteq H$ satisfying the constraints at all frequencies. Taking direct sums with bent functions extends the result to every odd $n \geq 375$. The support construction does not require balancedness. As a natural follow-up, we also obtain balanced functions with the same strict lower bound by changing the minimum possible number of values, using sampling without replacement. The subsets B and the sets of modified points are obtained by existence proofs.

Keywords: Boolean function, Gauss period, Nonlinearity, Discrepancy, Support.

1 Introduction

Constructing Boolean functions at large distance from every affine function is a basic problem in coding theory and cryptography [2]. For a Boolean function f on n variables, the nonlinearity $\mathrm{NL}(f)$ is its minimum Hamming distance from the affine Boolean functions. It is well-known that the maximum nonlinearity is also the covering radius of the first-order Reed-Muller code. Parseval's identity gives $\mathrm{NL}(f) \leq 2^{n-1}-2^{n/2-1}$. For even n, bent functions attain this bound [16]. For odd n, equality is excluded by the integrality of the Walsh coefficients. Concatenating two bent functions on $n-1$ variables gives nonlinearity

$$\mathcal{N}_{\mathrm{BC}}(n) = 2^{n-1}-2^{(n-1)/2}. \tag{1.1}$$

This is the Bent Concatenation (BC) bound: it is an attainable construction bound, not a universal upper bound.

In 1983, Patterson and Wiedemann [13] (see also [14]) constructed Boolean functions on 15 variables with nonlinearity 16276, exceeding the corresponding bent concatenation value 16256. Taking their direct sum with a bent function on disjoint variables gives, for every odd $n \geq 15$,

$$\begin{aligned} \mathcal{N}_{\mathrm{PW}}(n) &= 2^{n-1} - 27 \cdot 2^{(n-11)/2} \\ &= \mathcal{N}_{\mathrm{BC}}(n) + 20 \cdot 2^{(n-15)/2}. \end{aligned} \tag{1.2}$$

We call this the Patterson-Wiedemann (PW) direct-sum bound.

For small numbers of variables, Kavut, Maitra and Yücel obtained rotation symmetric Boolean functions on nine variables with nonlinearity 241 [7], exceeding the bent concatenation value 240. Kavut and Yücel subsequently obtained nonlinearity 242 in generalized rotation symmetric classes and related classes [8]. Direct sums with bent functions then give nonlinearity 996 on 11 variables and 4040 on 13 variables. For odd $n \geq 15$, continuing this construction from a nine-variable function of nonlinearity 242 gives $2^{n-1} - 28 \cdot 2^{(n-11)/2}$, which is still below (1.2). Recent pre-prints claim that the maximum nonlinearity of a 9-variable Boolean function is exactly 242. Independent pre-prints by Wang et al. [19] and Hou et al. [6] report this result. As a passing remark, we should also mention that Yuan et al. [21] and Lou et al. [12] have also very recently and independently determined that the maximum nonlinearity of 8-variable balanced Boolean functions is 116. This is a long-standing open problem since the work of [4]. The recent results discussed above leverage advances in computation, ranging from computational analysis of different cases to the use of AI tools to assist in developing proof strategies, or a combination of these approaches.

Coming back to the issue of additional nonlinearity over the bent concatenation bound, one should note that Kavut and Maitra [9] also constructed Patterson-Wiedemann type functions directly on 21 variables whose nonlinearity exceeds the bent concatenation bound. However, the nonlinearity achieved is lower than that obtained by the PW direct sum, namely $2^{21-1} - 2^{\frac{21-1}{2}} + 160$. In fact, it was noted in [10] that it might be possible to obtain a 21-variable function with nonlinearity $2^{21-1} - 2^{\frac{21-1}{2}} + 196$, though discovering such a function still remains elusive. In this direction, here we provide explicit parameters that improve on the PW direct-sum bound (1.2). In this regard, one should note that Schmidt [18] proved the asymptotic result

$$\max_{f:\mathbb{F}_2^n \to \mathbb{F}_2} \mathrm{NL}(f) = 2^{n-1} - (1 + o(1))2^{n/2-1} \qquad (n \to \infty),$$

settling an asymptotic conjecture of Patterson and Wiedemann. However, this asymptotic result does not provide explicit parameters for a fixed number of variables that improve on (1.2).

1.1 Organization and Contribution

We obtain such an improvement on 375 variables, i.e., our result explains how to obtain a Boolean function where the nonlinearity exceeds the direct sum of a 15-variable Patterson-Wiedemann function and a bent function on 360 variables.

The support has the form $B \cup \bigcup_{\substack{1 \leq i \leq 750 \\ g(\alpha^i)=1}} \alpha^i H$, where g is a prescribed 0, 1-valued function constant on each coset, and only $B \subseteq H$ remains to be chosen. Three distinct values of the Gauss periods determine the contribution of the prescribed cosets. Their values determine the offsets in a 0-1 linear constraint system. Using a minor modification of an argument of Reis and Song [15], we prove that this constraint system has a solution B. Theorems 3.5 and 4.2 give Boolean functions, including balanced ones, on 375 variables satisfying

$$\mathrm{NL}(f) > 2^{374} - 442367 \cdot 2^{168} = \mathcal{N}_{\mathrm{BC}}(375) + 81921 \cdot 2^{168}.$$

Theorems 3.6 and 4.3 extend this to every odd $n \geq 375$.

The remainder of this paper is organized as follows. Section 2 gives the necessary preliminaries. Section 3 presents the support construction, computes the contribution of the prescribed cosets, proves the existence of B, derives the nonlinearity bound and gives the direct-sum extension. Section 4 treats balanced functions separately.

2 Preliminaries

In this section, we recall and present the algebraic preliminaries. In this paper, additive characters are written as $(-1)^{\mathrm{Tr}_1^n(\cdot)}$, and multiplicative characters as powers of roots of unity.

2.1 Indicator functions and the trace functions

Let $\mathbb{F}_2 = \{0, 1\}$ be the field with two elements, and let $\mathbb{F}_2^n$ denote the vector space of binary n-tuples. An n-variable Boolean function is a map $f : \mathbb{F}_2^n \to \mathbb{F}_2$, and we use $\mathcal{B}_n$ to denote the set of all such functions. The basic representation of f is its truth table

$$f = \big[f(0,0,\ldots,0), f(1,0,\ldots,0), f(0,1,\ldots,0), \ldots, f(1,1,\ldots,1)\big].$$

We say that a Boolean function f is balanced if its truth table contains an equal number of ones and zeros, that is, if its Hamming weight equals 2^{n-1}, where the Hamming weight of f, wt (f), is defined as the size of the support $\mathrm{supp}(f) = \{x \in \mathbb{F}_2^n : f(x) = 1\}$. For a finite set A, the notation $|A|$ denotes its cardinality. For a real or complex number, $|z|$ denotes its absolute value or modulus. The indicator function of A and the Kronecker delta are defined by

$$\mathbf{1}_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases} \text{ and } \delta_{r,a} = \mathbf{1}_{\{a\}}(r) = \begin{cases} 1, & r = a, \\ 0, & r \neq a. \end{cases}$$

If $f = \mathbf{1}_A$ is a Boolean function, then A is its support and

$$(-1)^{f(x)} = 1 - 2\mathbf{1}_A(x) = \begin{cases} -1, & x \in A, \\ 1, & x \notin A. \end{cases} \tag{2.1}$$

The right-hand side is evaluated in $\mathbb{R}$. Equation (2.1) encodes Boolean values $0, 1$ by signs ± 1.

Choosing a basis over $\mathbb{F}_2$ identifies $\mathbb{F}_{2^n}$ with $\mathbb{F}_2^n$. We may therefore represent a Boolean function on n variables as a function on $\mathbb{F}_{2^n}$. The absolute trace $\mathrm{Tr}_1^n : \mathbb{F}_{2^n} \to \mathbb{F}_2$ is $\mathrm{Tr}_1^n(x) = \sum_{i=0}^{n-1} x^{2^i}$. The trace pairing $(u, x) \mapsto \mathrm{Tr}_1^n(ux)$ is nondegenerate [11, Chapter 2]. For each fixed $\omega \in \mathbb{F}_{2^n}$, we have $(-1)^{\mathrm{Tr}_1^n(\omega(x+y))} = (-1)^{\mathrm{Tr}_1^n(\omega x)}(-1)^{\mathrm{Tr}_1^n(\omega y)}$ and, therefore, $x \mapsto (-1)^{\mathrm{Tr}_1^n(\omega x)}$ is an additive character. The trace pairing also relates this notation to the usual coordinate inner product. For a basis $b_1, \ldots, b_n$, there is a trace-dual basis $b'_1, \ldots, b'_n$ such that $\mathrm{Tr}_1^n(b'_i b_j) = \delta_{i,j}$. Writing $x = \sum_{j=1}^n x_j b_j$ and $\omega = \sum_{i=1}^n \omega_i b'_i$ gives $\mathrm{Tr}_1^n(\omega x) = \sum_{i=1}^n \omega_i x_i$ in $\mathbb{F}_2$. Hence $(-1)^{\mathrm{Tr}_1^n(\omega x)}$ equals $(-1)^{\omega_1 x_1 + \cdots + \omega_n x_n}$ in these coordinates.

2.2 Walsh transforms, nonlinearity and direct sums

The Hamming distance between two Boolean functions is $d_H(f, g) = |\{x \in \mathbb{F}_{2^n} : f(x) \neq g(x)\}|$. The affine functions are precisely $x \mapsto \mathrm{Tr}_1^n(\omega x) + b$, where $\omega \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_2$. The nonlinearity $\mathrm{NL}(f)$ is the minimum Hamming distance from f to these functions. For $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$, its Walsh transform at a point $\omega \in \mathbb{F}_{2^n}$ is defined as

$$W_f(\omega) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x) + \mathrm{Tr}_1^n(\omega x)}.$$

The multiset of values of the Walsh transform is called the Walsh spectrum of f. For a real-valued function $\mathsf{h} : \mathbb{F}_{2^n} \to \mathbb{R}$, we also use the normalized transform

$$\widehat{\mathsf{h}}(\omega) = 2^{-n/2} \sum_{x \in \mathbb{F}_{2^n}} \mathsf{h}(x)(-1)^{\mathrm{Tr}_1^n(\omega x)}.$$

Note that $\widehat{\mathbf{1}_A}$ transforms the indicator function, whereas $2^{-n/2} W_{\mathbf{1}_A}$ transforms the sign function $1 - 2\mathbf{1}_A$. These are different quantities. For a Boolean function $f \in \mathcal{B}_n$, the nonlinearity $\mathrm{NL}(f)$ is its minimum Hamming distance from the affine Boolean functions. It can be easily verified that $\mathrm{NL}(f) = 2^{n-1} - \frac{1}{2} \max_{\omega \in \mathbb{F}_{2^n}} |W_f(\omega)|$. Moreover, the Walsh transform of a Boolean function $f \in \mathcal{B}_n$ can be computed as

$$W_f(\omega) = \begin{cases} 2^n - 2\mathrm{wt}(f), & \omega = 0, \\ -2 \displaystyle\sum_{x \in \mathrm{supp}(f)} (-1)^{\mathrm{Tr}_1^n(\omega x)}, & \omega \neq 0. \end{cases} \tag{2.2}$$

By Parseval's identity, for any n-variable Boolean function, we have $\sum_{\omega \in \mathbb{F}_{2^n}} {W_f}^2(\omega) = 2^{2n}$. So we have $\mathrm{NL}(f) \leq 2^{n-1} - 2^{n/2-1}$. A Boolean function is called bent if $|W_f(\omega)| = 2^{n/2}$ for every ω. There are no bent functions for odd n, since $2^{n/2}$ is then not an integer.

Let f and b be functions on n and $2m$ disjoint variables, respectively. Their direct sum is $F(x, z) = f(x) + b(z)$, with addition modulo 2. Separating the two sums gives

$$W_F(\omega, \nu) = W_f(\omega) W_b(\nu).$$

If b is bent, then $|W_b(\nu)| = 2^m$, and consequently

$$\mathrm{NL}(F) = 2^{n+2m-1} - 2^m\big(2^{n-1} - \mathrm{NL}(f)\big). \tag{2.3}$$

2.3 Gauss periods and Gauss sums in characteristic two

Let an odd integer $q > 1$ divide $2^n - 1$, and let α be a primitive element of $\mathbb{F}_{2^n}$, so that its multiplicative order is $2^n - 1$. Set $N_n = (2^n - 1)/q$, $H = \langle \alpha^q \rangle$, $C_r = \alpha^r H = \{\alpha^{r+q\ell} : 0 \leq \ell < N_n\}$ with $0 \leq r < q$. Here $\langle \alpha^q \rangle$ is the multiplicative subgroup generated by α^q. The subgroup has order $N_n = (2^n - 1)/q$ and has q cosets, each of cardinality N_n. Coset indices are read modulo q: the elements α^a and α^b belong to the same coset if and only if $a \equiv b \pmod{q}$.

Definition 2.1. Put $\zeta = e^{2\pi \mathrm{i}/q}$, where $\mathrm{i}^2 = -1$. The Gauss periods of order q are

$$\eta_r = \sum_{x \in C_r} (-1)^{\mathrm{Tr}_1^n(x)} = \sum_{\ell=0}^{N_n - 1} (-1)^{\mathrm{Tr}_1^n(\alpha^{r+q\ell})}, \ 0 \leq r < q. \tag{2.4}$$

The corresponding Gauss sums are

$$G_j = \sum_{k=0}^{2^n - 2} \zeta^{jk} (-1)^{\mathrm{Tr}_1^n(\alpha^k)}, \ 0 \leq j < q. \tag{2.5}$$

We can see that a Gauss period is the difference between two trace counts in a coset, that is, $\eta_r = |\{x \in C_r : \mathrm{Tr}_1^n(x) = 0\}| - |\{x \in C_r : \mathrm{Tr}_1^n(x) = 1\}|$.

Gauss sums are root-of-unity weighted sums of these integers and allow their algebraic evaluation. Gauss periods in turn give the Walsh contributions of the cosets. The factor ζ^{jk} is the value of a multiplicative character at α^k. Since $|\mathbb{F}_{2^n}^*|$ is odd, every homomorphism from this group to $\{1, -1\}$ is trivial. The nontrivial multiplicative characters therefore require roots of unity other than ± 1.

Proposition 2.2. Let $q > 1$ be an odd divisor of $2^n - 1$, and let α be a primitive element of $\mathbb{F}_{2^n}$. Use the notation $N_n, H, C_r, \zeta, \eta_r, G_j$ of Definition 2.1 and the paragraph preceding it, all the indices should be read modulo q. Every period η_r is an integer with $\eta_r \equiv N_n \pmod 2$.

The periods and Gauss sums form a discrete Fourier transform pair:

$$G_j = \sum_{r=0}^{q-1} \eta_r \zeta^{jr}, \quad \eta_r = \frac{1}{q} \sum_{j=0}^{q-1} G_j \zeta^{-jr}. \tag{2.6}$$

The Gauss sums satisfy

$$\begin{gathered} G_0 = -1, \quad G_{2j} = G_j, \quad G_{-j} = \overline{G_j}, \\ |G_j| = 2^{n/2} \quad (1 \le j < q). \end{gathered} \tag{2.7}$$

Multiplication by an element of C_s shifts the period index, and the periods have the following total and sum of squares:

$$\begin{aligned} &\sum_{x \in C_r} (-1)^{\mathrm{Tr}_1^n(\omega x)} = \eta_{r+s} \qquad (\omega \in C_s), \\ &\sum_{r=0}^{q-1} \eta_r = -1, \qquad \sum_{r=0}^{q-1} \eta_r^2 = 2^n - N_n. \end{aligned} \tag{2.8}$$

Proof. Group k by its residue class modulo q to obtain the first identity in (2.6). The second follows from

$$\sum_{j=0}^{q-1} \zeta^{ja} = \begin{cases} q, & a = 0 \pmod q, \\ 0, & a \neq 0 \pmod q. \end{cases}$$

Note that $\sum_{x \in \mathbb{F}_{2^n}^*} (-1)^{\mathrm{Tr}_1^n(\omega x)} = -1$ for any $\omega \neq 0$. This gives $G_0 = -1$. Since $\mathrm{Tr}_1^n(x^2) = \mathrm{Tr}_1^n(x)$ and squaring permutes the field, $G_{2j} = G_j$. The additive factor is real, so complex conjugation replaces j by $-j$.

To prove the absolute value identity, expand $|G_j|^2$ and write the two nonzero variables as $\alpha^r \alpha^k$ and α^k:

$$|G_j|^2 = \sum_{r=0}^{2^n-2} \zeta^{jr} \sum_{k=0}^{2^n-2} (-1)^{\mathrm{Tr}_1^n((\alpha^r+1)\alpha^k)}.$$

The inner sum equals $2^n - 1$ for $r = 0$ and -1 otherwise. For $j \neq 0 \pmod q$, $\sum_{r=0}^{2^n-2} \zeta^{jr} = 0$, giving $|G_j|^2 = 2^n$. Integrality and parity follow directly from (2.4). Multiplication by $\omega \in C_s$ maps C_r bijectively to C_{r+s}, proving the first identity in (2.8). Finally, we apply Parseval's identity to the real-valued function $\mathbf{1}_H$. Its transform equals $2^{-n/2} N_n$ at zero and $2^{-n/2} \eta_s$ on C_s, so $N_n^2 + N_n \sum_{s=0}^{q-1} \eta_s^2 = 2^n N_n$. These are standard identities in the theory of Gauss sums over finite fields [11, Chapter 5]. □

2.4 Quadratic residue signs and Gauss sums

In this subsection, let $q > 3$ be prime with $q \equiv 3 \pmod 4$. Let $\mathcal{R}$ be the set of nonzero quadratic residues modulo q, and write $-\mathcal{R} = \{-r : r \in \mathcal{R}\}$. These sets partition the nonzero residue classes. Define the integers

$$e_r = \begin{cases} 0, & r = 0, \\ 1, & r \in \mathcal{R}, \\ -1, & r \in -\mathcal{R}. \end{cases} \tag{2.9}$$

A quadratic residue has the form $a^2 \pmod q$ with $a \not\equiv 0 \pmod q$. This classification concerns the coset indices, squaring in the field of characteristic two is itself a permutation. We have

$e_{-r} = -e_r$ and $\sum_{r=0}^{q-1} e_r = 0$. The standard quadratic Gauss sum evaluation [11, Chapter 5] gives

$$\sum_{r=0}^{q-1} e_r \zeta^{-jr} = \begin{cases} 0, & j = 0, \\ -\mathrm{i}\sqrt{q}, & j \in \mathcal{R}, \\ \mathrm{i}\sqrt{q}, & j \in -\mathcal{R}. \end{cases}$$

Suppose $\mathrm{ord}_q(2) = (q-1)/2$, where $\mathrm{ord}_q(2)$ is the least positive integer d such that $2^d \equiv 1 \pmod q$. Then $\langle 2 \rangle = \mathcal{R}$ has order $(q-1)/2$ in the multiplicative group of nonzero residues modulo q. By (2.7), G_j is constant on $\mathcal{R}$ and takes the conjugate value on $-\mathcal{R}$.

Theorem 2.3. Let $q > 3$ be prime, $q \equiv 3 \pmod 4$, and $n = (q-1)/2 = \mathrm{ord}_q(2)$. Let h be the class number of $\mathbb{Q}(\sqrt{-q})$. For G_1 in (2.5), there are odd integers a, b such that

$$G_1 = 2^{(n-h)/2-1}(a + b\mathrm{i}\sqrt{q}), \text{ where } a^2 + qb^2 = 2^{h+2}, \qquad a2^{(n-h)/2} \equiv -2 \pmod q.$$

This is the prime-order, characteristic-two case of [18, Lemma 5] (see also [20, Theorem 4.1]). The class number is the number of reduced primitive positive definite binary quadratic forms of discriminant $-q$ [3].

Lemma 2.4. Under the assumptions of Theorem 2.3, a primitive element α can be chosen so that $\mathrm{Im}\, G_1 > 0$.

Proof. The integer b in Theorem 2.3 is odd, so $\mathrm{Im}\, G_1 \neq 0$. The inverse of a primitive element is primitive. Writing $G_1(\alpha)$ to indicate the choice and replacing k by $-k$ modulo $2^n - 1$ gives

$$G_1(\alpha^{-1}) = \sum_{k=0}^{2^n-2} \zeta^{-k}(-1)^{\mathrm{Tr}_1^n(\alpha^k)} = \overline{G_1(\alpha)}.$$

Exactly one of α and α^{-1} therefore gives the required sign of the imaginary part. Thus one may start with any primitive element and replace it by its inverse if $\mathrm{Im}\, G_1 < 0$. □

This choice fixes the signs of the coset indices, not a unique primitive element or a particular representation of the field. All subsequent coset indices use the same choice.

We use the following unequal-threshold form of the argument in [15, Proposition 3.6].

Lemma 2.5 (Linear feasibility for sign vectors)**.** Let $N > 4$ and $m \geq 1$ be integers, let $a_1, \ldots, a_m \in [-1, 1]^N$, and let $\Delta_1, \ldots, \Delta_m > 0$. Put $\phi(t) = 5/\sinh(t^2/3)$ for $t > 0$. If

$$\sum_{j=1}^{m} \phi\left(\frac{\Delta_j}{\sqrt{N}}\right) \leq N,$$

then there exists $x \in \{-1, 1\}^N$ such that $\langle a_j, x \rangle \leq \Delta_j + 4$ for every j.

Proof. We set $R(z) = N - \|z\|_2^2$ and $s(z) = R(z) - 4$. We use the domain K and potential Φ of [15, equations (3)-(4)], replacing the threshold in each row constraint and corresponding summand by Δ_j, and summing over the given m rows. We retain $\phi(0) = +\infty$ and $\Phi = 0$ when $s = 0$. By [15, Lemma 3.3(d)], we have

$$\Phi(0) = \frac{1}{N-4} \sum_j \phi\left(\frac{\Delta_j}{\sqrt{N-4}}\right) < \frac{1}{N} \sum_j \phi\left(\frac{\Delta_j}{\sqrt{N}}\right) \leq 1.$$

The compactness argument of [15, Proposition 3.6] gives a minimizer $z \in K$ with $\Phi(z) < \infty$. The scalar estimates and termwise differentiation in [15, Lemma 3.3 and proof of Lemma 3.5] apply separately to each row gap, while [15, Lemma 3.4] allows arbitrary row families. Hence the same minimizer argument gives $R(z) = 4$ and $\langle a_j, z \rangle \leq \Delta_j$. Rounding z to its nearest sign vector x gives $\langle a_j, x - z \rangle \leq \|x - z\|_1 \leq R(z) = 4$, proving the claim. □

3 Construction and nonlinearity

In this section, we construct the support of a function on 375 variables with nonlinearity greater than that of the direct sum of a 15-variable Patterson-Wiedemann function and a bent function on 360 variables. We seek a highly nonlinear function in an odd number of variables $n \geq 21$ with support $B \cup \bigcup_{i \in S} \alpha^i H$, where H is a subgroup of $\mathbb{F}_{2^n}^*$ and $B \subseteq H$ remains to be chosen and S specifies the values on the other multiplicative cosets. The fixed spectral contribution should be both explicitly computable and sufficiently small to leave room for the choice of B. We therefore consider the prime-order index-two family $q > 3$, $n = \mathrm{ord}_q(2) = (q-1)/2$ odd. The nonzero labels modulo q split into the quadratic-residue and nonresidue Frobenius orbits. Write $G_1 = A_G + \mathrm{i}B_G\sqrt{q}$ as in Theorem 2.3, with integer coefficients. The nontrivial Gauss sums take this value on the quadratic-residue orbit and its conjugate on the nonresidue orbit. With e_i denoting the Legendre symbol as in (2.9), the inverse Fourier transform is

$$\eta_i = \frac{-1 - A_G}{q} + B_G e_i + A_G \delta_{i,0}. \tag{3.1}$$

Thus two integers and a Legendre-symbol table determine the contribution of any label set in every nonzero frequency class. More precisely, for labels $u_0 = 0$ and $u_i \in \{0, 1\}$, the fixed signed contribution outside H is $P_u(r) = \sum_{i=1}^{q-1}(1 - 2u_i)\eta_{i+r}$. Removing label p and adding label v changes it by

$$\Delta P_u(r) = 2\eta_{p+r} - 2\eta_{v+r}.$$

Every frequency class can therefore be updated exactly without repeating the Gauss-sum evaluation. Few label changes make the update inexpensive and whether the peak decreases must be checked from the updated values. The remaining contribution from B is subsequently controlled by applying Lemma 2.5, and need not be constant within a frequency coset.

We next screen for a small fixed contribution, using the initial templates $S = c + \mathcal{R}$ or $S = c - \mathcal{R}$ with $0 \notin S$ and $|S| = (q-1)/2$. Among the 31 basic prime-order index-two candidates with $20 \leq n \leq 500$, the first dimension for which at least one such template satisfies $\max_r |P_u(r)|/2^{n/2} < 1.193239$ is $n = 375$, with $q = 751$. The smallest such peak for $n < 375$ is $17730/\sqrt{2^{27}} > 1.193239$, attained at $n = 299$. The unmodified template at 375 has peak $388/\sqrt{2^{17}}$. This is a finite comparison of the stated templates, including zero translations, rather than a minimality claim over all labels or constructions. For $q = 751$, choose the orientation $\mathrm{Im}\, G_1 > 0$ of Lemma 2.4 and then we obtain $G_1 = 2^{179}(-361 + \mathrm{i}\sqrt{751})$. Starting from $3 - \mathcal{R}$, the exchange $505 \mapsto 130$ preserves the label weight and reduces the fixed peak from $388 \cdot 2^{179}$ to $384 \cdot 2^{179}$. The resulting full distribution of the fixed contribution is used in the proof of Theorem 3.4 to verify the hypothesis of Lemma 2.5.

3.1 Basic results

Throughout this section, coset indices are read modulo 751, and the trace is Tr_1^{375} as in Section 2. Put $\mathcal{R} = \{2^j \bmod 751 : 0 \leq j < 375\}$ and $\zeta = e^{2\pi\mathrm{i}/751}$. Choose a primitive element α of $\mathbb{F}_{2^{375}}$ such that $\mathrm{Im}\, G_1 > 0$, as in Lemma 2.4, where G_1 is defined by (2.5) with $q = 751$ and $n = 375$. Let $H = \langle \alpha^{751} \rangle$ and $N = |H| = (2^{375} - 1)/751$.

Definition 3.1. On the representatives α^i, $0 \leq i \leq 750$, define

$$g(\alpha^i) = \begin{cases} 1, & i \in \big((3 - \mathcal{R}) \setminus \{505\}\big) \cup \{130\}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

Extend g to a 0, 1-valued function on $\mathbb{F}_{2^{375}}^*$ by setting $g(\alpha^i h) = g(\alpha^i)$ for $h \in H$. Thus g is constant on each coset.

Lemma 3.1. The number 751 is prime, $\mathrm{ord}_{751}(2) = 375$, and the class number of $\mathbb{Q}(\sqrt{-751})$ is 15. The norm and congruence conditions in Theorem 2.3 give $a = -361$ and $|b| = 1$. The function in (3.2) satisfies

$$g(1) = 0, \qquad \sum_{i=1}^{750} g(\alpha^i) = 375, \qquad \sum_{i=1}^{750} (-1)^{g(\alpha^i)} = 0. \tag{3.3}$$

Let $g_i = g(\alpha^i)$ and define e_i by (2.9). The integers

$$\kappa_r = 361(2g_r - 1 + \delta_{r,0}) + e_r + 2\sum_{x=0}^{750} e_x g_{r-x} \tag{3.4}$$

have the distribution in Table 1.

Proof. It can be easily checked that 751 is a prime, and $2^{375} \equiv 1$, $2^{125} \equiv 678$ and $2^{75} \equiv 569 \pmod{751}$. Since the only prime divisors of 375 are 3 and 5, the latter two congruences exclude every proper divisor as the order. Thus $\mathrm{ord}_{751}(2) = 375$, and $\mathcal{R}$ consists of the 375 nonzero quadratic residues.

By the theory of reduced positive definite binary quadratic forms [3], the class number for discriminant -751 is the number of reduced forms satisfying $|b| \le a \le c$, $b^2 - 4ac = -751$ and $\gcd(a, b, c) = 1$, with $b \ge 0$ when $|b| = a$ or $a = c$. Since $a \le \sqrt{751/3} < 16$, only $1 \le a \le 15$ need be considered. The complete list is

$$(1, 1, 188),\ (2, \pm 1, 94),\ (4, \pm 1, 47), (5, \pm 3, 38),\ (8, \pm 7, 25),$$
$$(10, \pm 7, 20), (10, \pm 3, 19),\ (13, \pm 9, 16).$$

There are 15 forms. For the norm equation $a^2 + 751b^2 = 2^{17}$, it is enough to check odd $|b|$ between 1 and 13. Only $|b| = 1$ gives an odd square a^2, namely 361^2. The congruence $a2^{180} \equiv -2 \pmod{751}$ then gives $a = -361$. Equivalently, $361 \cdot 2^{179} \equiv 1 \pmod{751}$.

Computations modulo 751 give $0 \notin 3 - \mathcal{R}$, $505 \in 3 - \mathcal{R}$ and $130 \notin 3 - \mathcal{R}$. Thus $g_0 = 0$ and exactly 375 of the values g_i equal 1, proving (3.3). A computer verification of the 751 integer values in (3.4) gives the complete distribution in Table 1. It also gives $\kappa_0 = -6$, $\sum_{r=0}^{750} \kappa_r = 0$, $\max_{0 \le r \le 750} |\kappa_r| = 384$ and $\sum_{r=0}^{750} \kappa_r^2 = 750 \cdot 2^{17}$. □

Table 1: Distribution of the integers κ_r.

Value	−368	−366	−364	−362	−360	−358	−6
Multiplicity	34	52	112	88	42	46	1
Value	356	358	360	362	364	366	384
Multiplicity	52	48	88	87	46	54	1

For $0 \le i \le 750$, write

$$\eta_i = \sum_{x \in \alpha^i H} (-1)^{\mathrm{Tr}_1^{375}(x)}, \qquad A_i = \sum_{r=1}^{750} (-1)^{g(\alpha^r)} \eta_{r+i}.$$

Theorem 3.2. The Gauss periods are

$$\eta_i = \begin{cases} (-1 - 270750 \cdot 2^{179})/751, & i = 0, \\ (-1 + 1112 \cdot 2^{179})/751, & i \in \mathcal{R}, \\ (-1 - 390 \cdot 2^{179})/751, & i \in -\mathcal{R}. \end{cases} \tag{3.5}$$

The signed sums satisfy $A_i = 2^{179}\kappa_{-i}$, and hence

$$|A_i| \le 384 \cdot 2^{179} < 1.0607 \cdot 2^{375/2}. \tag{3.6}$$

For every $\omega \in \alpha^i H$,

$$-2 \sum_{\substack{1 \le r \le 750 \\ g(\alpha^r)=1}} \sum_{x \in \alpha^r H} (-1)^{\mathrm{Tr}_1^{375}(\omega x)} = 1 + A_i + \eta_i.$$

The union of the 375 selected cosets has exactly $375N$ elements.

Proof. Lemma 3.1 and Theorem 2.3 give $G_1 = 2^{179}(-361 \pm \mathrm{i}\sqrt{751})$. The choice $\mathrm{Im}\, G_1 > 0$ gives $G_1 = 2^{179}(-361 + \mathrm{i}\sqrt{751})$. Then G_j takes this value on $\mathcal{R}$ and its conjugate on $-\mathcal{R}$. Substitute these values into (2.6) and use the quadratic Gauss sum evaluation to obtain (3.5), or equivalently

$$\eta_i = \frac{-1 + 361 \cdot 2^{179}}{751} + 2^{179} e_i - 361 \cdot 2^{179} \delta_{i,0}. \tag{3.7}$$

The constant term cancels by (3.3), giving

$$2^{-179} A_i = \sum_{r=1}^{750} (-1)^{g_r} (e_{r+i} - 361 \delta_{r+i,0}) = \kappa_{-i}.$$

The last equality uses $(-1)^{g_r} = 1 - 2g_r$, $g_0 = 0$ and $e_{-r} = -e_r$. Table 1 gives (3.6). Finally, we use the period translation identity and $\sum_{r=0}^{750} \eta_r = -1$ to obtain

$$-2 \sum_{\substack{1 \le r \le 750 \\ g_r = 1}} \eta_{r+i} = A_i - \sum_{r=1}^{750} \eta_{r+i} = 1 + A_i + \eta_i.$$

The 375 selected cosets are pairwise disjoint, so their union has cardinality $375N$. □

Using (3.7), integer computation also verifies $A_i = 2^{179}\kappa_{-i}$ for $0 \le i \le 750$, $\sum_{i=0}^{750} \eta_i = -1$ and $\sum_{i=0}^{750} \eta_i^2 = 2^{375} - N$.

Definition 3.3. With η_i and A_i as above, the constraint problem is to choose a single subset $B \subseteq H$ satisfying

$$\big|N - 2|B|\big| \le 1.193239 \cdot 2^{375/2} + 4, \tag{3.8}$$

$$\left| \eta_i + A_i - 2 \sum_{x \in B} (-1)^{\mathrm{Tr}_1^{375}(\omega x)} \right| \le 1.193239 \cdot 2^{375/2} + 4 \tag{3.9}$$

for every $0 \le i \le 750$ and every $\omega \in \alpha^i H$. The first constraint concerns zero and the second concerns every nonzero point.

Theorem 3.4. The constraint problem in Definition 3.3 has a solution $B \subseteq H$.

Proof. We put $M = 2^{375}$ and $C = 1193239/10^6$. We index the N coordinates by $h \in H$ and the rows by $\omega \in \mathbb{F}_{2^{375}}$, and set

$$a_{\omega,h} = (-1)^{\mathrm{Tr}_1^{375}(\omega h)}, \qquad d_0 = 0, \qquad d_\omega = A_i \quad (\omega \in \alpha^i H).$$

For each ω, we use the two rows a_ω and $-a_\omega$ in Lemma 2.5, with respective thresholds

$$\Delta_\omega^+ = C\sqrt{M} - d_\omega, \qquad \Delta_\omega^- = C\sqrt{M} + d_\omega.$$

Table 2: Scaled offset magnitudes and their multiplicities among nonzero frequency classes.

m	6	356	358	360	362	364	366	368	384
ν_m	1	52	94	130	175	158	106	34	1

Both thresholds are positive since $|d_\omega|/\sqrt{M} \le 384/\sqrt{131072} < C$. From Table 1, we see that the number of nonzero frequencies ω for which $|d_\omega| = m2^{179}$ is $N\nu_m$, where the values of m and ν_m are listed in Table 2. These multiplicities satisfy $\sum_m \nu_m = 751$. As $M/N > 751$ and ϕ is decreasing, the normalized sum in that lemma is bounded by

$$\begin{aligned}&\frac{1}{N}\sum_{\omega}\left[\phi\left(\frac{\Delta_\omega^+}{\sqrt{N}}\right)+\phi\left(\frac{\Delta_\omega^-}{\sqrt{N}}\right)\right]\\&\quad<\frac{2}{N}+\sum_m \nu_m\left[\phi\left(\sqrt{751}\left(C-\frac{m}{\sqrt{131072}}\right)\right)+\phi\left(\sqrt{751}\left(C+\frac{m}{\sqrt{131072}}\right)\right)\right]<1.\end{aligned}$$

Here the two rows corresponding to zero frequency contribute less than $2/N$ in total because $\phi(C\sqrt{751}) < 1$. Section 3.2 verifies that the nine-class sum is less than 0.999856, while $2/N < 10^{-6}$. The lemma therefore supplies $x \in \{-1,1\}^H$ with

$$|\langle a_\omega, x\rangle + d_\omega| \le C\sqrt{M} + 4 \qquad (\omega \in \mathbb{F}_{2^{375}}).$$

We take $B = \{h \in H : x_h = -1\}$. At zero, $\langle a_0, x\rangle = N - 2|B|$. For $\omega \in \alpha^i H$, $\langle a_\omega, x\rangle = \eta_i - 2\sum_{h\in B}(-1)^{\mathrm{Tr}_1^{375}(\omega h)}$. These are exactly the required constraints. □

3.2 Verification of the numerical bounds

We set $M = 2^{375}$, $N = (M-1)/751$, $C = 1193239/10^6$, and $\phi(t) = 5/\sinh(t^2/3)$ for $t > 0$. We verify the following bounds using only rational arithmetic. We put

$$a = 6678434726570384492722, \qquad L = \frac{a}{2^{64}}, \qquad U = \frac{a+1}{2^{64}}.$$

By direct integer comparison, we obtain $L^2 < 131072 < U^2$. For each scaled offset magnitude m in Table 2, we define

$$z_m^- = \min\left\{32, \frac{\left\lfloor 10^{12}\frac{751}{3}(C - m/L)^2\right\rfloor}{10^{12}}\right\},$$
$$z_m^+ = \min\left\{32, \frac{\left\lfloor 10^{12}\frac{751}{3}(C + m/U)^2\right\rfloor}{10^{12}}\right\}.$$

All the quantities $C - m/L$ are positive. Thus $z_m^\pm$ are positive lower bounds for $\frac{751}{3}(C \pm m/\sqrt{131072})^2$. The positive Taylor polynomial

$$S(z) = \sum_{j=0}^{63}\frac{z^{2j+1}}{(2j+1)!} < \sinh z \qquad (z > 0)$$

therefore gives the strict upper bounds in the last column of Table 3, obtained by rounding $\nu_m(5/S(z_m^-) + 5/S(z_m^+))$ upwards to multiples of 10^{-9}.

Since $M/N > 751$ and ϕ decreases, the normalized contribution from all nonzero frequencies is consequently less than 0.999856. At zero frequency, $751C^2/3 > 32$ and $S(32) > 5$ show that each of the two potential terms is less than 1. Their total normalized contribution is less than $2/N < 10^{-6}$. The full normalized potential is therefore less than $0.999857 < 1$, as required.

Table 3: Rational upper bounds for the paired potential contributions.

m	ν_m	Strict upper bound for $\nu_m \sum_{\pm} \phi\Big(\sqrt{751}(C \pm m/\sqrt{131072})\Big)$
6	1	0.000000001
356	52	0.008418549
358	94	0.026989444
360	130	0.065193753
362	175	0.150959898
364	158	0.230890417
366	106	0.258430833
368	34	0.136198212
384	1	0.122774042
	Sum	0.999855149

The integer deficit follows from the exact comparison

$$8 \cdot 442367^2 \cdot 10^{12} - (1193239 \cdot 2^{20} + 10^6)^2 = 110677890982703104 > 0.$$

Taking positive square roots of both sides of the resulting inequality and rearranging, we obtain $442367/2^{37/2} > C + 2^{-20}$. Since $5 + 2^{100} < 2^{167} < 2^{-20}\sqrt{M}$, we also obtain

$$C\, 2^{375/2} + 5 + 2^{100} < 442367 \cdot 2^{169}. \tag{3.10}$$

Finally, we have

$$2^{374} - 442367 \cdot 2^{168} = 2^{374} - 2^{187} + 81921 \cdot 2^{168}.$$

For $n = 375 + 2k$, multiplying the gain by 2^k gives $81921 \cdot 2^{(n-39)/2}$, whereas the Patterson-Wiedemann gain is

$$20 \cdot 2^{(n-15)/2} = 81920 \cdot 2^{(n-39)/2}.$$

All the comparisons above can be readily verified by computer using exact rational and integer arithmetic.

3.3 Main construction

Construction 1. Let α be the primitive element of $\mathbb{F}_{2^{375}}$ fixed in Section 3.1, and put $H = \langle \alpha^{751} \rangle$ and $N = |H| = (2^{375} - 1)/751$. Define the Boolean function $f_B : \mathbb{F}_{2^{375}} \to \mathbb{F}_2$ by its support

$$\operatorname{supp}(f_B) = B \cup \bigcup_{\substack{1 \le i \le 750 \\ g(\alpha^i) = 1}} \alpha^i H, \tag{3.11}$$

where g is given in Definition 3.1 and B is any solution of the constraint problem in Definition 3.3.

3.4 Nonlinearity

Theorem 3.5. Every function f_B in Construction 1 satisfies

$$\mathrm{NL}(f_B) > 2^{374} - 442367 \cdot 2^{168} = \mathcal{N}_{\mathrm{BC}}(375) + 81921 \cdot 2^{168} > \mathcal{N}_{\mathrm{PW}}(375). \tag{3.12}$$

Proof. Theorem 3.4 guarantees the existence of a subset B for Construction 1. We first consider the values of the Walsh transform at points $\omega \neq 0$. For $\omega \in \alpha^i H$, equations (3.11) and (2.2), together with Theorem 3.2, give

$$W_{f_B}(\omega) = 1 + A_i + \eta_i - 2\sum_{x\in B}(-1)^{\mathrm{Tr}_1^{375}(\omega x)}.$$

By (3.9), we obtain

$$\max_{\omega\neq 0}|W_{f_B}(\omega)| \leq 1.193239 \cdot 2^{375/2} + 5. \tag{3.13}$$

We now consider the value of the Walsh transform at the point $\omega = 0$. The 375 prescribed cosets and B are pairwise disjoint, so

$$|\operatorname{supp}(f_B)| = 375N + |B|,$$
$$W_{f_B}(0) = 2^{375} - 2(375N + |B|) = 1 + N - 2|B|.$$

By (3.8), we have

$$|W_{f_B}(0)| \leq 1.193239 \cdot 2^{375/2} + 5.$$

The exact margin verified in Section 3.2, equation (3.10), gives

$$\max_{\omega}|W_{f_B}(\omega)| < 884734 \cdot 2^{168}. \tag{3.14}$$

The Walsh characterization of nonlinearity now proves (3.12). Since $442367 = 27 \cdot 2^{14} - 1$, its lower-bound threshold exceeds the PW value by 2^{168}. □

3.5 Direct sums with bent functions

Theorem 3.6. For every odd $n \geq 375$, there exists a Boolean function F on n variables satisfying

$$\begin{aligned}\mathrm{NL}(F) &> 2^{n-1} - 2^{(n-1)/2} + 81921 \cdot 2^{168} \cdot 2^{(n-375)/2} \\ &= \mathcal{N}_{\mathrm{BC}}(n) + 81921 \cdot 2^{(n-39)/2} > \mathcal{N}_{\mathrm{PW}}(n).\end{aligned}$$

Proof. Write $n = 375+2m$. Take a subset B from Theorem 3.4 and define f_B by Construction 1. Its nonlinearity is bounded by Theorem 3.5. For $m \geq 1$, take any bent function b on $2m$ variables and put

$$F(x, z) = f_B(x) + b(z) \pmod 2.$$

For $m = 0$, take $F = f_B$. Equation (2.3) gives

$$\begin{aligned}\mathrm{NL}(F) &= 2^{n-1} - 2^m\big(2^{374} - \mathrm{NL}(f_B)\big) \\ &> 2^{n-1} - 442367 \cdot 2^{168+m} = \mathcal{N}_{\mathrm{BC}}(n) + 81921 \cdot 2^{168+m}.\end{aligned}$$

Since $20 \cdot 2^{(n-15)/2} = 81920 \cdot 2^{(n-39)/2}$, the displayed lower bound exceeds the PW value by $2^{(n-39)/2}$. □

We have specified the cosets, calculated their contribution and proved the existence of a single subset B. Computing all of B, finding a short membership criterion and reducing the number of variables remain open tasks for this construction.

We summarize the nonlinearity gains of the ordinary constructions in Table 4. If an odd-dimensional function on u variables has $\mathrm{NL}(f) = \mathcal{N}_{\mathrm{BC}}(u) + x$, its bent direct sums have

$$\mathrm{NL}(f \oplus b) = \mathcal{N}_{\mathrm{BC}}(n) + x\, 2^{(n-u)/2}. \tag{3.15}$$

For any odd $u \le v \le n$, the identity

$$x\,2^{(n-u)/2} = \big(x\,2^{(v-u)/2}\big)2^{(n-v)/2}$$

merely changes the reference dimension. A strict improvement requires a new v-variable function with gain $y > x\,2^{(v-u)/2}$. Such a function exists for all sufficiently large odd v by Schmidt's asymptotic theorem [18, Theorems 1 and 2], even if the new function is required to be balanced. Indeed, the normalized deficit $\gamma = (2^{u-1} - \mathrm{NL}(f))/2^{u/2-1}$ is strictly greater than 1 by Parseval and the integrality of the Walsh coefficients when u is odd. The asymptotic theorem supplies a balanced function with normalized deficit below γ and its integer gain is then such a y. Taking bent direct sums gives the strict improvement for every odd $n \ge v$. The assertion is for sufficiently large v, not for every $v > u$.

Table 4: Nonlinearity written as $\mathcal{N}_{\mathrm{BC}}(n)$ plus an integer gain. The first three rows give exact gains for the indicated seeds and their bent direct sums. The last row gives a strict lower bound.

Seed (u, NL)	Odd n	Gain over $\mathcal{N}_{\mathrm{BC}}(n)$	Gain at $n = 375$
$(9, 241)$ [7]	$n \ge 9$	$2^{(n-9)/2}$	$32768 \cdot 2^{168}$
$(9, 242)$ [8]	$n \ge 9$	$2 \cdot 2^{(n-9)/2}$	$65536 \cdot 2^{168}$
$(15, 16276)$ [13]	$n \ge 15$	$20 \cdot 2^{(n-15)/2}$	$81920 \cdot 2^{168}$
This work, $u = 375$	$n \ge 375$	$> 81921 \cdot 2^{(n-39)/2}$	$> 81921 \cdot 2^{168}$

4 Balanced functions by changing a small proportion of values

The preceding construction does not require balancedness. In this section, we start with any B from Construction 1 and change some values of f_B to obtain a balanced function. The number of changes is small relative to 2^{375} and we do not claim that these points can be explicitly listed in practice.

4.1 Distance to balancedness and a quantitative modification lemma

For any $B \subseteq H$,

$$W_{f_B}(0) = 1 + N - 2|B|, \qquad |\operatorname{supp}(f_B)| = 375N + |B|.$$

Thus f_B is balanced if and only if $|B| = (N+1)/2$. Its minimum Hamming distance from a balanced function is

$$d_{\mathrm{bal}}(f_B) = \frac{|W_{f_B}(0)|}{2} = \frac{|1 + N - 2|B||}{2}.$$

This is the absolute difference between the support size and 2^{374}, attained by changing the required number of values on the more frequent side. For B in Construction 1, the proof of Theorem 3.5 gives

$$d_{\mathrm{bal}}(f_B) \le \tfrac{1}{2}\big(1.193239 \cdot 2^{375/2} + 5\big) < 2^{375/2}.$$

The worst-case bound of 2 for each point's effect on a Walsh coefficient is insufficient to preserve our strict improvement. Instead, sampling without replacement controls the effect of one set of changes at every point.

Lemma 4.1. Let $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$, $M = 2^n$, $n \ge 1$, and $d = |W_f(0)|/2$. If $d = 0$, then f is balanced. If $d > 0$, there is a balanced function $\widetilde{f}$ such that $d_H(f, \widetilde{f}) = d$ and

$$\max_{\substack{\omega \in \mathbb{F}_{2^n} \\ \omega \ne 0}} |W_{\widetilde{f}}(\omega)| < \left(1 - \frac{d}{M/2 + d}\right) \max_{\substack{\omega \in \mathbb{F}_{2^n} \\ \omega \ne 0}} |W_f(\omega)| + \sqrt{8d \log(4M)}. \tag{4.1}$$

Here log denotes the natural logarithm, and d is the minimum possible number of changes required for balancedness.

Proof. It suffices to treat $d > 0$. Put $\sigma = \operatorname{sgn} W_f(0) \in \{-1, 1\}$ and

$$A = \{x \in \mathbb{F}_{2^n} : (-1)^{f(x)} = \sigma\}, \qquad a = |A| = M/2 + d.$$

Choose uniformly without replacement a subset $D \subseteq A$ of exactly d points, and put $\widetilde{f} = f + \mathbf{1}_D$ modulo 2. Changing d values on the more frequent side makes both sides have size $M/2$. Every such choice therefore makes the function balanced.

For $\omega \neq 0$, the identity $\mathbf{1}_A = (1 + \sigma(-1)^f)/2$ and character orthogonality give

$$\sum_{x \in A} (-1)^{\mathrm{Tr}_1^n(\omega x)} = \frac{\sigma}{2} W_f(\omega).$$

Moreover,

$$W_{\widetilde{f}}(\omega) = W_f(\omega) - 2\sigma \sum_{x \in D} (-1)^{\mathrm{Tr}_1^n(\omega x)},$$

so

$$\mathbb{E} W_{\widetilde{f}}(\omega) = \left(1 - \frac{d}{a}\right) W_f(\omega). \tag{4.2}$$

The sampling comparison in [5, Theorem 4] yields Hoeffding's inequality without replacement [1, Proposition 1.2]. By combining the upper- and lower-tail bounds, we find that a sum of $d > 0$ samples from a finite population in $[-1, 1]$ deviates from its mean by at least $t > 0$ with probability at most $2\exp(-t^2/(2d))$. Applied to the character sum, it gives

$$\mathbb{P}\Big(|W_{\widetilde{f}}(\omega) - \mathbb{E} W_{\widetilde{f}}(\omega)| \geq u\Big) \leq 2e^{-u^2/(8d)}.$$

Take $u = \sqrt{8d \log(4M)}$ and apply the union bound to the $M-1$ nonzero points. The probability that any constraint fails is at most

$$2(M-1)e^{-u^2/(8d)} = \frac{M-1}{2M} < \frac{1}{2}.$$

Thus a single D makes every deviation smaller than u. Combining this with (4.2) and $0 \leq 1 - d/a \leq 1$ proves (4.1). The rows need not be independent. The difference in support sizes proves that d is minimal. □

4.2 Balanced functions on 375 variables and their direct sums

Theorem 4.2. Take any B from Construction 1 and the associated function f_B. There is a subset $D \subseteq \mathbb{F}_{2^{375}}$ of cardinality $|D| = |1 + N - 2|B||/2$ such that $\widetilde{f}_B = f_B + \mathbf{1}_D$ modulo 2 is balanced and

$$\mathrm{NL}(\widetilde{f}_B) > 2^{374} - 442367 \cdot 2^{168} = \mathcal{N}_{\mathrm{BC}}(375) + 81921 \cdot 2^{168} > \mathcal{N}_{\mathrm{PW}}(375). \tag{4.3}$$

Consequently, for every odd $n \geq 375$, there exists a balanced Boolean function on n variables whose nonlinearity exceeds the Patterson-Wiedemann direct-sum bound.

Proof. Again put $M = 2^{375}$. If $W_{f_B}(0) = 0$, take $D = \varnothing$ and apply Theorem 3.5. Otherwise put $d = |W_{f_B}(0)|/2 < \sqrt{M}$ and apply Lemma 4.1. We set $u = \sqrt{8d \log(4M)}$. Since $\log(4M) = 377 \log 2 < 377$, we have

$$u^2 = 8d \log(4M) < 3016\sqrt{M} < 2^{200}.$$

Here $3016 < 2^{12}$ and $2^{12}\sqrt{M} = 2^{399/2} < 2^{200}$. Hence, we obtain $u < 2^{100}$. Combining this estimate with the bound (3.13) and the exact comparison (3.10), we obtain

$$\max_{\substack{\omega\in\mathbb{F}_{2^{375}}\\ \omega\neq 0}} |W_{\widetilde{f}_B}(\omega)| < \max_{\omega\neq 0} |W_{f_B}(\omega)| + 2^{100} < 884734 \cdot 2^{168}.$$

Since $W_{\widetilde{f}_B}(0) = 0$, the Walsh characterization of nonlinearity gives (4.3). □

The modification set D lies on the side where f_B takes its more frequent value. If $W_{f_B}(0) > 0$, the new support is $\mathrm{supp}(f_B) \cup D$. On the other hand, if $W_{f_B}(0) < 0$, it is $\mathrm{supp}(f_B) \setminus D$. Thus $\widetilde{f}_B$ need not retain the union-of-cosets structure, and the origin may be modified. We have proved the existence of one modification set satisfying all bounds and having the minimum cardinality, but have not listed its elements.

Theorem 4.3. For every $n = 375 + 2m$ with an integer $m \geq 0$, there is a balanced Boolean function F on n variables such that

$$\begin{aligned} \mathrm{NL}(F) &> 2^{n-1} - 2^{(n-1)/2} + 81921 \cdot 2^{168} \cdot 2^{(n-375)/2} \\ &= \mathcal{N}_{\mathrm{BC}}(n) + 81921 \cdot 2^{(n-39)/2} > \mathcal{N}_{\mathrm{PW}}(n). \end{aligned} \tag{4.4}$$

Proof. For $m = 0$, use Theorem 4.2. For $m \geq 1$, take its balanced function $\widetilde{f}_B$ and any bent function b on $2m$ disjoint variables. Each slice $x \mapsto \widetilde{f}_B(x) + b(z)$ is balanced, so $F(x, z) = \widetilde{f}_B(x) + b(z)$ is balanced. Equation (2.3) gives $\mathrm{NL}(F) > 2^{n-1} - 442367 \cdot 2^{168+m}$, which is the asserted bound. □

We now compare known constructions of balanced functions in odd numbers of variables. Reported high nonlinearities for balanced functions include 240 in nine variables and 16272 in fifteen variables [2, Table 3.2][17] and we do not assert that these values are proven global maxima. Their bent direct sums have gains 0 and $16 \cdot 2^{(n-15)/2}$ over $\mathcal{N}_{\mathrm{BC}}(n)$, respectively. One can instead combine an unbalanced function with a balanced function on an even number of variables. For a chosen balanced h_e on even $e \geq 2$ variables, write

$$d_e = 2^{e-1} - \mathrm{NL}(h_e), \qquad \epsilon_e = d_e - 2^{e/2-1} > 0. \tag{4.5}$$

The strict inequality holds because a balanced function is not bent. The Walsh product identity shows that $f \oplus h_e$ is balanced and

$$\mathrm{NL}(f \oplus h_e) = 2^{u+e-1} - 2\big(2^{u-1} - \mathrm{NL}(f)\big)d_e. \tag{4.6}$$

This gives Table 5. The row using the nine-variable function of nonlinearity 242 is always strictly below the balanced fifteen-variable function plus bent functions whenever both constructions apply, since $2 \cdot 2^{(n-9)/2} = 16 \cdot 2^{(n-15)/2}$ and $\epsilon_{n-9} > 0$. The PW function of nonlinearity 16276 combined with h_e, where $e = n - 15$, is better than the balanced fifteen-variable function plus bent functions precisely when

$$\epsilon_e < 2^{e/2}/54. \tag{4.7}$$

It is always below the ordinary PW direct-sum value, and hence below our bound for $n \geq 375$.

4.3 A slightly different direct-sum approach

Dobbertin's construction [4, Proposition 8] produces a balanced function on $2s$ variables from any balanced function on s variables, for every integer $s \geq 1$, with deficit

$$d_{2s} = 2^{s-1} + d_s. \tag{4.8}$$

Table 5: Nonlinearities of balanced constructions, written as $\mathcal{N}_{\mathrm{BC}}(n)$ plus a gain. Here h_e is any chosen balanced even-dimensional function and ϵ_e is defined in (4.5). The first four rows provide exact gains for the specified functions and the last row gives a strict lower bound.

Construction	Odd n	Gain over $\mathcal{N}_{\mathrm{BC}}(n)$
Balanced $(9, 240)$ plus bent	$n \geq 9$	0
Balanced $(15, 16272)$ plus bent	$n \geq 15$	$16 \cdot 2^{(n-15)/2}$
Unbalanced $(9, 242)$ plus h_{n-9}	$n \geq 11$	$2 \cdot 2^{(n-9)/2} - 28\epsilon_{n-9}$
PW $(15, 16276)$ plus h_{n-15}	$n \geq 17$	$20 \cdot 2^{(n-15)/2} - 216\epsilon_{n-15}$
This work, balanced 375-variable function plus bent	$n \geq 375$	$> 81921 \cdot 2^{(n-39)/2}$

We use this construction without assuming the conjecture that it is optimal when an optimal balanced input is used. For a concrete comparison at $n = 375$, start with the balanced fifteen-variable function, for which $d_{15} = 112$. Then $d_{30} = 16496$ by (4.8). Combining the PW fifteen-variable function with this balanced thirty-variable function gives a balanced forty-five-variable function with $d_{45} = 216 \cdot 16496 = 3563136$. Three applications of (4.8) give a balanced function on 360 variables with

$$d_{360} = 2^{179} + 2^{89} + 2^{44} + 3563136, \qquad \epsilon_{360} = 2^{89} + 2^{44} + 3563136.$$

Consequently the PW function plus this balanced factor has nonlinearity

$$2^{374} - 27 \cdot 2^{182} - 216\big(2^{89} + 2^{44} + 3563136\big). \tag{4.9}$$

Since $216(2^{89}+2^{44}+3563136) < 2^{182}$, this is strictly better than $2^{374}-28\cdot 2^{182}$ from the balanced fifteen-variable function plus a bent function. It is nevertheless below the ordinary PW value $2^{374}-27\cdot 2^{182}$, whereas the nonlinearity of our balanced functions exceeds that value by more than 2^{168}. Using any better balanced even-dimensional factor only decreases ϵ_e and the comparisons above remain valid. This is an explicitly achievable factor, not a claim of optimality in 360 variables. Dobbertin's asymptotic conjecture for balanced functions is a different statement, proved by Schmidt [18, Theorem 2].

5 Conclusion and Future Directions

We have proved the existence of Boolean functions on 375 variables whose nonlinearity exceeds the value obtained by taking the direct sum of a 15-variable Patterson-Wiedemann function and a bent function on 360 variables. More precisely, our construction provides

$$\mathrm{NL}(f) > 2^{374} - 442367 \cdot 2^{168} = \mathcal{N}_{\mathrm{BC}}(375) + 81921 \cdot 2^{168},$$

that improves the Patterson-Wiedemann direct-sum value by more than 2^{168}. Taking direct sums with bent functions extends this improvement to every odd number of variables $n \geq 375$. The construction combines an algebraically prescribed union of multiplicative cosets with a subset of a multiplicative subgroup whose existence is established through a linear-feasibility criterion. The index-two Gauss-sum evaluation makes the fixed spectral contribution explicit, while a finite numerical verification establishes the potential bound required by the feasibility argument. We have also shown that the construction can be modified to obtain balanced functions without losing the stated strict nonlinearity bound. This balancing step uses sampling without replacement and concentration inequalities to control all nonzero Walsh coefficients simultaneously.

The results are existence statements rather than efficient algorithms for evaluating the constructed functions. In particular, the subset that satisfies the simultaneous Walsh constraints

is not explicitly listed, and the modification set used to obtain balancedness is also obtained non-constructively. Finding explicit descriptions of these sets, developing efficient membership tests, and reducing the computational complexity of the construction remain important open problems. More broadly, this work illustrates how algebraic information about Gauss periods can be combined with general linear-feasibility methods and probabilistic arguments to obtain explicit quantitative bounds for Boolean-function nonlinearity. Further investigation of these connections may lead to improved constructions and a better understanding of the gap between known direct-sum constructions and the asymptotic behavior of maximum nonlinearity.

References

[1] R. Bardenet and O.-A. Maillard. Concentration inequalities for sampling without replacement. *Bernoulli*, 21(3):1361–1385, 2015.

[2] C. Carlet. *Boolean Functions for Cryptography and Coding Theory*. Cambridge University Press, 2021.

[3] D. A. Cox. *Primes of the Form $x^2 + ny^2$: Fermat, Class Field Theory, and Complex Multiplication*. Wiley, second edition, 2013.

[4] Hans Dobbertin. Construction of bent functions and balanced Boolean functions with high nonlinearity. In Bart Preneel, editor, *Fast Software Encryption*, volume 1008 of *Lecture Notes in Computer Science*, pages 61–74. Springer, 1995. Proceedings of FSE 1994.

[5] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58(301):13–30, 1963.

[6] X.-d. Hou, E. Sabo, and F. Slaughter. The covering radius of RM(1,9) is precisely 242. Cryptology ePrint Archive, Paper 2026/2529, 2026. https://eprint.iacr.org/2026/2529.

[7] S. Kavut, S. Maitra, and M. D. Yücel. Search for Boolean functions with excellent profiles in the rotation symmetric class. *IEEE Trans. Inform. Theory*, 53(5):1743–1751, 2007.

[8] S. Kavut and M. D. Yücel. 9-variable Boolean functions with nonlinearity 242 in the generalized rotation symmetric class. *Inform. Comput.*, 208(4):341–350, 2010.

[9] S. Kavut and S. Maitra. Patterson-Wiedemann type functions on 21 variables with nonlinearity greater than bent concatenation bound. *IEEE Trans. Inform. Theory*, 62(4):2277–2282, 2016.

[10] S. Kavut, S. Maitra, and F. Özbudak. A super-set of patterson-wiedemann functions: Upper bounds and possible nonlinearities. *SIAM J. Discret. Math.*, 32(1):106–122, 2018.

[11] R. Lidl and H. Niederreiter. *Finite Fields*, volume 20 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, second edition, 1997.

[12] Y. Lou, Y. Xu, and Q. Wang. The maximum nonlinearity of 8-variable balanced Boolean functions is 116. arXiv:2610.17897v1, 2026. https://arxiv.org/abs/2610.17897v1.

[13] N. J. Patterson and D. H. Wiedemann. The covering radius of the $(2^{15}, 16)$ Reed-Muller code is at least 16276. *IEEE Trans. Inform. Theory*, 29(3):354–356, 1983. Correction: IEEE Trans. Inform. Theory 36(2) (1990), 443.

[14] N. J. Patterson and D. H. Wiedemann. Correction to - the covering radius of the $(2^{15}, 16)$ Reed-Muller code is at least 16276. *IEEE Trans. Inform. Theory*, 36(2):443, 1990. Correction: IEEE Trans. Inform. Theory 36(2) (1990), 443.

[15] V. Reis and Z. Song. Three standard deviations suffice while one does not. arXiv:2609.34471v1, 2026. Preprint, September 28, 2026. https://arxiv.org/abs/2609.34471v1.

[16] O. S. Rothaus. On "bent" functions. *J. Combin. Theory Ser. A*, 20(3):300–305, 1976.

[17] S. Sarkar and S. Maitra. Idempotents in the neighbourhood of Patterson-Wiedemann functions having Walsh spectra zeros. *Designs, Codes and Cryptography*, 49(1–3):95–103, 2008.

[18] K.-U. Schmidt. Asymptotically optimal Boolean functions. *J. Combin. Theory Ser. A*, 164:50–59, 2019.

[19] Qichun Wang, Yueying Lou, and Zifan Wang. The covering radius of the Reed-Muller code RM(1,9) is 242. arXiv:2610.17871v1, 2026. https://arxiv.org/abs/2610.17871v1.

[20] J. Yang and L. Xia. Complete solving of explicit evaluation of Gauss sums in the index 2 case. *Sci. China Math.*, 53(9):2525–2542, 2010.

[21] Y. Yuan, Y. Hu, Y. Zhang, L. Zhang, and W. Wu. Theoretical open problems in symmetric cryptography: Verifiable LLM-guided analysis. Cryptology ePrint Archive, Paper 2026/1687, 2026. https://eprint.iacr.org/2026/1687.