Cryptography and Communications
https://doi.org/10.1007/s12095-026-00911-0

RESEARCH

# The nonlinearity of restrictions of direct sums of monomials and of threshold functions

**Serge Feukoua[3] · Ana Sălăgean[1] · Pantelimon Stănică[2]**

Received: 30 October 2025 / Accepted: 19 June 2026
© The Author(s) 2026

**Abstract**
We investigate the behavior of the nonlinearity for Boolean functions when their inputs are restricted to affine subspaces of the original domain. Functions which maintain a relatively high nonlinearity are of interest for cryptographic applications. To capture this behavior, we introduce the notion of $k$-restriction nonlinearity, as the minimal nonlinearity achieved when restricting to affine spaces of codimension at most $k$. We give some results that facilitate the computation of this parameter for direct sums of functions and for functions of relatively low nonlinearity. These results are then applied to two important classes of functions. Firstly, for direct sums of monomials, we provide an explicit formula and a characterization, showing that the $k$-restriction nonlinearity equals the nonlinearity of the direct sum of monomials obtained from the original function by removing any $k$ non-linear monomials of lowest degree. In general, the nonlinearity decreases by only a small amount for such functions. Secondly, for threshold functions, we prove that under certain conditions on the parameters, the $k$-restriction nonlinearity equals the nonlinearity of a related threshold function in fewer variables, from which an explicit formula for the $k$-restriction nonlinearity is derived.

**Keywords** Boolean functions · Threshold · Symmetric · Affine spaces · Restriction · Nonlinearity

**Mathematics Subject Classification (2010)** 06E30 · 94C10 · 94A60 · 11T71

Serge Feukoua, Ana Sălăgean, and Pantelimon Stănică contributed equally to this work.

Most of this work was done while the first author was at Loughborough University, UK.

✉ Serge Feukoua
feukouaf@yahoo.fr

Ana Sălăgean
A.M.Salagean@lboro.ac.uk

Pantelimon Stănică
pstanica@nps.edu

[1] Department of Computer Science, Loughborough University, Loughborough, UK

[2] Department of Applied Mathematics, Naval Postgraduate School, Monterey, California 93943, USA

[3] Department of Fundamental Sciences, National Advanced School of Public Works, Yaoundé, Cameroon

Published online: 25 July 2026

## 1 Introduction

Boolean functions used in cryptography need to satisfy certain requirements in order to resist attacks. In particular they need to have high nonlinearity, high algebraic degree, high algebraic immunity etc. These parameters should ideally not decrease significantly when the Boolean function is restricted to a hyperplane (or an affine subspace of reasonably high dimension) contained in its original domain. This stability is crucial for ensuring robustness against *guess-and-determine attacks*, where an adversary constrains the function's input to a specific affine subspace, thereby reducing the complexity of the attack.

A recent study [3] investigated the stability of the algebraic degree of Boolean functions under such restrictions. In this paper, we continue this line of research by examining another fundamental cryptographic property, namely the *nonlinearity*.

A key insight from Canteaut et al. [1] is that for any *bent function* $f$ (a function achieving optimal nonlinearity), the restriction of $f$ to any hyperplane of $\mathbb{F}_2^n$ remains *nearly bent*. This means that within the class of bent functions, nonlinearity remains relatively stable under hyperplane restrictions.

For a Boolean function $f$, we introduce the notion of $k$-restriction nonlinearity, that captures the minimal possible nonlinearity that $f$ can have when restricted to affine spaces of codimension at most $k$ (see Definition 3). The normalized version of nonlinearity is used throughout, as it provides a fairer comparison between functions in different numbers of variables. Some general properties of the $k$-restriction nonlinearity are given in Section 3.

In Section 4 we examine this parameter for the case of functions which are direct sums of $m$ functions, i.e. $f = f_1 + \ldots + f_m$ such that $f_i$ have disjoint sets of variables. We show in Theorem 3 that in this case the $k$-restriction nonlinearity of $f$ is lower bounded by the nonlinearity of the direct sum obtained from the original function $f$ by removing $k$ of the summands $f_i$ of highest nonlinearity. We then apply these results to the class of *direct sums of monomials*. Previous results considered the nonlinearity of such functions when restricted to only certain affine spaces, namely the ones defined by equations of the form $x_i = 0$ or $x_i = 1$, see [9], or to the set of vectors of a specified Hamming weight (such sets are not affine spaces) – see [5]. Our result (Corollary 1) gives the explicit value of the $k$-restriction nonlinearity of a direct sum of monomials $f = f_1 + \ldots + f_m$, by showing that it is equal to the nonlinearity of $f = f_1 + \ldots + f_{m-k}$, where $f_i$ are ordered such that $\deg(f_1) \geq \ldots \geq \deg(f_m) \geq 2$ (any terms of degree 1 can be ignored for the purpose of computing nonlinearity). This also gives a full characterization of the $k$-restriction nonlinearity of all the quadratic functions, since all quadratic functions are EA-equivalent to direct sums of monomials. We note that for direct sums of monomials, the nonlinearity is in general relatively stable (except for some edge cases), i.e. the $k$-restriction nonlinearity is not much lower than the nonlinearity of the function.

In Section 5 we show that for functions $f$ of normalized nonlinearity at most $\frac{1}{4}$, a closest affine function $\ell_0$ has another useful property: $\ell_0$ is also a closest affine function for the restriction of $f$ to any hyperplane $H$. This considerably reduces the search space for determining the 1-restriction nonlinearity of $f$. A generalization of this result to arbitrary codimension $k$ is also given.

We apply then these results to threshold functions in Section 6. A threshold function $T_{d,n}$ in $n$ variables returns 1 if and only if the Hamming weight of its input is at least $d$. Its nonlinearity can be explicitly computed (see [4]), Threshold functions are an important subclass of symmetric functions (i.e. functions which are invariant to permutations of the variables). For most, but not all threshold functions we provide an explicit formula for the

$k$-restriction nonlinearity, by showing that it is equal to the nonlinearity of a related threshold function in $n-k$ variables. For the threshold functions not covered by these results, we give conjectures supported by computational data.

## 2 Preliminaries

We let $\mathbb{F}_2$ be the finite field with two elements and $\mathbb{F}_2^n$ be the vector space over $\mathbb{F}_2$ of dimension $n$. A function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is an $n$-variable Boolean function. It can be uniquely represented in *Algebraic Normal Form* (ANF), that is, as a polynomial function of degree at most one in each variable, $f(x_1, \ldots, x_n) = \sum_{I \subseteq \{1,\ldots,n\}} a_I \prod_{i \in I} x_i$, with $a_I \in \mathbb{F}_2$ (see e.g. Carlet [2]).

The *algebraic degree* of a Boolean function $f$, denoted by $\deg(f)$, is the degree of its ANF representation. Those functions of algebraic degree at most one (respectively, two) are called *affine* (respectively *quadratic*).

**Notation 1** *For every $n$-variable Boolean function $f$, we denote by* $\mathrm{Var}(f)$ *the set consisting of all the elements $i \in \{1, \ldots, n\}$ such that $x_i$ appears in at least one term with nonzero coefficient in the ANF of $f$.*

For every vector $v \in \mathbb{F}_2^m$, the Hamming weight $\mathrm{w}(v)$ of $v$ is the number of its non zero coordinates (i.e. the cardinality of the support set of $v$, $\{i \in \{1, \ldots, m\} \mid v_i \neq 0\}$). The Hamming distance, $\mathrm{d}(v, v')$, between two vectors $v, v' \in \mathbb{F}_2^m$ is the number of positions where they differ, i.e. the cardinality of $\{i \mid v_i \neq v'_i\}$. Note that $\mathrm{d}(v, v') = \mathrm{w}(v - v') = \mathrm{w}(v + v')$. The *Hamming weight* $\mathrm{w}(f)$ of a Boolean function $f$ with domain $\mathbb{F}_2^n$ is the Hamming weight of its truth table, or equivalently the cardinality of the support of $f$, $\{x \in \mathbb{F}_2^n \mid f(x) \neq 0\}$. The *Hamming distance* between two functions $f_1, f_2$, denoted by $\mathrm{d}(f_1, f_2)$, is the *Hamming distance* between their truth tables, or equivalently, $\mathrm{d}(f_1, f_2) = \mathrm{w}(f_1 + f_2)$. The nonlinearity $\mathrm{nl}(f)$ of a Boolean function $f$ is the minimum Hamming distance between $f$ and affine functions.

We will work with functions in different numbers of variables, say two functions $f_1$ and $f_2$, in $n_1$ and $n_2$ variables, respectively. Comparing their weight or nonlinearity directly would not be fair, as their weight and their distance to an affine function (or any function) can be up to $2^{n_1}$ for $f_1$ and up to $2^{n_2}$ for $f_2$. For a fairer comparison, we will normalize the weight and the nonlinearity. Namely, for vectors $v, v' \in \mathbb{F}_2^m$ the normalized weight and distance are defined as $\mathrm{nw}(v) = \frac{1}{m}\mathrm{w}(v)$ and $\mathrm{nd}(v, v') = \frac{1}{m}\mathrm{d}(v, v')$ (i.e. dividing by the length of the vector). Similarly, the normalized weight, distance and nonlinearity for functions $f, f_1, f_2$ in $n$ variables are defined as $\mathrm{nw}(f) = \frac{1}{2^n}\mathrm{w}(f)$ and $\mathrm{nd}(f_1, f_2) = \frac{1}{2^n}\mathrm{d}(f_1, f_2)$ and $\mathrm{nnl}(f) = \frac{1}{2^n}\mathrm{nl}(f)$, respectively (i.e. dividing by the length $2^n$ of the truth table).

Recall that the Walsh-Hadamard transform of a Boolean function $f$ is the function $W_f : \mathbb{F}_2^n \to \mathbb{R}$ defined as $W_f(u) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+u \cdot x}$, for all $u \in \mathbb{F}_2^n$, where "$\cdot$" denotes the usual inner product. Again, we will use the normalized version:

**Definition 1** The normalized Walsh-Hadamard transform of a Boolean function $f$ in $n$ variables, denoted here by $\mathrm{NW}_f$, is the function $\mathrm{NW}_f : \mathbb{F}_2^n \to \mathbb{R}$ defined as

$$\mathrm{NW}_f(a) = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+a \cdot x},$$

for all $a \in \mathbb{F}_2^n$.

The nonlinearity of a function is related to its Walsh-Hadamard transform, namely $\mathrm{nl}(f) = 2^{n-1} - \frac{1}{2}\max_{a\in\mathbb{F}_2^n}\left|W_f(a)\right|$ (see, for example [6]). Normalizing, we obtain

$$\mathrm{nnl}(f) = \frac{1}{2}\left(1 - \max_{a\in\mathbb{F}_2^n}\left|\mathrm{NW}_f(a)\right|\right). \tag{1}$$

The (normalized) nonlinearity satisfies the bounds $\mathrm{nl}(f) \leq 2^{n-1} - 2^{n/2-1}$ and $\mathrm{nnl}(f) \leq \frac{1}{2}\left(1 - \frac{1}{2^{\frac{n}{2}}}\right)$, respectively; functions that attain the upper bound are called *bent functions* (in which case $n$ must be even). When $n$ is odd, functions $f$ with $\mathrm{nnl}(f) = \frac{1}{2}\left(1 - \frac{1}{2^{\frac{n-1}{2}}}\right)$ are called *nearly bent* functions.

***Remark 1*** It is well-known, and easy to check from the definition of the Walsh-Hadamard transform that if a function $f$ in $n$ variables is viewed as a function in $m > n$ variables, that is, as a function $g(x_1, \ldots, x_n, \ldots, x_m) = f(x_1, \ldots, x_n)$ we have $W_g(a, b) = \delta_0(b)2^{m-n}W_f(a)2^{m-n}W_f(a)$ (where $\delta_0$ is the Dirac symbol at 0 *i. e* the indicator of $\{0\}$) for all $a \in \mathbb{F}_2^n$, $b \in \mathbb{F}_2^{m-n}$. Consequently, $\mathrm{nl}(g) = 2^{m-n}\mathrm{nl}(f)$. However, the normalized version of the nonlinearity remains unchanged; namely, $\mathrm{nnl}(g) = \mathrm{nnl}(f)$ (for the normalized Walsh-Hadamard transform we have instead $\mathrm{NW}_g(a, b) = \delta_0(b)\mathrm{NW}_f(a)$) and $\mathrm{nnl}(g) = \mathrm{nnl}(f)$. Similarly, the normalized weight remains unchanged, $\mathrm{nw}(g) = \mathrm{nw}(f)$.

We recall the values of the normalized nonlinearity for some simple functions. When $f$ is a monomial function of algebraic degree $r \geq 2$, we have

$$\mathrm{nnl}(f) = \frac{1}{2^r}. \tag{2}$$

When a function $f$ is a direct sum of two functions $g$ and $h$ in $n_1$ and $n_2$ variables, respectively (i.e., $f = g + h$ and $\mathrm{Var}(g) \cap \mathrm{Var}(h) = \emptyset$), it is known that $W_f(a, b) = W_g(a)W_h(b)$ for all $a \in \mathbb{F}_2^{n_1}$, $b \in \mathbb{F}_2^{n_2}$ (see [6] for example). This means we also have $\mathrm{NW}_f(a, b) = \mathrm{NW}_g(a)\mathrm{NW}_h(b)$ and therefore,

$$\mathrm{nnl}(f) = \frac{1}{2}\left(1 - (1 - 2\mathrm{nnl}(g))(1 - 2\mathrm{nnl}(h))\right) = \mathrm{nnl}(g) + \mathrm{nnl}(h) - 2\mathrm{nnl}(g)\mathrm{nnl}(h). \tag{3}$$

More generally, when $f = f_1 + \cdots + f_m$ is a direct sum of $m$ functions, we have

$$\mathrm{nnl}(f) = \frac{1}{2}\left(1 - \prod_{i=1}^{m}(1 - 2\mathrm{nnl}(f_i))\right). \tag{4}$$

Note that in this case we have, for any subset $I \subseteq \{1, \ldots, m\}$,

$$\mathrm{nnl}(f) \geq \mathrm{nnl}\left(\sum_{i\in I} f_i\right). \tag{5}$$

When $f = f_1 + \cdots + f_m$ is a direct sum of $m$ monomials of degree at least two, Eqs. 4 and 2 yield

$$\mathrm{nnl}(f) = \frac{1}{2}\left(1 - \prod_{i=1}^{m}\left(1 - \frac{1}{2^{\deg(f_i)-1}}\right)\right). \tag{6}$$

When computing nonlinearity, it is convenient to view functions up to a suitable equivalence relation, namely EA-equivalence, which we will recall below. First recall that an affine

automorphism of $\mathbb{F}_2^n$ is a mapping $\varphi$ of the form $\varphi(x) = Mx + a$, where $x$, $a$ and $\varphi(x)$ are viewed as column vectors, $x^T = (x_1, \ldots, x_n)$, $a^T = (a_1, \ldots, a_n) \in \mathbb{F}_2^n$ and $M$ is an $n \times n$ invertible matrix over $\mathbb{F}_2$.

**Definition 2** Two Boolean functions $f$ and $g$ in $n$ variables are said to be affine equivalent if there exists an affine automorphism $\varphi$ of $\mathbb{F}_2^n$, such that $f = g \circ \varphi$, where '$\circ$' is the composition. The functions $f$ and $g$ are said to be EA-equivalent (extended affine equivalent), and we write then $f \sim g$, if there exists an affine automorphism $\varphi$ of $\mathbb{F}_2^n$ and an affine Boolean function $\ell$ in $n$ variables, such that $f = g \circ \varphi + \ell$.

Note that the algebraic degree, as well as the (normalized) Walsh-Hadamard transform and (normalized) nonlinearity are invariant under EA-equivalence.

The functions of algebraic degree 2 are fully characterized up to affine equivalence.

**Lemma 1** [7, Chapter 15, Theorem 8] *Every quadratic non-affine function in n variables is EA-equivalent to* $x_1x_2 + \cdots + x_{2m-1}x_{2m}$, *for some* $m \leq \frac{n}{2}$.

The restriction of a function $f$ to an affine space $A$ will be denoted by $f_{|A}$. Any system of $k$ equations that define an affine space of codimension $k$ can be transformed by Gaussian elimination into $k$ equations of the form: $x_{i_1} = a_{i_1}(y), \ldots, x_{i_k} = a_{i_k}(y)$, where $i_1, \ldots, i_k$ are distinct and $a_j(y)$ are affine functions in $y = (y_1, \ldots, y_{n-k})$, where $\{y_1, \ldots, y_{n-k}\} = \{x_1, \ldots, x_n\} \setminus \{x_{i_1}, \ldots, x_{i_k}\}$. The ANF of $f_{|A}$ can be obtained from the ANF of $f$ by substituting the variables $x_{i_1}, \ldots, x_{i_k}$ with $a_{i_1}(y), \ldots, a_{i_k}(y)$; we obtain thus a function in the $n-k$ variables $\{y_1, \ldots, y_{n-k}\}$, whose expression depends on the equations used for defining $A$, although all choices yield functions which are affine equivalent to each other.

The behavior of restrictions with respect to EA-equivalence is as follows.

**Proposition 1** [3, Lemma 4] *Let* $f, g$ *be n-variable functions such that* $f \sim g$, *i.e.* $g = f \circ \varphi + \ell$ *for some affine automorphism* $\varphi$ *of* $\mathbb{F}_2^n$ *and some affine function* $\ell$. *For any affine space A of* $\mathbb{F}_2^n$ *we have* $g_{|A} \sim f_{|\varphi(A)}$.

## 3 The *k*-restriction nonlinearity

In this work, we are interested in how the normalized nonlinearity of a function changes when we restrict the function to an affine subspace. It turns out that it can decrease, increase, or stay unchanged.

***Example 1*** For any function $f$ in $n$ variables $x_1, \ldots, x_n$, if there is one variable, say $x_1$, such that $f$ does not depend on $x_1$, then the restriction of $f$ to the hyperplane $H$ defined by the equation $x_1 = 0$ (or any affine equation containing $x_1$) has the same ANF as $f$, and therefore $\mathrm{nnl}(f_{|H}) = \mathrm{nnl}(f)$.

For the function $f(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_1x_3 + x_1x_2x_3$, one can check by direct computation that $\mathrm{nnl}(f) = \frac{1}{8}$; when the hyperplane $H$ is defined by the equation $x_1 = 0$ we have $f_{|H}(x_2, x_3) = x_2x_3$ and $\mathrm{nnl}(f_{|H}) = \frac{1}{4} > \mathrm{nnl}(f)$; for the hyperplane $H'$ defined by the equation $x_1 = 1$ we have $f_{|H'}(x_2, x_3) = x_2 + x_3$ and $\mathrm{nnl}(f_{|H'}) = 0 < \mathrm{nnl}(f)$.

We are particularly interested in functions for which the nonlinearity does not decrease much on any of the subspaces of a given codimension $k$. We are therefore led to define the following notion.

**Definition 3** Let $f$ be a Boolean function in $n$ variables and $0 \leq k < n$. The $k$-restriction nonlinearity of $f$ is defined as

$$\mathrm{Rnnl}_k(f) = \min_{\mathrm{codim}(A) \leq k} (\mathrm{nnl}(f_{|A})),$$

where $A$ are affine spaces of $\mathbb{F}_2^n$.

It is well-known that the nonlinearity nl is invariant under EA-equivalence; for any fixed $k$, any invertible affine transformation of $\mathbb{F}_2^n$ gives a one-to-one mapping on the set of affine subspaces of codimension $k$ of $\mathbb{F}_2^n$. This, in conjunction with Proposition 1, means that $\mathrm{Rnnl}_k$ is invariant under EA-equivalence:

**Proposition 2** *Let $f, g$ be two Boolean functions in $n$ variables. If $f \sim g$ then $\mathrm{Rnnl}_k(f) = \mathrm{Rnnl}_k(g)$ for any $0 \leq k \leq n$.*

Note that

$$\mathrm{nnl}(f) = \mathrm{Rnnl}_0(f) \geq \mathrm{Rnnl}_1(f) \geq \mathrm{Rnnl}_2(f) \geq \ldots \geq \mathrm{Rnnl}_{n-1}(f) = 0.$$

Other than the trivial case of affine (including constant) functions, which have nonlinearity zero, we do not know if there are any functions with $\mathrm{nnl}(f) = \mathrm{Rnnl}_1(f)$. We will be interested in functions for which $\mathrm{Rnnl}_1(f)$ is not much lower than $\mathrm{nnl}(f)$.

The parameter $\mathrm{Rnnl}_1(f)$ can immediately be computed for functions $f$ which have a single monomial in their ANF, or, more generally, can be written in the form $f(x) = \ell(x)g(x)$ with $\ell(x)$ an affine function; the restriction of $f$ to the hyperplane defined by the equation $\ell(x) = 0$ is the identically zero function, and therefore $\mathrm{Rnnl}_1(f) = 0$.

Next we consider the bent functions, as they have maximal nonlinearity. In Canteaut et al. [1, Theorem 5.2] it is shown that a Boolean function $f$ in $n$ variables is bent if and only if for every hyperplane $H$ of $\mathbb{F}_2^n$, the restrictions of $f$ to $H$ and $\mathbb{F}_2^n \setminus H$ (viewed as Boolean functions on $\mathbb{F}_2^{n-1}$) are nearly bent. Consequently, for any bent function $f$ we have

$$\mathrm{Rnnl}_1(f) = \frac{1}{2}\left(1 - \frac{1}{2^{\frac{n}{2}-1}}\right) = \mathrm{nnl}(f) - \frac{1}{2^{\frac{n}{2}+1}}. \tag{7}$$

This means that, except for very small values of $n$, the nonlinearity of bent functions is relatively stable under restrictions to hyperplanes.

***Remark 2*** Any affine space $A$ of codimension $k$ can be written as the intersection of $k$ hyperplanes, $A = H_1 \cap \ldots \cap H_k$; moreover $f_{|A} \sim (((f_{|H_1})_{|H_2}) \ldots)_{|H_k}$. When trying to compute $\mathrm{Rnnl}_k(f)$ for $k \geq 2$, one could be tempted to construct a space $A$ that minimizes $\mathrm{nnl}(f_{|A})$ by using a "greedy" approach to repeatedly pick hyperplanes that minimize the nonlinearity of the restriction at each step. For example, for $k = 2$, one could pick first a hyperplane $H_1$ such that $\mathrm{nnl}(f_{|H_1})$ is minimal among all hyperplanes (and therefore $\mathrm{nnl}(f_{|H_1}) = \mathrm{Rnnl}_1(f)$). Then one could pick a hyperplane $H_2$ which minimizes $\mathrm{nnl}((f_{|H_1})_{|H_2})$, i.e. $\mathrm{nnl}((f_{|H_1})_{|H_2}) = \mathrm{Rnnl}_1(f_{|H_1})$. One could then hope that $\mathrm{Rnnl}_1(f_{|H_1}) = \mathrm{Rnnl}_2(f)$; in other words the restriction to the space of codimension two $A = H_1 \cap H_2$ reaches the global minimum nonlinearity. However this approach does not always work. A counterexample is the function $f : \mathbb{F}_2^6 \to \mathbb{F}_2$ with truth table

$$f = [0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1,$$
$$1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1].$$

For this function, $\mathrm{Rnnl}_2(f) = 0$ and $\mathrm{Rnnl}_1(f) = 0.1875$. One of the hyperplanes where the latter is achieved, is the hyperplane $H_1$ defined by $x_2 + x_5 + x_6 = 0$. However, for this

minimal hyperplane $H_1$ and any choice of hyperplane $H_2$, we have $\mathrm{nnl}(f_{|H_1\cap H_2}) > 0$; the minimum is actually $\mathrm{Rnnl}_1(f_{|H_1}) = 0.0625$, which is different from $\mathrm{Rnnl}_2(f) = 0$. Thus, starting from this optimal hyperplane for codimension 1, it is impossible to reach an optimal codimension-2 restriction, showing that the greedy approach fails.

***Remark 3*** Alternatively, one could wonder whether all spaces $A$ of codimension $k = 2$ that minimize the nonlinearity of the restriction, i.e. $\mathrm{nnl}(f_{|A}) = \mathrm{Rnnl}_2(f)$, can be written as the intersection of two hyperplanes $A = H_1 \cap H_2$ where at least one of the hyperplanes, say $H_1$ is such that $H_1$ minimizes the nonlinearity of the restriction of $f$ to hyperplanes, i.e. $\mathrm{nnl}(f_{|H_1}) = \mathrm{Rnnl}_1(f)$. This is not true in general. For example, consider the symmetric Boolean function $S_{3,4} : \mathbb{F}_2^4 \to \mathbb{F}_2$ defined by

$$S_{3,4}(x_1, x_2, x_3, x_4) = \begin{cases} 1 & \text{if } \mathrm{w}(x_1, x_2, x_3, x_4) = 3, \\ 0 & \text{otherwise.} \end{cases}$$

For this function, $\mathrm{Rnnl}_1(S_{3,4}) = 0$ is achieved by the hyperplane defined by the equation $x_1 + x_2 + x_3 + x_4 = 0$, and $\mathrm{Rnnl}_2(S_{3,4}) = 0$ is achieved (among other spaces) by the codimension-2 space $A_0 = H_A \cap H_B$ where $H_A : x_1 = 0$ and $H_B : x_2 = 0$. However, neither of these component hyperplanes achieves the minimum nonlinearity of the restriction (which is $\mathrm{Rnnl}_1(S_{3,4}) = 0$); namely, $\mathrm{nnl}(S_{3,4|H_A}) = 0.125 \neq 0$ and $\mathrm{nnl}(S_{3,4|H_B}) = 0.125 \neq 0$. Thus, a minimal codimension-2 space can be formed from two hyperplanes, neither of which is minimal in codimension 1.

## 4 Direct sums

Let $f = f_1 + \cdots + f_m$ be a direct sum of $m \geq 2$ monomials, i.e. $\mathrm{Var}(f_i) \cap \mathrm{Var}(f_j) = \emptyset$ for all $i \neq j$. Without loss of generality we can assume $\deg(f_1) \geq \deg(f_2) \geq \ldots \geq \deg(f_m) \geq 2$ (any affine terms can be ignored due to invariance to EA-equivalence, see Proposition 2). Consider first hyperplanes $H$ defined by an equation $x_i = 0$ with $i \in \mathrm{Var}(f_j)$. Substituting $x_i$ by 0 in $f$ we immediately obtain $f_{|H} = f_1 + \cdots + f_{j-1} + f_{j+1} + \ldots + f_m$; this is again a direct sum, so using (6), we can compute $\mathrm{nnl}(f_1 + \cdots + f_{j-1} + f_{j+1} + \ldots + f_m)$, and see that the minimum over all values of $j$ is attained for $j = m$ (i.e. $i \in \mathrm{Var}(f_m)$), where we have

$$\mathrm{nnl}(f_1 + \cdots + f_{m-1}) = \frac{1}{2}\left(1 - \prod_{i=1}^{m-1}\left(1 - \frac{1}{2^{\deg(f_i)-1}}\right)\right). \tag{8}$$

Similarly, for an affine space of codimension $k \leq m$ defined by equations of the form $x_{i_1} = 0, \ldots, x_{i_k} = 0$, one can show that the restriction $f_{|A}$ is equal to the direct sum of those monomials of $f$ which do not contain any of the variables $x_{i_1}, \ldots, x_{i_k}$. The minimum nonlinearity among all such direct sums is achieved when $x_{i_j}$ appears in $f_{m-k+j}$, in which case the nonlinearity of the restriction $f_{|A}$ is

$$\mathrm{nnl}(f_1 + \cdots + f_{m-k}) = \frac{1}{2}\left(1 - \prod_{i=1}^{m-k}\left(1 - \frac{1}{2^{\deg(f_i)-1}}\right)\right). \tag{9}$$

The value above is obviously an upper bound for $\mathrm{Rnnl}_k(f)$:

$$\mathrm{Rnnl}_k(f) \leq \frac{1}{2}\left(1 - \prod_{i=1}^{m-k}\left(1 - \frac{1}{2^{\deg(f_i)-1}}\right)\right). \tag{10}$$

We will prove in Corollary 1 that this bound is in fact the exact value of $\mathrm{Rnnl}_k(f)$, that is, there are no spaces $A$ of codimension at most $k$ (defined by equations which are not necessarily of the form $x_i = 0$) such that $f_{|A}$ has nonlinearity strictly lower than the bound in Eq. 10.

We will actually prove a more general result, Theorem 3, regarding a lower bound for the $k$-restriction nonlinearity for direct sums of functions (not necessarily direct sums of monomials). To this end, we need first a technical Lemma, which includes an algorithm for making certain choices during Gaussian elimination on a system of linear equations. Note that if one is interested only in restrictions to hyperplanes (codimension one), one can skip directly to Theorem 3, as the Lemma below holds trivially in that case.

**Lemma 2** *Let $f = f_1 + \cdots + f_m$ be a direct sum of the functions $f_1, \ldots, f_m$. Let $A$ be any affine space of codimension $k$ with $1 \leq k \leq m$. Then there are $k$ affine equations that define $A$ and have the following additional properties: they are of the form $x_{i_j} = a_j(y)$, $j = 1, \ldots, k$, where $i_1, \ldots, i_k$ are distinct and $y = (y_1, \ldots y_{n-k})$ with $\{y_1, \ldots, y_{n-k}\} = \{x_1, \ldots, x_n\} \backslash \{x_{i_1}, \ldots, x_{i_k}\}$; moreover, writing $f = g + h$ where $g$ is the sum of those functions $f_i$ which contain at least one of the variables $x_{i_1}, \ldots, x_{i_k}$ (i.e. $\mathrm{Var}(f_i) \cap \{i_1, \ldots, i_k\} \neq \emptyset$), we have that substituting the variables $x_{i_1}, \ldots, x_{i_k}$ by $a_1(y), \ldots, a_k(y)$ in $g$ yields a function which is affine in the variables of $h$.*

***Proof*** Let $\mathtt{E}_1, \ldots, \mathtt{E}_k$ be any linearly independent affine equations that define $A$. We will perform Gaussian elimination on these equations, making careful choices at each step, as described in Algorithm 1. The intuition is that for each of the equations $\mathtt{E}_j$, in turn, with $j = 1, \ldots, k$, we pick a variable $x_{i_j}$, and use the equation $\mathtt{E}_j$ to substitute $x_{i_j}$ in all the other equations. When we choose the variable $x_{i_j}$, we always favor variables which belong to an "unused" function among $f_1, \ldots, f_m$, which did not have any of its variables picked already.

**Algorithm 1** Gaussian elimination for affine space restrictions for direct sums.

**Input:** A list of $k$ linearly independent affine equations $\mathtt{E}_1, \ldots, \mathtt{E}_k$ that define an affine space $A$; a function $f = f_1 + \cdots + f_m$ in variables $x_1, \ldots, x_n$, which is a direct sum of the polynomials $f_i$ (we denote by $\mathtt{Polynomial}(x_i)$ the unique element of $\{f_1, \ldots, f_m\}$ that contains the variable $x_i$)

**Output:** $k$ affine equations $\mathtt{E}_1, \ldots, \mathtt{E}_k$ that define $A$, of the form $x_{i_j} = a_j(y_1, \ldots y_{n-k})$, $j = 1, \ldots, k$, where $i_1, \ldots, i_k$ are distinct and $\{y_1, \ldots, y_{n-k}\} = \{x_1, \ldots, x_n\} \setminus \{x_{i_1}, \ldots, x_{i_k}\}$; the equations satisfy the property that when substituting $x_{i_1}, \ldots, x_{i_j}$ in $g$, where $g$ is equal to the sum of the polynomials in the set $\{\mathtt{Polynomial}(x_{i_1}), \ldots, \mathtt{Polynomial}(x_{i_k})\}$, the resulting function is linear in the variables $\{x_i : i \in \{1, \ldots, n\} \setminus \mathrm{Var}(g)\}$

```
UnusedPolynomials ← {f_1, ..., f_m}
for j = 1 to k do
    // Choose a variable with non-zero coefficient in equation E_j
    if ∃ variable x_ℓ in E_j such that Polynomial(x_ℓ) ∈ UnusedPolynomials then
        Choose such an x_ℓ, and set i_j ← ℓ
        UnusedPolynomials ← UnusedPolynomials \ {Polynomial(x_{i_j})}
    else
        Choose any variable x_ℓ with non-zero coefficient in E_j, and set i_j ← ℓ
    end if
    // Isolate x_{i_j} in equation E_j and substitute into all other equations
    Write equation E_j as x_{i_j} = a_j(x_1, ..., x_{i_j − 1}, x_{i_j + 1}, ..., x_n)
    for t ∈ {1, ..., k} \ {j} do
        Substitute x_{i_j}
        by a_j(x_1, ..., x_{i_j − 1}, x_{i_j + 1}, ..., x_n) into E_t
    end for
end for
return the equations E_1, ..., E_k
```

We now prove the correctness of Algorithm 1. It is not difficult to show the following statements: the equations returned by the algorithm define the same space $A$ as the input equations; the variables $x_{i_1}, \ldots, x_{i_k}$ are distinct; each of these variables $x_{i_t}$, for $1 \leq t \leq k$, only appears in one of the equations $\mathtt{E}_1, \ldots, \mathtt{E}_k$ in the output of the algorithm, namely in the left hand side of the equation $\mathtt{E}_t$.

The more technical part of the proof involves showing that the equations returned by the algorithm satisfy the last property required for the output of Algorithm 1. Namely, let $g$ be the sum of the polynomials in the set $\{\mathtt{Polynomial}(x_{i_1}), \ldots, \mathtt{Polynomial}(x_{i_k})\}$, where $\mathtt{Polynomial}(x_i)$ denotes the unique element of $\{f_1, \ldots, f_m\}$ that contains the variable $x_i$, and let $x^{(2)} = \{x_i : i \in \{1, \ldots, n\} \setminus \mathrm{Var}(g)\}$. We have to show that when the equations produced by the algorithm are used to substitute $x_{i_1}, \ldots, x_{i_j}$ in $g$, the resulting function is linear in the variables $x^{(2)}$. Consider a function $f_u$ that is one of the summands in $g$. Note that $f_u$ does not contain any of the variables $x^{(2)}$, but it contains at least one of the variables $x_{i_1}, \ldots, x_{i_k}$, say it contains $x_{i_{t_1}}, \ldots, x_{i_{t_s}}$, with $1 \leq t_1 < t_2 < \ldots < t_s \leq k$. To compute $(f_u)_{|A}$, we substitute $x_{i_{t_1}}, \ldots, x_{i_{t_s}}$ by $a_{t_1}(y), \ldots, a_{t_s}(y)$ in $f_u$. We have to show that any product of some of the functions $a_{t_1}(y), \ldots, a_{t_s}(y)$ has total degree at most one in the variables $x^{(2)}$. It suffices to show that the product of all of them, namely $\prod_{i=1}^{s} a_{t_i}(y)$ has total degree at most one in the variables $x^{(2)}$. If $s = 1$, this is obvious, as $a_{t_1}$ is affine. Now assume $s \geq 2$. We will show that none of the affine functions $a_{t_2}, \ldots, a_{t_s}$ contain any of the variables in $x^{(2)}$ (note that $a_{t_1}$ might contain variables in $x^{(2)}$). Assume for a contradiction that there is a variable $x_\ell \in x^{(2)}$ which appears in one of $a_{t_2}, \ldots, a_{t_s}$, say it appears in $a_{t_2}$. Since $x_\ell \in x^{(2)}$, we know $\mathtt{Polynomial}(x_\ell) \notin \{\mathtt{Polynomial}(x_{i_1}), \ldots, \mathtt{Polynomial}(x_{i_k})\}$. Consider the point in the algorithm where the **for** loop reached $j = t_2$. At this point, at the start of the loop, $\mathtt{UnusedPolynomials} = \{f_1, \ldots, f_m\} \setminus \{\mathtt{Polynomial}(x_{i_1}), \ldots, \mathtt{Polynomial}(x_{i_{t_2-1}})\}$. Therefore, $\mathtt{Polynomial}(x_\ell) \in \mathtt{UnusedPolynomials}$. On the other hand, since $t_1 < t_2$ and $\mathtt{Polynomial}(x_{i_{t_1}}) = \mathtt{Polynomial}(x_{i_{t_2}}) = f_u$, at this point we have $\mathtt{Polynomial}(x_{i_{t_2}}) \notin \mathtt{UnusedPolynomials}$. Therefore the algorithm would not choose $x_{i_{t_2}}$ when a preferable choice of $x_\ell$ is available, which is a contradiction. This concludes the proof of the correctness of the algorithm. □

**Theorem 3** *Let $f = f_1 + f_2 + \cdots + f_m$ be a direct sum of functions $f_i$. Without loss of generality we can assume $0 < \mathrm{nnl}(f_1) \leq \ldots \leq \mathrm{nnl}(f_m)$. Let $k$ be an integer, with $1 \leq k \leq m$. Then we have:*

$$\mathrm{Rnnl}_k(f) \geq \mathrm{nnl}(f_1 + f_2 + \cdots + f_{m-k}). \tag{11}$$

***Proof*** Let $A$ be an affine space of codimension $k$. To compute $f_{|A}$, we first compute a system of equations for $A$ as described in Lemma 2 (using Algorithm 1) and then substitute the variables $x_{i_1}, \ldots, x_{i_k}$ in $f$ by $a_1(y), \ldots, a_k(y)$. We write $f = g + h$ as in Lemma 2 (recall $g$ is the sum of those functions $f_i$ which contain at least one of the variables $x_{i_1}, \ldots, x_{i_k}$), and define $x^{(1)} = \{x_i : i \in \mathrm{Var}(g) \setminus \{i_1, \ldots, i_k\}\}$ and $x^{(2)} = \{x_i : i \in \mathrm{Var}(h)\}$. Denote by $n_1, n_2$ the number of variables in $x^{(1)}, x^{(2)}$, respectively. Obviously, $n_1 + n_2 = n - k$. We obtain for $f_{|A}$ the following:

$$f_{|A}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = g_{|A}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) + h(\mathbf{x}^{(2)}). \tag{12}$$

We know from Lemma 2 that $g_{|A}$ is affine in the variables in $\mathbf{x}^{(2)}$, and therefore (12) can be written as

$$f_{|A}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = g_1(\mathbf{x}^{(1)}) + G(\mathbf{x}^{(1)}) \cdot \mathbf{x}^{(2)} + h(\mathbf{x}^{(2)}), \tag{13}$$

for some Boolean function $g_1$ and some vector $G$ of $n_2$ Boolean functions in the variables $\mathbf{x}^{(1)}$.

The normalized Walsh transform of $f_{|A}$ at $(\mathbf{u}^{(1)}, \mathbf{u}^{(2)})$ with $\mathbf{u}^{(1)} \in \mathbb{F}_2^{n_1}$, $\mathbf{u}^{(2)} \in \mathbb{F}_2^{n_2}$ is

$$\begin{aligned} \mathrm{NW}_{f_{|A}}(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}) &= \frac{1}{2^{n-k}} \sum_{\mathbf{x}^{(1)} \in \mathbb{F}_2^{n_1}} \sum_{\mathbf{x}^{(2)} \in \mathbb{F}_2^{n_2}} (-1)^{f_{|A}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) + \mathbf{u}^{(1)} \cdot \mathbf{x}^{(1)} + \mathbf{u}^{(2)} \cdot \mathbf{x}^{(2)}} \\ &= \frac{1}{2^{n_1}} \sum_{\mathbf{x}^{(1)} \in \mathbb{F}_2^{n_1}} (-1)^{g_1(\mathbf{x}^{(1)}) + \mathbf{u}^{(1)} \cdot \mathbf{x}^{(1)}} S(\mathbf{x}^{(1)}, \mathbf{u}^{(2)}, G), \end{aligned} \tag{14}$$

where

$$\begin{aligned} S(\mathbf{x}^{(1)}, \mathbf{u}^{(2)}, G) &= \frac{1}{2^{n_2}} \sum_{\mathbf{x}^{(2)} \in \mathbb{F}_2^{n-r}} (-1)^{(G(\mathbf{x}^{(1)}) + \mathbf{u}^{(2)}) \cdot \mathbf{x}^{(2)} + h(\mathbf{x}^{(2)})} \\ &= \mathrm{NW}_h(G(\mathbf{x}^{(1)}) + \mathbf{u}^{(2)}). \end{aligned} \tag{15}$$

From (14) and (15) we obtain

$$\begin{aligned} |\mathrm{NW}_{f_{|A}}(\mathbf{u}^{(1)}, \mathbf{u}^{(2)})| &\le \frac{1}{2^{n_1}} \sum_{\mathbf{x}^{(1)} \in \mathbb{F}_2^{n_1}} |S(\mathbf{x}^{(1)}, \mathbf{u}^{(2)}, G)| \\ &= \frac{1}{2^{n_1}} \sum_{\mathbf{x}^{(1)} \in \mathbb{F}_2^{n_1}} |\mathrm{NW}_h(G(\mathbf{x}^{(1)}) + \mathbf{u}^{(2)})| \\ &\le \max_{\mathbf{x}^{(2)} \in \mathbb{F}_2^{n_2}} |\mathrm{NW}_h(\mathbf{x}^{(2)})|. \end{aligned}$$

Therefore

$$\max_{(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}) \in \mathbb{F}_2^{n-k}} |\mathrm{NW}_{f_{|A}}(\mathbf{u}^{(1)}, \mathbf{u}^{(2)})| \le \max_{\mathbf{x}^{(2)} \in \mathbb{F}_2^{n_2}} |\mathrm{NW}_h(\mathbf{x}^{(2)})|.$$

and consequently

$$\mathrm{nnl}(f_{|A}) \ge \mathrm{nnl}(h).$$

In other words, denoting by $I$ the subset of $\{1, \ldots, n\}$ such that $h = \sum_{i \in I} f_i$, we proved that for any affine subspace $A$ of codimension $k$ there is a subset $I \subseteq \{1, \ldots, n\}$ of cardinality at least $n - k$ such that

$$\mathrm{nnl}(f_{|A}) \ge \mathrm{nnl}(\sum_{i \in I} f_i). \tag{16}$$

Combining the inequality above with the definition of $\mathrm{Rnnl}_k$ (Definition 3), we obtain

$$\begin{aligned} \mathrm{Rnnl}_k(f) &\ge \min_{|I| \ge n-k} \mathrm{nnl}(\sum_{i \in I} f_i) \\ &= \min_{|I| = n-k} \mathrm{nnl}(\sum_{i \in I} f_i) \\ &= \mathrm{nnl}(\sum_{i=1}^{m-k} f_i), \end{aligned}$$

where the second equality uses relation (5) and the last equality uses (4) and the fact that $0 < \mathrm{nnl}(f_1) \le \ldots \le \mathrm{nnl}(f_m)$. □

For direct sums of monomials, applying Theorem 3 and (10) we obtain the exact value of the $k$-restriction nonlinearity:

**Corollary 1** *Let $f = f_1 + f_2 + \cdots + f_m$ be a direct sum of monomials $f_i$ with $\deg(f_1) \geq \ldots \geq \deg(f_m) \geq 2$, and let $k$ be an integer, with $k \geq 1$. Then we have:*

$$\begin{aligned}\mathrm{Rnnl}_k(f) &= \mathrm{nnl}(f_1 + f_2 + \cdots + f_{m-k}) \\ &= \frac{1}{2}\left(1 - \prod_{i=1}^{m-k}\left(1 - \frac{1}{2^{\deg(f_i)-1}}\right)\right) \\ &= \mathrm{nnl}(f) - \frac{1}{2}\prod_{i=1}^{m-k}\left(1 - \frac{1}{2^{\deg(f_i)-1}}\right)\left(1 - \prod_{i=m-k+1}^{m}\left(1 - \frac{1}{2^{\deg(f_i)-1}}\right)\right).\end{aligned}$$

Note that the last equality in the theorem above means that in general the nonlinearity decreases by only a relatively small amount when restricting a direct sum of monomials to a hyperplane, or a space of small codimension. For example, if $f$ is a direct sum of 5 monomials of degree 3, we have $\mathrm{nnl}(f) \approx 0.38$ and $\mathrm{Rnnl}_1(f) \approx 0.34$, a decrease of about 0.04; for restrictions to spaces of codimension 2, we have $\mathrm{Rnnl}_1(f) \approx 0.29$, a further decrease of 0.05. Of course, there are some edge cases where the decrease is more significant, for example when $k = 1$, the degree $\deg(f_m)$ is much lower than $\deg(f_{m-1})$ and $m$ is relatively low. For example, for $m = 5$, $\deg(f_1) = \cdots = \deg(f_4) = 10$ and $\deg(f_5) = 2$, the nonlinearity decreases from $\mathrm{nnl}(f) = 0.2519$ to $\mathrm{Rnnl}_1(f) \approx 0.0039$, a decrease of 0.2480.

As all quadratic function are EA-equivalent to direct sums of monomials (as recalled in Lemma 1), Corollary 1 gives the values of $\mathrm{Rnnl}_k(f)$ for all quadratic $f$.

**Corollary 2** *Let $f$ be a quadratic function in $n$ variables, and let $m \leq \frac{n}{2}$ be the unique integer such that $f(x_1, \ldots, x_n) \sim x_1x_2 + \cdots + x_{2m-1}x_{2m}$. Let $k$ be an integer with $1 \leq k \leq m$. Then we have*

$$\mathrm{Rnnl}_k(f) = \frac{1}{2}\left(1 - \frac{1}{2^{m-k}}\right) = \mathrm{nnl}(f) - \frac{2^k - 1}{2^{m+1}}.$$

***Proof*** The result follows directly from Corollary 1. We also present an alternative proof here. Due to the invariance of $\mathrm{Rnnl}_k$ under EA-equivalence, we can assume without loss of generality that $f(x_1, \ldots, x_n) = x_1x_2 + x_3x_4 + \cdots + x_{2m-1}x_{2m}$. From (10), we see that $\mathrm{Rnnl}_k(f) \leq \frac{1}{2}\left(1 - \frac{1}{2^{m-k}}\right)$.

If $k = m$, we obtain

$\mathrm{Rnnl}_k(f) = 0 = \frac{1}{2}\left(1 - \frac{1}{2^{m-m}}\right)$.

Now assume $k \leq m - 1$, and let $A$ be an arbitrary space of codimension $k$. From Carlet et al. [3, Lemma 10], $f_{|A}$ is EA-equivalent to $x_1x_2 + x_3x_4 + \cdots + x_{2(m-t)-1}x_{2(m-t)}$ where $1 \leq t \leq k$. Thanks to Relation (6), we have

$$\mathrm{nnl}(f_{|A}) = \frac{1}{2}\left(1 - \frac{1}{2^{m-t}}\right) \geq \frac{1}{2}\left(1 - \frac{1}{2^{m-k}}\right)$$

and consequently $\mathrm{Rnnl}_k(f) = \frac{1}{2}\left(1 - \frac{1}{2^{m-k}}\right)$. The second equality follows from the fact that, by Eq. 6, we have $\mathrm{nnl}(f) = \frac{1}{2}\left(1 - \frac{1}{2^m}\right)$. □

***Remark 4*** In Remark 2 we saw that for constructing a space $A$ such that $\mathrm{nnl}(f_{|A})$ is minimal, a "greedy" approach does not always work. Note that Corollary 1 shows that when $f$ is a direct sums of monomials, the "greedy" approach does work.

## 5 Boolean functions with low nonlinearity

We are considering functions with low nonlinearity; we start with functions with nonlinearity at most $\frac{1}{4}$. These can be characterized by the following known and straightforward result:

**Lemma 3** *Let $f$ be a Boolean function. The following statements are equivalent:*
*(i)* $\mathrm{nnl}(f) \leq \frac{1}{4}$.
*(ii) There exists an affine function $\ell_0$ at distance at most $\frac{1}{4}$ from $f$.*
*Moreover, if any of these conditions is satisfied, then* $\mathrm{nnl}(f) = \mathrm{nd}(f, \ell_0)$.
*In the case of strict inequality, we have that* $\mathrm{nnl}(f) < \frac{1}{4}$ *if and only if there exists a unique affine function $\ell_0$ with* $\mathrm{nd}(f, \ell_0) < \frac{1}{4}$.

***Proof*** If we assume (i) holds, then (ii) follows from the definition of nonlinearity. Now assume (ii) holds. Then $\mathrm{nd}(f, \ell) \geq \frac{1}{4}$ for all other affine functions $\ell$ different from $\ell_0$, and the inequality is strict if $\mathrm{nnl}(f) < \frac{1}{4}$ (this follows from the triangle inequality, $\mathrm{nd}(f, \ell) \geq \mathrm{nd}(\ell_0, \ell) - \mathrm{nd}(f, \ell_0)$, and the fact that the distance between any two affine functions is at least $\frac{1}{2}$). Consequently, $\ell_0$ is an affine function closest to $f$, so $\mathrm{nnl}(f) = \mathrm{nd}(f, \ell_0) \leq \frac{1}{4}$. This shows that (i) holds. □

In particular, Lemma 3 shows that if the weight of a function $f$ is at most $\frac{1}{4}$, then $\mathrm{nnl}(f) = \mathrm{nw}(f)$. Similarly, if $\mathrm{nw}(f) \geq \frac{3}{4}$, we have that $\mathrm{nw}(f + 1) \leq \frac{1}{4}$, so $\mathrm{nnl}(f) = \mathrm{nnl}(f + 1) = \mathrm{nw}(f + 1) = 1 - \mathrm{nw}(f)$.

The following result shows that, for functions of nonlinearity at most $\frac{1}{4}$, a closest affine function $\ell_0$ described in Lemma 3 has another useful property: $\ell_0$ is also a closest affine function for the restriction of $f$ to any hyperplane $H$. This considerably reduces the search space for determining $\mathrm{Rnnl}_1(f)$.

**Proposition 4** *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ with* $\mathrm{nnl}(f) \leq \frac{1}{4}$. *If $\ell_0$ is an affine function such that* $\mathrm{nd}(f, \ell_0) \leq \frac{1}{4}$, *then*

$$\mathrm{Rnnl}_1(f) = \min_H \mathrm{nd}(f_{|H}, (\ell_0)_{|H}),$$

*where the minimum is taken over all hyperplanes $H$. In particular, if* $\mathrm{nw}(f) \leq \frac{1}{4}$, *then*

$$\mathrm{Rnnl}_1(f) = \min_H \mathrm{nw}(f_{|H}).$$

***Proof*** We prove first the particular case of $\mathrm{nw}(f) \leq \frac{1}{4}$, i.e. $\ell_0$ is the identically zero function. Note that, by Lemma 3, $\mathrm{nnl}(f) = \mathrm{nw}(f)$. For any hyperplane $H$ of $\mathbb{F}_2^n$, its complement $\overline{H} = \mathbb{F}_2^n \setminus H$ is also a hyperplane. We have therefore $\mathrm{w}(f_{|H}) + \mathrm{w}(f_{|\overline{H}}) = \mathrm{w}(f)$, and normalizing the weight (note that $f$ is a function in $n$ variables whereas $f_{|H}$ and $f_{|H'}$ are functions in $n - 1$ variables) we obtain

$$\mathrm{nw}(f_{|H}) + \mathrm{nw}(f_{|\overline{H}}) = 2\mathrm{nw}(f). \tag{17}$$

Let us denote $a = \min_H \mathrm{nw}(f_{|H})$ and $b = \max_H \mathrm{nw}(f_{|H})$. From (17) we see that if the restriction of $f$ to a particular hyperplane $H_0$ achieves the minimum weight, i.e. $\mathrm{nw}(f_{|H_0}) = a$, then its restriction to $\overline{H_0}$ achieves the maximum weight $\mathrm{nw}(f_{|\overline{H_0}}) = b$; using (17), we obtain

$$a + b = 2\mathrm{nw}(f) \leq \frac{1}{2}. \tag{18}$$

In particular, this means $a \leq \mathrm{nw}(f) = \mathrm{nnl}(f)$. From the definition of nonlinearity and the definition of $\mathrm{Rnnl}_1(f)$, we see that

$$\mathrm{Rnnl}_1(f) = \min(\mathrm{nnl}(f), \min_{H,\ell} \mathrm{nw}(f_{|H} + \ell)),$$

where $H$ ranges over all affine hyperplanes of $\mathbb{F}_2^n$ and $\ell$ ranges over all affine (including constant) functions over $\mathbb{F}_2^{n-1}$. To prove that $\mathrm{Rnnl}_1(f) = a$, it suffices therefore to prove that $\mathrm{nw}(f_{|H} + \ell) \geq a$ for all $H$ and $\ell$.

When $\ell$ is the constant zero function, this is immediate. When $\ell$ is the constant one function, we have $\mathrm{nw}(f_{|H} + \ell) = 1 - \mathrm{nw}(f_{|H}) \geq 1 - b$ and $1 - b > \frac{1}{2} - b \geq a$ by Eq. 18.

When $\ell$ is an affine non-constant function, we have that

$$\begin{aligned}
\mathrm{nw}(f_{|H} + \ell) &= \mathrm{nw}(f_{|H}) + \mathrm{nw}(\ell) - 2\mathrm{nw}(\ell f_{|H}) \\
&= \mathrm{nw}(f_{|H}) + \frac{1}{2} - 2\mathrm{nw}(\ell f_{|H}) \\
&\geq \frac{1}{2} - \mathrm{nw}(f_{|H}) \\
&\geq \frac{1}{2} - b \geq a,
\end{aligned}$$

where the first inequality used the fact that $\mathrm{nw}(\ell f_{|H}) \leq \mathrm{nw}(f_{|H})$, and the last inequality used (18).

For the general case, when the closest affine function $\ell_0$ is not necessarily the identically zero function, we simply define $g = f + \ell_0$ and apply the result above to $g$, noting that $\mathrm{nw}(g) = \mathrm{nd}(f, \ell_0) \leq \frac{1}{4}$ and $\mathrm{nd}(f_{|H}, (\ell_0)_{|H}) = \mathrm{nw}(g_{|H})$. □

When we consider restrictions to affine subspaces $A$ of codimensions $k$ for arbitrary $k$, the following result gives a sufficient condition for a closest affine function to $f$ to also be a closest affine function to all the restrictions $f_{|A}$.

**Proposition 5** *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and let $1 \leq k < n$. Assume there is an affine function $\ell_0$ satisfying the the following two conditions:*

$$\mathrm{Rnnl}_{k-1}(f) \geq \min_{\mathrm{codim}(A)=k} \mathrm{nd}(f_{|A}, (\ell_0)_{|A})$$

*and*

$$\min_{\mathrm{codim}(A)=k} \mathrm{nd}(f_{|A}, (\ell_0)_{|A}) + \max_{\mathrm{codim}(A)=k} \mathrm{nd}(f_{|A}, (\ell_0)_{|A}) \leq \frac{1}{2},$$

*where $A$ ranges over all affine spaces of $\mathbb{F}_2^n$ of codimension exactly $k$. Then we have*

$$\mathrm{Rnnl}_k(f) = \min_{\mathrm{codim}(A)=k} \mathrm{nd}(f_{|A}, (\ell_0)_{|A}).$$

*In particular, if $f$ satisfies the following two conditions:*

$$\mathrm{Rnnl}_{k-1}(f) \geq \min_{\mathrm{codim}(A)=k} \mathrm{nw}(f_{|A})$$

*and*

$$\min_{\mathrm{codim}(A)=k} \mathrm{nw}(f_{|A}) + \max_{\mathrm{codim}(A)=k} \mathrm{nw}(f_{|A}) \leq \frac{1}{2},$$

*where $A$ ranges over all affine spaces of $\mathbb{F}_2^n$ of codimension exactly $k$, then we have*

$$\mathrm{Rnnl}_k(f) = \min_{\mathrm{codim}(A)=k} \mathrm{nw}(f_{|A}).$$

***Proof*** We will first prove the particular case of this Proposition, namely the case where $\ell_0$ is the identically zero function. Let us denote $a = \min_{\mathrm{codim}(A)=k} \mathrm{nw}(f_{|A})$ and $b = \max_{\mathrm{codim}(A)=k} \mathrm{nw}(f_{|A})$; the hypothesis states that $a + b \leq \frac{1}{2}$ and $\mathrm{Rnnl}_{k-1}(f) \geq a$.

From the definition of $\mathrm{Rnnl}_k$, we have

$$\mathrm{Rnnl}_k(f) = \min(\mathrm{Rnnl}_{k-1}(f), \min_{\ell, A, \mathrm{codim}(A)=k} \mathrm{nw}(f_{|A} + \ell)),$$

where $\ell$ ranges over all affine (including constant) functions over $\mathbb{F}_2^{n-k}$.

To prove that $\mathrm{Rnnl}_k(f) = a$, it suffices therefore to prove that $\mathrm{nw}(f_{|A} + \ell) \geq a$ for all $A$ with $\mathrm{codim}(A) = k$ and all affine $\ell$ over $\mathbb{F}_2^{n-k}$.

**Case 1: $\ell$ is the constant zero function.** This is immediate since $\mathrm{nw}(f_{|A}+\ell) = \mathrm{nw}(f_{|A}) \geq a$.

**Case 2: $\ell$ is the constant one function.** We have

$$\mathrm{nw}(f_{|A} + \ell) = 1 - \mathrm{nw}(f_{|A}) \geq 1 - b \geq \frac{1}{2} - b \geq a,$$

where the last inequality uses the fact that $a + b \leq \frac{1}{2}$ by hypothesis.

**Case 3: $\ell$ is an affine non-constant function.** We have that

$$\begin{aligned} \mathrm{nw}(f_{|A} + \ell) &= \mathrm{nw}(f_{|A}) + \mathrm{nw}(\ell) - 2\mathrm{nw}(\ell f_{|A}) \\ &= \mathrm{nw}(f_{|A}) + \frac{1}{2} - 2\mathrm{nw}(\ell f_{|A}) \\ &\geq \frac{1}{2} - \mathrm{nw}(f_{|A}) \geq \frac{1}{2} - b \geq a, \end{aligned}$$

where the first inequality used the fact that $\mathrm{nw}(\ell f_{|A}) \leq \mathrm{nw}(f_{|A})$, and the last inequality used the hypothesis $a + b \leq \frac{1}{2}$.

For the general case, when the affine function $\ell_0$ is not necessarily the identically zero function, we simply define $g = f + \ell_0$ and apply the result above to $g$, noting that $\mathrm{nd}(f_{|H}, (\ell_0)_{|H}) = \mathrm{nw}(g_{|H})$. □

***Remark 5*** Note that when $k = 1$, Proposition 4 and its proof show that functions $f$ with $\mathrm{nnl}(f) \leq \frac{1}{4}$ satisfy the conditions in Proposition 5. In this sense, Proposition 5 can be viewed as a generalization of Proposition 4.

## 6 Symmetric functions and threshold functions

In this section we use the results of the previous section to determine the $k$-restriction nonlinearity of most threshold functions. We start will some preliminaries on symmetric functions, as threshold functions are symmetric.

Recall that a function is symmetric if it is invariant under any permutation of the variables, i.e., if $f(x_1, \ldots, x_n) = f(x_{\pi(1)}, \ldots, x_{\pi(n)})$ for any permutation $\pi$ on $\{1, \ldots, n\}$. For Boolean functions, this is equivalent to the fact that $f$ outputs the same value for all the inputs having the same Hamming weight. We start with the following property.

**Lemma 4** *Let $f$ be a symmetric function in $n$ variables and $A$ an affine space of $\mathbb{F}_2^n$. Then for all permutations $\varphi$ of the $n$ variables $\{x_1, \ldots, x_n\}$, we have*

$$\mathrm{nnl}(f_{|A}) = \mathrm{nnl}(f_{|\varphi(A)}).$$

***Proof*** Let $\varphi$ be a permutation of the $n$ variables $\{x_1, \ldots, x_n\}$. By setting $g = f \circ \varphi$, we obtain $g = f$ (since $f$ is symmetric) and from Proposition 1, we have $g_{|A} = f_{|A} \sim f_{|\varphi(A)}$. By Proposition 2, the result follows. □

We will denote by $H_{i,\epsilon}$ the hyperplane defined by the equation $x_1 + \cdots + x_i + \epsilon = 0$, where $i \in \{1, 2, \ldots, n\}$ and $\epsilon \in \{0, 1\}$.

***Remark 6*** Thanks to Lemma 4, in order to compute the 1-restriction nonlinearity of a symmetric function $f$ in $n$ variables, instead of considering its restrictions to all the $2(2^n - 1)$ hyperplanes, it suffices to consider its restriction to the $2n$ hyperplanes of the form $H_{i,\epsilon}$, the remaining ones being obtained via some permutation of the variables.

Recall that for $n$ odd, it is known [8] that the maximum nonlinearity of a symmetric function on $n$ variables is $2^{n-1} - 2^{\frac{n-1}{2}}$. As recalled in Section 2, if $n$ is even, the maximum nonlinearity of any function, including symmetric functions, is $2^{n-1} - 2^{\frac{n}{2}-1}$, attained by symmetric bent functions. We will not need that here, but all of these symmetric functions with optimal nonlinearity are quadratic [6].

Now we address the class of threshold functions, which is a particular class of symmetric Boolean functions.

**Definition 4** Let $n$ and $d$ be two integers with $0 \leq d \leq n$. The threshold function of parameters $d, n$ is the Boolean function in $n$ variables denoted by $T_{d,n}$ and defined by:

$$T_{d,n}(x) = \begin{cases} 1 \; if \; \mathrm{w}(x) \geq d \\ 0 \; otherwise. \end{cases}$$

The weight of a threshold function is

$$\mathrm{nw}(T_{d,n}) = \frac{1}{2^n} \sum_{i=d}^{n} \binom{n}{i} = \frac{1}{2^n} \sum_{i=0}^{n-d} \binom{n}{i}, \tag{19}$$

and the nonlinearity of a threshold function is given in Carlet and Méaux [4] as follows (we use the normalized version):

$$\mathrm{nnl}(T_{d,n}) = \begin{cases} \frac{1}{2} - \frac{1}{2^n}\binom{n-1}{\frac{n-1}{2}} & \text{if } d = \frac{n+1}{2} \\ \frac{1}{2^n}\sum_{i=0}^{n-d}\binom{n}{i} = \mathrm{nw}(T_{d,n}) & \text{if } d > \frac{n+1}{2} \\ \frac{1}{2^n}\sum_{i=0}^{d-1}\binom{n}{i} = 1 - \mathrm{nw}(T_{d,n}) & \text{if } d < \frac{n+1}{2}. \end{cases} \tag{20}$$

Note that the cases of $d = n$ and $d = 0$ are trivial, as $T_{n,n}$ is equal to the monomial $x_1 \cdots x_n$ and $T_{0,n}$ is the constant function 1. We will assume therefore $1 < d < n$ (the case $d = 1$ being also trivial).

It is easy to verify that $T_{d,n}(x) = 1 + T_{n-d+1,n}(\mathbf{1} + x)$, where $\mathbf{1} = (1, 1, \ldots, 1)$ (see for example [4, Property 2]), which means $T_{d,n} \sim T_{n-d+1,n}$. Combined with the fact that $\mathrm{Rnnl}_k$ is invariant under EA-equivalence (Proposition 2), we obtain:

**Corollary 3** *For all $n \geq d \geq 1$ and all $k \geq 0$ we have:*

$$\mathrm{Rnnl}_k(T_{d,n}) = \mathrm{Rnnl}_k(T_{n-d+1,n}).$$

It is therefore sufficient to determine $\mathrm{Rnnl}_k(T_{d,n})$ for $d \geq \frac{n+1}{2}$. We will first concentrate on the case $d > \frac{n+1}{2}$.

We first study the case where the threshold function is restricted to any affine space of codimension $k$ defined by the system $x_i = \epsilon_i$, where $i \in \{1, 2, ..., k\}$ and $\epsilon_i \in \{0, 1\}$, and we have:

**Lemma 5** *Let $n$ and $d$ be integers with $1 < d < n$. Let $\epsilon_1, \ldots, \epsilon_k \in \{0, 1\}$, and let $A$ be the affine space of codimension $k$ defined by the system $x_i = \epsilon_i$, where $i \in \{1, 2, ..., k\}$. Let $d' = d - \sum_{i=1}^{k} \epsilon_i$ (with the sum being over integers). Then, $(T_{d,n})_{|A} \sim T_{d',n-k}$, and therefore, we have*

$$\mathrm{nnl}((T_{d,n})_{|A}) = \begin{cases} \frac{1}{2} - \frac{1}{2^{n-k}} \binom{n-k}{\frac{n-k}{2}} & \text{if } d' = \frac{n-k+1}{2} \\ \frac{1}{2^{n-k}} \sum_{t=d'}^{n-k} \binom{n-k}{t} = \mathrm{nw}(T_{d',n-k}) & \text{if } d' > \frac{n-k+1}{2} \\ \frac{1}{2^{n-k}} \sum_{t=0}^{d'-1} \binom{n-k}{t} = 1 - \mathrm{nw}(T_{d',n-k}) & \text{if } d' < \frac{n-k+1}{2}. \end{cases}$$

***Proof*** Let $y = (x_{k+1}, \ldots, x_n)$ be the vector of $n-k$ free variables that parameterize the affine space $A$. The restricted function $(T_{d,n})_{|A}$ is a Boolean function in these $n-k$ variables. By definition, for any input $y \in \mathbb{F}_2^{n-k}$, the value of the restricted function is given by evaluating the original function $T_{d,n}$ on the corresponding vector $x \in A$. This vector $x$ is formed by setting the first $k$ coordinates to $(\epsilon_1, \ldots, \epsilon_k)$ and the remaining $n-k$ coordinates to $y$. Let us denote this specific vector $x(y) = (\epsilon_1, \ldots, \epsilon_k, x_{k+1}, \ldots, x_n)$.

The threshold function $T_{d,n}$ is defined by the Hamming weight of its input. So, for a given $y$,

$$(T_{d,n})_{|A}(y) = T_{d,n}(x(y)) = 1 \iff \mathrm{w}(x(y)) \geq d.$$

The Hamming weight of the concatenated vector $x(y)$ is the sum of the weights of its parts,

$$\mathrm{w}(x(y)) = \mathrm{w}(\epsilon_1, \ldots, \epsilon_k) + \mathrm{w}(x_{k+1}, \ldots, x_n) = \left( \sum_{i=1}^{k} \epsilon_i \right) + \mathrm{w}(y).$$

Substituting this back into the condition for the restricted function, we get

$$(T_{d,n})_{|A}(y) = 1 \iff \left( \sum_{i=1}^{k} \epsilon_i \right) + \mathrm{w}(y) \geq d.$$

Rearranging the terms, we find the condition on the weight of $y$,

$$(T_{d,n})_{|A}(y) = 1 \iff \mathrm{w}(y) \geq d - \sum_{i=1}^{k} \epsilon_i.$$

The condition becomes $\mathrm{w}(y) \geq d'$. This is precisely the definition of the threshold function $T_{d',n-k}$ acting on the input vector $y \in \mathbb{F}_2^{n-k}$.

Therefore, we have shown that $(T_{d,n})_{|A}$ is identical to the function $T_{d',n-k}$ (up to a relabeling of variables from $\{x_{k+1}, \ldots, x_n\}$ to $\{x_1, \ldots, x_{n-k}\}$). As the normalized nonlinearity is invariant under EA-equivalence,

$$\mathrm{nnl}((T_{d,n})_{|A}) = \mathrm{nnl}(T_{d',n-k}).$$

The formula for the nonlinearity of a threshold function is given by Eq. 20. Applying this formula to the function $T_{d',n-k}$ in $n-k$ variables yields the result stated in the lemma. □

***Remark 7*** Thanks to Lemmas 4, 5 remains true if the affine space $A$ is replaced by the affine space $A'$ given by the system of equations $x_{j_i} = \epsilon_i$, where $i \in \{1, , \ldots, k\}$ and $1 \leq j_1 < j_2 < \cdots < j_k \leq n$.

We now study the restrictions of $T_{d,n}$ to affine subspaces of codimension $k$. In the special case $k = 1$, we will also record a sharper characterization of the equality cases, which will be used later.

**Theorem 6** *Let $d, n, k$ be positive integers with $1 \leq k \leq d < n$ and $1 < d \leq n-k$. For any affine space $A$ of codimension exactly $k$, we have*

$$\frac{1}{2^{n-k}}\sum_{i=d}^{n-k}\binom{n-k}{i} = \mathrm{nw}(T_{d,n-k}) \leq \mathrm{nw}((T_{d,n})_{|A}) \leq \mathrm{nw}(T_{d-k,n-k}) \leq \frac{1}{2^{n-k}}\sum_{i=d-k}^{n-k}\binom{n-k}{i}.$$

*Moreover, the lower bound is achieved if $A$ is an affine space of codimension $k$ defined by the system $x_i = 0$ with $i \in \{1, \ldots, k\}$, or by any permutation of variables followed by this form. The upper bound is achieved if $A$ is an affine space of codimension $k$ defined by the system $x_i = 1$ with $i \in \{1, \ldots, k\}$, or by any permutation of variables followed by this form. In particular,* $\min_{\mathrm{codim}(A)=k} \mathrm{nw}((T_{d,n})_{|A}) = \frac{1}{2^{n-k}}\sum_{i=d}^{n-k}\binom{n-k}{i}$ *and* $\max_{\mathrm{codim}(A)=k} \mathrm{nw}((T_{d,n})_{|A}) = \frac{1}{2^{n-k}}\sum_{i=d-k}^{n-k}\binom{n-k}{i}$.

*When $k = 1$, for any hyperplane $H$, we obtain* $\mathrm{nw}(T_{d,n-1}) \leq \mathrm{nw}((T_{d,n})_{|H}) \leq \mathrm{nw}(T_{d-1,n-1})$. *Moreover, the lower bound is achieved if and only if $H$ is defined by $x_1 + \cdots + x_n = ((d-1) \bmod 2)$ or by an equation of the form $x_i = 0$, and the upper bound is achieved if and only if $H$ is defined by $x_1 + \cdots + x_n = (d \bmod 2)$ or by an equation of the form $x_i = 1$.*

***Proof*** By Lemma 4, it suffices to consider affine spaces $A$ of codimension $k$ written in canonical form $x_i = a_i(x_{k+1}, \ldots, x_n) + \epsilon_i$, $i \in \{1, \ldots, k\}$, where each $a_i$ is linear and each $\epsilon_i \in \{0, 1\}$. The ANF of $(T_{d,n})_{|A}$ can be obtained by substituting $x_1, \ldots, x_k$ by $a_1(x_{k+1}, \ldots, x_n) + \epsilon_1, \ldots, a_k(x_{k+1}, \ldots, x_n) + \epsilon_k$ in the ANF of $T_{d,n}$. We have then $(T_{d,n})_{|A}(x_{k+1}, \ldots, x_n) = T_{d,n}(a_1(x_{k+1}, \ldots, x_n) + \epsilon_1, \ldots, a_k(x_{k+1}, \ldots, x_n) + \epsilon_k, x_{k+1}, \ldots, x_n)$. Therefore, $(T_{d,n})_{|A}(x_{k+1}, \ldots, x_n) = 1$ if and only if one of the following $k+1$ mutually exclusive conditions holds:

- $\mathrm{w}(x_{k+1}, \ldots, x_n) \geq d$;
- $\mathrm{w}(x_{k+1}, \ldots, x_n) = d-1$ and $\mathrm{w}\,(a_1(x_{k+1}, \ldots, x_n) + \epsilon_1, \ldots, a_k(x_{k+1}, \ldots, x_n) + \epsilon_k) \geq 1$;

$$\vdots$$

- $\mathrm{w}(x_{k+1}, \ldots, x_n) = d-k$ and $\mathrm{w}\,(a_1(x_{k+1}, \ldots, x_n) + \epsilon_1, \ldots, a_k(x_{k+1}, \ldots, x_n) + \epsilon_k) = k$.

Denoting by $N_{d-j,A}$ the number of vectors $(x_{k+1}, \ldots, x_n)$ of weight $d-j$ such that $\mathrm{w}\,(a_1(x_{k+1}, \ldots, x_n) + \epsilon_1, \ldots, a_k(x_{k+1}, \ldots, x_n) + \epsilon_k) \geq j$, we have

$$\mathrm{nw}((T_{d,n})_{|A}) = \frac{1}{2^{n-k}}\left(\sum_{j=d}^{n-k}\binom{n-k}{j} + \sum_{j=1}^{k} N_{d-j,A}\right).$$

Clearly,

$$0 \leq N_{n-j,A} \leq \binom{n-k}{d-j}, \tag{21}$$

so we obtain $\frac{1}{2^{n-k}}\sum_{j=d}^{n-k}\binom{n-k}{j} \leq \mathrm{w}((T_{d,n})_{|A}) \leq \frac{1}{2^{n-k}}\sum_{j=d-k}^{n-k}\binom{n-k}{j}$.

The lower bound is attained when all $\epsilon_i = 0$, that is, for spaces of the form $x_1 = \cdots = x_k = 0$, and the upper bound is attained when all $\epsilon_i = 1$, that is, for spaces of the form $x_1 = \cdots = x_k = 1$. By symmetry, the same holds after any permutation of variables. For the case of arbitary $\mathrm{nw}((T_{d,n})_{|A}) = \frac{1}{2^{n-k}}\left(\sum_{j=d}^{n-k}\binom{n-k}{j} + \sum_{k}^{j=1} N_{d-j,A}\right)$,

$a_1$ we have $N_{d-j,A} = |\{(x_{k+1}, \ldots, x_n) | W(x_{k+1}, \ldots, x_n) = n - j, w(a_1(x_{k+1}, \ldots, x_n) + \epsilon_1, \ldots, a_k(x_{k+1}, \ldots, x_n) + \epsilon_k) \geq j\}|$. This proves the stated extremal values

$$\min_{\operatorname{codim}(A)=k} \operatorname{nw}((T_{d,n})_{|A}) = \frac{1}{2^{n-k}} \sum_{j=d}^{n-k} \binom{n-k}{j}$$

and

$$\max_{\operatorname{codim}(A)=k} \operatorname{nw}((T_{d,n})_{|A}) = \frac{1}{2^{n-k}} \sum_{j=d-k}^{n-k} \binom{n-k}{j}.$$

When $k = 1$, we can say more. Let $H$ be a hyperplane. Since $T_{d,n}$ is symmetric, it is enough to consider hyperplanes $H_{i,\epsilon}$ defined by

$$x_1 + \cdots + x_i + \epsilon = 0, \qquad i \in \{1, \ldots, n\},\ \epsilon \in \{0, 1\}.$$

Substituting $x_1 = x_2 + \cdots + x_i + \epsilon$, we obtain $(T_{d,n})_{|H_{i,\epsilon}}(x_2, \ldots, x_n) = 1$ if and only if one of the following mutually exclusive conditions holds:

- $\operatorname{w}(x_2, \ldots, x_n) \geq d$;
- $\operatorname{w}(x_2, \ldots, x_n) = d - 1$ and $x_2 + \cdots + x_i + \epsilon = 1$.

Therefore

$$\operatorname{nw}((T_{d,n})_{|H_{i,\epsilon}}) = \frac{1}{2^{n-1}} \left( \sum_{j=d}^{n-1} \binom{n-1}{j} + N_{i,\epsilon} \right),$$

where $N_{i,\epsilon}$ is the number of vectors $(x_2, \ldots, x_n)$ of weight $d-1$ such that $x_2 + \cdots + x_i + \epsilon = 1$. Thus

$$0 \leq N_{i,\epsilon} \leq \binom{n-1}{d-1}. \tag{22}$$

The minimum is attained exactly when $N_{i,\epsilon} = 0$. This happens in two cases:

- if $i = 1$ and $\epsilon = 0$, i.e. for the coordinate hyperplane $x_1 = 0$;
- if $i = n$ and $\epsilon = (d - 1) \bmod 2$, because then for every vector of weight $d - 1$ we have

  $$x_2 + \cdots + x_n \equiv d - 1 \pmod 2,$$

  so $x_2 + \cdots + x_n + \epsilon = 0$ for all such vectors.

For $1 < i < n$, among the vectors $(x_2, \ldots, x_n)$ of weight $d - 1$ there are some for which $x_2 + \cdots + x_i$ is even and some for which it is odd, so in this case

$$0 < N_{i,\epsilon} < \binom{n-1}{d-1}.$$

Hence the lower bound is achieved if and only if $H$ is given by $x_i = 0$ for some $i$, or by

$$x_1 + \cdots + x_n = ((d - 1) \bmod 2).$$

Similarly, the maximum is attained exactly when

$$N_{i,\epsilon} = \binom{n-1}{d-1}.$$

This happens in two cases:

- if $i = 1$ and $\epsilon = 1$, i.e. for the coordinate hyperplane $x_1 = 1$;

- if $i = n$ and $\epsilon = d \bmod 2$, because then for every vector of weight $d-1$, $x_2+\cdots+x_n+\epsilon = 1$.

Again, for $1 < i < n$ the inequalities in Eq. 22 are strict. Therefore the upper bound is achieved if and only if $H$ is given by $x_i = 1$ for some $i$, or by $x_1 + \cdots + x_n = (d \bmod 2)$. This proves the additional claim for $k = 1$. □

In order to apply Proposition 4 to threshold functions, we will need to examine, using (20), which threshold functions have nonlinearity at most $\frac{1}{4}$. To this end, for a given integer $n \geq 2$, we need to consider which integers $v$ satisfy the following inequality

$$\sum_{i=0}^{v} \binom{n}{i} \leq 2^{n-2}. \tag{23}$$

Denote by $B(n)$ the highest value of $v$ for which the inequality (23) holds. Using the well-known identities $\binom{n}{i} = \binom{n}{n-i}$ and $\sum_{i=0}^{n} \binom{n}{i} = 2^n$, we see that $B(n) < \lfloor \frac{n}{2} \rfloor$. We will also need the following straightforward combinatorial identity, for any $1 \leq v \leq n$:

$$\sum_{i=0}^{v} \binom{n+1}{i} = \sum_{i=0}^{v} \left( \binom{n}{i-1} + \binom{n}{i} \right) = \binom{n}{v} + 2\sum_{i=0}^{v-1} \binom{n}{i} \tag{24}$$

which also yields the inequalities

$$2\sum_{i=0}^{v-1} \binom{n}{i} < \sum_{i=0}^{v} \binom{n+1}{i} < 2\sum_{i=0}^{v} \binom{n}{i}. \tag{25}$$

Using (25), we see that $B(n) \leq B(n+1)$. Some immediate computations for $n = 6, 7, \ldots, 21$ give the values 1, 2, 2, 2, 3, 3, 4, 4, 5, 5, 6, 6, 7, 7, 7, 8 for $B(n)$.

Using Theorem 6, Proposition 4 and Corollary 3, we obtain the following result.

**Theorem 7** *Let $d, n$ be integers with $1 < d < n$. Define $B(n)$ as the largest integer $v$ such that $\sum_{i=0}^{v} \binom{n}{i} \leq 2^{n-2}$. The following hold:*

*(i) If $d \geq n - B(n)$ (which is equivalent to $\mathrm{nw}(T_{d,n}) \leq \frac{1}{4}$, and also implies $\mathrm{nnl}(T_{d,n}) \leq \frac{1}{4}$ and $d > \frac{n+1}{2}$) we have*

$$\begin{aligned} \mathrm{Rnnl}_1(T_{d,n}) &= \mathrm{nw}((T_{d,n})_{|H_{1,0}}) = \mathrm{nw}(T_{d,n-1}) = \mathrm{nnl}(T_{d,n-1}) \\ &= \frac{1}{2^{n-1}} \sum_{i=d}^{n-1} \binom{n-1}{i} \\ &= \mathrm{nnl}(T_{d,n}) - \frac{1}{2^n} \binom{n-1}{d-1}. \end{aligned} \tag{26}$$

*(ii) If $d \leq B(n) + 1$ (which is equivalent to $\mathrm{nw}(T_{d,n}) \geq \frac{3}{4}$, and also implies $\mathrm{nnl}(T_{d,n}) \leq \frac{1}{4}$ and $d < \frac{n+1}{2}$) we have*

$$\begin{aligned} \mathrm{Rnnl}_1(T_{d,n}) &= \mathrm{nnl}(T_{d-1,n-1}) \\ &= \frac{1}{2^{n-1}} \sum_{i=0}^{d-2} \binom{n-1}{i} \\ &= \mathrm{nnl}(T_{d,n}) - \frac{1}{2^n} \binom{n-1}{d-1}. \end{aligned} \tag{27}$$

***Proof*** Since $B(n) < \lfloor \frac{n}{2} \rfloor$, one can check that $d \geq n - B(n)$ implies $d > \frac{n+1}{2}$ (and therefore $\mathrm{nnl}(T_{d,n}) = \mathrm{nw}(T_{d,n})$, by Eq. 20) and $d \leq B(n) + 1$ implies $d < \frac{n+1}{2}$ (and therefore $\mathrm{nnl}(T_{d,n}) = 1 - \mathrm{nw}(T_{d,n})$, by Eq. 20).

To prove (i) we apply Proposition 4. First, we verify that the conditions of the proposition are satisfied. Since $n - d \leq B(n)$, by the definition of $B(n)$ we have $\sum_{i=0}^{n-d} \binom{n}{i} \leq 2^{n-2}$, which is equivalent to $\sum_{i=d}^{n} \binom{n}{i} \leq 2^{n-2}$, or in normalized form, using (19), $\mathrm{nw}(T_{d,n}) \leq \frac{1}{4}$. Therefore, by Proposition 4, we have

$$\mathrm{Rnnl}_1(T_{d,n}) = \min_H \mathrm{nw}((T_{d,n})_{|H}).$$

By Theorem 6, we have

$$\min_H \mathrm{nw}((T_{d,n})_{|H}) = \mathrm{nw}((T_{d,n})_{|H_{1,0}}) = \mathrm{nw}(T_{d,n-1}) = \frac{1}{2^{n-1}} \sum_{i=d}^{n-1} \binom{n-1}{i}.$$

For the equality $\mathrm{nw}(T_{d,n-1}) = \mathrm{nnl}(T_{d,n}) - \frac{1}{2^n}\binom{n-1}{d-1}$, we use equations (20), (19) and (24). Namely, since $d > \frac{n+1}{2}$, we have

$$\begin{aligned} \mathrm{nnl}(T_{d,n}) &= \frac{1}{2^n} \sum_{i=0}^{n-d} \binom{n}{i} = \frac{1}{2^n}\binom{n-1}{n-d} + \frac{1}{2^{n-1}} \sum_{i=0}^{n-d-1} \binom{n-1}{i} \\ &= \frac{1}{2^n}\binom{n-1}{d-1} + \mathrm{nw}(T_{d,n-1}). \end{aligned}$$

The proof of (i) is complete. Item (ii) follows by using (i) and Corollary 3. □

***Example 2*** The 1-restriction nonlinearity $\mathrm{Rnnl}_1(T_{d,n})$ of the threshold function $T_{d,n}$ is determined in Table 1 using Theorem 7 above, for selected values of $n$ with $d = n - 4$. The case $n = 7, d = 3$ uses part (ii) of the theorem (with $d \leq B(n) + 1$), while the cases $n = 12$ through $n = 20$ use part (i) of the theorem (with $d \geq n - B(n)$).

**Table 1** Nonlinearity and 1-restriction nonlinearity for $T_{d,n}$ (with $d = n - 4$)

| $n$ | $d$ | $\mathrm{nnl}(T_{d,n})$ | $\mathrm{Rnnl}_1(T_{d,n})$ |
|---|---|---|---|
| 7 | 3 | $\frac{29}{128} \approx 0.22656$ | $\frac{7}{64} \approx 0.10938$ |
| 12 | 8 | $\frac{397}{2048} \approx 0.19385$ | $\frac{29}{256} \approx 0.11328$ |
| 13 | 9 | $\frac{1093}{8192} \approx 0.13342$ | $\frac{299}{4096} \approx 0.07300$ |
| 14 | 10 | $\frac{1471}{16384} \approx 0.08978$ | $\frac{189}{4096} \approx 0.04614$ |
| 15 | 11 | $\frac{1941}{32768} \approx 0.05923$ | $\frac{235}{8192} \approx 0.02869$ |
| 16 | 12 | $\frac{2517}{65536} \approx 0.03841$ | $\frac{9}{512} \approx 0.01758$ |
| 17 | 13 | $\frac{1607}{65536} \approx 0.02452$ | $\frac{697}{65536} \approx 0.01064$ |
| 18 | 14 | $\frac{253}{16384} \approx 0.01544$ | $\frac{417}{65536} \approx 0.00636$ |
| 19 | 15 | $\frac{1259}{131072} \approx 0.00961$ | $\frac{247}{65536} \approx 0.00377$ |
| 20 | 16 | $\frac{1549}{262144} \approx 0.00591$ | $\frac{145}{65536} \approx 0.00221$ |

Theorem 7 covers most, but not all functions $T_{d,n}$ with $d \neq \frac{n+1}{2}$. For example, using the values of $B(n)$ computed above, we see that for $n = 6, \ldots, 12$ it covers all such functions except $T_{4,6} \sim T_{3,6}$, $T_{5,8} \sim T_{4,8}$, $T_{6,9} \sim T_{4,9}$, $T_{6,10} \sim T_{5,10}$, $T_{7,11} \sim T_{5,11}$, $T_{7,12} \sim T_{6,12}$ (computations of the nonlinearity of some of these functions is given in the Appendix). The ratio of functions not covered by the theorem out of all functions $T_{d,n}$ with $d \neq \frac{n+1}{2}$ becomes vanishingly small as $n$ grows, as explained in the remark below.

***Remark 8*** Recall the well-known approximation for the sum of the first $v$ binomial coefficients, when $v \leq n/2$, namely $\sum_{i=0}^{v} \binom{n}{i} \leq 2^{nH_2(\frac{v}{n})}$, where $H_2(p) = -p\log_2(p) - (1-p)\log_2(1-p)$ with $p \in [0, 1]$ is the entropy function (see, for example [7]). Therefore, for inequality (23) to hold, it is sufficient that $nH_2(\frac{v}{n}) \leq n-2$, that is, $H_2(\frac{v}{n}) \leq 1 - \frac{2}{n}$. Defining $B'(n)$ as the largest integer $v \leq n/2$ for which $H_2\left(\frac{v}{n}\right) \leq 1 - \frac{2}{n}$, we see that $B'(n)$ is a lower bound for $B(n)$. Numerical computations for $n = 6, \ldots, 21$ give the values of 1, 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 5, 5, 5, 6, 6 for $B'(n)$. Recalling the behavior of the function $H_2(p)$, which reaches the maximum of $H_2(p) = 1$ if and only if $p = 0.5$, we see that $\lim_{n\to\infty}\left(1 - \frac{2}{n}\right) = 1$, and therefore $\lim_{n\to\infty} \frac{B'(n)}{n} = 0.5$. The proportion of functions $T_{d,n}$, which are covered by Theorem 7 equals $\frac{2(B(n)+1)}{n}$. Taking into account that $B'(n) \leq B(n) < \frac{n}{2}$ and $\lim_{n\to\infty} \frac{B'(n)}{n} = 0.5$, we obtain $\lim_{n\to\infty} \frac{2(B(n)+1)}{n} = 1$. In other words, the proportion of functions not covered by Theorem 7 becomes vanishingly small as $n$ grows.

Computer calculations for small $n$ show that (26) and (27) in Theorem 7 actually hold for all $T_{d,n}$ with $d \neq \frac{n+1}{2}$, not just for the cases $d \geq n - B(n)$ and $d \leq B(n) + 1$. This suggests the following conjecture.

**Conjecture 1** *Let $1 < d \leq n$ with $d \neq \frac{n+1}{2}$. For all $d > \frac{n+1}{2}$, we have* $\mathrm{Rnnl}_1(T_{d,n}) = \mathrm{nnl}(T_{d,n-1})$. *For all $d < \frac{n+1}{2}$, we have* $\mathrm{Rnnl}_1(T_{d,n}) = \mathrm{nnl}(T_{d-1,n-1})$.

We now consider the case of $d = \frac{n+1}{2}$. Numerical computations for small values of $n$ suggest the following conjecture.

**Conjecture 2** *For $d = \frac{n+1}{2}$, we have* $\mathrm{Rnnl}_1(T_{d,n}) = \mathrm{nnl}((T_{d,n})_{|H_{2,0}}) = \mathrm{nd}((T_{d,n})_{|H_{2,0}}, \ell)$, *where $\ell$ is the linear function defined as $\ell(x_2, \ldots, x_n) = x_2$ and $H_{2,0}$ is the hyperplane define by the equation $x_1 + x_2 = 0$.*

The last quantity appearing in the conjecture above can be explicitly computed for any threshold function and we do so in the next result.

**Proposition 8** *The distance from $(T_{d,n})_{|H_{2,0}}$ to the linear function $\ell$ defined as $\ell(x_2, \ldots, x_n) = x_2$ is* $\mathrm{nd}((T_{d,n})_{|H_{2,0}}, \ell) = \frac{1}{2} - \frac{1}{2^{n-1}}\binom{n-1}{d-1}$.

***Proof*** The restriction of $T_{d,n}$ to the hyperplane $H_{2,0}$, defined by the equation $x_1 = x_2$, can be obtained by substituting $x_1$ with $x_2$ in $T_{d,n}$. Therefore $(T_{d,n})_{|H_{2,0}}(x_2, \ldots, x_n) = 1$ if and only if $\mathrm{w}(x_2, x_2, x_3, \ldots, x_n) \geq d$. The number of tuples $(x_2, \ldots, x_n)$ for which $(T_{d,n})_{|H_{2,0}}(x_2, \ldots, x_n) = 1$ and $\ell(x_2, \ldots, x_n) = x_2 = 0$ is therefore the number of tuples $(x_3, \ldots, x_n)$ with $\mathrm{w}(x_3, \ldots, x_n) \geq d$, i.e. $\sum_{i=d}^{n}\binom{n-2}{i}$ tuples. The number of tuples $(x_2, \ldots, x_n)$ for which $(T_{d,n})_{|H_{2,0}}(x_2, \ldots, x_n) = 0$ and $\ell(x_2, \ldots, x_n) = x_2 = 1$ is the number of tuples $(x_3, \ldots, x_n)$ with $\mathrm{w}(1, 1, x_3, \ldots, x_n) < d$, or equivalently $\mathrm{w}(x_3, \ldots, x_n) \leq d - 3$ i.e. $\sum_{i=0}^{d-3}\binom{n-2}{i}$ tuples. In total, therefore

$$\mathrm{d}((T_{d,n})_{|H_{2,0}}, \ell) = \sum_{i=0}^{d-3}\binom{n-2}{i} + \sum_{i=d}^{n}\binom{n-2}{i}$$

$$= 2^{n-2} - \binom{n-2}{d-2} - \binom{n-2}{d-1}$$
$$= 2^{n-2} - \binom{n-1}{d-1},$$

which, after dividing by $2^{n-1}$ for normalization, gives the required value for $\mathrm{nd}((T_{d,n})_{|H_{2,0}}, \ell)$. □

**Corollary 4** *If Conjecture 2 holds, then for the threshold function $T_{d,n}$ with $d = \frac{n+1}{2}$, we have*

$$\mathrm{Rnnl}_1\left(T_{\frac{n+1}{2},n}\right) = \frac{1}{2} - \frac{1}{2^{n-1}}\binom{n-1}{\frac{n-1}{2}} = \mathrm{nnl}\left(T_{\frac{n+1}{2},n}\right) - \frac{1}{2^n}\binom{n-1}{\frac{n-1}{2}}.$$

*If Conjectures 1 and 2 hold, then*

$$\mathrm{Rnnl}_1(T_{d,n}) = \mathrm{nnl}(T_{d,n}) - \frac{1}{2^n}\binom{n-1}{d}$$

**Corollary 5** *for all $1 < d < n$.*

Theorem 7 can be generalized as follows to restrictions of threshold functions to spaces of arbitrary codimension $k$.

**Theorem 9** *Let $d, n, k$ be integers such that $1 < d < n$, $1 \le k \le d$ and $k \le n - d$. Define $B(m)$ as the largest integer $v$ such that $\sum_{i=0}^{v}\binom{m}{i} \le 2^{m-2}$.*

*(i) If $d \ge n - B(n-k+1)$ (which also implies $d > \frac{n+1}{2}$), we have*

$$\mathrm{Rnnl}_k(T_{d,n}) = \mathrm{nw}((T_{d,n})_{|A_0^k}) = \mathrm{nw}(T_{d,n-k}) = \mathrm{nnl}(T_{d,n-k})$$
$$= \frac{1}{2^{n-k}}\sum_{i=d}^{n-k}\binom{n-k}{i}$$
$$= \mathrm{Rnnl}_{k-1}(T_{d,n}) - \frac{1}{2^{n-k+1}}\binom{n-k}{d-1},$$

*where $A_0^k$ is the affine space of codimension $k$ defined by the system of equations $x_i = 0$ with $i \in \{1, \ldots, k\}$.*

*(ii) If $d \le B(n-k+1)+1$ (which also implies $d < \frac{n+1}{2}$), we have*

$$\mathrm{Rnnl}_k(T_{d,n}) = \mathrm{nnl}(T_{d-k,n-k})$$
$$= \frac{1}{2^{n-k}}\sum_{i=0}^{d-k-1}\binom{n-k}{i}$$
$$= \mathrm{Rnnl}_{k-1}(T_{d,n}) - \frac{1}{2^{n-k+1}}\binom{n-k}{d-1}.$$

***Proof*** Since $B(n-k+1) \le B(n) < \lfloor\frac{n}{2}\rfloor$, one can check that $d \ge n - B(n-k+1)$ implies $d > \frac{n+1}{2}$ (and therefore $\mathrm{nnl}(T_{d,n}) = \mathrm{nw}(T_{d,n})$, by Eq. 20) and $d \le B(n-k+1)+1$ implies $d < \frac{n+1}{2}$ (and therefore $\mathrm{nnl}(T_{d,n}) = 1 - \mathrm{nw}(T_{d,n})$, by Eq. 20).

To prove (i), we proceed by induction on $k$. The base case $k = 1$ is established by Theorem 7. For the inductive step, assume the result holds for $k-1$, and that $d \ge n - B(n-k+1)$. According to Proposition 5, it suffices to verify the following two conditions for affine spaces of codimension exactly $k$. □

**Condition 1** We need to show that

$$\mathrm{Rnnl}_{k-1}(T_{d,n}) \geq \min_{\mathrm{codim}(A)=k} \mathrm{nw}((T_{d,n})_{|A}). \tag{28}$$

Since $n-d \leq B(n-k+1) \leq B(n-k+2)$, by the inductive hypothesis we have

$$\mathrm{Rnnl}_{k-1}(T_{d,n}) = \frac{1}{2^{n-k+1}} \sum_{i=d}^{n-k+1} \binom{n-k+1}{i}.$$

By Theorem 6, we have

$$\min_{\mathrm{codim}(A)=k} \mathrm{nw}((T_{d,n})_{|A}) = \frac{1}{2^{n-k}} \sum_{i=d}^{n-k} \binom{n-k}{i}.$$

Substituting these values in Eq. 28, this means we need to prove

$$\frac{1}{2^{n-k+1}} \sum_{i=d}^{n-k+1} \binom{n-k+1}{i} \geq \frac{1}{2^{n-k}} \sum_{i=d}^{n-k} \binom{n-k}{i}. \tag{29}$$

Multiplying both sides by $2^{n-k+1}$, changing the index of summation and using the usual combinatorial identity $\binom{m}{t} = \binom{m}{m-t}$, the inequality above is equivalent to

$$\sum_{i=0}^{n-d-k+1} \binom{n-k+1}{i} \geq 2 \sum_{i=0}^{n-d-k} \binom{n-k}{i},$$

which follows from (25).

**Condition 2** We need to show that

$$\min_{\mathrm{codim}(A)=k} \mathrm{nw}((T_{d,n})_{|A}) + \max_{\mathrm{codim}(A)=k} \mathrm{nw}((T_{d,n})_{|A}) \leq \frac{1}{2}. \tag{30}$$

Using Theorem 6, this is equivalent to

$$\frac{1}{2^{n-k}} \left( \sum_{i=d}^{n-k} \binom{n-k}{i} + \sum_{i=d-k}^{n-k} \binom{n-k}{i} \right) \leq \frac{1}{2}.$$

Multiplying by $2^{n-k}$, changing the index of summation and using the usual identity $\binom{m}{t} = \binom{m}{m-t}$, this is equivalent to

$$\sum_{i=0}^{n-d-k} \binom{n-k}{i} + \sum_{i=0}^{n-d} \binom{n-k}{i} \leq 2^{n-k-1}. \tag{31}$$

Using (24), the left hand side of the inequality above becomes

$$\begin{aligned} \sum_{i=0}^{n-d-k} \binom{n-k}{i} + \sum_{i=0}^{n-d} \binom{n-k}{i} &\leq \sum_{i=0}^{n-d-1} \binom{n-k}{i} + \sum_{i=0}^{n-d} \binom{n-k}{i} \\ &= \binom{n-k}{n-d} + 2 \sum_{i=0}^{n-d-1} \binom{n-k}{i} \\ &= \sum_{i=0}^{n-d} \binom{n-k+1}{i}. \end{aligned} \tag{32}$$

Since we assumed $n - d \leq B(n - k + 1)$, by the definition of $B(.)$ we have

$$\sum_{i=0}^{n-d} \binom{n-k+1}{i} \leq 2^{n-k-1},$$

which, combined with Eq. 32, proves (31). Therefore Condition 2 is satisfied.

Item (ii) follows from (i) and Corollary 3.

***Remark 9*** The threshold functions covered by Theorem 9 are a subset of the threshold functions covered by Theorem 7. While Theorem 7 covers all $T_{d,n}$ with weights in the intervals $[0, \frac{1}{4}] \cup [\frac{3}{4}, 1]$, a computation using (24) shows that the condition $d \geq n - B(n - k + 1)$ in Theorem 9($i$) is equivalent to $\mathrm{nw}(T_{d,n}) \leq \frac{1}{4} - \sum_{i=0}^{k-2} \frac{1}{2^{n-i}} \binom{n-i-1}{d-i-1}$. Similarly, the condition $d \leq B(n-k+1)+1$ in Theorem 9($ii$) is equivalent to $\mathrm{nw}(T_{d,n}) \geq \frac{3}{4} + \sum_{i=0}^{k-2} \frac{1}{2^{n-i}} \binom{n-i-1}{d-i-1}$.

***Remark 10*** In Remark 2 we saw that for constructing a space $A$ such that $\mathrm{nnl}(f_{|A})$ is minimal, a "greedy" approach does not always work. Note that for those threshold functions for which Theorem 9 applies, the "greedy" approach does work.

## 7 Conclusion and future work

In this work, we introduced the concept of the $k$-restriction normalized nonlinearity, $\mathrm{Rnnl}_k(f)$, to formally measure the stability of a Boolean function's normalized nonlinearity when restricted to arbitrary affine subspaces of codimension at most $k$. This metric is crucial for assessing resistance against cryptanalytic attacks that operate by constraining the input space.

Our primary contribution is the comprehensive characterization of $\mathrm{Rnnl}_k$ for two important classes of functions. For the first class, direct sums of monomials, we established an explicit formula for $\mathrm{Rnnl}_k(f)$ (Corollary 1) showing that the nonlinearity remains relatively stable under restriction. Since any quadratic function in $n$ variables is EA-equivalent to a direct sum of $m$ monomials for some $m \leq n/2$, we also obtain an explicit formula for all quadratic functions, namely $\mathrm{Rnnl}_k(f) = \frac{1}{2}\left(1 - \frac{1}{2^{m-k}}\right)$. This confirms that for quadratic functions, $\mathrm{Rnnl}_k(f)$ decreases predictably and minimally.

For the second class, threshold functions ($T_{d,n}$, which return 1 if and only if the weight of the input is at least $d$), we proved in Theorem 7 that when the normalized nonlinearity of the function is at most 0.25 we have $\mathrm{Rnnl}_1(T_{d,n}) = \mathrm{nnl}(T_{d,n}) - \frac{1}{2^n}\binom{n-1}{d-1}$. This covers all threshold functions except those where the threshold $d$ is equal or close to $\frac{n+1}{2}$. The precise characterization for these remaining boundary cases remains an open challenge. Based on extensive computational analysis of these functions, we posited Conjectures 1 and 2 to predict the exact $\mathrm{Rnnl}_1$ values for these remaining parameters; if these conjectures hold, the relation $\mathrm{Rnnl}_1(T_{d,n}) = \mathrm{nnl}(T_{d,n}) - \frac{1}{2^n}\binom{n-1}{d-1}$ holds for all threshold functions. Establishing a formal proof for these conjectures would represent a conclusive end to the characterization of the 1-restriction normalized nonlinearity for all threshold functions. Furthermore, extending this analysis to $\mathrm{Rnnl}_k(T_{d,n})$ for $k \geq 2$ (for which we covered only certain values of $d, n$ in Theorem 9), and applying the $\mathrm{Rnnl}_k$ metric to other cryptographically relevant classes of Boolean functions, such as algebraic immunity-optimal functions, would be of significant interest.

## Appendix: Restricted nonlinearity for some threshold functions

We give below tables of the 1-restricted nonlinearity for some threshold functions. In the tables, $H_{i,\epsilon}$ denotes the hyperplane given by the equation $x_1 + \cdots + x_i + \epsilon = 0$, where $i \in \{1, 2, \ldots, n\}$ and $\epsilon \in \{0, 1\}$. The notation $\frac{\mu}{\nu}$ denotes the pair of values $\mu = 2^{n-1}\mathrm{nnl}\left((T_{d,n})_{|H_{i,0}}\right)$ and $\nu = 2^{n-1}\mathrm{nnl}\left((T_{d,n})_{|H_{i,1}}\right)$. The last two columns give respectively the minimal nonlinearity under restrictions to hyperplanes and the nonlinearity of the original function (Tables 2, 3, and 4).

**Table 2** $2^7\mathrm{nnl}((T_{5,8})_{|H_{i,0}})$, $2^7\mathrm{nnl}((T_{5,8})_{|H_{i,1}})$, $2^7\mathrm{Rnnl}_1(T_{5,8})$, $2^7\mathrm{nnl}(T_{5,8})$

| $T_{5,8}$ \ $H_{i,\epsilon}$ | $\frac{H_{1,0}}{H_{1,1}}$ | $\frac{H_{2,0}}{H_{2,1}}$ | $\frac{H_{3,0}}{H_{3,1}}$ | $\frac{H_{4,0}}{H_{4,1}}$ | $\frac{H_{5,0}}{H_{5,1}}$ | $\frac{H_{6,0}}{H_{6,1}}$ | $\frac{H_{7,0}}{H_{7,1}}$ | $\frac{H_{8,0}}{H_{8,1}}$ | $2^7\mathrm{Rnnl}_1\left(T_{5,8}\right)$ | $2^7\mathrm{nnl}\left(T_{5,8}\right)$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $T_{5,8}$ | $\frac{29}{44}$ | $\frac{29}{44}$ | $\frac{45}{44}$ | $\frac{45}{44}$ | $\frac{45}{44}$ | $\frac{45}{44}$ | $\frac{29}{44}$ | $\frac{29}{44}$ | 29 | 93/2 |

**Table 3** $2^8\mathrm{nnl}((T_{5,9})_{|H_{i,0}})$, $2^8\mathrm{nnl}((T_{5,9})_{|H_{i,1}})$, $2^8\mathrm{Rnnl}_1(T_{5,9})$, $2^8\mathrm{nnl}(T_{5,9})$.

| $T_{5,9}$ \ $H_{i,\epsilon}$ | $\frac{H_{1,0}}{H_{1,1}}$ | $\frac{H_{2,0}}{H_{2,1}}$ | $\frac{H_{3,0}}{H_{3,1}}$ | $\frac{H_{4,0}}{H_{4,1}}$ | $\frac{H_{5,0}}{H_{5,1}}$ | $\frac{H_{6,0}}{H_{6,1}}$ | $\frac{H_{7,0}}{H_{7,1}}$ | $\frac{H_{8,0}}{H_{8,1}}$ | $\frac{H_{9,0}}{H_{9,1}}$ | $2^8\mathrm{Rnnl}_1\left(T_{5,9}\right)$ | $2^8\mathrm{nnl}\left(T_{5,9}\right)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $T_{5,9}$ | $\frac{93}{93}$ | $\frac{58}{88}$ | $\frac{93}{93}$ | $\frac{90}{88}$ | $\frac{93}{93}$ | $\frac{90}{88}$ | $\frac{93}{93}$ | $\frac{58}{88}$ | $\frac{93}{93}$ | 58 | 93 |

**Table 4** $2^9\mathrm{nnl}((T_{6,10})_{|H_{i,0}})$, $2^9\mathrm{nnl}((T_{6,10})_{|H_{i,1}})$, $2^9\mathrm{Rnnl}_1(T_{6,10})$, $2^9\mathrm{nnl}(T_{6,10})$

| $T_{6,10}$ \ $H_{i,\epsilon}$ | $\frac{H_{1,0}}{H_{1,1}}$ | $\frac{H_{2,0}}{H_{2,1}}$ | $\frac{H_{3,0}}{H_{3,1}}$ | $\frac{H_{4,0}}{H_{4,1}}$ | $\frac{H_{5,0}}{H_{5,1}}$ | $\frac{H_{6,0}}{H_{6,1}}$ | $\frac{H_{7,0}}{H_{7,1}}$ | $\frac{H_{8,0}}{H_{8,1}}$ | $\frac{H_{9,0}}{H_{9,1}}$ | $\frac{H_{10,0}}{H_{10,1}}$ | $2^9\mathrm{Rnnl}_1\left(T_{6,10}\right)$ | $2^9\mathrm{nnl}\left(T_{6,10}\right)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $T_{6,10}$ | $\frac{130}{186}$ | $\frac{130}{186}$ | $\frac{190}{186}$ | $\frac{190}{186}$ | $\frac{190}{190}$ | $\frac{190}{190}$ | $\frac{190}{186}$ | $\frac{190}{186}$ | $\frac{186}{130}$ | $\frac{186}{130}$ | 130 | 193 |

**Acknowledgements** The authors would like to thank the editors for efficiently handling our paper and the reviewers for their careful reading, beneficial comments, and constructive suggestions to improve our paper. The first two-named authors were supported by the Engineering and Physical Sciences Research Council, UK [EPSRC grant EP/W03378X/1]. The third-named author (P.S.) would like to thank the second-named author (A.S.) and the University of Loughborough for the invitation and the excellent working conditions while this paper was being written.

**Author Contributions** These authors contributed equally to this work.

**Funding** The first two authors were supported by the Engineering and Physical Sciences Research Council, UK [EPSRC grant EP/W03378X/1]

**Data Availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing Interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give

appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

## References

1. Canteaut, A., Carlet, C., Charpin, P., Fontaine, C.: On cryptographic properties of the cosets of $R(1, m)$. IEEE Trans. Inf. Theory **47**, 1494–1513 (2001)
2. Carlet, C.: Boolean Functions for Cryptography and Error Correcting Codes, Cambridge University Press (2021)
3. Carlet, C., Feukoua, S., Sălăgean, A.: The stability of the algebraic degree of Boolean functions when restricted to affine spaces. Designs, Codes Cryptogr, pp. 1–34 (2025)
4. Carlet, C., Méaux, P.: A complete study of two classes of Boolean functions: direct sums of monomials and threshold functions. IEEE Trans. Inf. Theory **68**(5), 3404–3425 (2021)
5. Carlet, C., Méaux, P., Rotella, Y.: Boolean functions with restricted input and their robustness; application to the FLIP cipher. IACR Trans. Symmetric Cryptology **3**, 192–227 (2017)
6. Cusick, T.W., Stănică, P.: Cryptographic Boolean functions and applications, Academic Press (2017). (2nd ed.)
7. MacWilliams, F. J., Sloane, N. J. A: The Theory of Error-Correcting Codes. North-Holland (1978)
8. Maitra, S., Sarkar, P.: Maximum Nonlinearity of Symmetric Boolean Functions on Odd Number of Variables. IEEE Trans. Inf. Theory **48**(59), 2626–2630 (2002)
9. Méaux, P., Journault, A., Standaert, F.X., Carlet, C.: Towards stream ciphers for efficient FHE with low-noise ciphertexts. In: Fischlin, M., Coron, J.S. (eds.) Adv. Crypt. - EUROCRYPT 2016. LNCS, vol. 9665, pp. 311–343. Springer, Berlin, Heidelberg (2016)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.