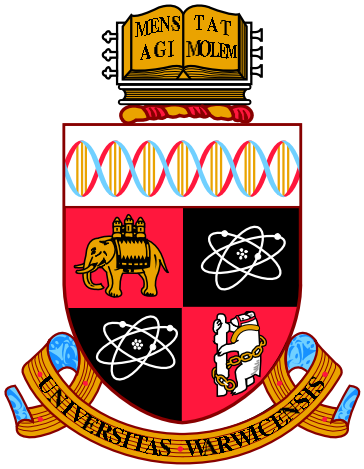


Combinatorial and learning properties of threshold and related functions

by

Elena Zamaraeva

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Mathematics Institute

September 2020

Contents

Acknowledgments

First and foremost I would like to express my sincere gratitude to my supervisors professors Vadim Lozin and Nikolai Zolotykh for their continued guidance and support, as well as their detailed proofreading of this thesis. Professor Vadim Lozin welcomed me in the Mathematics Institute and has always been incredibly patient and generous with his time, resources, and knowledge. I am particularly grateful for the many interesting problems he suggested and the invaluable insight I gained through our mathematical conversations.

The previous contribution of professor Nikolai Zolotykh to this area laid the cornerstone of my work and I am grateful for the support and generous encouragement he has been providing me since he supervised my Master thesis.

Additionally, I would like to thank professor Joviša Žunić for his insightful advice and fruitful collaboration. His work motivated the research described in Chapters 4 and 5.

Finally, I would like to thank my family for the support throughout my entire Ph.D study. I thank my mother, my sister and my brother for their support and care every day from the distance. I am grateful to my husband, Viktor for his indubitable belief in me. I dedicate this Ph.D thesis to my two lovely children, Andrew and Natalia who are the pride and joy of my life.

Declarations

Chapter 2 is based on a joint work with Vadim Lozin, Igor Razgon, Viktor Zamaraev, and Nikolai Zolotykh [38, 39]. Chapters 4 and 5 are based on a joint work with Joviša Žunić. Except for that, I declare that, to the best of my knowledge, the material contained in this dissertation is original and my own work except where otherwise indicated, cited, or commonly known. The material in this dissertation is submitted to the University of Warwick for the degree of Doctor of Philosophy, and has not been submitted to any other university or for any other degree.

Abstract

This thesis is devoted to the study of threshold and related functions over two different domains.

In the first part, we consider Boolean threshold and linear read-once functions. We show that a positive function f of n variables has exactly $n + 1$ extremal points if and only if it is linear read-once. The class of linear read-once functions is also known to be the intersection of the classes of read-once and threshold functions. Generalizing this result we show that the class of linear read-once functions is the intersection of read-once and Chow functions. Then, we characterize the class of linear read-once functions by means of minimal forbidden subfunctions within the universe of read-once and the universe of threshold functions. Furthermore, we prove that the subclass of threshold functions with the minimum specification number does not coincide with the class of linear read-once functions thereby disproving a conjecture of Anthony et al. from 1995 [6]. We propose techniques that we believe might be useful to characterize the subclass of threshold functions with the minimum specification number. We also found all threshold functions up to 6 variables from this subclass.

In the second part, we turn to the k-threshold functions over a two-dimensional rectangular grid, i.e. the functions representable as the conjunctions of k threshold functions. In [34] a bijection between non-constant threshold functions over this domain and prime segments was established. This result was used in [34] to estimate the number of threshold functions asymptotically. No asymptotic formula for the number of k-threshold functions was known for $k > 1$. We consider the case $k = 2$ and characterize 2-threshold functions via the pairs of oriented prime segments with specific properties. We apply this characterization to derive an asymptotic formula for the number of 2-threshold functions depending on the size of the grid. This result also improves a trivial upper bound on the number of

k-threshold functions for $k > 2$.

In the third part, we study specifying sets of k-threshold functions. First, we consider the class of k-threshold functions with non-fixed k, i.e. the union of all k-threshold functions for all k. We prove that a function in this class has a unique minimal specifying set with respect to the class of k-threshold functions with non-fixed k and provide the structural characterization of this set. For two-dimensional k-threshold functions we refine the given structure and derive a bound on the size of the minimal specifying set. Then we fix the parameter $k = 2$ and analyze the size and the number of minimal specifying sets of two-dimensional 2-threshold functions. In particular, we construct a sequence of 2-threshold functions over a squared grid of size $m \times m$ for which the number of minimal specifying sets grows as $\Theta(m^2)$. We also show that if a two-dimensional 2-threshold function has a unique representation as a conjunction of two threshold functions, then its specification number is at most 9. The results of this part of the thesis were obtained prior the Ph.D. study and are therefore provided as a supplementary material.

Chapter 1

Introduction

Threshold functions naturally arise in various theoretical and applied studies. Informally, a threshold function defines a partition of a given set of points (domain) into two parts via separating linear inequality. In this thesis we will focus on threshold and related functions, defined over a discrete set of points in the d-dimensional space $\mathbb{R}^d$.

Let S be a discrete set of points in $\mathbb{R}^d$. A function f that maps S to $\{0,1\}$ is called *threshold* (or *linearly separable* or a *halfspace*) if there exist d *weights* $w_1,\ldots,w_d \in \mathbb{R}$ and a *threshold* $t \in \mathbb{R}$ such that, for every point $(x_1,\ldots,x_d) \in S$,

$$f(x_1,\ldots,x_d) = 1 \iff \sum_{i=1}^{d} w_i x_i \geq t.$$

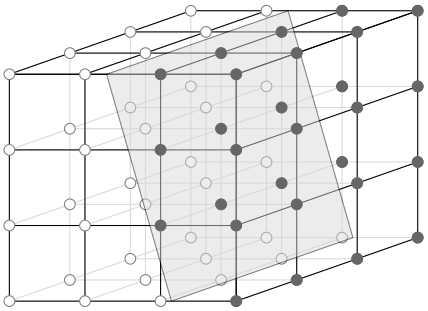
The inequality $w_1x_1 + \ldots + w_dx_d \leq t$ is called a *threshold inequality* representing the function f. The hyperplane $w_1x_1 + \ldots + w_dx_d = t$ is called a *separating hyperplane* for the function f. It is not hard to see that there are uncountably many different threshold inequalities (and separating hyperplanes) representing a given threshold function.

A function f that maps S to $\{0,1\}$ is called k*-threshold* (or a k*-halfspace*) if there exist at most k threshold functions $f_1,\ldots,f_k$ such that f is the logical conjunction of the functions $f_1,\ldots,f_k$, i.e.

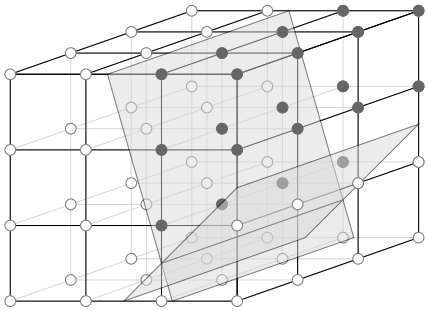
$$f = f_1 \wedge \ldots \wedge f_k.$$

We say that the functions $f_1,\ldots,f_k$ or a system of inequalities corresponding to the functions *define* the k-threshold function f. A k-threshold function is called *proper* k*-threshold* if it is not $(k-1)$-threshold.

Remark. In the literature, various generalizations of linear threshold functions are studied. For example, a *degree-*d *polynomial threshold function* is a function $f : \{-1,1\}^n \to \{-1,1\}$ expressible as $f(x) = \mathrm{sgn}(p(x))$, where p is a multivariate degree-d polynomial with real coefficients, and sgn is -1 for negative arguments and 1 otherwise (see e.g. [17]). Another example constitute so-called k*-valued threshold functions* separating a given set of



(a) A threshold function with threshold inequality $7x_1 + 2x_2 \geq 15$.



(b) A 2-threshold function with system of threshold inequalities $\begin{cases} 7x_1 + 2x_2 \geq 15, \\ -2x_1 + 2x_2 \geq -3. \end{cases}$

Figure 1.1: The black points are true points of the corresponding threshold and 2-threshold functions defined over $\mathbb{Z}_4^3$.

points into $k+1$ subsets by k parallel hyperplanes (see e.g. [43]).

We emphasize that in the current work generalization of threshold functions is not due to increase of the degree of a separating structure or the number of thresholds, but due to increase of the number of linear separating structures.

In applications of threshold functions, a Boolean hypercube or the d-dimensional integer hypercube $\mathbb{Z}_n^d = \{0, 1, \ldots, n-1\}^d$ of size n are common domains. Fig. 1.1 illustrates the partition of $\mathbb{Z}_4^3$ by threshold and 2-threshold functions.

An unceasing interest to threshold and k-threshold functions is due to their relevance in many areas of computer science, such as machine learning, digital geometry, and computer vision. In [4] Angluin considered a model of concept learning with specific kinds of queries, including membership and equivalence queries. In this model a domain X and a concept class $C \subseteq 2^X$ are known to both the learner (or learning algorithm) and the teacher. The goal of the learner is to identify an unknown target concept $T \in C$ that has been fixed by the teacher. To this end, the learner may ask the teacher membership queries "*does an element x belong to T?*", to which the teacher answers "*yes*" or "*no*"; or the learner may ask the equivalence queries "*is $T' \in C$ the target concept T?*", to which the teacher answers "*yes*" or provides a counterexample, i.e. an element x which belongs to either T or T' but not to both. The learning complexity of a learning algorithm with respect to a concept class C is the minimum number of queries sufficient for the algorithm to identify any concept in C. The learning complexity of a concept class C is defined as the minimum learning complexity of a learning algorithm with respect to C over all learning algorithms which learn C using membership queries, equivalence queries or both.

In terms of Angluin's model, $\{0,1\}$-valued functions defined over a set of points S can be considered as characteristic functions of the concepts. Here S is the domain and a function f mapping S to $\{0,1\}$ defines a concept represented by its set of ones.

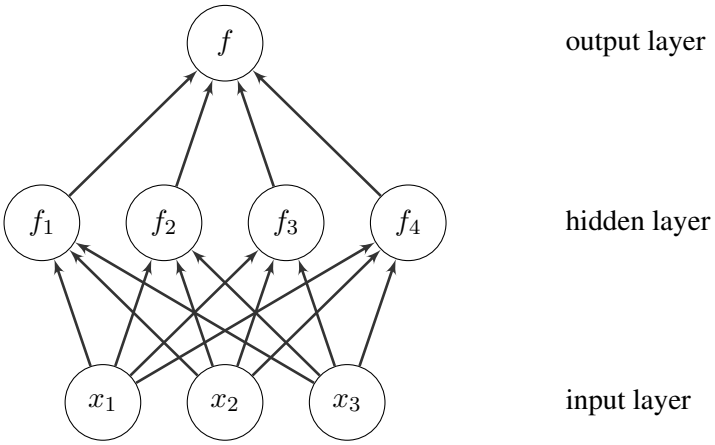


Figure 1.2: A 2-layer network with 3 inputs, 4 nodes in the hidden layer and one output corresponds to a 4-threshold function $f(x_1, x_2, x_3) = f_1 \wedge f_2 \wedge f_3 \wedge f_4$.

Lower bounds on the complexity of learning with equivalence queries for threshold, k-threshold functions, and some related geometric objects were derived in [40]. In [27] the complexity of learning with membership queries was studied for threshold functions over $\mathbb{Z}_n^d$. In [12] the authors provided an efficient algorithm of learning with membership queries for k-threshold functions over the two-dimensional grid. Learning of Boolean k-threshold functions was studied, for instance, in [10, 28, 36, 33].

In the realm of learning with membership queries, *teaching* (*specifying*) *sets* and *essential points* play a central role. Let C be a class of functions mapping S to $\{0, 1\}$, and let f be a function from C. A set of points $T \subseteq S$ is called a *teaching* or *specifying set for f with respect to C* if no other function from C coincides with f in all points of T. The number of points in a minimum specifying set for f (with respect to C) is called the *specification number of f* (*with respect to C*). Clearly, the target function f cannot be identified without learning the values of all points in a specifying set for f. Therefore, the specification number of f is a lower bound on the complexity of learning with membership queries.

A point $x \in S$ is called *essential for f with respect to C* if there exists a function $g \in C$ such that $g(x) \neq f(x)$ and g coincides with f on $S \setminus \{x\}$. It is easy to see that the set of essential points of f is a subset of any specifying set of f. Furthermore, it is known that the set of essential points for a threshold function is a specifying set by itself (see e.g. [6, 44]). The specification number and essential points of Boolean threshold functions were studied in [6]. Minimal specifying sets and the specification number of threshold functions over $\mathbb{Z}_n^d$ were investigated in [3, 44, 45, 50]. The specifying sets of k-threshold functions were studied by the author in [48, 49] and the corresponding results are presented in Appendix B.

In the theory of neural networks, a feed-forward 2-layer neural network plays a

central role. The nodes from the hidden and output layers of this network often represent threshold functions. Moreover, it is known that every Boolean function can be expressed as a composition of threshold functions (see [15]), and hence, 2-layer feed-forward neural networks with threshold functions as hidden and output nodes can compute any Boolean function. For instance, a k-threshold function over d variables corresponds to a 2-layer network with d inputs, k nodes in the hidden layer representing k threshold functions, and the output node representing the conjunction function (see Fig. 1.2). The connections between Boolean functions and artificial neural networks are surveyed in [5].

There is one-to-one correspondence between k-threshold functions over $\mathbb{Z}^d$ and integer (digital) polytopes with vertices in $\mathbb{Z}^d$ if we do not restrict the parameter k. Therefore, the study of k-threshold functions might be useful in the study of integer polytopes, which are among the most important objects in integer linear programming and digital geometry.

In digital geometry, the problem of polyhedral separability can be formulated in terms of k-threshold functions as follows: given a domain S, a finite set of points $T \subseteq S$, and a positive integer k, does there exist a k-threshold function f over S such that T is the set of ones of f? The problem of polyhedral separability is widely investigated (see [19, 41, 11, 9, 18, 8, 22, 23]). In particular, in [11] the authors studied bilinear separation which is closely related to 2-threshold functions, and the papers [22, 23] are devoted to the polyhedral separability problem in two- and three-dimensional spaces.

The thesis has two self-contained parts. We outline below each of these parts.

1.1 Boolean threshold and linear read-once functions

The first part consists of Chapters 2 and 3 and was motivated by the following conjecture of Anthony, Brightwell, and Shawe-Taylor [6].

Conjecture 1. *Linear read-once functions are the only Boolean threshold functions with the minimum possible specification number.*

In Chapter 2 we formulate and prove some weaker statement than Conjecture 1. Namely, we show that linear read-once functions are the only positive threshold functions with the minimum possible number of extremal points. Even stronger, we show that linear read-once functions are the only positive functions with the minimum possible number of extremal points. Furthermore, in Chapter 2 we also establish a relation between the classes of linear read-once, read-once, and Chow functions, by showing that the first one is the intersection of the other two. Finally, we find the set of minimal read-once functions which are not linear read-once and the set of minimal threshold functions which are not linear read-once.

In Chapter 3 we disprove Conjecture 1 by providing an infinite sequence of counterexamples. On the other hand, in the attempt to characterize the subclass of threshold

functions with the minimum possible specification number we obtain some positive results. First, we show that this class is free from self-dual functions. Then, we provide two procedures that extend a threshold function with the minimum specification number in such a way that the resulting function has more variables and also has the minimum specification number. Furthermore, we observe that symmetric variables play special role in the subclass of threshold functions with the minimum specification number. We finish the chapter by enumerating all non-canalyzing threshold functions of up to 5 variables with the minimum specification number. Appendix A provides the list of the those functions of 6 variables.

1.2 2-threshold functions over a rectangular grid

In the second part of the thesis we turn to a two-dimensional integer grid with non-necessarily equal "width" and "height" and consider 2-threshold functions over this domain. A useful characterization of two-dimensional threshold functions via oriented prime segments was provided in [34] to estimate asymptotically the number of threshold functions and was used in subsequent works (see [1, 35, 42]). The asymptotic formula for the number of threshold functions was later improved in [1], [2], and [25], however no formulas were known for the number of k-threshold functions for any $k > 1$.

In Chapter 4 we provide a characterization of 2-threshold functions establishing the bijection between almost all pairs of oriented prime segments with certain properties and almost all 2-threshold functions. In Chapter 5 we apply this structural result to derive an asymptotic formula for the number of two-dimensional 2-threshold functions. The obtained formula also improves trivial upper bounds on the number of k-threshold functions for $k > 2$.

1.3 k-threshold functions and their specifying sets

Appendix B contains results obtained by the author of the thesis prior to starting Ph.D at Warwick and published in [48] and [49]. These results reveal relations between essential points of k-threshold functions and their specifying sets. In particular, we describe the structure of the specifying sets of two-dimensional 2-threshold functions and identify a subclass of functions with specification number bounded by a constant. We include these results for completeness, because they are closely related to Chapters 4 and 5.

Chapter 2

Linear read-once and related Boolean functions

2.1 Introduction

Linear read-once functions constitute a remarkable subclass of several classes of Boolean functions. Within the universe of threshold functions the importance of linear read-once functions is due to the fact that they attain the minimum value of the specification number, i.e. of the number of Boolean points that uniquely specify a function in this universe (see [6]). To study the range of values of specification number of threshold functions one can be restricted to positive threshold functions depending on all their variables, in which case the functions can be completely specified by their sets of extremal points, i.e. maximal zeros and minimal ones. In other words, the specification number of a positive threshold function is upper bounded by the number of its extremal points. For a linear read-once function of n variables, these numbers coincide and equal $n + 1$. In 1995 Anthony et al. [6] conjectured that for all other threshold functions the specification number is strictly greater than $n + 1$.

Despite the fact that the conjecture is not true (the counterexamples will be provided in Chapter 3) we show that the set of extremal points satisfies the statement of the conjecture, i.e. a positive threshold Boolean function depending on all its n variables has $n+1$ extremal points if and only if it is linear read-once. Moreover, not only does this result hold for positive threshold functions but also for all positive functions.

In [20], it was shown that the class of linear read-once functions is the intersection of the classes of read-once and threshold functions. Generalizing this result we show that the class of linear read-once functions is the intersection of read-once and Chow functions. We also find the set of minimal read-once functions which are not linear read-once and the set of minimal threshold functions which are not linear read-once.

The organization of the chapter is as follows. All preliminary information related to the topic of the chapter, including definitions and notation, is presented in Section 2.2.

Section 2.3 is devoted to the number of extremal points in positive functions. In Section 2.4 we show that the class of linear read-once functions is the intersection of the classes of read-once and Chow functions, and identify the set of minimal read-once functions which are not linear read-once. In Section 2.5 we provide the set of minimal threshold functions which are not linear read-once.

2.2 Preliminaries

Let $B = \{0, 1\}$. For a *Boolean n-dimensional hypercube* B^n we define *sub-hypercube* $B^n(x_{i_1} = \alpha_1, \ldots, x_{i_k} = \alpha_k)$ as the set of all points of B^n for which coordinate i_j is equal to α_j for every $j = 1, \ldots, k$. For a point $\mathbf{x} \in B^n$ we denote by $\overline{\mathbf{x}}$ the point in B^n with $(\overline{\mathbf{x}})_i = 1$ if and only if $(\mathbf{x})_i = 0$ for every $i \in [n]$.

For a Boolean function $f = f(x_1, \ldots, x_n)$ on B^n, $k \in [n]$, and $\alpha_k \in \{0, 1\}$ we denote by $f_{|x_k=\alpha_k}$ the Boolean function on B^{n-1} defined as follows:

$$f_{|x_k=\alpha_k}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) = f(x_1, \ldots, x_{k-1}, \alpha_k, x_{k+1}, \ldots, x_n).$$

For $i_1, \ldots, i_k \in [n]$ and $\alpha_1, \ldots, \alpha_k \in \{0, 1\}$ we denote by $f_{|x_1=\alpha_1,\ldots,x_k=\alpha_k}$ the function $(f_{|x_1=\alpha_1,\ldots,x_{k-1}=\alpha_{k-1}})_{|x_k=\alpha_k}$. We say that $f_{|x_1=\alpha_1,\ldots,x_k=\alpha_k}$ is the *restriction* of f to $x_1 = \alpha_1, \ldots, x_k = \alpha_k$. We also say that a Boolean function g is a *restriction* (or *subfunction*) of a Boolean function $f \in B^n$ if there exist $i_1, \ldots, i_k \in [n]$ and $\alpha_1, \ldots, \alpha_k \in \{0, 1\}$ such that $g = f_{|x_1=\alpha_1,\ldots,x_k=\alpha_k}$.

It is known (see e.g. Theorem 9.3 in [15]) that the class of threshold functions is closed under taking restrictions, i.e. any restriction of a threshold function is again a threshold function.

A variable x_k is called *irrelevant* for f if $f_{|x_k=1} \equiv f_{|x_k=0}$, i.e., $f_{|x_k=1}(\mathbf{x}) = f_{|x_k=0}(\mathbf{x})$ for every $\mathbf{x} \in B^{n-1}$. Otherwise, x_k is called *relevant* for f. If x_k is irrelevant for f we also say that f *does not depend on* x_k.

By $\preccurlyeq$ we denote a partial order over the set B^n, induced by inclusion in the power set lattice of the n-set. In other words, $\mathbf{x} \preccurlyeq \mathbf{y}$ if $(\mathbf{x})_i = 1$ implies $(\mathbf{y})_i = 1$. In this case we will say that $\mathbf{x}$ is *below* $\mathbf{y}$. When $\mathbf{x} \preccurlyeq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$ we will sometimes write $\mathbf{x} \prec \mathbf{y}$. We denote by $\vee$ and $\wedge$ the logical disjunction and conjunction respectively. We also often omit the operator $\wedge$ and denote conjunction by mere juxtaposition.

We say that a Boolean formula is in disjunctive normal form (DNF) if it is a disjunction consisting of one or more conjunctive clauses, each of which is a conjunction of one or more literals (variables or their negations). A Boolean formula is in conjunctive normal form (CNF) if it is a conjunction consisting of one or more disjunctive clauses, each of which is a disjunction of one or more literals. A DNF (resp. CNF) is called *minimal* if it has the minimum possible number of clauses among all DNFs (resp. CNF) representing the

same Boolean function.

Two Boolean functions f and g are *congruent*, if they are identical up to renaming (without identification) and/or negation of variables.

Definition 2.2.1. A Boolean function f is called *positive* (also known as *positive monotone* or *increasing*) if $f(\mathbf{x}) = 1$ and $\mathbf{x} \preccurlyeq \mathbf{y}$ imply $f(\mathbf{y}) = 1$. We say that a Boolean formula is *positive* if it does not contain the operation of negation.

It is clear that a positive Boolean function admits representation via a positive Boolean formula. For a positive Boolean function f, the set of its false points forms a down-set and the set of its true points forms an up-set of the partially ordered set $(B^n, \preccurlyeq)$. We denote by

Z^f the set of maximal false points,

U^f the set of minimal true points.

We will refer to a point in Z^f as a *maximal zero of* f and to a point in U^f as a *minimal one of* f. A point will be called an *extremal point of* f if it is either a maximal zero or a minimal one of f. We denote by $r(f)$ the number of extremal points of f.

Let $k \in \mathbb{N}, k \geq 2$. A Boolean function f on B^n is *k-summable* if, for some $r \in \{2, \ldots, k\}$, there exist r (not necessarily distinct) false points $\mathbf{x}_1, \ldots, \mathbf{x}_r$ and r (not necessarily distinct) true points $\mathbf{y}_1, \ldots, \mathbf{y}_r$ such that $\sum_{i=1}^{r} \mathbf{x}_i = \sum_{i=1}^{r} \mathbf{y}_i$ (where the summation is over $\mathbb{R}^n$). A function is *asummable* if it is not k-summable for all $k \geq 2$.

Theorem 2 ([21])**.** *A Boolean function is a threshold function if and only if it is asummable.*

Definition 2.2.2. A Boolean function f is called *read-once* if it can be represented by a Boolean formula using the operations of conjunction, disjunction, and negation in which every variable appears at most once. We say that such a formula is a *read-once formula* for f.

Example 3. *The boolean formulas $x_1 \vee x_2 \wedge x_3$ and $x_1 \wedge x_2 \vee \overline{x_1} \wedge x_3$ are read-once and non-read-once respectively, the same holds for the functions representable by these formulas. However, a non-read-once formula can represent a read-once function. For example, the formula $x_1 \wedge x_2 \vee x_1 \wedge x_3$ is non-read-once, while the corresponding function $f(x_1, x_2, x_3) = x_1 \wedge x_2 \vee x_1 \wedge x_3$ is read-once, because it is also representable by the read-once formula $x_1 \wedge (x_2 \vee x_3)$.*

Definition 2.2.3. A read-once function f is *linear read-once* (*lro*) if it is either a constant function, or it can be represented by a *nested formula* defined recursively as follows:

1. both literals x and $\overline{x}$ are nested formulas;

2. $x \vee t$, $x \wedge t$, $\overline{x} \vee t$, $\overline{x} \wedge t$ are nested formulas, where x is a variable and t is a nested formula that contains neither x, nor $\overline{x}$.

Example 4. *The both functions* $f(x_1, x_2, x_3, x_4) = x_1 \vee x_2 \wedge (x_3 \vee x_4)$ *and* $g(x_1, x_2, x_3, x_4) = x_1 \wedge x_2 \vee x_3 \wedge x_4$ *are read-once, but only* f *is linear read-once.*

In [6], lro functions depending on all variables have been called *nested*. It is not difficult to see that an lro function f is positive if and only if the nested formula representing f does not contain negations.

In [20], it has been shown that the class of lro functions is precisely the intersection of threshold and read-once functions.

Definition 2.2.4. A Boolean function $f = f(x_1, \ldots, x_n)$ is called *canalyzing*[1] if there exists $i \in [n]$ such that $f_{|x_i=0}$ or $f_{|x_i=1}$ is a constant function.

It is easy to see that if f is a positive canalyzing function then $f_{|x_i=0} \equiv \mathbf{0}$ or $f_{|x_i=1} \equiv \mathbf{1}$, for some $i \in [n]$. In Example 4 the function f is canalyzing as $f_{|x_1=1} \equiv \mathbf{1}$, and the function g is non-canalyzing.

Let $\mathfrak{T}_n$ be the class of threshold Boolean functions of n variables.

Definition 2.2.5. A set of points $S \subseteq B^n$ is a *specifying set for* a threshold function f of n variables if the only threshold function consistent with f on S is f itself. In this case we also say that S *specifies* f in the class threshold functions. The minimal cardinality of a specifying set for f in $\mathfrak{T}_n$ is called the *specification number of* f (in $\mathfrak{T}_n$) and denoted $\sigma_{\mathfrak{T}_n}(f)$.

It was shown in [29] and later in [6] that the specification number of a threshold function of n variables is at least $n + 1$.

Theorem 5 ([29, 6])**.** *For any threshold Boolean function* f *of* n *variables* $\sigma_{\mathfrak{T}_n}(f) \geq n+1$.

Also, in [6] it was shown that the lower bound is attained for lro functions.

Theorem 6 ([6])**.** *For any lro function* f *depending on all its* n *variables,* $\sigma_{\mathfrak{T}_n}(f) = n + 1$.

Moreover, the same paper proves that the lower bound is only attained for functions with no irrelevant variables.

Theorem 7 ([6])**.** *Suppose* $f \in \mathfrak{T}_n$ *depends on exactly* k *variables. Then*

$$\sigma_{\mathfrak{T}_n}(f) \geq 2^{(n-k)}(k+1).$$

[1] The notion of canalyzing functions was introduced in [31] and is widely used in biological applications of Boolean networks. In [32, 30, 37] linear read-once functions are called *nested canalyzing functions* and studied as a special case of canalyzing functions.

In estimating the specification number of a threshold Boolean function $f \in \mathfrak{T}_n$ it is often useful to consider essential points of f defined as follows.

Definition 2.2.6. A point $\mathbf{x}$ is *essential* for f (with respect to the class $\mathfrak{T}_n$), if there exists a function $g \in \mathfrak{T}_n$ such that $g(\mathbf{x}) \neq f(\mathbf{x})$ and $g(\mathbf{y}) = f(\mathbf{y})$ for every $\mathbf{y} \in B^n$, $\mathbf{y} \neq \mathbf{x}$.

Clearly, any specifying set for f contains all essential points for f. It turns out that the essential points alone are sufficient to specify f in $\mathfrak{T}_n$ [14]. Therefore, we have the following well-known result.

Theorem 8 ([14]). *The specification number $\sigma_{\mathfrak{T}_n}(f)$ of a function $f \in \mathfrak{T}_n$ is equal to the number of essential points of f.*

Moreover, there is also a strong connection between essential and extremal points of a positive threshold function.

Theorem 9 ([6]). *Let f be a positive threshold function from $\mathfrak{T}_n$ depending on all its variables, then the set of essential points of f is a subset of the set of extremal points of f.*

The following result is a restriction of Theorem 4 in [50] (proved for threshold functions of many-valued logic) to the case of Boolean threshold functions.

Theorem 10 ([50]). *A true point of a Boolean threshold function f is essential if and only if there is a separating hyperplane containing it.*

Thus, the set of all essential ones (*resp.* zeros) of $f \in \mathfrak{T}_n$ is the union of all points in B^n belonging to at least one separating hyperplane for the function f (*resp.* $\overline{f}$).

2.3 Positive functions and the number of extremal points

It was observed in [6] that in the study of specification number of threshold functions, one can be restricted to positive functions. To prove Theorem 6, the authors of [6] first showed that for a positive threshold function f depending on all its variables the set of extremal points specifies f. Then they proved that for any positive lro function f of n relevant variables the number of extremal points is $n + 1$.

In addition to proving Theorem 6, the authors of [6] also conjectured that lro functions are the only functions with the specification number $n + 1$ in the class $\mathfrak{T}_n$.

Conjecture 11 ([6]). *If $f \in \mathfrak{T}_n$ has the specification number $n + 1$, then f is linear read-once.*

In this section, we show that this conjecture becomes a true statement if we replace 'specification number' by 'number of extremal points'.

Theorem 12. *Let $f = f(x_1, \ldots, x_n)$ be a positive function with $k \geq 0$ relevant variables. Then the number of extremal points of f is at least $k + 1$. Moreover f has exactly $k + 1$ extremal points if and only if f is linear read-once.*

We will prove Theorem 12 by induction on n. The statement is easily verifiable for $n = 1$. Let $n > 1$ and assume that the theorem is true for functions of at most $n - 1$ variables. In the rest of the section we prove the statement for n-variable functions. Our strategy consists of three major steps. First, we prove the statement for canalyzing functions in Section 2.3.2. This case includes lro functions. Then, in Section 2.3.3, we prove the result for non-canalyzing functions f such that for each variable x_i both restrictions $f_{|x_i=0}$ and $f_{|x_i=1}$ are canalyzing. Finally, in Section 2.3.4, we consider the case of non-canalyzing functions f depending on a variable x_i such that at least one of the restrictions $f_{|x_i=0}$ and $f_{|x_i=1}$ is non-canalyzing. In Section 2.3.1, we introduce some terminology and prove a preliminary result.

2.3.1 A property of extremal points

We say that a maximal zero (*resp.* minimal one) $\mathbf{y}$ of $f(x_1, \ldots, x_n)$ *corresponds to a variable* x_i if $(\mathbf{y})_i = 0$ (*resp.* $(\mathbf{y})_i = 1$). It is not difficult to see that for any relevant variable x_i, there exists at least one minimal one and at least one maximal zero corresponding to x_i. We say that an extremal point of f *corresponds to a set S of variables* if it corresponds to at least one variable in S.

Lemma 13. *For every set S of k relevant variables of a positive function f, there exist at least $k + 1$ extremal points corresponding to this set.*

Proof. Let S be a minimal counterexample and let P be the set of extremal points corresponding to the variables in S. Without loss of generality we assume that S consists of the first k variables of the function, i.e. $S = \{x_1, \ldots, x_k\}$. Due to the minimality of S we may also assume that $|P| = k$ and for every proper subset S' of S there exist at least $|S'| + 1$ extremal points corresponding to S'. This implies, by Hall's Theorem of distinct representatives [24], that there exists a bijection between S and P mapping variable x_i to a point $\mathbf{a}^i \in P$ *corresponding* to x_i.

Let $\mathbf{a}$ be any maximal zero in P. We denote by $\mathbf{b}$ the point which coincides with $\mathbf{a}$ in all coordinates beyond the first k, and for each $i \in \{1, 2, \ldots, k\}$ we define the i-th coordinate of $\mathbf{b}$ to be 1 if $\mathbf{a}^i$ is a maximal zero, and to be 0 if $\mathbf{a}^i$ is a minimal one.

Assume first that $f(\mathbf{b}) = 0$ and let $\mathbf{c}$ be any maximal zero above $\mathbf{b}$ (possibly $\mathbf{b} = \mathbf{c}$). If $(\mathbf{c})_1 = \ldots = (\mathbf{c})_k = 1$, then $\mathbf{a} \prec \mathbf{c}$, contradicting that $\mathbf{a}$ is a maximal zero. Therefore, $(\mathbf{c})_i = 0$ for some $1 \leq i \leq k$ and hence $\mathbf{c}$ is a maximal zero corresponding to $x_i \in S$. Moreover, $\mathbf{c}$ is different from any maximal zero $\mathbf{a}^j \in P$ because the j-th coordinate of $\mathbf{a}^j \in P$ is 0, while the j-th coordinate of $\mathbf{c}$ is 1.

Suppose now that $f(\mathbf{b}) = 1$ and let $\mathbf{c}$ be any minimal one below $\mathbf{b}$ (possibly $\mathbf{b} = \mathbf{c}$). If $(\mathbf{c})_1 = \ldots = (\mathbf{c})_k = 0$, then $\mathbf{c} \prec \mathbf{a}$, contradicting the positivity of f. Therefore, $(\mathbf{c})_i = 1$ for some $1 \leq i \leq k$ and hence $\mathbf{c}$ is a minimal one corresponding to $x_i \in S$. Moreover, $\mathbf{c}$ is different from any minimal one $\mathbf{a}^j \in P$ because the j-th coordinate of $\mathbf{a}^j \in P$ is 1, while the j-th coordinate of $\mathbf{c}$ is 0.

A contradiction in both cases shows that there is no counterexamples to the statement of the lemma. □

2.3.2 Canalyzing functions

Lemma 14. *Let $f = f(x_1, \ldots, x_n)$ be a positive canalyzing function with $k \geq 0$ relevant variables. Then the number of extremal points of f is at least $k+1$. Moreover f has exactly $k+1$ extremal points if and only if f is lro.*

Proof. The case $k = 0$ is trivial, and therefore we assume that $k \geq 1$.

Let x_i be a variable of f such that $f_{|x_i=0} \equiv \mathbf{0}$ (the case $f_{|x_i=1} \equiv \mathbf{1}$ is similar). Let $f_0 = f_{|x_i=0}$ and $f_1 = f_{|x_i=1}$. Clearly, x_i is a relevant variable of f, otherwise $f \equiv \mathbf{0}$, that is, $k = 0$. Since every relevant variable of f is relevant for at least one of the functions f_0 and f_1, we conclude that f_1 has $k-1$ relevant variables.

The equivalence $f_0 \equiv \mathbf{0}$ implies that for every extremal point $(\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n)$ of f_1, the corresponding point $(\alpha_1, \ldots, \alpha_{i-1}, 1, \alpha_{i+1}, \ldots, \alpha_n)$ is extremal for f. For the same reason, there is only one extremal point of f with the i-th coordinate being equal to 0, namely, the point with all coordinates equal to 1, except for the i-th coordinate. Hence, $r(f) = r(f_1) + 1$.

1. If f_1 is lro, then f is also lro, since f can be expressed as $x_i \wedge f_1$. By the induction hypothesis $r(f_1) = k$ and therefore $r(f) = k + 1$.

2. If f_1 is not lro, then f is also not lro, which is easy to see. By the induction hypothesis $r(f_1) > k$ and therefore $r(f) > k + 1$.

□

2.3.3 Non-canalyzing functions with canalyzing restrictions

In this section, we study non-canalyzing positive functions such that for each variable x_i both restrictions $f_{|x_i=0}$ and $f_{|x_i=1}$ are canalyzing.

First we remark that all variables of those functions are relevant. Indeed, if such a function has an irrelevant variable then the function is canalyzing.

Claim 15. *Let $f = f(x_1, \ldots, x_n)$ be a positive non-canalyzing function such that for each variable x_i both restrictions $f_{|x_i=0}$ and $f_{|x_i=1}$ are canalyzing. Then all variables of f are relevant.*

Proof. Let x_i be irrelevant, then $f_{|x_i=0} \equiv f_{|x_i=1}$. But $f_{|x_i=0}$, $f_{|x_i=1}$ are canalyzing, hence there exists $p \in [n]$ such that $f_{|x_i=0,\, x_p=0} \equiv f_{|x_i=1,\, x_p=0} \equiv \mathbf{0}$ or $f_{|x_i=0,\, x_p=1} \equiv f_{|x_i=1,\, x_p=1} \equiv \mathbf{1}$. In the former case $f_{|x_p=0} \equiv \mathbf{0}$, in the latter case $f_{|x_p=1} \equiv \mathbf{1}$. In any case f is canalyzing. Contradiction. □

Claim 16. *Let $f = f(x_1, \ldots, x_n)$ be a positive non-canalyzing function such that for each variable x_i both restrictions $f_{|x_i=0}$ and $f_{|x_i=1}$ are canalyzing. Then for each i,*

(a) there exists a maximal zero that contains 0's in exactly two coordinates one of which is i,

(b) there exists a minimal one that contains 1's in exactly two coordinates one of which is i.

Proof. Fix an i and denote $f_0 = f_{|x_i=0}$, $f_1 = f_{|x_i=1}$. Since f_0 is canalyzing, there exists $p \in [n]$ such that $f_{0|x_p=0} \equiv \mathbf{0}$ or $f_{0|x_p=1} \equiv \mathbf{1}$. We claim that the latter case is impossible. Indeed, the positivity of f and $f_{0|x_p=1} \equiv \mathbf{1}$ imply $f_{1|x_p=1} \equiv \mathbf{1}$, and therefore $f_{|x_p=1} \equiv \mathbf{1}$. This contradicts the assumption that f is non-canalyzing. Thus, $f_{0|x_p=0} \equiv \mathbf{0}$. Now we claim that the Boolean point $\mathbf{y}$ with exactly two 0's in coordinates i and p is a maximal zero. Indeed, if f in at least one of three points above $\mathbf{y}$ is 0, then, by positivity of f, $f_{|x_i=0} = 0$ or $f_{|x_p=0}$, which contradicts the assumption that f is non-canalyzing.

Similarly, one can show that $f_{1|x_r=1} \equiv \mathbf{1}$ for some $r \in [n]$ implying that the Boolean point with exactly two 1's in coordinates i and r is a minimal one. □

Claim 17. *Let $f = f(x_1, \ldots, x_n)$ be a positive non-canalyzing function such that for each variable x_i both restrictions $f_{|x_i=0}$ and $f_{|x_i=1}$ are canalyzing. Then there is a minimal one $\mathbf{y}$ of Hamming weight 2 such that $\overline{\mathbf{y}}$ is a maximal zero, unless $n = 4$ in which case f has 6 extremal points.*

Proof. Consider a graph G_0 (*resp.* G_1) with vertex set $[n]$ every edge ij of which represents a maximal zero (*resp.* minimal one) that contains 0's (*resp.* 1's) in exactly two coordinates i and j. By Claim 16, every vertex in G_0 is covered by an edge and every vertex in G_1 is covered by an edge. From this it follows in particular that each graph G_0, G_1 has at least $\lceil n/2 \rceil$ edges.

In terms of the graphs G_0 and G_1, the claim is equivalent to saying that G_0 and G_1 have a common edge. It is not difficult to see that for $n \leq 3$ the graphs G_1 and G_0 necessarily have a common edge. Let us show that this is also the case for $n \geq 5$.

Assume that G_0 and G_1 have no common edges, i.e. every edge of G_0 is a non-edge (a pair of non-adjacent vertices) in G_1. Let us prove that

(*) every edge ij of G_0 forms a vertex cover in G_1, i.e. every edge of G_1 shares a vertex with either i or j (and not with both according to our assumption).

Indeed, let ij be an edge of G_0 and assume that G_1 contains an edge pq such that p is different from i, j and q is different from i, j. Then the minimal one corresponding to the edge pq of G_1 is below the maximal zero corresponding to the edge ij of G_0. This contradicts the positivity of f and proves (*).

Consider an edge ij in G_0. Since $n \geq 5$, then G_0 has at least 3 edges, hence from (*) we get that at least one of i, j covers at least two edges of G_1, say i covers ip and iq. Let ps be an edge of G_0 covering p. If $s \neq q$, then ps does not cover the edge iq of G_1 which contradicts to (*). If $s = q$, let t be any vertex different from i, j, p, q. The vertex t must be covered by some edge tr in G_1. If r is different from i, j then tr does not cover ij in G_0. If r is different from p, q then tr does not cover pq in G_0. In both cases we get a contradiction to (*), hence for $n \geq 5$ the graphs G_0 and G_1 necessarily have a common edge and hence the result follows in this case.

It remains to analyze the case $n = 4$. Up to renaming variables, the only possibility for G_0 and G_1 to avoid a common edge is for G_0 to have edges 12 and 34 and for G_1 to have edges 13 and 24. In other words, $(0, 0, 1, 1)$ and $(1, 1, 0, 0)$ are maximal zeros and $(1, 0, 1, 0)$ and $(0, 1, 0, 1)$ are minimal ones. By positivity, this completely defines the function f, except for two points $(0, 1, 1, 0)$ and $(1, 0, 0, 1)$. However, regardless of the value of f in these points, both of them are extremal and hence f has 6 extremal points. □

Claim 18. *Let $f = f(x_1, \ldots, x_n)$ be a positive non-canalyzing function such that for each variable x_i both restrictions $f_{|x_i=0}$ and $f_{|x_i=1}$ are canalyzing. Let* $\mathbf{y}$ *be a minimal one of Hamming weight 2 such that* $\overline{\mathbf{y}}$ *is a maximal zero. Denote the two coordinates of* $\mathbf{y}$ *containing 1's by i and s, and let $f_0 = f_{|x_i=0}$ and $f_1 = f_{|x_i=1}$.*

(a) Variable x_s is relevant for both functions f_0 and f_1.

(b) If a point $\mathbf{a} = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n) \in B^{n-1}$ is an extremal point of f_{α_i} for some $\alpha_i \in \{0, 1\}$, then $\mathbf{a}' = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \ldots, \alpha_{n-1}) \in B^n$ is an extremal point of f.

Proof. First, we note that since $\mathbf{y}$ is a minimal one, $f_{1|x_s=1} \equiv \mathbf{1}$. Similarly, since $\overline{\mathbf{y}}$ is a maximal zero, $f_{0|x_s=0} \equiv \mathbf{0}$.

To prove (a), suppose to the contrary that f_0 does not depend on x_s. Then $f_{0|x_s=1} \equiv f_{0|x_s=0} \equiv \mathbf{0}$, and therefore $f_0 \equiv \mathbf{0}$, which contradicts the assumption that f is non-canalyzing. Similarly, one can show that x_s is relevant for f_1.

Now we turn to (b) and prove the statement for $\alpha_i = 1$. For $\alpha_i = 0$ the arguments are symmetric.

Assume first that $\alpha_s = 1$. Since $\mathbf{y}$ is a minimal one, we have $f_1(\mathbf{b}) = 1$ for all $\mathbf{b} = (\beta_1, \ldots, \beta_{i-1}, \beta_{i+1}, \ldots, \beta_n)$ with $\beta_s = 1$. Due to the extremality of $\mathbf{a}$, all its components besides α_s are zeros. It follows that $\mathbf{a}' = \mathbf{y}$, which is a minimal one by assumption.

It remains to assume that $\alpha_s = 0$. Let $\mathbf{a}$ be a maximal zero for the function f_1. If $\mathbf{a}'$ is not a maximal zero for f, then there is $\mathbf{a}'' \succ \mathbf{a}'$ with $f(\mathbf{a}'') = 0$. Since $\mathbf{a}'' \succ \mathbf{a}'$ and $\alpha_i = 1$, the i-th component of $\mathbf{a}''$ is 1. By its removal, we obtain a zero of f_1 that is strictly above $\mathbf{a}$ in contradiction to the minimality of the latter.

Let $\mathbf{a}$ be a minimal one for the function f_1. If $\mathbf{a}'$ is not a minimal one for f, then there is $\mathbf{a}'' \prec \mathbf{a}'$ with $f(\mathbf{a}'') = 1$. The i-th component of $\mathbf{a}''$ must be 0, since otherwise by its removal we obtain a one for f_1 strictly below $\mathbf{a}$. Also, the s-th component of $\mathbf{a}''$ must be 0, since this component equals 0 in $\mathbf{a}$. But then $\mathbf{a}'' \preceq \overline{\mathbf{y}}$ with $f(\mathbf{a}'') = 1$ and $f(\overline{\mathbf{y}}) = 0$, a contradiction. □

Lemma 19. *Let $f = f(x_1, \ldots, x_n)$ be a positive non-canalyzing function such that for each variable x_i both restrictions $f_{|x_i=0}$ and $f_{|x_i=1}$ are canalyzing. Then the number of extremal points of f is at least $n + 2$.*

Proof. By Claim 17 we may assume that there is a minimal one $\mathbf{y}$ that contains 1's in exactly two coordinates, say i and s, such that $\overline{\mathbf{y}}$ is a maximal zero. Denote $f_0 = f_{|x_i=0}$ and $f_1 = f_{|x_i=1}$.

Let P, P_0, and P_1 be the sets of relevant variables of f, f_0, and f_1, respectively. By Claim 15, P is the set of all variables. Since any relevant variable of f is relevant for at least one of the functions f_0, f_1 and, by Claim 18 (a), x_s is a relevant variable of both of them, we have

$$n = |P| \leq |P_0 \cup P_1| + 1 = |P_0| + |P_1| - |P_0 \cap P_1| + 1 \leq |P_0| + |P_1|.$$

By Lemma 14, $r(f_0) \geq |P_0| + 1$, $r(f_1) \geq |P_1| + 1$. Finally, by Claim 18 (b) the number $r(f)$ of extremal points of f is at least $r(f_0) + r(f_1) \geq |P_0| + |P_1| + 2 \geq n + 2$. □

2.3.4 Non-canalyzing functions containing a non-canalyzing restriction

Due to Lemmas 14 and 19 it remains to show the bound for a positive non-canalyzing function $f = f(x_1, \ldots, x_n)$ such that for some $i \in [n]$ at least one of $f_0 = f_{|x_i=0}$ and $f_1 = f_{|x_i=1}$ is non-canalyzing. Let k be the number of relevant variables of f and let us prove that the number of extremal points of f is at least $k + 2$.

Consider two possible cases:

(a) x_i is a irrelevant variable of f;

(b) x_i is a relevant variable of f.

In case (a) the function $f_{|x_i=0} \equiv f_{|x_i=1}$ is non-canalyzing and has the same number of extremal points and the same number of relevant variables as f. By induction, the number of extremal points of f is at least $k + 2$.

Now let us consider case (b). Assume without loss of generality that $i = n$, and let $f_0 = f_{|x_n=0}$ and $f_1 = f_{|x_n=1}$. We assume that f_0 is non-canalyzing and prove that f has at least $k + 2$ extremal points, where k is the number of relevant variables of f. The case when f_0 is canalyzing, but f_1 is non-canalyzing is proved similarly.

Let us denote the number of relevant variables of f_0 by m. Clearly, $1 \leq m \leq k-1$. Exactly $k - 1 - m$ of k relevant variables of f are irrelevant for the function f_0. Note that these $k - 1 - m$ variables are necessarily relevant for the function f_1. By the induction hypothesis, the number $r(f_0)$ of extremal points of f_0 is at least $m + 2$.

We introduce the following notation:

C_0 – the set of maximal zeros of f corresponding to x_n;

P_0 – the set of all other maximal zeros of f, i.e., $P_0 = Z^f \setminus C_0$;

C_1 – the set of minimal ones of f corresponding to x_n;

P_1 – the set of all other minimal ones of f, i.e., $P_1 = U^f \setminus C_1$.

For a set $A \subseteq B^n$ we will denote by A^* the restriction of A to the first $n - 1$ coordinates, i.e., $A^* = \{(\alpha_1, \ldots, \alpha_{n-1}) \mid (\alpha_1, \ldots, \alpha_{n-1}, \alpha_n) \in A \text{ for some } \alpha_n \in \{0, 1\}\}$.

By definition, the number of extremal points of f is

$$r(f) = |C_0| + |P_1| + |C_1| + |P_0| = |C_0^*| + |P_1^*| + |C_1^*| + |P_0^*|. \tag{2.1}$$

We want to express $r(f)$ in terms of the number of extremal points of f_0 and f_1. For this we need several observations. First, we observe that if $(\alpha_1, \ldots, \alpha_{n-1}, \alpha_n)$ is an extremal point for f, the point $(\alpha_1, \ldots, \alpha_{n-1})$ is extremal for f_{α_n}. Furthermore, we have the following straightforward claim.

Claim 20. *P_1^* is the set of minimal ones of f_0 and P_0^* is the set of maximal zeros of f_1.*

In contrast to the minimal ones of f_0, the set of maximal zeros of f_0 in addition to the points in C_0^* may contain extra points, which we denote by N_0^*. In other words, $Z^{f_0} = C_0^* \cup N_0^*$. Similarly, besides C_1^*, the set of minimal ones of f_1 may contain additional points, which we denote by N_1^*. That is, $U^{f_1} = C_1^* \cup N_1^*$.

Claim 21. *The set N_0^* is a subset of the set P_0^* of maximal zeros of f_1. The set N_1^* is a subset of the set P_1^* of minimal ones of f_0.*

Proof. We will prove the first part of the statement, the second one is proved similarly. Suppose to the contrary that there exists a point $\mathbf{a} = (\alpha_1, \ldots, \alpha_{n-1}) \in N_0^* \setminus P_0^*$, which is a maximal zero for f_0, but is not a maximal zero for f_1. Notice that $f_1(\mathbf{a}) = 0$, as otherwise $(\alpha_1, \ldots, \alpha_{n-1}, 0)$ would be a maximal zero for f, which is not the case, since $\mathbf{a} \notin C_0^*$. Since $\mathbf{a}$ is not a maximal zero for f_1, there exists a maximal zero $\mathbf{b} \in B^{n-1}$ for f_1 such

that $\mathbf{a} \prec \mathbf{b}$. But then we have $f_0(\mathbf{b}) = 1$ and $f_1(\mathbf{b}) = 0$, which contradicts the positivity of function f. □

From Claim 20 we have $r(f_0) = |Z^{f_0} \cup U^{f_0}| = |C_0^*| + |N_0^*| + |P_1^*|$, which together with (2.1) and Claim 21 imply

$$\begin{aligned} r(f) &= |C_0^*| + |P_1^*| + |C_1^*| + |P_0^*| = |C_0^*| + |P_1^*| + |C_1^*| + |N_0^*| + |P_0^* \setminus N_0^*| \\ &= r(f_0) + |C_1^*| + |P_0^* \setminus N_0^*|. \end{aligned} \tag{2.2}$$

Using the induction hypothesis we conclude that $r(f) \geq m+2+|C_1^*|+|P_0^* \setminus N_0^*|$. To derive the desired bound $r(f) \geq k+2$, in the rest of this section we show that $C_1^* \cup P_0^* \setminus N_0^*$ contains at least $k - m$ points.

Claim 22. *Let x_i, $i \in [n-1]$, be a relevant variable for f_1, which is irrelevant for f_0. Then every maximal zero for f_1 corresponding to x_i belongs to $P_0^* \setminus N_0^*$ and every minimal one for f_1 corresponding to x_i belongs to C_1^*.*

Proof. Let $\mathbf{x} \in N_0^*$ and assume $(\mathbf{x})_i = 0$. Then by changing in $\mathbf{x}$ the i-th coordinate from 0 to 1 we obtain a point $\mathbf{x}'$ with $f_0(\mathbf{x}') = 1 \neq f_0(\mathbf{x})$, since $\mathbf{x}$ is a maximal zero for f_0. This contradicts the assumption that x_i is irrelevant for f_0. Therefore, $(\mathbf{x})_i = 1$ and hence no maximal zero for f_1 corresponding to x_i belongs to N_0^*, i.e. every maximal zero for f_1 corresponding to x_i belongs to $P_0^* \setminus N_0^*$

Similarly, one can show that no minimal one for f_1 corresponding to x_i belongs to N_1^*, i.e. every minimal one for f_1 corresponding to x_i belongs to C_1^*. □

Recall that there are exactly $k-1-m$ variables that are relevant for f_1 and irrelevant for f_0. Lemma 13 implies that there are at least $k - m$ extremal points for f_1 corresponding to these variables. By Claim 22, all these points belong to the set $C_1^* \cup P_0^* \setminus N_0^*$. This conclusion establishes the main result of this section.

Lemma 23. *Let $f = f(x_1, \ldots, x_n)$ be a positive non-canalyzing function with k relevant variables such that for some $i \in [n]$ at least one of the restrictions $f_0 = f_{|x_i=0}$ and $f_1 = f_{|x_i=1}$ is non-canalyzing. Then the number of extremal points of f is at least $k + 2$.*

2.4 Chow and read-once functions

An important class of Boolean functions was introduced in 1961 by Chow [13] and is known nowadays as *Chow functions*. This notion can be defined as follows.

Definition 2.4.1. The Chow parameters of a Boolean function $f(x_1, \ldots, x_n)$ are the $n + 1$ integers $(w_1(f), w_2(f), \ldots, w_n(f), w(f))$, where $w(f)$ is the number of true points of f

and $w_i(f)$ is the number of true points of f where x_i is also true. A Boolean function f is a Chow function if no other function has the same Chow parameters as f.

In this section, we look at the intersection of the classes of Chow and read-once functions and show that this is precisely the class of lro functions. Thus, our result generalizes a result from [20] showing that the class of lro functions is the intersection of the classes of read-once and threshold functions.

There are two read-once functions that play a crucial role in our characterization of read-once Chow functions:

$$g_1 = g_1(x, y, z, u) = (x \vee y) \wedge (z \vee u),$$

$$g_2 = g_2(x, y, z, u) = (x \wedge y) \vee (z \wedge u).$$

Lemma 24. *Functions g_1, g_2 and all the functions obtained from them by negating some variables are not Chow.*

Proof. Function g_1 is not Chow, because g_1 is different from $(x \vee z) \wedge (y \vee u)$ (e.g. they have different values at the point $x = 1$, $y = 0$, $z = 1$, $u = 0$), but both functions have the same Chow parameters $(6, 6, 6, 6, 9)$. In a similar way, one can show that neither g_2 nor any function obtained from g_1 or g_2 by negating some variables is Chow. □

The following lemma shows that the class of Chow functions is closed under taking restrictions.

Lemma 25. *If $f(x_1, \ldots, x_n)$ is a Chow function, then any restriction of f is also Chow.*

Proof. Suppose to the contrary that f has a restriction which is not a Chow function, namely,

$$g = g(x_{i_{k+1}}, \ldots, x_{i_n}) := f_{x_{i_1}=\alpha_1,\ldots,x_{i_k}=\alpha_k},$$

for some $i_1, \ldots, i_n \in [n]$, $\alpha_1, \ldots, \alpha_k \in \{0, 1\}$ and g is not a Chow function. Then there exists a function $g' = g'(x_{i_{k+1}}, \ldots, x_{i_n})$ with the same Chow parameters as g. We define function $f'(x_1, \ldots, x_n)$ as follows:

$$f'(x_1, \ldots, x_n) = \begin{cases} f(x_1, \ldots, x_n) & \text{if } (x_{i_1}, \ldots, x_{i_k}) \neq (\alpha_1, \ldots, \alpha_k), \\ g'(x_{k+1}, \ldots, x_n) & \text{if } (x_{i_1}, \ldots, x_{i_k}) = (\alpha_1, \ldots, \alpha_k). \end{cases}$$

Since $w(g) = w(g')$, we conclude that $w(f) = w(f')$. Similarly, for every $i \in \{i_{k+1}, \ldots, i_n\}$ the equality $w_i(g) = w_i(g')$ implies $w_i(f) = w_i(f')$. Consequently, f and f' have the same Chow parameters, which contradicts the fact that f is Chow. □

Lemma 26. *Any canalyzing read-once function f, which is not lro, has a non-constant non-canalyzing read-once function as a restriction.*

Proof. Let f be a minimum counterexample to the claim. Since f is canalyzing, there exists $\alpha, \beta \in \{0,1\}$ such that $f_{|x_i=\alpha} \equiv \beta$. We assume that $\alpha = \beta = 1$, i.e. $f_{|x_i=1} \equiv \mathbf{1}$, in which case $f = x_i \vee f_{|x_i=0}$ (the other cases are similar).

Clearly, $f_{|x_i=0}$ is read-once, since any restriction of a read-once function is read-once. Also, $f_{|x_i=0}$ is not lro, since otherwise f is lro, and hence $f_{|x_i=0}$ is not a constant function. Since f is a counterexample, $f_{|x_i=0}$ is canalyzing and has no non-constant non-canalyzing read-once restrictions. But then we have a contradiction to the minimality of f. □

Theorem 27. *For a read-once function f the following statements are equivalent:*

(1) *f is an lro function;*

(2) *f is a Chow function;*

(3) *every restriction of f is congruent neither to g_1 nor to g_2.*

Proof. It is known that all lro functions are threshold [20] and all threshold functions are Chow [13]. Therefore, (1) implies (2).

To prove that (2) implies (3), we observe that by Lemma 25 any restriction of f is Chow. This together with Lemma 24 imply the conclusion.

Finally, to prove that (3) implies (1), we show that if f is non-lro, then it has as a restriction a function congruent to g_1 or g_2. Without loss of generality we assume that f is positive, non-canalyzing, otherwise we would rename some variables and/or consider a non-constant non-canalyzing restriction of f which is guaranteed by Lemma 26.

Since f is a read-once function, there exist read-once functions f_1 and f_2 such that either $f = f_1 \wedge f_2$ or $f = f_1 \vee f_2$ and the sets of relevant variables of f_1 and f_2 are disjoint. We let $f = f_1 \vee f_2$, since the other case can be proved similarly. Let F_1 and F_2 be simplified read-once formulas of f_1 and f_2 respectively, in particular, the negation operation, if any, is only applied to individual variables. Suppose, one of the formulas F_1 and F_2, say F_1, does not contain a conjunction. Then for any relevant variable x_i of f_1 we have $f_{|x_i=1} \equiv \mathbf{1}$, which contradicts the assumption that f is non-canalyzing. Hence, both formulas F_1 and F_2 necessarily contain conjunctions. This means that there exist $i_1, \ldots, i_n \in [n]$, $\alpha_5, \ldots, \alpha_n \in \{0,1\}$ such that

$$f_{1|x_{i_5}=\alpha_5,\ldots,x_{i_k}=\alpha_k} = x_{i_1} \wedge x_{i_2}$$

and

$$f_{2|x_{i_{k+1}}=\alpha_{k+1},\ldots,x_{i_n}=\alpha_n} = x_{i_3} \wedge x_{i_4},$$

where $\{x_{i_5}, \ldots, x_{i_k}\}$ and $\{x_{i_{k+1}}, \ldots, x_{i_n}\}$ are the sets of relevant variables of the functions

f_1 and f_2, respectively. Consequently

$$\begin{aligned} f_{|x_{i_5}=\alpha_5,\ldots,x_{i_n}=\alpha_n} &= f_{1|x_{i_5}=\alpha_5,\ldots,x_{i_k}=\alpha_k} \vee f_{2|x_{i_{k+1}}=\alpha_{k+1},\ldots,x_{i_n}=\alpha_n} \\ &= (x_{i_1} \wedge x_{i_2}) \vee (x_{i_3} \wedge x_{i_4}). \end{aligned}$$

□

2.5 Minimal non-lro threshold functions

For $n \geq 3$, denote by g_n the function defined by its DNF

$$g_n(x_1, \ldots, x_n) = x_1x_2 \vee x_1x_3 \vee \ldots \vee x_1x_n \vee x_2 \ldots x_n.$$

It is well known that a positive function has a unique minimal CNF and DNF. Moreover, there is one-to-one correspondence between the clauses in the minimal DNF (resp. CNF) of a positive function and its minimal ones (resp. maximal zeros) (see, for instance, [15]). In the below lemma we will use this property of a positive function to retrieve its extremal points.

Lemma 28. *For any $n \geq 3$, the function g_n is positive, non-lro, and threshold, depending on all its variables, and the specification number of g_n is $2n$.*

Proof. Clearly, g_n is positive and depends on all its variables. Also, it is easy to verify that g_n is not canalyzing, and therefore g is non-lro.

Now, we claim that the minimal CNF of g_n is

$$(x_1 \vee x_2)(x_1 \vee x_3) \ldots (x_1 \vee x_n)(x_2 \vee x_3 \vee \cdots \vee x_n).$$

Indeed, the equivalence of the minimal DNF and the minimal CNF can be directly checked by expanding the latter and applying the absorption law:

$$\begin{aligned} &(x_1 \vee x_2)(x_1 \vee x_3) \ldots (x_1 \vee x_n)(x_2 \vee x_3 \vee \cdots \vee x_n) \\ &= (x_1 \vee x_2x_3 \ldots x_n)(x_2 \vee x_3 \vee \cdots \vee x_n) \\ &= x_1x_2 \vee x_1x_3 \vee \cdots \vee x_1x_n \vee x_2x_3 \ldots x_n. \end{aligned}$$

From the minimal DNF and the minimal CNF of g_n we retrieve the minimal ones

$$\begin{aligned} \mathbf{x}_1 &= (1, 1, 0, \ldots, 0), \\ \mathbf{x}_2 &= (1, 0, 1, \ldots, 0), \\ &\ldots\ldots\ldots\ldots\ldots \\ \mathbf{x}_{n-1} &= (1, 0, 0, \ldots, 1), \\ \mathbf{x}_n &= (0, 1, 1, \ldots, 1), \end{aligned}$$

and maximal zeros of g_n

$$\begin{aligned}\mathbf{y}_1 &= (0,0,1,\ldots,1),\\ \mathbf{y}_2 &= (0,1,0,\ldots,1),\\ &\ldots\ldots\ldots\ldots\ldots\\ \mathbf{y}_{n-1} &= (0,1,1,\ldots,0),\\ \mathbf{y}_n &= (1,0,0,\ldots,0),\end{aligned}$$

respectively (see Theorems 1.26, 1.27 in [15]).

It is easy to check that all minimal ones $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ satisfy the equation

$$(n-2)x_1 + x_2 + x_3 + \cdots + x_n = n-1,$$

and all maximal zeros $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ satisfy the equation

$$(n-2)x_1 + x_2 + x_3 + \cdots + x_n = n-2.$$

Hence $(n-2)x_1 + x_2 + x_3 + \cdots + x_n \geq n-1$ is a threshold inequality representing the function g_n.

It remains to prove that g_n has $2n$ essential points. First, the inequality $\sigma_{\mathfrak{T}_n}(g_n) \geq 2n$ follows from Claim 33 and the fact that g_n is a self-dual function (self-dual functions and the proof of Claim 33 are presented in Section 3.3). Second, since g_n is a positive threshold function depending on all variables, the set of its extremal points specifies g_n, and hence $\sigma_{\mathfrak{T}_n}(g_n) \leq 2n$. These two facts imply $\sigma_{\mathfrak{T}_n}(g_n) = 2n$. □

It is not difficult to see that g_n is a *minimal* threshold function which is not lro, i.e. any restriction of g_n is an lro function. Moreover, the same is true for any function congruent to g_n, since the negation of a variable or renaming of variables of a threshold function results in a threshold function. We denote the set of all functions congruent to g_n for all n by $\mathcal{G}$ and show in what follows that there are no other minimal threshold functions which are not lro.

Theorem 29. *A threshold function f is lro if and only if it does not contain any function from $\mathcal{G}$ as a restriction.*

Proof. Stetsenko proved in [46] that the set of all minimal not read-once functions consists of the functions congruent to one of the following:

$$\begin{aligned}&g_n(x_1,\ldots,x_n) = x_1(x_2 \vee \ldots \vee x_n) \vee x_2 \ldots x_n && (n \geq 3),\\ &h^1_n(x_1,\ldots,x_n) = x_1 \ldots x_n \vee \overline{x_1} \ldots \overline{x_n} && (n \geq 2),\\ &h^2_n(x_1,\ldots,x_n) = x_1(x_2 \vee x_3 \ldots x_n) \vee x_2\overline{x_3} \ldots \overline{x_n} && (n \geq 3),\\ &h^3(x_1,\ldots,x_5) = x_1(x_3x_4 \vee x_5) \vee x_2(x_3 \vee x_4x_5);\\ &h^4(x_1,\ldots,x_4) = x_1(x_2 \vee x_3) \vee x_3x_4.\end{aligned}$$

Let us show that all functions in this list, except g_n, are 2-summable, hence are not threshold.

- For the function h_n^1 we have:

$$h_n^1(1,0,\ldots,0) = h_n^1(0,1,\ldots,1) = 0,$$
$$h_n^1(0,0,\ldots,0) = h_n^1(1,1,\ldots,1) = 1$$

  and

$$(1,0,\ldots,0) + (0,1,\ldots,1) = (0,0,\ldots,0) + (1,1,\ldots,1).$$

- For the function h_n^2 we have:

$$h_n^2(1,0,0,\ldots,0) = h_n^2(0,1,1,\ldots,1) = 0,$$
$$h_n^2(0,1,0,\ldots,0) = h_n^2(1,0,1,\ldots,1) = 1$$

  and

$$(1,0,\ldots,0) + (0,1,\ldots,1) = (0,1,0,\ldots,0) + (1,0,1,\ldots,1).$$

- For the function h^3 we have:

$$h^3(0,0,1,1,1) = h^3(1,1,0,0,0) = 0,$$
$$h^3(0,1,1,0,0) = h^3(1,0,0,1,1) = 1$$

  and

$$(0,0,1,1,1) + (1,1,0,0,0) = (0,1,1,0,0) + (1,0,0,1,1).$$

- For h^4 we have:

$$h^4(1,0,0,1) = h^4(0,1,1,0) = 0,$$
$$h^4(1,1,0,0) = h^4(0,0,1,1) = 1$$

  and

$$(1,0,0,1) + (0,1,1,0) = (1,1,0,0) + (0,0,1,1).$$

Since the functions h_n^1, h_n^2, h^3, h^4 are not threshold, f does not contain as a restriction any function congruent to any of them. If, additionally, f contains no function from $\mathcal{G}$ as a restriction, then f is read-once and hence is lro. If f contains a function from $\mathcal{G}$ as a restriction, then f is not read-once and hence is not lro. □

2.6 Conclusion

In this chapter we proved a number of results related to the class of linear read-once functions. We showed that the class of linear read-once functions coincides with the subclass of positive Boolean functions depending on all variables with the minimum possible number of extremal points. Furthermore, we also proved that this class is the intersection of

the classes of read-once and Chow functions. Finally, we characterized the class of linear read-once functions by means of minimal forbidden subfunctions within the universe of read-once functions and the universe of threshold functions. These results witness the importance of the class of linear read-once functions as a subclass of the mentioned classes of Boolean functions.

Chapter 3

Boolean threshold functions with minimum specification number

3.1 Introduction

In Section 2.5 we characterized the class of linear read-once functions within the universe of threshold functions by the set $\mathcal{G}$ of minimal functions which are not linear read-once. We showed that all functions in $\mathcal{G}$ depending on n variables have specification number $2n$, which can be viewed as an argument supporting Conjecture 1. Nevertheless, in this chapter we disprove the conjecture by providing a counterexample (Section 3.2) and address the problem of characterizing the set $\mathcal{T}_n$ of threshold functions depending on n variables with the minimum specification number $n + 1$. Furthermore, we investigate the question of whether this set can by described recursively similarly to the class of linear read-once functions. In Section 3.3 we show that the specification number of self-dual functions of n variables is at least $2n$, and hence the set $\mathcal{T}_n$ does not contain any self-dual function. In Section 3.4 we introduce the operation of extension on a variable, which together with operations of elementary conjunction and disjunction with a new variable can be used to obtain all functions in $\mathcal{T}_n$ from the functions in $\mathcal{T}_{n-1}$ for every $n \leq 5$. Section 3.5 is devoted to another operation with similar properties, which applies to threshold functions with symmetric variables. In Section 3.6 and Appendix A we enumerate and analyze non-linear read-once functions in $\mathcal{T}_n$ for all $n \leq 6$.

3.2 Non-canalyzing threshold functions with minimum specification number

The following counterexample to Conjecture 1 is a generalized version of the counterexample provided in [38] and [39].

Theorem 30. *Let n and k be natural numbers such that $3 \leq k \leq n-1$ and let $f_{n,k}(x_1, \ldots, x_n)$ be a Boolean function defined by its DNF*

$$x_1x_2 \vee x_1x_3 \vee \cdots \vee x_1x_k \vee x_2x_3 \ldots x_n.$$

Then $f_{n,k}$ is a positive, non-lro, threshold function, depending on all its variables, and the specification number of $f_{n,k}$ is $n+1$.

Proof. Clearly, $f_{n,k}$ depends on all its variables, it is positive, not canalyzing, and therefore $f_{n,k}$ is non-lro. Let us show that $f_{n,k}$ is a threshold function. In the same way as in Lemma 28 we find the minimal ones

$$\begin{aligned}
\mathbf{x}_1 &= (1,1,0,\ldots,0,0,\ldots,0),\\
\mathbf{x}_2 &= (1,0,1,\ldots,0,0,\ldots,0),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\mathbf{x}_{k-1} &= (1,0,0,\ldots,1,0,\ldots,0),\\
\mathbf{x}_k &= (0,1,1,\ldots,1,1)
\end{aligned}$$

and the maximal zeros

$$\begin{aligned}
\mathbf{y}_1 &= (0,0,1,\ldots,1,1),\\
\mathbf{y}_2 &= (0,1,0,\ldots,1,1),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\mathbf{y}_{n-2} &= (0,1,1,\ldots,0,1),\\
\mathbf{y}_{n-1} &= (0,1,1,\ldots,1,0),\\
\mathbf{z} &= (z_1,\ldots,z_n), \text{where } z_i = 0 \text{ iff } i \in \{2,\ldots,k\}
\end{aligned}$$

respectively. We use these points to construct a threshold inequality for $f_{n,k}$. We distinguish two cases. First, if $k = n-1$ then it is easy to check that the inequality

$$(2n-5)x_1 + 2(x_2 + x_3 + \cdots + x_{n-1}) + x_n \geq 2n-3$$

holds in all minimal ones and does not hold in all minimal zeros of $f_{n,n-1}$, and hence it is a threshold inequality for the threshold function $f_{n,n-1}$. Similarly, if $k < n-1$, then the inequality

$$((k-1)(n-k+1)-1)x_1 + \sum_{i=2}^{k}(n-k+1)x_i + \sum_{i=k+1}^{n} x_i \geq k(n-k+1)-1 \quad (3.1)$$

is a threshold inequality for $f_{n,k}$.

It remains to show that $f_{n,k}$ has $n+1$ essential points. Since $f_{n,k}$ depends on all its variables, every essential point of $f_{n,k}$ is extremal. Therefore, since $f_{n,k}$ has $n+k$ extremal

points, it suffices to prove that $k-1$ of them are not essential. We will show that the points $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{k-1}$ are not essential. On the contrary, suppose that there exists a threshold function f' that differs from $f_{n,k}$ only in the point $\mathbf{y}_i$, $i \in [k-1]$, i.e. $f'(\mathbf{y}_i) = 1$ and $f'(\mathbf{x}) = f_{n,k}(\mathbf{x})$ for every $\mathbf{x} \neq \mathbf{y}_i$. From $\overline{\mathbf{y}_{n-1}} \preccurlyeq \mathbf{z}$ and $f_{n,k}(\mathbf{z}) = 0$ we conclude that $f_{n,k}(\overline{\mathbf{y}_{n-1}}) = 0$, and therefore $\mathbf{x}_i + \mathbf{y}_i = \mathbf{y}_{n-1} + \overline{\mathbf{y}_{n-1}}$ implies that f' is 2-summable, which contradicts our assumption. Hereby, $k-1$ of $n+k$ extremal points of $f_{n,k}$ are not essential, and the specification number of $f_{n,k}$ achieves its lower bound which is $n+1$. □

We observe that the functions described in the theorem contain, as restrictions, functions from the set $\mathcal{G}$, for instance,

$$f_{n,k|x_{k+1}=1,\ldots,x_n=1} = g_k.$$

Therefore we have the following

Corollary 31. *The set of threshold functions with minimum specification number is not closed under taking restrictions.*

This corollary also shows that specification number is not monotone with respect to restrictions, i.e. by restricting a function specification number can increase.

3.3 Self-dual threshold functions

Definition 3.3.1. The *dual* of a Boolean function f is the function f^d defined by formula

$$f^d(x_1, \ldots, x_n) = \overline{f}(\overline{x_1}, \ldots, \overline{x_n}).$$

A Boolean function f is called *self-dual* if $f^d = f$.

It is known that there are self-dual functions in the class of threshold functions, e.g., minimal non-lro threshold functions from Section 2.5. However, in this section we show that the class $\mathcal{T}_n$ does not contain self-dual functions.

Lemma 32. *Let $f(x_1, \ldots, x_n)$ be a positive self-dual threshold function and $i \in [n]$. Let also*

$$a_1x_1 + \cdots + a_{i-1}x_{i-1} + a_{i+1}x_{i+1} + \cdots + a_nx_n \geq a_0$$

be a threshold inequality of the restriction $f_{x_i=0}$. Then there exists a positive ϵ such that the following inequality is a threshold inequality for f:

$$a_1x_1 + \cdots + a_{i-1}x_{i-1} + (2a_0 - \sum_{j \in [n]\setminus\{i\}} a_j - \epsilon)x_i + a_{i+1}x_{i+1} + \cdots + a_nx_n \geq a_0.$$

Proof. Assume without loss of generality that $i = n$ and denote $f_0 = f_{|x_n=0}, f_1 = f_{|x_n=1}$. We will show that there exists $\epsilon > 0$ such that

$$a_1x_1 + \cdots + a_{n-1}x_{n-1} + (2a_0 - \sum_{j\in[n-1]} a_j - \epsilon)x_n \geq a_0 \tag{3.2}$$

is a threshold inequality for f. Consider the threshold inequality for f_0

$$a_1x_1 + \cdots + a_{n-1}x_{n-1} \geq a_0 \tag{3.3}$$

and denote $A = \sum_{j=1}^{n-1} a_j$. We claim that the following inequality holds in all true points of f_1 and only in them:

$$a_1x_1 + \cdots + a_{n-1}x_{n-1} > A - a_0. \tag{3.4}$$

Indeed,

$$f_1(\mathbf{x}) = 1 \Leftrightarrow f_0(\overline{\mathbf{x}}) = 0 \Leftrightarrow a_1 \cdot (\overline{\mathbf{x}})_1 + \cdots + a_{n-1} \cdot (\overline{\mathbf{x}})_{n-1} < a_0,$$

where

$$a_1 \cdot (\overline{\mathbf{x}})_1 + \cdots + a_{n-1} \cdot (\overline{\mathbf{x}})_{n-1} = A - (a_1 \cdot (\mathbf{x})_1 + \cdots + a_{n-1} \cdot (\mathbf{x})_{n-1}).$$

Since B^n is a discrete set of points, there exists a positive ϵ such that the following inequality is equivalent to inequality (3.4):

$$a_1x_1 + \cdots + a_{n-1}x_{n-1} \geq A - a_0 + \epsilon. \tag{3.5}$$

To complete the proof, we notice that inequality (3.2) is equal to (3.3) in the points with zero n-th coordinate and equal to (3.5) in all other points, and hence inequality (3.2) is a threshold inequality for f, as claimed. □

Claim 33. *Let $f(x_1, \ldots, x_n)$ be a self-dual threshold function. Then f has at least $2n$ essential points.*

Proof. Theorem 7 implies the statement for f with irrelevant variables, so we assume that f depends on all its variables. Without loss of generality, we further assume that f is a positive function and denote $f_0 = f_{|x_n=0}, f_1 = f_{|x_n=1}$. First, we will show that for any essential point $(\alpha_1, \ldots, \alpha_{n-1})$ of f_0 the point $(\alpha_1, \ldots, \alpha_{n-1}, 0)$ is essential for f. Let $\mathbf{x}$ be an essential one of f_0. From Theorem 10 it follows that there exists a threshold inequality $a_1x_1 + \cdots + a_{n-1}x_{n-1} \geq a_0$ for f_0 such that $a_1(\mathbf{x})_1 + \cdots + a_{n-1}(\mathbf{x})_{n-1} = a_0$. By Lemma 32 the inequality

$$a_1x_1 + \cdots + a_{n-1}x_{n-1} + (2a_0 - \sum_{j=1}^{n-1} a_j - \epsilon)x_n \geq a_0 \tag{3.6}$$

is a threshold inequality for f for some positive ϵ. Since we have equality in (3.6) in the point $\mathbf{x}' = ((\mathbf{x})_1, \ldots, (\mathbf{x})_{n-1}, 0)$, we conclude that $\mathbf{x}'$ is an essential one of f. Following the same arguments we obtain that for every essential zero $\mathbf{y}$ of f_0 the point $\mathbf{y}' = ((\mathbf{y})_1, \ldots, (\mathbf{y})_{n-1}, 0)$ is an essential zero for f. As f_0 is a function of $n-1$ variables and has at least n essential points, the function f has at least n essential points with zero n-th coordinate.

Finally, by symmetry of self-dual functions, $\mathbf{x}$ is an essential point of f if and only if $\overline{\mathbf{x}}$ is, hence f has the same number of essential points with one n-th coordinate as the number of essential points with zero n-th coordinate, and the statement follows. □

Corollary 34. *$\mathcal{T}_n$ does not contain self-dual functions.*

3.4 The extension of a threshold function on a variable

By definition any lro function of n variables for $n > 1$ can be obtained from an lro function of $n-1$ variables as the conjunction or disjunction of this function and a new variable, we will refer to this operation as *adding a variable*. In [6] it was shown that operation of adding a variable increases the specification number of a function by one. Formally, the following lemma was proved in [6]:

Lemma 35 ([6])**.** *Let $f(x_1, \ldots, x_n)$ be a threshold function depending on all its variables. Then the functions $f'(x_1, \ldots, x_n, y) = y \vee f$ and $f''(x_1, \ldots, x_n, y) = y \wedge f$ both have specification number $\sigma_{\mathfrak{T}_n}(f) + 1$.*

Since the class of lro functions can be constructed recursively by the operations of adding a variable starting from the constant functions, Lemma 35 implies that any lro function depending on all its variables has specification number one more than the number of variables. It is natural to ask whether the recursive definition of the class of lro functions can be generalized to the whole class $\mathcal{T}_n$. This section is devoted to some results in this direction.

Definition 3.4.1. Let $f(x_1, \ldots, x_n)$ be a positive Boolean function, $i \in [n]$, and let y be a new variable. The (x_i, y)-*extension* of f is the function

$$f^{(x_i,y)}(x_1, \ldots, x_n, y) = x_i(y \vee f_{|x_i=1}) \vee y f_{|x_i=0}.$$

We say that $f^{(x_i,y)}$ can be obtained from f by extension on the variable x_i. To illustrate the relation between adding a variable and extension on a variable operations we will make use of restriction graphs.

Definition 3.4.2. Let $f = f(x_1, \ldots, x_n)$ be a Boolean function and $S = \{x_{i_1}, \ldots, x_{i_k}\}$ be a set of variables of f. We say that a graph G is the S-*restriction graph* for f if its vertex

(a) The $\{y\}$-restriction graph for the function $f' = f \vee y$.

(b) The $\{y\}$-restriction graph for the function $f'' = f \wedge y$.

(c) The $\{x_i, y\}$-restriction graph for the function $f^{(x_i,y)}$. The dashed ellipses correspond to the $\{y\}$-restriction graphs for $f_{|x_i=1} \vee y$ and $f_{|x_i=0} \wedge y$.

Figure 3.1: The restriction graphs for the functions obtained from a given positive Boolean function $f = f(x_1, \ldots, x_n)$ by the operations of adding a variable and extension on the variable x_i for some $i \in [n]$.

set is the set of all restrictions of f to S and for any $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k \in \{0, 1\}$ two vertices $f_{|x_{i_1}=\alpha_1,\ldots,x_{i_k}=\alpha_k}$ and $f_{|x_{i_1}=\beta_1,\ldots,x_{i_k}=\beta_k}$ are connected by an edge if and only if vectors $(\alpha_1, \ldots, \alpha_k)$ and $(\beta_1, \ldots, \beta_k)$ differ in the exactly one coordinate.

If we look at the $\{y\}$-restriction graphs for the functions $f' = f \vee y$ (Fig. 3.1a) and $f'' = f \wedge y$ (Fig. 3.1b), we observe that both of them have a constant function as a vertex. It is due to the property of the functions obtained by the operation of adding a variable that one of the restrictions to the added variable is a constant function. Then we consider the $\{x_i, y\}$-restriction graph of $f^{(x_i,y)}$ (Fig. 3.1c) and notice that two of the four vertices of the graph are also constant functions. Moreover, the graph can be split into two subgraphs that are very similar to those on Figures 3.1a and 3.1b.

The operations of adding a variable and extension on a variable are in certain relation reflected in Fig. 3.1c and the following equations:

$$f^{(x_i,y)}_{|x_i=1} = f_{|x_i=1} \vee y,$$
$$f^{(x_i,y)}_{|x_i=0} = f_{|x_i=0} \wedge y.$$

In other words, a restriction of $f^{(x_i,y)}$ on a variable x_i is obtained from the corresponding restriction of the original function via operation of adding a variable.

Furthermore, below we establish more similarities between the two operations by showing that similarly to the operation of adding a variable the operation of extension on a variable applied to a threshold function results in a threshold function and, more importantly, also increases specification number by at most one. We start with a few preliminary statements.

The following lemma can be considered as a criteria for a function to be obtained from a given function by extension on a variable.

Lemma 36. *Let $f(x_1, \ldots, x_n)$ and $g(x_1, \ldots, x_n, x_{n+1})$ be Boolean functions, and let $i \in$*

$[n]$. *Then g is the (x_i, x_{n+1})-extension of f if and only if*

$$g(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n, \overline{\alpha_i}) = f(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n)$$

and

$$g(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n, \alpha_i) = \alpha_i$$

for any $\alpha_1, \ldots, \alpha_n \in \{0, 1\}$.

Proof. Let f' be the (x_i, x_{n+1})-extension of f, we will show that $f' \equiv g$. Indeed, considering the restrictions of the functions on the variables x_i and x_{n+1}, we notice that they are equal:

$$\begin{aligned} f'_{|x_i=0,x_{n+1}=0} &= 0 &&= g_{|x_i=0,x_{n+1}=0}, \\ f'_{|x_i=1,x_{n+1}=1} &= 1 &&= g_{|x_i=1,x_{n+1}=1}, \\ f'_{|x_i=1,x_{n+1}=0} &= f_{x_i=1} &&= g|_{x_i=1,x_{n+1}=0}, \\ f'_{|x_i=0,x_{n+1}=1} &= f_{x_i=0} &&= g_{|x_i=0,x_{n+1}=1}. \end{aligned}$$

□

Claim 37. *Let $f(x_1, \ldots, x_n)$ be a positive Boolean function. If there exist k ones $\mathbf{x}_1, \ldots, \mathbf{x}_k$ and k zeros $\mathbf{y}_1, \ldots, \mathbf{y}_k$ of f such that*

$$(a_1, \ldots, a_n) = \mathbf{x}_1 + \cdots + \mathbf{x}_k \preccurlyeq \mathbf{y}_1 + \cdots + \mathbf{y}_k = (b_1, \ldots, b_n),$$

then f is k-summable. If, in addition, $b_i = k$ or $a_i = 0$ for some $i \in [n]$ then $f_{|x_i=1}$ or $f_{|x_i=0}$ is k-summable respectively.

Proof. To prove the first part of the statement, we observe that if $(a_1, \ldots, a_n) = (b_1, \ldots, b_n)$ then $k \geq 2$ and f is k-summable by definition. Further, if $(a_1, \ldots, a_n) \prec (b_1, \ldots, b_n)$ we can switch some coordinates of the points $\mathbf{x}_1, \ldots, \mathbf{x}_k$ from zeros to ones to obtain k points $\mathbf{x}'_1, \ldots, \mathbf{x}'_k$ such that

$$\mathbf{x}'_1 + \cdots + \mathbf{x}'_k = (b_1, \ldots, b_n).$$

Since f is a positive function and $\mathbf{x}_j \preccurlyeq \mathbf{x}'_j$ for each $j \in [k]$ we have $f(\mathbf{x}'_1) = \cdots = f(\mathbf{x}'_k) = 1$, and hence $k \geq 2$ and f is k-summable.

Now, let $b_i = k$ for some $i \in [n]$, the case $a_i = 0$ can be proved similarly. Without loss of generality, we assume $i = n$, and hence $(\mathbf{y}_j)_n = 1$ for each $j \in [k]$. Therefore, if we consider the restriction $f_1 = f_{|x_n=1}$, then $f_1((\mathbf{y}_j)_1, \ldots, (\mathbf{y}_j)_{n-1}) = f(\mathbf{y}_j) = 0$. Next, since f is positive, for each $\mathbf{x}_j$ for $j \in [k]$ we have $f_1((\mathbf{x}_j)_1, \ldots, (\mathbf{x}_j)_{n-1}) = f(\mathbf{x}_j) = 1$. Therefore, f_1 is k-summable for the same reason as f is. □

Claim 38. *Let $f(x_1, \ldots, x_n)$ be a positive threshold function and $f(\mathbf{x}) = 0$ for some point $\mathbf{x}$. The point $\boldsymbol{x}$ is an inessential zero of f if and only if for some positive m and $k \geq m$*

there exist k not necessarily distinct zeros $\mathbf{z}_1, \ldots, \mathbf{z}_k$ and $k - m$ not necessarily distinct ones $\mathbf{z}_{k+1}, \ldots, \mathbf{z}_{2k-m}$ of f such that

$$\mathbf{z}_1 + \cdots + \mathbf{z}_k = \mathbf{z}_{k+1} + \cdots + \mathbf{z}_{2k-m} + m \cdot \mathbf{x} \tag{3.7}$$

and $\mathbf{x} \notin \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$.

Proof. Denote by g the Boolean function equal to f in all points except $\mathbf{x}$. First, assume equation (3.7) holds for some k zeros and $k-m$ ones of f, then g is k-summable, and hence $\mathbf{x}$ is inessential for f.

Now, assume that $\mathbf{x}$ is an inessential point of f, then g is k-summable for some $k \geq 2$, therefore there exist not necessarily distinct zeros $\mathbf{z}_1, \ldots \mathbf{z}_k$ and not necessarily distinct ones $\mathbf{z}_{k+1}, \ldots, \mathbf{z}_{2k}$ of g such that

$$\mathbf{z}_1 + \cdots + \mathbf{z}_k = \mathbf{z}_{k+1} + \cdots + \mathbf{z}_{2k}. \tag{3.8}$$

Consider two possible cases:

- $\mathbf{x} \notin \{\mathbf{z}_{k+1}, \ldots, \mathbf{z}_{2k}\}$. As f and g differ only in $\mathbf{x}$ and $\mathbf{x} \notin \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$, equation (3.8) shows that f is k-summable, a contradiction.
- $\mathbf{x} \in \{\mathbf{z}_{k+1}, \ldots, \mathbf{z}_{2k}\}$. Relation (3.8) implies (3.7) for some positive m.

□

The following claim is symmetric to Claim 38 and can be proved similarly.

Claim 39. *Let $f(x_1, \ldots, x_n)$ be a positive threshold function and $f(\mathbf{x}) = 1$ for some point $\mathbf{x}$. The point $\mathbf{x}$ is an inessential one of f if and only if for some positive m and $k \geq m$ there exist k not necessarily distinct ones $\mathbf{z}_1, \ldots, \mathbf{z}_k$ and $k - m$ not necessarily distinct zeros $\mathbf{z}_{k+1}, \ldots, \mathbf{z}_{2k-m}$ of f such that*

$$\mathbf{z}_1 + \cdots + \mathbf{z}_k = \mathbf{z}_{k+1} + \cdots + \mathbf{z}_{2k-m} + m \cdot \mathbf{x}$$

and $\mathbf{x} \notin \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$.

Claim 40. *Let $f(x_1, \ldots, x_n)$ be a positive threshold function and m and $k \geq m$ be some positive numbers. Let $\mathbf{z}_1, \ldots, \mathbf{z}_k$ be not necessarily distinct zeros and $\mathbf{z}_{k+1}, \ldots, \mathbf{z}_{2k-m}$ not necessarily distinct ones of f. If*

$$\mathbf{z}_{k+1} + \cdots + \mathbf{z}_{2k-m} + m \cdot \mathbf{x} \preccurlyeq \mathbf{z}_1 + \cdots + \mathbf{z}_k \tag{3.9}$$

for some point $\mathbf{x}$ such that $f(\mathbf{x}) = 0$ and $\mathbf{x} \notin \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$, then $\mathbf{x}$ is an inessential point of f.

Proof. As in the previous claim we denote by g the function equal to f in all points except $\mathbf{x}$. By Claim 37 the function g is k-summable, and hence $\mathbf{x}$ is not an essential point of f. $\square$

The main result of the section consists of two parts. First, in the following lemma we prove that the operation of extension on a variable applied to a threshold function results in a threshold function. Then, we show that the operation of extension on a variable increases the specification number of a function by at most one.

Lemma 41. *Let $f(x_1, \ldots, x_n)$ be a positive threshold function and let $n > 1$. The extension of f on a variable is a threshold function.*

Proof. Without loss of generality we prove the lemma for the extension on the variable x_1, i.e. we will show that the (x_1, x_{n+1})-extension of f is a threshold function. Denote $f_0 = f_{|x_1=0}, f_1 = f_{|x_1=1}$. Then, by definition,

$$f^{(x_1,x_{n+1})}(x_1, \ldots, x_n, x_{n+1}) = x_1(x_{n+1} \vee f_1) \vee x_{n+1} f_0.$$

To obtain a contradiction, assume $f^{(x_1,x_{n+1})}$ is not threshold and let k be the minimum number such that $f^{(x_1,x_{n+1})}$ is k-summable. Then there exist k not necessarily distinct zeros $\mathbf{y}_1, \ldots, \mathbf{y}_k$ and k not necessarily distinct ones $\mathbf{z}_1, \ldots, \mathbf{z}_k$ of $f^{(x_1,x_{n+1})}$ such that

$$\mathbf{y}_1 + \cdots + \mathbf{y}_k = \mathbf{z}_1 + \cdots + \mathbf{z}_k = (a_1, \ldots, a_n) \tag{3.10}$$

for some non-negative integer $a_1, \ldots, a_n$. Since

$$f^{(x_1,x_{n+1})}(0, \alpha_2, \ldots, \alpha_n, 0) = 0$$

and

$$f^{(x_1,x_{n+1})}(1, \alpha_2, \ldots, \alpha_n, 1) = 1$$

for any $\alpha_2, \ldots, \alpha_n \in \{0, 1\}$, we conclude that for every $i \in [k]$ at least one of $(\mathbf{y}_i)_1$ and $(\mathbf{y}_i)_{n+1}$ is equal to 0, and at least one of $(\mathbf{z}_i)_1$ and $(\mathbf{z}_i)_{n+1}$ is equal to 1. Therefore

$$k \geq \sum_{i=1}^{k} ((\mathbf{y}_i)_1 + (\mathbf{y}_i)_{n+1}) = a_1 + a_n = \sum_{i=1}^{k} ((\mathbf{z}_i)_1 + (\mathbf{z}_i)_{n+1}) \geq k,$$

and hence

$$(\mathbf{y}_i)_1 = \overline{(\mathbf{y}_i)_{n+1}}, (\mathbf{z}_i)_1 = \overline{(\mathbf{z}_i)_{n+1}} \tag{3.11}$$

for every $i \in [k]$.

Equations (3.11) and Lemma 36 imply

$$f((\mathbf{y}_i)_1, \ldots, (\mathbf{y}_i)_n) = f^{(x_1,x_{n+1})}(\mathbf{y}_i)$$

and

$$f((\mathbf{z}_i)_1, \ldots, (\mathbf{z}_i)_n) = f^{(x_1, x_{n+1})}(\mathbf{z}_i)$$

for every $i \in [k]$, which together with equation (3.10) show that f is k-summable, a contradiction.

□

Theorem 42. *Let $f(x_1, \ldots, x_n)$ be a positive function from $\mathcal{T}_n$ and $n > 1$. The extension of f on a variable belongs to $\mathcal{T}_{n+1}$.*

Proof. Without loss of generality we prove the statement for the (x_1, x_{n+1})-extension $f^{(x_1, x_{n+1})}$. Denote $f_0 = f_{|x_1=0}, f_1 = f_{|x_1=1}$, then

$$f^{(x_1, x_{n+1})}(x_1, \ldots, x_n, x_{n+1}) = x_1(x_{n+1} \vee f_1) \vee x_{n+1} f_0.$$

By Lemma 41, the function $f^{(x_1, x_{n+1})}$ is threshold, so we only need to show that its specification number is $n + 2$.

Assume first that at least one of f_0 and f_1 is a constant function. To show that $f^{(x_1, x_{n+1})} \in \mathcal{T}_{n+1}$, we consider four cases:

1. $f_1 \equiv 0$, then $f \equiv 0$, contradicting the assumption that $f \in \mathcal{T}_n$.
2. $f_1 \equiv 1$, then $f = x_1 \vee f_0$. As $f \in \mathcal{T}_n$ it depends on all its variables and Lemma 35 yields $\sigma_{\mathfrak{T}_{n-1}}(f_0) + 1 = \sigma_{\mathfrak{T}_n}(f)$ and, consequently, $\sigma_{\mathfrak{T}_{n-1}}(f_0) = n$. Moreover, $f^{(x_1, x_{n+1})} = x_1 \vee x_{n+1} f_0$, and, by Lemma 35, its specification number is $\sigma_{\mathfrak{T}_{n-1}}(f_0) + 2$, i.e. $f^{(x_1, x_{n+1})} \in \mathcal{T}_{n+1}$.
3. $f_0 \equiv 0$, this case can be handled in the same way as the previous one.
4. $f_0 \equiv 1$, then $f \equiv 1$, contradicting the assumption that $f \in \mathcal{T}_n$.

Assume now that both f_0 and f_1 are non-constant functions. Consider an inessential zero $\mathbf{y}$ of f, we will show that $\mathbf{y}' = ((\mathbf{y})_1, \ldots, (\mathbf{y})_n, \overline{(\mathbf{y})_1})$ is an inessential zero of $f^{(x_1, x_{n+1})}$ (the case of an inessential one of f can be proved similarly). By Lemma 36, we have $f^{(x_1, x_{n+1})}(\mathbf{y}') = f(\mathbf{y}) = 0$. Since $\mathbf{y}$ is an inessential zero of f, by Claim 38, for some m and k ($0 < m \leq k$) there exist k not necessarily distinct zeros $\mathbf{z}_1, \ldots, \mathbf{z}_k$ and $k - m$ not necessarily distinct ones $\mathbf{z}_{k+1}, \ldots, \mathbf{z}_{2k-m}$ of f such that

$$\mathbf{z}_1 + \cdots + \mathbf{z}_k = \mathbf{z}_{k+1} + \cdots + \mathbf{z}_{2k-m} + m \cdot \mathbf{y} = (a_1, \ldots, a_n)$$

for some non-negative integer $a_1, \ldots, a_n$. Consider the points $\mathbf{z}'_i = ((\mathbf{z}_i)_1, \ldots, (\mathbf{z}_i)_n, \overline{(\mathbf{z}_i)_1})$ for $i \in [2k - m]$. By Lemma 36, we have $f(\mathbf{z}_i) = f^{(x_1, x_{n+1})}(\mathbf{z}'_i)$ for each $i \in [2k - m]$. Moreover,

$$\mathbf{z}'_1 + \cdots + \mathbf{z}'_k = \mathbf{z}'_{k+1} + \cdots + \mathbf{z}'_{2k-m} + m \cdot \mathbf{y}' = (a_1, \ldots, a_n, k - a_1).$$

Hence $\mathbf{y}'$ is an inessential zero of $f^{(x_1,x_{n+1})}$ by Claim 38. This implies that the number of essential points of $f^{(x_1,x_{n+1})}$ with distinct first and last coordinates does not exceed the number of essential points of f, i.e., $n+1$.

We complete the proof if we show that $f^{(x_1,x_{n+1})}$ has at most one essential point with the same first and last coordinates. For this purpose, we first show that only two points with the given property are extremal points of $f^{(x_1,x_{n+1})}$, namely, $(0,1,\ldots,1,0)$ is a maximal zero of $f^{(x_1,x_{n+1})}$ and $(1,0,\ldots,0,1)$ is its minimal one. Indeed, the point $(0,1,\ldots,1,0)$ is a maximal zero of $f^{(x_1,x_{n+1})}$ because $f^{(x_1,x_{n+1})}(0,1,\ldots,1,0)=0$ and

$$f^{(x_1,x_{n+1})}(1,1,\ldots,1,0) = f_1(1,\ldots,1) = 1,$$
$$f^{(x_1,x_{n+1})}(0,1,\ldots,1,1) = f_0(1,\ldots,1) = 1.$$

Both equations hold as f_0 and f_1 are non-constant positive functions. The proof for the point $(1,0,\ldots,0,1)$ is similar. All other points with the equal first and last coordinates are either below the maximal zero $(0,1,\ldots,1,0)$ or above the minimal one $(1,0,\ldots,0,1)$, and therefore they are not extremal points.

It remains to show that one of the points $(1,0,\ldots,0,1)$ and $(0,1,\ldots,1,0)$ is inessential. We observe that $f^{(x_1,x_{n+1})}$ is not a self-dual function, otherwise f would be self-dual, contradicting Corollary 34. Since $f^{(x_1,x_{n+1})}$ is not self-dual, there exists a point $\mathbf{x}$ such that $f^{(x_1,x_{n+1})}(\mathbf{x}) = f^{(x_1,x_{n+1})}(\overline{\mathbf{x}})$. The equation

$$\mathbf{x} + \overline{\mathbf{x}} = (0,1,\ldots,1,0) + (1,0,\ldots,0,1) = (1,\ldots,1),$$

combined with Claims 38 and 39 leads to the conclusion that regardless of the value of $f^{(x_1,x_{n+1})}$ in the point $\mathbf{x}$, at least one of the points $(1,0,\ldots,0,1)$ and $(0,1,\ldots,1,0)$ is inessential for $f^{(x_1,x_{n+1})}$. Hence $f^{(x_1,x_{n+1})}$ has at most $n+2$ essential points and belongs to $\mathcal{T}_{n+1}$, as claimed. □

Example 43. *The function $f_{n,k}$ from Theorem 30 is obtained from the linear read-once function*

$$f(x_1,x_2,x_{k+1},\ldots,x_n) = x_2(x_1 \vee x_{k+1}\ldots x_n)$$

by applying $k-2$ times the operation of extension on the variable x_1:

$$\begin{aligned}
(x_2(x_1 \vee x_{k+1}\ldots x_n))^{(x_1,x_3)} &= x_1(x_2 \vee x_3) \vee x_2x_3x_{k+1}\ldots x_n,\\
(x_1(x_2 \vee x_3) \vee x_2x_3x_{k+1}\ldots x_n)^{(x_1,x_4)} &= x_1(x_2 \vee x_3 \vee x_4) \vee x_2x_3x_4x_{k+1}\ldots x_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
(x_1(x_2 \vee x_3 \vee \ldots \vee x_{k-1}) \vee x_2\ldots x_{k-1}x_{k+1}\ldots x_n)^{(x_1,x_k)} &\\
&= x_1(x_2 \vee \ldots \vee x_k) \vee x_2\ldots x_n = f_{n,k}.
\end{aligned}$$

3.5 Symmetric variables extension of a function from $\mathcal{T}_n$

The relations between the operations of adding a variable and extension on a variable motivated us to investigate further in this direction. Applying the operation of adding a variable to different restrictions of a threshold function, we have obtained one more interesting result.

Definition 3.5.1. Let $f(x_1, \ldots, x_n)$ be a Boolean function, and let $i, j \in [n]$ be two distinct indices. We say that the variables x_i and x_j are *symmetric* if

$$f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n).$$

Definition 3.5.2. Let $f(x_1, \ldots, x_n)$ be a positive Boolean function, let x_i and x_j be its symmetric variables for some distinct $i, j \in [n]$, and let y be a new variable. We define the (x_i, x_j, y)-*s-extension* (*symmetric variables extension*) of f as follows:

$$f^{(x_i,x_j,y)}(x_1, \ldots, x_n, y) = x_i x_j y f_{|x_i=1,x_j=1} \vee (x_i \vee x_j \vee y) f_{|x_i=1,x_j=0} \vee f_{|x_i=0,x_j=0}.$$

We say that $f^{(x_i,x_j,y)}$ can be obtained from f by the operation of symmetric variables extension on the variables x_i and x_j. We also observe that the variables x_i, x_j, y are symmetric in the function $f^{(x_i,x_j,y)}$, and hence the operation of symmetric variables extension increases the number of symmetric variables in the resulting function.

To illustrate the operation of symmetric variables extension we compare the $\{x_i, x_j\}$-restriction graph G_0 for the function f and the $\{x_i, x_j, y\}$-restriction graph G_1 for its symmetric extension $f^{(x_i,x_j,y)}$ (see Fig. 3.2). We observe that the vertices of G_0 and G_1 consist of three Boolean functions: $f_{00} = f_{|x_i=0,x_j=0}$, $f_{11} = f_{|x_i=1,x_j=1}$, and $f_{10} = f_{|x_i=1,x_j=0} = f_{|x_i=0,x_j=1}$ (the latter equality is due to the symmetry of x_i and x_j).

As x_i, x_j, and y are symmetric variables for $f^{(x_i,x_j,y)}$ we also have

$$f^{(x_i,x_j,y)}_{x_i=1,x_j=0,y=0} = f^{(x_i,x_j,y)}_{x_i=0,x_j=1,y=0} = f^{(x_i,x_j,y)}_{x_i=0,x_j=0,y=1}$$

and

$$f^{(x_i,x_j,y)}_{x_i=1,x_j=1,y=0} = f^{(x_i,x_j,y)}_{x_i=1,x_j=0,y=1} = f^{(x_i,x_j,y)}_{x_i=0,x_j=1,y=1}.$$

Moreover, from the the definition of symmetric extension it follows that all these restrictions are equal and coincide with f_{10} and that also $f^{(x_i,x_j,y)}_{x_i=0,x_j=0,y=0} = f_{00}$ and $f^{(x_i,x_j,y)}_{x_i=1,x_j=1,y=1} = f_{11}$. Hence, the set of vertices of G_1 correspond to the same Boolean functions f_{00}, f_{10}, f_{11} as the set of vertices of G_0.

If we consider the function obtained from f by k applications of the symmetric variables extension operation on the same pair of variables x_i, x_j, the corresponding $\{x_i, x_j, y_1, \ldots, y_k\}$-restriction graph (Fig. 3.2) will look similarly. Namely, the functions f_{11} and f_{00} will be in the top and the bottom of the graph respectively and all $k+1$ internal

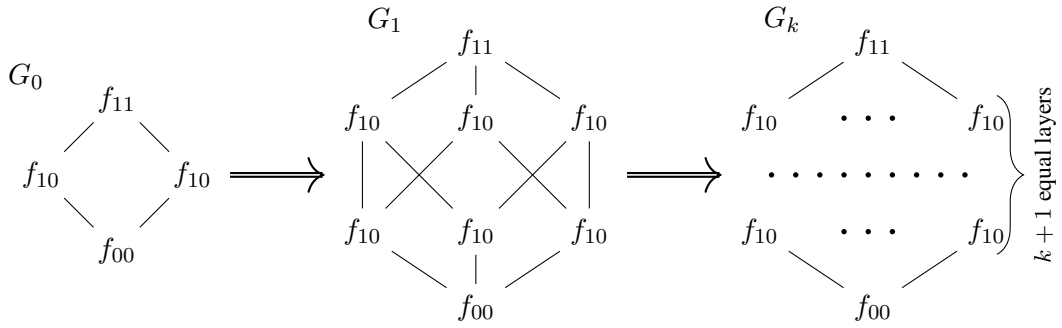


Figure 3.2: The graph G_0 is the $\{x_i, x_j\}$-restriction graph for a function f with symmetric variables x_i and x_j, where $f_{00} = f_{|x_i=0,x_j=0}$, $f_{10} = f_{|x_i=1,x_j=0} = f_{|x_i=0,x_j=1}$, $f_{11} = f_{|x_i=1,x_j=1}$. The graph G_1 is the $\{x_i, x_j, y\}$-restriction graph for the (x_i, x_j, y)-s-extension of f. The graph G_k is the $\{x_i, x_j, y_1, \ldots, y_k\}$-restriction graph for the function obtained from f by k symmetric variables extension operations, where $y_1, \ldots, y_k$ are new variables.

"layers" will be filled by the copies of the function f_{10}.

In contrast to the operation of extension on a variable, the operation of symmetric variables extension does not necessarily leave the function in the class of threshold functions.

Example 44. *Consider the threshold function* $f(x_1, \ldots, x_5) = x_1x_2 \vee (x_1 \vee x_2)x_3x_4x_5$ *and consider its* (x_1, x_2, x_6)*-s-extension:*

$$f^{(x_1,x_2,x_6)} = x_1x_2x_6 \vee (x_1 \vee x_2 \vee x_6)x_3x_4x_5.$$

The function $f^{(x_1,x_2,x_6)}$ *is* 2*-summable as*

$$f^{(x_1,x_2,x_6)}(1,1,0,0,0,1) = f^{(x_1,x_2,x_6)}(0,1,1,1,1,0) = 1,$$
$$f^{(x_1,x_2,x_6)}(0,1,1,0,1,1) = f^{(x_1,x_2,x_6)}(1,1,0,1,0,0) = 0,$$

and

$$(1,1,0,0,0,1) + (0,1,1,1,1,0) = (0,1,1,0,1,1) + (1,1,0,1,0,0),$$

hence $f^{(x_1,x_2,x_6)}$ *is not threshold.*

Although, the operation of symmetric variables extension does not always preserve the property of being threshold, when it does, it increases the number of essential points by at most one. To prove this fact, we first provide a few auxiliary statements.

Theorem 45 ([6])**.** *Let* $f(x_1, \ldots, x_n)$ *be a function in* $\mathfrak{T}_n$*. Then any specifying set of* f *contains* $n + 1$ *points in general position, and possibly some others.*

Theorem 45, in particular, implies that all essential points of a function in $\mathcal{T}_n$ are in general position.

Lemma 46. *Let $f(x_1, \ldots, x_n) \in \mathcal{T}_n$, and let x_i and x_j be distinct symmetric variables of f. Then the set of essential points of f has exactly 2 points with different i-th and j-th coordinates and these points only differ in these two coordinates.*

Proof. Without loss of generality we assume $i = 1$ and $j = 2$. We first observe that f has at least one essential point with distinct values in the first and second coordinates, otherwise the $n + 1$ essential points of f would not be in general position, contradicting Theorem 45.

Next, let $\mathbf{a} = (\alpha_1, \overline{\alpha_1}, \alpha_3, \ldots, \alpha_n)$ be an essential point of f for some $\alpha_1, \alpha_3, \ldots, \alpha_n \in \{0, 1\}$, then by symmetry of x_1 and x_2 the point $\mathbf{a}' = (\overline{\alpha_1}, \alpha_1, \alpha_3, \ldots, \alpha_n)$ is also essential for f. We claim that there is no other essential points of f with distinct first and second coordinates.

On the contrary, suppose that there exists an essential point $\mathbf{b} = (\beta_1, \overline{\beta_1}, \beta_3, \ldots, \beta_n)$ for some $\beta_1, \beta_3, \ldots, \beta_n \in \{0, 1\}$ such that $\mathbf{b} \notin \{\mathbf{a}, \mathbf{a}'\}$. Again, the point $\mathbf{b}' = (\overline{\beta_1}, \beta_1, \beta_3, \ldots, \beta_n)$ is also an essential point of f. However, depending on the values of α_1 and β_1 we have either

$$\mathbf{a} + \mathbf{b}' = \mathbf{a}' + \mathbf{b}$$

or

$$\mathbf{a} + \mathbf{b} = \mathbf{a}' + \mathbf{b}'.$$

In both cases the points $\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}'$ are not in general position, which contradicts Theorem 45. □

In the lemma below we will use the following notation: for a Boolean vector $\mathbf{a} = (\alpha_1, \ldots, \alpha_m)$ and a set of Boolean numbers $\beta_1, \ldots, \beta_n \in \{0, 1\}$ the $(m + n)$-dimensional vector $(\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n)$ will be denoted by $(\mathbf{a}, \beta_1, \ldots, \beta_n)$.

Lemma 47. *Let $f(x_1, \ldots, x_n)$ be a threshold function with symmetric variables x_i and x_j for some distinct $i, j \in [n]$ and let f' be its (x_i, x_j, x_{n+1})-s-extension. If f' is a threshold function, then for any inessential point $\mathbf{a} = (\alpha_1, \ldots, \alpha_n)$ of f and $\alpha_{n+1} \in \{0, 1\}$ the point $\mathbf{a}' = (\mathbf{a}, \alpha_{n+1})$ is inessential for f' if $f(\mathbf{a}) = f'(\mathbf{a}')$ and $(\alpha_{n-1}, \alpha_n, \alpha_{n+1}) \in \{(0, 0, 0), (1, 1, 1), (1, 0, 0), (0, 1, 0), (1, 0, 1), (0, 1, 1)\}$.*

Proof. Under the conditions stated above, suppose $f(\mathbf{a}) = f'(\mathbf{a}') = 0$, the case $f(\mathbf{a}) = f'(\mathbf{a}') = 1$ can be proved similarly. Since $\mathbf{a}$ is inessential for f, by Claim 38, there exist $k - m$ ones $\mathbf{x}_1, \ldots, \mathbf{x}_{k-m}$ of f and k zeros $\mathbf{y}_1, \ldots, \mathbf{y}_k$ of f for some $k \geq 2$ and $m \in [k - 1]$ such that

$$\mathbf{x}_1 + \cdots + \mathbf{x}_{k-m} + m \cdot \mathbf{a} = \mathbf{y}_1 + \cdots + \mathbf{y}_k = (b_1, \ldots, b_n),$$

where $b_1, \ldots, b_n \in [0, k]$. Denote the following sets of points

$$\begin{aligned}
&X = \{\mathbf{x}_1, \ldots, \mathbf{x}_{k-m}\},\\
&Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_k\},\\
&X_{11} = \{(x_1 \ldots, x_n) \in X | x_{n-1} = x_n = 1\},\\
&X_{00} = \{(x_1 \ldots, x_n) \in X | x_{n-1} = x_n = 0\},\\
&X_{10} = \{(x_1 \ldots, x_n) \in X | x_{n-1} \neq x_n\},\\
&Y_{11} = \{(x_1 \ldots, x_n) \in Y | x_{n-1} = x_n = 1\},\\
&Y_{00} = \{(x_1 \ldots, x_n) \in Y | x_{n-1} = x_n = 0\},\\
&Y_{10} = \{(x_1 \ldots, x_n) \in Y | x_{n-1} \neq x_n\}.
\end{aligned}$$

Without loss of generality we assume

$$\begin{aligned}
&X_{00} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{|X_{00}|}\},\\
&X_{10} = \{\mathbf{x}_{|X_{00}|+1}, \ldots, \mathbf{x}_{|X_{00}|+|X_{10}|}\},\\
&X_{11} = \{\mathbf{x}_{|X_{00}|+|X_{10}|+1}, \ldots, \mathbf{x}_{k-m}\},\\
&Y_{00} = \{\mathbf{y}_1, \ldots, \mathbf{y}_{|Y_{00}|}\},\\
&Y_{10} = \{\mathbf{y}_{|Y_{00}|+1}, \ldots, \mathbf{y}_{|Y_{00}|+|Y_{10}|}\},\\
&Y_{11} = \{\mathbf{y}_{|Y_{00}|+|Y_{10}|+1}, \ldots, \mathbf{y}_k\}.
\end{aligned}$$

Since

$$f'(x_1, \ldots, x_{n-2}, 1, 1, 1) = f(x_1, \ldots, x_{n-2}, 1, 1),$$

we have

$$f'(\mathbf{x}, 1) = 1 \text{ for all } \mathbf{x} \in X_{11}.$$

Similarly, we obtain

$$\begin{aligned}
&f'(\mathbf{x}, 0) = 1 \text{ for all } \mathbf{x} \in X_{00} \cup X_{10},\\
&f'(\mathbf{y}, 0) = 0 \text{ for all } \mathbf{y} \in Y_{00},\\
&f'(\mathbf{y}, 1) = 0 \text{ for all } \mathbf{y} \in Y_{10} \cup Y_{11}.
\end{aligned}$$

Using these $k - m$ ones of f' and k zeros of f' together with Claim 40 we will prove that $\mathbf{a}'$ is an inessential point of f'. Indeed, since

$$\begin{gathered}
(\mathbf{x}_1, 0) + \cdots + (\mathbf{x}_{|X_{00}|+|X_{10}|}, 0) + (\mathbf{x}_{|X_{00}|+|X_{10}|+1}, 1) + \cdots + (\mathbf{x}_{k-m}, 1) + m \cdot \mathbf{a}'\\
= (b_1, \ldots, b_n, |X_{11}| + \alpha_{n+1} m)
\end{gathered}$$

and

$$(\mathbf{y}_1, 0) + \cdots + (\mathbf{y}_{|Y_{00}|}, 0) + (\mathbf{y}_{|Y_{00}|+1}, 1) + \cdots + (\mathbf{y}_k, 1) = (b_1, \ldots, b_n, |Y_{10}| + |Y_{11}|),$$

the desired conclusion will follow from Claim 40 if we show

$$|X_{11}| + \alpha_{n+1}m \leq |Y_{10}| + |Y_{11}|. \tag{3.12}$$

To this end, we observe that all the points of Y_{00} have zero in the $(n-1)$-th and n-th coordinates, hence $|Y_{10}| + |Y_{11}| = |Y| - |Y_{00}| \geq \max(b_{n-1}, b_n)$. Therefore, (3.12) holds if the inequality below does:

$$|X_{11}| + \alpha_{n+1}m \leq \max(b_{n-1}, b_n). \tag{3.13}$$

To obtain the latter inequality we consider three cases separately:

1. $\alpha_{n+1} = 0$, i.e. $(\alpha_{n-1}, \alpha_n, \alpha_{n+1}) \in \{(0,0,0), (1,0,0), (0,1,0)\}$ and $\alpha_{n+1}m = 0$. Since X_{11} is the set of points where both the $(n-1)$-th and n-th coordinates are ones, we have $|X_{11}| \leq \min(b_{n-1}, b_n) \leq \max(b_{n-1}, b_n)$.

2. $\alpha_n = \alpha_{n+1} = 1$, i.e. $(\alpha_{n-1}, \alpha_n, \alpha_{n+1}) \in \{(0,1,1), (1,1,1)\}$ and $\alpha_{n+1}m = m$. Since the n-th coordinate of every point in $X_{11} \cup \{\mathbf{a}\}$ is one, we conclude $|X_{11}| + m \leq b_n \leq \max(b_{n-1}, b_n)$.

3. $(\alpha_{n-1}, \alpha_n, \alpha_{n+1}) = (1,0,1)$. The inequality $|X_{11}| + m \leq b_{n-1} \leq \max(b_{n-1}, b_n)$ can be shown by the same arguments as in the previous case.

Inequality (3.13) holds in all three cases, thus, by Claim 40, the point $\mathbf{a}'$ is inessential for f'. □

We are now in a position to prove the main result of the section.

Theorem 48. *Let $f(x_1, \ldots, x_n)$ be a positive function in $\mathcal{T}_n$, let x_i and x_j be distinct symmetric variables of f, and let f' be the (x_i, x_j, y)-s-extension of f. If f' is threshold then it belongs to $\mathcal{T}_{n+1}$.*

Proof. Without loss of generality we assume $i = n-1$, $j = n$. Let S be the set of essential points of f and let $S_{\alpha\beta} = \{(x_1, \ldots, x_n) \in S | x_{n-1} = \alpha, x_n = \beta\}$. Lemma 46 implies that the sets S_{01} and S_{10} both consist of one point, and hence

$$|S_{00}| + |S_{11}| = n + 1 - |S_{01}| - |S_{10}| = n - 1.$$

From this equation and Lemma 47 it follows that the set of essential points of f' has at most $n-1$ points with equal $(n-1)$-th, n-th and $(n+1)$-th coordinates.

Now we turn to the points with non-equal $(n-1)$-th, n-th and $(n+1)$-th coordinates. Let $(\alpha_1, \ldots, \alpha_{n-2}, 0, 1)$ be the only point in S_{01} for some $\alpha_1, \ldots, \alpha_{n-1} \in \{0,1\}$, then by Lemma 46 we have $S_{10} = \{(\alpha_1, \ldots, \alpha_{n-2}, 1, 0)\}$. Next, by Lemma 47 and pairwise

symmetry of the variables x_{n-1}, x_n and x_{n+1} the only points with non-equal $(n-1)$-th, n-th and $(n+1)$-th coordinates which can be essential for f' are the points

$$(\alpha_1, \ldots, \alpha_{n-2}, 1, 0, 0),$$
$$(\alpha_1, \ldots, \alpha_{n-2}, 0, 1, 0),$$
$$(\alpha_1, \ldots, \alpha_{n-2}, 0, 0, 1),$$
$$(\alpha_1, \ldots, \alpha_{n-2}, 1, 1, 0),$$
$$(\alpha_1, \ldots, \alpha_{n-2}, 1, 0, 1),$$
$$(\alpha_1, \ldots, \alpha_{n-2}, 0, 1, 1).$$

We claim that only three of the above points can be extremal for f', and therefore essential. Indeed, f' has the same value in all of them, and hence the first three points cannot be maximal zeros and the last three points cannot be minimal ones. Summarizing results, f' has at most $n-1$ essential points with equal $(n-1)$-th, n-th and $(n+1)$-th coordinates and at most three other essential points, giving at most $n+2$ essential points in total, and therefore $f' \in \mathcal{T}_{n+1}$. □

By definition, the operation of symmetric variables extension is only applicable to functions with symmetric variables. However, we believe that in the class $\mathcal{T}_n$ this property is not rare. We support this by the following several observations. First, it is easy to see, that after applying the conjunction operation to a Boolean function twice, the new variables of the resulting function are symmetric. The same is true for the disjunction operation. In fact, almost all positive linear read-once functions have symmetric variables. This statement is formulated in the following claim.

Claim 49. *Let $f(x_1, \ldots, x_n)$ be a positive linear read-once function without symmetric variables. Then f is either a constant or a single variable.*

Second, the operation of extension of a function applied twice on the same variable, produces the function with symmetric variables:

$$f(x_1, \ldots, x_n)^{(x_i, x_{n+1})(x_i, x_{n+2})} = x_i(x_{n+1} \vee x_{n+2} \vee f_{x_i=1}) \vee x_{n+1}x_{n+2}f_{x_i=0}.$$

Finally, the operation of symmetric variables extension also increases the number of symmetric variables.

3.6 Enumeration of $\mathcal{T}_n$ for $n \leq 6$

We finish the chapter by characterizing the functions from $\mathcal{T}_n$ for $n \leq 6$.

Claim 50. *All functions from $\mathcal{T}_n$ for $n \leq 3$ are linear read-once.*

Proof. The statement follows from Theorem 29 and the fact that the set $\mathcal{G}$ from Theorem 29 consists of the functions of at least 3 variables. □

Since a canalyzing function from $\mathcal{T}_n$ can be reduced to a function in $\mathcal{T}_{n-1}$, we will restrict our attention to non-canalyzing functions in $\mathcal{T}_n$. The following two claims were obtained by enumerating all the functions in $\mathcal{T}_n$ for the corresponding n. The code is written in Wolfram Language and provided in Appendix A.

Claim 51. *The function $f(x_1, x_2, x_3, x_4) = x_1(x_3 \vee x_4) \vee x_2x_3x_4$ is a unique non-congruent non-canalyzing function in $\mathcal{T}_4$ up to dualization.*

We observe that the function from Claim 51 is (x_1, x_4)-extension of the linear read-once function $f(x_1, x_2, x_3) = (x_1 \vee x_2)x_3$.

Claim 52. *There are exactly 7 non-congruent non-canalyzing functions in $\mathcal{T}_5$ up to dualization and each of them can be obtained from a function in $\mathcal{T}_4$ via the operation of extension on a variable. These 7 non-canalyzing functions are as follows:*

1. $f_{5,4} = x_1(\boldsymbol{x_2} \vee x_3 \vee x_4) \vee \boldsymbol{x_2}x_3x_4x_5 = f_{4,2}^{(x_1,x_2)}(x_1, x_3, x_4, x_5)$,
2. $f_{5,3} = x_1(\boldsymbol{x_2} \vee x_3) \vee \boldsymbol{x_2}x_3x_4x_5 = (x_3(x_1 \vee x_4x_5))^{(x_1,x_2)}$,
3. $x_1(\boldsymbol{x_2} \vee x_3 \vee x_4x_5) \vee \boldsymbol{x_2}x_3x_4 = f_{4,3}^{(x_1,x_2)}(x_3, x_1, x_4, x_5)$,
4. $x_1(\boldsymbol{x_2} \vee x_3 \vee x_4x_5) \vee \boldsymbol{x_2}x_3 = (x_1x_4x_5 \vee x_3)^{(x_1,x_2)}$,
5. $x_1(x_2 \vee x_3x_4x_5) \vee x_2x_4x_5 = x_2(\boldsymbol{x_1} \vee x_4x_5) \vee \boldsymbol{x_1}x_3x_4x_5 = ((x_2 \vee x_3)x_4x_5)^{(x_2,x_1)}$,
6. $x_1(x_2 \vee x_3x_5) \vee x_2x_5(x_3 \vee x_4) = x_2(\boldsymbol{x_1} \vee x_5(x_3 \vee x_4)) \vee \boldsymbol{x_1}x_3x_5 = (x_5(x_2x_4 \vee x_3))^{(x_2,x_1)}$,
7. $x_1(\boldsymbol{x_2} \vee x_3x_4 \vee x_3x_5 \vee x_4x_5) \vee \boldsymbol{x_2}x_3(x_4 \vee x_5) = f_{4,3}^{(x_1,x_2)}(x_3, x_4, x_5, x_1)$.

Claims 50, 51, and 52 imply the following

Claim 53. *For $n \leq 5$ the class $\mathcal{T}_n$ can be defined recursively using the operations of adding a variable, extension on a variable and symmetric variables extension.*

The following example demonstrates that Claim 53 does not hold for $\mathcal{T}_6$:

$$f(x_1, \ldots, x_6) = x_1(x_2 \vee x_3 \vee x_4 \vee x_5) \vee x_5(x_3 \vee x_4) \vee x_3x_4(x_6 \vee x_2).$$

Indeed, it is easy to check that the above function f cannot be obtained from any function in $\mathcal{T}_5$ using only the operations of adding a variable, extension on a variable and symmetric variables extension.

Appendix A contains all non-congruent non-canalyzing functions from $\mathcal{T}_6$ up to dualization.

3.7 Conclusion

In this chapter we showed the existence of threshold Boolean functions of n variables, which are not linear read-once and for which the specification number is at its lowest bound $n+1$. This leaves open the problem of characterizing the set of all functions with minimum specification number. We also made the first steps towards such a characterization via the set of operations which would transform an original function from $\mathcal{T}_n$ to a function from $\mathcal{T}_{n+1}$ and would be enough to describe all non-canalyzing functions in $\mathcal{T}_{n+1}$. Specifically, we introduced the operations of extension on a variable and symmetric variables extension which cover all functions from $\mathcal{T}_n$ for small n. Although general case remains open, the obtained operations can generate plenty of non-canalyzing functions from $\mathcal{T}_n$ for any n, which we believe will be useful for further progress in characterizing functions in $\mathcal{T}_n$ for arbitrary n.

Chapter 4

A characterization of 2-threshold functions via prime segments

4.1 Introduction

Threshold functions admit various representations and usually the choice of specific description depends on restrictions of a particular application. The most natural way of defining threshold functions is via threshold inequalities. However, for a given threshold function there are continuously many defining threshold inequalities, and given two linear inequalities it is not obvious whether they define the same threshold function or not.

Another way of describing threshold functions is via essential points. The set of essential points of a threshold function f together with the values of f in all these points uniquely identifies f in the class of threshold functions. Moreover, the set of essential points can be used to obtain a threshold inequality for f in linear time. To this end, it suffices to solve the system of linear inequalities where coefficients are the coordinates of essential points and the variables are the coefficients of a threshold inequality.

Example 54. *Consider a threshold function f over $\mathbb{Z}_3^2$ defined by its essential ones $(0,0),(2,2)$ and essential zeros $(0,1),(1,2)$ (see Fig. 4.1). Let $w_1x_1+w_2x_2 \geq t$ be a threshold inequality for f which we want to find. It holds in the true points of f and does not hold in the false points. Since the set of essential points specifies f, any solution of the following system of inequalities corresponds to a threshold inequality for f:*

$$\begin{cases} 0 \geq t, \\ 2w_1 + 2w_2 \geq t, \\ w_2 < t, \\ w_1 + 2w_2 < t. \end{cases} \tag{4.1}$$

It is easy to see that $w_1 = 2, w_2 = -2, t = -1$ is one of the solutions of (4.1) and

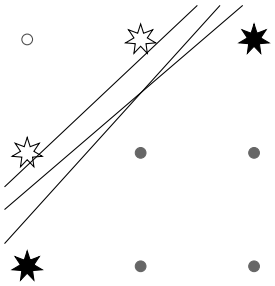


Figure 4.1: The black and white stars denote essential ones and zeros of f respectively. Any solution of (4.1) corresponds to a line which separates the sets of zeros and ones of f.

$2x_1 - 2x_2 \geq -1$ *is a threshold inequality for* f.

Although the set of essential points is a practical way of specifying threshold functions, for 2-threshold functions this approach does not work. Indeed, in contrast to threshold functions, the set of essential points of a 2-threshold function does not always specify it. Example 100 in Appendix B demonstrates this fact.

A useful characterization of two-dimensional threshold functions via oriented prime segments was provided in [34] to estimate their number. In this chapter we consider 2-threshold functions over a two-dimensional integer grid $\mathcal{G}_{m,n} = \{0, 1, \ldots, m-1\} \times \{0, 1, \ldots, n-1\}$ for natural m and n. We provide a characterization for 2-threshold functions over $\mathcal{G}_{m,n}$ by establishing a bijection between almost all pairs of oriented prime segments with certain properties and almost all 2-threshold functions.

The organization of the chapter is as follows. All preliminary information related to the chapter, including definitions and notation, can be found in Section 4.2. In Section 4.3 we describe and adapt to our purposes the bijection between oriented prime segments and non-constant threshold functions from [34]. In Section 4.4 we introduce special pairs of oriented prime segments, which we call *proper* pairs, and establish one-to-one correspondence between almost all proper pairs and almost all 2-threshold functions.

4.2 Preliminaries

In this and the following chapters we denote points on the plane by capital letters A, B, C, etc. For two sets of points S_1, S_2 we denote by $d(S_1, S_2)$ the (Euclidean) distance between the sets, that is, the minimum distance between two points $A \in S_1$ and $B \in S_2$. When a set consists of a single point we omit $\{\}$ and write simply $d(A, S_2)$ or $d(A, B)$ to denote the distance between point A and set S_2 or the distance between the points A and B, respectively. For two distinct points A, B we denote by $\ell(AB)$ the line which passes through these points. Furthermore, for a convex polygon $\mathcal{P}$ we denote by Area$(\mathcal{P})$ the area of $\mathcal{P}$.

A point $A = (x, y)$ is *integer*, if both of its coordinates x and y are integers. Two points A, B are called *adjacent* if they are integers and there is no other integer points on

AB. A segment with adjacent endpoints is called *prime*.

The convex hull of a set of points $X \subseteq \mathbb{R}^d$ is denoted by $\text{Conv}(X)$. We say that X is *in convex position* if all elements of X are vertices of its convex hull, i.e. $X = \text{Vert}(\text{Conv}(X))$. For a function $f : \mathbb{Z}_n^d \to \{0, 1\}$ we denote by $M_0(f)$ and $M_1(f)$ the sets of zeros and ones of f respectively. We also denote by $P(f)$ the convex hull of $M_1(f)$, that is $P(f) = \text{Conv}(M_1(f))$.

4.2.1 Segments, triangles, quadrilaterals and their orientation

We often denote a *convex* polygon by a sequence of its vertices in either clockwise or counterclockwise order. For example, by AB, ABC, and $ABCD$ we denote, respectively, the segment with endpoints A, B, the triangle with vertices A, B, C, and the convex quadrilateral with vertices A, B, C, D and edges AB, BC, CD, DA. When the order of vertices is important, we call the polygon or segment *oriented* and add an arrow in the notation, that is, $\overrightarrow{AB}$, $\overrightarrow{ABC}$, $\overrightarrow{ABCD}$ denote the oriented segment, the oriented triangle, and the oriented convex quadrilateral, respectively.

Let $A = (a_1, a_2), B = (b_1, b_2), C = (c_1, c_2)$ be distinct points on the plane. It is a basic fact that A, B, C are collinear if and only if $\Delta = 0$, where

$$\Delta = \begin{vmatrix} a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \\ c_1 & c_2 & 1 \end{vmatrix}.$$

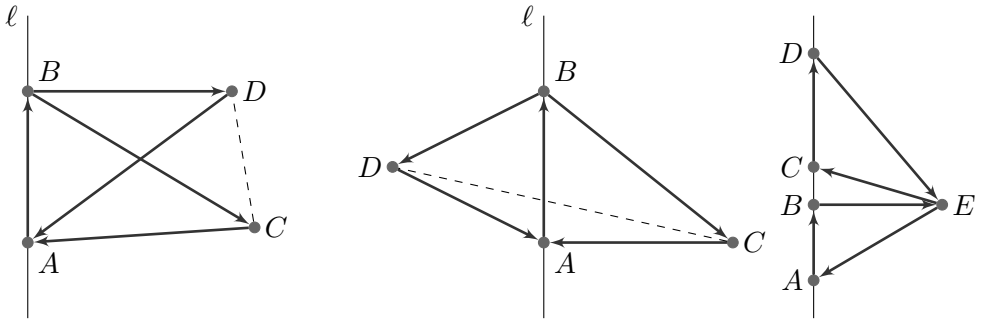
The oriented triangle $\overrightarrow{ABC}$ is called *clockwise* if $\Delta < 0$ and *counterclockwise* if $\Delta > 0$. Geometrically, an oriented triangle $\overrightarrow{ABC}$ is clockwise (resp. counterclockwise) if its vertices A, B, C, in order, rotate clockwise (resp. counterclockwise) around the triangle's center. Some properties of oriented triangles easily follow from the definition:

Claim 55. *Let ℓ be a line and let A, B be two distinct points on ℓ. Then for any two points $C, D \notin \ell$ the orientations of the triangles $\overrightarrow{ABC}$ and $\overrightarrow{ABD}$ are the same if and only if $\ell \cap CD = \emptyset$ (see Fig. 4.2a and 4.2b).*

Claim 56. *Let $\overrightarrow{AB}, \overrightarrow{CD}$ be two collinear segments with the same orientation. Then for any $E \notin \ell(AB)$ the triangles $\overrightarrow{ABE}$ and $\overrightarrow{CDE}$ have the same orientation (see Fig. 4.2c).*

Claim 57. *Let A, B, C, D be four distinct points such that $\overrightarrow{ABD}$, $\overrightarrow{BCD}$, $\overrightarrow{CAD}$ are clockwise (resp. counterclockwise) triangles. Then $\overrightarrow{ABC}$ is a clockwise (resp. counterclockwise) triangle.*

Proof. We will prove the statement for clockwise triangles, the counterclockwise case is symmetric. Denote $\mathcal{P} = \text{Conv}(\{A, B, C, D\})$. First, we show that D is not a vertex of $\mathcal{P}$. Suppose, to the contrary, that D is a vertex of $\mathcal{P}$, then two of the segments CD, BD, AD are



(a) CD does not intersect ℓ and $\overrightarrow{ABC}$ and $\overrightarrow{ABD}$ have the same orientation.

(b) CD intersects ℓ and $\overrightarrow{ABC}$ and $\overrightarrow{ABD}$ have the opposite orientations.

(c) $\overrightarrow{ABC}$ and $\overrightarrow{CDE}$ have the same orientation.

Figure 4.2: The orientation of the triangles depending on the positions of points

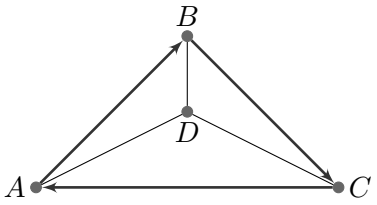


Figure 4.3: $\overrightarrow{ABC}$ has the same orientation as $\overrightarrow{ABD}$, $\overrightarrow{BCD}$, and $\overrightarrow{CAD}$.

edges of $\mathcal{P}$. The triangle $\overrightarrow{CAD}$ is clockwise, hence the triangle $\overrightarrow{CDA}$ is counterclockwise and the points A and B are separated by $\ell(CD)$, and therefore CD is not an edge of $\mathcal{P}$. Similarly, the opposite orientations of the triangles $\overrightarrow{ABD}$ and $\overrightarrow{BDC}$ imply that BD is not an edge of $\mathcal{P}$. The above contradicts the assumption that two of the segments CD, BD, AD are edges of $\mathcal{P}$, and therefore D is not a vertex of $\mathcal{P}$ and $\mathcal{P}$ is the triangle with vertices A, B, C. Finally, since D is an interior point of $\mathcal{P}$, the points C and D lie on the same side from $\ell(AB)$, hence the triangles $\overrightarrow{ABD}$ and $\overrightarrow{ABC}$ have the same orientation, i.e. $\overrightarrow{ABC}$ is clockwise, as required (see Fig. 4.3). □

It is clear, that for a given convex oriented quadrilateral $\overrightarrow{ABCD}$ the orientation of the triangles $\overrightarrow{ABC}$, $\overrightarrow{BCD}$, $\overrightarrow{CDA}$, and $\overrightarrow{DAB}$ is the same and determines the orientation of $\overrightarrow{ABCD}$. Moreover, the opposite is also true.

Claim 58. *Let $\overrightarrow{ABC}$, $\overrightarrow{BCD}$, $\overrightarrow{CDA}$, $\overrightarrow{DAB}$ be clockwise (resp. counterclockwise) triangles. Then* Conv$(\{A, B, C, D\})$ *is a quadrilateral with edges AB, BC, CD, and DA and the orientation of $\overrightarrow{ABCD}$ is clockwise (resp. counterclockwise).*

Proof. Clearly, A, B, C, and D are pairwise distinct points. Let $\mathcal{P} = \text{Conv}(\{A, B, C, D\})$. Since $\overrightarrow{ABC}$ and $\overrightarrow{DAB}$ are triangles with the same orientation we conclude that C and D lie on the same side of $\ell(AB)$, and therefore $\ell(AB)$ is a tangent to $\mathcal{P}$ and AB is an edge of $\mathcal{P}$. By similar arguments each of the segments BC, CD, and DA is an edge of $\mathcal{P}$, hence $\mathcal{P}$

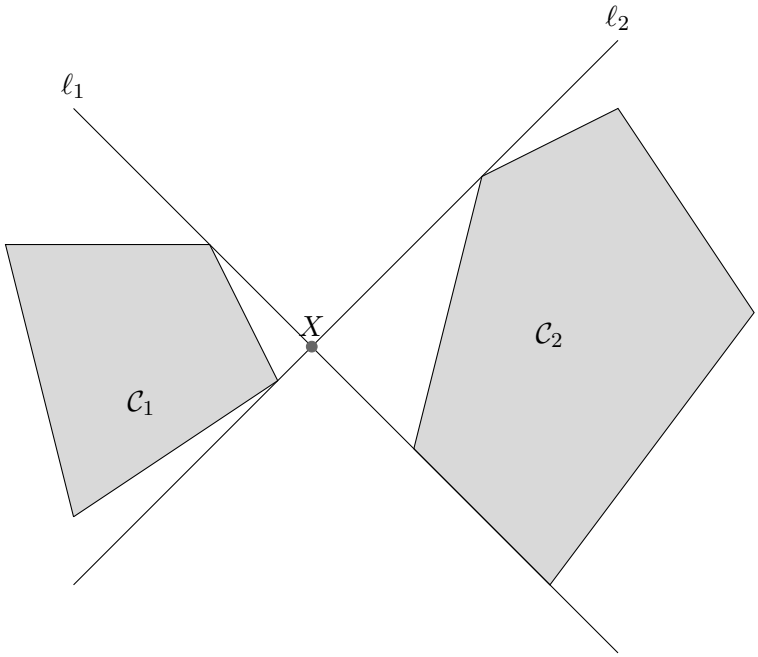


Figure 4.4: ℓ_1 is a right tangent from X to $\mathcal{C}_1$ and to $\mathcal{C}_2$, and a right inner common tangent for $\mathcal{C}_1$ and $\mathcal{C}_2$. ℓ_2 is a left tangent from X to $\mathcal{C}_1$ and to $\mathcal{C}_2$, and a left inner common tangent for $\mathcal{C}_1$ and $\mathcal{C}_2$.

is a quadrilateral. Finally, the orientation of the triangles implies that $\overrightarrow{ABCD}$ has the same orientation as the orientation of the triangles. □

4.2.2 Convex sets and their tangents

Let $\mathcal{C}$ be a convex set. A convex polygon $\mathcal{P}$ is called *circumscribed* about $\mathcal{C}$ if for every edge AB of $\mathcal{P}$ the line $\ell(AB)$ is a tangent to $\mathcal{C}$ and $AB \cap \mathcal{C} \neq \emptyset$.

Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two disjoint convex sets. A line ℓ is called an *inner common tangent* to $\mathcal{C}_1$ and $\mathcal{C}_2$ if it is a tangent to both of them, and $\mathcal{C}_1$ and $\mathcal{C}_2$ are separated by ℓ.

Let ℓ be a tangent to a convex set $\mathcal{C}$, and let X be a point in $\ell \setminus \mathcal{C}$. Then ℓ is called a *right* (resp. *left*) *tangent* from X to $\mathcal{C}$ if for any points $Y \in \mathcal{C} \cap \ell$ and $Z \in \mathcal{C} \setminus \ell$ the triangle $\overrightarrow{XYZ}$ is counterclockwise (resp. clockwise). The following claim is a simple consequence of the above definition.

Claim 59. *Let ℓ be a right (resp. left) tangent from a point X to a convex set $\mathcal{C}$, and let $Y \in \ell$. Then ℓ is a right (resp. left) tangent from Y to $\mathcal{C}$ if and only if $XY \cap \mathcal{C} = \emptyset$.*

Let ℓ be an inner common tangent to two disjoint convex sets $\mathcal{C}_1$ and $\mathcal{C}_2$, and let A, B be two points such that $A \in \mathcal{C}_1 \cap \ell$ and $B \in \mathcal{C}_2 \cap \ell$. Then ℓ is called a *right* (resp. *left*) *inner common tangent* to $\mathcal{C}_1$ and $\mathcal{C}_2$ if ℓ is a right (resp. left) tangent from A to $\mathcal{C}_2$, and a right (resp. left) tangent from B to $\mathcal{C}_1$ (see Fig. 4.4). It is easy to see that any pair of disjoint convex sets has exactly one right and exactly one left inner common tangent.

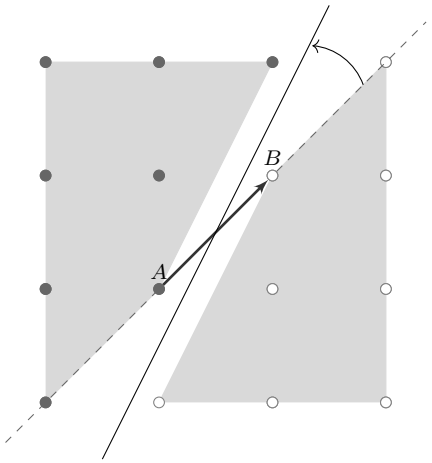


Figure 4.5: $\overrightarrow{AB}$ defines the threshold function $f_{\overrightarrow{AB}}$ where Conv$(M_1(f_{\overrightarrow{AB}}))$ and Conv$(M_0(f_{\overrightarrow{AB}}))$ are the left and right grey regions respectively. The dashed line is the left inner common tangent to Conv$(M_0(f_{\overrightarrow{AB}}))$ and Conv$(M_1(f_{\overrightarrow{AB}}))$. The solid line is a separating line for $f_{\overrightarrow{AB}}$.

4.3 Oriented prime segments and threshold functions

Definition 4.3.1. Let A and B be two adjacent points in $\mathcal{G}_{m,n}$. We say that $\overrightarrow{AB}$ *defines* a function $f : \mathcal{G}_{m,n} \to \{0, 1\}$ if:

1. $f(A) = 1, f(B) = 0$;
2. for any $X \in \mathcal{G}_{m,n} \cap \ell(AB)$ we have $f(X) = 1$ if and only if $d(A, X) < d(B, X)$;
3. for any $X \in \mathcal{G}_{m,n} \setminus \ell(AB)$ we have $f(X) = 1$ if and only if $\overrightarrow{ABX}$ is a counterclockwise triangle.

The function defined by $\overrightarrow{AB}$ will be denoted as $f_{\overrightarrow{AB}}$.

The following statement is an immediate consequence of Definition 4.3.1.

Claim 60. *Let $\overrightarrow{AB}$ be a prime segment in $\mathcal{G}_{m,n}$ and let $f = f_{\overrightarrow{AB}}$ be the function over $\mathcal{G}_{m,n}$ defined by $\overrightarrow{AB}$. Then for any $C \in \ell(AB) \cap \mathcal{G}_{m,n}$ we have either $f(C) = 1$ and $A \in BC$ or $f(C) = 0$ and $B \in AC$.*

In [34] authors, in different terms, showed that a function $f_{\overrightarrow{AB}}$ defined by an oriented prime segment $\overrightarrow{AB}$ is threshold and the line $\ell(AB)$ is an inner common tangent to the convex hulls of the sets of ones and zeros of f. For the convenience, the following theorem partly repeats the result from [34], thus adapting it to our purposes and making our exposition self-contained.

Theorem 61. *Let A and B be two adjacent points in $\mathcal{G}_{m,n}$ and let $f = f_{\overrightarrow{AB}}$. Then*

(1) f is a threshold function;

(2) A and B are essential points of f;

(3) $\ell(AB)$ is the left inner common tangent to Conv$(M_1(f))$ *and* Conv$(M_0(f))$.

Proof. First we prove (1). Indeed, if we consider the line $\ell(AB)$ and turn it counterclockwise slightly around the middle of the segment AB to not intersect any integer points then we obtain a separating line for f, hence f is a threshold function (see Fig. 4.5).

Let us now prove (2). Consider the line $\ell(AB)$ and turn it counterclockwise slightly around the point A to not intersect any integer points except A. The obtained line separates $M_1(f) \setminus \{A\}$ and $M_0(f) \cup \{A\}$, and witnesses that the function that differs from f in the unique point A is threshold. Therefore, the point A is essential for f. Similarly, one can show that B is also essential for f.

Now we prove (3). First, it is easy to see that $\ell(AB)$ is a tangent to both Conv$(M_1(f))$ and Conv$(M_0(f))$. Furthermore, since Conv$(M_1(f))$ and Conv$(M_0(f))$ are separated by $\ell(AB)$, we conclude that $\ell(AB)$ is an inner common tangent for Conv$(M_1(f))$ and Conv$(M_0(f))$. Now, by Definition 4.3.1, for any $X \in M_1(f) \setminus \ell(AB)$ the triangle $\overrightarrow{BAX}$ is clockwise, and for any $X \in M_2(f) \setminus \ell(AB)$ the triangle $\overrightarrow{ABX}$ is clockwise. Hence, $\ell(AB)$ is a left tangent from B to Conv$(M_1(f))$ and from A to Conv$(M_2(f))$, i.e. $\ell(AB)$ is a left inner common tangent for Conv$(M_1(f))$ and Conv$(M_0(f))$. □

In [34] authors also proved a bijection between oriented prime segments and non-constant threshold functions:

Theorem 62 ([34])**.** *There is one-to-one correspondence between oriented prime segments in $\mathcal{G}_{m,n}$ and non-constant threshold functions over $\mathcal{G}_{m,n}$.*

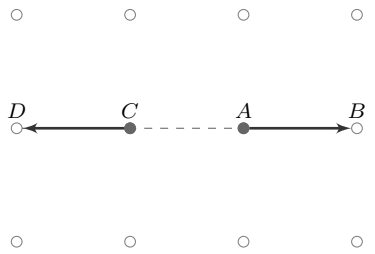
Corollary 63. *Let f be a non-constant threshold function over $\mathcal{G}_{m,n}$. Then there exists a unique prime segment AB with $A, B \in \mathcal{G}_{m,n}$ such that $f = f_{\overrightarrow{AB}}$.*

4.4 Pairs of oriented prime segments and 2-threshold functions

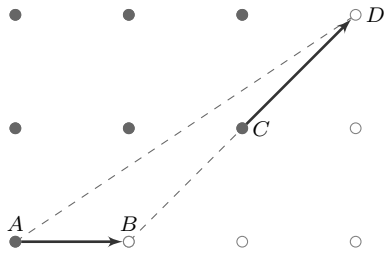
Since a 2-threshold function is the conjunction of two threshold functions, the defining threshold functions via oriented prime segments can be naturally extended to 2-threshold functions.

Definition 4.4.1. We say that a pair of oriented prime segments $\overrightarrow{AB}, \overrightarrow{CD}$ in $\mathcal{G}_{m,n}$ *defines* a 2-threshold function f over $\mathcal{G}_{m,n}$ if

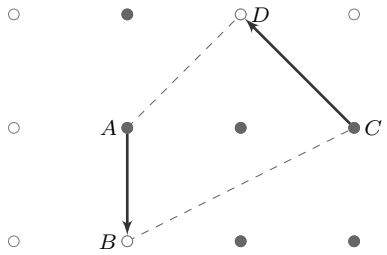
$$f = f_{\overrightarrow{AB}} \wedge f_{\overrightarrow{CD}}.$$



(a) $AC \subset BD$



(b) $C \in BD$ and $\overrightarrow{ABD}$ is a counterclockwise triangle



(c) $\overrightarrow{ABCD}$ is a convex counterclockwise quadrilateral.

Figure 4.6: Black points are the true points of $f = f_{\overrightarrow{AB}} \wedge f_{\overrightarrow{CD}}$ where $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is a proper pair of segments.

A 2-threshold function can be expressed as the conjunction of different pairs of threshold functions, therefore there is no bijection between pairs of oriented prime segments and non-constant 2-threshold functions. However, we may impose some restrictions on the pairs of oriented prime segments to exclude redundant pairs of segments defining the same function.

Definition 4.4.2. We say that a pair of oriented segments $\overrightarrow{AB}, \overrightarrow{CD}$ is *proper* if the segments are prime and

$$f_{\overrightarrow{CD}}(A) = f_{\overrightarrow{CD}}(B) = f_{\overrightarrow{AB}}(C) = f_{\overrightarrow{AB}}(D) = 1.$$

Claim 64. *Let $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ be a proper pair of segments. Then $A \neq D, C \neq B$, and $B \neq D$.*

Proof. The statement follows from the inequalities $f_{\overrightarrow{CD}}(A) \neq f_{\overrightarrow{CD}}(D)$, $f_{\overrightarrow{AB}}(C) \neq f_{\overrightarrow{AB}}(B)$, and $f_{\overrightarrow{AB}}(B) \neq f_{\overrightarrow{AB}}(D)$. □

The following theorem provides the criteria for a pair of oriented prime segments to be proper.

Theorem 65. *The pair of prime segments $\overrightarrow{AB}, \overrightarrow{CD}$ is proper if and only if one of the following holds:*

(1) $AC \subset BD$*;*

(2) $A \in BD$ *and* $\overrightarrow{CDB}$ *is a counterclockwise triangle or* $C \in BD$ *and* $\overrightarrow{ABD}$ *is counterclockwise triangle;*

(3) $\overrightarrow{ABCD}$ *is a counterclockwise quadrilateral.*

Proof. Clearly Conv$(\{A, B, C, D\})$ has at least 2 and at most 4 vertices. The proof of the theorem is split up into Lemmas 66, 67, and 68 according to the number of vertices of Conv$(\{A, B, C, D\})$. □

The following lemma treats the case where Conv$(\{A, B, C, D\})$ is a segment.

Lemma 66. *A pair of collinear prime segments* $\overrightarrow{AB}, \overrightarrow{CD}$ *is proper if and only if* $AC \subset BD$*;*

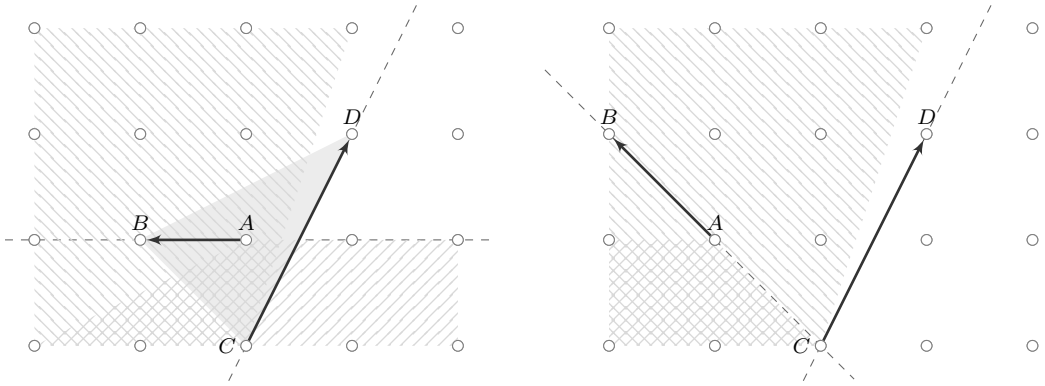
Proof. Let $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ be a proper pair of collinear prime segments (see Fig. 4.6a). Then using Claim 60 we derive from $f_{\overrightarrow{AB}}(D) = f_{\overrightarrow{CD}}(B) = 1$ the inclusion $A, C \in BD$.

Conversely, let $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ be a pair of collinear prime segments with $AC \subset BD$. The primality of the segments implies that $A \in BC$ and $C \in AD$. Therefore, by Claim 60, we have $f_{\overrightarrow{AB}}(C) = f_{\overrightarrow{CD}}(A) = f_{\overrightarrow{AB}}(D) = f_{\overrightarrow{CD}}(B) = 1$, and hence the pair $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is proper, as required. □

Lemma 67. *Let* $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ *be a pair of prime segments such that* Conv$(\{A, B, C, D\})$ *is a triangle. Then the pair is proper if and only if either* $\overrightarrow{CDB}$ *is a counterclockwise triangle with* $A \in BD$ *or* $\overrightarrow{ABD}$ *is a counterclockwise triangle with* $C \in BD$*.*

Proof. First, assume $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is a proper pair of prime segments with Conv$(\{A, B, C, D\})$ being a triangle. There are four cases to consider:

1. $D \in \overrightarrow{ABC}$. We claim that this case is impossible. Indeed, if D belongs to the triangle $\overrightarrow{ABC}$, then D belongs neither to BC nor to AC, as otherwise, by Claim 60, at least one of $f_{\overrightarrow{CD}}(A)$ and $f_{\overrightarrow{CD}}(B)$ would be zero, contradicting the assumption that $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is proper. Therefore, $\ell(CD)$ separates A and B, which contradicts $f_{\overrightarrow{CD}}(A) = f_{\overrightarrow{CD}}(B)$.
2. $B \in \overrightarrow{CDA}$. This case is impossible by similar arguments as in Case 1.
3. $C \in \overrightarrow{ABD}$. We show in this case that $\overrightarrow{ABD}$ is a counterclockwise triangle and $C \in BD$ (see Fig. 4.6b). The former follows from $f_{\overrightarrow{AB}}(D) = 1$. To prove the latter, suppose to the contrary that $C \notin BD$. Then $\ell(BD)$ does not intersect AC, and hence, by Claim 55, the orientations of the triangles $\overrightarrow{BDC}$ and $\overrightarrow{BDA}$ are the same. Since the orientation of $\overrightarrow{BDA}$ is the same as that of $\overrightarrow{ABD}$, we conclude that the orientation of $\overrightarrow{BCD}$ is counterclockwise, and therefore the orientation of $\overrightarrow{CDB}$ is clockwise, which contradicts $f_{\overrightarrow{CD}}(B) = 1$.



(a) A, B, C, D are in general position. $\mathcal{P}$ is the grey triangle.

(b) A, B, and C are collinear.

Figure 4.7: The stripped regions are $\text{Conv}(M_1(f_{\overrightarrow{AB}}))$ and $\text{Conv}(M_1(f_{\overrightarrow{CD}}))$. The grid region is $\text{Conv}(M_1(f_{\overrightarrow{AB}})) \cap \text{Conv}(M_1(f_{\overrightarrow{CD}}))$.

4. $A \in \overrightarrow{CDB}$. In this case arguments similar to the analysis of Case 3 show that $\overrightarrow{CDB}$ is a counterclockwise triangle and $A \in BD$.

Assume now that $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is a pair of prime segments such that $\overrightarrow{ABD}$ is a counterclockwise triangle and $C \in BD$. The case where $\overrightarrow{CDB}$ is a counterclockwise triangle with $A \in BD$ is symmetric and we omit the details. Since $C \in BD$, the orientation of $\overrightarrow{ABC}$ and $\overrightarrow{CDA}$ is the same as the orientation of $\overrightarrow{ABD}$, i.e. counterclockwise. Consequently, $f_{\overrightarrow{AB}}(D) = f_{\overrightarrow{AB}}(C) = f_{\overrightarrow{CD}}(A) = 1$. Furthermore, by Claim 60, we have $f_{\overrightarrow{CD}}(B) = 1$, and therefore the pair $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is proper. □

Lemma 68. *Let $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ be a pair of prime segments, such that A, B, C, and D are in convex position. Then the pair is proper if and only if AB, BC, CD, DA are edges of* $\text{Conv}(\{A, B, C, D\})$ *and the orientation of $\overrightarrow{ABCD}$ is counterclockwise.*

Proof. First let $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ be a proper pair of prime segments. It follows from $f_{\overrightarrow{AB}}(C) = f_{\overrightarrow{AB}}(D) = f_{\overrightarrow{CD}}(A) = f_{\overrightarrow{CD}}(B) = 1$ that the triangles $\overrightarrow{ABC}$, $\overrightarrow{ABD}$, $\overrightarrow{CDA}$, and $\overrightarrow{CDB}$ are counterclockwise. Therefore, by Claim 58, AB, BC, CD, DA are edges of $\text{Conv}(\{A, B, C, D\})$ and the orientation of $\overrightarrow{ABCD}$ is counterclockwise, as required (see Fig. 4.6c).

Conversely, let $\overrightarrow{ABCD}$ be a counterclockwise quadrilateral. By definition, the triangles $\overrightarrow{ABC}$, $\overrightarrow{BCD}$, $\overrightarrow{CDA}$, $\overrightarrow{DAB}$ are counterclockwise. Therefore

$$f_{\overrightarrow{CD}}(B) = f_{\overrightarrow{CD}}(A) = f_{\overrightarrow{AB}}(C) = f_{\overrightarrow{AB}}(D) = 1,$$

and hence the pair $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is proper. □

The following claim is related to the property of non-proper pairs of oriented prime segments.

Claim 69. *Let $\overrightarrow{AB}, \overrightarrow{CD}$ be distinct prime segments in $\mathcal{G}_{m,n}$ such that $f_{\overrightarrow{AB}}(C) = 1$, $f_{\overrightarrow{AB}}(D) = 0$, and $f_{\overrightarrow{CD}}(A) = 1$. Then $f_{\overrightarrow{CD}}(B) = 1$, the points B, C, D are not collinear, and $A \in \overrightarrow{BCD}$.*

Proof. First we claim that the points A, B, C, D are not collinear. Suppose to the contrary, that they are collinear. Then, by Claim 60, we have $A \in BC$, $B \in AD$, and $C \in AD$, which imply that either $A = C$ or $A = B$. The latter is not possible as AB is a prime segment. Therefore $A = C$ and $B \in CD$. Since CD is prime and $C = A \neq B$, we conclude that $B = D$ and $\overrightarrow{AB} = \overrightarrow{CD}$, which contradicts the assumption of the statement.

Assume now that A, B, C, D do not lie on the same line. From $f_{\overrightarrow{AB}}(C) \neq f_{\overrightarrow{AB}}(D)$ it follows that $\ell(AB)$ intersects CD. Suppose three of the points A, B, C, D are collinear. We will consider four cases:

1. *A, C, B are collinear, i.e. $CD \cap \ell(AB) = C$* (see Fig. 4.7b). By Claim 60, we have $A \in BC$ and hence $A \in \overrightarrow{BCD}$. To show $f_{\overrightarrow{CD}}(B) = 1$ we observe that the segments $\overrightarrow{AB}$ and $\overrightarrow{CB}$ are collinear and have the same orientation, and therefore, by Claim 56, the triangles $\overrightarrow{ABD}$ and $\overrightarrow{CBD}$ have the same orientation. Since $f_{\overrightarrow{AB}}(D) = 0$, the triangle $\overrightarrow{ABD}$ is clockwise, and hence $\overrightarrow{CDB}$ is counterclockwise and $f_{\overrightarrow{CD}}(B) = 1$.

2. *A, B, D are collinear, i.e. $CD \cap \ell(AB) = D$.* We will prove that this case is impossible by showing that $\overrightarrow{CDA}$ is a clockwise triangle, which contradicts $f_{\overrightarrow{CD}}(A) = 1$. By Claim 60, we have $B \in AD$, and therefore the segments $\overrightarrow{AB}$ and $\overrightarrow{AD}$ are collinear and have the same orientation. Hence, by Claim 56, the triangles $\overrightarrow{ABC}$ and $\overrightarrow{ADC}$ have the same orientation. Namely, since $f_{\overrightarrow{AB}}(C) = 1$, we conclude that both triangles are counterclockwise. Consequently, $\overrightarrow{CDA}$ is clockwise, as desired.

3. *A, C, D are collinear, i.e. $CD \cap \ell(AB) = A$.* Since CD is prime and $f_{\overrightarrow{CD}}(A) = 1$, we conclude that $A = C$ and hence the first case takes place.

4. *C, B, D are collinear, i.e. $CD \cap \ell(AB) = B$.* Since CD is prime and $f_{\overrightarrow{AB}}(D) = 0$, we conclude that $B = D$ and hence the second case takes place.

Assume finally that A, B, C, D are in general position and denote $\mathcal{P} = \text{Conv}(\{A, B, C, D\})$ (see Fig. 4.7a). We consider the oriented triangles $\overrightarrow{CDA}$, $\overrightarrow{ABC}$, $\overrightarrow{BAD}$, and $\overrightarrow{CDB}$. It follows from the assumptions of the claim that the first three triangles are counterclockwise. Therefore, by Claim 57, the triangle $\overrightarrow{CDB}$ is also counterclockwise, and hence $f_{\overrightarrow{CD}}(B) = 1$.

It remains to show that A belongs to the triangle $\overrightarrow{BCD}$, i.e. $\mathcal{P} = \overrightarrow{BCD}$. Suppose, to the contrary, $\mathcal{P} \neq BCD$. Then A is a vertex of $\mathcal{P}$ and two of the segments AC, AB, and AD are edges of $\mathcal{P}$. We will arrive to a contradiction by showing that neither AB nor AD

can be an edge of $\mathcal{P}$. Indeed, if AB is an edge of $\mathcal{P}$, then C and D are not separated by $\ell(AB)$, which contradicts $f_{\overrightarrow{AB}}(C) \neq f_{\overrightarrow{AB}}(D)$. Furthermore, if AD is an edge of $\mathcal{P}$, then B and C are not separated by $\ell(AD)$, and hence the triangles $\overrightarrow{DAC}$ and $\overrightarrow{DAB}$ have the same orientation. However, the triangle $\overrightarrow{DAC}$ is counterclockwise as $f_{\overrightarrow{CD}}(A) = 1$, and the triangle $\overrightarrow{DAB}$ is clockwise as $f_{\overrightarrow{AB}}(D) = 0$. Contradiction. □

Corollary 70. *Under the conditions of Claim 69 the intersection $\ell(AB) \cap CD$ is a point X and $A \in XB$.*

Theorem 65 implies a sequence of useful statements about 2-threshold functions. The first of them leads to the conclusion that the 2-threshold function defined by a pair of oriented segments is *proper* whenever the pair is *proper*.

Claim 71. *Let $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ be a proper pair of segments. Then $AC \cap BD \neq \emptyset$.*

Proof. By Theorem 65, one of the following statements is true:

(1) $AC \subset BD$; in this case $AC \cap BD = AC$.

(2) $A \in BD$ and $\overrightarrow{CDB}$ is a counterclockwise triangle or $C \in BD$ and $\overrightarrow{ABD}$ is counterclockwise triangle; then $AC \cap BD = A$ or $AC \cap BD = C$ respectively.

(3) $\overrightarrow{ABCD}$ is a convex counterclockwise quadrilateral, hence AC and BD are diagonals, and therefore they intersect.

In all cases we have $AC \cap BD \neq \emptyset$, as required. □

The claim proves that the convex hulls of the sets of true and false points of a function defined by a proper pair of segments intersect, and hence the function is not threshold.

Corollary 72. *Every* proper *pair of oriented segments in $\mathcal{G}_{m,n}$ defines a* proper *2-threshold function over $\mathcal{G}_{m,n}$.*

Corollary 73. *Let $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ be a proper pair of collinear segments that define a 2-threshold function f over $\mathcal{G}_{m,n}$. Then $M_1(f) = AC \cap \mathcal{G}_{m,n}$ (see Fig. 4.6a).*

Corollary 74. *Let $\{\overrightarrow{AB}, \overrightarrow{AD}\}$ be a proper pair of segments that define a 2-threshold function f over $\mathcal{G}_{m,n}$. Then $M_1(f) = \{A\}$ (see Fig. 4.8).*

Corollary 75. *Let $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ be a proper pair of segments that define a 2-threshold function f over $\mathcal{G}_{m,n}$. Then $AB \cap CD \neq \emptyset$ if and only if $M_1(f) = \{A\}$ (see Fig. 4.8).*

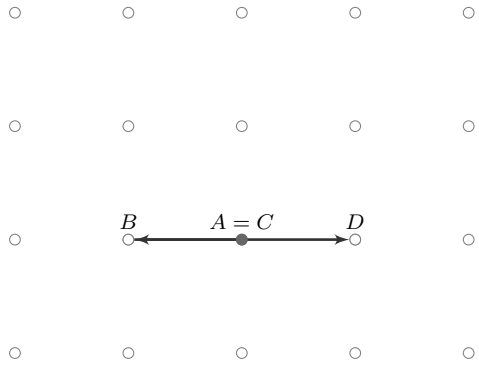


Figure 4.8: f is true in the unique point $A = C$.

4.4.1 Proper pairs of oriented segments and proper 2-threshold functions

In the following statements we will show that any proper 2-threshold function f can be defined by a proper pair of segments, and such a pair is unique if f has a true point on the boundary of the grid. We start with the existence of a proper pair of segments for f.

Theorem 76. *For any proper* 2*-threshold function* f *over* $\mathcal{G}_{m,n}$ *there exists a* proper *pair of oriented segments in* $\mathcal{G}_{m,n}$ *that defines* f.

Proof. Note that every proper 2-threshold function is a conjunction of two non-constant threshold functions, therefore it follows from Corollary 63 that there exists a pair of oriented prime segments that defines f. Let $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ be a pair of oriented prime segments defining f such that $|M_1(f_{\overrightarrow{AB}})| + |M_1(f_{\overrightarrow{CD}})|$ is minimized. We claim that $f_{\overrightarrow{CD}}(A) = f_{\overrightarrow{AB}}(C) = 1$. For the sake of contradiction, assume without loss of generality that $f_{\overrightarrow{CD}}(A) = 0$. By Theorem 61, the point A is essential for $f_{\overrightarrow{AB}}$, hence the function f', that differs from $f_{\overrightarrow{AB}}$ in the unique point A, is threshold. Since $A \in M_0(f_{\overrightarrow{CD}})$ and $M_1(f') = M_1(f_{\overrightarrow{AB}}) \setminus \{A\}$, we have

$$M_1(f') \cap M_1(f_{\overrightarrow{CD}}) = M_1(f_{\overrightarrow{AB}}) \cap M_1(f_{\overrightarrow{CD}}) = M_1(f),$$

and therefore $f = f' \wedge f_{\overrightarrow{CD}}$. By assumption f is proper, and hence f' is a non-constant threshold function. Consequently, by Corollary 63, there exists an oriented prime segment $\overrightarrow{A'B'}$ that defines f'. Therefore, the pair $\overrightarrow{A'B'}, \overrightarrow{CD}$ defines f. But $|M_1(f')| < |M_1(f_{\overrightarrow{AB}})|$, which contradicts the choice of $\overrightarrow{AB}, \overrightarrow{CD}$.

Since f is non-threshold, there exist $X, Y \in M_0(f)$ such that $XY \cap \text{Conv}(M_1(f)) \neq \emptyset$. Indeed, otherwise $\text{Conv}(M_0(f))$ and $\text{Conv}(M_1(f))$ would be disjoint, and therefore separable by a line. Hence, for any pair of prime segments $\overrightarrow{AB}, \overrightarrow{CD}$ that defines f neither $f_{\overrightarrow{AB}}$ nor $f_{\overrightarrow{CD}}$ can be false in both X, Y. Furthermore, since $X, Y \in M_0(f)$, we conclude that one of the points is a false point of $f_{\overrightarrow{AB}}$ and a true point of $f_{\overrightarrow{CD}}$, and the other point is a true point of $f_{\overrightarrow{AB}}$ and a false point of $f_{\overrightarrow{CD}}$.

Let $\mathcal{X}$ be the family of *ordered* pairs of segments $\overrightarrow{AB}, \overrightarrow{CD}$ defining f such that

$X \in M_0(f_{\overrightarrow{AB}}) \cap M_1(f_{\overrightarrow{CD}})$ and $Y \in M_1(f_{\overrightarrow{AB}}) \cap M_0(f_{\overrightarrow{CD}})$. Denote

$$M_X = \bigcap_{(\overrightarrow{AB},\overrightarrow{CD})\in\mathcal{X}} M_0(f_{\overrightarrow{AB}}) \cap M_1(f_{\overrightarrow{CD}}).$$

$$M_Y = \bigcap_{(\overrightarrow{AB},\overrightarrow{CD})\in\mathcal{X}} M_1(f_{\overrightarrow{AB}}) \cap M_0(f_{\overrightarrow{CD}}).$$

Notice that each of M_X and M_Y is the intersections of convex sets that have a common element, and therefore both M_X and M_Y are non-empty and convex. Moreover, since $M_X, M_Y \subset M_0(f)$, both Conv(M_X) and Conv(M_Y) are disjoint from Conv$(M_1(f))$.

Let ℓ_X be the left inner common tangent to Conv$(M_1(f))$ and Conv(M_X). Let $A^* \in$ Conv$(M_1(f)) \cap \ell_X$, $B^* \in$ Conv$(M_X) \cap \ell_X$ be such that A^*B^* is of minimum length. We claim that A^*B^* is a prime segment. To prove this, we show first that Conv$(M_1(f) \cup M_X)$ contains no integer points other than points in $M_1(f) \cup M_X$. Indeed, let $(\overrightarrow{AB}, \overrightarrow{CD})$ be a pair of segments from $\mathcal{X}$, and suppose there exists an integer point Z in Conv$(M_1(f) \cup M_X)$ that belongs neither to $M_1(f)$ nor to M_X. Notice, by definition, $M_X \subset M_1(f_{\overrightarrow{CD}})$ and $M_1(f) \subset M_1(f_{\overrightarrow{CD}})$, which implies that Conv$(M_1(f) \cup M_X) \subseteq$ Conv$(M_1(f_{\overrightarrow{CD}}))$. Consequently, if $f_{\overrightarrow{AB}}(Z) = 1$ we have $Z \in M_1(f)$, and if $f_{\overrightarrow{AB}}(Z) = 0$ we have $Z \in M_X$, a contradiction. Now, any segment with endpoints in $M_1(f) \cup M_X$ belongs to Conv$(M_1(f) \cup M_X)$, hence if there is an integer point Z in the interior of A^*B^* then $Z \in M_1(f) \cup M_X$, which contradicts the minimality of A^*B^*. Similarly, considering the left inner common tangent ℓ_Y to Conv$(M_1(f))$ and Conv(M_Y), the two points $C^* \in M_1(f) \cap \ell_Y$, $D^* \in M_Y \cap \ell_Y$ at minimum distance define a prime segment C^*D^*. Fig. 4.9 illustrates M_X, M_Y, A^*, B^*, C^*, and D^*.

Let now $f^* = f_{\overrightarrow{A^*B^*}} \wedge f_{\overrightarrow{C^*D^*}}$ be the 2-threshold function defined by $\{\overrightarrow{A^*B^*}, \overrightarrow{C^*D^*}\}$. In the rest of the proof we will show that $f = f^*$ and the pair $\{\overrightarrow{A^*B^*}, \overrightarrow{C^*D^*}\}$ is proper. To establish the former we will prove that $M_1(f) = M_1(f^*)$.

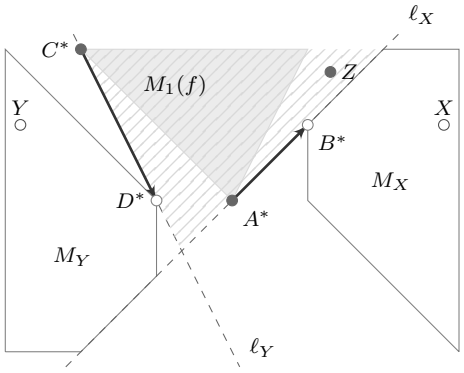
First we show that $M_1(f) \subseteq M_1(f^*)$. Indeed, by definition, $\ell(\overrightarrow{A^*B^*}) = \ell_X$ is a left tangent from B^* to Conv$(M_1(f))$, and therefore $M_1(f) \subseteq M_1(f_{\overrightarrow{A^*B^*}})$. Similarly, we have $M_1(f) \subseteq M_1(f_{\overrightarrow{C^*D^*}})$, and therefore $M_1(f) \subseteq M_1(f_{\overrightarrow{A^*B^*}}) \cap M_1(f_{\overrightarrow{C^*D^*}}) = M_1(f^*)$.

Now, let us show that $M_1(f^*) \subseteq M_1(f)$. Assume, to the contrary, $M_1(f^*) \setminus M_1(f) \neq \emptyset$ and let Z be a point in $M_1(f^*) \setminus M_1(f)$. In particular, we have $Z \notin M_X \cup M_Y$. We observe that $f(Z) = 0$ and $Z \notin M_Y$ imply that there exists a pair $(\overrightarrow{AB}, \overrightarrow{CD}) \in \mathcal{X}$ such that $Z \in M_0(f_{\overrightarrow{AB}})$, and therefore $M_X \cup \{Z\} \subseteq M_0(f_{\overrightarrow{AB}})$ and

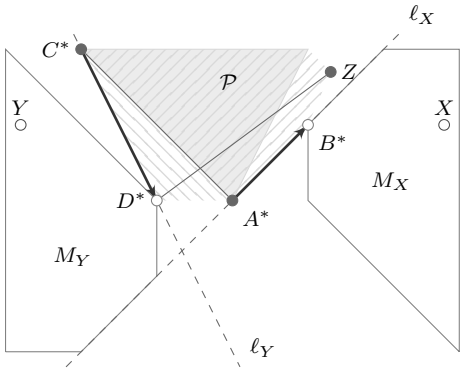
$$\text{Conv}(M_X \cup \{Z\}) \cap \text{Conv}(M_1(f)) = \emptyset. \tag{4.2}$$

Similarly, it can be shown that

$$\text{Conv}(M_Y \cup \{Z\}) \cap \text{Conv}(M_1(f)) = \emptyset. \tag{4.3}$$



(a) All integer points of the stripped region are exactly the true points of f^*. Z is chosen outside of Conv$(M_1(f))$ and such that $f^*(Z) = 1$.

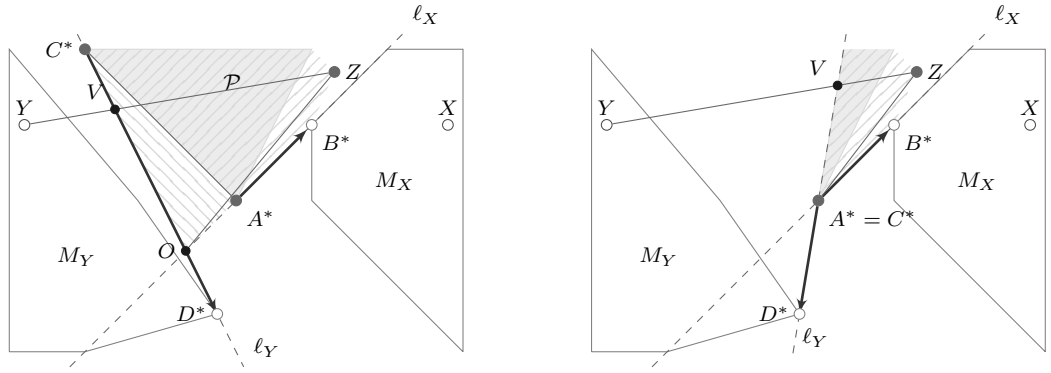


(b) The pair $\overrightarrow{A^*B^*}$, $\overrightarrow{C^*D^*}$ is proper. The stripped region is $\mathcal{P}$. S_1 and S_2 have the different pattern orientation. The segment D^*Z intersects A^*C^*.

Figure 4.9: The white polygons are Conv(M_X) and Conv(M_Y). The grey polygon is Conv$(M_1(f))$.

We will consider two cases depending on whether $\{\overrightarrow{A^*B^*}, \overrightarrow{C^*D^*}\}$ is a proper pair or not. We start with the case of proper pair, in which case we have $f_{\overrightarrow{A^*B^*}}(D^*) = f_{\overrightarrow{C^*D^*}}(B^*) = 1$ (see Fig. 4.9a). First we claim that $A^* \neq C^*$. Indeed, otherwise, by Corollary 74, we would have $M_1(f^*) = \{A^*\}$, and therefore since $M_1(f) \subseteq M_1(f^*)$ and $M_1(f^*) \setminus M_1(f) \neq \emptyset$, we would conclude that f is the constant-zero function, contradicting the assumption that f is a proper 2-threshold function. Let us now denote $\mathcal{P} = \text{Conv}(M_1(f^*) \cup \{B^*, D^*\})$. From $M_1(f) \cup \{D^*\} \subseteq M_1(f_{\overrightarrow{A^*B^*}})$ and $A^*, B^* \in \ell_X$ it follows that ℓ_X is a tangent to $\mathcal{P}$ where A^* is a tangent point. Analysis similar to the above implies that ℓ_Y is a tangent to $\mathcal{P}$ and C^* is a tangent point. Consequently, all points of $\mathcal{P} \setminus A^*C^*$ are separated by the segment A^*C^* into two parts, which we denote as S_1 and S_2 (see Fig. 4.9b). By Claim 71, the segments A^*C^* and B^*D^* intersect, and hence B^* and D^* are in different parts, say $B^* \in S_1$ and $D^* \in S_2$. We now claim that Z belongs to one of the parts S_1 and S_2. To see this, we first observe that $Z \in M_1(f^*) \subseteq \mathcal{P}$. Furthermore, since Z belongs to $M_0(f)$, it does not belong to A^*C^*, and hence the claim. Now, assume without loss of generality $Z \in S_1$, and therefore D^*Z intersects A^*C^*. Since $D^* \in M_Y$ and $A^*C^* \subseteq \text{Conv}(M_1(f))$, we conclude that $\text{Conv}(M_Y \cup \{Z\}) \cap \text{Conv}(M_1(f)) \neq \emptyset$, which contradicts (4.3).

Suppose now that the pair $\{\overrightarrow{A^*B^*}, \overrightarrow{C^*D^*}\}$ is not proper, which implies that $f_{\overrightarrow{A^*B^*}}(D^*) = 0$ or $f_{\overrightarrow{C^*D^*}}(B^*) = 0$. There is no loss of generality in assuming $f_{\overrightarrow{A^*B^*}}(D^*) = 0$ (see Fig. 4.10a). Then Claim 69 yields $f_{\overrightarrow{C^*D^*}}(B^*) = 1$. Let $A^* \neq C^*$, the case $A^* = C^*$ will be considered separately. From $f_{\overrightarrow{A^*B^*}}(C^*) \neq f_{\overrightarrow{A^*B^*}}(D^*)$ it follows that ℓ_X intersects C^*D^*. We denote $O = \ell_X \cap C^*D^*$ and consider $\mathcal{P} = \text{Conv}(M_1(f^*) \cup \{B^*, O\})$. As in the previous case it can be verified that ℓ_X, ℓ_Y are tangents to $\mathcal{P}$, and



(a) $A^* \neq C^*$, $OZ \subseteq \text{Conv}(M_Y \cup \{Z\})$. (b) $A^* = C^*$, $A^* \in \text{Conv}(M_Y \cup \{Z\})$.

Figure 4.10: The pair $\{\overrightarrow{A^*B^*}, \overrightarrow{C^*D^*}\}$ is not proper and $f_{\overrightarrow{A^*B^*}}(D^*) = 0$. The grey region is $\text{Conv}(M_1(f))$. The stripped region is $\mathcal{P}$. S_1 and S_2 have the different pattern orientation.

therefore, A^* and C^* are tangent points. Thus the points of $\mathcal{P} \setminus A^*C^*$ are separated by A^*C^* into two parts, which we denote as S_1 and S_2. We next prove that O and B^* are in different parts. For this purpose, we consider the triangle $\overrightarrow{B^*C^*D^*}$, and Claim 69 implies $A^* \in \overrightarrow{B^*C^*D^*}$. It is easily seen that $OB^* = \overrightarrow{B^*C^*D^*} \cap \ell(A^*B^*)$, hence $A^* \in OB^*$, and therefore O and B^* belong to the different parts, say $B^* \in S_1$ and $O \in S_2$. Clearly, $Z \in \mathcal{P} \setminus A^*C^*$, and therefore either $Z \in S_1$ or $Z \in S_2$. The latter would contradict (4.2), so we assume the former holds, which in turn implies $OZ \cap A^*C^* \neq \emptyset$. To obtain a contradiction with (4.3) we will show $OZ \subseteq \text{Conv}(M_Y \cup \{Z\})$. To this end we first observe that ℓ_Y intersects YZ because $f_{\overrightarrow{C^*D^*}}(Y) \neq f_{\overrightarrow{C^*D^*}}(Z)$. Let V be the intersection point of YZ and ℓ_Y. Now from $f_{\overrightarrow{A^*B^*}}(Y) = f_{\overrightarrow{A^*B^*}}(Z) = 1$ it follows that $V \in \text{Conv}(M_1(f_{\overrightarrow{A^*B^*}}))$. Since $D^* \in M_0(f_{\overrightarrow{A^*B^*}})$, we conclude that ℓ_X intersects D^*V and $O \in D^*V$. But $D^*V \subseteq \overrightarrow{YD^*Z} \subseteq \text{Conv}(M_Y \cup \{Z\})$, and therefore $O \in \text{Conv}(M_Y \cup \{Z\})$ and $OZ \subseteq \text{Conv}(M_Y \cup \{Z\})$, leading to a contradiction. Suppose now that $A^* = C^*$ (see Fig. 4.10b). By replacing O with A^*, and using arguments similar to the above one can show that $A^* \in \overrightarrow{YD^*Z}$ and $A^*Z \subseteq \text{Conv}(M_Y \cup \{Z\})$, which contradicts (4.3). The contradictions in all the cases imply that $M_1(f^*) \setminus M_1(f) = \emptyset$, and hence $f = f^*$.

We have shown that $\{\overrightarrow{A^*B^*}, \overrightarrow{C^*D^*}\}$ defines f. It remains to prove that $\{\overrightarrow{A^*B^*}, \overrightarrow{C^*D^*}\}$ is a proper pair of segments. Since $B^* \in M_X$ and $B^* \in M_0(f_{\overrightarrow{A^*B^*}})$, the definition of M_X implies that $f_{\overrightarrow{C^*D^*}}(B^*) = 1$. Similarly, from $D^* \in M_Y$ and $D^* \in M_0(f_{\overrightarrow{C^*D^*}})$ we conclude $f_{\overrightarrow{A^*B^*}}(D^*) = 1$. Finally, the equality $f_{\overrightarrow{A^*B^*}}(C^*) = f_{\overrightarrow{C^*D^*}}(A^*) = 1$ follows from $A^*, C^* \in M_1(f)$. Hence $\{\overrightarrow{A^*B^*}, \overrightarrow{C^*D^*}\}$ is a proper pair of segments that defines f, as claimed. □

Lemma 77. *Let f be a $\{0,1\}$-valued function over $\mathcal{G}_{m,n}$ with a unique true point $X = (x_1, x_2)$ such that either $x_1 \in \{0, m-1\}$ or $x_2 \in \{0, n-1\}$, but not both. Then f is a*

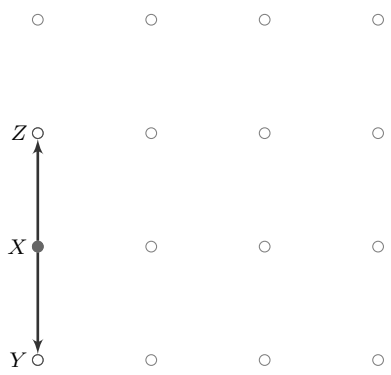


Figure 4.11: $\{\overrightarrow{XY}, \overrightarrow{XZ}\}$ defines a 2-threshold function f such that $M_1(f) = \{X\}$.

2-*threshold function with a unique proper pair of segments defining* f.

Proof. Due to symmetry it is enough to consider the case $x_1 = 0$ and $x_2 \in \{1, \ldots, n-2\}$. We will show that $\{\overrightarrow{XY}, \overrightarrow{XZ}\}$, where $Y = (0, x_2 - 1)$, $Z = (0, x_2 + 1)$, is the desired pair (see Fig. 4.11). In [48] it was proved that any $\{0,1\}$-function containing one true point is k-threshold for any $k \geq 2$, hence f is a 2-threshold function. From Theorem 65 and Corollary 74 it follows that the pair $\{\overrightarrow{XY}, \overrightarrow{XZ}\}$ is proper and defines f. Now, let us prove that there is no other proper pair of segments that defines f.

Let $\{\overrightarrow{XY'}, \overrightarrow{XZ'}\}$ be a proper pair segments that defines f. We will show that $\{Y', Z'\} = \{Y, Z\}$. First, $f(Z) = 0$ implies that $f_{\overrightarrow{XY'}}(Z) = 0$ or $f_{\overrightarrow{XZ'}}(Z) = 0$. Without loss of generality we assume $f_{\overrightarrow{XZ'}}(Z) = 0$. Since both $\overrightarrow{XZ}$ and $\overrightarrow{XZ'}$ are prime, we conclude that either $Z' = Z$ or $\overrightarrow{XZ'Z}$ is a clockwise triangle. For the sake of contradiction, let us assume the latter holds. By definition of a clockwise triangle,

$$\begin{vmatrix} 0 & x_2 & 1 \\ z_1 & z_2 & 1 \\ 0 & x_2 + 1 & 1 \end{vmatrix} = z_1 < 0,$$

where $Z' = (z_1, z_2)$. But this contradicts $z_1 \geq 0$, hence $Z' = Z$. Now let us show that $Y' = Y$. Indeed, as $\{\overrightarrow{XY'}, \overrightarrow{XZ}\}$ is a proper pair, by definition, $Y' \in M_1(f_{\overrightarrow{XZ}}) = \{(0,0), (0,1), \ldots, (0, x_2)\}$, and therefore, since $\overrightarrow{XY'}$ is prime and $X = (0, x_2)$, we conclude that $Y' = (0, x_2 - 1) = Y$. □

Theorem 78. *For any proper 2-threshold function f over $\mathcal{G}_{m,n}$ that contains true points on the boundary of $\mathcal{G}_{m,n}$ there exists a* unique *proper pair of oriented segments in $\mathcal{G}_{m,n}$ that defines f.*

Proof. By Theorem 76, there exists at least one proper pair of oriented segments that defines f. Suppose, for the sake of contradiction, that there are two different proper pairs of oriented segments defining f, which we denote as $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ and $\{\overrightarrow{A'B'}, \overrightarrow{C'D'}\}$ respectively.

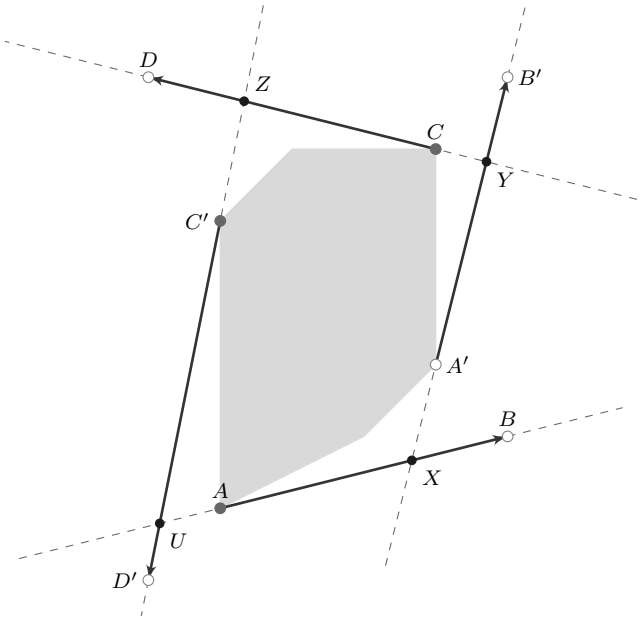


Figure 4.12: The grey region is Conv$(M_1(f))$, which is included in Conv$(\{X, Y, Z, U\})$.

First we will prove that

$$\{\overrightarrow{AB}, \overrightarrow{CD}\} \cap \{\overrightarrow{A'B'}, \overrightarrow{C'D'}\} = \emptyset. \tag{4.4}$$

Suppose, to the contrary, that $\overrightarrow{AB} = \overrightarrow{A'B'}$, then $\overrightarrow{CD} \neq \overrightarrow{C'D'}$. Since $f_{\overrightarrow{AB}}(D) = f_{\overrightarrow{AB}}(D') = 1$ and $f(D) = f(D') = 0$, we have $f_{\overrightarrow{C'D'}}(D) = f_{\overrightarrow{CD}}(D') = 0$. Furthermore, $f(C) = f(C') = 1$ implies $f_{\overrightarrow{C'D'}}(C) = f_{\overrightarrow{CD}}(C') = 1$. On the other hand, by Claim 69, the equations $f_{\overrightarrow{C'D'}}(C) = 1, f_{\overrightarrow{CD}}(C') = 1, f_{\overrightarrow{CD}}(D') = 0$ imply $f_{\overrightarrow{C'D'}}(D) = 1$, a contradiction.

Now we will look more closely at the functions $f_{\overrightarrow{AB}}, f_{\overrightarrow{CD}}, f_{\overrightarrow{A'B'}}$, and $f_{\overrightarrow{C'D'}}$. Since $f(B) = 0$ we have either $f_{\overrightarrow{A'B'}}(B) = 0$ or $f_{\overrightarrow{C'D'}}(B) = 0$. Without loss of generality we assume $f_{\overrightarrow{A'B'}}(B) = 0$. From $f_{\overrightarrow{AB}}(A') = 1, f_{\overrightarrow{A'B'}}(A) = 1, f_{\overrightarrow{A'B'}}(B) = 0$, and Claim 69 it follows that the points A, B, B' are not collinear and $f_{\overrightarrow{AB}}(B') = 1$. The latter together with the fact that $f(B') = 0$ imply $f_{\overrightarrow{CD}}(B') = 0$. By Corollary 70, the line $\ell(A'B')$ intersects AB in a unique point, which we denote by X, and $A' \in XB'$.

Analysis similar to above shows that $f_{\overrightarrow{CD}}(B') = 0$ implies $f_{\overrightarrow{C'D'}}(D) = 0$ and that the line $\ell(CD)$ intersects $A'B'$ in a unique point, which we denote by Y, and $C \in YD$. In turn, the equation $f_{\overrightarrow{C'D'}}(D) = 0$ implies $f_{\overrightarrow{AB}}(D') = 0$ and the intersection of $\ell(C'D')$ and CD in a unique point denoted by Z, and $C' \in ZD'$. Finally, the equation $f_{\overrightarrow{AB}}(D') = 0$ implies that $\ell(AB)$ intersects $C'D'$ in a unique point denoted by U, and $A \in UB$.

In the rest of the proof we will show that $M_1(f) \subseteq$ Conv$(\{X, Y, Z, U\})$ and that X, Y, Z, U are interior points of Conv$(\mathcal{G}_{m,n})$, which will lead to a contradiction (see Fig. 4.12). We will consider four different cases.

Case 1. *The points X, Y, Z, U are pairwise distinct.* First we will show that $\text{Conv}(\{X, Y, Z, U\})$ is a counterclockwise quadrilateral with the edges XY, YZ, ZU, and UX (see Fig. 4.12). Applied to $f_{\overrightarrow{AB}}, f_{\overrightarrow{C'D'}}$, Claim 69 yields $A \in \overrightarrow{BC'D'}$, and hence $A \in UB$. The latter together with $X \in AB$ imply that $\overrightarrow{AB}$ and $\overrightarrow{UX}$ have the same orientation. By similar arguments, $\overrightarrow{A'B'}$ and $\overrightarrow{XY}$, $\overrightarrow{CD}$ and $\overrightarrow{YZ}$, and $\overrightarrow{C'D'}$ and $\overrightarrow{ZU}$ have the same orientation respectively. Now we observe that the assumption $Y \neq Z$ implies $Z \notin \ell(A'B')$. Therefore, since $f_{\overrightarrow{A'B'}}(C) = f_{\overrightarrow{A'B'}}(D) = 1$ and $Z \in CD$, the triangle $\overrightarrow{A'B'Z}$ is counterclockwise. Hence, by Claim 56, the triangle $\overrightarrow{XYZ}$ is counterclockwise. By similar arguments, the triangles $\overrightarrow{YZU}$, $\overrightarrow{ZUX}$, $\overrightarrow{UXY}$ are counterclockwise. Consequently, by Claim 58, $\text{Conv}(\{X, Y, Z, U\})$ is a quadrilateral $XYZU$ with edges XY, YZ, ZU, UX.

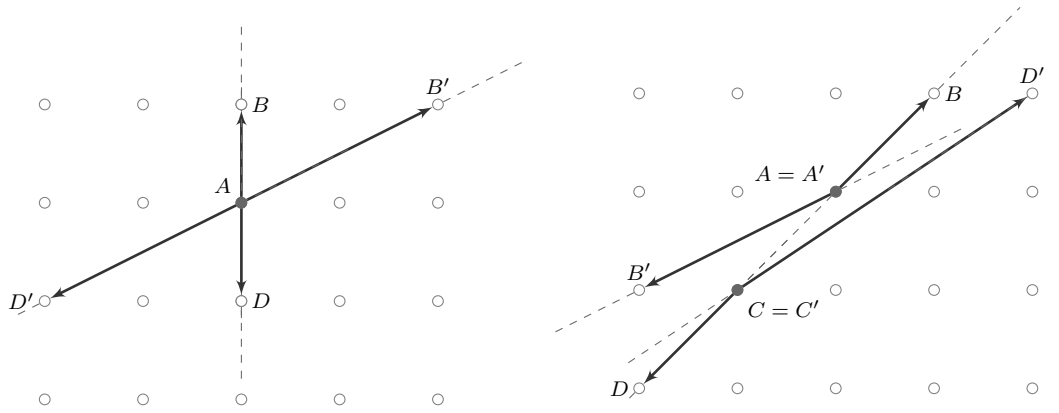
Next, the inclusion $\text{Conv}(M_1(f)) \subseteq XYZU$ follows from the fact that $XYZU$ is a polygon circumscribed about $\text{Conv}(M_1(f))$. Indeed, each of the lines $\ell(A'B') = \ell(XY)$, $\ell(CD) = \ell(YZ)$, $\ell(C'D') = \ell(ZU)$, and $\ell(AB) = \ell(UX)$ is a tangent to $\text{Conv}(M_1(f))$, and $A' \in XY \cap \text{Conv}(M_1(f))$, $C \in YZ \cap \text{Conv}(M_1(f))$, $C' \in ZU \cap \text{Conv}(M_1(f))$, $A \in UX \cap \text{Conv}(M_1(f))$.

It remains to prove that all the points X, Y, Z, and U are interior points of $\text{Conv}(\mathcal{G}_{m,n})$, i.e. $X, Y, Z, U \notin B(\mathcal{G}_{m,n})$, where

$$B(\mathcal{G}_{m,n}) = \{0, m-1\} \times [0, n-1] \cup [0, m-1] \times \{0, n-1\}.$$

We will prove that $X \notin B(\mathcal{G}_{m,n})$, for the other three points the arguments are similar. Suppose, to the contrary, that $X \in B(\mathcal{G}_{m,n})$. Since $X \in AB$ and $A \in UB$, we have $X \in UB$. We claim that X is an interior point of UB. Indeed, $X \neq U$ by the assumption. Furthermore, the equality $X = B$ would imply $A' \in BB'$, which is not possible as $f_{\overrightarrow{A'B'}}$ is a threshold function and $f_{\overrightarrow{A'B'}}(B) = 0, f_{\overrightarrow{A'B'}}(A') = 1, f_{\overrightarrow{A'B'}}(B') = 0$. Now, since both U and B belong to $\text{Conv}(\mathcal{G}_{m,n})$, and X is an interior point of UB and a boundary point of $\text{Conv}(\mathcal{G}_{m,n})$, we conclude that $\ell(UB) = \ell(AB)$ is a tangent to $\text{Conv}(\mathcal{G}_{m,n})$. We will arrive to a contradiction by showing that $\ell(AB)$ separates D and D'. First, we observe that $D' \notin \ell(AB)$, as otherwise we would have $U = D'$ and $A \in D'B$, which is not possible as $f_{\overrightarrow{AB}}$ is threshold and $f_{\overrightarrow{AB}}(B) = 0, f_{\overrightarrow{AB}}(A) = 1, f_{\overrightarrow{AB}}(D') = 0$. Consequently, $\overrightarrow{ABD'}$ is a clockwise triangle. On the other hand, the triangle $\overrightarrow{ABD}$ is counterclockwise as the pair $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is proper. Therefore, $\ell(AB)$ separates D and D'. This contradiction proves that X does not belong to $B(\mathcal{G}_{m,n})$.

Case 2. *$X = Z$ or $Y = U$.* Suppose $X = Z$. Then from $X \in AB$ and $Z \in CD$ it follows that AB and CD intersect. However, $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is a proper pair of segments, and, by Corollary 75, we have $M_1(f) = \{A\}$ (see Fig. 4.13a). Since f is a proper 2-threshold function, A is not a vertex of $\text{Conv}(\mathcal{G}_{m,n})$, and therefore Lemma 77 implies $A \in \{1, \ldots, m-2\} \times \{1, \ldots, n-2\}$, as required. The case $Y = U$ is symmetric and we omit the details.



(a) $M_1(f) = \{A\}$, $A = A' = C = C' = X = Y = Z = U$.

(b) $M_1(f) = \{A, C\}$, $A = A' = X = Y, C = C' = Z = U$.

Figure 4.13: Examples of 2-threshold functions with two distinct proper pairs of segments.

Case 3. $|\{X, Y, Z, U\}| = 3$, $X \neq Z$, *and* $Y \neq U$. Let $X = Y$, using the same arguments as in Case 1 it can be shown that $\overrightarrow{XZU}$ is a triangle circumscribed about $\text{Conv}(M_1(f))$, and that none of X, Z, and U lies on the boundary of $\mathcal{G}_{m,n}$. The cases $X = U$, $Y = Z$, and $Z = U$ are symmetric and we omit the details.

Case 4. $|\{X, Y, Z, U\}| = 2$ *and* $X \neq Z, Y \neq U$. Then either $X = Y$ and $U = Z$ or $X = U$ and $Y = Z$. The two cases are symmetric and therefore we consider only one of them, namely, $X = Y, U = Z$. First we will show that $\text{Conv}(M_1(f)) = AC$. Indeed, from $X \in AB$, $Y \in A'B'$, and $A' \in \overrightarrow{ABB'}$ it follows that $X = Y = A'$, and hence $A = A'$ as AB is prime. Moreover, $Y \in \ell(CD)$ together with $Y = A$ imply that A, C, D are collinear points, and hence $\text{Conv}(\{A, B, C, D\})$ has at most three vertices. Then, by Theorem 65, either $A \in BD$ or $C \in BD$ or both. All cases lead to the conclusion that A, B, C, D are collinear, and, by Corollary 73, we have $\text{Conv}(M_1(f)) = AC$ (see Fig. 4.13b).

Now, it remains to show that $A, C \notin B(\mathcal{G}_{m,n})$. Conversely, suppose $A \in B(\mathcal{G}_{m,n})$ or $C \in B(\mathcal{G}_{m,n})$. Without loss of generality we assume the former, which in turn implies that $\ell(AB)$ is a tangent to $\text{Conv}(\mathcal{G}_{m,n})$ as A is an interior point of BD and $B, D \in \mathcal{G}_{m,n}$. We will arrive to a contradiction by showing that $\ell(AB)$ separates B' and D'. For this we observe that neither $\overrightarrow{A'B'}$ nor $\overrightarrow{C'D'}$ belongs to $\ell(AB)$. Indeed, as by Theorem 65 $AC \subset BD$, the inclusion $A'B' \subset \ell(AB)$ would imply that $A'B'$ coincides either with AB or with CD, and the inclusion $C'D' \subset \ell(AB)$ would imply that $C'D'$ coincides either with AB or with CD. In each of the cases we would have a contradiction with (4.4). This observation together with the fact that $A', C' \in M_1(f) \subseteq AC \subset \ell(AB)$ imply that neither B' nor D' belongs to $\ell(AB)$. Consequently, as $f_{\overrightarrow{AB}}$ takes different values in B' and D' we conclude that $\ell(AB)$ separates B' and D', as required. □

In the remainder of this section we will deal with functions with the unique true

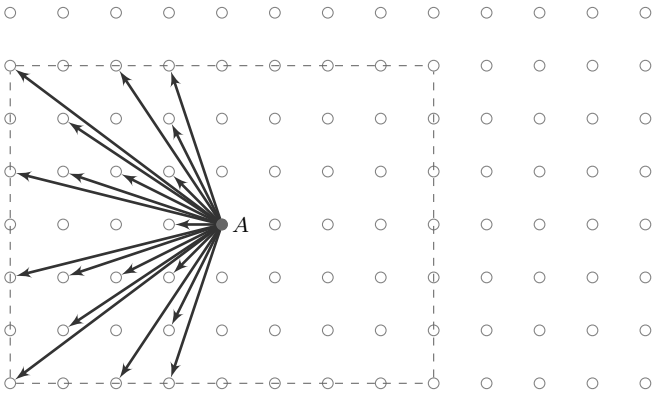


Figure 4.14: For $A = (4, 3)$, all the proper pairs of segments belong to the subgrid with the dashed boundary and A in the center. The possible choices of B are drawn on the left half of the subgrid.

point, i.e. *singleton-functions*. It is obvious that functions with the unique true point which coincides with one of the corners of the grid are threshold, hence we will not consider them. All other singleton-functions with the true point on the boundary of the grid were handled in Lemma 77. Thus we only need to look at the singleton functions with the true point in the interior of the grid.

Claim 79. *Let f be a $\{0,1\}$-valued function over $\mathcal{G}_{m,n}$ with a unique true point $A = (a_1, a_2)$ such that $a_1 \in \{1, \ldots, m-2\}$ and $a_2 \in \{1, \ldots, n-2\}$. Then f is a 2-threshold function, and the number of proper pairs of segments defining f is at most*

$$\frac{3}{\pi^2}mn + O(m \log n).$$

Proof. Let $M_1(f) = \{A\}$, without loss of generality we assume

$$a_1 \leq \frac{m-1}{2}, a_2 \leq \frac{n-1}{2}. \tag{4.5}$$

Let $\overrightarrow{AB}$ and $\overrightarrow{AD}$ be distinct prime segments. By Theorem 65, the pair $\{\overrightarrow{AB}, \overrightarrow{AD}\}$ is proper if and only if both segments belong to the same line. Hence, if $\{\overrightarrow{AB}, \overrightarrow{AD}\}$ is proper, then $d(\overrightarrow{AB}) = d(\overrightarrow{AD})$, and therefore all the considered pairs of segments belong to a subgrid of size $(2a_1 + 1) \times (2a_2 + 1)$. Next, we notice that for any given proper pair $\{\overrightarrow{AB}, \overrightarrow{AD}\}$ the points B and D are symmetric to each other with respect to A. Therefore it is enough to estimate the number of choices for B. Let $B = (b_1, b_2), D = (d_1, d_2)$. The only proper pair with $b_1 = d_1$ is the pair where $\{B, D\} = \{(a_1, a_2 + 1), (a_1, a_2 - 1)\}$, so we can exclude this case and assume $b_1 \neq d_1$. By symmetry, we may also assume $b_1 < d_1$. Putting all together and using a standard number-theoretical formula (5.17) (stated in the next chapter)

we derive the number of possible choices for B (see Fig. 4.14):

$$\sum_{b_1=0}^{a_1-1} \sum_{\substack{b_2=0 \\ (b_1-a_1)\perp(b_2-a_2)}}^{2a_2+1} 1 = \sum_{p=1}^{a_1} \sum_{\substack{q=-a_2 \\ p\perp q}}^{a_2+1} 1 = \frac{12}{\pi^2} a_1 a_2 + O(a_1 \log a_2),$$

where $p \perp q$ denotes that p and q are coprime. The target estimation follows from the latter by replacing a_1, a_2 with their upper bound (4.5). □

4.5 Conclusion

In this chapter we introduced the notion of proper pairs of segments and revealed the relation between them and proper 2-threshold functions. We proved that a 2-threshold function with true points on the boundary of the grid has a unique proper pair of segments that defines the function. The relationship between non-singleton 2-threshold functions with no true points on the boundary of the grid and proper pairs of segments remains unclear. From empirical observations we have the following conjectures regarding these functions depending on the shape of the convex hull of ones.

Conjecture 80. *Let f be a $\{0,1\}$-valued function over $\mathcal{G}_{m,n}$ such that* $\mathrm{Conv}(M_1(f))$ *is a segment with endpoints in the interior of the grid. Then f is a 2-threshold function with at most two different proper pairs of segments defining it.*

Conjecture 81. *Let f be a 2-threshold function over $\mathcal{G}_{m,n}$ such that* $\mathrm{Conv}(M_1(f))$ *has non-zero area and $M_1(f)$ is contained in the interior of the grid. Then f has a unique proper pair of segments defining it.*

This conjectures are based on the following observations:

1. All considered functions satisfying the conditions of any of the above conjectures have exactly one proper pair of segments which define them and do not belong to the same line.

2. Some of the functions satisfying Conjecture 80 also have a proper pair of segments which belong to the same line as the true points of the function. It is easy to see that a function f from the conjecture will have this additional proper pair of segments if and only if the line containing $\mathrm{Conv}(M_1(f))$ have common points with $\mathcal{G}_{m,n}$ in both directions from $\mathrm{Conv}(M_1(f))$. Fig. 4.13b illustrates the example of a function with such additional proper pair of segments.

Whether these conjectures hold or not, in the following chapter we will show that the proportion of the number of proper pairs of segments corresponding to the functions from the conjectures is negligible, in the sense that it does not affect the asymptotics of the number of 2-threshold functions.

Chapter 5

The asymptotics of the number of 2-threshold functions

5.1 Introduction

Denote by $t_k(m,n)$ the number of k-threshold functions over a two-dimensional rectangular grid $\mathcal{G}_{m,n} = \{0,1,\ldots,m-1\} \times \{0,1,\ldots,n-1\}$. Throughout the chapter we will write $t(m,n)$ instead of $t_1(m,n)$, as the former is a common notation in the literature. The asymptotics of the number of threshold functions for square grids was first obtained in [34]:

$$t(n,n) = \frac{6}{\pi^2} n^4 + O(n^3 \log n),$$

and for arbitrary rectangular grids in [1]:

$$t(m,n) = \frac{6}{\pi^2} m^2 n^2 + O(m^2 n \log n + m n^2 \log\log n),$$

where $m < n$ is assumed.

An improvement was found in [2]:

$$t(m,n) = \frac{6}{\pi^2} m^2 n^2 + O(m n^2 \log m),$$

see also [51]. The current best known formula was obtained in [25]:

$$t(m,n) = \frac{6}{\pi^2} m^2 n^2 + O(m n^2).$$

An important point to note here is that all the above results are based on the relation between non-constant threshold functions and (oriented) prime segments.

Based on the above estimation a trivial upper bound on the number of k-threshold functions for a fixed $k > 1$ is

$$t_k(m,n) \le \binom{t(m,n)}{k} = \frac{t(m,n)^k}{k!} + O\left(t(m,n)^{k-1}\right)$$
$$= \frac{6^k}{\pi^{2k}k!} m^{2k}n^{2k} + O\left(m^{2k-1}n^{2k}\right). \tag{5.1}$$

No asymptotics was known for the number of k-threshold functions for any $k > 1$. In this chapter we use the characterization of 2-threshold functions from the previous chapter to estimate the number of 2-threshold functions asymptotically. More specifically, the main result of the chapter is the following theorem.

Theorem 82.
$$t_2(m,n) = \frac{25}{12\pi^4} m^4n^4 + o(m^4n^4). \tag{5.2}$$

In Section 5.2 we show that almost all 2-threshold functions in $\mathcal{G}_{m,n}$ are in one-to-one correspondence with pairs of prime segments in convex position. To do this, we first establish a bijection between pairs of prime segments in convex position and proper pairs of segments $\overrightarrow{AB}, \overrightarrow{CD}$ such that Conv$(\{A,B,C,D\})$ is a quadrilateral. Second, we show that the latter objects are in one-to-one correspondence with almost all 2-threshold functions. Section 5.3 is devoted to the estimation of the number of pairs of prime segments in convex position. In Section 5.4 we use the obtained formula to improve the upper bound in (5.1) for $k > 2$.

5.2 From 2-threshold functions to pairs of segments in convex position

In this section we will reduce the estimation of the number of 2-threshold functions to the estimation of pairs of prime segments *in convex position*, i.e. pairs of segments that are opposite sides of a convex quadrilateral.

The following claim is a convenient necessary and sufficient condition for a pair of segments to be in convex position.

Claim 83. *Segments AB and CD are in convex position if and only if*

$$\begin{cases} \ell(AB) \cap CD = \emptyset, \\ \ell(CD) \cap AB = \emptyset. \end{cases} \tag{5.3}$$

Proof. Clearly, if AB and CD are in convex position, then (5.3) holds. To prove the converse, we observe that (5.3) implies that Conv$(\{A,B,C,D\})$ is not a segment or triangle, hence it is a convex quadrilateral with vertices A, B, C, and D. Moreover, $AB \cap CD = \emptyset$,

and hence the segments are neither diagonals nor adjacent edges, and consequently they are opposite edges of the quadrilateral Conv($\{A, B, C, D\}$). □

Let $q(m, n)$ be the total number of proper pairs of oriented segments in $\mathcal{G}_{m,n}$, and let $p(n, m)$ be the number of those of them, which are in convex position. It turns out, that $p(n, m)$ is asymptotically equal to the number of 2-threshold functions.

Theorem 84.

$$t_2(m, n) = p(m, n) + O\big(m^3n^3(m + n)\big). \tag{5.4}$$

Proof. The proof is split into two steps. First we prove that

$$t_2(m, n) = q(m, n) + O\big(m^2n^2(m + n)^2\big), \tag{5.5}$$

and then we show that

$$q(m, n) = p(m, n) + O\big(m^3n^3(m + n)\big). \tag{5.6}$$

The proof of (5.5) is based on the following two claims.

Claim 85. *Let $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ be a proper pair of segments in $\mathcal{G}_{m,n}$, and let $f = f_{\overrightarrow{AB}} \wedge f_{\overrightarrow{CD}}$ be the 2-threshold function defined by $\{\overrightarrow{AB}, \overrightarrow{CD}\}$. If f does not have true points on the boundary of the grid, i.e. $M_1(f) \subseteq \{1, \dots, m - 2\} \times \{1, \dots, n - 2\}$, then the distances $d(A, \ell(CD))$ and $d(B, \ell(CD))$ do not exceed one.*

Proof. The statement is obvious for $\ell(AB) = \ell(CD)$, so we assume that AB and CD are not collinear.

Let us first assume that $\ell(AB)$ and $\ell(CD)$ are not parallel and denote by O the intersection point of the two lines. We start by showing that there exists a point $X \in \ell(AB) \cap B(\mathcal{G}_{m,n})$ such that $AB \subseteq OX$. Indeed, since $f(A) = 1$, the point A is an interior point of Conv($\mathcal{G}_{m,n}$), and hence the line $\ell(AB)$ intersects $B(\mathcal{G}_{m,n})$ in exactly two points, which we denote by X and Y. Furthermore, as $\ell(CD)$ does not separate A and B, we have either $AB \subseteq OX$ or $AB \subseteq OY$. Without loss of generality assume $AB \subseteq OX$. Let $Z \in B(\mathcal{G}_{m,n})$ be the closest point to X such that $f_{\overrightarrow{AB}}(Z) = 1$. Clearly, $d(X, Z) \leq 1$. The assumption $M_1(f) \subseteq \{1, \dots, m - 2\} \times \{1, \dots, n - 2\}$ implies that $f(Z) = 0$, and therefore $f_{\overrightarrow{CD}}(Z) = 0$. Hence, either $Z \in \ell(CD)$ or the triangle $\overrightarrow{CDZ}$ is clockwise. The former implies that $d(X, \ell(CD)) \leq 1$. The latter leads to the same conclusion, if we notice that the triangle $\overrightarrow{CDX}$ is counterclockwise as X and A lie on the same side of $\ell(CD)$, and hence $\ell(CD)$ intersects XZ. Finally, since $A, B \in OX$, we conclude that $\max\{d(A, \ell(CD)), d(B, \ell(CD))\} \leq d(X, \ell(CD)) \leq 1$, as required.

The proof for parallel $\ell(AB)$ and $\ell(CD)$ is similar and uses the fact that the distance from any point of $\ell(AB)$ to $\ell(CD)$ is the same. □

Claim 86. *There are $O(m^2n^2(m+n)^2)$ proper pairs of segments $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ in $\mathcal{G}_{m,n}$ such that the 2-threshold function defined by $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ does not have a true point on the boundary of $\mathcal{G}_{m,n}$.*

Proof. There are at most mn ways to choose each of C and D. Given the segment CD, by Claim 85, each of A and B lies at distance at most one from $\ell(CD)$. Since there are $O(m+n)$ such points, we conclude that there are $O(m^2n^2(m+n)^2)$ desired pairs of segments. □

Let $t_2'(m,n)$ denote the number of proper 2-threshold functions over $\mathcal{G}_{m,n}$. Since $t_2(m,n) = t_2'(m,n) + t(m,n)$ and $t(m,n) = O(m^2n^2)$, to prove (5.5), it is enough to show that

$$t_2'(m,n) = q(m,n) + O\big(m^2n^2(m+n)^2\big). \tag{5.7}$$

For this, we first notice that, by Corollary 72, every proper pair of oriented segments in $\mathcal{G}_{m,n}$ defines a proper 2-threshold function. Furthermore, by Claim 86, only $O(m^2n^2(m+n)^2)$ of these pairs define 2-threshold functions with no true points on the boundary of $\mathcal{G}_{m,n}$. Finally, by Theorem 78, for any proper 2-threshold function that contains true points on the boundary of $\mathcal{G}_{m,n}$ there exists a *unique* proper pair of oriented segments in $\mathcal{G}_{m,n}$ that defines the function, and equation (5.7) follows.

To prove (5.6), we will show that the number of proper pairs of segments $\overrightarrow{AB}, \overrightarrow{CD}$ in $\mathcal{G}_{m,n}$ such that $\mathrm{Conv}(\{A,B,C,D\})$ is a segment or triangle is $O(m^3n^3(m+n))$. If $\mathrm{Conv}(\{A,B,C,D\})$ is a segment, then all of the four points A, B, C, and D lie on the same line. If $\mathrm{Conv}(\{A,B,C,D\})$ is a triangle, then, by Theorem 65, three of the points lie on the same line. In both cases there are three collinear points, say A, B, C. There are $O(m^2n^2)$ ways to choose two of these three points. Given two fixed points, there are at most $\max\{m-2, n-2\} = O(m+n)$ ways to choose the third one. For the fourth point, whether it lies on the same line with A, B, C or not, there are $O(mn)$ choices. Hence, altogether there are $O(m^3n^3(m+n))$ proper pairs of segments $\overrightarrow{AB}, \overrightarrow{CD}$ in $\mathcal{G}_{m,n}$ such that $\mathrm{Conv}(\{A,B,C,D\})$ is a segment or triangle, which implies (5.6). □

The relation between proper pairs of segments in convex position and pairs of non-oriented prime segments in convex position is revealed in the following theorem.

Theorem 87. *There is one-to-one correspondence between pairs of (non-oriented) prime segments in convex position and proper pairs of oriented segments in convex position.*

Proof. To prove the claim, we establish a bijective mapping between the two sets of pairs of segments. Clearly, if $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is a proper pair of segments in convex position, then AB and CD are prime and in convex position.

Now, let AB and CD be prime segments in convex position, then $\mathrm{Conv}(\{A,B,C,D\})$ is a quadrilateral, and AB and CD are two of its four edges. Assume,

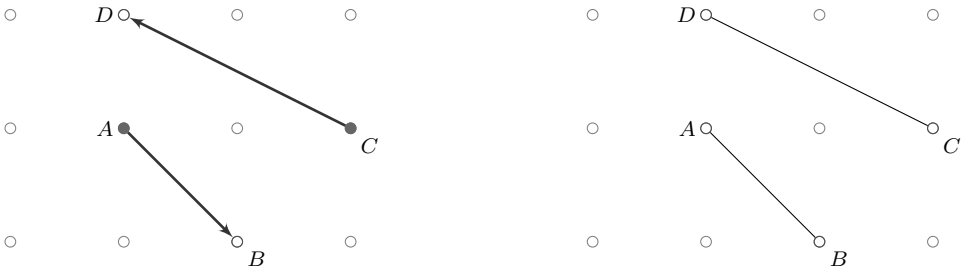


Figure 5.1: The proper pair of segments $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ in convex position and the corresponding pair of prime segments AB, CD in convex position.

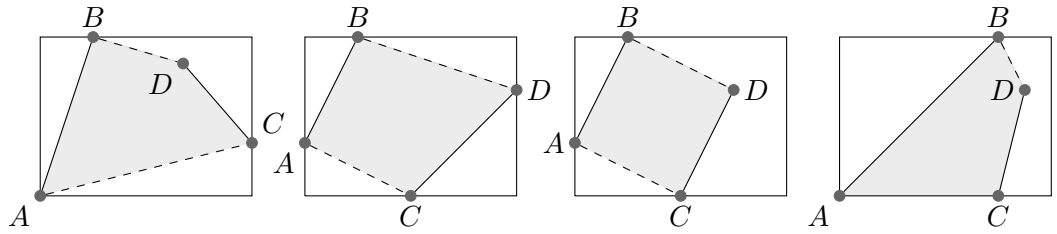


Figure 5.2: $\{AB, CD\}$ is a pair of segments in convex position. The grey shape is Conv$(\{A, B, C, D\})$. The rectangle is circumscribed about Conv$(\{A, B, C, D\})$ in (a) and (b) and not circumscribed in (c) and (d).

without loss of generality, that the other two edges are CD and DA (see Fig. 5.1). There are two oriented quadrilaterals corresponding to Conv$(\{A, B, C, D\})$, namely, $\overrightarrow{ABCD}$ and $\overrightarrow{DCBA}$, and these quadrilaterals have opposite orientations. Without loss of generality, we may assume that $\overrightarrow{ABCD}$ is the counterclockwise one, and hence, $\{\overrightarrow{AB}, \overrightarrow{CD}\}$ is a unique proper pair of segments in convex position corresponding to AB, CD. □

Due to the bijection established in Theorem 87, $p(m, n)$ denotes both the number of proper pairs of oriented segments and the number of pairs of prime segments in convex position.

5.3 The number of pairs of prime segments in convex position

In what follows we will extensively use rectangles with horizontal and vertical sides circumscribed about the convex quadrilaterals (see Fig. 5.2).

Denote by $\mathcal{R}_{u,v}$ a $u \times v$ rectangle Conv$(\mathcal{G}_{u+1,v+1})$ for natural numbers u and v. Denote by $Z(u, v)$ the set of pairs of prime segments $\{AB, CD\}$ in convex position such that $\mathcal{R}_{u,v}$ is circumscribed about Conv$(\{A, B, C, D\})$.

Theorem 88.

$$p(m,n) = \sum_{u=1}^{m}\sum_{v=1}^{n}(m-u)(n-v)|Z(u,v)|. \tag{5.8}$$

Proof. First for every convex quadrilateral with vertices in $\mathcal{G}_{m,n}$ there exists a unique rectangle with sides parallel to the sides of Conv($\mathcal{G}_{m,n}$) circumscribed about it. Hence, the statement follows from the fact that there are exactly $(m-u)(n-v)$ rectangles in $\mathcal{G}_{m,n}$ with sides of length u and v that are parallel to the sides of Conv($\mathcal{G}_{m,n}$). □

Let $Z_i(u,v) \subseteq Z(u,v)$ be the set of those pairs of segments AB, CD in $Z(u,v)$, for which exactly i points in $\{A,B,C,D\}$ are vertices of $\mathcal{R}_{u,v}$. Clearly, $Z(u,v)$ is the disjoint union of $Z_i(u,v), i = 0,1,2,3,4$, and therefore

$$|Z(u,v)| = \sum_{i=0}^{4}|Z_i(u,v)|. \tag{5.9}$$

Our next step is to estimate the cardinality of $Z_i(u,v)$ for every $i \in \{0,1,2,3,4\}$. The cases $i \in \{4,3\}$ are easy and we consider them below. The cases $i \in \{0,1,2\}$ are more involved and we treat them independently in Sections 5.3.2–5.3.4.

Lemma 89. $|Z_3(u,v)| + |Z_4(u,v)| = O(uv)$.

Proof. By definition, for any pair of segments $\{AB,CD\} \in Z_3(u,v) \cup Z_4(u,v)$ at least three of the endpoints of the segments are vertices of $\mathcal{R}_{u,v}$. Therefore, since there is a constant number of ways to map 3 of the endpoints of the segments to the vertices of $\mathcal{R}_{u,v}$, and there are $O(uv)$ ways to place the fourth point in $\mathcal{G}_{u+1,v+1}$, we conclude the lemma. □

5.3.1 Number theoretic preliminaries

In the subsequent sections we will use the following formulas. For the n-th harmonic number:

$$\sum_{i=1}^{n}\frac{1}{i} = \log n + \gamma + O\left(\frac{1}{n}\right) = \log n + O(1), \tag{5.10}$$

where γ is the Euler-Mascheroni constant.

For a fixed natural k the asymptotics of the sum of k-th powers can be estimated as

$$\sum_{i=1}^{n} i^k = \frac{n^{k+1}}{k+1} + O(n^k). \tag{5.11}$$

For a positive integer q the Euler function $\phi(q)$ is the number of positive integers that are coprime and less or equal to q. Some sums regarding the Euler function are as follows:

$$\sum_{x=1}^{n}\phi(x)\log(x) = \frac{3}{\pi^2}n^2\log n - \frac{3}{2\pi^2}n^2 + o(n^2). \tag{5.12}$$

The general formula for the power of x follows:

$$\sum_{x=1}^{n} \phi(x)x^k = \frac{6}{\pi^2}\frac{n^{k+2}}{(k+2)} + O(n^{k+1}\log n), \tag{5.13}$$

where k is integer.

Also, for a fixed natural q and integer $k \geq 0$ we have

$$\sum_{\substack{p=1 \\ p\perp q}}^{n} p^k = \frac{\phi(q)}{q}\frac{n^{k+1}}{(k+1)} + O(n^k 2^{w(q)}), \tag{5.14}$$

where $w(q)$ is the number of different prime divisors of q and

$$\sum_{q=1}^{n} O(2^{w(q)}) = O(n\log n). \tag{5.15}$$

For the negative powers of p we have

$$\sum_{\substack{p=1 \\ p\perp q}}^{n} \frac{1}{p} = \frac{\phi(q)}{q}\log n + \frac{\phi(q)}{q}\gamma + \frac{\phi(q)}{q}O\left(\frac{1}{n}\right) - \sum_{d|q}\mu(d)\frac{1}{d}\log d, \tag{5.16}$$

where $d|q$ means that d is a divisor of q.

More details about the derivations of the previous sums are provided in [26] and [7].

The following sum is obtained from (5.14) and (5.13):

$$\sum_{p=1}^{m}\sum_{\substack{q=1 \\ q\perp p}}^{n} 1 = \frac{6}{\pi^2}mn + O(m\log n). \tag{5.17}$$

The Möbius function μ_n is defined as

$$\mu_n \equiv \begin{cases} 0 & \text{if } n \text{ has one or more repeated prime factors} \\ 1 & \text{if } n = 1 \\ (-1)^k & \text{if } n \text{ is a product of } k \text{ distinct primes.} \end{cases}$$

For the Möbius function μ_n we have ([7]):

$$\sum_{n|k}\mu_n = \begin{cases} 1 & \text{if } k = 1 \\ 0 & \text{if } k > 1 \end{cases} \tag{5.18}$$

and

$$\sum_{n-1}^{\infty} \frac{\mu(n)}{n^2} = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}, \tag{5.19}$$

where $\zeta(n)$ is the Riemann zeta function.

5.3.2 The number of pairs of segments with two corner points

In this section we estimate $|Z_2(u,v)|$, i.e. the number of pairs of segments $\{AB, CD\}$ in $Z(u,v)$, for which exactly 2 points in $\{A,B,C,D\}$ are vertices of $\mathcal{R}_{u,v}$.

Let $\{X,Y\} = \{A,B,C,D\} \cap \text{Vert}(\mathcal{R}_{u,v})$. We consider the partition of $Z_2(u,v)$ into the following three subsets:

1. $Z_2^a(u,v)$ is the subset of $Z_2(u,v)$ such that X and Y are adjacent vertices of $\mathcal{R}_{u,v}$, i.e. XY is a side of $\mathcal{R}_{u,v}$.

2. $Z_2^b(u,v)$ is the subset of $Z_2(u,v)$ such that X and Y are opposite vertices of $\mathcal{R}_{u,v}$ and belong to the same segment.

3. $Z_2^c(u,v)$ is the subset of $Z_2(u,v)$ such that X and Y are opposite vertices of $\mathcal{R}_{u,v}$ and belong to the different segments.

Clearly,

$$|Z_2(u,v)| = |Z_2^a(u,v)| + |Z_2^b(u,v)| + |Z_2^c(u,v)|.$$

Let us show that the first summand does not affect the asymptotics of the sum which will be proved to be $\Theta(u^2v^2)$.

Lemma 90. $|Z_2^a(u,v)| = O(u^2v + uv^2)$.

Proof. Since $\mathcal{R}_{u,v}$ is circumscribed about $\text{Conv}(AB \cup CD)$ and two of the points A, B, C, D belong to the same side of $\mathcal{R}_{u,v}$, at least one of the other two points belongs to the opposite side of $\mathcal{R}_{u,v}$. Therefore there are $O(u+v)$ ways to place this point. Furthermore, there are $O(uv)$ ways to place the fourth point in $\mathcal{R}_{u,v}$, which implies the desired estimate. □

Lemma 91. *Let AB and CD be segments with endpoints in $\mathcal{G}_{m,n}$. Then AB and CD are in convex position if and only if A, B, C, and D are in general position, the triangle $\overrightarrow{ABD}$ has the same orientation as $\overrightarrow{ABC}$, and the triangle $\overrightarrow{CDA}$ has the same orientation as $\overrightarrow{CDB}$.*

Proof. We will prove the lemma by showing that its conditions are equivalent to those of Claim 83. First we claim that the equation $\ell(AB) \cap CD = \emptyset$ is equivalent to the statement that the points in both sets $\{A,B,C\}$ and $\{A,B,D\}$ are in general position and the orientations of $\overrightarrow{ABC}$ and $\overrightarrow{ABD}$ are the same. Indeed, $\overrightarrow{ABC}$ and $\overrightarrow{ABD}$ are triangles if and only if $C, D \notin \ell(AB)$. Moreover, the orientations of $\overrightarrow{ABC}$ and $\overrightarrow{ABD}$ are the same if and only if $\ell(AB)$ does not separate C and D.

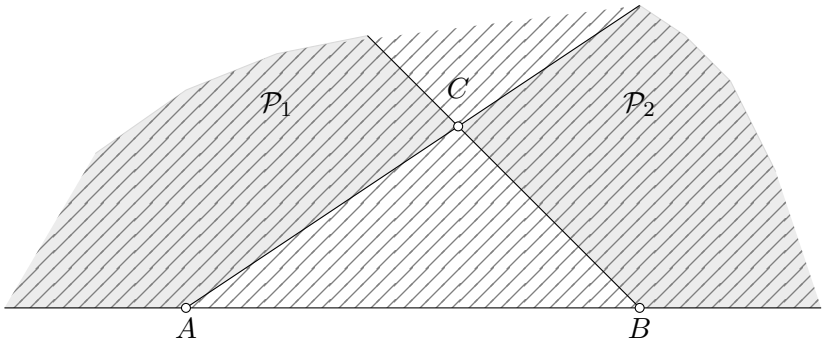


Figure 5.3: The stripped region is the area of points X such that the triangle $\overrightarrow{ABX}$ has the same orientation as $\overrightarrow{ABC}$. The region $\mathcal{P}_1$ is the set of points X such that $\overrightarrow{CXA}$, $\overrightarrow{CXB}$ are counterclockwise. The region $\mathcal{P}_2$ is the set of points X such that $\overrightarrow{CXA}$ and $\overrightarrow{CXB}$ are clockwise. The union $\mathcal{P}_1 \cup \mathcal{P}_2$ is the admissible region for D under fixed points A, B, and C.

Using similar arguments, one can establish the equivalence of the equation $\ell(CD) \cap AB = \emptyset$ and the statement that the points in both sets $\{C, D, A\}$ and $\{C, D, B\}$ are in general position and the orientations of triangles $\overrightarrow{CDA}$ and $\overrightarrow{CDB}$ are the same. □

We will employ Lemma 91 to describe the admissible region for the point D under fixed points A, B, C such that AB and CD are in convex position. Figure 5.3 illustrates the admissible region for D. It follows from Lemma 91 that for a segment AB and a point $C \notin \ell(AB)$ the segments AB and CD are in convex position if and only if D belongs to the interior of $\mathcal{P}_1 \cup \mathcal{P}_2$.

For a polygon $\mathcal{P}$, denote by $\mathcal{L}(\mathcal{P})$ the number of integer points in $\mathcal{P}$, i.e.

$$\mathcal{L}(P) = |\mathbb{Z}^2 \cap \mathcal{P}|.$$

For a polygon $\mathcal{P}$ and a point A denote by $Prime(\mathcal{P}, A)$ the number of integer points $X \in \mathcal{P}$ such that AX is a prime segment. If A is the origin $O = (0, 0)$ we simply write $Prime(\mathcal{P})$.

Lemma 92. *Let $\mathcal{R}_{u,v}$ be circumscribed about a triangle ABC. Then*

$$Prime(ABC, A) = \frac{6}{\pi^2}\text{Area}(ABC) + O(u + v). \tag{5.20}$$

Proof. Without loss of generality we assume that A coincides with the origin $O = (0, 0)$. Denote by $i \cdot ABC$ the triangle ABC scaled for a given factor $i > 0$, i.e.

$$i \cdot ABC \ = \ \left\{(i \cdot x, i \cdot y) \in \mathbb{Z}^2 | (x, y) \in ABC \right\}.$$

We start with the equation

$$\begin{aligned}\mathcal{L}(ABC) &= Prime\,(ABC) + Prime\left(\frac{1}{2}\cdot ABC\right) + Prime\left(\frac{1}{3}\cdot ABC\right) + \ldots \\ &= \sum_{j=1}^{u+v} Prime\left(\frac{1}{j}\cdot ABC\right)\end{aligned}$$

and the consequent one

$$\begin{aligned}\mathcal{L}\left(\frac{1}{p}\cdot ABC\right) &= \sum_{q=1}^{\infty} Prime\left(\frac{1}{p\cdot q}\cdot ABC\right) \\ &= \sum_{q=1}^{(u+v)/p} Prime\left(\frac{1}{p\cdot q}\cdot ABC\right).\end{aligned}$$

Using (5.18) we proceed with

$$\begin{aligned}Prime(ABC) &= \sum_{l=1}^{u+v} Prime\left(\frac{1}{l}\cdot ABC\right)\left(\sum_{k|l}\mu(k)\right) \\ &= \sum_{h=1}^{u+v}\mu(h)\sum_{i=1}^{(u+v)/h} Prime\left(\frac{1}{hi}\cdot ABC\right) \\ &= \sum_{h=1}^{u+v}\mu(h)\mathcal{L}\left(\frac{1}{h}\cdot ABC\right).\end{aligned}$$

From [16] we have

$$\mathcal{L}\,(ABC) = \text{Area}(ABC) + O\,(u+v)\,,$$

and hence

$$\begin{aligned}\sum_{h=1}^{u+v}\mu(h)\mathcal{L}\left(\frac{1}{h}\cdot ABC\right) &= \sum_{h=1}^{u+v}\mu(h)\frac{1}{h^2}\left(\text{Area}(ABC) + O\,(u+v)\right) \\ &= \text{Area}(ABC)\left(\sum_{h=1}^{\infty}\frac{\mu(h)}{h^2} - \sum_{h=u+v+1}^{\infty}\frac{\mu(h)}{h^2}\right) + O\,(u+v) \\ &\overset{(5.19)}{=} \frac{6}{\pi^2}\text{Area}(ABC) + O\left(\frac{\text{Area}(ABC)}{u+v} \;+\; u+v\right) \\ &= \frac{6}{\pi^2}\text{Area}(ABC) \;+\; O(u+v).\end{aligned}$$

□

Corollary 93. *Let $\mathcal{R}_{u,v}$ be circumscribed about a triangle ABC. Then the number of*

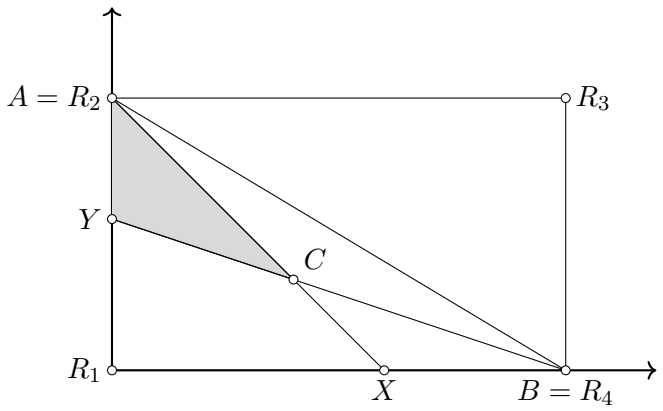


Figure 5.4: $C \in ABR_1$, the grey triangle is the admissible region for D

internal points X of ABC such that AX is a prime segment is

$$\frac{6}{\pi^2}\text{Area}(ABC) + O(u+v). \tag{5.21}$$

The lemma and the corollary above will be applied in the rest of the section and in Section 5.3.3 in the following way. Lemma 92 implies that the set of points D such that AB and CD are in convex position is contained in some *admissible region* $\mathcal{P}$, which is either a triangle with C being its vertex or a pair of triangles that have a unique common point, which is C and a vertex of each of them. In both cases we use Lemma 92 (or its corollary) to estimate the number of possible points D in $\mathcal{P}$ such that CD is a prime segment.

Lemma 94.

$$|Z_2^b(u,v)| = \begin{cases} \frac{1}{\pi^2}u^2v^2 + O(u^2v + uv^2) & \textit{if } u \perp v, \\ 0 & \textit{otherwise}. \end{cases}$$

Proof. First, we notice that for non-coprime u and v the diagonal of $\mathcal{R}_{u,v}$ is not a prime segment and the set $Z_2^b(u,v)$ is empty.

Let now u and v be coprime. Let us denote the vertices of $\mathcal{R}_{u,v}$ by R_1, R_2, R_3, and R_4 as in Fig. 5.4, and consider a pair $\{AB, CD\}$ from $Z_2^b(u,v)$. Without loss of generality we assume that AB is a diagonal of $\mathcal{R}_{u,v}$, i.e. either $AB = R_2R_4$ or $AB = R_1R_3$. Let us assume that $AB = R_2R_4$. Since, by definition, CD does not intersect AB, either $CD \in ABR_1$ or $CD \in ABR_3$. Let us assume that $CD \in ABR_1$ and let $C = (c_1, c_2)$, $D = (d_1, d_2)$. Clearly, $c_1 \neq d_1$ as otherwise $\ell(CD)$ would intersect AB. Without loss of generality we assume that $c_1 > d_1$. Let us denote

$$X = \ell(AC) \cap R_1B = \left(\frac{vc_1}{v - c_2}, 0\right)$$

and

$$Y = \ell(BC) \cap R_1 A = \left(0, \frac{uc_2}{u - c_1}\right).$$

It follows from Lemma 91 that AB and CD are in convex position if and only if D is an interior point of ACY. By Lemma 92, the number of choices for point D such that CD is a prime segment is

$$\frac{6}{\pi^2}\text{Area}(ACY) + o\left(c_1\left(v - c_2\right)\right) = \frac{6}{\pi^2}\text{Area}(ACY) + O(u+v).$$

Hence, summing up over all possible choices for the point C in ABR_1 and multiplying by 4 to take into account the cases of $C \in ABR_3$ and $AB = R_1R_3$ we derive

$$\begin{aligned}|Z_2^b(u,v)| &= 4\sum_{c_1=1}^{u-1}\sum_{c_2=1}^{\left\lfloor\frac{v(u-c_1)}{u}\right\rfloor}\left(\frac{6}{\pi^2}\text{Area}(ACY) + O(u+v)\right)\\ &= \frac{24}{\pi^2}\sum_{c_1=1}^{u-1}\sum_{c_2=1}^{\left\lfloor\frac{v(u-c_1)}{u}\right\rfloor}\text{Area}(ACY) + O(u^2v + uv^2),\end{aligned}$$

where

$$\text{Area}(ACY) = \frac{c_1 \cdot d(A,Y)}{2} = \frac{1}{2}c_1\left(v - \frac{uc_2}{u-c_1}\right) = \frac{1}{2}\left(vc_1 - \frac{uc_1c_2}{u-c_1}\right).$$

Therefore, we have

$$\begin{aligned}|Z_2^b(u,v)| &= \frac{12}{\pi^2}\sum_{c_1=1}^{u-1}\sum_{c_2=1}^{\left\lfloor\frac{v(u-c_1)}{u}\right\rfloor}\left(vc_1 - \frac{uc_1}{u-c_1}c_2\right) + O(u^2v+uv^2)\\ &\overset{(5.11)}{=} \frac{12}{\pi^2}\sum_{c_1=1}^{u-1}\left(\frac{v^2c_1(u-c_1)}{u} - \frac{uc_1}{u-c_1}\left(\frac{v^2(u-c_1)^2}{2u^2} + O\left(\frac{v(u-c_1)}{u}\right)\right)\right)\\ &+ O(u^2v+uv^2)\\ &= \frac{12}{\pi^2}\sum_{c_1=1}^{u-1}\left(\frac{v^2c_1(u-c_1)}{2u}\right) + O(u^2v+uv^2)\\ &= \frac{6v^2}{\pi^2}\sum_{c_1=1}^{u-1}\left(c_1 - \frac{c_1^2}{u}\right) + O(u^2v+uv^2)\\ &= \frac{6v^2}{\pi^2}\left(\frac{(u-1)^2}{2} - \frac{(u-1)^3}{3u} + O(u)\right) + O(u^2v+uv^2)\\ &= \frac{1}{\pi^2}u^2v^2 + O(u^2v+uv^2).\end{aligned}$$

□

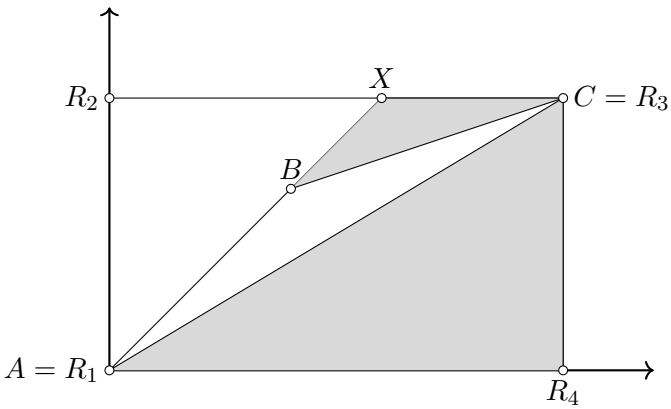


Figure 5.5: The point B belongs to AR_2R_3, the grey triangles form the admissible region for D.

Lemma 95. $|Z_2^c(u,v)| = \frac{42}{\pi^4}u^2v^2 + o(u^2v^2)$.

Proof. Consider a pair $\{AB, CD\}$ from $Z_2^c(u,v)$. Without loss of generality we assume $\{A,B,C,D\} \cap \text{Vert}(\mathcal{R}_{u,v}) = \{A,C\}$, that is either $\{A,C\} = \{R_1,R_3\}$ or $\{A,C\} = \{R_2,R_4\}$. The cases are symmetric, and hence it suffices to consider one of them, say $\{A,C\} = \{R_1,R_3\}$. Without loss of generality we assume $A = R_1$ and $C = R_3$ as in Fig. 5.5. The point $B = (b_1,b_2)$ belongs to one of the triangles ACR_2 and ACR_4. Due to symmetry, we assume without loss of generality $B \in ACR_2$, in which case we have $b_2 > \frac{vb_1}{u}$. Let us denote

$$X = \ell(AB) \cap CR_2 = \left(\frac{vb_1}{b_2}, v\right).$$

It follows from Lemma 91 that AB and CD are in convex position if and only if D is an interior point of $ACR_4 \cup BCX$ or an interior point of one of the segments CX, AR_4, or CR_4. By Lemma 92 and Corollary 93, the number of possible choices for D such that CD is a prime segment for a fixed B is

$$\frac{6}{\pi^2}\big(\text{Area}(ACR_4) + \text{Area}(BCX)\big) + O(u+v),$$

where

$$\text{Area}(ACR_4) = \frac{uv}{2}$$

and

$$\begin{aligned}\text{Area}(BCX) = \frac{(v-b_2)\cdot d(C,X)}{2} &= \frac{1}{2}(v-b_2)\left(u - \frac{vb_1}{b_2}\right)\\ &= \frac{1}{2}\left(uv + vb_1 - ub_2 - \frac{v^2b_1}{b_2}\right).\end{aligned}$$

Therefore, summing over all possible choices of B and multiplying by 4 to take into account the cases $B \in ACR_4$ and $\{A, C\} = \{R_2, R_4\}$ we derive:

$$\begin{aligned}|Z_2^c(u,v)| =& 4\sum_{b_1=1}^{u-1}\sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1\right\rfloor \\ b_2\perp b_1}}^{v} \frac{6}{\pi^2}(\text{Area}(ACR_4)+\text{Area}(BCX)+O(u+v))\\ =& \frac{12}{\pi^2}\sum_{b_1=1}^{u-1}\sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1\right\rfloor \\ b_2\perp b_1}}^{v}\left(2uv+vb_1-ub_2-\frac{v^2b_1}{b_2}\right)+O(u^2v+uv^2).\end{aligned} \tag{5.22}$$

We will estimate the asymptotics of different summands of (5.22) separately.

1. Estimation of $\sum\sum 2uv$.

Using formulas (5.14), (5.13), and (5.15) we obtain

$$\sum_{b_1=1}^{u-1}\sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1\right\rfloor \\ b_2\perp b_1}}^{v} 1 = \sum_{b_1=1}^{u-1}\left(v\frac{\phi(b_1)}{b_1}-\frac{v}{u}\phi(b_1)+O\left(2^{w(b_1)}\right)\right) \tag{5.23}$$

$$\begin{aligned}&= v\left(\frac{6}{\pi^2}u+O(\log u)\right)-\frac{v}{u}\left(\frac{3}{\pi^2}u^2+O(u\log u)\right)+O(u\log u)\\ &= \frac{3}{\pi^2}uv+O(v\log u)+O(u\log u).\end{aligned} \tag{5.24}$$

Changing the order of summation in the above sum, we deduce the same result, but with a slightly different error term:

$$\sum_{b_1=1}^{u-1}\sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1\right\rfloor \\ b_2\perp b_1}}^{v} 1 = \sum_{b_2=1}^{v}\sum_{\substack{b_1=1 \\ b_1\perp b_2}}^{\left\lceil \frac{ub_2}{v}-1\right\rceil} 1$$

$$\begin{aligned}&= \sum_{b_2=1}^{v}\left(\frac{\phi(b_2)}{b_2}\left\lceil\frac{ub_2}{v}-1\right\rceil+O\left(2^{w(b_2)}\right)\right)\\ &= \sum_{b_2=1}^{v}\left(\frac{u}{v}\phi(b_2)+O\left(\frac{\phi(b_2)}{b_2}\right)+O\left(2^{w(b_2)}\right)\right)\\ &= \frac{u}{v}\left(\frac{3}{\pi^2}v^2+O\left(v\log v\right)\right)+O(v)+O\left(v\log v\right)\\ &= \frac{3}{\pi^2}uv+O(u\log v)+O(v\log v).\end{aligned} \tag{5.25}$$

Finally, denoting $\alpha = \max(u,v)$ and $\beta = \min(u,v)$ we derive from (5.24) and (5.25)

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} 1 = \frac{3}{\pi^2}uv + O(\alpha \log \beta) + O(\beta \log \beta)$$

$$= \frac{3}{\pi^2}uv + O(u \log v + v \log u), \tag{5.26}$$

and hence

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} 2uv = \frac{6}{\pi^2}u^2v^2 + O(u^2 v \log v + uv^2 \log u). \tag{5.27}$$

2. Estimation of $\sum\sum vb_1$.

Using formulas (5.23) and (5.13) we obtain

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} b_1 = \sum_{b_1=1}^{u-1} b_1 \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} 1$$

$$= \sum_{b_1=1}^{u-1} \left(v\phi(b_1) - \frac{v}{u} b_1 \phi(b_1) + O\left(b_1 2^{w(b_1)} \right) \right)$$

$$= \frac{3}{\pi^2}u^2 v + O(uv \log u) - \frac{v}{u}\left(\frac{2}{\pi^2}u^3 + O(u^2 \log u) \right) + O(u^2 \log u)$$

$$= \frac{1}{\pi^2}u^2 v + O(uv \log u) + O(u^2 \log u). \tag{5.28}$$

Again, changing the order of summation in the above sum, we deduce the same result with a different error term:

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} b_1 = \sum_{b_2=1}^{v} \sum_{\substack{b_1=1 \\ b_1 \perp b_2}}^{\left\lceil \frac{ub_2}{v}-1 \right\rceil} b_1$$

$$= \sum_{b_2=1}^{v} \left(\frac{\phi(b_2)}{2b_2} \left\lceil \frac{ub_2}{v} - 1 \right\rceil^2 + \frac{ub_2}{v} O\left(2^{w(b_2)} \right) + O\left(2^{w(b_2)} \right) \right)$$

$$= \sum_{b_2=1}^{v} \left(\frac{u^2}{2v^2} b_2 \phi(b_2) + \frac{u}{2v}\phi(b_2) + \frac{ub_2}{v} O\left(2^{w(b_2)} \right) + O\left(2^{w(b_2)} \right) \right)$$

$$= \frac{1}{\pi^2}u^2 v + O(u^2 \log v) + O(uv \log v). \tag{5.29}$$

Finally, denoting $\alpha = \max(u,v)$ and $\beta = \min(u,v)$ we derive from (5.28) and

(5.29)

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} b_1 = \frac{1}{\pi^2} u^2 v + u\left(O(\alpha \log \beta) + O(\beta \log \beta)\right) = \frac{1}{\pi^2} u^2 v + o(u^2 v),$$

and hence

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} v b_1 = \frac{1}{\pi^2} u^2 v^2 + o(u^2 v^2) \tag{5.30}$$

3. Estimation of $\sum\sum ub_2$.

Using formulas (5.14), (5.15), and (5.13) we obtain:

$$\begin{aligned}
\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} b_2 &= \sum_{b_1=1}^{u-1} \left(\frac{\phi(b_1) v^2}{2b_1} - \frac{\phi(b_1) b_1 v^2}{2u^2} + O\left(v 2^{w(b_1)}\right) \right) \\
&= \frac{v^2}{2} \sum_{b_1=1}^{u-1} \left(\frac{\phi(b_1)}{b_1} - \frac{1}{u^2} b_1 \phi(b_1) \right) + O(uv \log u) \\
&= \frac{v^2}{2} \left(\frac{6}{\pi^2} u + O(\log u) - \frac{1}{u^2} \left(\frac{2}{\pi^2} u^3 + O(u^2 \log u) \right) \right) + O(uv \log u) \\
&= \frac{2}{\pi^2} uv^2 + O(uv \log u) + O(v^2 \log u). \qquad (5.31)
\end{aligned}$$

Similarly to the previous case, by changing the order of summation, one can show that

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} b_2 = \frac{2}{\pi^2} uv^2 + O(uv \log v) + O(v^2 \log v),$$

which together with (5.31) imply

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} b_2 = \frac{2}{\pi^2} uv^2 + o(uv^2),$$

and hence

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} ub_2 = \frac{2}{\pi^2} u^2 v^2 + o(u^2 v^2). \tag{5.32}$$

4. Estimation of $\sum\sum \frac{v^2 b_1}{b_2}$.
Using formulas (5.16), (5.12), (5.13), and the fact that $\log \lfloor x \rfloor = \log x + O\left(\frac{1}{x}\right)$ for $x \geq 1$, we obtain

$$\begin{aligned}
&\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} \frac{b_1}{b_2} = \sum_{b_1=1}^{u-1} b_1 \left(\sum_{\substack{b_2=1 \\ b_2 \perp b_1}}^{v} \frac{1}{b_2} - \sum_{\substack{b_2=1 \\ b_2 \perp b_1}}^{\left\lfloor \frac{vb_1}{u} \right\rfloor} \frac{1}{b_2} \right) \\
&= \sum_{b_1=1}^{u-1} b_1 \left(\frac{\phi(b_1)}{b_1} \log v + \frac{\phi(b_1)}{b_1} O\left(\frac{1}{v}\right) - \frac{\phi(b_1)}{b_1} \log \left\lfloor \frac{vb_1}{u} \right\rfloor - \frac{\phi(b_1)}{b_1} O\left(\frac{u}{vb_1}\right) \right) \\
&= \sum_{b_1=1}^{u-1} \phi(b_1) \left(\log v - \log \frac{vb_1}{u} + O\left(\frac{1}{v}\right) + O\left(\frac{u}{vb_1}\right) \right) \\
&= \sum_{b_1=1}^{u-1} \phi(b_1) \left(\log u - \log b_1 + O\left(\frac{1}{v}\right) + O\left(\frac{u}{vb_1}\right) \right) \\
&= \frac{3}{\pi^2} u^2 \log u + O(u \log^2 u) - \frac{3}{\pi^2} u^2 \log u + \frac{3}{2\pi^2} u^2 + o(u^2) + O\left(\frac{u^2}{v}\right) \\
&= \frac{3}{2\pi^2} u^2 + o(u^2) + O\left(\frac{u^2}{v}\right),
\end{aligned} \tag{5.33}$$

and hence

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v} \frac{v^2 b_1}{b_2} = \frac{3}{2\pi^2} u^2 v^2 + o(u^2 v^2). \tag{5.34}$$

Finally, combining (5.22), (5.27), (5.30), (5.32), and (5.34) we derive

$$\begin{aligned}
|Z_2^c(u,v)| &= \frac{12}{\pi^2} \left(\frac{6}{\pi^2} u^2 v^2 + \frac{1}{\pi^2} u^2 v^2 - \frac{2}{\pi^2} u^2 v^2 - \frac{3}{2\pi^2} u^2 v^2 + o(u^2 v^2) \right) \\
&= \frac{42}{\pi^4} u^2 v^2 + o(u^2 v^2).
\end{aligned}$$

□

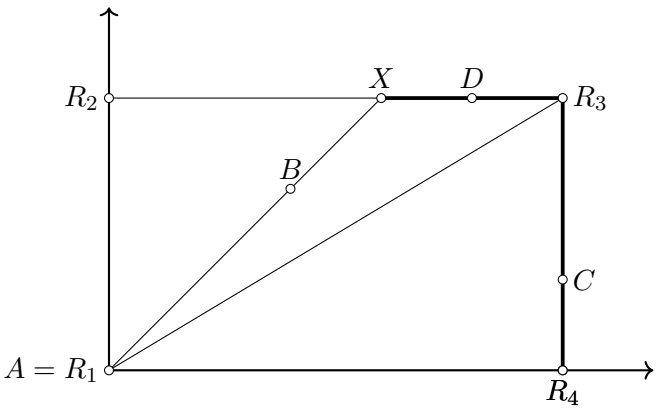


Figure 5.6: The case $B \in AR_2R_3$. The points C and D belong to the segments R_3R_4 and XR_3 respectively, where $X = \ell(AB) \cap R_2R_3$.

5.3.3 The number of pairs of segments with one corner point

There is no loss of generality in assuming $\{A, B, C, D\} \cap \text{Vert}(\mathcal{R}_{u,v}) = \{A\}$ for every pair $\{AB, CD\} \in Z_1(u, v)$. We consider the partition of $Z_1(u, v)$ into the following two subsets:

1. $Z_1^a(u, v)$ the set of those pairs $\{AB, CD\}$ in which the point B is an interior point of $\mathcal{R}_{u,v}$;

2. $Z_1^b(u, v)$ the set of those pairs $\{AB, CD\}$ in which the point B belongs to the boundary of $\mathcal{R}_{u,v}$.

In the rest of the section we estimate the sizes of these sets in separate lemmas.

Lemma 96. $|Z_1^a(u, v)| = \frac{72}{\pi^4}u^2v^2 + o(u^2v^2)$.

Proof. Due to symmetry, for a corner point R of $\mathcal{R}_{u,v}$ the number of pairs $\{AB, CD\} \in Z_1^a(u, v)$, where A coincides with R, is the same for every $R \in \{R_1, R_2, R_3, R_4\}$. Therefore, it is enough to estimate the number of pairs where A coincides with a fixed corner point of $\mathcal{R}_{u,v}$, and we assume that $A = R_1$.

Since B is an interior point of $\mathcal{R}_{u,v}$ and neither C nor D is a corner point of $\mathcal{R}_{u,v}$, we conclude that one of C and D belongs to the interior of R_2R_3 and the other belongs to the interior of R_3R_4. Without loss of generality, we assume that C is an interior point of R_3R_4 and D is an interior point of R_2R_3.

Under the above assumptions, we will first estimate the number of pairs in $Z_1^a(u, v)$ in which $B = (b_1, b_2)$ belongs to the triangle AR_2R_3. Notice that the latter assumption is equivalent to the inequality $\frac{b_1}{b_2} \leq \frac{u}{v}$. Let us denote

$$X = \ell(AB) \cap R_2R_3 = \left(\frac{vb_1}{b_2}, v\right).$$

It follows from Lemma 91 that AB and CD are in convex position if and only if D is an interior point of XR_3 (see Fig. 5.6). Therefore, by denoting $D = (d_1, v)$ and $C = (u, c_2)$, the number of desired prime pairs segments can be expressed as

$$\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1 \right\rfloor \\ b_2 \perp b_1}}^{v-1} \sum_{c_2=1}^{v-1} \sum_{\substack{d_1=\left\lfloor \frac{vb_1}{b_2}+1 \right\rfloor \\ (u-d_1)\perp(v-c_2)}}^{u-1} 1. \tag{5.35}$$

We start by estimating the contribution of the latter two sums.

$$\begin{aligned}
\sum_{c_2=1}^{v-1} \sum_{\substack{d_1=\left\lfloor \frac{vb_1}{b_2}+1 \right\rfloor \\ (u-d_1)\perp(v-c_2)}}^{u-1} 1 &= \sum_{c_2'=1}^{v} \sum_{\substack{d_1'=1 \\ d_1' \perp (c_2')}}^{u-\left\lfloor \frac{vb_1}{b_2}+1 \right\rfloor} 1 + O(u) \qquad (5.36)\\
&\overset{(5.14)}{=} \sum_{c_2'=1}^{v} \left(\frac{\phi(c_2')}{c_2'} \left(u - v\frac{b_1}{b_2} + O(1) \right) + O\left(2^{w(c_2')} \right) \right) + O(u) \\
&\overset{(5.13)}{=} \left(\frac{6v}{\pi^2} + O(\log v) \right) \left(u - v\frac{b_1}{b_2} + O(1) \right) + O(v \log v) + O(u) \\
&= \frac{6v}{\pi^2} \left(u - v\frac{b_1}{b_2} \right) + O(u \log v) + O\left(\frac{b_1}{b_2} v \log v \right) + O(v \log v) \\
&= \frac{6v}{\pi^2} \left(u - v\frac{b_1}{b_2} \right) + O(v \log v) + O(u \log v).
\end{aligned}$$

By changing the order of summation in (5.36), one can show that

$$\sum_{c_2=1}^{v-1} \sum_{\substack{d_1=\left\lfloor \frac{vb_1}{b_2}+1 \right\rfloor \\ (u-d_1)\perp(v-c_2)}}^{u-1} 1 = \frac{6v}{\pi^2} \left(u - v\frac{b_1}{b_2} \right) + O(u \log v + v \log u).$$

Now, plugging in the above result to (5.35) and using formulas (5.26) and (5.33) we obtain:

$$\begin{aligned}
&\sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1\right\rfloor \\ b_2\perp b_1}}^{v-1} \sum_{c_2=1}^{v-1} \sum_{\substack{d_1=\left\lfloor \frac{vb_1}{b_2}+1\right\rfloor \\ (u-d_1)\perp(v-c_2)}}^{u-1} 1 \\
&= \sum_{b_1=1}^{u-1} \sum_{\substack{b_2=\left\lfloor \frac{vb_1}{u}+1\right\rfloor \\ b_2\perp b_1}}^{v-1} \left(\frac{6v}{\pi^2}\left(u - v\frac{b_1}{b_2}\right) + O(u\log v + v\log u)\right) \\
&= \frac{6v}{\pi^2}\left(u\left(\frac{3}{\pi^2}uv + O(u\log v + v\log u)\right) - v\left(\frac{3}{2\pi^2}u^2 + o(u^2) + O\left(\frac{u^2}{v}\right)\right)\right) \\
&+ O(u^2 v\log v + uv^2\log u) \\
&= \frac{9}{\pi^4}u^2v^2 + v^2 o(u^2) + O(u^2 v\log v + uv^2\log u).
\end{aligned}$$

Note that the obtained estimation is symmetric with respect to u and v, which implies that the number of pairs in $Z_1^a(u,v)$ in which B belongs to AR_3R_4 has the same asymptotics. Therefore, taking into account additionally all symmetric cases corresponding to the location of A, we finally conclude that

$$|Z_1^a(u,v)| = \frac{72}{\pi^4}u^2v^2 + o(u^2v^2).$$

□

Lemma 97.

$$|Z_1^b(u,v)| = \frac{6v^2}{\pi^2}\sum_{\substack{b_1=1 \\ b_1\perp v}}^{u}(2u-b_1) + \frac{6u^2}{\pi^2}\sum_{\substack{c_2=1 \\ c_2\perp u}}^{v}(2v-c_2) + O(u^2v + uv^2).$$

Proof. As in the proof of Lemma 96, due to symmetry, for a corner point R of $\mathcal{R}_{u,v}$ the number of pairs $\{AB, CD\} \in Z_1^b(u,v)$, where A coincides with R, is the same for every $R \in \{R_1, R_2, R_3, R_4\}$. Therefore, it is enough to estimate the number of pairs where A coincides with a fixed corner point of $\mathcal{R}_{u,v}$, and we assume that $A = R_1$.

It is easy to see that if B is an internal point of R_1R_2 or R_1R_4, then one of C and D belongs to the interior of R_2R_3 and the other belongs to the interior of R_3R_4. Therefore, taking into account primality of AB, the number of pairs, in which B is an internal point of R_1R_2 or R_1R_4, is $O(uv)$. The latter does not affect the asymptotics, and without loss of generality we assume from now on that B is an internal point of one of the sides R_2R_3 and R_3R_4.

Suppose first that B is an internal point of R_2R_3, i.e. $B = (b_1, v)$ for some $0 < b_1 < u$, and $b_1 \perp v$. Then $AB \cap R_3R_4 = \emptyset$, and hence $CD \cap R_3R_4 \neq \emptyset$, which implies

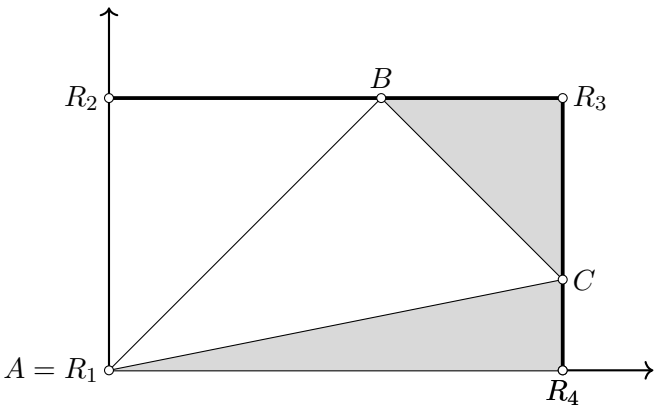


Figure 5.7: The points B and C belong to the segments R_2R_3 and R_3R_4 respectively, the grey triangles form the admissible area for D.

that either C or D belongs to the interior of R_3R_4. Without loss of generality we assume the former, i.e. $C = (u, c_2)$ for some $0 < c_2 < v$ (see Fig. 5.7). Under these assumptions, Lemma 91 implies that AB and CD are in convex position if and only if the point D belongs to $BCR_3 \setminus (BC \cup \{R_3\})$ or $ACR_4 \setminus (AC \cup \{R_4\})$. Therefore, using Lemma 92 and Corollary 93, we conclude that the number of such pairs of prime segments is

$$\begin{aligned}
&\sum_{\substack{b_1=1\\ b_1 \perp v}}^{u-1} \sum_{c_2=1}^{v-1} \frac{6}{\pi^2} \left(\text{Area}(BCR_3) + \text{Area}(ACR_4) + O(u+v)\right) \\
&= \frac{3}{\pi^2} \sum_{\substack{b_1=1\\ b_1 \perp v}}^{u} \sum_{c_2=1}^{v} ((u-b_1)(v-c_2) + uc_2) + O(u^2v + uv^2) \\
&= \frac{3}{\pi^2} \sum_{\substack{b_1=1\\ b_1 \perp v}}^{u} \sum_{c_2=1}^{v} (uv - vb_1 + b_1c_2) + O(u^2v + uv^2) \\
&= \frac{3}{\pi^2} \sum_{\substack{b_1=1\\ b_1 \perp v}}^{u} \left((uv - vb_1)v + b_1 \left(\frac{v^2}{2} + O(v)\right)\right) + O(u^2v + uv^2) \\
&= \frac{3v^2}{\pi^2} \sum_{\substack{b_1=1\\ b_1 \perp v}}^{u} \left(u - \frac{b_1}{2}\right) + O(u^2v + uv^2).
\end{aligned}$$

By symmetry, the number of pairs in which B is an internal point of R_3R_4 is

$$\frac{3u^2}{\pi^2} \sum_{\substack{c_2=1\\ c_2 \perp u}}^{v} \left(v - \frac{c_2}{2}\right) + O(u^2v + uv^2).$$

Putting all together and taking into account the symmetric cases corresponding to

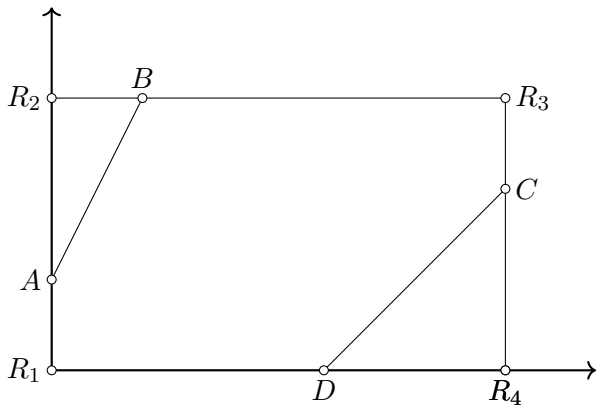


Figure 5.8: Each of A, B, C, and D belongs to a unique side of $\mathcal{R}_{u,v}$. The endpoints of the same segment belong to the adjacent sides of $\mathcal{R}_{u,v}$.

the location of A, we finally conclude that

$$|Z_1^b(u,v)| = \frac{6v^2}{\pi^2} \sum_{\substack{b_1=1 \\ b_1 \perp v}}^{u} (2u - b_1) + \frac{6u^2}{\pi^2} \sum_{\substack{c_2=1 \\ c_2 \perp u}}^{v} (2v - c_2) + O(u^2 v + uv^2).$$

□

We note that in Lemma 97 we deliberately did not compute a closed-form asymptotic, as the obtained formula will be crucial later to obtain a better error term.

5.3.4 The number of pairs of segments with no corner points

In this section we estimate the size of $Z_0(u,v)$, i.e. the number of those pairs of segments in $Z(u,v)$ none of whose endpoints is a corner of $\mathcal{R}_{u,v}$.

Lemma 98. $|Z_0(u,v)| = \frac{72}{\pi^4} u^2 v^2 + O(u^2 v \log v)$.

Proof. Let $\{AB, CD\}$ be an arbitrary pair in $Z_0(u,v)$. The fact that none of the points A, B, C, and D is a corner of $\mathcal{R}_{u,v}$ implies that each of the sides of $\mathcal{R}_{u,v}$ contains exactly one of these points. Furthermore, since AB and CD are in convex position, we conclude that the endpoints of the same segment belong to the adjacent sides of $\mathcal{R}_{u,v}$. Therefore, without loss of generality we can assume $A \in R_1R_2$ and $C \in R_3R_4$, in which case either $B \in R_2R_3$ and $D \in R_1R_4$, or $B \in R_1R_4$ and $D \in R_2R_3$. The two cases are symmetric and we assume the former one, i.e. $B \in R_2R_3, D \in R_1R_4$ (see Fig. 5.8). Let us denote $A = (0, a_2), C = (u, c_2), B = (b_1, v)$, and $D = (d_1, 0)$.

Under the above assumptions the segments AB and CD are prime if and only if

$(v-a_2) \perp b_1$ and $(u-d_1) \perp c_2$, and the number of such pairs is

$$\sum_{a_2=1}^{v-1} \sum_{\substack{b_1=1 \\ b_1 \perp (v-a_2)}}^{u-1} \sum_{c_2=1}^{v-1} \sum_{\substack{d_1=1 \\ (u-d_1) \perp c_2}}^{u-1} 1 = \sum_{a_2'=1}^{v-1} \sum_{\substack{b_1=1 \\ b_1 \perp a_2'}}^{u-1} \sum_{c_2=1}^{v-1} \sum_{\substack{d_1'=1 \\ d_1' \perp c_2}}^{u-1} 1.$$

Using formula (5.17) we obtain

$$\sum_{a_2'=1}^{v-1} \sum_{\substack{b_1=1 \\ b_1 \perp a_2'}}^{u-1} \sum_{c_2=1}^{v-1} \sum_{\substack{d_1'=1 \\ d_1' \perp c_2}}^{u-1} 1 = \sum_{a_2'=1}^{v-1} \sum_{\substack{b_1=1 \\ b_1 \perp a_2'}}^{u-1} \left(\frac{6}{\pi^2} uv + O(u \log v) \right) = \frac{36}{\pi^4} u^2 v^2 + O(u^2 v \log v).$$

Finally, taking into account the symmetric case of $B \in R_1 R_4$ and $D \in R_2 R_3$, we derive the desired result

$$|Z_0(u,v)| = \frac{72}{\pi^4} u^2 v^2 + O(u^2 v \log v).$$

□

5.3.5 Summarizing results

In the following theorem we prove the main result of the chapter by putting everything together.

Theorem 99.

$$p(m,n) = \frac{25}{12\pi^4} m^4 n^4 + o(m^4 n^4).$$

Proof. First, using (5.9) and Lemmas 89 and 90, we expand formula (5.8) as follows:

$$\begin{aligned}
p(m,n) &= \sum_{u=1}^{m-1} (m-u) \sum_{v=1}^{n-1} (n-v) \cdot |Z(u,v)| \\
&= \sum_{u=1}^{m-1} (m-u) \sum_{v=1}^{n-1} (n-v) \cdot \Big(|Z_4(u,v)| + |Z_3(u,v)| + |Z_2(u,v)| \\
&\qquad + |Z_1(u,v)| + |Z_0(u,v)| \Big) \\
&= \sum_{u=1}^{m-1} (m-u) \sum_{v=1}^{n-1} (n-v) \cdot \Big(|Z_2^b(u,v)| + |Z_2^c(u,v)| + |Z_1^a(u,v)| \\
&\qquad + |Z_1^b(u,v)| + |Z_0(u,v)| \Big) + o(m^4 n^4).
\end{aligned}$$

Next, we proceed by estimating different parts of the above sum separately.

1. Estimation of $\sum(m-u)\sum(n-v)\cdot|Z_2^b(u,v)|$.
Using Lemma 94 and formulas (5.14), (5.15), and (5.13), we obtain

$$\begin{aligned}
&\sum_{u=1}^{m-1}(m-u)\sum_{v=1}^{n-1}(n-v)\cdot|Z_2^b(u,v)| \\
&=\sum_{u=1}^{m-1}(m-u)\sum_{\substack{v=1\\ v\perp u}}^{n-1}(n-v)\cdot|Z_2^b(u,v)| \\
&=\sum_{u=1}^{m-1}(m-u)\sum_{\substack{v=1\\ v\perp u}}^{n-1}(n-v)\left(\frac{1}{\pi^2}u^2v^2+O(u^2v+uv^2)\right) \qquad (5.37)\\
&=\frac{1}{\pi^2}\sum_{u=1}^{m}(mu^2-u^3)\sum_{\substack{v=1\\ v\perp u}}^{n}(nv^2-v^3)+O(m^4n^3+m^3n^4) \\
&=\frac{1}{\pi^2}\sum_{u=1}^{m}(mu^2-u^3)\left(\frac{\phi(u)}{3u}n^4-\frac{\phi(u)}{4u}n^4+O\left(n^3 2^{w(u)}\right)\right)+O(m^4n^3+m^3n^4) \\
&=\frac{1}{12\pi^2}\sum_{u=1}^{m}\left(mn^4u\phi(u)-n^4u^2\phi(u)+O\left(mn^3u^2 2^{w(u)}\right)\right)+O(m^4n^3+m^3n^4) \\
&=\frac{1}{12\pi^2}\left(\frac{2m^4n^4}{\pi^2}-\frac{3m^4n^4}{2\pi^2}+O\left(m^4n^3\log m\right)\right)+O(m^4n^3+m^3n^4) \\
&=\frac{1}{24\pi^4}m^4n^4+O(m^4n^3\log m+m^3n^4).
\end{aligned}$$

Now, symmetry of formula (5.37) implies also the estimation with a symmetric error term

$$\sum_{u=1}^{m-1}(m-u)\sum_{v=1}^{n-1}(n-v)\cdot|Z_2^b(u,v)|=\frac{1}{24\pi^4}m^4n^4+O\left(m^3n^4\log n+m^4n^3\right).$$

Finally, comparing the two estimations one can derive

$$\sum_{u=1}^{m-1}(m-u)\sum_{v=1}^{n-1}(n-v)\cdot|Z_2^b(u,v)|=\frac{1}{24\pi^4}m^4n^4+o(m^4n^4). \qquad (5.38)$$

2. Estimation of $\sum(m-u)\sum(n-v)\Big(|Z_2^c(u,v)|+|Z_1^a(u,v)|+|Z_0(u,v)|\Big)$.
Using Lemmas 95, 96, 98, and formula (5.11), we obtain

$$\begin{aligned}
&\sum_{u=1}^{m-1}(m-u)\sum_{v=1}^{n-1}(n-v)\Big(|Z_2^c(u,v)|+|Z_1^a(u,v)|+|Z_0(u,v)|\Big)\\
&=\sum_{u=1}^{m-1}(m-u)\sum_{v=1}^{n-1}(n-v)\left(\frac{42}{\pi^4}u^2v^2+\frac{72}{\pi^4}u^2v^2+\frac{72}{\pi^4}u^2v^2+o(u^2v^2)\right)\\
&=\frac{186}{\pi^4}\sum_{u=1}^{m}(m-u)u^2\sum_{v=1}^{n}(n-v)v^2+o(m^4n^4)\\
&=\frac{186}{\pi^4}\sum_{u=1}^{m}(m-u)u^2\left(\frac{n^4}{3}-\frac{n^4}{4}+O(n^3)\right)+o(m^4n^4)\\
&=\frac{31}{2\pi^4}n^4\sum_{u=1}^{m}(m-u)u^2+o(m^4n^4)=\frac{31}{24\pi^4}m^4n^4+o(m^4n^4).
\end{aligned}\tag{5.39}$$

3. Estimation of $\sum(m-u)\sum(n-v)\cdot|Z_1^b(u,v)|$.
Using Lemma 97 we derive

$$\begin{aligned}
&\sum_{u=1}^{m-1}(m-u)\sum_{v=1}^{n-1}(n-v)\cdot|Z_1^b(u,v)|\\
&=\sum_{u=1}^{m-1}(m-u)\sum_{v=1}^{n-1}(n-v)\left(\frac{6v^2}{\pi^2}\sum_{\substack{b_1=1\\ b_1\perp v}}^{u}(2u-b_1)+\frac{6u^2}{\pi^2}\sum_{\substack{c_2=1\\ c_2\perp u}}^{v}(2v-c_2)+O(u^2v+uv^2)\right)\\
&=\frac{6}{\pi^2}\sum_{u=1}^{m}(m-u)\sum_{v=1}^{n}(n-v)v^2\sum_{\substack{b_1=1\\ b_1\perp v}}^{u}(2u-b_1)\\
&+\frac{6}{\pi^2}\sum_{v=1}^{n}(n-v)\sum_{u=1}^{m}(m-u)u^2\sum_{\substack{c_2=1\\ c_2\perp u}}^{v}(2v-c_2)+o(m^4n^4).
\end{aligned}\tag{5.40}$$

We notice that the first of the summands in the latter formula is obtained from the second one by swapping u with v, b_1 with c_2, and m with n, hence it suffices to find a closed-form estimation only for one of them, say for the first one.

Using formula (5.14) we obtain

$$\sum_{v=1}^{n}(n-v)v^2\sum_{\substack{b_1=1\\ b_1\perp v}}^{u}(2u-b_1)=\sum_{v=1}^{n}(n-v)v^2\left(2\frac{\phi(v)}{v}u^2-\frac{\phi(v)}{2v}u^2+O\left(u2^{w(v)}\right)\right)$$
$$=\frac{3}{2}\sum_{v=1}^{n}\left(u^2nv\phi(v)-u^2v^2\phi(v)+O\left(unv^2 2^{w(v)}\right)\right)$$
$$=\frac{3}{2}\left(u^2n\frac{2}{\pi^2}n^3-u^2\frac{3}{2\pi^2}n^4\right)+O(u^2n^3\log n)+O\left(un^4\log n\right)$$
$$=\frac{3}{4\pi^2}u^2n^4+n^2\left(O(u^2n\log n)+O\left(un^2\log n\right)\right). \tag{5.41}$$

By changing the order of summation in the above sum, we deduce the same result with a different error term:

$$\sum_{v=1}^{n}(n-v)v^2\sum_{\substack{b_1=1\\ b_1\perp v}}^{u}(2u-b_1)=\sum_{b_1=1}^{u}(2u-b_1)\sum_{\substack{v=1\\ v\perp b_1}}^{n}(n-v)v^2$$
$$=\sum_{b_1=1}^{u}(2u-b_1)\left(\frac{\phi(b_1)}{3b_1}n^4-\frac{\phi(b_1)}{4b_1}n^4+O\left(n^3 2^{w(b_1)}\right)\right)$$
$$=\frac{1}{12}\sum_{b_1=1}^{u}\left(2un^4\frac{\phi(b_1)}{b_1}-n^4\phi(b_1)+O\left(n^3u2^{w(b_1)}\right)\right)$$
$$=\frac{1}{12}\left(2un^4\frac{6}{\pi^2}u+O(un^4\log u)-n^4\frac{3}{\pi^2}u^2+O(un^4\log u)+O(n^3u^2\log u)\right)$$
$$=\frac{3}{4\pi^2}u^2n^4+n^2\left(O(un^2\log u)+O\left(nu^2\log u\right)\right). \tag{5.42}$$

Comparing the error terms in (5.41) and (5.42) we obtain

$$\sum_{v=1}^{n}(n-v)v^2\sum_{\substack{b_1=1\\ b_1\perp v}}^{u}(2u-b_1)=\frac{3}{4\pi^2}u^2n^4+o(u^2n^4).$$

Using the obtained formula and formula (5.11) we proceed

$$\begin{aligned}\frac{6}{\pi^2}\sum_{u=1}^{m}(m-u)\sum_{v=1}^{n}(n-v)v^2\sum_{\substack{b_1=1\\ b_1\perp v}}^{u}(2u-b_1) &= \frac{6}{\pi^2}\sum_{u=1}^{m}(m-u)\left(\frac{3}{4\pi^2}u^2n^4+o(u^2n^4)\right)\\ &= \frac{9n^4}{2\pi^4}\sum_{u=1}^{m}(mu^2-u^3)+o(m^4n^4)\\ &= \frac{9n^4}{2\pi^4}\left(\frac{m^4}{3}-\frac{m^4}{4}\right)+o(m^4n^4)\\ &= \frac{3}{8\pi^4}m^4n^4+o(m^4n^4).\end{aligned}$$

Due to symmetry, the second summand in formula (5.40) has the same asymptotics, and therefore

$$\sum_{u=1}^{m-1}(m-u)\sum_{v=1}^{n-1}(n-v)\cdot|Z_1^b(u,v)| = \frac{3}{4\pi^4}m^4n^4+o(m^4n^4). \tag{5.43}$$

Finally, plugging in (5.38), (5.39), and (5.43) into the initial formula we obtain

$$p(m,n) = \frac{1}{24\pi^4}m^4n^4+\frac{31}{24\pi^4}m^4n^4++\frac{3}{4\pi^4}m^4n^4+o(m^4n^4) = \frac{25}{12\pi^4}m^4n^4+o(m^4n^4).$$

□

Theorems 99 and 84 imply Theorem 82.

5.4 The number of k-threshold functions for $k > 2$

The obtained asymptotic formula for the number of 2-threshold functions can be used to improve the trivial upper bound (5.1) on the number of k-threshold functions for $k \geq 3$. Indeed, since a k-threshold function can be seen as a conjunction of several 2-threshold functions and at most one threshold function, we have:

$$\begin{aligned}t_k(m,n) \leq \binom{t_2(m,n)}{\frac{k}{2}} &= \frac{t_2(m,n)^{\frac{k}{2}}}{\frac{k}{2}!}+o\left(m^{2k}n^{2k}\right)\\ &= \frac{5^k}{12^{\frac{k}{2}}\pi^{2k}\frac{k}{2}!}m^{2k}n^{2k}+o\left(m^{2k}n^{2k}\right)\end{aligned} \tag{5.44}$$

for even k and

$$\begin{aligned} t_k(m,n) &\leq \binom{t_2(m,n)}{\lfloor \frac{k}{2} \rfloor} \frac{t(m,n)}{k} = \frac{t_2(m,n)^{\lfloor \frac{k}{2} \rfloor} t(m,n)}{\lfloor \frac{k}{2} \rfloor ! k} + o\left(m^{2k} n^{2k}\right) \\ &= \frac{5^{k-1}6}{12^{\lfloor \frac{k}{2} \rfloor} \pi^{2k} \lfloor \frac{k}{2} \rfloor ! k} m^{2k} n^{2k} + o\left(m^{2k} n^{2k}\right) \end{aligned} \tag{5.45}$$

for odd k. Since for even k

$$\frac{5^k}{12^{\frac{k}{2}} \pi^{2k} \frac{k}{2}!} \leq \frac{6^k}{\pi^{2k} k!}$$

if and only if $k \leq 22$, and for odd k

$$\frac{5^{k-1}6}{12^{\lfloor \frac{k}{2} \rfloor} \pi^{2k} \lfloor \frac{k}{2} \rfloor ! k} \leq \frac{6^k}{\pi^{2k} k!}$$

if and only if $k \leq 23$, we conclude that the upper bounds in (5.44) and (5.45) improve the trivial estimation (5.1) for every $k \leq 23$.

5.5 Conclusion

A natural question is whether the approach we used to asymptotically enumerate 2-threshold functions can be generalized to higher order threshold functions, say to 3-threshold functions. One difference between 2-threshold and 3-threshold functions that might be an obstacle towards such a generalization is an observation that while almost all 2-threshold functions have a true point on the boundary of the grid, this does not hold for 3-threshold. This property of 2-threshold functions was crucial in our analysis.

Another natural question is to what extent the error term in the asymptotic formula (5.2) can be improved.

Appendix A

Non-canalyzing functions of 6 variables with the minimum specification number

Below we provide all non-congruent non-canalyzing functions from $\mathcal{T}_6$ up to dualization.

1. $x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_1x_2x_4x_6 \vee x_1x_2x_5x_6 \vee x_1x_3x_4x_5 \vee x_1x_3x_4x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$
2. $x_1x_2x_3x_6 \vee x_1x_2x_4x_6 \vee x_1x_2x_5x_6 \vee x_1x_3x_4x_5 \vee x_1x_3x_4x_6 \vee x_1x_3x_5x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_2x_3x_5x_6 \vee x_4x_5x_6$
3. $x_1x_2x_3x_5 \vee x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_1x_2x_4x_6 \vee x_1x_2x_5x_6 \vee x_1x_3x_4x_5 \vee x_2x_3x_4x_5 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$
4. $x_1x_2x_3x_4 \vee x_1x_2x_3x_5 \vee x_1x_5x_6 \vee x_2x_3x_6 \vee x_2x_4x_5 \vee x_2x_4x_6 \vee x_2x_5x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$
5. $x_1x_2x_4x_6 \vee x_1x_2x_5x_6 \vee x_1x_3x_4x_5 \vee x_1x_3x_4x_6 \vee x_1x_3x_5x_6 \vee x_1x_4x_5x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_2x_3x_5x_6 \vee x_2x_4x_5x_6 \vee x_3x_4x_5x_6$
6. $x_1x_2x_3x_4 \vee x_1x_2x_3x_5 \vee x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_1x_5x_6 \vee x_2x_4x_6 \vee x_2x_5x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$
7. $x_1x_2x_3x_4 \vee x_1x_2x_3x_5 \vee x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_1x_2x_4x_6 \vee x_1x_3x_4x_5 \vee x_2x_3x_4x_5 \vee x_2x_5x_6 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$
8. $x_1x_2x_3x_5 \vee x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_1x_2x_4x_6 \vee x_1x_2x_5x_6 \vee x_1x_3x_4x_5 \vee x_1x_3x_4x_6 \vee x_1x_3x_5x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_2x_3x_5x_6 \vee x_4x_5x_6$
9. $x_1x_2x_3x_4 \vee x_1x_2x_3x_5 \vee x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_1x_2x_4x_6 \vee x_1x_2x_5x_6 \vee x_1x_3x_4x_5 \vee x_1x_3x_4x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

10. $x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_1x_2x_4x_6 \vee x_1x_2x_5x_6 \vee x_1x_3x_4x_5 \vee x_1x_3x_4x_6 \vee x_1x_3x_5x_6 \vee x_1x_4x_5x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_2x_3x_5x_6 \vee x_2x_4x_5x_6 \vee x_3x_4x_5x_6$

11. $x_1x_2x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

12. $x_1x_2x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

13. $x_1x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_5x_6$

14. $x_1x_2x_3x_6 \vee x_4x_5 \vee x_4x_6 \vee x_5x_6$

15. $x_1x_2x_4x_5 \vee x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

16. $x_1x_3x_4x_5 \vee x_2x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

17. $x_1x_2x_3x_4x_5 \vee x_3x_6 \vee x_4x_6 \vee x_5x_6$

18. $x_1x_2x_3x_6 \vee x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

19. $x_1x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

20. $x_1x_3x_4x_5 \vee x_2x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

21. $x_1x_2x_3x_4x_5 \vee x_2x_3x_6 \vee x_4x_6 \vee x_5x_6$

22. $x_1x_2x_3x_6 \vee x_2x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

23. $x_1x_2x_3x_4x_5 \vee x_2x_4x_6 \vee x_3x_4x_6 \vee x_5x_6$

24. $x_1x_2x_4x_6 \vee x_2x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

25. $x_1x_3x_4x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_5x_6$

26. $x_1x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

27. $x_1x_2x_3x_5 \vee x_3x_6 \vee x_4x_5 \vee x_4x_6 \vee x_5x_6$

28. $x_1x_2x_6 \vee x_3x_4x_5 \vee x_3x_6 \vee x_4x_6 \vee x_5x_6$

29. $x_1x_2x_4x_5 \vee x_3x_4x_5 \vee x_3x_6 \vee x_4x_6 \vee x_5x_6$

30. $x_1x_2x_3x_4x_5 \vee x_2x_6 \vee x_3x_6 \vee x_4x_6 \vee x_5x_6$

31. $x_1x_2x_6 \vee x_2x_3x_4x_5 \vee x_3x_6 \vee x_4x_6 \vee x_5x_6$

32. $x_1x_2x_3x_5 \vee x_2x_3x_6 \vee x_4x_5 \vee x_4x_6 \vee x_5x_6$

33. $x_1x_3x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_6 \vee x_4x_6 \vee x_5x_6$

34. $x_1x_2x_3x_6 \vee x_2x_4x_5 \vee x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

35. $x_1x_2x_4x_5 \vee x_2x_3x_6 \vee x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

36. $x_1x_4x_6 \vee x_2x_3x_4x_5 \vee x_2x_4x_6 \vee x_3x_4x_6 \vee x_5x_6$

37. $x_1x_3x_4x_5 \vee x_2x_3x_4x_5 \vee x_2x_3x_6 \vee x_4x_6 \vee x_5x_6$

38. $x_1x_2x_4x_5 \vee x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

39. $x_1x_3x_4x_5 \vee x_2x_3x_4x_5 \vee x_2x_4x_6 \vee x_3x_4x_6 \vee x_5x_6$

40. $x_1x_2x_3x_6 \vee x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

41. $x_1x_2x_3x_4x_5 \vee x_2x_5x_6 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

42. $x_1x_2x_5x_6 \vee x_2x_3x_4x_5 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

43. $x_1x_3x_4x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

44. $x_1x_2x_3x_5 \vee x_2x_4x_5 \vee x_3x_4x_5 \vee x_3x_6 \vee x_4x_6 \vee x_5x_6$

45. $x_1x_2x_3x_4 \vee x_2x_3x_5 \vee x_3x_6 \vee x_4x_5 \vee x_4x_6 \vee x_5x_6$

46. $x_1x_2x_3x_5 \vee x_2x_3x_6 \vee x_2x_4x_5 \vee x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

47. $x_1x_2x_3x_6 \vee x_2x_4x_5 \vee x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

48. $x_1x_2x_4x_5 \vee x_2x_3x_6 \vee x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

49. $x_1x_5x_6 \vee x_2x_3x_4x_5 \vee x_2x_5x_6 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

50. $x_1x_2x_4x_6 \vee x_2x_5x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

51. $x_1x_3x_4x_5 \vee x_2x_3x_4x_5 \vee x_2x_5x_6 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

52. $x_1x_2x_4x_5 \vee x_1x_2x_4x_6 \vee x_1x_3x_4x_5 \vee x_2x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

53. $x_1x_2x_4x_6 \vee x_1x_3x_4x_5 \vee x_1x_3x_4x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_5x_6$

54. $x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_1x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

55. $x_1x_2x_3x_5 \vee x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

56. $x_1x_2x_4x_6 \vee x_1x_2x_5x_6 \vee x_1x_3x_4x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

57. $x_1x_2x_5x_6 \vee x_1x_3x_4x_6 \vee x_1x_3x_5x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_2x_3x_5x_6 \vee x_4x_5x_6$

58. $x_1x_2x_4x_5 \vee x_1x_2x_6 \vee x_1x_3x_6 \vee x_2x_3x_6 \vee x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

59. $x_1x_2x_4x_5 \vee x_1x_2x_6 \vee x_1x_3x_4x_5 \vee x_2x_3x_4x_5 \vee x_3x_6 \vee x_4x_6 \vee x_5x_6$

60. $x_1x_2x_3x_5 \vee x_2x_3x_6 \vee x_2x_4x_5 \vee x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

61. $x_1x_2x_3x_5 \vee x_1x_2x_3x_6 \vee x_1x_4x_5 \vee x_2x_4x_5 \vee x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

62. $x_1x_2x_3x_5 \vee x_1x_2x_4x_5 \vee x_1x_3x_6 \vee x_2x_3x_6 \vee x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

63. $x_1x_2x_3x_4 \vee x_2x_3x_5 \vee x_2x_3x_6 \vee x_2x_4x_5 \vee x_3x_4x_5 \vee x_4x_6 \vee x_5x_6$

64. $x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_1x_2x_4x_6 \vee x_2x_5x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

65. $x_1x_2x_3x_5 \vee x_1x_2x_4x_5 \vee x_1x_4x_6 \vee x_2x_3x_6 \vee x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6 x_1x_2x_3x_5 \vee x_1x_2x_3x_6 \vee x_1x_4x_6 \vee x_2x_4x_5 \vee x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

66. $x_1x_2x_4x_5 \vee x_1x_2x_4x_6 \vee x_1x_3x_4x_5 \vee x_2x_3x_4x_5 \vee x_2x_5x_6 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

67. $x_1x_2x_4x_6 \vee x_1x_3x_4x_5 \vee x_1x_3x_4x_6 \vee x_2x_3x_4x_5 \vee x_2x_3x_4x_6 \vee x_2x_5x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

68. $x_1x_2x_3x_4 \vee x_1x_2x_3x_5 \vee x_1x_4x_6 \vee x_2x_3x_6 \vee x_2x_4x_5 \vee x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

69. $x_1x_2x_3x_5 \vee x_1x_2x_3x_6 \vee x_1x_2x_4x_5 \vee x_2x_4x_6 \vee x_2x_5x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_3x_5x_6 \vee x_4x_5x_6$

70. $x_1x_4x_6 \vee x_2x_3x_4 \vee x_2x_3x_5 \vee x_2x_3x_6 \vee x_2x_4x_5 \vee x_2x_4x_6 \vee x_3x_4x_5 \vee x_3x_4x_6 \vee x_5x_6$

The following code written in Wolfram Language was used to enumerate all threshold functions from T_n for $n \leq 6$. First, all non-congruent positive functions were enumerated.

```
monotoneFunctions[varsNum_] := Module[{},
  expressions = {False, True};
  If[varsNum == 1, expressions, 0];
  vars = Array[x, varsNum];
  xn = Last[vars];
  permutVars = Permutations[vars];
  (*enumerate all distinct monotone functions*)
  Do[
   newExs = List[];
   Do[
    Do[
     If[BooleanConvert[ex1 || ex2, "DNF"] == ex2,
      newEx = BooleanConvert[ex1 || ex2 && xi, "DNF"];
      AppendTo[newExs, newEx]]
     , {ex2, expressions}]
    , {ex1, expressions}];
   expressions = newExs
```

```
    , {xi, vars}];
   Export["monotone(5)all.txt", expressions, "List"];
   Print["Delete exs with < n vars"];
   Print[TimeObject[Now]];
   (*delete all expressions with < n variables*)
   i = 1;
   While[i <= Length[expressions],
    If[Length[BooleanVariables[Part[expressions, i]]]
    < varsNum,
     expressions = Drop[expressions, {i}], i = i + 1]];
   (*delete all isomorphic expressions*)

   Print["Delete isomorphic exs"];
   Print[TimeObject[Now]];
   Do[
    If[i >= Length[expressions] - 1, Break[]];
    testEx = Part[expressions, i];
    permutations = allPermutations[testEx, vars, permutVars];
    permutations = Drop[permutations, {1}];
    If[Length[permutations] == 0, Continue[]];
    Do[
     Do[
      fEx = findExpression[expressions, testPerm, j];
      If[fEx > 0,
         expressions = Drop[expressions, {fEx}]; Break[], 0];
      If[j >= Length[expressions], Break[]]
      , {j, i + 1, Length[expressions]}]
     , {testPerm, permutations}];
    If[i >= Length[expressions] - 1, Break[]]
    , {i, Length[expressions]}];
   Print[Length[expressions]];
   expressions
   ]

findExpression[list1_, expression1_, index1_] :=
 Module[{expression = expression1,
         res,
         item,
         list = list1,
```

```
        listLength,
        index = index1},
  res = False;
  listLength = Length[list];
  If[listLength - index < 0, Return[res]];
  Do[
   item = Part[list, i];
   If[ TautologyQ[Equivalent[expression, item]],
       res = i; Break[]],
       {i, index, listLength}];
  Return[res]
  ]

allPermutations[expression1_, vars1_, permutVars1_] :=
 Module[{expression = expression1,
         vars = vars1,
         expressions,
         permutVars = permutVars1,
         permutCount, varsCount,
         xn,
         p,
         permutList,
         permut,
         fEx,
         ex},
  varsCount = Length[vars];
  If[varsCount == 0
  || Length[BooleanVariables[expression]] == 0,
   Return[{expression}]];
  xn = Last[vars];
  permutCount = Length[permutVars];
  permutList = List[];
  Do[
   permut = List[];
   Do[
    AppendTo[permut, Part[vars, i] -> Part[p, i]]
    , {i, varsCount}];
   AppendTo[permutList, permut]
   , {p, permutVars}];
```

```
  expressions = List[];
  Do[
   ex = Replace[expression, p, {0, 10}];
   fEx = findExpression[expressions, ex, 1];
   If[fEx, 0, AppendTo[expressions, ex]]
   , {p, permutList}];
  Return[expressions];
  ]

findIsomorphic[list1_, expression1_, vars_, permutVars_] :=
 Module[{expression = expression1,
         permutations,
         item,
         list = list1,
         res},
  permutations
       = allPermutations[expression, vars, permutVars];
  res = False;
  Do[
   If[findExpression[list, item], res = True; Return[res]],
   {item, permutations}];
  Return[res]
  ]
```

Then all obtained positive functions were checked whether they are threshold or not. For threshold functions the set of essential points was found. The functions with the minimum specification number were moved to the target list.

```
(*returns minimal ones (if value_ = 1) or maximal zeros
(if value_ = 0) for the threshold function
corresponding to expression_*)

extremalpoints[expression1_, value_] :=
 Module[{expression = expression1,
         points},
  expression = BooleanConvert[expression,
                 If[value > 0, "DNF", "CNF"]];
  vars = Sort[BooleanVariables[expression]];
  points =
   If[value > 0,
```

```
    (expression) /. Or -> List, (expression) /.
     And -> List];
  If[! VectorQ[points], points = {points}];
  If[Length[points] == 0, points = {expression}, 0];
  vectors = List[];
  Do[
   ones = Sort[BooleanVariables[i]];
   list = List[];
   Do[
    AppendTo[list,
     If [value > 0 , Boole[MemberQ[ones, j]],
      Boole[! MemberQ[ones, j]]]], {j, vars}
    ];
   AppendTo[vectors, list], {i, points}
   ];
  vectors
  ]

sum[vars_, coefs_] := Module[{},
  list = vars * coefs;
  Total[list]
  ]

(*check if expression_ corresponds to
a threshold function*)

isthreshold[expression1_, zeros1_, ones1_, print1_] :=
 Module[{expression = expression1,
         ones = ones1,
         zeros = zeros1,
         coefs,
         print = print1},
  expression = BooleanMinimize[expression];
  result = True;
  If[TautologyQ[expression ] || TautologyQ[! expression],
   Return[result]];
  If[Length[zeros] <= 0,
     zeros = extremalpoints[expression, 0], 0];
  If[Length[ones] <= 0,
```

```
    ones = extremalpoints[expression, 1], 0];
  vars = Sort[BooleanVariables[expression]];
  coefs = Array[a, Length[vars]];
  ineqList = List[];
  Do[AppendTo[ineqList, i > 0], {i, coefs}];
  AppendTo[ineqList, b > 0];
  Do[AppendTo[ineqList, sum[i, coefs] >= b], {i, ones}];
  Do[AppendTo[ineqList, sum[i, coefs] < b], {i, zeros}];
  AppendTo[coefs, b];
  equations = FindInstance[ineqList, coefs, Reals];
  If[Length[equations] > 0, result = True;
   If[print,
    Print["Coefficients of threshold inequality \
a[1]x[1]+...+a[n]x[n]>=b:"]; Print[equations]],
    result = False];
  result
  ]

(*check if expression corresponds to
       a threshold function*)

isthreshold[expression1_, print1_] :=
 Module[{expression = expression1,
         ones,
         zeros,
         print = print1},
  expression = BooleanMinimize[expression];
  zeros = extremalpoints[expression, 0];
  ones = extremalpoints[expression, 1];
  isthreshold[expression, zeros, ones, print]
  ]

(*check if point_ is essential for the threshold
function corresponding to expression_*)

isessential[expression1_, zeros1_, ones1_, point_] :=
 Module[{expression = expression1,
        zeros = zeros1,
        ones = ones1},
```

```
  expressiond = expression;
  minterm = True;
  vars = Sort[BooleanVariables[expression]];
  varsNum = Length[vars];
  Do[If[Part[point, i] > 0,
        minterm = minterm && Part[vars, i],
    1], {i, varsNum}];
  minterm = BooleanMinimize[minterm];
  If[MemberQ[zeros, point],
     expressiond = expression || minterm,
   If[MemberQ[ones, point],
    (*If the point is a minimal one*)
    expressiond = False;
    Do[
     If[MemberQ[BooleanVariables[minterm], i], 0,
      expressiond = expressiond || minterm && i]
     , {i, vars}];
    Do[
     If [point == i, Continue[], 0];
     term1 = True;
     Do[If[Part[i, j] > 0,
     term1 = term1 && Part[vars, j], 1], {j,
       varsNum}];
     expressiond = expressiond || term1, {i, ones}],
     Return False]];
  isthreshold[BooleanMinimize[expressiond], False]
  ]

(*returns essential points of the function
corresponding to the expression_*)

essentialpoints[expression1_, zeros1_, ones1_] :=
 Module[{expression = expression1,
         ones = ones1,
         zeros = zeros1},
  essenPoints = List[];
  Do[If[isessential[expression, zeros, ones, i],
    AppendTo[essenPoints, i], 0]
   , {i, zeros}];
```

```
  Do[If[isessential[expression, zeros, ones, i],
    AppendTo[essenPoints, i], 0]
   , {i, ones}];
  essenPoints
  ]

(* checks if expression_ corresponds to
a threshold function.
If yes, then returns extremal and essential points*)

analyze[expression1_] :=
 Module[{expression = expression1, zeros, ones},
  expression = BooleanMinimize[expression];
  numVars = Length[BooleanVariables[expression]];
  ones = extremalpoints[expression, 1];
  zeros = extremalpoints[expression, 0];
  threshold = False;
  If[isthreshold[expression, zeros, ones, True] != False,
   threshold = True,
   threshold = False];
  Print["Is f threshold?"];
  Print[threshold];
  If[! threshold, 0,
   essenpoints = essentialpoints[expression, zeros, ones];
   Print["Is f has minimum specification number?"];
   Print[Length[essenpoints] <= numVars + 1];
   Print["Extremal points number:"];
   Print[Length[ones] + Length[zeros]];
   Print["Essential points number:"];
   Print[Length[essenpoints]];
   Print["Minimal ones:"];
   Print[Column[ones]];
   Print["Maximal zeros:"];
   Print[Column[zeros]];
   Print["Essential points:"];
   Print[Column[essenpoints]];

   ];
  ]
```

```
(*Insert the expression inside "analyze[...]" below
instead of the example expression*)
analyze[x1 && x2 || x3 || x4]
```

A sample of the output of the given code is following:

```
Coefficients of threshold inequality
a[1]x[1]+...+a[n]x[n]>=b:
{{a[1]->1,a[2]->1,a[3]->2,a[4]->2,b->2}}
Is f threshold?
True
Does f have the minimum specification number?
True
Extremal points number:
5
Essential points number:
5
Minimal ones:
{1,1,0,0}
{0,0,1,0}
{0,0,0,1}
Maximal zeros:
{0,1,0,0}
{1,0,0,0}
Essential points:
{0,1,0,0}
{1,0,0,0}
{1,1,0,0}
{0,0,1,0}
{0,0,0,1}
```

Appendix B

k-threshold functions and their specifying sets

This appendix contains results of the author related to two-dimensional 2-threshold and k-threshold functions from [48] and [49]. According to the notation in these papers we denote by $\mathfrak{T}(d,n,k)$ the class of k-threshold functions over $\mathbb{Z}_n^d$ and by $\mathfrak{T}(d,n,*)$ the class of k-threshold functions over $\mathbb{Z}_n^d$ for arbitrary k, i.e.

$$\mathfrak{T}(d,n,*) = \bigcup_{k\geq 1} \mathfrak{T}(d,n,k).$$

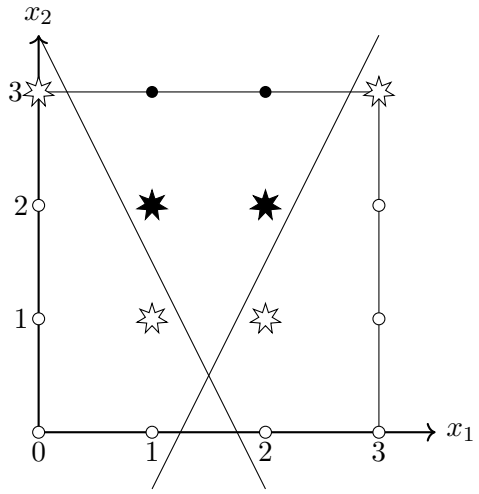
If $k=1$ we will write $\mathfrak{T}(d,n)$.

Let $\mathcal{C}$ be a class of $\{0,1\}$-valued functions over some domain and let $f \in \mathcal{C}$. The *teaching dimension* of a class $\mathcal{C}$ is defined as

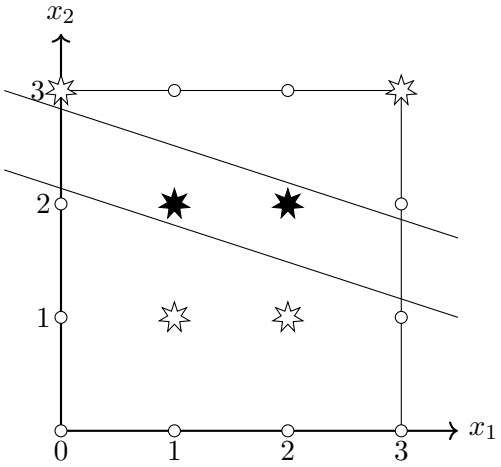
$$\sigma(\mathcal{C}) = \max_{f\in\mathcal{C}} \sigma_{\mathcal{C}}(f),$$

where $\sigma_{\mathcal{C}}(f)$ is the specification number of f with respect to $\mathcal{C}$. The teaching dimension of a class of functions is an important learning property of the class. In machine learning, the main goal of a learning algorithm with membership queries is to find any specifying set of a target function f with respect to a concept class $\mathcal{C}$. The algorithm succeeds if it queried the values of the function in all points of some specifying set of the function. Therefore the teaching dimension of the class $\mathcal{C}$ is a lower bound on the learning complexity of this class.

Denote by $J(f,\mathcal{C})$ the number of minimal specifying sets of f with respect to the class $\mathcal{C}$. It is known, that the set of essential points of a threshold function is a specifying set of this function. Together with the simple observation that any specifying set of a function contains all its essential points, this imply that any threshold function have a unique minimal specifying set, that is $J(f,\mathfrak{T}(d,n)) = 1$. The situation becomes different for k-threshold functions when $k \geq 2$. We illustrate this difference in the following example.



(a) f is a 2-threshold function with $M_1(f) = \{(1,2),(2,2),(1,3),(2,3)\}$.



(b) g is a 2-threshold function with $M_1(g) = \{(1,2),(2,2)\}$. g and f agree on S.

Figure B.1: The stars are the essential points of f. The black elements are the true points of the corresponding function.

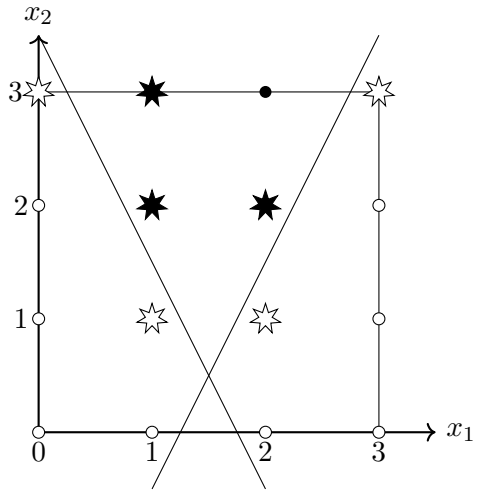
Example 100. Let f be a 2-threshold function over $\mathbb{Z}_4^2$ such that

$$M_1(f) = \{(1,2),(1,3),(2,2),(2,3)\}.$$

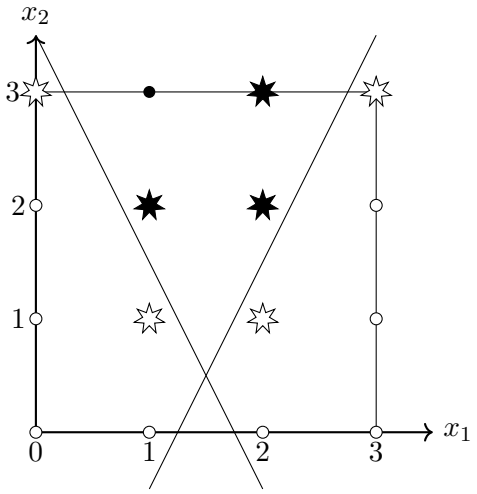
The set of essential points of f is $S = \{(1,1),(1,2),(2,1),(2,2),(0,3),(3,3)\}$. This set is not a specifying set because there exists a function $g \in \mathfrak{T}(2,4,2)$ such that $M_1(g) = \{(1,2),(2,2)\}$ and g agrees with f on S (see Fig. B.1). However, if we add any of the two points $(1,3)$ and $(2,3)$ to S, then we obtain a minimal specifying set of f (see Fig. B.2) with respect to $\mathfrak{T}(2,4,2)$, and therefore $J(f,\mathfrak{T}(2,4,2)) \geq 2$.

In this appendix we study combinatorial and structural properties of specifying sets of k-threshold functions for $k \geq 2$. In particular, we construct a sequence of functions from $\mathfrak{T}(2,n,2)$ for which the number of minimal specifying sets grows as $\Omega(n^2)$. On the other hand, we show that any k-threshold function f has a unique minimal specifying set with respect to $\mathfrak{T}(d,n,*)$ coinciding with the set of essential points of f. In addition, we give a general structural description of minimal specifying sets of such functions. For functions in $\mathfrak{T}(2,n,*)$ we refine the given structure and derive a bound on the size of the minimal specifying sets. Finally, we show that any two-dimensional 2-threshold function that has a unique pair of defining threshold functions has specification number at most 9.

The organization of the appendix is as follows. In Section B.1 we consider essential points of the conjunction of arbitrary $\{0,1\}$-valued functions $f_1,\ldots,f_k$ and their connection with essential points of these functions. In the beginning of Section B.2 we show that in general a k-threshold function can have more than one minimal specifying set. The main result of Section B.2 (Theorem 107) states that a minimal specifying set of a k-threshold



(a) $S \cup \{(1,3)\}$ is a minimal specifying set



(b) $S \cup \{(2,3)\}$ is a minimal specifying set

Figure B.2: The stars denote the points of the minimal specifying sets which consist of the set of essential points $S = \{(1,1),(2,1),(1,2),(2,2),(0,3),(3,3)\}$ and one more point.

function with respect to $\mathfrak{T}(d,n,*)$ is unique and coincides with the set of its essential points. The structure of this set is given as well. In Section B.3 we consider the class $\mathfrak{T}(2,n,*)$ and for a function f in the class we prove an upper bound on the size of the set of essential points. Section B.4 is devoted to the two-dimensional 2-threshold functions with a unique defining pair of threshold functions, where we show that for each of these functions any of its minimal specifying sets is of size at most 9 and there can be $\Omega(n^2)$ of different minimal specifying sets.

B.1 The set of essential points of a $\{0,1\}$-valued functions conjunction

Since a k-threshold function is a conjunction of k threshold functions, it is interesting to investigate the connection between essential points of threshold functions $f_1,\ldots,f_k$ and essential points of their conjunction. In this section we prove several claims that establish this relationship. For a natural $k>1$ and a class $\mathcal{C}$ of $\{0,1\}$-valued functions we denote by $\mathcal{C}^k$ the class of functions which can be presented as the conjunction of k functions from $\mathcal{C}$.

We denote the set of essential points of f with respect to the class $\mathcal{C}$ by $S(f,\mathcal{C})$ or by $S(f)$ when $\mathcal{C}$ is clear. Let also $S_\nu(f) = S(f) \cap M_\nu(f)$.

Claim 101 ([48]). *Let $\mathcal{C}$ be a class of $\{0,1\}$-valued functions over a domain X and $f_1,\ldots,f_k \in \mathcal{C}$. Then for the function $f = f_1 \wedge \cdots \wedge f_k$ and each $i \in [k]$ we have*

$$S_1(f_i,\mathcal{C}) \cap M_1(f) \subseteq S_1(f,\mathcal{C}^k).$$

Claim 102 ([48]). *Let $\mathcal{C}$ be a class of $\{0,1\}$-valued functions over a domain X and $f_1, \ldots, f_k \in \mathcal{C}$. Then for the function $f = f_1 \wedge \cdots \wedge f_k$ and each $i \in [k]$ we have*

$$S_0(f_i, \mathcal{C}) \cap \bigcap_{j \neq i} M_1(f_j) \subseteq S_0(f, \mathcal{C}^k).$$

Claim 103 ([48]). *Let $\mathcal{C}$ be a class of $\{0,1\}$-valued functions over a domain X and $f \in \mathcal{C}^k$. If there exists a unique set $f_1, \ldots, f_k \in \mathcal{C}$ such that $f = f_1 \wedge \cdots \wedge f_k$, then for each $i \in [k]$ we have*

$$S(f_i, \mathcal{C}) \subseteq \bigcap_{j \neq i} M_1(f_j).$$

Corollary 104 ([48]). *Let $\mathcal{C}$ be a class of $\{0,1\}$-valued functions over a domain X and $f \in \mathcal{C}^k$. If there exists a unique set $f_1, \ldots, f_k \in \mathcal{C}$ such that $f = f_1 \wedge \cdots \wedge f_k$ then*

$$\bigcup_{i=1}^{k} S_0(f_i, \mathcal{C}) \subseteq S_0(f, \mathcal{C}^k)$$

and

$$\bigcup_{i=1}^{k} S_1(f_i, \mathcal{C}) \subseteq S_1(f, \mathcal{C}^k).$$

B.2 Specifying sets for functions in $\mathfrak{T}(d, n, *)$

In this section we prove that for $k, d \geq 2$ the teaching dimension of $\mathfrak{T}(d, n, k)$ is n^d and present functions attaining this bound. Then we consider the class $\mathfrak{T}(d, n, *)$ and show that for a function $f \in \mathfrak{T}(d, n, *)$ the set of its essential points with respect to $\mathfrak{T}(d, n, *)$ is also a specifying set, and therefore it is a unique minimal specifying set of f with respect to $\mathfrak{T}(d, n, *)$.

Lemma 105 ([48]). *Let $f : \mathbb{Z}_n^d \to \{0,1\}$ be a function such that $1 \leq |\text{Vert}(P(f))| \leq 2$ and $P(f) \cap M_0(f) = \emptyset$. Then f is k-threshold for any $k \geq 2$.*

In [4] it was established that the teaching dimension of a class containing the empty set and N singleton sets is at least N. This result and Lemma 105 give us the teaching dimension for $\mathfrak{T}(d, n, k)$ for $k \geq 2$:

Corollary 106. *$\sigma(\mathfrak{T}(d, n, k)) = n^d$ for every $k \geq 2$.*

For a polytope P denote by $B(P)$ the set of integer points on the border of P and by $\text{Int}(P)$ the set of internal integer points of P. For $f \in \mathfrak{T}(d, n, *)$ denote by $D(f)$ the set $\{x \in M_0(f) : \text{Conv}(P(f) \cup \{x\}) \cap M_0(f) = \{x\}\}$.

Theorem 107 ([48]). *Let $f \in \mathfrak{T}(d, n, *)$ and $d, n \geq 2$. Then*

$$S(f, \mathfrak{T}(d, n, *)) = \begin{cases} \mathbb{Z}_n^d, & M_1(f) = \emptyset; \\ \text{Vert}(P(f)) \cup D(f), & M_1(f) \neq \emptyset; \end{cases}$$

and $S(f, \mathfrak{T}(d, n, *))$ is a unique minimal specifying set of f.

Lemma 108 ([48]). *Let $f \in \mathfrak{T}(d, n, k)$ for some $d, k \geq 2$ and let $M_1(f) = \{x'\}$. Then*

$$S(f, \mathfrak{T}(d, n, k)) = \{x'\} \cup \{x \in \mathbb{Z}_n^d : \text{GCD}(|x_1 - x_1'|, \ldots, |x_d - x_d'|) = 1\},$$

$S(f, \mathfrak{T}(d, n, k))$ is a unique specifying set of f with respect to $\mathfrak{T}(d, n, k)$, and

$$|S(f, \mathfrak{T}(d, n, k))| = \Theta(n^d).$$

B.3 Specifying sets for functions in $\mathfrak{T}(2, n, *)$

In the previous section we proved that for a function in $\mathfrak{T}(d, n, *)$, where $d \geq 2$, the set of its essential points is also the unique minimal specifying set. In this section we consider the class $\mathfrak{T}(2, n, *)$ and describe the structure of the set of essential points for a function in this class. We also give an upper bound on the size of this set.

Let us consider an arbitrary function $f \in \mathfrak{T}(2, n, *)$. Note that $P(f)$ can be the empty set, a point, a segment or a polygon. Let $P(f)$ be a segment or a polygon, that is $|M_1(f)| > 1$, and let $a_1x_1 + a_2x_2 = a_0$ be the edge equation for an edge e of $P(f)$. Without loss of generality we may assume that $\text{GCD}(a_1, a_2) = 1$. Denote by *edge inequality* for edge e the inequality $a_1x_1 + a_2x_2 \leq a_0$ or/and $a_1x_1 + a_2x_2 \geq a_0$ if it holds for all points of $P(f)$. Note that if $P(f)$ is a segment, then it has one edge but two edge inequalities corresponding to the edge. If $P(f)$ is a polygon, then it has exactly one edge inequality for each edge. Hence, the number of edge inequalities for $P(f)$ is equal to the number of its vertices.

Let f be a function from $\mathfrak{T}(2, n, *)$ with $|M_1(f)| > 1$ and let

$$a_{i1}x_1 + a_{i2}x_2 \leq a_{i0}, \quad i = 1, \ldots, |\text{Vert}(P(f))|$$

be edge inequalities for $P(f)$. The *extended edge inequality* for an edge e of $P(f)$ is $a_1x_1 + a_2x_2 \leq a_0 + 1$, where $a_1x_1 + a_2x_2 \leq a_0$ is the corresponding edge inequality for e. By $P'(f)$ we denote the following extension of $P(f)$

$$\{x = (x_1, x_2) : a_{i1}x_1 + a_{i2}x_2 \leq a_{i0} + 1, \quad i = 1, \ldots, |\text{Vert}(P(f))|\}.$$

We also denote

$$\Delta P(f) = P'(f) \setminus P(f).$$

It follows from the definition that $P'(f)$ contains $P(f)$, and for every straight line ℓ' containing an edge of $P'(f)$ there exists an edge in $P(f)$ belonging to the closest parallel to the ℓ' straight line which contains integer points.

If P is a polygon then denote by $\mathcal{P}(P)$ the perimeter of P.

The following claim uses the Pick's formula (see [47]) for the area of a convex polygon P with integer vertices:

$$\text{Area}(P) = \text{Int}(P) + \frac{B(P)}{2} - 1.$$

Claim 109 ([48]). *Let $f \in \mathfrak{T}(2, n, *)$ and $\text{Area}(P(f)) > 0$. Then $D(f) = \Delta P(f) \cap M_0(f)$.*

Corollary 110 ([48]). *Let $f \in \mathfrak{T}(2, n, *)$ and $\text{Area}(P(f)) > 0$. Then*

$$S(f, \mathfrak{T}(2, n, *)) = (\Delta P(f) \cap M_0(f)) \cup \text{Vert}(P(f)).$$

Corollary 111 ([48]). *Let $f \in \mathfrak{T}(2, n, *)$ and $\text{Area}(P(f)) > 0$. Then*

$$S(f, \mathfrak{T}(2, n, *)) = O(n).$$

The following claim establishes the relationship between the perimeters of $P(f)$ and $P'(f)$ to help us to estimate the size of the set of essential points of a function in $\mathfrak{T}(2, n, *)$.

Claim 112 ([49]). *Let P and P' be a convex polygon with integer vertices and its extension respectively. Then*

$$\mathcal{P}(P') < 5\mathcal{P}(P) + \frac{4}{\sin \min\limits_{v \in \text{Vert}(P)} \frac{q(v,P)}{2}}.$$

Corollary 113 ([49]). *Let P be a convex polygon with integer vertices, E be a rectangle with integer vertices such that $P \subseteq E$. If P' is the extension of P then*

$$length(b(P') \cap E)) < 5\mathcal{P}(P) + \frac{4}{\sin \frac{q_{min}(P,E)}{2}} + 8,$$

where $q_{min}(P, E) = \min\limits_{v \in \text{Vert}(P) \setminus b(E)} q(v, P)$.

In what follows we will use $q_{min}(P, E)$ to estimate the teaching dimension of a k-threshold function. In the case, when $E = \mathbb{Z}_n^2$ we will write $q_{min}(P)$.

Theorem 114 ([49]). *Let $f \in \mathfrak{T}(2, n, *)$ and $\text{Area}(P(f)) > 0$. Then*

$$|S(f, \mathfrak{T}(2, n, *))| = O\left(\min\left(n, \mathcal{P}(P(f)) + \frac{1}{q_{min}(P(f))}\right)\right).$$

Example 115 ([48]). *Consider a function $f \in \mathfrak{T}(2, 12, *)$ (see Fig. B.3). The grey set is $\Delta P(f)$. The black stars are the points of* Vert$(P(f))$ *and the white stars are the points of $\Delta P(f) \cap M_0(f)$.*

The estimation has been given for the functions with at most one true point and for the functions f with non-null area of $P(f)$. The functions f for which $P(f)$ is a segment are considered in the following lemma.

Lemma 116 ([49]). *Let $f \in \mathfrak{T}(2, n, *)$, $|M_1(f)| > 1$. If there exist integer numbers a, b, c such that* GCD$(a, b) = 1$ *and $ax_1 + bx_2 + c = 0$ for each $(x_1, x_2) \in M_1(f)$, then*

$$S(f) = \{v_1, v_2\} \cup \{v_1 \pm (b, -a), v_2 \pm (b, -a)\} \cap M_0(f) \cup \\ \cup \{(x_1, x_2) \in M_0(f) : |ax_1 + bx_2 + c| = 1\},$$

where Vert$(P(f)) = \{v_1, v_2\}$ *and*

$$|S(f)| \leq \frac{2n}{\max(|a|, |b|)} + 4.$$

From Lemma 108, Corollary 111, and Lemma 116 it follows

Theorem 117 ([49]). *If $f \in \mathfrak{T}(2, n, *)$, then*

$$\sigma_{\mathfrak{T}(2,n,*)}(f) = \begin{cases} \Theta(n^2), & |M_1(f)| \leq 1, \\ O(n), & |M_1(f)| > 1. \end{cases}$$

Claim 118 ([48]). *Let $f \in \mathfrak{T}(2, n, *)$ and $M_1(f) > 1$. Then f is a* $|$Vert$(P(f))|$*-threshold function and the sets of essential points of f with respect to $\mathfrak{T}(2, n, *)$ and $\mathfrak{T}(2, n, |$*Vert$(P(f))| + 1)$ *coincide.*

Example 119 ([48]). *Consider a function $f \in \mathfrak{T}(2, n, 4)$ with $M_1(f) = \{(1, 1), (1, 2), (2, 1)\}$ (see Fig. B.4). We have* Vert$(P(f)) = \{(1, 1), (1, 2), (2, 1)\}$ *and f is a 3-threshold function. Further,*

$$\Delta P(f) \cap \mathbb{Z}_4^2 = \{(0, 0), (1, 0), (2, 0), (3, 0), (0, 1), (3, 1), (0, 2), (2, 2), (3, 2), (0, 3), (1, 3)\},$$

and hence $S(f, \mathfrak{T}(2, n, *)) = \mathbb{Z}_4^2 \setminus \{(3, 2), (2, 3), (3, 3)\}$ = $S(f, \mathfrak{T}(2, n, 4))$.

B.4 Specification number of two-dimensional 2-threshold functions

In this section we consider 2-threshold functions over $\mathbb{Z}_n^2$ and their specification number. We will split the class of 2-threshold functions into two main parts and estimate their spec-

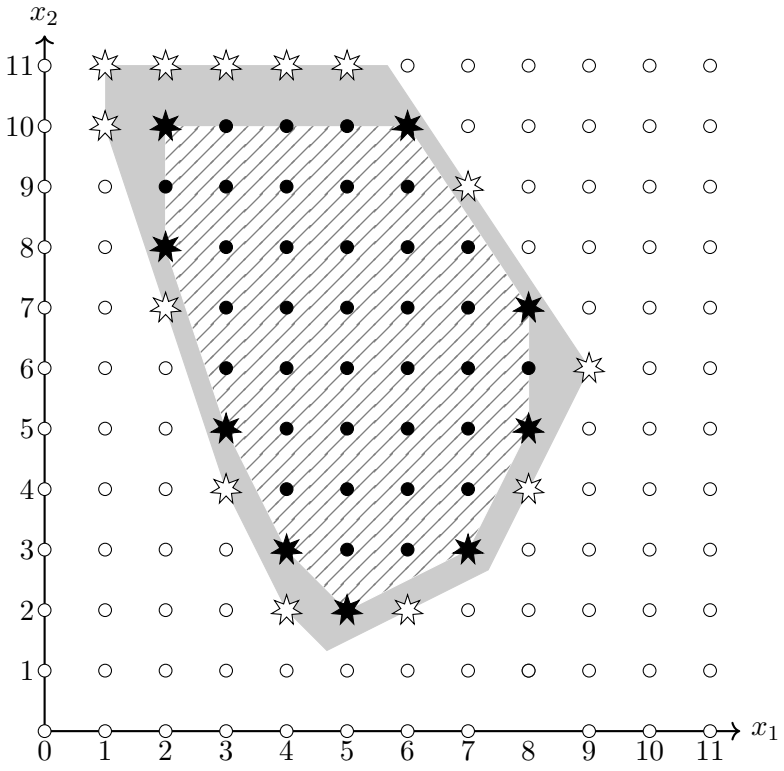


Figure B.3: The grey shape is $\Delta P(f)$, the stripped region is $P(f)$.

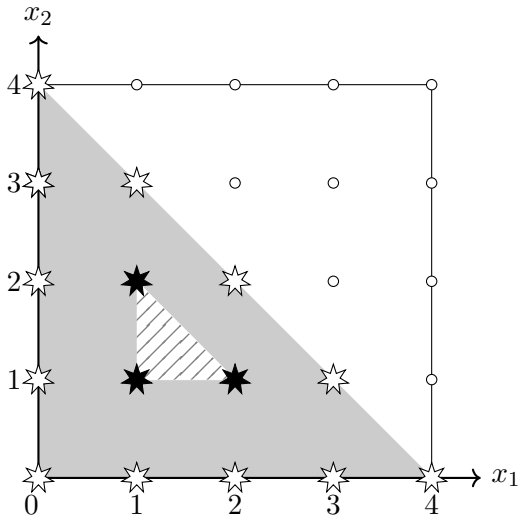


Figure B.4: The grey shape is $\Delta P(f)$, the stripped region is $P(f)$.

ification number separately.

Let f be a threshold function over $\mathbb{Z}_n^2$ and let a_0, a_1, a_2 be real numbers which are not all zero. We call the line $a_1x_1 + a_2x_2 = a_0$ an *i-separator* (or just *separator*) of f if there exists $i \in \{0, 1\}$ such that

$$x = (x_1, x_2) \in M_i(f) \iff a_1x_1 + a_2x_2 \leq a_0.$$

For example, a separating line of f defines a 1-separator of f. Let us prove some properties of separators of threshold functions.

It is known [3] that $|S(g)| \in \{3, 4\}$ and $|S_1(g)|, |S_0(g)| \in \{1, 2\}$ for any $g \in \mathfrak{T}(2, n)$ and the 1-valued essential points of g are adjacent vertices of $P(g)$.

Claim 120 ([48]). *Let f be a threshold function over $\mathbb{Z}_n^2$. For any $i \in \{0, 1\}$ there exists an i-separator of f which contains all points of $S_i(f)$.*

Claim 121 ([48]). *Let f be a threshold function over $\mathbb{Z}_n^2$ and let ℓ be an i-separator for f for some $i \in \{0, 1\}$. Then* Vert(Conv($\ell \cap \mathbb{Z}_n^2$)) $\subseteq S_i(f)$.

Theorem 122 ([48]). *Let $f \in \mathfrak{T}(2, n, 2)$, $M_1(f) \cap B(\text{Conv}(\mathbb{Z}_n^2)) \neq \emptyset$, and let f_1, f_2 be threshold functions such that $f = f_1 \wedge f_2$,*

$$S(f_1) \cap M_0(f_2) = \emptyset$$

and

$$S(f_2) \cap M_0(f_1) = \emptyset.$$

Then $\{f_1, f_2\}$ is a unique pair of functions defining f and

$$\sigma_{\mathfrak{T}(2,n,2)}(f) \leq 9.$$

Remark 123 ([48]). *Theorem 122 also holds when the domain is a convex subset of $\mathbb{Z}_n^2$.*

Corollary 124 ([48]). *Let $f \in \mathfrak{T}(2, n, 2)$ and there is a unique set of threshold functions $\{f_1, f_2\}$ defining f. If $M_1(f) \cap B(\text{Conv}(\mathbb{Z}_n^2)) \neq \emptyset$, then*

$$\sigma_{\mathfrak{T}(2,n,2)}(f) \leq 9.$$

The following theorem proves that the number of minimal specifying sets of 2-threshold functions can grow as $\Omega(n^2)$.

Theorem 125 ([48]).

$$\max_{f \in \mathfrak{T}(2,n,2)} J(f, \mathfrak{T}(2, n, 2)) = \Omega(n^2).$$

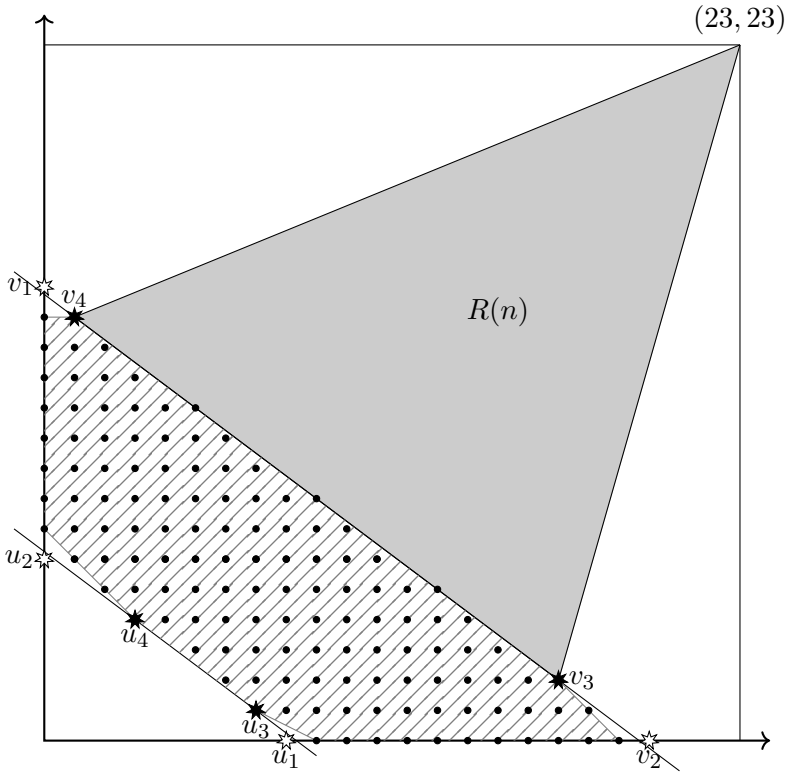


Figure B.5: An example of $f^{(24)}$, the black points are the points of $M_1(f)$, the grey region is $R(n)$, the stripped region is $P(f)$.

Proof. The sketch of the proof is following. Let

$$m = m(n) = \left\lfloor \frac{n-1}{4} \right\rfloor.$$

For $n \geq 21$ let $f^{(n)} \in \mathfrak{T}(2, n, 2)$ be defined by the following system of inequalities:

$$\begin{cases} -3x_1 - 4x_2 \leq -25, \\ 3x_1 + 4x_2 \leq 12m - 1. \end{cases}$$

It is easy to see that $f^{(n)}$ has 8 essential points and this set of points with any point from the triangle $R(n) \subset \mathrm{Conv}(\mathbb{Z}_n^2)$ (see Fig. B.5) forms a minimal specifying set of $f^{(n)}$. Further it is not hard to prove that $R(n)$ consists of $\Omega(n^2)$ integer points. Hence, the number of minimal specifying sets of $f^{(n)}$ grows as $\Omega(n^2)$. □

B.5 Conclusion

In this appendix we investigated structural and combinatorial properties of essential points and specifying sets of k-threshold functions.

We proved that for each $k \geq 2$ the class $\mathfrak{T}(d, n, k)$ contains functions with a minimal specifying set of size $\Theta(n^d)$. We considered two-dimensional 2-threshold functions and proved that the set of essential points of such a function is not necessary a minimal specifying set. Moreover we showed that the number of minimal specifying sets can grow as $\Omega(n^2)$. Also for two-dimensional 2-threshold functions we showed that any function, that has a unique pair of threshold functions defining it, has at most 9 elements in a minimal specifying set. It would be interesting to estimate the proportion of the functions with this property in the class of 2-threshold functions.

Bibliography

[1] D. M. Acketa and J. Žunić. On the maximal number of edges of convex digital polygons included into an $m \times m$-grid. *Journal of Combinatorial Theory, Series A*, 69(2): 358–368, 1995.

[2] M. A. Alekseyev. On the number of two-dimensional threshold functions. *SIAM Journal on Discrete Mathematics*, 24(4): 1617–1631, 2010.

[3] M. A. Alekseyev, M. G. Basova, and N. Zolotykh. On the minimal teaching sets of two-dimensional threshold functions. *SIAM Journal on Discrete Mathematics*, 29(1): 157–165, 2015.

[4] D. Angluin. Queries and concept learning. *Machine learning*, 2(4): 319–342, 1998.

[5] M. Anthony. Boolean Functions and Artificial Neural Networks. *CDAM Research Report LSE-CDAM-2003-01*, 2003.

[6] M. Anthony, G. Brightwell, and J. Shawe-Taylor. On specifying Boolean functions by labelled examples. *Discrete Applied Mathematics*, 61(1):1–25, 1995.

[7] T. M. Apostol. *Introduction to Analytic Number Theory*, Springer, Berlin, 1976

[8] A. Astorino and A. Fuduli. Support vector machine polyhedral separability in semisupervised learning. *Journal of Optimization Theory and Applications*, 154(3): 1039–1050, 2013.

[9] A. Astorino and M. Gaudioso. Polyhedral separability through successive LP. *Journal of Optimization Theory and Applications*, 112(2): 265–293, 2002.

[10] E. B. Baum. Neural net algorithms that learn in polynomial time from examples and queries. *IEEE Transactions on Neural Networks*, 2(1): 5–19, 1991.

[11] K. P. Bennett and O. L. Mangasarian. Bilinear separation of two sets in n-space. *Computational Optimization and Applications*, 2(3): 207–227, 1993.

[12] W. J. Bultman and W. Maass. Fast identification of geometric objects with membership queries. *Information and Computation* 118(1): 48–64, 1995.

[13] C.K. Chow. On the characterization of threshold functions. *IEEE Symposium on Switching Circuit Theory and Logical Design*: 34–48, 1961.

[14] T. M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, EC-14(3): 326–334, 1965.

[15] Y. Crama and P. L. Hammer. Boolean functions: Theory, algorithms, and applications. *Cambridge University Press*: 2011.

[16] H. Davenport. On a principle of Lipschitz. *Journal of the London Mathematical Society*: 1(3), 179–183, 1951.

[17] I. Diakonikolas, D. M. Kane, and J. Nelson. Bounded independence fools degree-2 threshold functions. *2010 IEEE 51st Annual Symposium on Foundations of Computer Science. IEEE*: 11–20, 2010.

[18] M. M. Dundar, M. Wolf, S. Lakare, M. Salganicoff, and V. C. Raykar. Polyhedral classifier for target detection: a case study: colorectal cancer. *Proceedings of the 25th international conference on Machine learning*: 288–295, 2008.

[19] H. Edelsbrunner and F. Preparata. Minimum polygonal separation. *Information and Computation*, 77(3): 218–232, 1988.

[20] T. Eiter, T. Ibaraki, and K. Makino. Decision lists and related Boolean functions. *Theoretical Computer Science*, 270(1): 493–524, 2002.

[21] C. C. Elgot. Truth functions realizable by single threshold organs. In *Proceedings of the Second Annual Symposium on Switching Circuit Theory and Logical Design*, SWCT: 225–245, 1961.

[22] Y. Gérard. About the decidability of polyhedral separability in the lattice $\mathbb{Z}^d$. *Journal of Mathematical Imaging and Vision*, 59(1): 52–68, 2017.

[23] Y. Gérard. Recognition of digital polyhedra with a fixed number of faces is decidable in dimension 3. *Discrete Geometry for Computer Imagery*, 279–290, 2017.

[24] P. Hall. On Representatives of Subsets, *J. London Math. Soc*, 10 (1): 26–30, 1935.

[25] P. Haukkanen and J. K. Merikoski. Asymptotics of the number of threshold functions on a two-dimensional rectangular grid. *Discrete Applied Mathematics*, 161: 13–18, 2013.

[26] G. H. Hardy and E. M. Wright. An introduction to the theory of numbers. *Oxford University Press*, 1979.

[27] T. Hegedüs. Geometrical concept learning and convex polytopes. *Proceedings of the 7th Annual ACM Conference on Computational Learning Theory*, ACM Press, New York, 228–236, 1994.

[28] T. Hegedüs and P. Indyk. On learning disjunctions of zero-one threshold functions with queries. *Algorithmic Learning Theory. ALT 1997*, 1316: 446–460, 1997.

[29] S.-T. Hu. Threshold logic. *University of California Press, Berkeley*, 1965.

[30] A. S. Jarrah, B. Raposa, and R. Laubenbacher. Nested canalyzing, unate cascade, and polynomial functions. *Physica D: Nonlinear Phenomena*, 233(2): 167–174, 2007.

[31] S. Kauffman. The Origins of order: self-organization and selection in evolution. *Oxford University Press, New York, Oxford*, 1993.

[32] S. Kauffman, C. Peterson, B. Samuelsson, and C. Troein. Random Boolean network models and the yeast transcriptional network. *Proceedings of the National Academy of Sciences*, 100(25): 14796–14799, 2003.

[33] A. R. Klivans, R. O'Donnell, and R. A. Servedio. Learning intersections and thresholds of halfspaces. *Journal of Computer and System Sciences* , 68(4): 808–840, 2004.

[34] J. Koplowitz, M. Lindenbaum, and A. Bruckstein. The number of digital straight lines on an $N \times N$ grid. *IEEE Transactions on Information Theory*, 36(1): 192–197, 1990.

[35] J. Koplowitz and M. Lindenbaum. A new parameterization of digital straight lines. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 8: 847–852, 1991.

[36] S. Kwek and L. Pitt. PAC Learning intersections of halfspaces with membership queries. *Algorithmica*, 22(1-2): 53–75, 1998.

[37] Y. Li, J. O. Adeyeye, D. Murrugarra, B. Aguilar, and R. Laubenbacher. Boolean nested canalizing functions: A comprehensive analysis. *Theoretical Computer Science*, 481: 24–36, 2013.

[38] V. Lozin, I. Razgon, V. Zamaraev, E. Zamaraeva, and N. Zolotykh. Specifying a positive threshold function via extremal points. *Proceedings of the 28th International Conference on Algorithmic Learning Theory*: 208–222, 2017.

[39] V. Lozin, I. Razgon, V. Zamaraev, E. Zamaraeva, and N. Zolotykh. Linear read-once and related Boolean functions. *Discrete Applied Mathematics*, 250: 16–27, 2018.

[40] W. Maass and G. Turán. Algorithms and lower bounds for On-Line learning of geometric concepts. *Machine Learning*, 14: 251–269, 1994.

[41] N. Megiddo. On the complexity of polyhedral separability. *Discrete & Computational Geometry*, 3(4): 325–337, 1988.

[42] A. Ngom, I. Stojmenović, and J. Žunić. On the number of multilinear partitions and the computing capacity of multiple-valued multiple-threshold perceptrons. *IEEE Transactions on Neural Networks*, 14(3): 469–477, 2003.

[43] S. Olafsson and Y.S. Abu-Mostafa. The capacity of multilevel threshold functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(2): 277–281, 1988.

[44] V. Shevchenko and N. Zolotykh. Lower bounds for the complexity of learning half-spaces with membership queries. *International Conference on Algorithmic Learning Theory*, 1501: 61–71, 1998.

[45] V. Shevchenko and N. Zolotykh. On the complexity of deciphering the threshold functions of k-valued logic. *Doklady. Mathematics.*, 58(2): 268–270, 1998.

[46] V. Stetsenko. On almost bad Boolean bases. *Theoretical computer science*, 136(2): 419–469, 1994.

[47] J. Trainin. An elementary proof of Pick's theorem. *The Mathematical Gazette*, 91(522): 536–540, 2007.

[48] E. Zamaraeva. On teaching sets of k-threshold functions. *Information and Computation*, 251: 301–313, 2016.

[49] E. Zamaraeva. On teaching sets for 2-threshold functions of two variables. *Journal of Applied and Industrial Mathematics*, 11(1): 130–144, 2017.

[50] N. Zolotykh and V. Shevchenko. Estimating the complexity of deciphering a threshold functions in a k-valued logic. *Computational mathematics and mathematical physics*, 39(2): 328–334, 1999.

[51] J. Žunić. Note on the number of two-dimensional threshold functions. *SIAM Journal on Discrete Mathematics*, 25: 1266–1268, 2011.