FACULTY OF ARTS
Charles University

BACHELOR THESIS

Vít Fojtík

Lower Bounds on Boolean Formula Size

Dolní odhady velikosti Booleovských formulí

Department of Logic

Prague 2019

Thesis supervisor: Mgr. Pavel Hrubeš, Ph.D.

I would like to thank my supervisor, Dr. P. Hrubeš, for making sense of all my disorgainzed ideas, for always pointing me in a sensible direction, and especially for his plentiful, detailed remarks.

I declare that I carried out this bachelor thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date .... signature of the author

Title: Lower Bounds on Boolean Formula Size

Author: Vít Fojtík

Department: Department of Logic

Supervisor: Mgr. Pavel Hrubeš, Ph.D. , Institute of Mathematics of the Czech Academy of Sciences

Abstract: The aim of this thesis is to study methods of constructing lower bounds on Boolean formula size. We focus mainly on formal complexity measures, generalizing the well-known Krapchenko measure to a class of graph measures, which we thereafter study. We also review one of the other main approaches, using random restrictions of Boolean functions. This approach has yielded the currently largest lower bounds. Finally, we mention a program for finding super-polynomial bounds based on the KRW conjecture.

Keywords: Boolean formula, formal complexity measure, lower bounds

Abstrakt: Cílem této práce je studovat metody konstrukce dolních odhadů velikosti Booleovských formulí. Soustředíme se zde především na formální míry složitosti, přičemž zobecníme známou Krapchenkovu míru na třídu grafových měr, které následně studujeme. Zabýváme se také dalším z hlavních přístupů, využívající náhodné restrikce Booleovských funkcí. Na závěr zmíníme program pro nalezení super-polynomiálních odhadů založený na KRW doměnce.

Klíčová slova: Booleovská formule, formální míra složitosti, dolní odhady

Contents

List of Abbreviations

$\mathbb{N}$	The set of all natural numbers including 0
$\mathbb{R}^+$	The set of all non-negative real numbers
$\mathbb{Z}_2$	The finite field of size 2
$\log$	The binary logarithm
$f[A]$	The image of A under f
$f^{-1}(b)$	The fiber of b under f
^{T}	Transposition
$\mathbb{P}$	Probability
$\mathbb{E}$	Expected value

Introduction

Boolean formula size is a measure of complexity of a problem corresponding to the time required by parallel computations. One of the open problems in computational complexity is to find an explicit Boolean function with super-polynomial formula size. Even though it has been proved that such functions exist (and they even form the majority of all Boolean functions), no attempt to prove a super-polynomial bound on some function has so far been successful.

In this thesis, we study two of the main streams in construction of formula size lower bounds. Chapter 1 introduces the necessary background on Boolean formulas and reviews some examples of functions and their bounds. In Chapter 2 we study the notion of a formal complexity measure and generalize a common measure to a class of complexity measures. Thereafter, in Chapter 3 we look at an alternative way of thinking about complexity measures and review some boundaries of the complexity-measure approach. Finally, Chapter 4 reviews the other main approach to formula size and the progress towards finding super-polynomial bounds.

1. Boolean formula complexity

In this chapter we introduce Boolean functions, circuits and their measures of complexity, and finally Boolean formulas. Thereafter, we review some of the functions most common in complexity theory. We mostly uphold the conventions and notation of Wegener [1987].

Definition 1. *A Boolean function is any function $f : \{0,1\}^n \to \{0,1\}$ for some $n \in \mathbb{N}$. By B_n we denote the set of all n-dimensional Boolean functions, $B_n = \{f : \{0,1\}^n \to \{0,1\}\}$.*

In the theory of computation, Boolean functions are typically viewed as a mathematical representation of some task or problem, where we study the properties of some computational model solving this task—in our case, the size Boolean circuits and formulas.

Partially defined functions In some contexts it is useful to consider functions defined only on a subset of $\{0,1\}^n$.

Definition 2. *A partially defined Boolean function is any function $f : \{0,1\}^n \to \{0,1,?\}$ for some $n \in \mathbb{N}$. By B_n^* we denote the set of all n-dimensional partially defined Boolean functions, $B_n^* = \{f : \{0,1\}^n \to \{0,1,?\}\}$.*

For $f, g \in B_n^*$ we say that g is an extension of f, $f \subseteq g$, if for all $x \in \{0,1\}^n$ it holds that if $f(x) \in \{0,1\}$, then $f(x) = g(x)$.

Remark. Here, "$f(x) = ?$" is interpreted as "f is undefined on x". This notation is equivalent to defining partially defined functions as $f : M \subseteq \{0,1\}^n \to \{0,1\}$.

1.1 Boolean circuits and formulas

In this section we define Boolean circuits and formulas and formulate some basic findings about their size.

1.1.1 Boolean circuits

Boolean circuits model computation on a fixed number of input variables. Apart from the inputs, they consist of interconnected gates, each corresponding to some Boolean function. The set of admissible functions is called the basis. For the purposes of this thesis, we consider only the De Morgan basis, $\{\neg, \vee, \wedge\}$. The following definition is a reformulation of the one of Wegener [1987].

Definition 3. *Let $n \in \mathbb{N}$ be fixed. A (De Morgan) Boolean circuit is a weakly connected directed acyclic graph $S = (V, E)$ such that each $v \in V$ has a label P_v that is one of the following:*

- *$\vee$, or $\wedge$. In that case, v has two predecessors in E.*
- *$\neg$. Then, v has one predecessor in E.*
- *x_i for some $i \leq n$, 0, or 1. In that case, v has no predecessors in E.*

In the first two cases, v is called an inner gate. In the third case, v is called a leaf or an input.

Remark. Predecessors of v in E are all $u \in V$ such that $(u, v) \in E$.

Definition 4. *Let $S = (V, E)$ be a circuit on n variables and let $(y_1, \ldots, y_n) \in \{0,1\}^n$. We define inductively the value of the circuit on $y_1, \ldots, y_n$ on gate $v \in V$, denoted $\overline{v}(y_1, \ldots, y_n)$:*

- *If $P_v = 1$, $\overline{v}(y_1, \ldots, y_n) = 1$, if $P_v = 0$, $\overline{v}(y_1, \ldots, y_n) = 0$.*
- *If $P_v = x_i$, $\overline{v}(y_1, \ldots, y_n) = y_i$.*
- *If $P_v = \neg$ and u is the predecessor of v, $\overline{v}(y_1, \ldots, y_n) = 1 - \overline{u}(y_1, \ldots, y_n)$.*
- *If $P_v = \vee$ and v_1, v_2 are the predecessors of v,*

  $$\overline{v}(y_1, \ldots, y_n) = \max(\overline{v_1}(y_1, \ldots, y_n), \overline{v_2}(y_1, \ldots, y_n)),$$

  if $P_v = \wedge$ and v_1, v_2 are the predecessors of v,

  $$\overline{v}(y_1, \ldots, y_n) = \min(\overline{v}(y_1, \ldots, y_n), \overline{v}(y_1, \ldots, y_n)).$$

We say that the circuit computes a (partially defined) function f, if there exists a gate $v \in V$ such that for all $y \in \{0,1\}^n$ (such that $f(y) \in \{0,1\}$) $f(y) = \overline{v}\,(y)$.

Remark. The definition of the value is sound due to the fact that S is acyclic and therefore has a topological ordering that can be used for induction.

Bases In the thesis we consider exclusively circuits on the De Morgan basis $\{\neg, \wedge, \vee\}$. The two other most commonly used bases are the monotone basis $\{\wedge, \vee\}$ and the full basis B_2. Unlike the other two, the monotone basis is incomplete in the sense that there exist functions which are not computed by any $\{\wedge, \vee\}$-circuits. Since every monotone circuit is a De Morgan circuit and every De Morgan circuit is a B_2-circuit, finding some bounds is surely hardest for the B_2 basis and easiest for monotone circuits.

Circuit size and depth Among circuits computing the same function, the most natural measure of efficiency is the size—the number of gates. Another basic characteristic of a circuit is the length of the longest chain of gates connected by the predecessor relation, called the depth.

Definition 5. *Let $S = (V, E)$ be a circuit and let V_I be the set of all inner gates, $V_I = \{v \in V \mid P_v \in \{\neg, \vee, \wedge\}\}$. We define the size of S as $C(S) = |V_I|$ and the depth of S, $U(S)$, as the number of edges in the longest path in S.*

For $f \in B_n$ denote by $C(f)$ the smallest size of a circuit computing f and by $U(f)$ the least depth of a circuit computing f.

Remark. Circuit size can be interpreted as time of computation in serial computing, while depth corresponds to time of parallel computation.

1.1.2 Boolean formulas

Bounded fan-out In Boolean formulas, we bound the fan-out of each gate, which is the number of its successors in the predecessor relation. This can be interpreted as considering computation in which we cannot reuse intermediate results.

Definition 6. *Let $S = (V, E)$ be a circuit and $v \in V$. The fan-out of v is the number of $u \in V$ such that $(v, u) \in E$.*

De Morgan formulas Boolean formulas are usually defined as circuits such that each gate has fan-out at most 1. However, by repeated application the De Morgan laws, for every formula on the De Morgan basis there exists a formula computing the same function such that its $\neg$-gates have only variables as their predecessor, and that the number of leaves is equal in both formulas.

Since most sources assume De Morgan formulas in this form, we give a corresponding definition.

Definition 7. *A (De Morgan) Boolean formula is a weakly connected directed acyclic graph $S = (V, E)$ such that each $v \in V$ has fan-out at most 1 and has a label P_v that is one of the following:*

- *$\vee$, or $\wedge$. In that case, v has two predecessors in E.*
- *x_i for some $i \leq n$, $\neg x_i$ for some $i \leq n$, 0, or 1. In that case, v has no predecessors in E.*

The value of a formula is defined similarly as for circuits. The size of S is defined as $L^*(S) = |V_I|$. The number of leaves of S is $L(S) = L^*(S) + 1$.

For $f \in B_n$ we define the formula size (or complexity) of f, $L(f)$, as the least number of leaves of a formula computing f.

Remark. Formulas correspond to binary trees. The definition of the number of leaves of a formula is based on the fact that any binary tree with k inner vertices has $k + 1$ leaves.

The convention of using the number of leaves as formula size is due to the fact that it is often simpler to reason about than the number of gates, particularly when working with sub-formulas of a formula. In the following, we will refer to L both as the number of leaves and as the size.

Representation by a propositional formula Boolean formulas can be represented by propositional (De Morgan) formulas. The corresponding propositional formula for a Boolean formula consisting of a single gate v is

- T, for $P_v = 1$,
- $\perp$, for $P_v = 0$,
- x_i, for $P_v = x_i$.

A Boolean formula consisting of a $\vee$-gate connecting subformulas S_1 and S_2 is represented by $\varphi_1 \vee \varphi_2$, where φ_1 represents S_1 and φ_2 represents S_2.

The number of leaves corresponds to the number of occurrences of variables in the propositional formula representation. The value of a Boolean formula on some input is the value of the propositional formula under a corresponding valuation.

The importance of formula size As already mentioned, the formula complexity of a function can be interpreted as time required by serial computation that is not allowed to reuse intermediate results. However, the following theorem gives formula complexity a more important role.

Theorem 1 (Spira [1971])**.** *There exists a $k \in \mathbb{N}\backslash\{0\}$ such that for all $n \in \mathbb{N}\backslash\{0\}$ and for all $f \in B_n$*

$$\log (L(f)) \leq U(f) \leq k \log (L(f))$$

Remark. This theorem connects formula size to general circuit depth. Therefore, studying formula size gives us insight into time requirements of parallel computations.

1.2 Examples of Boolean functions

In this section we review some Boolean functions commonly used in formula complexity, along with bounds on their formula size.

1.2.1 Basic Boolean functions

Projections Projections are functions that give one of the input variables as their output.

Definition 8. *The i-th n-dimensional projection $p_i^n \in B_n$ is defined for all $(x_1, \dots, x_n) \in \{0,1\}^n$ as $p_i^n(x_1, \dots, x_n) = x_i$.*

Remark. When context allows it, we will use simply p_i instead of p_i^n. A similar convention will be applied for all following functions.

Projections and their negations are the simplest (non-constant) functions in B_n. Their complexity is clearly 1, since they are computed by a formula consisting of a single gate labeled x_i.

The parity function Some of the first bounds on formula size of a particular function were found for the parity function (Subbotovskaya [1961], Khrapchenko [1971]). It returns 1 iff the number of input variables equal to 1 is odd.

Definition 9. *The n-dimensional parity function, $\bigoplus_n \in B_n$, is defined for all $(x_1, \dots, x_n) \in \{0,1\}^n$ as*

$$\bigoplus_n (x_1, \dots, x_n) = x_1 \oplus \cdots \oplus x_n,$$

where $\oplus$ denotes addition in $\mathbb{Z}_2$.

Remark. It was proved by Khrapchenko [1971] that the formula size of parity is at least n^2 up to a constant. We cover this result in more detail in Chapter 2. The following simple construction gives an upper bound on the formula size of parity.

Proposition 2. *Let $n \in \mathbb{N}$. If $n = 2^b$, $L(\bigoplus) \leq n^2$. Otherwise, $L(\bigoplus) \leq 4n^2$.*

Proof. To prove the first part, it is enough to find formula of size n^2 computing parity. We proceed by induction on b. If $b = 0$ ($n = 1$), parity is equal to the first 1-dimensional projection and is therefore computed by a formula consisting of a single gate labeled x_1.

For $n = 2^{b+1}$, let S_1 and S_2 be optimal formulas computing the parity function on $x_1, \dots, x_{\frac{n}{2}}$ and on $x_{\frac{n}{2}+1}, \dots, x_n$, respectively. By the induction hypothesis, $L(S_1) \leq \frac{n^2}{4} \geq L(S_2)$. Let φ_1 and φ_2 be the propositional formula representations of S_1 and S_2. Then, $(\varphi_1 \wedge \neg\varphi_2) \vee (\neg\varphi_1 \wedge \varphi_2)$ represents (after propagating $\neg$ into the subformulas) a formula computing the parity function on $x_1, \dots, x_n$ and the number of its leaves is equal to $4 \cdot \max(L(S_1), L(S_2)) \leq n^2$.

The second part can be proved from the first using the fact that for any $n \in \mathbb{N} \setminus \{0\}$ there exists a b such that $n \leq 2^b < 2n$. From the first part, there exists a formula S computing parity on 2^b inputs of size at most $2^{2b} \leq 4n^2$. By substituting all the gates in S labeled by x_k, $k > n$, for 0-gates, we get a formula computing the parity function on n variables of size at most $4n^2$.

□

Modulo functions A natural generalization of the parity function is a function that returns 1 iff the number of input variables equal to 1 equals 1 modulo k.

Definition 10. *The n-dimensional modulo k function, MOD_k^n, is defined for all $(x_1, \dots, x_n) \in \{0,1\}^n$ as*

$$MOD_k^n(x_1, \dots, x_n) = \begin{cases} 1 & \text{if } \sum_{i=1}^n x_i \equiv 1 \text{ (mod } k\text{)}, \\ 0 & \text{otherwise.} \end{cases}$$

Remark. In Chapter 2, we show quadratic lower bounds on $L(\mathrm{MOD}_k)$ for $k = 3$ and 4.

The majority function Another frequently appearing function is majority, a function that returns 1 iff at least half of the input variables are equal to 1.

Definition 11. *The n-dimensional majority function, MAJ_n, is defined for all $(x_1, \dots, x_n) \in \{0,1\}^n$ as*

$$MAJ_n(x_1, \dots, x_n) = \begin{cases} 1 & \text{if } \sum_{i=1}^n x_i \geq \frac{n}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

Remark. Despite its simple definition and the fact that it has been studied from the beginnings of formula complexity, the majority function still has a large gap between the best lower and upper bounds on its formula size—Gál et al. [2018] report the current best (De Morgan) bounds as $n^{3.91}$ (upper, due to Sergeev [2016]) and n^2 (lower, proved in Chapter 2), both up to a constant. Because of this gap, majority has been proposed as a candidate function for super-quadratic formula size (for example by Ueno [2015]).

Threshold functions Again, we can generalize this notion to a function giving 1 iff at least k inputs are equal to 1.

Definition 12. *The n-dimensional k-th threshold function, T_k^n, is defined for all $(x_1, \ldots, x_n) \in \{0,1\}^n$ as*

$$T_k^n\left(x_1, \ldots, x_n\right) = \begin{cases} 1 & \text{if } \sum_{i=1}^n x_i \geq k, \\ 0 & \text{otherwise.} \end{cases}$$

Remark. In Chapter 2, we prove a bound of the form $k\left(n-k+1\right)$.

The element distinctness function All of the functions above (except for projections) are symmetric, meaning their output does not depend on the order of inputs. As an example of a non-symmetric function, the element distinctness function divides the input into blocks of equal length and checks if any two blocks are equal.

Definition 13. *Let $n = 2^k \cdot 2k$. The n-dimensional element distinctness function, ED_n, is defined for all $(x_1, \ldots, x_{2^k}) \in \{0,1\}^{2^k \times 2k}$ ($x_i \in \{0,1\}^{2k}$) as*

$$ED_n\left(x_1, \ldots, x_{2^k}\right) = \begin{cases} 1 & \text{if } \forall i \neq j \leq 2^k\text{: } x_i \neq x_j, \\ 0 & \text{otherwise.} \end{cases}$$

Remark. A method introduced by Nechiporuk [1966] gives a bound of $\frac{n^2}{\log n}$ up to a constant for the ED_n function. In Chapter 2, we use the function to show the limitations of our methods.

1.2.2 Operations on Boolean functions

There are many ways to combine functions into a new one. For example, we can insert the outputs of some functions into a function from B_2 or B_1.

Definition 14. *For $f, g \in B_n$ we define $\neg f = 1 - f$ and $f \vee g = \max(f, g)$.*

Theorem 3. *Let $f, g \in B_n$. Then,*

(i) $L\left(\neg f\right) = L\left(f\right)$,

(ii) $L\left(f \vee g\right) \leq L\left(f\right) + L\left(g\right)$.

Proof. (i) Let $S = (V, E)$ be an optimal formula for f. We define a formula $\neg S$ by switching all $\vee$-gates for $\wedge$-gates (and vice versa) and all input literals for their negations. Then, $\neg S$ computes $\neg f$ (by the De Morgan laws) and its size is equal to $L(S)$. Therefore,

$$L(\neg f) \leq L(\neg S) = L(S) = L(f).$$

Because $\neg\neg f = f$, $L\left(\neg f\right) \geq L\left(f\right)$.

(ii) Let S^f and S^g be optimal formulas for f and g, respectively, with propositional representations φ^f, φ^g. Consider the formula S represented by $\varphi^f \vee \varphi^g$. S clearly computes $f \vee g$ and therefore its size is at least $L\left(f \vee g\right)$, that is

$$L\left(f \vee g\right) \leq L\left(S\right) = L\left(f\right) + L\left(g\right).$$

□

Definition 15. *Let $f, g \in B_n$ and $h \in B_2$, we define $h(f,g) \in B_n$ for all $x \in \{0,1\}^n$ as*

$$h(f,g)(x) = h(f(x), g(x))$$

Remark. It does not always hold that

$$L(h(f,g)) \leq L(f) + L(g).$$

For example, let $n = 2$ and take f as the first projection, p_1, g as the second projection, p_2, and h as $\oplus$ (addition in $\mathbb{Z}_2$). Then,

$$L(f) + L(g) = 2.$$

In order for the inequality to hold, the $\oplus$ would have to be computed by a formula of size at most two, that is, with one inner gate. Only such formulas correspond to disjunctions and conjunctions of literals, none of which compute $\bigoplus$ (both the disjunction and the conjunction are constant on three out of four input combinations, while $\bigoplus$ returns 1 on two of them).

However, based on Theorem 3 we can prove the following result. It is a special case of a well known theorem stated in the next subsection (Theorem 5).

Proposition 4. *Let $h \in B_2$ and $f, g \in B_n$. Then,*

$$L(h(f,g)) \leq L(h) \cdot \max(L(f), L(g)).$$

Proof. Assume $L(f) \geq L(g)$. Let S^h, S^f and S^g be optimal formulas for h, f and g, respectively. Define formula S by replacing all instances of x_1 in S^h by S^f, all instances of $\neg x_1$ by $\neg S^f$ and similarly for x_2 and S^g. S clearly computes $h(f,g)$. We know that

$$L\left(\neg S^f\right) = L\left(S^f\right) \geq L(S^g) = L(\neg S^g).$$

Because for each leaf we substitute a formula of size at most $L(f)$ and the total number of all leaves is $L(h)$, we get

$$L(h(f,g)) \leq L(S) \leq L(h) \cdot L(f).$$

□

Remark. The proposition could easily be generalized to $h \in B_k$ for any k.

Notice that if we take for example $f = g$, the resulting complexity can be much smaller than $L(h) \cdot L(f)$. This is due to the fact that f and g are allowed to share input variables. In the next subsection, we consider a similar situation with the difference that f and g each have their own set of variables.

1.2.3 Composite functions

Composition of functions is a central concept in modern techniques for bounding formula size. Most of the current super-quadratic bounds were proved for some sort of composite functions. The idea of composition is to take outputs of independent copies of the same function as the input for another function.

Definition 16. *Let $b, m \in \mathbb{N} \setminus \{0\}$, $f \in B_b$, $g \in B_m$, and $n = bm$. The composition of f and g is $f \circ g \in B_n$ defined for all $x = (x_{1,1}, \dots, x_{1,m}, \dots, x_{b,1}, \dots, x_{b,m}) \in \{0,1\}^n$ as*

$$(f \circ g)(x) = f\left(g\left(x_{1,1}, \dots, x_{1,m}\right), \dots, g\left(x_{b,1}, \dots, x_{b,m}\right)\right).$$

Reformulating Proposition 4 in terms of composition, we get a well known simple result.

Theorem 5. *Let $f \in B_b$ and $g \in B_m$. Then,*

$$L(f \circ g) \leq L(f) \cdot L(g).$$

The proof is essentially the same as for Proposition 4.

The KRW conjecture The converse inequality of Theorem 5 is far more interesting. A program for proving super-polynomial lower bounds on formula size of specific functions has been proposed by Karchmer et al. [1995] that is based on proving what is now generally known as the KRW conjecture,

$$L(f \circ g) \geq L(f) \cdot L(g),$$

up to a constant. It has been proved for several special cases of g, and bounds on specific functions have been found that hold under the conjecture. We will review these results in more detail in Chapter 4.

The Andreev function The first super-quadratic bound was proved by Andreev [1987] for a function that is not composite itself, but the proof is based on composite functions. It uses a function that interprets a part of its input as a truth table of some function (in lexicographic order) and applies a transformation of the rest of the input to this function.

Definition 17. *Assume that $n = 2^b$ and at the same time $n = b \cdot m$ (that is, $b = 2^l$). For $y = (y_{1,1}, \dots y_{1,m}, \dots y_{b,1}, \dots y_{b,m}) \in \{0,1\}^n$ define $g(y) = k \in \mathbb{N}$ as the natural number whose (lexicographic) binary representation is $\left(\bigoplus_m \left(y_{1,1}, \dots, y_{m,1}\right), \dots, \bigoplus_m \left(y_{1,b}, \dots, y_{m,b}\right)\right)$.*

The $2n$-dimensional Andreev function, $A_{2n} \in B_{2n}$, is defined for all $(x, y) = (x_0, \dots, x_{n-1}, y_{1,1}, \dots y_{1,m}, \dots y_{b,1}, \dots y_{b,m}) \in \{0,1\}^{2n}$ as

$$A_{2n}(x) = x_{g(y)}.$$

Remark. The first bound on this function of $n^{2.5-o(1)}$ up to a constant was proved by Andreev [1987]. Later, this bound was improved to $n^{3-o(1)}$ by Håstad [1993] and $\frac{n^3}{(\log n)^2 \log \log n}$ up to a constant by Tal [2014]. We review some of the techniques applied in Chapter 4.

2. Formal complexity measures

One of the predominant approaches in construction of lower bounds on formula size uses the abstract notion of a formal complexity measure. This method dates back to Khrapchenko [1971], who implicitly used a complexity measure to prove a quadratic bound for the parity function. A formal complexity measure is any function on B_n that is at most 1 on all projections p_i and that behaves similarly to L for operations on functions.

Definition 18. *Let $n \in \mathbb{N} \setminus \{0\}$ be fixed. A function $m : B_n \to \mathbb{R}^+$ is a formal complexity measure, if*

- *$\forall i \leq n$: $m(p_i) \leq 1$,* *(Normalization)*
- *$\forall f \in B_n$: $m(f) = m(\neg f)$,* *(Symmetry)*
- *$\forall f, g \in B_n$: $m(f \vee g) \leq m(f) + m(g)$.* *(Subadditivity)*

We denote by C_n the set of all formal complexity measures on B_n.

The following well-known theorem states, that L is the greatest complexity measure.

Theorem 6. *For all $n \in \mathbb{N} \setminus \{0\}$, $L|B_n$ is a formal complexity measure, $L|B_n \in C_n$. Additionally, for all $m \in C_n$ and all $f \in B_n$, $m(f) \leq L(f)$.*

Proof. Symmetry and subadditivity of L follows from Theorem 3, normalization is trivial (remark after Definition 8).

We prove the second part by induction on $L(f)$. If $L(f) = 1$, f is a projection or its negation and by normalization (and symmetry): $m(f) \leq 1 = L(f)$. Let $L(f) = k$ and assume for all g such that $L(g) < k$ that $L(g) \geq m(g)$. Assume the optimal formula for f consists of a $\vee$-gate connecting subformulas S_1, S_2, computing functions f_1, f_2. Then,

$$\begin{aligned} L(f) &= L(S) \\ &= L(S_1) + L(S_2) \\ &\geq L(f_2) + L(f_2) \\ &\geq m(f_1) + m(f_2) \\ &\geq m(f_1 \vee f_2) \\ &= m(f). \end{aligned}$$

□

Remark. As a consequence of Theorem 6, any lower bound on any complexity measure is a lower bound on L. We can therefore proceed by constructing measures that are easier to compute (or bound) than L.

Properties of C_n Before continuing on to specific measures, we review a few simple properties of measures in general that follow from the definition.

Proposition 7. *Let $n \in \mathbb{N} \setminus \{0\}$. Then,*

(i) $\forall m \in C_n \forall f, g \in B_n$*:* $m(f \wedge g) \leq m(f) + m(g)$*,*

(ii) $\forall m_1, m_2 \in C_n \forall \lambda_1, \lambda_2 \in [0,1], \lambda_1 + \lambda_2 = 1$*:* $(\lambda_1 m_1 + \lambda_2 m_2) \in C_n$*,*

(iii) $\forall m_1, m_2 \in C_n$*:* $\max(m_1, m_2) \in C_n$*.*

Remark. Statements (ii) and (iii) can be generalized to any finite number of measures.

In the definition of formal complexity measures, the symmetry condition is sometimes replaced by statement (i)—this weaker condition still satisfies Theorem 6. We will see a similar situation in Chapter 3.

Normalization Any real function on B_n satisfying the symmetry and subadditivity conditions can be divided by the maximal value of a projection to obtain a complexity measure.

Proposition 8. *Let $\widetilde{m} : B_n \to \mathbb{R}^+$ be a function satisfying the symmetry and subadditivity conditions. Then, $m : B_n \to \mathbb{R}^+$ defined for all $f \in B_n$ as*

$$m(f) = \frac{\widetilde{m}(f)}{\max_{i \leq n} \widetilde{m}(p_i)}$$

is a formal complexity measure.

2.1 The Krapchenko measure

The Krapchenko measure is one of the most renowned complexity measures in formula complexity, mainly due to its relative computational simplicity. It formalizes the idea that functions that assign different values to inputs close to each other are complex—here, distance of inputs is interpreted by way of the Hamming metric.

Definition 19. *The Hamming metric, $d_H : (\{0,1\}^n)^2 \to \mathbb{N}$, is defined for all $(x,y) \in (\{0,1\}^n)^2$ as the number of coordinates in which x and y differ,*

$$d_H(x,y) = |x \oplus^n y|_1,$$

where $\oplus^n$ denotes addition in $\mathbb{Z}_2^n$ and $|z|_1$ equals the number of coordinates of z equal to 1*.*

For $A, B \subseteq \{0,1\}^n$ and $k \leq n$ we denote

$$H_k(A,B) = \{(x,y) \in A \times B \mid d_H(x,y) = k\}.$$

Remark. The graph $(\{0,1\}^n, H_1(\{0,1\}^n, \{0,1\}^n)))$, whose vertices are n-vectors on $\{0,1\}$ and edges are pairs of vectors of Hamming distance 1, is commonly referred to as the n-dimensional (unit) hypercube.

Definition 20. *Let $n \in \mathbb{N} \setminus \{0\}$. For all $A, B \subseteq \{0,1\}^n$ disjoint we define*

$$P_1\left(A,B\right) = \frac{|H_1\left(A,B\right)|^2}{|A \times B|}.$$

The n-dimensional Krapchenko measure, $K : B_n \to \mathbb{R}^+$, is defined for all $f \in B_n$ as

$$K(f) = \max\{P_1\left(A,B\right) \mid A \subseteq f^{-1}(0) \ \&\ B \subseteq f^{-1}(1)\}.$$

Lemma 9. *For all $a, b, c, d \in \mathbb{R}^+$,*

$$\frac{(a+b)^2}{c+d} \leq \frac{a^2}{c} + \frac{b^2}{d}.$$

Proof. Multiplying both sides by the denominators, the claim is equivalent to

$$cd(a+b)^2 \leq a^2d(c+d) + b^2c(c+d),$$

which can be rewritten as

$$a^2cd + 2abcd + b^2cd \leq a^2cd + a^2d^2 + b^2c^2 + b^2cd,$$

which is equivalent to

$$0 \leq a^2d^2 - 2abcd + b^2c^2 = (ad - bc)^2.$$

This is always true, so the claim holds.

□

Remark. As the following theorem states, the Krapchenko measure is indeed a formal complexity measure. We reproduce the proof from Wegener [1987], which we generalize in Section 2.3.

Theorem 10. *The Krapchenko measure K is a formal complexity measure.*

Proof. Let $i \leq n$ and let $A \subseteq p_i^{-1}(0)$ and $B \subseteq p_i^{-1}(1)$. For each $x \in A$ there exists only one $y \in \{0,1\}^n$ such that $y_i \neq x_i = 1$ and $d_H(x,y) = 1$. This implies that $|H_1(A,B)| \leq \min\left(|A|, |B|\right)$, so

$$P_1(A,B) \leq \frac{\min\left(|A|,|B|\right)^2}{|A||B|} \leq 1.$$

Therefore, $K\left(p_i\right) \leq 1$.

Symmetry follows from the fact that $|H_1(A,B)| = |H_1(B,A)|$.

Let $f, g \in B_n$ and let $A \subseteq f^{-1}(0)$, $B \subseteq f^{-1}(1)$ be optimal for $K(f \vee g)$, so that $K(f \vee g) = P_1(A,B)$. Then, $f[A] = \{0\} = g[A]$ and there exist B_1, B_2 disjoint such that $B_1 \cup B_2 = B$ and $f[B_1] = \{1\} = g[B_2]$. We get $|H_1(A,B)| = |H_1(A,B_1)| + |H_1(A,B_2)|$ and $|A \times B| = |A|\left(|B_1| + |B_2|\right)$. By Lemma 9,

$$\begin{aligned}
K\left(f \vee g\right) &= P_1\left(A,B\right) \\
&= \frac{1}{|A|}\frac{\left(|H_1(A,B_1)| + |H_1(A,B_2)|\right)^2}{|B_1| + |B_2|} \\
&\leq \frac{|H_1(A,B_1)|^2}{|A \times B_1|} + \frac{|H_1(A,B_2)|^2}{|A \times B_2|} \\
&= P_1(A,B_1) + P_1(A,B_2).
\end{aligned}$$

Because $A \subseteq f^{-1}(0) \cap g^{-1}(0)$ and $B_1 \subseteq f^{-1}(1)$, $B_2 \subseteq g^{-1}(1)$, we get $P_1(A, B_1) + P_1(A, B_2) \leq K(f) + K(g)$.

□

Corollary. *For all $n \in \mathbb{N} \setminus \{0\}$ and all $f \in B_n$*

$$K(f) \leq L(f).$$

Computing the Krapchenko measure We mention some simple observations that will be useful in the following examples. For any A and B, $H_1(A, B)$ can be computed as the sum over $x \in A$ of the number of neighbors of x in B ($y \in B$ is a neighbor of x if its distance from x is equal to 1),

$$H_1(A, B) = \sum_{x \in A} |\{y \in B \mid d_H(x, y) = 1\}| = \sum_{x \in A} |H_1(\{x\}, B)|.$$

If for all $x \in A$: $|H_1(\{x\}, B)| = c$, then $H_1(A, B) = c|A|$.

If $|A| = |B|$, $P_1(A, B)$ equals to the square of the average over x in A of the number of neighbors of x in B. That is,

$$P_1(A, B) = \left(\frac{1}{|A|}\left(\sum_{x \in A} |H_1(\{x\}, B)|\right)\right)^2.$$

Parity The Krapchenko measure provides a simple way to prove a quadratic bound for parity.

Proposition 11. *For all $n \in \mathbb{N} \setminus \{0\}$,*

$$L\left(\bigoplus_n\right) \geq n^2.$$

Proof. Set $A = \bigoplus^{-1}(0)$, $B = \bigoplus^{-1}(1)$. For all $x \in A$, all n neighbors of x are in B (changing any one bit of x changes its parity). Furthermore, $|A| = |B| = 2^{n-1}$. So,

$$P_1(A, B) = \left(\frac{1}{|A|}\left(\sum_{x \in A} n\right)\right)^2 = n^2.$$

□

Remark. In most cases, we do not necessarily need to find the exact value of $K(f)$. Since $P_1(A, B)$ is a lower bound on $K(f)$ for all (suitable) A, B, it is sufficient to find A and B giving a large $P_1(A, B)$. Here, however, we have found the maximal value of P_1 (follows from theorem 14).

Propositions 2 and 11 together imply that for $n = 2^b$, formula size of parity is exactly n^2.

Threshold functions Now we apply the same technique to a general threshold function, T_k^n.

Definition 21. *For all $n \in \mathbb{N} \setminus \{0\}$ and all $k \leq n$, the n-dimensional k-slice, $S_k^n \subseteq \{0,1\}^n$, is the set of elements of $\{0,1\}^n$ with exactly k coordinates equal to 1,*

$$S_k^n = \{x \in \{0,1\}^n \mid |x|_1 = k\}.$$

Remark. Symmetric Boolean functions can be defined as being constant on all slices. For all n and k the size of the k-slice is equal to the number of ways to select k out of n coordinates, that is

$$|S_k^n| = \binom{n}{k}.$$

For any $x \in S_k^n$, the number of neighbors of x in S_{k+1}^n is equal to $n-k$ (the number of 0-coordinates that can be increased to 1) and the number of its neighbors in S_{k-1}^n is equal to k (the number of 1-coordinates that can be decreased to 0).

Proposition 12. *For all $n \in \mathbb{N} \setminus \{0\}$ and all $k \leq n$,*

$$L\left(T_k^n\right) \geq k(n-k+1).$$

As a special case for $k = \lceil \frac{n}{2} \rceil$ (the majority function),

$$L(MAJ_n) \geq \begin{cases} \frac{n^2}{4} + \frac{n}{2} & \text{if } n \text{ is even,} \\ \frac{n^2}{4} + \frac{n}{2} + \frac{1}{4} & \text{if } n \text{ is odd.} \end{cases}$$

Proof. Let $n \in \mathbb{N} \setminus \{0\}$ and $k \leq n$, $k > 0$. Set $A = S_{k-1}^n$ and $B = S_k^n$. Then, $T_k[A] = \{0\}$ and $T_k[B] = \{1\}$. For any $x \in A$, the number of neighbors of x in B is equal to $n-(k-1)$, so $|H_1(A,B)| = \binom{n}{k-1}(n-k+1)$. Therefore,

$$P_1\left(A,B\right) = \frac{\left(\binom{n}{k-1}(n-k+1)\right)^2}{\binom{n}{k-1}\binom{n}{k}} = \left((n-k+1)\right)^2 \frac{k}{n-k+1} = k(n-k+1).$$

□

Modulo 3 We show one more example, bounding the MOD_3 function.

Proposition 13. *Let $n = 3m$, then*

$$L(MOD_3) \geq \frac{n^2}{2}.$$

Proof. Set $A = \{x \in \{0,1\}^n \mid |x|_1 \mod 3 = 0 \vee |x|_1 \mod 3 = 2\}$ and $B = \{y \in \{0,1\}^n \mid |y|_1 \mod 3 = 1\}$. Any $y \in B$ has all of its n neighbors in A. Furthermore,

$$|B| = \sum_{k=0}^{n/3} \binom{n}{3k+1} = \frac{1}{3}\left(2^n - o(n)\right).$$

Then, $|A| \approx \frac{2}{3}2^n$, so we get

$$P_1\left(A, B\right) = \frac{\left(n\frac{1}{3}2^n\right)^2}{\frac{1}{3}2^n\frac{2}{3}2^n} = \frac{n^2}{2}.$$

□

Limitations of K We have seen that the Krapchenko measure provides a simple way of obtaining lower bounds on formula size. However, the bounds it provides cannot be more than quadratic.

Theorem 14. *Let $n \in \mathbb{N} \setminus \{0\}$ and $f \in B_n$. Then,*

$$K(f) \leq n^2.$$

Proof. Since each element of $\{0,1\}^n$ has n Hamming neighbors and since for all A, B disjoint: $|H_1(A,B)| = \sum_{x\in A}|H_1(\{x\},B)| = \sum_{y\in B}|H_1(A,\{y\})|$, it follows that $H_1(A,B) \leq \min(n|A|, n|B|)$. This implies a quadratic upper bound on the Krapchenko measure of any function, because

$$P_1(A,B) \leq \frac{\left(\min\left(n|A|, n|B|\right)\right)^2}{|A||B|} \leq n^2.$$

□

The Krapchenko measure on partially defined functions We can view P_1 as a measure on partially defined Boolean functions in the following sense: If we extend the definition of a complexity measure to all partially defined functions, the function $\kappa : B_n^* \to \mathbb{R}^+$, $\kappa(\widetilde{f}) = P_1(\widetilde{f}^{-1}(0), \widetilde{f}^{-1}(1))$ would be a complexity measure. Furthermore, there is a correspondence between K and κ:

$$\begin{aligned} \forall f \in B_n : K(f) &= \max\{\kappa(\widetilde{f}) \mid \widetilde{f} \in B_n^* \ \&\ \widetilde{f} \subseteq f\}, \\ \forall \widetilde{f} \in B_n^* : \kappa(\widetilde{f}) &\leq \min\{K(f) \mid f \in B_n \ \&\ \widetilde{f} \subseteq f\}. \end{aligned}$$

The first line follows from the definition of K, the second from the fact that for any $f \supseteq \widetilde{f}$, we can set $A = \widetilde{f}^{-1}(0)$ and $B = \widetilde{f}^{-1}(1)$). On the second line, equality may hold under additional conditions, but there exist partial functions with $\kappa(\widetilde{f}) = 0$ (while $K(f)$ is always positive).

In Chapter 3 we generalize the idea of measures defined for all partially defined functions and their duality with complexity measures.

2.2 Hamming distance measures

In this section we introduce a set of complexity measures based on the Hamming metric.

Definition 22. *Let $n \in \mathbb{N} \setminus \{0\}$ and $k \leq n$. For all $A, B \subseteq \{0,1\}^n$ disjoint we define*

$$P_k\left(A, B\right) = \frac{|H_k\left(A, B\right)|^2}{|A \times B|}.$$

We define $\widetilde{K}_k : B_n \to \mathbb{R}^+$ for all $f \in B_n$ as

$$\widetilde{K}_k\left(f\right) = \max\{P_k\left(A, B\right) \mid A \subseteq f^{-1}(0) \;\&\; B \subseteq f^{-1}(1)\}.$$

Remark. Even though $\widetilde{K}_k$ satisfies the symmetry and subadditivity conditions (which will follow from a more general theorem in section 2.3), it needs to be normalized in order to become a measure.

Theorem 15. *For all $n \in \mathbb{N} \setminus \{0\}$, all $k \leq n$ and all $i \leq n$,*

$$\widetilde{K}_k\left(p_i\right) = \binom{n-1}{k-1}^2.$$

Proof. Set $A = p_i^{-1}(0)$, $B = p_i^{-1}(1)$. For any $x \in A$, its k-neighbors (neighbors in the $H_k(\{0,1\}^n, \{0,1\}^n)$ relation) in B are all y that differ in the i-th coordinate and in $k-1$ of the $n-1$ remaining coordinates, so $|H_k(\{x\}, B)| = \binom{n-1}{k-1}$.

$$P_k(A, B) = \frac{\left(|A|\binom{n-1}{k-1}\right)^2}{|A||B|} = \binom{n-1}{k-1}^2.$$

Hence, $\widetilde{K}_k\left(p_i\right) \geq \binom{n-1}{k-1}^2$.

On the other hand, for all $A \subseteq p_i^{-1}(0)$ and $B \subseteq p_i^{-1}(1)$, $|H_k(A, B)| \leq \min\left(\binom{n-1}{k-1}|A|, \binom{n-1}{k-1}|B|\right)$, so

$$P_k(A, B) \leq \frac{\min\left(|A|\binom{n-1}{k-1}, |B|\binom{n-1}{k-1}\right)^2}{|A||B|} \leq \binom{n-1}{k-1}^2.$$

Therefore, $\widetilde{K}_k\left(p_i\right) \leq \binom{n-1}{k-1}^2$.

□

Definition 23. *The n-dimensional Hamming k-measure, $K_k : B_n \to \mathbb{R}^+$, is defined for all $f \in B_n$ as*

$$K_k\left(f\right) = \frac{\widetilde{K}_k\left(f\right)}{\binom{n-1}{k-1}^2}.$$

Modulo 4 We will use the Hamming 2-measure to find a bound for the MOD_4 function.

Example 1. For simplicity assume $n = 4m$. Set $A = \{x \in \{0,1\}^n \mid |x|_1 \mod 4 = 3\}$ and $B = \{y \in \{0,1\}^n \mid |y|_1 \mod 4 = 1\}$. Let x be an element of the slice S_k^n, where $k \mod 4 = 3$. Then, x has $\binom{k}{2}$ 2-neighbors in S^n_{k-2} and $\binom{n-k}{2}$ 2-neighbors in S^n_{k+2}. So, $|H_2(S^n_k, B)| = \binom{n}{k}\left(\binom{k}{2} + \binom{n-k}{2}\right) = \binom{n}{k}(\frac{1}{2}n(n-1) - kn + k^2)$.

$$\begin{aligned} H_2(A,B) = & \left(\sum_{k=0}^{n/4} \binom{n}{4k+3}\left(\binom{4k+3}{2} + \binom{n-4k-3}{2}\right)\right) \\ = & \frac{1}{4}(2^{n-1}n(n-1) - 2^{n-1}n^2 + 2^{n-2}n(n+1) + o(n)) \\ \approx & \; 2^{n-3}n^2. \end{aligned}$$

We have $|A| = |B| = 2^{n-2}$, so

$$\frac{P_2(A,B)}{(n-1)^2} = \frac{(2^{n-3}n^2)^2}{(n-1)^2(2^{n-2})^2} = \frac{1}{4}(n^2 + 2n + o(1)).$$

Example 2. For comparison we find a bound for the same function using K. Set $A = \{x \in \{0,1\}^n \mid |x|_1 \mod 4 = 0 \vee |x|_1 \mod 4 = 2\}$ and $B = \{y \in \{0,1\}^n \mid |y|_1 \mod 4 = 1\}$. Any element of B has all of its neighbors in A and $|A| = 2^{n-1}$, $|B| = 2^{n-2}$. Therefore,

$$P_1(A,B) = \frac{\left(2^{n-2}n\right)^2}{2^{n-2}2^{n-1}} = \frac{n^2}{2}.$$

Remark. Not only was the second computation much simpler, it also yielded a slightly better bound.

Proposition 16. *For all $n \in \mathbb{N} \setminus \{0\}$,*

$$L\left(MOD^n_4\right) \geq \frac{n^2}{2}.$$

The following theorem describes limitations of Hamming measures.

Theorem 17. *Let $n \in \mathbb{N} \setminus \{0\}$, $k \leq n$ and $f \in B_n$. Then,*

$$K_k(f) \leq \frac{n^2}{k^2}.$$

Proof. Every $x \in \{0,1\}^n$ has $\binom{n}{k}$ k-neighbors, which means that $|H_k(A,B)| \leq \min\left(\binom{n}{k}|A|, \binom{n}{k}|B|\right)$. Therefore,

$$\frac{P_k(A,B)}{\binom{n-1}{k-1}^2} \leq \left(\frac{\binom{n}{k}}{\binom{n-1}{k-1}}\right)^2 = \left(\frac{n}{k}\right)^2.$$

□

Remark. This implies that K_k cannot give super-linear bounds for large k.

2.3 Graph measures

We can view $K_k(f)$ as counting edges of a graph of elements of distance k from each other, $G_k = (\{0,1\}^n, H_k(\{0,1\}^n, \{0,1\}^n))$, intersected with the complete bipartite graph $(\{0,1\}^n, f^{-1}(0) \times f^{-1}(1))$. In this section we generalize this idea to arbitrary graphs.

Definition 24. *Let $n \in \mathbb{N} \setminus \{0\}$ and let $G = (\{0,1\}^n, E)$ be an undirected graph on $\{0,1\}^n$. For all $A, B \subseteq \{0,1\}^n$ disjoint we define*

$$P_G(A,B) = \frac{|E \cap (A \times B)|^2}{|A \times B|}.$$

We define $\widetilde{K_G} : B_n \to \mathbb{R}^+$ for all $f \in B_n$ as

$$\widetilde{K_G}(f) = \max\{P_G(A,B) \mid A \subseteq f^{-1}(0) \,\&\, B \subseteq f^{-1}(1)\}.$$

Proposition 18. *For all $n \in \mathbb{N} \setminus \{0\}$ and $G = (\{0,1\}^n, E)$, $\widetilde{K_G}$ is symmetric and subadditive.*

Proof. The proof is an adaptation of the proof of Theorem 10. Symmetry follows from the fact that G is undirected, so $|E \cap (A \times B)| = |E \cap (B \times A)|$.

Let $f, g \in B_n$ and let $A \subseteq f^{-1}(0)$, $B \subseteq f^{-1}(1)$ be optimal for $\widetilde{K_G}(f \vee g)$, so that $\widetilde{K_G}(f \vee g) = P_G(A,B)$. Then, $f[A] = \{0\} = g[A]$ and there exist B_1, B_2 disjoint such that $B_1 \cup B_2 = B$ and $f[B_1] = \{1\} = g[B_2]$. We get $|E \cap (A \times B)| = |E \cap (A \times B_1)| + |E \cap (A \times B_2)|$ and $|A \times B| = |A|\,(|B_1| + |B_2|)$. By Lemma 9,

$$\begin{aligned}
\widetilde{K_G}(f \vee g) &= P_G(A,B) \\
&= \frac{1}{|A|}\frac{(|E \cap (A \times B_1)| + |E \cap (A \times B_2)|)^2}{|B_1| + |B_2|} \\
&\leq \frac{|E \cap (A \times B_1)|^2}{|A \times B_1|} + \frac{|E \cap (A \times B_2)|^2}{|A \times B_2|} \\
&= P_G(A, B_1) + P_G(A, B_2).
\end{aligned}$$

Because $A \subseteq f^{-1}(0) \cap g^{-1}(0)$ and $B_1 \subseteq f^{-1}(1)$, $B_2 \subseteq g^{-1}(1)$, we get $P_G(A,B_1) + P_G(A,B_2) \leq \widetilde{K_G}(f) + \widetilde{K_G}(g)$.

□

Definition 25. *Let $p_G = \max\{\widetilde{K_G}(p_i) \mid i \leq n\}$. The complexity measure induced by G, K_G, is defined for all $f \in B_n$ as*

$$K_G(f) = \frac{\widetilde{K_G}(f)}{p_G}.$$

Remark. The following is a corrolary of Proposition 18 along with Proposition 8.

Theorem 19. *For all $n \in \mathbb{N} \setminus \{0\}$ and all $G = (\{0,1\}^n, E)$, K_G is a formal complexity measure.*

Modulo 4 We now revisit the MOD_4 function, using a graph measure to find another bound.

Example 3. Let $n = 2m$ and let e_i be the i-th unit vector. Define $E = \{(x, x + e_{2i-1} + e_{2i}) \mid i \leq \frac{n}{2}\}$, $G = (\{0,1\}^n, E)$. G is a subgraph of H_2 in which all neighbors differ in one coordinate pair out of $\frac{n}{2}$ fixed pairs. Since each x has only one neighbor that differs in the i-th coordinate for all i, $p_G = 1$.

Let $A = \{x \in \{0,1\}^n \mid |x|_1 \mod 4 = 3\}$ and $B = \{y \in \{0,1\}^n \mid |y|_1 \mod 4 = 1\}$. For all $x \in A$ and $i \leq \frac{n}{2}$, $x + e_{2i-1} + e_{2i}$ is an element of B iff $x_{2i-1} = x_{2i}$. For i fixed, the number of all such x in the slice S^n_k is $\binom{n-2}{k-2} + \binom{n-2}{k}$. So,

$$|E \cap (A \times B)| = \sum_{k=0}^{n/4} \frac{n}{2} \left(\binom{n-2}{4k+3-2} + \binom{n-2}{4k+3} \right) = \frac{n}{2}(2^{n-3} - o(n)).$$

Because $|A| = |B| = 2^{n-3}$, we get

$$P_G(A, B) = \frac{n^2}{4} - o(1).$$

Remark. Again, we obtained a bound of $\frac{n^2}{4}$, which is weaker than the $\frac{n^2}{2}$ given by K. The graph from the last example cannot give larger bounds than $\frac{n^2}{4}$, because each of its vertices has $\frac{n}{2}$ neighbors (see the proof of Theorem 17).

Element distinctness Since it is not symmetric, the element distinctness function (Definition 13) cannot be effectively bounded by way of counting over slices as in most of the previous examples.

Example 4. Let $n = 2^k 2k$. For all $x \in \{0,1\}^n$ we denote by $x = (x_1, \ldots, x_{2^k})$ the partition of x into 2^k vectors of length $2k$. For $i \neq j \leq 2^k$ define $\widetilde{x}_{i,j}$ as the vector created by substituting x_j for x_i in x. Finally, we define $E = \{(x, \widetilde{x}_{i,j}) \mid i \neq j \leq 2^k \ \&\ x_i \neq x_j\}$, $G = (\{0,1\}^n, E)$.

Let us consider the r-th projection and let $i = r \div 2k$, $l = r \mod 2k$. For any x such that $(x_i)_l = 0$, if x_i is equal to some x_j, then all y such that for all $s \neq i$: $y_s = x_s$ and $(y_i)_l = 1$ are neighbors of x. It follows that $p_G \leq (2^{2k-1})^2$.

Define $B = \{y \mid \forall i \neq j : y_i \neq y_j\}$, all 2^k-tuples of distinct $2k$-vectors, and $A = \{\widetilde{y}_{i,j} \mid y \in B\}$, which equals $\{x \mid \exists i \neq j\ (x_i = x_j \ \&\ \forall \{r, s\} \neq \{i, j\} : x_r \neq x_s)\}$, the set of all 2^k-tuples of $2k$-vectors such that exactly one pair of vectors is identical. Any $y \in B$ has $2^k(2^k - 1)$ neighbors in A (the number of ways to choose a j and an i for substituting y_j for y_i). The size of B equals to the number of all 2^k-tuples of distinct $2k$-vectors, $|B| = 2^{2k}(2^{2k} - 1) \ldots (2^{2k} - 2^k + 1) = \frac{2^{2k}!}{(2^{2k}-2^k)!}$. Similarly, $|A| = \binom{2^k}{2} \frac{2^{2k}!}{(2^{2k}-2^k+1)!}$. Therefore,

$$P_G(A, B) = \frac{\left(2^k(2^k - 1)\frac{2^{2k}!}{(2^{2k}-2^k)!}\right)^2}{\frac{2^{2k}!}{(2^{2k}-2^k)!}\binom{2^k}{2}\frac{2^{2k}!}{(2^{2k}-2^k+1)!}} = 2(2^{2k} - 2^k + 1)2^k(2^k - 1).$$

Normalizing the result, we get

$$\frac{P_G(A, B)}{p_G} \geq \frac{2(2^{2k} - 2^k + 1)2^k(2^k - 1)}{(2^{2k-1})^2} \approx 8 - o(1).$$

So, despite the promising $P_G(A, B)$ we have not been able to find a bound on the size of ED_n due to the huge value of p_G.

Example 5. We now try to obtain a bound using the Krapchenko measure. Set $A = \{x \mid \exists i \neq j\, (x_i = x_j \;\&\; \forall \{r,s\} \neq \{i,j\} : x_r \neq x_s)\}$ and $B = \{y \mid \forall i \neq j : y_i \neq y_j\}$ (same as in the previous example).

Denote $U = \{(x,l) \in A \times \{1,\dots,n\} \mid x + e_l \in A \;\&\; \exists j \neq l \div 2k : x_j = x_{l \div 2k}\}$, the set of all pairs of a vector and a coordinate such that the coordinate belongs to one of the two identical subvectors and changing the coordinate will create a new pair of identical subvectors. Let $U_l = \{x \mid (x,l) \in U\}$.

For any $2k$-vector z, the number of all vectors with distance from z at least two is $2^{2k} - 2k - 1$. The number of all $(2^k - 2)$-tuples of distinct vectors with distance from z at least two is $\frac{(2^{2k}-2k-1)!}{(2^k-2)!}$. Finally, the number of all (2^k-2)-tuples of distinct vectors such that at least one has distance one from z and none is identical to z is the number of all (2^k-2)-tuples of distinct vectors non-equal to z minus the number of all (2^k-2)-tuples of distinct vectors with distance from z at least two, $\frac{(2^{2k}-1)!}{(2^k-2)!} - \frac{(2^{2k}-2k-1)!}{(2^k-2)!}$. It follows that

$$|U_l| = (2^k-1)2^{2k}\left(\frac{(2^{2k}-1)!}{(2^k-2)!} - \frac{(2^{2k}-2k-1)!}{(2^k-2)!}\right)$$

and $|U| = 2^k 2k |U_l|$. Then,

$$\begin{aligned} |H_1(A,B)| &= 4k|A| - |U| \\ &= 4k\binom{2^k}{2}\frac{2^{2k}!}{(2^{2k}-2^k+1)!} - 4k\binom{2^k}{2}2^{2k}\frac{(2^{2k}-1)! - (2^{2k}-2k-1)!}{(2^k-2)!} \\ &\approx 4k\binom{2^k}{2}\frac{2^{2k}!}{(2^{2k}-2^k+1)!} \end{aligned}$$

and we get

$$P_1(A,B) \geq \frac{\left(4k\binom{2^k}{2}\frac{2^{2k}!}{(2^{2k}-2^k+1)!}\right)^2}{\binom{2^k}{2}\frac{\left(2^{2k}\right)!\left(2^{2k}\right)!}{\left(2^{2k}-2^k\right)!\left(2^{2k}-2^k+1\right)!}} = \frac{8k^2 2^k(2^k-1)}{2^{2k}-2^k+1} = 8k^2(1-o(1)) \approx 2\,(\log n)^2$$

We can see that using formal complexity measures to construct bounds on non-symmetric functions can be complicated and not very effective. We might slightly improve the bound if we used $B = \{y \mid \forall i \neq j : y_i \neq y_j \;\&\; \exists i \neq j : d_H(x_i, x_j) = 1\}$ instead.

The role of K In all the examples, the Krapchenko measure proved more effective than any other graph measure. Even though we believe the Krapchenko measure might be the greatest graph measure, we have not been able to prove it, nor find a counterexample.

Limitations of graph measures In Chapter 3 we show that all graph measures give at most quadratic bounds (Theorem 27).

Theorem 20. *There exists $a \in \mathbb{R}^+$ such that for all $n \in \mathbb{N} \setminus \{0\}$, all graphs $G = (\{0,1\}^n, E)$ and all $f \in B_n$*

$$K_G(f) \leq an^2.$$

3. Combinatorial rectangles

Combinatorial rectangles provide an alternative way of thinking about Boolean functions. This point of view was first applied by Rychkov [1985]. In this chapter we introduce combinatiorial rectangles and their measures and show their correspondence to Boolean functions and formal complexity measures. Thereafter, we review the results of Karchmer et al. [1992] and Hrubeš et al. [2010], giving an upper bound of $\frac{9}{8}(n^2+n)$ for a class of rectangle measures, including all graph-induced measures.

Definition 26. *Let $S_0, S_1 \subseteq \{0,1\}^n$, $S = S_0 \times S_1$ is an n-dimensional combinatiorial rectangle, if $S_0 \cap S_1 = \emptyset$. A subrectangle of S is a subset of S that is a rectangle. We denote the set of all n-dimensional rectangles by R_n and the set of all subrectangles of some $S \in R_n$ by $R_n(S)$.*

Example 6. Let $f \in B_n$, we can define the rectangle of f as $S_f = f^{-1}(0) \times f^{-1}(1)$. Since $f^{-1}(0) \cap f^{-1}(1) = \emptyset$, S_f is a combinatorial rectangle.

Definition 27. *Let $f \in B_n$ and $S = S_0 \times S_1 \in R_n$. We say that f separates S if $f[S_0] \subseteq \{0\}$ and $f[S_1] \subseteq \{1\}$.*

Remark. Equivalently, f separates S if $S \subseteq S_f$.

Definition 28. *A rectangle $M \in R_n$ is monochromatic if it is separated by some projection or its negation. We denote by M_n the set of all monochromatic rectangles and by $M_n(S)$ the set of all monochromatic subrectangles of S, for any $S \in R_n$.*

Remark. $M \in R_n$ is monochromatic iff there exists some $i \leq n$ and $b \in \{0,1\}$ such that for all $x \in M_0$: $x_i = b$ and for all $y \in M_1$: $y_i = 1 - b$.

Remark. For a partially defined Boolean function $\widetilde{f} \in B_n$ we can define $S_{\widetilde{f}} = \widetilde{f}^{-1}(0) \times \widetilde{f}^{-1}(1)$. Then, for each $S \in R_n$ there exists $\widetilde{f} \in B_n$ such that $S = S_{\widetilde{f}}$, namely

$$\widetilde{f}(x) = \begin{cases} 1 & \text{if } x \in S_1, \\ 0 & \text{if } x \in S_0, \\ ? & \text{otherwise.} \end{cases}$$

This is a one-to-one correspondence, except for constant partial functions, which all correspond to the empty rectangle $\emptyset$.

3.1 Rectangle measures

In this section we introduce an analogue of formal complexity measures for combinatorial rectangles and review some basic examples of rectangle measures.

Definition 29. *A function $\mu : R_n \to \mathbb{R}^+$ is a rectangle measure if*

- *$\forall M \in M_n$: $\mu(M) \leq 1$,* *(Normalization)*
- *$\forall S = S_0 \times S_1 \in R_n$: $\mu(S_0 \times S_1) = \mu(S_1 \times S_0)$,* *(Symmetry)*

- *$\forall S \in R_n\ \forall S^1, S^2 \in R_n$ disjoint such that $S^1 \cup S^2 = S$: $\mu(S) \leq \mu(S^1) + \mu(S^2)$.* *(Subadditivity)*

Remark. Here, subadditivity is a stronger condition than the one we considered for complexity measures, as it encompasses both $\vee$-subadditivity and $\wedge$-subadditivity: any $S = S_0 \times S_1$ can be divided into two disjoint subrectangles in two ways, as $(S_0 \times S_1^1) \cup (S_0 \times S_1^2)$, $S_1^1 \cap S_1^2 = \emptyset$, or $(S_0^1 \times S_1) \cup (S_0^2 \times S_1)$, $S_0^1 \cap S_0^2 = \emptyset$. The first corresponds to disjunction, the second to conjunction. Under symmetry, however, the notions are equivalent.

Symmetry is not a necessary condition for rectangle measures, but we keep it in order for the definition to correspond to that of complexity measures. Without symmetry it corresponds to the more general definition of formal complexity measures mentioned after Proposition 7.

Relating rectangle and complexity measures The following theorem states that any formal complexity measure induces a rectangle measure and vice versa.

Theorem 21. *Let $n \in \mathbb{N} \setminus \{0\}$.*

(i) For all formal complexity measures $m \in C_n$, μ_m defined for all $S \in R_n$ as

$$\mu_m(S) = \min\{m(f) \mid f \in B_n \text{ separates } S\}$$

is a rectangle measure and for all $f \in B_n$: $\mu_m(S_f) = m(f)$.

(ii) For all rectangle measures μ, m_μ defined for all $f \in B_n$ as

$$m_\mu(f) = \max\{\mu(S) \mid f \text{ separates } S \in R_n\}$$

is a formal complexity measure and for all $f \in B_n$: $\mu(S_f) = m_\mu(f)$.

Proof. (i) Normalization: Let $M \in M_n$ be monochromatic, separated by a projection or its negation q. Since m is normalized and symmetric, $m(q) \leq 1$, so

$$\mu_m(M) \leq m(q) \leq 1.$$

Symmetry: Follows from the symmetry of m and the fact that f separates $S_0 \times S_1$ iff $\neg f$ separates $S_1 \times S_0$.

Subadditivity: Let $S^1, S^2 \in R_n$ form a disjoint partition of $S \in R_n$ and assume $S_0 = S_0^1 = S_0^2$ (otherwise use S_1). Let f and $g \in B_n$ be optimal separating functions for S^1 and S^2, respectively, that is $\mu_m(S^1) = m(f)$ and $\mu_m(S^2) = m(g)$. Then, $f \vee g$ separates S, so

$$\mu_m(S^1) + \mu_m(S^2) = m(f) + m(g) \geq m(f \vee g) \geq \mu_m(S).$$

$\mu_m(S_f) = m(f)$ follows from the fact that f is the only function separating S_f.

(ii) Normalization: For $i \leq n$, the optimal rectangle $S \in R_n$ separated by p_i (such that $m_\mu(p_i) = \mu(S)$) is separated by a projection, which makes it monochromatic. Therefore,

$$m_\mu(p_i) \leq \mu(S) \leq 1.$$

Symmetry: Follows from the symmetry of μ and the fact that f separates $S_0 \times S_1$ iff $\neg f$ separates $S_1 \times S_0$.

Subadditivity: Let $S = S_0 \times S_1$ be the optimal rectangle for $f \vee g$. There exist $S_1^1, S_1^2 \subseteq S_1$ disjoint such that $S_1 = S_1^1 \cup S_1^2$, $f[S_1^1] \subseteq \{1\}$ and $g[S_1^2] \subseteq \{1\}$. That is, f separates $S^1 = S_0 \times S_1^1$ and g separates $S^2 = S_0 \times S_1^2$. Furthermore, S^1 and S^2 form a disjoint partition of S. So,

$$m_\mu(f \vee g) = \mu(S) \leq \mu(S^1) + \mu(S^2) \leq m_\mu(f) + m_\mu(g).$$

Again, $\mu(S_f) = m_\mu(f)$ follows from the fact that f is the only function separating S_f.

□

Theorem 22. *Let us denote $\Lambda = \mu_L$, the rectangle measure induced by L. Then, for all rectangle measures μ and all $S \in R_n$:*

$$\Lambda(S) \geq \mu(S).$$

Proof. The proof is an extension of the proof of Theorem 6. We proceed by induction on $\Lambda(S) \in \mathbb{N}$. If $\Lambda(S) = 1$, S is separated by a function computed by a formula of size 1, that is by a projection or its negation, which means S is monochromatic and $\mu(S) \leq 1$.

Now let $\Lambda(S) = n > 1$ and let the following hold: for all $R \in R_n$, if $\Lambda(R) < n$, then $\Lambda(R) \geq \mu(R)$. Let f be the optimal function separating S, that is, $\Lambda(S) = L(f)$, and let F be the optimal formula for f. Without loss of generality, the last gate in F is a $\vee$-gate connecting subformulas F_1 and F_2. Let F_1 and F_2 compute functions f_1 and f_2, respectively. Then, for $S = S_0 \times S_1$, there exists a disjoint partition $S_1 = S_1^1 \cup S_1^2$ such that $f_1[S_1^1] \subseteq \{1\}$ and $f_2[S_1^2] \subseteq \{1\}$, so f_1 separates $S^1 = S_0 \times S_1^1$ and f_2 separates $S^2 = S_0 \times S_1^2$. Altogether, we get

$$\begin{aligned} \mu(S) &\leq \mu(S^1) + \mu(S^2) \\ &\leq \Lambda(S^1) + \Lambda(S^2) \\ &\leq L(f_1) + L(f_2) \\ &\leq L(F_1) + L(F_2) \\ &= L(F) \\ &= \Lambda(S). \end{aligned}$$

□

Remark. This means Λ plays the same role for rectangle measures as L does for complexity measures. Λ was introduced by Karchmer and Wigderson [1990] as a measure denoting the minimal number of leaves in a communication protocol for a Karchmer–Wigderson game on a rectangle S.

The partition number An important example of a rectangle measure is the partition number, the minimal number of disjoint monochromatic rectangles covering a rectangle.

Definition 30. *For $n \in \mathbb{N} \setminus \{0\}$, the partition number, $D : R_n \to \mathbb{N}$, is defined for all $S \in R_n$ as*

$$D(S) = \min\{k \mid \exists M^1, \ldots, M^k \in M_n \textit{ disjoint: } S = \bigcup_{i=1}^{k} M^i\}.$$

Remark. The partition number is maximal on a subset of measures that satisfy a stronger form of subadditivity.

Definition 31. *A function $\mu : R_n \to \mathbb{R}^+$ is strongly subadditive, if for all $S \in R_n$ and all $S^1, \ldots, S^k \in R_n$ disjoint such that $S = \bigcup_{i=1}^{k} S^i$ it holds that*

$$\mu(S) \leq \sum_{i=1}^{k} \mu(S^i).$$

Theorem 23 (Rychkov [1985])**.** *For all $n \in \mathbb{N} \setminus \{0\}$, D is a strongly subadditive rectangle measure. Additionally, for all strongly subadditive rectangle measures μ and all $S \in R_n$,*

$$\mu(S) \leq D(S).$$

Remark. D is a rectangle measure, so $D(S_f) \leq L(f)$ holds for all f. On the other hand, Aho et al. [1983] showed that $\log L(f) \leq (\log D(S_f))^2$, which implies that D can yield super-polynomial bounds. Therefore, strongly additive measures may suffice for finding large bounds. The gap between L and D has been intensely studied. For example, a rectangle measure has been constructed by Ueno [2010] that can surpass D on some rectangles, meaning the gap is probably non-trivial.

The Krapchenko rectangle measure and graph measures We close this section by reformulating the Krapchenko measure and graph measures for rectangles.

Definition 32. *For $n \in \mathbb{N} \setminus \{0\}$, the n-dimensional Krapchenko rectangle measure, κ_n, is defined for all non-empty $S = S_0 \times S_1 \in R_n$ as*

$$\kappa_n(S) = \frac{|H_1(S_0, S_1)|^2}{|S|}$$

Remark. The definition of K can be reinterpreted as $K = m_\kappa$, meaning for all $f \in B_n$: $K(f) = \max\{\kappa(S) \mid f \text{ separates } S \in R_n\}$.

Definition 33. *For $n \in \mathbb{N} \setminus \{0\}$ and $G = (\{0,1\}^n, E)$ an undirected graph on $\{0,1\}^n$, we denote*

$$\widetilde{\delta}_G(S) = \frac{|E \cap S|^2}{|S|}.$$

The rectangle measure induced by G, δ_G, is defined for all $S \in R_n$ as

$$\delta_G(S) = \frac{\widetilde{\delta}_G(S)}{\max_{M \in M_n} \widetilde{\delta}_G(M)}.$$

3.2 Convexity

It was proved by Karchmer et al. [1992] that a certain rectangle measure, the fractional partition number, is at most quadratic. Later, Hrubeš et al. [2010] showed that the fractional partition number is the largest in a set of measures called convex measures and that this set encompasses many of the most common measures.

Definition 34. *Let us denote by χ_S the characteristic function of any $S \in R_n$, that is, for all $e \in \{0,1\}^n \times \{0,1\}^n$*

$$\chi_S(e) = \begin{cases} 1 & \text{if } e \in S, \\ 0 & \text{otherwise.} \end{cases}$$

Let $n \in \mathbb{N} \setminus \{0\}$, $S, S^1, \ldots, S^k \in R_n$ and $r_1, \ldots, r_k \in [0,1]$. We say that $S^1, \ldots, S^k$ is a fractional cover of S with weights $r_1, \ldots r_k$, denoted $S = \sum_{i=1}^k r_i S^i$, if for all $e \in \{0,1\}^n \times \{0,1\}^n$

$$\chi_S(e) = \sum_{i=1}^k r_i \chi_{S^i}(e).$$

Remark. Equivalently, $S^1, \ldots, S^k$ is a fractional cover of S if $S^1, \ldots, S^k$ are subrectangles of S and for all $e \in S$

$$\sum_{i | e \in S^i} r_i = 1.$$

Definition 35. *A function $\mu : R_n \to \mathbb{R}^+$ is convex if for all $S \in R_n$, all $S^1, \ldots, S^k \in R_n$ and all $r_1, \ldots, r_k \in [0,1]$ such that $S = \sum_{i=1}^k r_i S^i$*

$$\mu(S) \leq \sum_{i=1}^k r_i \mu(S^i).$$

Definition 36. *For $n \in \mathbb{N} \setminus \{0\}$, the fractional partition number, $D^* : R_n \to \mathbb{R}^+$, is defined for all $S \in R_n$ as*

$$D^*(S) = \min\{\sum_{i=1}^k r_i \mid \exists M^1, \ldots, M^k \in M_n : S = \sum_{i=1}^k r_i M^i\}.$$

Theorem 24 (Karchmer et al. [1992])**.** *For all $n \in \mathbb{N} \setminus \{0\}$, D^* is a convex rectangle measure. Additionally, for all convex rectangle measures μ and all $S \in R_n$,*

$$\mu(S) \leq D^*(S).$$

Remark. Convexity implies strong subadditivity, since any disjoint partition $S = \bigcup_{i=1}^k S^i$ is a fractional partition with all weights equal to 1. Therefore, $D^*(S) \leq D(S)$ for all $S \in R_n$.

The following theorem, bounding convex measures, was first proved by Karchmer et al. [1992]. We present the version of Hrubeš et al. [2010] with a slightly better constant.

Theorem 25 (Hrubeš et al. [2010])**.** *For all $n \in \mathbb{N} \setminus \{0\}$ and all $S \in R_n$,*

$$D^*(S) \le \frac{9}{8}(n^2 + n).$$

Corollary. *For all convex rectangle measures μ and all $S \in R_n$,*

$$\mu(S) \le \frac{9}{8}(n^2 + n).$$

Example 7. Let us show that κ is convex. Let $S \in R_n$, $S = \sum_{i=1}^k r_i S^i$. Then, $|S| = \sum_{i=1}^k r_i |S^i|$ and $|H_1(S_0, S_1)| = \sum_{i=1}^k r_i |H_1(S_0^i, S_1^i)|$ (see the example after definition 37 for proof). Since for all $a_1, \dots, a_k, b_1, \dots, b_k$:

$$\frac{\left(\sum_{i=1}^k a_i\right)^2}{\sum_{i=1}^k b_i} \le \sum_{i=1}^k \frac{a_i^2}{b_i}$$

(can be proved by induction from Lemma 9), we get

$$\begin{aligned}
\kappa(S) &= \frac{|H_1(S_0, S_1)|^2}{|S|} \\
&= \frac{\left(\sum_{i=1}^k r_i |H_1(S_0^i, S_1^i)|\right)^2}{\sum_{i=1}^k r_i |S^i|} \\
&\le \sum_{i=1}^k \frac{\left(r_i |H_1(S_0^i, S_1^i)|\right)^2}{r_i |S^i|} \\
&= \sum_{i=1}^k r_i \frac{|H_1(S_0^i, S_1^i)|^2}{|S^i|} \\
&= \sum_{i=1}^k r_i \kappa(S^i).
\end{aligned}$$

Sufficient conditions for convexity Several conditions were formulated by Hrubeš et al. [2010] under which a measure is convex and therefore cannot give stronger than quadratic bounds. We apply one of them to graph measures.

Definition 37. *A function $\mu : R_n \to \mathbb{R}^+$ is additive, if for all $S \in R_n$, all $S^1, \dots, S^k \in R_n$ and all $r_1, \dots, r_k \in [0, 1]$ such that $S = \sum_{i=1}^k r_i S^i$*

$$\mu(S) = \sum_{i=1}^k r_i \mu(S^i).$$

Example 8. For any graph $G = (\{0,1\}^n, E)$, the function $\nu_G : S \mapsto |E \cap S|$ is additive. Let $S = \sum_{i=1}^{k} r_i S^i$, then

$$\begin{aligned}
|E \cap S| &= \sum_{e \in (\{0,1\}^n)^2} \chi_{E \cap S}(e) \\
&= \sum_{e \in (\{0,1\}^n)^2} \chi_E(e)\chi_S(e) \\
&= \sum_{e \in (\{0,1\}^n)^2} \chi_E(e) \left(\sum_{i=1}^{k} r_i \chi_{S^i}(e) \right) \\
&= \sum_{i=1}^{k} r_i \left(\sum_{e \in (\{0,1\}^n)^2} \chi_E(e)\chi_{S^i}(e) \right) \\
&= \sum_{i=1}^{k} r_i |E \cap S^i|.
\end{aligned}$$

As a corollary, the function $\nu : S \mapsto |S|$ is additive.

The following is a special case of a theorem by Hrubeš et al. [2010] and it gives a criterion for convexity of measures composed of additive functions.

Theorem 26. *Let $F : \mathbb{R}^+ \to \mathbb{R}^+$ be convex and nondecreasing and let $\mu_1, \mu_2 : R_n \to \mathbb{R}^+$ be additive and symmetric. Then, $\mu : R_n \to \mathbb{R}^+$ defined for all $S \in R_n$ as*

$$\mu(S) = \mu_2(S) F\left(\frac{\mu_1(S)}{\mu_2(S)} \right)$$

is subadditive, symmetric and convex.

Example 9. Let us finally show that graph measures are convex. For any graph $G = (\{0,1\}^n, E)$, define $\nu_G(S) = |E \cap S|$, $\nu(S) = |S|$. We have seen that ν_G and ν are additive and they are also clearly symmetric. Set $F(x) = x^2$, F is convex and nondecreasing on $\mathbb{R}^+$. Then,

$$\tilde{\delta}_G(S) = \frac{|E \cap S|^2}{|S|} = \nu(S) F\left(\frac{\nu_G(S)}{\nu(S)} \right)$$

is convex, so δ_G is also convex.

Theorem 27. *For all $n \in \mathbb{N} \setminus \{0\}$, all graphs $G = (\{0,1\}^n, E)$ and all $S \in R_n$*

$$\delta_G(S) \leq \frac{9}{8}(n^2 + n).$$

Remark. Theorem 20 is a corollary of theorem 27.

4. Super-quadratic bounds

One of the other main streams in lower bounds on formula size, apart from formal complexity measures, is based on the shrinkage of formulas under random restrictions. This method has yielded the current largest bounds, which are also the only super-quadratic bounds. In this chapter we give a brief overview of the main results and frontiers of this approach.

4.1 Random restrictions

The method of random restrictions was first used by Subbotovskaya [1961] to prove a bound of $n^{1.5}$ for parity the parity function. In this section we introduce the key points of the method and illustrate it on some examples.

Definition 38. *A restriction on n variables is an element of $Q_n = \{0, 1, *\}^n$. For $\rho \in Q_n$ and $x \in \{0,1\}^n$ we define $\rho(x)$ by*

$$\rho(x)_i = \begin{cases} x_i & \text{if } \rho_i = *, \\ \rho_i & \text{otherwise.} \end{cases}$$

For $f \in B_n$ and $\rho \in Q_n$ we define the restriction of f by ρ, $f|_\rho : \{0,1\}^n \to \{0,1\}$, for all $x \in \{0,1\}^n$ as

$$f|_\rho(x) = f(\rho(x)).$$

We denote by Q_k^n the set of all n-dimensional restrictions leaving k variables unrestricted, $Q_k^n = \{\rho \in Q_n \mid |\rho|_* = k\}$.

Even though there exist many variations, the following is the most commonly used definition of random restrictions.

Definition 39. *For $p \in (0,1)$, an n-dimensional random p-restriction is defined as a random vector $\boldsymbol{\varrho} = (\varrho_1, \ldots, \varrho_n)^T \in Q_n$ such that $\varrho_1, \ldots, \varrho_n$ are independent identically distributed random variables satisfying*

$$\mathbb{P}(\varrho_1 = *) = p,$$

$$\mathbb{P}(\varrho_1 = 0) = \mathbb{P}(\varrho_1 = 1) = \frac{1-p}{2}.$$

The following theorem, commonly referred to as the shrinkage lemma, is central to the method of random restrictions. We present the version proved by Subbotovskaya [1961] and Håstad [1993]. It states that the formula size of a random restriction of a function decreases from the original formula on average by a factor of p^2.

Theorem 28. *There exists $c > 0$ such that for all $p \in (0,1)$, all random p-restrictions $\boldsymbol{\varrho}$ and all $f \in B_n$*

$$\mathbb{E}\left[L\left(f|_{\boldsymbol{\varrho}}\right)\right] \leq c\left(p^2 L(f)\right).$$

Remark. Subbotovskaya [1961] proved a version of this lemma with $p^{1.5}$ on the right hand side. Later, Håstad [1993] proved the actual value of the exponent to be to 2.

For the purpose of our examples we will use a simplified version of random restrictions with a corresponding shrinkage lemma.

Definition 40. *Let $k \leq n$. The n-dimensional k-variable random restriction is a random vector $\boldsymbol{\varrho}_k$ distributed uniformly on Q_k^n, that is, for all $\rho \in Q_k^n$*

$$\mathbb{P}\left(\boldsymbol{\varrho} = \rho\right) = \frac{1}{|Q_k^n|} = \frac{1}{\binom{n}{k}2^{n-k}}.$$

Theorem 29. *For all $k \leq n$ and all $f \in B_n$*

$$\mathbb{E}\left[L\left(f|_{\boldsymbol{\varrho}_k}\right)\right] \leq c\left(\frac{k}{n}\right)^2 L(f).$$

Example 10. Let us show a simple application of the shrinkage lemma to the parity function by Subbotovskaya [1961]. Consider for any $\rho \in Q_1^n$ the restriction $\bigoplus|_\rho$. It is equal to some projection or its negation and therefore has formula size 1. We get

$$c\left(\frac{1}{n}\right)^2 L\left(\bigoplus\right) \geq \mathbb{E}\left[L\left(\bigoplus|_{\boldsymbol{\varrho}_1}\right)\right] = \sum_{\rho\in Q_k^n} \frac{1}{|Q_k^n|} = 1.$$

Multiplying both sides by n^2,

$$L\left(\bigoplus\right) \geq cn^2.$$

Example 11. We can try to use the shrinkage lemma to improve our bounds on the majority function. Assume $n = 2m$. For all $k \leq n$ and all $\rho \in Q_k^n$, $Maj|_\rho$ is equal to some threshold (or constant) function. Namely, let $i = |\rho|_1$, then $Maj|_\rho = T_{m-i}^k$ for $i \leq \min(n-k, m)$ and $i \geq \max(0, m-k)$ (otherwise it is constant). Since $K(f) \leq L(f)$ for all f,

$$\mathbb{E}\left[K\left(f|_{\boldsymbol{\varrho}_k}\right)\right] \leq \mathbb{E}\left[L\left(f|_{\boldsymbol{\varrho}_k}\right)\right] \leq c\left(\frac{k}{n}\right)^2 L(f).$$

In Chapter 2 we proved that $K(T_k^n) \geq k(n-k+1)$, so for k and ρ as above, $K(Maj|_\rho) = K(T_{m-i}^k) \geq (m-i)(k-m+i+1)$. Furthermore,

$$\mathbb{P}(|\boldsymbol{\varrho}_k|_1 = i) = \frac{\binom{n}{k}\binom{n-k}{i}}{\binom{n}{k}2^{n-k}} = \frac{\binom{n-k}{i}}{2^{n-k}}.$$

Altogether, we get

$$c\left(\frac{k}{n}\right)^2 L(Maj) \geq \mathbb{E}\left[K\left(Maj|_{\boldsymbol{\varrho}_k}\right)\right] \geq \sum_{i=\max(0,m-k)}^{\min(n-k,m)} \frac{\binom{n-k}{i}}{2^{n-k}}(m-i)(k-m+i+1).$$

Denote the right hand term by b_k,

$$L(Maj) \geq c\frac{n^2}{k^2}b_k = c\frac{n^2}{k^2}\sum_{i=\max(0,\frac{n}{2}-k)}^{\min(n-k,\frac{n}{2})} \frac{\binom{n-k}{i}}{2^{n-k}}\left(\frac{n}{2}-i\right)\left(k-\frac{n}{2}+i+1\right).$$

We computed the bounds $\frac{n^2}{k^2}b_k$ for several values of k. These are shown in table 4.1. We can see that this method has not yielded better bounds on $L(Maj)$ than we already have.

k	$\frac{n^2}{k^2}b_k$
n	$\frac{n^2}{4}+\frac{n}{2}$
$n-1$	$\frac{n^2}{4}-\frac{1}{4}$
$\frac{n}{2}$	$\frac{n^2}{16}-o(n^2)$
2	$\binom{n-1}{\frac{n}{2}-1}2^{2-n}$
1	$\binom{n-1}{\frac{n}{2}-1}2^{1-n}$

Table 4.1: Bounds on Maj found combining the shrinkage lemma and Krapchenko bounds.

4.2 Super-quadratic functions

In this section we give a brief overview of functions for which super-quadratic bounds have been proved.

The Andreev bound The first super-quadratic bound by Andreev [1987] was proved using random restrictions. The proof uses the idea that the formula size of the Andreev function is greater or equal to the formula size of the composition of parity with any function on $\log n$ bits. That is, for all $f \in B_{\log n}$

$$L(A_{2n}) \geq L(f \circ \bigoplus).$$

We know that there exists $f \in B_{\log n}$ with exponential formula size (in $\log n$). Assuming the KRW conjecture, we would get

$$L(f \circ \bigoplus) = 2^{\log n}\left(\frac{n}{\log n}\right)^2 = \frac{n^3}{(\log n)^2}.$$

Instead, Andreev uses the fact that the parity function is non-constant under all restrictions (of less than n variables). Therefore, $f \circ (\bigoplus|_\rho)$ provides f with as many different inputs as $f \circ \bigoplus$. This idea can be formalized by a lower bound on the expected value in the shrinkage lemma.

Variations of the Andreev function Several variations of the Andreev functions have also yielded super-quadratic bounds. A bound of $\frac{n^3}{\log n(\log\log n)^2}$ was proved by Tal [2017] for a function similar to the Andreev function that instead of specifying the whole truth table as a part of the input uses an error-correcting code to determine the function. This bound is slightly greater than Tal's $\frac{n^3}{(\log n)^2 \log\log n}$ for the Andreev function (Tal [2014]). Bogdanov [2018] showed the bound $\frac{n^3}{\log n(\log\log n)^2}$ for any element of a special class of random Boolean functions.

Gál et al. [2018] proved an almost-cubic lower bound for a variation of the Andreev function that uses majority instead of the parity function. They even showed, that any function that coincides with majority on the middle two slices (which includes parity) can be used to obtain a similar bound. Instead of simple random restrictions, they used "staged" random restrictions that iteratively restrict variables in a way that keeps the restricted majority function non-constant.

Towards the KRW conjecture While it is deemed unlikely that shrinkage could provide larger than cubic bounds (Dinur and Meir [2016]), proving the KRW conjecture would provide explicit super-polynomial bounds. The conjecture essentially states that the naive way (Proposition 4) of constructing a formula for a composite function is the optimal way.

Karchmer et al. [1995] proposed to prove the conjecture for compositions with the universal relation U, a weaker version than for functions. For $U \circ U$ it was proved by Edmonds et al. [2001] and later Gavinsky et al. [2014] proved the conjecture for $f \circ U$, f arbitrary. The most recent major result was achieved by Dinur and Meir [2016], who proved the conjecture for $f \circ \bigoplus$, f arbitrary. The results of Gál et al. [2018] can be also viewed as a step towards proving the conjecture for compositions with majority.

Conclusion

We have given an overview of two of the main approaches in construction of lower bounds on (De Morgan) Boolean formula size. The first approach, using formal complexity measures, has not so far yielded any super-quadratic bounds. Furthermore, there has not been as much research of this approach in the last few years, especially since it has been proved that only non-convex measures can give noteworthy bounds.

The second approach uses mainly random restrictions and composition, and has been able to produce almost cubic bounds. On the other hand, the research conducted in this approach is somewhat less systematic than for complexity measures. It is generally believed that understanding composition of Boolean function would lead to super-polynomial bounds, mainly through the KRW conjecture.

An open problem of this thesis is whether K is the largest graph measure (or at least the largest Hamming measure). However, since all of the measures involved are convex, the consequences of the problem are not particularly useful.

Bibliography

A. V. Aho, J. D. Ullman, and M. Yannakakis. On notions of information transfer in VLSI circuits. In *Proceedings of the fifteenth annual ACM symposium on Theory of computing*, pages 133–139. ACM, 1983.

A. E. Andreev. A method for obtaining efficient lower bounds for monotone complexity. *Algebra and Logic*, 26(1):1–18, 1987.

A. Bogdanov. Small bias requires large formulas. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

I. Dinur and O. Meir. Toward the KRW composition conjecture: Cubic formula lower bounds via communication complexity. *Computational Complexity*, pages 1–88, 2016.

J. Edmonds, R. Impagliazzo, S. Rudich, and J. Sgall. Communication complexity towards lower bounds on circuit depth. *Computational Complexity*, 10(3):210–246, 2001.

A. Gál, A. Tal, and A. Trejo Nuñez. Cubic formula size lower bounds based on compositions with majority. In *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

D. Gavinsky, O. Meir, O. Weinstein, and A. Wigderson. Toward better formula lower bounds: an information complexity approach to the KRW composition conjecture. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 213–222. ACM, 2014.

J. Håstad. The shrinkage exponent is 2. In *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*, pages 114–123. IEEE, 1993.

P. Hrubeš, S. Jukna, A. Kulikov, and P. Pudlák. On convex complexity measures. *Theoretical Computer Science*, 411(16-18):1842–1854, 2010.

M. Karchmer and A. Wigderson. Monotone circuits for connectivity require super-logarithmic depth. *SIAM Journal on Discrete Mathematics*, 3(2):255–265, 1990.

M. Karchmer, E. Kushilevitz, and N. Nisan. Fractional covers and communication complexity. In *[1992] Proceedings of the Seventh Annual Structure in Complexity Theory Conference*, pages 262–274. IEEE, 1992.

M. Karchmer, R. Raz, and A. Wigderson. Super-logarithmic depth lower bounds via the direct sum in communication complexity. *Computational Complexity*, 5(3-4):191–204, 1995.

V. M. Khrapchenko. Method of determining lower bounds for the complexity of P-schemes. *Mathematical Notes*, 10(1):474–479, 1971.

E. I. Nechiporuk. A Boolean function. *Engl. transl. in Sov. Phys. Dokl.*, 10: 591–593, 1966.

K. L. Rychkov. A modification of Khrapchenko's method and its application to bounding the complexity of P-networks for coding functions. *Methods of Discrete Analysis in the Theory of Graphs and Circuits, Institut Matematiki SOAN SSSR, Novosibirsk*, pages 91–98, 1985.

I. S. Sergeev. Complexity and depth of formulas for symmetric Boolean functions. *Moscow University Mathematics Bulletin*, 71(3):127–130, 2016.

P. Spira. On time-hardware complexity tradeoffs for boolean functions. In *Proceedings of the 4th Hawaii Symposium on System Sciences, 1971*, pages 525–527, 1971.

B. A. Subbotovskaya. Realization of linear functions by formulas using or, and, not. In *Doklady Akademii Nauk*, volume 136, pages 553–555. Russian Academy of Sciences, 1961.

A. Tal. Shrinkage of De Morgan formulae by spectral techniques. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 551–560. IEEE, 2014.

A. Tal. Formula lower bounds via the quantum method. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1256–1268. ACM, 2017.

K. Ueno. Breaking the rectangle bound barrier against formula size lower bounds. In *International Symposium on Mathematical Foundations of Computer Science*, pages 665–676. Springer, 2010.

K. Ueno. Candidate boolean functions towards super-quadratic formula size. *IEICE TRANSACTIONS on Information and Systems*, 98(3):524–531, 2015.

I. Wegener. *The complexity of Boolean functions.* BG Teubner, 1987.