On Convex Complexity Measures

P. Hrubeš[a,1], S. Jukna[b,c,2], A. Kulikov[d,3], P. Pudlák[e,*,4]

[a]*School of Mathematics, Institute for Advanced Study, Princeton NJ*
[b]*Institute of Mathematics and Computer Science, Vilnius, Lithuania*
[c]*Institut für Informatik, Universität Frankfurt, Germany*
[d]*Steklov Institute of Mathematics at St. Petersburg*
[e]*Mathematical Institute, Prague*

Abstract

Khrapchenko's classical lower bound n^2 on the formula size of the parity function f can be interpreted as designing a suitable measure of sub-rectangles of the combinatorial rectangle $f^{-1}(0) \times f^{-1}(1)$. Trying to generalize this approach we arrived at the concept of *convex measures*. We prove the negative result that convex measures are bounded by $O(n^2)$ and show that several measures considered for proving lower bounds on the formula size are convex. We also prove quadratic upper bounds on a class of measures that are not necessarily convex.

Key words: Boolean formula; Complexity measure; Combinatorial rectangle; Convexity; Rank; Matrix norm

1. Introduction

Most proofs of lower bounds on the formula size can be viewed as inventing suitable *formal complexity measures* of boolean functions which can be non-trivially bounded from below at some explicitly given boolean function $f : \{0,1\}^n \to \{0,1\}$. Such measures are real valued functions defined on all boolean functions and satisfying certain conditions. Formal complexity measures were introduced by Paterson. He showed that Khrapchenko's n^2 lower bound on the formula size of the parity function [9] can be recast in this formalism (see e.g. [19], Sect. 8.8). Generalizing Khrapchenko's argument for the parity function, Rychkov [16] proved $\Omega(n^2)$ lower bounds for error correcting codes. All these results are for the de Morgan basis $\neg, \vee, \wedge$. In principle this approach should give lower bounds for every basis, but no results for other bases have been obtained in this manner. In this paper we will only consider the de Morgan basis.

Apparently, Rychkov [16] was the first who explicitly related formulas with combinatorial rectangles. In order to obtain larger lower bounds, Razborov [14] proposed to look at rectangles as

matrices over some field and introduce appropriate measures on sub-rectangles in terms of the corresponding sub-matrices. Razborov studied the measures based on the rank of matrices. He showed in [15] that the rank can only give linear lower bounds for the de Morgan basis, but it gives super-polynomial lower bounds for the monotone basis $\vee, \wedge$.

More recently, a number of various matrix norms have been proposed for proving lower bounds on communication complexity and formula size [10, 11, 13]. Unfortunately up to now none of the proposed measures was able to prove more than quadratic lower bounds. Therefore it is necessary to explain this failure before we attempt to break the n^2 barrier for lower bounds based on formal complexity measures.

In previous papers some limitations of the method used therein were proved. Here we will introduce another general concept, *convex measures*. The reason for introducing this concept is to capture a large class of measures that are defined using matrices on the rectangle $f^{-1}(0) \times f^{-1}(1)$. We will prove that such measures are always at most $O(n^2)$ and show that some measures considered before are of this type, hence the upper bound also applies to them. Our upper bound on convex measures is based on the upper bound on the fractional cover number of Karchmer, Kushilevitz, and Nisan [7]. Using a different technique we will also prove quadratic, and even linear upper bounds on some other measures related to convex measures.

The main message of this paper is that we must use non-convex measures in order to beat the n^2 bound. Non-convexity, however, is only necessary but not sufficient property: we show that the measure based on matrix rank introduced by Razborov in [14] is not convex but still, as shown by Razborov in [15], cannot even yield super-linear lower bounds. The only super-quadratic lower bounds, the lower bound $\Omega(n^{2.5-o(1)})$ of Andreev [2] and the lower bound $\Omega(n^{3-o(1)})$ of Håstad [4], have not been translated into the formalism of measures yet; these bounds were obtained using Subbotovskaya's idea of random restrictions [18].

2. Basic concepts

Let n be a fixed positive integer, let $\mathcal{F}_n$ denote the set of all boolean functions $f : \{0,1\}^n \to \{0,1\}$. Literals are boolean variables and their negations. We will consider boolean formulas with tight negations over the de Morgan basis $\{\wedge, \vee, \neg\}$. That is, inputs are literals, and the gates are $\wedge$ and $\vee$. The *size* of such a formula is the number of input literals. The *formula size complexity*, $L(f)$, of a boolean function f is the smallest size of a formula computing f. A function $m : \mathcal{F}_n \to \mathbb{R}$ is called *a formal complexity measure* of boolean functions if it satisfies the following inequalities:

(a) *Normalization*: the measure of each literal is at most 1;
(b) *Subadditivity*: $m(g \vee h) \leq m(g) + m(h)$ and $m(g \wedge h) \leq m(g) + m(h)$, for every $g, h \in \mathcal{F}_n$.

It follows, by induction, that for every formal complexity measure m, we have that $L(f) \geq m(f)$ for all boolean functions f. On the other hand, L is a formal complexity measure, hence we are not losing anything by using formal complexity measures. The hope is that while it is hard to compute $L(f)$, we may be able to handle other complexity measures. With this goal in mind, the following larger class of measures—so-called rectangle measures—were considered by many authors.

Let $U_n = \{0,1\}^n \times \{0,1\}^n$. In this paper we shall define an n-dimensional *combinatorial rectangle*, or just a *rectangle*, to be a non-empty Cartesian product $S = S^0 \times S^1$ such that $S \subseteq U_n$ and $S^0 \cap S^1 = \emptyset$. (Note that U_n itself is *not* a rectangle.) The sets S^0 and S^1 are called *sides* of the rectangle $S = S^0 \times S^1$. A *sub-rectangle* of S is a subset $R \subseteq S$ which itself forms a rectangle. Vector

pairs $e = (x, y)$ with $x \neq y$ will be referred to as *edges*. A boolean function $f : \{0,1\}^n \to \{0,1\}$ *separates* the rectangle $S = S^0 \times S^1$ if

$$f(x) = \begin{cases} 0 & \text{for } x \in S^0, \\ 1 & \text{for } x \in S^1. \end{cases}$$

If the sets S^0 and S^1 form a partition of $\{0,1\}^n$, then the rectangle $S = S^0 \times S^1$ is called a *full rectangle*. Note that there is a one-to-one correspondence between boolean functions $f : \{0,1\}^n \to \{0,1\}$ and full rectangles that have the form

$$S_f := f^{-1}(0) \times f^{-1}(1) \, ;$$

but there are many more rectangles than boolean functions.

An important class of rectangles are monochromatic rectangles which are the rectangles that can be separated by single literals. That is, a rectangle $M = M^0 \times M^1$ is *monochromatic*, if there exists an $i \in \{1, \ldots, n\}$ and an $\varepsilon \in \{0,1\}$ such that for all $(x,y) \in M$, $x_i = \varepsilon$ and $y_i = 1 - \varepsilon$; here x_i is the i-th bit in x. The smallest monochromatic rectangles are single edges, i.e., rectangles of the form $M = \{(x,y)\}$ with $x \neq y$. The largest ones are the so-called *canonical* monochromatic rectangles

$$M_{i,\varepsilon} = \{(x,y) \in U_n \colon x_i = \varepsilon \text{ and } y_i = 1 - \varepsilon \} \, .$$

These $2n$ rectangles cover every rectangle.

Instead of rectangles within the whole set U_n, one can work only with rectangles $R \subseteq S$ within some fixed rectangle S, say, within the full rectangle $S_f = f^{-1}(0) \times f^{-1}(1)$ of a given boolean function f.

For the rest of this paper we shall assume that the dimension n and a rectangle S are fixed. We shall call S the ambient rectangle. In what follows, $\mathcal{R} = \mathcal{R}(S)$ will denote the set of all sub-rectangles and $\mathcal{M} = \mathcal{M}(S)$ the set of all monochromatic sub-rectangles of S.

2.1. Rectangle measures and communication complexity

A *rectangle function* is a real-valued function $\mu : \mathcal{R} \to \mathbb{R}$. Such a function is:

(i) *Normalized,* if $\mu(M) \leq 1$ for every monochromatic rectangle $M \in \mathcal{M}$.
(ii) *Subadditive,* if $\mu(R) \leq \mu(R_1) + \mu(R_2)$, for every rectangle $R \in \mathcal{R}$ and for every its partition into disjoint union of rectangles $R_1, R_2 \in \mathcal{R}$.

Rectangle functions satisfying both these conditions (normalization and subadditivity) are called *rectangle measures*. Note that we do not require such a measure be *monotone*, in the sense that $R \subseteq S$ must imply $\mu(R) \leq \mu(S)$. The first condition is usually achieved by normalization. That is, if a rectangle function ν is subadditive and $\nu(M) > 0$ for some monochromatic rectangle $M \in \mathcal{M}$, we obtain a measure by defining

$$\mu(R) = \frac{\nu(R)}{\max_M \nu(M)},$$

where M ranges over all monochromatic rectangles.

These two conditions already suffice for lower-bounding the formula size. Notice that rectangles can be decomposed into disjoint unions of two rectangles in two ways—vertically and horizontally. Subadditivity of rectangle measures corresponds to the two conditions of subadditivity (b) in the definition of formal complexity measures of boolean functions.

3

The connection between rectangle measures and formula size can be best seen in the framework of communication games, as introduced by Karchmer and Wigderson [8]: having a rectangle R, one of the players decomposes R either row-wise or column-wise, and the players continue the game on one of the sub-rectangles R_1 or R_2. Let $\Gamma(R)$ denote the minimal number of leaves in a communication protocol tree for a rectangle R in a Karchmer-Wigderson game. Then $L(f) = \Gamma(S_f)$ [8]. The measure $\Gamma(R)$ itself is a rectangle measure. Moreover, by induction on $\Gamma(R)$, it can be easily shown that $\Gamma(R) \geq \mu(R)$ holds for any rectangle measure μ. Hence, subadditive rectangle measures can reach $L(f)$ as well. The advantage, however, is that now we have a larger class of measures, and the subadditivity condition for rectangle measures is weaker requirement than that for boolean functions.

We keep this important observation as

Proposition 2.1. *For every boolean function f and every subadditive rectangle measure μ we have that $L(f) = \Gamma(S_f) \geq \mu(S_f)$.*

The two concepts—rectangle measures and formal complexity measures—are related as follows.

Observation 2.2. *If $m(f)$ is a formal complexity measure of boolean functions, then the rectangle function $\mu(R)$, defined by $\mu(R) := \min\{m(f) : f$ separates $R\}$, is a rectangle measure.*

2.2. Strongly subadditive measures and the partition number

A stronger condition than (ii) has also been considered:

(iii) *Strong subadditivity*: $\mu(R) \leq \sum_{i=1}^{m} \mu(R_i)$, for every rectangle R and every its partition into disjoint union of rectangles $R_i \subseteq R$.

In order to obtain a lower bound on $L(f)$ it suffices to require this property only for $R = S_f$. Note, however, that rectangle measures, satisfying the strong subadditivity condition (iii) may not achieve $L(f)$, because they lower bound a different quantity, namely, the *partition number* of rectangles defined by:

$$\mathsf{D}(R) = \min\{k : R \text{ can be decomposed into } k \text{ disjoint monochromatic rectangles}\}.$$

As observed by Rychkov [16], this measure was implicitly used already in Khrapchenko's proof [9]. Since $\mathsf{D}(R)$ is strongly subadditive, it is also subadditive. Hence, $L(f) \geq \mathsf{D}(S_f)$ for any boolean function f. But in the opposite direction we only know that $\log_2 L(f) \leq (\log_2 \mathsf{D}(S_f))^2$ [1]. Still, the latter inequality implies that boolean functions f in n variables such that $\mathsf{D}(S_f) \geq 2^{(1-o(1))\sqrt{n}}$ exist. Hence, in principle, the partition number $\mathsf{D}(S)$ can also achieve super-polynomial lower bounds on the formula size. The problem how large the gap $L(f)/\mathsf{D}(S_f)$ can actually be remains still open.

The measure $\mathsf{D}(R)$ has several nice properties.

Proposition 2.3. $\mathsf{D}(R)$ *is the largest strongly subadditive measure, i.e., $\mathsf{D}(R)$ is strongly subadditive and for every strongly subadditive measure μ, $\mu(R) \leq \mathsf{D}(R)$ for all rectangles R.*

We leave the proof to the reader as an easy exercise. Although $\mathsf{D}(R)$ is the largest strongly subadditive measure, it makes sense to study other strongly subadditive measures, because it is very difficult to compute $\mathsf{D}(R)$ for specific functions.

4

Other nice properties of $\mathsf{D}(R)$ include the following: it is defined independently of a particular boolean function, can be naturally extended from rectangles to all subsets $X \subseteq S$ and is monotonic with respect to set-inclusion. A consequence for lower bounds based on measures is that one can use measures with all these nice properties and still obtain exponential lower bounds.

However, we cannot stretch the good properties too far. In particular, it is essential that in the subadditivity conditions the rectangles in the partitions must be pairwise disjoint. Would we not require them to be disjoint, then $\mu(S) \leq 2n$ would hold for any n-dimensional rectangle S, just because each such rectangle can be covered by $2n$ canonical monochromatic rectangles. In the next section we will show another property, the convexity, that limits the values of measures satisfying it.

3. Convex measures and the fractional partition number

For a rectangle R, let χ_R be its indicator function, that is, $\chi_R(e) = 1$ for $e \in R$, and $\chi_R(e) = 0$ for $e \notin R$. Let R be a rectangle, $R_1, \ldots, R_m$ its sub-rectangles and $r_1, \ldots, r_m$ weights from $[0, 1]$ such that

$$\chi_R = \sum_{i=1}^{m} r_i \cdot \chi_{R_i}, \tag{1}$$

Then we say that the rectangles $R_1, \ldots, R_m$ with the weights $r_1, \ldots, r_m$ are a *fractional partition* of the rectangle R. This is equivalent to the condition that for every edge $e \in R$,

$$\sum_{i : e \in R_i} r_i = 1.$$

Notice that if all $r_i \in \{0, 1\}$ then a fractional partition is a partition. Instead of (1) we shall use the following simpler notation

$$R = \sum_i r_i R_i.$$

In this paper we are mainly interested in the following strengthening of the strong subadditivity condition (iii) for rectangle measures μ:

(iv) *Convexity*: A rectangle function μ is *convex* if, for every rectangle R and every fractional partition $R = \sum_i r_i R_i$,

$$\mu(R) \leq \sum_{i=1}^{m} r_i \cdot \mu(R_i). \tag{2}$$

Note that in results about convex measures it often suffices to assume condition (2) only for fractional partitions consisting of *monochromatic* rectangles. In particular this concerns Proposition 3.1 and Theorem 3.5.

Karchmer, Kushilevitz and Nisan in [7] introduced a modification of the partition number which they called deterministic fractional cover number. In this paper we will call it *fractional partition number* and denote it by $\mathsf{D}^*(R)$. To call it 'cover number' would be misleading, because it is important that one uses partitions, not general coverings. This measure is defined by:

$$\mathsf{D}^*(R) = \min \sum_i r_i,$$

such that R has a fractional partition with monochromatic rectangles $M_1, \ldots, M_m$ and weights $r_1, \ldots, r_m$.

The following is a fractional version of Proposition 2.3.

Proposition 3.1. D^* *is the largest convex measure, i.e.,* D^* *is convex and for every convex measure* μ, $\mu(R) \leq \mathsf{D}^*(R)$ *for all rectangles* R.

PROOF. First we will show that D^* is convex. Let $R = \sum_{j \in J} r_j R_j$ be a fractional partition of R and, for every j, let $R_j = \sum_{i \in I_j} s_{ij} M_{ij}$ be a fractional partition of R_j such that M_{ij} are monochromatic and $\mathsf{D}^*(R_j) = \sum_i s_{ij}$ (such fractional partitions exist by definition). Then, clearly, $R = \sum_{ij} r_j s_{ij} M_{ij}$ is a fractional partition of R into monochromatic rectangles. Hence

$$\mathsf{D}^*(R) \leq \sum_{ij} r_j s_{ij} = \sum_j r_j \mathsf{D}^*(R_j) \, .$$

Now we will show the second part. Let μ be a convex measure. Let $R = \sum_i r_i M_i$ be a fractional partition of R into monochromatic rectangles such that $\mathsf{D}^*(R) = \sum_i r_i$. Using convexity and normality of μ we get

$$\mu(R) \leq \sum_i r_i \mu(M_i) \leq \sum_i r_i = \mathsf{D}^*(R) \, .$$

$\square$

Theorem 3.2 ([7]). *For every n-dimensional rectangle S, $\mathsf{D}^*(S) \leq 4n^2$.*

Consequently every convex measure is bounded by $4n^2$. For the sake of completeness we will reproduce their proof. By more careful computation we will get the constant $\frac{9}{8}$ instead of 4. We will state and prove the bound for all convex measures.

Following Karchmer [6], and Karchmer, Kushilevitz and Nisan [7], associate with each subset $I \subseteq [n] = \{1, \ldots, n\}$ the following two *parity rectangles*.

$$P_{I,\varepsilon} = \left\{ x \in \{0,1\}^n \colon \oplus_{i \in I} x_i = \varepsilon \right\} \times \left\{ y \in \{0,1\}^n \colon \oplus_{i \in I} y_i = 1 - \varepsilon \right\}, \ \varepsilon = 0, 1 \, .$$

Hence, monochromatic rectangles correspond to the case when $|I| = 1$. There are exactly 2^{n+1} parity rectangles (including the empty one).

Lemma 3.3. *Every edge $(x,y) \in \{0,1\}^n \times \{0,1\}^n$ such that $x \neq y$ belongs to exactly 2^{n-1} parity rectangles.*

PROOF. For $I \subseteq [n]$, let $v_I \in \{0,1\}^n$ be its incidence vector. Let $e = (x,y) \in S$. Since $x \neq y$, the vector $x \oplus y$ is not a zero vector. Since each nonzero vector is orthogonal over $GF(2)$ to exactly half of the vectors in $\{0,1\}^n$, this implies that precisely 2^{n-1} of the vectors v_I are non-orthogonal to $x \oplus y$. This means that each edge e belongs to precisely 2^{n-1} of the sets $P_I = P_{I,0} \cup P_{I,1}$. Since $P_{I,0} \cap P_{I,1} = \emptyset$, we are done. $\square$

Lemma 3.4. *Let μ be a rectangle measure defined on S. Then for every $I \subseteq [n]$, $\varepsilon = 0, 1$, we have $\mu(P_{I,\varepsilon} \cap S) \leq \frac{9}{8} |I|^2$.*

PROOF. A parity rectangle $P_{I,\varepsilon}$ can be viewed as a rectangle corresponding to the parity function in $|I|$ variables, or its negation. As shown in [17, 12], parity of $n = 2^l + k$ variables can be computed by a formula of size $c(n) = 2^l(2^l + 3k)$. This gives $c(n) \leq \frac{9}{8}n^2$. To see that, observe that the function $y(y + 3x)/(y + x)^2$ for $x \in (0, y)$ reaches maximum at the point $x = y/3$, and it has then value $9/8$. Hence $\mu(S \cap P_{I,\varepsilon}) \leq \frac{9}{8}|I|^2$, since μ is a lower bound to the formula size. □

Theorem 3.5. *If μ is a convex rectangle measure then, for every n-dimensional rectangle S,*

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

PROOF. Let S be a rectangle and μ a convex measure. For $i = 1, \ldots n$, $\varepsilon = 0, 1$, let

$$\mathcal{R}_{i,\varepsilon} := \{P_{I,\varepsilon} \cap S \colon I \subseteq [n], |I| = i\}$$

and let $\mathcal{R}_{\mathrm{par}}$ be the union of all these $2n$ families of parity sub-rectangles of S. For counting reasons, we shall understand $\mathcal{R}_{\mathrm{par}}$ as a multi-set, elements of $\mathcal{R}_{\mathrm{par}}$ corresponding to different parity rectangles are considered different. Under this provision, Lemma 3.3 implies that every edge in S is contained in exactly 2^{n-1} elements of $\mathcal{R}_{\mathrm{par}}$. Hence $\mathcal{R}_{\mathrm{par}}$ form a fractional partition of S with each rectangle $R \in \mathcal{R}_{\mathrm{par}}$ of weight $r_R = 2^{-(n-1)}$. By the previous lemma, we know that $\mu(R) \leq ci^2$ for every $R \in \mathcal{R}_{i,\varepsilon}$, where $c = 9/8$. The convexity of μ implies that

$$\mu(S) \leq \sum_{R \in \mathcal{R}_{\mathrm{par}}} r_R \cdot \mu(R) = 2^{-(n-1)} \sum_{R \in \mathcal{R}_{\mathrm{par}}} \mu(R) = 2^{-(n-1)} \sum_{i,\varepsilon} \sum_{R \in \mathcal{R}_{i,\varepsilon}} \mu(R)$$

$$\leq 2^{-(n-1)} \sum_{i=1}^{n} \sum_{\varepsilon=0,1} \binom{n}{i} ci^2 = 2^{-(n-1)} 2c \sum_{i=1}^{n} \binom{n}{i} i^2 = 2^{-(n-2)} c \sum_{i=1}^{n} \binom{n}{i} i^2.$$

The identity $\binom{n}{k} \cdot k = n \cdot \binom{n-1}{k-1}$ gives

$$\sum_{i=1}^{n} \binom{n}{i} i^2 = n \cdot \sum_{i=1}^{n} \binom{n-1}{i-1} i = n \cdot \sum_{i=1}^{n} \binom{n-1}{i-1} + n \cdot \sum_{i=1}^{n} \binom{n-1}{i-1}(i-1)$$

$$= n \cdot \sum_{i=1}^{n} \binom{n-1}{i-1} + n \cdot \sum_{i=2}^{n} \binom{n-1}{i-1}(i-1)$$

$$= n \cdot \sum_{i=1}^{n} \binom{n-1}{i-1} + n(n-1) \cdot \sum_{i=2}^{n} \binom{n-2}{i-2}$$

$$= n2^{n-1} + n(n-1)2^{n-2} = (n^2 + n)2^{n-2}.$$

Hence, $\mu(S) \leq 2^{-(n-2)} c(n^2 + n) 2^{n-2} = c(n^2 + n)$. □

4. General construction of convex measures

In his seminal paper [9], Khrapchenko proved a general lower bound on formula size complexity of the form

$$L(f) \geq \frac{|\{(x, y) \in R \colon \mathrm{dist}(x, y) = 1\}|^2}{|R|},$$

7

where R is a sub-rectangle of S_f. Paterson (see, e.g., [19]) interpreted this formula as a formal complexity measure and reproved Khrapchenko's n^2 lower bound on the parity function in this formalism. We will call the measure

$$\kappa(R) = \frac{|R \cap Y|^2}{|R|}, \tag{3}$$

where $Y = \{(x, y) \colon \operatorname{dist}(x, y) = 1\}$ is the set of all vector pairs of Hamming distance 1, the *Khrapchenko measure.* One can also understand Rychkov's lower bounds on error correcting codes as lower bounds based on the Khrapchenko measure. There one uses pairs of distance at most $d + 1$ instead of 1 for codes of the minimal distance $2d + 1$.

We can interpret Khrapchenko's lower bound as follows. One starts with rectangle functions $s(R) = |R|$, $w(R) = |Y \cap R|$, which themselves do not give better than linear lower bounds. We define a new rectangle function $\mu(R) = F(w(R), s(R))$ by means of a real function $F(x, y) = x^2/y$, and it is this measure that allows us to prove quadratic lower bounds. In this scenario, subadditivity is guaranteed by properties of F. This suggests the possibility of obtaining a new rectangle measure from a set of rectangle measures by means of a function $F : \mathbb{R}^m \to \mathbb{R}$ in the hope that the new measure will be more apt to prove lower bounds. In this section, we observe that if F has nice properties then F will produce a subadditive function, but if F has too nice properties, it will produce a convex function.

Notice that the Khrapchenko measure has the form

$$\kappa(R) = s(R) \cdot \varphi\left(\frac{w(R)}{s(R)}\right)$$

with $w(R) = |R \cap Y|$, $s(R) = |R|$ and $\varphi(x) = x^2$. Subadditivity of κ stems from the fact that the used real function φ is convex. As will be stated in Corollary 4.3, convexity of φ implies that μ is a convex rectangle measure (if $w(R)$, $s(R)$ satisfy certain conditions), and hence μ cannot give better than quadratic lower bounds.

We will need another condition (stronger than convexity):

(vi) *Additivity:* $\alpha(R) = \alpha(R_1) + \alpha(R_2)$, for all rectangles $R, R_1, R_2 \in \mathcal{R}$ such that R is the disjoint union of R_1 and R_2.

Observe that, for every additive rectangle function α,

$$\alpha(R) = \sum_{e \in R} \alpha(e). \tag{4}$$

Thus an additive rectangle function is defined by a matrix on the ambient rectangle S. Examples of such rectangle functions are $|R|$ and $|R \cap Y|$ that appear in the definition of the Khrapchenko measure. The convexity of additive measures is a consequence of the following stronger property:

$$\alpha(R) = \sum_{i=1}^{m} r_i \cdot \alpha(R_i), \tag{5}$$

for every fractional partition $R = \sum_{i=1}^{m} r_i R_i$, which is an immediate consequence of (4).

The fractional partition number D^* was introduced in [7] in order to apply the linear programming duality for obtaining lower bounds on communication complexity of relations, in particular

for proving lower bounds on formula size complexity. Applying the duality for linear programs, one can write this measure as

$$\mathsf{D}^*(S) = \max_w \sum_{e \in S} w(e),$$

where the maximum is over all functions $w : S \to \mathbb{R}$ satisfying the constraint $\sum_{e \in M} w(e) \leq 1$ for all monochromatic rectangles M. Hence, in order to prove a lower bound $\mathsf{D}^*(S) \geq t$ it is enough to find at least one weight function $w : S \to \mathbb{R}$ such that $\sum_{e \in S} w(e) \geq t$, and the weight of each monochromatic rectangle does not exceed 1. In our terminology this means to find an additive measure w such that $w(S) \geq t$.

In other words, whenever a lower bound can be proved using a convex measure, it can be proved using an additive measure. However, in practice it may be easier to work with convex measures rather than additive ones. Karchmer, Kushilevitz, and Nisan found a surprisingly new proof of Khrapchenko's n^2 lower bound based on an additive measure. Their measure uses positive and negative values. As we will see, it is necessary to use negative values in order to obtain superlinear lower bounds. (This implies that D^* is not additive.)

We start with a simple observation.

Proposition 4.1. *Any non-negative linear combination of convex rectangle functions is a convex rectangle function.*

PROOF. Let $\mu_1, \ldots, \mu_n$ be convex rectangle functions. Let $\mu(R) = \sum_{i=1}^n a_i \cdot \mu_i(R)$ be their linear combination with all $a_i \geq 0$. Let $R = \sum_{j=1}^m r_j R_j$ be a fractional partition of R. The convexity of μ_i's implies that $\mu_i(R) \leq \sum_{j=1}^m r_j \cdot \mu_i(R_j)$, for all $i = 1, \ldots, n$. Then

$$\mu(R) = \sum_{i=1}^n a_i \cdot \mu_i(R) \leq \sum_{i=1}^n a_i \sum_{j=1}^m r_j \cdot \mu_i(R_j) = \sum_{j=1}^m r_j \sum_{i=1}^n a_i \cdot \mu_i(R_j) = \sum_{j=1}^m r_j \cdot \mu(R_j).$$

$\square$

Let $F : \mathbb{R}^m \to \mathbb{R}$ be a real function in m variables. We shall think of m-tuples of real numbers as vectors in $\mathbb{R}^m$. The results below can be extended to functions whose domain is a subset of R^m closed w.r.t. addition of vectors, and multiplication by positive real numbers. We say that F is *subadditive*, if any two non-negative numbers a and b, and any two vectors $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^m$,

$$F(a\vec{x} + b\vec{y}) \leq aF(\vec{x}) + bF(\vec{y}). \tag{6}$$

If this only holds for $a = b = 1$, then F is called *weakly subadditive*. What makes weakly subadditive function subadditive is the condition $F(a\vec{x}) \leq aF(\vec{x})$ for every $a > 0$.

Let now $s(R)$ and $w(R)$ be two rectangle functions. Having such rectangle functions and a real-valued function $F(x, y)$, we can consider induced rectangle functions. This can be easily extended to m-tuples of rectangle functions and for functions F on more than two variables.

Proposition 4.2. *Let $F(x, y)$ be a subadditive function, and $s(R)$ an additive rectangle function. Then the induced rectangle function $\mu_F(R) = F(w(R), s(R))$ is convex if*

1. *either $w(R)$ is additive,*
2. *or $w(R)$ is convex and $F(x, y)$ is nondecreasing in x.*

PROOF. To prove the first claim, assume that both $w(R)$ and $s(R)$ are additive, and let $\sum_i r_i R_i$ be a fractional partition of R. Set $w_i = w(R_i)$ and $s_i = s(R_i)$. By (5), we have that $w(R) = \sum_i r_i \cdot w_i$ and $s(R) = \sum_i r_i \cdot s_i$. Since F is a subadditive function, this yields

$$\mu_F(R) = F(w(R), s(R)) = F\left(\sum_i r_i w_i, \sum_i r_i s_i\right) \le r_i \sum_i F(w_i, s_i) = \sum_i \mu_F(R_i).$$

If $w(R)$ is only convex (not necessarily additive) but $F(x,y)$ is nondecreasing in x, then we can replace the second equality by inequality. □

Note that subadditivity of F guarantees subadditivity of μ_F, and hence μ_F can be (after appropriate normalization) used as a rectangle measure for proving lower bounds. But if F is also a subadditive function, μ_F will be convex and the lower bounds given by μ_F cannot exceed $O(n^2)$. However, there are many weakly subadditive real functions that are not subadditive function. It is not clear whether the function F can be chosen in such a way that μ_F will give better than quadratic lower bounds.

Say that a rectangle function $s(R)$ is *positive* if $s(R) > 0$ for every nonempty rectangle R (of our ambient rectangle).

Corollary 4.3. *Let a rectangle function μ be defined as follows:*

$$\mu(R) = s(R) \cdot \varphi\left(\frac{w(R)}{s(R)}\right), \tag{7}$$

where $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex real function and $s(R)$ is additive and positive rectangle function.

1. *If $w(R)$ is additive, then μ is convex.*
2. *If φ is nondecreasing and $w(R)$ is subadditive then μ is subadditive.*
3. *If φ is nondecreasing and $w(R)$ is convex then μ is convex.*

PROOF. It is sufficient to prove that the function $F(x,y) = y \cdot \varphi(x/y)$ is a subadditive function. The condition $F(ax, ay) \le aF(x,y)$ is immediate. (This is in fact equality and F is a norm.) Subadditivity of F is an application of Jensen's inequality:

$$\varphi\left(\frac{y_1 z_1 + y_2 z_2}{y_1 + y_2}\right) \le \frac{y_1 \varphi(z_1) + y_2 \varphi(z_2)}{y_1 + y_2}. \tag{8}$$

Assume $y_1, y_2 > 0$. Setting $z_i = x_i/y_i$, we obtain that

$$\varphi\left(\frac{x_1 + x_2}{y_1 + y_2}\right) \le \frac{y_1 \varphi\left(\frac{x_1}{y_1}\right) + y_2 \varphi\left(\frac{x_2}{y_2}\right)}{y_1 + y_2}.$$

Hence

$$(y_1 + y_2) \cdot \varphi\left(\frac{x_1 + x_2}{y_1 + y_2}\right) \le y_1 \cdot \varphi\left(\frac{x_1}{y_1}\right) + y_2 \cdot \varphi\left(\frac{x_2}{y_2}\right).$$

□

5. Polynomial rectangle measures

An important special case of measures considered above are rectangle measures μ of the form (7) based on functions $\varphi(x) = x^k$, $k \geq 1$. That is, rectangle measures

$$\mu(R) = \frac{w(R)^k}{|R|^{k-1}}, \tag{9}$$

where $w(R)$ is subadditive. Such a measure $\mu(R)$ will be called *polynomial measure of degree k*.

Note that $\varphi(x) = x^k$, $k \geq 1$, is a nondecreasing convex function for $x \geq 0$, and $\varphi(x)$ is a convex function on $\mathbb{R}$, if k is an even natural number. Hence Corollary 4.3 implies that the rectangle function $\mu(R)$ in (9) is subadditive, if

(i) $w(R)$ is subadditive and non-negative, or
(ii) $w(R)$ is additive and $k = a/b$ for integers $a \geq b > 0$ and a even.

If $\mu(R)$ is normalized, the condition (i) or (ii) guarantees that $\mu(R)$ is a rectangle measure. In the case (ii), $\mu(R)$ is also convex. In the case (i), $\mu(R)$ is convex if $w(R)$ is convex. Therefore, by Theorem 3.5, such polynomial measures can yield at most quadratic lower bounds.

On the other hand, every rectangle measure is a polynomial measure of degree one. This shows that polynomial measures can in principle give exponential lower bounds.

Quadratic lower bounds were proved by Khrapchenko [9] using polynomial measures of degree $k = 2$ with $w(R)$ additive and positive, as well as by Karchmer, Kushilevitz and Nisan [7] using polynomial measures of degree $k = 1$ with $w(R)$ additive but not non-negative.

5.1. Small degree: $1 \leq k < 2$ and additive weight

For $1 \leq k < 2$, polynomial measures with $w(R)$ *subadditive* and non-negative *can* give exponential lower bounds. To see this, consider the rectangle function $\mu(R) = w(R)^k/|R|^{k-1}$ with $w(R) = L(R)$ being the smallest size of a formula separating R. Hence, this weight function $w(R)$ is subadditive and non-negative, and $\mu(R)$ is normalized since $L(R)$ is normalized. Most boolean functions in n variables, and hence, most n-dimensional rectangles R require $L(R) \geq 2^{n(1-o(1))}$. For such rectangles R, measure $\mu(R)$ gets asymptotically close to the values

$$\frac{2^{kn}}{2^{2n(k-1)}} = 2^{n(2-k)}.$$

On the other hand, small degree measures are useless, if we require the weight function $w(R)$ be non-negative and *additive*.

Proposition 5.1. *Let $k \geq 1$ and $\mu(R) = w(R)^k/s(R)^{k-1}$ be a rectangle measure, where $s(R)$ is a positive monotone rectangle function. If the weight function $w(R)$ is additive and non-negative, then $\mu(R) \leq (2n)^k$ for any n-dimensional rectangle R.*

PROOF. The normalization condition $\mu(M) \leq 1$, for a monochromatic rectangle M implies that

$$w(M) \leq s(M)^{\frac{k-1}{k}}.$$

Since every n-dimensional rectangle can be (non-disjointly) covered by at most $2n$ canonical monochromatic rectangles $M_{i,\varepsilon}$, we have

$$w(R) \leq \sum_{i,\varepsilon} w(M_{i,\varepsilon}) \leq \sum_{i,\varepsilon} s(M_{i,\varepsilon})^{\frac{k-1}{k}} \leq 2n \cdot s(R)^{\frac{k-1}{k}} \,.$$

Dividing by $s(R)^{\frac{k-1}{k}}$ and raising to the power k we get the inequality. $\qquad\square$

Hence, if the used additive weight function $w(R)$ is *non-negative*, then no polynomial measure of degree $k < 2$ can even reach the n^2 lower bound (even if the function $s(S)$ is not necessarily additive). Note however that $w(R)$ being non-negative is here essential: for $k = 1$, additive measures can give quadratic lower bounds, if some edges are assigned *negative* weights [7].

5.2. Large degree: $k \geq 2$ and subadditive weight

We now show that every polynomial measure of degree $k > 2$, with $w(R)$ subadditive can give at most linear lower bounds.

Theorem 5.2. *Let $k \geq 2$ and let $w(R)$ be a subadditive rectangle function. Suppose that either $w(R)$ is nonnegative, or k is an integer. Suppose that the rectangle function defined by*

$$\mu(R) = \frac{w(R)^k}{|R|^{k-1}}$$

is normalized. Then, for any n-dimensional rectangle S, we have

1. *$\mu(S) \leq n^2$ if $k = 2$;*
2. *$\mu(S) = O(n)$ if $k > 2$.*

In the proof we will need the following technical lemma (whose proof is given an Appendix).

Lemma 5.3. *Let $a \geq 1$ and $\alpha \in [0, 1)$, and let $\xi(a)$ be the maximum, over all $x, y \in [0, 1]$, of*

$$h_a(x, y) = (xy)^\alpha + ((1 - x)(1 - y))^\alpha + a(x(1 - y))^\alpha + a((1 - x)y)^\alpha \,.$$

Then

(i) *$\xi(a) = \max\left\{ a(1 + a^{\frac{1}{\alpha - 1}})^{1-\alpha}, \; 2^{1-2\alpha}(1 + a) \right\}.$*

(ii) *If $\alpha = \frac{1}{2}$ then for every $d \geq 1$*

$$d + 1 \geq \xi(d) \,. \tag{10}$$

(iii) *If $\alpha > \frac{1}{2}$ then there exists a constant c such that for every $d \geq 1$,*

$$c \cdot (d + 1)^{1-\alpha} \geq \xi(c \cdot d^{1-\alpha}) \,. \tag{11}$$

We now turn to the actual proof of Theorem 5.2. Let $S = S^0 \times S^1$. Since μ is normalized, we have that

$$w(M) \leq |M|^{1-1/k} \tag{12}$$

for every monochromatic rectangle M.

Claim 5.4.

1. *If $k = 2$ then $w(S) \leq n \cdot |S|^{1/2}$.*
2. *If $k > 2$ then $w(S) \leq cn^{1/k}|S|^{1-1/k}$, for a constant c.*

Note that Theorem 5.2 is a direct consequence of this claim. In the case $k = 2$,

$$\mu(S) = \frac{w(S)^2}{|S|} \leq \frac{(n|S|^{1/2})^2}{|S|} = n^2,$$

and in the case $k > 2$,

$$\mu(S) = \frac{w(S)^k}{|S|^{k-1}} \leq \frac{(c \cdot n^{1/k}|S|^{1-1/k})^k}{|S|^{k-1}} = c^k n = O(n).$$

Hence, it remains to prove Claim 5.4. Let $\dim R = |\{i \colon \exists (x,y) \in R \colon x_i \neq y_i\}|$, and let

$$w(m, d) = \max\{w(R) \colon \dim R \leq d \text{ and } |R| = m\}.$$

Given a rectangle R with $\dim R = d+1$, we can split it into four disjoint rectangles, two monochromatic ones and two remaining ones of a smaller dimension. More exactly, if R is an $a \times b$ rectangle then, for some $x, y \in [0, 1]$, the monochromatic rectangles will be of sizes $ax \times by$ and $a(1-x) \times b(1-y)$, and the two remaining rectangles of size $ax \times (1-y)b$ and $a(1-x) \times by$. By (12), we have that

$$w(m, 1) \leq m^\alpha \qquad \text{where } \alpha := 1 - 1/k.$$

Since $w(R)$ is subadditive, we have a recurrent inequality

$$w(m, d+1) \leq \sup_{x,y \in [0,1]} \left((xym)^\alpha + ((1-x)(1-y)m)^\alpha + w(x(1-y)m, d) + w((1-x)ym, d)\right).$$

We want to upper bound $w(m, d)$. For this, it is sufficient to find a function g which satisfies $g(m, 1) \geq m^\alpha$ and

$$g(m, d+1) \geq \sup_{x,y \in [0,1]} \left((xym)^\alpha + ((1-x)(1-y)m)^\alpha + g(x(1-y)m, d) + g((1-x)ym, d)\right).$$

We look for a solution of the form

$$g(m, d) = m^\alpha \cdot h(d).$$

Hence $h(d)$ needs to satisfy the inequalities $h(1) \geq 1$ and

$$h(d+1) \geq \sup_{x,y \in [0,1]} \left((xy)^\alpha + ((1-x)(1-y))^\alpha + h(d)(x(1-y))^\alpha + h(d)((1-x)y)^\alpha\right).$$

Using the definition from Lemma 5.3, it is sufficient to have $h(d) \geq 1$ and

$$h(d+1) \geq \xi(h(d)).$$

Lemma 5.3 then asserts that for $\alpha = 1/2$ (i.e., $k = 2$) $h(d) = d$ is a solution, and for $\alpha \geq 1/2$ (i.e., $k > 2$), $h(d) = c \cdot d^{1-\alpha}$ is a solution. This completes the proof of Claim 5.4, and thus, the proof of Theorem 5.2. $\qquad \square$

13

6. More examples of measures

In this section we shall survey rectangle measures and show that several of the proposed measures are convex. Most rectangle measures are based on some matrix defined on S, i.e., a mapping $A : S \rightarrow F$, for some field. The idea of studying matrix parameters for proving lower bounds on formula size complexity is due to Razborov [14].

6.1. Khrapchenko-type measures

Khrapchenko's bound can be viewed as based on the matrix

$$A[x, y] = 1 \text{ if } d(x, y) = 1 \text{ and } 0 \text{ otherwise.} \tag{13}$$

Similarly Rychkov's lower bounds on codes of distance $2d + 1$ are based on matrices that have 1 for pairs of distance at most $d + 1$ and 0 otherwise.

There are several n^2 lower bounds for parity based on convex measures. One is the original Khrapchenko's bound, the other is the bound of Karchmer, Kushilevitz and Nisan that uses an additive measure. There is yet another convex measure that gives the same bound. Namely, let A be a real matrix defined on S. Then the rectangle function defined by

$$\phi_A(R^0 \times R^1) := \sum_{x \in R^0} \frac{(\sum_{y \in R^1} A[x, y])^2}{|R^1|}$$

is convex. Indeed, since the measure of a rectangle is the sum of the measures of its rows, it suffices to show convexity for rows. This follows from Corollary 4.3.3.

If S is the rectangle of the parity function and A as in (13), the function ϕ_A is normalized, hence measure, and $\phi_A(S) = n^2$. The measure ϕ_A for this special matrix A was introduced by Koutsoupias [10].

6.2. Matrix rank

Razborov in [14] used the rank of matrices to prove superpolynomial lower bounds on monotone formula size. Given an $n \times n$ matrix A (over some field), he associates with it the following measure for n-dimensional rectangles:

$$\mu_A(R) = \frac{\text{rank}(A_R)}{\max_M \text{rank}(A_M)}, \tag{14}$$

where A_R is the restriction of A to the rectangle R (obtained by setting to 0 all entries outside R), and the maximum is over all monochromatic sub-rectangles of R (or over all canonical monochromatic rectangles of the ambient rectangle S, as originally defined in [14]; Proposition 6.1 below holds under both definitions). If $\text{rank}(A_R) = 0$ then we set $\mu_A(R) = 0$.

Subadditivity of rank implies that these measures are subadditive. But it turns out that rank-based measures are not convex.

Proposition 6.1. *For any even integer n there is a $(0, 1)$ matrix A such that the measure μ_A is not convex.*

PROOF. Let n be even. Take a rectangle $R = R^0 \times R^1$ with $R^0 = \{x_1, \ldots, x_n\}$ and $R^1 = \{y_1, \ldots, y_n\}$ where $x_i = e_i$, $y_i = e_i + e_{i+1}$ and $e_i \in \{0,1\}^{n+1}$ is the ith unit vector. Let A be the complement of the $n \times n$ unit matrix. We define the fractional partition of the rectangle R as follows.

For every $i \in [n]$ we take the size-1 rectangle $R_i = \{(x_i, y_i)\}$ and give it weight $r_i = 1$. To cover the rest of the rectangle R, we use rectangles $R_I = \{(x_i, y_j) \colon i \in I, j \notin I\}$ for all $I \subseteq [n]$ of size $|I| = n/2$, and give them weight

$$r_I = \left(4 - \frac{4}{n}\right) \binom{n}{n/2}^{-1}.$$

This is a fractional partition, because rectangle R_I contains $n^2/4$ of the $n^2 - n$ ones in A and there are $\binom{n}{n/2}$ such rectangles.

For every $i \in [n]$ we have that $\mu_A(R_i) = 0$ since we have only 0's on the diagonal of A. For every subset I of $[n]$ we have that $\mu_A(R_I) = 1$ since there are no 0's outside the diagonal, implying that A_{R_I} is an all-1 matrix. Hence, on the right hand side of the corresponding inequality (2) for convexity we have the sum of n zeros (the ranks of the size one matrices on the diagonal) and $\binom{n}{n/2}$ terms each being at most $4\binom{n}{n/2}^{-1}$, implying that the right hand sums to at most 4. On the other hand, since $\mathrm{rank}(A)$ is n or $n-1$ (which depends on n and the field), on the left hand side we have $\mu_A(R) \geq (n-1)/2$: by the construction of R, no monochromatic sub-rectangle M of R can hit the diagonal in more than one entry, implying that $\mathrm{rank}(A_M) \leq 2$. $\square$

We have shown that, for some measures μ_A, the convexity inequality (2) fails badly: the right hand side is constant whereas the left had side is $\Omega(n)$. Since the measures μ_A based on the rank are not convex, Theorem 3.5 does not apply for them. Still, Razborov in [15] proved that these measures belong to the class of so-called submodular measures, and none of them can yield larger than $O(n)$ lower bound.

6.3. Matrix norms

Interesting measures can be obtained from matrix norms. A mapping $A \mapsto \|A\|$ is a matrix norm if it satisfies all the properties of vector norms:

(i) $\|A\| \geq 0$ with equality if and only if $A = 0$;
(ii) $\|rA\| = |r| \cdot \|A\|$ for all numbers r and all matrices A, and
(iii) $\|A + B\| \leq \|A\| + \|B\|$ for all matrices A and B.

(Often, the sub-multiplicativity $\|A \cdot B\| \leq \|A\| \cdot \|B\|$ is also required, but we do not require this here.) In particular, every matrix norm is a subadditive function in the sense of Section 4, and the rectangle function $\mu(R) = \|A_R\|$ is convex. By Corollary 4.3.3, if φ is a non-decreasing convex real function and s is an additive rectangle function, then the rectangle function

$$\mu(R) = s(R) \cdot \varphi\left(\frac{\|A_R\|}{s(R)}\right), \tag{15}$$

is also convex, and hence cannot give better than $O(n^2)$ lower bounds. We give two examples of measures that appear in the literature.

6.3.1. Factorization norm

Factorization norm $\gamma_2(A)$, is mainly used in Banach space theory. Linial and Shraibman used this norm to prove lower bounds on the quantum communication complexity [13]. It has several equivalent definitions one of which is:

$$\gamma_2(A) = \max_{\|B\|_2=1} \|A \circ B\|_2,$$

where $A \circ B$ is the Hadamard (i.e. component wise) product of matrices and

$$\|A\|_2 = \max_{u,v \neq 0} \frac{|u^t A v|}{\|u\| \cdot \|v\|}$$

is the spectral norm of A; here, $\|u\| = (\sum_i u_i^2)^{1/2}$ is the Euclidean norm of vector u. Since γ_2 is a norm, any rectangle measure of the form (15) that uses γ_2 can yield at most quadratic lower bounds.

6.3.2. Spectral norm and its square

Recently, the spectral norm of matrices was used to introduce a number of rectangle measures. They are based on the rectangle functions

$$\sigma_A(R) = \|A_R\|_2^2$$

for particular matrices A. One can show that this function is subadditive, hence, if we normalize it we obtain a measure.

Koutsoupias [10] first introduced this function for the "distance-one" matrix A defined by (13), and showed that $L(f) \geq \sigma_A(S_f)$ holds for this matrix. (Note that for these matrices σ_A is normalized.)

Barnum, Saks and Szegedy [3] introduced a parameter of boolean functions defined by

$$\mathsf{SA}(f) := \max_{A \neq 0} \frac{\|A\|_2}{\max_i \|A_i\|_2},$$

where A ranges over all nonzero non-negative matrices on S_f and $A_i[x,y] = A[x,y]$ if $x_i \neq y_i$ and 0 otherwise. They used it for lower bounds on quantum query complexity. Laplante, Lee and Szegedy [11] proved that $\mathsf{SA}(f)^2$ is a lower bound on formula size complexity. Høyer, Lee and Špalek [5] proved that the non-negativity restriction can be removed both for lower bounds on quantum query complexity and lower bounds on formula size.

Clearly, SA is connected with σ_A via the equality

$$\mathsf{SA}(f)^2 = \max_{A \neq 0} \frac{\sigma_A(S_f)}{\max_{M \in \mathcal{M}(S_f)} \sigma_A(M)}.$$

Hence one can derive some properties of SA from the properties of σ_A. In particular, Theorem 7 of Lee's paper [12] shows that $\sigma_A(R)$ is a convex rectangle function in our sense. We will show that this fact is a consequence of Corollary 4.3.

Proposition 6.2 ([12]). *For every matrix A, the rectangle function $\sigma_A(R)$ is convex.*

PROOF. Let u, v be vectors such that $u^t A v = \|A\|_2$. For a rectangle $R = X \times Y$, let u_R denote u restricted to X and let v_R denote v restricted to Y. Then the measure $\sigma_A(R)$ has the form

$$\sigma_A(R) = s(R) \cdot \varphi\left(\frac{w(R)}{s(R)}\right),$$

where $\varphi(x) = x^2$, $s(R) = \|u_R\|^2 \cdot \|v_R\|^2$ and $w(R) = |u_R^t A_R v_R|$. The rectangle function

$$s(R) = \|u_R\|^2 \cdot \|v_R\|^2 = \left(\sum_{x \in X} u[x]^2\right) \cdot \left(\sum_{y \in Y} v[y]^2\right) = \sum_{(x,y) \in R} u[x]^2 v[y]^2$$

is additive. So, by Corollary 4.3(3), it is enough to verify that $w(R) = |u_R^t A_R v_R|$ is a convex rectangle function. To do this, let $R = \sum_k r_k R_k$ be a fractional partition of R. Then

$$|u_R^t A_R v_R| = \left|\sum_{(x,y) \in R} u_R^t[x] A_R[x, y] v_R[y]\right| = \left|\sum_k r_k \sum_{(x,y) \in R_k} u_{R_k}^t[x] A_{R_k}[x, y] v_{R_k}[y]\right|$$

$$\leq \sum_k r_k \left|\sum_{(x,y) \in R_k} u_{R_k}^t[x] A_{R_k}[x, y] v_{R_k}[y]\right| = \sum_k r_k |u_{R_k}^t A_{R_k} v_{R_k}|.$$

Hence, $w(R)$ is convex, and we are done. $\qquad\square$

The convexity of $\sigma_A(R)$ together with Theorem 3.5 implies that $\mathsf{SA}(f)^2 \leq 9/8n^2$. Note, however, that in [11] Laplante, Lee and Szegedy proved a little more: $\mathsf{SA}(f)^2 \leq n^2$.

7. Open problems

Problem 7.1. Can rectangle functions $\mu_F(R) = F(w(R), s(R))$ with $F(x, y)$ subadditive and both $w(R)$ and $s(R)$ additive yield super-quadratic lower bounds?

Problem 7.2. Is it possible to generalize the quadratic upper bound of Theorem 5.2 to measures of the form

$$\mu(R) = \frac{w(R)^k}{s(R)^{k-1}},$$

where $s(R)$ is an arbitrary additive and positive measure?

We only have such upper bounds for $w(R)$ subadditive and $s(R) = |R|$, or $w(R)$ convex and $s(R)$ additive and positive. The problem is to find a common generalization of these two cases.

Problem 7.3. Is it possible to prove super-polynomial lower bounds on monotone formulas using convex measures?

This is equivalent to the problem of [7] whether the monotone fractional covering number can be super-polynomial.

Problem 7.4. Prove a super-quadratic lower bound using formal complexity measures.

Interpreting Andreev's [2] or Håstad's [4] proof in terms of measures may be a way to make progress in lower bounds on the formula size complexity.

17

References

[1] A. Aho, J. Ullman, M. Yannakakis, On notions of information transfer in VLSI circuits, in: Proc. of 15th ACM STOC (1983) pp. 133–139.

[2] A.E. Andreev, On a method for obtaining more than quadratic effective lower bounds for the complexity of π-schemes, Moscow Univ. Math. Bull. 42(1) (1987) 63–66.

[3] H. Barnum, M. Saks, M. Szegedy, Quantum decision trees and semidefinite programming, in: Proc. 18th IEEE Conference on Computational Complexity, (2003) pp. 179–193.

[4] J. Håstad, The shrinkage exponent is 2, SIAM J. on Comput. 27 (1998) 48–64.

[5] P. Høyer, T. Lee, R. Spalek, Negative weights make adversaries stronger, in: Proc. of 39th Ann. ACM Symp. on Theory of Computing (2007) pp. 526–535.

[6] M. Karchmer, Communication Complexity: A New Approach to Circuit Depth, MIT Press, 1989.

[7] M. Karchmer, E. Kushilevitz, N. Nisan, Fractional covers and communication complexity, SIAM J. on Discrete Math. 8(1) (1995) 76–92.

[8] M. Karchmer, A. Wigderson, Monotone circuits for connectivity require super-logarithmic depth SIAM J. on Discrete Math. 3 (1990) 255–265.

[9] M. V. Khrapchenko, A method of obtaining lower bounds for the complexity of π-schemes, Math. Notes Acad. Sci. USSR 10 (1972) 474–479.

[10] E. Koutsoupias, Improvements on Khrapchenko's theorem, Theoret. Comput. Sci. 116 (1993) 399–403.

[11] S. Laplante, T. Lee, M. Szegedy, The quantum adversary method and formula size lower bounds, Computational Complexity 15(2) (2006) 163–196.

[12] T. Lee, A new rank technique for formula size lower bounds, in: Lect. Notes in Comput. Sci., vol. 4393, Springer, Berlin, 2007, pp. 145–156.

[13] N. Linial, A. Shraibman, Lower bounds in communication complexity based on factorization norms, Random Structures and Algorithms, 2008. to appear.

[14] A. A. Razborov, Applications of matrix methods to the theory of lower bounds in computational complexity, Combinatorica 10(1) (1990) 81–93.

[15] A. A. Razborov, On submodular complexity measures, in: Boolean Function Complexity, London Math. Soc. Lecture Note Series 169 (1992) 76–83.

[16] K. L. Rychkov, A modification of Khrapchenko's method and its applications to lower bounds for π-schemes of code functions, Metody Diskretnogo Analiza 42 (1985) 91–98. (in Russian).

[17] S. V. Yablonskii, Realization of linear functions in the class of Π-schemes, Doklady Akademii Nauk SSSR 94(5) (1954) 805–806. (in Russian).

[18] B. A. Subbotovskaya (1961), Realizations of linear functions by formulas using $+, ., -$, Doklady Akademii Nauk SSSR, 136(3) (1961) 553–555 (in Russian). English translation in: Soviet Mathematics Doklady 2 (1961), 110–112.

[19] I. Wegener, The Complexity of Boolean Functions, Wiley-Teubner, 1987.

Appendix: Proof of Lemma 5.3

To prove the first claim (i), let $a \geq 1$ and $\alpha \in [0, 1)$ be given. Our goal is to determine

$$\xi(a) = \max_{x,y \in [0,1]} h_a(x, y),$$

where

$$h_a(x, y) = (xy)^\alpha + ((1-x)(1-y))^\alpha + a(x(1-y))^\alpha + a((1-x)y)^\alpha.$$

The function $h(x, y) := h_a(x, y)$ is continuous and hence it attains maximum on the square $P = [0, 1] \times [0, 1]$. The maximum can be reached either in the interior of P, or on the boundary. The boundary itself consists of the corners and the sides of P. We consider these cases separately.

The corners. We obtain

$$h(0, 0) = h(1, 1) = 1, \qquad h(0, 1) = h(1, 0) = a.$$

18

The sides. Set $y := 1$ and let us determine critical points of $h(x, 1)$ on $(0, 1)$. Setting the x-derivative of $h(x, 1)$ to 0 gives

$$x^{\alpha-1} - a(1-x)^{\alpha-1} = 0.$$

Hence the only critical point is at

$$x = \frac{a^{\frac{1}{\alpha-1}}}{1 + a^{\frac{1}{\alpha-1}}},$$

and the value of $h(x)$ is

$$a(1 + a^{\frac{1}{\alpha-1}})^{1-\alpha}.$$

The other cases are symmetric.

The interior. Since $h(x, y) = h(1 - x, 1 - y)$, h has a critical point at $(x, y) = (1/2, 1/2)$. The value of $h(x, y)$ at this point is

$$2^{1-2\alpha}(1 + a).$$

There are no other critical points, since the x-partial derivative is strictly monotone in x and hence it can have at most one zero.

Altogether we get

$$\max_P h(x, y) = \max\left\{1, a, a(1 + a^{\frac{1}{\alpha-1}})^{1-\alpha}, 2^{1-2\alpha}(1 + a)\right\}. \tag{16}$$

Since $a \geq 1$, this gives $\max_P h(x, y) = \max\left\{a(1 + a^{\frac{1}{\alpha-1}})^{1-\alpha}, 2^{1-2\alpha}(1 + a)\right\}$.

To prove the second claim (ii), let $\alpha = \frac{1}{2}$. Then $\xi(a) = \max\left\{a(1 + a^{-2})^{\frac{1}{2}}, (1 + a)\right\}$, and we must show that

$$d + 1 \geq \max\left\{d(1 + d^{-2})^{\frac{1}{2}}, 1 + d\right\},$$

which is immediate.

To prove the last claim (iii), let $\alpha > \frac{1}{2}$. We must find $c \geq 1$ such that

$$\begin{aligned}
c \cdot (d + 1)^{1-\alpha} &\geq c \cdot d^{1-\alpha}(1 + (c \cdot d^{1-\alpha})^{\frac{1}{\alpha-1}})^{1-\alpha}, \\
c \cdot (d + 1)^{1-\alpha} &\geq 2^{1-2\alpha}(1 + c \cdot d^{1-\alpha}).
\end{aligned}$$

The first inequality is satisfied by any $c \geq 1$. Since $1 - \alpha > 0$, it is equivalent to

$$d + 1 \geq d \cdot (1 + (c \cdot d^{1-\alpha})^{\frac{1}{\alpha-1}})$$

and hence to $d + 1 \geq d + c^{\frac{1}{\alpha-1}}$ resp. to $c^{\frac{1}{1-\alpha}} \geq 1$. The second inequality will be satisfied, if

$$c \cdot ((d + 1)^{1-\alpha} - 2^{1-2\alpha} \cdot d^{1-\alpha}) \geq 2^{1-2\alpha}.$$

We have

$$c \cdot ((d + 1)^{1-\alpha} - 2^{1-2\alpha} \cdot d^{1-\alpha}) \geq c \cdot d^{1-\alpha}(1 - 2^{1-2\alpha}).$$

Our assumption $\alpha > 1/2$ implies $2^{1-2\alpha} < 1$, and it is sufficient to set

$$c = \frac{2^{1-2\alpha}}{1 - 2^{1-2\alpha}} = \frac{1}{2^{2\alpha-1} - 1}.$$

If $\alpha \in (\frac{1}{2}, 1)$ then $c > 1$. $\qquad\square$