On Convex Complexity Measures
(Draft — do not distribute!)

P. Hrubeš* S. Jukna† A. Kulikov‡ P. Pudlák§

April 6, 2009

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 subrectangles 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.

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 nontrivially 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 [6] can be recast in this formalism (see e.g. [14], Sect. 8.8). Generalizing Khrapchenko's argument for the parity function, Rychkov [13] 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.

In order to obtain larger lower bounds, Razborov [11] proposed to look at rectangles as matrices over some field and introduce appropriate measures on subrectangles in terms of the corresponding submatrices. Razborov studied the measures based on the rank of matrices.

*School of Mathematics, Institute for Advanced Study, Princeton NJ. Email: pahrubes@centrum.cz Partially supported by NSF grant CCF 0832797.

†Institute of Mathematics, Vilnius. Email: jukna@thi.informatik.uni-frankfurt.de Partially supported by a DFG grant SCHN 503/4-1.

‡Steklov Institute of Mathematics at St. Petersburg. Email: kulikov@logic.pdmi.ras.ru Partially supported by RFBR (grant 08-01-00640-a), RAS Program for Fundamental Research ("Modern Problems of Theoretical Mathematics"), and grant NSh-2203-2003-1 of the President of Russian Federation for Leading Scientific Schools Support.

§Mathematical Institute, Prague. Partially supported by grants A1019401 and 1M0021620808.

He showed that the rank can only give linear lower bounds for the De Morgan basis, but it gives superpolynomial 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 [7, 8, 10]. 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. The only superquadratic lower bound, the lower bound $\Omega(n^{3-o(1)})$ of [2], has not been translated into the formalism of measures yet.

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 [4]. Using a different technique we will also prove linear and quadratic upper bounds on some other measures that are closely related to convex measures.

2 Basic concepts

Let n be a fixed positive integer, let $\mathcal{F}$ denote the set of all boolean functions $f : \{0,1\}^n \to \{0,1\}$. Literals are boolean variables and their negations. Let $L(f)$ denote *the formula size complexity* of f in the De Morgan basis, i.e., the minimal number of occurrences of literals in a formula expressing f using connectives $\{\vee, \wedge\}$. A function $m : \mathcal{F} \to \mathbb{R}$ 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 $g, h \in \mathcal{F}$.

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 loosing 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 $S = S^0 \times S^1$. 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 of the form

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

but there are much more rectangles than boolean functions.

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 $\epsilon \in \{0,1\}$ such that for all $(x, y) \in M$, $x_i = \epsilon$ and $y_i = 1 - \epsilon$; 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

$$M_{i,\epsilon} = \{(x, y) \mid x_i = \epsilon \text{ and } y_i = 1 - \epsilon\}\,.$$

These $2n$ rectangles cover every rectangle.[COMMENT: SK: I would prefer to give a more formal definition here and in what follows. For example, what are (x, y) here? We are usually interested in subrectangles of S only. So, why not define $M_{i,\epsilon}(S) = \{(x, y) \in S \ldots$?]

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. In this case we will call R a *subrectangle* of S. In what follows, $\mathcal{R}(S)$ will denote the set of all subrectangles and $\mathcal{M}(S)$ the set of all monochromatic subrectangles of S.

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.

2.1 Subadditive measures and communication complexity

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

A *rectangle measure* is a rectangle function $\mu : \mathcal{R} \to \mathbb{R}$ satisfying the following two conditions:

(i) *Normalization*: $\mu(M) \leq 1$ for every monochromatic rectangle $M \in \mathcal{M}$.

(ii) *Subadditivity*: $\mu(R) \leq \mu(R_1) + \mu(R_2)$, for every rectangle $R \in \mathcal{R}$ and for every its partition into disjoint union of rectangles $R_1, R_2 \in \mathcal{R}$.

[COMMENT: pH. Changed "weak subadditivity" to "subadditivity", and the latter to "strong subadditivity". The usage was inconsistent anyway.] The first condition is usually achieved by normalization. That is, if a rectangle function ν is subadditive, 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.

The connection between rectangle measures can be best seen [COMMENT: pH. On the contrary, I would prefer not to get into this. Why dont we speak directly about formula size? SK: I still think it is better to leave this part here. Of course, we can relate formula size to the number of rectangles directly (as Rychkov did), but still. Note that Razborov also describes this connection in his paper.] in the framework of communication games, as introduced by Karchmer and Wigderson [5]: 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 subrectangles R_1 or R_2. Let $\Gamma(R)$ denote the minimal number of leaves in a tree like communication protocol for a rectangle R in a Karchmer-Wigderson game. Then $L(f) = \Gamma(S_f)$ [5]. 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 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 $\mu(R) := \min\{m(f) \mid f \text{ separates } R\}$, is a rectangle measure.*

2.2 Subadditive measures and the partition number

A more general condition than (ii) has also been considered:

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

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:

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

This measure was implicitly used already in Khrapchenko's proof [6]. 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 $\log_2 L(f) \leq (\log_2 \mathsf{D}(S_f))^2$ [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.

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 $X \subseteq S$, by

$$\mathsf{D}(X) = \min\{K \mid X \text{ can be covered by } K \text{ disjoint monochromatic rectangles}\},$$

and is monotonic with respect to set-inclusion.[COMMENT: SK: Is it really important for us that it can be extended in this way?] 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 and $R_1, \ldots, R_m$ be 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 $e \in R$,

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

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 $R = \sum_i r_i R_i$

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

Karchmer, Kushilevitz and Nisan in [4] 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. 1. 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

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

2. 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

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

□

Theorem 3.2 ([4])**.** $\mathsf{D}^* \leq 4n^2$.

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 [3], and Karchmer, Kushilevitz and Nisan [4], associate with each subset $I \subseteq [n] = \{1, \ldots, n\}$ the following two *parity rectangles*.

$$P_{I,\epsilon} = \{x \in \{0,1\}^n \mid \oplus_{i \in I}\, x_i = \epsilon\} \times \{y \in \{0,1\}^n \mid \oplus_{i \in I}\, y_i = 1 - \epsilon\}\,,\ \epsilon = 0, 1\,.$$

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

Lemma 3.3. *Every edge $(x, y) \in \{0,1\}^n \times \{0,1\}^n$ s.t. $x \neq y$ 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 $I \subseteq [n]$, $\epsilon = 0, 1$, we have $\mu(P_{I,\epsilon} \cap S) \leq \frac{9}{8}|I|^2$.*

Proof. A parity rectangle $P_{I,\epsilon}$ can be viewed as a rectangle corresponding to the parity function in $|I|$ variables, or its negation. If $n = 2^l + k$ then parity can be computed by a formula of size $c(n) = 2^l(2^l + 3k)$, see [9]. This gives

$$c(n) \leq \frac{9}{8}n^2.$$

To see that, observe that the function $\frac{y(y+3x)}{(y+x)^2}$, $x \in (0, y)$ reaches maximum at the point $x = \frac{y}{3}$, and it has value $\frac{y(y+3\frac{y}{3})}{(y+\frac{y}{3})^2} = \frac{9}{8}$. Hence $\mu(S \cap P_{I,\epsilon}) \leq \frac{9}{8}|I|^2$, since μ is a lower bound to the formula size. $\square$

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

$$\mu(S) \leq \frac{9}{8}n^2(1 + o(1)).$$

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

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

and let $\mathcal{R}_{\text{par}} := \bigcup_{i=1,\ldots,n,\epsilon=0,1} \mathcal{R}_{i,\epsilon}$. For technical reasons, we shall understand $\mathcal{R}_{\text{par}}$ as a multiset, elements of $\mathcal{R}_{\text{par}}$ 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 $\mathcal{R}_{\text{par}}$. Hence $\mathcal{R}_{\text{par}}$ form a fractional partition of S with each rectangle $R \in \mathcal{R}_{\text{par}}$ of weight $r_R = 2^{-(n-1)}$.

By the previous lemma, we know that for every $R \in \mathcal{R}_{i,\epsilon}$, $\mu(R) \leq c_p i^2$, where $c_p = \frac{9}{8}$. For every $i = 1, \ldots n$, $\epsilon = 0, 1$, $|\mathcal{R}_{i,\epsilon}| = \binom{n}{i}$.

The convexity of μ implies that [COMMENT: pH. An additional factor of 2 comes from $\epsilon = 0, 1$. SK: Yes, I agree that 2 was missing in the previous version.]

$$\mu(S) \leq \sum_{R \in \mathcal{R}_{\text{par}}} r_R \cdot \mu(R) = 2^{-(n-1)} \sum_{R \in \mathcal{R}_{\text{par}}} \mu(R) = 2^{-(n-1)} \sum_{i,\epsilon} \sum_{R \in \mathcal{R}_{i,\epsilon}} \mu(R)$$
$$\leq 2^{-(n-1)} \sum_{i=1}^{n} \sum_{\epsilon=0,1} \binom{n}{i} c_p i^2 = 2^{-(n-1)} 2c_p \sum_{i=1}^{n} \binom{n}{i} i^2 = 2^{-(n-2)} c_p \sum_{i=1}^{n} \binom{n}{i} i^2 .$$

The identity $\binom{n}{k} \cdot k = n \cdot \binom{n-1}{k-1}$ gives [COMMENT: The 2 disappears here.]

$$\begin{aligned}
\sum_{i=1}^{n} \binom{n}{i} i^2 = n \cdot \sum_{i=1}^{n} \binom{n-1}{i-1} i &= n \cdot \sum_{i=1}^{n} \binom{n-1}{i-1} + n \cdot \sum_{i=1}^{n} \binom{n-1}{i-1}(i-1) \\
&= n \cdot \sum_{i=1}^{n} \binom{n-1}{i-1} + n \cdot \sum_{i=2}^{n} \binom{n-1}{i-1}(i-1) \\
&= n \cdot \sum_{i=1}^{n} \binom{n-1}{i-1} + n(n-1) \cdot \sum_{i=2}^{n} \binom{n-2}{i-2} \\
&= n2^{n-1} + n(n-1)2^{n-2} = (n^2+n)2^{n-2}\,.
\end{aligned}$$

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

4 General construction of convex measures

In his seminal paper [6], Khrapchenko proved a general lower bound on formula size complexity of the form

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

where R is a subrectangle of S_f. Paterson 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 $Y = \{(x,y) \mid \mathrm{dist}(x,y) = 1\}$ is the set of all vector pairs of Hamming distance 1, the *Khrapchenko measure*. One can also interpret Rychkov's lower bounds on error correcting codes as lower bounds based on the Khrapchenko measure. There one uses pairs of distance $d+1$ instead of Y for codes of the minimal distance $2d+1$.[COMMENT: SK: it was pairs of distance d for codes of the minimal distance $d > 1$.]

We can interpret Khrapchenko's lower bound as follows. One starts with rectangle functions $s_1(R)$, $s_2(R)$, which themselves do not give better than linear lower bounds. We define a new rectangle function $\mu(R) = F(s_1(R), s_2(R))$ by means of a real function F, 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 some given 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 subadditve measure, but if F has too nice properties, it will produce a convex measure.

Notice that the Khrapchenko measure has the form

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

with $\mathsf{w}(R) = |R \cap Y|$, $\mathsf{s}(R) = |R|$ and $\varphi(x) = x^2$. Subadditivity of μ 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 $\mathsf{w}(R)$, $\mathsf{s}(R)$ 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 functions 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:

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

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 [4] in order to apply the linear programing 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.)

First, let us note two simple closure properties of subadditive and convex rectangle functions.

Proposition 4.1. *1. A non-negative linear combination of subadditive resp. of convex rectangle functions is a subadditive resp. convex rectangle function.*

2. If $\mu_z, z \in J$ is a set of subadditive resp. convex rectangle functions then $\max_{z \in J} \mu_z$ is a subadditive resp. convex rectangle function.

Proof. Immediate. (1. follows from the fact that $\sum_i a_i x_i$ is a non-decreasing subnorms and Proposition 4.2.) □

Let $F(\overline{x})$ be a real function in m variables. We shall think of m-tuples of real numbers as vectors in $\mathbb{R}^m$. We assume that the domain of F is closed w.r. to addition of vectors, and multiplication by a positive real number. We say that F is *subadditive*, if

$$F(\overline{x_1} + \overline{x_2}) \leq F(\overline{x_1}) + F(\overline{x_2})\,. \tag{5}$$

We say that it is *a subnorm*, if, in addition, for every $a > 0$

$$F(a \cdot \overline{x}) \leq aF(\overline{x})\,. \tag{6}$$

Proposition 4.2. *Let a rectangle function μ be defined as follows:*

$$\mu(R) = F(w_1(R), \ldots, w_p(R), s_1(R), \ldots, s_q(R))\,, \tag{7}$$

where F is subadditive and $s_i(R)$ are additive rectangle functions.[COMMENT: SK: I would say that this proposition is too heavy at the moment.]

1. *If the weights $w_i(R)$ are additive, then μ is subadditive. If $F(x_1 \ldots x_p, \overline{y})$ is nondecreasing in $x_i, i = 1, \ldots, p$ and $w_i(R)$ are subadditive then μ is subadditive.*
2. *If F is a subnorm and $w_i(R)$ are additive then μ is convex. If F is nondecreasing in $x_i, i = 1, \ldots, p$ and $w_i(R)$ are convex then μ is convex.*

Proof. We prove part 2., part 1. is easier. First note that if F is a subnorm and a_i are non-negative then

$$F(\sum a_i \overline{x_i}) \leq \sum a_i F(\overline{x_i})\,, \tag{8}$$

where the sum is over i s.t. a_i is nonzero. To prove the first claim of part 2., assume that both $w_i(R)$ and $s_i(R)$ are additive, and let $\sum_{i=1,\ldots,m} r_i R_i$ be a fractional partition of R. Set $s^i_j = s_i(R_j)$ and $w^i_j = w_i(R_j)$. By additivity, we have that $w_i(R) = \sum_j r_j \cdot w^i_j$ and $s_i(R) = \sum_j r_j \cdot s^i_j$. Hence

$$\begin{aligned}
\mu(R) &= F\left(w_1(R), \ldots, w_p(R), s_1(R), \ldots, s_q(R)\right) && (9)\\
&= F\left(\sum_j r_j \cdot w^1_j, \ldots, \sum_j r_j \cdot w^p_j, \sum_j r_j \cdot s^1_j, \ldots, \sum_j r_j \cdot s^q_j\right) \\
&\leq \sum_j r_j F(w^1_j, \ldots, w^p_j, s^1_j, \ldots, s^q_j) && (10)\\
&= \sum_j r_j \mu(R_j)\,.
\end{aligned}$$

If F is nondecreasing and $\mathsf{w}(R)$ is convex, then we can replace the equality in (9) by inequality, and the desired inequality $\mu(R) \leq \sum_i r_i \cdot \mu(R_i)$ still holds. □

Note that subadditivity of F guarantees subadditivity of μ, and hence μ can be (after appropriate normalisation) used as a rectangle measure for proving lower bounds. But if F is also a subnorm, μ will be convex and the lower bounds given by μ cannot exceed $O(n^2)$. However, there are many subadditive real functions that are not subnorms. It is not clear whether the function F can be chosen in such a way that μ will give better than quadratic lower bounds. [COMMENT: pH. Note that we can replace the definition of subnorm by a single weaker condition $F(\sum_i r_i\overline{x}_i) \leq (\sum_i r_i)\max_i F(\overline{x}_i)$, and the $O(n^2)$ upper bound still applies.]

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

$$\mu(R) = \mathsf{s}(R) \cdot \varphi\left(\frac{\mathsf{w}(R)}{\mathsf{s}(R)}\right), \tag{11}$$

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

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

Proof. It is sufficient to prove that the function $F(x, y) = y\phi(\frac{x}{y})$ is a subnorm. The condition $F(ax, ay) \leq aF(x, y)$ is immediate. (This is in fact equality and F is a norm.) Subadditivity of F is an application of Jensen's inequality.

$$\varphi\left(\frac{y_1z_1 + y_2z_2}{y_1 + y_2}\right) \leq \frac{y_1\varphi(z_1) + y_2\varphi(z_2)}{y_1 + y_2}. \tag{12}$$

Assume $y_1, y_2 > 0$. Setting $z_1 = \frac{x_1}{y_1}$, $z_2 = \frac{x_2}{y_2}$, we obtain that

$$\varphi(\frac{x_1 + x_2}{y_1 + y_2}) \leq \frac{y_1\varphi(\frac{x_1}{y_1}) + y_2\varphi(\frac{x_2}{y_2})}{y_1 + y_2}.$$

Hence

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

□

5 Polynomial rectangle measures

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

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

where $\mathsf{s}(R)$ is a positive additive rectangle function and $\mathsf{w}(R)$ is subadditive. We call them *polynomial measures of degree k*.

Since $\varphi = x^k, k \geq 1$ is a nondecreasing convex function, Theorem 4.3 implies that polynomial measures are subadditive measures. Moreover, if $\mathsf{w}(R)$ is convex then $\mu(R)$ is also convex. Therefore, by Theorem 3.5, polynomial measures with convex $\mathsf{w}(R)$ can yield at most quadratic lower bounds.

On the other hand, every subadditive measure is a polynomial measure of degree one. This shows that polynomial measures can in principle give exponential lower bounds. Similarly, polynomial measures can attain exponential values for every $1 \leq k < 2$. [COMMENT: SK: At this point, this is not obvious. Moreover, this is repeated (again, without explanation) in Remark . I would suggest to eliminate this paragraph and to add the corresponding explanation in the Remark.]

Using $\mathsf{w}(R)$ additive, quadratic lower bounds were proved with $k = 2$, by Khrapchenko [6], and with $k = 1$ by Karchmer, Kushilevitz and Nisan [4]. [COMMENT: "We shall show that, in a sense, the exponents $k = 1$ and $k = 2$ are the best possible." If we use w non-negative additive then $k = 2$ is best possible. If w can be negative, we do not have any statement for $k < 2$. pH]

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

Note that the normalization condition $\mu(M) \leq 1$, for a monotone rectangle M implies that

$$\mathsf{w}(M) \leq (\mathsf{s}(M))^{\frac{k-1}{k}} \tag{13}$$

Since every n-dimensional rectangle can be (non-disjointly) covered by at most $2n$ rectangles, we obtain that any polynomial measure of degree k cannot give lower bounds larger than $O(n^k)$, as long as the used weight function $\mathsf{w}(R)$ is *additive non-negative*. In what follows, let S_n stand for an *arbitrary* n-dimensional rectangle.

Proposition 5.1. *Let μ be a polynomial measure of degree $k \geq 1$. If the used weight function $\mathsf{w}(R)$ is additive and non-negative, then $\mu(S) \leq (2n)^k$.*

Proof. Using (13) we have

$$\mathsf{w}(S) \leq \sum_{i,\epsilon} \mathsf{w}(M_{i,\epsilon}) \leq \sum_{i,\epsilon} \mathsf{s}(M_{i,\epsilon})^{\frac{k-1}{k}} \leq 2n\mathsf{s}(S)^{\frac{k-1}{k}}.$$

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

□

Hence, if the used weight function $\mathsf{w}(R)$ is *non-negative*, then no polynomial measure of degree $k < 2$ can even reach the n^2 lower bound.

Remark 5.2. *The assumption that $\mathsf{w}(R)$ is non-negative cannot be omitted. An additive ($k = 1$) measure is used in [4] to prove a quadratic lower bound. The assumption that $\mathsf{w}()$ is additive also cannot be omitted. Every subadditive measure is a polynomial measure of degree $k = 1$. Hence polynomial measures of degree $k = 1$ can yield exponential lower bounds. The same applies to any $1 \leq k < 2$.* [COMMENT: SK: I would suggest here to explain that one gets a measure $L(R)^k/|R|^{k-1}$ in order to obtain an exponential lower bound.]

There exist polynomial measures with $1 \leq k \leq 2$ and w positive additive proving $\Omega(n^k)$ lower bounds: consider

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

where $\mathsf{w}(R)$ is the number of pairs $(x, y) \in R$ of Hamming distance 1, and c is a constant which guarantees the normalization condition. [COMMENT: SK: Why mot to give the exact value of c here? ($c = 2^{-(2-k)(n-1)}$) I think that we should also mention that this measure is used for parity.]

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

We now show that every polynomial measure of degree $k > 2$, with $\mathsf{w}(R)$ subadditive and $\mathsf{s}(R) = |R|$ can give at most linear lower bounds.

Theorem 5.3. *Let μ be a rectangle measure of the form*

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

where $\mathsf{w}(R)$ is a subadditive rectangle function.

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

We first prove the following lemma:

Lemma 5.4. *Let $a \geq 1$ and $\alpha \in [0, 1)$, and let*

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

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

1. $\xi(a) = \max\left\{a(1 + a^{\frac{1}{\alpha-1}})^{1-\alpha},\ 2^{1-2\alpha}(1 + a)\right\}$.
2. *If $\alpha = \frac{1}{2}$ then for every $d \geq 1$*
$$d + 1 \geq \xi(d)\,. \tag{14}$$
3. *If $\alpha > \frac{1}{2}$ then there exists a constant c s.t for every $d \geq 1$,*
$$c \cdot (d + 1)^{1-\alpha} \geq \xi(c \cdot d^{1-\alpha})\,. \tag{15}$$

Proof. 1. Let $a \geq 1$ be given. 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 those cases separately.

The corners. We obtain

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

The sides. Let $y := 1$ and let us determine critical points of $h(x, 1)$ on $(0, 1)$. The equation $\frac{\partial h(x,1)}{\partial x} = 0$ 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 $\max_P h(x,y) = \max\left\{a(1+a^{\frac{1}{\alpha-1}})^{1-\alpha}, 2^{1-2\alpha}(1+a)\right\}$.

2. 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.

3. Let $\alpha > \frac{1}{2}$. We must find $c \geq 1$ s.t.

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

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

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

$\alpha > \frac{1}{2}$ 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 $c > 1$. $\square$

Proof of Proposition 5.3. Let $S = S^0 \times S^1$. Since μ is normalized, we have that

$$\mathsf{w}(M) \leq |M|^{1-1/k} \tag{17}$$

for every monochromatic rectangle M. Let $\alpha := 1 - 1/k$. For $k \leq 2$, α is in the interval $(\frac{1}{2}, 1)$.

We will prove the following claim:

Claim 5.5. *1. If $k = 2$ then $\mathsf{w}(S) \leq n \cdot |S|^{\frac{1}{2}}$.*

2. If $k > 2$ then $\mathsf{w}(S) \le c|S|^\alpha n^{1-\alpha}$, where c is the constant from (15).

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

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

and in the case $k > 2$

$$\mu(S) = \frac{\mathsf{w}(S)^k}{|S|^{k-1}} \le \frac{(c \cdot |S|^\alpha n^{1-\alpha}))^k}{|S|^{k-1}} = c^k n = O(n)\,.$$

It remains to prove the claim. Let $\dim(R) = |\{i \mid \exists (x,y) \in R: \; x_i \ne y_i\}|$, and let

$$\mathsf{w}(m,d) = \max\{\mathsf{w}(R) \mid \dim(R) \le d \text{ and } |R| = m\}\,.$$

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 (17), we have that

$$\mathsf{w}(m,1) \le m^\alpha\,.$$

Since $\mathsf{w}(R)$ is subadditive, we have a recurrent inequality

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

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

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

We look for a solution of the form

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

Hence $g(d)$ is to satisfy the inequalities $g(1) \ge 1$ and

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

Using the definition from the previous lemma, it is sufficient to have $g(d) \ge 1$ and

$$g(d+1) \ge \xi(g(d)).$$

The lemma then asserts that for $\alpha = \frac{1}{2}$ (i.e., $k = 2$) $g(d) = d$ is a solution, and for $\alpha \ge \frac{1}{2}$ (i.e., $k > 2$), $g(d) = c \cdot d^{1-\alpha}$ is a solution. This completes the proof of the claim. $\square$

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 $A : S \to F$, for some field. The idea of studying matrix parameters for proving lower bounds on formula size complexity is due to Razborov [11].

Matrix rank

[COMMENT: pH. I removed the alternative proof of parity lower bound. If you want, add it back, but with a hint why is it interesting.]

Razborov used the rank of matrices to prove lower bounds on monotone formula size. We can view his approach as defining the subadditive measure

$$\mu(R) = \max_{A \neq 0} \frac{\operatorname{rank}(A_R)}{\max_{M \in \mathcal{M}} \operatorname{rank}(A_M)}, \tag{18}$$

where $A : S \to \mathbb{R}$ is a matrix corresponding to the ambient square and A_R its restriction to the square R. [COMMENT: SK: Note that there is also some ambiguity here. Do we mean $\mathcal{M}(\mathcal{S})$ or $\mathcal{M}(\mathcal{R})$ here? It seems that both are possible. I always kept in mind the first one.] Subadditivity of μ is a consequence of subadditivity of rank (and Proposition 4.1). Unfortunately, as Razborov proved in [12], this measure cannot yield nonlinear lower bounds in the full de Morgan basis. Specifically, he proved that the corresponding formal complexity measure defined on boolean functions satisfies the condition of *submodularity*

$$m(g \wedge h) + m(g \vee h) \leq m(g) + m(h)$$

and that all measures satisfying this condition are bounded by a linear function in n.

However, rank is not a convex complexity measure, as we show in the following:

Proposition 6.1. *Let $\mu(R)$ be defined as in (18). Then μ is not a convex rectangle measure.*

Proof. [COMMENT: pH. I do not find this proof informative. If you think the argument is correct, maybe we can delete it and leave it as an observation?] Let $f : \{0,1\}^3 \to \{0,1\}$ be the parity function with $n = 3$ variables. $S = S_f$ is a 4×4 rectangle. It is covered by 6 canonical rectangles $M_{i,\epsilon}, i = 1, 2, 3, \epsilon = 0, 1$.

Let us first show that $\mu(S) = 4$. It is sufficient to find a 4×4 matrix A s.t. $\frac{\operatorname{rank}(A)}{\max_{M \in \mathcal{M}} \operatorname{rank}(A_M)} = 4$. For $\sigma \in \{0,1\}^3$, let $\sigma^\star \in \{0,1\}^3$ be defined by $\sigma^\star_i = 1 - \sigma_{i+1}$, where $i+1$ is addition modulo three. Let $A_{\sigma,\mu} = 1$, if $\mu = \sigma^\star$ and $A_{\sigma,\mu} = 0$ otherwise. Since the map $\sigma \to \sigma^\star$ is a bijection between $f^{-1}(0)$ and $f^{-1}(1)$, A is a permutation matrix and hence $\operatorname{rank}(A) = 4$. Let us show that for every monochromatic rectangle M, $\operatorname{rank}(A_M) \leq 1$. Without loss of generality assume that $M \subseteq M_{1,\epsilon}$ with $\epsilon = 0, 1$. Assume that $\langle \sigma, \sigma^\star \rangle \in M$. Then $\sigma_1 = \epsilon$, $\sigma^\star_1 = 1 - \epsilon$. Definition of $\sigma^\star$ gives $\sigma^\star_2 = 1 - \epsilon$. Hence $\sigma^\star_3 = 1$, since the parity of $\sigma^\star$ must be odd. Hence also $\sigma_2 = 0$ and $\sigma_3 = \epsilon$. This shows that A_M contains at most one non-zero entry and so $\operatorname{rank}(A_M) \leq 1$.

Let $\mathcal{R}$ be the set of canonical rectangles $\mathcal{M}$, together with the rectangles of the form $\mathcal{S} = \{P_{I,\epsilon} \cap S \mid |I| = 2, \epsilon = 0, 1\}$. $|\mathcal{M}|, |\mathcal{S}| = 6$. Every edge in S is contained in exactly three rectangles in $\mathcal{R}$. If $\langle \sigma_1, \sigma_2 \rangle \in S$ then σ_1 and σ_2 differ in an odd number of bits, i.e., one or three. In the latter case, the edge is contained in three rectangles from $\mathcal{M}$, in the former in one canonical rectangle and two from $\mathcal{S}$. (Or apply Lemma 3.3, excluding S from the partition.) Hence $\mathcal{R}$ forms a fractional cover with weights $1/3$. Since μ is normalised, $\mu(R) \le 1$ for $R \in \mathcal{M}$. Every $R \in \mathcal{S}$ is 2×2 rectangle and hence $\mu(R) \le 2$. This implies

$$\sum_{R\in\mathcal{R}} \frac{1}{3}\mu(R) \le \frac{1}{3}\Big(\sum_{R\in\mathcal{M}} 1 + \sum_{R\in\mathcal{S}} 2\Big) = \frac{1}{3}(6 + 2\cdot 6) = 3 < 4\,.$$

$\square$

[COMMENT: SK: As Pavel P. pointed out there is a miscalculation in the last line. It is not clear to me whether it is possible to close this hole at the moment. But I think that we can use the proof from a recent Stasys' email. It works only for the original Razborov's measure (without maximizing over all A), but it is ok in my opinion.]

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) $\|A\| \ge 0$ with equality if and only if $A = 0$;

(ii) $\|rA\| = |r| \cdot \|A\|$ for all numbers r and all matrices A, and

(iii) $\|A + B\| \le \|A\| + \|B\|$ for all matrices A and B.

In particular, every matrix norm is a subnorm 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

$$\mu(R) = s(R) \cdot \varphi\left(\frac{\|A_R\|}{s(R)}\right), \tag{19}$$

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

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 [10]. 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. componentwise) product of matrices. Since γ_2 is a norm, any rectangle measure of the form (19) that uses γ_2 can yield at most quadratic lower bounds.

Spectral norm. Recall that the spectral norm $\|A\|_2$ of a matrix A can be defined by

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

Laplante, Lee and Szegedy [8] introduced the parameter of boolean functions f defined by

$$\mathsf{sumPI}(f) := \max_{A\neq 0} \frac{\|A_R\|_2^2}{\max_M \|A_M\|_2^2}\,.$$

They showed that in spite of a rather nonconstructive definition, $\mathsf{sumPI}(f)$ is computable in polynomial time from the truth table of f. They also proved that

$$\mathsf{sumPI}(f) \leq \mathsf{D}(S_f) \text{ and } \mathsf{sumPI}(f) \leq n^2.$$

This follows (with a worse constant $\frac{9}{8}$) from the fact that sumPI is a convex rectangle function, as we state in the following proposition: [COMMENT: pH. I did not see the merit of the original proposition. On the other hand, I could not immediately see why sumPI is convex.]

Proposition 6.2. sumPI *is a convex rectangle function and hence* $\mathsf{sumPI}(S) = O(n^2)$.

Proof. By Proposition 4.1, it suffices to show that $\mu_1(R) = \|A_R\|_2^2$ is convex for any fixed matrix $A \neq 0$, and by the definition of spectral norm, that

$$\mu_2(R) = \frac{|u_R^t A_R v_R|^2}{|u_R|_2^2 |v_R|_2^2}$$

is convex for any fixed non-zero vectors u, v. Note that μ_2 can be written as

$$\mu_2(R) = |u_R|_2^2 |v_R|_2^2 \varphi(\frac{|u_R^t A_R v_R|}{|u_R|_2^2 |v_R|_2^2}),$$

with $\varphi(x) = x^2$. The rectangle function $|u_R^t A_R v_R|$ is convex, $|u_R|_2^2 |v_R|_2^2)$ additive, and we can apply Corollary (4.3) $\square$

Laplante, Lee and Szegedy [8] also introduced another rectangle measure $\mathsf{maxPI}(f)$, also based on the spectral norm of matrices, and showed that $(\mathsf{sumPI}(f))^2 \leq (\mathsf{maxPI}(f)^2) \leq \mathsf{D}(S_f)$. [COMMENT: I do not understand this long discussion. If maxPI is defined this way, it is useless for proving lower bounds. Maybe the original definition is more informative? Besides, maxPI is not a rectangle measure, and hence does not fit into our framework. Can the one who wrote this and read the paper do something about it? pH]

Before considering $\mathsf{maxPI}(f)$, we shall explain a tricky way of normalizing rectangle functions. Suppose we want to bound the partition number $\mathsf{D}(R)$ of a given rectangle R from below. Let μ be a subadditive rectangle function (i.e., if $\{M_j\}_{j\in J}$ is a partition of R into monochromatic rectangles, then $\mu(R) \leq \sum_{j\in J} \mu(M_j)$). In order to obtain a rectangle measure, we normalize μ by dividing it by $\max_M \mu(M)$, where M ranges over *all* monochromatic rectangles. If we have a set $\{\mu_i\}_{i\in I}$ of subadditive rectangle functions, then there is a more sophisticated way of normalization. Define a rectangle function μ by

$$\mu(R) := \min_{\{M_j\}_{j\in J}} \max_{i\in I} \frac{\mu_i(R)}{\max_j \mu_i(M_j)},$$

where the minimum is over all partitions $\{M_j\}_{j\in J}$ of R into monochromatic rectangles. That is, this time we normalize the measures $\{\mu_i\}_{i\in I}$ not with respect to *all* monochromatic rectangles, but rather with respect to an apparently smaller set of such rectangles participating in a given partition $\{M_j\}_{j\in J}$ of R. Notice the resulting rectangle function μ need not be normalized or subadditive, yet it bounds the measure $\mathsf{D}(R)$ from below. Indeed, let $\{M_j^*\}_{j\in J^*}$ be a partition of R into the least number of monochromatic rectangles. Then

$$\mu(R) \leq \max_i \frac{\mu_i(R)}{\max_j \mu_i(M_j^*)} \leq \max_i \frac{\sum_j \mu_i(M_j^*)}{\max_j \mu_i(M_j^*)} \leq |J^*| = \mathsf{D}(R).$$

Now, $\mathsf{maxPI}(f)$ can be defined from the spectral norm applying such a normalization to rectangle functions $\mu_A(R) = \|A\|_2$:

$$\mathsf{maxPI}(f) := \min_{\{M_j\}_{j\in J}} \max_{A} \frac{\|A_{S_f}\|_2}{\max_j \|A_{M_j}\|_2}. \tag{20}$$

(It should be noted that originally this function was defined in a completely different way.) Laplante, Lee and Szegedy [8] proved an upper bound $(\mathsf{maxPI}(f))^2 \leq n^2$. We do not know if a quadratic upper bound can be proved using convexity.

7 Open problems

Problem 7.1. *Can we find a continuous function $F : \mathbb{R}^m \to \mathbb{R}$ s.t. F is subadditive and the rectangle measure $\mu(R) = F(s_1(R), \dots, s_m(R))$ can give better than quadratic lower bounds, with $s_i(R)$ additive? (cf. Section 4)*

Problem 7.2. *Is it possible to generalize the quadratic upper bound of Theorem 5.3 to measures of the form*

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

where $\mathsf{s}(R)$ is an additive positive measure?

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

Problem 7.3. *Is it possible to prove superpolynomial lower bounds on monotone formulas using convex measures?*

This is equivalent to the problem of [4] whether the monotone fractional covering number can be superpolynomial.

Problem 7.4. *Prove the cubic lower bound of [2] using formal complexity measures.*

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

References

[1] V. Aho, A. 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, 1983.

[2] J. Håstad. The shrinkage exponent is 2. *SIAM Journal on Computing*, 27:48–64, 1998.

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

[4] M. Karchmer, E. Kushilevitz, and N. Nisan. Fractional covers and communication complexity. *SIAM Journal on Discrete Mathematics*, 8(1):76–92, 1995.

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

[6] M. V. Khrapchenko, A method of obtaining lower bounds for the complexity of π-schemes. *Math. Zamet.*, 10(1):83–92, 1971. (in Russian). English translation in *Math. Notes Acad. Sci. USSR* 10:474–479, 1972.

[7] E. Koutsoupias. Improvements on Khrapchenko's theorem. *Theoret. Comput. Sci.*, 116:399–403, 1993.

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

[9] T. Lee. A new rank technique for formula size lower bounds. In *Proceedings of the 24th Annual Symposium on Theoretical Aspects of Computer Science, Lecture Notes in Computer Science*, volume 4393, pages 145–156, 2007.

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

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

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

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

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