SIAM J. DISC. MATH.
Vol. 8, No. 1, pp. 76–92, February 1995

006

FRACTIONAL COVERS AND COMMUNICATION COMPLEXITY*

MAURICIO KARCHMER†, EYAL KUSHILEVITZ‡, AND NOAM NISAN§

Abstract. It is possible to view communication complexity as the minimum solution of an *integer programming* problem. This integer programming problem is relaxed to a *linear programming* problem and from it information regarding the original communication complexity question is deduced. A particularly appealing avenue this opens is the possibility of proving *lower* bounds on the communication complexity (which is a minimization problem) by exhibiting *upper* bounds on the maximization problem defined by the *dual* of the linear program.

This approach works very neatly in the case of nondeterministic communication complexity. In this case a special case of Lovász's fractional cover measure is obtained. Through it the *amortized* nondeterministic communication complexity is completely characterized. The power of the approach is also illustrated by proving lower and upper bounds on the nondeterministic communication complexity of various functions.

In the case of deterministic complexity the situation is more complicated. Two attempts are discussed and some results using each of them are obtaied. The main result regarding the first attempt is negative: one cannot use this method for proving superpolynomial lower bounds for formula size. The main result regarding the second attempt is a "direct-sum" theorem for two-round communication complexity.

Key words. communication complexity, linear programming bound

AMS subject classifications. 68, 94A15, 94A29, 94A49

1. Introduction.

Many combinatorial optimization problems can be expressed as integer programming problems. Relaxing an integer programming problem to a linear programming problem often gives useful information regarding the original one. In this paper we apply this technique to the study of *communication complexity.*

We consider communication complexity in the wide context of computing relations: we have two players P_1 and P_2, holding n-bit input strings, x and y respectively. They wish to find a value z satisfying a relation $R(x, y, z)$.[1] The goal of the players is to communicate as few bits as possible. This general communication complexity problem contains as special cases the communication complexity of functions, as defined by Yao [Y79] (and studied in numerous works later on), and the relations defined by Karchmer and Wigderson [KW88], which are important because of their close relationship with boolean circuit depth.

It is convenient to count the number of different histories of the protocol. It is well known (see [K89]) that the logarithm of this quantity is equal (up to a constant factor) to the communication complexity. In the case of relations corresponding to circuit depth of boolean functions, this measure gives exactly the formula size. We

* Received by the editors October 13, 1992; accepted for publication (in revised form) December 27, 1993. A preliminary version of this paper appeared as an invited paper in *Proc. 7th IEEE Structure in Complexity Theory*, June 1992, pp. 262–274.

† Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 (mauricio@math.mit.edu). This author's research was supported by National Science Foundation grant NSF-CCR-90-10533.

‡ Department of Computer Science, Technion, Haifa 32000, Israel (eyalk@cs.technion.ac.il). This author's research was done while the author was at Aiken Computation Lab., Harvard University, and was supported by National Science Foundation grant NSF-CCR-90-07677.

§ Department of Computer Science, Hebrew University, Jerusalem 91904, Israel (noam@cs.huji.ac.il). This author's research was supported by the Wolfson Research awards administered by the Israel Academy of Sciences and Humanities and by USA–Israel BSF 89-00126.

[1] We assume that such z always exists. Alternatively, we can say that if there is no such z then any output is legal.

may thus view the communication problem as a covering problem: cover the whole space of possible inputs by possible histories. As an integer programming problem this becomes the following: assign 0-1 weights to the possible histories of a communication protocol such that each possible input is covered with weight 1.

We formalize this integer programming problem and then study the linear programming relaxation of it. Two of the most intriguing features of this approach are

- It allows one to study the *dual* linear programming problem. In particular, one can give *lower bounds* to the original problems by providing *upper bounds* to their dual problems.
- It turns out that the linear programming relaxation often has "direct sum" properties; i.e., the complexity of solving two independent problems simultaneously is exactly equal to the sum of the separate complexities. These results then imply similar results for the original complexity measure.

In this paper we study three different formalizations and relaxations. The first formalization deals with the *nondeterministic case.* It is presented first since it is the most elegant and successful case. Two formalizations for the *deterministic* case are also presented, neither of them without problems.

Our first formalization, for nondeterministic communication complexity, is studied in §2. The main results we obtain in this case are

- The linear relaxation gives exactly Lovász's "fractional cover" measure [L75]. On the other hand it has a natural interpretation in communication complexity terms.
- The linear relaxation is always very close to the "true" nondeterministic communication complexity.
- We get direct-sum results for nondeterministic communication complexity, re-proving and strengthening the recent results of [FKN91]. In particular we show that the linear relaxation completely characterizes the *amortized* nondeterministic communication complexity.
- Various known upper and lower bounds are given new simple proofs using the linear programming relaxations.
- Some connections are shown with the private-coins vs. public-coins question in randomized communication complexity.

Our second formalization, studied in §3, considers the "natural" approach to describing deterministic communication complexity as integer and then linear programs. Our main concern in this section is with communication complexity of relations that correspond to boolean circuit depth and formula size (as in [KW88]). The main results we obtain using this approach are

- We give a surprising new proof of Khrapchenko's quadratic lower bound [K71] on the formula size of the parity function.
- We show that this approach *cannot* give superquadratic lower bounds for the formula size of any boolean function. In particular, the solution of the integer program may be vastly different from the solution of the linear program.
- We give some indication that for *monotone* circuit depth and formula size, this approach may yield exponential lower bounds.

The basic failure of the "natural" approach to deterministic communication complexity led us to consider the third formalization, discussed in §4. This formalization uses a round-by-round approach to communication complexity, and we were only able to obtain results for one-round and two-round protocols. The main results we obtain are

- Direct-sum results for deterministic one-round and two-round communication complexity.
- En route we generalized some of Lovász's results regarding fractional covers to the weighted case, results that may be of independent interest. (Generalizations of some of the results were already known [C79].)

The three different sections of this paper are technically nearly independent of each other. Each section contains an introduction which describes the formalization studied in the section and mentions the basic results obtained.

2. Nondeterministic complexity.

2.1. Introduction. In the nondeterministic model of communication complexity the two players may act nondeterministically, but once they reach an answer, they must be sure of its correctness. It is well known that the nondeterministic communication complexity of a relation R, denoted $C_N(R)$, is simply the logarithm (base 2) of the number of monochromatic rectangles needed to cover the matrix associated with the function.

DEFINITION 2.1. *Given a relation $R \subseteq \{0,1\}^n \times \{0,1\}^n \times Z$ we denote by M_R the matrix representing this relation. That is, each row of M_R corresponds to an input x of P_1, and each column corresponds to an input y of P_2. The entry (x,y) contains the set of all z's that satisfy $R(x,y,z)$. A* rectangle *of M_R is a submatrix of the form $A \times B$ where $A, B \subseteq \{0,1\}^n$. A rectangle $A \times B$ is called* monochromatic *if there exists some element z which is a member of all entries of the rectangle.*

DEFINITION 2.2. *The* nondeterministic cover number *of a relation R, denoted $N(R)$, is the minimum number of monochromatic rectangles that cover M_R, allowing overlaps.*

We associate with every relation R a hypergraph $H_R = (V, E)$ as follows. The vertices of H_R are all possible inputs (i.e., $V = \{0,1\}^n \times \{0,1\}^n$). The hyperedges are all monochromatic rectangles. We can write the nondeterministic cover number as an integer programming problem. Let R be a relation, and let $H_R = (V, E)$ be the corresponding hypergraph. A nondeterministic cover of R can be viewed as a *boolean* function $\phi : E \to \{0,1\}$, such that

$$\forall v \in V \; : \sum_{e \in E \; : \; v \in e} \phi(e) \geq 1.$$

The cover number $N(R)$ is defined as $\min_\phi \sum_{e \in E} \phi(e)$ where ϕ ranges over all nondeterministic covers of H_R.[2] We now define the relaxation of $N(R)$.

DEFINITION 2.3. *A* nondeterministic fractional cover *of H_R is a* real *function $\phi : E \to [0,1]$, such that*

$$\forall v \in V \; : \sum_{e \in E \; : \; v \in e} \phi(e) \geq 1.$$

The fractional cover number *of R, denoted $N^*(R)$, is defined as $\min_\phi \sum_{e \in E} \phi(e)$ where ϕ ranges over all nondeterministic fractional covers of H_R.*

[2] Note that in fact all the definitions of the various cover numbers do not make any use of the special structure of the hypergraphs of the form H_R, and therefore can be generalized to any hypergraph (as in [L75]). For making the exposition more clear we concentrate on hypergraphs of the form H_R. In the technical part of this section we will be interested in the cover numbers of other hypergraphs as well.

As mentioned, this definition is just a special case of Lovász's definition of fractional covers [L75], and therefore we may apply his more general results. In particular, Lovász shows (see Theorem 2.6 below) that the fractional cover number N^* can never be much smaller than the cover number N (clearly, $N^*(R) \leq N(R)$, for every R). Thus, the linear program will give us much information regarding the original nondeterministic communication complexity problem. We use this approach to obtain some very simple proofs of (basically known) *upper* and *lower* bounds to nondeterministic communication complexity.

We now give a simple interpretation for the fractional cover number in the case of communication complexity: a simple way to give a lower bound to $N(R)$ is to give an upper bound to the size of any monochromatic rectangle. This can of course be done relative to any distribution P on $X \times Y$: let $\text{Bound}_P(R) = \max_e Pr_P(e)$, where e ranges over all monochromatic rectangles of R. It is clear that for any distribution P, $1/\text{Bound}_P(R)$ is a lower bound for $N(R)$. It turns out that the best bound one can obtain this way is exactly $N^*(R)$.

LEMMA 2.4. $N^*(R) = \max_P \frac{1}{\text{Bound}_P(R)}$, *where* P *ranges over all probability distributions on* $X \times Y$.

Proof. By the same argument as above, $N^*(R) \geq \frac{1}{\text{Bound}_P(R)}$, for every P. Therefore one direction follows immediately. For the second direction we use the primal-dual theorem for linear programming. The dual of the linear program defining $N^*(R)$ is

$$\max_{\xi} \sum_{x,y} \xi(x,y),$$

where ξ is any real function $\xi : V \to [0,1]$, such that

$$\forall e \in E \quad \sum_{(x,y)\in e} \xi(x,y) \leq 1.$$

The lemma can be verified by associating with every ξ a distribution P_ξ, $P_\xi(x,y) = \xi(x,y)/\sum_{x',y'} \xi(x',y')$. □

The main result we obtain regarding the fractional cover number is that this measure captures completely the cost of solving simultaneously several problems on independent inputs. In particular, we show that N^* is multiplicative with respect to the direct sum.

THEOREM. *Let* $R, R_1, \ldots, R_k$ *be arbitrary relations. Then,*

- $\log N^*(R) \leq C_N(R) \leq \log N^*(R) + O(\log n)$,
- $\prod_i N^*(R_i) = N^*(R_1 \times \cdots \times R_k)$

These results immediately imply the "direct sum" results in [FKN91]. In fact the following corollary gives the underlying reason for these "direct sum" results. Let k be an integer and let R^k denote the "direct sum" of R k times. Namely, to compute R^k we need to compute simultaneously R on k independent inputs. Denote by $\tilde{C}_N(R)$ the *amortized communication complexity* of R [FKN91], i.e., $\limsup_{k\to\infty} C_N(R^k)/k$.

COROLLARY. *Let* R *be any relation. Then,* $\tilde{C}_N(R) = \log N^*(R) \geq C_N(R) - O(\log n)$.

A similar theorem holds for the "one-sided" version of nondeterministic complexity of boolean *functions*, where the players have to be sure about the output only if they output 1. In this case, the problem is to cover the 1's of the function using 1-monochromatic rectangles. (We denote by $C_{NP}(f), \tilde{C}_{NP}(f)$, and $NP^*(f)$ the analogues of $C_N(f), \tilde{C}_N(f)$, and $N^*(f)$, for this case.)

To demonstrate the tightness of our results we exhibit a simple function which has a large gap between $NP(f)$ and $NP^*(f)$, and thus also between $C_{NP}(f)$ and $\tilde{C}_{NP}(f)$. Let $NE(x,y)$ be the nonequality function giving "1" iff the n-bit strings x and y are not equal. We show

$$C_{NP}(NE) = \Theta(\log n) \text{ but } \tilde{C}_{NP}(NE) = \Theta(1).$$

It is interesting to note that for this function the complexity difference between $C_{NP}(f)$ and $\tilde{C}_{NP}(f)$ mirrors the complexity difference between the "private-coins" and "public-coins" variants of *randomized* complexity. We explain this phenomena (that randomization in the public-coins model is more powerful than nondeterminism) and prove that while the (one-sided error) randomized complexity in the public-coins model can be smaller than $C_{NP}(f)$, it is always at least $\tilde{C}_{NP}(f) = \log NP^*(f)$.

2.2. Direct sums. We start by defining the product of two hypergraphs.

DEFINITION 2.5. *Given two hypergraphs H_1 and H_2 we define their* product $H_1 \times H_2$ *by* $V(H_1 \times H_2) = V(H_1) \times V(H_2)$ *and* $E(H_1 \times H_2) = \{e_1 \times e_2 | e_1 \in E(H_1), e_2 \in E(H_2)\}$.

The following result of Lovász is crucial.

THEOREM 2.6 ([L75]). *Let H, H_1, and H_2 be any hypergraphs then*

1. $N^*(H) \geq \frac{N(H)}{\ln |V(H)|}$;
2. $N^*(H_1 \times H_2) = N^*(H_1) \cdot N^*(H_2)$.

The first statement directly yields the following corollary.

COROLLARY 2.7. *Let R be any relation. Then,* $\log N^*(R) \leq C_N(R) \leq \log N^*(R) + \log n + O(1)$.

DEFINITION 2.8. *Given two relations R and S, their* direct sum, *denoted $R \times S$, is the problem of solving both R and S simultaneously on independent inputs.*

Note that usually, for two relations R and S, the hypergraph $H_{R \times S}$ is not the same as the hypergraph obtained by the product $H_R \times H_S$. However, the following lemma claims that both have the same nondeterministic fractional cover number.

LEMMA 2.9. *Let R and S be two relations. Then $N^*(H_{R \times S}) = N^*(H_R \times H_S)$.*

Proof. For proving that $N^*(H_{R \times S}) \leq N^*(H_R \times H_S)$, note that if e_R is a monochromatic rectangle of M_R and e_S is a monochromatic rectangle of M_S, then $e_R \times e_S$ is a monochromatic rectangle of $M_{R \times S}$. Therefore, $E(H_R \times H_S) \subseteq E(H_{R \times S})$. This implies that every nondeterministic fractional cover ϕ defined for $H_R \times H_S$ can be extended with zeroes to a nondeterministic fractional cover of $H_{R \times S}$, and thus the inequality follows.

On the other hand, given the optimal nondeterministic fractional cover ϕ defined on $E(H_{R \times S})$, we can take every hyperedge $e = X \times Y \subseteq V(H_R) \times V(H_S)$ with $\phi(e) > 0$, and define X_R, X_S, Y_R, and Y_S to be the projections of X and Y on the first and second coordinates respectively (i.e., the projections on $V(H_R)$ and $V(H_S)$ respectively). Now, define $e_R = X_R \times Y_R$ and $e_S = X_S \times Y_S$. These are monochromatic rectangles of M_R and M_S (respectively) and thus $e_R \times e_S$ is a hyperedge of $H_R \times H_S$. Define, for every e, $\phi'(e_R \times e_S) = \phi(e)$ (if more than one hyperedge correspond to the same e_R, e_S then $\phi'(e_R \times e_S)$ is the sum of $\phi(e)$ for all those e's). We get that ϕ' is a nondeterministic fractional cover of $H_R \times H_S$ (since the monochromatic rectangle $e_R \times e_S$ contains the monochromatic rectangle e) and therefore $N^*(H_R \times H_S) \leq N^*(H_{R \times S})$. □

We can now get the following set of "direct sum" results.

THEOREM 2.10. *Let $R, R_1, \ldots, R_k$ be the arbitrary relations:*

- $N^*(\times_{i=1}^k R_i) = \prod_{i=1}^k N^*(R_i)$;
- $\sum_{i=1}^k \log N^*(R_i) \le C_N(\times_{i=1}^k R_i) \le \sum_{i=1}^k \log N^*(R_i) + \log kn + O(1)$;
- $\tilde{C}_N(R) = \log N^*(R)$.

Proof. The lower bounds on C_N follow from Theorem 2.6 and Lemma 2.9. The upper bounds follow from Theorem 2.6 and Corollary 2.7. The bound for $\tilde{C}_N$ is obtained by taking k copies of R and letting k approach infinity. □

By using Corollary 2.7 we can eliminate N^* from the statement, getting as a corollary the somewhat weaker results of [FKN91].

COROLLARY 2.11. *Let $R, R_1, \ldots, R_k$ be the arbitrary relations:*

- $\sum_{i=1}^k C_N(R_i) - k \log n - O(1) \le C_N(\times_{i=1}^k R_i) \le \sum_{i=1}^k C_N(R_i)$
- $C_N(R) - \log n - O(1) \le \tilde{C}_N(R) \le C_N(R)$

2.3. One-sided nondeterministic complexity. In the case that boolean functions are computed, one is frequently only interested in the "NP"-version of nondeterministic complexity, i.e., where the players need only be sure of the answer in the case where $f(x, y) = 1$. We denote this complexity by $C_{NP}(f)$. It is not difficult to see that the corresponding covering problem is simply to cover all the 1-inputs of f by 1-monochromatic rectangles.

It is straightforward to carry over all of our results to this case as well, where in the direct sum of f and g, we need only cover the joint 1's of f and g, i.e., cover the 1's of $f \wedge g$. In particular we get the following corollaries.

THEOREM 2.12. *Let $f_1, \ldots, f_k$ be any k functions. Then*

$$\sum_{i=1}^k C_{NP}(f_i) - k \log n - O(1) \le C_{NP}(f_1 \wedge \cdots \wedge f_k) \le \sum_{i=1}^k C_{NP}(f_i).$$

The following example shows how the above results can be used for proving lower bounds on the nondeterministic communication complexity.

Example 1. Let the "disjointness" function be defined as follows: $\mathrm{DISJ}_n(x, y)$ is defined for every $x, y \in \{0,1\}^n$ as 1 if there is no index i such that $x_i = y_i = 1$ (and 0 otherwise). Clearly,

$$\mathrm{DISJ}_n(x, y) = \bigwedge_{i=1}^n \mathrm{DISJ}_1(x_i, y_i).$$

Therefore, $NP(\mathrm{DISJ}_n) \ge NP^*(\mathrm{DISJ}_n) = (NP^*(\mathrm{DISJ}_1))^n = 2^n$, where the last equality follows by noting that $NP^*(\mathrm{DISJ}_1) = 2$. Thus we have

$$C_N(\mathrm{DISJ}_n) = C_{NP}(\mathrm{DISJ}_n) = n.$$

2.4. Fractional covers and randomized complexity. The following theorem relates $NP^*(f)$ to $C_{R-\mathrm{pub}}(f)$—the communication complexity of computing f by a *probabilistic one-side error protocol* (i.e., a protocol that might err only if $f(x, y) = 1$ with probability smaller than, say, $\frac{1}{2}$) in the *public coins* model.[3] It is known that $C_{R-\mathrm{pub}}(f)$ is smaller than $C_{R-\mathrm{priv}}(f)$ (one-sided error protocols in the *private-coins* model) by at most an additive factor of $\log n$. Clearly, $C_{NP}(f) \le C_{R-\mathrm{priv}}(f)$.[4] We

[3] In the public coins model, instead of flipping coins locally, the two parties share a string of random coins. For a formal definition of the model and some results on the relations between the public coins and private coins models, see [N91].

[4] The parties "guess" good random coins and run the randomized protocol.

already proved that $\log NP^*(f)$ is smaller than $C_{NP}(f)$ by at most an additive term of $\log n$. To complete the picture we give the following theorem.

THEOREM 2.13. *Let f be a function. Then* $\log NP^*(f) \leq C_{R-\text{pub}}(f) + 1$.

Proof. Given F, a probabilistic one-sided error protocol for f in the public-coins model, we will construct a fractional cover for the 1's of M_f, as needed. Let r be a possible (public) random string and let $p(r)$ be its probability. Fixing r, then F is just a deterministic protocol and therefore induces a cover of the 1's of M_f by at most $2^{C_{R-\text{pub}}(f)}$ monochromatic rectangles. We add to the cover all the rectangles in which the output is "1." As the protocol has only one-sided errors then these rectangles cover only "1"-entries. With each such rectangle e we associate a value $\phi(e) = 2p(r)$. We repeat this process for every possible random string r. We claim that the obtained cover is what we aim for. First, note that for every (x, y) such that $f(x, y) = 1$ we have

$$\sum_{e:\ (x,y)\in e} \phi(e) = 2 \cdot \text{Prob}(F \text{ outputs 1 on input } (x, y)) \geq 2 \cdot \frac{1}{2} = 1.$$

Finally, note that the cover we construct satisfies

$$\sum_e \phi(e) \leq \sum_r 2p(r) \cdot 2^{C_{R-\text{pub}}(f)} = 2 \cdot 2^{C_{R-\text{pub}}(f)}. \qquad \square$$

We now show that the above theorem can be used to estimate $NP^*(f)$.

Example 2. Let the "nonequality" function be defined as follows: $NE_n(x, y)$ is defined for every $x, y \in \{0,1\}^n$ as 1 if $x \neq y$ and 0 otherwise. It is known that $C_D(NE_n) = n$, and that $C_N(NE_n) = \Theta(\log n)$.[5] On the other hand, $C_{R-\text{pub}}(NE_n) = O(1)$.[6] By Theorem 2.13, we get that $NP^*(NE_n) = O(1)$. (Note that for this function $C_{R-\text{pub}}$ is less than NP.)

In the following example we show how to use these techniques to derive nontrivial *upper* bounds on the nondeterministic communication complexity. Interestingly, this is done without describing explicitly protocols that compute the functions.

Example 3. Let n be a perfect square. Let f_n be the following function: view each input string as $\sqrt{n}$ substrings of length $\sqrt{n}$ (i.e., $x = \overline{x}_1\overline{x}_2 \ldots \overline{x}_{\sqrt{n}}$ and $y = \overline{y}_1\overline{y}_2 \ldots \overline{y}_{\sqrt{n}}$, where $\overline{x}_i, \overline{y}_i \in \{0,1\}^{\sqrt{n}}$, for every i). Let f_n be defined as follows: $f_n(x, y)$ is 1 if there exists an i such that $\overline{x}_i = \overline{y}_i$. This function was studied in [MS82], [F87], and a (tight) $O(\sqrt{n})$ upper bound was proved for its nondeterministic communication complexity, using a complex protocol. Here we give a very simple proof for this upper bound. Clearly, $C_{NP}(f_n) = O(\sqrt{n})$.[7] Therefore to prove that $C_N(f_n) = O(\sqrt{n})$ it is enough to prove that $C_{NP}(\overline{f}_n) = O(\sqrt{n})$.[8] For this, we write $\overline{f}_n(x) = \bigwedge_{i=1}^{\sqrt{n}} NE_{\sqrt{n}}(\overline{x}_i, \overline{y}_i)$. Therefore, $NP^*(\overline{f}_n) = (NP^*(NE_{\sqrt{n}}))^{\sqrt{n}}$ which equals by the previous example to $(O(1))^{\sqrt{n}}$. This implies that $C_{NP}(\overline{f}_n) = O(\sqrt{n})$.

[5] P_1 "guesses" an index i and sends the index i together with x_i to P_2.

[6] The parties can view the public random string as a n-bit vector b and exchange the inner product of b with x and y.

[7] P_1 "guesses" i and sends i and $\overline{x}_i$ to P_2 who checks whether $\overline{x}_i = \overline{y}_i$.

[8] The trivial upper bound for $C_{NP}(\overline{f}_n)$ is $O(\sqrt{n}\log n)$: P_1 "guesses" for every $1 \leq i \leq \sqrt{n}$ an index in which $\overline{x}_i$ differs from $\overline{y}_i$. It sends to P_2 all those indices with their values.

3. Deterministic complexity: disjoint cover.

3.1. Introduction. As shown, our approach works well for *nondeterministic* complexity. In the case of computing *functions* we do get some nontrivial information regarding *deterministic* complexity, as [AUY83] showed that there can be at most a quadratic gap between deterministic and nondeterministic complexity. In the case of *relations* we may get no information at all, as the gap between deterministic and nondeterministic complexity can be exponential. However, we will show that the suggested approach leads to some results.

A natural approach to present deterministic communication complexity as a covering problem is simply to forbid overlap of any two rectangles in the monochromatic cover.

DEFINITION 3.1. *The* deterministic cover number *of a relation R, denoted $D(R)$, is the minimum number of monochromatic rectangles in a* disjoint (*nonoverlapping*) *cover of the set of inputs.*

As opposed to the nondeterministic case where the nondeterministic complexity, $C_N(R)$, was always $\Theta(\log N(R))$, it is still an open problem whether the deterministic complexity, $C_D(R)$, is always $\Theta(\log D(R))$. However, it is still true that $\log D(R) \leq C_D(R)$. Furthermore, it is implicit in [AUY83] that $\log D(R) \geq \sqrt{C_D(R)}$, and thus these two measures are quite close. Let us also mention that if R_g is a relation associated with the circuit depth of a boolean function g (à la [KW88]) then $D(R_g)$ yields a lower bound to the formula size complexity.

Therefore it is important to understand the measure $D(R)$. This measure has the advantage of being more combinatorial than $C_D(R)$. As previously, we can express $D(R)$ as an integer program. For a relation R, let $H_R = (V, E)$ be the corresponding hypergraph. A *deterministic cover* of H_R is a *boolean* function $\phi : E \to \{0, 1\}$, such that

$$\forall v \in V \ : \sum_{e \in E \ : \ v \in e} \phi(e) = 1.$$

The *deterministic cover number* of R, denoted $D(R)$, is $\min_\phi \sum_{e \in E} \phi(e)$ where ϕ is a deterministic cover. Again, we can relax the integrality condition. Thus, we get the following definition.

DEFINITION 3.2. *A deterministic* fractional *cover of the hypergraph H_R is a* real *function $\phi : E \to [0, 1]$, such that*

$$\forall v \in V \ : \sum_{e \in E \ : \ v \in e} \phi(e) = 1.$$

$D^*(R)$ is defined as $\min_\phi \sum_{e \in E} \phi(e)$ where ϕ is a deterministic fractional cover.

Our goal is to prove lower bounds on $D(R)$ by proving lower bounds for $D^*(R)$. For this, we look at the dual linear program which is defined as follows. Let R be a relation, and let H_R be the corresponding hypergraph. Then, by the *duality* theorem,

$$D^*(R) = \max_w \sum_{v \in V} w(x, y),$$

where w ranges over all real functions that satisfy

$$\forall e \in E \ \sum_{v \in E} w(x, y) \leq 1.$$

It is important to notice that, as opposed to the nondeterministic case, there may be a huge gap between $D(R)$ and $D^*(R)$ (we will present an example below). Still, lower bounds for $D^*(R)$ do give lower bounds to $D(R)$. Our first result does exactly that, giving a new proof to Khrapchenko's quadratic lower bound for the formula size of the parity function, by proving a lower bound for $D^*(R_{\oplus_n})$, where $R_{\oplus_n}$ is the relation associated (à la [KW88]) with parity.[9] The new proof is achieved by exhibiting an *upper* bound to the dual problem.

THEOREM. *Let $R_{\oplus_n}$ be the relation associated with the parity function (as above) then $D^*(R_{\oplus_n}) = \Theta(n^2)$.*

Our major result regarding this approach is negative though. We show that this method cannot prove superquadratic lower bounds to the formula size of *any* boolean function.

THEOREM. *Let f be any boolean function and let R_f be the relation associated with it. Then $D^*(R_f) = O(n^2)$.*

Note that for most boolean functions f, $D(R_f) = 2^{\Theta(n)}$. This result specifically suggests that anyone aiming to prove lower bounds for the circuit depth of boolean functions should abandon this approach. However, we do give some indication that proving lower bounds to *monotone* circuit depth might be possible using this approach. This gives another example of the big difference between monotone and nonmonotone computation.

3.2. Khrapchenko's lower bound. Khrapchenko [K71] gives the only known general lower bound for search problems. Let $R \subseteq X \times Y \times Z$ be any relation, and let M be the corresponding matrix. Let $A \subseteq X \times Y$ be any set with the following properties:

1. $\forall (x,y) \in A$, $|M_{x,y}| = 1$.
2. $\forall x \in X$ and $z \in Z$ there is at most one $y \in Y$ such that $(x,y) \in A$ and $M_{x,y} = \{z\}$.
3. $\forall y \in Y$ and $z \in Z$ there is at most one $x \in X$ such that $(x,y) \in A$ and $M_{x,y} = \{z\}$.

THEOREM 3.3 ([K71]). $D(R) \geq \frac{|A|^2}{|X| \cdot |Y|}$.

As an example of an application of Theorem 3.3 consider the matrix of $R_{\oplus_n}$ indexed by $\{x : \oplus_i x_i = 1\} \times \{y : \oplus_i y_i = 0\}$ and whose (x,y) entry is $\{i : x_i \neq y_i\}$.

COROLLARY 3.4. $D(R_{\oplus_n}) \geq n^2$.

Proof. Let $A = \{(x,y) \in X \times Y \text{ such that } d(x,y) = 1\}$.[10] It is easy to see that A has the required properties and provides the desired lower bound. □

We prove a slight strengthening of this result.

THEOREM 3.5. $D^*(R) \geq \frac{|A|^2}{|X| \cdot |Y|}$.

COROLLARY 3.6. $D^*(R_{\oplus_n}) \geq n^2$, *with equality for $n = 2^k$.*

We give here the proof of the corollary. The same ideas with some technical algebraic calculations can be used to prove Theorem 3.5 in its full generality. In §3.2.1, we give some general heuristics that can help in such proofs. In §3.2.2, we use these heuristics for proving the corollary.

3.2.1. Heuristics for proving an upper bound for the dual. To prove a lower bound for $D^*(R)$, we only have to *exhibit* a solution to the dual program. For

[9] The relation associated with a function f, denoted R_f, consists of all triples (x,y,i) such that $f(x) = 1$, $f(y) = 0$, and $x_i \neq y_i$.

[10] $d(x,y) = |\{i : x_i \neq y_i\}|$.

this, we have at our disposal the powerful paradigm of *trial and error*. The following heuristics can be quite helpful in our quest for a solution for the dual problem. After presenting these heuristics we will use them for the proof of Corollary 3.6.

- Let $\Pi(H_R)$ be the automorphism group of H_R. A solution $\vec{w}$ for the dual problem is *invariant* under $\Pi(H_R)$ if $w_{(x,y)} = w_{\pi(x,y)}$ for every (x,y) and $\pi \in \Pi(H_R)$. Similarly, we can define invariant solutions for D^*. A symmetrization argument can be used to show that, without loss of generality, the optimal solutions to both D^* and its dual are invariant under $\Pi(H_R)$. This clearly reduces the size of both linear programs.
- Intuitively, it is worthwhile to give (x,y) a positive weight if it does not appear in many monochromatic rectangles. This is because such a positive weight does not affect many rectangles. Conversely, if (x,y) appears in lots of monochromatic rectangles then we could benefit by making $w_{(x,y)}$ negative, thus helping many rectangles without lowering by much the value of D^*.
- Having decided which pairs (x,y) will get positive weights, we could test this decision by asking whether the following modified version of M has the same D^*:

$$\tilde{M}_{x,y} = \begin{cases} M_{x,y} & \text{if } w_{(x,y)} \text{ is positive,} \\ Z & \text{otherwise,} \end{cases}$$

  where Z is the set of all possible solutions. In a sense, this means that we have to assign positive weights to the hardest pairs. Conversely, if we have a solution for the dual problem, we will get information about the *core* of the problem.
- Given an optimal solution to D^*, the theory of linear programming tells us which of the inequalities of the dual have to be saturated. In particular, if for a given monochromatic rectangle e, $\phi(e)$ is positive then the corresponding inequality in the dual has to be saturated. That is, the sum of the weights of the entries in e have to add up to one. Given that we suspect that a given solution to D^* is optimal, we can use this information to try to construct a suitable solution for the dual.

3.2.2. Proof of Corollary 3.6. We will assume that $n = 2^k$. For general n the corollary follows from the theorem. The upper bound follows from the protocol attaining $\Gamma(R_{\oplus_n}) \leq n^2$, where Γ denotes the number of different histories in the protocol. We describe it here: let $I = \{1, \ldots, n/2\}$. The players start by exchanging the parities of their vectors on I. That is, $\oplus_{i \in I} x_i$ and $\oplus_{i \in I} y_i$. The players then continue recursively in either I or $[n] \setminus I$ depending on whether $\oplus_{i \in I} x_i \neq \oplus_{i \in I} y_i$ or not. It is easy to see that this is a correct protocol with n^2 different histories.

The lower bound will follow by providing a specific function w for the dual problem which add up to n^2. At this point we could provide w and finish in two more lines. Instead, we will reason using our heuristics and derive the desired solution. We therefore can get away with some informality.

First, we start with a belief that the upper bound just described is optimal. If so, we know that the inequalities associated with the chosen rectangles have to be saturated. We therefore have to understand better our upper bound. Let $A = \{(x,y) \in X \times Y \text{ such that } d(x,y) = 1\}$. A closer look at the protocol reveals that each of its histories is followed by the same number of entries from A. Furthermore, each history defines a square rectangle with exactly one entry from A in each row and column. Note that the set of histories partition $M_{R_{\oplus_n}}$.

Following our first heuristic, $w(x,y)$ will depend only on $d(x,y)$. Following our second heuristic, it is worthwhile to give entries from A a positive weight. Let us try

$$w(x,y) = \begin{cases} a & \text{if } (x,y) \in A, \\ -b & \text{otherwise,} \end{cases}$$

for some a and b. We will finish the proof if we find a and b which respect all inequalities and saturate those associated with histories from our upper bound. This is because we have n^2 saturated rectangles which partition the whole matrix.

Let us look now at the monochromatic rectangles. In each one there is at most one entry from A in every row or column. Therefore, the heaviest rectangles are the square ones with exactly one entry from A in each row and column. For a $k \times k$ such square we have the inequality

$$ka - k(k-1)b \leq 1$$

with equality when $k = |A|/n^2 = N$ (the size of the rectangles in the *optimal* solution). Writing the above inequalities as $-bk^2 + (a+b)k - 1 \leq 0$ we have one root of the left-hand side, namely $k = N$. Noticing that the inequality is only restricted to integral k, we can let the second root be $N-1$ and solve for a and b. This finishes our proof. For the skeptic, we provide the final values $a = 2/N$ and $b = 1/N(N-1)$. □

3.3. The linear programming bound and boolean relations. Let $f : \{0,1\}^n \mapsto \{0,1\}$ be a boolean function and let R_f be indexed by $f^{-1}(1) \times f^{-1}(0)$, and for $(x,y) \in f^{-1}(1) \times f^{-1}(0)$ let the corresponding entry be $\{i : x_i \neq y_i\}$. We call relations of the form R_f *boolean relations*. For example, $R_{\oplus_n}$ is a boolean relation. The relevance of this definition comes from the following theorem.

THEOREM 3.7 ([KW88]). *For every f, $d(f) = C(R_f)$ and $L(f) = \Gamma(R_f)$.*

Here, $d(f)$ and $L(f)$ are the *depth* and *formula size* of f respectively. For definitions of circuits and related material concerning the above theorem see [BS90], [K89].

Let U_n be the relation indexed by $\{0,1\}^n \times \{0,1\}^n$ and for $(x,y) \in \{0,1\}^n \times \{0,1\}^n$, $x \neq y$, let the corresponding entry be $\{i : x_i \neq y_i\}$. If $x = y$, the corresponding entry remains undefined. The following claim is trivial and explains why we call U_n the *universal relation* [K89].

CLAIM 3.8. *For every $f : \{0,1\}^n \mapsto \{0,1\}$, $D(R_f) \leq D(U_n)$ and $D^*(R_f) \leq D^*(U_n)$.*

We will show that the best lower bound for boolean relations attainable via the linear programming bound is very weak by proving the following theorem.

THEOREM 3.9. *$D^*(U_n) = O(n^2)$.*

Proof. We will use some of the ideas behind the logarithmic randomized protocol for U_n [K89]. For $S \subseteq [n]$, let $A_S = \{x : \oplus_{i \in S} x_i = 1\} \times \{y : \oplus_{i \in S} y_i = 0\}$ and $B_S = \{x : \oplus_{i \in S} x_i = 0\} \times \{y : \oplus_{i \in S} y_i = 1\}$. The upper bound of Corollary 3.6 implies that both $D(A_S)$ and $D(B_S)$ are at most n^2. Let P_S be an optimal partition of A_S and B_S into monochromatic rectangles.

It is easy to see that for every $x \neq y$, $(x,y) \in A_S \cup B_S$ for exactly half the subsets S. We will give a weight of $2^{-(n-1)}$ to every rectangle from $\cup_{S \subseteq [n]} P_S$. Each pair (x,y) with $x \neq y$ is covered by 2^{n-1} such rectangles with total unit weight. Also, the total weight is $2^n \cdot (2n^2) \cdot 2^{-(n-1)} = 4n^2$. □

3.4. The linear programming bound and monotone boolean relations. Let $f : \{0,1\}^n \mapsto \{0,1\}$ be a *monotone* boolean function. We denote by $\min(f)$ and $\max(f)$ the set of minterms and maxterms of f respectively (see [K89] for definitions). Note that each minterm and each maxterm, as a set of variables, intersect. Let R_f^m be a relation indexed by $\min(f) \times \max(f)$ and for $(p,q) \in \min(f) \times \max(f)$ let the corresponding entry be $p \cap q$. We call matrices of the form R_f^m *monotone boolean relations.* The following theorem is the monotone analogue of Theorem 3.7.

THEOREM 3.10 ([KW88]). *For every monotone function f, $d_m(f) = C(R_f^m)$ and $L_m(f) = \Gamma(R_f^m)$.*

Here, d_m and L_m denote monotone depth and monotone formula size (see [K89] for definitions).

As in the preceding section, we define a *monotone universal relation.* Let U_n^m be a relation indexed by $\mathcal{P}([n]) \times \mathcal{P}([n])$ and for $p,q \in \mathcal{P}([n])$ with $p \cap q \neq \emptyset$ let the corresponding entry be $p \cap q$. If $p \cap q = \emptyset$ the entry of U_n^m remains undefined. We have the following claim.

CLAIM 3.11. *For every monotone function $f : \{0,1\}^n \mapsto \{0,1\}$ we have $D(R_f^m) \leq D(U_n^m)$ and $D^*(R_f^m) \leq D^*(U_n^m)$.*

The main reason to define universal relations is to try on them new ideas to prove lower bounds. The fact that $D^*(U_n) = O(n^2)$ means that the best lower bound for boolean relations attainable with the linear programming bound is at most quadratic. The following theorem gives evidence to the fact that the linear programming bound may give exponential lower bounds for monotone boolean relations.

THEOREM 3.12. *$D^*(U_n^m) \geq d^n$ for some constant $d > 1$.*

Proof. We will construct a feasible solution to the dual problem whose value is the desired bound. Following our first heuristic, $w(p,q)$ depends only on $|p \cap q|$. It is natural to try the Hadamard matrix $((-1)^{|p\cap q|})_{p,q\in\mathcal{P}([n])}$ as the sign pattern for our weights. We want to give a positive weight to those entries (p,q) with $|p \cap q| = 1$ and we must give zero weights to the undefined entries. Let us try

$$w(p,q) = \begin{cases} 0 & \text{if } p \cap q = \emptyset, \\ -(-1)^{|p\cap q|}c & \text{otherwise,} \end{cases}$$

for some constant c to be specified later. Using the fact that

$$\sum_{p,q\in\mathcal{P}([n])} (-1)^{|p\cap q|} = 2^n$$

we get

$$\sum_{p,q\in\mathcal{P}([n])} w(p,q) = c \cdot |\{(p,q) : p \cap q = \emptyset\}| = c \cdot (3^n - 2^n).$$

We now look at the monochromatic rectangles. Let $R_i = \{p : i \in p\} \times \{q : i \in q\}$. Every monochromatic rectangle is a subrectangle of R_i for some i. The pattern of signs of weights of entries from R_i constitute an $2^{n-1} \times 2^{n-1}$ Hadamard matrix. We will use the following lemma of Lindsey [ES74, p. 88] which says that minors of a Hadamard matrix are *balanced.*

LEMMA 3.13. *Let H be an $N \times N$ Hadamard matrix and let S and T be subsets of rows and columns respectively. Then,*

$$\sum_{i\in S, j\in T} H_{i,j} \leq \sqrt{N \cdot |S| \cdot |T|}\,.$$

In our case, we use the lemma to show that for every subrectangle R of R_i

$$\sum_{(p,q)\in R} w_{(p,q)} \leq c \cdot \sqrt{2^{n-1}2^{n-1}2^{n-1}} \leq c \cdot \sqrt{2^{3n}} \leq c \cdot (\sqrt{8})^n$$

which is less than 1 if $c = (\sqrt{8})^{-n}$. We have therefore found a feasible solution whose value is $(\sqrt{8})^{-n}(3^n - 2^n) \geq d^n$ for any $d < 3/\sqrt{8}$. □

4. Deterministic communication: two rounds.

4.1. Introduction. The previous approach tried to look at the protocol globally, and failed. Our next approach deals with the protocol in a round-by-round fashion. We will associate a covering problem with every round of the protocol. Unfortunately, we are not able to carry our results to protocols having an arbitrary number of rounds, but only succeed for one-round and two-round protocols.

To best explain our approach let us first limit ourselves to one-round protocols.[11] Intuitively, in a one-round protocol, P_2 partitions the columns of the matrix in a way that enables P_1 to decide on the answer. Formally, we associate with any relation R the following covering problem. Let X denote the space of all possible inputs to P_1, and Y the space of inputs to P_2. A set $A \subseteq Y$ is called *compatible* if for every $x \in X$ there exists an answer z that is a legal answer for all $y \in A$ (i.e., such that $R(x, y, z)$ holds for all $y \in A$). $D_1(R)$ is defined to be the minimum number of compatible sets that are needed in order to cover Y.

It is not difficult to see that $\log D_1(R)$ gives the *one-round communication complexity* of R (denoted by $C_{D_1}(R)$). It is also not difficult to see that in this case the disjoint and nondisjoint covers are the same, and thus when we relax the integer problem to a linear one, giving $D_1{}^*(R)$, we will be able to use Lovász's results regarding fractional covers. At this point we will already be able to reprove the "direct sum" results for one-round deterministic complexity obtained in [FKN91], specifically $C_{D_1} \geq \tilde{C}_{D_1} \geq C_{D_1} - O(\log n)$. In [KRW91] it was conjectured that for every R, $\tilde{C}_D(R) \geq C_D(R) - O(\log n)$. In [FKN91] it was proved that $\tilde{C}_D(R) \geq \sqrt{C_D(R)} - O(\log n)$. Here we show that the conjecture is true for two-round protocols.

We look at two-round protocols[12] in the following way: in the first round, P_1 partitions the rows of the matrix, and then the parties continue with a one-round protocol on the subdomain. This can be expressed as the following *weighted* covering problem. Our aim is to cover X, where we are allowed to use any subset of X in the cover, and the cost of using a subset $A \subseteq X$ is the one-round complexity of solving R given that $x \in A$, denoted $D_1(A)$.

DEFINITION 4.1. *A* cover *of X is a* boolean *function $\phi : P(X) \to \{0, 1\}$, such that*

$$\forall x \in X \ : \sum_{A \in P(X) \ : \ x \in A} \phi(A) \geq 1.$$

[11] P_2 sends to P_1 a single message, and P_1 then needs to compute the answer.

[12] P_1 sends a message to P_2, who sends another message to P_1, who computes the answer.

The weighted cover number *of R, denoted $D_2(R)$, is defined as*

$$D_2(R) = \min_{\phi} \sum_{A \in P(Y)} \phi(A) D_1(A),$$

where ϕ is a cover.

It is not difficult to see that $\log D_2(R)$ is equal (up to a constant factor) to the two-round deterministic complexity, $C_{D_2}(R)$. Again, we relax the integrality conditions and look at the resulting linear program giving $D_2{}^*$. We can now no longer use Lovász's results, as we have a "weighted" covering problem. This problem was already considered by Chvátal [C79] who extended the first part of Theorem 2.6 to the "weighted" case. We prove that the second part of Theorem 2.6 can be generalized as well. We believe that these generalizations are of independent interest. Using these generalization we can prove a direct-sum result for two-round communication complexity. In particular, let $\tilde{C}_{D_2}(R)$ denote the amortized two-round communication complexity of R.

THEOREM 4.2. *For every two relations R and S,*

- $D_2(R)D_2(S)/\text{poly}(n) \leq D_2(R \times S) \leq D_2(R)D_2(S)$;
- $\tilde{C}_{D_2}(R) = \Theta(C_{D_2}(R)) - O(\log n)$.

4.2. Weighted fractional covers. In this subsection we present the new notion of *weighted fractional covers.* This notion will be later used in the proof of Theorem 4.2.

DEFINITION 4.3. *Let H be a hypergraph and let w be a* weight function *defined on $E(H)$ such that $w(e) \geq 1$, for every hyperedge e. Given a deterministic/nondeterministic integral/fractional cover ϕ the weight function w gives it a weight $w(\phi) = \sum_{e \in E(H)} w(e)\phi(e)$. The* weighted cover numbers *$D(H,w), D^*(H,w), N(H,w)$, and $N^*(H,w)$ are defined as the minimum of $w(\phi)$ over all appropriate covers ϕ. (Note that the original definitions, as presented in §2.1, are special cases of the new definitions with $w \equiv 1$.)*

The next theorem is an extension of Theorem 2.6.

THEOREM 4.4. *Let H, H_1, and H_2 be any hypergraphs, and let w, w_1, and w_2 be weight functions on $E(H)$, $E(H_1)$, and $E(H_2)$ (respectively) that give weights ≥ 1 for every hyperedge (i.e., $w(e) \geq 1$, for all $e \in E(H)$). Then*

1. $N^*(H,w) \geq \frac{N(H,w)}{\ln |V(H)|}$.
2. $N^*(H_1 \times H_2, w_1 \times w_2) = N^*(H_1,w_1) \cdot N^*(H_2,w_2)$, *where $w_1 \times w_2$ is defined as $w_1 \times w_2(e_1 \times e_2) = w_1(e_1) \cdot w_2(e_2)$.*

Proof. Part (1) of the theorem was proved in [C79].[13] To prove (2) we first prove that $N^*(H_1 \times H_2, w_1 \times w_2) \leq N^*(H_1,w_1) \cdot N^*(H_2,w_2)$. Let ϕ_1 and ϕ_2 be the optimal-weight fractional covers for H_1 and H_2 (i.e., those that give the minimum for $N^*(H_1,w_1)$ and $N^*(H_2,w_2)$ respectively). Define $\phi(e_1 \times e_2) = \phi_1(e_1) \cdot \phi_2(e_2)$. We show that ϕ is a nondeterministic fractional cover of $H_1 \times H_2$ and that its weight is the multiplication of the weights of ϕ_1 and ϕ_2. Clearly, ϕ is a function from $E \triangleq E(H_1 \times H_2)$ to $[0,1]$. In addition, every vertex $(v_1,v_2) \in V(H_1 \times H_2)$ is covered as needed:

$$\sum_{e_1 \times e_2 \in E \,:\, (v_1,v_2) \in e_1 \times e_2} \phi(e_1 \times e_2) = \sum_{e_1 \in E_1 \,:\, v_1 \in e_1} \sum_{e_2 \in E_2 \,:\, v_2 \in e_2} \phi(e_1 \times e_2)$$

[13] In fact, the result stated in [C79] is somewhat different than the one stated here. However, the proof in [C79] immediately implies the result stated here.

$$= \sum_{e_1 \in E_1 \,:\, v_1 \in e_1} \sum_{e_2 \in E_2 \,:\, v_2 \in e_2} \phi_1(e_1)\phi_2(e_2)$$
$$= \sum_{e_1 \in E_1 \,:\, v_1 \in e_1} \phi_1(e_1) \sum_{e_2 \in E_2 \,:\, v_2 \in e_2} \phi_2(e_2)$$
$$\geq 1 \cdot 1 = 1.$$

Thus, ϕ is a legal cover and we get

$$\begin{aligned} N^*(H_1 \times H_2, w_1 \times w_2) &\leq \sum_{e_1 \times e_2 \in E} w_1 \times w_2(e_1 \times e_2) \cdot \phi(e_1 \times e_2) \\ &= \sum_{e_1 \times e_2 \in E} w_1(e_1) w_2(e_2) \phi_1(e_1) \phi_2(e_2) \\ &= \sum_{e_1 \in E_1} w_1(e_1)\phi_1(e_1) \cdot \sum_{e_2 \in E_2} w_2(e_2)\phi_2(e_2) \\ &= N^*(H_1, w_1) \cdot N^*(H_2, w_2). \end{aligned}$$

For proving the other direction, that is $N^*(H_1 \times H_2, w_1 \times w_2) \geq N^*(H_1, w_1) \cdot N^*(H_2, w_2)$, it is convenient to use again the dual program

$$N^*(H, w) = \max \left\{ \vec{1}^T \Phi | A\Phi \leq w, \Phi \geq \vec{0} \right\}.$$

We can think about every such vector Φ as a real function defined over $V(H)$. Let Φ_1 and Φ_2 be the functions that give the maximum for $N^*(H_1, w_1)$ and $N^*(H_2, w_2)$ (respectively), at the above linear program. Define $\Phi(v_1, v_2) = \Phi_1(v_1) \cdot \Phi_2(v_2)$. We show that Φ satisfies the conditions in the linear program for $H_1 \times H_2$ and that its value (i.e, $\sum_{v \in V(H_1 \times H_2)} \Phi(v)$) is the multiplication of the values of Φ_1 and Φ_2. Clearly, Φ is a nonnegative function defined over $V \stackrel{\triangle}{=} V(H_1 \times H_2)$. In addition, for every hyperedge $e_1 \times e_2 \in E(H_1 \times H_2)$

$$\begin{aligned} \sum_{(v_1,v_2) \in V \,:\, (v_1,v_2) \in e_1 \times e_2} \Phi(v_1, v_2) &= \sum_{v_1 \in V_1 \,:\, v_1 \in e_1} \sum_{v_2 \in V_2 \,:\, v_2 \in e_2} \Phi(v_1, v_2) \\ &= \sum_{v_1 \in V_1 \,:\, v_1 \in e_1} \sum_{v_2 \in V_2 \,:\, v_2 \in e_2} \Phi_1(v_1)\Phi_2(v_2) \\ &= \sum_{v_1 \in V_1 \,:\, v_1 \in e_1} \Phi_1(v_1) \sum_{v_2 \in V_2 \,:\, v_2 \in e_2} \Phi_2(v_2) \\ &\leq w_1(e_1) \cdot w_2(e_2) \\ &= w_1 \times w_2(e_1 \times e_2). \end{aligned}$$

Thus,

$$\begin{aligned} N^*(H_1 \times H_2, w_1 \times w_2) &\geq \sum_{(v_1,v_2) \in V} \Phi(v_1, v_2) \\ &= \sum_{(v_1,v_2) \in V} \Phi_1(v_1)\Phi_2(v_2) \\ &= \sum_{v_1 \in V_1} \Phi_1(v_1) \cdot \sum_{v_2 \in V_2} \Phi_2(v_2) \\ &= N^*(H_1, w_1) \cdot N^*(H_2, w_2). \end{aligned}$$

This completes the proof of the theorem. □

4.3. Proof of Theorem 4.2. Now, we can come back to the proof of Theorem 4.2. To analyze the two-round deterministic communication complexity of a relation R, we define the following hypergraph H_R^2. The vertices are again all the pairs in $\{0,1\}^n \times \{0,1\}^n$. The hyperedges are all the rectangles of the form $A \times \{0,1\}^n$, where $A \subseteq \{0,1\}^n$. Now, for each such hyperedge e we define its weight to be $D_1(e)$, the one-way deterministic communication complexity of computing R on the subdomain e.

Note that for every relation R, the definitions of H_R^2 and of D_1 imply that $D(H_R^2, D_1) = N(H_R^2, D_1)$ and $D^*(H_R^2, D_1) = N^*(H_R^2, D_1)$.

We are interested in the relations between $D(H_{R\times S}^2, D_1)$ and $D(H_R^2, D_1), D(H_S^2, D_1)$. Again, it can be easily verified that $D(H_{R\times S}^2, D_1) \leq D(H_R^2, D_1) \cdot D(H_S^2, D_1)$. For proving connections in the opposite direction, let us concentrate for a while on the case of computing functions. We need the following lemma.

LEMMA 4.5. *Let $e \in E(H_{f\times g}^2)$. That is, $e = A \times (\{0,1\}^n \times \{0,1\}^n)$, where $A \subseteq (\{0,1\}^n \times \{0,1\}^n)$. Let A_f and A_g be the projection of A on the first and second coordinates (respectively). Then,*

1. $D_1((A_f \times A_g) \times (\{0,1\}^n \times \{0,1\}^n)) = D_1(A_f \times \{0,1\}^n) \cdot D_1(A_g \times \{0,1\}^n)$;
2. $D_1(e) = D_1((A_f \times A_g) \times (\{0,1\}^n \times \{0,1\}^n))$.

Proof. (1) was proved in [FKN91]. We now prove (2): As $A \subseteq A_f \times A_g$ then one direction is trivial. Therefore, it is enough to prove that for any $B \subseteq (\{0,1\}^n \times \{0,1\}^n)$, if the submatrix $A \times B$ is constant in each row then so is the bigger submatrix $(A_f \times A_g) \times B$. By the definitions if $A \times B$ in constant in each row then for every $(x_1, x_2) \in A$ and every $(y_1, y_2), (y_1', y_2') \in B$ we have $f \times g((x_1, x_2), (y_1, y_2)) = f \times g((x_1, x_2), (y_1', y_2'))$. In particular, this means that for every $x_1 \in A_f, x_2 \in A_g$ and every $(y_1, y_2), (y_1', y_2') \in B$ we have $f(x_1, y_1) = f(x_1, y_1')$ and $g(x_2, y_2) = g(x_2, y_2')$, which gives us what we need. □

The following theorem is an analogue of Lemma 2.9.

THEOREM 4.6. *Let f and g be two functions. Then $N^*(H_{f\times g}^2, D_1) = N^*(H_f^2 \times H_g^2, D_1 \times D_1)$.*

Proof. The proof that $N^*(H_{f\times g}^2, D_1) \leq N^*(H_f^2 \times H_g^2, D_1 \times D_1)$ is similar to the proof of Theorem 2.9 (second direction), together with the first part of Lemma 4.5 that guarantees that $D_1(e_f \times e_g) = D_1(e_f) \cdot D_1(e_g)$.

The proof that $N^*(H_{f\times g}^2, D_1) \geq N^*(H_f^2 \times H_g^2, D_1 \times D_1)$ is similar to the proof of Theorem 2.9 (first direction), together with the second part of Lemma 4.5 that guarantees that $D_1(e) = D_1(e_f) \cdot D_1(e_g)$. □

Using the last two theorems, we get

$$D(H_f^2, D_1) \cdot D(H_g^2, D_1) \geq D(H_{f\times g}^2, D_1) \geq \frac{D(H_f^2, D_1) \cdot D(H_g^2, D_1)}{cn^2},$$

for some constant c.

Let us now briefly discuss the case of computing general relations and not necessarily functions. The equality in Lemma 4.5 part (1) does not hold anymore. However, by [FKN91] the two sides cannot be too far. As a result, Theorem 4.6 is changed as well and it claims: let R and S be two relations. Then

$$\frac{N^*(H_R^2 \times H_S^2, D_1 \times D_1)}{\ln |V(H_{R\times S})|} \leq N^*(H_{R\times S}^2, D_1) \leq N^*(H_R^2 \times H_S^2, D_1 \times D_1),$$

which implies

$$D(H^2_R, D_1) \cdot D(H^2_S, D_1) \geq D(H^2_{R\times S}, D_1) \geq \frac{D(H^2_R, D_1) \cdot D(H^2_S, D_1)}{cn^4}.$$

REFERENCES

[AUY83] A. V. AHO, J. D. ULLMAN, AND M. YANNAKAKIS, *On notions of information transfer in* VLSI *circuits*, Proc. 15th Annual ACM Symp. on Theory of Computing, 1989, pp. 133–139.

[BS90] R. BOPPANA AND M. SIPSER, *The Complexity of finite functions*, in Handbook of Theoretical Computer Science (Vol. A), J. van Leeuwen, ed., Elsevier Science Publishers, 1990, pp. 759–804.

[C79] V. CHVÁTAL, *A greedy heuristic for set-covering problem*, Math. Oper. Res., 4 (1979), pp. 233–235.

[ES74] P. ERDOS AND J. SPENCER, *Probabilistic Methods in Combinatorics*, Academic Press, New York, 1974.

[FKN91] T. FEDER, E. KUSHILEVITZ, AND M. NAOR, *Amortized communication complexity*, Proc. 32nd Annual Symposium on Foundations of Computer Science, 1991, pp. 239–248; SIAM J. Comput., 25 (1995), to appear.

[F87] M. FURER, *The power of randomness for communication complexity*, Proc. 19th Annual ACM Sympos. on Theory of Computing, (1987), pp. 178–181.

[GH89] M. GOLDMANN AND J. HASTAD, *A simple lower bound for monotone clique using a communication game*, Inform. Process. Lett., 41 (1992), pp. 221–226.

[K71] V. KHRAPCHENKO, *A method of determining lower bounds for the complexity of π-schemes*, Math. Notes Acad. Sci. USSR, (1971), pp. 474–479.

[K89] M. KARCHMER, *Communication Complexity: A New Approach to Circuit Depth*, The MIT Press, Cambridge, MA, 1989.

[KRW91] M. KARCHMER, R. RAZ, AND A. WIGDERSON, *On proving super-logarithmic depth lower bounds via the direct sum in communication complexity*, in Proc. 6th IEEE Structure in Complexity Theory Annual Conference, (1991), pp. 299–304.

[KW88] M. KARCHMER AND A. WIGDERSON, *Monotone circuits for connectivity require super-logarithmic depth*, SIAM J. Discrete Math., 3 (1990), pp. 255–265.

[L75] L. LOVÁSZ, *On the ratio of optimal integral and fractional covers*, Discrete Math., 13 (1975), pp. 383–390.

[MS82] K. MEHLHORN AND E. M. SCHMIDT, *Las Vegas is better than determinism in VLSI and distributive computing*, Proc. 14th Annual ACM Sympos. on Theory of Computing, (1982), pp. 330–337.

[N91] I. NEWMAN, *Private vs. common random bits in communication complexity*, Inform. Process. Lett. 39 (1991), pp. 67–71.

[Y79] A. C.-C. YAO, *Some complexity questions related to distributive computing*, Proc. 11th Annual ACM Symp. on Theory of Computing, 1979, pp. 209–213.