Direct Product Results and the GCD Problem, in Old and New Communication Models

Itzhak Parnafes* Ran Raz† Avi Wigderson ‡

Abstract

This paper contains several results regarding the communication complexity model and the 2-prover games model, which are based on interaction between the two models:

1. *We show how to improve the rate of exponential decrease in the parallel repetition theorem of [Ra] in terms of the communication complexity of the verifier's predicate.*
2. *We apply the improved parallel repetition theorem of 2-prover games to derive, for the first time, a direct product theorem for communication complexity.*

The second derivation uses a common generalization of the two models, which is independently interesting. We initiate a study of its power by considering the GCD problem, and some variations of it, which exhibit a power gap between the new model and the classical communication complexity model. This gap is partly based on the following upper bounds: Given n-bit inputs x and y to Alice and Bob respectively, they can achieve the tasks below with very high probability using only $O(n/\log n)$ communication bits:

1. *Decide if $GCD(x, y) = 1$.*
2. *If $GCD(x, y) = 1$, compute numbers a (by Alice) and b (by Bob), satisfying $a \cdot x + b \cdot y = 1$.*

Observe that the outputs in the second task are in general of length $\Omega(n)$. A complete analysis of the communication complexity of these two problems (in several models and modes) is given.

*Weizmann Institute of Science

†Weizmann Institute of Science. Partially supported by an American-Israeli BSF grant 95-00238

‡The Hebrew University. Partially supported by an American-Israeli BSF grant 92-00106, by the Wolfson Research Awards, administered by the Israel Academy of Sciences and Humanities, and by the Alfred P. Sloan Foundation

1 Introduction

This paper is naturally divided into a "structural" and "concrete" parts. The first part deals with several results of the "direct product" type (which holds for every function). The second part analyses the complexity of particular problems related to the GCD function, in some new and old communication models and modes.

1.1 Direct Product Theorems

Direct sum and direct product problems, conjectures and results are an important paradigm in understanding the power and "sophistication" of a computational model, and have been studied for a variety of models.

In a direct sum problem, on which we shall not elaborate (see [IRW] for more details and references) one asks the following question: If a model requires cost c (in some complexity measure) to solve a given problem on one input x, how costly would it be to solve it on k "independent" inputs $x_1, \cdots, x_k$? A direct sum result holds if no significant savings can be obtained by combining computations, i.e., if the model requires cost about $c \cdot k$ to solve all k instances. Several models obey this rule, while for others surprising savings are possible despite the independence of different inputs. We just mention here that for communication complexity (the model of interest here) the general question is still open, although some nontrivial bounds and applications are known [FKN, EIRS, KRW].

In a direct product problem, a dual view is taken, in which we fix the cost (and model), and study the probability p that a random input (over some input distribution) is solved correctly. A direct product result holds if for fixed p the probability that the model solves[1] k independent instances correctly drops exponentially with k. Even when a direct product theorem holds, an important parameter is the *rate* of exponential decrease $R \leq 1$ (which should depend on the model and cost, but not on k). Such a theorem holds with rate R if for fixed p the success probability to solve k independent inputs is $p^{\Omega(Rk)}$. Several such results were proven already, and we briefly and informally describe them below. All of them hold for every input distribution and every function (or relation) computed.

- **Boolean Circuits:** In this standard model the cost is the circuit's size. A direct product theorem is known with rate $R = 1 - \epsilon$ for any cost $c^{\Theta(1)}$ as long as $k << \log c$ (and this is best possible for relativizing results). It can be derived from the stronger XOR Lemma of Yao [Ya] (see e.g. [NRS]). It is also implicit in [ABG] and explicit in [GNW].
- **Boolean Decision Trees:** This model attempts to compute a Boolean function on input x after querying (adaptively) at most d (= the cost = tree depth) input bits. [NRS] proved a direct product theorem with rate $R = 1$ for every d. A weaker theorem was independently proved in [IRW].
- **One-Way Communication Complexity:** In this model Bob is attempting to compute a function $f(x, y)$ after receiving the input y, and *one* message of at most c communication bits (= the cost) from Alice, who got the input x. A direct product theorem with rate $R = 1$ was proved in [IRW] for every c.
- **2-Prover Games:** Here Alice and Bob, who receive x and y respectively, are trying to satisfy a 4-way predicate $Q(x, y, u, v)$ by independently sending (to a verifier) messages u (by Alice) and v (by Bob). There is no explicit cost in the original description of this model, but an upper bound t on the message lengths (= $|u| + |v|$ = the communication complexity from Alice and Bob to the verifier) may serve as cost. The Parallel Repetition Theorem (PRT) of [Ra] gives a direct product theorem with rate $R = 1/t$ for this model. This bound is nearly tight, as [FV] proved that the best rate possible is at most $(\log t)/t$.

The main result of this part of the paper is a direct product theorem for the general communication complexity model (formally given as Theorem 3.1 in Section 3), where, unlike the One-Way case, the players may exchange c bits using many rounds:

THEOREM 1: For every input distribution and every function, communication complexity of cost c has a direct product theorem with rate $1/c$.

The proof follows by a simple reduction to the PRT. A direct reduction, however, gives only rate of 2^{-c}. To improve it, we first improve the rate in the PRT. At first sight this may seem impossible, due to the near tightness of the rate in PRT, mentioned above. However, t is simply not the "right" cost measure. We observe that the proof of the PRT allows the parameter t to be replaced by a possibly much smaller parameter, related to the communication complexity of the verifier's predicate Q. For each *fixed* input pair x, y to Alice and Bob, imagine the communication complexity of testing if a pair u, v satisfies $Q(x, y, u, v) = 1$. Let $\bar{c}(Q)$ be the maximum of this complexity over all pairs x, y. Clearly, $\bar{c}(Q) \leq t$. Moreover, in some cases it is exponentially smaller. We prove (formally in Theorem 5.1 of Section 5, and actually for the stronger parameter $\bar{\nu}(Q)$ defined there):

THEOREM 2: The parallel repetition theorem holds with rate $1/\bar{c}(Q)$.

The reduction for obtaining the direct product theorem for general communication complexity (THEOREM 1) from the improved direct product theorem for 2-prover games (THEOREM 2) uses a common generalization of both models, and an extension of THEOREM 2 for this new model. Intuitively, the new model is a 2-Prover model, in which Alice and Bob may communicate some c bits (as in the usual communication complexity model) before sending their messages to the verifier. The formal definition as well as the generalization of THEOREM 2 for this model appears in Section 6.

Feeling that the new model is of independent interest, we turn to analyze it and its relation to classical communication complexity in the next part of the paper.

1.2 A New Model of Communication Complexity

In Yao's classical model [Y], Alice and Bob respectively get (n-bit) inputs x, y, and after communicating should derive a *common* answer s. Their task is to try to satisfy some 3-way predicate $R(x, y, s)$. Most commonly $s = f(x, y)$ for some function f.

The new model specifies that after communicating, Alice and Bob should derive their own *individual* answers u, v respectively. Their task is to try to satisfy a 4-way predicate $Q(x, y, u, v)$. Clearly, the common answer model can be thought of as a special case of the

[1] This requires a precise definition, for which there are several natural choices – see Section 3

individual answer model by setting $Q(x, y, u, v) = 1$ iff $(u = v) \wedge (R(x, y, u) = 1)$ (i.e., forcing the players to give the same answer).

Conversely, given any problem (i.e., a predicate Q) for the individual model, we can construct a problem (i.e., a predicate R) for the common model by setting $R(x, y, (u, v)) = Q(x, y, u, v)$, (i.e., forcing both players to know both answers). With the last transformation, we can always think of a 4-way predicate Q as a problem for both models.

Clearly, the above discussion implies that for every computational mode (deterministic, probabilistic, non-deterministic etc.), every Q is at least as easy in the individual model as in the common model. This raises the question of power gaps between these models.

Our main results are for the bounded error probabilistic communication complexity, that is, the minimum number of communication bits Alice and Bob need, to achieve error probability $< 1/3$ (over the random coins), in satisfying Q on the worst input pair x, y. Denote this by $r(Q)$ for the common model and by $\hat{r}(Q)$ for the individual model.

1.3 Communication Complexity of GCD in Old and New Models

The problems we study are related to the Greatest Common Divisor problem (GCD). Since the deterministic communication complexity of all the problems considered will be $\Theta(n)$, we will focus our attention on the probabilistic case.

Let $GCD(x, y)$ denote the gcd of the two integers (each of which is smaller than 2^n), represented by the n-bit strings x, y. Let RP be the relative primeness function, i.e., $RP(x, y) = 1$ if $GCD(x, y) = 1$ and $RP(x, y) = 0$ otherwise. Let $EGCD$ (i.e., Extended-GCD) be a 4-way predicate defined by $EGCD(x, y, u, v) = 1$ iff either $RP(x, y) = 0$ or $u \cdot x + v \cdot y = GCD(x, y) = 1$. (we put no explicit restriction on the size of the answers u, v).

We give tight bounds on the probabilistic (and deterministic) communication complexity of this relation in both models (recall again that in the common model both players must agree on u, v while in the individual model Alice should know u and Bob should know v). The theorem below comprises Theorems 7.4 and 7.5 in Section 7.

THEOREM 3:

- $r(EGCD) = \Theta(n)$.
- $\hat{r}(EGCD) = \Theta(n/\log n)$.

Except for the trivial upper bound (the first one), all other bounds (in both models) are obtained via probabilistic reductions to computing RP in the common model, for which we prove:

THEOREM 4: $r(RP) = \Theta(n/\log n)$.

The question of whether the logarithmic gap between the two models, achieved for $EGCD$ in THEOREM 3, is the largest possible has two answers: The trivial answer is no, as some simple examples exhibit a linear gap. In Section 7 we explain why such examples are not interesting, and provide a natural condition on relations (related to the measure $\bar{c}(Q)$ controlling the rate in THEOREM 2) which make the gap problem more meaningful. Under this condition, it is open if larger than logarithmic gaps are possible.

2 (Classical) Communication Complexity

We work with the standard communication complexity model of Yao [Y]. In this model, we consider three finite sets X,Y, and S, and a function $f : X \times Y \to S$. Intuitively, two players, Alice and Bob, respectively get $x \in X$ and $y \in Y$, and their goal is to compute the value of $f(x, y)$ by exchanging bits of information.

Formally, a protocol P (over X, Y, S) is a binary tree in which

1. Every internal vertex in odd levels (in which A sends a bit) is labeled by a Boolean function of X (into one of the vertex' children).
2. Every internal vertex in even levels (in which B sends a bit) is labeled by a Boolean function of Y (into one of the vertex' children).
3. Every leaf is labeled by an element of S.

Clearly, every input (x, y) uniquely defines a leaf reached by the protocol, and we let $P(x, y)$ denote the label of this leaf. The complexity of P is the depth of the tree. Thus a c-protocol is one whose depth is at most c. There are several standard measures of communication complexity which will be of interest for this paper:

- **Deterministic Complexity:** We denote by $c(f)$ the minimum c for which there is a c-protocol P computing f, that is, a protocol for which $P(x, y) = f(x, y)$, for all inputs $(x, y) \in X \times Y$.
- **Probabilistic Complexity:** A probabilistic c-protocol is simply a probability distribution over c-protocols. The (bounded error) probabilistic complexity $r(f)$ is defined as the minimum c, for which

there exists such a distribution, such that if P is a deterministic c-protocol taken at random according to that distribution then $Pr[P(x,y) = f(x,y)] \geq 2/3$, for all inputs $(x,y) \in X \times Y$.

- **Distributional Complexity:** Here $X \times Y$ is endowed with a probability measure μ. For every integer c, we denote by $p_c(f)$ the largest value of the probability $Pr[P(x,y) = f(x,y)]$ (with respect to μ), taken over all deterministic c-protocols P.
- **Nondeterministic Complexity and Fractional Cover Complexity:** These measures are defined for the case $S = \{0,1\}$. Intuitively, the players are trying to "verify" that their input is a 1 of f. Formally, call a *1-rectangle* any function $g : X \times Y \to \{0,1\}$ such that $g(x,y) = 1$ iff $(x,y) \in X' \times Y'$ for some subsets $X' \subseteq X$ and $Y' \subseteq Y$. A *fractional cover* for f is a finite sequence (α_j, g_j), where the α's are nonnegative reals, and the g's are 1-rectangles, satisfying

  1. $\sum_j \alpha_j \cdot g_j(x,y) = 0$ for every x, y with $f(x,y) = 0$.
  2. $\sum_j \alpha_j \cdot g_j(x,y) \geq 1$ for every x, y with $f(x,y) = 1$.

  The fractional cover complexity $\nu(f)$ is the logarithm of the smallest $\sum_j \alpha_j$, over all fractional covers for f. A cover is integral if all $\alpha_j = 1$. The nondeterministic complexity $n(f)$ is simply the logarithm of the smallest number of elements in an integral cover. Note that $\nu(f) \leq n(f) \leq c(f)$, and that the gaps can be quite large. E.g. for the inequality function, $f(x,y) = 1$ iff $x \neq y$, over $X = Y = \{0,1\}^n$, we have $c(f) = n$, $n(f) = \log n$, and $\nu(f) = 2$.

3 Direct Product of Communication Complexity Problems

We will be mainly interested in the *distributional complexity* of functions. Thus, a communication problem $f : X \times Y \to S$ will be endowed with a probability distribution μ on $X \times Y$ (which is not necessarily a product distribution). Since our results will hold for any distribution, we shall not specify μ explicitly.

The direct product problem is to study how well can f be computed on k independent inputs. For a set W and integer k, let W^k denote the Cartesian product of W with itself k times. Let $\bar{x} = (x_1, \cdots, x_k)$, and $\bar{y} = (y_1, \cdots, y_k)$. Define $f^{\otimes k} : X^k \times Y^k \to S^k$ by

$$f^{\otimes k}(\bar{x}, \bar{y}) = (f(x_1, y_1), \cdots, f(x_k, y_k)).$$

We stress that the input distribution associated with $f^{\otimes k}$ is $\mu^{\otimes k}$, the k-fold product of μ with itself.

Two natural ways to define the direct product problem appear respectively in [IRW] and [NRS]. We will choose the second, after briefly discussing the first.

One natural way to compute $f^{\otimes k}$ is by a single protocol [IRW]. Thus we wish to compare $p_d(f^{\otimes k})$ to $p_c(f)$, where the depth d for $f^{\otimes k}$ is related to the allowed depth c for f. In particular, for what functions d of c and k do we have the following exponential decrease ? : for every function f and integers c, k,

$$p_{d(k,c)}(f^{\otimes k}) = p_c(f)^{\Omega(k)}.$$

This was not known to hold even for "minimal choice", $d = c$, — except in the restricted model of one-way communication complexity [IRW] — (but follows from our results here). For the "maximal choice", $d = ck$, we can construct a simple example f, for which there is no decrease at all for $k = c + 1$, i.e., $p_c(f) = p_{ck}(f^{\otimes k}) = 1/2$.

Another natural way to compute $f^{\otimes k}$ (which was introduced in [NRS] for decision trees, and we adopt here) is to allow k different c-protocols for the k different outputs of $f^{\otimes k}$, where each is allowed to depend on **all** inputs (i.e., we have k different binary trees, with functions labeling the internal vertices of each depending on X^k or Y^k, and not only on the corresponding coordinate). Define a (c,k)-protocol over X^k, Y^k, S^k, as a collection of such k c-protocols over X^k, Y^k, S. Define $p_{(c,k)}(f^{\otimes k})$ to be the maximum of

$$Pr\left[\cap_{i=1}^k P_i(\bar{x}, \bar{y}) = f(x_i, y_i)\right],$$

over all (c,k)-protocols. It is easy to see that:

Proposition 3.1 *For every function f, and integers c, k,*

$$p_c(f^{\otimes k}) \leq p_{(c,k)}(f^{\otimes k}) \leq p_{ck}(f^{\otimes k}).$$

We will prove here (in Section 6.3) that for fixed f, $p_{(c,k)}$ is exponentially decreasing.

Theorem 3.1 *For every function f, and integers c, k we have*

$$p_{(c,k)}(f^{\otimes k}) \leq p_c(f)^{\Omega(k/c)}.$$

Observe that the *rate* of this exponential decay is $1/c$. Whether it can be replaced by an absolute constant is a very interesting question.

Before proceeding we remark that every such exponential decrease result has direct application to the help bit problem (see e.g. [ABG, Cai, NW, NRS]), which we will not define formally here. Loosely speaking, the players A and B cannot correctly compute $f^{\otimes k}$ by a c-protocol even after obtaining arbitrary $o(k)$ bits of information on the entire input $(\bar{x}, \bar{y})$.

4 Parallel Repetition of Standard Games

A game G consists of four finite sets X, Y, U, V, with a probability measure $\mu : X \times Y \to R^+$, and a predicate $Q : X \times Y \times U \times V \to \{0, 1\}$. A protocol for G consists of a function $u : X \to U$, and a function $v : Y \to V$. The value of the protocol is defined as

$$Pr[Q(x, y, u(x), v(y)) = 1]$$

(where (x, y) are chosen according to μ). The value of the game $w(G)$ is defined to be the maximal value of all protocols for G.

Note that G can be thought of as a zero-rounds communication model as follows: Alice and Bob receive respectively x, y (chosen according to the distribution μ). They each respond with one message (from U, V respectively) to a referee (verifier). Their task is to maximize the probability that Q is satisfied. Note that $t = \log |U||V|$ is an upper bound on the (zero-round) communication complexity of every such protocol.

The (parallel repetition) game $G^{\otimes k}$ captures the direct product problem for this model. It consists of the sets X^k, Y^k, U^k, V^k, with the product measure

$$\mu^{\otimes k}(\bar{x}, \bar{y}) = \prod_{i=1}^{k} \mu(x_i, y_i),$$

and the predicate

$$Q^{\otimes k}(\bar{x}, \bar{y}, \bar{u}, \bar{v}) = \prod_{i=1}^{k} Q(x_i, y_i, u_i, v_i),$$

where $\bar{x} = (x_1, \cdots, x_k) \in X^k, \bar{y} = (y_1, \cdots, y_k) \in Y^k, \bar{u} = (u_1, \cdots, u_k) \in U^k, \bar{v} = (v_1, \cdots, v_k) \in V^k$.

The main result of [Ra] is proving that for a fixed game, the value of $G^{\otimes k}$ decreases exponentially with k.

Theorem 4.1 *[Parallel Repetition Theorem [Ra]] For all games G,*

$$w(G^{\otimes k}) \leq w(G)^{\Omega(k/t)}.$$

Note that the *rate* of this exponential decay is $1/t$, which is the length of the messages sent by the players. [FV] recently gave an example of a game G with

$$w(G^{\otimes k}) \geq w(G)^{O(k/(t/\log t))}.$$

This example shows that the rate in this theorem cannot be replaced by a global constant, and that in some cases $1/t$ is almost best possible.

5 Improving the Rate in Parallel Repetition

As mentioned above, the result of [FV] shows that the rate, as a function of the players' message length t, is nearly optimal. Our improvement will give a rate which is a function of another parameter, which may be arbitrarily smaller than t (even a constant independent of t). This parameter comes from viewing the acceptance predicate Q of a game G, with inputs x, y, as a communication problem over $U \times V$. While t corresponds to the trivial upper bound on the (even zero-rounds) deterministic communication complexity, our parameter will be the fractional cover complexity (recall the definition in Section 2) of these problems.

More formally, given a game G, we define for each pair $(x, y) \in X \times Y$ a communication problem $Q_{x,y} : U \times V \to \{0, 1\}$, by $Q_{x,y}(u, v) = Q(x, y, u, v)$. Recalling the definitions of deterministic, nondeterministic and fractional cover complexities (from Section 2), let

- $\bar{c}(Q) = Max_{x,y}\, c(Q_{x,y})$.
- $\bar{n}(Q) = Max_{x,y}\, n(Q_{x,y})$.
- $\bar{\nu}(Q) = Max_{x,y}\, \nu(Q_{x,y})$.

All of these parameters can replace t for controlling the rate in the Parallel Repetition Theorem. Since for every Q we have $\bar{\nu}(Q) \leq \bar{n}(Q) \leq \bar{c}(Q)$, we state it for the stronger $\bar{\nu}(Q)$, which is denoted equivalently by $\bar{\nu}(G)$.

Theorem 5.1 *For all games G,*

$$w(G^{\otimes k}) \leq w(G)^{\Omega(k/\bar{\nu}(G))}.$$

Proof: (Sketch) The proof is by a simple modification of the proof given in [Ra]. Let us describe the essential ideas in that proof:

Let $(\bar{u}, \bar{v})$ be any protocol for $G^{\otimes k}$. Denote by q_i the random variable $Q(x_i, y_i, u_i(\bar{x}), v_i(\bar{y}))$ (under the distribution $\mu^{\otimes k}$). We write $Pr[A]$ for the probability of the event A under the distribution $\mu^{\otimes k}$. Our task is to upper bound the value of the protocol $(\bar{u}, \bar{v})$, namely

$$Pr[\Pi_{i=1}^{k} q_i = 1].$$

To facilitate induction on k we will estimate

$$Pr[\Pi_{i=1}^{k} q_i = 1 | A],$$

where $A = \hat{X} \times \hat{Y}$ with $\hat{X} \subseteq X^k$ and $\hat{Y} \subseteq Y^k$. Let $C_G(k, r)$ be the maximum of the above probability, over all such subsets A with

$$-\log \mu^{\otimes k}(A) \leq r,$$

and over all protocols $(\bar{u},\bar{v})$.

Most effort in [Ra] is devoted to proving that, as long as A is "large enough", then in some coordinate i the probability $Pr[q_i = 1|A]$ is strictly smaller than 1. Precisely

Theorem 5.2 *[Ra] If* $-\log \mu^{\otimes k}(A) < \delta \cdot k$ *then for some* i,

$$Pr[q_i = 1|A] < w', \tag{1}$$

(where $w' < 1$, *and* δ, w' *depend only on* G *and not on* k*).*

Using this we prove by induction on k,

Theorem 5.3 *For every* k, r *with* $r \le \delta \cdot k$ *we have*

$$C_G(k,r) \le (w')^{(\delta k - r)/\bar{\nu}(G)}.$$

Clearly, this proves Theorem 5.1 by setting $r = 0$.

Assume w.l.o.g. that the inequality (1) holds for $i = k$. Then we can write (for any protocol $(\bar{u},\bar{v})$)

$$Pr[\Pi_{i=1}^{k} q_i = 1|A] = \tag{2}$$

$$Pr[q_k = 1|A] \cdot Pr[\Pi_{i=1}^{k-1} q_i = 1|A, q_k = 1] \le$$
$$w' \cdot Pr[\Pi_{i=1}^{k-1} q_i = 1|A, q_k = 1].$$

To use induction we partition the event $q_k = 1$ into Cartesian products. Here comes the generalization of the original proof, in which these Cartesian products were chosen to be the simplest ones: each is determined by a tuple $(x_k, y_k, u_k(\bar{x}), v_k(\bar{y}))$ for which Q is 1. Here we will consider any partition, in fact a fractional cover, to 1-rectangles of Q:

Formally, for each $(x,y) \in X \times Y$, let $(\alpha_j^{x,y}, g_j^{x,y})$ be a fractional cover of $Q_{x,y}$. Since each $g_j^{x,y}$ is a 1-rectangle in $U \times V$, and by the fact that the function u_k (resp. v_k) is only a function of X^k (resp. Y^k), we can write

$$Pr[\Pi_{i=1}^{k-1} q_i = 1|A, q_k = 1] \le \tag{3}$$

$$\sum_{x,y,j} Pr[(x_k, y_k) = (x,y)|A, q_k = 1] \cdot$$

$$\alpha_j^{x,y} \cdot Pr[\Pi_{i=1}^{k-1} q_i = 1|A, B_j^{x,y}],$$

where each $B_j^{x,y}$ is the rectangle in $X^k \times Y^k$ induced by fixing $x_k = x, y_k = y$ and fixing the k-th coordinate of the answers, i.e., the pair $(u_k(\bar{x}), v_k(\bar{y}))$, to be a pair in the rectangle induced by $g_j^{x,y}$.

Since each $B_j^{x,y}$ determines (x_k, y_k), the fact that we have a product measure enables reducing the number of coordinates by 1, remaining with the right product measure. Moreover, the k-input $(\bar{u},\bar{v})$ induces protocols for each part of the partition, which are now on $k-1$ inputs. The extra information given by each part is determined by the size of the relevant 1-rectangle. Finally, taking the worst case over all (x,y) values of the k-th coordinate, and recalling that $\bar{\nu}(G) = max_{x,y} \nu(Q_{x,y})$, we infer the inductive formula

$$C_G(k,r) \le w' \sum \alpha_j C_G(k-1, r+\beta_j), \tag{4}$$

where $\sum_j \alpha_j \beta_j \le \bar{\nu}(G)$.

Using simple convexity arguments as in [Ra] this proves Theorem 5.3. □

6 The New Model – Games with Communication

In this section we introduce communication complexity to standard games, resulting in a model that can be thought of as a generalization of both. Later, we extend the parallel repetition theorem for the new model.

6.1 The New Model

As in standard games, a game G consists of four sets X, Y, U, V, with a probability measure $\mu : X \times Y \to R^+$, and a predicate $Q : X \times Y \times U \times V \to \{0,1\}$. As before, Alice and Bob receive respectively x, y (chosen according to the distribution μ). As before, they each respond with one message (from U, V respectively), and they are trying to maximize the probability that Q is satisfied. However, before they respond, they are allowed to communicate between them c bits, according to a communication protocol. Since after exchanging c bits of information each player may have some information about the input of the other, it is expected that in some cases the value of the game is higher than before.

More formally, a protocol P (for Alice and Bob) is a protocol over X, Y, S, as defined in Section 2, with $S = U^X \times V^Y$, (i.e., pairs of functions $u : X \to U$ and $v : Y \to V$). In words, P is a standard protocol, with each leaf z labeled by a pair of functions (u_z, v_z), which the players use to construct their answers in the standard game, (i.e., the answers are $u_z(x), v_z(y)$).

The value of the protocol is defined as $Pr[Q(x, y, u_z(x), v_z(y)) = 1]$, where (x,y) are chosen according to μ, and z is the leaf reached by Alice having input x, and Bob having input y, when using the protocol P. For every integer c, the value of the game, $w_c(G)$, is defined to be the maximal value of all protocols (for G) of depth c, over X, Y, S. Note that this definition of games is a generalization of standard games, as well as a generalization of communication complexity games.

6.2 A Parallel Repetition Theorem for the New Model

In parallel repetition in the new model we have, as before, k copies of the same game G, defined on k independent inputs. Alice and Bob get the input $(\bar{x}, \bar{y}) = ((x_1, \cdots, x_k), (y_1, \cdots, y_k))$, chosen according to the k-fold product of μ with itself. They each respond with k messages (from U, V respectively), and they are trying to maximize the probability that Q is satisfied on all the coordinates. Of course, before the players give their answers, they are allowed to apply a communication protocol (of specified depth c). In fact, we allow k different c-protocols P_i over $X^k, Y^k, U^X \times V^Y$, for the k different answers of the two players. Note that each protocol P_i is allowed to depend on all inputs, not just the the i-th one.

Formally, define a (c, k)-protocol as a collection of such k c-protocols. The value of the protocol is the probability of success on all the coordinates simultaneously, i.e., $Pr[\forall i Q(x_i, y_i, u_i(\bar{x}), v_i(\bar{y})) = 1]$, where $u_i(\bar{x}), v_i(\bar{y})$ denote the answers of the two players (correspondingly), given by the function corresponding to the leaf z_i reached by P_i on inputs $\bar{x}, \bar{y}$.

Define the value of the direct product game, $w_{(c,k)}(G^{\otimes k})$, to be the maximum value over all (c, k)-protocols for the game $G^{\otimes k}$. The following theorem gives our generalization of the parallel repetition theorem. Recall the definition of $\bar{\nu}(G)$ in the previous section.

Theorem 6.1 *For all games G,*

$$w_{(c,k)}(G^{\otimes k}) \le w_c(G)^{\Omega(k/(c+\bar{\nu}(G)))}.$$

Proof: The proof is by a simple reduction to Theorem 5.1:

Let P be a c-protocol for the game G. Since c is fixed, P is given by the functions describing for each player what bit to send at every internal node z of the protocol tree, and by the functions u_z, v_z for every leaf z. Thus, an input x to Alice determines a vector $A(x)$, which for each internal vertex z in an odd level of the tree, contains Alice's message (in $\{0, 1\}$) on x at this vertex, and for each leaf z contains Alice's answer on x at this leaf (i.e., $u_z(x)$). Similarly we define $B(y)$ for Bob.

Call the set of possible vectors for Alice $\mathcal{A}$, and for Bob $\mathcal{B}$. Observe that the two vectors $A \in \mathcal{A}, B \in \mathcal{B}$ determine:

1. A unique leaf z of P (that has a path in "$A \cup B$" to the root), and
2. The answers given at this leaf.

Denote this leaf by $z(A, B)$, and the corresponding answers by $u(A, B), v(A, B)$. When $A = A(x)$, and $B = B(y)$, these answers are just the answers of P on x, y.

Thus x, y, A, B determine the value of the predicate Q of the game G. This value is just $Q(x, y, u(A, B), v(A, B))$. This observation is the basis for the reduction we use.

Given the game G, we will build an equivalent standard game (without communication), $\tilde{G}_c$, such that:

1. $w_c(G) = w(\tilde{G}_c)$.
2. $w_{(c,k)}(G^{\otimes k}) = w((\tilde{G}_c)^{\otimes k})$.
3. $\bar{\nu}(\tilde{G}_c) \le c + \bar{\nu}(G)$.

The proof of the theorem hence follows by Theorem 5.1.

The standard game $\tilde{G}_c$ will simulate the game G in only one round. The sets X, Y, and the measure μ, of the game $\tilde{G}_c$, stay the same as the ones of G. The sets of answers for the game $\tilde{G}_c$ are defined to be $\tilde{U}_c = \mathcal{A}$ and $\tilde{V}_c = \mathcal{B}$. The predicate, $\tilde{Q}_c$, for the game $\tilde{G}_c$ is defined to be $\tilde{Q}_c(x, y, A, B) = Q(x, y, u(A, B), v(A, B))$.

In words, the two players in the game $\tilde{G}_c$ receive the inputs x, y (as before). The first responds with a vector A and the second with a vector B. The two answers A, B (together) determine the value of $u_z(x), v_z(y)$ (the answers for the original game G). The predicate is satisfied iff the predicate of the original game G is satisfied on these answers.

Thus, the games G, and $\tilde{G}_c$ are equivalent. The discussion above shows that every c-protocol for the game G defines an equivalent standard protocol (without communication) for the game $\tilde{G}_c$. The opposite is also true: every protocol for $\tilde{G}_c$ defines an equivalent protocol for G, since from $\{A(x), B(y) : x \in X, y \in Y\}$ one can explicitly recover the communication protocol for G. These two protocols have the same value. Therefore, we have $w_c(G) = w(\tilde{G}_c)$.

An analogous argument shows that $w_{(c,k)}(G^{\otimes k}) = w((\tilde{G}_c)^{\otimes k})$.

In order to prove that $\bar{\nu}(\tilde{G}_c) \le c + \bar{\nu}(G)$, we will show that for every x, y,

$$\nu(\tilde{Q}_{x,y}) \le c + \nu(Q_{x,y}).$$

The set $\tilde{U}_c \times \tilde{V}_c = \mathcal{A} \times \mathcal{B}$ can be partitioned according to the leaf $z(A, B)$, reached by the vectors A, B. By a standard argument, each set in this partition is a product set (that is, a subset of $\mathcal{A}$ times a subset of $\mathcal{B}$), and can be written as $\mathcal{A}(z) \times \mathcal{B}(z)$. The restriction of $\tilde{Q}_{x,y}$ to $\mathcal{A}(z) \times \mathcal{B}(z)$ is equivalent to the original predicate $Q_{x,y}$ (since it considers only the answers $u_z(x), v_z(y)$ given in A, B on the relevant leaf z). Therefore, denoting this restriction by $Q^z_{x,y}$, we have $\nu(Q^z_{x,y}) = \nu(Q_{x,y})$. Since the number of leaves is at most 2^c, the number of 1-rectangles needed to describe $\tilde{Q}_{x,y}$ is at most 2^c times the number of those needed to describe $Q_{x,y}$. Therefore $\nu(\tilde{Q}_{x,y}) \leq c + \nu(Q_{x,y})$, and $\bar{\nu}(\tilde{G}_c) \leq c + \bar{\nu}(G)$.

□

6.3 A Direct Product Theorem for Communication Complexity

In this subsection we prove Theorem 3.1. Indeed, this theorem becomes now a special case of Theorem 6.1, after making the appropriate definitions:

Fix $f : X \times Y \rightarrow S$ (and the input distribution μ). Define a game $G = G_f$ as follows: The sets X, Y and the input distribution μ are the same. Set $U = S$ and $V = S$. Define $Q = Q_f$ by $Q(x, y, u, v) = 1$ iff $u = v = f(x, y)$. It is trivial to see that for every $(x, y) \in X \times Y$ we have $\nu(Q_{x,y}) = 0$, since $Q_{x,y}$ is just a 1-rectangle. Furthermore, for any integer c we clearly have, by the definition of Q, that $p_c(f) = w_c(G)$ as well as $p_{(c,k)}(f^{\otimes k}) = w_{(c,k)}(G^{\otimes k})$. By Theorem 6.1 we are done.

7 Communication Complexity Revisited

7.1 Definitions and Basic Facts

Here we return to view the new model more as an extension of the classical communication complexity model, rather than of 2-prover games. The definition below is equivalent to the one in the previous section.

Let Q be a relation over X, Y, U, V. It will serve us to define two different communication problems: one for the classical, *common* answer model, and one for the new, *individual* answer model. In both, Alice and Bob respectively receive $x \in X$ and $y \in Y$, and follow a fixed c-protocol. After communication their tasks differ. In the common answer model, both Alice and Bob have to agree on a pair $(u, v) \in U \times V$. In the individual model, Alice computes $u \in U$ and Bob $v \in V$.

The deterministic communication complexity for both models are defined in the standard way: Let $c(Q)$ be the smallest c such that for every x, y, Alice and Bob's common answer (u, v), obtained after following a c-protocol, satisfies $Q(x, y, u, v) = 1$. Similarly, $\hat{c}(Q)$ is the smallest c with the same requirement for the individual answer model.

A probabilistic c-protocol is a distribution over deterministic c-protocols (resulting in the standard definition of a joint random string shared by the two players). Again, we define the probabilistic bounded error communication complexities $r(Q), \hat{r}(Q)$ in the standard way, as the minimum c for which, after following a probabilistic protocol, the probability of satisfying $Q(x, y, u, v)$ is at least $2/3$ for every input pair x, y. It is easy to see that in the new model, just like the old one, the constant $2/3$ has no special importance, and any γ with $1/2 < \gamma < 1$ does not effect these measures by more than a constant factor.

Comment: For functions $f : X \times Y \rightarrow S$ we will continue to use the measures $c(f)$ and $r(f)$ defined for the common model in Section 2 - indeed the above can be taken as their generalization for relations.

It is trivial that the new individual model is at least as strong as the classical common model.

Proposition 7.1 *For every relation Q we have $\hat{c}(Q) \leq c(Q)$ and $\hat{r}(Q) \leq r(Q)$.*

The functions we will study in the next sections are related to the Greatest Common Divisor (GCD) problem. Let $RP : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}$ be the Relative Primeness function, with $RP(x, y) = 1$ iff $GCD(x, y) = 1$, where the inputs are considered as n-bit integers.

It is well known that one can always express $GCD(x, y)$ as an integer combination of x, y. Indeed, such a combination is the output of the extended Euclidean algorithm. This is the basis for the relation $EGCD \subseteq \{0,1\}^n \times \{0,1\}^n \times \mathcal{Z} \times \mathcal{Z}$ (where $\mathcal{Z}$ denotes the integers), defined by $EGCD(x, y, u, v) = 1$ iff either $RP(x, y) = 0$ or $ux + vy = 1$. Note, that we require the players to compute the combination only if the inputs are relatively prime, and that we do not restrict the size of the integer coefficients u, v, despite the fact that small ($O(n)$ length) ones always exist.

Standard arguments show that the deterministic communication complexity of all those problems is linear, that is, $c(RP), c(EGCD), \hat{c}(EGCD)$ are all $\Theta(n)$.

For proving our lower and upper bounds in the probabilistic case, we will need some basic tools from probabilistic communication complexity. Let $DISJ : \{0,1\}^m \times \{0,1\}^m \rightarrow \{0,1\}$, where the inputs represent subsets of an m-element set, be defined by $DISJ(S, T) = 1$ iff $S \cap T = \emptyset$. A fundamental probabilistic lower bound, of many applications, was proven by Kalyanasundaram and Schnitger [KS] (and simplified by Razborov [R]) for the disjointness function.

Theorem 7.1 *[KS, R] $r(DISJ) = \Omega(m)$.*

This result is clearly best possible. However, the trivial upper bound of $O(m)$ can be greatly improved if the input sets S, T are much smaller than m. Let $DISJ_k$ be the restriction of $DISJ$ to subsets of size k. It is trivial that $c(DISJ_k) \leq k \log m + 1$. A useful result of Hastad and Wigderson [HW] shows that the probabilistic complexity can be made completely independent of m, the universe size. Moreover, this algorithm produces a witness to justify a disjointness answer.

Theorem 7.2 *[HW] For every $k \leq m$, $r(DISJ_k) = O(k)$. Moreover, on disjoint inputs S, T, this probabilistic $O(k)$-protocol produces a subset W (output by both players) of $[m]$ satisfying (with high probability): $S \subset W$, $T \subset [m] \setminus W$.*

7.2 RP in the Common Model

Theorem 7.3 $r(RP) = \Theta(n/\log n)$.

Proof: Throughout this proof we use elementary bounds on the distribution of primes, without spelling them out.

First, let us prove the lower-bound, that is, the probabilistic communication complexity of the problem is $\Omega(\frac{n}{\log n})$. This statement comes from a simple reduction to the set disjointness function: Let p_i denote the i-th prime. For a subset $S \subseteq [m]$, let $N_S = \Pi_{i \in S} p_i$ (which is a number of length $O(m \log m)$ bits). On inputs S and T, Alice and Bob can construct respectively N_S and N_T. Obviously, the two sets $S, T \subset [m]$ are disjoint iff $RP(N_S, N_T) = 1$.

Thus by Theorem 7.1, this puts a lower bound of $\Omega(m)$ for a GCD problem with $m \log m$ bits inputs. Set $n = m \log m$ to get $r(RP) = \Omega(\frac{n}{\log n})$.

For the upper bound, we will use the reverse reduction to the set disjointness function, but now with $m = 2^n$: Given a number $N < 2^n$, let the subset $S_N \subset [m]$ be the set of prime divisors of N. Note that since every such number has at most $k \stackrel{\text{def}}{=} n/\log n$ distinct prime divisors, $|S_N| \leq k$. On inputs x, y, Alice and Bob can construct respectively S_x and S_y.

By Theorem 7.2 there exists a probabilistic protocol for $DISJ_k$, which uses $O(k)$ bits, and gives the answer with high probability. Alice and Bob can apply this protocol to S_x, S_y. They answer "$RP(x, y) = 1$" iff that protocol answers "disjoint". By the above discussion, this is the right answer with high probability.

□

7.3 EGCD and the Gap Between the Models

Theorem 7.4 $r(EGCD) = \Theta(n)$.

Proof: The upper bound is trivial. For the lower bound, let us show that a protocol that results in both parties knowing the coefficients must exchange $\Omega(n)$ (indeed, $n - o(n)$) bits.

If this is not the case, a simple protocol can be described, by which each party computes the other's input, and uses only $O(n/\log n)$ extra communication bits, whenever the inputs are relatively prime. This protocol will run the randomized protocol for computing $RP(x, y)$, which takes $O(\frac{n}{\log n})$ bits. If $RP(x, y) = 1$, then running the assumed protocol, each party has only to (privately) solve a linear equation in order to get the other party's input.

Needless to say, no protocol allowing each party to compute the other's input can exchange less than $\Omega(n)$ bits, even when the inputs are restricted to be prime numbers to begin with.

□

Theorem 7.5 $\hat{r}(EGCD) = \Theta(n/\log n)$.

Proof: On relatively prime inputs x and y, Alice has to compute u, and Bob has to compute v, such that

$$u \cdot x + v \cdot y = 1.$$

The lower bound claims that this feat can not be done with less than $\Omega(\frac{n}{\log n})$ bits of communication, even when a small probability of error is allowed. This is true, because otherwise we get a better protocol for RP as follows: After computing the coefficients, the two parties will just check probabilistically if the linear combination is 1 (using some efficient, standard probabilistic equality protocol with $O(\log n)$ communication bits), and answer accordingly. This contradicts Theorem 7.3.

Now, let us show how the linear coefficients can be computed using $O(\frac{n}{\log n})$ bits of communication: First, the players apply the RP protocol. If $RP(x, y) = 0$, they output $u = 0$ and $v = 0$ respectively (in fact, any answer will do).

Otherwise, using once again the reduction to set disjointness and the protocol of Theorem 7.2, the players can produce a set W containing all the prime divisors of x, but none of the prime divisors of y.

Let r be the n-th power of the product of all primes in W, and let s be the n-th power of the product of all primes in the complement of W. Then r, s are mutually primes, known to both players, and satisfying that x divides r, and y divides s.

Thus, r, s have linear coefficients a and b, such that $a \cdot r + b \cdot s = 1$. Now, Alice and Bob each have what they were looking for: Alice will output $v = a\frac{r}{x}$, and Bob will output $u = b\frac{s}{y}$ as the required linear coefficients.

□

7.4 Interesting vs. trivial gaps

We conclude this section by attempting to better understand the sources of power gaps between the common and individual models. For this, consider the following relation ID, which exhibits an infinite gap between the models. Define $ID \subset (\{0,1\}^n)^4$ by $ID(x,y,u,v) = 1$ iff $(x = u) \wedge (y = v)$. The following is trivial.

Proposition 7.2 *$c(ID) = n$ and $\hat{c}(ID) = 0$.*

In trying to understand what makes such examples trivial, and examples like $EGCD$ interesting, let us go back to the measures controlling the rate of decrease in the parallel repetition theorem, $\bar{\nu}(Q), \bar{n}(Q), \bar{c}(Q)$ defined in Section 5. They relate to the worst case (over the choice of x, y), complexity (resp. fractional, nondeterministic, deterministic) of the communication problem $Q_{x,y}$. Note, that this problem essentially asks the players, for a *fixed* pair x, y, to verify that the individual answers u, v satisfy the predicate $Q(x,y,u,v)$.

It is easy to verify that all these measures are very small (0 or 1) for $Q = ID$, but are $\Omega(n)$ for $Q = EGCD$. While it is still unclear to us which one of these, if any, is the right "non-triviality" measure, we challenge the reader to find relations Q with (say) $\bar{n}(Q) = \Omega(n)$ satisfying:
1) $\hat{c}(Q)/c(Q) = o(1)$, 2) $\hat{r}(Q)/r(Q) = o(1/\log n)$.

References

[ABG] A. Amir, R. Beigel, W. Gasarch, "Some connections between bounded query classes and nonuniform complexity", *5th Structures in Complexity Theory Conference*, 1990.

[Cai] J. Cai, "Lower bounds for constant depth circuits in the presence of help bits", *30th FOCS*, pp. 532–537, 1989.

[EIRS] J. Edmonds, R. Impagliazzo, S. Rudich, J. Sgall, "Communication complexity towards lower bounds on circuit depth", *32nd FOCS* , pp. 249–257, 1991.

[FV] U. Feige, O. Verbitsky, "Error Reduction by Parallel Repetition - a Negative Result", *11th Annual IEEE Conference on Computational Complexity, 96.*

[FKN] T. Feder, E. Kushilevitz, M. Naor, "Amortized Communication Complexity", *32nd FOCS*, pp. 239–248, 1991.

[GNW] O. Goldreich, N. Nisan, A. Wigderson, "On Yao's XOR lemma", *ECCC* TR 95-050, 1995.

[HW] J. Hastad, A. Wigderson, "The probabilistic communication complexity of disjointness of k sets is $O(k)$", Unpublished Manuscript, 1990.

[IRW] R. Impagliazzo, R. Raz. A. Wigderson, "A Direct Product Theorem", *Proc. of the 9th Structures in Complexity conference*, pp. 88–96, 1994.

[KKN] M. Karchmer, E. Kushilevitz, N. Nisan, "Fractional Covers and Communication Complexity", *7th Structures in Complexity Theory Conference*, pp. 262–274, 1992.

[KRW] M. Karchmer, R. Raz, A. Wigderson, "On Proving Super-Logarithmic Depth Lower Bounds via the Direct Sum in Communication Complexity", *Structures in Complexity Theory '91*, pp. 299-304 (1991).

[KS] B. Kalyanasundaram and G. Schnitger "The Probabilistic Communication Complexity of Set Intersection", *Proceedings Structure in Complexity Theory* pp.41-49, 1987.

[L] L.Lovász, "On the ratio of optimal integral and fractional cover", *Discrete Mathematics* , 13, pp. 383-390, 1975.

[NRS] N. Nisan, S. Rudich M. Saks, "Products and Help Bits in Decision Trees", *35th FOCS*, pp. 318–329, 1994.

[NW] N. Nisan, A. Wigderson, " Rounds in Communication Complexity Revisited", *SIAM Journal on Computing*, Vol 22, No. 1, 1993.

[R] A. A. Razborov, "On the Distributional Complexity of Disjointness", *Proc. 17th ICALP*, pp. 249–253, 1990.

[Ra] R. Raz, "A Parallel Repetition Theorem", to appear in *SIAM Journal on Computing*. Preliminary version in *27th STOC*, pp. 447-456, 1995.

[Y] A. C.-C. Yao, "Some complexity questions related to distributive computing", *Proceedings of* 11^{th} *STOC*, pp. 209-213 (1979).

[Ya] A.C. Yao, "Theory and Application of Trapdoor Functions", in *23st FOCS*, pages 80–91, 1982.