Electronic Colloquium on Computational Complexity, Report No. 86 (2010)

On a Theorem of Razborov

Henning Wunderlich*

April 28, 2010

Abstract

In an unpublished Russian manuscript Razborov proved that a matrix family with high rigidity over a finite field would yield a language outside the polynomial hierarchy in communication complexity.

We present an alternative proof that strengthens the original result in several ways. In particular, we replace rigidity by the strictly weaker notion of toggle rigidity.

It turns out that Razborov's astounding result is actually a corollary of a slight generalization of Toda's First Theorem in communication complexity, and that matrix rigidity over a finite field is a lower-bound method for bounded-error modular communication complexity.

We also give evidence that Razborov's strategy is a promising one by presenting a protocol with few alternations for the inner product function mod two and by discussing problems possibly outside the communication complexity version of the polynomial hierarchy.

1. Introduction

In communication complexity (Kushilevitz & Nisan 1997) communication models are studied where several players want to cooperatively solve a problem. The resource under consideration is *communication*, i.e., the number of communicated bits. In general, the players have to communicate because the input is distributed among them. The arguably simplest communication model is Yao's model (Yao 1979) where two players Alice and Bob want to compute the value $f(x, y)$ of a function $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$. Here, Alice has $x \in \mathcal{X}$ and Bob has $y \in \mathcal{Y}$ and they may send each other messages (bits) according to a fixed protocol. Enriching this model with resources like randomization, guessing, or alternation leads to various variants, and studying the relative power of these models is a recurring theme in communication complexity.

One approach, that brings the relative power of the different modes of comunication into prominence, is a structural one, initiated by Babai *et al.* (1986), where communication complexity classes like $\mathbf{P}^{\mathrm{cc}}$, $\mathbf{NP}^{\mathrm{cc}}$, $\mathbf{MOD}_k\mathbf{P}^{\mathrm{cc}}$, $\mathbf{PP}^{\mathrm{cc}}$, $\mathbf{PH}^{\mathrm{cc}}$, $\mathbf{PSPACE}^{\mathrm{cc}}$, etc. are defined analogously to the ones in the Turing-machine setting. While much more is known about the inclusion relationships of these communication complexity classes than about such relationships between classical ones, it is still a challenging open problem whether or not the polynomial hierarchy $\mathbf{PH}^{\mathrm{cc}}$ is a proper subset of polynomial space $\mathbf{PSPACE}^{\mathrm{cc}}$.

In an unpublished Russian manuscript Razborov (1989) proposed a possible strategy to solve this problem, based on a remarkable theorem saying that a matrix family with high rigidity over a finite field would yield a language outside the polynomial hierarchy in communication complexity. His theorem refers to the concept of matrix rigidity, which measures how many entries in a matrix one has to change in order to reduce the rank of the matrix below a given bound. Rigidity was originally introduced by Valiant (1977) for proving circuit lower bounds. Today, it plays an important role in several branches of theoretical computer science (see e.g., Lokam 2009).

The purpose of this article is to present an alternative proof of Razborov's Theorem that strengthens the original result in several ways. Namely, instead of using Valiant's definition of rigidity, we define and use the weaker notion of *toggle rigidity*. The improvements to Razborov's result are thus twofold:

*Universität Ulm, Fakultät für Ingenieurwissenschaften und Informatik, Institut für Theoretische Informatik, Oberer Eselsberg, D-89069 Ulm, Email: henning.wunderlich@uni-ulm.de

ISSN 1433-8092

First of all, the definition of toggle rigidity involves an arbitrary probability distribution on the matrix entries enabling us to put high weight on hard parts of the matrix and low weight on easier parts. In contrast, Valiant's definition corresponds to the equiponderant case, where a fixed uniform distribution is used.

Secondly, toggle rigidity severely restricts the allowed changes for rank reduction. Here, only toggling values between 0 and 1 is allowed, while in the classical definition a matrix entry can be replaced with an arbitrary field element. We give an example showing that toggle rigidity is indeed strictly weaker than classical rigidity.

In addition, we remark that over the field of real numbers high lower bounds for restricted versions of rigidity have already been obtained by Lokam (2001). These results imply high lower bounds for toggle rigidity over the field of real numbers.

Hence, as the paper at hand provides toggle rigidity as a new tool, equipped with toggle rigidity, it might be easier to prove high lower bounds in the future.

As an English translation of Razborov's work is unfortunately not available, a comparison between the techniques used in his proof and the ones used here could not be made.

Toda (1991) proved two theorems that caused astonishment. First of all, he showed that in the Turing-machine setting the polynomial hierarchy is contained in a certain complexity class, $\mathrm{BP} \cdot \oplus\mathbf{P}$. We refer to this result as his *First Theorem*. His *Second Theorem* states that the latter class is contained in $\mathbf{P}(\mathbf{PP})$, the Turing closure of probabilistic polynomial time.

We show that these theorems also hold in communication complexity, the reason being that Razborov's result will turn out to be a corollary of a slight generalization of Toda's First Theorem in communication complexity. Consequently, matrix rigidity (over a finite field) is actually a lower-bound method for bounded-error modular communication complexity, a measure based on randomization and guessing where the players accept an input with bounded error according to a modular acceptance mode.

Furthermore, we give evidence that Razborov's strategy is a promising one by presenting a protocol with few alternations for the inner product function mod two, and by discussing candidate problems possibly outside the communication complexity version of the polynomial hierarchy. This will lead us into the field of sparse quasi-random graphs (Chung & Graham 2002).

Outline. In Section 2 below we fix notation and give basic definitions used throughout this article. To set the stage for the proof of our main result and to make this article more self-contained we survey parts of communication complexity in Section 3 and structural complexity in Section 4. The informed reader, on tiptoes with expectation, may safely skip these and turn right to Section 5, where Toda's Theorems are transferred to communication complexity, followed by Section 6 containing the main result, a strengthening of Razborov's Theorem. In the final Section 7 we discuss Razborov's strategy.

2. Preliminaries

We fix notation and give basic definitions used throughout this article.

We denote with $[n]$ the set $\{1, \ldots, n\}$ of the first n natural numbers. For a set S we write $\binom{S}{k}$ for the set of all subsets of S with cardinality k. For a real number r we denote with *floor* r, $\lfloor r \rfloor$, the largest integer not exceeding r, and with *ceiling* r, $\lceil r \rceil$, the smallest integer greater than or equal to r. The logarithm to the basis 2 is denoted with log. For a prime power q, we denote with $\mathbb{F}_q$ the finite field with q elements.

Occasionally, in order to avoid ugly case distinctions we use *Iverson's bracket* $[P]$ defined on predicates P, which evaluates to 1, if P is true, and to 0 otherwise.

We define matrices with arbitrary finite index sets for rows and columns. Accordingly, a matrix M over $\mathcal{Z}$ is just a map $M\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ for finite sets $\mathcal{X}$ and $\mathcal{Y}$. We write $M_{x,y}$ for M's entry in row $x \in \mathcal{X}$ and column $y \in \mathcal{Y}$.

We only work with the *binary alphabet* $\mathbb{B} \coloneqq \{\mathtt{0}, \mathtt{1}\}$. The length of a word $x \in \mathbb{B}^*$ is denoted by $|x|$. For two words $x, y \in \mathbb{B}^*$ the word y is a *prefix of* x, if there exists $z \in \mathbb{B}^*$ such that $x = yz$. A set $S \subseteq \mathbb{B}^*$ is *prefix-free* if for all distinct $x, y \in S$ we have that y is not a prefix of x. An example of a *prefix-free encoding* of x is $\overline{x} \coloneqq \mathtt{0}^{|x|}\mathtt{1}x$. In order to encode pairs of words $x, y \in \mathbb{B}^*$ we use the

pairing function $\langle x, y\rangle \coloneqq \overline{x}y$. For a mathematical object o contained in an at most countable set we denote with $\langle o\rangle$ a suitable prefix-free encoding of o.

Functions with range $\mathbb{B}$ are called *Boolean functions*.

For an excellent introduction to graph theory we refer the reader to Diestel (2005). Given a *graph* G we write $V(G)$ to denote its *nodes* (*vertices*) and $E(G)$ to denote its *edges*. As usual, the *adjacency matrix of* G, A^G, is defined by

$$A^G_{x,y} \coloneqq [\{x, y\} \in E(G)] \quad \text{, for all } x, y \in V(G) \ .$$

For subsets $X, Y \subseteq V(G)$ we define the *set of edges between* X *and* Y as

$$E_G(X, Y) \coloneqq \big\{\{x, y\} \in E(G) \,\big|\, x \in X, y \in Y\big\} \ .$$

Finally, we define

$$e_G(X, Y) \coloneqq |E_G(X, Y)| + |E_G(X \cap Y, X \cap Y)|$$

as the number of edges with one endpoint in X and the other one in Y. If an edge belongs to the intersection $X \cap Y$, then it is counted twice in $e_G(X, Y)$.

3. Communication Complexity

In this section we give a comparably short introduction to parts of communication complexity. In particular, we describe Yao's model and state some important basic results that are used later. We refer the reader to Kushilevitz & Nisan (1997) for an excellent introduction to the field of communication complexity.

3.1. Deterministic protocols. In his seminal work, Yao (1979) introduced a simple communication model. In *Yao's model*, there are two players (parties) Alice and Bob with unlimited computational power, who want to cooperatively compute a function $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$, where $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are finite sets. Both have complete information about f but receive only parts of the input. Alice is given $x \in \mathcal{X}$, Bob is given $y \in \mathcal{Y}$, and they exchange messages (bits) in order to compute $f(x, y)$. The players communicate according to a fixed *(deterministic) protocol* Π (over domain $\mathcal{X} \times \mathcal{Y}$ with range $\mathcal{Z}$) that specifies how the communication is carried out. At each stage of the computation, the protocol must determine whether the run has terminated. In this case, it must specify the output value. Otherwise, it must specify the player who speaks next. Each message sent by a player must solely depend on the player's input and the messages communicated so far, because this is the only "information" the player has about the inputs.

There do exist different formalizations of the notion *protocol* depending on the applications the protocol designer has in mind. Since we want to analyze protocols, we formalize them via binary trees. See e.g., Hromkovic (2000) for a definition that is equivalent but different from ours. It is known that this combinatorial view on protocols has many advantages. In particular, it allows us to prove high lower complexity bounds in this model in contrast to many other computation models. A formal definition of protocols is given below.

Definition 3.1 (Deterministic protocol). *A* deterministic protocol Π (over domain $\mathcal{X} \times \mathcal{Y}$ with range $\mathcal{Z}$) *is a labeled directed finite binary tree (*protocol tree*). Each leaf ℓ is labeled by an output value $z_\ell \in \mathcal{Z}$. If v is an inner node of Π, then it has a left and a right child v_0 and v_1, respectively, and the arc from v to v_b is labeled by $b \in \mathbb{B}$. The node v is labeled either by a function $a_v\colon \mathcal{X} \to \mathbb{B}$ or by a function $b_v\colon \mathcal{Y} \to \mathbb{B}$. The root of the protocol tree of Π is denoted by* $\mathrm{root}(\Pi)$*, the set of nodes by* V_Π*, and the set of leaves by* L_Π*, respectively.*

Let Alice have $x \in \mathcal{X}$, and let Bob have $y \in \mathcal{Y}$. When they communicate according to a protocol Π, they start at $\mathrm{root}(\Pi)$. The nodes of the protocol tree of Π can be interpreted as (common) "computation states". If both players are in such a state v during the run of the protocol, then one of two things can happen: If v is a leaf, the communication ends and both players know the output value z_v. If v is an inner node, we say that *Alice speaks*, if v is labeled by a_v. In this case,

Alice sends the bit $b := a_v(x)$ and both players change their computation state to v_b, analogously, if *Bob speaks*.

We say that an input (x, y) *reaches* a node v of Π if the players arrive at v when running the protocol on the respective input. We denote by $R_v \subseteq \mathcal{X} \times \mathcal{Y}$ the set of inputs reaching v.

The concatenation of the messages communicated during a run of a protocol Π on input (x, y) is called *transcript* and is denoted by $\Pi(x, y)$.

For each input (x, y) the execution of a deterministic protocol Π leads to exactly one output value. This defines a function f_Π. We say that Π *computes a function* $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$, if $f = f_\Pi$.

Having now introduced a computation model and a resource, it is time to define corresponding cost and complexity measures.

Definition 3.2 (Worst-case deterministic communication cost). *Let Π be a deterministic protocol over domain $\mathcal{X} \times \mathcal{Y}$. The* worst-case deterministic communication cost of Π, $\mathrm{D}(\Pi)$, *is defined as*

$$\mathrm{D}(\Pi) := \max_{(x,y)\in\mathcal{X}\times\mathcal{Y}} |\Pi(x, y)| \ .$$

Definition 3.3 (Closeness). *Let f and g be two functions defined on the same domain D, let μ be a probability distribution on D, and let $\epsilon \geq 0$ be a real number. The functions f and g are* (μ, ϵ)-close, *if $\mu(f \neq g) := \mu\{z \in D \mid f(z) \neq g(z)\} \leq \epsilon$.*

Definition 3.4 (Distributional error). *Let $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function, let μ be a probability distribution on $\mathcal{X} \times \mathcal{Y}$, let $\epsilon \geq 0$ be a real number, and let Π be a deterministic protocol over domain $\mathcal{X} \times \mathcal{Y}$ with range $\mathcal{Z}$. We say that Π* computes *f* with (μ, ϵ)-distributional error, *if f and f_Π are (μ, ϵ)-close.*

Definition 3.5 (Worst-case deterministic communication complexity). *Let $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function, let μ be a probability distribution on $\mathcal{X} \times \mathcal{Y}$, and let $\epsilon > 0$ be a real number.*

(i) *The* worst-case deterministic communication complexity of f, $\mathrm{D}(f)$, *is the minimum worst-case deterministic communication cost of a deterministic protocol computing f.*

(ii) *The* worst-case (μ, ϵ)-distributional deterministic communication complexity of f, $\mathrm{D}^\mu_\epsilon(f)$, *is the minimum worst-case deterministic communication cost of a deterministic protocol computing f with (μ, ϵ)-distributional error.*

The following property should be obvious:

Observation 3.6. *Let $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function, let μ be a probability distribution on $\mathcal{X} \times \mathcal{Y}$, and let $\epsilon > 0$ be a real number. Then we have*

$$\mathrm{D}^\mu_\epsilon(f) = \min\{\mathrm{D}(\tilde{f}) \mid \tilde{f} \text{ and } f \text{ are } (\mu, \epsilon)\text{-close}\} \ .$$

3.2. Randomized protocols. We refer the reader to Motwani & Raghavan (1995) for an excellent introduction to the exciting field of randomized algorithms. Often, randomized algorithms are both simpler and faster than every known deterministic algorithm solving the same problem. The same applies if one adds randomness to Yao's deterministic two-player model. Randomized communication complexity was also defined in the seminal paper of Yao (1979). In the randomized model, the players are allowed to "toss coins" during the execution of a protocol, and the messages they send each other may also depend on the outcomes of the coin tosses. Consequently, the messages, the transcript and the computed function become random variables. We distinguish between two types of randomized protocols, namely "public-coin" and "private-coin" ones. In a public-coin protocol, Alice and Bob share a common public coin whose outcomes are known to both players. In a private-coin protocol, each player has its own random coin to flip. We want to emphasize that Alice cannot see Bob's coin flips and vice versa. While the latter model seems more realistic

than the public-coin model, it was shown by Newman (1991) that the models are essentially the same. We note that a randomized protocol can be interpreted as a probability distribution over deterministic protocols.

Definition 3.7 (Randomized protocol).

(i) *A* randomized public-coin protocol Π (over domain $\mathcal{X} \times \mathcal{Y}$ with range $\mathcal{Z}$) *is defined as a pair* $\Pi \coloneqq (\Pi', \mathrm{C})$, *where* C *is a random variable over a finite set* $\mathcal{C}$, *and* Π' *is a deterministic protocol over domain* $(\mathcal{X} \times \mathcal{C}) \times (\mathcal{Y} \times \mathcal{C})$ *with range* $\mathcal{Z}$. *The random variable* C *is called the* common coin.

(ii) *A* randomized private-coin protocol Π (over domain $\mathcal{X} \times \mathcal{Y}$ with range $\mathcal{Z}$) *is defined as a triple* $\Pi \coloneqq (\Pi', \mathrm{A}, \mathrm{B})$, *where* A *and* B *are random variables over finite sets* $\mathcal{A}$ *and* $\mathcal{B}$, *respectively, and* Π' *is a deterministic protocol over domain* $(\mathcal{X} \times \mathcal{A}) \times (\mathcal{Y} \times \mathcal{B})$ *with range* $\mathcal{Z}$. *The random variable* A *is called* Alice's coin, B *is called* Bob's coin.

Definition 3.8 (Computed function). *Let* Π *be a randomized protocol over domain* $\mathcal{X} \times \mathcal{Y}$ *with range* $\mathcal{Z}$. *We define the* function f_Π computed by Π *as the random variable* $f_\Pi \coloneqq ((x, y) \mapsto f_{\Pi'}(x, \mathrm{C}, y, \mathrm{C}))$ *over* $\mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$, *if* $\Pi = (\Pi', \mathrm{C})$ *is a randomized public-coin protocol, and as* $f_\Pi \coloneqq ((x, y) \mapsto f_{\Pi'}(x, \mathrm{A}, y, \mathrm{B}))$, *if* $\Pi = (\Pi', \mathrm{A}, \mathrm{B})$ *is a randomized private-coin protocol.*

Definition 3.9 (Transcript). *Let* Π *be a randomized protocol. Given an input* (x, y), *we define the* transcript $\Pi(x, y)$ *as the random variable* $\Pi(x, y) \coloneqq \Pi'(x, \mathrm{C}, y, \mathrm{C})$, *if* $\Pi = (\Pi', \mathrm{C})$ *is a randomized public-coin protocol, and as* $\Pi(x, y) \coloneqq \Pi'(x, \mathrm{A}, y, \mathrm{B})$, *if* $\Pi = (\Pi', \mathrm{A}, \mathrm{B})$ *is a randomized private-coin protocol.*

Definition 3.10 (Worst-case randomized communication cost). *Let* Π *be a randomized protocol over domain* $\mathcal{X} \times \mathcal{Y}$. *The* worst-case randomized communication cost of Π, $\mathrm{R}(\Pi)$, *is defined as*

$$\mathrm{R}(\Pi) \coloneqq \max_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \max_{c \in \mathcal{C}} |\Pi'(x, c, y, c)| \ ,$$

if $\Pi = (\Pi', \mathrm{C})$ *is a randomized public-coin protocol with common coin* C *defined over the finite set* $\mathcal{C}$, *and as*

$$\mathrm{R}(\Pi) \coloneqq \max_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \max_{a \in \mathcal{A}, b \in \mathcal{B}} |\Pi'(x, a, y, b)| \ ,$$

if $\Pi = (\Pi', \mathrm{A}, \mathrm{B})$ *is a randomized private-coin protocol with Alice's coin* A *defined over* $\mathcal{A}$ *and Bob's coin* B *defined over* $\mathcal{B}$ *for finite sets* $\mathcal{A}$ *and* $\mathcal{B}$, *respectively.*

Definition 3.11 (error ϵ). *Let* $f \colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ *be a function, and let* $\epsilon \geq 0$ *be a real number. A randomized protocol* Π *over domain* $\mathcal{X} \times \mathcal{Y}$ *with range* $\mathcal{Z}$ computes f with error ϵ, *if for all* $(x, y) \in \mathcal{X} \times \mathcal{Y}$ *we have*

$$\Pr\left[f_\Pi(x, y) \neq f(x, y)\right] \leq \epsilon \ .$$

In the literature, a randomized protocol that computes a function with *bounded error*, i.e., with an error bounded away from one half by a constant, is sometimes called *Monte Carlo protocol*.

Definition 3.12 (Worst-case randomized communication complexity). *Let* f *be a function, and let* $\epsilon > 0$ *be a real number. The* worst-case randomized public-coin ϵ-error communication complexity of f, $\mathrm{R}^{\mathrm{pub}}_\epsilon(f)$, *is defined as the minimum worst-case randomized communication cost of a randomized public-coin protocol computing* f *with error* ϵ.

For randomized private-coin protocols one can define the complexity measure $\mathrm{R}^{\mathrm{priv}}_\epsilon(f)$ analogously to the one for public-coin protocols.

As mentioned above, public- and private-coin complexities are not far apart. For a proof of the following result, see e.g., (Kushilevitz & Nisan 1997, p. 33, Theorem 3.14; p. 34, Exercise 3.15).

FACT 3.13 (Newman). *Let $f\colon \mathbb{B}^n \times \mathbb{B}^n \to \mathbb{B}$ be a Boolean function. For every $\epsilon > 0$ and $\delta > 0$ we have*

$$\mathrm{R}^{\mathrm{priv}}_{\epsilon+\delta}(f) \leq \mathrm{R}^{\mathrm{pub}}_{\epsilon}(f) + \mathcal{O}(\log n + \log(1/\delta)) \quad .$$

Newman's result also shows that the size of the probability space can be restricted to $2^{\mathrm{polylog}(n)}$ for inputs of size n, if one allows a small increase in communication cost and error probability.

An important technique in the theory of randomized algorithms is *probability amplification*, i.e., one reduces the error probability of a randomized algorithm to an arbitrarily small constant by running the algorithm on the same input several times with independent coin tosses and then taking the majority vote of the outcomes. This can be done for randomized protocols, too.

The following fact can be found in (Köbler *et al.* 1993, p. 70, Lemma 2.14). We make use of this Chernoff-like result in Theorem 6.18.

FACT 3.14 (Probability amplification). *Let E be an event that occurs with probability $\frac{1}{2} + \epsilon$, $0 < \epsilon \leq \frac{1}{2}$. Then E occurs within t independent trials (t odd) at least $t/2$ times with probability at least*

$$1 - \frac{1}{2} \cdot \left(1 - 4 \cdot \epsilon^2\right)^{t/2} \quad .$$

Many lower-bound methods for randomized communication complexity are based on the following simple application of *Yao's Minimax-principle* (see e.g., Kushilevitz & Nisan 1997, p. 36, Theorem 3.20 or Yao 1983) relating randomized and distributional complexity.

FACT 3.15 (Yao). *For every Boolean function f and every $\epsilon > 0$ we have*

$$\mathrm{R}^{\mathrm{pub}}_{\epsilon}(f) = \max_{\mu} \mathrm{D}^{\mu}_{\epsilon}(f) \quad .$$

3.3. Counting protocols. Analogously to the Turing-machine model, one can add the power of counting to Yao's model. The concept of counting means that the players can make nondeterministic guesses during a computation. As on different guesses the output values may be different, one has to specify an *acceptance mode*, a predicate that tells us which inputs are considered to be accepted based on the number of accepting and rejecting computations.

There are several possibilities to define counting protocols (via proof systems, covers, etc.). We choose the following variant.

DEFINITION 3.16 (Counting protocol). *A* counting protocol (over domain $\mathcal{X} \times \mathcal{Y}$) *is a deterministic protocol over domain $(\mathcal{X} \times \mathbb{B}^{g_A}) \times (\mathcal{Y} \times \mathbb{B}^{g_B})$ with range $\mathbb{B}$, where $g_A, g_B \geq 0$ are natural numbers denoting the lengths of the* guess strings.

Note that one could define counting protocols using abstract guess sets instead of $\mathbb{B}^{g_A}$ and $\mathbb{B}^{g_B}$, respectively. We do not do this here, because the above definition corresponds more closely with the definition of complexity class operators in Section 5. In addition, it is important to bound the number of guess bits used. If we had used an abstract guess set, we would have to encode the guess elements as strings and then we were back in $\mathbb{B}^{g_*}$.

As in the case of randomized protocols, we should distinguish between "public-guess" and "private-guess" counting protocols. A *public-guess counting protocol* is only defined for inputs $(\langle x, w\rangle, \langle y, w\rangle)$, where w is the public-guess string of Alice and Bob, while a *private-guess counting protocol* is defined for all inputs $(\langle x, w_A\rangle, \langle y, w_B\rangle)$. Here, w_A denotes Alice's guess, while w_B denotes Bob's.

In the sequel, we refrain from distinguishing between these two variants of counting protocols, because we only consider counting protocols with guess strings bounded polylogarithmically in the length of the "real" input (x, y). In this case, these variants are essentially equivalent.

Definition 3.17. *For a counting protocol Π we denote with*

$$\begin{aligned} \mathrm{acc}_\Pi(x,y) &:= \left|\{\Pi((x,w_A),(y,w_B)) \mid f_\Pi((x,w_A),(y,w_B)) = 1\}\right| \text{ and} \\ \mathrm{rej}_\Pi(x,y) &:= \left|\{\Pi((x,w_A),(y,w_B)) \mid f_\Pi((x,w_A),(y,w_B)) = 0\}\right| \end{aligned}$$

the number of accepting and rejecting transcripts *of Π on input (x,y), respectively.*

Definition 3.18 (Computed function). *Given a counting protocol Π and an acceptance mode Ξ, the* function computed by Π in acceptance mode Ξ, f_Π^Ξ, *is defined as*

$$f_\Pi^\Xi(x,y) := [\Xi(\mathrm{acc}_\Pi(x,y), \mathrm{rej}_\Pi(x,y))] \ .$$

We say that Π computes f in acceptance mode Ξ, *if $f_\Pi^\Xi = f$.*

We list the most prominent acceptance modes:

$$\begin{aligned} \mathrm{N}^1(acc, rej) &:= (acc > 0) \ , \\ \mathrm{N}^0(acc, rej) &:= (rej = 0) \ , \\ \mathrm{PP}(acc, rej) &:= (acc > rej) \ , \\ \mathrm{MOD}_k(acc, rej) &:= (acc \bmod k = 1) \ , k \geq 2, \\ \oplus\mathrm{P}(acc, rej) &:= \mathrm{MOD}_2(acc, rej) \ . \end{aligned}$$

N^1 is the *nondeterministic*, N^0 the *co-nondeterministic*, PP the *probabilistic*, MOD_k the *mod-k*, and $\oplus$P the *parity acceptance mode.*

Definition 3.19 (Worst-case communication cost). *We define the* worst-case communication cost *of a (public- or private-guess) counting protocol as the worst-case deterministic communication cost, when viewed as a deterministic protocol, plus the lengths of the guess strings.*

Definition 3.20 (Counting complexities).

(i) *The* nondeterministic communication complexity of f, $\mathrm{N}^1(f)$, *is defined as the minimum worst-case communication cost of a counting protocol computing f in nondeterministic acceptance mode.*

(ii) *The* co-nondeterministic communication complexity of f, $\mathrm{N}^0(f)$, *is defined as the minimum worst-case communication cost of a counting protocol computing f in co-nondeterministic acceptance mode.*

(iii) *The* probabilistic communication complexity of f, $\mathrm{PP}(f)$, *is defined as the minimum worst-case communication cost of a counting protocol computing f in probabilistic acceptance mode.*

(iv) *Let $k \geq 2$ be a natural number. The* mod-k communication complexity of f, $\mathrm{MOD}_k(f)$, *is defined as the minimum worst-case communication cost of a counting protocol computing f in mod-k acceptance mode.*

(v) *The* parity communication complexity of f, $\oplus\mathrm{P}(f)$, *is defined as* $\mathrm{MOD}_2(f)$.

Analogously to distributional deterministic communication complexity, we define a distributional mod-k communication complexity. For this measure one can prove an analogous Minimax-statement (Observation 3.29) for the bounded-error mod-k communication complexity (Definition 3.27) as for bounded-error randomized communication complexity (Fact 3.15).

Definition 3.21 (Computed function). *Let $f\colon \mathcal{X}\times\mathcal{Y}\to\mathbb{B}$ be a Boolean function, let $k\geq 2$ be a natural number, let μ be a probability distribution on $\mathcal{X}\times\mathcal{Y}$, and let $\epsilon>0$ be a real number. We say that a counting protocol Π* computes *f* in mod-k acceptance mode with (μ,ϵ)-distributional error, *if f and $f_\Pi^{\mathrm{MOD}_k}$ are (μ,ϵ)-close.*

Definition 3.22 (Distributional modular communication complexity).
*The (μ,ϵ)-*distributional mod-k communication complexity *of f, $\mathrm{MOD}^{\mu}_{k,\epsilon}(f)$, is defined as the minimum worst-case communication cost of a counting protocol computing f in mod-k acceptance mode with (μ,ϵ)-distributional error.*

Similar to Observation 3.6 we have

Observation 3.23. *Let f be a Boolean function, let $k\geq 2$ be a natural number, and let $\epsilon>0$ be a real number. Then we have*

$$\mathrm{MOD}^{\mu}_{k,\epsilon}(f)=\min\{\mathrm{MOD}_k(\tilde{f})\mid \tilde{f}\text{ and }f\text{ are }(\mu,\epsilon)\text{-close}\}\ .$$

We will see later that interesting effects occur when one combines counting with randomization.

Definition 3.24 (Randomized counting protocol). *A* (public-coin) randomized counting protocol *Π (over domain $\mathcal{X}\times\mathcal{Y}$) is a probability distribution over counting protocols, i.e., $\Pi\coloneqq(\{\Pi_a\}_{a\in A},\alpha)$, where α is a random variable with values in a set A, and each Π_a, $a\in A$, is a counting protocol over domain $\mathcal{X}\times\mathcal{Y}$.*

Definition 3.25 (Computed function). *Let $f\colon \mathcal{X}\times\mathcal{Y}\to\mathbb{B}$ be a Boolean function, let $k\geq 2$ be a natural number, and let $\epsilon>0$ be a real number. A randomized counting protocol Π* computes *f* in mod-k acceptance mode with error ϵ, *if for all $(x,y)\in\mathcal{X}\times\mathcal{Y}$ we have*

$$\Pr_{\alpha}\left[f^{\mathrm{MOD}_k}_{\Pi_\alpha}(x,y)\neq f(x,y)\right]\leq\epsilon\ .$$

In this case, we call Π an ϵ-error mod-k protocol for *f.*

Definition 3.26 (Communication cost). *The* communication cost of a randomized counting protocol *$\Pi\coloneqq(\{\Pi_a\}_{a\in A},\alpha)$ is defined as the maximum communication cost of the counting protocols Π_a that have non-zero weight under the probability distribution on A induced by α.*

Definition 3.27 (Bounded-error modular comunication complexity).
Let f be a Boolean function, let $k\geq 2$ be a natural number, and let $\epsilon>0$ be a real number. The (public-coin) ϵ-error mod-k communication complexity *of f, $\mathrm{BP}\cdot\mathrm{MOD}^{\mathrm{pub}}_{k,\epsilon}(f)$, is defined as the minimum communication cost of an ϵ-error mod-k protocol for f.*

Observation 3.28. *For every Boolean function f, every natural number $k\geq 2$ and every real number $\epsilon>0$ the ϵ-error mod-k communication complexity of f can be upper-bounded by*

$$\mathrm{BP}\cdot\mathrm{MOD}^{\mathrm{pub}}_{k,\epsilon}(f)\leq\min\{\mathrm{D}(f),\mathrm{MOD}_k(f),\mathrm{R}^{\mathrm{pub}}_{\epsilon}(f)\}\ .$$

An adaptation of Fact 3.15 yields

Observation 3.29. *For every Boolean function f, every natural number $k\geq 2$ and every real number $\epsilon>0$ we have*

$$\mathrm{BP}\cdot\mathrm{MOD}^{\mathrm{pub}}_{k,\epsilon}(f)=\max_{\mu}\mathrm{MOD}^{\mu}_{k,\epsilon}(f)\ .$$

3.4. Alternating protocols. The concept of alternation was originally defined for the Turing-machine model as a generalization of nondeterminism. Alternation can be translated to Yao's model. This was done in (Babai *et al.* 1986, p. 339) by defining players East, West, North, and South. We give an equivalent definition of alternating protocols.

In an alternating protocol the players may guess bits. Each state of the protocol is either *rejecting* (0), *accepting* (1), *existential* ($\exists$), or *universal* ($\forall$).

If a player guesses a bit in an existential state, then this guess is called existential; universal guesses are defined similarly.

A formal definition is given below.

DEFINITION 3.30 (Alternating protocol). *An* alternating protocol (over domain $\mathcal{X}\times\mathcal{Y}$) *is a labeled binary tree, where leaves ℓ are labeled by $z_\ell \in \mathbb{B}$ and inner nodes v are labeled by $\mathrm{Q}_v \in \{\exists,\forall\}$ and by functions $a_v\colon \mathcal{X}\to\{0,1,*\}$ or $b_v\colon \mathcal{Y}\to\{0,1,*\}$, respectively. Each inner node v has two children v_0 and v_1.*

If in a run of an alternating protocol the players are in common state v labeled by a_v, we say that Alice *guesses universally*, if $a_v(x) = *$ and $\mathrm{Q}_v = \forall$, and that she *guesses existentially*, if $a_v(x) = *$ and $\mathrm{Q}_v = \exists$, analogously, if it is Bob's turn.

DEFINITION 3.31 (Computed function). *Given an alternating protocol Π over domain $\mathcal{X}\times\mathcal{Y}$, the* function computed by Π, $f_\Pi\colon \mathcal{X}\times\mathcal{Y}\to\mathbb{B}$, *is defined as follows. We associate with each node v of Π a function $f_v\colon \mathcal{X}\times\mathcal{Y}\to\mathbb{B}$. For a leaf ℓ we define $f_\ell(x,y) \coloneqq z_\ell$. For an inner node v labeled by a_v we define*

$$f_v(x,y) \coloneqq \begin{cases} f_{v_c}(x,y) & \text{, if } c \coloneqq a_v(x) \in \mathbb{B} \ , \\ [\mathrm{Q}_v c \in \mathbb{B}\colon f_{v_c}(x,y) = 1] & \text{, if } a_v(x) = * \ , \end{cases}$$

similarly, for inner nodes labeled by b_v. Finally, we define f_Π as the function computed at the root.

An alternating protocol Π computes a function f, *if $f_\Pi = f$.*

DEFINITION 3.32 (Alternating communication cost). *For an alternating protocol Π, the* worst-case alternating communication cost of Π, $\mathrm{A}(\Pi)$, *is defined as the maximum length of a path from the root to a leaf in the protocol tree of Π.*

DEFINITION 3.33 (Alternating communication complexity). *The* worst-case alternating communication complexity of f, $\mathrm{A}(f)$, *is defined as the minimum worst-case alternating communication cost of an alternating protocol computing f.*

We say that an alternating protocol has k alternations if starting in an existential state the maximum number of alternations between existential and universal states on every path from the root to a leaf of the protocol tree is bounded by k. With $\mathrm{A}^k(f)$ we denote the restriction of $\mathrm{A}(f)$ to alternating protocols with k alternations.

3.5. Lower bounds. The complexity measures introduced in the preceding subsections are hard to calculate, because in general it is extremely expensive to enumerate all protocols computing a function in order to find one with minimal communication cost. This is why for each complexity measure M one tries to find combinatorial measures $M' \le M$ that are easily computable and (hopefully) close to M.

For worst-case (co-)nondeterministic communication complexity such a combinatorial measure is the *rectangle-size method*, and, as a special case, the *fooling-set method*. For definitions, applications and proofs we refer the reader to (Kushilevitz & Nisan 1997, Sections 1.3 and 2.4).

Many lower bound methods have been developed for randomized communication complexity. The most prominent ones are the *discrepancy method* (Kushilevitz & Nisan 1997, p. 38, Section 3.5), the *Fourier method* of Raz (1995), the *ϵ-monochromatic rectangle-size method* (or *corruption*

method, see e.g., Beame *et al.* 2006), and the *factorization-norm method* of Linial & Shraibman (2007). The latter work investigates an approximate γ_2-norm, γ_2^α, and shows that most known lower bounds for bounded-error randomized communication complexity are actually lower bounds for bounded-error quantum communication complexity. Klauck (2001) gave a characterization of the PP communication complexity via the discrepancy method. Alternative characterizations via *margin complexity* and the equivalent γ_2^∞*-measure* were obtained in Linial & Shraibman (2009a).

The most important method for worst-case deterministic communication complexity, the *rank method*, was introduced in Mehlhorn & Schmidt (1982). The basic idea is to consider a function $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ as a matrix M^f of dimensions $|\mathcal{X}| \times |\mathcal{Y}|$. The rows of M^f are indexed by the elements of $\mathcal{X}$ and the columns are indexed by the elements of $\mathcal{Y}$. The (x,y)-entry $M^f_{x,y}$ of M^f is simply defined as $f(x,y)$.

A basic fact is that a deterministic protocol computing f partitions the input space $\mathcal{X} \times \mathcal{Y}$ into monochromatic (i.e., f-constant) rectangles $R_1, \ldots, R_t$, where a *(combinatorial) rectangle* is a set $R = A \times B$, $A \subseteq \mathcal{X}$, $B \subseteq \mathcal{Y}$. As the rectangles form a partition of the input space, the matrix M^f can be written as a sum $\sum z_i \cdot M_i$ of t rank-one matrices $M_1, \ldots, M_t$, where $(M_i)_{x,y} := [(x,y) \in R_i]$, i.e., M_i has value one on the rectangle inputs and zero otherwise. Thus, $\mathbb{F}\text{-rank}(M^f) \leq t$ for every suitable field $\mathbb{F}$ with $\mathcal{Z} \subseteq \mathbb{F}$.

The above considerations yield

Fact 3.34 (Mehlhorn & Schmidt). *For every real-valued function f we have*

$$\log \mathbb{R}\text{-rank}\left(M^f\right) \leq \mathrm{D}(f) \ .$$

It is still open whether the rank method is polynomially tight for Boolean functions. (For arbitrary functions one can show exponential gaps.)

Open Question 3.35 (Logarithmic-rank conjecture). *Do we have*

$$\mathrm{D}(f) \leq \left(\log \mathbb{R}\text{-rank}\left(M^f\right)\right)^{\mathcal{O}(1)}$$

for every Boolean function f?

Non-constant gaps have been shown in Raz & Spieker (1993) and Nisan & Wigderson (1995).

For primes q, the logarithmic-rank conjecture for mod-q communication complexity was proved in (Damm *et al.* 2004, Proposition 5.3). Using a different characterization via intersection graphs the same result had been established earlier implicitly by Pudlák & Rödl (1994) for $q = 2$.

Fact 3.36. *Let f be a Boolean function, and let q be a prime. Then we have*

$$\frac{\log \mathbb{F}_q\text{-rank}\left(M^f\right)}{q-1} \leq \mathrm{MOD}_q(f) \leq \log \mathbb{F}_q\text{-rank}\left(M^f\right) + \mathcal{O}(1) \ .$$

We will use this important fact twice: first of all, in the derivation of a lower bound-method for the bounded-error modular communication complexity in Section 6, and secondly, when we consider the modular communication complexity of problems based on quasi-random graph families in Section 7.

4. Structural Communication Complexity

The field of *structural complexity theory* is so broad and rich that we do not make any attempt to give an overview of this field or at least to list the most important results. As an excuse, we would like to cite (Hemaspaandra & Ogihara 2002, p. 263), where they say that

> "it would be impossible to define or collect the field's most important theorems"

in their appendix (Appendix A Rogues' Gallery of Complexity Classes) that has a size of 40 pages. Instead, for thorough introductions to (parts of) structural complexity we refer the reader to the excellent monographs of Balcázar *et al.* (1990, 1995); Du & Ko (2000); Köbler *et al.* (1993); Schöning (1986). Good surveys on a variety of topics in this field can be found in Selman (1988); Selman & Hemaspaandra (1997), especially on counting complexity in Schöning (1988) and Fortnow (1997), respectively.

To a complexity theorist, structure is meaning. In order to understand computational resources and their relationships one groups families of problems into *complexity classes* that can be solved with a certain computational power stemming from the resources one has added to the model under consideration.

Classically, in structural complexity theory one considers the Turing-machine model and models of Boolean and algebraic circuits. Depending on whether or not one admits the resources randomization, counting or alternation to the Turing machine model, one obtains "standard" complexity classes like *deterministic polynomial time*, **P**, *nondeterministic polynomial time*, **NP**, *conondeterministic polynomial time*, **coNP**, *bounded error probabilistic polynomial time*, **BPP**, *unbounded error probabilistic polynomial time*, **PP**, *parity polynomial time* $\oplus\mathbf{P}$, (*Parity-P* for short), *the polynomial-time hierarchy*, $\mathbf{PH} = \bigcup_{k\geq 0} \mathbf{\Sigma}_k^p$, and *polynomial space*, **PSPACE**. By their very definition one obtains a set of standard inclusions (see Table 4.1).

P	$\subseteq$	**BPP**	$\subseteq$	**PP**	$\subseteq$	**PSPACE** ,
P	$\subseteq$	**NP**, **coNP**	$\subseteq$	**PH**	$\subseteq$	**PSPACE** ,
P	$\subseteq$	**PP**, $\oplus$**P**	$\subseteq$	**PSPACE** .		

Table 4.1: Standard inclusions

Many complexity classes are (or can be formulated as) *counting classes*. These classes are based on polynomial-time Turing machines that can guess bits together with a fixed *acceptance mode* Ξ. Let $\mathrm{acc}_T(x)$ and $\mathrm{rej}_T(x)$ denote the number of accepting and rejecting computations of a Turing machine T on input x, respectively. Then an input is accepted by T in acceptance mode Ξ, if $\Xi(\mathrm{acc}_T(x), \mathrm{rej}_T(x))$ is true. A prominent example of such a counting class is $\oplus\mathbf{P}$, defined by Papadimitriou & Zachos (1983), where it was shown that $\oplus\mathbf{P}(\oplus\mathbf{P}) = \oplus\mathbf{P}$. Here, the acceptance mode gives "true", if the number of accepting computations is odd. The class $\#\mathbf{P}$ contains all functions $\mathrm{acc}_T(x)$, where T is a polynomial-time Turing machine that can guess bits.

The classes **PH** and **PSPACE** can be defined via the concept of alternation: problems in **PH** are decidable by polynomial-time alternating Turing machines with a constant number of alternations, problems in **PSPACE** with an efficient number (polynomial in the input size). An alternating Turing machine can guess bits universally and existentially. An input is accepted, if all successor configurations of a universal guess are accepting, and if for every existential guess there exists an accepting successor configuration.

One can define operators on complexity classes, e.g., the useful BP-operator, which was defined by Schöning (1989). Using the BP-operator and a relativized version of the so-called *Valiant-Vazirani Lemma* (Valiant & Vazirani 1986), Toda was able to prove his celebrated theorems (Toda 1990, p. 38, Corollary 3.2.8(2) and Theorem 3.2.7) establishing the inclusions

$$\mathbf{PH} \subseteq \mathrm{BP} \cdot \oplus\mathbf{P} \subseteq \mathbf{P}(\#\mathbf{P}) = \mathbf{P}(\mathbf{PP}) \ .$$

They tell us that counting (mod 2) plus the use of a random source is at least as powerful as the whole polynomial-time hierarchy **PH**, and that the same applies to the closure of **PP** under polynomial-time Turing reductions. See also Schöning (1991) for a proof sketch diaphanously presenting the main ideas.

Research in *structural communication complexity* started with the work of Babai *et al.* (1986), where some analogies between the Turing-machine classes mentioned above and corresponding communication complexity classes $\mathbf{P}^{\mathrm{cc}}$, $\mathbf{NP}^{\mathrm{cc}}$, $\mathbf{PP}^{\mathrm{cc}}$, $\oplus\mathbf{P}^{\mathrm{cc}}$, $\mathbf{PSPACE}^{\mathrm{cc}}$, $\mathbf{PH}^{\mathrm{cc}} = \bigcup_{k\geq 0} \mathbf{\Sigma}_k^{\mathrm{cc}}$, etc. were shown. Interestingly, while (almost) nothing is known about the standard classes in the

$$\begin{array}{lllllllll} \mathbf{P} & \subsetneq (?) & \mathbf{BPP} & \subsetneq (?) & \mathbf{PP} & \subsetneq (?) & \mathbf{PSPACE} & , \\ \mathbf{P} & \subsetneq (?) & \mathbf{NP}, \mathbf{coNP} & \subsetneq (?) & \mathbf{PH} & \subsetneq (?) & \mathbf{PSPACE} & , \\ \mathbf{P} & \subsetneq (?) & \mathbf{PP}, \oplus\mathbf{P} & \subsetneq (?) & \mathbf{PSPACE} & , & & \end{array}$$

and what about the pairs **NP** vs. **coNP**(?), **NP** vs. **PP**(?), **NP** vs. $\oplus$**P**(?), or **PP** vs. $\oplus$**P**(?), ...

Table 4.2: Unknown inclusion relationships

Turing-machine model (see Table 4.2), almost everything is known about the inclusion relationships between the respective communication complexity classes (see Table 4.3). One of the few exceptions is the long-standing open problem, whether or not the polynomial hierarchy is strictly contained in polynomial space. Besides Razborov, several authors proposed interesting strategies to tackle this problem: one is by Sherstov (2008b), who proved that showing a suitable upper bound on the *statistical query dimension* (a notion from learning theory) of the circuit class $\mathbf{AC}^0$ would separate the respective communication complexity classes. Another one is by Lokam (2001), who showed that high lower bounds on *weak* forms of rigidity over the field of *real* numbers would yield the desired separation. Adapting Lokam's proof by replacing rank with the γ_2-norm, Linial & Shraibman (2009a) were able to replace Lokam's weak rigidity with *mc-rigidity*, an approximate version of margin complexity. The latter result is not surprising when compared to Toda's Second Theorem, because margin complexity characterizes the class $\mathbf{PP}^{\mathrm{cc}}$, and an approximate version thus corresponds to "something" like the class $\mathrm{BP} \cdot \mathbf{PP}^{\mathrm{cc}}$.

$$\begin{array}{lllllllll} \mathbf{P}^{\mathrm{cc}} & \subsetneq & \mathbf{BPP}^{\mathrm{cc}} & \subsetneq & \mathbf{PP}^{\mathrm{cc}} & \subsetneq & \mathbf{PSPACE}^{\mathrm{cc}} & , \\ \mathbf{P}^{\mathrm{cc}} & \subsetneq & \mathbf{NP}^{\mathrm{cc}}, \mathbf{coNP}^{\mathrm{cc}} & \subsetneq & \mathbf{PH}^{\mathrm{cc}} & , & & \\ \mathbf{P}^{\mathrm{cc}} & \subsetneq & \mathbf{PP}^{\mathrm{cc}}, \oplus\mathbf{P}^{\mathrm{cc}} & \subsetneq & \mathbf{PSPACE}^{\mathrm{cc}} & . & & \end{array}$$

The following pairs are incomparable:
$(\mathbf{NP}^{\mathrm{cc}}, \mathbf{coNP}^{\mathrm{cc}})$, $(\mathbf{NP}^{\mathrm{cc}}, \mathbf{PP}^{\mathrm{cc}})$, $(\mathbf{NP}^{\mathrm{cc}}, \oplus\mathbf{P}^{\mathrm{cc}})$, $(\mathbf{PP}^{\mathrm{cc}}, \oplus\mathbf{P}^{\mathrm{cc}})$.

Table 4.3: Known inclusion relationships

For more ground work, in particular on closure properties, the Boolean communication hierarchy, or counting communication complexity classes like $\mathbf{MOD}_k\mathbf{P}^{\mathrm{cc}}$, see Halstenberg & Reischuk (1990) or Damm *et al.* (2004). Klauck (2003) established separation results between the classes $\mathbf{MA}^{\mathrm{cc}}$ and $\mathbf{NP}^{\mathrm{cc}}$, $\mathbf{MA}^{\mathrm{cc}}$ and $\mathbf{APP}^{\mathrm{cc}}$, and $\mathbf{APP}^{\mathrm{cc}}$ and $\mathbf{PP}^{\mathrm{cc}}$, respectively. In recent research, Buhrman *et al.* (2007) showed $\mathbf{\Sigma}_2^{\mathrm{cc}}, \mathbf{\Pi}_2^{\mathrm{cc}} \not\subseteq \mathbf{PP}^{\mathrm{cc}}$. This was improved to $\mathbf{\Sigma}_2^{\mathrm{cc}}, \mathbf{\Pi}_2^{\mathrm{cc}} \not\subseteq \mathbf{UPP}^{\mathrm{cc}}$ by Razborov & Sherstov (2008).

Formal languages are defined a bit differently here than in the Turing-machine world because of the distributive nature of communication complexity. The *set of pairs of strings of equal length* is denoted by $\mathbb{B}^{**} := \{(x, y) \mid x, y \in \mathbb{B}^*, |x| = |y|\}$. A *(formal) language* L is a subset of $\mathbb{B}^{**}$, its *n-bit section* L_n is the set of all pairs $(x, y) \in L$ of n-bit words x, y.

A *communication complexity class* is a set of languages. As our bounds on communication will use floors, ceilings, and logarithms, the set of polynomials is not expressive enough, and we have to define $\mathbf{poly} := \{f \colon \mathbb{R}^+ \to \mathbb{R}^+ \mid \exists \text{ polynomial } p \colon f \leq p\}$, the *set of functions with polynomial growth*.

A family of Boolean functions $f := (f_n)_{n\in\mathbb{N}}$, $f_n \colon \mathbb{B}^n \times \mathbb{B}^n \to \mathbb{B}$, can be considered as a family of characteristic functions that defines a language $L_f := \{(x, y) \in \mathbb{B}^{**} \mid f_{|x|}(x, y) = 1\}$. For the other direction, a language L defines a family of functions $\chi^L := (\chi^{L_n})_{n\in\mathbb{N}}$, where $\chi^{L_n} \colon \mathbb{B}^n \times \mathbb{B}^n \to \mathbb{B}$, $\chi^{L_n}(x, y) := [(x, y) \in L_n]$.

In the sequel, we often do not distinguish between languages and characteristic function families. In particular, for a complexity measure M we write $M(L_n)$, where it should correctly read $M(\chi^{L_n})$.

We call a protocol over domain $\mathbb{B}^n \times \mathbb{B}^n$ an *n-bit* protocol. A protocol family $(\Pi_n)_{n\in\mathbb{N}}$ of n-bit protocols Π_n *decides* a language L if each Π_n computes the characteristic function of L_n.

In each structural theory, the standard set of complexity classes is defined based on the standard set of complexity measures (deterministic, randomized, nondeterministic, etc.) and a notion of *efficiency*. In structural communication complexity, if a problem can be solved with communication complexity polylogarithmically in the input length, then we consider this as efficient.

Definition 4.1 (Some standard classes).

$$\begin{aligned}
\mathbf{P}^{\mathrm{cc}} &:= \{L \subseteq \mathbb{B}^{**} \mid \exists p \in \mathbf{poly}\colon \mathrm{D}(L_n) \le p(\log n)\}\ ,\\
\mathbf{BPP}^{\mathrm{cc}} &:= \{L \subseteq \mathbb{B}^{**} \mid \exists p \in \mathbf{poly}\colon \mathrm{R}^{\mathrm{pub}}_{1/3}(L_n) \le p(\log n)\}\ ,\\
\mathbf{PP}^{\mathrm{cc}} &:= \{L \subseteq \mathbb{B}^{**} \mid \exists p \in \mathbf{poly}\colon \mathrm{PP}(L_n) \le p(\log n)\}\ ,\\
\mathbf{NP}^{\mathrm{cc}} &:= \{L \subseteq \mathbb{B}^{**} \mid \exists p \in \mathbf{poly}\colon \mathrm{N}^1(L_n) \le p(\log n)\}\ ,\\
\mathbf{coNP}^{\mathrm{cc}} &:= \{L \subseteq \mathbb{B}^{**} \mid \exists p \in \mathbf{poly}\colon \mathrm{N}^0(L_n) \le p(\log n)\}\ ,\\
\mathbf{MOD}_k\mathbf{P}^{\mathrm{cc}} &:= \{L \subseteq \mathbb{B}^{**} \mid \exists p \in \mathbf{poly}\colon \mathrm{MOD}_k(L_n) \le p(\log n)\}\ ,\\
\oplus\mathbf{P}^{\mathrm{cc}} &:= \mathbf{MOD}_2\mathbf{P}^{\mathrm{cc}}\ .
\end{aligned}$$

In the Turing-machine model the complexity classes **PSPACE** and **PH** are defined based on the resource "space". The important observation that these classes can be defined via alternating Turing machines opened the possibility to define analogous classes in structural communication complexity.

Definition 4.2 (Alternating classes).

$$\begin{aligned}
\mathbf{PSPACE}^{\mathrm{cc}} &:= \{L \subseteq \mathbb{B}^{**} \mid \exists p \in \mathbf{poly}\colon \mathrm{A}(L_n) \le p(\log n)\}\ ,\\
\mathbf{PH}^{\mathrm{cc}} &:= \bigcup_{k\ge 0} \mathbf{\Sigma}^{\mathrm{cc}}_k\ ,\\
\mathbf{\Sigma}^{\mathrm{cc}}_0 := \mathbf{P}^{\mathrm{cc}}\ ,\ \mathbf{\Sigma}^{\mathrm{cc}}_{k+1} &:= \{L \subseteq \mathbb{B}^{**} \mid \exists p \in \mathbf{poly}\colon \mathrm{A}^k(L_n) \le p(\log n)\}\ , k \ge 0\ .
\end{aligned}$$

From the plethora of function classes we only need the class *Sharp-P*, $\#\mathbf{P}^{\mathrm{cc}}$, in the sequel. It contains all function families $\mathrm{acc}_\Pi := (\mathrm{acc}_{\Pi_n})_{n\in\mathbb{N}}$ defined by protocol families $\Pi := (\Pi_n)_{n\in\mathbb{N}}$ of n-bit counting protocols Π_n that are efficient, i.e., there exists a $p \in \mathbf{poly}$ such that for all n the communication cost of Π_n is bounded by $p(\log n)$.

An important concept in structural complexity is *relativization*. Analogous to oracle Turing machines one can define *oracle protocols*. A deterministic, randomized, counting, or alternating protocol Π over $\mathcal{X} \times \mathcal{Y}$ is an *oracle protocol* with oracle family $O = (O_m)_{m\in\mathbb{N}}$, if Π contains *oracle nodes* in its protocol tree. Associated with an oracle node v are two functions $a_v\colon \mathcal{X} \to \mathbb{B}^{m_v}$ and $b_v\colon \mathcal{Y} \to \mathbb{B}^{m_v}$. If Alice and Bob reach an oracle node v during a computation on input $(x, y) \in X \times \mathcal{Y}$, they compute by themselves $x' := a_v(x)$ and $y' := b_v(y)$, respectively, and call O_{m_v} on (x', y'). The oracle node v has exactly $|\mathbf{range}(O_{m_v})|$ many successors. Alice and Bob continue the computation on one of them according to the returned value $O_{m_v}(x', y')$. The communication cost for each oracle call is $\lceil \log |\mathbf{range}(O_{m_v})| \rceil$. *Relativized communication complexity classes* are defined via efficient oracle protocol families, where for each oracle node v the *query length* m_v is bounded by $2^{\mathrm{polylog}(n)}$. For example, $\mathbf{P}^{\mathrm{cc}}(L')$ contains all languages L which can be decided by an efficient protocol family $(\Pi_n)_{n\in\mathbb{N}}$ of deterministic n-bit oracle protocols Π_n with oracle family $(L'_m)_{m\in\mathbb{N}}$.

Reductions play a central role in structural complexity. In Babai *et al.* (1986) different kinds of reductions were defined analogously to the Turing-machine model. In structural communication complexity, many-one reductions defined below are also called *rectangular reductions*.

Definition 4.3 (Reductions). *Let L and L' be languages.*

(i) *L is* many-one reducible *to L', if there exist a bound $b \in \mathbf{poly}$ and a family of function pairs $\{(f_n, g_n)\}_{n\in\mathbb{N}}$, $f_n, g_n\colon \mathbb{B}^n \to \mathbb{B}^{\lceil 2^{b(\log n)}\rceil}$, such that for all n-bit input pairs (x, y) we have*

$$(x, y) \in L \iff (f_n(x), g_n(y)) \in L'\ .$$

(ii) *L is* Turing reducible *to L', if $L \in \mathbf{P}^{\mathrm{cc}}(L')$.*

(iii) *L is* majority reducible *to L', if there exist bounds $b, t \in \mathbf{poly}$ and a family of function pairs $\{(f_n, g_n)\}_{n\in\mathbb{N}}$, $f_n, g_n \colon \mathbb{B}^n \to \mathbb{B}^*$, such that for all n-bit input pairs (x, y) we have*

$$\begin{aligned} f_n(x) &= \langle x_1, \ldots, x_\ell \rangle \ , \\ g_n(y) &= \langle y_1, \ldots, y_\ell \rangle \ , \end{aligned}$$

where $\ell := \lceil t(\log n) \rceil$, $|x_i| = |y_i| \leq \lceil 2^{b(\log n)} \rceil$ and

$$(x, y) \in L \iff (x_i, y_i) \in L' \text{ for the majority of the indices } i \in [\ell].$$

(iv) *L is* conjunctively reducible *to L', if there exist bounds $b, t \in \mathbf{poly}$ and a family of function pairs $\{(f_n, g_n)\}_{n\in\mathbb{N}}$, $f_n, g_n \colon \mathbb{B}^n \to \mathbb{B}^*$, such that for all n-bit input pairs (x, y) we have*

$$\begin{aligned} f_n(x) &= \langle x_1, \ldots, x_\ell \rangle \ , \\ g_n(y) &= \langle y_1, \ldots, y_\ell \rangle \ , \end{aligned}$$

where $\ell := \lceil t(\log n) \rceil$, $|x_i| = |y_i| \leq \lceil 2^{b(\log n)} \rceil$ and

$$(x, y) \in L \iff (x_i, y_i) \in L' \text{ for all indices } i \in [\ell].$$

5. Toda's Theorems

In this section we prove a slight generalization of Toda's remarkable theorems (see e.g., Toda 1991) in the setting of communication complexity.[1]

In order to formulate and prove the respective statements we must first define several complexity class operators and state some of their properties.

Crucial to a translation of these proofs is the consequent use of complexity class operators, because relativizing statements like the Lemma of Valiant & Vazirani (1986) seems to be impossible in communication complexity.

Definition 5.1 (Complexity class operators). *For a language L and a bound $p \in \mathbf{poly}$ we define*

$$\begin{aligned} \forall^p(L) &:= \left\{ (x, y) \in \mathbb{B}^{**} \;\middle|\; \forall w \in \mathbb{B}^{\lceil p(\log |x|) \rceil} \colon (\langle x, w \rangle, \langle y, w \rangle) \in L \right\} \ , \\ \exists^p(L) &:= \left\{ (x, y) \in \mathbb{B}^{**} \;\middle|\; \exists w \in \mathbb{B}^{\lceil p(\log |x|) \rceil} \colon (\langle x, w \rangle, \langle y, w \rangle) \in L \right\} \ , \\ \mathrm{MOD}_k^p(L) &:= \left\{ (x, y) \in \mathbb{B}^{**} \;\middle|\; |\{ w \in \mathbb{B}^{\lceil p(\log |x|) \rceil} \mid (\langle x, w \rangle, \langle y, w \rangle) \in L \}| \bmod k \neq 0 \right\} \ , \\ \oplus^p(L) &:= \mathrm{MOD}_2^p(L) \ . \end{aligned}$$

For a communication complexity class $\mathcal{C}$ we define

$$\begin{aligned} \mathrm{co} \cdot \mathcal{C} &:= \left\{ \overline{L} \mid L \in \mathcal{C} \right\} \ , \\ \forall \cdot \mathcal{C} &:= \{ \forall^p(L) \mid L \in \mathcal{C}, p \in \mathbf{poly} \} \ , \\ \exists \cdot \mathcal{C} &:= \{ \exists^p(L) \mid L \in \mathcal{C}, p \in \mathbf{poly} \} \ , \\ \mathrm{MOD}_k \cdot \mathcal{C} &:= \{ \mathrm{MOD}_k^p(L) \mid L \in \mathcal{C}, p \in \mathbf{poly} \} \ , \\ \oplus \cdot \mathcal{C} &:= \mathrm{MOD}_2 \cdot \mathcal{C} \ . \end{aligned}$$

Furthermore, we define the communication complexity version of the BP-operator introduced in Schöning (1989).

A language L is in $\mathrm{BP} \cdot \mathcal{C}$ if there exist a language $L' \in \mathcal{C}$ and a bound $q \in \mathbf{poly}$ such that for all n-bit input pairs (x, y) we have

$$\begin{aligned} (x, y) \in L &\implies \left| \left\{ r \in \mathbb{B}^{\lceil q(\log n) \rceil} \;\middle|\; (\langle x, r \rangle, \langle y, r \rangle) \in L' \right\} \right| / 2^{\lceil q(\log n) \rceil} \geq 2/3 \ , \\ (x, y) \notin L &\implies \left| \left\{ r \in \mathbb{B}^{\lceil q(\log n) \rceil} \;\middle|\; (\langle x, r \rangle, \langle y, r \rangle) \in L' \right\} \right| / 2^{\lceil q(\log n) \rceil} \leq 1/3 \ . \end{aligned}$$

[1] In an unpublished manuscript, Lokam (1996) claimed without proof that Toda's Second Theorem holds in the setting of communication complexity.

A careful reader familiar with randomized communication complexity might wonder why the operators above are defined in a *public-coin style*, i.e., both players get the same witness/random string. Indeed, one can define the operators such that each player gets his/her own witness/random string (*private-coin style*). These definitions are equivalent, if the operators are simulated by a protocol. Alice can guess Bob's witness and send it to him, or she can send him her random string, because the length of witnesses/random strings is bounded polylogarithmically in the length of the input.

Furthermore, one can ask if such bounds on the witness length are really necessary. The answer to this question depends on the operator under consideration. Let $\mathrm{BP}^{\mathrm{unbd}}\cdot$ and $\exists^{\mathrm{unbd}}\cdot$ denote the *unbounded* versions of the BP- and $\exists$-operator, respectively, i.e., the witness bound q is not restricted to be chosen from **poly**. On the one hand, Newman's Theorem (Fact 3.13) shows how to replace a large probability space by a small one. This implies that there is no essential difference between $\mathrm{BP}\cdot$ and $\mathrm{BP}^{\mathrm{unbd}}\cdot$, i.e., under natural conditions on the communication complexity class $\mathcal{C}$ we have $\mathrm{BP}^{\mathrm{unbd}}\cdot\mathcal{C} = \mathrm{BP}\cdot\mathcal{C}$. On the other hand, it is easy to see that every counting class is contained in $\exists^{\mathrm{unbd}}\cdot\forall\cdot\mathbf{P}^{\mathrm{cc}}$. Let $\mathcal{C}$ be based on an acceptance mode μ, let $L\in\mathcal{C}$, and let Π be an efficient counting protocol for L. On an input (x,y) of length $n=|x|$, Alice existentially guesses a string v containing all $2^{\mathrm{polylog}(n)}$ outputs of Π on input (x,y). If the guess string v obeys the acceptance mode μ, Alice and Bob check the correctness of Alice's guess. They universally guess witnesses w_A, w_B, respectively, simulate $\Pi((x,w_A),(y,w_B)) =: t$, and accept iff the t-th bit of v equals the output of Π.

The following observation shows that the operators are defined in the "right" way, and that the names given to them are compatible with the names of classical communication complexity classes, if the operators are applied to $\mathbf{P}^{\mathrm{cc}}$.

Observation 5.2 (Compatibility).

$$\begin{aligned}\mathbf{NP}^{\mathrm{cc}} &= \exists\cdot\mathbf{P}^{\mathrm{cc}}\ , & \mathbf{MOD}_k\mathbf{P}^{\mathrm{cc}} &= \mathrm{MOD}_k\cdot\mathbf{P}^{\mathrm{cc}}\ ,\\ \mathbf{coNP}^{\mathrm{cc}} &= \forall\cdot\mathbf{P}^{\mathrm{cc}}\ , & \oplus\mathbf{P}^{\mathrm{cc}} &= \oplus\cdot\mathbf{P}^{\mathrm{cc}}\ ,\\ \mathbf{BPP}^{\mathrm{cc}} &= \mathrm{BP}\cdot\mathbf{P}^{\mathrm{cc}}\ .\end{aligned}$$

Observation 5.3. *For every natural number $k\geq 2$ we have*

$$\mathrm{BP}\cdot\mathbf{MOD}_k\mathbf{P}^{\mathrm{cc}} = \{L \mid \exists p\in\mathbf{poly}\colon \mathrm{BP}\cdot\mathrm{MOD}^{\mathrm{pub}}_{k,1/3}(L_n)\leq p(\log n)\}\ .$$

We observe the following properties of the communication complexity class operators. The respective proofs are so easy that we omit most of them for brevity.

Observation 5.4 (Probability amplification). *Let $\mathcal{C}$ be a communication complexity class closed under majority reductions, and let $b\in\mathbf{poly}$. If a language L is in $\mathrm{BP}\cdot\mathcal{C}$, then there exist a language $L'\in\mathcal{C}$ and a bound $q\in\mathbf{poly}$ such that for all n-bit input pairs (x,y) we have*

$$\begin{aligned}(x,y)\in L &\implies \left|\left\{r\in\mathbb{B}^{\lceil q(\log n)\rceil}\;\middle|\;(\langle x,r\rangle,\langle y,r\rangle)\in L'\right\}\right|/2^{\lceil q(\log n)\rceil}\geq 1-2^{-b(\log n)}\ ,\\ (x,y)\notin L &\implies \left|\left\{r\in\mathbb{B}^{\lceil q(\log n)\rceil}\;\middle|\;(\langle x,r\rangle,\langle y,r\rangle)\in L'\right\}\right|/2^{\lceil q(\log n)\rceil}\leq 2^{-b(\log n)}\ .\end{aligned}$$

Observation 5.5 (Inclusion). *Let $\mathcal{C}$ be a communication complexity class that is closed under many-one reductions. Then for every operator $\mathrm{Op}\in\{\forall,\exists,\mathrm{MOD}_k,\oplus,\mathrm{BP}\}$ we have $\mathcal{C}\subseteq\mathrm{Op}\cdot\mathcal{C}$.*

Observation 5.6 (Monotonicity). *Let $\mathcal{C}$ and $\mathcal{D}$ be two communication complexity classes such that $\mathcal{C}\subseteq\mathcal{D}$. Then for every operator $\mathrm{Op}\in\{\mathrm{co},\forall,\exists,\mathrm{MOD}_k,\oplus,\mathrm{BP}\}$ we have $\mathrm{Op}\cdot\mathcal{C}\subseteq\mathrm{Op}\cdot\mathcal{D}$.*

OBSERVATION 5.7 (Idempotency). *Let $\mathcal{C}$ be a communication complexity class that is closed under many-one reductions. Then for every operator $\mathrm{Op} \in \{\forall, \exists, \oplus\}$ we have $\mathrm{Op} \cdot \mathrm{Op} \cdot \mathcal{C} = \mathrm{Op} \cdot \mathcal{C}$.*

Under certain closure properties of the communication complexity class $\mathcal{C}$ one could prove the idempotency of the MOD_q-operator for a prime q. We prove a stronger statement in the following lemma:

LEMMA 5.8 (Lowness). *For every prime q we have*

$$\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}\left(\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}\right) = \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \ .$$

In particular, $\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$ is closed under complementation, as well as Turing, majority, and conjunctive reductions.

The above result is well-known in the Turing-machine setting. For the case $q = 2$, it was proved in Papadimitriou & Zachos (1983). For an arbitrary prime q, this "lowness" result is a consequence of Fermat's Little Theorem (see Beigel & Gill 1992).

PROOF. First of all, we prove that $\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$ is closed under complement. Let L be a language in $\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$. Then there exists an efficient protocol family $\Pi \coloneqq (\Pi_n)_{n\in\mathbb{N}}$ of n-bit counting protocols Π_n deciding L in mod-q acceptance mode. From Π we construct an efficient protocol family $\Pi' \coloneqq (\Pi'_n)_{n\in\mathbb{N}}$ deciding L in mod-q acceptance mode such that $\mathrm{acc}_{\Pi'_n} = \mathrm{acc}_{\Pi_n}^{q-1}$. Note that by Fermat's Little Theorem we have

$$\begin{aligned} \mathrm{acc}_{\Pi_n}(x,y) \bmod q = 0 &\implies \mathrm{acc}_{\Pi'_n}(x,y) \bmod q = 0 \ , \\ \mathrm{acc}_{\Pi_n}(x,y) \bmod q \neq 0 &\implies \mathrm{acc}_{\Pi'_n}(x,y) \bmod q = 1 \ . \end{aligned}$$

From Π' we construct an efficient protocol family $\overline{\Pi} \coloneqq \left(\overline{\Pi}_n\right)_{n\in\mathbb{N}}$ deciding $\overline{L}$ in mod-q acceptance mode. On an n-bit input Alice guesses a bit b and sends it to Bob. If $b = \mathtt{1}$, the corresponding subtree of the protocol contains $q-1$ accepting transcripts. (For example, Alice sends a guess string w of length $q-1$ to Bob and they accept iff w is of the form $\mathtt{0}^i\mathtt{1}0^{q-2-i}$, $0 \leq i \leq q-2$.) If $b = \mathtt{0}$ they execute Π'_n.

Note that $\mathrm{acc}_{\overline{\Pi}_n} = (q-1) + \mathrm{acc}_{\Pi'_n}$ implying

$$\begin{aligned} (x,y) \in L &\implies \mathrm{acc}_{\overline{\Pi}_n}(x,y) \bmod q = (q-1)+1 \bmod q = 0 \ , \\ (x,y) \notin L &\implies \mathrm{acc}_{\overline{\Pi}_n}(x,y) \bmod q = (q-1)+0 \bmod q \neq 0 \ . \end{aligned}$$

Thus, $\overline{L}$ is in $\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$.

Now, we prove the lemma. Let L be a language in $\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}\left(\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}\right)$. Then there exists an efficient protocol family $\Pi \coloneqq (\Pi_n)_{n\in\mathbb{N}}$ deciding L in mod-q acceptance mode, where each n-bit counting protocol Π_n has oracle access to a language L' in $\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$. There exist efficient protocol families $\Pi^b \coloneqq \left(\Pi^b_n\right)_{n\in\mathbb{N}}$, $b \in \{0,1\}$, of n-bit counting protocols Π^b_n deciding $\overline{L'}$ for $b = 0$ and L' for $b = 1$, respectively, in mod-q acceptance mode. We assume that the Fermat trick has been applied to them such that the number of accepting transcripts is either 0 or 1 mod q.

We construct an efficient protocol family $\Pi' \coloneqq (\Pi'_n)_{n\in\mathbb{N}}$ deciding L in mod-q acceptance mode as follows. On an n-bit input (x,y) Alice and Bob simulate Π_n. At each oracle node instead of calling the oracle Alice guesses the oracle answer and sends it to Bob. At the end, if Π_n rejects they reject the input, too. Otherwise, let $v_1, \ldots, v_k$ be the visited oracle nodes with attached function pairs a_i, b_i, $i \in [k]$, and let $c_1, \ldots, c_k$ be the guessed oracle answers. For each $i \in [k]$ Alice and Bob execute Π^{c_i} on input $(a_i(x), b_i(y))$. They accept (x,y) iff all executions are accepting. □

The idempotency of the BP-operator follows from its probability amplification property (Observation 5.4).

OBSERVATION 5.9 (Idempotency of BP$\cdot$). *We have $\mathrm{BP} \cdot \mathrm{BP} \cdot \mathcal{C} = \mathrm{BP} \cdot \mathcal{C}$ for every communication complexity class $\mathcal{C}$ closed under majority reductions.*

Observation 5.10 (co· vs. ∃·, ∀· and BP·). *Let $\mathcal{C}$ be a communication complexity class. We have* $\mathrm{co}\cdot\exists\cdot\mathcal{C} = \forall\cdot\mathrm{co}\cdot\mathcal{C}$, $\mathrm{co}\cdot\forall\cdot\mathcal{C} = \exists\cdot\mathrm{co}\cdot\mathcal{C}$, *and* $\mathrm{co}\cdot\mathrm{BP}\cdot\mathcal{C} = \mathrm{BP}\cdot\mathrm{co}\cdot\mathcal{C}$.

Definition 5.11 (Intersection & union). *Let $\mathcal{C}$ and $\mathcal{D}$ be communication complexity classes. The class $\mathcal{C}$ is* closed under $\mathcal{D}$-intersection *iff for all $A \in \mathcal{C}$ and $B \in \mathcal{D}$ we have $A \cap B \in \mathcal{C}$, and it is* closed under $\mathcal{D}$-union *iff for all $A \in \mathcal{C}$ and $B \in \mathcal{D}$ we have $A \cup B \in \mathcal{C}$.*

Definition 5.12 (Normal class). *We call a communication complexity class $\mathcal{C}$* normal *iff it is closed under $\mathbf{P}^{\mathrm{cc}}$-intersection, $\mathbf{P}^{\mathrm{cc}}$-union, and many-one reductions, and if it contains $\mathbf{P}^{\mathrm{cc}}$.*

Swapping lemmata are well-known in the field of structural complexity theory. Below, we give a proof of a lemma of this type for the sake of completeness. The main ingredient is the probability amplification property of the BP-operator (Observation 5.4).

Lemma 5.13 (Swapping). *Let $\mathcal{C}$ be a communication complexity class closed under majority reductions. Then* $\exists\cdot\mathrm{BP}\cdot\mathcal{C} \subseteq \mathrm{BP}\cdot\exists\cdot\mathcal{C}$.

Proof. Let L be a language in $\exists\cdot\mathrm{BP}\cdot\mathcal{C}$. Then there exist a language L' in $\mathrm{BP}\cdot\mathcal{C}$ and a bound $p' \in \mathbf{poly}$ such that $L = \exists^{p'}(L')$. As $L' \in \mathrm{BP}\cdot\mathcal{C}$ and $\mathcal{C}$ is closed under majority reductions we use probability amplification to obtain a language L'' in $\mathcal{C}$ and a bound $p'' \in \mathbf{poly}$ such that

$$\begin{aligned}(\langle x,w\rangle,\langle y,w\rangle) \in L' &\implies \Pr_r\left[(\langle\langle x,w\rangle,r\rangle,\langle\langle y,w\rangle,r\rangle) \in L''\right] \geq 1 - 2^{-\ell'_n-2} \text{ , and}\\ (\langle x,w\rangle,\langle y,w\rangle) \notin L' &\implies \Pr_r\left[(\langle\langle x,w\rangle,r\rangle,\langle\langle y,w\rangle,r\rangle) \in L''\right] \leq 2^{-\ell'_n-2}\end{aligned}$$

for every n-bit input pair (x,y) and witness w. Here, $\ell'_n := \lceil p'(\log n)\rceil$, and the random string r is uniformly drawn from $\mathbb{B}^{\ell''_n}$, where $\ell''_n := \lceil p''(\log n)\rceil$.

Furthermore, we define

$$L''' := \left\{(\langle\langle x,r_1\rangle,w_1\rangle,\langle\langle y,r_2\rangle,w_2\rangle) \,\middle|\, (\langle\langle x,w_1\rangle,r_1\rangle,\langle\langle y,w_2\rangle,r_2\rangle) \in L''\right\} .$$

This language is in $\mathcal{C}$, because $\mathcal{C}$ is closed under many-one reductions. Hence, $L'''' := \exists^{p'}(L''') \in \exists\cdot\mathcal{C}$. It is easy to see that

$$\begin{aligned}(x,y) \in L &\implies \exists w\colon (\langle x,w\rangle,\langle y,w\rangle) \in L'\\ &\implies \Pr_r\left[\exists w\colon (\langle\langle x,w\rangle,r\rangle,\langle\langle y,w\rangle,r\rangle) \in L''\right] \geq \frac{3}{4}\\ &\implies \Pr_r\left[\exists w\colon (\langle\langle x,r\rangle,w\rangle,\langle\langle y,r\rangle,w\rangle) \in L'''\right] \geq \frac{3}{4}\\ &\implies \Pr_r\left[(\langle x,r\rangle,\langle y,r\rangle) \in L''''\right] \geq \frac{3}{4} ,\end{aligned}$$

and that

$$\begin{aligned}(x,y) \notin L &\implies \forall w\colon (\langle x,w\rangle,\langle y,w\rangle) \notin L'\\ &\implies \Pr_r\left[\exists w\colon (\langle\langle x,w\rangle,r\rangle,\langle\langle y,w\rangle,r\rangle) \in L''\right] \leq 2^{\ell'_n}\cdot 2^{-\ell'_n-2}\\ &\implies \Pr_r\left[(\langle x,r\rangle,\langle y,r\rangle) \in L''''\right] \leq \frac{1}{4} .\end{aligned}$$

We conclude $L \in \mathrm{BP}\cdot\exists\cdot\mathcal{C}$. □

The Lemma of Valiant & Vazirani (1986) is a classical result in structural complexity theory. Valiant and Vazirani observed that if one randomly (using randomness (R)) adds certain clauses $\psi(R)$ to a satisfiable SAT-formula ϕ, then with non-negligible probability $\phi \wedge \psi(R)$ has a unique satisfying assignment. As "1" is an odd number and SAT is complete for the class **NP**, we can rephrase the statement in terms of complexity classes.

LEMMA 5.14 (Valiant & Vazirani). $\mathbf{NP} \subseteq \mathrm{RP} \cdot \oplus\mathbf{P}$.

Here, $\mathrm{RP} \cdot \mathcal{C}$ denotes the closure of $\mathcal{C}$ under randomized many-one reductions with one-sided error.

Is it possible to make an analogous statement in the setting of communication complexity? Indeed, it is. The set intersection function, SI, and the inner product function mod two, IP, correspond to SAT and $\oplus$SAT, respectively.

DEFINITION 5.15 (Set intersection). *The* set intersection function *is defined as* $\mathrm{SI} := (\mathrm{SI}_n)_{n\in\mathbb{N}}$, *where* $\mathrm{SI}_n(x,y) := [\,\exists i \in [n] \colon x_i = y_i = \mathtt{1}\,]$ *for all* $x = x_1 \cdots x_n$ *and* $y = y_1 \cdots y_n$.

DEFINITION 5.16 (Inner product function mod two). *The* inner product function mod two, $\mathrm{IP} := (\mathrm{IP}_n)_{n\in\mathbb{N}}$, *is defined as* $\mathrm{IP}_n(x,y) := \sum_{i\in[n]} x_i y_i \bmod 2$ *for all* $x = x_1 \cdots x_n$ *and* $y = y_1 \cdots y_n$.

On inputs $x = x_1 \cdots x_n$ and $y = y_1 \cdots y_n$ Alice and Bob randomly reduce SI_n to IP_n as follows. First of all, they randomly choose a natural number k. The "right" k would obey $2^{k-2} \leq |S| < 2^{k-1}$, where $S := \{i \in [n] \mid x_i = y_i = \mathtt{1}\}$. Then they randomly choose a pairwise independent hash function $h \colon [n] \to \{0,1\}^k$ that selects a subset $S_h := \{i \in [n] \mid h(i) = 0^k\}$ of the indices $[n]$. They call IP_n on $x' = x'_1 \cdots x'_n$ and $y' = y'_1 \cdots y'_n$, where $x'_i := x_i$ for $i \in S_h$, and $x'_i := \mathtt{0}$ otherwise; analogously for y'_i. With non-negligible probability there is a unique index $i \in S_h$ satisfying $x'_i = y'_i = \mathtt{1}$. Thus, we have obtained

LEMMA 5.17 (Valiant & Vazirani). $\mathbf{NP}^{\mathrm{cc}} \subseteq \mathrm{RP} \cdot \oplus\mathbf{P}^{\mathrm{cc}}$.

As we have seen, it is no problem to prove a Valiant-Vazirani Lemma in communication complexity. But what about the relativized version?

OPEN QUESTION 5.18. *Let A be a language. Do we have*

$$\mathbf{NP}^{\mathrm{cc}}(A) \subseteq \mathrm{RP} \cdot \oplus\mathbf{P}^{\mathrm{cc}}(A) \ ?$$

In the setting of communication complexity, relativization seems to destroy the possibility to construct an efficient reduction. Let $\Pi^A := (\Pi^A_n)_{n\in\mathbb{N}}$ be an efficient oracle protocol family for a language $L \in \mathbf{NP}^{\mathrm{cc}}(A)$. Then Π^A_n may have $2^{\mathrm{polylog}(n)}$ many oracle nodes. Hence, the different oracle answers might lead to $2^{2^{\mathrm{polylog}(n)}}$ many different partitions of the input space. A simple many-one reduction via characteristic vectors does not seem to work.

This problem can be circumvented by the use of complexity class operators. It is possible to prove Toda's Theorem in the setting of communication complexity via the respective complexity class operators and the following *operator-theoretical version* of the Valiant-Vazirani Lemma (see Wunderlich 2009).

LEMMA 5.19 (Valiant & Vazirani). *Let $\mathcal{C}$ be a normal communication complexity class closed under conjunctive reductions. Then $\exists \cdot \mathcal{C} \subseteq \mathrm{BP} \cdot \oplus \cdot \mathcal{C}$.*

As we want to prove a slight generalization of Toda's First Theorem, we slightly adapt the Valiant-Vazirani Lemma.

LEMMA 5.20. *For every prime q we have*

$$\exists \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \subseteq \mathrm{BP} \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \ .$$

PROOF. The proof is an adaptation of an algebraic proof due to Fortnow in (Fortnow 1997, p. 88, Lemma 3.12). Let L be a language in $\exists \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$. There exist a language $L' \in \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$ and a bound $p \in \mathbf{poly}$ such that $L = \exists^p(L')$. Define $\ell_n := \lceil p(\log n) \rceil$. We fix an input $(x,y) \in L$, $|x| = |y| = n$. Let $S := \{w \in \mathbb{B}^{\ell_n} \mid (\langle x,w\rangle, \langle y,w\rangle) \in L'\}$ be the set of witnesses of (x,y) and $d := |S|$ its size. We will pick a natural number m such that $\log(2\ell_n d) < m \leq \log(4\ell_n d)$ and

encode the witnesses as polynomials over $F \coloneqq \mathbb{F}_{2^m}$, the finite field with 2^m elements. We will then consider pairs $(a, b) \in F^2$ and show that for a sizable fraction of them there will be exactly one polynomial p representing a witness such that $p(a) = b$. The statement follows by choosing m, a and b at random.

For a string $s = s_1 \cdots s_\ell$ we define the polynomial $p_s(X) \coloneqq \sum_{i=1}^{\ell} s_i X^{i-1}$. We fix a witness w in S. An element a of F is called w-good, if for all witnesses $w' \neq w$ in S we have $p_w(a) \neq p_{w'}(a)$. Since p_w and $p_{w'}$ can agree on at most ℓ_n elements, there are at least $|F| - \ell_n d$ many w-good elements in F. Consider the set A_w containing all pairs $(a, p_w(a))$ for w-good elements a. The sets A_w and $A_{w'}$ are disjoint for different strings w and w'. Define $A \coloneqq \bigcup_{w \in S} A_w$. Then $|A| \geq d(|F| - \ell_n d)$. We define the language L'' in $\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$ by

$$L'' \coloneqq \big\{ (\langle\langle x, r\rangle, w\rangle, \langle\langle y, r\rangle, w\rangle) \,\big|\, n \coloneqq |x| = |y|, r = \langle m^*, a, b\rangle, m^* \in [2\ell_n], \\ a, b \in \mathbb{F}_{2^{m^*}}, |w| = \ell_n, p_w(a) = b, (\langle x, w\rangle, \langle y, w\rangle) \in L' \big\} \ ,$$

where $r = \langle m^*, a, b\rangle$ means that we use r as an encoding of a natural number m^* and field elements a and b. Furthermore, we define

$$L''' \coloneqq \mathrm{MOD}_q^p(L'') \in \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \, (\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}) = \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \ .$$

If $(x, y) \notin L$ then for all w and r the pair $(\langle\langle x, r\rangle, w\rangle, \langle\langle y, r\rangle, w\rangle)$ is not in L'', and thus $(x, y) \notin L'''$.

If $(x, y) \in L$ then with probability $1/2\ell_n$ we have $m = m^*$ as $m \leq \log(4\ell_n d) \leq 2\ell_n$. In case $m = m^*$ the size of A is at least $\ell_n d^2$, the size of F^2 is at most $16\ell_n^2 d^2$. If we choose (a, b) at random in F^2 we have a $1/16\ell_n$ chance of being in A. Consequently, for a fixed input (x, y) the probability of choosing r at random such that $m = m^*$ and $(a, b) \in A$ is at least $1/32\ell_n^2$. In this case, there is exactly one witness w for $(\langle x, r\rangle, \langle y, r\rangle)$ showing $(x, y) \in L'''$.

The class $\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$ is closed under majority reductions by Lemma 5.8. Thus, probability amplification is possible, and we finally obtain $L \in \mathrm{BP} \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$. □

Below, we provide a characterization of the polynomial hierarchy based on complexity class operators.

Observation 5.21 (Polynomial hierarchy). *Each level of the polynomial hierarchy* $\mathbf{PH}^{\mathrm{cc}}$ *is expressible as*

$$\mathbf{\Sigma}_0^{\mathrm{cc}} = \mathbf{P}^{\mathrm{cc}} \ , \ \mathbf{\Sigma}_{k+1}^{\mathrm{cc}} = \exists \cdot \mathrm{co} \cdot \mathbf{\Sigma}_k^{\mathrm{cc}} \ , \ k \geq 0.$$

Toda's First Theorem 5.22. *For every prime* q *we have*

$$\mathbf{PH}^{\mathrm{cc}} \subseteq \mathrm{BP} \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \ .$$

Proof. The proof is analogous to the one in the Turing-machine setting. We prove $\mathbf{\Sigma}_k^{\mathrm{cc}} \subseteq \mathrm{BP} \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$ by induction on k:
Case $k = 0$: This follows from $\mathbf{\Sigma}_0^{\mathrm{cc}} = \mathbf{P}^{\mathrm{cc}} \subseteq \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$ and the inclusion poperty of the BP-operator (Observation 5.5).
Case $k \to k + 1$: We have

$$\begin{aligned}
&(5.23) && \mathbf{\Sigma}_{k+1}^{\mathrm{cc}} = \exists \cdot \mathrm{co} \cdot \mathbf{\Sigma}_k^{\mathrm{cc}} \\
&(5.24) && \subseteq \exists \cdot \mathrm{co} \cdot \mathrm{BP} \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \\
&(5.25) && = \exists \cdot \mathrm{BP} \cdot \mathrm{co} \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \\
&(5.26) && = \exists \cdot \mathrm{BP} \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \\
&(5.27) && \subseteq \mathrm{BP} \cdot \exists \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \\
&(5.28) && \subseteq \mathrm{BP} \cdot \mathrm{BP} \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \\
&(5.29) && = \mathrm{BP} \cdot \mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}} \ .
\end{aligned}$$

(5.23) By Observation 5.21.

(5.24) By the induction hypothesis for $\mathbf{\Sigma}_k^{cc}$ and monotonicity (Observation 5.6) of the operators co· and $\exists\cdot$.

(5.25) By commutativity of co· and BP-operator (Observation 5.10).

(5.26) By closure under complement of $\mathbf{MOD}_q\mathbf{P}^{cc}$ (Lemma 5.8).

(5.27) By the Swapping Lemma (Lemma 5.13). This can be applied, because $\mathbf{MOD}_q\mathbf{P}^{cc}$ is closed under majority reductions (Lemma 5.8).

(5.28) By the Valiant-Vazirani Lemma (Lemma 5.20) and monotonicity of the BP-operator (Observation 5.6).

(5.29) By idempotency of the BP-operator (Observation 5.9). This holds, because $\mathbf{MOD}_q\mathbf{P}^{cc}$ is closed under majority reductions (Lemma 5.8).

□

For the Turing-machine model the fact below was established in Angluin (1980).

FACT 5.30 (Angluin). $\mathbf{P}^{cc}(\mathbf{PP}^{cc}) = \mathbf{P}^{cc}(\#\mathbf{P}^{cc})$.

PROOF. The proof is analogous to the one in the Turing-machine setting. Alice and Bob can compute every $\#\mathbf{P}^{cc}$-function f by binary search with polylog communication asking oracle queries to $\mathrm{Graph}_{\leq}(f) \in \mathbf{PP}^{cc}$, where $\mathrm{Graph}_{\leq}(f) := \{(\langle x, v\rangle, \langle y, v\rangle) \mid (v)_2 \leq f(x,y)\}$, and $(v)_2$ is the binary value of the string v. □

TODA'S SECOND THEOREM 5.31. $\mathrm{BP}\cdot\oplus\mathbf{P}^{cc} \subseteq \mathbf{P}^{cc}(\#\mathbf{P}^{cc})$.

PROOF. The proof is analogous to the one in the Turing-machine setting. If $\Pi := (\Pi_n)_{n\in\mathbb{N}}$ is an efficient family of counting protocols with $\mathrm{acc}_\Pi := (\mathrm{acc}_{\Pi_n})_{n\in\mathbb{N}}$ in $\#\mathbf{P}^{cc}$, and if we choose $p \in \mathbf{poly}$, then there exists an efficient family of counting protocols $\Pi' := (\Pi'_n)_{n\in\mathbb{N}}$ such that $\mathrm{acc}_{\Pi'_n}(x,y) = (1 + \mathrm{acc}_{\Pi_n}(x,y)^{\lceil p(\log n)\rceil})^{\lceil p(\log n)\rceil}$. The family $\mathrm{acc}_{\Pi'}$ is in $\#\mathbf{P}^{cc}$, because the class $\#\mathbf{P}^{cc}$ contains all constant functions and is closed under addition and multiplication. □

We close this section with a corollary summing up the previous results.

COROLLARY 5.32. *For every prime q we have obtained the following chain of inclusions:*

$$\mathbf{PH}^{cc} \subseteq \mathrm{BP}\cdot\mathbf{MOD}_q\mathbf{P}^{cc} \subseteq \mathbf{PSPACE}^{cc} \ .$$

Furthermore,

$$\mathbf{PH}^{cc} \subseteq \mathrm{BP}\cdot\oplus\mathbf{P}^{cc} \subseteq \mathbf{P}^{cc}(\#\mathbf{P}^{cc}) = \mathbf{P}^{cc}(\mathbf{PP}^{cc}) \subseteq \mathbf{PSPACE}^{cc} \ .$$

6. Razborov's Theorem

The first difficulty one encounters when one tries to separate the polynomial hierarchy from polynomial space in communication complexity is that we do not have any measures/lower-bound methods for alternating communication complexity. Fortunately, we have Toda's First Theorem,

$$\mathbf{PH}^{cc} \subseteq \mathrm{BP}\cdot\mathbf{MOD}_q\mathbf{P}^{cc} \subseteq \mathbf{PSPACE}^{cc} \ , q \text{ prime},$$

telling us that the classes $\mathrm{BP}\cdot\mathbf{MOD}_q\mathbf{P}^{cc}$ are in a "sandwich" position between the two alternating classes. It is an important observation that these classes are *not* based on the concept of alternation, enabling us to derive measures, namely *approximate ranks*, that even characterize these classes. This will be done in Section 6.1. Finally, as to each notion of approximate rank there is an equivalent notion of matrix rigidity, we are able to reprove and strengthen Razborov's Theorem in Section 6.2.

6.1. Approximate rank. We define notions of *approximate* $\mathbb{F}$-rank for fields $\mathbb{F}$, which to the author's knowledge have not appeared in the published literature before. In particular, we are interested in *approximate toggle ranks*, because they characterize bounded-error modular communication complexity.

DEFINITION 6.1 (Approximate $\mathbb{F}$-rank). *Let $\mathbb{F}$ be a field, let M be a matrix over $\mathbb{F}$ with row index set $\mathcal{X}$ and column index set $\mathcal{Y}$, let μ be a probability distribution on $\mathcal{X} \times \mathcal{Y}$, and let $\epsilon \geq 0$ be a real number. The* (μ, ϵ)-approximate $\mathbb{F}$-rank of M *is defined as*

$$\mathbb{F}\text{-rank}_\epsilon^\mu(M) \coloneqq \min\{\mathbb{F}\text{-rank}(\tilde{M}) \mid \mu(\tilde{M} \neq M) \leq \epsilon, \tilde{M} \text{ a matrix over } \mathbb{F}\} \ .$$

Here, $\mu(\tilde{M} \neq M) \coloneqq \mu\{(x,y) \mid \tilde{M}_{x,y} \neq M_{x,y}\}$.

The ϵ-approximate $\mathbb{F}$-rank of M *is defined as*

$$\mathbb{F}\text{-rank}_\epsilon^*(M) \coloneqq \max_\mu \mathbb{F}\text{-rank}_\epsilon^\mu(M) \ .$$

DEFINITION 6.2 (Approximate toggle $\mathbb{F}$-rank). *Let $\mathbb{F}$ be a field, let M be a Boolean matrix with row index set $\mathcal{X}$ and column index set $\mathcal{Y}$, let μ be a probability distribution on $\mathcal{X} \times \mathcal{Y}$, and let $\epsilon \geq 0$ be a real number. The* (μ, ϵ)-approximate toggle $\mathbb{F}$-rank of M *is defined as*

$$\text{toggle}\,\mathbb{F}\text{-rank}_\epsilon^\mu(M) \coloneqq \min\{\mathbb{F}\text{-rank}(\tilde{M}) \mid \mu(\tilde{M} \neq M) \leq \epsilon, \tilde{M} \text{ a Boolean matrix}\} \ .$$

The ϵ-approximate toggle $\mathbb{F}$-rank of M *is defined as*

$$\text{toggle}\,\mathbb{F}\text{-rank}_\epsilon^*(M) \coloneqq \max_\mu \mathbb{F}\text{-rank}_\epsilon^\mu(M) \ .$$

Clearly, approximate $\mathbb{F}_2$-rank and approximate toggle $\mathbb{F}_2$-rank coincide. To become more familiar with these notions of approximate ranks, we provide two examples. The first one is perhaps surprising.

EXAMPLE 6.3. Let U denote the uniform distribution. For every $n \times n$-matrix N over $\mathbb{F}_2$ with full rank n we have

$$\mathbb{F}_2\text{-rank}_{1/n^2}^U(N) = n - 1 \ .$$

This can be seen by looking at the determinant of N. First-row-expansion gives

$$1 = |N| = \bigoplus_{i=1}^n N_{1,i} \cdot \left|N^{(1,i)}\right| \ .$$

Here, $N^{(1,i)}$ denotes the matrix N with row 1 and column i deleted.

Thus, there exists an odd set I of indices such that

$$N_{1,i} \cdot \left|N^{(1,i)}\right| = 1$$

for all $i \in I$. Complementing a single $(1, i_0)$-entry of N for an $i_0 \in I$ reduces the rank by one. $\Diamond$

Proving high lower bounds for approximate ranks seems to be a challenging open problem. Nevertheless, it might be easier to prove lower bounds for approximate toggle ranks, because toggling values from 0 to 1 or from 1 to 0 is far more restrictive than replacing entries with arbitrary field elements.

The next example shows that our intuiton is right, i.e., for primes $q > 2$, approximate toggle rank is indeed a strictly weaker notion.

EXAMPLE 6.4. Again, let U denote the uniform distribution. Consider the following matrix:

$$M \coloneqq \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \ .$$

On the one hand, it holds

$$\operatorname{toggle} \mathbb{F}_3\text{-rank}^U_{1/9}(M) = 3 \ ,$$

because for every $x \in \{0,1\}$ we have

$$\begin{vmatrix} x & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & x & 1 \\ 1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & x \end{vmatrix} = 2 - x$$

and

$$\begin{vmatrix} 0 & x & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 1 & x \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ x & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & x \\ 1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ x & 1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & x & 0 \end{vmatrix} = 1 + x \ .$$

On the other hand, by setting $x \coloneqq 2$ the above calculation also shows that

$$\mathbb{F}_3\text{-rank}^U_{1/9}(M) = 2 \ .$$

◊

The next theorem states that the logarithm of approximate toggle $\mathbb{F}_q$-rank is a measure characterizing the bounded-error mod-q communication complexity of a function.

THEOREM 6.5 (Characterization). *Let f be a Boolean function, let q be a prime, and let $\epsilon \geq 0$ be a real number. Then we have*

$$\begin{aligned} \mathrm{BP} \cdot \mathrm{MOD}^{\mathrm{pub}}_{q,\epsilon}(f) &\geq \frac{\log\left(\operatorname{toggle} \mathbb{F}_q\text{-rank}^*_\epsilon\left(M^f\right)\right)}{q-1} \ , \\ \mathrm{BP} \cdot \mathrm{MOD}^{\mathrm{pub}}_{q,\epsilon}(f) &\leq \log\left(\operatorname{toggle} \mathbb{F}_q\text{-rank}^*_\epsilon\left(M^f\right)\right) + \mathcal{O}(1) \ . \end{aligned}$$

PROOF. For the lower bound, we observe that

$$\begin{aligned} (q-1) \cdot \mathrm{BP} \cdot \mathrm{MOD}^{\mathrm{pub}}_{q,\epsilon}(f) &= (q-1) \cdot \max_\mu \mathrm{MOD}^\mu_{q,\epsilon}(f) && (6.6) \\ &= \max_\mu \min_{\tilde{f} \colon \mu(\tilde{f} \neq f) \leq \epsilon} (q-1) \cdot \mathrm{MOD}_q(\tilde{f}) && (6.7) \\ &\geq \max_\mu \min_{\tilde{f} \colon \mu(\tilde{f} \neq f) \leq \epsilon} \log \mathbb{F}_q\text{-rank}\left(M^{\tilde{f}}\right) && (6.8) \\ &= \log \max_\mu \min_{\tilde{f} \colon \mu(\tilde{f} \neq f) \leq \epsilon} \mathbb{F}_q\text{-rank}\left(M^{\tilde{f}}\right) \\ &= \log \max_\mu \min_{\tilde{M} \colon \mu(\tilde{M} \neq M^f) \leq \epsilon} \mathbb{F}_q\text{-rank}(\tilde{M}) \\ &= \log \max_\mu \left(\operatorname{toggle} \mathbb{F}_q\text{-rank}^\mu_\epsilon\left(M^f\right)\right) \\ &= \log\left(\operatorname{toggle} \mathbb{F}_q\text{-rank}^*_\epsilon\left(M^f\right)\right) \ , \end{aligned}$$

where (6.6) holds by Observation 3.29, (6.7) by Observation 3.23, and (6.8) by Fact 3.36, respectively. The upper bound can be derived similarly using

$$\mathrm{MOD}_q(\tilde{f}) \leq \log \mathbb{F}_q\text{-rank}(M^{\tilde{f}}) + \mathcal{O}(1)$$

of Fact 3.36. □

6.1.1. Digression: approximate real rank. The same argument with $\mathbb{R}$-rank shows

THEOREM 6.9. *Let f be a Boolean function, and let $\epsilon > 0$ be a real number. Then we have*

$$\mathrm{R}^{\mathrm{pub}}_{\epsilon}(f) \geq \log\left(\mathrm{toggle}\,\mathbb{R}\text{-rank}^{*}_{\epsilon}\left(M^{f}\right)\right) \ .$$

Thus, we have obtained that the logarithm of the approximate toggle $\mathbb{R}$-rank is a measure for bounded-error public-coin randomized communication complexity. Note that if the logarithmic-rank conjecture (Open Question 3.35) holds, then bounded-error public-coin randomized communication complexity and the logarithm of approximate toggle $\mathbb{R}$-rank are polynomially tight. It would be very interesting to know if one of the known lower-bound methods developed for bounded-error public-coin randomized communication complexity yields exponentially better lower bounds than the logarithm of approximate toggle $\mathbb{R}$-rank, because this would disprove the logarithmic-rank conjecture.

6.2. Matrix rigidity. The concept of *(matrix) rigidity* was introduced by Valiant (1977) as a tool to derive lower bounds in circuit complexity. (For an introduction, we refer the reader to Codenotti 2000.) A matrix has high rigidity, if small perturbations, i.e., changes of a small number of entries in the matrix, do not lower the rank much. Proving a strong enough lower bound on the rigidity of a matrix implies a non-trivial lower bound, i.e., a superlinear size or a superlogarithmic depth, on the complexity of any linear circuit computing the set of linear forms associated with it.

In this subsection we establish an explicit connection between measures of communication complexity and matrix rigidity leading to a stengthening of Razborov's Theorem.

The formal definition of matrix rigidity is given below for the sake of completeness.

DEFINITION 6.10 (Hamming weight). *For a matrix M over a field $\mathbb{F}$ we define the* Hamming weight of M, $\mathrm{wt}(M)$, *as the number of nonzero entries in M.*

DEFINITION 6.11 (Rigidity). *Let M be a matrix over a field $\mathbb{F}$, and let r be a natural number. The* (matrix) rigidity $\mathrm{R}^{\mathbb{F}}_{M}$ of M *is defined as*

$$\mathrm{R}^{\mathbb{F}}_{M}(r) \coloneqq \min\{\mathrm{wt}(\tilde{M} - M) \mid \mathbb{F}\text{-rank}(\tilde{M}) \leq r, \tilde{M} \text{ a matrix over } \mathbb{F}\} \ ,$$

i.e., the minimum number of entries in M that must be changed in order to reduce the rank to r.

In order to obtain a close correspondence between rigidity and approximate rank, we generalize rigidity via arbitrary probability distributions on matrix entries.

DEFINITION 6.12 (Rigidity). *Let $\mathbb{F}$ be a field, let M be a matrix over $\mathbb{F}$ with row index set $\mathcal{X}$ and column index set $\mathcal{Y}$, let μ be a probability distribution on $\mathcal{X} \times \mathcal{Y}$, and let r be a natural number. We define*

$$\mathbb{F}\text{-rigidity}^{\mu}_{M}(r) \coloneqq \min\{\mu(M \neq \tilde{M}) \mid \mathbb{F}\text{-rank}(\tilde{M}) \leq r, \tilde{M} \text{ a matrix over } \mathbb{F}\} \ ,$$
$$\mathbb{F}\text{-rigidity}^{*}_{M}(r) \coloneqq \max_{\mu} \mathbb{F}\text{-rigidity}^{\mu}_{M}(r) \ .$$

Note that $\mathrm{R}^{\mathbb{F}}_{M}(r) = |\mathcal{X} \times \mathcal{Y}| \cdot \mathbb{F}\text{-rigidity}^{U}_{M}(r)$, where U denotes the uniform distribution on $\mathcal{X} \times \mathcal{Y}$.

Most importantly, we introduce a new variant defined for Boolean matrices we call *toggle rigidity*, where only toggling values between `0` and `1` is allowed for rank reduction. This notion corresponds to approximate toggle rank (Definition 6.2).

DEFINITION 6.13 (Toggle rigidity). *Let $\mathbb{F}$ be a field, let M be a Boolean matrix with row index set $\mathcal{X}$ and column index set $\mathcal{Y}$, let μ be a probability distribution on $\mathcal{X} \times \mathcal{Y}$, and let r be a natural number. We define*

$$\mathrm{toggle}\,\mathbb{F}\text{-rigidity}^{\mu}_{M}(r) \coloneqq \min\{\mu(M \neq \tilde{M}) \mid \mathbb{F}\text{-rank}(\tilde{M}) \leq r, \tilde{M} \text{ a Boolean matrix}\} \ ,$$
$$\mathrm{toggle}\,\mathbb{F}\text{-rigidity}^{*}_{M}(r) \coloneqq \max_{\mu} \mathrm{toggle}\,\mathbb{F}\text{-rigidity}^{\mu}_{M}(r) \ .$$

It follows right from the definition that rigidities are monotonically decreasing functions. In addition, we observe that matrix rigidity and approximate rank are equivalent concepts.

Observation 6.14 (Equivalence). *Let $\mathbb{F}$ be a field, let M be a matrix over $\mathbb{F}$, let μ be a probability distribution on M's entries, and let $\epsilon \geq 0$ be a real number. Then*

$$\mathbb{F}\text{-rank}_{\epsilon}^{\mu}(M) \leq r \iff \mathbb{F}\text{-rigidity}_{M}^{\mu}(r) \leq \epsilon \ .$$

If M is Boolean, then

$$\operatorname{toggle}\mathbb{F}\text{-rank}_{\epsilon}^{\mu}(M) \leq r \iff \operatorname{toggle}\mathbb{F}\text{-rigidity}_{M}^{\mu}(r) \leq \epsilon \ .$$

Main Theorem 6.15. *Let $\mathbb{F}$ be a finite field, let $f \coloneqq (f_n)_{n\in\mathbb{N}}$ be a family of Boolean functions $f_n \colon \mathbb{B}^n \times \mathbb{B}^n \to \mathbb{B}$, let $(\mu_n)_{n\in\mathbb{N}}$ be a family of probability distributions μ_n on $\mathbb{B}^n \times \mathbb{B}^n$, let $r(n) \geq 2^{(\log n)^{\omega(1)}}$ be a sequence of natural numbers, and let $\epsilon(n) \geq 1/2^{(\log n)^{\mathcal{O}(1)}}$ be a sequence of real numbers. If*

$$\operatorname{toggle}\mathbb{F}\text{-rigidity}_{M^{f_n}}^{\mu_n}\big(r(n)\big) > \epsilon(n)$$

then the language L_f associated with f is not in the polynomial hierarchy $\mathbf{PH}^{\mathrm{cc}}$.

Proof. It suffices to prove the theorem for finite fields $\mathbb{F}_q$ with q prime, since $\mathbb{F}_{q^n}\text{-rank}(A) \leq \mathbb{F}_q\text{-rank}(A)$ for every matrix A over $\mathbb{F}_q$ and every $n \geq 1$. By probability amplification (Fact 3.14),

$$\mathrm{BP}\cdot\mathrm{MOD}_{q,\epsilon(n)}^{\mathrm{pub}}(f_n) \leq (\log n)^{\mathcal{O}(1)} \cdot \mathrm{BP}\cdot\mathrm{MOD}_{q,1/3}^{\mathrm{pub}}(f_n) \ .$$

By Theorem 6.5,

$$\log\left(\operatorname{toggle}\mathbb{F}_q\text{-rank}_{\epsilon(n)}^{*}\left(M^{f_n}\right)\right) \leq (q-1)\cdot\mathrm{BP}\cdot\mathrm{MOD}_{q,\epsilon(n)}^{\mathrm{pub}}(f_n) \ .$$

By the assumption and Observation 6.14,

$$2^{(\log n)^{\omega(1)}} \leq r(n) \leq \operatorname{toggle}\mathbb{F}_q\text{-rank}_{\epsilon(n)}^{*}\left(M^{f_n}\right) \ .$$

This yields $\mathrm{BP}\cdot\mathrm{MOD}_{q,1/3}^{\mathrm{pub}}(f_n) \geq (\log n)^{\omega(1)}$. □

As a corollary we obtain

Razborov's Theorem 6.16. *Let $\mathbb{F}$ be a finite field, and let $f \coloneqq (f_n)_{n\in\mathbb{N}}$ be a family of Boolean functions. If*

$$\mathrm{R}_{M^{f_n}}^{\mathbb{F}}\left(2^{(\log n)^{\omega(1)}}\right) > \frac{2^{2n}}{2^{(\log n)^{\mathcal{O}(1)}}} \ ,$$

then the language L_f associated with f is not in the polynomial hierarchy $\mathbf{PH}^{\mathrm{cc}}$.

Note that the strengthening of Razborov's Theorem (Main Theorem 6.15) is twofold:

First of all, by placing an arbitrary probability distribution on the matrix entries instead of using the uniform distribution one can put high weight on hard parts of the matrix and low weight on easier parts. This has proven useful for other communication complexity measures. For example, Babai *et al.* (1986) showed that $\mathrm{D}_{\epsilon}^{\mu}(\mathrm{SI}_n) \leq \mathcal{O}(\sqrt{n}\log n)$ for every product distribution μ, while there exists a non-product distribution ν such that $\mathrm{D}_{\epsilon}^{\nu}(\mathrm{SI}_n) \geq \Omega(n)$. (For this fact, see Kalyanasundaram & Schnitger 1992 and Razborov 1992.) Sherstov (2008b) even showed exponential gaps for an explicit function family and arbitrary gaps for a non-explicit function family in Sherstov (2008a). Such effects might occur in the rigidity setting as well.

Secondly, we think that the use of the toggle variant of rigidity instead of classical rigidity will play an important role in future progress on proving high lower bounds. Recall that we have already proved a separation between approximate rank and approximate toggle rank in Example 6.4 giving us higher lower bounds when using toggle rigidity instead of classical rigidity by Observation 6.14.

In addition, high lower bounds for restricted versions of rigidity over the field of real numbers have already been obtained by Lokam (2001). These results imply high lower bounds for toggle rigidity over the field of real numbers.

6.2.1. Digression: concentration of measure. As the name suggests concentration-of-measure results state that a measure puts high weight on a specific set of small size. Several concentration-of-measure results have been established for communication complexity measures, see e.g., Alon *et al.* (1985); Linial & Shraibman (2009b); Orlitsky & El Gamal (1990) showing that most functions have high complexity. The same is true for bounded-error modular communication complexity by an old result from (Valiant 1977, p. 172–173, Theorem 6.4(ii)), where it was shown that over a finite field most Boolean matrices have high rigidity.

FACT 6.17 (Valiant). *Let q be a prime. For all natural numbers n and r, a $(1-1/n)$-fraction of all Boolean $n \times n$-matrices M has rigidity*

$$\mathrm{R}_M^{\mathbb{F}_q}(r) \geq \frac{(n-r)^2 - 2n(\log_q 2) - \log n}{2\log_q n + 1} \ ,$$

if $r < n - \sqrt{2n(\log_q 2) + \log n}$.

THEOREM 6.18. *Let q be a prime. For n sufficiently large, a $(1-1/2^n)$-fraction of all Boolean functions $f\colon \mathbb{B}^n \times \mathbb{B}^n \to \mathbb{B}$ has bounded-error mod-q communication complexity*

$$\mathrm{BP} \cdot \mathrm{MOD}_{q,1/4}^{\mathrm{pub}}(f) \geq \Omega\left(\frac{n}{\log n}\right) \ .$$

PROOF. By Fact 6.17 there exists a constant c such that for n sufficiently large a $(1-1/2^n)$-fraction of all Boolean functions $f\colon \mathbb{B}^n \times \mathbb{B}^n \to \mathbb{B}$ has rigidity

$$\mathrm{R}_{M^f}^{\mathbb{F}_q}(r) \geq c \cdot \frac{(2^n - r)^2}{n} \ ,$$

if $r \leq 2^{n-1}$. Fix such a function f. Define $b(n) \coloneqq \mathrm{BP} \cdot \mathrm{MOD}_{q,1/4}^{\mathrm{pub}}(f)$, $t(n) \coloneqq 6 \cdot \log(2n/c)$, and $\epsilon(n) \coloneqq \frac{1}{2} \cdot (\frac{3}{4})^{t(n)/2}$. By probability amplification (Fact 3.14), we have

$$\mathrm{BP} \cdot \mathrm{MOD}_{q,\epsilon(n)}^{\mathrm{pub}}(f) \leq t(n)b(n) \ .$$

Assume for a contradiction that $b(n) < \frac{n-1}{(q-1)t(n)}$. Let U be the uniform distribution on $\mathbb{B}^n \times \mathbb{B}^n$. Define

$$r_U(n) \coloneqq \mathbb{F}_q\text{-rank}_{\epsilon(n)}^U\left(M^f\right) \leq 2^{(q-1)\mathrm{BP}\cdot\mathrm{MOD}_{q,\epsilon(n)}^{\mathrm{pub}}(f)} =: r(n) \ .$$

Then $\mathrm{R}_{M^f}^{\mathbb{F}_q}\left(r_U(n)\right) \leq \epsilon(n) \cdot 2^{2n}$ by Observation 6.14. By monotonicity we have

$$\frac{1}{2} \cdot \left(\frac{3}{4}\right)^{t(n)/2} \geq \frac{\mathrm{R}_{M^f}^{\mathbb{F}_q}\left(r(n)\right)}{2^{2n}} \geq \frac{c}{n} \cdot \left(1 - 2^{t(n)b(n)(q-1)-n}\right)^2 \geq \frac{c}{4n} \ ,$$

contradicting $(\frac{3}{4})^{t(n)/2} = (\frac{27}{64})^{\log(2n/c)} < (\frac{1}{2})^{\log(2n/c)} = \frac{c}{2n}$. We conclude

$$\mathrm{BP} \cdot \mathrm{MOD}_{q,1/4}^{\mathrm{pub}}(f) = b(n) \geq \frac{n-1}{(q-1)t(n)} \ .$$

□

7. Discussion

Concerning proofs of high lower bounds for matrix rigidity the reader may abandon himself to despair after glancing his eye over existing results. Despite considerable efforts by many researchers, see e.g., Cheraghchi (2005); Codenotti *et al.* (2000); Friedman (1993); Lokam (2000, 2001); Midrijanis (2005); Pudlák (1994); Pudlák & Rödl (1994); Shokrollahi *et al.* (1997); de Wolf (2006), no

explicit construction of a rigid family of matrices over finite fields is known. (For infinite fields Lokam 2006 was able to derive quadratic lower bounds for the rigidity of (in an algebraic sense) explicit matrix families using the concept of (generalized) Smolensky-Shoup-dimension.)

Thus, it is justified to ask if Razborov's strategy to separate $\mathbf{PH}^{cc}$ from $\mathbf{PSPACE}^{cc}$ is a promising one. It could be the case that $\mathrm{BP} \cdot \mathbf{MOD}_q\mathbf{P}^{cc} = \mathbf{PSPACE}^{cc}$. Then his strategy fails even if $\mathbf{PH}^{cc} \subsetneq \mathrm{BP} \cdot \mathbf{MOD}_q\mathbf{P}^{cc}$. In the next subsection we give evidence that this scenario does not occur by presenting a protocol with few alternations for the inner product function mod two suggesting $\mathrm{BP} \cdot \oplus\mathbf{P}^{cc} \subsetneq \mathbf{PSPACE}^{cc}$. In the subsequent subsection we ask what properties possible candidates outside the polynomial hierarchy must have that possibly enable us to prove high lower bounds for them.

7.1. A protocol with few alternations for the inner product function mod two. We want to develop an alternating protocol with few alternations for the inner product function mod two.

For the moment, let L_{IP} denote the language corresponding to IP. It is complete for the class $\oplus\mathbf{P}^{cc}$ under many-one reductions. This is one of many reasons why the inner product function mod two has been studied extensively:

In (Kushilevitz & Nisan 1997, p. 12, Exercise 1.25) it was shown that $\mathrm{R}_0(\mathrm{IP}_n) \geq \mathrm{N}^0(\mathrm{IP}_n) \geq n-1$ using the rectangle-size method. This implies $L_{\mathrm{IP}} \notin \mathbf{coNP}^{cc}$. The lower bound $\mathrm{R}_0^{\mathrm{pub}}(\mathrm{IP}_n) \geq n-1$ for the public-coin model was shown in (Dietzfelbinger & Wunderlich 2007, p. 249, Example 3.7).

The distributional communication complexity of IP was studied in Chor & Goldreich (1988) improving on a result of Vazirani (1987). See also (Babai *et al.* 1986, p. 345, Lemma 9.3, Corollary 9.4). A proof similar to Chor & Goldreich (1988) was given in (Kushilevitz & Nisan 1997, p. 39, Example 3.29; p. 40, Exercise 3.30) that shows $\mathrm{R}_{\frac{1}{2}-\epsilon}(\mathrm{IP}_n) \geq n - \mathcal{O}(\log \frac{1}{\epsilon})$ using the discrepancy method. This implies $L_{\mathrm{IP}} \notin \mathbf{BPP}^{cc}$. Klauck (2003) showed a strong connection between majority covers and the discrepancy method. Hence, the result above actually gives $\mathrm{PP}(\mathrm{IP}_n) = \Theta(n)$. This implies $L_{\mathrm{IP}} \notin \mathbf{PP}^{cc}$. In the work of Forster (2002), a linear lower bound was established in the unbounded-error communication complexity model, implying even $L_{\mathrm{IP}} \notin \mathbf{UPP}^{cc}$.

LEMMA 7.1. *Let $x = x_1 \cdots x_n$ and $y = y_1 \cdots y_n$ be inputs. Divide them into an odd number k of blocks, i.e., $x = x^{(1)} \cdots x^{(k)}$ and $y = y^{(1)} \cdots y^{(k)}$. Then for $b \in \{0,1\}$ the following are equivalent:*

(i) $\mathrm{IP}_n(x,y) = b$.

(ii) *There exists an odd number of blocks indexed by elements in $S \subseteq [k]$ such that $\mathrm{IP}(x^{(i)}, y^{(i)}) = b$ for $i \in S$ and $\mathrm{IP}(x^{(j)}, y^{(j)}) = 1 - b$ for $j \in \overline{S}$.*

PROOF. (ii) $\Longrightarrow$ (i): The cardinality of $\overline{S}$ is even. Thus, we have

$$\begin{aligned}\mathrm{IP}_n(x,y) &= \sum_{i\in[k]} \mathrm{IP}(x^{(i)}, y^{(i)}) \bmod 2 \\ &= \sum_{i\in S} \mathrm{IP}(x^{(i)}, y^{(i)}) + \sum_{j\in\overline{S}} \mathrm{IP}(x^{(j)}, y^{(j)}) \bmod 2 \\ &= |S| \cdot b + |\overline{S}| \cdot (1-b) \bmod 2 = b \ .\end{aligned}$$

(i) $\Longrightarrow$ (ii): Define $S := \{i \in [k] \mid \mathrm{IP}(x^{(i)}, y^{(i)}) = b\}$. By the assumption, we have $b = \mathrm{IP}_n(x,y) = |S| \cdot b + |\overline{S}| \cdot (1-b) \bmod 2$. If $b = 0$ then $|\overline{S}|$ is even, implying $|S|$ odd. If $b = 1$ then $|S| \bmod 2 = 1$. □

The simple lemma above leads to a "divide and conquer"-strategy to compute the inner product function mod two with few alternations. This is implemented in protocol $I_k(s,t,b)$ (Algorithm 1).

OBSERVATION 7.2. *On n-bit inputs $x = x_1 \cdots x_n$ and $y = y_1 \cdots y_n$ the protocol*

$$I_k(s,t,b) \text{ accepts} \iff \mathrm{IP}_{t-s+1}(x_s \cdots x_t, y_s \cdots y_t) = b \ .$$

Consequently, the protocol $I_k(1,n,1)$ computes $\mathrm{IP}_n(x,y)$.

PROOF. The correctness of the protocol follows from Lemma 7.1 by induction on $t - s + 1$. □

There are two alternations in each round of the protocol, and the number of rounds is bounded by $t = \log n / \log k$. If we choose an odd natural number k of size $(\log n)^{\mathcal{O}(1)}$, then the communication cost in each round is $\mathcal{O}(k)$ bits, and the number of alternations is $\mathcal{O}(\log n / \log\log n)$, substantially less than allowed. Recall that $\mathbf{PSPACE}^{cc}$ was defined as the class of languages which can be recognized by protocols using $(\log n)^{\mathcal{O}(1)}$ communication and $(\log n)^{\mathcal{O}(1)}$ many alternations. In particular, the number of alternations is allowed to be proportional to the communication cost.

We consider this as evidence that the class $\oplus\mathbf{P}^{cc}$ is much "easier" than the class $\mathbf{PSPACE}^{cc}$, because the $\oplus\mathbf{P}^{cc}$-complete problem L_{IP} needs so few alternations. Finally, we conjecture that even the class $\mathrm{BP} \cdot \oplus\mathbf{P}^{cc}$ is much "easier" than $\mathbf{PSPACE}^{cc}$, because Schöning's generalization $\mathrm{BP} \cdot \mathcal{C} \subseteq \exists \cdot \forall \cdot \mathcal{C} \cap \forall \cdot \exists \cdot \mathcal{C}$ of the classical result of Lautemann, which is easily transferred into the communication complexity context (see e.g., Babai *et al.* 1986), tells us that randomization with bounded error can be replaced with just two additional alternations.

Algorithm 1: Protocol $I_k(s, t, b)$

Input: Alice has $x = x_1 \cdots x_n$ and Bob has $y = y_1 \cdots y_n$
Data: Both know s, t, b and the odd natural number k
if $(k \geq t - s + 1)$ **then**
  begin
    `/* Trivial protocol: Alice sends her input; both compute the value by themselves. */`
    Alice and Bob compute $b' := \mathrm{IP}_{t-s+1}(x_s \cdots x_t, y_s \cdots y_t)$ using the trivial protocol;
    `/* They return 1, if` b `equals` b'`, and 0 otherwise: */`
    return $(b == b')$;
  end
else
  begin
    `/* Alice guesses the following strings and sends them to Bob: */`
    Guess existentially $S \subseteq [k], |S|$ *odd*;
    Guess universally $i \in S$;
    Guess universally $j \in \overline{S}$;
    Guess universally $h \in \{i, j\}$;
    `/* Both compute for themselves (no communication)`

$$d := t - s + 1 \ , \qquad s_1 := s + (h-1) \cdot B \ ,$$
$$B := \lceil d/k \rceil \ , \qquad t_1 := \min\{t, s + h \cdot B\} \ ,$$
$$b_1 := \begin{cases} b & , h = i \ , \\ 1 - b & , h = j \ . \end{cases}$$

    `*/`
    return $I_k(s_1, t_1, b_1)$;
  end
end

7.2. Quasi-random graphs. We investigate a new connection between communication complexity and the fascinating field of *quasi-random graphs* (see e.g., Chung *et al.* 1989). We think that problems based on adjacency questions about quasi-random graph families have high bounded-error modular communication complexity, and thus are good candidates for separating the polynomial hierarchy from polynomial space.

While Chung & Tetali (1993) have shown that high communication complexity leads to quasi-randomness, we prove that under certain conditions the opposite direction also holds. Unfortunately, we cannot prove lower bounds for the bounded-error modular communication complexity,

but we are able to show that the modular communication complexity of such problems is lower-bounded by $\log(1/P(n)) - \mathcal{O}(1)$, where $P(n)$ denotes the edge density of the graph family. Thus, known constructions of sparse quasi-random graph families like Erdös-Renyi graphs (defined below) yield many explicit problems provably outside the classes $\mathbf{MOD}_q\mathbf{P}^{\mathrm{cc}}$, q prime.

7.2.1. Basic definitions. We consider D-regular graph families $\mathcal{G} \coloneqq (G_n)_{n\geq 1}$ on N nodes, i.e., each G_n is a $D(n)$-regular graph on $N(n)$ nodes. We define the *edge density of* $\mathcal{G}$ by

$$P(n) \coloneqq \frac{|E(G_n)|}{\left|\binom{V(G_n)}{2}\right|} = 2 \cdot \frac{D(n)}{N(n)-1} \quad .$$

We only consider graph families with $P(n)N(n) \to \infty$ for $n \to \infty$. A graph family is *dense*, if $D(n) = \Theta(N(n))$, and *sparse*, if $D(n) = o(N(n))$.

Graph families define problems in communication complexity. For a graph G let EDGE_G denote the Boolean function such that the communication matrix of EDGE_G equals the adjacency matrix of G. (In other words, Alice has $x \in V(G)$, Bob has $y \in V(G)$, and they want to know if $\{x, y\} \in E(G)$.) Then, a graph family $\mathcal{G} \coloneqq (G_n)_{n\in\mathbb{N}}$ defines a family $\mathrm{EDGE}_{\mathcal{G}} \coloneqq (\mathrm{EDGE}_{G_n})_{n\in\mathbb{N}}$ of Boolean functions.

Definition 7.3 (Discrepancy). *A graph family $\mathcal{G}$ has the* discrepancy property **DISC(1)**, *if for all subsets $X, Y \subseteq V(G_n)$ we have*

$$\left|e_{G_n}(X,Y) - P(n)|X||Y|\right| = o\big(P(n)(N(n))^2\big) \quad .$$

Graph families with the discrepancy property have been thoroughly studied in the theory of *quasi-random graphs*. For space reasons, we do not make any attempt to give an introduction to this fascinating field, but we refer the reader to e.g., Chung & Graham (2002) and Krivelevich & Sudakov (2006) as possible starting points.

From here on, we call a graph family *quasi-random*, if it has the discrepancy property.

7.2.2. Almost superregular problems. In our opinion, what makes sparse quasi-random graph families amenable to high lower bounds in communication complexity are their *superregularity properties*.

Definition 7.4 (Almost superregular). *Let $A, B\colon \mathbb{N} \to \mathbb{N}$ be functions, and let $M \coloneqq (M_n)_{n\in\mathbb{N}}$ be a family of matrices $M_n\colon \mathcal{X}_n \times \mathcal{Y}_n \to \mathbb{F}$ over a field $\mathbb{F}$ such that $|\mathcal{X}_n| = |\mathcal{Y}_n| =: N(n)$. We call the family M* almost (A, B)-superregular over $\mathbb{F}$*, if for every $A(n) \times A(n)$-submatrix K of M_n we have $\mathbb{F}\text{-rank}(K) \geq B(n)$.*

Furthermore, we call a function family $(f_n)_{n\in\mathbb{N}}$ almost (A, B)-superregular over $\mathbb{F}$ *if the corresponding family $(M^{f_n})_{n\in\mathbb{N}}$ of communication matrices is almost (A, B)-superregular over $\mathbb{F}$.*

Of course, the definition above only makes sense for $B \leq A$. *Superregular matrices (over $\mathbb{F}$)* were defined in Valiant (1977) as matrices such that every quadratic submatrix has full rank (over $\mathbb{F}$). Thus, for $N(n) \coloneqq 2^n$, a family of superregular $N(n) \times N(n)$-matrices over $\mathbb{F}$ is almost (A, A)-superregular over $\mathbb{F}$ for every function A.

Valiant himself constructed integer matrices superregular over the field of rational numbers (see e.g., Lokam 2009, p. 22, Theorem 2.12). Unfortunately, families of Boolean superregular matrices over a fixed finite field do not exist.

Theorem 7.5. *For every fixed finite field $\mathbb{F}$ there do not exist families of superregular matrices.*

Proof. First of all, we consider Boolean families over a fixed finite field $\mathbb{F}$. Assume that there exists a superregular family $(M_n)_{n\in\mathbb{N}}$ of Boolean $n \times n$-matrices. We define $K(n) \coloneqq n$, and $N(n) \coloneqq 2^{K(n)+2} \log K(n)$. Choose an arbitrary natural number $n_0 \geq 21$. The matrix $W \coloneqq M_{N(n_0)}$ is an edge coloring of the biclique $K_{N(n_0),N(n_0)}$. Bipartite Ramsey theory (see e.g., the result of Conlon 2008 or prior work) tells us that in this big biclique there exists a small biclique $K_{K(n_0),K(n_0)}$

that is monochromatic under W. This means that W contains an $n_0 \times n_0$-submatrix T consisting of zeros only, or ones only. Thus, $\mathbb{F}$-rank$(T) \in \{0, 1\}$ in contradiction to the superregularity property implying $\mathbb{F}$-rank$(T) = n_0 \geq 21$.

The same argument can be applied to arbitrary matrices over $\mathbb{F}$. Then, one has to use $|\mathbb{F}|$-colorings instead of 2-colorings and the bound N is higher. □

In contrast, almost (A, B)-superregular matrix families do exist for certain functions A and B over every finite field. As the reader might have already guessed, such families are given by the adjacency matrices of sparse quasi-random graph families.

Almost superregularity over the field of real numbers can be elegantly proven via spectral techniques. For this, we define

Definition 7.6 (Matrix norms). *Let A be a complex $n \times n$-matrix.*

(i) *The* spectral norm of A *is defined as* $||A|| \coloneqq \max_{x \neq 0} ||Ax||/||x||$.

(ii) *The* Frobenius norm of A *is defined as* $||A||_{\mathrm{F}} \coloneqq \sqrt{\sum_{i,j} |A_{i,j}|^2}$.

Definition 7.7 (Approximate Hamming weight). *Let $\theta > 0$ be a real number. For a Boolean $n \times n$-matrix A we define the* θ-approximate Hamming weight of A, $\widetilde{\mathrm{wt}}_\theta(A)$, *as the minimum Hamming weight of a $(\theta n) \times (\theta n)$-submatrix of A.*

In other words, we consider all $(\theta n) \times (\theta n)$-submatrices, count the number of ones in them, and take the minimum.

Lemma 7.8. *Let $f \coloneqq (f_n)_{n \in \mathbb{N}}$, $f_n \colon \mathbb{B}^n \times \mathbb{B}^n \to \mathbb{B}$, be a family of Boolean functions, and let $N(n) \coloneqq 2^n$. Then for every constant real number $\theta > 0$ the family f is almost $\left(\theta N, \frac{\widetilde{\mathrm{wt}}_\theta\left(M^{f_n}\right)}{||M^{f_n}||^2}\right)$-superregular over $\mathbb{R}$.*

Proof. A basic fact from linear algebra (see e.g., Lokam 2001) is that for every submatrix B of a matrix A we have $\mathbb{R}$-rank$(B) \geq ||B||_{\mathrm{F}}^2/||A||^2$. Note that for a Boolean matrix B we have $||B||_{\mathrm{F}}^2 = \mathrm{wt}(B)$. □

Theorem 7.9. *Let $\mathcal{G} \coloneqq (G_n)_{n \in \mathbb{N}}$ be a D-regular quasi-random graph family on N nodes with edge density P. For every constant real number $\theta > 0$ the family* $\mathrm{EDGE}_{\mathcal{G}}$ *is almost $\left(\theta N, \Omega\left(\theta^2/P\right)\right)$-superregular over $\mathbb{R}$.*

Proof. Let $M_n \coloneqq A^{G_n}$. First of all, $||M_n|| = D(n)$, because G_n is $D(n)$-regular. Let B be a $(\theta N(n)) \times (\theta N(n))$-submatrix of M_n that realizes $\widetilde{\mathrm{wt}}_\theta(M_n)$. By the discrepancy property, we have

$$\begin{aligned} \mathrm{wt}(B) &= (1 + o(1)) \cdot P(n) \cdot (\theta N(n))^2 \\ &\approx 2\theta^2 D(n) N(n) \ \text{, for } n \text{ sufficiently large.} \end{aligned}$$

Applying Lemma 7.8 yields the lower bound. □

The following theorem shows that the result above can be generalized to any field, in particular, finite fields. Of course, it is proved in a different way, because we do not have spectral techniques over finite fields.

THEOREM 7.10. *Let $\mathbb{F}$ be a field, and let $\mathcal{G} \coloneqq (G_n)_{n\in\mathbb{N}}$ be a D-regular quasi-random graph family on N nodes with edge density P. For every constant real number $\theta > 0$ the family* $\mathrm{EDGE}_{\mathcal{G}}$ *is almost $\left(\theta N, \Omega\left(\theta^2/P\right)\right)$-superregular over $\mathbb{F}$.*

PROOF. Let n be sufficiently large. Let $M_n \coloneqq A^{G_n}$, and define $A \coloneqq \theta N$. Consider an arbitrary $A(n) \times A(n)$-submatrix T of M_n. We want to show that T has a high $\mathbb{F}_q$-rank. Let U and W be the subsets of $V(G_n)$ of size $A(n)$ that correspond to the rows and columns of T, respectively. By the discrepancy property of $\mathcal{G}$ we have

$$e_{G_n}(U,W) \approx P(n) \cdot (A(n))^2 \approx 2\theta^2 D(n) N(n) \ .$$

Let $W' \subseteq W$ contain all $v \in W$ such that $e_{G_n}(U,v) \geq 1$. Then $|W'| \approx 2\theta^2 N(n)$, because of

$$|W'| D(n) \geq e_{G_n}(U,W') = e_{G_n}(U,W) \approx 2\theta^2 D(n) N(n) \ .$$

Let T' be the submatrix of T, where the columns are restricted to W'. We successively permute rows and columns of T' in order to obtain a "stair" of ones starting with the stairhead in the upper left corner and going down, where each stair has length $\leq D(n)$. Hence, the number of stairs is at least $\approx (2\theta^2 N(n))/D(n)$ implying that $B(n) \coloneqq \mathbb{F}\text{-rank}(T) \geq \mathbb{F}\text{-rank}(T') \approx 2\theta^2 \frac{D(n)}{N(n)}$, for n sufficiently large. We conclude that the family $\mathrm{EDGE}_{\mathcal{G}}$ is almost (A,B)-superregular over $\mathbb{F}$.

Now, we permute T': Take the first column $v_1 \in W'$. By definition of W' there exists a row $u_1 \in U$ that is a neighbor of v_1. Take u_1 as the new first row. It has $t_1 \leq D(n)$ neighbors $v_1, \ldots, v_{t_1}$ in W'. Permute the columns such that these neighbors form the first t_1 columns of T'. We created the first stair. Again, take another column $v_{t_1+1} \in W' - \{v_1, \ldots, v_{t_1}\}$ and continue this process to create the next stairs. This can be done at least $|W'|/D(n)$ many times. □

7.2.3. Lower bounds. The results obtained so far yield strong lower bounds for worst-case deterministic and modular communication complexity.

Given two function families $f \coloneqq (f_n)_{n\in\mathbb{N}}$ and $g \coloneqq (g_n)_{n\in\mathbb{N}}$, we call $f_n \colon \mathcal{X}'_n \times \mathcal{Y}'_n \to \mathcal{Z}_n$ a *large subfunction* of $g_n \colon \mathcal{X}_n \times \mathcal{Y}_n \to \mathcal{Z}_n$, if there exists a constant real number $\theta > 0$ such that f_n is the restriction of g_n to $\mathcal{X}'_n \times \mathcal{Y}'_n$ for sets $\mathcal{X}'_n \subseteq \mathcal{X}_n$, $|\mathcal{X}'_n| \geq \theta|\mathcal{X}_n|$, and $\mathcal{Y}'_n \subseteq \mathcal{Y}_n$, $|\mathcal{Y}'_n| \geq \theta|\mathcal{Y}_n|$, respectively.

THEOREM 7.11. *For a quasi-random D-regular graph family $\mathcal{G} \coloneqq (G_n)_{n\in\mathbb{N}}$ with edge density P we have*

$$\mathrm{D}(\mathrm{EDGE}_{G_n}) \geq \log\left(\frac{1}{P(n)}\right) - \mathcal{O}(1) \ . \tag{7.12}$$

This also holds for every family of large subfunctions of $\mathrm{EDGE}_{\mathcal{G}}$.

PROOF. Follows from Theorem 7.9 and Fact 3.34 (real-rank lower bound). □

Interestingly, the right hand side looks like an entropic quantity.

This lower bound cannot be tight for worst-case deterministic communication complexity, because it is actually a lower bound for modular communication complexity.

THEOREM 7.13. *Let q be a prime. For a quasi-random D-regular graph family $\mathcal{G} \coloneqq (G_n)_{n\in\mathbb{N}}$ with edge density P we have*

$$\mathrm{MOD}_q(\mathrm{EDGE}_{G_n}) \geq \frac{1}{q-1} \cdot \log\left(\frac{1}{P(n)}\right) - \mathcal{O}(1) \ . \tag{7.14}$$

This also holds for every family of large subfunctions of $\mathrm{EDGE}_{\mathcal{G}}$.

PROOF. Follows from Theorem 7.10 and Fact 3.36 (modular-rank lower bound). □

There is a variety of constructions of sparse quasi-random graph families that have appeared in the literature. We exemplify our lower-bound method with the so-called Erdös-Renyi graphs that arise from finite geometries.

Definition 7.15 (Erdös-Renyi graphs). *Let p be a prime power. We define the p-th* Erdös-Renyi graph*, $\mathcal{ER}_p$, as follows: Let $V(\mathcal{ER}_p)$ be the points of the projective plane over $\mathbb{F}_p$. Nodes $x = (x_0, x_1, x_2)$ and $y = (y_0, y_1, y_2)$ are adjacent iff we have $x_0y_0 + x_1y_1 + x_2y_2 = 0$ in $\mathbb{F}_p$.*

Fact 7.16. *The p-th Erdös-Renyi graph has $|V(\mathcal{ER}_p)| = (p^3 - 1)/(p - 1) = p^2 + p + 1$ many nodes. It is $D(p)$-regular with $D(p) \coloneqq (p^2 - 1)/(p - 1)$, and it has the discrepancy property for $\tilde{P}(p) \coloneqq (p + 1)/(p^2 + p + 1)$.*

As $P(2n) \approx \tilde{P}(N(n)) = \Theta(1/N(n))$, we obtain

Corollary 7.17. *For a prime q,* $\mathrm{MOD}_q(\mathrm{EDGE}_{\mathcal{ER}_{2^n}}) \geq \frac{n}{q-1} - \mathcal{O}(1)$.

We note that similar high lower bounds can be obtained for explicit families based on *Delsarte-Goethals-Turyn graphs*, *generalized Erdös-Renyi graphs* (defined over the projective geometry of dimension $t \geq 2$), certain *incidence graphs* of *generalized m-gons*, *Ramanujan graphs*, or *projective-norm graphs*, see (Krivelevich & Sudakov 2006, p. 22–29) for details.

We expect that such lower bounds also hold for bounded-error modular communication complexity, because these games on graphs are hard everywhere, i.e., the modular communication complexity stays high even when restricting to a large subfunction. We leave this as an open question.

Open Question 7.18. *Let $\mathcal{G} \coloneqq (G_n)_{n\in\mathbb{N}}$ be a quasi-random D-regular graph family with edge density P. Do we have*

$$\mathrm{BP} \cdot \mathrm{MOD}^{\mathrm{pub}}_{q,\epsilon}(\mathrm{EDGE}_{G_n}) \geq \Omega\left(\log \frac{1}{P(n)}\right) \ ?$$

Note the following trivial upper bound (see e.g., Linial & Shraibman 2009b, Thm. 35)

$$\mathrm{BP} \cdot \mathrm{MOD}^{\mathrm{pub}}_{q,\epsilon}(\mathrm{EDGE}_{G_n}) \leq \mathrm{R}^{\mathrm{pub}}_{\epsilon}(\mathrm{EDGE}_{G_n}) \leq \mathcal{O}(\log D(n))$$

that holds for every D-regular graph family $\mathcal{G} \coloneqq (G_n)_{n\in\mathbb{N}}$. Therefore, possible candidates outside the polynomial hierarchy must not be too sparse.

Acknowledgement

I would like to express my deep gratitude to Martin Dietzfelbinger, Uwe Schöning and Jacobo Torán for many fruitful discussions and their support. In addition, I would like to sincerely thank Stefan Arnold for his careful reading of the manuscript and helpful comments.

References

Noga Alon, Peter Frankl & Vojtech Rödl (1985). Geometrical Realization of Set Systems and Probabilistic Communication Complexity. In *26th Annual Symposium on Foundations of Computer Science, 21–23 October 1985, Portland, Oregon, USA*, 277–280. IEEE Computer Society.

Dana Angluin (1980). On Counting Problems and the Polynomial-Time Hierarchy. *Theor. Comput. Sci.* **12**, 161–173.

László Babai, Peter Frankl & Janos Simon (1986). Complexity classes in communication complexity theory (preliminary version). In *27th Annual Symposium on Foundations of Computer Science, FOCS 1986, 27–29 October 1986, Toronto, Ontario, Canada*, 337–347. IEEE Computer Society.

José L. Balcázar, Josep Díaz & Joaquim Gabarró (1990). *Structural Complexity II.* Texts in Theoretical Computer Science, An EATCS Series. Springer-Verlag, 1st edition.

José L. Balcázar, Josep Díaz & Joaquim Gabarró (1995). *Structural Complexity I.* Texts in Theoretical Computer Science, An EATCS Series. Springer-Verlag, 2nd edition.

Paul Beame, Toniann Pitassi, Nathan Segerlind & Avi Wigderson (2006). A Strong Direct Product Theorem for Corruption and the Multiparty Communication Complexity of Disjointness. *Computational Complexity* **15**(4), 391–432.

Richard Beigel & John Gill (1992). Counting Classes: Thresholds, Parity, Mods, and Fewness. *Theor. Comput. Sci.* **103**(1), 3–23.

Harry Buhrman, Nikolai K. Vereshchagin & Ronald de Wolf (2007). On Computation and Communication with Small Bias. In *22nd Annual IEEE Conference on Computational Complexity (CCC 2007), 13–16 June 2007, San Diego, California, USA*, 24–32. IEEE Computer Society.

Mahdi Cheraghchi (2005). On Matrix Rigidity and the Complexity of Linear Forms. *Electronic Colloquium on Computational Complexity (ECCC)* (070).

Benny Chor & Oded Goldreich (1988). Unbiased Bits from Sources of Weak Randomness and Probabilistic Communication Complexity. *SIAM J. Comput.* **17**(2), 230–261.

Fan Chung & Ronald Graham (2002). Sparse quasi-random graphs. *Combinatorica* **22**, 217–244.

Fan R. K. Chung, Ronald L. Graham & Richard M. Wilson (1989). Quasi-random graphs. *Combinatorica* **9**(4), 345–362.

Fan R. K. Chung & Prasad Tetali (1993). Communication Complexity and Quasi Randomness. *SIAM J. Discrete Math.* **6**(1), 110–123.

Bruno Codenotti (2000). Matrix rigidity. *Linear Algebra and its Applications* **304**(1–3), 181–192.

Bruno Codenotti, Pavel Pudlák & Giovanni Resta (2000). Some structural properties of low-rank matrices related to computational complexity. *Theor. Comput. Sci.* **235**(1), 89–107.

David Conlon (2008). A new upper bound for the bipartite Ramsey problem. *Journal of Graph Theory* **58**(4), 351–356.

Carsten Damm, Matthias Krause, Christoph Meinel & Stephan Waack (2004). On relations between counting communication complexity classes. *J. Comput. Syst. Sci.* **69**(2), 259–280.

Reinhard Diestel (2005). *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, 3rd edition.

Martin Dietzfelbinger & Henning Wunderlich (2007). A characterization of average case communication complexity. *Inf. Process. Lett.* **101**(6), 245–249.

Ding-Zhu Du & Ker-I Ko (2000). *Theory of Computational Complexity*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., 1st edition.

Jürgen Forster (2002). A linear lower bound on the unbounded error probabilistic communication complexity. *J. Comput. Syst. Sci.* **65**(4), 612–625.

Lance Fortnow (1997). *Counting Complexity*, 81–107. In Selman & Hemaspaandra (1997).

Joel Friedman (1993). A note on matrix rigidity. *Combinatorica* **13**(2), 235–239.

Bernd Halstenberg & Rüdiger Reischuk (1990). Relations between Communication Complexity Classes. *J. Comput. Syst. Sci.* **41**(3), 402–429.

Lane A. Hemaspaandra & Mitsunori Ogihara (2002). *The Complexity Theory Companion*. Texts in Theoretical Computer Science, An EATCS Series. Springer-Verlag.

Juraj Hromkovic (2000). *Communication Complexity and Parallel Computing*. Texts in Theoretical Computer Science – An EATCS Series. Springer-Verlag.

Bala Kalyanasundaram & Georg Schnitger (1992). The Probabilistic Communication Complexity of Set Intersection. *SIAM J. Discrete Math.* **5**(4), 545–557.

Hartmut Klauck (2001). Lower Bounds for Quantum Communication Complexity. In *42nd Annual Symposium on Foundations of Computer Science, October 14–17, Las Vegas Nevada, USA*, 288–297. IEEE.

Hartmut Klauck (2003). Rectangle Size Bounds and Threshold Covers in Communication Complexity. In *18th Annual IEEE Conference on Computational Complexity,7–10 July 2003, Aarhus, Denmark*, 118–134. IEEE Computer Society.

Johannes Köbler, Uwe Schöning & Jacobo Torán (1993). *The Graph Isomorphism Problem – Its Structural Complexity*. Birkhäuser Boston.

Michael Krivelevich & Benny Sudakov (2006). Pseudo-random graphs. In *More sets, graphs and numbers*, E. Győri, G. O. H. Katona & Laszlo Lovász, editors, volume 15 of *Bolyai Soc. Math. Studies*, 199–262. Springer-Verlag.

Eyal Kushilevitz & Noam Nisan (1997). *Communication Complexity*. Cambridge University Press.

Nathan Linial & Adi Shraibman (2009a). Learning Complexity vs Communication Complexity. *Combinatorics, Probability & Computing* **18**(1–2), 227–245.

Nati Linial & Adi Shraibman (2007). Lower bounds in communication complexity based on factorization norms. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing, San Diego, California, USA, June 11–13, 2007*, David S. Johnson & Uriel Feige, editors, 699–708. ACM.

Nati Linial & Adi Shraibman (2009b). Lower bounds in communication complexity based on factorization norms. *Random Struct. Algorithms* **34**(3), 368–394.

Satyanarayana V. Lokam (1996). Spectral Methods for Matrix Rigidity with Applications to Size-Depth Trade-offs and Communication Complexity. Electronically available at CiteSeerX, http://130.203.133.150/viewdoc/summary?doi=10.1.1.52.4411.

Satyanarayana V. Lokam (2000). On the rigidity of Vandermonde matrices. *Theor. Comput. Sci.* **237**(1–2), 477–483.

Satyanarayana V. Lokam (2001). Spectral Methods for Matrix Rigidity with Applications to Size-Depth Trade-offs and Communication Complexity. *J. Comput. Syst. Sci.* **63**(3), 449–473.

Satyanarayana V. Lokam (2006). Quadratic Lower Bounds on Matrix Rigidity. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 15–20, 2006, Proceedings*, Jin-Yi Cai, S. Barry Cooper & Angsheng Li, editors, volume 3959 of *Lecture Notes in Computer Science*, 295–307. Springer-Verlag.

Satyanarayana V. Lokam (2009). Complexity Lower Bounds using Linear Algebra. *Foundations and Trends in Theoretical Computer Science* **4**(1-2), 1–155.

Kurt Mehlhorn & Erik Meineche Schmidt (1982). Las Vegas Is better than Determinism in VLSI and Distributed Computing (Extended Abstract). In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing, 5–7 May 1982, San Francisco, California, USA*, 330–337. ACM.

Gatis Midrijanis (2005). Three lines proof of the lower bound for the matrix rigidity. *CoRR* **abs/cs/0506081**.

Rajeev Motwani & Prabhakar Raghavan (1995). *Randomized Algorithms*. Cambridge University Press.

Ilan Newman (1991). Private vs. Common Random Bits in Communication Complexity. *Inf. Process. Lett.* **39**(2), 67–71.

Noam Nisan & Avi Wigderson (1995). On Rank vs. Communication Complexity. *Combinatorica* **15**(4), 557–565.

Alon Orlitsky & Abbas El Gamal (1990). Average and randomized communication complexity. *IEEE Transactions on Information Theory* **36**(1), 3–16.

Christos H. Papadimitriou & Stathis Zachos (1983). Two remarks on the power of counting. In *Theoretical Computer Science, 6th GI-Conference, Dortmund, Germany, January 5–7, 1983, Proceedings*, Armin B. Cremers & Hans-Peter Kriegel, editors, volume 145 of *Lecture Notes in Computer Science*, 269–276. Springer-Verlag.

Pavel Pudlák (1994). Communication in Bounded Depth Circuits. *Combinatorica* **14**(2), 203–216.

Pavel Pudlák & Vojtech Rödl (1994). Some combinatorial-algebraic problems from complexity theory. *Discrete Mathematics* **136**(1–3), 253–279.

Ran Raz (1995). Fourier Analysis for Probabilistic Communication Complexity. *Computational Complexity* **5**(3/4), 205–221.

Ran Raz & Boris Spieker (1993). On the "log rank"-Conjecture in Communication Complexity. In *34th Annual Symposium on Foundations of Computer Science, FOCS 1993, 3–5 November 1993, Palo Alto, California, USA*, 168–176. IEEE Computer Society.

Alexander Razborov (1989). On Rigid Matrices (in Russian). Technical report, Steklov Mathematical Institute. Electronically available at http://people.cs.uchicago.edu/~razborov/rigid.pdf.

Alexander A. Razborov (1992). On the Distributional Complexity of Disjointness. *Theor. Comput. Sci.* **106**(2), 385–390.

Alexander A. Razborov & Alexander A. Sherstov (2008). The Sign-Rank of AC^O. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25–28, 2008, Philadelphia, PA, USA*, 57–66. IEEE Computer Society.

Uwe Schöning (1986). *Complexity and Structure*, volume 211 of *Lecture Notes in Computer Science*. Springer-Verlag.

Uwe Schöning (1988). The power of counting. In Selman (1988), 204–223.

Uwe Schöning (1989). Probabilistic Complexity Classes and Lowness. *J. Comput. Syst. Sci.* **39**(1), 84–100.

Uwe Schöning (1991). Recent Highlights in Structural Complexity Theory (invited talk). In *SOFSEM'91, Nizké Tratry (CSFR)*, Conference Proceedings, 205–216. Springer-Verlag.

Alan L. Selman (editor) (1988). *Complexity Theory Retrospective*. Foundations of Computing. Springer-Verlag.

Alan L. Selman & Lane A. Hemaspaandra (editors) (1997). *Complexity Theory Retrospective II*. Foundations of Computing. Springer-Verlag.

Alexander A. Sherstov (2008a). Communication Complexity under Product and Nonproduct Distributions. In *Proceedings of the 23rd Annual IEEE Conference on Computational Complexity, CCC 2008, 23–26 June 2008, College Park, Maryland, USA*, 64–70. IEEE Computer Society.

Alexander A. Sherstov (2008b). Halfspace Matrices. *Computational Complexity* **17**(2), 149–178.

Mohammad Amin Shokrollahi, Daniel A. Spielman & Volker Stemann (1997). A Remark on Matrix Rigidity. *Inf. Process. Lett.* **64**(6), 283–285.

Seinosuke Toda (1990). *Computational Complexity of Counting Complexity Classes*. Ph.D. thesis, Dept. of Comput. Sci. & Inf. Mat., Univ. of Electro-Commun., Tokyo.

Seinosuke Toda (1991). PP is as Hard as the Polynomial-Time Hierarchy. *SIAM J. Comput.* **20**(5), 865–877.

Leslie G. Valiant (1977). Graph-Theoretic Arguments in Low-Level Complexity. In *Mathematical Foundations of Computer Science 1977, 6th Symposium, Tatranska Lomnica, Czechoslovakia, September 5–9, 1977, Proceedings*, Jozef Gruska, editor, volume 53 of *Lecture Notes in Computer Science*, 162–176. Springer-Verlag.

Leslie G. Valiant & Vijay V. Vazirani (1986). NP is as Easy as Detecting Unique Solutions. *Theor. Comput. Sci.* **47**(3), 85–93.

Umesh V. Vazirani (1987). Strong communication complexity or generating quasirandom sequences form two communicating semi-random sources. *Combinatorica* **7**(4), 375–392.

Ronald de Wolf (2006). Lower Bounds on Matrix Rigidity Via a Quantum Argument. In *Automata, Languages and Programming, 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10–14, 2006, Proceedings, Part I*, Michele Bugliesi, Bart Preneel, Vladimiro Sassone & Ingo Wegener, editors, volume 4051 of *Lecture Notes in Computer Science*, 62–71. Springer-Verlag.

Henning Wunderlich (2009). On Toda's Theorem in Structural Communication Complexity. In *SOFSEM 2009: Theory and Practice of Computer Science, 35th Conference on Current Trends in Theory and Practice of Computer Science, Spindleruv Mlýn, Czech Republic, January 24–30, 2009. Proceedings*, Mogens Nielsen, Antonín Kucera, Peter Bro Miltersen, Catuscia Palamidessi, Petr Tuma & Frank D. Valencia, editors, volume 5404 of *Lecture Notes in Computer Science*, 609–620. Springer-Verlag.

Andrew Chi-Chih Yao (1979). Some Complexity Questions Related to Distributive Computing (Preliminary Report). In *Conference Record of the Eleventh Annual ACM Symposium on Theory of Computing, 30 April–2 May, 1979, Atlanta, Georgia, USA*, 209–213. ACM.

Andrew Chi-Chih Yao (1983). Lower Bounds by Probabilistic Arguments (Extended Abstract). In *24th Annual Symposium on Foundations of Computer Science, FOCS 1983, 7–9 November 1983, Tucson, Arizona, USA*, 420–428. IEEE Computer Society.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de