Randomized feasible interpolation and monotone circuits with a local oracle

Jan Krajíček

Faculty of Mathematics and Physics
Charles University in Prague

Abstract

The feasible interpolation theorem for semantic derivations from K. (1997) [16] allows to derive from some short semantic derivations (e.g. in resolution) of the disjointness of two NP sets U and V a small communication protocol (a general dag-like protocol in the sense of K. (1997) [16]) computing the Karchmer-Wigderson multi-function $KW[U,V]$ associated with the sets, and such a protocol further yields a small circuit separating U from V. When U is closed upwards the protocol computes the monotone Karchmer-Wigderson multi-function $KW^m[U,V]$ and the resulting circuit is monotone. K. (1998) [18] extended the feasible interpolation theorem to a larger class of semantic derivations using the notion of a real communication complexity (e.g. to the cutting planes proof system CP).

In this paper we generalize the method to a still larger class of semantic derivations by allowing randomized protocols. We also introduce an extension of the monotone circuit model, monotone circuits with a local oracle (CLOs), that does correspond to communication protocols for $KW^m[U,V]$ making errors. The new randomized feasible interpolation thus shows that a short semantic derivation (from a certain class of derivations larger than in the original method) of the disjointness of U, V, U closed upwards, yields a small randomized protocol for $KW^m[U,V]$ and hence a small monotone CLO separating the two sets.

This research is motivated by the open problem to establish a lower bound for proof system R(LIN/$\mathbf{F}_2$) operating with clauses formed by linear Boolean functions over $\mathbf{F}_2$. The new randomized feasible interpolation applies to this proof system and also to (the semantic versions of) cutting planes CP, to small width resolution over CP of K. (1998) [17] (system R(CP)) and to random resolution RR of Buss, Kolodziejczyk and Thapen [5]. The method does not yield yet lengths-of-proofs lower bounds; for this it is necessary to establish lower bounds for randomized protocols or for monotone CLOs.

Consider a propositional proof system R(LIN/$\mathbf{F}_2$) that operates with clauses of linear equations over $\mathbf{F}_2$ and combines the rules of both resolution and linear

equational calculus. A line C in a proof has the form

$$\{f_1, \ldots, f_k\}$$

with $f_i \in \mathbf{F}_2[x_1, \ldots, x_n]$ linear polynomials and the intended meaning is that an assignment $x := a \in \{0,1\}^n$ to variables makes C true if and only if one of $f_i = 1$ becomes true, i.e. the truth value of C is computed by Boolean formula

$$\bigvee_{i \leq k} f_i$$

in the language with $\bigvee, \oplus, 0, 1$. We often leave the outside brackets $\{,\}$ out when writing clauses. For $L \subseteq C$ define $\sum L := \sum_{f \in L} f$.

The rules of R(LIN/$\mathbf{F}_2$) are the following four:

$$\frac{}{h, h+1} \qquad \frac{C}{C, f} \qquad \frac{C, 0}{C} \qquad \frac{C, g \quad C, h}{C, g+h+1} \ .$$

We shall call the rules $\mathbf{F}_2$-*axiom*, *weakening*, *contraction* and *the binary rule*, respectively. This proof system (albeit defined slightly differently but polynomially equivalently, denoted Res-Lin there) has been considered already by Itsykson and Sokolov [10] who proved an exponential lower bound for tree-like proofs. They also showed that the semantic version of the system (in the sense of semantic derivations of [16]) is p-equivalent to the syntactic version, whether tree-like or dag-like. This paper is motivated by the problem to establish a lower bound for unrestricted (i.e. dag-like) R(LIN/$\mathbf{F}_2$) proofs.

Proof systems combining resolution or, more generally, logical reasoning with algebraic reasoning were considered earlier by several authors: [17] defined proof systems $R(CP)$ and $LK(CP)$ extending cutting plane by a logic reasoning and proved an exponential lower bound for a subsystem of $R(CP)$, Hirsch and Kojevnikov [8, 13] considered resolution over a system for linear programing and Kojevnikov [13] improved upon a bound in [17]. Raz and Tzameret [26] studied resolution over linear equations with integral coefficients and proved a lower bound for a class of its proofs, and Alekhnovich et.al. [1] defined polynomial calculus with resolution PCR which extends PC in a way that incorporates resolution (lines of proofs are polynomials, however).

There is also a link to the well-known open problem to establish lower bounds for constant depth Frege systems in DeMorgan language augmented by a connective counting modulo a prime, the so called $AC^0[p]$-Frege systems. The strongest subsystem of such a system for which a lower bound is known is a low degree polynomial calculus operating with polynomials formed from AC^0-formulas, [15]. The lower bound problem for R(LIN/$\mathbf{F}_2$) seems interesting also because the top proof system is logical. Note that Buss, Kolodziejczyk and Zdanowski [6] proved that, in fact, the $AC^0[p]$-Frege system collapses (with a quasi-polynomial blow-up in proof size) to a proof system operating with clauses of conjunctions of low degree polynomials.

Our approach is to use feasible interpolation for semantic derivations from [16] but we need to generalize it first to allow small errors. The generalization we

develop here allows randomized communication protocols with errors (protocols in the sense of [16]) for computing the Karchmer-Wigderson multi-function. Protocols making no errors correspond to separating circuits but protocols with errors do not yield separating circuits making some error. Instead we introduce an extension of the circuit model, circuits with a local oracle (CLO), that does correspond to protocols with errors.

Tree-like protocols with errors for $KW^m[U, V]$ yield monotone separating formulas with a local oracle and subsume the ordinary Karchmer-Wigderson (1988) [12] protocols pictured as binary trees. A lower bound in this case is known (cf. [9, 18] for examples based on the bipartite perfect matching problem and Hall's theorem). Further, monotone CLOs efficiently simulate monotone real circuits (Section 6) and any two disjoint sets can be separated by a small non-monotone CLO (Lemma 2.3 and the remark at the end of Section 3). To establish a lower bound for monotone CLOs separating two NP sets, one closed upwards, is an open problem.

To be able to apply randomized feasible interpolation to R(LIN/$\mathbf{F}_2$) we use the approximation method of Razborov [30] and Smolensky [33] in order to reduce the linear width (defined in Section 4) in a general not too long proof at the expense of introducing an error (cf. Section 5). The new method may have further applications and, in particular, it applies to the semantic versions of cutting planes CP, to small width resolution over cutting planes R(CP), and to random resolution RR. The method on its own does not yield yet lengths-of-proofs lower bounds; for this it is necessary to establish lower bounds for randomized protocols or for monotone CLOs. Some partial results about monotone CLOs are obtained in [21].

The paper is organized as follows. Section 1 recalls some notions and results from [16]. In Section 2 we define the concept of randomized protocols and use it to formulate randomized feasible interpolation. In Section 3 we introduce circuits with a local oracle (CLO) and prove that they correspond to protocols with errors and that, in particular, randomized protocols yield CLOs. In Section 4 we introduce the linear width of R(LIN/$\mathbf{F}_2$) proofs and discuss the case when it is small. Randomized feasible interpolation is proved for R(LIN/$\mathbf{F}_2$) in Section 5 and for CP and small width R(CP) in Section 6. The lower bound problem for monotone CLOs (and hence for randomized protocols computing the monotone Karchmer-Wigderson multi-function for some pair of sets) is discussed in Section 7. The paper is concluded by a few remarks in Section 8. A proof complexity background can be found in [14, 24].

1 Feasible interpolation preliminaries

The general feasible interpolation theorem from [16] for semantic derivations uses communication complexity. One considers two disjoint NP sets $U, V \subseteq \{0,1\}^n$ and the Karchmer-Wigderson multi-function whose valid values on a pair $(u, v) \in U \times V$ is any coordinate in which u, v differ. The aim is to extract from a short proof of the disjointness of U, V some upper bound on the computational

complexity of this multi-function in some computational model. Proving then a computational complexity lower bound for the model allows to infer a length-of-proofs lower bound. The original set-up (and the one most frequently used) derives from the proof data the existence of a small circuit separating U and V. In the monotone case one can use then known strong lower bounds for monotone circuits, for example Alon and Boppana [2].

When the construction of [16] is applied to tree-like proofs it leads to familiar protocols for communication that are pictured as binary trees, cf.[12]. However, for applications to general, dag-like, proofs one needs a more general notion of a protocol defined in [16, Def.2.2]. The key fact, allowing to prove some lower bounds, is that similarly as small tree-like communication protocols correspond to small formulas separating U and V (by Karchmer and Wigderson [12]), the more general protocols used in [16] correspond to small separating circuits.

Let us now recall formally relevant definitions and facts from [16]. A multi-function on $U \times V$ with values in some set $I \neq \emptyset$ is a ternary relation $R \subseteq U \times V \times I$ such that for all $(u, v) \in U \times V$ there is $i \in I$ such that $R(u, v, i)$. Some value for (u, v) from its domain can be computed by two players, one receiving u and the other one v, exchanging bits of information until they agree on a valid value i. The communication complexity of R, $CC(R)$, is the minimal number of bits they need to exchange (in an optimal protocol) in the worst case.

The Karchmer-Wigderson multi-function $KW[U, V]$ of a particular interest is defined for two disjoint sets $U, V \subseteq \{0,1\}^n$: a valid value of $KW(u, v)$ on pair $(u, v) \in U \times V$ is any $i \in [n]$ such that $u_i \neq v_i$. The monotone version of this function $KW^m[U, V]$ is defined when U is closed upwards (or V downwards) and a valid value on (u, v) is any $i \in [n]$ such that $u_i = 1 \wedge v_i = 0$.

Given two disjoint $U, V \subseteq \{0,1\}^n$ and $R \subseteq U \times V \times I$ a multi-function, [16, Def.2.2] defines a protocol for R to be a 4-tuple $\mathbf{P} = (G, \mathsf{lab}, F, S)$ satisfying the following conditions:

(P1) G is a directed acyclic graph that has one source (the in-degree 0 node called the *root*) denoted $\emptyset$.

(P2) The nodes with the out-degree 0 are *leaves* and they are labelled by the mapping lab by elements of I.

(P3) $S(u, v, x)$ is a function (the *strategy*) that assigns to a node $x \in G$ and a pair $u \in U$ and $v \in V$ node $S(u, v, x)$ accessible by an edge from x.

(P4) For every $u \in U$ and $v \in V$, $F(u, v) \subseteq G$ is a set (called the *consistency condition*) satisfying:

  (a) $\emptyset \in F(u, v)$,

  (b) $x \in F(u, v) \rightarrow S(u, v, x) \in F(u, v)$,

  (c) if $x \in F(u, v)$ is a leaf and $\mathsf{lab}(x) = i$, then $R(u, v, i)$ holds.

We say that $\mathbf{P}$ is *tree-like* iff G is a tree.

The complexity of $\mathbf{P}$ is measured by its *size*, which is the cardinality of G, and by the following notion: The *communication complexity* of $\mathbf{P}$, denoted $CC(\mathbf{P})$, is the minimal t such that for every $x \in G$ the communication complexity for the players (one knowing u and x, the other one v and x) to decide $x \in_? F(u,v)$ or to compute $S(u,v,x)$ is at most t.

The interpolation theorem in [16] was formulated using the notion of a *semantic derivation* ([16, Def. 4.1]): A sequence of sets $D_1, \ldots, D_k \subseteq \{0,1\}^N$ is a semantic derivation of D_k from $A_1, \ldots, A_m \subseteq \{0,1\}^N$ if each D_i is either one of A_j's or contains $D_{j_1} \cap D_{j_2}$, for some $j_1, j_2 < i$. A semantic derivation is a *refutation* of $A_1, \ldots, A_m$ iff $D_k = \emptyset$.

We shall introduce now a general set-up for our investigation of interpolation and we shall refer to it the whole paper. We assume the following conditions for parameters and sets, and introduce the following notation:

$$N = n + s + r \ , \ N, n \geq 1 \ . \tag{1}$$

$$A_1, \ldots, A_m \subseteq \{0,1\}^{n+s} \ \text{ and } \ B_1, \ldots, B_\ell \subseteq \{0,1\}^{n+r} \ . \tag{2}$$

From the total N variables, n represent an input a from $\{0,1\}^n$, s variables represent a potential witness b for the membership of a in U and r variables represent a potential witness c for the membership of a in V (U and V are defined below). For $A \subseteq \{0,1\}^{n+s}$ define

$$\tilde{A} := \bigcup_{(a,b) \in A} \{(a,b,c) \mid c \in \{0,1\}^r\} \tag{3}$$

and for $B \subseteq \{0,1\}^{n+r}$ define:

$$\tilde{B} := \bigcup_{(a,c) \in B} \{(a,b,c) \mid b \in \{0,1\}^s\} \ . \tag{4}$$

where a, b, c range over $\{0,1\}^n$, $\{0,1\}^s$ and $\{0,1\}^r$, respectively. Define:

$$U = \{u \in \{0,1\}^n \mid \exists b \in \{0,1\}^s; (u,b) \in \bigcap_{j \leq m} A_j\} \tag{5}$$

and

$$V = \{v \in \{0,1\}^n \mid \exists c \in \{0,1\}^r; (v,c) \in \bigcap_{j \leq \ell} B_j\} \ . \tag{6}$$

We shall also refer to the following monotonicity condition. For all $u, u' \in \{0,1\}^n$ and $b \in \{0,1\}^s$:

$$(u,b) \in \bigcap_{j \leq m} A_j \wedge u' \geq u \longrightarrow (u',b) \in \bigcap_{j \leq m} A_j \ . \tag{7}$$

The complexity of sets in a semantic derivation is measured by the following notion of (monotone) communication complexity of subsets of $\{0,1\}^N$ defined

in [16]. For $D \subseteq \{0,1\}^N$, $u, v \in \{0,1\}^n$, $q^u \in \{0,1\}^s$ and $r^v \in \{0,1\}^r$ consider four tasks:

1. Decide whether $(u, q^u, r^v) \in D$.
2. Decide whether $(v, q^u, r^v) \in D$.
3. If $(u, q^u, r^v) \in D \not\equiv (v, q^u, r^v) \in D$ find $i \leq n$ such that $u_i \neq v_i$.
4. If $(u, q^u, r^v) \in D$ and $(v, q^u, r^v) \notin D$ either find $i \leq n$ such that

$$u_i = 1 \wedge v_i = 0$$

or decide that there is some u' satisfying

$$u' \geq u \wedge (u', q^u, r^v) \notin D \ .$$

The *communication complexity* $CC(D)$ of D is the minimal t such that the tasks 1.-3. can be solved by the players, one knowing u, q^u and the other one knowing v, r^v, exchanging at most t bits. The *monotone communication complexity w.r.t.* U of D, denoted $\mathrm{MCC}_U(D)$, is the minimal $t \geq CC(D)$ such that also the task 4. can be solved by the players exchanging at most t bits.

Now we are ready to recall a fact about the existence of protocols from the proof of [16, Thm.5.1].

Theorem 1.1 ([16])

Assume the set-up conditions (1)-(6) and assume that $\pi = D_1, \ldots, D_k$ is a semantic refutation of the sets $\tilde{A}_1, \ldots, \tilde{A}_m, \tilde{B}_1, \ldots, \tilde{B}_\ell$. Let $t \geq 1$ be such that $t \geq CC(D_i)$ for all $i \leq k$.

Then there is a protocol for $KW[U, V]$ of size $k + 2n$ and of communication complexity $O(t)$. The protocol has k inner vertices, the sets in π, and additional $2n$ vertices, the leaves, labelled by all possible formulas $u_i = 1 \wedge v_i = 0$ and $u_i = 0 \wedge v_i = 1$.

If condition (7) is also satisfied and $MCC_U(D_i) \leq t$ for all $i \leq k$ then there is a protocol for $KW^m[U, V]$ of size $k + n$ and of communication complexity $O(t)$.

Further, the consistency condition F is defined in both the monotone and the non-monotone cases identically as:

$$D \in F(u, v) \quad \textit{iff} \quad (v, q^u, r^v) \notin D$$

for D in π, and

$$x \in F(u, v) \quad \textit{iff} \quad \mathsf{lab}(x) \textit{ is valid for } u, v$$

for x a leaf.

Moreover, if π is tree-like, so is G.

2 Randomized feasible interpolation for semantic derivations

First we generalize protocols to allow a randomization and some error.

Definition 2.1 *A* randomized protocol *for multi-function $R \subseteq U \times V \times I$ with error $\epsilon > 0$ is a random variable $(\mathbf{P_r})_\mathbf{r}$ where each $\mathbf{P_r}$ is a 4-tuple satisfying conditions (P1), (P2), (P3) and (P4a) defining protocols and instead of conditions (P4b) and (P4c) it satisfies:*

(P4b') For every $(u,v) \in U \times V$,

$$Prob_\mathbf{r}[\exists x,\ x \in F_\mathbf{r}(u,v) \wedge S_\mathbf{r}(u,v,x) \notin F_\mathbf{r}(u,v)] \ \leq \ \epsilon \ .$$

(P4c') For every $(u,v) \in U \times V$,

$$Prob_\mathbf{r}[\exists leaf\ x,\ x \in F_\mathbf{r}(u,v) \wedge \mathsf{lab}_\mathbf{r}(x) = i \wedge \neg R(u,v,i)] \ \leq \ \epsilon \ .$$

The size of $(\mathbf{P_r})_\mathbf{r}$ is $\max_\mathbf{r} size(\mathbf{P_r})$ and the communication complexity of $(\mathbf{P_r})_\mathbf{r}$ is $\max_\mathbf{r} CC(\mathbf{P_r})$. We say that $(\mathbf{P_r})_\mathbf{r}$ is tree-like if each $\mathbf{P_r}$ is.

We note a simple observation.

Lemma 2.2 *For any randomized protocol $(\mathbf{P_r})_\mathbf{r}$ for multi-function $R \subseteq U \times V \times I$ of size S, communication complexity t and error ϵ there exists a randomized protocol $(\tilde{\mathbf{P}}_\mathbf{r})_\mathbf{r}$ for multi-function R of size at most $2S$ (with at most S leaves), communication complexity at most $3t$ and error ϵ such that $(P4b)$ never fails, i.e. the probability in $(P4b')$ is 0.*

Proof :

Introduce for each inner node $x \in G_\mathbf{r}$ a new leaf node $\tilde{x}$, label it arbitrarily (e.g. $u_1 = 1 \wedge v - 1 = 0$), and define a new strategy $\tilde{S}_\mathbf{r}$ that first checks if

$$x \in F_\mathbf{r}(u,v) \rightarrow S_\mathbf{r}(u,v,x) \in F_\mathbf{r}(u,v)$$

is true and if so it uses $S_\mathbf{r}$, otherwise it sends x into $\tilde{x}$ and the failure of the condition is the definition of $\tilde{x} \in \tilde{F}_\mathbf{r}(u,v)$.

q.e.d.

In connections with interpolation we are interested in the situation when the multi-function is the Karchmer-Wigderson one. It makes sense to consider only the monotone case $KW^m[U,V]$ as the next lemma recalls.

Lemma 2.3 (Raz and Wigderson [27]) *Let U, V be any two disjoint subsets of $\{0,1\}^n$. Then for any $\epsilon > 0$ there is a tree-like randomized protocol $(\mathbf{P_r})_\mathbf{r}$ computing $KW[U,V]$ of size $S = (n + \epsilon^{-1})^{O(1)}$, communication complexity $t = O(\log n + \log(\epsilon^{-1}))$ and error ϵ.*

In particular, for $\epsilon = n^{-\Omega(1)}$ the size is $S = n^{O(1)}$ and the communication complexity is $t = O(\log n)$.

Proof :

A randomized protocol computing $KW[U,V]$ is determined by $\log(\epsilon^{-1})$ subsets $I \subseteq [n]$. The players exchange the parity of the bits in their respective strings belonging to the first such I, then to the second, etc. until they find I for which the parity differs. Then they find a valid value for $KW[U,V]$ by binary search. If they do not find such I, they declare an error. This gives a randomized protocol of size polynomial in n, ϵ^{-1}, with communication complexity $2(\log n + \log(\epsilon^{-1}))$, and error ϵ.

q.e.d.

Now we introduce a notion that we will use in the context of semantic derivations. Let $X \in \{0,1\}^N$ and let $\mathcal{Y} = (Y_{\mathbf{r}})_{\mathbf{r}}$ be a random distribution on subsets of $\{0,1\}^N$, and let $\delta > 0$. We say that $\mathcal{Y}$ is a *δ-approximation* of X iff for all $w \in \{0,1\}^N$:

$$\mathrm{P}rob_{\mathbf{r}}[w \in X \triangle Y_{\mathbf{r}}] \ \leq \ \delta$$

where $X \triangle Y$ is the symmetric difference.

Working in the set-up (1)-(6) the sets X and $Y_{\mathbf{r}}$ are subsets of $\{0,1\}^N$ and the definitions of CC and MCC_U apply to them. With this in mind we further define that the *(monotone) communication complexity* of $\mathcal{Y}$ is at most t if this is true for all $Y_{\mathbf{r}}$, and that the *δ-approximate (monotone) communication complexity* of X is at most t if there is a δ-approximation $\mathcal{Y}$ of X with this property.

Theorem 2.4

Assume the set-up conditions (1)-(7). Let $\pi : D_1, \ldots, D_k = \emptyset$ be a semantic refutation of sets $\tilde{A}_1, \ldots, \tilde{A}_m, \tilde{B}_1, \ldots, \tilde{B}_\ell$ such that the δ-approximate monotone communication complexity of every D_i is at most t.

Then there is a randomized protocol $(\mathbf{P}_{\mathbf{r}})_{\mathbf{r}}$ for $KW^m[U,V]$ of size at most $k+n$, communication complexity $O(t)$ and of error at most $3\delta k$.

Moreover, if the refutation π is tree-like then also $(\mathbf{P}_{\mathbf{r}})_{\mathbf{r}}$ is tree-like.

Proof :

Take the protocol $\mathbf{P} = (G, \mathsf{lab}, F, S)$ provided by Theorem 1.1. Its strategy S and the consistency condition F are defined in terms of sets D_i. In particular, for any $(u,v) \in U \times V$ and $x \in G$ an inner node, both the value of $S(u,v,x)$ and the truth value of $x \in F(u,v)$ are defined from at most 3 truth values of statements of the form $(u, q^u, r^v) \in D_i$ or $(v, q^u, r^v) \in D_i$ for some specific indices $i \leq k$ determined by x, where q^u and r^v depend just on u and v, respectively.

Not knowing anything about the monotone communication complexity of the sets D_i we cannot estimate the communication complexity of $\mathbf{P}$. At this point we use the δ-approximations of the sets D_i. If $(E^i_{\mathbf{s}})_{\mathbf{s}}$ are δ-approximations of D_i, $i \leq k$, let the space of samples $\mathbf{r}$ for $\mathbf{P}_{\mathbf{r}}$ be the product of the sample spaces of these k δ-approximations and define $S_{\mathbf{r}}$ and $F_{\mathbf{r}}$ as S and F before but using the particular sets $E^i_{\mathbf{s}}$ (with $\mathbf{s}$ determined by $\mathbf{r}$) in place of the sets D_i. In particular, $D_i \in F_{\mathbf{r}}(u,v)$ iff $(v, q^u, r^v) \notin E^i_{\mathbf{s}}$. Further, put $G_{\mathbf{r}} := G$ and $\mathsf{lab}_{\mathbf{r}} := \mathsf{lab}$.

For any given $(u,v) \in U \times V$ and $x \in G$ the (truth) value of $S_{\mathbf{r}}$ and $F_{\mathbf{r}}$ differs from S and F respectively with probability at most 3δ. Hence for (u,v) the error in conditions (P4b') and (P4c ') is at most $\epsilon := 3\delta k$.

q.e.d.

We describe yet another type of semantic refutations that also yields randomized protocols.

Theorem 2.5 *Assume the set-up conditions (1)-(7). Let $e \geq 1$, $\epsilon > 0$ and let $(C_{\mathbf{r}})_{\mathbf{r}}$ be a random distribution on $(\mathcal{P}(\{0,1\}^N))^e$, $C_{\mathbf{r}} = (C^1_{\mathbf{r}}, \ldots, C^e_{\mathbf{r}})$, such that each $\bigwedge_{i \leq e}(C^i_{\mathbf{r}})_{\mathbf{r}}$ is an ϵ-approximation of $\{0,1\}^N$.*

Assume that for all samples $\mathbf{r}$ there is a semantic refutation $\pi_{\mathbf{r}}$ of

$$\tilde{A}_1, \ldots, \tilde{A}_m, \tilde{B}_1, \ldots, \tilde{B}_\ell, C^1_{\mathbf{r}}, \ldots, C^e_{\mathbf{r}}$$

with k lines, and such that the monotone communication complexity of all sets in $\pi_{\mathbf{r}}$ is at most t.

Then there is a randomized protocol for $KW^m[U,V]$) of size at most $k+n+e \leq 2k+n$, communication complexity $O(t)$ and of error at most ϵ.

Moreover, if the refutation π is tree-like then also $(\mathbf{P}_{\mathbf{r}})_{\mathbf{r}}$ is tree-like.

Proof :

The construction of $\mathbf{P} = (G, \mathsf{lab}, F, S)$ in Theorem 1.1 yields G whose inner nodes correspond to lines of the refutation and leaves are extra n nodes. The construction uses the fact that for $(u,v) \in U \times V$ the strings q^u and r^v are chosen so that $(u, q^u, r^v) \in \tilde{A}_i$ and $(v, q^u, r^v) \in \tilde{B}_j$ for all $i \leq m$ and $j \leq \ell$. In particular, each initial set $\tilde{A}_1, \ldots, \tilde{A}_m, \tilde{B}_1, \ldots, \tilde{B}_\ell$ contains either (u, q^u, r^v) or (v, q^u, r^v).

In the presence of the new initial clauses $C^i_{\mathbf{r}}$ this is no longer true and it may happen that both (u, q^u, r^v) and (v, q^u, r^v) are outside of some $C^i_{\mathbf{r}}$.

We define $\mathbf{P}_{\mathbf{r}}$ as follows. Each $G_{\mathbf{r}}$ has e extra leaves y_i labelled arbitrarily (say $u_1 = 1 \wedge v_1 = 0$ for the definiteness) and the strategy $S_{\mathbf{r}}(u,v,x)$ sends node x corresponding to $C^i_{\mathbf{r}}$ to y_i if

$$(v, q^u, r^v) \notin C^i_{\mathbf{r}} \tag{8}$$

and the same condition defines when $y_i \in F_{\mathbf{r}}(u,v)$.

As $\bigwedge_{i \leq e}(C^i_{\mathbf{r}})_{\mathbf{r}}$ is an ϵ-approximation of $\{0,1\}^N$, (8) happens with probability at most ϵ in total.

q.e.d.

3 Monotone circuits with a local oracle

Our aim in this section is to define a generalization of the circuit model that corresponds to protocols with errors computing $KW^m[U,V]$. We restrict ourselves to the monotone case due to Lemma 2.3 (see also the remark at the end of this section).

A monotone *circuit with a local oracle* (monotone CLO, briefly) separating U from V is determined by the following data:

1. a monotone Boolean circuit $D(x_1, \ldots, x_n, y_1, \ldots, y_e)$ with inputs $\overline{x}$ and $\overline{y}$,
2. a set $\mathcal{R}$ of combinatorial rectangles $U_j \times V_j \subseteq U \times V$, for $j \leq e$, called *oracle rectangles* of the CLO,

and satisfying the following condition:

3. for all monotone Boolean functions $f_j : \{0,1\}^n \to \{0,1\}$, $j \leq e$, such that
$$f_j(U_j) \subseteq \{1\} \quad \text{and} \quad f_j(V_j) \subseteq \{0\}$$
the function
$$C(\overline{x}) \ := \ D(\overline{x}, f_1(\overline{x}), \ldots, f_e(\overline{x}))$$
separates U from V:
$$C(U) = \{1\} \quad \text{and} \quad C(V) = \{0\} \ .$$

The *size* of the CLO is the size of D and its *locality* is
$$\frac{|\bigcup_{j \leq e} U_j \times V_j|}{|U \times V|}$$
(we assume both U, V are non-empty). Note that C defines a monotone Boolean function for any choice of monotone functions f_j.

The proof of the following lemma expands a bit upon a proof by Razborov [31].

Lemma 3.1 *Assume that $(\mathbf{P_r})_\mathbf{r}$ is a randomized protocol for $KW^m[U,V]$ of size s, communication complexity t and error ϵ.*

Then there is a monotone circuit with a local oracle separating U from V of size $s2^{O(t)}$ and locality ϵ.

Proof :

Assume $(\mathbf{P_r})_\mathbf{r}$ is a randomized protocol satisfying the hypothesis of the lemma, with $\mathbf{P_r} = (G_\mathbf{r}, \mathsf{lab}_\mathbf{r}, F_\mathbf{r}, S_\mathbf{r})$. By Lemma 2.2 we may assume that each $G_\mathbf{r}$ makes errors only in leaves, i.e. violates possibly only the condition (P4c) of Section 1 in the sense of (P4c') of Definition 2.1. This may increase the size and the communication complexity proportionally but that does not change the form $s2^{O(t)}$ of the upper bound.

By averaging there must be some sample $\mathbf{r}$ such that $\mathbf{P_r}$ makes an error for at most ϵ-part of all pairs $U \times V$. Fix one such protocol $(G, \mathsf{lab}, F, S) := \mathbf{P_r}$ for the rest of the proof. We may also assume that the communication of the players deciding that a leaf a is in $F(u,v)$ ends with each player sending the value of the i-th bit of u or v, respectively, where $i = \mathsf{lab}(a)$. That is, they both know at the end whether an error occurred for (u,v) and the set of these erroneous pairs for which $a \in F(u,v)$ is a disjoint union of combinatorial rectangles.

For a vertex a of G and a string $w \in \{0,1\}^t$ denote:

- $R_{a,w}$ the rectangle $U_{a,w} \times V_{a,w}$, some $U_{a,w} \subseteq U$ and $V_{a,w} \subseteq V$, of pairs $(u,v) \in U \times V$ such that the communication of the players deciding $a \in_? F(u,v)$ evolves according to w and ends with the affirmation of the membership,

- k_a: the number of nodes in G that can be reached from node a by a directed path (so $k_a = 1$ for a a leaf, while $k_\emptyset \leq s$ for the root $\emptyset$).

Assume

$$R_1 := U_1 \times V_1, \ldots, R_e := U_e \times V_e \ , \quad \text{for } j \leq e \tag{9}$$

enumerate all rectangles $R_{a,w}$ where a is a leaf and $(u,v) \in R_{a,w}$ iff $a \in F(u,v)$ and the players decided this with communication w but $\mathsf{lab}(a)$ is incorrect for (u,v), i.e. an error occurs for (u,v) at a.

Claim 1: *For all $a \in G$ and $w \in \{0,1\}^t$ there is a size $\leq k_a 2^{O(t)}$ monotone circuit with a local oracle separating $U_{a,w}$ from $V_{a,w}$ such that its oracle rectangles are included among (9). The constant implicit in the exponent $O(t)$ is independent of a.*

For a, w we shall denote by $D_{a,w}, \mathcal{R}_{a,w}$ a monotone CLO that is claimed to exists; the set $\mathcal{R}_{a,w}$ is the set of its oracle rectangles. We shall establish the claim by induction on k_a.

If $k_a = 1$, a is a leaf. Take arbitrary rectangle $U_{a,w} \times V_{a,w}$. Either $i = \mathsf{lab}(a)$ is correct on the rectangle, then $D_{a,w}$ is just the input x_i and $\mathcal{R}_{a,w} = \emptyset$, or not, and then $D_{a,w} = y_j$ and $\mathcal{R}_{a,w} = \{R_j\}$ where $U_{a,w} \times V_{a,w}$ is R_j in the enumeration (9).

Assume $k_a > 1$ and let $w \in \{0,1\}^t$. For $u \in U_{a,w}$ let $u^* \in \{0,1\}^{4^t}$ be a vector whose bits u^*_ω are parameterized by $\omega = (\omega_1, \omega_2) \in \{0,1\}^t \times \{0,1\}^t$ and such that:

- $u^*_\omega = 1$ iff there is a $v \in V_{a,w}$ such that the communication of the players computing $S(u,v,a)$ evolves according to ω_1 and the computation of $S(u,v,a) \in_? F(u,v)$ evolves according to ω_2 (note that it has to end with the affirmation that $S(u,v,a) \in F(u,v)$.

Define $v^*_\omega \in \{0,1\}^{4^t}$ dually:

- $v^*_\omega = 0$ iff there is a $u \in U_{a,w}$ such that the communication of the players computing $S(u,v,a)$ evolves according to ω_1 and the computation of $S(u,v,a) \in_? F(u,v)$ evolves according to ω_2.

Let $U^*_{a,w}$ and $V^*_{a,w}$ be the sets of all these vectors u^* and v^*, respectively.

Claim 2: *There is a monotone formula $\varphi_{a,w}$ in 4^t variables z_{ω_1,ω_2} and of size $2^{O(t)}$ separating $U^*_{a,w}$ from $V^*_{a,w}$.*

Claim 2 follows from a theorem of Karchmer and Wigderson [12]: the players can find a coordinate ω in which $u^*_\omega = 1$ and $v^*_\omega = 0$ by first computing $S(u,v,a)$ (getting thus ω_1) and then deciding $S(u,v,a) \in_? F(u,v)$ (obtaining thus ω_2). The strings $u \in U_{a,w}, v \in V_{a,w}$ yielding u^*, v^* need not to be unique but that is

not needed; it suffices that each player has a canonical way to pick one such u or v, respectively.

For $\omega_1 \in \{0,1\}^t$ let a_{ω_1} be the node $S(u,v,a)$ computed for some u, v with communication ω_1. Then define a monotone circuit with a local oracle by setting:

$$D_{a,w} \ := \ \varphi_{a,w}(\ldots,\ z_{\omega_1,\omega_2}/D_{a_{\omega_1},\omega_2},\ldots)$$

and:

$$\mathcal{R}_{a,w} := \bigcup_{(\omega_1,\omega_2)} \mathcal{R}_{a_{\omega_1},\omega_2} \ .$$

As $k_{a_{\omega_1}} < k_a$, the induction hypothesis implies that all $D_{a_{\omega_1},\omega_2}$ work correctly on all $U_{a_{\omega_1},\omega_2} \times V_{a_{\omega_1},\omega_2}$. Thus, by the definition of the formula $\varphi_{a,w}$, the circuit $D_{a,w}$ works also correctly.

This concludes the proof of Claim 1 and of the theorem (which follows from the claim by taking for a the root of G). The bound ϵ to the locality comes from our choice to start with a protocol making an error for at most an ϵ-part of $U \times V$.

q.e.d.

It may be worthwhile to remark that the oracle rectangles of the CLO constructed in the proof can be divided into $O(s)$ blocks (corresponding to different leaves) such that the rectangles in each block are disjoint (they correspond to different communication histories).

The particular CLO is constructed from a particular $\mathbf{P_r}$ chosen by averaging. However, we could construct a CLO for each $\mathbf{P_r}$ and instead of estimating the locality of the one CLO estimate the probability that a pair (u,v) gets into an oracle rectangle. We do not pursue this generality further here but we state it formally as it may play a role in an eventual lower bound argument for randomized protocols.

Lemma 3.2 *Assume that $(\mathbf{P_r})_\mathbf{r}$ is a randomized protocol for $KW^m[U,V]$ of size s, communication complexity t and error ϵ.*

Then there is a distribution $(C_\mathbf{r})_\mathbf{r}$ over monotone circuits with a local oracle separating U from V, each of size $s2^{O(t)}$ and such that for any pair $(u,v) \in U \times V$:

$$Prob_\mathbf{r}[(u,v) \ \text{ is in an oracle rectangle of } C_\mathbf{r}] \ \leq \ \epsilon \ .$$

The next two lemmas establish a form of converse of Lemma 3.1. Let U^{min} be the set of $\leq$-minimal elements of U and V^{max} the set of $\leq$-maximal elements of V. In particular, no two elements of U^{min} (or of V^{max}), respectively, are comparable and hence any partial Boolean function on U^{min} (or on V^{max}) can be extended to a monotone one on $U^{min} \cup V^{max}$.

Lemma 3.3 *Assume $D, \{U_j \times V_j\}_{j \leq e}$ is a monotone CLO separating U^{min} from V^{max}, of size s and locality μ.*

Then there is a protocol (G, lab, F, S) *for* $KW^m[U^{min}, V^{max}]$ *of size* s*, communication complexity* 2 *and making an error for at most* $s \cdot \mu^{1/2}$*-part of* $U^{min} \times V^{max}$*.*

Proof :

For each $j \leq e$, the measure of $U_j \times V_j$ in $U^{min} \times V^{max}$ is less than μ and hence

(i) either $|U_j|/|U^{min}| < \mu^{1/2}$,

(ii) or $|V_j|/|V^{max}| < \mu^{1/2}$.

Define a monotone Boolean function f_j that is identically 1 on U_j, identically 0 on V_j, and for a string from $\{0,1\}^n \setminus (U_j \cup V_j)$ it equals to 0 in the case (i) or to 1 in the case (ii).

Put $C(\overline{x}) := D(\overline{x}, f_1(\overline{x}), \ldots, f_e(\overline{x}))$. Define a protocol (G, lab, F, S) as follows:

- the vertices of G are the nodes of D, the root is the output node and the edges lead from a node of D to its two input nodes,
- for a node a of G corresponding to a subcircuit E of D, define the consistency condition by:

  $$a \in F(u,v) \text{ iff } (E(u, f_1(u), \ldots, f_e(u)) = 1 \wedge E(v, f_1(v), \ldots, f_e(v)) = 0) ,$$

- the strategy finds an input into E that is also in $F(u,v)$,
- the labeling lab assigns to input nodes x_i of D the value i and to input nodes y_j an arbitrary value, say 1.

An error can occur only at the labeling of the input nodes corresponding to a variable y_j. Because such a node is in $F(u,v)$, it must hold that $f_j(u) = 1$ and $f_j(v) = 0$. In both cases (i) and (ii) considered in the definition of f_j the measure of the rectangle of such pairs (u,v) is less than $\mu^{1/2}$ and there are at most $e \leq s$ of them. This proves the lemma.

q.e.d.

Next we show, for the sake of a completeness of the discussion, that one can get a better estimate of the error of the protocol if one allows Boolean functions (and circuits) to have also a third value between 0 and 1. Denote the third value 1/2 and define the conjunction and the disjunction on $\{0, 1/2, 1\}$ as the minimum and the maximum, respectively. Call such functions and circuits *3-valued*. We shall say that $D, \{U_j \times V_j\}_{j \leq e}$ is a monotone 3-valued CLO separating U form V if the condition 3. in the definition of the CLO is obeyed even w.r.t. to all monotone 3-valued functions f_j.

Lemma 3.4 *Assume $D, \{U_j \times V_j\}_{j \leq e}$ is a monotone 3-valued CLO separating U^{min} from V^{max}, of size s a locality μ.*

Then there is a protocol (G, lab, F, S) for $KW^m[U^{min}, V^{max}]$ of size s, communication complexity 2 and making an error for at most μ-part of $U \times V$.

Proof :

The construction of (G, lab, F, S) is similar to that in the proof of Lemma 3.3 but we define the functions f_j differently: f_j equals to 1 on U_j, to 0 on V_j and to 1/2 everywhere else.

With this definition the analysis at the end when an error occurs for a pair (u, v) at a node corresponding to y_j leads as before to a rectangle of (u, v) such that $f_j(u) = 1 \wedge f_j(v) = 0$ but that is now simply $U_j \times V_j$. Hence the measure of the set of pairs for which an error occurs is at most the locality of the CLO.

q.e.d.

Let us conclude the section with a couple of remarks. The first one is that monotone CLOs simulate efficiently monotone real circuits of [23] (circuits allowing any non-decreasing real functions at gates); we shall show this in Lemma 6.2. The second remark[1] is that general, non-monotone, CLOs are very strong: any two disjoint subsets of $\{0,1\}^n$ can be separated by a polynomial size CLO (in fact, a formula with a local oracle) with polynomially small locality. This is seen as follows: take the randomized protocol from Lemma 2.3 and turn it into a non-monotone (dropping in the definition the condition of monotonicity of oracle functions f_j) CLO of size $poly(n, \epsilon^{-1})$ and locality ϵ separating U from V by the construction of Lemma 3.1.

4 Linear width

The *linear width* of an R(LIN/$\mathbf{F}_2$)-clause C is the number of fs in it; we shall denote it lw(C). For a set Φ of R(LIN/$\mathbf{F}_2$)-clauses denote by $\Phi \vdash_w C$ the fact that C can be derived in R(LIN/$\mathbf{F}_2$) from Φ by a proof whose all lines have linear width at most w.

When the linear width is small the clauses have small communication complexity (in the sense of Section 1) and Theorem 1.1 yields a small monotone protocol and that yields lower bounds (cf. [16, Sec.7]).

Unfortunately, general R(LIN/$\mathbf{F}_2$) refutations need not to have small linear width. It is easy to prove a lower bound on the linear width of an R(LIN/$\mathbf{F}_2$) refutation by translating it into a polynomial calculus PC refutation and by appealing to degree lower bounds for that system. In particular, to an R(LIN/$\mathbf{F}_2$)-clause $C = \{f_1, \ldots, f_k\}$ assign polynomial over $\mathbf{F}_2$ $p_C := \Pi_{i \leq k}(1 - f_i)$: C is satisfied by $a \in \{0,1\}^n$ iff $p_C(a) = 0$. An R(LIN/$\mathbf{F}_2$)-refutation π of a set Φ of R(LIN/$\mathbf{F}_2$) clauses can be then straightforwardly translated into a PC refutation π' of the set of polynomials

$$p_C \ , \quad C \in \Phi$$

[1] I owe this remark to Igor C. Oliveira.

such that the degree of π' is bounded above by the linear width of π. In particular, the weakening rule and the binary rule translate into the multiplication and the addition rules of PC, respectively.

To illustrate this lower bound argument let us consider as a specific example the set $\neg PHP_n$ of R(LIN/$\mathbf{F}_2$) clauses:

- $1 - x_{ij}, x_{kj}$, for $i \neq k$ and any j,
- $1 - x_{ij}, x_{ik}$, for any i and $j \neq k$,
- $\sum_j x_{ij}$, any i,

with variables x_{ij}, $i \in [n+1], j \in [n]$. The linear width of these clauses is 1. However, the set of polynomials p_C for $C \in \neg PHP_n$ is precisely the set for which the degree $n/2$ lower bound for PC refutations was established by Razborov [32].

We shall employ the approximation method in Section 5 to reduce in a sense the linear width. This construction introduces, however, some error into derivations (modelled in one of the constructions by new initial clauses to be called $Ax(\pi, \mathbf{r})$) and this prevents the simple reduction to PC described above.

5 Randomized feasible interpolation for R(LIN/$\mathbf{F}_2$)

In this section we use the Razborov-Smolensky approximation method [30, 33] to reduce in a sense the linear width of not too large R(LIN/$\mathbf{F}_2$) refutations.

Theorem 5.1 *Assume the set-up conditions (1)-(7) and assume that sets $A_1, \ldots, A_m, B_1, \ldots, B_\ell$ are defined by R(LIN/$\mathbf{F}_2$)-clauses.*

Let π be an R(LIN/$\mathbf{F}_2$) refutation of (the clauses defining) these sets with k steps. Let $w \geq 1$ be any parameter.

Then there is a randomized protocol for $KW^m[U, V]$) of size at most $k + n$, communication complexity $O(w \log n)$ and of error at most $3 \cdot 2^{-w}k$.

Moreover, if the refutation π is tree-like then also G is tree-like.

Proof :

Let D be any $R(LIN)$-clause, i.e. a clause formed by some linear polynomails. Following [30, 33] define a 2^{-w}-approximation $(Y_{\mathbf{s}})_{\mathbf{s}}$ of D by the following process:

- Using the sample $\mathbf{s}$ pick independently at random $L_1, \ldots, L_w \subseteq D$,
- put $Y_{\mathbf{s}}$ to be the set defined by $\bigvee_{j \leq w} \sum L_j$,

($\sum L_j$ is the sum of all linear polynomials in L_j).

Claim: *Let D be an R(LIN/$\mathbf{F}_2$)-clause of linear width w. Then $MCC_U(D) = O(w \log n)$.*

Let us write the w linear functions forming D in a matrix form as:

$$Ax + By + Cz + E \ .$$

The U-player sends Au and Bq^u and the V-player sends Av and Cr^v, $4w$ bits in total. After this they know the truth values of $(u, q^u, r^v) \in D$ and $(v, q^u, r^v) \in D$ and if they differ they can use the binary search on a differing row in Au and Av to find i for which $u_i \neq v_i$ ($2logn$ bits in total).

It remains to estimate the communication complexity of the task 4. from the definition of MCC_U under the assumption that $(u, q^u, r^v) \in D$ and $(v, q^u, r^v) \notin D$, i.e.:

$$Au + Bq^u + Cr^v + E \neq \overline{0} \ \text{ and } \ Av + Bq^u + Cr^v + E = \overline{0} \ .$$

In particular, $Au \neq Av$.

The players will attempt to put A in a reduced-row echelon form but by a specific process. The U-player sends $i_1 \in [n]$ ($\log n$ bits) such that $u_{i_1} = 0$ and the x_{i_1}-column in A is non-zero. The players then both separately transform A using the elementary row and column operations in some canonical way to a unique matrix A^1 whose first column corresponds to x_{i_1} and $A^1_{1,1} = 1$ and all other entries in the first column are 0.

In the second step they apply the same process to A^1, not using x_{i_1}. That is, the U-player sends $\log n$ bits identifying some $i_2 \in [n]$, $i_2 \neq i_1$, such that $u_{i_2} = 0$ and the x_{i_2}-column in A^1 has a non-zero element in one of the rows $2, \ldots, w$. Then they again separately transform A^1 into A^2 with the first two columns corresponding to x_{i_1} and x_{i_2} and the left-upper corner 2×2 submatrix being the identity matrix I_2 and all other entries in the first two columns being 0.

They proceed analogously as long as it is possible. Two cases may occur:

(i) After $t \leq w$ steps A^t is in the row-reduced echelon form: the left-upper corner $t \times t$ submatrix being the identity matrix I_t and all other entries in the first t columns being 0, and all rows $t+1, \ldots, w$ are zero.

(ii) After some step $t < w$ A^t is not in the row-reduced echelon form but the U-player has nothing to choose: there is no $i \neq i_1, \ldots, i_t$ such that the x_i-column in A^t has a non-zero element in one of the rows $t+1, \ldots, w$ and $u_i = 0$.

In Case (i) we can switch the values of some u_i, $i \in \{i_1, \ldots, i_t\}$, from 0 to 1 to get $u' \geq u$ such that $A^t u' = A^t v$ and hence $(u', q^u, r^v) \notin D$.

In Case (ii) the rows $t+1, \ldots, w$ need not to be zero but $A_{ij} \neq 0$ for $i, j > t$ implies that $u_i = 1$ (thinking of the i-th column as corresponding to x_i). If for one such i $v_i = 0$, the V-player sends the $\log n$ bits to identify it; they found i such that $u_i = 1 \wedge v_i = 0$. If all such $v_i = 1$ then $Gu = Gv$ where G is the $(w-t) \times n$ matrix consisting of the last $w - t$ rows of A^t. Writing the first t rows of A^t as (I_t, H), where H is a $t \times (n-t)$ matrix, we see we can find

some $u' \geq u$ changing only some u_i, $i \in \{i_1, \ldots, i_t\}$, from 0 to 1 such that $(I_t, H)u' = (I_t, H)v$ and hence also $A^t u' = A^t v$ and $(u', q^u, r^v) \notin D$.

In all cases the players solved the task 4. and they exchanged $O(w \log n)$ bits at most.

Applying Theorem 2.4 concludes the proof of the theorem.

q.e.d.

We now give an alternative proof of the randomized feasible interpolation for R(LIN/$\mathbf{F}_2$), referring to Theorem 2.5 this time. It is more laborious and gives somewhat worse bounds on the size of the resulting protocols but it may be useful in connections with the problem of resolution over low degree polynomial calculus that we shall discuss in the Section 8, and it also puts R(LIN/$\mathbf{F}_2$) in a direct relation with the random R of [5] (see Section 8).

Let π be an R(LIN/$\mathbf{F}_2$) refutation of $\Phi := A_1, \ldots, A_m, B_1, \ldots, B_\ell$ and let $w \geq 1$ be a parameter to be specified later. In this situation we perform the following random process $\mathbf{r}$ and transform π to an R(LIN/$\mathbf{F}_2$) refutation $\pi(\mathbf{r})$ of Φ extended by a set $Ax(\pi, \mathbf{r})$ of extra clauses:

1. For each $C \in \pi$ pick independently at random subsets $L_1, \ldots, L_\ell \subseteq C$ and form clause $C^{\mathbf{r}} := \{\sum L_1, \ldots, \sum L_\ell\}$.

2. For each $C \in \pi$, $C = f_1, \ldots, f_k$, add to the set $Ax(\pi, \mathbf{r})$ the following k clauses:
$$C^{\mathbf{r}}, f_j + 1 \quad \text{, for } j = 1, \ldots, k \ .$$

3. Transform π into $\pi(\mathbf{r})$, following the construction below, summarized in Lemma 5.2.

Clauses in 2. formalize that $f_j = 1$ implies that $C^{\mathbf{r}} = 1$. Before we describe $\pi(\mathbf{r})$ we need to establish a few simple facts.

Claim 1: *For all assignments $a \in \{0,1\}^n$: $C^{\mathbf{r}}(a) = 1$ implies $C(a) = 1$. For any $a \in \{0,1\}^n$ the probability that $C^{\mathbf{r}}(a) = 0 \wedge C(a) = 1$ is at most 2^{-w}.*

Claim 2: *(a) For any g, h: $g + h \vdash_2 g, h$.*
(b) For any $C \in \pi$ and $g \in C^{\mathbf{r}}$: $g \vdash_{|C|} C$.

In part (a): derive from $g + h$ clause $g, g + h$ and also an $\mathbf{F}_2$-axiom $g, g + 1$ from which g, h follows by the binary rule and contraction. In part (b): if $g = f_{j_1} + \ldots + f_{j_v}$ use part (a) to derive from g clause $f_{j_1} + \ldots + f_{j_{v-1}}, f_{j_v}$, and then repeat this to remove from the sum all f_js to get the clause $f_{j_1}, \ldots, f_{j_v}$ from which C follows by the weakening rule.

Claim 3: *Let $C \in \pi$, $C = f_1, \ldots, f_k$, and let $g = f_{j_1} + \ldots + f_{j_v}$ be an arbitrary sum of a non-empty subset of C (i.e. not necessarily in $C^{\mathbf{r}}$). Then*

$$Ax(\pi, \mathbf{r}), \{g\} \ \vdash_{w+3} C^{\mathbf{r}} \ .$$

By Claim 2(a) derive in linear width 2 from g clause $f_{j_1} + \ldots + f_{j_{v-1}}, f_{j_v}$ and combine this by the binary rule and contraction with clause $f_{j_v} + 1, C^{\mathbf{r}}$ from $Ax(\pi, \mathbf{r})$ to get

$$f_{j_1} + \ldots + f_{j_{v-1}}, C^{\mathbf{r}}$$

in linear width bounded by $w + 3$. Then repeat the same process to remove from the sum polynomials $f_{j_{v-1}}, f_{j_{v-2}}, \ldots, f_{j_1}$ to end up just with $C^{\mathbf{r}}$.

Claim 4: *Assume*

$$\frac{C}{C, h}$$

is an inference in π. Then

$$Ax(\pi, \mathbf{r}), C^{\mathbf{r}} \vdash_{2w+2} (C, h)^{\mathbf{r}} .$$

Assume $C^{\mathbf{r}} = \{g_1, \ldots, g_w\}$ where each g_i is a sum of some polynomials from C and thus also from C, h. So repeating Claim 3 w-times to remove $g_w, g_{w-1}, \ldots, g_1$ we derive $(C, h)^{\mathbf{r}}$. The linear width is at most $w + 3$ (from Claim 3) plus $w - 1$ (for side polynomials $g_1, \ldots, g_{w-1}$), i.e. at most $2w + 2$ in total.

Claim 5: *Assume*

$$\frac{C, g \quad C, h}{C, g + h + 1}$$

is an inference in π. Then

$$Ax(\pi, \mathbf{r}), (C, g)^{\mathbf{r}}, (C, h)^{\mathbf{r}}) \vdash_{2w+3} (C, g + h + 1)^{\mathbf{r}} .$$

We proceed as in Claim 4 and attempt to derive from $Ax(\pi, \mathbf{r}), (C, g)^{\mathbf{r}}$ clause $(C, g + h + 1)^{\mathbf{r}}$. The only obstacle to doing so is when the polynomial g occurs in a sum in $(C, g)^{\mathbf{r}}$: in that case we leave it as a side polynomial. That is, from $(C, g)^{\mathbf{r}}$ we derive $(C, g + h + 1)^{\mathbf{r}}, g$ in linear width at most $2w + 2 + 1 = 2w + 3$.

Analogously from $(C, h)^{\mathbf{r}}$ derive $(C, g + h + 1)^{\mathbf{r}}, h$ and then by the binary rule

$$(C, g + h + 1)^{\mathbf{r}}, g + h + 1 .$$

From that we get the wanted $(C, g + h + 1)^{\mathbf{r}}$ using the axiom

$$(g + h + 1) + 1, (C, g + h + 1)^{\mathbf{r}}$$

from $Ax(\pi, \mathbf{r})$, the binary rule and a contraction.

The following lemma follows form the last two claims.

Lemma 5.2 *Let π be an R(LIN/$\mathbf{F}_2$) refutation of $A_1, \ldots, A_m, B_1, \ldots, B_\ell$ consisting of k clauses and of linear width w_0. Let $w \geq 1$ be an arbitrary parameter. Then for a random $\mathbf{r}$ there is an R(LIN/$\mathbf{F}_2$)-refutation $\pi(\mathbf{r})$ of*

$$\Phi, Ax(\pi, \mathbf{r})$$

of linear width bounded above by

$$w' \ := \ 2w + 3$$

and with at most $O(ww_0k)$ clauses.

Proof :

The bound to the linear width follows from the last two claims, using also that

$$\Phi, Ax(\pi, \mathbf{r}) \ \vdash_{w'} \ \Phi^{\mathbf{r}} \ .$$

The bound to the number of clauses follows by inspecting that in both Claims 4 and 5 the constructed derivations have $O(ww_0)$ clauses.

q.e.d.

We used in this construction the syntactic version of R(LIN/$\mathbf{F}_2$) rather than the semantic one in order to generate explicitly the sets $Ax(\pi, \mathbf{r})$.

Now we can apply Theorem 2.5. The values of parameters appearing in that theorem are:

- $\epsilon := 2^{-w}k$: the conjunction of axioms in $Ax(\pi, \mathbf{r})$ corresponding to any one clause in π are 2^{-w}-approximations of $\{0,1\}^N$ (Claim 1).
- Number of steps: $O(ww_0k)$.
- Monotone communication complexity: $O(w \log n)$.

Theorem 5.3 *Assume the set-up conditions (1)-(7) and assume that sets $A_1, \ldots, A_m, B_1, \ldots, B_\ell$ are defined by R(LIN/$\mathbf{F}_2$)-clauses.*

Let π be an R(LIN/$\mathbf{F}_2$)-refutation of (the clauses defining) these sets with k steps and of the linear width bounded by w_0.

Then for every $w \geq 1$ there is a randomized protocol $(\mathbf{P}_\mathbf{r})_\mathbf{r}$ for $KW^m[U, V]$ of size at most $O(ww_0k)+n$, communication complexity $O(w \log n)$ and of error at most $2^{-w}k$.

Moreover, if the refutation π is tree-like then also G is tree-like.

Using Lemma 3.1 we can turn Theorems 5.1 and 5.3 into statements about separating monotone CLOs (we use Theorem 5.1 in the corollary).

Corollary 5.4

Assume the set-up conditions (1)-(7) and assume that sets $A_1, \ldots, A_m, B_1, \ldots, B_\ell$ are defined by R(LIN/$\mathbf{F}_2$)-clauses.

Let π be an R(LIN/$\mathbf{F}_2$) refutation of (the clauses defining) these sets with k steps. Let $w \geq 1$ be any parameter.

Then there is a monotone CLO of size at most $(k+n)2^{O(w \log n)}$ and of locality at most $3 \cdot 2^{-w}k$ separating U from V.

Moreover, if the refutation π is tree-like then the monotone CLO is a formula.

6 Randomized feasible interpolation for CP

Following [18] call a semantic derivation *CP-like* iff the proof steps are defined by integer linear inequalities. CP-like derivations were interpolated in [18] by protocols but their complexity was measured in terms of the real game defined there: players send each a real number to a referee and he announces how are these ordered. The real communication complexity of a multi-function R, $\mathrm{CC}^{\mathbf{R}}(R)$, is the minimal number of rounds (of sending numbers to the referee in an optimal protocol) needed to compute a valid value for R in the worst case. We can use this notion to measure the communication complexity of our protocols $\mathbf{P}$ and define $\mathrm{CC}^{\mathbf{R}}(\mathbf{P})$ analogously to how $\mathrm{CC}(\mathbf{P})$ was defined. We will not recall details as we will use here only the relation of the real communication complexity to the well-established probabilistic communication complexity.

Let R be a multi-function defined on $U \times V \subseteq \{0,1\}^n \times \{0,1\}^n$ and let $C^{pub}_{\epsilon}(R)$ be the probabilistic communication complexity of a multi-function R with public coins and error $\epsilon > 0$. The following equality was derived in [18, L.1.6] from a result of Nisan [22]. For $\epsilon < \frac{1}{2}$ it holds

$$C^{pub}_{\epsilon}(R) \leq \mathrm{CC}^{\mathbf{R}}(R) \cdot O(\log n + \log \epsilon^{-1}) \ . \tag{10}$$

We will use [18, Thm.3.3].

Theorem 6.1 *Assume the set-up conditions (1)-(7). Assume that the sets $A_1, \ldots, A_m$ and $B_1, \ldots, B_\ell$ are defined by integer linear inequalities and that there is a CP-like refutation π of $\tilde{A}_1, \ldots, \tilde{A}_m, \tilde{B}_1, \ldots, \tilde{B}_\ell$ that has k steps.*

Then for any $\epsilon < 1/k$ there is a randomized protocol for $KW^m[U,V]$ of size $k+n$, communication complexity $O(\log(n/\epsilon))$ and of error at most ϵk.

Moreover, if the refutation π is tree-like then also G is tree-like.

Proof :

Theorem 3.3. of [18] shows that there is a protocol for $KW[U,V]$ (resp. for $KW^m[U,V]$) of the stated size and with the real communication complexity $O(1)$. Then (10) implies that that protocol can be simulated by a randomized protocol of communication complexity $O(\log(n/\epsilon))$ which, for given u, v, computes at every node x the strategy function and the consistency condition with error at most ϵ. Hence the total error is estimated by ϵk. This entails the theorem.

q.e.d.

Note that analogously to Corollary 5.4 this can be turned into a statement about separating monotone CLOs. However, it is more direct to use the argument from the preceding proof to show that monotone CLOs efficiently simulate monotone real circuits of Pudlák [23] which do separate pairs U, V by the interpolation theorem established there.

Lemma 6.2 *Assume $U, V \subseteq \{0,1\}^n$ and U is closed upwards (or V downwards). Let C be a monotone real circuit of size s separating U from V.*

Then for every $0 < \epsilon < \frac{1}{2}$ there is a monotone CLO D separating U from V, having size $s(\frac{n}{\epsilon})^{O(1)}$ and locality $\mu \leq s\epsilon$.

In particular, for any $\mu > 0$ there is a monotone CLO separating U from V with locality $\leq \mu$ and size $(ns\mu^{-1})^{O(1)}$.

Proof :

Circuit C yields a protocol for $KW^m[U, V]$ of size s and real communication complexity $O(1)$: the graph of the protocol is C turned upside down (output is the root), the consistency condition $F(u, v)$ consists of subcircuits E where $E(u) > E(v)$, and the strategy is defined so that the consistency condition is preserved.

As in the proof of Theorem 6.1 the protocol can be turned into a randomized protocol of size s, communication complexity $O(\log(n/\epsilon))$ and error at most $s\epsilon$. The required monotone CLO then exists by Lemma 3.1.

The particular case is obtained by setting $\epsilon := s/\mu$.

q.e.d.

Let us remark that the constructions underlying Theorem 6.1 and Lemma 6.2 apply also to the proof system R(CP) of [17] operating with clauses formed by CP-inequalities and yield a small separating CLO for small width. In particular, if each clause in an R(CP)-refutation has size at most w then the (monotone) real communication complexity is at most w and this yields a monotone separating CLO of the size as in Lemma 6.2 for $w = O(\log(n/\epsilon))$.

7 The lower bound problem for monotone CLOs

This section is devoted to a discussion of the problem to establish a lower bound for monotone circuits with a local oracle separating two sets U and V (obeying all set-up conditions (1) - (7)). This would imply via Lemma 3.1 also a lower bound for randomized protocols for $KW^m[U, V]$ and hence a length-of-proofs lower bound for R(LIN/$\mathbf{F}_2$).

We shall consider the classical pair of disjoint sets of graphs having a large clique and of graphs colorable by a small number of colors. Let $n_0 \geq \omega > \xi \geq 1$ and put $n := \binom{n_0}{2}$. We shall identify in this context $[n]$ with the set of unordered pairs of distinct elements from $[n_0]$; we think of each such pair as denoting a potential edge in a graph with vertices $[n_0]$.

Take for $U \subseteq \{0,1\}^n$ the set $Clique_{n_0,\omega}$ of all graphs on $[n_0]$ that contain a clique of size ω. We shall also denote by $Clique_{n_0,\omega}(p,q)$ the set of the following clauses in atoms p_{ij}, $i \neq j \in [n_0]$, and q_{ui}, $u = 1,\ldots,\omega$ and $i \in [n_0]$ (hence there are $s = \omega \cdot n_0$ q-atoms):

- $\bigvee_{i\in[n_0]} q_{ui}$, one for each $u \in [\omega]$,
- $\neg q_{ui} \vee \neg q_{vi}$, one for all $u < v \in [\omega]$ and $i \in [n_0]$,
- $\neg q_{ui} \vee \neg q_{vj} \vee p_{ij}$, one for all $u < v \in [\omega]$ and $i \neq j \in [n_0]$.

Sets A_i from the set-up condition (2) are the sets defined by these clauses.

The set $V \subseteq \{0,1\}^n$ will be the set of graphs on $[n_0]$ that are ξ-colorable. We shall denote it $Color_{n_0,\xi}$ and by $Color_{n_0,\xi}(p,r)$ the set of the following clauses in the p-atoms and atoms r_{ia}, $i \in [n_0]$ and $a \in [\xi]$ (there are $n_0 \cdot \xi$ r-atoms):

- $\bigvee_{a\in[\xi]} r_{ia}$, one for each $i \in [n_0]$,
- $\neg r_{ia} \vee \neg r_{ib}$, one for all $a < b \in [\xi]$ and $i \in [n_0]$,
- $\neg r_{ia} \vee \neg r_{ja} \vee \neg p_{ij}$, one for all $a \in [\xi]$ and $i \neq j \in [n_0]$.

Sets B_j from the set-up condition (2) are the sets defined by these clauses.

If we identify a truth assignment $w \in \{0,1\}^n$ to the p-atoms with graph G_w on $[n_0]$, truth assignments to q_{ui} satisfying $Clique_{n_0,\omega}(w,q)$ correspond to injective (multi-)maps from $[\omega]$ onto a clique in G_w and analogously truth assignments to r_{ia} making $Color_{n_0,\xi}(w,r)$ true correspond to colorings of G_w by ξ colors. Thus if $\omega > \xi$ the sets U and V are disjoint and its is easy to see that they, together with the clauses above, satisfy the set-up conditions (1)-(7) from Section 1.

Let us first note that a lower bound for a monotone CLO with oracle rectangles inside $U \times V$ can be derived as an easy consequence of a theorem of Jukna [11, Thm.3], generalizing an earlier result by Yao [34]. In particular, [11, Thm.3] states that there is no small (polynomial size) monotone circuit computing the characteristic function χ_U of U for $\omega = (n_0/\log n_0)^{2/3}$ even if the circuits are allowed to use at gates arbitrary monotone Boolean functions as long as all their min-terms have size $o(\omega)$. In the case of a monotone CLO with oracle rectangles $U_j \times V_j$ we can take for all functions f_j the disjunction f of all conjunctions

$$\lceil X \rceil \;:=\; \bigwedge_{i\neq j\in X} p_{ij} \tag{11}$$

where sets $X \subseteq [n_0]$ run over all sets of vertices of size $\xi + 1$. Clearly f is identically 1 on U and 0 on V and hence if, say, $\xi = \omega^{1/2}$, Jukna's [11, Thm.3] applies. However, this is not good enough: we want a stronger lower bound but more importantly we need a lower bound for monotone CLOs separating U from V and not just for those computing χ_U.

The classical result of Alon and Boppana [2], strengthening Razborov's [29] lower bound, offers such a lower bound for ordinary monotone circuits.

Theorem 7.1 (Alon and Boppana [2, Thm.3.11])
Assume that $3 \leq \xi < \omega$ *and* $\sqrt{\xi}\omega \leq \frac{n_0}{8 \log n_0}$. *Then any monotone circuit separating* $Clique_{n_0,\omega}$ *from* $Color_{n_0,\xi}$ *must have the size at least* $2^{\Omega(\sqrt{\xi})}$.

It appears possible that the same lower bound holds also for monotone CLOs with a small constant locality. Alluding to Boppana and Sipser [3, L.4.2] we prove at least the following partial result for monotone CLOs of the restricted form

$$D \;:=\; \bigvee_{i \leq a} (\lceil X_i \rceil \wedge C_i(\overline{y})) \tag{12}$$

where

12.1 $|X_i| \leq \lfloor \xi^{1/2} \rfloor$ and $\lceil X_i \rceil$ is defined as in (11) using variables x_{ij} in place of p_{ij},

12.2 $C_i(\overline{y})$ is a monotone circuit of an arbitrary size not containing the x-variables,

12.2 the size a of the disjunction is arbitrary.

Lemma 7.2 *Assume that* $4 \leq \xi < \omega$ *and that* n_0 *is large enough. Then no monotone circuit with a local oracle* D *of the form (12), satisfying conditions 12.1-3 and with locality* $\mu \leq \frac{1}{16}$ *separates* $Clique_{n_0,\omega}$ *from* $Color_{n_0,\xi}$.

The proof of the lemma will be summarized after Lemma 7.4.

A CLO separating $U(= Clique_{n_0,\omega})$ from $V(= Color_{n_0,\xi})$ separates also U^{min} from V^{max}. Note that elements of U^{min} are graphs consisting of a clique of size ω and having no other edges and elements of V^{max} are ξ-partite graphs with all possible edges among the different parts. These two sets are called in [2, 3] *positive* and *negative* examples, respectively. In fact, for the counting purposes the negative examples are represented as ξ-colorings of $[n_0]$, each coloring determining the maximal graph for which it is still a graph coloring.

Let $D(\overline{x}, \overline{y}), \mathcal{R}$ be a monotone CLO of the form (12), satisfying 12.1-3, with locality μ and with e oracle rectangles $U_j \times V_j$. Let

$$Bad \;:=\; \bigcup_{j \leq e} U_j \times V_j \;\subseteq\; U^{min} \times V^{max}.$$

We know that $|Bad| \leq \mu \cdot |U^{min} \times V^{max}|$.

In the argument we shall consider other rectangles inside $U^{min} \times V^{max}$ and y-variables attached to them. Let us introduce the following notation. For $U' \subseteq U^{min}$ and $V' \subseteq V^{max}$ let $y[U', V']$ be a new variable. Its *valid interpretation* is any monotone Boolean function $h : \{0,1\}^n \to \{0,1\}$ that is 1 on U' and 0 on V'. Two specific valid interpretations of the y-variables are:

- $\mathcal{F}_U$-interpretation: each $y[U', V']$ is interpreted by the Boolean function that is 1 on U' and 0 everywhere else on $U^{min} \cup V^{max}$,

- $\mathcal{F}_V$-interpretation: each $y[U', V']$ is interpreted by the Boolean function that is 0 on V' and 1 everywhere else on $U^{min} \cup V^{max}$,

(we only care for values on $U^{min} \cup V^{max}$). Let $E(\overline{x}, \overline{y})$ be a monotone circuit involving also some of the y-variables and let $\mathcal{F}$ be a valid interpretation of the y-variables. Then

$$E(\overline{x}, \mathcal{F})$$

denotes the Boolean function obtained by substituting for each y-variable in E the function interpreting it in $\mathcal{F}$.

Lemma 7.3 *Let $E(\overline{x}, \overline{y})$ be a monotone circuit. It holds on $U^{min} \cup V^{max}$:*

1. *For any valid interpretation $\mathcal{F}$:*

$$E(\overline{x}, \mathcal{F}_U) \leq E(\overline{x}, \mathcal{F}) \leq E(\overline{x}, \mathcal{F}_V) \ .$$

2. *For $\mathcal{F} = \mathcal{F}_U, \mathcal{F}_V$:*

$$(y[U_1, V_1] \vee y[U_2, V_2])(\mathcal{F}) \ = \ y[U_1 \cup U_2, V_1 \cap V_2](\mathcal{F}) \ .$$

3. *For $\mathcal{F} = \mathcal{F}_U, \mathcal{F}_V$:*

$$(y[U_1, V_1] \wedge y[U_2, V_2])(\mathcal{F}) \ = \ y[U_1 \cap U_2, V_1 \cup V_2](\mathcal{F}) \ .$$

4. *If both $U_1 \times V_1$ and $U_2 \times V_2$ are subsets of Bad, so are $U_1 \cup U_2 \times V_1 \cap V_2$ and $U_1 \cap U_2 \times V_1 \cup V_2$.*

Proof :

Parts 1 and 4 are obvious. Let χ_W be the characteristic function of $W \subset \{0,1\}^n$. For Part 2:

$$(y[U_1, V_1] \vee y[U_2, V_2])(\mathcal{F}_U) = \chi_{U_1} \vee \chi_{U_2} = \chi_{U_1 \cup U_2} = y[U_1 \cup U_2, V_1 \cap V_2](\mathcal{F}_U)$$

and

$$(y[U_1, V_1] \vee y[U_2, V_2])(\mathcal{F}_V) = \chi_{\setminus V_1} \vee \chi_{\setminus V_2} = \chi_{\setminus (V_1 \cap V_2)} = y[U_1 \cup U_2, V_1 \cap V_2](\mathcal{F}_V) \ .$$

Part 3 is analogous.

q.e.d.

We shall argue that either $D(\overline{x}, \mathcal{F}_U)$ rejects a lot of U^{min} or that $D(\overline{x}, \mathcal{F}_V)$ accepts a lot of V^{max}. The choice to evaluate how well D works on U^{min} using the interpretation $\mathcal{F}_U$ and on V^{max} using $\mathcal{F}_V$ gives us (due to Part 1 of Lemma 7.3) the best chance to detect errors.

Note that $\lceil X \rceil$ is equivalent to

$$\lceil X \rceil y[U^{min}, \emptyset]$$

under the two extreme interpretations as $y[U^{min}, \emptyset]$ is 1 on U^{min} under $\mathcal{F}_U$ and 1 on both U^{min} and V^{max} under $\mathcal{F}_V$. So we could have allowed in (12) also stand-alone terms $\lceil X \rceil$ and if we defined $\lceil \emptyset \rceil := 1$ also stand-alone y-variables.

Lemma 7.4 *Assume $\mu \le 1/16$. Then for any monotone CLO E of the form*

$$E = \bigvee_{i \le a} (\lceil X_i \rceil \wedge y[U_i, V_i])$$

where a is arbitrary, $|X_i| \le \lfloor \xi^{1/2} \rfloor$ and all rectangles $U_i \times V_i$ are subsets of Bad it holds:

1. *Either $E(\mathcal{F}_V)$ accepts at least $1/4$ of V^{max},*
2. *or $E(\mathcal{F}_U)$ rejects at least $3/4$ of U^{min}.*

Proof :

If E is the empty disjunction, it is constantly zero and the second option occurs.

If not, note that as all rectangles $U_i \times V_i$ are subsets of Bad, their measure in $U^{min} \times V^{max}$ at most μ. Hence at least one of its sides U_i or V_i has the measure at most $\mu^{1/2}$ in U^{min} or V^{max}, respectively. Now consider two cases:

1. There is a term $\lceil X_i \rceil \wedge y[U_i, V_i]$ in E with V_i having the measure at most $\mu^{1/2}$ in V^{max},
2. not 1.

Denote $\ell := \max_{j \le a} |X_j|$; we have $\ell \le \lfloor \xi^{1/2} \rfloor$.

In the first case the term $\lceil X_i \rceil \wedge y[U_i, V_i](\mathcal{F}_V)$ accepts at least the fraction of

$$[1 - \frac{\binom{\ell}{2}}{\xi}] - \mu^{1/2} \ge [\frac{3}{4} - \frac{\binom{\ell}{2}}{\xi}] \ge \frac{1}{4}$$

elements $v \in V^{max}$: the first term is the same estimate as in [3, L.4.2], the second accounts for the elements of V_i.

In the second case use $\mathcal{F}_U$: all $y[U_i, V_i](\mathcal{F}_U)$ are 1 only inside U_i and hence E accepts at most the subset $\bigcup_i U_i$ of U^{min}. But for each u from this union the pair $(u, v) \in Bad$ for at least a fraction of $\mu^{1/2}$ of elements v of V^{max}. Hence the measure of the union is at most $\mu^{1/2} \le \frac{1}{4}$.

q.e.d.

Now we can derive Lemma 7.2. By parts 2 and 3 of Lemma 7.3, each subcircuit $C_i(\overline{y})$ of D is equivalent under both $\mathcal{F}_U$ and $\mathcal{F}_V$ to some $y[U_i, V_i]$ such that, by part 4 of that lemma, $U_i \times V_i \subseteq Bad$. Hence Lemma 7.4 applies.

Let us remark that there is a certain discrepancy in the sizes when protocols are turned to CLOs in Lemma 3.1 and CLOs are transformed into protocols in Lemma 3.3. Thus even if the lower bound for monotone CLOs was not valid one could still try the tight 3-valued version of Lemma 3.4.

8 Concluding remarks

We remark without elaborating it that Theorem 2.5 yields a randomized feasible interpolation[2] for the random resolution system proposed informally by Dantchev and defined formally by Buss, Kolodziejczyk and Thapen [5, Sec.5.2]. Pudlák and Thapen [25] consider more variants of the definition and they prove a feasible interpolation for the tree-like case. According to the definition from Buss et.al. [5] an ϵ-random resolution refutation distribution of a set of clauses Φ is a random distribution $(\pi_{\mathbf{r}})_{\mathbf{r}}$ of resolution refutations of $\Psi \cup \Delta_{\mathbf{r}}$, where $\Delta_{\mathbf{r}}$ are sets of clauses such that any fixed truth assignment fails to satisfy $\bigwedge \Delta_{\mathbf{r}}$ with the probability at most ϵ. In other words, if $X_{\mathbf{r}}$ is the set of assignments satisfying all clauses in $\Delta_{\mathbf{r}}$ then $(X_{\mathbf{r}})_{\mathbf{r}}$ is an ϵ-approximation of the universe of all assignments. The number of steps in such a random refutation is the maximal number of steps among all $\pi_{\mathbf{r}}$.

$R(LIN/\mathbf{F}_2)$ can be generalized to a proof system $R(PC_d/\mathbf{F}_2)$, resolution over degree d PC, operating with clauses formed by degree $\leq d$ polynomials over $\mathbf{F}_2$; just add an extra rule

$$\frac{C, g}{C, gh + h + 1}$$

corresponding to the multiplication rule of polynomial calculus PC (cf. Clegg, Edmonds and Impagliazzo [7]). Both processes from Section 5 of reducing the width of clauses in a proof work analogously as for $R(LIN/\mathbf{F}_2)$. For definiteness let us now consider the construction underlying Lemma 5.2. The clauses $C^{\mathbf{r}} = \{g_1, \ldots, g_w\}$ can be themselves replaced by a single polynomial $1 - \Pi_{j \leq w}(1 - g_j)$ of degree $\leq wd$. Hence the process can be repeated any fixed number of times and thus, in fact, it can be applied to $AC^0[2]$-formulas and $AC^0[2]$-Frege proofs instead of $R(PC_d/\mathbf{F}_2)$-proofs only. This would result in a semantic PC-refutation of the original set of clauses augmented by additional initial polynomials (analogous to axioms $Ax(\pi, \mathbf{r})$) of degree $w^{O(1)}$ which yields also a syntactic PC-refutation of the same set of clauses and of the same degree by Buss et.al.[4, Thm.2.6]. A similar reduction can be obtained also by using the characterization of $AC^0[2]$-Frege proofs via the so called extended Nullstellensatz proofs of Buss et.al.[4] and removing the extension axioms there by a random assignment to the extension variables at the expense of introducing the new initial polynomials. However, if monotone CLOs separating $Clique_{n_0,\omega}$ and $Color_{n_0,\xi}$ from Section 7 must be indeed large, randomized feasible interpolation will not work in this situation as constant depth Frege systems admit short proofs of the weak pigeonhole principle and hence also of the disjointness of the sets $Clique_{n_0,\omega}$ and $Color_{n_0,\xi}$ (when $\omega \geq 2\xi$). Note also that $R(PC_d/\mathbf{F}_2)$ even without the extra axioms p-simulates $R(d)$, a proof systems operating with d-DNFs (cf. [19]), which is known to be fairly strong (it corresponds to bounded arithmetic theory $T^2_2(\alpha)$ for d poly-logarithmic in n, cf.[19]).

[2] A different one than [20].

Acknowledgements:
I thank Michal Garlík for pointing out a missing $\log n$ factor in Section 5, to Igor C. Oliveira and Pavel Pudlák for comments on drafts of a part of the paper and to Neil Thapen for discussions about related topics.

References

[1] M. Alekhnovich, E. Ben-Sasson, A. A. Razborov and A. Wigderson, Pseudorandom Generators in Propositional Proof Complexity, *SIAM Journal on Computing*, **34(1)**, (2004), pp.67-88.

[2] N. Alon and R. Boppana, The monotone circuit complexity of Boolean functions, *Combinatorica*, **7(1)**, (1987), pp.1-22.

[3] R. Boppana and M. Sipser, The complexity of finite functions, in: Handbook of Theoretical Computer Science, (1990), pp.759–804. Elsevier Science Publishers.

[4] S. R. Buss, R. Impagliazzo, J. Krajíček, P. Pudlák, A. A. Razborov, and J. Sgall: Proof complexity in algebraic systems and bounded depth Frege systems with modular counting, *Computational Complexity*, **6(3)**, (1996/1997), pp.256-298.

[5] S. R. Buss, L. A. Kolodziejczyk and N. Thapen, Fragments of approximate counting, *J. of Symbolic Logic*, Vol **79:2**, (2014), pp.496-525.

[6] S. R. Buss, L. A. Kolodziejczyk and K. Zdanowski, Collapsing modular counting in bounded arithmetic and constant depth propositional proofs, *Transactions of the AMS*, **367**, (2015), pp.7517-7563.

[7] M. Clegg, J. Edmonds, and R. Impagliazzo, Using the Groebner basis algorithm to find proofs of unsatisfiability, in: *Proc. 28^{th} Annual ACM Symp. on Theory of Computing*, (1996), pp. 174-183. ACM Press.

[8] E. Hirsch and A. Kojevnikov, Several notes on the power of Gomory-Chvatal cuts, *Annals of Pure and Applied Logic*, **141**, (2006), pp.429-436.

[9] R. Impagliazzo, T. Pitassi and A. Urquhart, Upper and lower bounds for treelike cutting planes proofs, in: Proc. of the 9th Annual IEEE *Symposium on Logic in Computer Science*, Piscataway, NJ, IEEE Computer Science Press, (1994), pp.220228.

[10] D. Itsykson and D. Sokolov, Lower bounds for splittings by linear combinations, in: Proc. MFCS, Eds. E. Csuhaj-Varju, M. Dietzfelbinger, Z. Esik, LN in CS, Springer, Vol.**8635**(2014), pp.372-383.

[11] S. Jukna, Monotone circuits and local computations, in: *Proc. of 31st Conf. of Lithuanian Math. Soc.*, (1990).

[12] M. Karchmer and A. Wigderson, Monotone circuits for connectivity require super - logarithmic depth, in: *Proc.* 20^{th} *Annual ACM Symp. on Theory of Computing*, (1988), pp.539-550. ACM Press.

[13] A. Kojevnikov, Improved lower bounds for tree-like resolution over linear inequalities, in Proc. of the 10th International Conference on Theory and Applications of Satisfiability Testing (SAT), Eds. J. Marques-Silva, K. A. Sakallah, LN in CS, Springer, Vol.**4501**, (2007), pp.70-79.

[14] J. Krajíček, *Bounded arithmetic, propositional logic, and complexity theory*, Encyclopedia of Mathematics and Its Applications, Vol. **60**, Cambridge University Press, (1995).

[15] J. Krajíček, Lower bounds for a proof system with an exponential speed-up over constant-depth Frege systems and over polynomial calculus, in: Eds. I.Prívara, P. Růžička, 22nd Inter. Symp. *Mathematical Foundations of Computer Science* (Bratislava, August '97), Lecture Notes in Computer Science 1295, Springer-Verlag, (1997), pp.85-90.

[16] J. Krajíček, Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic, *J. Symbolic Logic*, **62(2)**, (1997), pp. 457-486.

[17] J. Krajíček, Discretely ordered modules as a first-order extension of the cutting planes proof system, *J. Symbolic Logic*, **63(4)**, (1998), pp.1582-1596.

[18] J. Krajíček, Interpolation by a game, *Mathematical Logic Quarterly*, **44(4)**, (1998), pp.450-458.

[19] J. Krajíček, On the weak pigeonhole principle, *Fundamenta Mathematicae*, Vol.**170(1-3)**, (2001), pp.123-140.

[20] J. Krajíček, A feasible interpolation for random resolution, to appear in *Logical methods in Computer Science*, preprint April 2016 available at ArXiv: `https://arxiv.org/abs/1604.06560`

[21] J. Krajíček and I. C. Oliveira, On monotone circuits with local oracles and clique lower bounds, submitted.

[22] N. Nisan, The communication complexity of the threshold gates, in: *Combinatorics, P. Erdös is Eighty*, **Vol. 1**, Eds. Miklós et.al., Bolyai Math. Soc., (1993), pp.301-315.

[23] P. Pudlák, Lower bounds for resolution and cutting planes proofs and monotone computations, *J. Symbolic Logic*, **62**, (1987), pp.981-998.

[24] P. Pudlák, The lengths of proofs, in: Handbook of Proof Theory, S.R. Buss ed., Elsevier, (1998), pp.547-637.

[25] P. Pudlák and N. Thapen, Random resolution refutations, preprint available at `http://eccc.hpi-web.de/report/2016/175/`.

[26] R. Raz and I. Tzameret, Resolution over Linear Equations and Multilinear Proofs, *Annals of Pure and Applied Logic*, **155(3)**, (2008), pp. 194-224.

[27] R. Raz and A. Wigderson, Probabilistic Communication Complexity of Boolean Relations (Extended Abstract), in: Proc. 30th Found. of Comp. Sci. (FOCS), (1989), pp.562-567.

[28] R. Raz and A. Wigderson, Monotone circuits for matching require linear depth, *J. of Assoc. for Computing Machinery*, **39(3)**, (1992), pp.736744.

[29] A. A. Razborov, Lower bounds on the monotone complexity of some Boolean functions, *Soviet Mathem. Doklady*, **31**, (1985), pp.354-357.

[30] A. A. Razborov, Lower bounds on the size of bounded depth networks over a complete basis with logical addition, *Matem. Zametki*, **41(4)**, (1987), 598-607.

[31] A. A. Razborov, Unprovability of lower bounds on the circuit size in certain fragments of bounded arithmetic, *Izvestiya of the R.A.N.*, **59(1)**, (1995), pp.201-224.

[32] A. A. Razborov, Lower Bounds for the Polynomial Calculus, *Computational Complexity*, **7(4)**, (1998), pp.291-324.

[33] R. Smolensky, Algebraic methods in the theory of lower bounds for Boolean circuit complexity, in: *Proc. 19th Ann. ACM Symp. on Th. of Computing*, (1987), pp. 77-82.

[34] A. C.-C. Yao, Circuits and local computation, in: *Proc. of the 21st annual ACM symposium on Theory of computing*, ACM, New York, (1989), pp.186-196.

Mailing address:
Department of Algebra
Faculty of Mathematics and Physics
Charles University
Sokolovská 83, Prague 8, CZ - 186 75
The Czech Republic
`krajicek@karlin.mff.cuni.cz`