58th Annual IEEE Symposium on Foundations of Computer Science

Random formulas, monotone circuits, and interpolation

Pavel Hrubeš
Institute of Mathematics of AVCR
Prague, Czech Republic
pahrubes@gmail.com

Pavel Pudlák
Institute of Mathematics of AVCR
Prague, Czech Republic
pudlak@math.cas.cz

***Abstract*—We prove new lower bounds on the sizes of proofs in the Cutting Plane proof system, using a concept that we call *unsatisfiability certificate*. This approach is, essentially, equivalent to the well-known feasible interpolation method, but is applicable to CNF formulas that do not seem suitable for interpolation. Specifically, we prove exponential lower bounds for random k-CNFs, where k is the logarithm of the number of variables, and for the Weak Bit Pigeon Hole Principle. Furthermore, we prove a monotone variant of a hypothesis of Feige [12]. We give a superpolynomial lower bound on monotone real circuits that approximately decide the satisfiability of k-CNFs, where $k = \omega(1)$. For $k \approx \log n$, the lower bound is exponential.**

***Keywords*-Cutting Planes; random formulas; interpolation**

I. Introduction

In proof complexity, we want to identify tautologies which are hard to prove in a given proof system. For most proof system, the Frege system for example, it is not even known whether such hard tautologies exist. In others, the known hard instances have very specific form tailored to beat the system – this is the case of bounded depth Frege or Cutting Planes. There is a general consensus among researchers in computational complexity that random k-CNF formulas are hard instances both for decision procedures and for propositional proofs. The seminal paper of Chvátal and Szemerédi [9] gives exponential lower bounds on Resolution refutations of random formulas of suitable density. This was simplified in [6]; similar results have been proven for a few other proof systems, such as Polynomial Calculus [5], [1]. There are proof systems for which we do have lower bound methods (the aforementioned Cutting Planes and bounded depth Frege), yet no superpolynomial lower bounds are known for random formulas. Thus one of the central problems in proof complexity is to prove lower bounds on the refutations of random k-CNFs in stronger systems, ideally for $k = 3$. In this paper we make some progress by proving an exponential lower bound for random k-CNFs in Cutting Planes, with $k \approx \log n$.

Cutting Planes is a proof system designed to show that a given set of linear inequalities has no $0, 1$-solution. This is achieved by simple syntactic operations, namely, we can add two inequalities, multiply by a positive constant, and round fractional constants if the coefficients are integers. A system of inequalities has no $0, 1$-solution iff we can thus obtain the inequality $0 \geq 1$. As a decision procedure for integer linear programming, Cutting Planes was defined by Gomory and Chvátal [16], [8]. As a proof system, Cutting Planes was later proposed in [11]. An exponential lower-bounds for this system were proved by Haken and Cook [18], and Pudlák [25]. In [13], it was noted that in this lower bound, the structure of the inference rules is irrelevant: it works also for the so-called Semantic Cutting Planes, where we can derive from two inequalities any inequality which is their valid consequence. The hard tautologies in [25] have a very specific form, and thus it seemed impossible to extend the technique to random k-CNFs. Nevertheless, we will prove:

Theorem 1. *For every constant $c > 1$, if $\mathcal{C}$ is a random k-CNF with n variables and $m = O(n2^k)$ clauses, where $k \geq c \log n$, then, with high probability, every Cutting Planes (or Semantic Cutting Planes) refutation of $\mathcal{C}$ has size $2^{n^{\Omega(1)}}$.*

The theorem is most interesting for k small, such as $k \approx \log n$; if k is very large, it becomes trivial. Concerning the parameter m, one can easily show that for $m \geq cn2^k$, where c is a sufficiently large constant, a random CNF is unsatisfiable with high probability.[1] A result similar to Theorem 1 has recently been obtain by Fleming et al. in [14].

A general technique for proving proof complexity lower bounds is the so-called *feasible interpolation*. On a high level, given a tautology A of certain form, we can associate with it a computational problem P so that if A has a short proof (in our proof system) then P is easy to compute (in a specific model). Hence, it is a method to reduce hardness of proof to hardness of computation. This strategy was first used by Krajíček [23] to obtain lower-bounds for Resolution from *monotone Boolean* circuit lower bounds, proved by Razborov [28] and improved by Alon and Boppana [2]. Feasible interpolation is also the only known technique to obtain the lower bounds for Cutting Planes. In order to implement the method, Pudlák [25] considered a stronger computational model, *monotone real circuit*, and proved an exponential lower-bound for this model. This illustrates

[1] For small m, when a random CNF is satisfiable, the theorem holds trivially.

0272-5428/17 $31.00 © 2017 IEEE
DOI 10.1109/FOCS.2017.20

one limitation of the interpolation technique: stronger proof systems lead to stronger models of computation, for which it may be hard or impossible to prove lower bounds (and strong systems like Frege are believed not to have interpolation even by general, non-monotone, Boolean circuits). Another limitation is that interpolation is applicable to formulas of a very specific form, a form that random formulas do not have. This apparently strictly limits the range of applications of this method.

In this paper, we modify the interpolation technique so that it is applicable it to a wider range of tautologies/contradictions. Our approach is based on the concept of *unsatisfiability certificate*, which we now define. Let

$$\mathcal{C} = \{C_1, \ldots, C_m\} \tag{1}$$

be a CNF with m clauses. Let X_0, X_1 be a partition of its variables into two disjoint sets. Hence, every clause C_i in $\mathcal{C}$ can be written as $C_i = C_i^0 \cup C_i^1$, where C_i^0 depends on variables from X_0 and C_i^1 on variables from X_1 only.

Definition. *A monotone Boolean function $F : \{0,1\}^m \to \{0,1\}$ will be called* an X_0, X_1-unsatisfiability certificate for $\mathcal{C}$ *(or simply* certificate for $\mathcal{C}$*), if for every $A \subseteq [m]$ the following hold:*

$$\text{if } \{C_i^1 : i \in [m] \setminus A\} \text{ is satisfiable then } F(A) = 1\,, \tag{2}$$

$$\text{if } \{C_i^0 : i \in A\} \text{ is satisfiable then } F(A) = 0\,. \tag{3}$$

One can easily see that an unsatisfiability certificate exists iff $\mathcal{C}$ is unsatisfiable, hence the name. The importance of this concept stems from the fact that lower bounds on circuit complexity of unsatisfiability certificates imply lower bounds on the size refutations of $\mathcal{C}$. This depends on the type of circuits and proofs; here we are interested in monotone real circuits and Cutting Planes proofs. Hence Theorem 1 is a corollary of the following lower bound on the complexity of certificates.

Theorem 2. *Let $c > 1$ be a constant and let $n \geq 1$ be given. Let X_0, X_1 be a partition of $2n$ variables into two sets of equal size. If $\mathcal{C}$ is a random k-CNF with $O(n2^k)$-clauses, variables $X_0 \cup X_1$, and $k \geq c \log n$, then, with high probability, every certificate for $\mathcal{C}$ requires monotone real circuits of size $2^{n^{\Omega(1)}}$.*

In our approach, the main obstacle to extending Theorem 1 to constant k is that we do not have monotone circuit lower bounds on certificates in this regime. Note that one has[2] a polynomial *upper* bound on certificates of random k-CNFs with $c'n$ clauses, where c' is sufficiently large with respect to k, while we believe such formulas are hard for Cutting Planes.

Another proof complexity technique, intimately related to interpolation, is based on communication complexity. Consider a CNF as in (1), and consider the associated two-player game: Player I has an assignment α_1 to the variables X_1, Player II an assignment α_0 to the variables X_0, and they are supposed to agree on a clause $C_i \in \mathcal{C}$ which is false in the assignment $\alpha_0 \cup \alpha_1$. In some systems, we can obtain super-polynomial lower bounds on proof size from super-logarithmic lower bounds on the communication complexity of the associated game. This strategy works for the tree-like versions of Resolution, Cutting Planes [19], and Lovász-Schrijver and similar systems [4]. Note that the paper of Impagliazzo et al. [19] uses randomized communication, and Beame et al. [4] uses, in addition, multi-party communication; a key ingredient of both are communication lower-bounds for Disjointness [27], [24], [30]. The advantage of the communication approach is that no specific form of the CNF is required; the disadvantage, that it applies only to tree-like proofs. In some sense, our notion of unsatisfiability certificate mimics the game-theoretic viewpoint: the definition of F is set up so that the monotone Karchmer-Wigderson game for F (cf. [22]) solves the associated two-player game. Hence, communication lower bounds can be seen as bounds on monotone circuit *depth* of F. However, talking about the circuit *size* of F has the advantage of being applicable to general, non-tree-like, proofs. This holds at least in Cutting Planes – it is an interesting challenge to develop reasonable notions of interpolation or unsatisfiability certificate for other systems, corresponding to randomized/multiparty communication.

We will give another example of an interesting hard formula for Cutting Planes, the weak *Bit Pigeon Hole Principle*, BPHP_N^M. The usual Pigeon Hole Principle, PHP_n^m, is a contradictory CNF which asserts that there is a total injective map from $[m]$ to $[n]$, $m > n$. The variables of PHP indicate whether the i-th pigeon is mapped to the j-th hole. A classical result of Haken [17] says that PHP_n^{n+1} requires exponential Resolution refutations; this was later improved for the weak PHP by Raz [26], who showed that PHP_n^m requires Resolution refutation of size exponential in n, no matter how large m is. However, PHP_n^{n+1} does have polynomial size Cutting Planes proofs. The Bit Pigeon Hole Principle is formulated differently. It asserts that there is a one-to-one function $f : [M] \to [N]$, but the variables of BPHP correspond to the bit representation of the graph of the alleged function f. We show:

Theorem 3. *Every Cutting Planes refutation of the Weak Bit Pigeon Hole Principle BPHP_N^M, $M > N$, has size $2^{\Omega(N^{1/8})}$.*

Prior to this paper, no lower bounds were known even for the strong version where $M = N + 1$.

Finally, we will consider a purely computational problem of approximately deciding the satisfiability of random k-CNFs. While deciding satisfiability of 3-CNFs is a well-known **NP**-complete problem, it is not known if the problem is **NP**-hard when we only want to decide it for most

[2] see Remark 18, and Sections III-A, V for further discussion

formulas. In [12] Feige proposed the conjecture that there is no polynomial time algorithm that accepts all satisfiable 3-CNFs and rejects *almost all* unsatisfiable 3-CNFs of a particular density. (The clause density of interest is such that almost all 3-CNFs are unsatisfiable.) It would be interesting to give a piece of evidence for the conjecture by proving a superpolynomial lower bound at least for monotone Boolean circuits. We cannot prove this, but we can prove such a result for k-CNFs where k is a function of the number of variables that goes to infinity. In fact, our lower bound is for a more general kind of monotone circuits, the real monotone circuits.

In more detail, for a given n and k, let $C_1, \ldots, C_N$ be the set of all k-clauses in n variables, where $N := \binom{n}{k}2^k$. Note that k-satisfiability problem can be viewed as a *monotone* Boolean function in N variables: identifying the variables with k-clauses, the function which accepts all unsatisfiable k-CNFs and rejects satisfiable k-CNFs is clearly monotone. Let us denote this function $\text{UNSAT}_{k,n}$. Note that exponential lower bounds on the monotone circuit size of $\text{UNSAT}_{3,n}$ (hence also $\text{UNSAT}_{k,n}$, $k \geq 3$) follow from the known monotone circuit lower bounds. A Boolean function $f : \{0,1\}^N \to \{0,1\}$ is said to *δ-approximate $\text{UNSAT}_{k,n}$ on* $\binom{[N]}{t}$, if the following hold[3] :

(i) for every $A \subseteq [N]$, if $\{C_i : i \in A\}$ is satisfiable then $f(A) = 0$;

(ii) $f(A) = 1$ for at least δ-fraction of t-element subsets A of $[N]$.

Theorem 4. *Let $0 < \delta, \epsilon \leq 1$ be constants. Let $N = \binom{n}{k}2^k$, $n \leq t \leq N^{1-\epsilon}$, and $k \leq \log n$. Assume that f is a monotone function which δ-approximates $\text{UNSAT}_{k,n}$ on $\binom{[N]}{t}$. Then for all k and n sufficiently large, f requires monotone real circuits of size $n^{2^{\Omega(k)}}$. If on the other hand $k \geq \log n$, the bound becomes $2^{n^{\Omega(1)}}$.*

The lower bound is super-polynomial in N if $k \to \infty$.

Organization of the paper.: After Preliminaries, we discuss the concept of unsatisfiability certificate in Section III, its connection with interpolation and communication complexity. In Section IV, we deal with the required monotone lower bounds, proving Theorems 2,3 and 4.

II. Preliminaries

CNFs, random k-CNFs: A *literal* is a variable or its negation. A *clause* is a set of literals, not containing a variable and its negation; k-clause is a clause of size k. A CNF is a set of clauses; k-CNF is a set of k-clauses. An assignment satisfies a clause C, if it satisfies some literal in C, and it satisfies a CNF $\mathcal{C}$ if it satisfies every clause in $\mathcal{C}$. $C(\alpha)$ will denote the Boolean value of C in the assignment α.

A random k-CNF $\mathcal{C}$ in n variables and m clauses is obtained by picking k-clauses $C_1, \ldots, C_m$ uniformly and independently at random from the $\binom{n}{k}2^k$ possible k-clauses. We choose clauses *with repetition* so $|\mathcal{C}|$ may be smaller than m.

Throughout, we use the following notation. For a clause C, let $\text{Var}(C)$ be the set of variables that appear in C, positively or negatively. For a set of variables X, let $C \restriction X$ be the part of C depending on X, namely, $C \restriction X := C \cap \{x, \neg x : x \in X\}$. If $\mathcal{C}$ is a CNF $\{C_1, \ldots C_m\}$, $\mathcal{C} \restriction X$ denotes the CNF $\{C_1 \restriction X, \ldots, C_m \restriction X\}$. For a sequence of events $E(1), E(2), \ldots$ indexed by n, we say that $E(n)$ *holds with high probability (w.h.p.)* if $\lim_{n\to\infty} \Pr[E(n)] = 1$.

The satisfiability threshold: It is well-known that as the number of clauses m increases there is a transition from the case when almost all formulas are satisfiable to the case when almost all unsatisfiable. The transition occurs on a very short interval [15]. This phenomenon is well understood for constant k (although it is still open if the threshold really exists, i.e., the interval converges to a point), see [10] and the references therein. We do not know if this problem has also been studied for k as an increasing function of n. In this paper, we only need a lower bound that guarantees that a random formula is unsatisfiable w.h.p. By the union bound, the probability that a random k-CNF with m clauses is satisfied is upper-bounded by $(1-2^{-k})^m 2^n < \mathrm{e}^{m2^{-k}}2^n$. Hence if $m \geq (\ln 2 + \epsilon)2^k n$, for some constant $\epsilon > 0$, then a random formula is unsatisfiable with high probability. Note that if $k = O(\log n)$, this bound is polynomial in n and the number of k-clauses is quasi-polynomial.

Monotone functions and circuits, monotone real circuits: Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. Identifying a subset of $[n]$ with its characteristic function, we will view inputs of f as subsets of $[n]$. We say that f is *monotone*, if $A \subseteq B$ implies $f(A) \leq f(B)$ for every $A, B \subseteq [n]$. A Boolean function $f(x_1, \ldots, x_m)$ is a *monotone projection* of $g(x_1, \ldots, x_m)$, if $f = g(z_1, \ldots, z_m)$ where $z_1, \ldots, z_m \in \{0, 1, x_1, \ldots, x_m\}$. A *monotone circuit* is a Boolean circuit which uses only the binary gates $\wedge, \vee$ and constants $0, 1$. A *monotone real circuit* (defined in [25]) computes a Boolean function, but the intermediary steps in the computation can be real numbers. The only restriction is that the gates consist of binary monotone real functions. (A function $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is monotone if $x_1 \leq y_1, x_2 \leq y_2$ implies $g(x_1, x_2) \leq g(y_1, y_2)$). The size of a circuit is the number of its gates. A monotone Boolean circuit can be viewed as a special case of a monotone real circuit. It was shown by Rosenbloom in [29] that monotone real circuits are strictly more powerful.

Cutting Planes: Cutting Planes is a proof system designed to show that a given set of linear inequalities $\mathcal{L}$ has no $0,1$-solution. A Cutting Planes refutation starts from

[3]Note that we cannot add the condition $|A| = t$ to (i) because monotone Boolean circuits for slice functions are as powerful as general Boolean circuits, cf. [31]. We could fix the size of A in (i), but it must a value essentially larger than t.

the inequalities in $\mathcal{L}$, produces new inequalities by means of simple syntactic rules (namely, adding two inequalities and the "rounding-up" rule), until it reaches the contradictory inequality $0 \geq 1$. Since all our results hold true for a stronger version, the *Semantic Cutting Planes system,* we only define the latter one. (The semantic version is properly stronger, as shown in [13].) For a linear inequality L of the form $\sum_{i=1}^{n} a_i x_i \geq a_0$, where $a_0, \ldots, a_n \in \mathbb{R}$, we say that a Boolean assignment $\alpha \in \{0,1\}^n$ *satisfies* L, if $\sum_{i=1}^{n} a_i \alpha_i \geq a_0$. We say that an inequality L_3 *semantically follows* from L_1, L_2, if every assignment which satisfies both L_1 and L_2 satisfies also L_3. Let $\mathcal{L}$ be a set of linear inequalities. A *Semantic Cutting Planes refutation* of $\mathcal{L}$ is a sequence $L_1, \ldots, L_s$ such that: (1) every L_i is either an element of $\mathcal{L}$, or it semantically follows from (at most) two previous inequalities, (2) L_s is the inequality $0 \geq 1$. The *size* of the refutation is s. A clause can be identified with a linear inequality, and a CNF with a set of linear equalities. For example, the clause $\{x, y, \neg z\}$ would be written as $x + y + (1 - z) \geq 1$. Hence, we can talk about Semantic Cutting Planes refutations of CNFs.

III. Monotone functions as certificates of unsatisfiability

In this section, we give some comments about unsatisfiability certificates, as defined in the Introduction. The first observation is:

Proposition 5. *A CNF $\mathcal{C}$ is unsatisfiable if and only if a monotone certificate for $\mathcal{C}$ exists (for every partition X_0, X_1 of the variables).*

Proof: Let $\mathcal{C}$ be as in (1). Note that $\mathcal{C}$ is satisfiable iff there exist $A \subseteq [m]$ such that both $A_0 := \{C_i^0 : i \in A\}$ and $A_1 := \{C_i^1 : i \in [m] \setminus A\}$ are satisfiable. If $\mathcal{C}$ is satisfiable, a certificate for F would have to have $F(A) = 0$ (by (3)) and $F(A) = 1$ (by (2)). If $\mathcal{C}$ is unsatisfiable, define F as follows: $F(A) = 0$, if A_0 is satisfiable, and $F(A) = 1$ otherwise. Then the condition (3) holds. Condition (2): if A_1 is satisfiable then A_0 is not, and so $F(A) = 1$. ■

In the proof, we have shown that F, defined by:

$$F(A) = 0 \text{ iff } \{C_i^0 : i \in A\} \text{ is satisfiable,} \qquad (4)$$

is an X_0, X_1-certificate for $\mathcal{C}$. This function is a restriction of the function $\text{UNSAT}_{k,n}$.

We note that the definition of unsatisfiability certificate can be more succinctly stated as follows. For $\mathcal{C}$ as in (1), an X_0, X_1-certificate for $\mathcal{C}$ is a monotone Boolean function such that, for every assignment α_0 to X_0 and α_1 to X_1:

$$\begin{aligned} F(C_1^0(\alpha_0), \ldots, C_m^0(\alpha_0)) &= 0\,, \\ F(\neg C_1^1(\alpha_1), \ldots, \neg C_m^1(\alpha_1)) &= 1\,. \end{aligned}$$

A. Communication complexity

The function F can be more easily understood in terms of its Karchmer-Wigderson game. For a monotone function $f : \{0,1\}^m \to \{0,1\}$, recall its monotone KW-game: Player I has input $A \subseteq [m]$ such that $f(A) = 1$, Player II an input B with $f(B) = 0$, and they are supposed to find a coordinate $i \in A \setminus B$. For an unsatisfiability certificate for $\mathcal{C}$, this boils down to the following game.

X_0, X_1-game for $\mathcal{C}$: Player I on has an assignment α_1 to the variables X_1, Player II an assignment α_0 to the variables X_0, and they are supposed to agree on a clause $C_i \in \mathcal{C}$ which is false in the assignment $\alpha_0 \cup \alpha_1$.

Assume that we have an X_0, X_1-unsatisfiability certificate F for $\mathcal{C}$, and a protocol P for its monotone KW game. Then the protocol also solves the X_0, X_1-game for $\mathcal{C}$. For, given α_0 and α_1, Player I and II can each first generate the input $A := \{i : C_i^1(\alpha_1) = 0\}$ and $B := \{i : C_1^0(\alpha_0) = 1\}$. By definition of F, $F(A) = 1$ and $F(B) = 0$. Running the protocol P on A, B, the players find $i \in A \setminus B$ which corresponds to an unsatisfied clause. Conversely, given a protocol for the X_0, X_1-game, it can be modified to a protocol for some X_0, X_1-certificate for $\mathcal{C}$.

These observations give:

Remark 6. *Let $D(\mathcal{C})$ be the smallest depth of a monotone Boolean circuit computing an X_0, X_1-certificate for $\mathcal{C}$. Let $CC(\mathcal{C})$ be the deterministic communication complexity of the X_0, X_1-game for $\mathcal{C}$. Then, up to a constant factor, $\log(D(\mathcal{C}))$ and $CC(\mathcal{C})$ are the same.*

We can also point out that *randomized* communication complexity is a lower bound on monotone *real* circuit depth of F. Since we are interested in the size of proofs and circuits, rather than depth, we will not use either fact in the sequel. However, the game theoretic viewpoint is especially useful when considering upper-bounds. For example, consider an unsatisfiable system of linear equations over GF_2, $L_1 = a_1, \ldots, L_m = a_m$, where L_i are linear functions in variables $X_0 \cup X_1$ and $a_i \in \{0,1\}$. Then, given assignments to the variables X_0 and X_1, the two players can identify an unsatisfied equation using only $O(\log m)$ bits of communication (an exercise). Hence, for Tseitin's formulas, or any system of unsatisfiable inequalities over GF_2 presented as a polynomial size CNF, there is a simple communication protocol of depth $O(\log n)$. Consequently, unsatisfiability certificates for such formulas can be computed by circuits of logarithmic depth (and polynomial size).

B. Proof complexity; a comparison with interpolation

The notion of unsatisfiability certificate is closely related to feasible interpolation, a concept introduced by Krajíček [23]. In the interpolation setting, we have a partition of the variables into *three* parts Y, X_0, X_1 and an unsatisfiable CNF

$$\mathcal{D} = \mathcal{D}_0(Y, X_0) \cup \mathcal{D}_1(Y, X_1)\,, \qquad (5)$$

where the formulas $\mathcal{D}_e$ may only contain variables from the displayed sets. We want to find a function $f : \{0,1\}^{|Y|} \to \{0,1\}$ such that, for every $\alpha \in \{0,1\}^{|Y|}$,

(i) If $\mathcal{D}_0(\alpha, X_0)$ is satisfiable then $f(\alpha) = 0$,
(ii) If $\mathcal{D}_1(\alpha, X_1)$ is satisfiable then $f(\alpha) = 1$.

The function f is called an *interpolant* of $\mathcal{D}_0$ and $\mathcal{D}_1$.

To have a specific example in mind, a classical choice is the *Clique-Coloring Principle.* Think of the common variables Y as representing edges of n-vertex graph. Then $\text{Clique}_n^k(Y, X_1)$ is a CNF formula asserting that X_1 defines a clique of size k in Y. This can be written as a polynomial-size CNF, which is monotone in Y (i.e., no variable in Y is negated). Similarly $\text{Color}_n^k(Y, X_0)$ is a small CNF formula asserting that X_0 defines a k-coloring of Y. Then $\text{Color}_n^k(Y, X_0) \wedge \text{Clique}_n^{k+1}(Y, X_1)$ is unsatisfiable. An interpolant is a function which accepts graphs with $(k+1)$-clique and rejects k-colorable graphs. Improving the seminal result of Razborov [28], it was shown in [2] that, for a suitable choice of k, every such f requires monotone circuit of size $2^{n^{\Omega(1)}}$. In [25] this was extended to monotone real circuits.

This in turn provides lower-bounds on sizes of refutations in some proof systems:

Theorem 7 ([23], [25], [13])**.** *Let $\mathcal{D}$ be as in (5) where $\mathcal{D}_1$ is monotone in Y. Assume that $\mathcal{D}$ has a resolution refutation with s proof lines. Then there is an an interpolant of $\mathcal{D}_0$ and $\mathcal{D}_1$ of a monotone circuit size $O(s)$. For Cutting Planes (and Semantic Cutting Planes) refutations, the same holds with monotone real circuit of size $O(s + |Y| \cdot |\mathcal{D}|)$.*

We now explain how to convert unsatisfiability certificates to interpolants. Let $\mathcal{C}$ be as in (1). Introduce m fresh variables $Y = \{y_1, \ldots, y_m\}$, and consider the CNF $\mathcal{D}$:

$$C_1^0 \cup \{\neg y_1\}, \ldots, C_m^0 \cup \{\neg y_m\}, C_1^1 \cup \{y_1\}, \ldots, C_m^1 \cup \{y_m\},$$

and let $\mathcal{D}_0(X_0, Y)$ be the first m clauses, and $\mathcal{D}_1(X_1, Y)$ the rest. Then one can see that every monotone interpolant of $\mathcal{D}_0$ and $\mathcal{D}_1$ is also an X_0, X_1-certificate for $\mathcal{C}$.

This implies that the above interpolation theorem applies also to certificates of unsatisfiability:

Theorem 8. *Assume that a CNF $\mathcal{C}$ has a resolution refutation with s proof lines. Then, for every partition of the variables, $\mathcal{C}$ has an unsatisfiability certificate computable by a monotone circuit of size $O(s)$. For Cutting Planes (and Semantic Cutting Planes) refutations, the same holds with monotone real circuit of size $O(s + |\mathcal{C}|^2)$.*[4]

Proof: $\mathcal{C}$ can be obtained from $\mathcal{D}$ by resolving $C_i^0 \cup \{\neg y_i\}$ and $C_i^1 \cup \{y_i\}$ for every $i \in [m]$. Hence $\mathcal{D}$ has a refutation with $O(s)$ proof lines (w.l.o.g., we can assume $m \leq s$). Since the variables Y appear only negatively in $\mathcal{D}_0$ (and only positively in $\mathcal{D}_1$), there exists an interpolant f of $\mathcal{D}_0$ and $\mathcal{D}_1$ of monotone circuit size $O(s)$.This gives an X_0, X_1-certificate of monotone circuit size $O(s)$. The case of Cutting Planes is similar. ■

[4]In fact, the term $|\mathcal{C}|^2$ is redundant, as would be revealed by a more careful argument. This is irrelevant for our applications.

One can prove Theorem 8 directly, without recourse to Theorem 7. This way, one can demonstrate a tighter connection between unsatisfiability certificates and proofs. Let us outline the construction in the case of Resolution.

Given a minimal resolution refutation of $\mathcal{C}$, a monotone circuit M computing a certificate F for $\mathcal{C}$ can be directly constructed as follows: the graph of M is the same as the graph of the refutation. Initial clauses C_i are replaced by y_i, the variables of the function F. For the inner gates, we write $\wedge$ if the proof resolves a variable in X_0, and $\vee$ if the proof resolves a variable in X_1.

To prove that this, indeed, gives us an unsatisfiability certificate, we need a generalization of this concept. This more general version will also be used in Theorem 2. For a set of assignments Γ, a CNF is called Γ-satisfiable if it has a satisfying assignment in Γ. For $e \in \{0,1\}$, let Γ_e be a set of Boolean assignments to the variables X_e. We will say that F is Γ_0, Γ_1-*certificate for* $\mathcal{C}$, if conditions (2) and (3) hold when "satisfiable" is replaced with "Γ_1-satisfiable" and "Γ_0-satisfiable", respectively.

Now one can easily prove by induction on the depth that the subcircuit of M determined by a node v is a Γ_0, Γ_1-unsatisfiability certificate for $\mathcal{C}$ where

$$\Gamma_e = \{\alpha \ : \ (C \restriction X_e)(\alpha) = 0\},\ e \in \{0,1\},$$

and where C is the clause that is at the node v in the Resolution proof.

One can also see that Theorem 7 can be viewed as a consequence Theorem 8. That is, we now explain how to obtain an interpolant from a certificate. First, a simple lemma (the proof is straightforward and we omit it):

Lemma 9. *Let F be an X_0, X_1-certificate for $\mathcal{C}$. For $e \in \{0,1\}$, let Γ_e be a set of Boolean assignments to X_e. Assume that $\mathcal{C}' \subseteq \mathcal{C}$ is such that every $C \in \mathcal{C} \setminus \mathcal{C}'$ is satisfied by every assignment from $\Gamma_0 \times \Gamma_1$. Then some Γ_0, Γ_1-certificate for $\mathcal{C}'$ is a projection of F.*

Suppose now $\mathcal{D}$ is as in (5). Assume that $\mathcal{D}_1(Y, X_1)$ is monotone in Y, where $Y = \{y_1, \ldots, y_m\}$. Introduce fresh variables $z_1, \ldots, z_m$ and let $\mathcal{C}$ be the CNF

$$\mathcal{D}_0(Y, X_0) \cup \mathcal{D}_1(Z, X_1) \cup \{\{\neg z_i, y_i\} : i \in [m]\}.$$

If $\mathcal{D}$ is unsatisfiable then so is $\mathcal{C}$. Let $X_0' := X_0 \cup Y$ and $X_1' := X_1 \cup Z$. Let Γ_0 be the satisfying assignments of $\mathcal{D}_0(Y, X_0)$, and Γ_1 the satisfying assignments of $\mathcal{D}_1(Z, X_1)$. Then one can see that a Γ_0, Γ_1-certificate for $\mathcal{C}$ is a monotone interpolant of $\mathcal{D}_0(Y, X_0)$ and $\mathcal{D}_1(Z, X_1)$. Hence a monotone interpolant of $\mathcal{D}_0(Y, X_0)$ and $\mathcal{D}_1(Z, X_1)$ is a projection of X_0', X_1'-certificate for $\mathcal{C}$.

IV. The monotone lower bounds

In this section, we explain our lower bound technique and illustrate it on a simpler version of Theorem 2 in subsection IV-B. We then prove Theorems 2, 4, and 3.

A. The lower bound criterion

Our lower bounds use standard techniques from monotone circuit lower bounds. We use the machinery presented in a monograph of Jukna [21], which was developed in [7], [20].

Let $f : \{0,1\}^m \to \{0,1\}$ be a monotone function. Again, we view inputs of f as subset of $[m]$. We call $A \subseteq [m]$ an *accepting input*, if $f(A) = 1$. We call it a *rejecting input*, if $f([m] \setminus A) = 0$. Let $\mathcal{F}$ be a *multiset* whose elements are subsets of $[m]$. For a parameter r, let $\sharp_r(\mathcal{F})$ be the maximum number of elements of $\mathcal{F}$, including multiplicities, containing a fixed r-element set. Namely, the maximum size of the multiset $\{A \in \mathcal{F} : B \subseteq A\}$ over all r-element sets B.

Lemma 10 (Jukna)**.** *Let f be a monotone Boolean function in m variables and $2 \le r_0, r_1 \le m$. Let $\mathcal{F}_1$ be a multiset of accepting inputs and $\mathcal{F}_0$ a multiset of rejecting inputs of f. Then every monotone real circuit computing f has size at least the minimum of*

$$\frac{|\mathcal{F}_1| - (r_0-1)\sharp_1(\mathcal{F}_1)}{(2r_0)^{r_1+1} \cdot \sharp_{r_1}(\mathcal{F}_1)} \quad \textit{and} \quad \frac{|\mathcal{F}_0|}{(2r_1)^{r_0+1} \cdot \sharp_{r_0}(\mathcal{F}_0)}$$

Proof: This follows from Theorem 9.21 in [21] in the same way as Theorem 9.18 follows from Theorem 9.17 in that book. The only difference is that we work with multisets, rather than sets. ∎

The asymmetry in the lemma can make computations cumbersome. We prefer to work with a slightly relaxed, but simpler lower bound:

$$\min_{e \in \{0,1\}} \frac{|\mathcal{F}_e|}{(2r_{1-e})^{2r_e} \cdot \sharp_{r_e}(\mathcal{F}_e)}\,, \tag{6}$$

assuming $(r_0 - 1)\sharp_1(\mathcal{F}_1) \le |\mathcal{F}_1|/2$.

In our applications, we will usually consider $\mathcal{F}$ of the following form. Let $\mathcal{C} = \{C_1, \ldots, C_m\}$ be a CNF in n variables and Γ a set of Boolean assignments to the variables. Suppose that $f : \{0,1\}^m \to \{0,1\}$ rejects every subset of $\mathcal{C}$ satisfiable in Γ, namely $f(A) = 0$ whenever $\{C_i : i \in A\}$ is satisfied by an assignment from Γ. This is the situation in Theorem 4, as well as the definition of unsatisfiability certificate (where Γ is simply the set of all assignments). For an assignment $\alpha \in \Gamma$, let $A_\alpha \subset [m]$ be the set of indices of clauses falsified by α, namely

$$A_\alpha := \{i \in [m] : C_i(\alpha) = 0\}\,.$$

Then every clause in $[m] \setminus A_\alpha$ is satisfied by α and hence A_α is a rejecting input of f. Let $\mathcal{F}(\mathcal{C}, \Gamma)$ be the multiset of A_α's, for all assignments $\alpha \in \Gamma$.

In order to apply Lemma 10, we want to estimate $\sharp_r(\mathcal{F}(\mathcal{C}, \Gamma))$. This will follow from simple expansion properties of clauses in $\mathcal{C}$. Given $1 \le r \le m$, define

$$\mathrm{Ex}_r(\mathcal{C}) := \min_{T \subseteq \mathcal{C}, |T| = r} \frac{|\bigcup_{C \in T} \mathrm{Var}(C)|}{r}\,.$$

Lemma 11. *For f and $\sharp_r(\mathcal{F}(\mathcal{C}, \Gamma))$ as discussed, $\mathcal{F}(\mathcal{C}, \Gamma)$ is a multiset of rejecting inputs of f. Moreover,*

$$\sharp_r(\mathcal{F}(\mathcal{C}, \Gamma)) \le 2^{n - r\mathit{Ex}_r(\mathcal{C})}\,.$$

Proof: We have already explained that $\mathcal{F}(\mathcal{C}, \Gamma)$ consists of rejecting inputs, and it remains to prove the inequality. It is enough to take Γ as the set of all assignments. Given an $T \subseteq \mathcal{C}$ of size r, we want to estimate the number assignments α such that $T \subseteq A_\alpha$. In other words, every clause $C \in T$, is false under α. Let S be the set of such assignments. If $\alpha \in S$, then we must have $\alpha(\ell) = 0$ for every literal ℓ appearing in some $C \in T$. Hence if two clauses in T contain both a variable and its negation, S is empty. Otherwise, all assignments in S have the same value on variables in $\bigcup_{C \in T} \mathrm{Var}(C)$. Hence

$$|S| \le 2^{n - |\bigcup_{C \in T} \mathrm{Var}(C)|} \le 2^{n - r\mathrm{Ex}_r(\mathcal{C})}\,,$$

as required. ∎

B. Unsatisfiability certificates for bipartite CNFs

In this subsection, we give a lower bound for random *bipartite* CNFs (defined below). The proof is easier than the proof for general random CNFs, it works for small k, and also for CNFs of large clause density. The proof of Theorem 2 itself is deferred to the next section.

Let us fix a partition X_0, X_1 of $2n$ variables with $|X_0| = |X_1| = n$. A clause of width $2k$ is called *balanced*, if it contains exactly k literals from X_0 (and X_1). A random bipartite CNF is obtained by independently choosing m balanced clauses uniformly and independently at random. We want to show that an X_0, X_1-certificate for a random bipartite CNF requires large monotone circuit.

The basic ingredient is the next lemma.

Lemma 12. *Let $\mathcal{C}$ be a random k-CNF in n variables and m clauses. Assume that $m \le (n/40rk)^{k(1-\epsilon)}$ where $0 < \epsilon \le 1/2$ is fixed and $r = r(n)$, $k = k(n)$ are parameters with $\lim_{n\to\infty} rk = \infty$. Then, with high probability,*

$$\mathit{Ex}_r(\mathcal{C}) \ge \epsilon k\,.$$

Proof: Given a set Z of variables and $i \in [m]$,

$$\Pr[\mathrm{Var}(C_i) \subseteq Z] = \binom{|Z|}{k}\binom{n}{k}^{-1} \le \left(\frac{\mathrm{e}|Z|}{n}\right)^k\,.$$

Hence, for a fixed $A \subseteq [m]$ of size r,

$$\Pr[|\bigcup_{i \in A} \mathrm{Var}(C_i)| < \epsilon rk] \le \left(\frac{\mathrm{e}n}{\lfloor \epsilon rk \rfloor}\right)^{\lfloor \epsilon rk \rfloor} \left(\frac{\mathrm{e}\lfloor \epsilon rk \rfloor}{n}\right)^{rk} \le \left(\frac{20rk}{n}\right)^{rk(1-\epsilon)}\,.$$

Since there are at most m^r of the sets A, we have

$$\Pr[\mathrm{Ex}_r(\mathcal{C}, X) < \epsilon k\,] \le \left(m \left(\frac{20rk}{n} \right)^{k(1-\epsilon)} \right)^r ,$$

which tends to zero if $m \le (n/40rk)^{k(1-\epsilon)}$. ∎

The main theorem about bipartite CNFs is as follows.

Theorem 13. *Let $k = k(n)$ be a function of n such that $\lim_{n\to\infty} k = \infty$ and $k \le \log n$. Assume that $m \le n^{k(1-\epsilon)}$ for some constant $\epsilon > 0$. Let $\mathcal{C}$ be a random bipartite $2k$-CNF in $2n$ variables with m clauses. Then, w.h.p., every X_0, X_1-certificate for $\mathcal{C}$ requires monotone real circuit of size $2^{2^{\delta k}}$, where $\delta > 0$ is a constant depending on ϵ. If on the other hand, if $k \ge 10 \log n$ and $m \le 2^{ck}$, where c is an arbitrarily large constant, the lower bound is $2^{\Omega(n)}$.*

Note that the main bound becomes super-polynomial if $k \ge c' \log\log n$, where c' is a constant depending on ϵ.

Proof: Let $\mathcal{C}$ be as assumed and let F be an unsatisfiability certificate. We want to apply the bound from Lemma 10 with $r_0 = r_1 := r$. Let Γ_e, $e \in \{0,1\}$, be the set of all Boolean assignments to X_e. Let $\mathcal{F}_0 := \mathcal{F}(\mathcal{C} \restriction X_0, \Gamma_0)$, as in Lemma 11. It is a multiset of rejecting inputs. Similarly, one can see that $\mathcal{F}_1 := \mathcal{F}(\mathcal{C} \restriction X_1, \Gamma_1)$ is a multiset set of accepting inputs of F. We have

$$|\mathcal{F}_0|, |\mathcal{F}_1| = 2^n .$$

Set ϵ' small enough: $\epsilon' := \epsilon/4$ suffices. Let $r := \lfloor 2^{\epsilon' k/4} \rfloor$. In the sequel we will omit the integer part signs. This is possible, because we assume that k is sufficiently large and thus the resulting error can be compensated by slightly changing the constants. (This also concerns the proofs below.) The previous lemma gives $\mathrm{Ex}_r(\mathcal{C} \restriction X_e) \ge \epsilon' k$ for both $e \in \{0,1\}$ – this is because $\mathcal{C} \restriction X_e$ is a random k-CNF in n variables and

$$(n/40rk)^{k(1-\epsilon')} \ge (n/40k2^{k\epsilon'/4})^{k(1-\epsilon')} \ge \\ \ge (n/n^{\epsilon'})^{k(1-\epsilon')} \ge n^{k(1-\epsilon)} \ge m ,$$

for n sufficiently large. Hence,

$$\sharp_r(\mathcal{F}_0), \sharp_r(\mathcal{F}_1) \le 2^{n-\epsilon' k r} .$$

Since $\sharp_1(\mathcal{F}_1) = 2^{n-k}$, we have $(r-1)\sharp_1(\mathcal{F}_1) \le |\mathcal{F}_1|/2$. Hence the bound in (6) becomes

$$\frac{|\mathcal{F}_e|}{(4r^2)^r \sharp_r(\mathcal{F}_e)} \ge \left(\frac{2^{\epsilon' k}}{4r^2} \right)^r \ge 2^r = 2^{2^{\epsilon' k/4}} ,$$

for k large enough.

For the "moreover" part, apply Lemma 12 with $\epsilon := 1/2$. Set $r := \alpha n/k$ where α is a small enough positive constant chosen so that the assumption of the lemma is satisfied. The lower-bound is again

$$\left(\frac{2^{\epsilon k}}{4r^2} \right)^r \ge \left(\frac{2^{k/2} k^2}{4\alpha^2 n^2} \right)^r \ge \left(2^{k/4} \right)^r = 2^{\alpha n/4} .$$

∎

Corollary 14. *Assuming the same conditions for k, m and n, w.h.p. a random bipartite $2k$-CNF in $2n$ variables with m clauses requires Semantic Cutting Planes proofs of size $2^{2^{\Omega(k)}}$. If on the other hand $k \ge 10 \log n$ and $m = 2^{O(k)}$, the lower bound is $2^{\Omega(n)}$.*

C. Proof of Theorem 2

Let X_0, X_1 be a partition of $2n$ variables with $|X_0| = |X_1| = n$. Let $\mathcal{C}$ be a random k-CNF in variables $X_0 \cup X_1$ with m clauses. We will assume that $\log n \le k \le n$ and $m = O(n2^k)$.

Fix a constant $\delta \in (0, 1/2)$. For $e \in \{0,1\}$, we will call a clause C an *X_e-clause*, if it contains at least $(1-\delta)k$ variables from X_e. A clause which is neither X_0-clause nor X_1-clause will be called *a mixed clause*. Such a clause has at least δk variables from both X_0 and X_1. Let

$$\mathcal{D}_e := \{C_i \restriction X_e : C_i \text{ is } X_e\text{-clause}\} , \text{ for } e \in \{0,1\} ,$$
$$\mathcal{C}_{mix} := \{C_i : C_i \text{ is mixed}\} .$$

Let $H(\delta) := -\delta \log_2(\delta) - (1-\delta)\log_2(1-\delta)$ denote the binary entropy of δ.

Lemma 15. *For some constant $c_1 > 0$,*

$$\Pr[|\mathit{Var}(C_i) \cap X_e| \ge (1-\delta)k] \le c_1 \cdot 2^{-k(1-H(\delta))} .$$

Proof: For $r \le k$,

$$\Pr[|\mathrm{Var}(C_i) \cap X_e| = r] = \binom{n}{r}\binom{n}{k-r}\binom{2n}{k}^{-1} .$$

This can be rewritten as

$$\binom{k}{r}\binom{2n-k}{n-r}\binom{2n}{n}^{-1}$$

Using $\binom{2n}{n} \sim 2^{2n}(\pi n)^{-1/2}$ and $\binom{2n-k}{n-r} = O(2^{2n-k}(\pi(2n-k))^{-1/2})$, we have

$$[\Pr[|\mathrm{Var}(C_i) \cap X_e| = r] \le c_1 \binom{k}{r} 2^{-k} ,$$

for some constant c_1. Hence,

$$\Pr[|\mathrm{Var}(C_i) \cap X_e| \ge (1-\delta)k] \\ \le c_1 2^{-k} \sum_{r=0}^{\delta k} \binom{k}{r} \le c_1 2^{-k} 2^{H(\delta)k} ,$$

using the well-known upper bound on the sum of binomial coefficients [3]. ∎

Lemma 16. *Let $e \in \{0,1\}$. There exists a constant c_2 that depends only on the constant in $m = O(n2^k)$ such that the following happen w.h.p.:*

(i) $\mathcal{C}$ contains at most $c_2 n 2^{H(\delta)k}$ X_e-clauses.

(ii) For every variable $x \in X_e$, x is contained in at most $c_2 k 2^{H(\delta)k}$ X_e-clauses.

(iii) If $r = \omega(1)$ and $rk \le c_2 n$, then $Ex_r(\mathcal{C}_{mix}) \ge \delta k/2$.

Proof: Part (i) is an application of the Chernoff bound: the expected number of X_e-clauses is at most $c_1 m 2^{-k(1-H(\delta))} = O(n 2^{H(\delta)k})$.

(ii) Given $x \in X_e$, the probability that x is contained in $\ge s$ X_e-clauses is at most

$$\binom{m}{s}(c_1 2^{-k(1-H(\delta))}k/n)^s \le (\mathrm{e} c_1 m 2^{-k(1-H(\delta))}k/ns)^s\,.$$

If $s = 3c_1 2^{-k(1-H(\delta))}km/n = O(k2^{H(\delta)k})$, the bound is $(\mathrm{e}/3)^s$, hence the probability goes to zero as $k \to \infty$.

(iii) is a straightforward application of Lemma 12, with $\epsilon = 1/2$ and replacing k by δk. This is because a random δk clause in the variables X_e can be obtained by first generating a random mixed clause, and picking a random δk-subset from its restriction to X_e. Note that the condition $m \le (n/40rk)^{k(1-\epsilon)}$ from Lemma 12 is satisfied, because we assume $k \ge \log n$ and we can pick the constant c_2 sufficiently small. ∎

Lemma 17. *Assume $k \ge c \log n$, where $c > 1$ is a constant. Then there exists $\delta \in (0, 1/2)$ so that, w.h.p., both $\mathcal{D}_0$ and $\mathcal{D}_1$ have at least 2^{n-1} satisfying assignments.*

Proof: This is an application of the Lovász Local Lemma. Set δ so that $(1 - \delta - H(\delta))c > 1$. Assume that $\mathcal{C}$ is fixed, but satisfies i and ii from the previous lemma.

Pick an assignment α to the variables X_e uniformly at random. Given a clause D in $\mathcal{D}_e$, let E_D be the event that α does not satisfy D. This happens with probability $\le p$ where $p := 2^{-(1-\delta)k}$. Two events E_D and $E_{D'}$ are independent if D and D' do not share a variable. Hence, E_D is independent of all but d other events, where $d := c_2 k^2 2^{H(\delta)k}$. To every E_D, we assign the weight $q := 2p$. The condition of the Local Lemma $p \le q(1-q)^d$ is satisfied, because $qd = c_2 k^2 2^{(H(\delta)+\delta-1)k} \to 0$ as $k \to \infty$, which follows from the assumption $(1 - \delta - H(\delta))c > 1$. Let $s := c_2 n 2^{kH(\delta)}$ be the expected size of $\mathcal{D}_e$. Hence, by the Local Lemma, the probability that $\mathcal{D}_e$ is satisfied is at least $(1-q)^s$, which approaches 1. This is because $qs = 4n2^{(H(\delta)+\delta-1)k}$ and we assumed that $(1 - H(\delta) - \delta)k \ge c' \log n$ with $c' > 1$. ∎

We now proceed to prove Theorem 2. Assume that F is a certificate for $\mathcal{C}$. Fix δ from Lemma 17. For $e \in \{0,1\}$, let Γ_e be the set of satisfying assignments of $\mathcal{D}_e$. We first restrict F to talk only about the mixed clauses. Namely, by Lemma 9, there is a Γ_0, Γ_1-certificate g for $\mathcal{C}_{mix}$ which is a projection of F.

In order to lower-bound circuit size of g, we again use the bound from Lemma 10 with $r_0 = r_1 = r$. By the previous lemma, we can assume $|\Gamma_e| \ge 2^{n-1}$. Hence

$$|\mathcal{F}_e| \ge 2^{n-1}\,, e \in \{0,1\}\,.$$

Moreover, we have $\sharp_1(\mathcal{F}_e) \le 2^{n-\delta k}$.

Before proceeding with the proof, note that if k is large, e.g., $k \ge \sqrt{n}$, we automatically get the lower-bound $2^{n^{\Omega(1)}}$. For, the trivial lower-bound for g is $2^{\delta k - 1}$. This "trivial bound" is obtained by noting that $\min_{e\in\{0,1\}} \frac{|\mathcal{F}_e|}{\sharp_1(\mathcal{F}_e)}$ is a lower-bound on the number of variables g depends on.

So let us assume $k \le \sqrt{n}$. Set $r := n^{\delta/8}$. By Lemma 16, part (iii), $\mathrm{Ex}_r(\mathcal{C}_{mix} \restriction X_e) \ge \delta k/2$. Hence, by Lemma 11,

$$\sharp_r(\mathcal{F}_e) \le 2^{n - \delta k r/2}\,,\ e \in \{0,1\}\,.$$

Finally, (6) gives the lower bound

$$\frac{|\mathcal{F}_e|}{(2r)^{2r} \cdot \sharp_r(\mathcal{F}_e)} \ge \left(\frac{2^{\delta k/2}}{4r^2}\right)^r \ge \left(\frac{n^{\delta/2}}{4r^2}\right)^r \ge 2^{n^{\delta/8}}\,,$$

for n sufficiently large. This finishes the proof of Theorem 2.

The clause density can be slightly improved, but not too much:

Remark 18. *Theorem 2 holds for $m = O(n2^{(1+\epsilon)k})$ and $\epsilon > 0$ small enough. If $m \ge c' n 2^{2k}$ for general n, k and c' large enough then, w.h.p., $\mathcal{C}$ has X_0, X_1-certificate with monotone circuit of size $O(m)$.*

The first statement can be proved by inspecting the proof. There is sufficient leeway to change the constants to allow the larger bound on m. The second part follows by noting that a large CNF will contain an unsatisfiable subset in the variables X_0 only. Then the X_0, X_1-communication game has a logarithmic depth protocol; hence the X_0, X_1-certificate has a monotone circuit of size $O(m)$. In more detail: if $\mathcal{C}$ contains an unsatisfiable subset $\{C_i : i \in A\}$ which only depends on the variables X_0, then, given assignments α_0, α_1, Player II can simply send to Player I the name of the clause in A not satisfied by α_0, which takes $O(\log |A|)$ bits.

D. Proof of Theorem 4

Lemma 19. *Assume that f δ-approximates UNSAT$_{k,n}$ on $\binom{[N]}{t}$. Pick a random m-element subset A of $[N]$, for a given $t \le m \le N$. Then, with probability at least $\delta/2$, f accepts at least $\delta/2$-fraction of t-element subsets of A.*

Proof: The expected number of t-element subsets of a randomly chosen m-element set accepted by f is $\ge \delta\binom{m}{t}$. The lemma then follows from Markov's inequality. ∎

We also need a minor modification of Lemma 12. The lemma is stated and proved for random formulas where we allow repetitions in the process of selecting clauses. The lemma is also true, if we select clauses without repetition:

Lemma 20. *Lemma 12 holds when $\mathcal{C}$ is a random CNF with exactly m distinct clauses.*

We now proceed to prove Theorem 4. Assume that f is a monotone function which δ-approximates UNSAT$_{k,n}$ on $\binom{[N]}{t}$, where $n \le t \le N^{1-\epsilon}$ and $N = 2^k\binom{n}{k}$. We will also assume that $k \le n^{\epsilon/8}$: every unsatisfiable k-CNF has at least 2^k clauses. Hence, the function f must depend on at least

2^k variables, otherwise it would be constant, which gives the bound $2^{n^{\Omega(1)}}$ if k is large.

Set ϵ' sufficiently small, $\epsilon' := \epsilon/8$ suffices. Set $m := t^{1+4\epsilon'}$. This guarantees that $m \le N^{1-\epsilon/2} \le (2n)^{(1-\epsilon/2)k}$. Let A be a random m-element subset of $[N]$. Let g be the restriction of f to subsets of A. Using Lemma 10, we want to prove a lower bound on the circuit size of g. We set

$$r_0 := n^{\epsilon'}\ ,\ r_1 := 2^{\epsilon' k/2-3}\,.$$

As $\mathcal{F}_1$, the set of accepting inputs, we take the set of all t-element subsets of A accepted by g. By the previous lemma, with a constant positive probability,

$$|\mathcal{F}_1| \ge \delta' \binom{m}{t}\,, \quad \text{where } \delta' := \delta/2\,.$$

Furthermore,

$$\sharp_{r_1}(\mathcal{F}_1) \le \binom{m-r_1}{t-r_1} \le \binom{m}{t}\cdot\left(\frac{t}{m}\right)^{r_1} \le \binom{m}{t} n^{-4\epsilon' r_1}\,.$$

Hence we have

$$\frac{|\mathcal{F}_1|}{(2r_0)^{2r_1}\cdot\sharp_{r_1}(\mathcal{F}_1)} \ge \delta'\left(\frac{n^{4\epsilon'}}{4r_0^2}\right)^{r_1} \ge \delta' n^{\epsilon' 2^{\epsilon' k/2-3}}\,.$$

$\mathcal{F}_0$ is obtained as follows. View A as the k-CNF $\mathcal{C} := \{C_i, i \in A\}$. Since g rejects all satisfiable subsets of $\mathcal{C}$, we can take $\mathcal{F}_0$ as the multiset $\mathcal{F}(\mathcal{C},\Gamma)$ as in Lemma 11, where Γ is the set of all Boolean assignments. Then

$$|\mathcal{F}_0| = 2^n\,.$$

We apply Lemma 20 to show that, with high probability, $\mathrm{Ex}_{r_0}(\mathcal{C}) \ge \epsilon' k$. The assumption of Lemma 12 is satisfied, as

$$(n/40r_0k)^{k(1-\epsilon')} \ge (n/(40n^{\epsilon'}n^{\epsilon'}))^{k(1-\epsilon')} \ge (n/40^2)^{(1-3\epsilon')k} \ge (2n)^{k(1-\epsilon/2)} \ge m.$$

Hence, by Lemma 11,

$$\sharp_{r_0}(\mathcal{F}_0) \le 2^{n-\epsilon' k r_0}\,.$$

Thus we have

$$\frac{|\mathcal{F}_0|}{(2r_1)^{2r_0}\cdot\sharp_{r_0}(\mathcal{F}_1)} \ge \left(\frac{2^{\epsilon' k}}{4r_1^2}\right)^{r_0} \ge 2^{r_0} = 2^{n^{\epsilon'}}\,.$$

Finally, (6) gives a lower bound which is the minimum of

$$\delta' n^{\epsilon' 2^{\epsilon' k/2-3}}\ ,\ 2^{n^{\epsilon'}}\,.$$

If $k \le \log n$, we obtain a lower bound of the form $n^{2^{\Omega(k)}}$; for large k, we go for the latter bound.

E. Theorem 3: the Weak Bit Pigeonhole Principle

We now show how one can apply unsatisfiability certificates to the Weak Bit Pigeon Hole Principle, proving Theorem 3.

Let $M > N = 2^n$. The CNF formula BPHP_N^M has Mn variables $x_{i,k}$, $i \in [M]$, $k \in [n]$. If we think of them as representing $M \times n$ Boolean matrix, the principle asserts that all the rows of this matrix are distinct. More exactly, let $[x_i \ne x_j]$ be the unique $2n$-CNF asserting that the vectors $(x_{i,1},\ldots,x_{i,n})$ and $(x_{j,1},\ldots,x_{j,n})$ are distinct. The CNF has N clauses. Then BPHP_N^M is the union of $M(M-1)/2$ CNFs

$$[x_i \ne x_j]\,,\ i < j \in [M]\,.$$

Since we assume $M > N$, the formula is unsatisfiable. It has exactly $NM(M-1)/2$ clauses. One can also think of the variables as giving bit-representation of an injection f from $[M]$ to $\{0,1\}^n$ – hence the name the *Bit* Pigeonhole Principle.

We want to show that for some partition of the variables, BPHP requires X_0, X_1-certificates of exponential monotone complexity. This will also imply an exponential Cutting Planes lower bound. For simplicity of presentation, we assume that n *is even.* The partition is as follows, X_0 consists of the first $n/2$ columns of the matrix, and X_1 of the last $n/2$ columns, namely:

$$X_0 := \{x_{i,k} : i \in [M], k \in \{1,\ldots,n/2\}\}\,,$$
$$X_1 := \{x_{i,k} : i \in [M], k \in \{n/2+1,\ldots,m\}\}\,.$$

Assume that F is X_0, X_1-certificate for BPHP. F has $NM(M-1)/2$ variables corresponding to clauses of BPHP. Denote the variables $y_{i,j,k}$ where $i < j \in [M]$ and $y_{i,j,k}$ corresponds to the k-th clause in $[x_i \ne x_j]$. Let $H : \{0,1\}^{M(M-1)/2} \to \{0,1\}$ be obtained by replacing $y_{i,j,k}$ with $y_{i,j}$ for every k. Hence H is a projection of F. We view its inputs as M-vertex graphs. We claim that H *distinguishes graphs G with $\chi(G) \le \sqrt{N}$ from graphs with* $\chi(\bar{G}) \le \sqrt{N}$, namely,

$$\text{if } \chi(G) \le \sqrt{N} \text{ then } H(G) = 0\,, \tag{7}$$
$$\text{if } \chi(\bar{G}) \le \sqrt{N} \text{ then } H(G) = 1\,. \tag{8}$$

Here, $\chi(G)$ is the chromatic number of G and $\bar{G}$ its complement.

To see (7), assume that $\chi(G) \le \sqrt{N}$. Then the CNF

$$\bigcup_{(i,j)\in G} ([x_i \ne x_j] \restriction X_0)$$

is satisfiable: indeed, it is satisfied by the bit-representation of some $\sqrt{N}$-coloring of G. Hence, by the definition of unsatisfiability certificate, $H(G) = 0$. The condition (8) is dual.

Thus we have reduced the lower bound on BPHP to the following:

Theorem 21. *For $M > q^2$, every monotone function that distinguishes M-vertex graphs G with $\chi(G) \leq q$ from graphs with $\chi(\bar{G}) \leq q$ has monotone real circuit complexity $2^{\Omega(q^{1/4})}$.*

The lower bound on the proofs of BPHP in Cutting Planes (Theorem 3) now follows from Theorem 21 and Theorem 7. It now remains to prove the theorem.

Proof of Theorem 21: Assume that F is a monotone function that distinguishes M-vertex graphs G with $\chi(G) \leq q$ from graphs with $\chi(\bar{G}) \leq q$. We want to show that F requires large monotone real circuits. One option how to proceed would be to reduce the clique-coloring function to F using monotone projections. A better bound follows from a direct application of the standard lower bound method, which we will do here. The proof is an adaptation of the forn the clique-coloring problem presented in [21].

We first modify Lemma 10. Let μ be a real valued function defined on subsets of a finite set A. We say that μ is a *legal measure,* if μ is nonnegative,

$$\mu(S) \leq \mu(S \cup \{a\}) \leq \mu(S) + 2 \quad \text{and} \quad |S| \leq \mu(S)^2$$

for every $S \subseteq A$ and $a \in A$. Let, furthermore, $\mathcal{F}$ be a multiset of subsets of A. Then we define

$$\#_r^\mu(\mathcal{F}) := \max_{S,\mu(S)=r} |\{X \in \mathcal{F} \mid S \subseteq X\}|,$$

where the size is counted with multiplicities. The following is again implicit in [21][5]:

Lemma 22 ([21])**.** *Let F be a monotone Boolean function in m variables and $2 \leq r \leq m$. Let $\mathcal{F}_1$, $\mathcal{F}_0$ be multisets of accepting and rejecting inputs of F, respectively. Then every monotone real circuit computing F has size at least the minimum of*

$$\frac{|\mathcal{F}_1| - r^2 \cdot \sharp_1(\mathcal{F}_1)}{(2r)^{4r} \cdot \sharp_r^\mu(\mathcal{F}_1)} \quad \textit{and} \quad \frac{|\mathcal{F}_0|}{(2r)^{4r} \cdot \sharp_r^\mu(\mathcal{F}_0)}.$$

Given a mapping $f : [M] \to [q]$, we define a graph G_f by connecting i with j, $i \neq j$, by an edge iff $f(i) = f(j)$. We take the multisets

$$\mathcal{F}_0 = \mathcal{F}_1 := \{G_f \mid f : [M] \to [q]\},$$

where each graph is counted with the multiplicity given by the number of mappings f that define it. We take a natural measure, which has been used by Jukna and some authors before him, defined as follows. For a set of edges S, $\mu(S)$ is the number of edges of a spanning forest of the graph defined by S. Equivalently, $\mu(S)$ is the number of vertices covered by S minus the number of connected components of S. The standard argument (which is the observation that the probability that random q-coloring colors two vertices by the same color with probability q and that these events are independent for edges in a forest) gives us

$$\frac{|\mathcal{F}_e|}{\sharp_r^\mu(\mathcal{F}_e)} \geq q^r \quad \text{for } e \in \{0,1\}.$$

If we take $r := \lfloor (q/32)^{1/4} \rfloor$, the second fraction in the lemma turns out to be exponential in $2^{\Omega(q^{1/4})}$. Now it only remains to show that $r^2 \cdot \sharp_1(\mathcal{F}_1) \leq |\mathcal{F}_1|/2$. The probability that a random graph G_f contains a fixed edge (i,j) is the probability that f colors i and j by the same color, which is $1/q$. Hence $r^2 \cdot \sharp_1(\mathcal{F}_1) \leq r^2 q^{-1} |\mathcal{F}_1| < |\mathcal{F}_1|/2$ by the choice of r.

V. Open problems

The most interesting problem is to extend our lower bounds to the case of a constant k. That is, prove super-polynomial lower bounds on Cutting Planes refutations of random k-CNFs for a constant k, or at least tree-like Cutting Planes. And, if possible, extend Theorems 2 and 4 in this direction. The following are the simplest problems we do not know how to solve (recall the X_0, X_1-game defined in Section III-A):

(a) Let $\mathcal{C}$ be a random 3-CNF of a suitable clause density above the unsatisfiabilty threshold. Does the X_0, X_1-game for $\mathcal{C}$ require deterministic communication complexity of $n^{\Omega(1)}$? If so, does $\mathcal{C}$ require unsatisfiability certificate of exponential monotone circuit size?
(b) The same questions about random bipartite 6-CNF.

Let us explain some aspects of the problems. First, for a random CNF of large-enough clause density, both the games in (a) and (b) have small *probabilistic* communication complexity. This is because, w.h.p., every assignment will make false a constant fraction of clauses in $\mathcal{C}$, allowing the two players to pick a false clause at random. As tempting as it is, this means that we cannot solve problems (a) or (b) by means of a randomized reduction to Disjointness. Furthermore, by Remark 18, in (a) the "suitable density" must be fairly small, otherwise the (deterministic) communication complexity is logarithmic. Without proof[6], we note the following:

Remark 23. *(i) There exists an explicit unsatisfiable 3-CNF such that every X_0, X_1-certificate requires monotone real circuits of size $2^{n^{\Omega(1)}}$ (for some partition X_0, X_1).*
(ii) Assume that $\mathcal{C}$ is such that every clause in $\mathcal{C}$ contains at most 2 variables from X_0. Then $\mathcal{C}$ has a certificate with monotone circuit of size polynomial in $|\mathcal{C}|$ and depth $O(\log |\mathcal{C}|^2)$. Hence, the X_0, X_1-game for $\mathcal{C}$ has communication complexity $O(\log |\mathcal{C}|^2)$.

[5]Since our function is selfdual, we have stated the lemma with only one measure μ and only one parameter r.

[6]Hint. In part (i) represent the Clique-Coloring principle as a 3-CNF. In part (ii), recall that, as in (4), an X_0, X_1-certificate can be obtained as a monotone projection of $\mathrm{UNSAT}_{2,n}$, where the latter is closely related to s,t-connectivity.

Part (i) shows that exponential lower bounds on unsatisfiability certificates can be proved even for $k = 3$, for a specific $\mathcal{C}$. On the other hand, (ii) shows that using the method of this paper, it is impossible to prove a superpolynomial lower bound for random bipartite 4-CNFs. It seems, however, that for $k \geq 6$, the bottleneck is only in the lower bound methods for monotone real circuits or communication games.

ACKNOWLEDGEMENT

We thank Neil Thapen and Nicola Galesi for useful discussions. This research was supported by ERC grant FEALORA 339691.

REFERENCES

[1] M. Alekhnovich and A. Razborov. Lower bounds for the polynomial calculus: non-binomial case. In *FOCS*, pages 190–199, 2001.

[2] N. Alon and R. B. Boppana. The monotone circuit complexity of boolean functions. *Combinatorica*, 7(1):1–22, 1987.

[3] R. B. Ash. *Information Theory*. Dover Publications, Inc., 1990.

[4] P. Beame, T. Pitassi, and N. Segerlind. Lower bounds for Lovász-Schrijver systems and beyond follow from multiparty communication complexity. *SIAM J. Comput.*, 37(3):845–869, 2007.

[5] E. Ben-Sasson and R. Impagliazzo. Random CNFs are hard for the polynomial calculus. In *FOCS*, pages 415–421, 1999.

[6] E. Ben-Sasson and A. Wigderson. Short proofs are narrow; resolution made simple. *J. ACM*, 48(2):149–169, 2001.

[7] C. Berg. *On Oracles and Circuit – Topics in Computational Complexity*. PhD thesis, KTH, 1997.

[8] V. Chvátal. Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete Mathematics*, 4(4):305–337, 1973.

[9] V. Chvátal and E. Szemeredi. Many hard examples for resolution. *J. ACM*, 35(4):759–768, 1988.

[10] A. Coja-Oghlan and K. Panagiotou. The asymptotic k-sat threshold. *Advances in Mathematics*, 288:985–1068, 2016.

[11] W. Cook, C. R. Coullard, and G. Turán. On the complexity of cutting-plane proofs. *Discrete Applied Mathematics*, 18(1):25–38, 1987.

[12] U. Feige. Relations between average case complexity and approximation complexity. In *STOC*, pages 534–543, 2002.

[13] Y. Filmus, P. Hrubeš, and M. Lauria. Semantic Versus Syntactic Cutting Planes. In *STACS*, 2016.

[14] N. Fleming, D. Pankratov, T. Pitassi, and R. Robere. Random CNFs are hard for cutting planes. *ECCC*, 2017.

[15] E. Friedgut. Sharp thresholds of graph properties, and the k-sat problem. 1998.

[16] R. E. Gomory. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the AMS*, 64(5):275–278, 1958.

[17] A. Haken. The intractability of resolution. *Theoretical Computer Science*, 39:297–308, 1985.

[18] A. Haken and S. A. Cook. An exponential lower bound for the size of monotone real circuits. *J. Comput. Syst. Sci.*, 58(2):326–335, April 1999.

[19] R. Impagliazzo, T. Pitassi, and A. Urquhart. Upper and lower bounds for tree-like cutting planes proofs. In *9th IEEE Symposium on Logic in Computer Science*, pages 220–228, 1994.

[20] S. Jukna. Finite limits and monotone computations: The lower bounds criterion. In *CCC*, pages 302–, 1997.

[21] S. Jukna. *Boolean Function Complexity: Advances and Frontiers*. Springer Publishing Company, Inc., 2012.

[22] M. Karchmer and A. Wigderson. Monotone circuits for connectivity require super-logarithmic depth. In *STOC*, 1988.

[23] J. Krajíček. Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic. *J. Symb. Log.*, 62(2):457–486, 1997.

[24] T. Lee and A. Shraibman. Disjointness is hard in the multiparty number-on-the-forehead model. *CCC*, 2008.

[25] P. Pudlák. Lower bounds for resolution and cutting plane proofs and monotone computations. *The Journal of Symbolic Logic*, 62(3):981–998, 1997.

[26] R. Raz. Resolution lower bounds for the weak pigeonhole principle. *J. ACM*, 51(2):115–138, March 2004.

[27] A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.

[28] A.A. Razborov. Lower bounds on the monotone complexity of some boolean functions. *Soviet Mathematics Doklady*, 31:354–357, 1985.

[29] A. Rosenbloom. Monotone real circuits are more powerful than monotone boolean circuits. *IPL*, 61(3):161–164, 1997.

[30] A. A. Sherstov. The multiparty communication complexity of set disjointness. In *STOC*, pages 525–548, 2012.

[31] I. Wegener. The complexity of boolean functions, 1987.