Electronic Colloquium on Computational Complexity, Report No. 119 (2012)

Pseudo-partitions, Transversality and Locality: A Combinatorial Characterization for the Space Measure in Algebraic Proof Systems*

Ilario Bonacina† and Nicola Galesi‡

Dipartimento di Informatica
Sapienza Università di Roma

September 24, 2012

Abstract

We devise a new combinatorial characterization for proving space lower bounds in algebraic systems like Polynomial Calculus (Pc) and Polynomial Calculus with Resolution (Pcr). Our method can be thought as a Spoiler-Duplicator game, which is capturing boolean reasoning on polynomials instead that clauses as in the case of Resolution. Hence, for the first time, we move the problem of studying the space complexity for algebraic proof systems in the range of 2-players games, as is the case for Resolution.

A very simple case of our method allows us to obtain all the currently known space lower bounds for Pc/Pcr (CT_n, PHP^m_n, BIT-PHP^m_n, XOR-PHP^m_n). The way our method applies to all these examples explains how and why all the known examples of space lower bounds for Pc/Pcr are an application of the method originally given by [2] that holds for set of contradictory polynomials having high degree. Our approach unifies in a clear way under a common combinatorial framework and language the proofs of the space lower bounds known so far for Pc/Pcr.

More importantly, using our approach in its full potentiality, we answer to the open problem [2, 34] of proving space lower bounds in Polynomial Calculus and Polynomials Calculus with Resolution for the polynomial encoding of randomly chosen k-CNF formulas. Our result holds for $k \geq 4$. Then, as proved for Resolution in [12], also in Pc and in Pcr refuting a random k-CNF over n variables requires high space measure of the order of $\Omega(n)$. Our method also applies to the Graph-PHP^m_n, which is a PHP^m_n defined over a constant (left) degree bipartite expander graph. We develop a common language for the two examples.

*This work was supported by the project "the Limits of Theorem Proving" granted by John Templeton Foundation

†ilariobonacina@gmail.com

‡galesi@di.uniroma1.it

ISSN 1433-8092

Contents

1 Extended Abstract

Proof complexity is a research field initiated by Cook and Reckhow [27] that studies the complexity of proving (alternatively refuting) propositional tautologies (alternatively contradictions) in different logical propositional proof systems. The historical motivation for investigating the complexity of proofs is the P vs. NP question. A proof system is said to be *polynomially bounded* if for every tautology $x \in TAUT$ there is a proof $\pi(x)$ of size at most *polynomial* in the size of x. As observed in [27], one way of establishing co-NP$\neq$ NP , and hence P$\neq$ NP would be to prove that there are no polynomially bounded proof systems. One suggested approach to this problem is that of studying proof limits in always stronger proof systems. But proving that NP $\neq$ co-NP showing incrementally that examples of proof systems are not polynomially bounded seems unlikely. Rarely a universal statement is proved by proving all its instances. Nevertheless proving these lower bounds we may hope to uncover hidden computational hardness assumptions and then try to reduce the conjecture to some more approachable problem [38]. This is what is known as the *Cook's Program* in Proof Complexity. Among the most studied proof systems there are the logical systems of Resolution [45, 19] and algebraic proof systems like Polynomial Calculus [26] or Polynomial Calculus with Resolution [2].

1.1 High-level Motivations

1.1.1 Theoretical investigation of Space measure

As remarked by Razborov [43], proof complexity plays the same role in the field of feasible proofs of the role played by the Boolean Circuits/Turing Machine in the field of efficient computations. Hence *Proof Size* in Proof Complexity should be view as *Circuit-Size/Running-Time* in circuit complexity. It is then no surprise that, as for efficient computations we consider *memory occupation* as a measure of efficiency, a notion of *Proof Space Measure* was introduced also for proof systems ([32, 2]) and since then studied and investigated in depth in this field, especially for resolution ([32, 2, 12, 31, 15, 16, 40, 41, 34] among many others). As seen there is a vast bibliography on the space measure for the system of resolution. On the other hand Polynomial Calculus, though being a very well-studied proof systems when considering the size and degree complexity of a proof [26, 23, 24, 44, 42, 13, 37, 3, 36, 35], is still at the beginning of the investigation of the space measure [2, 34]. The reason being that current lower bounds techniques for Resolution space do not hold for algebraic systems that deal with polynomials.

The main motivation of our work is to contribute to the development of the theoretical study of the space complexity measure for propositional proof systems and specifically in algebraic proof systems. We design a new combinatorial characterization for proving space lower bounds in algebraic systems like Polynomial Calculus (Pc) and Polynomial Calculus with Resolution (Pcr). Our approach unifies in a clear way under a common combinatorial framework the proofs of *all* the space lower bounds known so far for Pc/Pcr (CT_n, PHP^m_n, BIT-PHP^m_n, XOR-PHP^m_n). Moreover we answer to the open problem [2, 34] of proving space lower bounds in Pc and Pcr for the polynomial encoding of randomly chosen k-CNF formulas.

1.1.2 Finite Model Theory and Proof Complexity

Atserias [4] discovered a very interesting connection between the fields of finite model theory and propositional proof complexity. This connection was capturing the following informal reasoning: if a formula is hard to refute in Resolution, then for a bounded player should be hard to distinguish it from a satisfiable formula. This was the base for the result that encodings of combinatorial principles as propositional tautologies are hard-to-prove could serve for logical non-expressibility result via combinatorial games. The second link between finite model theory and propositional proof complexity is the tight connection between the number of pebbles needed by the an adversary (Duplicator) to win the existential-pebble game and the concept of width in Resolution. Two crucial facts relate pebble games to resolution proof complexity measures. First Feder and Vardi [33], observed that the satisfiability problem of a k-CNF formula can be identified with the homomorphism problem on relational structures. Then existential-pebble games provide a purely combinatorial characterization of resolution width. Second Ben-Sasson and Galesi [12] invented a 2-player Matching Game to study space lower bounds in Resolution k-CNF formulas. The Matching Game is essentially an existential pebble game which indeed was used by Atserias [4] to establish the connection between Finite Model Theory and Proof Complexity. The main observation is that winning strategies for the adversary in the Matching Game can be described in terms of a class of homomorphisms characterizing Duplicator winning strategies in Ehrenfeucht-Fraïssé games. As an application of this combinatorial characterization Atserias [4] and Atserias and Dalmau [5] got the impressive result relating the space and the width in resolution showing that space is lower bounded by width.

Our work can be view as a first step towards a 2-players game characterization for algebraic systems, i.e. dealing with polynomials, instead that with clauses. Our main definition (k-extendibility) characterizes the winning strategies for an adversary as a class of combinatorial objects. While for resolution the class of homomorphism is in fact a class of partial bounded boolean assignments, in our case we have *Admissible Configurations*, which are pairs containing a partition of a subset of the variables (*pseudo-partitions*) and a whole class of assignments fulfilling some locality properties (*locally modifiability*). Our main definition should also be compared with the definition, given by Esteban, et al. in [31] of wining strategies for getting space lower bounds in $Res(k)$, that is a Resolution system on k-DNF.

1.1.3 SAT-Solvers and Theorem Provers

The *satisfiabilty* problem and the study of complexity measure related to SAT-solvers and theorem provers have recently been matter of research in proof complexity. From a proof complexity point of view an interesting feature of the modern SAT-solvers is that they are still based on the Davis-Putnam-Logemann-Loveland or DPLL procedure [30, 29] augmented with clause learning [9, 39] or similar techniques. It is well-known that the DPLL algorithm applied on UNSAT formulas produce a (tree-like) resolution refutations of that formula. This is the reason why there is a growing interests in studying (theoretically) the complexity of logical proof systems SAT-solver algorithms [9, 6, 22, 14, 21]. Indeed studying the complexity of proofs in such systems allows to understand the potential and limitations of such algorithms for SAT-solving or theorem proving.

It is well-known that the main problems of modern SAT-solvers is that of rapidly accessing huge amount of informations. Typically this algorithms downgrade since they have to access millions of clauses stored into secondary memory and check for assignments on these clauses. Then one of the main bottleneck for these algorithms is represented by the memory occupation.

In proof complexity studying (theoretically, but driven by concrete applied industrial problems) proof size and proof space, one wants to understand how the resources of time and space are linked and how they can be optimized. We could say that the final aim might be that of studying theoretically the limit of applied SAT-solver algorithms and hopefully that of finding some theoretical results indicating how to overcome applied problems (see [9, 6, 22, 14, 21] among many other works in the area).

Polynomial Calculus (Pc) is a proof systems having its algebraic base on the Gröbner Basis Algorithm. Then Pc is surely one of the proof systems that have some hope of producing new insights onto the field of SAT-solvers. For instance Clegg et al., [26] showed an algorithm (based on Gröbner Algorithm) to find in polynomial time in the minimal degree required, a Pc refutation of a set of polynomials. For this reason at that time there was quite some hope polynomial calculus could give raise to better SAT-solvers than those based on Resolution. There are PC-based solvers such as PolyBoRi [20], but in general they seem to be an order of magnitude slower than state-of-the-art solvers.

Our work contributes to better understand theoretically the space measure in Polynomial Calculus. We think that our discrete combinatorial characterization of the space in algebraic proof systems might open the way to better understand how to encode polynomials and devise algorithms (Theorem provers or SAT-solvers) working on polynomials but using discrete combinatorial concepts.

1.2 Proof systems and Space Measure in Proof Complexity

We denote by x a Boolean variable. A literal l is either a variable or its negation. A clause $C = (l_1 \vee \ldots \vee l_k)$ is a disjunction of literals, a term $T = (l_1 \wedge \ldots \wedge l_k)$ is a conjunction of literals. We think of clauses and terms as sets, so that the ordering of the literals is irrelevant and no literals are repeated. We denote the empty clause, i.e. the clause containing no literals, by $\Box$. A clause (term) containing at most k literals is called a k-clause (k-term). A CNF formula $F = C_1 \wedge \ldots \wedge C_m$ is a conjunction of clauses, and a DNF formula is a disjunction of terms. We will think of CNF and DNF formulas as sets of clauses and terms, respectively. A k-CNF formula is a CNF formula consisting of k-clauses, and a k-DNF formula consists of k-terms. A clause C is a clause over a set of variables V if the set of variables it mentions is a subset of V . We similarly define terms, CNFs and DNFs over V .

The *Resolution system* is a refutation system for the set of all unsatisfiable CNF. Resolution uses as its only rule the *Resolution rule*

$$\frac{\{x\} \cup C \qquad \{\neg x\} \cup D}{C \cup D}$$

for clauses C, D and a variable x. The aim in Resolution is to demonstrate unsatisfiability of a clause set by deriving the empty clause. If in a derivation every derived clause is used at most once as a prerequisite of the Resolution rule, then the derivation is called *tree-like*, otherwise it is *dag-like*. The *size* of a Resolution proof is the number of its clauses. The *width* of a clause is its number of literals. The width of a proof is the maximal width of a close in the proof.

Polynomial Calculus (Pc) is a refutational system defined in [26], and based on the ring $\mathbb{F}[x_1, \ldots, x_n]$ of polynomials. Given $p \in \mathbb{F}[x_1, \ldots, x_n]$ we always consider equations of the form $p = 0$, and we simply denote them as p. The equations are intended to hold on $\{0, 1\}^n$ thus the system contains the following logical axioms:

$$x_i^2 - x_i, \quad i \in [n] \ \textit{(Boolean Axioms).}$$

Moreover it has two rules. For any $\alpha, \beta \in \mathbb{F}$, p, q polynomials and variable x:

$$\frac{p \qquad q}{\alpha p + \beta q} \ \textit{(Linear Combination),} \qquad \frac{p}{xp} \ \textit{(Multiplication).}$$

A Pc proof of a polynomial g from a set of initial polynomials $f_1, \ldots, f_m$ (denoted by $f_1, \ldots, f_m \vdash g$) is a sequence of polynomials where each one is either an initial one, a logical axiom, or it is obtained applying one of the rules to previously derived polynomials. A Pc refutation is a proof of the polynomial 1.

Pc is a complete proof system, in the sense that a polynomial g has a Pc proof from a set of polynomials E iff $g(\vec{x}) = 0$ for every $\vec{x} \in \{0, 1\}^n$ which is a common root of E. Moreover E has no common $\{0, 1\}$ solutions (we call E contradictory) iff $1 \in \mathit{Span}(E \cup \{x_i^2 - x_i\}_{i\in[n]})$. Completeness of Pc comes as a corollary of Hilbert's Nullstellensatz (see [28]) and from complete algorithms based on Gröebner bases [26].

We remark here that when we work in Polynomial Calculus, we implicitly assume that the polynomials $\{x_i^2 - x_i\}_{i\in[n]}$ are always included in the set of initial polynomials.

Given a Pc proof Π, the *degree* of Π, $deg(\Pi)$, is the maximal degree of a polynomial in the proof; the *size* of Π, $S(\Pi)$, is the number of monomials in the proof, the *length* of Π, $|\Pi|$, is the number of lines in the proof.

Following what done in [32, 2] for studying space complexity in Resolution and general sequential proof systems, we view a proof in Pc as similar to a non-deterministic Turing machine computation, with a working memory where all derivation steps are saved and a special read-only input tape from which the initials polynomials being refuted (the axioms) can be downloaded. Thus the length of a proof is essentially the time of the computation while the space measures the memory consumption. Following [2] we have:

Definition (Memory Configuration). *A* Memory Configuration *in* Pc *is a set of polynomials. Given a set $F = \{f_1, \ldots, f_m\}$ of initials polynomials and a polynomial g, a* Pc *proof Π of $f_1, \ldots, f_m \vdash g$ can be view as sequence of memory configurations $\Pi = \{\mathcal{C}_0, \ldots, \mathcal{C}_l\}$ such that: $\mathcal{C}_0 = \emptyset$, $\mathcal{C}_l$ contains g and for all $i \in [l]$, $\mathcal{C}_i$ is obtained by $\mathcal{C}_{i-1}$ by one of the following three rules:*

Axiom Download *$\mathcal{C}_i = \mathcal{C}_{i-1} \cup \{p\}$, where p is some initial polynomial $f_j \in F$ or some boolean axiom.*

Inference Adding *$\mathcal{C}_i = \mathcal{C}_{i-1} \cup \{p\}$, where p si some polynomial inferred by using one of the rule of the calculus applied on polynomials occurring in $\mathcal{C}_{i-1}$.*

Erasure $\mathcal{C}_i = \mathcal{C}_{i-1} \setminus \{p\}$, *for some* $p \in \mathcal{C}_{i-1}$.

Following [2] we define the the *space* measure for Pc.

Definition (Space Measure). [1] *The* space *of a* Pc *memory configuration* $\mathcal{C}$, $Sp(\mathcal{C})$ *is the number of distinct monomials occurring in* $\mathcal{C}$. *The space of a* Pc *proof* Π, $Sp(\Pi)$), *is the maximal space of a memory configuration in* Π. *The space of proving* g *from* $\{f_1 \ldots, f_m\}$ *in* Pc $Sp(\{f_1 \ldots, f_m\} \vdash g)$, *is the minimal space over all possible* Pc *proofs of* g *from* $\{f_1 \ldots, f_m\}$.

The standard polynomial translation tr of CNF formulas into polynomials is defined as follows:

$$tr(x) = x \qquad tr(\neg x) = (1-x) \qquad tr(\bigvee_{i=1}^{n} l_i) = \prod_{i=1}^{n} tr(l_i)$$

When we refer to Pc refutations of some family of CNF formulas we always mean refutations of the family of polynomial translation of the CNF formulas.

Polynomial Calculus with Resolution (Pcr) [2] is a refutational system which extends Pc to polynomials in the ring $\mathbb{F}[x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n]$, where $\bar{x}_1, \ldots, \bar{x}_n$ are new formal variables. Pcr includes the axioms and rules of Pc plus a new set of logical axioms defined by

$$1 - x_i - \bar{x}_i \quad i \in [n]$$

to force $\bar{x}$ variables to have the opposite values of x variables. The standard polynomial translation tr of CNF formulas into polynomials in $\mathbb{F}[x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n]$ is the following:

$$tr(x) = x \qquad tr(\neg x) = \bar{x} \qquad tr(\bigvee_{i=1}^{n} l_i) = \prod_{i=1}^{n} tr(l_i)$$

When we refer to Pcr refutations of some family of CNF formulas we always mean refutations of the family of polynomial obtained by the translation above applied to the CNFs.

We extend to Pcr the definitions of proof, refutation, degree, size and length and space given for Pc. Observe that using the linear transformation $\bar{x} \mapsto 1 - x$, any Pcr refutation can be converted into a Pc refutation without increasing the degree. As noticed above such transformation could cause an exponential increase in size. When in the next we refer to space we omit to say if we are in Pc or Pcr.

1.3 Partial Assignments

Let V be a set of variables, we say that an application $\alpha : V \longrightarrow \{0, 1, \star\}$ is a *partial (boolean) assignment* over V. The *domain* of α is $dom(\alpha) = \alpha^{-1}(\{0, 1\})$. If $x \in dom(\alpha)$ we say that α is *assigning* a value to x.

We denote with $\emptyset$ the empty set and the partial assignment with empty domain, i.e. the assignment mapping each variable to $\star$. It will be clear from the context if we are talking about sets or assignments.

Given two partial assignments α and β such that $\alpha(x) = \beta(x)$ for each $x \in dom(\alpha) \cap dom(\beta)$. We define the partial assignment $\alpha \cup \beta$:

$$\alpha \cup \beta(x) = \begin{cases} \alpha(x) & \text{if } x \in dom(\alpha), \\ \beta(x) & \text{if } x \in dom(\beta), \\ \star & \text{otherwise.} \end{cases}$$

We say that a partial assignment β *extends* another partial assignment α if $dom(\alpha) \subseteq dom(\beta)$ and for all $x \in dom(\alpha)$, $\beta(x) = \alpha(x)$. We write $\alpha \subseteq \beta$.

Given a partial assignment α and $A \subseteq V$ we define the restriction $\alpha\restriction_A$:

$$\alpha\restriction_A (x) = \begin{cases} \alpha(x) & \text{if } x \in A, \\ \star & \text{otherwise.} \end{cases}$$

[1] This same definition can be given for Resolution substituting number of distinct monomials with number of clauses.

Given a clause C (or a polynomial P) we can substitute each variable x appearing in C (or P) with the value α is assigning to x, if $x \in \mathit{dom}(\alpha)$, or leave x untouched if $x \notin \mathit{dom}(\alpha)$. We denote the result of this operation with $\alpha(C)$.

If $x \notin \mathit{dom}(\alpha)$ we emphasize that $\alpha(x^2 - x) = x^2 - x \neq 0$.

DEFINITION 1.1 ($\models$ for formulas). *Let C be a boolean formula and α a partial assignment over the variables appearing in C. We say that α* models *C, $\alpha \models C$ if $\alpha(C) = 1$.*

Let $\mathcal{A}$ be a family of partial assignments, $\mathcal{A} \models C$ if for each $\alpha \in \mathcal{A}$ $\alpha \models C$.

If we are in PCR, i.e. we have a set of variables V and $\overline{V} = \{\overline{x} \mid x \in V\}$, and α is a partial assignment over V we can define clearly a partial assignment α^* over $V \cup \overline{V}$ extending α such that if $x \in \mathit{dom}(\alpha)$ then $\alpha^*(x + \overline{x} - 1) = 0$. Clearly is possible to do that defining

$$\alpha^*(\overline{x}) = \begin{cases} 1 - \alpha(x) & \text{if } x \in \mathit{dom}(\alpha), \\ \star & \text{otherwise.} \end{cases}$$

In the following we *always* suppose we are working with partial assignments over $V \cup \overline{V}$ of this sort. We do that referring explicitly only to the variables in V but every time we have a partial assignment α over V we implicitly are referring to the assignment α^*.

DEFINITION ($\models$ for polynomials). *Let V a set of variables and $\mathbb{F}[V]$ a ring of polynomials. Let I be a proper ideal in $\mathbb{F}[V]$ and p be a polynomial in $\mathbb{F}[V]$ and α a partial assignment. We say that α* models *p, $\alpha \models_I p$ if $\alpha(p) \in I$. If I it's clear from the context we'll omit the subscript.*

Let $\mathcal{A}$ be a family of partial assignments, $\mathcal{A} \models_I p$ if for each $\alpha \in \mathcal{A}$ $\alpha \models_I p$. Analogously if we have a family of polynomials.

DEFINITION. *Let V a set of variables and $\mathbb{F}[V]$ a ring of polynomials. Let I be a proper ideal in $\mathbb{F}[V]$ and α a partial assignment we'll say that α* respects *I if*

$$\forall p \in I \quad \alpha(p) \in I.$$

1.4 Random k-CNFs in Proof Complexity

It is well-known that in circuit complexity simple counting arguments show that a random function is hard to compute. In studying the complexity of a given proof system it is natural to ask what is the proof complexity of a tautology taken at random. However we do not have a definition of what is a random tautology. Still, in some cases, if we restrict our attention only to certain kinds of tautologies we can deduce information on their random behavior. An easy calculation shows that for a high enough constant, with high probability a random k-CNF formula is unsatisfiable. Let us introduce the definition of a random CNF.

Let n,m and k be positive natural numbers and let $X = \{x_1, \dots, x_n\}$ be a set of variables. Let $\mathcal{F}(n, m, k)$ be the set of all k-CNF formulas on X with exactly m clauses each defined on k literals on distinct variables. Alternatively, $\mathcal{F}(n, m, k)$ can be described as the result of repeating m times independently the following experiment: choose exactly k variables from X, and negate each variable independently with probability $1/2$. The ratio m/n is denoted by Δ, and is called the clause density. Usually, Δ is fixed to a constant and therefore is determined by n . We are interested in studying the asymptotic properties of a randomly chosen formula $F \sim \mathcal{F}(n, m, k)$ as n approaches to infinity. It is well known that when the clause density exceeds a certain constant θ_k that only depends on k, a randomly chosen formula is almost surely unsatisfiable. The question of the existence and value of a satisfiability threshold constant is an important open problem in combinatorics, and for more information on this subject, see e.g. [1]. In this work we are interested only in the region in which F is unsatisfiable with high probability, then we always consider fixed Δ, then $\mathcal{F}(n, m, k)$ can be made dependent only on n, Δ and k and denoted as $\mathcal{F}(n, \Delta, k)$.

The proof size of unsatisfiable random CNFs has been widely studied in proof complexity. Chvatal and Szemeredi in their seminal paper [25] showed that with high probability, any random k-CNF over n variables and Δn clauses for $\Delta = O(1)$, requires exponentially long Resolution proofs to be refuted. The importance of their work was in showing that in fact Resolution is a very weak proof system, because in some sense almost all unsatisfiable k-CNF require exponential size proofs to be refuted. Their lower bound was later improved and simplified by Beame and Pitassi in [10], and finally improved up to a ratio by Beame, Karp, Pitassi and Saks in [8], and reformulated in terms of a general technique based on the width by Ben-Sasson and Wigderson in [18].

The degree complexity in Pc/Pcr of refuting random k-CNF was established by Ben-Sasson and Impagliazzo in [13] for polynomials over fields with characteristic different form 2 and then for any field by Alekhnovich and Razborov in [3]. These works proved that with high probability refuting (unsatisfiable) random k-CNF in Pc/Pcr requires linear degree.

Concerning space it is known that in Resolution random k-CNF for $\Delta = O(1)$ requires space $\Omega(n)$. This was a result of Ben-Sasson and Galesi answering to a question posed in [2, 32]. Both [2, 34] posed the question of proving linear lower bounds for the space of refuting random k-CNF in Pc/Pcr. In this work we answer to this question, for the case $k \geq 4$.

1.5 State of the Art and Previous Work

The first work concerning space in proof complexity was that of Esteban and Toran [32], where, following suggestions of Haken and Kleine-Büning, they introduce the space measure for Resolution and gave space lower bounds for some examples of formulas, including the Pigeonhole principle. Alekhnovich et al. in [2] extended the definition of space to all proof systems and gave many examples of lower and upper bounds for different proof systems. In particular they introduce the space measure for Pc and Pcr and prove a lower bounds for the family of the compete tautologies CT_n and the PHP^m_n. In a subsequent work Ben-Sasson and Galesi [12] gave space lower bounds is Resolution for randomly chosen k-CNF formula, answering to open problems posed in [2, 32].

In another series of works [40, 41, 16] studied the relation between length and space in resolution. Recently strong trade-offs between length and space were obtained in [17] and very recently [7] gave trade-off results for formulas that require even superlinear space if length is optimal.

In the polynomial calculus (Pc) proof system introduced by Clegg et al. [26], clauses are interpreted as multilinear polynomial. The minimal refutation size of a formula in this proof system is related to the maximal degree of the polynomials appearing in the refutation [26, 37], and a number of strong lower bounds on proof size have been obtained by proving degree lower bounds for Pc/Pcr, for instance, [26, 23, 24, 44, 42, 13, 37, 3, 36, 35].

The study of space measure for Pc/Pcr started with the work [2]. It was immediately clear that the situation is very different from Resolution and that similar techniques to the ones developed for Resolution will not work. The simple examples of CT_n (Complete Tautologies) is already significative. CT_n is the CNF contradiction obtained by excluding any satisfying assignment for any possible clauses over n variables. While in Resolution [32] is straightforward to see that space n is necessary and sufficient, [2] proved that space $2/3n + 6$ is already sufficient in Pc/Pcr to refute CT_n, pointing out a significative difference between Resolution and Pcr. [2] develop a technique to obtain space lower bound for set of contradictory polynomials of high degree. They proved that if a set of unsatisfiable polynomials have minimal degree n, then the space to refute them in Pcr is at least $n/4$. Their result also applied to a polynomial encoding of the PHP^m_n (over multivalued variables) having only monomials of degree n.

Recently [34] gave the first space lower bounds for Pcr for class of formulas not having high initial degree. But their result holds for two encoding of variants of the Pigeonhole principle: the Bit Pigeonhole Principle, $BPHP^m_n$, and the XOR Pigeonhole Principle, $XPHP^m_n$, having respectively logarithmic and constant (4) initial degree. In these work anyway the analysis for the two lower bounds is very similar to the case of [2]. In our work we will clarify quite precisely why these lower bounds are obtained essentially the same way as the ones in [2].

1.6 Contributions and Innovations

The first contribution of this work is a new method for proving space lower bounds in PCR and PCR. In our approach is not anymore a structural property of the formula (to have high initial degree) that allow one to get lower bounds. But is a semantic argument, similar to that used in Resolution. It is known [5] that in Resolution “space is lower bounded by the width” and hence width lower bounds imply space lower bounds. This connection is obtained by a characterization of the width and the space through winning strategies of the adversary in a Spoiler-Duplicator k-existential game. We characterize precisely (Definition of k-extendibility) how long an adversary (Duplicator) can answer to a player (Spoiler) downloading polynomials into the memory without falling into a contradiction. This idea is the same as the one used in Resolution both in the characterization of the space by Asterias and Dalmau [5] or by Esteban et al. in [31] where they independently introduced the notion of k-dynamical satisfiability to study space lower bounds in Resolution or $Res(k)$.

The idea is that of finding for a CNF F to be refuted a combinatorial characterization of the winning strategies of the adversary. In the case of Resolution these winning strategies are families $\mathcal{F}$ of of bounded-domain partial assignments (bounded-domain partial homomorphism in the language of Atserias [4, 5]) which preserve two properties: (1) closure under sub-assignments; and (2) assignments in $\mathcal{F}$ with not too big domain, can be extended to bounded-domain new assignments, still in $\mathcal{F}$, which do not create a contradiction in the unsatisfiable CNF to refute.

In our case we devise a similar characterization. Instead of having families of bounded-domain assignments, we have families $\mathcal{F}$ of pairs formed by two elements: (1) partitions of subsets of the variables of bounded-cardinality (*pseudo-partitions* $\mathcal{Q}$); and (2) families of assignments which preserve a locality property over elements of the pseudo-partition ($\mathcal{Q}$-locally modifiable). As in the case of Resolution these families preserve a closure under restrictions; and (2) an extendibility property for ”small”-cardinality pseudo-partitions (condition 3 of k-extendibility).

We consider fixed a set V of variables, a ring of polynomials $\mathbb{F}[V]$, a contradictory set of polynomials ϕ included in $\mathbb{F}[V]$ and a proper ideal I in $\mathbb{F}[V]$.

DEFINITION (pseudo-partition). *A* pseudo-partition *over V is a collection of* disjoint *sets $\mathcal{Q} = \{Q_1, \dots, Q_t\}$, such that each $Q_i \subseteq V$. We use the notation $\cup\mathcal{Q}$ to denote the set of variables occurring in all elements of $\mathcal{Q}$.*

DEFINITION (transversal set). *Let $\mathcal{Q} = \{Q_1, \dots, Q_t\}$ be a pseudo-partition over V. We say that a set $A \subseteq V$ of variables is* transversal *to $\mathcal{Q}$ if $\forall Q_i \in \mathcal{Q}\ |Q_i \cap A| \leq 1$.*

We now introduce a class of *relevant* assignments with respect to pseudo-partitions. In the rest of the paper we are going to deal always with assignments from this class. First we need some notations.

DEFINITION. *Let $\mathcal{H}$ be family of assignments all with domain B, and let $A \subseteq B$. We define $\mathcal{H}\restriction_A = \{\alpha\restriction_A \mid \alpha \in \mathcal{H}\}$. If we have that $\mathcal{Q}$ is a pseudo-partition s.t. $\cup\mathcal{Q} \subseteq B$ we'll write $\mathcal{H}\restriction_{\mathcal{Q}}$ to indicate $\mathcal{H}\restriction_{\cup\mathcal{Q}}$.*

DEFINITION ($\mathcal{Q}$-locally-modifiable family of assignments). *Let $\mathcal{Q} = \{Q_1, \dots, Q_t\}$ be a pseudo-partition over V. A family of assignments $\mathcal{H}$ is* $\mathcal{Q}$-locally-modifiable *(we abbreviate by* $\mathcal{Q}$-lm*) with respect to I if and only if:*

1. $\forall \alpha \in \mathcal{H} \quad dom(\alpha) = \cup\mathcal{Q}$,
2. $\forall A \in \mathcal{Q}$, $\forall x \in A$, *there are $\alpha_0, \alpha_1 \in \mathcal{H}$ such that $\alpha_1(x) = 1$, $\alpha_0(x) = 0$ and $\alpha_0 \equiv \alpha_1$ over each Q_i different from A.*
3. *for each $B \subseteq \mathcal{Q}$, $\alpha \in \mathcal{H}\restriction_B$ and $\beta \in \mathcal{H}\restriction_{\mathcal{Q}\setminus B}$ imply that $\alpha \cup \beta \in \mathcal{H}$.*
4. $\forall A \in \mathcal{Q}$ $\mathcal{H}\restriction_A$ *is a family of partial assignments respecting I.*

DEFINITION (Admissible configurations). *Let V be a set of variables. An* admissible configuration *with respect to I is a pair $(\mathcal{Q}, \mathcal{H})$ such that: (1) $\mathcal{Q}$ is a pseudo-partition over V and (2) $\mathcal{H}$ is $\mathcal{Q}$-lm with respect to I.*

To compare admissible configurations we introduce a partial order.

Definition 1.2 ($\preceq$). *$(\mathcal{Q}, \mathcal{H}) \preceq (\mathcal{Q}', \mathcal{H}')$ if and only if (1) $\mathcal{Q} \subseteq \mathcal{Q}'$, and (2) $\mathcal{H}' \restriction_{\mathcal{Q}} = \mathcal{H}$.*

The next definition is our main definition and encloses the core of our lower bound proof in Theorem 3.1. This definition should be compared with definition of winning strategies for the Duplicator in the paper by Atserias and Dalmau [5] (Definition 2) or definition about winning strategies (Definition 28) in the paper by Esteban et al. [31].

Definition (k-extendibility / Winning strategies). *A non-empty family $\mathcal{F}$ of admissible configurations $(\mathcal{Q}, \mathcal{H})$ is k-*extendible *for ϕ with respect to I if and only if:*

1. $|\mathcal{Q}| \leq k$,
2. $\forall \mathcal{Q}' \subseteq \mathcal{Q}\ \ (\mathcal{Q}', \mathcal{H}\restriction_{\mathcal{Q}'}) \in \mathcal{F}$.
3. *if* $|\mathcal{Q}| < k$, *then* $\forall a \in \phi\ \ \exists (\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$ *such that:*
   (a) $(\mathcal{Q}, \mathcal{H}) \preceq (\mathcal{Q}', \mathcal{H}')$,
   (b) $\mathcal{H}' \models_I a$, *i.e* $\forall \alpha \in \mathcal{H}'\ \ \alpha(a) \in I$,
   (c) $|\mathcal{Q}'| \leq |\mathcal{Q}| + 1$.

Our main theorem places a precise link between finding a k-extendible family for an unsatisfiable CNF and the space needed to refute its translation as a set of polynomials in Pc or Pcr.

Theorem (Main Theorem). *Let ϕ be a contradictory set of polynomials in $\mathbb{F}[V]$ and I a proper ideal in that ring. Suppose that there exists a non-empty k-extendible family of admissible configurations $\mathcal{F}$ for ϕ with respect to I. Then the $Sp(\phi \vdash 1) \geq k/4$.*

The second contribution of the paper is the following: our Main Theorem allow us to re-obtain under unique combinatorial framework and technique all the known space lower bound for Pc/Pcr known so far. All these lower bounds are obtained by the Main Theorem showing concrete examples of extendible families of admissible configurations of the right dimension. It is worth to mention, in our opinion, that the way we obtain these lower bounds is using only a limited part of the strength of the definition of k-extendibility. We discuss this issues in more details in the next subsection. Here is sufficient to say that in the winning strategies we provide for the known cases, we use only a very specific type of pseudo-partitions: they are subsets of a fixed (full) partition of the variables.

Theorem ([34, 2]). *$Sp(CT_n \vdash 1) \geq n/4$, $Sp(PHP^m_n \vdash 1) \geq n/4$, $Sp(XPHP^m_n \vdash 1) \geq (n-1)/4$, $Sp(BPHP^m_n \vdash 1) \geq n/8$.*

As a third, and probably main contribution, we answer to the open problem [2, 34] of proving space lower bound for random k-CNF in Pc/Pcr. In this case we use our Main Theorem in its full potential. In building a $\Omega(n)$-exendible family of admissible configurations for a random k-CNF it is no longer sufficient to look only at *full* partitions of the variables, but we really have to deal with pseudo-partitions. One combinatorial ingredient of the construction of this family of configurations is the Matching Game of Ben-Sasson and Galesi [12] (simplified in [4]). But differently from their case, where they deal only with *matchings* in bipartite graphs, here we have to handle *multiple* matchings in bipartite graphs. Hence we extend the Matching Game to the case of multiple matchings.

Dealing with multiple matchings instead of matchings implies that to prove the required expansion property we need left degree at least 4 in the incidence bipartite graph associated to a random k-CNF. Our Theorem then hold for $k \geq 4$ (see also next subsection).

Theorem (space lower bound for random k-CNF). *Let $k \geq 4$ be any integer, $\epsilon > 0$ any constant and $\Delta \geq 1$. Let $F \sim \mathcal{F}(n, \Delta, k)$. There exists a constant $c = c_{k,\Delta,\epsilon}$, $c \geq 1$, such that with high probability $Sp(F \vdash 1) \geq \frac{n}{4c}$.*

Finally we prove an analogous result, and this is our fourth contribution, for the so-called Graph-Pigeonhole principle, which is a Pigeonhole principle defined over an expander bipartite graph with constant left degree. Also this theorem is proved through the Multiple Matching Game.

THEOREM (space lower bound for $\mathcal{G}$-PHP). *There exists a constant degree $d \geq 3$ bipartite graph $\mathcal{G} = (U \cup V, E)$ with $|U| = n+1$ and $|V| = n$, such that $Sp(\mathcal{G} - PHP \vdash 1) \geq \Omega(n/d)$.*

1.7 Main Ideas, Notions and Techniques

To explain our approach to the problem we start by describing a high level proof of the main theorem, which is common to all space lower bound proofs known so far, also in Resolution.

The proofs of the space lower bound theorem are based on the following idea: inductively for each memory configuration $\mathcal{C}_i$, find a bounded boolean function M_i (in the case of PC/PCR a 2-CNF such that its number of clauses $|M_i|$ is less than $2Sp(\mathcal{C}_i)$) which implies the memory configuration $\mathcal{C}_i$. Such proofs include always two key ingredients: (1) a Locality Lemma ([18, 32, 2, 12, 34]) that informally says that if a configuration $\mathcal{C}_i$ is satisfiable, then it will be satisfied by an M_i properly bounded in the space of $\mathcal{C}_i$; (2) a combinatorial property that allows to keep the memory configuration still satisfiable by a M_i when we download an axiom (both logical or belonging to the formula to refute) in the memory configuration and the space used is still not too much.

One important issue in the Locality Lemma for PCR of [2] and [34] is that the 2-CNF M_i should be formed by distinct variables. In the lower bound argument it is important to keep a sort of *independence* of the variables mentioned in the 2-CNF. In our approach this independence is realized through the elements of the pseudo-partition. We require that the variables in the 2-CNF all belongs to different elements of the pseudo-partition. Moreover we also require a *transversality* of the 2-CNF with respect to the pseudo-partition. That is we require that at most one variable for element of the partition can be hit in the 2-CNF. We consider the following two definitions.

DEFINITION (transversal set). *Let $\mathcal{Q}$ be a pseudo-partition over V. We say that a set $A \subseteq V$ of variables is* transversal *to $\mathcal{Q}$ if $\forall Q_i \in \mathcal{Q}\ |Q_i \cap A| \leq 1$.*

DEFINITION (transversal 2CNF). *Let M be a 2CNF in the variables V, we say that M is a* 2CNF transversal to a pseudo-partition $\mathcal{Q}$ *defined on V if $Var(M)$ is a transversal set to $\mathcal{Q}$ and moreover $\mathcal{Q}_M = \mathcal{Q}$, where $\mathcal{Q}_M = \{Q_i \in \mathcal{Q} \mid Q_i \cap Var(M) \neq \emptyset\}$, i.e. M hits once each element in $\mathcal{Q}$.*

If a 2-CNF M implies a memory configuration $\mathcal{C}$, this means that every assignment satisfying M also satisfy $\mathcal{C}$. In our case everything is filtered by the pseudo-partition (and by transversality with respect to it). We then model everything with respect to pseudo-partitions and in particular to admissible configurations.

Since the 2-CNF is transversal with respect to the pseudo-partition, also the assignments to it will be transversal to the pseudo-partition.

DEFINITION. *Let $\mathcal{Q}$ be a pseudo-partition over V, M a 2CNF and P a set of polynomials. We say that $M \models_I^{(\mathcal{Q},\mathcal{H})} P$ if and only if M is transversal to $\mathcal{Q}$, $\mathcal{H}$ is $\mathcal{Q}$-lm with respect to I and*

$$\forall \alpha \in \mathcal{H}\ (\alpha \models M \longrightarrow \alpha \models_I P).$$

LEMMA (Locality Lemma). *Let P be a set of polynomials, $\mathcal{Q}$ a pseudo-partition and $\mathcal{H}$ a $\mathcal{Q}$-lm family of assignments. Let M be a 2CNF transversal to $\mathcal{Q}$. If $M \models_I^{(\mathcal{Q},\mathcal{H})} P$, then there exists a pseudo-partition $\mathcal{Q}' \subseteq \mathcal{Q}$ and there exists a 2CNF M' transversal to $\mathcal{Q}'$ such that:*

- $M' \models_I^{(\mathcal{Q}',\mathcal{H}\restriction_{\mathcal{Q}'})} P$ *and*
- $|M'| \leq 2Sp(P)$[2].

[2] $Sp(P)$ is the number of distinct monomials appearing in P.

At this point, for the reader who knows the proofs of space lower bound theorem in [2, 34], should be clear that while in their case, what is really modelling the space measure is the number of distinct variables mentioned in the 2-CNF (divided by 2), in our case what is important is the number of elements in the pseudo-partitions (divided by 2). This means that while an adversary can find admissible configurations associated to the memory configurations where the number of elements in the pseudo-partitions are keep bounded, then the memory configuration will be still implied by a proper 2-CNF.

The inductive property on the number of configurations $\mathcal{C}_i$ we prove in our Main Theorem is the following:

There exists a pseudo-partition $\mathcal{Q}^i$, a 2CNF M_i transversal to $\mathcal{Q}^i$ and a family of assignments $\mathcal{H}_i$, $\mathcal{Q}^i$-lm, such that the following holds:

1. $M_i \models_I^{(\mathcal{Q}^i,\mathcal{H}_i,)} \mathcal{C}_i$;
2. $|M_i| \leq 2Sp(\mathcal{C}_i)$,
3. $(\mathcal{Q}^i, \mathcal{H}_i) \in \mathcal{F}$.

What is missed a this point is only to understand what is the combinatorial property that guarantees us to be able to download axioms in the memory but still keep the pseudo-partitions in the admissible configurations of bounded size. This property is the k-extendibility property. The third property guarantees us that if the pseudo-partition has cardinality strictly smaller than k then for each axiom we are still able to find another admissible configuration that satisfies that axiom through its associated set of locally modifiable assignments and has a most one element more in the pseudo-partition. Hence under the hypothesis that the space is $< k/4$, using $k-$extendibility, we can prove the inductive property which, on the other hand, implies the contradiction that the final configuration is satisfiable.

Our main theorem does not depend on the degree of the monomials in the set of polynomials to refute. This is an essential feature to get lower bounds for the space of refuting families of polynomials with small degree. But the theorem applies also to cases in which initials monomials are of high degree (as in the case of PHP^m_n), but giving in this form slightly worse results of what is currently known (see Section 4 for details). To get the best possible lower bound we tune our Theorem in order to apply it in his full strength also to such cases. First we introduce the notion of transversal monomial.

Definition (transversal monomial). *Let $\mathcal{Q}$ be a pseudo-partition over V, we say that a monomial m is* transversal *to $\mathcal{Q}$ if $Var(m)$ is a transversal set to $\mathcal{Q}$ and $deg(m) = |\mathcal{Q}|$, i.e. m is touching each element in $\mathcal{Q}$ once.*

Corollary. *Let $\phi = \psi \cup \mu$ a contradictory set of polynomials. Suppose that:*

1. *exists a non-empty k-extendible family of admissible configurations $\mathcal{F}$ over ψ with respect to the ideal $I = \{0\}$ and*
2. *every polynomial in μ is a* monomial *with degree at least k transversal to each pseudo-partition named in $\mathcal{F}$.*

Then, $Sp(\phi \vdash 1) \geq k/4$.

Let us discuss briefly the case of the PHP^m_n. The variables are x_{ij} for all $i \in [m]$ and $j \in [n]$. The axioms in PHP^m_n are: (1) $\neg x_{ij} \vee \neg x_{i'j}$ for all $i \neq i'$ and for all $j \in [n]$; (2) $x_{i1} \vee x_{i2} \vee \ldots \vee x_{in}$ for all $i \in [m]$.

As *full* partition we choose $\mathcal{P} = \{P_1, \ldots, P_n\}$, where $P_j = \{x_{ij} \mid i \in [m]\}$. We want to apply the previous Corollary. As ψ we choose all the logical axioms plus all the axioms in (1). As μ we choose all the axioms in (2).

We define $\mathcal{F}$ as the family of all the pairs $(\mathcal{Q}, \mathcal{H})$ such that $\mathcal{Q} \subseteq \mathcal{P}$ and $\mathcal{H}$ is the family of all the partial assignments of domain $\cup\mathcal{Q}$ satisfying the axioms in ψ with variables in $\cup\mathcal{Q}$, i.e. injective assignments of some pigeons into the holes.

Proposition. *The family $\mathcal{F}$ defined above is n-extendible for ψ.*

Let us discuss our approach with this example in hand. The main point in the argument [2] was that assuming (by contradiction) that the space used is small, then each time we download a high degree axiom in the memory, we are sure to find at least two *new* variables we can use to build a new 2-CNF that implies the new memory configuration. In our case, instead, we have that axioms of high degree are *transversal* to each element of the *full* partition of the variables. Since the space is (by contradiction) small, then we are able to find two elements of the partition where two find (in each of them) a new variable that we can use to form our new 2-CNF that implies the new memory configuration. As one can notice, this is easy in the case the PHP_n^m, since the axioms are of high degree and then, defining the proper partition, they hit all the element of the partition. Under the assumption the space is small, this guarantees us to find always new elements of the partitions where to pick new variables. As a matter of fact the same reasoning, that is *fixing a constant full partition of the variables*, is valid to get the lower bounds also for $BPHP_n^m$ and $XPHP_n^m$, which do not have high initial degree. In this sense these cases are an application of the method used by [2] for CT_n and PHP_n^m. This is since the particular syntactical properties of these encodings.

If, as in the case of random k-CNF or of the Graph-PHP, we do not have high degree initial axioms, then this reasoning is not valid anymore. Fixing a full partition at the begining is not useful anymore. We need another way of capturing the idea that "small space memory configurations can be satisfied by 2-CNF". The way we implement this is as follows:

- We use the (multiple) matching game to characterize at each step of the proofs what are the variables involved in a possible 2-CNF that implies the memory configuration. This is not new, since Ben-Sasson and Galesi in [12] where doing exactly this for Resolution. But instead of multiple matchings they had simple matchings and instead of 2-CNF they have assignments (i.e. 1-CNF).
- We handle, by the mechanism of the pseudo-partition and admissible configurations, the changing of the variables involved from one memory configuration to the other as modeled by the multiple Matching Game. Hence pseudo-partitions and the associated families of locally modifiable assignments might change in passing from one memory configuration to the sequent one.

In particular for random k-CNF and the Graph-PHP we dinamically maintain a property, the (r, s)-*double matching property* (see below), that allows us to identify dinamically for each memory configuration a set of initial clauses we are satisfying (in addition to the actual memory configuration). Moreover we can keep that set of initial clauses satisfied by using variables that we can consider "independent" and we capture this notion of independence by using pseudo-partitions and locally modifiable families of assignments.

Definition ((r, s)-double matching property). *Let $r \leq s$, $\mathcal{G} = (U \cup V, E)$ a bipartite graph and $A \subseteq U$ of size at most $r \leq s$ and $B \subseteq V \cap N_{\mathcal{G}}(A)$. We say that $(\mathcal{G}, A, B)$ has the* (r, s)-double matching property *if for every $C \subseteq U \setminus A$, if $|C| = s - |A|$ then there exists a 2-matching of C into $V \setminus B$.*

The idea behind this definition is to focus on the extension of an existing multiple matching, i.e. how in the Matching Game Duplicator can continue the game, hiding all the details on how Spoiler and Duplicator arrived to that configuration of the game but focusing only on the current configuration. In the previous definition the sets A and B play the role of the actual configuration of the game: we are not interested in how is constructed the multiple matching inside the sets A and B (and we inductively construct that multiple matching) to extend. The aim of that definition is to guarantee Duplicator that no matters how he and Spoiler arrived to a configuration, Duplicator can always make his move. Clearly this is a game very close to the Matching Game developed in [12, 4]. See Section 5 for all the details of how we succeed to use this game on a bipartite graph to obtain an $\Omega(n)$-extendible family for the Random k-CNF and Graph-PHP. The only detail of that construction we want to focus is the definition of expansion we need to dinamically maintain the (r, s)-double matching property. That is a stronger notion of expansion than the ones usually used (see [12, 4]) because we need to provide the existence of a *multiple* matchings (at least double matchings). The precise notion of expansion we use is the following.

Definition ((s, ϵ)-bipartite expansion). *Let $\mathcal{G} = (U \cup V, E)$ a bipartite graph. We say that $\mathcal{G}$ is an (s, ϵ)-*bipartite expander *if*

$$\forall A \subseteq U, \ |A| \leq s \longrightarrow |N_{\mathcal{G}}(A)| \geq (2 + \epsilon)|A|.$$

Due to this stronger requirement on the expansion we obtain our lower bound for random k-CNF for $k \geq 4$.

1.8 Future Directions and Research

We think that our characterization of the space in PC/PCR can open the way to a more precise characterization of the space, and we do not exclude the degree, of PC/PCR proofs in terms of 2-Player games like variants of the existential pebble games for Resolution like Ehrenfeucht-Fraïssé games. We find very attractive the idea that, as was done in Resolution by Atserias and Dalmau in [5], to find a precise combinatorial characterization of the degree and proving some relations between space and degree, similar to the one between width and space for Resolution. We think that our work and our notion of k-extendibility is a first step in this direction. So far there is no results that seems to exclude that "space might be lower bounded by degree" in PC/PCR. As was done for random k-CNF for DATALOG by Asterias [4], our game characterization of boolean reasoning with polynomials can suggest non-expressibility results in stronger logic appropriate to this kind of reasoning.

To work in this direction it might be useful to prove lower bounds for other classes of tautologies for which we known they require high degree. In particular we think to Tseitin Tautologies (Buss et al. in [23] proves that they require high degree) and Linear ordering principle on Graphs GOP_n (Galesi and Lauria [35] recently proved they require high degree in PC/PCR) or GT_n. We think that our characterization could work also for this case provided we have the right definition of graph underlying the principle.

Another issue concern the possibilities of using a similar characterization of the space to try prove space lower bounds in other more powerful systems. Nothing for instance is known about space complexity in Cutting Planes and Lovasz-Schriver proof systems. We think that also in this case our work can be a starting point to try to come up with similar ideas to prove space lower bounds in these systems.

Another natural open problem arising form our work is to study the variable space for PC/PCR for all the principles we prove space lower bounds for. We think that the same steps of [2] together with our approach based on transversality and pseudo-partitions one can hopefully prove quadratic lower bounds for variable space in all these cases.

Finally our work leaves open to study the case of the PC/PCR space for random 3-CNF. A solution to this question comes from the solution to the problem of showing the existence of a bipartite graph with left degree equals 3 having a sufficiently good expansion property (for instance with an expansion factor of $(2 + \epsilon)$ if one considers our definition of expansion).

2 Preliminary Definitions

We denote by x a Boolean variable. A literal l is either a variable or its negation. A clause $C = (l_1 \vee \ldots \vee l_k)$ is a disjunction of literals, a term $T = (l_1 \wedge \ldots \wedge l_k)$ is a conjunction of literals. We think of clauses and terms as sets, so that the ordering of the literals is irrelevant and no literals are repeated. We denote the empty clause, i.e., the clause containing no literals, by $\square$. A clause (term) containing at most k literals is called a k-clause (k-term). A CNF formula $F = C_1 \wedge \ldots \wedge C_m$ is a conjunction of clauses, and a DNF formula is a disjunction of terms. We will think of CNF and DNF formulas as sets of clauses and terms, respectively. A k-CNF formula is a CNF formula consisting of k-clauses, and a k-DNF formula consists of k-terms. A clause C is a clause over a set of variables V if the set of variables it mentions is a subset of V . We similarly define terms, CNFs and DNFs over V .

2.1 Algebraic proof systems and complexity measures

Polynomial Calculus (Pc) is a refutational system defined in [26], and based on the ring $\mathbb{F}[x_1, \ldots, x_n]$ of polynomials. Given $p \in \mathbb{F}[x_1, \ldots, x_n]$ we always consider equations of the form $p = 0$, and we simply denote them as p. The equations are intended to hold on $\{0, 1\}^n$ thus the system contains the following logical axioms:

$$x_i^2 - x_i, \quad i \in [n].$$

Moreover it has two rules. For any $\alpha, \beta \in \mathbb{F}$, p, q polynomials and variable x:

$$\frac{p \qquad q}{\alpha p + \beta q} \textit{ (Linear Combination)}, \qquad \frac{p}{xp} \textit{ (Multiplication)}.$$

A Pc proof of a polynomial g from a set of initial polynomials $f_1, \ldots, f_m$ (denoted by $f_1, \ldots, f_m \vdash g$) is a sequence of polynomials where each one is either an initial one, a boolean axiom, or it is obtained applying one of the rules to previously derived polynomials. A Pc refutation is a proof of the polynomial 1.

Pc is a complete proof system, in the sense that a polynomial g has a Pc proof from a set of polynomials E iff $g(\vec{x}) = 0$ for every $\vec{x} \in \{0, 1\}^n$ which is a common root of E. Moreover E has no common $\{0, 1\}$ solutions (we call E contradictory) iff $1 \in Span(E \cup \{x_i^2 - x_i\}_{i \in [n]})$. Completeness of Pc comes as a corollary of Hilbert's Nullstellensatz (see [28]) and from complete algorithms based on Gröebner bases [26].

We remark here that when we work in Polynomial Calculus, we implicitly assume that the polynomials $\{x_i^2 - x_i\}_{i \in [n]}$ are always included in the set of initial polynomials.

Given a Pc proof Π, the *degree* of Π, $deg(\Pi)$, is the maximal degree of a polynomial in the proof; the *size* of Π, $S(\Pi)$, is the number of monomials in the proof, the *length* of Π, $|\Pi|$, is the number of lines in the proof.

Following what done in [32, 2] for studying space complexity in Resolution and general sequential proof systems, we view a proof in Pc as similar to a non-deterministic Turing machine computation, with a working memory where all derivation steps are saved and a special read-only input tape from which the initials polynomials being refuted (the axioms) can be downloaded. Thus the length of a proof is essentially the time of the computation while the space measures the memory consumption. Following [2] we have:

Definition 2.1 (Memory Configuration). *A* Memory Configuration *in* Pc *is a set of polynomials. Given a set $F = \{f_1, \ldots, f_m\}$ of initials polynomials and a polynomial g, a* Pc *proof Π of $f_1, \ldots, f_m \vdash g$ can be view as sequence of memory configurations $\Pi = \{\mathcal{C}_0, \ldots, \mathcal{C}_l\}$ such that: $\mathcal{C}_0 = \emptyset$, $\mathcal{C}_l$ contains g and for all $i \in [l]$, $\mathcal{C}_i$ is obtained by $\mathcal{C}_{i-1}$ by one of the following three rules:*

Axiom Download *$\mathcal{C}_i = \mathcal{C}_{i-1} \cup \{p\}$, where p is some initial polynomial $f_j \in F$ or some boolean axiom.*

Inference Adding *$\mathcal{C}_i = \mathcal{C}_{i-1} \cup \{p\}$, where p si some polynomial inferred by using one of the rule of the calculus applied on polynomials occurring in $\mathcal{C}_{i-1}$*

Erasure *$\mathcal{C}_i = \mathcal{C}_{i-1} \setminus \{p\}$, for some $p \in \mathcal{C}_{i-1}$*

Following [2] we define the the *space* measure for Pc.

Definition 2.2 (Space Measure). *The* space *of a* Pc *memory configuration* $\mathcal{C}$, $Sp(\mathcal{C})$ *is the number of distinct monomials occurring in* $\mathcal{C}$. *The space of a* Pc *proof* Π, $Sp(\Pi)$), *is the maximal space of a memory configuration in* Π. *The space of proving* g *from* $\{f_1 \dots, f_m\}$ *in* Pc $Sp(\{f_1 \dots, f_m\} \vdash g)$, *is the minimal space over all possible* Pc *proofs of* g *from* $\{f_1 \dots, f_m\}$.

The standard polynomial translation *tr* of CNF formulas into polynomials is defined as follows:

$$tr(x) = x \qquad tr(\neg x) = (1 - x) \qquad tr(\bigvee_{i=1}^{n} l_i) = \prod_{i=1}^{n} tr(l_i)$$

When we refer to Pc refutations of some family of CNF formulas we always mean refutations of the family of polynomial translation of the CNF formulas.

Notice that the translation from CNF formulas into polynomials can lead to an exponential increment of the number of monomials with respect to the number of clauses in the formula translated. Therefore measuring the space in Pc as the number of distinct monomials leads to some substantial differences with Resolution. To overcome similar problems and try to make a treatment of complexity measure for proofs in Polynomial Calculus (in particular for the space), [2] introduced an algebraic system merging together Resolution and Polynomial Calculus.

Polynomial Calculus with Resolution (Pcr) [2] is a refutational system which extends Pc to polynomials in the ring $\mathbb{F}[x_1, \dots, x_n, \bar{x}_1, \dots, \bar{x}_n]$, where $\bar{x}_1, \dots, \bar{x}_n$ are new formal variables. Pcr includes the axioms and rules of Pc plus a new set of logical axioms defined by

$$1 - x_i - \bar{x}_i \quad i \in [n]$$

to force $\bar{x}$ variables to have the opposite values of x variables. The standard polynomial translation *tr* of CNF formulas into polynomials in $\mathbb{F}[x_1, \dots, x_n, \bar{x}_1, \dots, \bar{x}_n]$ is the following:

$$tr(x) = x \qquad tr(\neg x) = \bar{x} \qquad tr(\bigvee_{i=1}^{n} l_i) = \prod_{i=1}^{n} tr(l_i)$$

When we refer to Pcr refutations of some family of CNF formulas we always mean refutations of the family of polynomial obtained by the translation above applied to the CNFs.

We extend to Pcr the definitions of proof, refutation, degree, size and length and space given for Pc. Observe that using the linear transformation $\bar{x} \mapsto 1 - x$, any Pcr refutation can be converted into a Pc refutation without increasing the degree. As noticed above such transformation could cause an exponential increase in size. When in the next we refer to space we omit to say if we are in Pc or Pcr. All our results work for both Pcr and for Pc. In particular they also work for a general purpose calculus called *Functional Calculus*, introduced in[2, 34]. We omit details in this version of the work.

2.2 Partial Assignments

Let V be a set of variables, we say that an application $\alpha : V \longrightarrow \{0, 1, \star\}$ is a *partial (boolean) assignment* over V. The *domain* of α is $dom(\alpha) = \alpha^{-1}(\{0, 1\})$. If $x \in dom(\alpha)$ we say that α is *assigning* a value to x.

We denote with $\emptyset$ the empty set and the partial assignment with empty domain, i.e. the assignment mapping each variable to $\star$. It will be clear from the context if we are talking about sets or assignments.

Given two partial assignments α and β such that $\alpha(x) = \beta(x)$ for each $x \in dom(\alpha) \cap dom(\beta)$. We define the partial assignment $\alpha \cup \beta$:

$$\alpha \cup \beta(x) = \begin{cases} \alpha(x) & \text{if } x \in dom(\alpha), \\ \beta(x) & \text{if } x \in dom(\beta), \\ \star & \text{otherwise.} \end{cases}$$

We say that a partial assignment β *extends* another partial assignment α if $dom(\alpha) \subseteq dom(\beta)$ and for all $x \in dom(\alpha)$, $\beta(x) = \alpha(x)$. We write $\alpha \subseteq \beta$.

Given a partial assignment α and $A \subseteq V$ we define the restriction $\alpha\restriction_A$:

$$\alpha\restriction_A (x) = \begin{cases} \alpha(x) & \text{if } x \in A, \\ \star & \text{otherwise.} \end{cases}$$

Given a clause C (or a polynomial P) we can substitute each variable x appearing in C (or P) with the value α is assigning to x, if $x \in dom(\alpha)$, or leave x untouched if $x \notin dom(\alpha)$. We denote the result of this operation with $\alpha(C)$.

If $x \notin dom(\alpha)$ we emphasize that $\alpha(x^2 - x) = x^2 - x \neq 0$.

Definition 2.3 ($\models$ for formulas). *Let C be a boolean formula and α a partial assignment over the variables appearing in C. We say that α* models *C, $\alpha \models C$ if $\alpha(C) = 1$.*

Let $\mathcal{A}$ be a family of partial assignments, $\mathcal{A} \models C$ if for each $\alpha \in \mathcal{A}$ $\alpha \models C$.

If we are in Pcr, i.e. we have a set of variables V and $\overline{V} = \{\overline{x} \mid x \in V\}$, and α is a partial assignment over V we can define clearly a partial assignment α^* over $V \cup \overline{V}$ extending α such that if $x \in dom(\alpha)$ then $\alpha^*(x + \overline{x} - 1) = 0$. Clearly is possible to do that defining

$$\alpha^*(\overline{x}) = \begin{cases} 1 - \alpha(x) & \text{if } x \in dom(\alpha), \\ \star & \text{otherwise.} \end{cases}$$

In the following we *always* suppose we are working with partial assignments over $V \cup \overline{V}$ of this sort. We do that referring explicitly only to the variables in V but every time we have a partial assignment α over V we implicitly are referring to the assignment α^*.

Definition 2.4 ($\models$ for polynomials). *Let V a set of variables and $\mathbb{F}[V]$ a ring of polynomials. Let I be a proper ideal in $\mathbb{F}[V]$ and p be a polynomial in $\mathbb{F}[V]$ and α a partial assignment. We say that α* models *p, $\alpha \models_I p$ if $\alpha(p) \in I$. If I it's clear from the context we'll omit the subscript.*

Let $\mathcal{A}$ be a family of partial assignments, $\mathcal{A} \models_I p$ if for each $\alpha \in \mathcal{A}$ $\alpha \models_I p$. Analogously if we have a family of polynomials.

Observation 2.1. *Let V a set of variables and $\mathbb{F}[V]$ a ring of polynomials. Let I be a proper ideal in $\mathbb{F}[V]$, $P \subset \mathbb{F}[V]$ a set of polynomials and α a partial assignment. If $\alpha \models_I P$ then $\alpha \models_I Span(P)$. So in particular if P is a contradictory set of polynomials we have that for every partial assignment α and for every proper ideal I $\alpha \not\models_I P$.*

Definition 2.5. *Let V a set of variables and $\mathbb{F}[V]$ a ring of polynomials. Let I be a proper ideal in $\mathbb{F}[V]$ and α a partial assignment we'll say that α* respects *I if*

$$\forall p \in I \quad \alpha(p) \in I.$$

It's clear that partial boolean assignments respect the proper ideal $Span(\{x_i^2 - x_i, x_i + \overline{x_i} - 1\}_{i=1,\dots,n})$.

2.3 Graph properties and notations

Let $\mathcal{G} = (U \cup V, E)$ be a bipartite graph of left degree at most d.

Definition 2.6 (multiple matching). *Let $\mathcal{G} = (U \cup V, E)$ be a bipartite graph and $\pi \subseteq E$. Let $\pi(u) = \{v \in V \mid (u, v) \in \pi\}$. We say that π is a* matching *of $A \subseteq U$ if*

1. *$\pi \subseteq A \times V$,*
2. *for every u and u' in A $\pi(u)$ and $\pi(u')$ are disjoint non-empty sets.*

If for every $u \in A$ $|\pi(u)| = 2$ *we say that* π *is a* 2-matching *of* A*. If for every* $u \in A$ $|\pi(u)| \geq 2$ *we say that* π *is a* multiple matching *of* A*.*

Given a set $A \subseteq U$ of nodes, we define $\pi(A) = \{\pi(u) \mid u \in A\}$ and we denote by $\cup\pi(A)$ the set of variables in $\pi(A)$.

We use the following notion of expansion on bipartite graphs.

Definition 2.7 ((s, ϵ)-bipartite expansion). *Let* $\mathcal{G} = (U \cup V, E)$ *a bipartite graph. We say that* $\mathcal{G}$ *is an* (s, ϵ)-bipartite expander *if*

$$\forall A \subseteq U,\ |A| \leq s \longrightarrow |N_{\mathcal{G}}(A)| \geq (2 + \epsilon)|A|.$$

Notice that our expansion factor is $(2 + \epsilon)$ instead than the usual $(1 + \epsilon)$.

We use the standard notation $N_{\mathcal{G}}(A)$ to indicate the neighborhood of A in the graph $\mathcal{G}$. We use the following apllication of Hall's Theorem proved in [2] (Corollary 4.16).

Lemma 2.1 ([2]). *Let* $\mathcal{G} = (U \cup V, E)$ *be a bipartite graph. For every set* $A \subseteq U$*, if* $|N_{\mathcal{G}}(A)| \geq 2|A|$*, then there is a 2-matching of* U *in* V*.*

Notice that we have that if $A \subseteq U$ is the smallest set such that we cannot find a 2-matching of A in $\mathcal{G}$, then we have that $|N_{\mathcal{G}}(A)| < 2|A|$. Moreover if $\mathcal{G} = (U \cup V, E)$ is a (s, ϵ)-bipartite expander then, from the previous lemma, every subset of U of size at most s admit a 2-matching.

2.3.1 The Matching Game

If a bipartite graph $\mathcal{G} = (U \cup V, E)$ is such that $|V| > |U|$, then there is no perfect matching of U into V. Ben-Sasson and Galesi [12] introduced a 2-player game *the Matching Game* to prove this claim, using "limited space". The two players are a Prover and a Disprover. Prover tries to prove that there is no matching from U to V, and Disprover tries to prove that such a matching exists. Each player has k fingers. In each round of the game, Prover may place a finger over an uncovered node in U or remove a finger from a covered node in U. If Prover places a finger over node $u \in U$, Disprover must place her corresponding finger over an uncovered node in $N_{\mathcal{G}}(u)$. If Prover removes a finger from a node in U, Disprover must remove her corresponding finger from V. The game is over when Disprover is not able to answer to a move of the Prover. In that case, we say that Prover wins the game. If Disprover can make the game go on forever, we say that Disprover wins the game. Notice that at every non-final round, the fingers placed on U determine a partial matching of U into V. The goal of Disprover is to maintain a partial matching forever. The Matching Game was used by Atserias in [4] where he gave a more compact treatment of main properties. In Section 5 we extend the Matching Game to deal with *multiple matchings* in bipartite graphs instead that simply matchings. We refer to the notation developed in [4].

3 A Combinatorial Characterization for the Space Measure

In this section we consider fixed a set V of variables, a ring of polynomials $\mathbb{F}[V]$, a contradictory set of polynomials ϕ included in $\mathbb{F}[V]$ and a proper ideal I in $\mathbb{F}[V]$.

3.1 Preserving Axioms Satisfiability: k-*extendibility*

Let V be the set of variables appearing in some contradictory set of polynomials ϕ. We start introducing the main notions we use in the paper.

Definition 3.1 (pseudo-partition). *A* pseudo-partition *over* V *is a collection of* disjoint *sets* $\mathcal{Q} = \{Q_1, \ldots, Q_t\}$*, such that each* $Q_i \subseteq V$*. We use the notation* $\cup\mathcal{Q}$ *to denote the set of variables occurring in all elements of* $\mathcal{Q}$*.*

Definition 3.2 (transversal set). *Let* $\mathcal{Q} = \{Q_1, \ldots, Q_t\}$ *be a pseudo-partition over* V*. We say that a set* $A \subseteq V$ *of variables is* transversal *to* $\mathcal{Q}$ *if* $\forall Q_i \in \mathcal{Q}$ $|Q_i \cap A| \leq 1$*.*

We now introduce a class of *relevant* assignments with respect to pseudo-partitions. In the rest of the paper we are going to deal always with assignments from this class. First we need some notations.

DEFINITION 3.3. *Let $\mathcal{H}$ be family of assignments all with domain B, and let $A \subseteq B$. We define $\mathcal{H}\restriction_A = \{\alpha\restriction_A \mid \alpha \in \mathcal{H}\}$. If we have that $\mathcal{Q}$ is a pseudo-partition s.t. $\cup\mathcal{Q} \subseteq B$ we'll write $\mathcal{H}\restriction_{\mathcal{Q}}$ to indicate $\mathcal{H}\restriction_{\cup\mathcal{Q}}$*

DEFINITION 3.4 ($\mathcal{Q}$-locally-modifiable family of assignments). *Let $\mathcal{Q} = \{Q_1, \ldots, Q_t\}$ be a pseudo-partition over V. A family of assignments $\mathcal{H}$ is* $\mathcal{Q}$-locally-modifiable *(we abbreviate by* $\mathcal{Q}$-lm*) with respect to I if and only if:*

1. $\forall\alpha \in \mathcal{H} \quad dom(\alpha) = \cup\mathcal{Q}$,
2. *$\forall A \in \mathcal{Q}$, $\forall x \in A$, there are $\alpha_0, \alpha_1 \in \mathcal{H}$ such that $\alpha_1(x) = 1$, $\alpha_0(x) = 0$ and $\alpha_0 \equiv \alpha_1$ over each $\mathcal{Q}_i$ different from A.*
3. *for each $B \subseteq \mathcal{Q}$, $\alpha \in \mathcal{H}\restriction_B$ and $\beta \in \mathcal{H}\restriction_{\mathcal{Q}\setminus B}$ imply that $\alpha \cup \beta \in \mathcal{H}$.*
4. *$\forall A \in \mathcal{Q}$ $\mathcal{H}\restriction_A$ is a family of partial assignments respecting I.*

The main properties of a set $\mathcal{H}$ of $\mathcal{Q}$-lm assignments are made up to guarantee a sort of independence of the assignments in each element of the psuedo-partition $\mathcal{Q}$ (property 2). Property 3 of previous Definition is a closure property that captures the way one has to build inductively a family of locally modifiable family of assignments. An explanation of this property is in next Lemma 3.1, which is the main tool one has to use in the construction of a locally modifiable class of assignments for a set of contradictory polynomials.

We now give some examples of locally modifiable class of assignments to illustrate better our definition. Assume to have a pseudo-partition $\mathcal{Q} = \{\{x, y\}, \{z\}\}$ and let us describe $\mathcal{H}$ as a table. If $\mathcal{H}$ was the following class of assignments:

x	y	z
0	0	1
1	1	1
0	0	0

then one can see that property 2 of the definition is preserved but not property 3. It is important to stress the fact that property 2 is requiring that for the assignments that cover different values (0 and 1) to a given variable in an element of the pseudo-partition (x f.i. in this examples) are *exactly* the same on variables in other elements of the pseudo-partitions, but can be different on other variables in the same element of the pseudo-partition of the variable considered. We recall that in this example property 3 does not hold. In order to have property 3 in $\mathcal{H}$ it is sufficient to add another assignment

x	y	z
0	0	1
1	1	1
0	0	0
1	1	0

Let us illustrate main properties of our definitions. We first observe how $\mathcal{Q}$-locally-modifiable families of assignments include a large fraction of all the possible partial assignments with domain $\cup\mathcal{Q}$. We are interested in (partial) assignments with domain transversal to a pseudo-partition $\mathcal{Q}$, i.e. assigning at most one variable in each element of $\mathcal{Q}$. According to Definition 3.2 we call this kind of assignments *transversal* to a pseudo-partition $\mathcal{Q}$. One useful observation is the following:

OBSERVATION 3.1. *Let α be a partial assignment over V* transversal *to a pseudo-partition $\mathcal{Q}$ over V. Let $\mathcal{H}$ be a $\mathcal{Q}$-lm with respect to I. Then there exists a $\beta \in \mathcal{H}$ that extends α.*

Proof. Let $\delta \in \mathcal{H}$ be an assignment such that $A_\delta = \{x \in dom(\alpha) \mid \alpha(x) \neq \delta(x)\}$ has the minimal size among all the possible assignments in $\mathcal{H}$. If, by contradiction, $A_\delta \neq \emptyset$ we can find a variable $x \in A_\delta$. By property 2 of locally-modifiable family we can find an assignment $\beta \in \mathcal{H}$ such that $\alpha(x) = \beta(x)$. Consider now $B \in \mathcal{Q}$ such that $x \in B$. By property 3 of locally-modifiable family we have that $\delta' = \beta\restriction_B \cup \delta\restriction_{\mathcal{Q}\setminus\{B\}}$ is in $\mathcal{H}$ but $A_{\delta'}$ has size one less than A_δ. In fact we have that $\beta\restriction_B$ is not assigning a value to any of the variables of $dom(\alpha)$ except for x because α is transversal to $\mathcal{Q}$. For the minimality of A_δ this is absurd, so we must have that $A_\delta = \emptyset$. ☐

Another straightforward observation concerns with closure property of local modifiability with respect to inclusions of pseudo-partitions.

Observation 3.2. *If $\mathcal{Q}' \subseteq Q$ and $\mathcal{H}$ is $\mathcal{Q}$-lm, then $\mathcal{H}' = \mathcal{H}\restriction_{\mathcal{Q}'}$ is $\mathcal{Q}'$-lm.*

The next Lemma is one of the main tools provided by locally-modifiable assignments. It captures the way we can build a class of locally-modifiable assignments starting from assignments local to a set of variables. In the next, in all examples of formulas for which we need to build a family of locally modifiable assignments, we are going to use this lemma.

Lemma 3.1. *Let $\mathcal{Q}$ be a pseudo-partition over V and $\mathcal{H}$ be $\mathcal{Q}$-lm with respect to I. Let A be a subset of the variables V such that $\cup\mathcal{Q} \cap A = \emptyset$ and Σ be a family of assignments $\{A\}$-lm with respect to I. Then*

$$\mathcal{H}' = \{\alpha \mid \exists\beta \in \mathcal{H}\ \exists\gamma \in \Sigma\ \alpha = \beta \cup \gamma\}$$

is $\mathcal{Q} \cup \{A\}$-lm with respect to I.

Proof. If $\mathcal{Q} = \emptyset$, then $\mathcal{H}' = \Sigma$ and by hypothesis is $\{A\}$-lm. Assume then that $\mathcal{H} \neq \emptyset$. The only thing non trivial to prove is part (3) of the definition of family locally-modifiable. Let $\mathcal{Q}' = \mathcal{Q} \cup \{A\}$, $B' \subseteq \mathcal{Q}'$, $\alpha' \in \mathcal{H}'\restriction_{B'}$ and $\beta' \in \mathcal{H}'\restriction_{\mathcal{Q}'\setminus B'}$: we have to prove that $\alpha' \cup \beta' \in \mathcal{H}'$. Notice that if B is empty then the claim is immediate by definition of $\mathcal{H}'$ otherwise w.l.o.g. we can suppose that $B' = B \cup \{A\}$. So $\alpha' = \alpha \cup \gamma$ with $\alpha \in \mathcal{H}\restriction_B$ and $\gamma \in \Sigma$. We have now that $\mathcal{H}'\restriction_{\mathcal{Q}'\setminus B'} = \mathcal{H}\restriction_{\mathcal{Q}\setminus B}$. So $\alpha \cup \beta' \in \mathcal{H}$, because $\mathcal{H}$ is $\mathcal{Q}$-lm and we obtain that $\alpha' \cup \beta' = \gamma \cup \alpha \cup \beta' \in \mathcal{H}'$, by definition of $\mathcal{H}'$. ☐

Pseudo-partitions and locally modifiable families of assignments are combinatorial objects that will play central role in our main theorem.

Definition 3.5 (Admissible configurations). *Let V be a set of variables. An* admissible configuration *with respect to I is a pair $(\mathcal{Q}, \mathcal{H})$ such that: (1) $\mathcal{Q}$ is a pseudo-partition over V and (2) $\mathcal{H}$ is $\mathcal{Q}$-lm with respect to I.*

Notice that the configuration $(\emptyset, \{\emptyset\})$ is admissible. To compare admissible configurations we introduce a partial order.

Definition 3.6 ($\preceq$). *$(\mathcal{Q}, \mathcal{H}) \preceq (\mathcal{Q}', \mathcal{H}')$ if and only if (1) $\mathcal{Q} \subseteq \mathcal{Q}'$, and (2) $\mathcal{H}'\restriction_{\mathcal{Q}} = \mathcal{H}$.*

For a polynomial p the inclusion in the ideal, by a local modifiable class of assignments, is preserved by the partial order on admissible configurations.

Lemma 3.2. *Let P be a set of polynomials over variables V, and let $(\mathcal{Q}, \mathcal{H})$ and $(\mathcal{Q}', \mathcal{H}')$ be admissible configurations over V such that $(\mathcal{Q}, \mathcal{H}) \preceq (\mathcal{Q}', \mathcal{H}')$. If $\mathcal{H} \models_I P$, then $\mathcal{H}' \models_I P$*

Proof. Let $\alpha \in \mathcal{H}'$, we have that $\alpha = \beta \cup \gamma$ with $\beta \in \mathcal{H}$ and $\gamma \in \mathcal{H}'\restriction_{\mathcal{Q}'\setminus\mathcal{Q}}$ (and by Observation 3.2 this family is locally modifiable), because $(\mathcal{Q}, \mathcal{H}) \preceq (\mathcal{Q}', \mathcal{H}')$. By property (4) of local modifiability γ respects I. By hypothesis we have that $\beta(P) \in I$, so $\alpha(P) = \gamma(\beta(P)) \in I$. ☐

The next definition is our main definition and encloses the core of our lower bound proof in Theorem 3.1. This definition should be compared with definition of winning strategies for the Duplicator in the paper by Atserias and Dalmau [5] (Definition 2) or definition about winning strategies (Definition 28) in the paper by Esteban et al. [31].

Definition 3.7 (k-extendibility / Winning strategies). *A non-empty family $\mathcal{F}$ of admissible configurations $(\mathcal{Q},\mathcal{H})$ is k-*extendible *for ϕ with respect to I if and only if:*

1. $|\mathcal{Q}| \leq k$,
2. $\forall \mathcal{Q}' \subseteq \mathcal{Q}\ \ (\mathcal{Q}', \mathcal{H}\lceil_{\mathcal{Q}'}) \in \mathcal{F}$.
3. *if* $|\mathcal{Q}| < k$, *then* $\forall a \in \phi\ \ \exists (\mathcal{Q}',\mathcal{H}') \in \mathcal{F}$ *such that:*
   (a) $(\mathcal{Q},\mathcal{H}) \preceq (\mathcal{Q}',\mathcal{H}')$,
   (b) $\mathcal{H}' \models_I a$, *i.e* $\forall \alpha \in \mathcal{H}'\ \alpha(a) \in I$,
   (c) $|\mathcal{Q}'| \leq |\mathcal{Q}| + 1$.

We observe that if in the property (2) of the previous definition we choose $\mathcal{Q}' = \emptyset$ we have, as a special case, that $(\emptyset, \{\emptyset\}) \in \mathcal{F}$.

Moreover we notice that in property (3) when $a \in I$ then it is sufficient to have $\mathcal{Q}' = \mathcal{Q}$ and $\mathcal{H}' = \mathcal{H}$. The interesting case is when $a \notin I$ and $\mathcal{H} \not\models_I a$. In this case we must have that $|\mathcal{Q}'| = |\mathcal{Q}| + 1$ since otherwise, by the partial order we have that $\mathcal{Q}' = \mathcal{Q}$ and $\mathcal{H} = \mathcal{H}'$ so $\mathcal{H} \models_i a$. This is a key point in the proof of main theorem.

3.2 Locality Lemma for 2-CNFs over Admissible Configurations

Let us first introduce the main notions for the Locality Lemma. Given a formula ψ in the variables V and a pseudo-partition $\mathcal{Q}$ over V, we denote by $\mathcal{Q}_\psi$ the elements of the partition $\mathcal{Q}$ hit by $Var(\psi)$, i.e. $\mathcal{Q}_\psi = \{Q_i \in \mathcal{Q} \mid Q_i \cap Var(\psi) \neq \emptyset\}$. In particular we'll use this notation for formulas M that are 2CNFs.

According to Definition 3.2 we give the definition of *transversal* 2CNF.

Definition 3.8 (transversal 2CNF). *Let M be a 2CNF in the variables V, we say that M is a* 2CNF transversal *to a pseudo-partition $\mathcal{Q}$ defined on V if $Var(M)$ is a transversal set to $\mathcal{Q}$ and moreover $\mathcal{Q}_M = \mathcal{Q}$.*

We required that $\mathcal{Q}_M = \mathcal{Q}$ to simplify some following notations and proofs. We use the notation $|M|$ for the number of clauses in M.

Let us consider the following symbol $\models_I^{(\mathcal{Q},\mathcal{H})}$ defined only if $(\mathcal{Q},\mathcal{H})$ is an admissible configuration.

Definition 3.9. *Let $\mathcal{Q}$ be a pseudo-partition over V, M a 2CNF and P a set of polynomials. We say that $M \models_I^{(\mathcal{Q},\mathcal{H})} P$ if and only if M is transversal to $\mathcal{Q}$, $\mathcal{H}$ is $\mathcal{Q}$-lm with respect to I and*

$$\forall \alpha \in \mathcal{H}\ (\alpha \models M \longrightarrow \alpha \models_I P).$$

Lemma 3.3 (Locality Lemma). *Let P be a set of polynomials, $\mathcal{Q}$ a pseudo-partition and $\mathcal{H}$ a $\mathcal{Q}$-lm family of assignments. Let M be a 2CNF transversal to $\mathcal{Q}$. If $M \models_I^{(\mathcal{Q},\mathcal{H})} P$, then there exists a pseudo-partition $\mathcal{Q}' \subseteq \mathcal{Q}$ and there exists a 2CNF M' transversal to $\mathcal{Q}'$ such that:*

- $M' \models_I^{(\mathcal{Q}',\mathcal{H}\lceil_{\mathcal{Q}'})} P$ *and*
- $|M'| \leq 2Sp(P)$[3].

Proof. Let us consider the bipartite graph $\mathcal{G} = (U \cup V, E)$, where U is the set of distinct monomials appearing in P, V is the set of clauses appearing in M and we have that $(m, C) \in E$ if and only if $Var(m) \cap \mathcal{Q}_C \neq \emptyset$. Let us choose a maximal set $\Gamma \subseteq U$ such that $|N_{\mathcal{G}}(\Gamma)| \leq 2|\Gamma|$. By Lemma 2.1, we have that $\overline{\Gamma} = V \setminus \Gamma$ admit a 2-matching into $U \setminus N_{\mathcal{G}}(\Gamma)$. Let $\pi = \{(m_i, C_{i,1}), (m_i, C_{i,2})\}$ be that 2-matching.

Let us see now how to construct the 2CNF M'. For each edge in π we choose a variable $x_{i,j}$, where $j = 1, 2$, such that $x_{i,j} \in Var(m_i) \cap Q_{C_{1,j}}$ and we consider $sat_{i,j} \in \{0, 1\}$ such that every assignment that

[3] $Sp(P)$ is the number of distinct monomials appearing in P.

maps $x_{i,j}$ into $sat_{i,j}$ maps the monomial m_i to 0. Let us choose also the variables $y_{i,j} \in Var(C_{i,j}) \setminus \mathcal{Q}_{x_{i,j}}$: we'll use later these variables. Let

$$M' = N_{\mathcal{G}}(\Gamma) \cup \{(x_{i,1}^{sat_{i,1}} \vee x_{i,2}^{sat_{i,2}}) \mid i \in \overline{\Gamma}\}.$$

Let $\mathcal{Q}'$ be $\mathcal{Q}\restriction_{Var(M')}$. Clearly we have that $\mathcal{Q}'$ is a pseudo-partition and that M' is a 2CNF transversal to $\mathcal{Q}'$. We set $\mathcal{H}' = \mathcal{H}\restriction_{Q'}$. We have that $|M'| \leq 2Sp(P)$, indeed:

$$|M'| = |N_{\mathcal{G}}(\Gamma)| + |\overline{\Gamma}| \leq 2|\Gamma| + |\overline{\Gamma}| \leq 2(|\Gamma| + |\overline{\Gamma}|) = 2Sp(P).$$

The only part of the lemma remaining to prove is that $M' \models_I^{(\mathcal{Q}',\mathcal{H}')} P$. So let $\alpha \in \mathcal{H}'$ such that $\alpha \models M'$: we have to prove that $\alpha \models_I P$. The strategy to do this is to find a $\beta \in \mathcal{H}$ st

- $\beta \models_I M$,
- $\beta(m) = \alpha(m)$ for each monomial m appearing in P.

Before going into the construction of β, let us suppose we have such a β and see how we conclude from that. We have by hypothesis that $M \models_I^{(\mathcal{H},\mathcal{Q})} P$, so, from the first property we have that $\beta \models P$. By the second property we have that α and β are coincident on the monomials in P so we must have that $\alpha \models P$.

Let us go into the construction of β. We have that $\mathcal{H}\restriction_{\mathcal{Q}\setminus\mathcal{Q}'}$ is $(\mathcal{Q} \setminus \mathcal{Q}')$-lm and we have that exists a γ transversal to $\mathcal{Q} \setminus \mathcal{Q}'$ such that $\alpha \cup \gamma \models M$ (because M is transversal). Then, by Observation 3.1, we have that exists $\tilde{\gamma} \in \mathcal{H}\restriction_{\mathcal{Q}\setminus\mathcal{Q}'}$ such that $\tilde{\gamma} \supseteq \gamma$. If we set $\beta = \alpha \cup \tilde{\gamma}$ we have by definition that $\beta \in \mathcal{H}$ and clearly $\beta \models M$.

Let us prove that $\beta(m) = \alpha(m)$ for each monomial m appearing in P. For each m_i with $i \in \overline{\Gamma}$ we have that $\alpha(m) = 0$, then clearly $\beta(m_i) = 0$ (because $\beta \supseteq \alpha$). Let us consider now the case $m \in \Gamma$. If $\alpha(m) \neq \beta(m)$ we must have that β is assigning some variable from m, so we must have that exists $y_{i,j}$ such that $\mathcal{Q}_{y_{i,j}} \cap Var(m) \neq \emptyset$. This is absurd because we have that $\mathcal{Q}_{y_{i,j}} \subseteq Q_{C_{i,j}}$, and then $Q_{C_{i,j}} \cap Var(m) \neq \emptyset$, so we should have the edge $(m, C_{i,j})$ in $\mathcal{G}$, but by construction $m \in \Gamma$ and $C_{i,j} \notin N_{\mathcal{G}}(\Gamma)$. □

3.3 Space Lower Bound Theorem

Let us consider the following straightforward observation.

OBSERVATION 3.3. *Let $\mathcal{Q}$ be a pseudo-partition over V, $\mathcal{H}$ $\mathcal{Q}$-locally modifiable, M a 2CNF transversal to $\mathcal{Q}$, P a set of polynomials, and a a polynomial. If $M \models_I^{(\mathcal{Q},\mathcal{H})} P$ and $\mathcal{H} \models_I a$, then $M \models_I^{(\mathcal{Q},\mathcal{H})} P \cup \{a\}$.*

THEOREM 3.1 (Main Theorem). *Let ϕ be a contradictory set of polynomials in $\mathbb{F}[V]$ and I a proper ideal in that ring. Suppose that there exists a non-empty k-extendible family of admissible configurations $\mathcal{F}$ for ϕ with respect to I. Then the $Sp(\phi \vdash 1) \geq k/4$.*

Proof. Let $\Pi = \mathcal{C}_1, \ldots, \mathcal{C}_s$ be a refutation of ϕ in PCR. Assume by contradiction that $Sp(\Pi) < k/4$. We prove by induction on i that there exists a pseudo-partition $\mathcal{Q}^i$, a 2CNF M_i transversal[4] to $\mathcal{Q}^i$ and a family of assignments $\mathcal{H}_i$, $\mathcal{Q}^i$-lm such that the following holds:

1. $M_i \models_I^{(\mathcal{Q}^i,\mathcal{H}_i,)} \mathcal{C}_i$,
2. $|M_i| \leq 2Sp(\mathcal{C}_i)$,
3. $(\mathcal{Q}^i, \mathcal{H}_i) \in \mathcal{F}$.

[4]Remember that by definitions of transversal 2CNF this means that $\mathcal{Q}^i = \mathcal{Q}_{M_i}$.

Before proving the statement by induction on i, we show that the inductive hypothesis leads to a contradiction. The inductiove property (1) implies that every memory configuration can be mapped into I (if $M_i = \emptyset$ we must have that $\mathcal{C}_i \subseteq I$). This is impossible since the last one contains the polynomial 1 so we must have that $1 \in I$ but by hypothesis I is proper.

For the base case we set: $\mathcal{Q}_0 = \emptyset$, $M_0 = \emptyset$ and $\mathcal{H}_0 = \{\emptyset\}$. (1) follows since for an assignment satisfy a memory configuration is an universal statement abount the polynomials in that configuration. So the empty assignment satisfy the empty memory configuration. (2) follows since $|M_0| = Sp(\mathcal{C}_0) = 0$; (3) follows since by definition $(\emptyset, \{\emptyset\}) \in \mathcal{F}$.

For the inductive case we distinguish three cases according with the rules to modify the memory:

In the ERASURE case, we apply the Locality Lemma with $M = M_i$, $\mathcal{Q} = \mathcal{Q}^i$, $\mathcal{H} = \mathcal{H}_i$ and $P = \mathcal{C}_{i+1}$ to get $\mathcal{Q}'$ and M' satisfying the conclusions of the Lemma. We set $M_{i+1} = M'$, $\mathcal{Q}^{i+1} = \mathcal{Q}'$, $\mathcal{H}_{i+1} = \mathcal{H}_i\restriction_{\mathcal{Q}'}$, (1) then follows by (1) of the Locality Lemma. (2) follows from (2) of the Locality Lemma. (3) follows from the property 2 of the definition of k-extendibility.

In the INFERENCE ADDING case, we set $\mathcal{Q}^{i+1} = \mathcal{Q}^i$, $M_{i+1} = M_i$ and $\mathcal{H}_{i+1} = \mathcal{H}_i$. The result follows since $\mathcal{C}_{i+1}$ is a subset of the ideal generated by $\mathcal{C}_i$ and $\mathcal{H}_i \models_I Span(\mathcal{C}_i)$ (by Observation 2.1). Clearly we have $Sp(\mathcal{C}_i) < Sp(\mathcal{C}_{i+1})$.

In the AXIOM DOWNLOAD case, i.e. $\mathcal{C}_{i+1} = \mathcal{C}_i \cup \{a\}$ with $a \in \phi$ we distinguish two cases.

If $\mathcal{H}_i \models_I a$, then we set $\mathcal{Q}^{i+1} = \mathcal{Q}^i$, $M_{i+1} = M_i$ and $\mathcal{H}_{i+1} = \mathcal{H}_i$. (1) follows by Observation 3.3, (2) since $Sp(\mathcal{C}_i) < Sp(\mathcal{C}_{i+1})$ and (3) immediately from the setting.

Assume now that $\mathcal{H}_i \not\models_I a$. We claim that $|\mathcal{Q}^i| < k - 1$. We know that $|M_i| \leq 2Sp(\mathcal{C}_i)$, and, by the assumption, that $Sp(\mathcal{C}_i) < k/4 - 1$ (the -1 is since at step $i + 1$ we are downloading and axiom more into the memory). Since M_i is transversal to $\mathcal{Q}^i$, then $|\mathcal{Q}^i| = 2|M_i|$ and hence $|\mathcal{Q}^i| < k - 4 < k - 1$. By the claim, we can use the extendibility property of $\mathcal{F}$ on $(\mathcal{Q}^i, \mathcal{H}_i)$ and a, to conclude that there exist a $(\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$, such that: (a) $(\mathcal{Q}^i, \mathcal{H}_i) \preceq (\mathcal{Q}', \mathcal{H}')$, (b) $\mathcal{H}' \models_I a$, and (c) $|\mathcal{Q}'| \leq |\mathcal{Q}^i| + 1$. We distinguish two cases according to whether $a \in I$ or not. If $a \in I$, then we set $M_{i+1} = M_i$, $\mathcal{Q}^{i+1} = \mathcal{Q}'$ and $\mathcal{H}_{i+1} = \mathcal{H}'$. (1) then follows by Observation 3.3 since $\mathcal{H}' \models_I a$. (2) straightforwardly by $Sp(\mathcal{C}_i) < Sp(\mathcal{C}_{i+1})$ and $M_{i+1} = M_i$, (3) by the fact that $(\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$ by extendibility.

The interesting case is when $a \notin I$ and by the remark after the definition of k-extendibility this means $|\mathcal{Q}'| = |\mathcal{Q}^i| + 1$. We reason as follows: First among all pairs $(\mathcal{Q}', \mathcal{H}')$ in $\mathcal{F}$ satisfying the properties (a) and (b) of extendibility and such that $|\mathcal{Q}'| = |\mathcal{Q}^i| + 1$ we choose one $(\tilde{\mathcal{Q}}, \tilde{\mathcal{H}})$ which maximizes the number of initial polynomials in ϕ mapped by $\tilde{\mathcal{H}}$ in I. Let us call $\tilde{\phi}$ this set of polynomials such that $\tilde{\mathcal{H}} \models_I \tilde{\phi}$. Clearly $a \in \tilde{\phi}$. By Observation 2.1 we can't have that $\tilde{\mathcal{H}} \models_I \phi$, because ϕ is contradictory and I is proper.

So we must have that $\tilde{\phi}$ is a proper subset of ϕ, then there exists a polynomial $b \in \phi$ such that $\tilde{\mathcal{H}} \not\models b$. Observe that $|\mathcal{Q}'| < k$ (since $|\mathcal{Q}^i| < k - 1$), hence we can apply the extendibility property for a second time on $(\mathcal{Q}', \mathcal{H}')$ and b. We then have a pair $(\mathcal{Q}'', \mathcal{H}'') \in \mathcal{F}$ such that (a.2) $(\mathcal{Q}', \mathcal{H}') \preceq (\mathcal{Q}'', \mathcal{H}'')$, (b.2) $\mathcal{H}'' \models_I b$ and (c.2) $|\mathcal{Q}''| \leq |\mathcal{Q}'| + 1$. First we claim that:

Claim 3.1. $|\mathcal{Q}''| = |\mathcal{Q}'| + 1$

Proof. Assume by contradiction that $|\mathcal{Q}''| = |\mathcal{Q}'|$, then $|\mathcal{Q}''| = |\mathcal{Q}^i| + 1$. But then, since $b \notin \tilde{\phi}$, $(\tilde{\mathcal{Q}}, \tilde{\mathcal{H}})$ would not be the configuration in $\mathcal{F}$ which satisfies the maximal number of initial polynomials in ϕ. □

Now we are ready to set our new parameters. $\mathcal{Q}^{i+1} = \mathcal{Q}''$, $\mathcal{H}_{i+1} = \mathcal{H}''$. To form M_{i+1} we choose two *new* variables x and y. x belongs to the new element (wrt $\mathcal{Q}^i$) of the pseudo-partition $\mathcal{Q}'$ (we are in this case) and y to the new element of the pseudo-partition $\mathcal{Q}''$ (wrt to $\mathcal{Q}'$), which is guaranteed by the previous Claim. Hence $M_{i+1} = M_i \wedge (x \vee y)$.

Property $M_{i+1} \models^*_{(\mathcal{Q}^{i+1}, \mathcal{H}_{i+1})} \mathcal{C}_i \cup \{a\}$, holds since, by Lemma 3.2, $\mathcal{H}'' \models \tilde{\phi}$ and hence, since $a \in \tilde{\phi}$, $\mathcal{H}_{i+1} \models_I a$. Property (2) follows since $|M_{i+1}| = |M_i| + 1 \leq 2Sp(\mathcal{C}_i) + 2 = 2Sp(\mathcal{C}_{i+1})$. (3) follows since $(\mathcal{Q}'', \mathcal{H}'')$ is an admissible configuration in $\mathcal{F}$. □

4 Re-obtaining known space lower bounds: an unified framework

In this section we show how to re-obtain the known results given for CT_n and PHP_n^m by Aleknovich et al. in [2] and the results for $BPHP_n^m$ and $XOR - PHP_n^m$ given by Filmus et al. in [34]. Currently these are the only space known lower bounds for algebraic systems. As we see, all these cases fall into a very easy application of our main theorem. The main point in all these examples is that we do not need real pseudo-partitions that are changing passing from one memory configuration to the next one. In all these cases pseudo-paritions are $f$ull partitions of the variables and they remain the same along all the proofs, without changing. Hence in these cases characterizing the k-extendible family of assignment sit will be almost immediate since i.e. referring to a part ions that always remain fixed along any refutations. As we will see in the next section for random k-CNF or the Graph-PHP_n this will be not anymore the case.

Our main theorem does not depend on the degree of the monomials in the set of polynomials to refute. This is an essential feature to get lower bounds for the space of refuting families of polynomials with small degree. Our Theorem applies also to cases in which initials monomials are of high degree (as in the case of the pigeon hole principle or the case of complete contradictions), but giving slightly worse results of what is currently known (see Section 4 for details). To get the best possible lower bound we tune our Theorem in order to apply it in his full strength also to such cases.

According to Definition 3.2 we introduce the notion of transversal monomial.

Definition 4.1 (transversal monomial). *Let $\mathcal{Q}$ be a pseudo-partition over V, we say that a monomial m is* transversal *to $\mathcal{Q}$ if $Var(m)$ is a transversal set to $\mathcal{Q}$ and $deg(m) = |\mathcal{Q}|$, i.e. m is touching each element in $\mathcal{Q}$ once.*

Corollary 4.1. *Let $\phi = \psi \cup \mu$ a contradictory set of polynomials. Suppose that:*

1. *exists a non-empty k-extendible family of admissible configurations $\mathcal{F}$ over ψ with respect to the ideal $I = \{0\}$ and*
2. *every polynomial in μ is a* monomial *with degree at least k transversal to each pseudo-partition named in $\mathcal{F}$.*

Then, $Sp(\phi \vdash 1) \geq k/4$.

Proof. The proof is the same of the previous theorem. We use the very same notations used before. The only part we have to show is how to prove the induction properties when we download an axiom from μ. So let $\mathcal{C}_{i+1} = \mathcal{C}_i \cup \{m\}$ with $m \in \mu$.

We already noticed that $|Q^i| < k-1$, then we have that exists a variable x in $Var(m)$ not in $\cup\mathcal{Q}^i$. (2) implies we can find $(\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$ such that $(\mathcal{Q}^i, \mathcal{H}^i) \preceq (\mathcal{Q}', \mathcal{H}')$, $\mathcal{H}' \models_I x^2 - x$ and $|\mathcal{Q}'| \leq |\mathcal{Q}| + 1$. Since $x \notin \cup\mathcal{Q}^i$ we must have that $|\mathcal{Q}'| = |\mathcal{Q}| + 1$. We have that $|\mathcal{Q}'| < k$ then again we can find a variable $y \in Var(m)$ but not in $\cup\mathcal{Q}'$. We use the k-extendibility again obtaining the pair $(\mathcal{Q}'', \mathcal{H}'') \in \mathcal{F}$. Reasoning exactly as above we can find a variable $y \in Var(m)$ and $y \notin \cup\mathcal{Q}'$ so we obtain $\mathcal{Q}''$ and $\mathcal{H}''$ exactly as above we have that $|\mathcal{Q}''| = |\mathcal{Q}'| + 1$.

We set $\mathcal{Q}^{i+1} = \mathcal{Q}''$, $\mathcal{H}_{i+1} = \mathcal{H}''$ and $M_{i+1} = M_i \wedge (x^{sat_x(m)} \vee y^{sat_y(m)})$, where $sat_x(m)$ e $sat_y(m)$ are the values we can give to x or y respectively to set m to zero. □

4.1 CT_n

CT_n is a contradiction in the variables $x_1, \ldots, x_n$. We recall that the axioms of CT_n are *all* the possible clauses in the n variables of width n. We choose the full partition $\mathcal{P} = \{P_1, \ldots, P_n\}$, where $P_i = \{x_i\}$. This is a trivial special case of the Corollary 4.1. Following the notations of that Corollary we set as ψ all the logical axioms and $\mu = CT_n$ and $I = \{0\}$. Then we choose as a family $\mathcal{F}$ the pairs $(\mathcal{Q}, \mathcal{H})$ where $\mathcal{Q} \subseteq \mathcal{P}$ and $\mathcal{H}$ *all* the possible partial assignments with domain $\cup\mathcal{Q}$.

Proposition 4.1. *The family $\mathcal{F}$ defined above is n-extendible for ψ.*

Proof. The restriction part is clear. The extension part goes as follows: let $a \in \psi$ and $(\mathcal{Q}, \mathcal{H}) \in \mathcal{F}$. We have that $Var(a) = \{x_i\}$. If $x_i \in \cup\mathcal{Q}$ we clearly have that $\mathcal{H} \models_I a$. If $x_i \notin \cup\mathcal{Q}$ we set $\mathcal{Q}' = \mathcal{Q} \cup \{P_i\}$ and $\mathcal{H}'$ all the partial assignments with domain $\cup\mathcal{Q}'$. By applying Lemma 3.1 have that $\mathcal{H}'$ is $\mathcal{Q}'$-lm and we clearly have that $(\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$, $(\mathcal{Q}', \mathcal{H}') \succeq (\mathcal{Q}, \mathcal{H})$, $\mathcal{H}' \models a$ and $|\mathcal{Q}'| = |\mathcal{Q}| + 1$. ☐

THEOREM 4.2 ([2]). *$Sp(CT_n \vdash 1) \geq n/4$.*

Proof. We proved that the family $\mathcal{F}$ is n-extendible for ψ and it's easy to see that $\mu = CT_n$ satisfy the requests of the Corollary 4.1. The result follows. ☐

We observe that for CT_n is possible to apply directly the Main Result (Theorem 3.1) but, it's easy to see, in that manner we obtain as lower bound $n/8$. We wrote the Corollary 4.1 to re-obtain exactly the known lower bound.

4.2 PHP_n^m

The variables are x_{ij} for all $i \in [m]$ and $j \in [n]$. The axioms in PHP_n^m are:

1. $\neg x_{ij} \vee \neg x_{i'j}$ for all $i \neq i'$ and for all $j \in [n]$;
2. $x_{i1} \vee x_{i2} \vee \ldots \vee x_{in}$ for all $i \in [m]$.

As global partition we choose $\mathcal{P} = \{P_1, \ldots, P_n\}$, where $P_j = \{x_{ij} \mid i \in [m]\}$. We want to apply again the Corollary 4.1. As ψ we choose all the logical axioms plus all the axioms in (1). As μ we choose all the axioms in (2) and $I = \{0\}$.

We define $\mathcal{F}$ as the family of all the pairs $(\mathcal{Q}, \mathcal{H})$ such that $\mathcal{Q} \subseteq \mathcal{P}$ and $\mathcal{H}$ is the family of all the partial assignments of domain $\cup\mathcal{Q}$ satisfying the axioms in ψ with variables in $\cup\mathcal{Q}$.

PROPOSITION 4.2. *The family $\mathcal{F}$ defined above is n-extendible for ψ.*

Proof. The restriction part is clear. The extension part goes as follows: let $a \in \psi$ and $(\mathcal{Q}, \mathcal{H}) \in \mathcal{F}$, with $|\mathcal{Q}| < n$. We have that there exists exactly *one* $P_j \in \mathcal{P}$ such that $Var(a) \cap P_j \neq \emptyset$. If $P_j \in \mathcal{Q}$ we clearly have that $\mathcal{H} \models a$. If $P_j \notin \mathcal{Q}$ we set $\mathcal{Q}' = \mathcal{Q} \cup \{P_j\}$ and Σ the family of assignments with domain P_j verifying all the axioms in ψ with variables in P_j. We define

$$\mathcal{H}' = \{\alpha \mid \exists\beta \in \mathcal{H}\ \exists\gamma \in \Sigma\ \alpha = \beta \cup \gamma\}.$$

By Lemma 3.1 we have that $\mathcal{H}'$ is $\mathcal{Q}'$-lm.

Finally it is easy to see that $(\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$ that $(\mathcal{Q}', \mathcal{H}') \succeq (\mathcal{Q}, \mathcal{H})$, $\mathcal{H}' \models a$ and $|\mathcal{Q}'| = |\mathcal{Q}| + 1$. ☐

THEOREM 4.3 ([2]). *$Sp(PHP_n^m \vdash 1) \geq n/4$.*

Proof. We proved that the family $\mathcal{F}$ is n-extendible for ψ and it's easy to see that μ satisfy the requests of the Corollary 4.1. The result follows. ☐

Similarly with what we say about CT_n, it's possible to apply directly the Main Result (Theorem 3.1) also to PHP_n^m but in that manner we obtain as lower bound $n/8$. We wrote the Corollary 4.1 to re-obtain exactly the known space lower bound for PHP_n^m.

4.3 $BPHP_n^m$

The *Bit Pigeon-Hole Principle* is the formalization of the Pigeon-Hole principle that uses variables x_{ij} with $i \in [m]$ and $j \in [\log n]$. The intuitive meaning of the variable $x_{ij} = 1$ is "the pigeon i goes to some hole h and the j-th bit of a binary representation of h is 1". Similarly for $x_{ij} = 0$.

The axioms of $BPHP_n^m$ are clauses telling us that two pigeons i and i' can't go into the same hole h because they differ on the some bit of the binary representation of h. More formally for each hole $h \in [n]$ we consider the binary expansion of h, $(\epsilon_1^h, \ldots, \epsilon_{\log(n)}^h)_2$. Then if we define $B_{i,i'}^h = \bigvee_{j=1}^{\log(n)} \left(x_{ij} \neq \epsilon_j^h \vee x_{i'j} \neq \epsilon_j^h\right)$,

$$BPHP_n^m := \left\{B_{i,i'}^h \mid h \in [n], i \neq i' \in [m]\right\}.$$

We choose a global partition of the variables $\mathcal{P} = \{P_1, \ldots, P_m\}$ where $P_i = \{x_{ij} \mid j \in [\log n]\}$. Our strategy is to apply the Main Result (Theorem 3.1) using the ideal $I = Span(\{x_i^2 - x_i, x_i + \overline{x_i} - 1\}_{i=1,\ldots,n})$.

Given a hole $h = (\epsilon_1^h, \ldots, \epsilon_{\log(n)}^h)_2$ we define the hole $\overline{h} = (1 - \epsilon_1^h, \ldots, 1 - \epsilon_{\log(n)}^h)_2$. We observe that we have a partition of the holes $\mathcal{S} = \{S_1, \ldots, S_{n/2}\}$, where each $S_j = \{h, \overline{h}\}$ for some hole h.

We'll use the notation $\{i \mapsto h\}$ where $i \in [m]$ and $h \in [n]$, referring to a partial assignment α with domain P_i and such that $\alpha(x_{ij}) = \epsilon_j^h$.

We are now ready to define the family $\mathcal{F}$. The pair $(\mathcal{Q}, \mathcal{H})$ is in $\mathcal{F}$ if and only if:

1. $|\mathcal{Q}| \leq n/2$,
2. for each $A \in \mathcal{Q}$ there exists $i, i' \in [m]$ such that $A = P_i \cup P_{i'}$ and moreover $\mathcal{Q}$ is a pseudo-partition,
3. for each $A = P_i \cup P_{i'} \in \mathcal{Q}$ we choose a set $S_A = \{h, \overline{h}\} \in \mathcal{S}$ (for different elements of $\mathcal{Q}$ we choose different elements of $\mathcal{S}$) and we define $\alpha_A = \{i \mapsto h\} \cup \{i' \mapsto \overline{h}\}$ and $\overline{\alpha}_A = \{i \mapsto \overline{h}\} \cup \{i' \mapsto h\}$.

$$\mathcal{H} = \{\alpha \mid dom(\alpha) = \cup\mathcal{Q} \wedge \forall A \in \mathcal{Q}\ (\alpha\restriction_A = \alpha_A \vee \alpha\restriction_A = \overline{\alpha}_A)\}.$$

PROPOSITION 4.3. *$\mathcal{F}$ is $n/2$-extendible for to $BPHP_n^m$.*

Proof. The restriction part is obvious. Let us see the extension part: let $(\mathcal{Q}, \mathcal{H}) \in \mathcal{F}$ such that $|\mathcal{Q}| < n/2$ and a an axiom. If $a \in I$ we are done. So suppose now that $a = B_{i,i'}^h$. Clearly if i, i' are pigeons named in $\mathcal{Q}$ we have that $\mathcal{H} \models B_{i,i'}^h$. Then we can suppose that at least one of i or i' is not mentioned into $\mathcal{Q}$. If both are not mentioned into $\mathcal{Q}$ we put $A = P_i \cup P_{i'}$. If only one of them is not mentioned into $\mathcal{Q}$, wlog i is not mentioned into $\mathcal{Q}$, we want to find another pigeon not mentioned into $\mathcal{Q}$. We have that the pigeons mentioned in $\mathcal{Q}$ are strictly less than n and by hypothesis we have $m > n$ pigeons so we can find another pigeon i'' not mentioned into $\mathcal{Q}$. In this case we put $A = P_i \cup P_{i''}$. We define now $\mathcal{Q}' = \mathcal{Q} \cup \{A\}$. The assignments in $\mathcal{H}$ are naming strictly less than $n/2$ elements of the partition of the holes $\mathcal{S}$ (because the number of elements in $\mathcal{S}$ named in $\mathcal{H}$ equals $|\mathcal{Q}|$). So we can find an S_A not named in $\mathcal{H}$ and then we can define α_A and $\overline{\alpha}_A$ for that $\mathcal{S}_A$. Let $\Sigma = \{\alpha_A, \overline{\alpha}_A\}$. We define now $\mathcal{H}'$ as the set of partial assignments $\alpha = \beta \cup \gamma$ such that $\beta \in \mathcal{H}$ and $\gamma \in \Sigma$.

CLAIM 4.1. *Σ satisfys the requirements of Lemma 3.1.*

Proof. Let x be a variable in A. If $\alpha_A(x) = 0$ we have that $\overline{\alpha}_A(x) = 1$. This is by definition of α_A and $\overline{\alpha}_A$ and by the particular form of the partition of the holes $\mathcal{S}$ we choose. ☐

We can apply Lemma 3.1 and obtain that $\mathcal{H}'$ is $\mathcal{Q}'$-lm. It's straightforward to see that $(\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$ and that $(\mathcal{Q}', \mathcal{H}') \succeq (\mathcal{Q}, \mathcal{H})$, $\mathcal{H}' \models_I B_{i,i'}^h$ and that $|\mathcal{Q}'| = |\mathcal{Q}| + 1$. ☐

THEOREM 4.4 ([34]). *$Sp(BPHP_n^m \vdash 1) \geq n/8$.*

Proof. By the previous Proposition and the Main Theorem. ☐

4.4 $XPHP_n^m$

Quoting [34] we start recalling what is the *XOR pigeonhole principle formula* $XPHP_n^m$. $XPHP_n^m$ has propositional variables $x_{i,j}$ for each $i \in [0,m)$ and $j \in [0,n]$. (Recall that $[0,m) = \{0,\dots,m-1\}$ and $[0,n] = \{0,\dots,n\}$.) We think of $[0,m)$ as a set of pigeons and $[0,n]$ as a set of hole indicators. Each pigeon i gives a 0 or 1 value to every hole indicator j, recorded in the variable $x_{i,j}$.

The hole indicators indicate assignments of pigeons to holes indirectly: a pigeon $i \in [0,m)$ is assigned to a hole $j \in [0,n)$ when $x_{i,j} \not\equiv x_{i,j+1}$ is true, that is when $x_{i,j}$ and $x_{i,j+1}$ have different truth values. This assignment need not be unique: the formula will only ensure that each pigeon is assigned to an odd number of holes.

The formula $XPHP_n^m$ asserts the following:

1. Every pigeon gives different values to the first and last hole indicators. That is, for each $i \in [0,m)$, $x_{i,0} \not\equiv x_{i,n}$:
$$\begin{gathered} x_{i,0} \lor x_{i,n} \\ \neg x_{i,0} \lor \neg x_{i,n} \end{gathered}$$

2. At most one pigeon is assigned to any given hole. That is, for all distinct $i, i' \in [0,m)$ and all $j \in [0,n)$, $(x_{i,j} \equiv x_{i,j+1}) \lor (x_{i',j} \equiv x_{i',j+1})$:
$$\begin{gathered} x_{i,j} \lor \neg x_{i,j+1} \lor x_{i',j} \lor \neg x_{i',j+1} \\ \neg x_{i,j} \lor x_{i,j+1} \lor \neg x_{i',j} \lor x_{i',j+1} \\ x_{i,j} \lor \neg x_{i,j+1} \lor \neg x_{i',j} \lor x_{i',j+1} \\ \neg x_{i,j} \lor x_{i,j+1} \lor x_{i',j} \lor \neg x_{i',j+1} \end{gathered}$$

$XPHP_n^m$ is the conjunction of all the previous clauses so $XPHP_n^m$ is a 4-CNF and for $m > n$ it is a contradiction. To see this notice that, by condition (1), for each pigeon $i \in [0,m)$ there must be at least one hole $j \in [0,n)$ for which i gives different values to indicators j and $j+1$; say that such a j is assigned to i. Since $n < m$, by the pigeonhole principle there must be some pair of distinct pigeons which are assigned the same hole. But this contradicts condition (2).

Let us fix the partition of the variables $\mathcal{P} = \{P_0,\dots,P_{m-1}\}$, where $P_i = \{x_{i,j} \mid j \in [0,n)\}$, and the ideal $I = \mathit{Span}(\{x_i^2 - x_i, x_i + \overline{x_i} - 1\}_{i=1,\dots,n})$. We consider the family $\mathcal{F}$ made up by the pairs $(\mathcal{Q},\mathcal{H})$ such that

1. $\mathcal{Q} \subseteq \mathcal{P}$,

2. $\mathcal{H}$ is an encoding of all the injective assignments of the pigeons named in $\mathcal{Q}$. Formally we define $\{i \mapsto j\}$ as the two partial assignments $\alpha_{i,j}$ and $\beta_{i,j}$ with domain P_i where
$$\alpha_{i,j}(x_{i,j'}) = \begin{cases} 0 & \text{if } j' \le j \\ 1 & \text{otherwise} \end{cases}$$
and $\beta_{i,j}(x_{i,j'}) = 1 - \alpha_{i,j}(x_{i,j'})$. Then $\mathcal{H}$ is made up by assignments $\{i \mapsto j_i\}$ where P_i is one of the elements in $\mathcal{Q}$ and all the j_i are distinct.

PROPOSITION 4.4. *The family $\mathcal{F}$ defined above is $(n-1)$-extendible of $XPHP_n^m$.*

Proof. The restriction part is clear. Let us focus on the extension part. Let $(\mathcal{Q},\mathcal{H}) \in \mathcal{F}$ with $|\mathcal{Q}| < n-2$ and a an initial axiom in $XPHP_n^m$. We want to find a pair $(\mathcal{Q}',\mathcal{H}') \in \mathcal{F}$ such that (1) $(\mathcal{Q}',\mathcal{H}') \succeq (\mathcal{Q},\mathcal{H})$, (2) $\mathcal{H}' \models_I a$, (3) $|\mathcal{Q}'| \le |\mathcal{Q}| + 1$.

Let us suppose first that $a = (x_{i,j} \equiv x_{i,j+1}) \lor (x_{i',j} \equiv x_{i',j+1})$. If both P_i and P_i' are in $\mathcal{Q}$ by definition $\mathcal{H} \models_I a$. So w.l.o.g. suppose that $P_i \notin \mathcal{Q}$. Define $\mathcal{Q}' = \mathcal{Q} \cup \{P_i\}$. We have that $\mathcal{H}$ is made up of an injective assignment of at most $n-2$ pigeons, so we can find a hole h different from j not assigned. We define

$$\Sigma = \{i \mapsto h\},$$

then

$$\mathcal{H}' = \{\alpha \mid \exists\beta \in \mathcal{H}\ \exists\gamma \in \Sigma\ (\alpha = \beta \cup \gamma)\}.$$

By Lemma 3.1 we have that $\mathcal{H}'$ is $\mathcal{Q}'$-lm and it's straightforward to see that $(\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$ and the properties (a), (b) and (c) hold.

Similarly if $a = (x_{i,0} \not\equiv x_{i,n})$ we proceed as before assigning the pigeon i somewhere if needed. □

THEOREM 4.5. *([34]) $Sp(XPHP^m_n \vdash 1) \geq (n-1)/4$.*

Proof. By the previous Proposition and the Main Theorem. □

This is only slightly worse than the result obtained in [34].

5 New Results: Space Lower Bounds for Random Formulas and Graph-PHP

We prove that to refute random k-CNF formulas over n variables (and the Graph-PHP) it will be required high space in PC/PCR. We are going to construct a family of $\Omega(n)$-extendible admissible configurations for random k-CNF, $k \geq 4$. The main tool we use is a variation of the *Matching Game* (see Section 2.3.1) which was devised in [12] to prove space lower bound for random k-CNF in Resolution. It was also used in [4] to prove indefinability of random k-CNF in certain fragments of first order logic. Differently from these cases, here we are dealing with *double* matchings in a bipartite graph, instead of simply matchings. This is making some difference in the argument. Nevertheless the proofs of the main properties are essentially similar to that of [12, 4], except for small details which are due mainly to the fact that the invariant property (the (r,s)-double matching property we define next) deals with double matchings.

DEFINITION 5.1 ((r,s)-double matching property). *Let $r \leq s$, $\mathcal{G} = (U \cup V, E)$ a bipartite graph and $A \subseteq U$ of size at most $r \leq s$ and $B \subseteq V \cap N_{\mathcal{G}}(A)$. We say that $(\mathcal{G}, A, B)$ has the* (r,s)-double matching property *if for every $C \subseteq U \setminus A$, if $|C| = s - |A|$ then there exists a 2-matching of C into $V \setminus B$.*

LEMMA 5.1 (extension lemma). *Let $\mathcal{G} = (U \cup V, E)$ be a bipartite graph of left degree at most d that is a (s,ϵ)-bipartite expander. Let $A \subseteq U$ and $B \subseteq V$ such that $(\mathcal{G}, A, B)$ has the (r,s)-double matching property with*

$$r \leq \frac{\epsilon s}{d^2(d-1) + \epsilon}$$

and $|A| < r$.

For each $u \in U \setminus A$ there exists two distinct nodes $v, v' \in N_{\mathcal{G}}(u) \cap (V \setminus B)$ such that $(\mathcal{G}, A \cup \{u\}, B \cup \{v, v'\})$ has the (r,s)-double matching property.

Proof. Let $N_{\mathcal{G}}(u) \cap (V \setminus B) = \{v_1, \ldots, v_l\}$. Clearly we have that $l \leq d$ because $\mathcal{G}$ has left degree at most d and $l \geq 2$ because of the (r,s)-double matching property on $(\mathcal{G}, A, B)$.

Let $A' = A \cup \{u\}$ and $B^{ij} = B \cup \{v_i, v_j\}$ with $v_i \neq v_j$. We note that $|A'| \leq r$ because $|A| < r$ and $|A'| = |A| + 1$. Let us suppose for sake of contradiction that for every $i \in \{1, \ldots, l\}$ $(\mathcal{G}, A', B^{ij})$ has not the (r,s)-double matching property. This means that for every $i \neq j$ we have a set $C^{ij} \subseteq U \setminus A'$ that does not admit a 2-matching into $V \setminus B^{ij}$ s.t. $|C^{ij}| = s - |A'|$. Let $D^{ij} \subseteq C^{ij}$ not admitting a 2-matching into $V \setminus B^{ij}$ of minimal size. Then, by Lemma 2.1, we have that

$$|N_{\mathcal{G}}(D^{ij}) \cap (V \setminus B^{ij})| < 2|D^{ij}|,$$

so we obtain that

$$(2+\epsilon)|D^{ij}| \leq |N_{\mathcal{G}}(D^{ij})| = |N_{\mathcal{G}}(D^{ij}) \cap (V \setminus B^{ij})| + |N_{\mathcal{G}}(D^{ij}) \cap B^{ij}| < 2|D^{ij}| + |B^{ij}|, \tag{1}$$

where the first inequality came from the expansion property of $\mathcal{G}$ since $|D^{ij}| \leq s - |A'| < s$. From this chain of inequalities we obtain immediately that

$$|B^{ij}| > \epsilon |D^{ij}|,$$

and, using the fact that $B^{ij} \subseteq N_{\mathcal{G}}(A')$, we have that $|B^{ij}| \leq d|A'|$. Putting all this inequalities together we have that

$$d|A'| > \epsilon |D^{ij}|.$$

Claim 5.1. *$\bigcup_{i \neq j} D^{ij} \cup \{u\}$ does not admit a 2-matching into $V \setminus B$.*

Proof. To prove this suppose by contradiction that there exists a 2-matching $\pi \subseteq E$ of that set into $V \setminus B$. Let $\pi(u) = \{v_h, v_k\}$. We have that $\pi(D^{hk}) \cap \pi(u) \neq \emptyset$, in fact $\pi(D^{hk}) \subseteq V \setminus B$ and, by construction, $\pi(D^{hk}) \not\subseteq V \setminus B^{ij}$. So we must have that $\pi(D^{hk}) \cap \{v_h, v_k\} \neq \emptyset$. We reach a contradiction observing that $u \notin D^{h,k}$ so we obtain two elements mapped by π in the same element. □

We have that $\bigcup_{ij} D^{ij} \cup \{u\} \subseteq U \setminus A$ and $(\mathcal{G}, A, B)$ by hypothesis has the double matching property, so we must have that

$$\left| \bigcup_{ij} D_{ij} \cup \{u\} \right| > s - |A|$$

so we have that there exists a pair of indexes i, j such that $|D^{ij}| > \frac{s-|A'|}{l(l-1)} \geq \frac{s-|A'|}{d(d-1)}$.

So we have obtained that

$$d|A'| > \epsilon \frac{s - |A'|}{d(d-1)}.$$

And from this we obtain that

$$|A'| > \frac{\epsilon s}{d^2(d-1) + \epsilon} = r.$$

But this is a contradiction by hypothesis. □

Lemma 5.2 (retraction lemma). *Let $\mathcal{G} = (U \cup V, E)$ be a bipartite graph of left degree at most d that is a (s, ϵ)-bipartite expander. Let $A \subseteq U$ and $B \subseteq V$ such that $(\mathcal{G}, A, B)$ has the (r, s)-double matching property. If $u \in A$ and $L \subseteq N_{\mathcal{G}}(u) \cap B$ such that $|L| \geq 2$ and $B \setminus L \subseteq N_{\mathcal{G}}(A \setminus \{u\})$ and*

$$r \leq \frac{\epsilon s}{d + \epsilon},$$

then $(\mathcal{G}, A \setminus \{u\}, B \setminus L)$ has the (r, s)-double matching property.

Proof. Let $A' = A \setminus \{u\}$ and $B' = B \setminus L$. Clearly $|A'| \leq r$ and $B' \subseteq N_{\mathcal{G}}(A')$. Let $C \subseteq U \setminus A'$ of size $s - |A'|$. We have two cases: or $u \in C$ or $u \notin C$.

$\underline{u \in C:}$ In this case we have that $C \setminus \{u\} \subseteq U \setminus A$, and has size $s - |A'| - 1 = s - |A|$, then we have that there exists a 2-matching of $C \setminus \{u\}$ into $V \setminus B$. We have now by hypothesis that $|L| \geq 2$ so we can find (u, v) and (u, w) in $\{u\} \times L$. So we can extend the 2-matching found for $C \setminus \{u\}$ to a 2-matching of C.

$\underline{u \notin C:}$ We have that for every $w \in C$ there exists a 2-matching of $C \setminus \{w\} \subseteq U \setminus A$ into $V \setminus B \subseteq V \setminus B'$. Then if C is not 2-matchable into $V \setminus B'$ it follows that C is not 2-matchable of minimal size. Using Lemma 2.1 we have that

$$|N_{\mathcal{G}}(C) \cap (V \setminus B')| < 2|C|,$$

and, using the fact that $\mathcal{G}$ is an (s, ϵ)-bipartite expander, and that $|C| = s - |A'| \leq s$,

$$(2 + \epsilon)|C| \leq |N_{\mathcal{G}}(C)| < 2|C| + |B'|.$$

So $|B'| > \epsilon|C|$. We have now that $|C| = s - |A'|$, and $|B'| \leq d|A'|$, so we obtain

$$|A'| > \frac{\epsilon s}{d+\epsilon} \geq r.$$

A contradiction. □

5.1 Random k-CNF

Let n,m and k be positive natural numbers and let $X = \{x_1, \ldots, x_n\}$ be a set of variables. Let $\mathcal{F}(n, m, k)$ be the set of all k-CNF formulas on X with exactly m clauses each defined on k literals on distinct variables. Alternatively, $\mathcal{F}(n, m, k)$ can be described as the result of repeating m times independently the following experiment: choose exactly k variables from X, and negate each variable independently with probability $1/2$. We will use this interpretation whenever it is convenient. The ratio m/n is denoted by Δ, and is called the clause density. Usually, Δ is fixed to a constant and therefore is determined by n . We are interested in studying the asymptotic properties of a randomly chosen formula $F \sim \mathcal{F}(n, m, k)$ as n approaches to infinity. It is well known that when the clause density exceeds a certain constant θ_k that only depends on k, a randomly chosen formula is almost surely unsatisfiable. We are interested only in the region in which F is unsatisfiable with high probability, then we always consider fixed $\Delta >> \theta_k$, then $\mathcal{F}(n, m, k)$ can be made dependent only on n, Δ and k and denoted as $\mathcal{F}(n, \Delta, k)$

Let $F = \bigwedge_{i=1}^{\Delta n} C_k \sim \mathcal{F}(n, \Delta, k)$ be a random k-CNF. Let us consider the associated bipartite graph $\mathcal{G}_F = (U \cup V, E)$ where U is the set of clauses appearing in F and V is the underlying set of variables appearing in F. As in [12, 4] we put $(C, x) \in E$ if the variable x is appearing in some literal of C. We observe that the graph $\mathcal{G}_F$ has left degree k. It is a well-known result (see [25, 10, 18, 12, 4] among several others) that if $F \sim \mathcal{F}(n, \Delta, k)$, then $\mathcal{G}_F$ is a good expander (at least when the expansion factor is $(1 + \epsilon)$). Since in this work we are dealing with 2-matchings, we are interested in an expansion factor of $(2 + \epsilon)$ (see Definition 2.7). Nevertheless we are able to prove that also in this case $\mathcal{G}_F$ is a good expander, provided $k \geq 4$. We comment on the case $k = 3$ in the conclusions. The proof of next theorem is standard and can be found for instance in [12]. Our proof contains exactly the same calculations with the only difference that to deal with an expansion factor of $(2 + \epsilon)$ in $\mathcal{G}_F$ we need to have $k \geq 4$.

Theorem 5.1 ([25, 10, 12, 18]). *For any $k \geq 4$ and any constant ϵ with $0 < \epsilon < k - 3$, there is a constant $\kappa = \kappa_{k,\epsilon}$ such that if $F \sim \mathcal{F}(n, \Delta, k)$, then with high probability $\mathcal{G}_F$ is a (s, ϵ)-bipartite expander, with $s = \frac{\kappa \cdot n}{\Delta^{\frac{1+\epsilon}{k-3-\epsilon}}}$.*

Proof. The same proof given in [12] works exacly in our context substituting each occurrence of $(1 + \epsilon)$ with $(2 + \epsilon)$ and the condition $k \geq 3$ with $k \geq 4$. □

Let us suppose that the graph $\mathcal{G}$ is an (s, ϵ)-bipartite expander. Notice that for all k and s,

$$\frac{\epsilon s}{k^2(k-1)+\epsilon} = \min\left\{s, \frac{\epsilon s}{k+\epsilon}, \frac{\epsilon s}{k^2(k-1)+\epsilon}\right\}.$$

Let $\tilde{r}$ be that minimum and $I = Span(\{x_i^2 - x_i, x_i + \overline{x_i} - 1\}_{i=1,\ldots,n})$. We define the family $\mathcal{F}$ as follow: the pair $(\mathcal{Q}, \mathcal{H})$ is in $\mathcal{F}$ is and only if there exists a multiple matching π of some $A \subseteq U$ such that:

1. $|A| \leq \tilde{r}$,
2. $(\mathcal{G}, A, \cup\pi(A))$ has the $(\tilde{r}, s)$-double matching property,
3. $\mathcal{Q} = \pi(A)$,
4. $\mathcal{H}$ is such that $(\mathcal{Q}, \mathcal{H})$ is an admissible configuration and for each clause $C \in A$ $\mathcal{H}\restriction_{\pi(C)} \models_I C$.

Clearly we have that $(\emptyset, \{\emptyset\}) \in \mathcal{F}$ so the family we defined is non-empty and all the pairs in $\mathcal{F}$ are admissible configurations.

THEOREM 5.2. *The family $\mathcal{F}$ defined above is $\tilde{r}$-extendible.*

Proof. Suppose we have a pair $(\mathcal{Q}, \mathcal{H}) \in \mathcal{F}$, i.e. we have the properties (1), (2), (3) and (4) listed above. Clearly we have that $|\mathcal{Q}| \leq \tilde{r}$, because $|\mathcal{Q}| = |A|$ and, by (1), $|A| \leq \tilde{r}$.

If we have a $\mathcal{Q}' \subseteq \mathcal{Q}$ we have to prove that $(\mathcal{Q}', \mathcal{H}\restriction_{\mathcal{Q}'}) \in \mathcal{F}$. Let $A' = \{u \in A \mid \pi(u) \in \mathcal{Q}'\}$, $\pi' = \pi\restriction_{A'}$ the *multiple* matching obtained as a restriction of π over A' and $\mathcal{H}' = \mathcal{H}\restriction_{\mathcal{Q}'}$. (1) is true since $|A'| \leq |A| \leq \tilde{r}$. (3) is true since $\mathcal{Q}' = \pi'(A')$. (4) follows since $\pi = \pi'$ over A' and for all $C \in A'$, $\pi(C) \in \mathcal{Q}'$ and then we have $\mathcal{H}'\restriction_{\pi'(C)} \models C$ for each $C \in A'$. The difficult part is to prove that $(\mathcal{G}, A', \cup\pi'(A'))$ has the $(\tilde{r}, s)$-double matching property. We remove one by one the clauses $C \in A \setminus A'$ by applying for each such C the *retraction* Lemma 5.2 with $u = C$ and $L = \pi(C)$. It is straightforward to see that such L fulfills the hypothesis of retraction Lemma 5.2.

To prove the extension property for the family $\mathcal{F}$, let us suppose that $|\mathcal{Q}| < \tilde{r}$ and that we have an axiom a and we want to prove that we can find a pair $(\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$ such that (a) $(\mathcal{Q}', \mathcal{H}') \succeq (\mathcal{Q}, \mathcal{H})$, (b) $\mathcal{H}' \models a$ and (c) $|\mathcal{Q}'| \leq |\mathcal{Q}| + 1$.

As usual we need to distinguish the case of $a \in I$ (in this case we don't have anything to do) or a being a clause in the graph $\mathcal{G}$. Let us consider the case $a = C$ being a clause of the random kCNF. If $C \in A$ clearly we have done. If $C \notin A$ by Lemma 5.1 we can find a two distinct vertexes $v, v' \in V \setminus \cup\pi(A)$ such that $v, v' \in N_{\mathcal{G}}(u)$ and $(\mathcal{G}, A \cup \{u\}, \cup\pi(A) \cup \{v, v'\})$ has the $(\tilde{r}, s)$-double matching property. So we define $A' = A \cup \{C\}$ and $\pi' = \pi \cup \{(C, v), (C, v')\}$. And we define $\mathcal{Q}' = \mathcal{Q} \cup \{\{v, v'\}\} = \pi'(A')$. The only thing left is to construct the family $\mathcal{H}'$. To do this first we define a family $\Sigma = \{\gamma, \overline{\gamma}\}$ such that

- $dom(\gamma) = dom(\overline{\gamma}) = \{v, v'\}$,
- $\gamma(v) = sat_C(v)$ and $\gamma(v') = 1 - sat_C(v')$, where $sat_C(x) \in \{0, 1\}$ is the value we have to set the variable x to satisfy C,
- $\overline{\gamma}(v) = 1 - sat_C(v)$ and $\overline{\gamma}(v') = sat_C(v')$.

Then we define
$$\mathcal{H}' = \{\alpha \mid \exists \beta \in \mathcal{H}\ (\alpha = \beta \cup \gamma \vee \alpha = \beta \cup \overline{\gamma})\}.$$
Notice that $\mathcal{H}'$ is $\mathcal{Q}'$-lm with respect to I by Lemma 3.1. It is straightforward to see that for $(\mathcal{Q}', \mathcal{H}')$ the properties (a), (b), (c) of extendibility hold. Hence $(\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$. ☐

In this versions of the work we are not interested in improving constants, so we omit detailed calculations that will instead follow in a subsequent version of the paper.

THEOREM 5.3 (space lower bound for random k-CNF). *Let $k \geq 4$ be any integer, $\epsilon > 0$ any constant and $\Delta \geq 1$. Let $F \sim \mathcal{F}(n, \Delta, k)$. There exists a constant $c = c_{k,\Delta,\epsilon}$, $c \geq 1$, such that with high probability $Sp(F \vdash 1) \geq \frac{n}{4c}$.*

Proof. Theorem 5.1 tells us that with high probability $\mathcal{G}_F$ is a (s, ϵ)-expander, with $s = \frac{\kappa \cdot n}{\Delta^{\frac{1+\epsilon}{k-3-\epsilon}}}$. Using definition of $\tilde{r}$ we have that there exists a a constant $c = c_{k,\Delta,\epsilon}$ such that with high probability the family of admissible configurations of Theorem 5.2 is $(\frac{n}{c})$-extendible family for F. The theorem then follows by Theorem 3.1. ☐

5.2 Graph-PHP

We recall the definition of the Graph-PHP in order to fix the notations we'll use. Let $\mathcal{G} = (U \cup V, E)$ a bipartite graph and U and V two disjoint sets of size respectively $n + 1$ and n. Clearly there is no perfect matching from U to V. This combinatorial principle is expressed as a conjunction over the variables $W = \{x_{u,v} \mid (u, v) \in E\}$. Intuitively setting the variable $x_{u,v}$ to 1 means that the pigeon $u \in U$ is mapped to $v \in V$. For every $u \in U$ let
$$P_u = \bigvee_{v:(u,v) \in E} x_{u,v}$$

and for all $(u,w)\in E$ and $(v,w)\in E$ let

$$H_w^{u,v} = \neg x_{u,w} \vee \neg x_{v,w}.$$

$\mathcal{G}$-PHP is the conjunction of all the previous clauses. We observe that if $\mathcal{G}$ has left degree d then $\mathcal{G}$-PHP is a d-CNF.

According with the general strategy we fix the partition $\mathcal{P} = \{H_1,\dots,H_n\}$, where $H_j = \{x_{ij} \mid (i,j)\in E\}$, i.e. we are partitioning the variables according to the hole the are referring. Let us start with a notation we'll use: suppose we have a (multiple) matching π of a set $A\subseteq U$. For every $u\in A$ we call

$$var(\pi)(u) = \bigcup_{i\in\pi(u)} H_j$$

and for each $B\subseteq A$

$$var(\pi)(B) = \{var(\pi)(u) \mid u\in B\}.$$

Let us suppose that the graph $\mathcal{G}$ is an (s,ϵ)-bipartite expander with left degree d and let

$$\tilde{r} = \min\left\{s, \frac{\epsilon s}{d+\epsilon}, \frac{\epsilon s}{d^2(d-1)+\epsilon}\right\} = \frac{\epsilon s}{d^2(d-1)+\epsilon}.$$

Let consider the ideal $I = Span\left(\{H_w^{u,v}\}\cup\{x_i^2 - x_i, x_i+\overline{x_i}-1\}_{i=1,\dots,n}\right)$. We want now to construct a family $\mathcal{F}$ that is $\tilde{r}$-extendible with respect to I. The family $\mathcal{F}$ is defined as follow: the pair $(\mathcal{Q},\mathcal{H})\in\mathcal{F}$ if and only if there exists a multiple matching π of some $A\subseteq U$ st

1. $|A|\leq\tilde{r}$,
2. $(\mathcal{G},A,\cup\pi(A))$ has the $(\tilde{r},s)$-double matching property,
3. $\mathcal{Q} = var(\pi)(A)$,
4. $\mathcal{H}$ is such that $(\mathcal{Q},\mathcal{H})$ is an admissible configuration, for each $u\in A$ $\mathcal{H}\restriction_{var(\pi)(u)}\models_I P_u$ and $\mathcal{H}\restriction_{var(\pi)(u)}$ respects I.

Clearly we have that $(\emptyset,\{\emptyset\})\in\mathcal{F}$ so the family we defined is non-empty.

Formally the definition of this family is very close to the definition we had for the random k-CNF, but the ideal used is different so the proof that the family above is well defined and $\tilde{r}$-extendible is somehow different from the proof we provided for the random formulas.

THEOREM 5.4. *The family $\mathcal{F}$ defined above is $\tilde{r}$-extendible.*

Proof. Let us suppose we have a pair $(\mathcal{Q},\mathcal{H})\in\mathcal{F}$, i.e. we have the properties (1), (2), (3) and (4) listed above. Clearly we have that $|\mathcal{Q}|\leq\tilde{r}$, because $|\mathcal{Q}|=|A|$ and, by (1), $|A|\leq\tilde{r}$.

If we have a $\mathcal{Q}'\subseteq\mathcal{Q}$ we have to prove that $(\mathcal{Q}',\mathcal{H}\restriction_{\mathcal{Q}'})\in\mathcal{F}$. Let $A' = \{u\in A\mid \pi(u)\in\mathcal{Q}'\}$ and $\pi' = \pi\restriction_{A'}$ the matching obtained as a restriction of π over A'. Clearly we have that $|A'|\leq|A|\leq\tilde{r}$, $\mathcal{Q}' = var(\pi')(A')$ and $\mathcal{H}\restriction_{var(\pi')(u)}\models P_u$ for each $u\in A'$. The difficult part is to prove that $(\mathcal{G},A',\pi'(A'))$ has the $(\tilde{r},s)$-double matching property. We remove one by one the clauses $C\in A\setminus A'$ by applying for each such C Lemma 5.2 with $u=C$ and $L=\pi(C)$. It is straightforward to see that such L fulfills the hypothesis of Lemma 5.2.

To prove the extension property let's suppose that $|\mathcal{Q}|<\tilde{r}$ and that we have an axiom a and we want to prove that we can find a pair $(\mathcal{Q}',\mathcal{H}')\in\mathcal{F}$ s.t. (a) $(\mathcal{Q}',\mathcal{H}')\succeq(\mathcal{Q},\mathcal{H})$, (b) $\mathcal{H}'\models_I a$ and (c) $|\mathcal{Q}'|\leq|\mathcal{Q}|+1$. As usual we need to distinguish teo cases: $a\in I$ (i.e a is a logical axiom or $a = H_w^{u,v}$ for some u,v,w) or $a = P_u$ for some $u\in U$. In the first case, $a\in I$, as usual we have nothing to do.

Let us consider the case $a = P_u$ for some $u\in U$. If $u\in A$ clearly we are done. If $u\notin A$ by Lemma 5.1 we can find two distinct vertexes $v,v'\in N_G(u)$ not in $\cup\pi(A)$ such that $(\mathcal{G},A\cup\{u\},\cup\pi(A)\cup\{v,v'\})$ has the $(\tilde{r},s)$-double matching property. So we define $A' = A\cup\{u\}$ and $\pi' = \pi\cup\{(u,v),(u,v')\}$. And we define $\mathcal{Q}' = \mathcal{Q}\cup\{H_v\cup H_{v'}\} = var(\pi')(A')$. We have now to construct the family $\mathcal{H}$. To do this first we define a family of partial assignments Σ as the set of all assignments of domain $H_v \cup H_{v'}$ satisfying all the axioms $H_v^{ww'}$ and $H_{v'}^{ww'}$, i.e. all the axioms stating the injectivity on the holes v and v' and setting $x_{u,v}$ to 1 and $x_{u,v'}$ to 0, or setting $x_{u,v}$ to 0 and $x_{u,v'}$ to 1. We observe that the assignments we put in Σ respect I. Then we define

$$\mathcal{H}' = \{\alpha \mid \exists \beta \in \mathcal{H}\ \exists \gamma \in \Sigma\ (\alpha = \beta \cup \gamma)\}.$$

By Lemma 3.1 we have that $\mathcal{H}'$ is $\mathcal{Q}'$-lm with respect to I. Moreover it's straightforward to see that $(\mathcal{Q}', \mathcal{H}') \in \mathcal{F}$ and that the properties (a), (b), (c) hold. □

Theorem 5.5 (space lower bound for $\mathcal{G}$-PHP). *There exists a constant degree $d \geq 3$ bipartite graph $\mathcal{G} = (U \cup V, E)$ with $|U| = n+1$ and $|V| = n$, such that $Sp(\mathcal{G} - PHP \vdash 1) \geq \Omega(n/d)$.*

Proof. We proceed as in [12]. A similar proof to that Ben-Sasson in his thesis [11] (Theorem 2.46) prove that there exists a degree d bipartite graph $\mathcal{G} = (U \cup V, E)$ with $|U| = n+1$ and $|V| = n$ which is a $(\Omega(n/d), 7d/8 - 2)$-expander (it is sufficient to set $\epsilon = 7d/8 - 2$ in his proof for his calculations to work with our expansion factor of $(2 + \epsilon)$). The Theorem then follows using definition of $\tilde{r}$, previous Theorem 5.4 and Main Theorem 3.1. □

6 Open Problems

We think that our characterization of the space in Pc/Pcr can open the way to a more precise characterization of the space, and we do not exclude the degree, of Pc/Pcr proofs in terms of 2-Player games like variants of the existential pebble games for Resolution like Ehrenfeucht-Fraïssé games. We find very attractive the idea that, as was done in Resolution by Atserias and Dalmau in [5], to find a precise combinatorial characterization of the degree and proving some relations between space and degree, similar to the one between width and space for Resolution. We think that our work and our notion of k-extendibility is a first step in this directions. So far there is no results that seems to exclude that "space might be lower bounded by degree" in Pc/Pcr. As was done for random k-CNF for DATALOG by Asterias [4], our game characterization of boolean reasoning with polynomials can suggest non-expressibility results in stronger logic appropriate to this kind of reasoning.

To work in this direction it might be useful to prove lower bounds for other classes of tautologies for which we known they require high degree. In particular we think to Tseitin Tautologies (Beame et al. in [23] proves that they require high degree) and Linear ordering principle on Graphs GOP_n (Galesi and Lauria [35] recently proved they require high degree in Pc/Pcr) or GT_n. We think that our characterization could work also for this case provided we have the right definition of graph underlying the principle.

Another issue concern the possibilities of using a similar characterization of the space to try prove space lower bounds in other more powerful systems. Nothing for instance is known about space complexity in Cutting Planes and Lovasz-Schriver proof systems. We think that also in this case our work can be a starting point to try to come up with similar ideas to prove space lower bounds in these systems.

Another natural open problem arising form our work is to study the variable space for Pc/Pcr for all the principles we prove space lower bounds for. We think that the same steps of [2] together with our approach based on transversality and pseudo-partitions one can hopefully prove quadratic lower bounds for variable space in all these cases.

Finally our work leaves open to study the case of the Pc/Pcr space for random 3-CNF. A solution to this question comes from the solution to the problem of showing the existence of a bipartite graph with left degree equals 3 having a sufficiently good expansion property (for instance with an expansion factor of $(2 + \epsilon)$ if one considers our definition of expansion).

7 Acknowledgements

The authors are greatly grateful to Roberto Grossi and to the Department of Computer Science of the University of Pisa for the very kind hospitality they offer. Part of this work was done while Nicola Galesi was visiting Roberto Grossi at the University of Pisa. Massimo Lauria was listening our results in a preliminary phase of this work and gave us some useful comments and insights, was reading carefully our paper and also found some small mistake. We thank him for his help and expertise.

Last but not least we want to thanks Jacob Nordström for discussing our result with us and inspiring comments about future research directions, Mladen Mikša and Mark Vinals Perez to have read very carefully this work.

References

[1] Dimitris Achlioptas. *Random Satisfiability*, volume 185 of *Frontiers in Artificial Intelligence and Applications*, chapter 8, pages 245–270. IOS Press, February 2009. 7

[2] Michael Alekhnovich, Eli Ben-Sasson, Alexander A. Razborov, and Avi Wigderson. Space complexity in propositional calculus. *SIAM J. Comput.*, 31(4):1184–1211, 2002. 1, 3, 5, 6, 8, 10, 11, 12, 13, 14, 15, 16, 18, 24, 25, 33

[3] Michael Alekhnovich and Alexander A. Razborov. Lower bounds for polynomial calculus: Non-binomial case. In *42nd Annual Symposium on Foundations of Computer Science*, pages 190–199, 2001. 3, 8

[4] Albert Atserias. On sufficient conditions for unsatisfiability of random formulas. *Journal of ACM*, 51(2):281–311, 2004. A preliminary version appeared in the Proceedings of the 17th IEEE Symposium on Logic in Computer Sience (LICS), pagine 325–334, 2002. 3, 4, 9, 10, 13, 14, 18, 28, 30, 33

[5] Albert Atserias and Víctor Dalmau. A combinatorial characterization of resolution width. *J. Comput. Syst. Sci.*, 74(3):323–334, 2008. 4, 9, 10, 14, 20, 33

[6] Albert Atserias, J. Fichte, and M. Thurley. Clause-learning algorithms with many restarts and bounded-width resolution. *Theory and Applications of Satisfiability Testing-SAT 2009*, pages 114–127, 2009. 4

[7] Paul Beame, Chris Beck, and Russell Impagliazzo. Time-space tradeoffs in resolution: Superpolynomial lower bounds for superlinear space. Technical Report TR11-149, ECCC-Elctronic Colloquun on Computational Complexity, 2011. 8

[8] Paul Beame, Richard M. Karp, Toniann Pitassi, and Michael E. Saks. On the complexity of unsatisfiability proofs for random k-cnf formulas. In *STOC*, pages 561–571, 1998. 8

[9] Paul Beame, Henry A. Kautz, and Ashish Sabharwal. Towards understanding and harnessing the potential of clause learning. *J. Artif. Intell. Res. (JAIR)*, 22:319–351, 2004. 4

[10] Paul Beame and Toniann Pitassi. Simplified and improved resolution lower bounds. In *37th Annual Symposium on Foundations of Computer Science*, pages 274 – 282. IEEE, 1996. 8, 30

[11] Eli Ben-sasson. *Expansion in Proof Complexity.* PhD thesis, Hebrew University, 2001. 33

[12] Eli Ben-Sasson and Nicola Galesi. Space complexity of random formulae in resolution. *Random Struct. Algorithms*, 23(1):92–109, 2003. 1, 3, 8, 10, 11, 13, 18, 28, 30, 33

[13] Eli Ben-Sasson and Russell Impagliazzo. Random CNFs are hard for the polynomial calculus. In *40th Annual Symposium on Foundations of Computer Science*, pages 415–421, 1999. 3, 8

[14] Eli Ben-Sasson and Jan Johannsen. Lower bounds for width-restricted clause learning on small width formulas. *Theory and Applications of Satisfiability Testing–SAT 2010*, pages 16–29, 2010. 4

[15] Eli Ben-Sasson and Jakob Nordström. A space hierarchy for k-dnf resolution. *Electronic Colloquium on Computational Complexity (ECCC)*, 16(047), 2009. 3

[16] Eli Ben-Sasson and Jakob Nordström. Understanding space in resolution: Optimal lower bounds and exponential trade-offs. *Electronic Colloquium on Computational Complexity (ECCC)*, 16(034), 2009. 3, 8

[17] Eli Ben-Sasson and Jakob Nordström. Understanding space in proof complexity: Separations and trade-ofss via substitutions. In *In Proceedings of the 2nd Symposium on Innovations in Computer Science (ICS '11),*, pages 401–416, 2011. 8

[18] Eli Ben-Sasson and Avi Wigderson. Short proofs are narrow - resolution made simple. *J. ACM*, 48(2):149–169, 2001. 8, 11, 30

[19] Archie Blake. *Canonical Expressions in Boolean Algebra*. PhD thesis, University of Chicago, 1938. 3

[20] Michael Brickenstein and Alexander Dreyer. Polybori: A framework for gröbner-basis computations with boolean polynomials. *J. Symb. Comput*, 44(9):1326–1345, 2009. 4

[21] Sam Buss, Maria Luisa Bonet, and Jan Johannsen. Improved separations of regular resolution from clause learning proof systems. *Submitted*, 2012. 4

[22] Sam Buss, Jan Hoffman, and Jan Johannsen. Resolution trees with lemmas - resolution refinements that characterize dll-algorithms with clause learning. *Logical Methods in Computer Science*, 4,:4:13, 2008. 4

[23] Samuel R. Buss, Dima Grigoriev, Russell Impagliazzo, and Toniann Pitassi. Linear gaps between degrees for the polynomial calculus modulo distinct primes. *J. Comput. Syst. Sci.*, 62(2):267–289, 2001. 3, 8, 14, 33

[24] Samuel R. Buss, Russell Impagliazzo, Jan Krajíček, Pavel Pudlák, Alexander A. Razborov, and Jiří Sgall. Proof complexity in algebraic systems and bounded depth frege systems with modular counting. *Computational Complexity*, 6(3):256–298, 1997. 3, 8

[25] Vasek Chvátal and Endre Szemerédi. Many hard examples for resolution. *J. ACM*, 35(4):759–768, 1988. 7, 30

[26] Matthew Clegg, Jeff Edmonds, and Russell Impagliazzo. Using the Gröebner basis algorithm to find proofs of unsatisfiability. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing*, pages 174–183, 1996. 3, 4, 5, 8, 15

[27] Stephen A. Cook and Robert A. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44:36–50, 1979. 3

[28] David Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms : An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3rd edition.* Springer, 2007. 5, 15

[29] Martin Davis, George Logemann, and Donald Loveland. A machine program for theorem-proving. *Commun. ACM*, 5:394–397, July 1962. 4

[30] Martin Davis and Hilary Putnam. A computing procedure for quantification theory. *J. ACM*, 7:201–215, July 1960. 4

[31] Juan Luis Esteban, Nicola Galesi, and Jochen Messner. On the complexity of resolution with bounded conjunctions. *Theor. Comput. Sci.*, 321(2-3):347–370, 2004. 3, 4, 9, 10, 20

[32] Juan Luis Esteban and Jacobo Torán. Space bounds for resolution. *Information and Computation*, 171(1):84–97, 2001. 3, 5, 8, 11, 15

[33] Tomás Feder and Moshe Y. Vardi. The computational structure of monotone monadic snp and constraint satisfaction: A study through datalog and group theory. *SIAM J. Comput.*, 28(1):57–104, 1998. 3

[34] Yuval Filmus, Massimo Lauria, Jakob Nordström, Neil Thapen, and Noga Zewi. Space complexity in polynomial calculus. In *IEEE Conference on Computational Complexity 2012*, 2012. To appear. 1, 3, 8, 10, 11, 12, 16, 24, 26, 27, 28

[35] Nicola Galesi and Massimo Lauria. On the automatizability of polynomial calculus. *Theory of Computing Systems*, 47(2):491–506, 2010. 3, 8, 14, 33

[36] Nicola Galesi and Massimo Lauria. Optimality of size-degree trade-offs for polynomial calculus. *ACM Transactions on Computational Logic*, 12(1):??, 2010. To appear. 3, 8

[37] Russell Impagliazzo, Pavel Pudlák, and Jiří Sgall. Lower bounds for the polynomial calculus and the gröbner basis algorithm. *Computational Complexity*, 8(2):127–144, 1999. 3, 8

[38] Jan Krajíček. Propositional proof complexity i. 2009. 3

[39] Joao P. Marques-Silva and Karem A. Sakallah. Grasp—a new search algorithm for satisfiability. In *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD '96),*, pages 220–227, 1996. 4

[40] Jakob Nordström. Narrow proofs may be spacious: separating space and width in resolution. In *STOC*, pages 507–516, 2006. 3, 8

[41] Jakob Nordström and Johan Håstad. Towards an optimal separation of space and length in resolution. In *STOC*, pages 701–710, 2008. 3, 8

[42] Pavel Pudlák and Jiří Sgall. Algebraic models of computation and interpolation for algebraic proof systems. *DIMACS series in Theoretical Computer Science*, 39:279–296, 1998. 3, 8

[43] Alexander Razborov. Pseudorandom generators hard for k-dnf resolution and polynomial calculus resolution. *Manuscript availbale at author's webpage*, 2003. 3

[44] Alexander A. Razborov. Lower bounds for the polynomial calculus. *Computational Complexity*, 7(4):291–324, 1998. 3, 8

[45] J. A. Robinson. A machine-oriented logic based on the resolution principle. *J. ACM*, 12:23–41, January 1965. 3

ECCC ISSN 1433-8092

http://eccc.hpi-web.de