7. Sum of squares lower bounds for k-SAT and k-XOR Part 1/2.

http://www.csc.kth.se/~lauria/sos14/

Lecturer: Massimo Lauria

In this lecture we define the Positivstellensatz Calculus proof system and we show a lower bound for the degree needed to refute random k-XOR formulas. We define the Positivstellensatz Calculus, its subsystem Binomial Calculus and prove a lower bound for the degree of random k-XOR formulas in Binomial Calculus. The proof of why this result leads to a lower bound for Positivstellensatz is in the next lecture.

Positivstellesatz Calculus ($\mathsf{PC}_>$*)*

Positivstellensatz Calculus has been defined first in [GV01], and in this and next lecture we will see a lower bound for this proof system due to [Gri01]. The version of the lower bound that we are going to see is actually a slightly different formulation that has been reproved later in [Sch08].

Let us consider the ordered field of the reals $\mathbb{R}$, a finite set of variables X and a finite set P of polynomial equations in the ring $\mathbb{R}[X]$.

A *derivation of* $p \geq 0$ *from* P *in* $\mathsf{PC}_>$ is a sequence of polynomial equations ending with $p' = 0$ such that $p = p' + \sum_i h_i^2$, where h_i are polynomials in $\mathbb{R}[X]$. Each polynomial equation in the sequence is either from P or is the result of an inference from polynomial equations appearing previously in the sequence according to the following inference rules:

$$\frac{q = 0}{xq = 0}\, x \in X, \qquad \frac{q = 0 \quad r = 0}{\alpha q + \beta r = 0}\, \alpha, \beta \in \mathbb{R}. \tag{1}$$

A *refutation of* P *in* $\mathsf{PC}_>$ is a derivation of $-1 \geq 0$ starting from P. The *degree* of a derivation of an inequality $p \geq 0$ is the maximum among the degrees of the intermediate polynomials appearing in the derivation of p', and among the degrees of the h_i^2's.

The *Binomial Calculus* (BC) is a particular case of the previous proof system. A *BC derivation* of some binomial equation $p = 0$[1] from some set of binomial equations Q is a $PC_<$ derivation of $p \geq 0$ from Q where each intermediate polynomial equation is actually a binomial equation and the $\sum_i h_i^2$ part is 0. Essentially the binomial calculus restrict the inference rules in (1) so that the premises and the consequent are all polynomials with at most two non-zero terms.

A *refutation of* Q in BC is a derivation in BC of $\alpha - 1 = 0$ for some $\alpha \in \mathbb{R}$, $\alpha \neq 1$. The very same notion of degree of $\mathsf{PC}_>$ apply here.

[1] We recall that a binomial is sum of two terms, each of them of the form $\alpha \prod x_i$ for some $\alpha \in \mathbb{R}$ and for some subset of variables x_i from X.

Random k-XOR formulas

A *random k-XOR formula* on n variables with m *parity constraints* is a distribution on affine systems $\mod 2$. Let $X = \{x_1, \ldots, x_n\}$ be the set of variables and $m, k \in \mathbb{N}$. Repeat independently at random m times the fol-

lowing process: sample uniformly at random $b \in \{0,1\}$ and $S \subset [n]$ of size k and from them build the parity constraint $\sum_{i \in S} x_i \equiv b \pmod 2$. [2].

We can associate to a random k-XOR formula a set of polynomial equations such that the formula has a boolean solution iff the set of polynomial equations has a solution.

The encoding of a single parity constraint $\sum_{i \in S} x_i \equiv b \pmod 2$ as a set of polynomial equations in $\mathbb{R}[X]$ is the following:

$$\left\{ \prod_{i \in S} (1 - 2x_i) = (-1)^b \right\}_S \cup \{x_i^2 = x_i\}_{i \in X}. \tag{2}$$

It is clear that every satisfying assignment for these polynomials equations is a satisfying assignment for the linear system, and vice versa.

In what follow we will use another polynomial encoding of the linear system, using a different set of variables $Y = \{y_1, \ldots, y_n\}$. In this case the parity constraint $\sum_{i \in S} x_i \equiv b \pmod 2$ has a solution iff the following set of polynomial equations in $\mathbb{R}[Y]$ has a common zero:

$$\left\{ \prod_{i \in S} y_i = (-1)^b \right\} \cup \{y_i^2 = 1\}_{i \in S}. \tag{3}$$

Obviously there is an (invertible) affine transformation from $\mathbb{R}[X]$ to $\mathbb{R}[Y]$ that preserves satisfiability, and that maps one set of polynomial equations into the other: $x_i \mapsto (y_i - 1)/2$.

Notice that the degree of $\mathsf{PC}_>$ refutations of an unsatisfiable set of parity constraints does not depend on whether the encoding is the one in (2) or the one in (3), because the mapping between the two encoding is linear. Any $\mathsf{PC}_>$ refutation of (2) can be translated into one for (3) preserving the degree, and vice versa.

The following is the main theorem of this and next lecture, and is due to (Grigoriev, 2001) and (Schoenebeck, 2008).[3]

Theorem 1. *For each $k \geq 3$ and $\delta > 0$ there exists α, such that a random k-XOR formula ϕ in n variables and Δn clauses, where $\Delta \geq 1 + \frac{\ln 2}{2\delta^2}$, with high probability has the following properties:*

1. *At most $\left(\frac{1}{2} + \delta\right) \Delta n$ parity constraints of ϕ can be simultaneously satisfied,*
2. *any $\mathsf{PC}_>$ refutation of ϕ requires degree αn.*

Proof of part 1 of the Theorem. Given ϕ we proceed by applying Chernoff Bound and then union bound. Lets fix an assignment $x \in \{0,1\}^n$ and let $C_i(x)$ be the random variable that is 1 if x satisfy the i-th parity constraint in ϕ and 0 otherwise. Hence $\sum_i C_i(x)$ is the number of linear constraints of ϕ satisfied by x. Then $\mathbb{E}[C_i(x)] = \frac{1}{2}$ and by linearity $\mathbb{E}[\sum_i C_i(x)] = \frac{1}{2}\Delta n$. By Chernoff Bound[4], for any $\delta > 0$,

$$\mathbb{P}\left[\sum_i C_i(x) \geq (\frac{1}{2} + \delta)\Delta n \right] \leq e^{-2\delta^2 \Delta n}.$$

[2] A very similar process is used to build *random* k-SAT formulas: pick uniformly at random a set $S \subset [n]$ of size k and a random mapping $b : S \to \{0,1\}$. From those build the clause $\bigvee_{i \in S} x_i^{b(i)}$, where $x^1 := x$ and $x^0 := \neg x$. Repeat independently at random this process m times and take the conjunction of the clauses you get.

[3] Dima Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1):613–622, 2001; and Grant Schoenebeck. Linear level lasserre lower bounds for certain k-csps. In *Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on*, pages 593–602. IEEE, 2008

[4] We use the (standard) following form of Chernoff Bound: let $X_1, \ldots, X_m$ be independent 0-1 random variables and $X = \sum_{i \in [m]} X_i$ then for every $\lambda > 0$

$$\mathbb{P}\left[X \geq \mathbb{E}[X] + \lambda\right] \leq e^{-\frac{2\lambda^2}{m}}.$$

Hence by union bound

$$\mathbb{P}\left[\exists x \in \{0,1\}^n \ \left(\sum_i C_i(x) \geq (\frac{1}{2}+\delta)\Delta n\right)\right] \leq 2^n \cdot e^{-2\delta^2\Delta n} \leq e^{-n}.$$

The last inequality comes from the assumption that $\Delta \geq (1+\ln 2)\frac{1}{2\delta^2}$. □

Before going deep into the proof of the second part of Theorem 1 we just state and prove an interesting corollary.

Corollary 2. *For each $k \geq 3$ and $\delta > 0$, there exists an α, such that with high probability for a random k-SAT formula ϕ with Δn clauses and $\Delta \geq (1+\ln 2)\frac{1}{2\delta^2}$:*

1. *At most $\left(\frac{2^k-1}{2^k}+\delta\right)\Delta n$ clauses of ϕ can be satisfied at the same time and*
2. *any $\mathsf{PC}_>$ refutation of ϕ requires degree at least αn.*

Proof. The proof of point 1 is exactly the same of the analogous point of Theorem 1. The only difference is that the expected value of the random variable representing the number of clauses satisfied changes to $\frac{2^k-1}{2^k}\Delta n$. The rest of the calculations are exactly the same.

Regarding the second point we just observe that from the random k-XOR distribution we can derive in degree $(k+1)$ random k-SAT. For each parity constraint $\sum_{i\in S} x_i \equiv b \pmod 2$ in random k-XOR we choose uniformly at random one of the clauses derivable from that constraint.[5] Half of the possible clauses with k literals in the variables $\{x_i\}_{i\in S}$ is derivable in degree $k+1$ from $\prod_{i\in S}(1-2x_i) = (-1)^b$. Which half depends on the uniformly chosen bit b. Therefore, given the set of variables S, this process sample uniformly at random a clause over S.

The random k-CNF formula we obtain in this way has the same distribution of random k-SAT. Hence for sufficiently large n it is not possible to derive in small degree random k-SAT, otherwise random k-XOR would have small degree refutations too but this is excluded by Theorem 1. □

The previous Theorem and the Corollary show in particular that after αn steps in the Lasserre hierarchy the integrality gap is $1/2+\delta$ for Max k-XOR and $\frac{2^k-1}{2^k}+\delta$ for Max k-SAT. This means that for both of those problems the integrality gap cannot be much better than $1/2$ or $\frac{2^k-1}{2^k}$ respectively.

[5] For example consider the parity constraint $x_1+x_2+x_3 \equiv 0 \pmod 2$, that has polynomial encoding as $(1-2x_1)(1-2x_2)(1-2x_3) = 1$, that is the same of $x_1+x_2+x_3-2(x_1x_2+x_1x_3+x_2x_3)+4x_1x_2x_3 = 0$. From this, multiplying by x_1, x_2 and x_3 we can derive (in degree 4)

$$x_1 - x_1x_2 - x_1x_3 + 2x_1x_2x_3 = 0,$$
$$x_2 - x_1x_2 - x_2x_3 + 2x_1x_2x_3 = 0,$$
$$x_3 - x_3x_2 - x_1x_3 + 2x_1x_2x_3 = 0.$$

Summing all those and subtracting the initial polynomial $x_1+x_2+x_3-2(x_1x_2+x_1x_3+x_2x_3)+4x_1x_2x_3 = 0$ we get $x_1x_2x_3 = 0$, that is the encoding of $\neg x_1 \vee \neg x_2 \vee \neg x_3$.

Proof of Theorem 1 (Part 2)

As the proof is quite long, we recap briefly its high level structure:

- Observe that to prove a degree lower bound for k-XOR formulas, it is irrelevant if we choose the encoding in (2) or (3). So, to make our life easier, we choose the encoding in (3).
- Up to a factor of 2, the degree required to refute (3) in Binomial Calculus is the same as the degree needed to refute it in $\mathsf{PC}_>$. We will se this in the next Lecture.

- We prove a degree lower bound in BC for the encoding (3) of a random k-XOR over $\mathbb{R}[Y]$.

The remaining part of this lecture is devoted to proving the last point above. We premise a Lemma about the structure of random k-XOR formulas. The proof is omitted but follows immediately from Proposition 22 in (Schoenebeck, 2008)[6].

[6] Grant Schoenebeck. Linear level lasserre lower bounds for certain k-csps. In *Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on*, pages 593–602. IEEE, 2008

Lemma 3. *Given constants $k \geq 3$, $\Delta > 0$ and $\gamma \in (0, k/2)$, there exists a value $\beta > 0$, such that for a random k-XOR formula with n variables and Δn parity constraints, the following holds with high probability*

1. *for each $\phi' \subseteq \phi$ if $|\phi'| \leq \beta n$ then ϕ' is satisfiable,*
2. *for each $\phi' \subseteq \phi$ if $|\phi'| \leq \frac{2}{3}\beta n$ then there are at least $\gamma|\phi'|$ variables appearing exactly once in ϕ'.*

The previous Lemma ensures that a random k-XOR has the combinatorial properties required to prove its hardness with high probability.

Theorem 4. *Given constants $k \geq 3$, $\Delta > 0$ and $\gamma \in (0, k/2)$, there exists $\alpha > 0$, such that with high probability, for a random k-XOR formula ϕ in n variables and Δn constraints, its encoding (3) requires BC refutations of degree at least αn.*

Proof. Let B the set of all binomial equations we can derive from ϕ in Binomial Calculus. We define a measure $\mu : B \to \mathbb{R}$ as follows:[7]

$$\mu(p) := \min\left\{|\phi'| \,:\, \phi' \subseteq \phi \,\wedge\, \phi' \models p = 0\right\}. \tag{4}$$

[7] Notation $\phi' \models p$ means that the set of parity constraints ϕ' *implies* the equation $p = 0$, i.e., the satisfying assignmnets of ϕ' are also satisfying assignments of $p = 0$.

Similarly, for an assignment β and a formula ϕ, $\beta \models \phi$ means that the assignment β satisfies all constraints in ϕ.

Clearly for each binomial p appearing in the encoding of ϕ we have that $\mu(p) = 1$. Measure μ is sub-additive with respect to the inference rules in (1). Indeed it is immediate from the definition of μ that if $\{p, q\} \models r$ then $\mu(r) \leq \mu(p) + \mu(q)$.

Let us now consider a refutation π of ϕ in BC, say ending with $\eta = 1$ for some $\eta \in \mathbb{R}$, $\eta \neq 1$. By Lemma 3 we have that $\mu(\eta = 1) > \beta n$.

By the sub-additivity of μ there must be binomial equation $q = 0$ in π such that

$$\frac{1}{3}\beta n < \mu(q = 0) \leq \frac{2}{3}\beta n.$$

To see that just consider the proof proof DAG going from the axioms (all with measure μ equal 1) to the contradiction (with measure μ greater than βn). Take $q = 0$ to be a minimal equation appearing in π such that $\mu(q = 0) > \frac{1}{3}\beta n$. Binomial equation $q = 0$ must have been inferred by previous equations in the proof. By the fact that $q = 0$ is the *first* binomial in π having big μ and by sub-additivity of μ we have also the other inequality $\mu(q = 0) \leq \frac{2}{3}\beta n$.

We want now to prove that q has large degree. Let $\phi' \subseteq \phi$ such that $\phi' \models q$. By the above inequality we have that $\frac{1}{3}\beta n < |\phi'| \leq \frac{2}{3}\beta n$, hence by Lemma 3 we have at least $\gamma|\phi'|$ variables occurring exactly once in ϕ'. If we prove that those variables have to appear also in q we are done: as q is a

binomial this means that $\deg(q) \geq \gamma|\phi'|/2 \geq \frac{1}{6}\beta n$. Then the parameter α of the statement of the Theorem is just $\frac{1}{6}\beta$.

We now prove that each variable that appears once in ϕ' has to appear in q too. Suppose by contradiction there is some variable y_i appearing once in ϕ' and not appearing in q. This variable appears only in one parity constraint of ϕ', say ℓ. Consider $\bar{\phi} = \phi' \setminus \{\ell\}$. By minimality of ϕ' there exists an assignment β such that $\beta \models \bar{\phi}$ and both $\beta(\ell)$ and $\beta(q = 0)$ are false. Then just take β^* an assignment that disagree with β only on the value given to y_i. This change makes $\beta^*(q = 0) = \beta(q = 0)$ stays false, as y_i does not appear in q, but makes $\beta^*(l = 0)$ true, as flipping a single value in a parity constraint flip also the truth value of the constraint. Hence $\beta^* \models \phi$ and $\beta^*(q = 0)$ is false, in contradiction with the fact that $\phi \models q = 0$. □

References

[Gri01] Dima Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1):613–622, 2001.

[GV01] Dima Grigoriev and Nicolai Vorobjov. Complexity of null-and positivstellensatz proofs. *Annals of Pure and Applied Logic*, 113(1):153–160, 2001.

[Sch08] Grant Schoenebeck. Linear level lasserre lower bounds for certain k-csps. In *Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on*, pages 593–602. IEEE, 2008.