A Note on Semi-Algebraic Proofs and Gaussian Elimination over Prime Fields

Albert Atserias
Universitat Politècnica de Catalunya

February 13, 2015

Abstract

In this note we show that unsatisfiable systems of linear equations with a constant number of variables per equation over prime finite fields have polynomial-size constant-degree semi-algebraic proofs of unsatisfiability. These are proofs that manipulate polynomial inequalities over the reals with variables ranging in $\{0, 1\}$. This upper bound is to be put in contrast with the known fact that, for certain explicit systems of linear equations over the two-element field, such refutations require linear degree and exponential size if they are restricted to so-called static semi-algebraic proofs, and even tree-like semi-algebraic and sums-of-squares proofs. Our upper bound is a more or less direct translation of an argument due to Grigoriev, Hirsch and Pasechnik (Moscow Mathematical Journal, 2002) who did it for a family of linear systems of interest in propositional proof complexity. We point out that their method is more general and can be thought of as simulating Gaussian elimination.

1 Semi-algebraic proofs

The proof system we consider is inspired by the Sherali-Adams and Lovász-Schrijver lift-and-project methods for combinatorial optimization [8, 4], seen as proof systems for deriving polynomial inequalities (see also [6, 3, 5]). In addition to the axioms

$$x_i \geq 0 \qquad 1 - x_i \geq 0 \qquad x_i^2 - x_i \geq 0 \qquad x_i - x_i^2 \geq 0$$

for formal variables $x_1, \ldots, x_n$, it has the following inference rules:

$$\frac{P(\mathbf{x}) \geq 0 \qquad Q(\mathbf{x}) \geq 0}{a \cdot P(\mathbf{x}) + b \cdot Q(\mathbf{x}) \geq 0} \qquad \frac{P(\mathbf{x}) \geq 0}{P(\mathbf{x}) \cdot x_i \geq 0} \qquad \frac{P(\mathbf{x}) \geq 0}{P(\mathbf{x}) \cdot (1 - x_i) \geq 0}$$

where $P(\mathbf{x})$ and $Q(\mathbf{x})$ are polynomials with rational coefficients and variables within $\mathbf{x} = (x_1, \ldots, x_n)$, and a and b are non-negative rational numbers. The first rule is called *positive linear combination* and the second and third rules are called *multiplication* or *lifting rules*.

It follows from [4] that if a system of linear inequalities over the reals in n variables does not have any solution in $\{0,1\}^n$, then the trivial contradiction $-1 \geq 0$ can be derived from the given inequalities, even if all polynomials are restricted to total degree two. In general, the length of such a proof could be exponential in a polynomial in n, but the shortest such proof is never worse than that. Here, by length we mean the number of derived inequalities. This and other complexity measures for proofs are defined next.

The *degree* of a proof is the maximum of the total degrees of the polynomials that appear in it. The *length* of a proof is the number of inferences. The *size* of a proof is the sum of the sizes of the polynomials that appear in it, where the size of a polynomial is the sum of the degrees of its monomials. A proof is *tree-like* if every derived inequality is used at most once as the hypothesis of another rule, i.e. the shape of the proof is a tree, with the hypotheses and the axioms at the leaves and the conclusion at the root. A *refutation* is a proof of $-1 \geq 0$. When we write an inequality $P(\mathbf{x}) \geq Q(\mathbf{x})$ what we really mean is $P(\mathbf{x}) - Q(\mathbf{x}) \geq 0$. Similarly, when we write an equation $P(\mathbf{x}) = Q(\mathbf{x})$ what we really mean is the set of the two inequalities $P(\mathbf{x}) - Q(\mathbf{x}) \geq 0$ and $Q(\mathbf{x}) - P(\mathbf{x}) \geq 0$.

2 Some facts about semi-algebraic proofs

For every linear form $L(\mathbf{x}) = \sum_{i=1}^{n} a_i x_i$ with rational coefficients and every integer c, let $D_c(L(\mathbf{x}))$ be the quadratic polynomial $(L(\mathbf{x}) - c) \cdot (L(\mathbf{x}) - c + 1)$. In words, the inequality $D_c(L(\mathbf{x})) \geq 0$ states that $L(\mathbf{x})$ does not fall in the open interval $(c-1, c)$. Such statements have short proofs of low degree:

Lemma 1 (Grigoriev, Hirsch, and Pasechnik [3])**.** *For every integer c and for every linear form $L(\mathbf{x}) = \sum_{i=1}^{n} a_i x_i$ with integer coefficients $a_1, \ldots, a_n$, the inequality $D_c(L(\mathbf{x})) \geq 0$ has a tree-like proof of length polynomial in $\max\{|a_i| : i = 1, \ldots, n\}$ and n, and degree at most 3.*

The next lemma states that polynomial equalities can be freely substituted. A similar statement appears in [3][Lemma 5.2]; our statement is slightly stronger.

Lemma 2. *Let $P(\mathbf{x})$, $Q(\mathbf{x})$, and $R(\mathbf{x}, y)$ be polynomials with variables as indicated, and let d be the degree of y in $R(\mathbf{x}, y)$. The equation $R(\mathbf{x}, P(\mathbf{x})) = R(\mathbf{x}, Q(\mathbf{x}))$ has a proof from $P(\mathbf{x}) = Q(\mathbf{x})$ of length bounded by a degree-d polynomial in the sizes of $P(\mathbf{x})$, $Q(\mathbf{x})$ and $R(\mathbf{x}, y)$, and degree at most linear in the degree of $R(\mathbf{x}, y)$ and d times the degrees of $P(\mathbf{x})$ and $Q(\mathbf{x})$.*

Proof. It suffices to prove the statement when $R(\mathbf{x}, y)$ is linear in y; the general statement in which y has degree $d \geq 1$ in $R(\mathbf{x}, y)$ follows from iterating the lemma on the polynomial $R'(\mathbf{x}, y_1, \ldots, y_d)$ obtained from $R(\mathbf{x}, y)$ by replacing each y^s by $\prod_{i=1}^{s} y_i$. Write every monomial in $R(\mathbf{x}, y)$ in the form $y \cdot M(\mathbf{x})$, where $M(\mathbf{x})$ is a monomial without y. The equality $P(\mathbf{x}) \cdot M(\mathbf{x}) = Q(\mathbf{x}) \cdot M(\mathbf{x})$ follows at once from $P(\mathbf{x}) = Q(\mathbf{x})$ by the multiplication rule. Adding up over all monomials of $R(\mathbf{x}, y)$ we get the result. To see the bound on the size note that, in case $R(\mathbf{x}, y)$ is linear in y, the size of $R(\mathbf{x}, P(\mathbf{x}))$ is bounded by the product of the sizes of $R(\mathbf{x}, y)$ and $P(\mathbf{x})$, and similarly for $R(\mathbf{x}, Q(\mathbf{x}))$. Iterating d times to handle the general $d \geq 1$ case we get the degree-d polynomial bound on the length of the proof. □

3 Two-element field

We identify the elements of the two-element field $\mathbb{F}_2$ with $\{0,1\}$. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be formal variables ranging over $\mathbb{F}_2$ or $\mathbb{Q}$, depending on the context. For every linear equation of the form $\mathbf{a}^{\mathrm{T}}\mathbf{x} = b$, where $\mathbf{a} \in \mathbb{F}_2^n$ and $b \in \mathbb{F}_2$, let $\mathcal{S}(\mathbf{a}, b)$ be the system of linear inequalities

$$\sum_{i \in T}(1 - x_i) + \sum_{i \in I \setminus T} x_i \geq 1 \qquad \text{for all } T \subseteq I \text{ such that } |T| \equiv 1 - b \mod 2,$$

where $I = \mathrm{supp}(\mathbf{a}) := \{i \in [n] : a_i \neq 0\}$ and $[n] := \{1, \ldots, n\}$. Note that $\mathcal{S}(\mathbf{a}, b)$ has exactly $2^{|I|-1}$ inequalities, and that it is satisfied in $\mathbb{Q}$ by a $\{0,1\}$-assignment to the $\mathbf{x}$-variables if and only if $\mathbf{a}^{\mathrm{T}}\mathbf{x} = b$ is satisfied in $\mathbb{F}_2$ by the same assignment. For a system of m linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$ as above, let $\mathcal{S}(\mathbf{A}, \mathbf{b}) := \bigcup_{i=1}^{m} \mathcal{S}(\mathbf{a}_i, b_i)$ as $\mathbf{a}_i$ ranges over the rows of $\mathbf{A}$ and b_i ranges over the components of $\mathbf{b}$. Note that this system has at most $m2^w$ inequalities, where w is the maximum number of non-zero components in the rows of $\mathbf{A}$.

For an equation of the form $\mathbf{a}^{\mathrm{T}}\mathbf{x} = b$ as above, an alternative way of writing the system $\mathcal{S}(\mathbf{a}, b)$ is by imposing the system of polynomial equalities

$$\prod_{i \in T} x_i \cdot \prod_{i \in I \setminus T} (1 - x_i) = 0 \qquad \text{for all } T \subseteq I \text{ such that } |T| \equiv 1 - b \mod 2.$$

In the following, for $I \subseteq [n]$ and $T \subseteq I$, let $M_T^I(\mathbf{x}) := \prod_{i \in T} x_i \prod_{i \in I \setminus T}(1 - x_i)$. Such polynomials are called *extended monomials*. We start by noting that

$$\sum_{T \subseteq I} M_T^I(\mathbf{x}) = \prod_{i \in I}(x_i + 1 - x_i) = 1. \tag{1}$$

We continue relating the two forms of expressing the equation $\mathbf{a}^{\mathrm{T}}\mathbf{x} = b$:

Lemma 3 (Grigoriev, Hirsch, and Pasechnik [3])**.** *Let $\mathbf{a} \in \{0,1\}^n$ and $b \in \{0,1\}$, and let $I = \mathrm{supp}(\mathbf{a})$. For every $T \subseteq I$ such that $|T| \equiv 1 - b \bmod 2$, the equation $M_T^I(\mathbf{x}) = 0$ has a tree-like proof from $S(\mathbf{a}, b)$ of length linear in $|I|$, and degree at most $|I|$.*

Proof. Let $t = |T|$ and assume without loss of generality that $T = \{1, \ldots, t\}$ and $I \setminus T = \{t + 1, \ldots, s\}$. First multiply $\sum_{i=1}^{t}(1 - x_i) + \sum_{i=t+1}^{s} x_i \geq 1$ by x_1. Then use $x_1 - x_1^2 = 0$ to get rid of the term $(1 - x_1) \cdot x_1$ on the left-hand side. Repeat for $x_2, \ldots, x_t$ to get $\sum_{i=t+1}^{s} x_i \cdot \prod_{j=1}^{t} x_j \geq \prod_{j=1}^{t} x_j$. From here, first multiply by $1 - x_{t+1}$ and then use $x_{t+1} - x_{t+1}^2 = 0$ to get rid of $(1 - x_{t+1}) \cdot x_{t+1} \cdot \prod_{j=1}^{t}(1 - x_j)$ on the left-hand side. Repeat for $x_{t+2}, \ldots, x_s$ to get $0 \geq \prod_{j=1}^{t} x_i \cdot \prod_{j=t+1}^{s}(1 - x_j)$. The converse inequality has a direct proof not even using any of the axioms in $\mathcal{S}(\mathbf{a}, b)$. □

The last lemma we need also refers to extended monomials:

Lemma 4. *Let $T \subseteq I \subseteq [n]$. Then the equation $\left(\sum_{i \in I} x_i - |T|\right) \cdot M_T^I = 0$ has a tree-like proof of length linear in $|I|$, and degree at most $|I| + 1$.*

Proof. Write M for M_T^I. For every $i \in I \setminus T$, using $x_i \cdot (1 - x_i) = x_i - x_i^2 = 0$ we get $x_i \cdot M = 0$. For every $i \in T$, using $x_i^2 - x_i = 0$ we get $x_i \cdot M = M$. Adding up we get $\sum_{i \in I} x_i \cdot M = |T| \cdot M$. □

Theorem 1. *Let* $\mathbf{A} \in \{0,1\}^{m \times n}$ *and* $\mathbf{b} \in \{0,1\}^m$. *If* $\mathbf{Ax} = \mathbf{b}$ *is unsatisfiable in* $\mathbb{F}_2$, *then* $S(\mathbf{A}, \mathbf{b})$ *has a (not necessarily tree-like) refutation of size polynomial in* n *and* 2^w, *and degree linear in* w, *where* w *is the maximum number of non-zero components in any of the rows of* $\mathbf{A}$.

Proof. Let $\mathbf{a}_1, \ldots, \mathbf{a}_m$ be the rows of $\mathbf{A}$, with $\mathbf{a}_j = (a_{j,1}, \ldots, a_{j,n})$. Assume $\mathbf{Ax} = \mathbf{b}$ is unsatisfiable in $\mathbb{F}_2$. Then the $\mathbb{F}_2$-rank of the matrix $[\mathbf{A} \mid \mathbf{b}]$ is bigger than the rank of $\mathbf{A}$. This means that there exists a subset of rows J such that $|J| \leq n$ and $\sum_{j \in J} \mathbf{a}_j = \mathbf{0}$ and $\sum_{j \in J} b_j = 1$ with arithmetic in $\mathbb{F}_2$. In order to simplify notation, we assume without loss of generality that $J = \{1, \ldots, |J|\}$.

For every $k \in \{0, \ldots, |J|\}$, let

$$L_k(\mathbf{x}) := \frac{1}{2} \left(\sum_{j=1}^{k} \sum_{i=1}^{n} a_{j,i} x_i + \sum_{j=k+1}^{|J|} b_j \right).$$

We provide proofs of $D_c(L_k(\mathbf{x})) \geq 0$ for every $c \in R_k := \{0, \ldots, (k+1) \cdot n\}$ by reverse induction on $k \in \{0, \ldots, |J|\}$.

The base case $k = |J|$ is a special case of Lemma 1. To see why note that the condition $\sum_{j \in J} \mathbf{a}_j = \mathbf{0}$ means that if arithmetic is in $\mathbb{Q}$ then $\sum_{j \in J} a_{j,i}$ is an even natural number for every $i \in [n]$. But then all the coefficients of

$$L_{|J|}(\mathbf{x}) = \frac{1}{2} \sum_{j=1}^{|J|} \sum_{i=1}^{n} a_{j,i} x_i = \sum_{i=1}^{n} \left(\frac{1}{2} \sum_{j=1}^{|J|} a_{j,i} \right) x_i$$

are integers. Hence Lemma 1 applies.

Suppose now that $0 \leq k \leq |J| - 1$ and that we have a proof of $D_d(L_{k+1}(\mathbf{x})) \geq 0$ available for every $d \in R_{k+1}$. Fix $c \in R_k$; our immediate goal is to give a proof of $D_c(L_k(\mathbf{x})) \geq 0$. As k is fixed, write $L(\mathbf{x})$ instead of $L_{k+1}(\mathbf{x})$, and let the equation $\mathbf{a}_{k+1}^{\mathrm{T}} \mathbf{x} = b_{k+1}$ be written as $\sum_{i \in I} x_i = b$, where $I = \mathrm{supp}(\mathbf{a})$. Note that $L(\mathbf{x}) = L_k(\mathbf{x}) + \frac{1}{2} \cdot \ell(\mathbf{x})$ where $\ell(\mathbf{x}) := -b + \sum_{i \in I} x_i$. Fix $T \subseteq I$ such that $|T| \equiv b \bmod 2$, and let $d = c + \frac{t-b}{2}$ where $t = |T|$. Note that $d \in R_{k+1}$ as $c \in R_k$ and $0 \leq t \leq n$ and $0 \leq b \leq 1$ are such that $t - b$ is even. Multiplying $D_d(L(\mathbf{x})) \geq 0$ by $M_T^I(\mathbf{x})$ we get

$$(L(\mathbf{x}) - d) \cdot (L(\mathbf{x}) - d + 1) \cdot M_T^I(\mathbf{x}) \geq 0. \tag{2}$$

Replacing $L(\mathbf{x}) = L_k(\mathbf{x}) + \frac{1}{2} \cdot \ell(\mathbf{x})$ in the factor $(L(\mathbf{x}) - d)$ and recalling $d = c + \frac{t-b}{2}$, this inequality can be written as

$$(L_k(\mathbf{x}) - c) \cdot (L(\mathbf{x}) - d + 1) \cdot M_T^I(\mathbf{x}) + (L(\mathbf{x}) - d + 1) \cdot \tfrac{1}{2} \cdot A(\mathbf{x}) \geq 0 \tag{3}$$

where $A(\mathbf{x}) := (\ell(\mathbf{x}) + b - t) \cdot M_T^I(\mathbf{x})$. By Lemma 4 we have a proof of $A(\mathbf{x}) = 0$, and hence of $(L(\mathbf{x}) - d + 1) \cdot \frac{1}{2} \cdot A(\mathbf{x}) = 0$. Composing with (3) we get a proof of

$$(L_k(\mathbf{x}) - c) \cdot (L(\mathbf{x}) - d + 1) \cdot M_T^I(\mathbf{x}) \geq 0. \tag{4}$$

The same argument applied to the factor $(L(\mathbf{x}) - d + 1)$ of this inequality gives

$$(L_k(\mathbf{x}) - c) \cdot (L_k(\mathbf{x}) - c + 1) \cdot M_T^I(\mathbf{x}) \geq 0. \tag{5}$$

This is precisely $D_c(L_k(\mathbf{x})) \cdot M_T^I(\mathbf{x}) \geq 0$. Adding up over all $T \subseteq I$ with $|T| \equiv b \bmod 2$ we get

$$D_c(L_k(\mathbf{x})) \cdot \sum_{\substack{T \subseteq I \\ |T| \equiv b}} M_T^I(\mathbf{x}) \geq 0. \tag{6}$$

By Lemma 3, from the inequalities for $\sum_{i \in I} x_i = b$ we get proofs of $M_T^I(\mathbf{x}) = 0$ for every $T \subseteq I$ such that $|T| \equiv 1 - b \bmod 2$. But then also of $D_c(L_k(\mathbf{x})) \cdot M_T^I(\mathbf{x}) = 0$ for every such T. Adding up and composing with (6) we get

$$D_c(L_k(\mathbf{x})) \cdot \sum_{T \subseteq I} M_T^I(\mathbf{x}) \geq 0$$

which is precisely $D_c(L_k(\mathbf{x})) \geq 0$ because $\sum_{T \subseteq I} M_T^I(\mathbf{x}) = 1$ by (1).

At this point we proved $D_c(L_0(\mathbf{x})) \geq 0$ for every $c \in R_0 = \{0, \ldots, n\}$. Recall now that $\sum_{j=1}^{|J|} b_j$ is odd, say $2q+1$, and at most n. In particular $q+1$ belongs to R_0 and $L_0(\mathbf{x}) = q + \frac{1}{2}$. Thus we have a proof of $D_{q+1}(L_0(\mathbf{x})) \geq 0$ where $D_{q+1}(L_0(\mathbf{x})) = -\frac{1}{2} \cdot \frac{1}{2} = -\frac{1}{4}$. Multiplying by 4 we get $-1 \geq 0$. □

4 Prime fields

Let p be a prime. We identify the elements of the field with p elements $\mathbb{F}_p$ with the integers $\{0, \ldots, p-1\}$. For every $i \in [n]$, let $\mathbf{x}_i = (x_i(0), \ldots, x_i(p-1))$ be formal variables ranging over $\mathbb{Q}$, and let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$. By imposing the constraints

$$\begin{array}{ll} x_i(0) + \cdots + x_i(p-1) = 1 & \text{for all } i \in [n] \\ x_i(j) - x_i(j)^2 = 0 & \text{for all } i \in [n] \text{ and } j \in \{0, \ldots, p-1\} \end{array}$$

each $\mathbf{x}_i$ is the indicator vector of some value in $\mathbb{F}_p$. Consequently we think of $\mathbf{x}_i$ as a formal variable ranging over $\mathbb{F}_p$. In the following, let $\mathcal{Z}$ be the set of equations $x_i(0) + \cdots + x_i(p-1) = 1$ as i ranges over $[n]$.

For every linear equation of the type $\mathbf{a}^{\mathrm{T}}\mathbf{x} = b$ where $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{F}_p^n$ and $b \in \mathbb{F}_p$, let $\mathcal{S}(\mathbf{a}, b)$ be the system of linear inequalities

$$\textstyle\sum_{i \in I}(1 - x_i(z_i)) \geq 1 \quad \text{for all } \mathbf{z} \in \mathbb{F}_p^I \text{ such that } \sum_{i \in I} a_i z_i \not\equiv b \mod p$$

where $I = \mathrm{supp}(\mathbf{a}) := \{i \in [n] : a_i \neq 0\}$. Observe that these are at most $p^{|I|}$ different inequalities. For a system $\mathbf{Ax} = \mathbf{b}$ of m linear equations as above, let $\mathcal{S}(\mathbf{A}, \mathbf{b}) := \bigcup_{i=1}^{m} \mathcal{S}(\mathbf{a}_i, b_i)$ as $\mathbf{a}_i$ ranges over the rows of $\mathbf{A}$ and b_i ranges over the components of $\mathbf{b}$.

In the following, for $I \subseteq [n]$ and $\mathbf{z} \in \mathbb{F}_p^I$, let

$$M_{\mathbf{z}}(\mathbf{x}) := \prod_{i \in I} \left(x_i(z_i) \cdot \prod_{\substack{\ell=0: \\ \ell \neq z_i}}^{p-1} (1 - x_i(\ell)) \right).$$

We start with the analogue of (1). This time we need to assume some axioms.

Lemma 5. *Let $I \subseteq [n]$. The equation $\sum_{\mathbf{z} \in \mathbb{F}_p^I} M_{\mathbf{z}}(\mathbf{x}) = 1$ has a tree-like proof from $\mathcal{Z}$ of length polynomial in $|I|$ and p, and degree linear in $|I|p$.*

Proof. Using $\sum_{\ell=0}^{p-1} x_i(\ell) = 1$ and $x_i(z_i) - x_i(z_i)^2 = 0$ we have

$$1 = \prod_{i \in I} \sum_{\ell=0}^{p-1} x_i(\ell) = \sum_{\mathbf{z} \in \mathbb{F}_p^k} \prod_{i \in I} x_i(z_i) = \sum_{\mathbf{z} \in \mathbb{F}_p^k} \prod_{i \in I} x_i(z_i)^p = \sum_{\mathbf{z} \in \mathbb{F}_p^k} \prod_{i \in I} \left(x_i(z_i) \cdot \prod_{\substack{\ell=0: \\ \ell \neq z_i}}^{p-1} (1 - x_i(\ell)) \right).$$

Use Lemma 2 to get an actual proof. □

Lemma 6. *Let $\mathbf{a} \in \mathbb{F}_p^n$ and $b \in \mathbb{F}_p$. For every $\mathbf{z} \in \mathbb{F}_p^I$ such that $\sum_{i \in I} a_i z_i \not\equiv b \bmod p$, where $I = \mathrm{supp}(\mathbf{a})$, the equation $M_{\mathbf{z}}(\mathbf{x}) = 0$ has a tree-like proof from $\mathcal{S}(\mathbf{a}, b) \cup \mathcal{Z}$ of length polynomial in $|I|$ and p, and degree at most $|I|p$.*

Proof. Without loss of generality, assume $I = \{1, \ldots, k\}$. Start at $\sum_{i=1}^{k} (1 - x_i(z_i)) \geq 1$ from $\mathcal{S}(\mathbf{a}, b)$, multiply by $x_1(z_1)$, and use $x_1(z_1) \cdot (1 - x_1(z_1)) = x_1(z_1) - x_1(z_1)^2 = 0$ to get $\sum_{i=2}^{k} (1 - x_i(z_i)) \geq x_1(z_1)$. Repeat with $x_2(z_2), \ldots, x_k(z_k)$ to get $0 \geq \prod_{i=1}^{k} x_i(z_i)$. Multiply by $(1 - x_i(\ell))$ for every $i \in I$ and $\ell \in \{0, \ldots, p-1\} \setminus \{z_i\}$ to get $0 \geq M_{\mathbf{z}}(\mathbf{x})$. The reverse inequality has a direct proof not even using any of the axioms from $\mathcal{S}(\mathbf{a}, b) \cup \mathcal{Z}$. □

The following is the analogue of Lemma 4:

Lemma 7. *Let $I \subseteq [n]$, $\mathbf{a} \in \mathbb{F}_p^I$, and $\mathbf{z} \in \mathbb{F}_p^I$. Then the equation*

$$\left(\sum_{i \in I} a_i \sum_{\ell=0}^{p-1} \ell x_i(\ell) - \sum_{i \in I} a_i z_i \right) \cdot M_{\mathbf{z}}(\mathbf{x}) = 0$$

has a tree-like proof of length polynomial in $|I|$ and p, and degree at most $|I|p + 1$.

Proof. Write M for $M_{\mathbf{z}}(\mathbf{x})$. For every $i \in I$ and every $\ell \in \{0, \ldots, p-1\} \setminus \{z_i\}$, using $x_i(\ell) \cdot (1 - x_i(\ell)) = x_i(\ell) - x_i(\ell)^2 = 0$ we get $a_i \ell x_i(\ell) \cdot M = 0$. For every $i \in I$, using $x_i(z_i)^2 - x_i(z_i) = 0$ we get $a_i z_i x_i(z_i) \cdot M = a_i z_i \cdot M$. Adding up we get what we want. □

Theorem 2. *Let $\mathbf{A} \in \mathbb{F}_p^{m\times n}$ and $\mathbf{b} \in \mathbb{F}_p^m$. If $\mathbf{Ax} = \mathbf{b}$ is unsatisfiable in $\mathbb{F}_p$, then $\mathcal{S}(\mathbf{A}, \mathbf{b}) \cup \mathcal{Z}$ has a (not necessarily tree-like) refutation of size polynomial in n and p^w, and degree linear in w, where w is the maximum number of non-zero components of the rows of $\mathbf{A}$.*

Proof. Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{F}_p^n$ be the rows of $\mathbf{A}$. Assume $\mathbf{Ax} = \mathbf{b}$ is unsatisfiable in $\mathbb{F}_p$. Then the $\mathbb{F}_p$-rank of the matrix $[\mathbf{A} \mid \mathbf{b}]$ is bigger than the rank of $\mathbf{A}$, which means that there exists a subset of rows $J \subseteq [m]$ and a vector of multipliers $\mathbf{y} = (y_j : j \in J) \in \mathbb{F}_p^J$ such that $|J| \leq n$ and $\sum_{j\in J} y_j \mathbf{a}_j = \mathbf{0}$ and $\sum_{j\in J} y_j b_j = 1$ with arithmetic in $\mathbb{F}_p$. In order to simplify notation, we assume without loss of generality that $J = \{1, \ldots, |J|\}$.

For every $k \in \{0, \ldots, |J|\}$, let

$$L_k(\mathbf{x}) := \frac{1}{p}\left(\sum_{j=1}^{k} y_j \sum_{i=1}^{n} a_{j,i} X_i + \sum_{j=k+1}^{|J|} y_j b_j\right),$$

where $X_i := \sum_{\ell=0}^{p-1} \ell \cdot x_i(\ell)$. We provide proofs of $D_c(L_k(\mathbf{x})) \geq 0$ for every $c \in R_k := \{0, \ldots, (k+1)p^2 n\}$ by reverse induction on $k \in \{0, \ldots, |J|\}$.

The base case $k = |J|$ is a special case of Lemma 1. To see why note that the condition $\sum_{j\in J} y_j \mathbf{a}_j = \mathbf{0}$ means that if arithmetic is in $\mathbb{Q}$ then $\sum_{j\in J} y_j a_{j,i}$ is an integer multiple of p for every $i \in [n]$. But then all the coefficients of

$$L_{|J|}(\mathbf{x}) = \frac{1}{p}\sum_{j=1}^{|J|} y_j \sum_{i=1}^{n} a_{j,i} X_i = \sum_{i=1}^{n}\sum_{\ell=0}^{p-1}\left(\frac{1}{p}\sum_{j=1}^{|J|} y_j a_{j,i}\right) x_i(\ell)$$

are integers. Hence Lemma 1 applies.

Suppose now that $0 \leq k \leq |J| - 1$ and that we have a proof of $D_d(L_{k+1}(\mathbf{x})) \geq 0$ available for every $d \in R_{k+1}$. Fix $c \in R_k$; our immediate goal is to give a proof of $D_c(L_k(\mathbf{x})) \geq 0$. As k is fixed, write $L(\mathbf{x})$ instead of $L_{k+1}(\mathbf{x})$, and also y instead of y_{k+1}, and let the equation $\mathbf{a}_{k+1}^{\mathrm{T}}\mathbf{x} = b_{k+1}$ be written as $\sum_{i\in I} a_i \mathbf{x}_i = b$, where $I = \mathrm{supp}(\mathbf{a})$. Note that $L(\mathbf{x}) = L_k(\mathbf{x}) + \frac{y}{p} \cdot \ell(\mathbf{x})$ where $\ell(\mathbf{x}) := -b + \sum_{i\in I} a_i X_i$.

Split $\mathbb{F}_p^I$ into $Z := \{\mathbf{z} \in \mathbb{F}_p^I : \sum_{i\in I} a_i z_i \equiv b \bmod p\}$ and $\overline{Z} := \mathbb{F}_p^I \setminus Z$. Fix $\mathbf{z} \in Z$ and let $t := \sum_{i\in I} a_i z_i$ with arithmetic in $\mathbb{Q}$. Let $d = c + \frac{(t-b)y}{p}$ and note that $d \in R_{k+1}$ as $c \in R_k$, $0 \leq t \leq p^2 n$, $0 \leq y \leq p-1$, and $0 \leq b \leq p-1$ are such that $t - b$ is an integer multiple of p. Multiplying $D_d(L(\mathbf{x})) \geq 0$ by $M_{\mathbf{z}}(\mathbf{x})$ we get

$$(L(\mathbf{x}) - d) \cdot (L(\mathbf{x}) - d + 1) \cdot M_{\mathbf{z}}(\mathbf{x}) \geq 0. \tag{7}$$

Replacing $L(\mathbf{x}) = L_k(\mathbf{x}) + \frac{y}{p} \cdot \ell(\mathbf{x})$ in the factor $(L(\mathbf{x}) - d)$ and recalling $d = c + \frac{(t-b)y}{p}$, this inequality can be written as

$$(L_k(\mathbf{x}) - c) \cdot (L(\mathbf{x}) - d + 1) \cdot M_{\mathbf{z}}(\mathbf{x}) + (L(\mathbf{x}) - d + 1) \cdot \tfrac{y}{p} \cdot A(\mathbf{x}) \geq 0 \tag{8}$$

where $A(\mathbf{x}) := (\ell(\mathbf{x}) + b - t) \cdot M_{\mathbf{z}}(\mathbf{x})$. By Lemma 7 we have a proof of $A(\mathbf{x}) = 0$, and hence of $(L(\mathbf{x}) - d + 1) \cdot \frac{y}{p} \cdot A(\mathbf{x}) = 0$. Composing with (8) we get a proof of

$$(L_k(\mathbf{x}) - c) \cdot (L(\mathbf{x}) - d + 1) \cdot M_{\mathbf{z}}(\mathbf{x}) \geq 0. \tag{9}$$

The same argument applied to the factor $(L(\mathbf{x}) - d + 1)$ of this inequality gives

$$(L_k(\mathbf{x}) - c) \cdot (L_k(\mathbf{x}) - c + 1) \cdot M_{\mathbf{z}}(\mathbf{x}) \geq 0. \tag{10}$$

This is precisely $D_c(L_k(\mathbf{x})) \cdot M_{\mathbf{z}}(\mathbf{x}) \geq 0$. Adding over Z we get

$$D_c(L_k(\mathbf{x})) \cdot \sum_{\mathbf{z} \in Z} M_{\mathbf{z}}(\mathbf{x}) \geq 0. \tag{11}$$

By Lemma 6, from the inequalities in $\mathcal{S}(\mathbf{a}_{k+1}, b_{k+1})$ we get proofs of $M_{\mathbf{z}}(\mathbf{x}) = 0$ for every $\mathbf{z} \in \overline{Z}$. But then also $D_c(L_k(\mathbf{x})) \cdot M_{\mathbf{z}}(\mathbf{x}) = 0$ for every such $\mathbf{z}$. Adding up and composing with (11) we get

$$D_c(L_k(\mathbf{x})) \cdot \sum_{\mathbf{z} \in Z \cup \overline{Z}} M_{\mathbf{z}}(\mathbf{x}) \geq 0$$

which is precisely $D_c(L_k(\mathbf{x})) \geq 0$ because $\sum_{\mathbf{z} \in \mathbb{F}_p^I} M_{\mathbf{z}}(\mathbf{x}) = 1$ by Lemma 5.

At this point we proved $D_c(L_0(\mathbf{x})) \geq 0$ for every $c \in R_0 = \{0, \ldots, p^2 n\}$. Recall now that $\sum_{j=1}^{|J|} y_j b_j$ is congruent to 1 mod p, say $pq + 1$, and smaller than $p^2 n$. In particular $q + 1$ belongs to R_0 and $L_0(\mathbf{x}) = q + \frac{1}{p}$. Thus we have a proof of $D_{q+1}(L_0(\mathbf{x})) \geq 0$ where $D_{q+1}(L_0(\mathbf{x})) = (\frac{1}{p} - 1) \cdot \frac{1}{p} = \frac{1-p}{p^2}$. Multiplying by $\frac{p^2}{p-1} > 0$ we get $-1 \geq 0$. □

5 Closing remarks

The upper bound in Theorem 1 is to be put in contrast with the lower bounds proved by Grigoriev [2] as rediscovered by Schoenebeck [7]. Those lower bounds hold for *static* semi-algebraic proofs, and even static *sums-of-squares* (SOS) proofs. In short, the static version of semi-algebraic proofs can be formulated as the restriction to proofs in which all applications of the multiplication rules must precede all applications of the positive linear combination rule. Static sums-of-squares proofs would be the same with the addition of axioms of the form $\sum_{i=1}^{m} P_i(\mathbf{x})^2 \geq 0$ for arbitrary polynomials $P_1, \ldots, P_m$. See [1] and subsequent work for some recent exciting applications of static sums-of-squares proofs to combinatorial optimization.

The above-mentioned lower bounds show that there exist explicit systems of linear equations $\mathbf{Ax} = \mathbf{b}$ with n variables and three variables per equation, that are unsatisfiable over the two-element field but for which any static semi-algebraic or sums-of-squares refutation must have degree $\Omega(n)$. This holds with respect to the same representation of linear systems that we use here. It can also be seen that their proof also yields an exponential $2^{\Omega(n)}$ lower bound in size and length. More strongly, from the size-degree trade-off results in [5] for tree-like proofs, such lower bounds on degree and size apply also to the tree-like restrictions of semi-algebraic proofs and sums-of-squares proofs. We note that static proofs may be assumed tree-like without any significant loss in degree, size or length, so this is a strengthening. Theorem 1 shows that such lower bounds do not extend to general, i.e. dag-like, semi-algebraic proofs.

References

[1] B. Barak, F. G. S. L. Brandão, A. W. Harrow, J. A. Kelner, D. Steurer, and Y. Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *44th Annual ACM Symposium on the Theory of Computing*, pages 307–326, 2012.

[2] D. Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1–2):613–622, 2001.

[3] D. Grigoriev, E. A. Hirsch, and D. V. Pasechnik. Complexity of semi-algebraic proofs. *Moscow Mathematical Journal*, 4(2):647–679, 2002.

[4] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, 1991.

[5] T. Pitassi and N. Segerlind. Exponential Lower Bounds and Integrality Gaps for Tree-like Lovász-Schirjver Procedures. *SIAM Journal on Computing*, 41(1):128–159, 2012.

[6] P. Pudlák. On the complexity of the propositional calculus. In *Sets and Proofs, Invited Papers from Logic Colloquium '97*, pages 197–218. Cambridge University Press, 1999.

[7] G. Schoenebeck. Linear Level Lasserre Lower Bounds for Certain k-CSPs. In *49th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 593–602, 2008.

[8] H. D. Sherali and W. P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3(3):411–430, 1990.