Tseitin's Tautologies and Lower Bounds for Nullstellensatz Proofs

D. Grigoriev

Departments of Mathematics and Computer Science
The Pennsylvania State University
University Park, PA 16802
dima@cse.psu.edu

We use the known linear lower bound for Tseitin's tautologies for establishing linear lower bounds on the degree of Nullstellensatz proofs (in the usual boolean setting) for explicitly constructed systems of polynomials of a constant (in our construction 6) degree. It holds over any field of characteristic distinct from 2. Previously, a linear lower bound was proved [14] for an explicitly constructed system of polynomials of a logarithmic degree.

Introduction.

In the theory of effective Nullstellensatz the double exponential upper bound \(d^{O(n)}\) [15] on the degrees for general ideals and \(d^{O(n)}\) [2], [8] for the unit ideal are well known (here \(d\) is the degree and \(n\) is the number of variables of the input polynomials). These bounds are known to be sharp due to [13] for the first bound and for the second bound due to the example of Lazard-Mora-Philippon (see [2]).

In the proof system theory (see e.g. [3], [7], [6], [4], [14], [10]) a similar question is studied when among input polynomials \(f_1, \ldots, f_k \in F[X_1, \ldots, X_n]\) necessarily the polynomials \(X_i^2 - X_i, 1 \leq i \leq n\) appear (let us call such a system of input polynomials a boolean system). Then the known methods [13], [5] for obtaining lower bounds on the degrees of \(g_1, \ldots, g_k \in F[X_1, \ldots, X_n]\) such that \(\sum f_i g_i = 1\) (the latter representation is called a boolean Nullstellensatz refutation), provided it does exist (i.e. \(f_1, \ldots, f_k\) have no common zero), fail.

Notice that one could assume all \(g_i\) to be multilinear, in particular, their degrees are at most \(n\). So, the goal is to establish the linear in \(n\) lower bounds on the degrees of \(g_1, \ldots, g_k\).

In [14] the first such method was designed which allowed to prove linear lower bounds (even for the polynomial calculus, being a more powerful proof system rather than the Nullstellensatz proofs) for a system of polynomials of a logarithmic degree which describes a modification of the into pigeon-hole principle (an exposition of this method see also in [10]). It holds over an arbitrary field. But for many other systems of polynomials the issue of lower bounds still remains open. Let us also mention that in the earlier papers [3], [7], [6], [4] the methods for obtaining somewhat weaker than linear bounds were exhibited.

It seems to be an interesting general question, how to obtain lower bounds for boolean Nullstellensatz refutations. In this paper we develop an approach which allows to produce explicitly a system of polynomials of degree 6 and to prove a linear lower bound on the degree of its boolean Nullstellensatz refutation. This approach borrows an idea from [13] to reduce the issue of Nullstellensatz refutations to Thue systems. First, we introduce and study (see section 1) boolean multiplicative Thue systems (basically, they consist of binomials necessarily containing among them the polynomials \(X_i^2 - 1, 1 \leq i \leq n\)). They extend slightly Tseitin's tautologies [16], [9], [17], [18]. We exploit the construction of the Tseitin's tautologies ([9], [17], [18]), based on expanders ([1], [11], [12]) and give a somewhat simpler proof of a linear lower bound for the case of used in section 1 notion of refutations (lemma 4). Relying on it, we first prove a linear lower degree bound for Nullstellensatz refutations for the systems which include the polynomials \(X_i^2 - 1, 1 \leq i \leq n\) (theorem 1) and therefore, for the more customary boolean case of the polynomials \(X_i^2 - X_i, 1 \leq i \leq n\) (corollary 1 in section 2).

Some shortcoming is that theorem 1 (and thereby, corollary 1) does not hold over fields of characteristic 2. To get rid of the latter restriction, at the end of section 2 we consider boolean Thue systems relative the polynomials \(X_i^2 - X_i, 1 \leq i \leq n\). Unfortunately, in this case the best established bound is merely \(\Omega(\log n)\).

In section 3 we consider boolean Nullstellensatz refu-
tions for the Knapsack problem over any infinite field and prove a linear lower bound for it. For zero characteristic fields a similar result for the subset sum problem was shown in [10].

1 Boolean Multiplicative Thue Proof Systems

Let $F$ be a field with characteristic distinct from 2.

**Definition 1** A boolean multiplicative Thue system over $F$ in variables $X_1, \ldots, X_n$ is a family $T$ which consists of terms of two types:

$$X_i^2 \text{ for all } 1 \leq i \leq n$$

$$\sigma X_j^{j_1} \cdots X_k^{j_k}, \quad j_1, \ldots, j_k \in \{0, 1\}, \quad \sigma \in \{-1, 1\}$$

The system $T$ is satisfiable if all the terms from (1), (2) equal to 1 for certain $X_1, \ldots, X_n$ (evidently, all $X_1, \ldots, X_n$ belong to $\{-1, 1\}$).

Using (1) repeatedly one can reduce each term $\sigma X_j^{j_1} \cdots X_k^{j_k}$ with integer $j_1, \ldots, j_k$ to the form (2), throughout this section we consider terms in this reduced form, then the multiplication of monomials $X_j^{j_1} \cdots X_k^{j_k}$ corresponds to the sum of their exponent vectors $(j_1, \ldots, j_k)$ over $GF(2)$.

**Definition 2** A refutation for $T$ is a sequence of (reduced) terms $m_1, \ldots, m_N$ such that $m_{i-1} m_i$ is one of the terms from (2) (after the reductions by (1)) for each $1 \leq i \leq N$ and $-m_N = m_1 = 1$. The degree of the refutation is the maximum of the degrees of $m_1, \ldots, m_N$.

Obviously, if there is a refutation then $T$ is not satisfiable. The completeness proof in the next lemma is standard, cf. e.g. lemma 5.3 [18].

**Lemma 1** a) If $T$ is not satisfiable then there is a refutation.

b) There is a polynomial-time (moreover, from NC) algorithm for testing satisfiability. 

**Proof** Consider the following linear system $L_T$ over $GF(2)$ in the variables $z_1, \ldots, z_n$. For each term of type (2) include in $L_T$ the following linear equation: $j_1 z_1 + \cdots + j_n z_n = \chi(\sigma)$, where $\chi(1) = 0$, $\chi(-1) = 1$. The system $L_T$ is solvable if and only if $T$ is satisfiable. If $L_T$ is not solvable then a suitable linear combination (or in other words, the sum of a subset of the set) of its equations gives 0 at the left side and 1 at the right side. Then the product of all the terms corresponding to this subset provides a refutation. Lemma is proved.

Boolean multiplicative Thue systems extend slightly Tseitin's tautologies [16], and a refutation could be viewed as a special form of resolutions, we need just this form for the lower degree bound on the Nullstellensatz refutations in the next section. We exploit the construction [9], [17], [18] of the Tseitin's tautologies, based on expanders with a linear degree lower bound and give for it a somewhat simpler proof for the sake of self-containment.

Remind (see e.g. [1]) that an expander $G_n$ is a bipartite graph with two parts of vertices $A = \{A_1, \ldots, A_n\}$, $B = \{B_1, \ldots, B_n\}$, where $|A| = |B| = n$ such that $G_n$ is regular and for some constant $c > 0$ (the calculations in [11], [1] show that one could take $c = (6\sqrt{3} - 5)/18$, but we will not use it) any subset $D \subseteq A$, contains at least $(1 + c \left(1 - \frac{1}{2n}\right))|D|$ adjacent vertices in $B$ (the roles of $A$ and $B$ could be interchanged).

The system $T_n$ under producing has $6n$ variables $X_1, \ldots, X_{6n}$. Every variable among $X_1, \ldots, X_{6n}$ we identify with a corresponding edge of $G_n$. To any vertex $A_i \in A$, $1 \leq i \leq n$ corresponds the monomial $X^{a_i}$ (having the degree 0) from $T_n$ (of type (2)), being the product of the edges incident to $A_i$, where $a_i \in (GF(2))^n$. Renumbering (if necessary) the variables one can assume that $X_1$ is incident to $B_1 \in B$. Then include in $T_n$ the monomials $X^{b_1}, \ldots, X^{b_n}$ (each of degree 6) similar as above and the monomial $X_1 X^{b_i}$ of degree 5. Finally, add to (2) the term $-X_1$. The obtained system $T_n$ is not satisfiable because the product of all its terms of type (2) equals to $-1$ (obviously, $X^{a_1} \cdots X^{a_n} = X^{b_1} \cdots X^{b_n} = X_1 X_{6n}$). Denote $X_1 X^{b_i} = X^{b_i}$. For any $GF(2)$-linear combination $a_1 a_1 + \cdots + \alpha_n a_n + \beta b_1 + \cdots + \beta_n b_n \in (GF(2))^n$ we define its degree as the number of ones (thus, the degree of the corresponding monomial $X^{a_1} X^{a_2} \cdots X^{a_n} (X^{b_1} \cdots X^{b_n})^2$) and its weight as the number of ones among $a_1, \ldots, a_n, \beta_1, \ldots, \beta_n$.

**Lemma 2** a) The vectors $a_1, \ldots, a_n, b_1, b_2, \ldots, b_n$ are linearly independent.

b) For any constant $0 < c_1 < 1$ there exists a constant $c_0 > 0$ such that any linear combination with a weight between $c_1 n$ and $(2 - c_1)n$ has the degree greater than $c_0 n$.

**Proof** a) Let $a = a_1 a_1 + \cdots + a_n a_n, b = \beta_1 b_1 + \cdots + \beta_n b_n$ and $a + b = 0$. Observe that the supports of the vectors $a_1, \ldots, a_n$ are pairwise disjoint, the same is valid for the vectors $b_1, b_2, \ldots, b_n$. Hence the weights of the vectors $a$ and $b$ are equal, denote this weight by $s$. Let $a_{\gamma_1} = \cdots = a_{\gamma_n} = \beta_{\delta_1} = \cdots = \beta_{\delta_n} = 1$. If $s = n$ then $a_1 + \cdots + a_n + b_1 + b_2 + \cdots + b_n = (1, 0, \ldots, 0)$, so one can assume that $s < n$. Applying the property of expanders to the set $\{A_{\gamma_1}, \ldots, A_{\gamma_n}\} \subseteq A$ we get a variable $X_j$ which occurs in the monomial $X^{a_1} \cdots X^{a_n}$ without occurring in $X^{b_1} \cdots X^{b_n}$. Thus, the $j$-th coordinate of the vector $a + b$ does not vanish.

b) Consider a linear combination $a + b$. Denote by $s$ the weight of $a$ and by $t$ the weight of $b$. Let for definiteness $s \geq t$ and $a_{\gamma_1} = \cdots = a_{\gamma_n} = \beta_{\delta_1} = \cdots = \beta_{\delta_n} = 1$ (the case $t \geq s$ is considered in a sim-
ilar way). Then \( \frac{2^n}{3} \leq s \). Applying the property of expanders to the set \( \{ A_1, \ldots, A_n \} \subseteq A \), we obtain at least \( e_{0,n} = e(1 - \frac{2^n}{3}) \leq \frac{2^n}{3} \) variables which occur in \( X^{e_{0,n}} \) and do not occur in \( X^{e_{0,n}} \).

For any boolean multiplicative Thue system \( T \) it is helpful to consider the following graph \( \mathcal{M}_T \) of monomials (cf. [5]). Its \( 2^n \) vertices are the monomials \( X^j \), and two monomials \( X^j, X^j' \) are linked by an edge if \( X^j X^j' \) is one of the monomials of type (2) (after reductions by (1)). For the produced above system \( T_n \) we call an edge in \( \mathcal{M}_{T_n} \) between \( X^j \) and \( X^j' \) distinguished if \( X^j X^j' = X_1 \).

**Lemma 3** Refutations for \( T_n \) correspond exactly to cycles in \( \mathcal{M}_{T_n} \), which pass through 1 and have an odd number of distinguished edges.

**Proof** Any refutation corresponds to a cycle (see definition 2) \( m_0, \ldots, m_N \) and since \( -1 = -m_0 = m_N = (m_0 m_1) (m_1 m_2) \cdots (m_{N-1} m_N) \), we conclude that among the edges \( m_m m_l \) of \( \mathcal{M}_{T_n} \) there are an odd distinguished ones. Conversely, any cycle with an odd number of distinguished edges provides a refutation.

**Lemma 4** The degree of any refutation in \( T_n \) is at least \( \Omega(n) \) (cf. e.g. lemma 5.9 [18]).

**Proof** Take any refutation \( m_0, \ldots, m_N \). Denote by \( v_1, \ldots, v_N \in (GF(2))^{2n} \) the exponent vectors of the (reduced) monomials \( (m_0 m_1), (m_1 m_2), \ldots, (m_{N-1} m_N) \) ignoring all distinguished edges and preserving the order of the rest ones. Lemma 3 implies that \( v_1 + \cdots + v_N = (1, 0, \ldots, 0) \). By lemma 2a) the weight of the vector \( v_1 + \cdots + v_{\ell} \) equals to \( 2n \). For a certain \( \ell \geq 1 \) the weight of the vector \( v_1 + \cdots + v_\ell \) equals to \( n \). The vector \( e_\ell \) is the exponent vector of a monomial \( m_\ell m_\ell' \) for a certain \( \ell \geq \ell \). Then the exponent vector of the monomial \( m_\ell = (m_0 m_1) (m_1 m_2) \cdots (m_\ell m_\ell') \) equals either to \( v_1 + \cdots + v_\ell \) or to \( v_1 + \cdots + v_\ell + (1, 0, \ldots, 0) \). Lemma 3b) entails a lower bound \( \Omega(n) \) on the degree of \( v_1 + \cdots + v_\ell \), that proves the lemma.

2 Lower Bound on Nullstellenensatz Proofs

In [13] (see also [5]) a connection between Thue systems and membership problem for Thue ideals was exploited, and a double exponential lower bound for the latter problem was ascertained. Our situation is different since we study refutations (rather than the ideal membership problem), for which in general a single exponential upper bound is known ([2], [8]).

Convert any Thue system \( T \) (see section 1) into a *boolean multiplicative* polynomial ideal \( P \subseteq F[X_1, \ldots, X_n] \) replacing each term \( \sigma m \) (where \( \sigma \in \{-1, 1\} \) and \( m \) is a monomial) in (1) or (2) by the binomial \( 1 - \sigma m \). Evidently, \( T \) is satisfiable if and only if \( P \) is satisfiable. Denote by \( P_n = (1 - X^2_1, \ldots, 1 - X^2_n, f_1, \ldots, f_l) \) the polynomial ideal converted from \( T_n \).

**Theorem 1** Any Nullstellensatz refutation for \( P_n \) has the degree \( \Omega(n) \) (over any field with the characteristic distinct from 2).

**Proof** Let \( 1 = \sum g_i (1 - X^{2i}) + \sum g_i f_j \). Consider a modified graph \( \mathcal{M}_{P_n} \) with the same (as \( \mathcal{M}_{T_n} \)) set of \( 2^n \) vertices (i.e. monomials). For each term \( u \in F \) occurring in \( g_j \) and \( f_j = 1 - \sigma m \) we draw an edge \((m', m') \) in \( \mathcal{M}_{P_n} \), endowed with the weight \( u \). Thus, the induced weight of its incident vertex \( m' \) equals to \( u \), and the induced weight of the vertex \( m' m \) equals to \( -\sigma u \). Clearly, no edges correspond to the polynomials \( 1 - X^2 \).

W.l.o.g. we can consider the connected component of \( \mathcal{M}_{P_n} \) which contains the monomial 1. Observe that for every vertex of \( \mathcal{M}_{P_n} \) (except for just 1) the sum of the induced weights by all the incident edges equals to 0, and for the vertex 1 this sum equals to 1. If the connected component has a cycle with an odd number of distinguished edges \((m' X, m') \) (they correspond to the polynomial \( 1 + X_1 \)) then there is a cycle with the same property passing through 1, and we complete the proof of the theorem applying lemmata 3,4.

Now suppose on the contrary that each cycle has an even number of the distinguished edges. Then one can partition all the vertices (of the connected component) into two parts \( V_0, V_1 \). The set \( V_0 \) consists of all the vertices reachable in \( \mathcal{M}_{P_n} \) from 1 by paths with an even number of the distinguished edges. Then any distinguished edge links a vertex from \( V_0 \) with a vertex from \( V_1 \). Any other edge has its incident vertices either both in \( V_0 \) or in \( V_1 \).

We partition the sum of the weights (see above) induced by all the edges into \( \Sigma_0 + \Sigma_1 \) over the vertices from \( V_0 \) and from \( V_1 \), respectively. Then each distinguished edge gives an equal contribution into both \( \Sigma_0 \) and \( \Sigma_1 \). Every other edge gives the zero contribution into both \( \Sigma_0 \) and \( \Sigma_1 \). Hence \( \Sigma_0 = \Sigma_1 \). But on the other hand \( \Sigma_0 = 0 \) and \( \Sigma_0 = 1 \), which contradicts the supposition and proves the theorem.

Now we obtain a similar lower bound for more customary (see e.g. [5], [7], [6], [4], [14], [10]) boolean polynomial ideals (i.e. the ideals containing polynomials \( X^2_i - X_i \), for all \( 1 \leq i \leq n \), rather than \( X^2_i - 1 \) as above). For each variable \( X_i \), \( 1 \leq i \leq n \) making the linear transformation \( X_i \rightarrow -2X_i + 1 \) we transform the polynomials \( X^2_i - 1 \rightarrow 4(X^2_i - X_i) \). Denote by \( P'_n \subseteq F[X_1, \ldots, X_n] \) the system obtained by this transformation from \( P_n \) (evidently, it consists of the polynomials of degrees at most 6). Notice that \( P'_n \) is not necessary a binomial system (unlike \( P_n \)).

**Corollary 1** Any Nullstellensatz refutation for \( P'_n \) has the degree \( \Omega(n) \) (again over any field with the characteristic distinct from 2).

One could also study boolean Thue systems relative
the polynomials \(X_i^2 - X_i, 1 \leq i \leq n\), rather than \(X_i^2 - 1\). Some advantage of these systems is that it is possible to consider them over arbitrary fields \(F\) unlike the multiplicative systems which were useless for the fields of characteristic 2 (see section 1).

**Definition 3** A boolean Thue system \(T'\) is a family of polynomials of two types

\[
X_i^2 - X_i, \quad 1 \leq i \leq n \tag{3}
\]

\[
\alpha' m^i - \alpha'' m^j \tag{4}
\]

where \(\alpha', \alpha'' \in F\), \(m^i, m^j\) are monomials.

**Definition 4** A refutation of \(T'\) is a sequence of reduced terms \(m_0 = 1, m_1, \ldots, m_N = 0\) such that for any \(1 \leq i \leq N\) there is a \(F\) and a monomial \(m\) such that \(m_{i-1} = \alpha' m^i m^j\), \(m_i = \alpha'' m m^j\) for an appropriate polynomial of type (4).

Unfortunately, the next obtained for this system lower bound (which one can show by a straightforward induction) is weaker than the bound from theorem 1.

**Proposition 1** Consider a boolean Thue system in the variables \(X_1, \ldots, X_n, Y_1, \ldots, Y_n\) with the following polynomials of type (4):

\[
X_1 Y_1 - X_1, Y_1 - X_1 Y_2, \ldots, Y_n - X_n - 1, X_n - Y_n
\]

Then any refutation of this system has the degree at least \(\Omega(\log n)\). Note that this bound is sharp.

Notice that one can test satisfiability of a boolean Thue system \(T'\) in polynomial time. Indeed, among the terms occurring in (4), there should be a nonzero element of \(F\) (otherwise, just zeros would satisfy this system). Let it be a binomial \(\alpha_1 m_1 - \alpha_2 m_2 \neq 0\). We start recursively augmenting a subset \(U\) of variables first including in it all the variables from \(m_1\) (the variables from \(U\) should attain the value 1 to satisfy \(T'\)). At a recursive step if all the variables from a term \(0 \neq \alpha' m^j\) belong to \(U\) (this holds in particular if \(0 \neq \alpha' m^j \in F\) then add to \(U\) all the variables from \(m^j\) unless \(a' \neq 0\), in this case the algorithm yields a refutation and terminates). Continue doing this way while \(U\) is augmented. When \(U\) can’t be augmented anymore, one can satisfy \(T'\) putting 1 for the variables from \(U\) and putting 0 for the rest of the variables.

3 Lower Bound on Nullstellensatz Proofs for the Knapsack Problem over an Infinite Field

In [10] the lower bound \([n/2]\) for Nullstellensatz proofs is shown for the polynomials \(\{X_i^2 - X_i, 1 \leq i \leq n\}, \sum_{1 \leq i \leq n} C_i X_i - m\) for any \(m\) and nonzero \(C_1, \ldots, C_n\), actually over any field of zero characteristic. Here we prove the lower bound \(n\) for the knapsack problem \(\sum_{1 \leq i \leq n} C_i X_i - 1\) for suitable \(C_1, \ldots, C_n\) over an arbitrary infinite field \(F\).

**Proposition 2** For almost any \(\gamma_1, \ldots, \gamma_n \in F\) any Nullstellensatz refutation for the system of polynomials \(\{X_i^2 - X_i, 1 \leq i \leq n, \sum_{1 \leq i \leq n} \gamma_i X_i - 1\}\) has the degree at least \(n\).

**Proof** Let \(1 = \sum g_i (X_i^2 - X_i) + g(\sum \gamma_i X_i - 1)\) for some \(\gamma_1, \ldots, \gamma_n \in F\). W.l.o.g. we can assume that \(g\) is multilinear. Suppose that \(\deg g < n\).

There exist not all zero constants \(c_0, \gamma \in \{0, 1\}^n\) (in fact, they lie in the prime subfield of \(F\)) such that \(\sum \gamma_i G(\gamma) = 0\) (this identity holds for any multilinear polynomial \(G\) of a degree less than \(n\)). Therefore, \(\sum \gamma_i c_0 = 0\), where \(\gamma = (\gamma_1, \ldots, \gamma_n)\). The latter expression does not vanish for almost any \(\gamma_1, \ldots, \gamma_n\) from an infinite field \(F\).

4 Further Research and Open Questions

1) Over a field \(F\) of characteristic 2 which contains the field \(GF(4)\) one could almost literally repeat the construction from section 1 and theorem 1 with the following minor changes. The monomials (1) we replace by \(X_{i}^2\), respectively, \(j_1, \ldots, j_2, 0 \in \{0, 1, 2\}\) in (2), \(\sigma \in \{1, \theta, \theta^3\}\) where \(\theta\) is a generator of the multiplicative group \((GF(4))^*\). In the construction of the Thue system \(T_{\theta}\) one views \(a_i, b_i\) as the vectors from \((GF(3))^3\), as the distinguished term in (2) takes \(\theta X_i\) (replacing \(-X_i\)). Then lemmas 1, 2 go through, in lemma 3 one should replace “odd” by “not divisible by 3” and the corresponding graph \(M_{\theta}\) has now \(3^n\) vertices. This leads to lemma 4. In the proof of theorem 1 we partition the vertices of the connected component (which contains 1) of the graph \(M_{\theta}\) (now the latter has \(3^n\) vertices) into 3 parts, regarding \(d \mod 3\), where \(d\) is the number of distinguished edges on a path from this vertex to 1.

But it is not clear how to prove corollary 1 if we’d like just to stick with the boolean equations \(X_i^2 - X_i\). Over the field \(GF(2)\) it is even less clear, how to conduct a similar to the above construction.

2) How to extend theorem 1 and corollary 1 to the polynomial calculus (4) ([14]).

3) How to obtain better (rather than logarithmic, see proposition 1) lower bounds for boolean Thue systems described at the end of section 2? It relates to “reversible” pebble games, in which it is allowed to propagate pebbles also backwards. Is it possible to adjust for them the known lower bounds (the best one is \(\Omega(n/\log n)\)) for the (customary) pebble games (due to S. Cook, W. Paul, R. Tarjan et al.)? But actually boolean Thue systems seem to be more powerful then pebble games, and it would be interesting to obtain for them a linear lower bound.
Acknowledgement. The author is thankful to Sasha Razborov for drawing attention to this area.

References