Lower Bounds of Static Lovász-Schrijver Calculus Proofs for Tseitin Tautologies

Dmitry Itsykson * Arist Kojevnikov †
St.Petersburg Department of
Steklov Institute of Mathematics,
27 Fontanka, 191023 St.Petersburg, Russia.
http://logic.pdmi.ras.ru/{~arist,~dmitrits}/

Abstract

We prove an exponential lower bound on the size of static Lovász-Schrijver proofs of Tseitin tautologies. We use several techniques, namely, translating static $\mathbf{LS}_+$ proof into *Positivstellensatz* proof of Grigoriev et al., extracting a "good" expander out of a given graph by removing edges and vertices of Alekhnovich et al., and proving linear lower bound on the degree of *Positivstellensatz* proofs for Tseitin tautologies.

1 Introduction

Expander graphs, that were introduced in the early 70s of previous century by Margulis, play significant role in the complexity theory. The first lower bound on the proof size in the resolution proof system was achieved on graphs in [21]. Later it was improved by using expanders in [22]. The recent result of Reingold [20] on the equivalence of two complexity classes, logspace and symmetric logspace ($L = SL$), is based on properties of expander graphs. We should also mention new simplified proof of PCP theorem using expanders [8]. In this work we investigate the following property of expanders in the framework of semialgebraic proof complexity: after removing small enough

*Supported in part by RFBR grant 06-01-00502, INTAS grant 04-77-7173, grant NSh-8464.2006.1 and Microsoft Research Inspire Programe.

†Supported in part by RFBR grants 05-01-00932, 06-01-00502 and INTAS grant 04-83-3836.

linear (with respect to number of vertices) number of edges the remaining graph remains enough "good" expander.

The complexity of semialgebraic proof systems is one of the rapidly developed in last years area of propositional complexity. Semialgebraic proof systems are proof systems for the language of unsolvable systems of polynomial inequalities with rational coefficients and propositional variables.

One of the first introduced semialgebraic proof system was Cutting Planes (**CP**). [9, 4, 7]. This system operates with linear inequalities with integer coefficients by rules of addition and rounding. A proof in **CP** is the derivation of contradiction $0 \geq 1$. Exponential lower bound for **CP** was proved by Pudlak in [19].

Another semialgebraic proof system Lovász-Schrijver (**LS**) [15, 14] operates with quadratic inequalities and uses the following rules: addition, multiplication by variable or its negation. We also consider stronger version of this system using as axiom the fact about nonnegativeness of the square of linear polynomial ($\mathbf{LS}_+$). Exponential lower bounds for both systems are unknown.

In this paper we prove exponential lower bound for static (and therefore for tree-like) propositional prove system $\mathbf{LS}_+$. The only known lower bound for static system $\mathbf{LS}_+$ was proved in [12] for system of linear inequalities "symmetric knapsack", that has not short representation as Boolean formula. In the paper [3] lower bound n^ϵ was proved for tree-like **LS** as propositional proof system.

The paper is organized as follows. Sect. 2 contains the necessary definitions. The proof of the main result is based on ideas of Theorem 9.3 from [12] and is divided into four parts. In Sect. 3 we prove that if a graph G with n vertices is a "good" expander then we can extract a "good" expander out of G after removing $O(n)$ vertices. (This part of the proof was not necessary in the [12] as there knapsack problem was considered.) In order to do this, we use the technique of [1]. Sect. 4 contains the transformation of the lower bound for *Positivstellensatz* into a Boolean degree lower bound for static $\mathbf{LS}_+$. In Sect. 5 we extend the linear lower bound on degree of the *Positivstellensatz* calculus [10] to linear lower bound on the Boolean degree of Tseitin tautologies in binomial form. Finally, in Sect. 6 we obtain exponential lower bounds for Tseitin tautologies in static and tree-like $\mathbf{LS}_+$ with squares.

2 Preliminaries

2.1 Proof systems

A *proof system* [6] for a language L is a polynomial-time computable function mapping words (treated as proof candidates) to L (whose elements are considered as theorems).

A *propositional proof system* is a proof system for the language **TAUT** of Boolean tautologies in disjunctive normal form (DNF). In order to compare proof system for any co-**NP**-complete language with propositional proof systems we need to fix a concrete reduction of **TAUT** to L.

An *algebraic proof system* is a proof system for the co-**NP**-hard language of unsolvable systems of polynomial equations: we are given several polynomials over a field $\mathbb{F}$ and the question is whether these polynomials have no common roots in $\mathbb{F}$. The polynomials are represented as sums of monomials $c \cdot x_1 \cdots x_s$, where $x_1, \ldots, x_s$ are variables and $c \in \mathbb{F}^* = \mathbb{F} \setminus \{0\}$ is a constant given in some reasonable (e.g., binary) notation.[1]

It is easy to see that this problem is co-**NP**-complete: it is possible to transform Boolean formula F in k-DNF with n variables into the set of polynomials $f_1, \ldots, f_m$ such that the system of polynomials

$$f_1, \ldots, f_m, x_1^2 - x_1, \ldots, x_n^2 - x_n \ ,$$

has not common roots if and only if F is tautology. Let us give example of such transformation: each clause C_i of formula $\neg F$ in k-CNF containing variables $x_{i_1}, \ldots, x_{i_s}$ we transform into a polynomial

$$f_i = (1 - l_{i_1}) \cdot \ldots \cdot (1 - l_{i_s}) \ , \tag{2.1}$$

where $l_{i_j} = x_{i_j}$ if the variable x_{i_j} occurs in C_i positively, and $l_{i_j} = (1 - x_{i_j})$ if it occurs negatively.

In *Polynomial Calculus* (**PC**) [5], one starts with the polynomial equation system $\mathcal{D}$ and derives new polynomials using the following two rules:

$$\frac{f = 0 \qquad g = 0}{f + g = 0} \qquad \text{and} \qquad \frac{f = 0}{f \cdot g = 0} \ .$$

A proof in this system is a derivation of $1 = 0$ from $\mathcal{D}$ using these rules.

[1] Unfortunately this representation does not permit to consider propositional formulas without restriction on the length of conjunctions in DNF as $(1 - x_1) \cdot \ldots \cdot (1 - x_n)$ contains exponential number of monomials. Therefore we consider algebraic proof systems for co-**NP**-complete language of tautologies in k-DNF.

We fix $\mathbb{F}$ as the field of rational numbers for all the following proof systems. *Positivstellensatz* (**PS**) [13] operates with polynomials over a real field. The proof $\mathcal{D}$ consists of polynomials $g_1, \ldots, g_{m+n}$ and $h_1, \ldots, h_s$ such that

$$\sum_{i=1}^{m+n} f_i g_i = 1 + \sum_{j=1}^{s} h_j^2 \ . \tag{2.2}$$

It is a "static" proof in the sense that it contains only one step. Note that the right-hand side of (2.2) is the derivation in **PC**.

A *semialgebraic proof system* operates with language of unsolvable systems of polynomial inequalities. They are much more powerful than algebraic proof systems. No nontrivial complexity lower bounds for some of them are known so far. Moreover, in semialgebraic systems there exist short proofs of many tautologies that are hard for other proof systems [12].

To define a propositional proof system working with inequalities, we translate each formula $\neg F$ with n variables in CNF into a system of linear inequalities such that F is a tautology if and only if the system has not solution in $\{0,1\}$-variables. For a formula F, we translate each clause C_i of $\neg F$ with variables $x_{j_1}, \ldots, x_{j_t}$, into the inequality

$$l_{i_1} + \ldots + l_{i_t} - 1 \geq 0 \ , \tag{2.3}$$

where $l_{i_j} = x_{i_j}$ if the variable x_{i_j} occurs positively in the clause, and $l_{i_j} = (1 - x_{i_j})$ if x_{i_j} occurs negatively. For every variable x_i, $1 \leq i \leq n$, we also add to the system $\mathcal{D}$ the inequalities

$$0 \leq x_i \leq 1 \ . \tag{2.4}$$

In Lovász-Schrijver proof system (**LS**) [15], one obtains the contradiction $0 \geq 1$ using the rules

$$\frac{f \geq 0 \qquad g \geq 0}{\lambda_f f + \lambda_g g \geq 0} \ , \qquad \frac{h \geq 0}{hx \geq 0} \ , \qquad \frac{h \geq 0}{h(1-x) \geq 0} \ ,$$

where $\lambda_f, \lambda_g \geq 0$, the polynomial h is linear and x is a variable. Also, the set of axioms (2.4) is extended by the inequalities

$$x_i^2 - x_i \geq 0 \ , \qquad \text{for every variable } x_i,\ 1 \leq i \leq n \ . \tag{2.5}$$

The system $\mathbf{LS}_+$ [15] has the same axioms and derivation rules as **LS** and the addition axiom

$$h^2 \geq 0 \ , \qquad \text{for every linear } h \ . \tag{2.6}$$

The proof is *tree-like* if the underlying directed acyclic graph, representing the implication structure of the proof, is a tree. That is, every inequality in the proof, except for the initial inequalities, is used at most one as an antecedent of an implication.

A proof of inequality system $\{f_1 \geq 0, \ldots, f_m \geq 0\}$ with n variables in *static* $\mathbf{LS}_+$ [12] consists of positive real coefficients $c_{i,s}$ and multisets $U^+_{i,s}, U^-_{i,s}$ determining the polynomials $g_{i,s} = c_{i,s} \cdot \prod_{k \in U^+_{i,s}} x_k \cdot \prod_{k \in U^-_{i,s}} (1 - x_k)$ such that

$$\sum_{i=1}^{M} f_i \sum_s g_{i,s} + \sum_{j=1}^{n} (x_j^2 - x_j) \sum_s g_{m+j,s} + \sum_{i=n+1}^{n'} h_i^2 \sum_s g_{m+i,s} = -1 \ . \quad (2.7)$$

Note that static proof systems like **PS** and static $\mathbf{LS}_+$ are not propositional proof systems in the sense of Cook and Reckhow [6], but are something more general, since there is no obvious way to verify (2.7) in deterministic polynomial time. However, they can be easily augmented to match the definition of proof systems, by including a proof of equality (2.7) based on the axioms of a ring (see F-NS of [11]). Clearly, any lower bound for the original system is valid for any augmented system as well.

2.2 Tseitin formulas

Let us consider undirected graph $G = (V, E)$, degrees of all vertices are not exceed d, $V' \subseteq V$. For each edge e we attach propositional variable x_e. For each vertex $v \in V'$ we write down $\bigoplus_{e \ni v} x_e = 1$, and for each vertex $v \in V \setminus V'$ write down $\bigoplus_{e \ni v} x_e = 0$.[2] The conjunction of all written formulas we denote as $T_G^{V'}$.

Formula $T_G^{V'}$ may be defined by the following set of clauses:

$$\bigvee_{e \in S_v \setminus S'_v} x_e \vee \bigvee_{e \in S'_v} \neg x_e, \quad (2.8)$$

for all vertices $v \in V'$ and all even cardinality subsets S'_v of set of edges S_v that contain vertex v and for all vertices $v \in V \setminus V'$ and all S'_v of odd cardinality.

We need in the following lemmas:

Lemma 2.1. *Formula $T_G^{V'}$ is unsatisfiable if cardinality of set V' is odd.*

Proof. Assume that there is satisfying assignment for formula $T_G^{V'}$. Let us consider graph $\widetilde{G}$ that contains all edges of G with value 1 of corresponding

[2]Here and after $\oplus$ denotes operation sum modulo 2 ("exclusion or")

variable in satisfying assignment. Graph $\widetilde{G}$ contains odd number of vertices with odd degree. Therefore sum of all degrees is odd. But sum of degrees is doubled number of edges. □

For all subsets V' of odd cardinality formula $\neg T_G^{V'}$ is called *Tseitin tautology*. We will call formula $T_G^{V'}$ *Tseitin formula* and will usually omit V' in notations.

Substitution of a variable value in Tseitin formula T_G corresponds to removing of an edge in the graph G. Tseitin formula remain to be Tseitin after substitution: if we substitute value 0 then set V' is not changed and if we substitute 1 then both ends of the edge simultaneously change the parity if the sum around and therefore parity of $|V'|$ is not changed. We use the following simplified notation: if ρ is partial substitution of variable of formula T_G then we denote graph with set of edges corresponding to unassigned variable of ρ as $G|_\rho$.

Lemma 2.2. *$G = (V, E)$ is connected graph. The formula $T_G^{V'}$ is satisfiable, if cardinality of set V' is even.*

Proof. One may find satisfying assignment of formula $T_G^{V'}$ using the following algorithm:

- Assign to all variables value 0. The number of vertices with broken condition of parity is even.
- While there is vertices with broken condition of parity do:
  - Choose two such vertices and flip value of edges on the path between them. Thus the number of vertices with broken condition of parity is decreased by 2.

□

Lemma 2.3. *Let us $G = (V, E)$ be connected graph, $U \subsetneq V$. Then conjunction of clauses from $T_G^{V'}$ corresponding to vertices of set U is satisfiable formula.*

Proof. Consider connected components $H_1, H_2, \ldots, H_k$ of a graph $G_U = (U, E_U)$, where $E_U = \{(u, v) \in E | u, v \in U\}$. If H_i contains even number of vertices from V', then by Lemma 2.2 the part of the formula corresponding to H_i may be satisfied (all external edges has value 0). If H_i contains odd vertices from V', one may repair parity by assigning value 1 to one of external edges. □

2.3 Expander graphs

For subsets I, I_1 of vertices and subset of edges $J \subseteq E$ we define *boundary operation* ∂:

$$\partial_{V\setminus I, E\setminus J}(I_1) \;=\; \{(v,v') \in E \setminus J : v \in I_1 \text{ and } v' \in (V \setminus I_1) \setminus I\} \ .$$

Here and in the rest of the paper we use $\partial_{V,E}(I)$ as short notation for $\partial_{V\setminus\emptyset, E\setminus\emptyset}(I)$. We say that a graph $G = (V,E)$ is an (r,d,c)-*expander* [17] if the maximal degree of any vertex is d, and for every set $X \subseteq V$ of cardinality at most r,

$$|\partial_{V,E}(X)| \geq c \cdot |X| \ .$$

Further we always consider Tseitin formulas based on expanders.

2.4 Boolean degree

For lower bounds on **PS** refutations the following *binomial representation of Tseitin formulas* was used [10]. To each edge of the graph G we assign a $\{1,-1\}$-variable y_k. The system Tb_G contains the equations

$$Y(v) = c_v \cdot \prod_{e \ni v} y_e = 1 \tag{2.9}$$

for each vertex $v \in V'$ with constant $c_v = -1$, for each vertex $v \in V \setminus V'$ with constant $c_v = 1$ and $y_e^2 = 1$.

The Boolean degree of monomial in $\{0,1\}$-variables is the number of different variables in it. In case of $\{1,-1\}$-variables *the Boolean degree* of monomial is the number of variables having odd degree in it. The Boolean degree of polynomial is the maximum of Boolean degrees of all monomials in it.

We may transform polynomial in $\{0,1\}$-variables into the polynomial in $\{1,-1\}$-variables by means of the following substitution: $x_i = \frac{1-y_i}{2}$. From the explicitly form of this substitution we may conclude the following lemma:

Lemma 2.4. *The Boolean degree of polynomial in $\{0,1\}$-variables does not increase after the transformation to $\{1,-1\}$-variables.*

3 Closure Operator on Expanders

In this section we describe cleaning procedure of expander graphs which is very similar to cleaning procedure of expander matrices from [2, 1].

For a (r, d, c)-expander graph $G = (V, E)$ and a subset of its edges $J \subseteq E$ we define an inference relation $\vdash_J$ on subsets of vertices $I, I_1 \subseteq V$:

$$I \vdash_J I_1 \qquad \stackrel{def}{\iff} \qquad (|I_1| \leq \frac{r}{2}) \wedge (\left|\partial_{V \setminus I, E \setminus J}(I_1)\right| < \frac{c}{2}|I_1|) \ .$$

For a subset of vertices I and a set of edges J we consider the following *cleaning* procedure:

- If there exists a nonempty $I_1 \subseteq V$, such that $I \vdash_J I_1$ and $I \cap I_1 = \emptyset$, then take such I_1 and add it to I.
- Repeat the cleaning step as long as it is applicable.

Let the *closure* $Cl(J)$ of J be the result of cleaning procedure applied to $\emptyset$ and J.

The notion of $Cl(J)$ is ambiguous and depends on choice of set I_1. We call as $Cl(J)$ the result of any correct cleaning procedure. We will use special cleaning procedure from the following Lemma.

Lemma 3.1. *Let sets of vertices $I_1, I_2, \ldots, I_k$ be pairwise disjoint, $|I_j| \leq r/2$ for all $1 \leq j \leq k$ and $|\partial_{V,E \setminus J} I_j| < c/2|I_j|$. Then there is cleaning procedure with the following property: $I_j \subseteq Cl(J)$ for all $1 \leq j \leq k$.*

Proof. For all j we have $\emptyset \vdash_J I_j$, therefore $I \vdash_J I_j$ for all sets I. We add sets in the order: $I_1, I_2, \ldots, I_k$. After it we add other sets in arbitrary order. □

Informally speaking if we remove from expander graph edges from J, $Cl(J)$ is precisely the set of vertex we need to remove from graph to make it expander (but with worse properties).

Lemma 3.2 ([1], Lemma 3.4). *Assume that a graph $G = (V, E)$ is an (r, d, c)-expander and J is a subset of its edges. Let $I' = Cl(J)$ and $J' = \{(v, x) \in E : v \in I' \text{ or } x \in I'\}$. Denote by $G' = (V \setminus I', E \setminus J')$ the graph that results from G by removing vertices corresponding to I' and edges corresponding to J'. If G' is non-empty then it is an $(r/2, d, c/2)$-expander.*

Proof. Follows immediately from the definition of Cl. □

In the next lemma we show that if we take J of small cardinality, then the graph G' from Lemma 3.2 is non-empty.

Lemma 3.3 ([1], Lemma 3.5). *Let a graph $G = (V, E)$ be an (r, d, c)-expander and $|J| < cr/4$. Then $|Cl(J)| < 2c^{-1}|J|$.*

Proof. Assume that $|Cl(J)| \geq 2c^{-1}|J|$. Consider the sequence $I_1, I_2, \ldots, I_s$ appearing in the cleaning procedure; i.e.,

$$I = I_1 \cup I_2 \cup \ldots \cup I_k \vdash_J I_{k+1} \quad k = 1, \ldots, s-1 \ .$$

Note that $I_i \cap I_j = \emptyset$ for all $i \neq j$. Denote by $C_t = \cup_{k=1}^{t} I_k$ the set of vertices derived in t steps.

Let T be the first t such that $C_t \geq 2c^{-1}|J|$. Note that $|J| \leq c|C_T|/2$, hence $|C_T| \leq 2c^{-1}|J| + r/2 \leq r$. Because of the expansion properties of G, $\partial_{V,E}(C_T) \geq c|C_T|$, which implies

$$|\partial_{V,E\setminus J}(C_T)| \geq c|C_T| - |J| \geq c|C_T|/2 \ . \tag{3.1}$$

On the other hand, every time we add some I_{t+1} to C_t during the cleaning procedure, by the subadditivity property for the boundary operator we add strictly less than $c/2|I_{t+1}|$ new elements to $\partial_{V,E\setminus J}(C_T)$. This implies $|\partial_{V,E\setminus J}(C_T)| < c|C_T|/2$, which contradicts (3.1). □

4 Simulation of Static $\mathbf{LS}_+$ in PS

In this section we transform a proof in static $\mathbf{LS}_+$ of the system of linear inequalities Ta_G into a **PS** proof of the system of binomial equations Tb_G with constant increase of Boolean degree.

Let us consider **PS** proof

$$1 + \sum_{j=1}^{M} h_j^2 = \sum_{i=1}^{n} f_i g_i, \tag{4.1}$$

of binomial system of equalities $P_T : f_i = 0, i = 1, \ldots, n$.

The Boolean degree of **PS** proof (4.1) is the maximum of Boolean degrees of polynomials $f_i g_i$, $1 \leq i \leq n$ and h_j^2, $1 \leq j \leq M$.

Let us define *the Boolean degree of static* $\boldsymbol{LS}_+$ *proof* as maximum of Boolean degrees of polynomials $g_{i,l}$ from proof (2.7) in static $\mathbf{LS}_+$.

Next two lemmas can be applied to a static $\mathbf{LS}_+$ proof P of arbitrary Boolean formula F, they show that P can be transformed into the **PS** proof of F with only constant increase of Boolean degree.

Fix a Boolean formula F with m clauses and n variables, let F^A be set of linear inequalities provided by translation (2.3) and F^M be set of equations provided by (2.1) from formula F.

Lemma 4.1. *In static* $\boldsymbol{LS}_+$*, every proof P of F^A can be transformed into a proof P' of the polynomial equation system F^M. Moreover, if $Bdeg(P) = k$*

and the number of variables in every inequality of F^A is at most d, then $Bdeg(P') \leq k + d$.

Proof. The proof P can be represented in the form

$$\sum_{i=1}^{m} f_i^A \sum_s g_{i,s} + \sum_{i=m+1}^{n} f_i \sum_s g_{i,s} = -1 \ , \tag{4.2}$$

where

$$g_{i,s} = c_{i,s} \prod_{t \in U_{i,s}^+} x_t \cdot \prod_{t \in U_{i,s}^-} (1 - x_t)$$

for appropriate multisets of variables $U_{i,s}^+, U_{i,s}^-$ and a positive $c_{i,s} \in \mathbb{Q}$.

We show that the translation of a clause $C_i = (l_1 \vee \ldots \vee l_{d_i})$, $i = 1, \ldots, m$ into an inequality $f_i^A = \sum_{t=1}^{d_i} l_t - 1 \geq 0$ can be represented as the translation of the clause C_i into an equation $f_i^M = \prod_{t=1}^{d_i}(1 - l_t) = 0$:

$$f_i^A = -f_i^M + \rho(l_1, \ldots, l_{d_i}) \ , \tag{4.3}$$

where the second summand $\rho(l_1, \ldots, l_{d_i})$ is nonnegative and equal to a sum of literal products. The induction base is $\rho(l_1) = 0 \geq 0$, the induction step is

$$\rho(l_1, \ldots, l_{d_i}) = \rho(l_1, \ldots, l_{d_i - 1})(1 - l_{d_i}) + \sum_{t=1}^{d_i - 1} l_t \cdot l_{d_i} \geq 0.$$

Let us replace each f_i^A in proof P by (4.2). As a result, we obtain the proof P':

$$\sum_{i=1}^{m} -f_i^M \sum_s g'_{i,s} + \sum_{i=m+1}^{n'} f_i \sum_s g'_{i,s} = -1 \ , \tag{4.4}$$

where

$$g'_{i,s} = c'_{i,s} \cdot \prod_{t \in U_{i,s}^+} x_t \cdot \prod_{t \in U_{i,s}^-} (1 - x_t)$$

for appropriate multisets $U_{i,s}^+, U_{i,s}^-$ and positive $c'_{i,s} \in \mathbb{Q}$.

Since the right-hand side of (4.3) has the Boolean degree at most d, the Boolean degree of the new refutation is at most $k + d$. □

Lemma 4.2. *Every static $\boldsymbol{LS}_+$ proof P of F^M can be transformed into **PS** proof P' of it. If $Bdeg(P) = k$ and $Bdeg(f_i) \leq d$, then $Bdeg(P') \leq k + d$.*

Proof. We use ideas from the proof of Lemma 9.3, [12]. The refutation P can be represented in the form

$$\sum_{i=1}^{n+m} f_i \sum_s g_{i,s} + \sum_{j=1}^{n'} h_{0,j}^2 \cdot g_{m+n+1,j} + \sum_{j=n'+1}^{n''} g_{m+n+1,j} = -1 \ ,$$

where f_i, $1 \leq i \leq m$ are translations of Boolean clauses, $f_{m+i} = x_i^2 - x_i$, $1 \leq i \leq n$ and $g_{i,s} = c_{i,s} \cdot \prod_{t \in U_{i,s}^+} x_t \cdot \prod_{t \in U_{i,s}^-} (1 - x_t)$ for appropriate multisets of variables $U_{i,s}^+, U_{i,s}^-$, positive real $c_{i,s}$, and linear $h_{0,j}$.

Let us replace each occurrence of x_e in $g_{m+n+1,j}$ by $(x_e - x_e^2) + x_e^2 = -f_{m+e} + x_e^2$ and each occurrence of $1 - x_e$ by $(x_e - x_e^2) + (1 - x_e)^2 = -f_{m+e} + (1 - x_e)^2$, expand the factors obtained, gather all the terms containing at least one of f_i and the products of squares. As a result, we obtain **PS** proof P' of the form

$$\sum_{i=1}^{n+m} f_i g_i + \sum_{j=1}^{n'''} h_j^2 = -1 \ ,$$

for appropriate polynomials g_i, h_j. The Boolean degrees of g_i, h_j are at most $\mathrm{Bdeg}(g_{i,s})$ and Boolean degrees of all f_i are at most d, so Boolean degree of P' is at most $k + d$. □

Next part of the reductions depends on Tseitin formula $T = T_G$ constructed according to graph $G = (V, E)$ and its representations as systems of linear inequalities, equations and binomials.

Lemma 4.3. *Every **PS** proof P of Tm_G can be transformed into a **PS** proof P' of Tb_G. The Boolean degree of P' is at most $Bdeg(P) + d$.*

Proof. Assume the proof P is as follows:

$$\sum_{v, S_v} f_{v,S_v}^M \cdot g_{v,S_v} + \sum_{e \in E} (x_e^2 - x_e) \cdot g_e = 1 + \sum_j h_j^2 \ .$$

First of all, we replace each occurrence of x_e by $(1 - y_e)/2$. Note that the substitution transforms each $x_e^2 - x_e = 0$ into $(y_e^2 - 1)/4 = 0$, and each F^M into

$$\prod_{e \in S_v \setminus S_v'} \frac{1 + y_e}{2} \cdot \prod_{e \in S_v'} \frac{1 - y_e}{2} = 0 \ . \tag{4.5}$$

Due to Lemma 2.4 the Boolean degree of the new proof is at most $\mathrm{Bdeg}(P)$.

Next, we multiply (4.5) and (2.9) for $v \in V'$ and use the reduction modulo ideal $\langle y_e^2 = 1 | e \in E \rangle$:

$$\prod_{e \in S_v \setminus S'_v} \frac{1+y_e}{2} \prod_{e \in S'_v} \frac{1-y_e}{2} (\prod_{e \ni v} y_e + 1) =$$

$$\prod_{e \in S_v \setminus S'_v} \frac{y_e + y_e^2}{2} \prod_{e \in S'_v} \frac{y_e - y_e^2}{2} + \prod_{e \in S_v \setminus S'_v} \frac{1+y_e}{2} \prod_{e \in S'_v} \frac{1-y_e}{2} =$$

$$\prod_{e \in S_v \setminus S'_v} \frac{y_e + 1}{2} \prod_{e \in S'_v} \frac{y_e - 1}{2} + \prod_{e \in S_v \setminus S'_v} \frac{1+y_e}{2} \prod_{e \in S'_v} \frac{1-y_e}{2} =$$

$$2 \cdot \prod_{e \in S_v \setminus S'_v} \frac{1+y_e}{2} \prod_{e \in S'_v} \frac{1-y_e}{2} \cdot$$

The set S'_v has even cardinality, so $\prod_{e \in S'_v} (y_e - 1) = \prod_{e \in S'_v} (1 - y_e)$. A similar equality holds for $v \in V \setminus V'$.

Now we can write down the transformed proof P':

$$\sum_{v, S'_v} (\prod_{e \ni v} y_e + 1) \cdot 2 \cdot {f'}^M_{v,S'_v} \cdot {g'}_{v,S'_v} + \sum_{e \in E} 2^{-2} \cdot (y_e^2 - 1) \cdot g'_e = 1 + \sum_j {h'}_j^2 \ ,$$

where the polynomials ${f'}^M_{v,S'_v}$, g'_e, ${h'}_j^2$ are obtained from f^M_{v,S'_v}, g_e, h_j^2 by applying the substitution $x_i = (1 - y_e)/2$.

The Boolean degree of each equation (2.9) is at most d, hence $\mathrm{Bdeg}(P') \leq \mathrm{Bdeg}(P) + d$. □

Lemma 4.4. *Every static* $\boldsymbol{LS}_+$ *proof of the* Ta_G *can be transformed into a* ***PS*** *proof* Tm_G*. We can bound the Boolean degree of the new proof by* $k + 3d$*, where* k *is the Boolean degree of the static* $\boldsymbol{LS}_+$ *proof.*

Proof. Fix a static $\mathbf{LS}_+$ proof P of Ta_G and apply Lemma 4.1 to obtain a static $\mathbf{LS}_+$ proof P' of the equation system Tm_G. Next, transform P' into a **PS** proof P'' of Tm_G by Lemma 4.2. Finally, due to Lemma 4.3 we can transform P'' into a **PS** proof P''' of system Tm_G. The Boolean degree of P''' is at most $k + 3d$. □

5 Linear Lower Bound on the Boolean Degree of the PS Proof of Tseitin Formulas

In this section we extend lower bound on the degree of the binomial Tseitin formulas derivations in **PS** to lower bound on the *boolean degree*.

In [10] was used different notion of expanders but it is easy to see that the result [10, Lemma 8] is also correct in the following form:

Theorem 5.1 ([10], Lemma 8). *For all d, c there is an ϵ_0 so that if n-vertex graph G is $(n/2, d, c)$-expander, then the degree of every **PS** derivation of the system of equalities Tb_G is at least $\epsilon_0 n$, where Tb_G is the binomial representation of the Tseitin formula based on the graph G.*

Lemma 5.2. *Lets f_i be the set of multilinear polynomials. Then every **PS** derivation $1 + \sum_i h_i^2 = \sum_j f_j g_j + \sum_t (x_t^2 - 1)\tilde{g}_t$ can be transformed into the derivation $1 + \sum_i h_i'^2 = \sum_j f_j g_j' + \sum_t (x_t^2 - 1)\tilde{g}_t'$ so that it's Boolean degree is not increased and degree of each variable in the polynomials $h_i'^2$, $f_j g_j'$ and $(x_t^2 - 1)\tilde{g}_t'$ is at most 2.*

Proof. We will show that for any variable x_q the derivation $1 + \sum_i h_i^2 = \sum_j f_j g_j + \sum_t (x_t^2 - 1)\tilde{g}_t$ can be transformed into the derivation $1 + \sum_i h_i'^2 = \sum_j f_j g_j' + \sum_t (x_t^2 - 1)\tilde{g}_t'$ so that it's Boolean degree is not increased and the degree of variable x_q in the polynomials $h_i'^2$, $f_j g_j'$ and $(x_t^2 - 1)\tilde{g}_t'$ is at most 2.

Assume that in all monomials of the polynomial p an algorithm α replaces the variable x_q in the even degree with 1 and in the odd degree with x_q.

We denote $h_i' = \alpha(h_j), g_j' = \alpha(g_j)$ and $\tilde{g}_t' = \alpha(\tilde{g}_t)$. The polynomial $1 + \sum_i h_i'^2$ is not necessary equals to $\sum_j f_j g_j' + \sum_{t \neq q} (x_t^2 - 1)\tilde{g}_t'$. The main reason is the following: before application of α two monomials were equal (may be with different coefficients), and they are not equal after the application. Note that parity of number of all appearances of variable x_q was not changed and degree of all the appearances after replacement is at most 2, therefore the only way that two monomials after replacement become not equal is that these monomials are of the type m and $x_q^2 m$. For all monomials of this type we add to the right hand side of the derivation $(x_q^2 - 1)m$ with corresponding coefficients.

Now we need to repeat with same operation for other variables.

□

Now we are ready to prove main result of this section:

Theorem 5.3. *For all d and c there is a positive number ϵ so that for all $n \in \mathbb{N}$ if n-vertex graph G is $(n/2, d, c)$-expander, then Boolean degree of any **PS** derivation of the system of equalities Tb_G has degree at least ϵn.*

Proof. Consider a derivation of the system Tb_G in **PS**:

$$1 + \sum_i h_i^2 = \sum_j f_j g_j + \sum_t (x_t^2 - 1)\tilde{g}_t.$$

By Lemma 5.2 we can transform it without increasing of boolean degree into the derivation $1 + \sum_i h_i'^2 = \sum_j f_j g_j' + \sum_k (x_t^2 - 1)\tilde{g}_t'$. Boolean degree of the

last derivation is at least half of the initial degree. By Theorem 5.1 degree is less then $\epsilon_0 n$ for some ϵ_0.

Therefore the Boolean degree of the initial derivation is at least $\frac{\epsilon_0}{2}n$. □

Lemma 5.4. *For any d and c there is an $0 < \epsilon < 1$ and $R \in \mathbb{N}$ so that the Boolean degree of any static $\boldsymbol{LS}_+$ refutation of Tseitin formula (2.8) with respect to (r, d, c)-expander G, where $r = n/2$, is at least ϵr for all $r > R$.*

Proof. Let P be a static $\mathbf{LS}_+$ proof of the formula (2.8) represented as the system of linear inequalities, and Boolean degree of P is k. We apply Lemma 4.4 and transform it to into a **PS** proof P' of the equation system (2.9) extended by all $y_e^2 - 1 = 0, e \in E$. The Boolean degree of P' is at most $k + 3d$.

Theorem 5.3 implies that there is $\epsilon' > 0$ depended only on c and d, such that $k + 3d \geq \epsilon' n$; the required statement follows from the fact that d is a constant. □

6 An Exponential Lower Bound on the Size of Static $\mathbf{LS}_+$ Refutation of Tseitin Formulas

In this section we use the idea of the proof of lower bound for static $\mathbf{LS}_+$ from [12].

Lemma 6.1 ([12], Lemma 9.2). *Let M denote the number of $g_{i,s}$ in (2.7) that have Boolean degrees at least k and N denote the number of different variables in (2.7). Then there is a variable x and a value $a \in \{0, 1\}$ such that the result of substitution $x = a$ in (2.7) contains at most $M(1 - k/(2N))$ nonzero polynomials $g_{i,s}|_{x=a}$ of Boolean degrees at least k.*

Proof. For each polynomial $g_{i,s}$ with Boolean degree at least k there is at least k substitutions so that $g_{i,s}$ become zero. There is $2N$ different substitutions of variables from the proof. Therefore there exists substitution $x := a$ so that at least $Mk/(2N)$ polynomials $g_{i,s}$ with Boolean degree at least k become zero. □

In the following theorem we use graphs with a positive expansion constant $c > 1$. For sufficiently large n there are such graphs of degree bounded by a constant (see, e. g. the proof in the Sect. 4 of [18] that for any d-regular graph $G = (V, E)$ and any subset of vertices $A \subseteq V$

$$\frac{|\partial A|}{|A|} \geq (d - \lambda_1)\frac{|V \setminus A|}{|V|} \ ,$$

where λ is the second eigenvalue of G. It follows that G is $(\frac{|V|}{2}, d, \frac{d-\lambda}{2})$-expander. As example of the graph with small second eigenvalue we use *Ramanujan graph*: is a d-regular graph satisfying $\lambda_1 \leq 2\sqrt{d-1}$ and use the explicit construction of Ramanujan graphs, Sect. 5 of [18] or [16].

We assume that partial assignment is an ordered set of substitutions of the form $x := a$, and we apply these substitutions in the given order. For example, if an assignment ρ already contains $x := 1$ we assume that $\rho \cup \{x := 0\}$ equals to ρ.

In Sect. 3 the operator Cl was defined for sets of edges. We extend it for use with partial assignments: $\overline{Cl}(\rho) = Cl(\{e \mid \rho(e) \text{ is set to 0 or 1}\})$.

Definition 6.2. *Let f be a mapping from partial assignments to their extensions. For a set $x_1, x_2, \ldots, x_\kappa$ of formula T_G variables and for a $\{0,1\}$-constants set $a_1, a_2, \ldots a_\kappa$ we define* sequence of partial assignments with respect to f *as follows: $\rho_0 = \emptyset$; $\rho_i = f(\rho_{i-1} \cup \{x_i := a_i\}), 1 \leq i \leq \kappa$.*

Definition 6.3. *An edge e is called a* bridge *in the undirected graph G if the removing of e from G increases the number of connected components in G.*

Theorem 6.4. *Let graph G with n vertices be (r, d, c)-expander with $c > 1$, $r = n/2$, $d \geq 4$, formula T_G be Tseitin formula with respect to G and $\kappa = \lceil \frac{cr}{13} \rceil$. Then there exists a mapping f such that for any x_i and a_i $(1 \leq i \leq \kappa)$ partial assignment ρ_κ, that is the last in the sequence of partial assignments with respect to f, can be extended to partial a assignment σ and formula $T_G|_\sigma$ is nontrivial Tseitin formula with respect to $(r/2, d, c/2)$-expander with number of vertices at least $\frac{3}{4}n$.*

Proof. We define mapping f as the first part of result of the following algorithm.

Algorithm $\mathcal{A}$.
Input: Assignment π.
Output: Assignment π' (an extension of π) and an assignment τ.

1. $\pi' := \pi$, $\tau := \emptyset$.
2. While $G|_{\pi'}$ contains bridges execute steps 3-6.
3. Let e be lexicographically first bridge in the graph $G|_{\pi'}$. Let e split connected component H on H_1 and H_2 (assume that $|H_1| \leq |H_2|$).
4. Choose value a in such a way that a formula $T_{H_1}|_{\pi' \cup \{x_e := a\}}$ becomes satisfiable (it can be done by Lemma 2.3 since H is connected graph).
5. $\pi' := \pi' \cup \{x_e := a\}$, $\tau := \tau \cup \{x_e := a\}$.

6. Extend π' by satisfying assignment of formula $T_{H_1}|_{\pi'}$.

7. Return π' and τ.

Using second part of result of the algorithm we define a mapping g. Informally speaking g is the part of assignment $f(\pi) \setminus \pi$ corresponding to the bridges.

We define a partial assignment τ_i as follows:

$$\tau_1 = g(\{x_1 := a_1\}),\ \tau_i = \tau_{i-1} \cup g(\rho_{i-1} \cup \{x_i := a_i\}),\ 2 \le i \le \kappa\ .$$

For convenience we also define τ':

$$\tau'_i = \tau_i \cup \{x_1 := a_1\} \cup \bigcup_{2 \le j \le i:\ x_j \notin \rho_{j-1}} \{x_j = a_j\},\ 1 \le i \le \kappa\ .$$

Assignment τ'_κ corresponds to bridges and substitutions $x_i := a_i$.

Lemma 6.5. *All graphs $G|_{\rho_i}, 1 \le i \le \kappa$ consist of one connected component and probably of some vertices of zero degree.*

Proof. By induction on i. The graph G doesn't contain bridges (otherwise a bridge connects two connected components H_1 and H_2, $|H_1| \le |H_2|$, $|H_1| \le n/2 = r$ and $1 = |\partial H_1| \ge \lceil c|H_1| \rceil > 1$). Therefore $G|_{\{x_1 := a_1\}}$ is connected graph. By the construction of f we get that $G|_{\rho_1}$ consists of one connected component and may contains some disconnected vertices.

Induction step. Assume that $G|_{\rho_i}$ consists of one connected component and probably of some vertices of zero degree. By construction $G|_{\rho_i}$ doesn't contain bridges, therefore $G|_{\rho_i \cup \{x_{i+1} := a_{i+1}\}}$ also consists of one connected component and of some vertices of zero degree. The application of f saves this property. □

Corollary 6.6. *Let $s = |\tau_\kappa|$. Then $G|_{\tau'_\kappa}$ contains exact $s + 1$ connected components H, $H^{(1)}, H^{(2)}, \dots H^{(s)}$, and all subformulas corresponding to $H^{(i)}$ are satisfied by the assignment ρ_κ.*

Proof. Each new component appears after removing a bridge from the graph. By the definition of τ_i, s is exact the number of removed bridges. □

By the construction of the algorithm $\mathcal{A}$ the size of the component $H^{(i)}$ is at most $r = n/2$, since $\mathcal{A}$ every time chooses smallest component. By the expansion property of the graph G: $|\partial H^{(i)}| \ge 3$ (this inequality is true if $|H^{(i)}| < 3$ since each vertex has degree at least 4 and if $|H^{(i)}| \ge 3$ we can estimate $|\partial H^{(i)}| \ge \lceil c|H^{(i)}| \rceil \ge 3$). Hence the size of assignment τ'_κ is at

least $\frac{3s}{2}$. On another hand the size of τ'_κ is at most $\kappa + s$ since τ'_κ contains exact s bridges and at most κ substitutions of the type $x_j := a_j$. Therefore $\kappa + s \geq |\tau'_\kappa| \geq 3s/2$ and $\kappa \geq s/2$. And we can bound the size of τ'_κ in the following way $|\tau'_\kappa| \leq \kappa + s \leq 3\kappa < \frac{cr}{4}$.

The size of $H^{(i)}$ is less then $r/2$, otherwise by the expansion property of the graph G: $|\partial H^{(i)}| \geq cr/2$, but the numbers of edges in the graphs $G_{\tau'_\kappa}$ and G differ by $|\tau'_\kappa| < \frac{cr}{4}$. By Lemma 3.1 there is cleaning procedure such that $H^{(i)} \subseteq \overline{Cl}(\tau'_\kappa)$ (since H_i is connected component in the graph G_κ).

By Lemma 3.3 $|\overline{Cl}(\tau'_\kappa)| \leq r/2 = n/4$. The closure $\overline{Cl}(\tau'_\kappa)$ consists of all $H^{(i)}$ and some strict subset L of vertices from the component H. The assignment ρ_κ satisfies all $H^{(i)}$. Since H is connected component in the graph $G|_{\rho_\kappa}$, ρ_κ can be extended to assignment σ, satisfying part of the formula $T_G|_{\rho_\kappa}$, that contains edges with at least one end in L. The assignment σ actually removes from G the set of vertices $\overline{Cl}(\tau'_\kappa)$ with incident edges. By Lemma 3.2 the graph G_σ is $(r/2, d, c/2)$-expander. □

Theorem 6.7. *Let G be (r, d, c)-expander, with $c > 1$, $r = n/2$, $d \geq 4$, and n is number of vertices. The degree of each vertex in G is at least 4. T_G is Tseitin formula with respect to G. Any static proof of formula T_G in $\boldsymbol{LS}_+$ has size $\exp(\Omega(n))$.*

Proof. Let P be a static $\mathbf{LS}_+$ proof of the T_G. We set $k = \lceil \frac{\epsilon n}{5} \rceil$, where ϵ is from Corollary 5.4 for an $(r/2, d, c/2)$-expander.

Let f be mapping from the Theorem 6.4. We define the sequence of assignments: $\rho_0 = \emptyset, \rho_1, \ldots, \rho_\kappa$, $\kappa = \lceil \frac{cr}{13} \rceil$. $\rho_i = f(\rho_{i-1} \cup \{x_i := a_i\})$, where $x_i := a_i$ are substitutions from Lemma 6.1 for the proof $P|_{\rho_{i-1}}$.

By the Theorem 6.4 there exists assignment σ such that σ extends ρ_κ and $P' = P|_\sigma$ is static proof of Tseitin formula with respect to $(r/2, d, c/2)$-expander.

Let M_0 denote the number of polynomials $g_{i,l}$ of degree at least k in P. Let us denote strictly positive constants $(1 - \epsilon/(5d))$ by D ($0 < \epsilon < 1$, therefore $0 < D < 1$) and $\frac{c}{26}$ by C.

Since each vertex has degree at most d, we can estimate the number of edges: $N = |E| \leq dn/2$. By Lemma 6.1, the refutation P' contains at most $M_0(1 - k/(2N))^\kappa \leq M_0(1 - \frac{\epsilon n/5}{dn})^\kappa \leq M_0 \cdot D^{Cn}$ nonzero polynomials $g'_{i,l}$ of degrees at least k. By Corollary 5.4 there is at least one polynomial $g'_{i,l}$ of degree at least $\epsilon n/4 > k$. Hence we have $M_0 \cdot D^{Cn} \geq 1$, i.e., $M_0 \geq (1/D)^{Cn}$, which proves the theorem. □

Corollary 6.8. *Any tree-like $\boldsymbol{LS}_+$ refutation of (2.8) for a connected d-regular $(r = n/2, d, c)$-expander G with n vertices and $c > 2$ has size $exp(\Omega(n))$.*

Proof. We can easily simulate any tree-like $\mathbf{LS}_+$ proof by a static $\mathbf{LS}_+$ proof and apply Theorem 6.7 afterwards. □

Acknowledgment

Authors are grateful to Dima Grigoriev, Edward A. Hirsch, Alexander S. Kulikov and Sergey I. Nikolenko for useful discussions and to anonymous referees for numerous comments that improved the quality of this paper.

References

[1] M. Alekhnovich, E. A. Hirsch, and D. Itsykson. Exponential lower bounds for the running time of DPLL algorithms on satisfiable formulas. *J. of Automated Reasoning*, 35:51–72, 2005.

[2] M. Alekhnovich and A. Razborov. Lower bounds for polynomial calculus: Non-binomial case. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science, FOCS'01*, pages 190–199, 2001.

[3] P. Beame, T. Pitassi, and N. Segerlind. Lower Bounds for Lovasz-Schrijver Systems and Beyond Follow from Multiparty Communication Complexity. In *Proceedings of the 32nd Annual Colloquium on Automata, Languages, and Programming, ICALP'05*, pages 1176–1188, 2005.

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

[5] M. Clegg, J. Edmonds, and R. Impagliazzo. Using the Groebner basis algorithm to find proofs of unsatisfiability. In *Proceedings of the 28th annual ACM symposium on Theory of computing, STOC'96*, pages 174–183, 1996.

[6] S. A. Cook and R. A. Reckhow. The Relative Efficiency of Propositional Proof Systems. *The Journal of Symbolic Logic*, 44(1):36–50, 1979.

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

[8] I. Dinur. The PCP theorem by gap amplification. In *Proceedings of the 38th annual ACM symposium on Theory of computing, STOCS'06*, pages 241–250, 2005.

[9] R. E. Gomory. An algorithm for integer solutions of linear programs. In R. L. Graves and P. Wolfe, editors, *Recent Advances in Mathematical Programming*, pages 269–302. McGraw-Hill, 1963.

[10] D. Grigoriev. Linear lower bound on degrees of Positivstellensatz Calculus proofs for the Parity. *Teoretical Computer Science*, 259:613–622, 2001.

[11] D. Grigoriev and E. A. Hirsch. Algebraic proof systems over formulas. *Teoretical Computer Science*, 303:83–102, 2003.

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

[13] D. Grigoriev and N. Vorobjov. Complexity of Null- and Positivestellensatz proofs. *Annals of Pure and Applied Logic*, 113(1-3):153–160, 2001.

[14] L. Lovász. Stable sets and polynomials. *Discrete Mathematics*, 124:137–153, 1994.

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

[16] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[17] G. A. Margulis. Explicit construction of concetrators. *Problemy Peredachi Informatsii*, 9(4):71–80, 1973. English translation: Problems of information transmission, pages 325–332, 1973.

[18] R. Murty. Ramanujan graphs. *Journal of the Ramanujan Math. Society*, 18(1):1–20, 2003.

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

[20] O. Reingold. Undirected st-connectivity in log-space. In *Proceedings of the 37th annual ACM symposium on Theory of computing, STOCS'05*, pages 376–385, 2005.

[21] G. S. Tseitin. On the complexity of derivation in the propositional calculus. *Zapiski nauchnykh seminarov LOMI*, 8:234–259, 1968. English translation: Consultants Bureau, N.Y., 1970, pp. 115–125.

[22] A. Urquhart. Hard examples for resolution. *J. ACM*, 34(1):209–219, 1987.