Resolution complexity of perfect matching principles for sparse graphs

Dmitry Itsykson[1], Mikhail Slabodkin[2], and Dmitry Sokolov[1]

September 15, 2014

[1]Steklov Institute of Mathematics at St.Petersburg
[2]St. Petersburg Academic University

Introduction

We construct a family of graphs G_n with the resolution complexity of the perfect matching principle $2^{\Omega(n)}$.

- First exponential lower bound for PMP in the form $2^{\Omega(n)}$, where n is the number of variables.
- Matches upper bound.
- Implies several known lower bounds ($\mathrm{PHP}_{n,m}$) and improves some of them (PMP_{K_n}).

Resolution proof system

Definition

$\varphi = C_1 \wedge C_2 \wedge \ldots \wedge C_k$ — unsatisfiable CNF.
Resolution proof: $C_{i_1}, C_{i_2}, \ldots, C_{i_l}$

1. $C_{i_l} = \bot$.
2. Every C_{i_j} is either contained in φ or is obtained using resolution rule:

$$\frac{x \vee A \qquad\qquad \neg x \vee B}{A \vee B}$$

Definition

A family of unsatisfiable formulas F_n is weaker than H_n if for some m for all clauses $C \in H_n$, C is an implication of $\bigwedge_{i=1}^{m} C_i$, where C_i is a clauses of F_n.

Pigeonhole principle

PHP$_n^m$: m pigeons, n holes. Variables $\{p_{i,j}\}\, i = 1..m, j = 1..n$.
PHP$_n^m$ is a conjunction of statements:

- Every pigeon is contained in at least one hole.

$$\bigwedge_i (p_{i,1} \vee p_{i,2} \vee \ldots \vee p_{i,m})$$

- Every hole contains at most one pigeon.

$$\bigwedge_j (\neg p_{1,j} \vee \neg p_{2,j}) \wedge (\neg p_{1,j} \vee \neg p_{3,j}) \wedge \ldots \wedge (\neg p_{m-1,j} \vee \neg p_{m,j})$$

— Haken, 1985: $2^{\Omega(n)}$ for $m = n + 1$.
— Razborov, 2001: $2^{\Omega(n^{\frac{1}{3}})}$ for any $m > n$.

G-PHP$_n^m$: restriction on a particular bipartite graph G.
— Ben-Sasson, Wigderson, 2001: $2^{\Omega(n)}$ for $m = O(n)$ and G is a bipartite constant degree expander.

FPHP_n^m and Perfect matching

FPHP_n^m: weakening of PHP_n^m,

- Every pigeon is contained in at most one hole.

— Razborov, 2001: lower bound $2^{\Omega\left(\frac{n}{(\log m)^2}\right)}$, which implies $2^{\Omega\left(n^{1/3}\right)}$.

PMP_G: for some graph $G(V,E)$ a formula PMP_G encodes that G has a perfect matching. We assign a binary variable x_e for all $e \in E$. PMP_G is the conjunction of the conditions:

- For all $v \in V$ at least one edge that incident to v has value 1: $\bigvee_{(v,u)\in E} x_{(v,u)}$.
- For any pair of edges e_1, e_2 incident to v at most one of them takes value 1, $\neg x_{e_1} \vee \neg x_{e_2}$.

— Razborov, 2004: resolution complexity is at least $2^{\frac{\delta(G)}{\log^2 n}}$, where $\delta(G)$ is the minimal degree and n is the number of vertices.

Results

Theorem 1

$\exists D$ such that $\forall C\ \forall n\ \forall m \in [n+1, Cn]$ there exists such bipartite $G(X, Y, E)$ such that

- *G is explicit with maximum degree $\leq D$, $|X| = m, |Y| = n$.*
- *PMP$_{G_{n,m}}$ is unsatisfiable and refutable in at least $2^{\Omega(n)}$.*

The number of variables in PMP$_{G_{n,m}}$ is $O(n)$, therefore the lower bound matches (up to an application of a polynomial) the trivial upper bound $2^{O(n)}$ that holds for every formula with $O(n)$ variables.

Theorem 1 corollaries

- $\mathrm{PMP}_{G_{n,m}}$ is weaker than $G_{m,n}-\mathrm{PHP}_n^m$, PHP_n^m and FPHP_n^m, therefore Theorem 1 implies the same lower bound for $G_{m,n}-$ PHP_n^m, PHP_n^m and FPHP_n^m.
- The resolution complexity of $\mathrm{PMP}_{K_{m,n}}$ is $2^{\Omega(n)}$ where $m = O(n)$, which improves $2^{\Omega(n/\log^2 n)}$ (Razborov, 2004) and matches the upper bound $n2^n$ that follows from the upper bound for PHP_n^{n+1}.
- The lower bound for the resolution complexity of PMP_{K_n} is $2^{\Omega(n)}$, which improves the lower bound $2^{\Omega(n/\log^2 n)}$ (Razborov, 2004).

Boundary expanders, refutation width

Definition

A bipartite graph G with parts X and Y is a (r, c)-*boundary expander* if $\forall A \subseteq X$, if $|A| \leq r$ then $|\delta(A)| \geq c|A|$, where $\delta(A)$ is the set of all vertices in Y that are connected with exactly one vertex in A;

Definition

Ben-Sasson, Wigderson, 2001:

- *Width of the clause* $w(C)$ is a number of literals in C.
- *Width of the formula* $w(\varphi)$ is a maximum width of the clause in it.
- $w(\varphi)$ is *refutable in width* w if there exists refutation with maximum width of the clauses w.

Theorem (Ben-Sasson, Wigderson)

For any k-CNF unsatisfiable formula φ with n variables the size of resolution proof is at least $2^{\Omega\left(\frac{(w-k)^2}{n}\right)}$, where w is a minimal width of a resolutional proof.

Width-size connection

Theorem 2

Let G be a (r, c)-boundary expander with parts X and Y such that there is a matching in G that covers all vertices from Y. Then the width of all resolution proofs of PMP_G is at least $cr/2$.

If degrees of all vertices are at most D, then the size of any resolution proof of PHP_G is at least $2^{\Omega\left(\frac{(cr/2-D)^2}{n}\right)}$, where n is the number of edges in G.

Lemma (Itsykson, Sokolov, 2011)

$\forall\, d\ \forall\, C$ *and* $\forall\, n$ *and* $m \in [n+1, Cn]$ *there is an explicit construction of* $(r, 0.4d)$*-boundary expander* $G(X, Y, E)$ *with* $|X| = m$*,* $|Y| = n$ *and* $r = \Omega(n)$ *such that all degrees are bounded by* d^2*.*

Now Theorem 2 and Lemma imply Theorem 1.

Generalization

- $G(V, E)$ is an undirected graph.
- h is a function $V \to \mathbb{N}$.
- variables $\{x_e\}$ correspond to E.
- $\Psi_G^{(h)}$: $\forall v \in V$ exactly $h(v)$ edges $e_{v,u}$ have value 1.
- PMP$_G$ is a particular case of $\Psi_G^{(h)}$ for $h \equiv 1$.

Theorem 3

$\forall\ d \in \mathbb{N}\ \forall\ n$ large enough and $\forall\ h : V \to \{1, 2, \ldots, d\}$, where $|V| = n$, there exists such explicit $G(V, E)$, that $\Psi_G^{(h)}$ is unsatisfiable and the refutation size for $\Psi_G^{(h)}$ is at least $2^{\Omega(n)}$.

Theorem 2 corollaries

- **Tseitin formulas**. Let $G(V, E)$ be an arbitrary and $f : V \to \{0, 1\}$; variables x_e of $T_G^{(f)}$ correspond to E.

$$T_G^{(f)} = \bigwedge_{v \in V} \left(\bigoplus_{(v,u) \in E} x_{(v,u)} = f(v) \right)$$

  Let $h(v) = 2 - f(v)$. By Theorem 3 there exists G with n vertices of degree at most D such that the size of any resolution proof of the formula Ψ_G^h is at least $2^{\Omega(n)}$. Every condition of $T_G^{(h)}$ may be derived from a condition of Ψ_G^h in 2^D steps. Thus resolution complexity of $T_G^{(f)}$ is at least $2^{\Omega(n)}$ (— Urquhart, 1987).
- **Complete graph**. Let $h : V \to \{0, 1, \ldots, d\}$ be defined on the graph K_n and let formula $\Psi_{K_n}^{(h)}$ be unsatisfiable. By Theorem 3 there exists G with n vertices of bounded degree that the size of any resolution proof of Ψ_G^h is at least $2^{\Omega(n)}$. Formula $\Psi_G^{(h)}$ can be obtained from $\Psi_{K_n}^{(h)}$ by substituting zeroes to some edges, therefore the size of the resolution proof of $\Psi_{K_n}^{(h)}$ is at least $2^{\Omega(n)}$.