Approximating Bounded Degree Instances of NP-Hard Problems*

Marek Karpinski†

Abstract

We present some of the recent results on computational complexity of approximating bounded degree combinatorial optimization problems. In particular, we present the best up to now known explicit nonapproximability bounds on the very small degree optimization problems which are of particular importance on the intermediate stages of proving approximation hardness of some other generic optimization problems.

1 Introduction

An interesting approximation hardness phenomenon of combinatorial optimization was discovered in [PY91] and [ALMSS92], to the effect that the bounded degree instances of several optimization problems are hard to approximate to within an arbitrary constant. This fact seemed to be a bit puzzling at the time as bounded degree instances of many optimization problems were known to have trivial approximation algorithms dramatically improving performances of the best known approximation algorithms on general instances. An interesting artifact on their complementary, i.e., dense, instances was also the existence of polynomial time approximation schemes (PTASs) for them [AKK95], [KZ97], see [K01]

†Dept. of Computer Science, University of Bonn, 53117 Bonn. Supported in part by DFG grants, DIMACS, and IST grant 14036 (RAND-APX), and by the Max-Planck Research Prize. Research partially done while visiting Department of Computer Science, Yale University. Email: marek@cs.uni-bonn.de.
for a survey. We discuss here explicit approximation lower bounds for bounded
degree instances with a very small bounds on degrees like 3 or 4, and also the
best known approximation algorithms on that instances. These instances have
turned out to be particularly important at the intermediate reduction stages for
proving hardness of approximation of some other important optimization prob-
lems, like Set Cover, some restricted versions of Traveling Salesman Problem, and
the problem of Sorting by Reversals motivated recently by molecular biology, cf.
[F98], [PY91], [PY93], [BK99], [FK99], [E99], [EK00]. We mention here some
interesting new results on asymptotic relations between hardness of approxima-
tion and bounds on a degree of instances [H00], [T01]. These results do not yield
though explicit lower approximation bounds for small degree instances needed in
applications mentioned before.

We survey in this paper the best known up to now explicit approximation
lower bounds for the small degree (number of variable occurrences) optimization
problems, like the problems of maximization or minimization of the sat-
isfiability of systems of linear equations mod 2, MAX-CUT, MAX- and MIN-
BISECTION, MAX-2SAT, MAXIMUM INDEPENDENT SET, and MINIMUM
NODE COVER [BK99], [BK01b]. We move on, and apply these results to get ex-
licit lower approximation bounds for the problem of Sorting by Reversals [BK99],
and the Traveling Salesman Problem with distances one and two [EK00]. Finally,
we mention recent improvement on approximation ratios of algorithms for small
degree MAX-CUT and MAX-BISECTION problems based on local enhancing
methods for semidefinite programming [FKL00a], [FKL00b], [KCL00].

2 Bounded Degree Maximization Problems

We are going to define basic optimization problems of this section.

• MAX-$k$-LIN2: Given a set of equations mod 2 with exactly $k$ variables
  per equation, construct an assignment maximizing the number of equations
  satisfied.

• $b$-OCC-MAX-$k$-LIN2: Given a set of equations mod 2 with exactly $k$ vari-
  ables per equation and the number of occurrences of each variable bounded
  by $b$, construct an assignment maximizing the number of equations satisfied.

• $b$-OCC-MAX-HYBRID-LIN2: Given a set of equations mod 2 with exactly
  two or three variables per equation, and the number of occurrences of each
  variable bounded by $b$, construct an assignment maximizing the number of
  equations satisfied.

• $b$-OCC-MAX-2SAT: Given a conjunctive normal form formula with two
  variables per clause, construct an assignment maximizing the number of
  clauses satisfied.
• \textit{d-MAX-CUT}: Given an undirected graph of degree bounded by \textit{d}, partition its vertices into two groups so as to maximize the number of edges with exactly one endpoint in each group.

• \textit{d-MIS}: Given an undirected graph of degree bounded by \textit{d}, construct a maximum cardinality subset of vertices such that no two vertices of it are adjacent.

We are going to display now approximation preserving reductions which reduce from the \textit{MAX-E2-LIN2} and the \textit{MAX-E3-LIN2} problems. The method of reductions depends on a new \textit{wheel-amplifier} construction of Berman and Karpinski [BK99] designed specially for bounded degree problems. This kind of amplifier has turned out to be more efficient than the standard expander amplifiers (cf. e.g., Arora and Lund [AL97]) for small degree, and number of occurrences, optimization problems.

We start with the following known inapproximability results of Håstad [H97].

\textbf{Theorem 1.} ([H97]) \textit{For any } \(0 < \epsilon < \frac{1}{2}\), \textit{it is NP-hard to decide whether an instance of MAX-E2-LIN2 with } \(16n\) \textit{equations has its optimum value above } \((12 - \epsilon)n\) \textit{or below } \((11 + \epsilon)n\).

\textbf{Theorem 2.} ([H97]) \textit{For any } \(0 < \epsilon < \frac{1}{2}\), \textit{it is NP-hard to decide whether an instance of MAX-E3-LIN2 with } \(2n\) \textit{equations has its optimum value above } \((2 - \epsilon)n\) \textit{or below } \((1 + \epsilon)n\).

In Berman and Karpinski [BK99] the following polynomial time randomized approximation preserving reductions were constructed:

• \(f_1 : \text{MAX-E2-LIN2} \to \text{3-OCC-MAX-E2-LIN2}\),

• \(f_2 : \text{MAX-E2-LIN2} \to \text{3-MAX-CUT}\),

• \(f_3 : \text{MAX-E3-LIN2} \to \text{3-OCC-MAX-HYBRID-LIN2}\),

The constructions for \(f_1\), and \(f_2\) use variants of \textit{wheel-amplifier} methods, whereas a construction for \(f_3\) uses certain 3-hypergraph extension of it. The following optimizing properties of \(f_1, f_2,\) and \(f_3\) were proven in [BK99].

\textbf{Theorem 3.} ([BK99]) \textit{For any } \(0 < \epsilon < \frac{1}{7}\), \textit{it is NP-hard to decide whether an instance of } \(f_1(\text{MAX-E2-LIN2}) \in \text{3-OCC-MAX-E2-LIN2}\) \textit{with } \(336\) \textit{edges has its optimum value above } \((332 - \epsilon)n\) \textit{or below } \((331 + \epsilon)n\).

A similar result can be proven for \(f_2\), and the 3-MAX-CUT-problem.
Theorem 4. ([BK99]) For any $0 < \epsilon < \frac{1}{2}$, it is NP-hard to decide whether an instance of $f_3(\text{MAX-E2-LIN2}) \in \text{3-MAX-CUT}$ with $336$ edges has its optimum value above $(332 - \epsilon)n$ or below $(331 + \epsilon)n$.

For $f_3$ and MAX-HYBRID-LIN2 we have

Theorem 5. ([BK99]) For any $0 < \epsilon < \frac{1}{2}$, it is NP-hard to decide whether an instance of $f_3(\text{MAX-E2-LIN2}) \in \text{3-OCC-MAX-HYBRID-LIN2}$ with $60n$ equations with exactly two variables and $2n$ equations with exactly three variables has its optimum value above $(62 - \epsilon)n$ or below $(61 + \epsilon)n$.

Theorem 4 can be also used to derive the following bound for 3-OCC-MAX-2SAT.

Theorem 6. ([BK99]) For any $0 < \epsilon < \frac{1}{2}$, it is NP-hard to decide whether an instance of 3-OCC-MAX-2SAT, with $2016n$ clauses has its optimum above $(2012 - \epsilon)n$ or below $(2011 + \epsilon)n$.

The 3-OCC-MAX-HYBRID-LIN2 problem and Theorem 5 can be used to derive lower bounds for 4-MIS problem, and using some more subtle construction, even for 3-MIS problem.

Theorem 7. ([BK99]) For any $0 < \epsilon < \frac{1}{2}$, it is NP-hard to decide whether an instance of 4-MIS with $152n$ nodes has its optimum value above $(74 - \epsilon)n$ or below $(73 + \epsilon)n$, and whether an instance of 3-MIS with $284n$ nodes has its optimum value above $(140 - \epsilon)n$ or below $(139 + \epsilon)n$.

The results above imply the following explicit nonapproximability results.

Corollary 1. For every $\epsilon > 0$, it is NP-hard to approximate:

1. 3-OCC-MAX-E2-LIN2 and 3-MAX-CUT to within a factor $332/331 - \epsilon$,
2. 3-OCC-MAX-HYBRID-LIN2 to within a factor $62/61 - \epsilon$,
3. 3-OCC-MAX-2SAT to within a factor $2012/2011 - \epsilon$,
4. 4-MIS to within a factor $74/73 - \epsilon$, and 3-MIS to within a factor $140/139 - \epsilon$.

The best to our current knowledge gaps between upper and lower approximation bounds are summarized in Table 1. The upper approximation bounds are from [GW94], [BF94], [BF95], [FG95], [FKL00a]. The technical results of this section will be used also later on in our paper.
TABLE 1:
Bounded Degree Maximization Problems

<table>
<thead>
<tr>
<th>Problem</th>
<th>Approx. Upper</th>
<th>Approx. Lower</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-OCC-MAX-E2-LIN2</td>
<td>1.0858</td>
<td>1.0030</td>
</tr>
<tr>
<td>3-OCC-MAX-IIYBRID-LIN2</td>
<td>2</td>
<td>1.0163</td>
</tr>
<tr>
<td>3-MAX-CUT</td>
<td>1.0858</td>
<td>1.0030</td>
</tr>
<tr>
<td>3-OCC-MAX-2SAT</td>
<td>1.0741</td>
<td>1.0005</td>
</tr>
<tr>
<td>3-MIS</td>
<td>1.2</td>
<td>1.0071</td>
</tr>
<tr>
<td>4-MIS</td>
<td>1.4</td>
<td>1.0136</td>
</tr>
</tbody>
</table>

3 Bounded Degree Minimization Problems

We are going to introduce now the following minimization problems.

- **d-Node Cover**: Given an undirected graph of degree bounded by $d$, construct a minimum cardinality subset of vertices such that each edge of a graph has at least one of its endpoints in it.

- **MIN-E$k$-LIN2**: Given a set of equations mod 2 with exactly $k$ variables per equation, construct an assignment minimizing the number of equations satisfied.

- **b-OCC-MIN-E$k$-LIN2**: Given a set of equations mod 2 with exactly $k$ variables per equation and the number of occurrences of each variable exactly equal to $b$, construct an assignment minimizing the number of equations satisfied.

- **MIN-BISECTION**: Given an undirected graph, partition the vertices into two equal halves so as to minimize the number of edges with exactly one endpoint in each half.

- **d-MIN-BISECTION**: Given a $d$-regular graph, partition the vertices into two equal halves so as to minimize the number of edges with exactly one endpoint in each half.

We will specialize now techniques of Section 2 to obtain lower approximation bounds on bounded degree minimization problems.
We start with a direct application of Theorem 7 towards d-Node Cover problem. For a given undirected graph $G = (V, E)$, and a maximum independent set $I$ of $G$, $V \setminus I$ is a minimum node cover of $G$. We take now an instance of 4-MIS with $152n$ nodes. It is NP-hard, for any $0 < \epsilon < \frac{1}{2}$, to decide whether 4-Node Cover has its optimum value above $(152 - 73 - \epsilon)n = (79 - \epsilon)n$ or below $(152 - 74 + \epsilon)n = (78 + \epsilon)n$. Similarly for 3-Node Cover. Thus we have

**Theorem 8.** For any $0 < \epsilon < \frac{1}{2}$, it is NP-hard to decide whether an instance of 4-Node Cover with $152n$ nodes has its optimum value above $(79 - \epsilon)n$ or below $(78 + \epsilon)n$, and whether an instance of 3-Node Cover with $284n$ has its optimum value above $(145 - \epsilon)n$ or below $(144 + \epsilon)n$.

Theorem 8 gives the following approximation lower bounds for 4-Node Cover and 3-Node Cover problems.

**Corollary 2.** For every $\epsilon > 0$, it is NP-hard to approximate

1. 3-Node Cover to within a factor $145/144 - \epsilon$,

2. 4-Node Cover to within a factor $79/78 - \epsilon$.

We turn now to the bounded occurrence minimum satisfiability of linear equations.

We need the following recent result of Dinur, Kindler, Raz and Safra [DKRS00] (see also [DKS98], [KST97]).

**Theorem 9.** ([DKRS00]) MIN-LIN2 is NP-hard to approximate to within a factor $n^{c/\log n}$ for some constant $c$.

MIN-LIN2 is equivalent to the well known Nearest Codeword problem (cf. [ABSS93]). Only very recently the first sublinear approximation ratio $O(n/\log n)$ algorithm was designed by Berman and Karpinski [BK01b].

We introduce now a notion of an $(r, t)$-approximation algorithm. For two functions $r$ and $t$, we call an approximation algorithm $A$ for an optimization problem $P$, an $(r(n), t(n))$-approximation algorithm if $A$ approximates $P$ within an approximation ratio $r(n)$ and $A$ works in $O(t(n))$ time for $n$ a size of an instance.

Berman and Karpinski [BK01b] proved the following result on the $(r, t)$-approximations of the 3-OCC-MIN-E3-LIN2 problem.

**Theorem 10.** ([BK01b]) There exists a constant $c$ such that if there exists an $(r(n), t(n))$-approximation algorithm for 3-OCC-MIN-E3-LIN2, then there exists an $(r(cn), t(cn))$-approximation algorithm for MIN-LIN2.

Theorem 9 entails now
Theorem 11. The problem 3-OCC-E3-LIN2 is NP-hard to approximate to within a factor $n^{\epsilon/\log n}$ for some constant $c$.

The 3-OCC-MIN-E3-LIN2 problem is equivalent to the exactly-3 bounded occurrence 3-ary Nearest Codeword problem (c.f. [KST97]), and therefore we have

Corollary 3. The 3-ary Nearest Codeword problem with the number of occurrences of each variable exactly equal to 3 is NP-hard to approximate to within a factor $n^{\epsilon/\log n}$ for some constant $c$.

We apply a similar technique for the problem of MIN-BISECTION. Here our result will be only relative to the approximation hardness of MIN-BISECTION, the status of which is wide open, and we know currently of no proof technique which excludes existence of a PTAS for that problem.

Somewhat surprisingly in that context, Berman and Karpinski [BK01b] proved the following result on approximation hardness of bounded degree version of MIN-BISECTION.

Theorem 12. ([BK01b]) If there exists an $(r(n), t(n))$-approximation algorithm for 3-MIN-BISECTION, then there exists an $(r(n^3), t(n^3))$-approximation algorithm for MIN-BISECTION.

The best currently known approximation algorithm for the MIN-BISECTION is of ratio $O(\log^2 n)$ due to Feige and Krauthgamer [FK00]. Any asymptotic improvement on approximation ratio $r$ for 3-regular graphs, say $r = o(\log^2 n)$, will entail, by Theorem 12, an improvement on an approximation ratio for the general MIN-BISECTION.

A similar technique can be also used to prove approximation hardness result for the general planar MIN-BISECTION of the planar MIN-BISECTION problem on 3-regular graphs.

4 Some Application

We are going to apply our previous results for some other generic optimization problems. The first problem is one of the most important problems in analysis of genome rearrangements, and it is being recently also motivated by other algorithmic problems of computational molecular biology.

- MIN-SBR (Sorting by Reversals): Given a permutation, construct a minimum length sequence of reversals (see for definitions [BP96]) which transforms it to the identity permutation.
We refer also to some other variants of Sorting by Reversals problems studied in [C99], called MSBR and Tree SBR (see the definitions there).

The proof technique used in [BK99] to prove explicit approximation lower bound of Theorem 7 for 4-MIS can be adapted to prove for the first time the inapproximability of MIN-SBR, and, in fact, also giving an explicit approximation lower bound for that problem.

**Theorem 13.** ([BK99]) For every $\epsilon > 0$, it is NP-hard to approximate MIN-SBR within a factor $1237/1236 - \epsilon$.

Caprara [C99] has used the above result to prove inapproximability of the both aforementioned problems, MSBR, and Tree SBR, and to compute the first explicit approximation lower bounds for those problems.

We turn now to another application of the results of Section 2. We denote by (1,2)-TSP the Traveling Salesman Problem with distances one and two, and its asymmetric version by (1,2)-ATSP (cf. [PY93], [V92]).

Engebretsen and Karpinski [EK00] has used recently the result on 3-OCC-MAX-HYBRID-LIN2 of Theorem 5 to prove the following explicit inapproximability result for (1,2)-ATSP problem.

**Theorem 14.** ([EK00]) For every $\epsilon > 0$, it is NP-hard to approximate (1,2)-ATSP within a factor $321/320 - \epsilon$.

The construction used by Engebretsen and Karpinski [EK00] could be also adapted to yield an explicit result for (1,2)-TSP.

**Theorem 15.** ([EK00]) For every $\epsilon > 0$, it is NP-hard to approximate (1,2)-TSP within a factor $743/742 - \epsilon$.

## 5 New Upper Approximation Bounds

The intricacy of proving the first explicit approximation lower bounds for small degree optimization problems, and the resulting huge gaps between upper and lower approximation bounds has stimulated research on improving approximation ratios for those problems as well as for some other generic problems.

The first gap for 3-MAX-CUT (and 3-OCC-MAX-E2-LIN2) was improved recently by Feige, Karpinski and Langberg [FKL00a], see Table 1. The technique of [FKL00a] is based on a new local enhancing method for semidefinite programs.

**Theorem 16.** ([FKL00a]) There exists a polynomial time algorithm approximating 3-MAX-CUT within a factor 1.0858.
We note that the best approximation ratio currently known for MAX-CUT problem on general graphs is 1.1383 ([GW94]), and the best known approximation lower bound is 1.0624 [H97]. We note also that for the semidefinite relaxation of MAX-CUT used in [GW94], the integrality gap is at least 1.1312, even for 2-regular graphs. Thus the bound of Theorem 16 beats the integrality bound even for 2-regular graphs.

We turn now to the special case of regular bounded degree graphs, and will investigate approximation algorithms for the MAX-CUT and MAX-BISECTION (partitioning of a graph into two halves so as to maximize a number of the edges between them).

Rd-MAX-CUT and Rd-MAX-BISECTION are the MAX-CUT and MAX-BISECTION problems, respectively, restricted to $d$-regular graphs.

Feige, Karpinski and Langberg [FKL00a], [FKL00b] were able to improve the best known approximation ratios for both bounded degree problems, Rd-MAX-CUT, and Rd-MAX-BISECTION. The best known approximation ratio for MAX-BISECTION on general graphs is 1.4266 [HZ00].

**Theorem 17.** ([FKL00a], [FKL00b]) There are polynomial time algorithms that approximate $R_d$-MAX-CUT and $R_d$-MAX-BISECTION problems within factor 1.0823 and 1.1991, respectively.

Using an additional local advancement method, Karpinski, Kowaluk and Lingas [KKL00], have further improved approximation ratios of the low degree Rd-MAX-BISECTION problems.

**Theorem 18.** ([KKL00]) There exists a polynomial time algorithm approximating $R_d$-MAX-BISECTION within a factor 1.1806.

Interestingly, the first improvements on approximation ratios of MAX-BISECTION on low degree planar graphs undertaken in [KKL00] has lead to design of the first PTASs for the general planar MAX-BISECTION as well as for other geometrically defined classes of graphs (see [JKLS01]).

On the lower bounds side, we note that the techniques of [BK99] yield also the best up to now explicit approximation lower bounds for $R_d$-MAX-CUT, and $R_d$-MAX-BISECTION problems equal to the lower approximation bound for 3-MAX-CUT problem of Section 2.

6 Summary of Approximation Results on Bounded Degree Minimization Problems

We present here (Table 2) the results of Section 3 and 4 on bounded degree minimization problems and the best to our knowledge gaps between upper and lower
approximation bounds on those problems. The upper approximation bounds are from [BF94], [BF95], [BK01b], [FK00], [BHK01], [V92], [PY93].

**TABLE 2:**
Bounded Degree and Weight Minimization Problems

<table>
<thead>
<tr>
<th>Problem</th>
<th>Approx. Upper</th>
<th>Approx. Lower</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-Node Cover</td>
<td>1.1666</td>
<td>1.0069</td>
</tr>
<tr>
<td>4-Node Cover</td>
<td>1.2857</td>
<td>1.0128</td>
</tr>
<tr>
<td>3-OCC-MIN-E3-LIN2</td>
<td>O(n/logn)</td>
<td>n$^{\Omega(1)/\log\log n}$</td>
</tr>
<tr>
<td>3-MIN-BISECTION</td>
<td>O(log$^2 n$)</td>
<td>Equal to MIN-BISECTION</td>
</tr>
<tr>
<td>MIN-SBR</td>
<td>1.375</td>
<td>1.0008</td>
</tr>
<tr>
<td>(1,2)-TSP</td>
<td>1.1667</td>
<td>1.0013</td>
</tr>
<tr>
<td>(1,2)-ATSP</td>
<td>1.4167</td>
<td>1.0031</td>
</tr>
</tbody>
</table>

7 Open Problems and Further Research

An obvious open problem is to improve on both the lower and upper approximation bounds of bounded degree optimization problems, especially on those with the very small degree bounds. The essential improvements on the explicit lower bounds for these problems might be of paramount difficulty though, but same time they are also of great interest. Any such improvement would have immediate effects on the explicit lower bounds for other optimization problems, as indicated in this paper. Perhaps somewhat easier undertaking would be improving on upper approximation bounds. Here essential improvements were already achieved on the problems like a small degree MAX-CUT, and MAX-BISECTION mentioned in Section 5. How about improvements on other bounded degree optimization problems?

References


