Electronic Colloquium on Computational Complexity, Report No. 97 (2015)

Towards Better Inapproximability Bounds for TSP: A Challenge of Global Dependencies *

Marek Karpinski †

Abstract. We present in this paper some of the recent techniques and methods for proving best up to now explicit approximation hardness bounds for metric symmetric and asymmetric *Traveling Salesman Problem* (TSP) as well as related problems of Shortest Superstring and Maximum Compression. We attempt to shed some light on the underlying paradigms and insights which lead to the recent improvements as well as some inherent obstacles for further progress on those problems.

1 Introduction

The metric *Traveling Salesman Problem* (TSP) is one of the best known and broadly studied combinatorial optimization problems. Nevertheless its approximation status remained surprisingly elusive and very resistant for any new insights even after several decades of research. Basically, there were no improvements of the approximation algorithm of Christofides [C76] for the general metric TSP, and also a very slow improvement of the explicit inapproximability bounds for that problem [PY93], [E03], [EK06], [PV06], [L12], [KS12], [KS13a], [KS13b], [KLS13], [AT15]. The attainable values of

*A version of this paper appears in Proc. 20th Symp. on Fundamentals of Computation Theory, FCT '15, August 17–19, 2015.

†Department of Computer Science and the Hausdorff Center for Mathematics, University of Bonn. Research supported by DFG grants and the Hausdorff grant EXC59-1. Email: marek@cs.uni-bonn.de

ISSN 1433-8092

the explicit inapproximability bounds, and especially the methods for proving them, could give valuable insights into the algorithmic nature of the problem at hand. Unfortunately, there is still a huge gap between upper and lower approximation bounds for TSP. The best upper bound stands at the moment still firmly at 50% (approximation ratio 1.5). There were however several improvements of underlying approximation ratios for special cases of metric TSP like $(1, 2)$-metric [BK06] and the Graphic TSP [MS11], [SV14]. Also the corresponding inapproximability bounds for those special instances were established in [E03], [EK06], [KS13a], [KS13b].

We discuss also some new results on related problems of the *Shortest Superstring*, *Maximum Asymmetric TSP* and *Maximum Compression Problem* (cf. [KS13a]).

In this paper we introduce an essentially different method from the earlier work of [PV06] to attack the general problem. This method uses some new ideas on small occurrence optimization. The inspiration for it came from the constructions used for restricted cases of TSP in [E03] and [EK06].

2 Underlying Idea

The general idea is to use somehow instances of metric TSP to solve approximately another instance of optimization problems with provable inapproximability bounds. Thus, establishing the possible approximation hardness barriers for the TSP-solver itself. The reduction must be approximation preserving, i.e. validate *goodness* of feasible solutions of a given problem and also validating *goodness* of a corresponding tour. The direction from the tour to a feasible solution would be crucial for that method. To give a simple illustration, we start from that optimization problem and construct an instance of the TSP. We have to establish now the correspondence between the solutions of the problem and the tours of TSP. That correspondence must satisfy the crucial property that the problem has a *good* solution if and only if TSP has a *short* tour. The second direction from TSP to the optimization problem seems conceptually at the first glance more difficult. That intuition is correct and the main effort has been devoted to that issue. The suitable optimization problem will be a specially tailored bounded occurrence optimization problem of Sections 4–6.

3 TSP and Related Problems

We are going to define now main optimization problems of the paper.

- Metric TSP (here TSP for short): Given a metric space (V, d) (usually given by a complete weighted graph or a connected weighted graph). Construct a shortest tour visiting every vertex exactly once.

- Asymmetric Metric TSP (ATSP): Given an asymmetric metric space (V, d), d may be asymmetric. Construct a shortest tour visiting every vertex exactly once.

- Graphic TSP: Given a connected graph $G = (V, E)$. Construct a shortest tour in the *shortest path* metric completion of G or equivalently construct a smallest Eulerian spanning multi-subgraph of G.

- Shortest Superstring Problem (SSP): Given a finite set of strings S. Construct a shortest superstring such that every string in S is a substring of it.

- Maximum Compression Problem (MCP): Given a finite set of strings S. Construct a superstring of S with maximum *compression* which is the difference between the sum of the lengths of the strings in S and the length of the superstring.

- Maximum Asymmetric Traveling Salesman Problem (MAX-ATSP): Given a complete directed graph with nonnegative weights. Construct a tour of maximum weight visiting every vertex exactly once.

4 Bounded Occurrence Optimization Problems

We introduce here a notion of bounded occurrence optimization playing important roles in our construction.

- MAX-E3-LIN2: Given a set of equations mod 2 with exactly 3 variables per equation. Construct an assignment maximizing the number of equations satisfied.

- 3-Occ-MAX-HYBRID-LIN2: Given a set of equations mod 2 with exactly 2 or 3 variables per equation and the number of occurrences of each variable being bounded by 3.

The approximation hardness of 3-Occ-MAX-HYBRID-LIN2 problem was proven for the first time by Berman and Karpinski [BK99] (see also [K01], [BK03]) by randomized reduction f from MAX-E3-LIN2 and the result of Håstad [H01] on that problem, f : MAX-E3-LIN2 $\to$ 3-Occ-MAX-HYBRID-LIN2.

Theorem 1 ([BK99])**.** *For every $0 < \varepsilon < \frac{1}{2}$, it is NP-hard to decide whether an instance of f(MAX-E3-LIN2) $\in$ 3-Occ-MAX-HYBRID-LIN2 with $60n$ equations with exactly two variables and $2n$ equations with exactly three variables has its optimum value above $(62-\varepsilon)n$ or below $(61+\varepsilon)n$.*

The above result will be used in the simulational approximation reduction g : 3-Occ-MAX-HYBRID-LIN2 $\to$ TSP to the instances of the metric TSP.

5 Bi-wheel Amplifier Graphs

We shortly describe here one of the main concepts of our construction and proofs, that is a concept of a bi-wheel amplifier introduced in [KLS13].

The construction extends the notion of a wheel amplifier of [BK99], [BK01] (we refer to [BK99] for the notions of *contact* and *checker* vertices).

A bi-wheel amplifier with $2n$ contact vertices is constructed in the following way. First we construct two disjoint cycles with each $7n$ vertices and we number the vertices by $0, 1, \ldots 7n-1$. The *contacts* will be the vertices with the numbers being a multiple of 7, while the remaining vertices will be *checkers*. We complete the construction by selecting at random a perfect matching from the checkers of one cycle to the checkers of the other cycle (see Figure 1).

We use a bi-wheel amplifier in a similar way to the standard wheel amplifier. The crucial difference is that the cycle edges will correspond to the equality constraints, and matching edges will correspond to inequality constraints. The contacts of one cycle will represent the positive appearance of the original variable, and the contacts of the others the negative ones. The

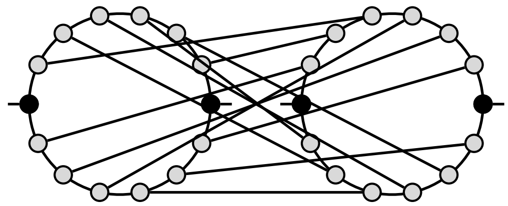

Figure 1: A bi-wheel amplifier with $n = 2$. The vertices ○ denote checkers and the vertices ● denote contacts.

main reason is that encoding inequality constraints will be more efficient then encoding equalities with TSP gadgets.

We can prove the following crucial result on bi-wheels.

Lemma 1 ([KLS13])**.** *With high probability, a bi-wheel is a* 3*-regular amplifier.*

We are going to apply Lemma 1 in the next section.

6 Preparation Lemma

We need the following preparation lemma to simplify our constructions.

We will call in the sequel the problem 3-Occ-MAX-HYBRID-LIN2 simply the HYBRID problem. Lemma 1 will be used now to prove the following lemma.

Lemma 2 ([KLS13])**.** *For any constant* $\varepsilon > 0$ *and* $b \in \{0, 1\}$*, there exists an instance* I *of the HYBRID problem with all variables occurring exactly three times and having* $21m$ *equations of the form* $x \oplus y = 0$*,* $9m$ *equations of the form* $x \oplus y = 1$ *and* m *equations of the form* $x \oplus y \oplus z = b$*, such that it is NP-hard to decide whether there is an assignment to variables which leaves at most* $\varepsilon \cdot m$ *equations unsatisfied, or every assignment to variables leaves at least* $(0.5 - \varepsilon)m$ *equations unsatisfied.*

Because of the above lemma we can assume, among other things, that equations with three variables in I are of the form, say, $x \oplus y \oplus z = 0$. For explicit constructions of simulating gadgets especially the bi-wheel amplifiers and improved gadgets for size-three equations we refer to [KLS13].

7 Instances of Metric TSP

We describe first an underlying idea of the construction of g. We start with instances of the HYBRID problem by constructing a graph with special gadgets structures representing equations. Here we use a new tool of *bi-wheel amplifier* graphs [KLS13], sketched shortly in Section 5.

Given an instance I of the HYBRID problem. Let us denote the corresponding graph by G_S. We analyse first the process of constructing a tour in G_S for a given assignment a to the variables of a HYBRID instance I.

Lemma 3 ([KLS13])**.** *If there is an assignment to the variables of an instance I of the HYBRID problem with $31m$ equations and ν bi-wheels which makes k equations unsatisfied, then there exists a tour in G_S which costs at most $61m + 2\nu + k + 2$.*

We have to prove also corresponding bounds for the opposite direction. Given a tour in G_S, we construct an assignment to the variables of the *associated* instance of the HYBRID problem.

Lemma 4 ([KLS13])**.** *If there exists a tour in G_S with cost $61m+k-2$, then there exists an assignment to the variables of the corresponding instance of the HYBRID problem which makes at most k equations unsatisfied.*

Lemma 3 and Lemma 4 entail now straightforwardly our main result.

Theorem 2 ([KLS13])**.** *The TSP problem is NP-hard to approximate to within any approximation ratio less than* $\mathbf{123/122}$.

On the upper approximation bound of this problem, the best approximation algorithm after more than three decades research is still Christofides [C76] algorithm with approximation ratio 1.50. This leaves a curious huge gap between upper bound 50% and lower bound of about 1% wide open.

8 Instances of Asymmetric TSP (ATSP)

We consider now asymmetric metric instances of TSP. There is no polynomial time constant approximation ratio algorithm known for that problem. The best known approximation algorithm achieves an approximation ratio

$O(\log n/\log\log n)$ [AGM+10]. It motivates a strong interest on inapproximability bounds for this problem. We establish here the best-of-now explicit inapproximability bounds, still constant and very far from the best upper approximation bound.

The plan of our attack is similar to the case of symmetric TSP. We construct here for an instance I of the HYBRID problem a directed, for this case, graph G_A using the constructions of bi-directed edges. The corresponding lemmas describe the opposite directions of the reductions: assignment to tour and tour to assignment.

Lemma 5 ([KLS13])**.** *Given an assignment to the variables of an instance I of the HYBRID problem with ν bi-wheels which makes k equations of I unsatisfied, then there exists a tour in G_A which costs at most $37m + 5\nu + 2m\lambda + 2\nu\lambda + k$ for a fixed constant $\lambda > 0$.*

Lemma 6 ([KLS13])**.** *If there exists a tour in G_A with cost $37m + k + 2\lambda m$, then there exists an assignment to the variables of the corresponding instance of the HYBRID problem which leaves at most k equations unsatisfied.*

Lemmas 5 and 6 entail now our main result of this section.

Theorem 3 ([KLS13])**.** *The ATSP problem is NP-hard to approximate to within any approximation ratio less than* **75/74**.

9 A Role of Weights

A natural question arises about the role of the magnitudes of weights necessary to carry good approximation reductions from the HYBRID problem. The bounded metric situations were studied for the first time in [EK06]. We define $(1, B)$-TSP ($(1, B)$-ATSP) as the TSP (ATSP) problem taking values from the set of integers $\{1, \ldots, B\}$. The case of $(1, 2)$-TSP is important for its connection to the Graphic TSP. Using a specialization of the methods of Section 7 and 8 we obtain the following explicit inapproximability bounds.

Theorem 4 ([KS12])**.** *It is NP-hard to approximate*

(1). the $(1, 2)$-*TSP to within any factor less than* $535/534$;

(2). the $(1, 4)$-*TSP to within any factor less than* $337/336$;

(3). *the* $(1,2)$*-ATSP to within any factor less than* $207/206$*;*

(4). *the* $(1,4)$*-ATSP to within any factor less than* $141/140$*.*

For the most restrictive case of $(1,2)$-TSP, there are better algorithms known [BK06] than the Christofides algorithm.

10 Graphic TSP

We consider now another restricted case of TSP called the Graphic TSP and an interesting for generic reasons case of Graphic TSP on cubic graphs. There was a significant progress recently in designing improved approximation algorithms for the above problems, cf. [MS11], [SV14]. Taking an inspiration from the technique on bounded metric TSP, we prove the following

Theorem 5 ([KS13b])**.** *The Graphic TSP is NP-hard to approximate to within any factor less than* $535/534$*.*

Theorem 6 ([KS13b])**.** *The Graphic TSP on cubic graphs is NP-hard to approximate to within any factor less than* $1153/1152$*.*

The result of Theorem 6 is also the first inapproximability result for the cubic graph instances (cf. [BSS+11]).

11 Some Applications

As a further application of our method, we prove new explicit inapproximability bounds for the Shortest Superstring Problem (SSP) and Maximum Compression Problem (MCP) improving the previous bounds by an order of magnitude [KS13a].

Theorem 7 ([KS13a])**.** *The SSP is NP-hard to approximate to within any factor less than* $333/332$*.*

Theorem 8 ([KS13a])**.** *The MCP is NP-hard to approximate to within any factor less than* $204/203$*.*

The best approximation algorithm for MCP reduces that problem to MAX-ATSP (cf., e.g. [KLS^{+}05]). On the other hand, MCP can be seen as a restricted version of the MAX-ATSP. This entails the NP-hardness of approximating MAX-ATSP to within any factor less than 204/203.

12 Summary of Main Results

We summarize here the best known explicit inapproximability results on the TSP problems (Table 1) and the applications (Table 2).

Table 1: Inapproximability results on the TSP problems

Problem	**Approximation Hardness Bound**
TSP	123/122 [KLS13]
Asymmetric TSP	75/74 [KLS13]
Graphic TSP	535/534 [KS13b]
Graphic TSP on cubic graphs	1153/1152 [KS13b]
$(1, 2)$-TSP	535/534 [KS12]
$(1, 4)$-TSP	337/336 [KS12]
$(1, 2)$-ATSP	207/206 [KS12]
$(1, 4)$-ATSP	141/140 [KS12]

13 Further Research

We presented a method for improving the best known explicit inapproximability bounds for TSP and some related problems. The method depends essentially on a new construction of bounded degree amplifiers. It is still a sensible direction to go for improving design of our new amplifiers and perhaps discovering new proof methods. Also, improving an explicit lower

Table 2: Inapproximability results on the related problems

Problem	**Approximation Hardness Bound**	
SSP	333/332	[KS13a]
MCP	204/203	[KS13a]
MAX-ATSP	204/203	[KS13a]

bound for the HYBRID problem would have immediate consequences toward inapproximability bound of the TSP. Here again the best upper approximation bounds are much higher (but within a couple of percentage points) from the currently provable lower approximation bounds. The TSP problem lacks good definability properties and is definitionally *globally dependent* on all its variables. Our methods used in the proofs were however local in that sense.

In order to get essentially better lower approximation bounds (if such are in fact possible) one should perhaps try to design some more global and perhaps some more weight dependent methods. One possible way would be perhaps to design directly a new *global* PCP construction for the TSP. This seems for the moment to be a very difficult undertaking because of the before mentioned definability properties of the problem.

References

[AT15] *Approximation Taxonomy of Metric TSP*, (2015), URL `http://theory.cs.uni-bonn.de/info5/tsp`

[AGM+10] A. Asadpour, M. Goemans, A. Mądry, S. Gharan and A. Saberi, *An* $O(\log n/ \log\log n)$*-Approximation Algorithm for the Asymmetric Traveling Salesman Problem*, Proc. 21st SODA (2010), pp. 379–389.

[BK99] P. Berman and M. Karpinski, *On Some Tighter Inapproximability Results*, Proc. 26th ICALP (1999), LNCS 1644, pp. 200–209.

[BK01] P. Berman and M. Karpinski, *Efficient Amplifiers and Bounded Degree Optimization*, ECCC TR01-053, 2001.

[BK03] P. Berman and M. Karpinski, *Improved Approximation Lower Bounds on Small Occurrence Optimization*, ECCC TR03-008, 2003.

[BK06] P. Berman and M. Karpinski, *8/7-approximation algorithm for* $(1,2)$*-TSP*, In Proc. 17th SODA (2006), pp. 641–648.

[BSS+11] S. Boyd, R. Sitters, S. van der Ster and L. Stougie, *TSP on Cubic and Subcubic Graphs*, In Proc. 15th IPCO (2011), LNCS 6655, pp. 65–77.

[C76] N. Christofides, *Worst-Case Analysis of a New Heuristic for the Traveling Salesman Problem*, Technical Report CS-93-13, Carnegie Mellon University, Pittsburgh, 1976.

[E03] L. Engebretsen, *An Explicit Lower Bound for TSP with Distances One and Two*, Algorithmica 35, pp. 301–318, 2003.

[EK06] L. Engebretsen and M. Karpinski, *TSP with Bounded Metrics*, J. Comput. Syst. Sci. 72, pp. 509–546, 2006.

[H01] J. Håstad, *Some Optimal Inapproximability Results*, J. ACM 48, pp. 798–859, 2001.

[KLS+05] H. Kaplan, M. Lewenstein, N. Shafrir and M. Sviridenko, *Approximation Algorithms for Asymmetric TSP by Decomposing Directed Regular Multigraphs*, J. ACM 52, pp. 602–626, 2005.

[K01] M. Karpinski, *Approximating Bounded Degree Instances of NP-Hard Problems*, In Proc. 13th FCT (2001), LNCS 2138, pp. 24–34.

[KLS13] M. Karpinski, M. Lampis and R. Schmied, *New Inapproximability Bounds for TSP*, In Proc. 24th ISAAC (2013), LNCS 8283, pp. 568–578; to appear in J. Comput. Syst. Sci.

[KS12] M. Karpinski and R. Schmied, *On Approximation Lower Bounds for TSP with Bounded Metrics*, CoRR abs/1201.5821, 2012.

[KS13a] M. Karpinski and R. Schmied, *Improved Lower Bounds for the Shortest Superstring and Related Problems* In Proc. 19th CATS (2013), CRPIT 141, pp.27-36.

[KS13b] M. Karpinski and R. Schmied, *Approximation Hardness of Graphic TSP on Cubic Graphs* CoRR abs/1304.6800, 2013, journal version RAIRO-Operations Research 49, pp. 651–668, 2015

[L12] M. Lampis, *Improved Inapproximability for TSP*, In Proc. 15th APPROX - RANDOM (2012), LNCS 7408, pp. 243–253; also in Theory of Computing 10, pp. 217–236, 2014.

[MS11] T. Mömke and O. Svensson, *Approximating Graphic TSP by Matchings*, In Proc. IEEE 52nd FOCS (2011), pp. 560–569.

[PV06] C. Papadimitriou and S. Vempala, *On the Approximability of the Traveling Salesman Problem*, Combinatorica 26, pp. 101–120, 2006.

[PY93] C. Papadimitriou and M. Yannakakis, *The traveling Salesman Problem with Distances One and Two*, Math. Oper. Res. 18, pp. 1–11, 1993.

[SV14] A. Sebő and J. Vygen, *Shorter Tours by Nicer Ears: 7/5-approximation for the Graph-TSP, 3/2 for the Path Version, and 4/3 for Two-Edge-Connected Subgraphs*, Combinatorica 34, pp. 1–34, 2014.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de