A priori TSP in the scenario model

Martijn van Ee[a], Leo van Iersel[b], Teun Janssen[b], René Sitters[a,c]

[a] *Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands*
[b] *Delft University of Technology, Mekelweg 4, 2628 CD Delft, Netherlands*
[c] *Centrum voor Wiskunde en Informatica (CWI), Science Park 123, 1098 XG Amsterdam, The Netherlands*

Abstract

In this paper, we consider the *a priori* traveling salesman problem in the scenario model. In this problem, we are given a list of subsets of the vertices, called *scenarios*, along with a probability for each scenario. Given a tour on all vertices, the resulting tour for a given scenario is obtained by restricting the solution to the vertices of the scenario. The goal is to find a tour on all vertices that minimizes the expected length of the resulting restricted tour. We show that this problem is already NP-hard and APX-hard when all scenarios have size four. On the positive side, we show that there exists a constant-factor approximation algorithm in three restricted cases: if the number of scenarios is fixed, if the number of missing vertices per scenario is bounded by a constant, and if the scenarios are nested. Finally, we discuss an elegant relation with an *a priori* minimum spanning tree problem.

Keywords: traveling salesman problem, *a priori* optimization, master tour, optimization under scenarios

1. Introduction

In universal and *a priori* routing, we extend our classical routing problems to the case that the set of clients is uncertain or changes regularly. Because reoptimizing over and over again might be inconvenient or impossible, we want to find a single tour. Given a tour and a set of clients, the active set, we shortcut the tour to the active set. In universal routing, the goal is to minimize the worst-case ratio of the value of the obtained solution and the deterministic optimal value. In *a priori* routing, we want to be good on average. The problem we consider in this paper is formally defined as follows. A preliminary version of this paper was published in [10].

In the *a priori* traveling salesman problem (*a priori* TSP) in the scenario model, we are given a complete graph $G = (V, E)$ with weights that form a metric and a set of scenarios $\mathcal{S}$ with $S_1, \ldots, S_m \subseteq V$. Scenario S_j has probability p_j of being the active set, where $\sum_j p_j = 1$. We begin by finding an ordering on V, called the first-stage tour. When an active set is released, the second-stage tour is obtained by shortcutting the first-stage tour on the vertices of the active set. The goal is to find a first-stage tour that minimizes the expected length of the second-stage tour. Throughout this paper, we assume that the edge weights obey the triangle inequality.

The *a priori* TSP has, for example, a direct application in the photo-lithography processes used in semi-conductor manufacturing to transfer the geometric pattern of a chip onto a wafer [9]. This is done by putting UV-light through a photomask on a photoresistant layer on top of the wafer. The entire wafer is not exposed at once, but one square at a time. If certain parts of the square do not need to be exposed, blades are moved in to block the UV-light. Moving the blades is a time-consuming, and hence costly, process. Since it often influences the total processing time of a wafer in the lithography machine, minimizing the distance reduces the processing time. The blading positions are defined in a file. The blading positions are obtained from this file by reading it from top to bottom and the positions are used by the machine in order of appearance. A product will visit the photolithography machine multiple times during its fabrication. Every time it will use the same file that defines its blading positions, but it will not use all blading positions defined in the file in every visit. For each visit, there is a given subset of the blading positions that has to be used. Hence minimizing the movement of the blades comes down to finding an ordering of the blading positions such that the sum over all visits of the total distance between the blading-positions is minimized. The authors of [9] show that this is precisely a form of the *a priori* TSP in the scenario model.

A priori TSP has already been considered in the independent decision and black-box model. In the independent decision model, vertex i is active with probability p_i, independent of the other vertices. Shmoys and Talwar [29] showed that a sample-and-augment approach gives a randomized 4-approximation, which can be derandomized to an 8-approximation algorithm. This factor was improved by Van Zuylen [30] to 6.5. In the black-box model, we have no knowledge on the probability distribution over the vertices, but we are able to sample from it, i.e., to query the probability of any subset of the vertices. Schalekamp and Shmoys [28] showed that one can obtain a randomized $O(\log n)$-approximation even without sampling. A deterministic $O(\log^2 n)$-approximation can be obtained by using the result for universal TSP [17]. It was shown by [16] that there is an $\Omega(\log n)$ lower bound for deterministic algorithms on general metrics. By using the result of [19] and Theorem 3 in [16], there is no deterministic algorithm with guarantee $o\left(\sqrt[6]{\log n/\log\log n}\right)$ for planar metrics. For randomized algorithms, no lower bound is known for the black-box model.

The above mentioned results give us the first results for *a priori* TSP in the scenario model. First of all, we inherit the randomized $O(\log n)$-approximation. Secondly, we know that a deterministic algorithm that does not use the information given in the scenarios will not achieve an approximation guarantee better than $O(\log n)$. The main question is whether we can use the scenarios to improve upon the $O(\log n)$ upper bound and which restrictions we can put on the scenarios in order to obtain constant-factor approximability. This question will be considered in this paper.

The scenario model has not been studied extensively for other optimization problems. Immorlica et al. [21] investigated scenario versions of Vertex Cover and Shortest Path. Ravi and Sinha [26] also looked at these problems and also defined scenario versions of Bin Packing, Facility Location and Set Cover. The problems in [26] differ from our setting in the sense that the weights used in the instance differ between scenarios. Further, the authors of [6] investigates a two-stage stochastic scheduling problem, where the set of jobs to be processed is uncertain. Finally, in [12], the classical scheduling problem of minimizing the makespan on two machines is considered in the *a priori* model with scenarios. It would be interesting to consider other stochastic combinatorial optimization problems in the *a priori* framework.

A priori TSP can be considered as a stochastic version of TSP. Alternatively, one could consider a robust version where we want to minimize the maximum length over all scenarios. We will refer to this problem as Min-Max TSP. When applicable, we will state to which extend the theorems for *a priori* TSP also hold for the Min-Max TSP. An easy observation is that the approximation ratios for universal TSP carry over directly to MinMax-TSP. Hence, we have an $O(\log^2 n)$-approximation algorithm.

In this paper, we will first examine the most natural lower bound that we call the master tour lower bound. We use this lower bound to show that there exists a constant-factor approximation algorithm for the problem if the number of scenarios is fixed. However, we also show that this lower bound cannot be used to improve upon the $O(\log n)$-approximation when the number of scenarios is unrestricted. We then look at several natural restrictions on the scenarios, namely small, big and nested scenarios. We give strong inapproximability results for small scenarios, a constant-factor approximation for big scenarios (where a constant number of vertices is missing per scenario) and a 9-approximation algorithm for nested scenarios. Finally, we show that there exists an elegant connection to an *a priori* minimum spanning tree problem. We end with a discussion on some open problems.

2. Master tour lower bound

In this section, we explore the master tour lower bound. Here, we use that the contribution of scenario S_j to the objective value of an optimal solution, denoted by OPT, is at least $p_j T_j^*$, where T_j^* is the length of the optimal tour on S_j, so $\text{OPT} \geq \sum_j p_j T_j^*$.

Two natural algorithms for *a priori* TSP in the scenario model are the following. For each scenario, find an α-approximate tour, where α is the best approximation ratio available for TSP, and sort the scenarios on their resulting tour lengths T_j. Rename the scenarios such that $T_1 \leq T_2 \leq \ldots \leq T_m$. Now traverse the tours $1, 2, \ldots, m$, skipping already visited vertices, resulting in tour τ_1. Alternatively, rename the scenarios such that $p_1 \geq p_2 \geq \ldots \geq p_m$ and traverse the tours $1, 2, \ldots, m$, skipping already visited vertices, resulting in tour τ_2. We get the following result.

Theorem 1. *Tours τ_1 and τ_2 are $(2m - \frac{1}{2})$-approximations for* a priori *TSP in the scenario model, where $m \geq 2$ is the number of scenarios.*

Proof. Let us analyze tour τ_1. Consider an arbitrary scenario S_j. Let D_j be the diameter of G restricted to S_j, so we have $T_j^* \geq 2D_j$. Note that when analyzing the contribution of scenario S_j, we only have to consider tours that contain vertices in S_j. Further, it might happen that two scenarios, say S_x and S_y, with $x, y < j$, $S_x \cap S_j \neq \emptyset$ and $S_y \cap S_j \neq \emptyset$, are disjoint and all scenarios S_z with $x < z < y$ have an empty intersection with S_j. In this case, we have to move from a vertex in S_x to a vertex in S_y. If $d(A, B)$ denotes the maximum distance between a vertex in A and a vertex in B, then this move costs us at most an extra $d(S_x \cap S_j, S_y \cap S_j) \leq D_j$. For $j = 1$, the contribution is just $p_1 T_1 \leq \alpha p_1 T_1^*$. For $j \geq 2$, the contribution of S_j to the objective value of our solution is at most

$$
\begin{aligned}
& p_j(T_1 + D_j + T_2 + \ldots + T_{j-1} + D_j + T_j) \\
\leq & p_j(jT_j + (j-1)D_j) \leq p_j\left(\alpha j T_j^* + (j-1)\frac{1}{2}T_j^*\right) = \left(\left(\alpha + \frac{1}{2}\right)j - \frac{1}{2}\right)p_j T_j^*.
\end{aligned}
$$

The objective value is at most

$$
\alpha p_1 T_1^* + \sum_{j=2}^{m}\left(\left(\alpha + \frac{1}{2}\right)j - \frac{1}{2}\right)p_j T_j^* \leq \left(\left(\alpha + \frac{1}{2}\right)m - \frac{1}{2}\right)\textsc{Opt}.
$$

Since the currently best approximation guarantee for TSP is 1.5 [7], we get a $(2m - \frac{1}{2})$-approximation algorithm. The analysis for τ_2 is similar and the proof is omitted here. ☐

Since in the proof of Theorem 1 we bound the length of each tour by $2m - 1$ times the optimal tour for that scenario, it is obvious that τ_1 and τ_2 are also $(2m - 1)$-approximations for Min-Max TSP.

It turns out that the master tour lower bound will not give a constant-factor approximation for *a priori* TSP on general metrics. This can be deduced from Theorem 2 in [16], which roughly states the following. Suppose you are given a d-regular Ramanujan graph G on n vertices with girth $g \geq \frac{2}{3}\log_{d-1} n$ [25] and consider the shortest-path metric induced by this graph. Take a random walk of length $70g$ in G and let S be the vertices visited in this walk. Now, consider a TSP-tour on the vertices of G. Theorem 2 in [16] states that for each of the first $g/2$ steps of the tour restricted to S, the probability that the edge has length $\Omega(\log n)$ is bounded from below by a constant.

Theorem 2. *There is a family of instance of* a priori *TSP in the scenario model such that* $\textsc{Opt} = \Omega(\log n)\sum_j p_j T_j^*$ *and* $\textsc{Opt} = \Omega(\log m)\sum_j p_j T_j^*$.

Proof. We use Theorem 2 from [16] as discussed above. Let G be a d-regular Ramanujan graph on n vertices with girth $g \geq \frac{2}{3}\log_{d-1} n$ and consider the shortest-path metric induced by this graph. The set of scenarios is the set of all vertex sets of walks of length $70g$. The probability p_j of scenario S_j is equal to the probability that S_j is the vertex set of a random walk of length $70g$. For a fixed first-stage tour, Theorem 2 in [16] states that in each of the first $g/2$ steps of the second-stage tour, there is a constant probability that the second-stage tour uses an edge of length $\Omega(\log n)$. This implies that the expected length of the first $g/2$ steps of the tour have expected length $\Omega(\log n)$. Since $T_j^* = O(g)$, the first $g/2$ steps are a constant fraction of all the steps and so the lower bound also holds for the entire tour. Hence, we have an instance such that $\text{OPT} = \Omega(\log n) \sum_j p_j T_j^*$. The number of scenarios is equal to the number of possible walks of length $70g$. This is equal to $n \cdot d^{70g} = O(nd^{\log n}) = O(n^{\log d + 1})$. Since d is a constant, this number is polynomially bounded. Hence, we have $\Theta(\log m) = \Theta(\log n)$, which gives us the second lower bound. □

A natural question one can ask is whether a given instance has an optimal value that is equal to the master tour lower bound. Stated differently, is there a tour such that if we shortcut on the vertices of a scenario, we get the optimal solution for that scenario? Deineko et al. [8] studied this problem for the case where every possible subset is a scenario. They called this the master tour problem and showed that it is polynomially solvable. We can reformulate the problem to the case where we are given a set of scenarios and we only have to be optimal for these scenarios. It turns out that this problem is Δ_2^p-complete [11].

3. Small scenarios

We start with showing that *a priori* TSP is still NP-hard when all scenarios are very small. We reduce from the Max Cut problem [14]. Here, we are given a graph $G = (V, E)$ and our goal is to find a set $X \subseteq V$ such that $|\delta(X)|$ is maximized, where $\delta(X)$ is the set $\{(i, j) \in E : i \in X, j \notin X\}$.

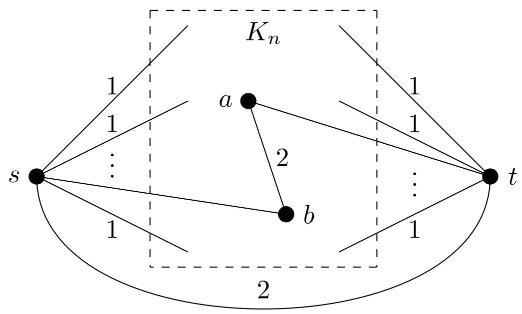

Figure 1: Graph G' as in the proof of Theorem 3.

Theorem 3. A priori *TSP is NP-hard even when* $|S_j| \leq 4$ *for all* j.

Proof. We reduce from the Max Cut problem. Given an instance $G = (V, E)$ of Max Cut, we create an instance of *a priori* TSP by making a complete graph G' on $V \cup \{s, t\}$. All edges with s or t as endpoint, except edge (s, t), have length 1 and all other edges have length 2 (see Figure 1). For every edge $(a, b) \in E$, we create a scenario $\{a, b, s, t\}$. All scenarios have equal probability. Note that the second-stage tour on a scenario either has a length of 4 or length 6. We say that a scenario is satisfied if its resulting tour has length 4. Hence, minimizing the expected length is equivalent to maximizing the number of satisfied scenarios. We will show that $\text{OPT}_{\text{TSP}} = 6|E| - 2\text{OPT}_{\text{CUT}}$, where OPT_{TSP} and OPT_{CUT} are the optimal sum (instead of the average) of tour lengths of *a priori* TSP in the created instance and the optimal value of Max Cut in the original instance respectively.

Suppose there is a cut, say $Q \subseteq V$, such that $\delta(Q)$ contains at least k edges. First, visit s. Then, visit the vertices of Q in arbitrary order. After that, we visit t. Finally, we visit the vertices not in Q in arbitrary order. It is easy to see that every scenario corresponding to an edge in $\delta(Q)$ has length 4, whereas other scenarios have length 6. Hence, there is a tour satisfying at least k scenarios.

On the other hand, suppose that we have a tour in G' satisfying at least k scenarios. Without loss of generality, the tour can be written as (s, R_1, t, R_2), where R_1 and R_2 are sequences of vertices. The only way to satisfy a scenario $\{a, b, s, t\}$ is by putting one vertex of $\{a, b\}$ in R_1 and one vertex in R_2. Hence, the k satisfied scenarios correspond to edges in the cut $\delta(R_1)$ which has size at least k. □

By adjusting the proof of Theorem 3, we can prove that the master tour problem with scenarios is NP-complete when $|S_j| \leq 5$. This is done by reducing from Set Splitting instead of Max Cut and using that 3-Set Splitting is NP-complete [24]. In 3-Set Splitting, we are given n elements and a collection Σ of sets containing three distinct elements. The question is whether we can partition the elements such that each set is splitted, i.e., there is a partition $(X, V \setminus X)$ such that neither $\sigma_1, \sigma_2, \sigma_3 \in X$ nor $\sigma_1, \sigma_2, \sigma_3 \in V \setminus X$ for all $\{\sigma_1, \sigma_2, \sigma_3\} \in \Sigma$. This also shows that Min-Max TSP is NP-hard when $|S_j| \leq 5$ for all j. Moreover, when $|S_j| \leq 5$ for all j, we cannot approximate Min-Max TSP within a factor of $\frac{4}{3}$, unless P=NP. This is because a splitted set will correspond to a scenario with tour length 6, whereas an unsplitted set corresponds to a scenario with tour length 8. The complexity of the master tour problem with scenarios is still open for $|S_j| \leq 4$.

Note that the graph we used in the proof of Theorem 3 is obtained by taking the metric completion of $K_{2,n}$. This graph is planar, bipartite and it has treewidth equal to 2. Deterministic TSP would be polynomially solvable on such a graph with bounded treewidth. Furthermore, there is a PTAS for deterministic TSP in planar graphs [2]. The next theorem shows that this is not the case for *a priori* TSP (since the proof uses the same graph as before, a metric completion of $K_{2,n}$). This theorem relies on the fact that Max Cut, given the unique games conjecture (UGC), cannot be approximated by a factor above the Goemans-Williamson [15] constant, i.e., approximately 0.878567, unless P=NP [23]. Without this conjecture, Håstad [20] showed that it cannot be approximated above a factor $\frac{16}{17}$, unless P=NP.

Theorem 4. *There is no* 1.0117*-approximation for* a priori *TSP with* $|S_j| \leq 4$*, unless P=NP. Assuming UGC, there is no* 1.0242*-approximation, unless P=NP.*

Proof. Consider the reduction from the proof of Theorem 3. As a result, we have $\text{OPT}_{\text{TSP}} = 6|E| - 2\text{OPT}_{\text{CUT}}$. If we have an $(1 + \alpha)$-approximation algorithm, we get a tour with total length at most $(1 + \alpha)(6|E| -$

2OPT_{CUT}). This implies that there are at least η satisfied scenarios, where

$$\begin{aligned} 4\eta + 6(|E| - \eta) &= (1+\alpha)(6|E| - 2\text{OPT}_{\text{CUT}}) \\ -2\eta &= -2(1+\alpha)\text{OPT}_{\text{CUT}} + 6\alpha|E| \\ \eta &= (1+\alpha)\text{OPT}_{\text{CUT}} - 3\alpha|E|. \end{aligned}$$

These correspond to edges in the cut, hence we have

$$\begin{aligned} \text{Size of cut} &\geq (1+\alpha)\text{OPT}_{\text{CUT}} - 3\alpha|E| \\ &\geq (1+\alpha)\text{OPT}_{\text{CUT}} - 6\alpha\text{OPT}_{\text{CUT}} \\ &= (1-5\alpha)\text{OPT}_{\text{CUT}}, \end{aligned}$$

where the second inequality follows from $\text{OPT}_{\text{CUT}} \geq |E|/2$. Hence, assuming P$\neq$NP, there is no $(1+\alpha)$-approximation for $1 - 5\alpha \geq \frac{16}{17}$, i.e., there is no 1.0117-approximation. If we also assume that the unique games conjecture holds, there is no $(1+\alpha)$-approximation for $1 - 5\alpha \geq 0.878567$, i.e., there is no 1.0242-approximation. □

Since graph G' in Figure 1 used in Theorem 4 is the metric completion of $K_{2,n}$, we get the following corollary.

Corollary 1. A priori *TSP in the scenario model on planar bipartite graphs does not admit a PTAS, unless P=NP.*

When $|S_j| \leq 6$, we can slightly strengthen the result of Theorem 4, by reducing from Max E4-Set Splitting, which cannot be approximated with a factor above $\frac{7}{8}$, unless P=NP [20]. This gives an inapproximability of 1.0265 when $|S_j| \leq 6$.

We can strengthen the inapproximability of *a priori* TSP by using strong results on Permutation CSP's [18]. The problem that we need we will call 4-Undirected Cyclic Ordering (4-UCO). To the best of our knowledge, the problem has never been considered. In this problem, we are given a ground set U and a set of 4-tuples Δ^{UCO} using elements from U. Our goal is to construct an ordering on U that maximizes the number of satisfied 4-tuples. We say that 4-tuple (a, b, c, d) is satisfied if one of the following sequences is a subsequence of the total ordering: (a, b, c, d), (b, c, d, a) , (c, d, a, b), (d, a, b, c), (d, c, b, a), (c, b, a, d), (b, a, d, c), (a, d, c, b). In other words, we get a collection of cycles and we want to find a cyclic ordering maximizing the number of cycles that can be embedded in it. For completeness, we first show that deciding whether all 4-tuples can be satisfied is NP-complete by using a reduction from Cyclic Ordering. In this problem, we are given a set of ordered triples Δ^{CO} of ground set U. The question is whether there exists a cyclic ordering on all elements such that each triple is ordered in the right direction. This problem is NP-complete [13].

Theorem 5. *4-Undirected Cyclic Ordering is NP-hard.*

Proof. Given an instance of Cyclic Ordering, we create elements a_1 and a_2 for every element $a \in U$ and three additional elements, x, y and z. For every element $a \in U$ we create 4-tuples (x, y, a_1, a_2), (x, z, a_1, a_2) and (y, z, a_1, a_2). For every triple in Δ^{CO}, we create one 4-tuple by splitting an arbitrary element. For example, we create 4-tuple (a_1, b_1, b_2, c_1) for triple (a, b, c).

If there exists a cyclic ordering, say $(a, b, \ldots, q)$, we can construct the following satisfying solution for 4-UCO: $(x, y, z, a_1, a_2, b_1, b_2, \ldots, q_1, q_2)$.

On the other hand, suppose that we have a satisfying solution for 4-UCO. Without loss of generality, we may assume that (x, y, a_1, a_2) is visited in this direction. We will show that x, y and z are visited consecutively. Suppose this is not the case and x, y and z are placed at different positions on the solution. This splits the solution into three segments. It is easy to see that for a any $u \in U$, we must have u_1 and u_2 in the same segment. Now, suppose that these elements are visited in the segment between x and y. This implies that the tour has to visit (x, u_2, u_1, y) in this order. However, this conflicts with 4-tuple (y, z, u_1, u_2). Similarly, placing u_1 and u_2 between y and z implies visiting (y, u_2, u_1, z) in this order. This conflicts with 4-tuple (x, y, u_1, u_2). Thus, we know that the solution visits x, y and z consecutively. We now fix the positions of u_1 for all $u \in U$ and we move u_2 to the position next to u_1. This does not conflict with any of the scenario's. The resulting arrangement of the u_1 vertices corresponds to an arrangement consistent with Δ^{CO}. $\square$

In [18], it is shown that every Permutation CSP of constant arity is approximation resistant. This means that, under the unique games conjecture, the best we can do is constructing a random ordering. Classical problems like Cyclic Ordering and Betweenness are in this class of problems. One can check that 4-UCO is also in this class. A corollary of the work of Guruswami et al. [18] is that for any $\epsilon > 0$ it is hard to distinguish between instances where at least a $(1 - \epsilon)$ fraction of the 4-tuples can be satisfied from instances where at most a $(\frac{1}{3} + \epsilon)$ fraction of the 4-tuples can be satisfied, assuming the unique games conjecture is true. The natural generalization of 4-UCO is 5-UCO. For this problem, there is no algorithm having a guarantee larger than $\frac{1}{12}$. This gives the following results.

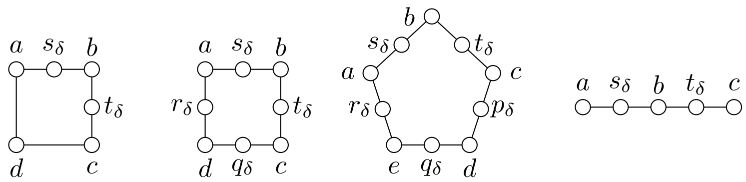

Figure 2: Gadgets used in proofs of Theorem 6 and 7.

Theorem 6. *Under UGC, there is no* α*-approximation for* a priori *TSP with*

(a) $\alpha < \frac{10}{9}$ *when* $|S_j| \leq 6$,

(b) $\alpha < \frac{4}{3}$ *when* $|S_j| \leq 8$*,*

(c) $\alpha < \frac{41}{30}$ *when* $|S_j| \leq 10$*,*

unless P=NP.

Proof. (a) Given an instance of 4-UCO, we create $|U| + 2|\Delta^{\text{UCO}}|$ vertices, one for each element of U and two for each 4-tuple in Δ^{UCO}. We create edges that correspond to 4-tuples in Δ^{UCO} in the following way. For 4-tuple $\delta = (a, b, c, d)$, we have vertices a, b, c, d and vertices s_δ and t_δ. We create edges $(a, s_\delta), (s_\delta, b), (b, t_\delta), (t_\delta, c), (c, d)$ and (d, a), as in Figure 2. The scenarios correspond to these six vertices for every tuple. Finally, the distances correspond to the shortest path distances in the created graph. A tuple is satisfied if and only if the tour restricted to the scenario has length 6. A solution satisfying $\frac{1}{3}$ of the scenarios has value at least $\frac{1}{3} \cdot 6 + \frac{2}{3} \cdot 7 = \frac{20}{3}$. A solution satisfying all scenarios has a value of 6. Since it is hard to distinguish between these two cases, we obtain an inapproximability of $\frac{20}{18} = \frac{10}{9}$ for *a priori* TSP with $|S_i| \leq 6$.

(b) We use a similar reduction. Instead of adding two vertices per tuple, we create four new vertices. In Figure 2, these vertices are called s_i, t_i, q_i and r_i. The scenarios will therefore have size 8. Again, a tuple is satisfied if and only if the tour restricted to the scenario has length 8. However, if we restrict the tour to a scenario corresponding to a non-satisfied tuple, it must have length at least 12. A similar calculation gives an inapproximability of $(\frac{1}{3} \cdot 8 + \frac{2}{3} \cdot 12)/8 = \frac{4}{3}$.

(c) We now reduce from 5-UCO. We add 5 dummy vertices for each scenario and place them between consecutive elements on the cycles. The scenarios will therefore have size 10. Again, a tuple is satisfied if and only if the tour restricted to the scenario has length 10. If we restrict the tour to a scenario corresponding to a non-satisfied tuple, it must have length at least 14. A similar calculation gives an inapproximability of $(\frac{1}{12} \cdot 10 + \frac{11}{12} \cdot 14)/10 = \frac{41}{30}$.

□

Finally, we note that by using twice the diameter of a scenario as a lower bound, we can show that taking an arbitrary tour as a solution is a $c/2$-approximation when $|S_j| \leq c$. A random tour gives a value of at most $(c^2 - 3c + 4)/(2c - 2)$ times the optimal value in expectation. This factor approaches $c/2$ for c large.

3.1. Path-version

One could also consider the path-version of *a priori* TSP. In fact, the application on photolithography is modeled as the path-version. It is easy to see that this problem is trivial when $|S_j| \leq 2$ for all j. If we delete t from the graph created in the reduction of Theorem 3, we can use this graph and the same reduction to show that the path-version of *a priori* TSP is NP-hard when $|S_j| \leq 3$. It is easy to see that this graph can be obtained by taking the metric completion of the star graph. Note that, we can also adjust Theorem 4 to the path-version which will give the same inapproximability result, i.e., there is no 1.0117-approximation, unless P=NP, and there is no 1.0242-approximation if we also assume that the UGC holds.

We can strengthen previous results by using hardness results for Betweenness. In this problem, we are given a set of triples Δ^{B} from elements of U. The triple (a, b, c) is satisfied if (a, b, c) or (c, b, a) is a subsequence of the total ordering. The goal is to find an ordering on U maximizing the number of satisfied triples. By [18], the best approximation ratio is $\frac{1}{3}$, assuming UGC. Without this conjecture, there is no approximation for Betweenness with a factor better than $\frac{1}{2}$, unless P=NP [3].

Theorem 7. *There is no $\frac{9}{8}$-approximation for* a priori *path-TSP with $|S_j| \leq 5$, unless P=NP. Assuming UGC, there is no $\frac{7}{6}$-approximation, unless P=NP.*

Proof. Given an instance of Betweenness, we create a graph with $|U| + 2|\Delta^{\mathrm{B}}|$ vertices. A scenario contains the elements used in a triple and two extra vertices. The edges are drawn in the following way. For triple $\delta = (a, b, c)$, we add edges $(a, s_\delta), (s_\delta, b), (b, t_\delta)$ and (t_δ, c) (Figure 2). A triple is satisfied if and only if the path restricted to the scenario has length 4. Assuming UGC, we get that there is no approximation algorithm with guarantee smaller than $(\frac{1}{3} \cdot 4 + \frac{2}{3} \cdot 5)/4 = \frac{7}{6}$ for *a priori* path-TSP with $|S_j| \leq 5$, unless P=NP. Without assuming UGC, there is no approximation algorithm with guarantee smaller than $(\frac{1}{2} \cdot 4 + \frac{1}{2} \cdot 5)/4 = \frac{9}{8}$, unless P=NP. □

4. Big scenarios

In this section, we investigate the special case of big scenarios, i.e., the case when each scenario has size at least $n - c$, for small c. One would expect that simply taking the optimal tour on the entire vertex set V would perform well on these instances. Here, we analyze this option. Let us denote $\textsc{Opt}(S)$ for the optimal value of a tour on $S \subseteq V$. Further, let $\textsc{Opt}(V)|_S$ denote the value of the optimal tour on V shortcutted to S. As before, let D_S denote the diameter of the graph restricted to S.

Lemma 1. *For $S \subseteq V$ and $c \leq n$ such that $|S| = n - c$, we have*

$$\textsc{Opt}(V)|_S \leq \textsc{Opt}(S) + cD_S.$$

Proof. When shortcutting the optimal tour on V to S we delete paths where only the endpoints are in S. Denote these paths by $\mathcal{P}_i$ for $i = 1, \ldots, c'$, with $c' \leq c$. Let L_i be the length of path $\mathcal{P}_i$ and let $|\mathcal{P}_i|$ be the number of internal vertices on path $\mathcal{P}_i$. We can extend the optimal tour on S to a tour on V by adding these paths. If $|\mathcal{P}_i| \geq 2$, we add $\mathcal{P}_i$ and an extra edge connecting the endpoints. If $|\mathcal{P}_i| = 1$, then we add the cheapest edge from this single internal vertex to a vertex in S twice, which costs us at most L_i. This results in

$$\textsc{Opt}(V) \leq \textsc{Opt}(S) + \sum_{i:|\mathcal{P}_i|\geq 2} (D_S + L_i) + \sum_{i:|\mathcal{P}_i|=1} L_i. \tag{1}$$

On the other hand, we can relate $\textsc{Opt}(V)$ and $\textsc{Opt}(V)|_S$ in the following way. Note that, when short-cutting, we delete each of the $\mathcal{P}_i$ and replace it by an edge between two vertices in S, which costs at most D_S. Hence, we have

$$\textsc{Opt}(V)|_S \leq \textsc{Opt}(V) + \sum_i (D_S - L_i). \tag{2}$$

Suppose there are c_2 paths with $|\mathcal{P}_i| \geq 2$ and c_1 paths with $|\mathcal{P}_i| = 1$. Note that $c \geq c_1 + 2c_2$. Combining Equations (1) and (2) we get

$$\begin{aligned}
\textsc{Opt}(V)|_S &\leq \textsc{Opt}(V) + \sum_{i:|\mathcal{P}_i|\geq 2} (D_S - L_i) + \sum_{i:|\mathcal{P}_i|=1} (D_S - L_i) \\
&\leq \textsc{Opt}(S) + 2 \sum_{i:|\mathcal{P}_i|\geq 2} D_S + \sum_{i:|\mathcal{P}_i|=1} D_S \\
&= \textsc{Opt}(S) + (2c_2 + c_1) D_S \leq \textsc{Opt}(S) + cD_S.
\end{aligned}$$

□

The inequality is tight for the graph in Figure 3 with $c = 2$. We can generalize this tight instance for $c \leq n/2$ by adding more diagonal paths.

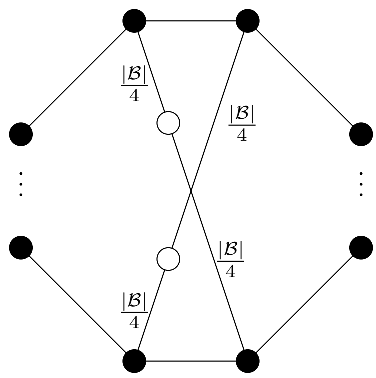

Figure 3: Instance for which inequality of Lemma 1 is asymptotically tight for $c = 2$, where $\mathcal{B}$ is the set of black (non-white) vertices.

Theorem 8. *The optimal solution on V is a $(1 + \frac{c}{2})$-approximation for* a priori *TSP with $|S_i| \geq n - c$, where $c \leq n$.*

Obviously, these results extend to the Min-Max TSP.

5. Nested scenarios

Let us now consider the case of nested scenarios, i.e., $S_1 \subseteq S_2 \subseteq \ldots \subseteq S_m$. Here, the following algorithm gives a constant-factor approximation. First, compute an 1.5-approximate tour T_j for scenario S_j for all j. Let $\alpha_1 = 1$. Next, for $h = 2, 3, \ldots$ let α_h be the largest number $k > \alpha_{h-1}$ for which $T_k \leq 2T_{\alpha_{h-1}}$. If no such k exists then let $\alpha_h = \alpha_{h-1} + 1$. The first-stage tour is obtained by visiting vertices in the order $T_{\alpha_1}, T_{\alpha_2}, \ldots$.

Theorem 9. *The algorithm above is a 9-approximation for nested scenarios.*

Proof. Consider scenario S_j. The last vertices of this scenario will be visited on the tour T_{α_h}, where h is the smallest index such that $\alpha_h \geq j$. Note that for any $h \geq 2$, we have $T_{\alpha_h} > 2T_{\alpha_{h-2}}$. Hence, we can decompose the concatenated tour up to T_{α_h} into two parts which correspond to even and odd h respectively, such that both parts have geometrically increasing tour lengths. The length of the concatenated tour up to T_{α_h} is therefore at most

$$2T_{\alpha_{h-1}} + 2T_{\alpha_h}.$$

If $\alpha_h = j$ then the length of the tour is at most $2T_{\alpha_{h-1}} + 2T_{\alpha_h} \leq 4T_{\alpha_h} = 4T_j \leq 6T_j^*$.

If $\alpha_h > j$, then $j > \alpha_{h-1}$. So, we must have $T_{\alpha_h} \leq 2T_{\alpha_{h-1}}$ and the length of the tour is at most $2T_{\alpha_{h-1}} + 2T_{\alpha_h} \leq 6T_{\alpha_{h-1}} \leq 9T^*_{\alpha_{h-1}} \leq 9T_j^*$. □

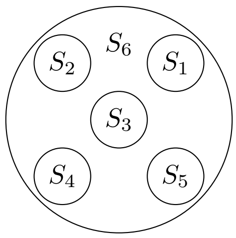

Figure 4: Star-like instance with 6 scenarios.

Finding a constant-factor approximation is still open for laminar scenarios, i.e., when for each i, j, either $S_i \cap S_j = \emptyset$ or $S_i \subseteq S_j$ or $S_j \subseteq S_i$. It is even open in the case when the scenarios have the following star-like structure (illustrated in Figure 4).

$$S_i \cap S_j = \emptyset \text{ for } i \neq j, i, j = 1, \ldots, m-1, \text{ and } S_m = \bigcup_{j=1}^{m-1} S_j. \tag{3}$$

It would be interesting if one could get a constant-factor approximation for these instances. Finally, observe that the Min-Max TSP for laminar scenarios reduces to standard TSP since the largest scenario determines the value of the solutions.

6. Relation with Minimum Spanning Tree problems

It would be nice to have a similar relation between *a priori* TSP and *a priori* MST as in the deterministic setting. We consider two versions of *a priori* MST. The first one is defined by Bertsimas [4], who called it *a priori* MST, although it seems more natural to call it *a priori* Steiner Tree. The second problem is defined by Boria et al. [5], who called it Probabilistic MST under Closest Ancestor (PMST-CA). In both problems, we have a graph $G = (V, E)$ and a probability distribution over subsets of vertices. The second problem is only defined on complete graphs and has a root r that is always active. The root is optional in the first problem. The goal is to construct a tree on the entire vertex set in the first stage. A subset S of the vertices, drawn according to the probability distribution, is revealed in the second stage. In the *a priori* MST, the second-stage tree will be obtained by deleting inactive vertices, provided that the remaining tree stays connected. In the PMST-CA, the second-stage tree only contains active vertices. This is done by taking an edge between an active vertex and its closest active ancestor in the rooted first-stage tree. In both problems, the goal is to construct a first-stage tour that minimizes the expected length of the second-stage tree.

Unfortunately, it turns out that the expected length of the optimal *a priori* MST defined by Bertsimas is not smaller than the optimal *a priori* TSP in general. The gap between the optimal values of *a priori* MST and *a priori* TSP can be arbitrarily large.

Theorem 10. *There are instances such that the optimal value of the* a priori *MST-solution is arbitrarily larger than the optimal value of the* a priori *TSP-solution.*

Proof. Take a 3-regular graph with girth g. Sachs [27] showed that these graphs exist. Define a scenario for each edge by the endpoints of the edge. All scenarios have the same probability. Any tour on this graph will be shortcutted to a tour of length 2 for each scenario, so the objective value of *a priori* TSP is 2. Consider the optimal *a priori* MST. Since this is a tree, it uses $n-1$ edges. If an edge is in the tree, the corresponding scenario gets value 1. If an edge is not in the tree, the corresponding scenario gets value at least $g-1$. Since there are $3n/2$ edges (and scenarios), we get at least the following objective value.

$$\left(\frac{3n/2-(n-1)}{3n/2}\right)(g-1)+\frac{n-1}{3n/2}=\frac{g+1}{3}+\frac{2g-4}{3n}\geq\frac{g+1}{3}.$$

Now, we can take g arbitrarily large, which makes the objective value arbitrarily large and hence the gap with the objective value of *a priori* TSP. □

Unlike the *a priori* MST, the PMST-CA can be used as a lower bound for *a priori* TSP. In fact, we only lose a factor 2. Note that this only works for the rooted case, since PMST-CA is defined with a root vertex.

Theorem 11. *If there is an α-approximation for the PMST-CA, then there is a 2α-approximation for the rooted* a priori *TSP, and vice versa.*

Proof. First, we show that the following inequalities are valid, where OPT_{MST} and OPT_{TSP} denote the optimal values of PMST-CA and *a priori* TSP respectively.

$$\text{OPT}_{\text{MST}} \leq \text{OPT}_{\text{TSP}} \leq 2\text{OPT}_{\text{MST}}.$$

The first inequality can be proven by taking the optimal *a priori* TSP-tour and deleting one edge. This gives a spanning tree on V, called T. If we look at a specific active set S, then the optimal *a priori* TSP-tour restricted to S will have exactly one edge less than before. Namely, if we delete edge (a, b) from tour $(1, \ldots, a, b, \ldots, n)$, only edge $(\max\{k \in S : k \leq a\}, \min\{k \in S : k \geq b\})$ will disappear from the restricted tour on S. Note that for active set S, the tour without this edge is the same as T shortcutted to S. Hence, this is a feasible solution for PMST-CA with cost no larger than the optimal value of *a priori* TSP, and the first inequality has been proven.

The second inequality is proven by doubling the optimal tree and shortcutting the obtained Eulerian tour. In each scenario, the cost of the edges is at most twice the cost of the edges in the tree restricted to the scenario.

Now, if there is an α-approximation for PMST-CA, we double the tree and shortcut the Eulerian tour to obtain a tour on V. This tour has a value of at most

$$2\alpha\text{OPT}_{\text{MST}} \leq 2\alpha\text{OPT}_{\text{TSP}}.$$

Given an α-approximation for *a priori* TSP, we take the tour and delete one edge. The resulting tree has a value of at most

$$\alpha\text{OPT}_{\text{TSP}} \leq 2\alpha\text{OPT}_{\text{MST}}.$$

□

Recall that there is a randomized 4-approximation for *a priori* TSP in the independent decision model [29]. There is also a deterministic 6.5-approximation [30] for this problem. Using Theorem 11, we obtain the following corollary.

Corollary 2. *There is a randomized* 8*-approximation and a deterministic* 13*-approxi-mation for PMST-CA in the independent decision model. There is also a* $O(\log n)$*-approximation in the black-box model.*

Unfortunately, Theorem 11 does not imply a 2-approximation for *a priori* TSP, since we can prove that PMST-CA is NP-hard in the scenario model. For this, we need the following lemma. This lemma holds for both the scenario and the independent decision model.

Lemma 2. *If PMST-CA is NP-hard in the non-metric case, then it is NP-hard in the metric case.*

Proof. One can turn a graph into a graph satisfying the triangle inequality by adding a sufficiently large number M to all distances. In the PMST-CA, this affects every solution by an additive constant equal to $\sum_S p(S)(|S|-1)M$, where $p(S)$ is the probability that set S is the active set. Hence, the complexity of the problem is preserved in the metric case. □

Boria et al. [5] showed that PMST-CA is NP-hard in the independent decision model, but only for the non-metric case. Using Lemma 2, we obtain the following corollary.

Corollary 3. *PMST-CA is NP-hard in the independent decision model, even if the triangle inequality is satisfied.*

Theorem 12. *PMST-CA in the scenario model is NP-hard.*

Proof. We reduce from the NP-complete problem Exact Cover by 3-Sets [22]. In this problem, we are given $3q$ elements, $X = \{x_1, \ldots, x_{3q}\}$, and m subsets, $Y = \{y_1, \ldots, y_m\}$, with $y_i \subseteq X$ and $|y_i| = 3$ for all i. The problem asks whether there are q sets that together cover all elements. Create the graph as in Figure 5. There are m scenarios with probability $1/m$. Define $S_i = X \cup \{r, s, y_i\}$.

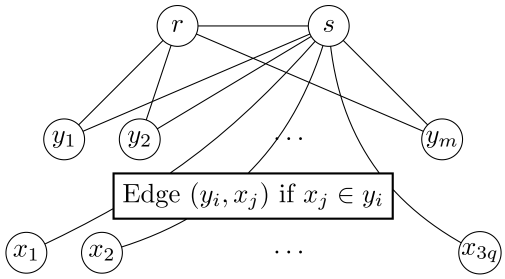

Figure 5: Graph used in proof of Theorem 12. Edges (r, s) and (r, y_i) have length 0. Edges (s, y_i) and (y_i, x_j) have length 1. Edges (s, x_j) have length 2. All other edges have length M, where M is a large number.

If there is an exact cover, then construct the following solution. If set y_i is chosen in the cover, then use edge (s, y_i) and the edges from vertex y_i to the corresponding elements of y_i. If set y_i is not in the cover, then use edge (r, y_i). Finally, use edge (r, s). For any y_i in the cover, consider the subtree containing s, y_i and the x_j's corresponding to elements from subset y_i. In scenario S_i, the resulting subtree has value 4. In all other scenarios, vertex y_i will not be present and this subtree will contain three edges from s to the vertices of the elements. Hence, this solution has expected value equal to $q(1/m \cdot 4 + (m-1)/m \cdot 6) = q(6 - 2/m)$.

Note that an optimal tree will never use edges with weight M or a combination of edges that enforce using an edge of weight M in the shortcut solution. This leaves five ways of connecting a specific set vertex y_i and element vertex x_j, where j is in set i, to r and s. The five subtrees are depicted in Figure 6.

Tree T_3 is dominated by T_1, since T_1 only has cost 2 for connecting x_j when y_i is inactive while T_3 always has cost 2. Similarly, T_4 is dominated by T_2 and T_5 is dominated by T_1. So, an optimal tree is a combination

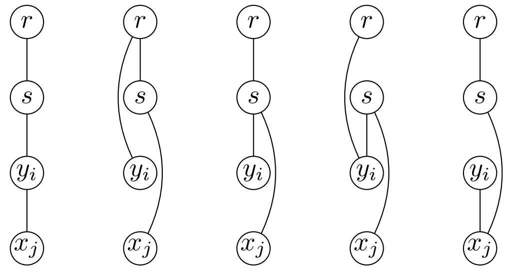

Figure 6: Subtrees T_1 up to T_5.

of T_1 and T_2. Suppose that the tree connects k set vertices to s which connect ℓ elements vertices. The other set vertices are connected to r whereas the other element vertices are connected to s. Number the k set vertices connected to s as $1, \ldots, k$ and say that set vertex i connects ℓ_i element vertices. This tree has an expected value of

$$\frac{1}{m}\sum_{i=1}^{k}((\ell_i + 1) + 2(3q - \ell_i)) + \frac{m-k}{m}6q = 6q + \frac{1}{m}(k - \ell),$$

which is equal to $q(6 - 2/m)$ if and only if $k = q$ and $\ell = 3q$. Hence, there is a tree with expected value at most $q(6 - 2/m)$ if and only if there is an exact cover. Using Lemma 2 completes the proof. $\square$

7. Conclusion

In this paper, we showed how to get constant-factor approximation algorithms for some well-structured instances of the *a priori* TSP. An interesting question that remains unanswered is whether there exists a constant-factor approximation for *a priori* TSP with laminar scenarios. More specifically, it is still open whether we can do this on star-like scenarios as defined in Equation (3). Next to restricted scenarios we also considered restricted metrics. In Section 3 we showed that there is no PTAS for planar bipartite graphs. We do not have such results in the Euclidean plane. It would be interesting to settle the approximability of the problem in this metric. It is easy to construct examples where the optimal solution crosses itself and hence the non-crossing property does not hold. This property was a crucial ingredient of the PTAS by Arora [1] for the deterministic problem. So far, we have not been able to show any lower bound or improve the upper bound for this special case.

We did not succeed in improving the $O(\log n)$-approximation for the general problem. In fact, we conjecture that there is no $o(\log n)$-approximation algorithm for *a priori* TSP in the scenario model in the general case.

Acknowledgments. We would like to thank Karen Aardal, Jan Driessen and Neil Olver for useful discussions. A part of the work by Teun Janssen has been performed in the project INTEGRATE "Integrated Solutions for Agile Manufacturing in High-mix Semiconductor Fabs", co-funded by grants from France, Italy, Ireland, The Netherlands and the ECSEL Joint Undertaking. Martijn van Ee and René Sitters are supported by the NWO Grant 612.001.215. Leo van Iersel was partially supported by the NWO, including Vidi grant 639.072.602, and partially by the 4TU Applied Mathematics Institute.

References

[1] Sanjeev Arora. Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems. *Journal of the ACM*, 45(5):753–782, 1998.

[2] Sanjeev Arora, Michelangelo Grigni, David R. Karger, Philip N. Klein, and Andrzej Woloszyn. A polynomial-time approximation scheme for weighted planar graph TSP. In *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 33–41. ACM/SIAM, 1998.

[3] Per Austrin, Rajsekar Manokaran, and Cenny Wenner. On the NP-hardness of approximating ordering-constraint satisfaction problems. *Theory of Computation*, 11:257–283, 2015.

[4] Dimitris Bertsimas. *Probabilistic combinatorial optimization problems*. PhD thesis, Massachusetts Institute of Technology, 1988.

[5] Nicolas Boria, Cécile Murat, and Vangelis Paschos. On the probabilistic min spanning tree problem. *Journal of Mathematical Modelling and Algorithms*, 11(1):45–76, 2012.

[6] Lin Chen, Nicole Megow, Roman Rischke, and Leen Stougie. Stochastic and robust scheduling in the cloud. In N. Garg, K. Jansen, A. Rao, and J.D.O. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2015*, volume 40 of *LIPIcs*, pages 175–186. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2015.

[7] Nicos Christofides. Worst-case analysis of a new heuristic for the travelling salesman problem. Technical report, DTIC Document, 1976.

[8] Vladimir G. Deineko, Rüdiger Rudolf, and Gerhard J. Woeginger. Sometimes travelling is easy: The master tour problem. *SIAM Journal of Discrete Mathematics*, 11(1):81–93, 1998.

[9] Jan Driessen and Teun Janssen. Minimizing the blading in lithography machines: an application of the *a priori* TSP problem. Unpublished manuscript, 2016.

[10] Martijn van Ee, Leo van Iersel, Teun Janssen, and René Sitters. *A priori* TSP in the scenario model. In K. Jansen and M. Mastrolili, editors, *Approximation and Online Algorithms: 14th International Workshop, Revised Selected Papers*, volume 10138 of *Lecture Notes in Computer Science*, pages 183–196. Springer, 2017.

[11] Martijn van Ee and René Sitters. On the complexity of master problems. In G.F. Italiano, G. Pighizzini, and D.T. Sannella, editors, *Proceedings of the 40th Symposium on Mathematical Foundations of Computer Science*, volume 9235 of *Lecture Notes in Computer Science*, pages 567–576. Springer, 2015.

[12] Esteban Feuerstein, Alberto Marchetti-Spaccamela, Frans Schalekamp, René Sitters, Suzanne van der Ster, Leen Stougie, and Anke van Zuylen. Scheduling over scenarios on two machines. In Z. Cai, A. Zelikovsky, and A.G. Bourgeois, editors, *Proceedings of the 20th International Conference on Computing and Combinatorics*, volume 8597 of *Lecture Notes in Computer Science*, pages 559–571. Springer, 2014.

[13] Zvi Galil and Nimrod Megiddo. Cyclic ordering is NP-complete. *Theoretical Computer Science*, 5(2):179–182, 1977.

[14] Michael R. Garey, David S. Johnson, and Larry J. Stockmeyer. Some simplified NP-complete graph problems. *Theoretical Computer Science*, 1(3):237 – 267, 1976.

[15] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.

[16] Igor Gorodezky, Robert D. Kleinberg, David B. Shmoys, and Gwen Spencer. Improved lower bounds for the universal and *a priori* TSP. In M.J. Serna, R. Shaltiel, K. Jansen, and J.D.P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 13th International Workshop, APPROX 2010, and 14th International Workshop, RANDOM 2010*, volume 6302 of *Lecture Notes in Computer Science*, pages 178–191. Springer, 2010.

[17] Anupam Gupta, Mohammad Taghi Hajiaghayi, and Harald Räcke. Oblivious network design. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 970–979. ACM, 2006.

[18] Venkatesan Guruswami, Johan Håstad, Rajsekar Manokaran, Prasad Raghavendra, and Moses Charikar. Beating the random ordering is hard: Every ordering CSP is approximation resistant. *SIAM Journal on Computing*, 40(3):878–914, 2011.

[19] Mohammad T. Hajiaghayi, Robert Kleinberg, and Tom Leighton. Improved lower and upper bounds for universal TSP in planar metrics. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 649–658. ACM, 2006.

[20] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001.

[21] Nicole Immorlica, David R. Karger, Maria Minkoff, and Vahab S. Mirrokni. On the costs and benefits of procrastination: Approximation algorithms for stochastic combinatorial optimization problems. In *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 691–700. SIAM, 2004.

[22] Richard M. Karp. Reducibility among combinatorial problems. *Complexity of Computer Computations*, pages 85–103, 1972.

[23] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007.

[24] László Lovász. Coverings and colorings of hypergraphs. In *Proceedings of the 4th Southeastern Conference on Combinatorics, Graph Theory and Computing*, pages 3–12, 1973.

[25] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramunajan graphs. *Combinatorica*, 8(3):261 – 277, 1988.

[26] R. Ravi and Amitabh Sinha. Hedging uncertainty: Approximation algorithms for stochastic optimization. *Mathematical Programming*, 108(1):97–114, 2006.

[27] Horst Sachs. Regular graphs with given girth and restricted circuits. *Journal of the London Mathematical Society*, 1(1):423–429, 1963.

[28] Frans Schalekamp and David B. Shmoys. Algorithms for the universal and *a priori* TSP. *Operations Research Letters*, 36(1):1–3, 2008.

[29] David B. Shmoys and Kunal Talwar. A constant approximation algorithm for the *a priori* traveling salesman problem. In A. Lodi, A. Panconesi, and G. Rinaldi, editors, *Proceedings of the 13th International Conference on Integer Programming and Combinatorial Optimization*, volume 5035 of *Lecture Notes in Computer Science*, pages 331–343. Springer, 2008.

[30] Anke van Zuylen. Deterministic sampling algorithms for network design. *Algorithmica*, 60(1):110–151, 2011.