Approximation Ineffectiveness of a Tour-Untangling Heuristic

Bodo Manthey[1] and Jesse van Rhijn[1]

[1]Department of Applied Mathematics, University of Twente

1 Introduction

The Travelling Salesperson Problem (TSP) is a classic example of an NP-hard combinatorial optimization problem [7]. One of its most studied variants is the Euclidean TSP, where the weight of an edge is given by the Euclidean distance between its endpoints. Even this restricted version is NP-hard [8].

Due to this hardness, practitioners often turn to approximation algorithms and heuristics for the TSP. One simple heuristic is 2-opt [1]. In this heuristic, one searches for a pair of edges in the tour that can be replaced by a different pair, such that the total length of the tour decreases. Although this heuristic performs well in practice [1, Chapter 8], it may require an exponential number of iterations to converge [6].

Interestingly, Van Leeuwen & Schoone showed that a restricted variant of 2-opt in which one only removes intersecting edges terminates in $O(n^3)$ iterations in the worst case [9]. For convenience, we refer to this variant as X-opt. More recently, da Fonseca et al. [5] analyzed this heuristic once more, extending the results to matching problems and showing a bound of $O(tn^2)$ for instances where all but t points are in convex position. Their work builds on previous work on computing uncrossing matchings [2].

Removing intersecting edges clearly shortens a tour. However, not all 2-opt iterations remove intersections. Indeed, 2-opt has proved extremely effective also for non-Euclidean TSP instances, where there is no notion of intersecting edges at all. This raises the question: can one use X-opt instead of 2-opt at minimal cost to the approximation guarantee, thereby ensuring an efficient heuristic for TSP instances in the plane? The approximation ratio of 2-opt has long been known to sit between $\Omega\left(\frac{\log n}{\log\log n}\right)$ and $O(\log n)$ for d-dimensional Euclidean instances [4], and has recently been settled to $\Theta\left(\frac{\log n}{\log\log n}\right)$ for 2-dimensional instances [3]. However, no previous work seems to have discussed all intersection-free tours.

We analyze this simpler case here, showing an approximation ratio of $\Omega(n)$ in the worst case and $\Omega(\sqrt{n})$ in the average case. This answers our previously raised question in the negative; in order to obtain a good approximation ratio, one must allow for iterations that improve the tour without removing intersections. Especially the average-case result stands in stark contrast to the average-case approximation ratio of 2-opt, which is $O(1)$ [4].

We augment our results with a numerical experiment, which presents a different picture. To within the precision we are able to achieve, our experiments indicate an expected average-case approximation ratio of X-opt of $O(1)$. We consider this evidence that the techniques we use to obtain the average-case bound of $\Omega(\sqrt{n})$, which are standard techniques used to perform probabilistic analyses of local search heuristics, fall short.

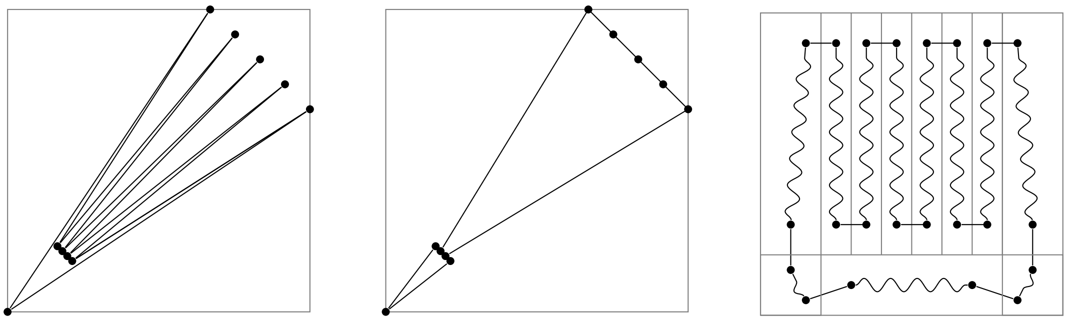

Figure 1: The construction used in Theorem 1. Left: a noncrossing tour of length $\Omega(n)$. Center: a tour of length $O(1)$. Right: The construction used in Theorem 3; the wavy lines represent Hamiltonian paths, while the straight lines represent edges.

2 Worst Case

We construct an instance in which there exists a noncrossing tour with length $\Omega(n)$, as well as a tour of constant length. The construction we use is depicted in Figure 1.

Theorem 1. *Let $n \in \mathbb{N}$ be even. For every $0 < \epsilon < 1$, there exists an instance of the Euclidean TSP in the plane where X-opt has approximation ratio at least $\frac{n}{2} \cdot (1 - \epsilon)$. In particular, the approximation ratio can be brought arbitrarily close to $\frac{n}{2}$.*

The construction used for Theorem 1 only holds for even n. For odd n, we can use a similar construction, but the approximation ratio then becomes $(n-1)/2 \cdot (1-\epsilon)$.

A simple argument shows that the approximation ratio given in Theorem 1 is essentially as bad as one can get in the metric TSP. Given any instance, let x and y be those points separated by the greatest distance. Any tour must travel from x to y and back to x again, so any tour is of length at least $2 \cdot \|x - y\|$. Moreover, every tour contains exactly n edges, so any tour has length at most $n \cdot \|x - y\|$. Hence, the approximation ratio of the metric TSP is at most $n/2$.

3 Average Case

We consider a standard average-case model wherein n points are placed uniformly and independently in the plane. We then construct a tour of length $\Omega(n)$ in expectation. We present our results in Theorem 3.

To simplify our arguments, it would be convenient if we could consider only a subset of all points, so that we can look at a linear-size sub-instance with nicer properties. Constructing a long noncrossing tour through this subset would be much easier.

Since an optimal tour through a set of points X is always at least as long as an optimal tour through a subset $Y \subset X$, it is tempting to conjecture that something similar holds for all noncrossing tours. Perhaps, given a noncrossing tour through $Y \subseteq X$, we can extend the tour to all of X without decreasing its length? Unfortunately, this turns out to be false.

Theorem 2. *There exists a set of points in the plane, together with a noncrossing tour T through all but one of these points, such that all noncrossing tours through all points have length strictly less than the length of T.*

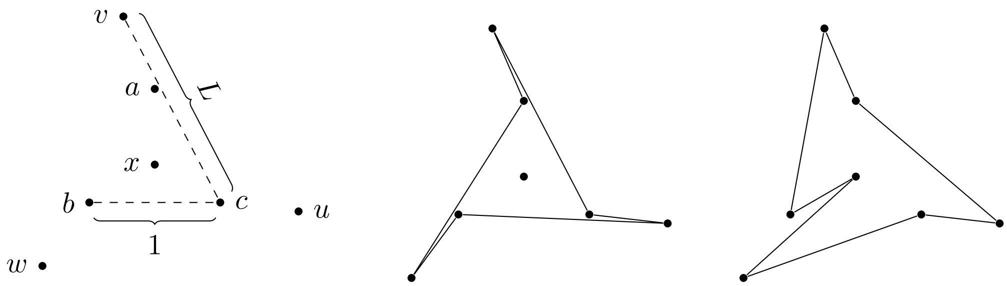

Figure 2: Left: instance described in Theorem 2. Center: the tour T from Theorem 2. Right: the longest possible noncrossing tour through all points in the instance.

Theorem 2 shows that in constructing a tour through a random instance, we must carefully make sure to take all points into account. Any points we leave behind could reduce the length of the tour if we attempt to add them after constructing a long subtour.

Theorem 3. *Suppose a TSP instance is formed by placing n points uniformly at random in the unit square. Then the expected approximation ratio of X-opt on this instance is $\Omega(\sqrt{n})$.*

Proof sketch. As the worst-case tour consists of many long almost-parallel edges, we construct a similar tour in random instances. We divide part of the unit square into many long parallel strips, and form noncrossing Hamiltonian paths within these strips. We connect paths of adjacent strips without creating intersections, forming a long Hamiltonian path through all strips together. The endpoints can then be connected if we leave out some space for points along which to form a connecting path. See Figure 1 for a schematic depiction of our construction. □

4 Practical Performance of Uncrossing Tours

We next compare Theorem 3 to a numerical experiment. We generate instances with n points sampled from the uniform distribution over $[0,1]^2$, and run X-opt on these instances. As a starting tour, we pick a tour from the uniform distribution on all tours. We compute the lengths of the locally optimal tours obtained from our implementation of X-opt, and average them for each fixed value of n we evaluate. We consider the simplest possible pivot rule: starting from an arbitrary edge e in the tour, we check whether e intersects with any other edge, performing an exchange when we find the first such edge. If we do not find such an intersecting edge, we move on to the next edge in the tour and repeat the process. By "next", we mean that we order the edges of a tour according to the permutation on the vertices by which we represent the tour.

Since the optimal tour length is $\Theta(\sqrt{n})$ with high probability, we compare the length of the tours we obtain with this function. Their ratio then serves as a proxy for the approximation ratio of X-opt. We perform this procedure for $n \in \{100, \ldots, 1000, 2000, \ldots, 10000\}$. For each value of n, we take $N = 16,000$ samples. The results are shown in Figure 3.

5 Discussion

Although the results we presented in Theorem 1 and Theorem 3 are rather negative for X-opt, the numerical experiments of Section 4 paint a much more optimistic picture. The heuristic appears to be much more efficient in practice than our lower bounds suggest. Indeed, while Theorem 3

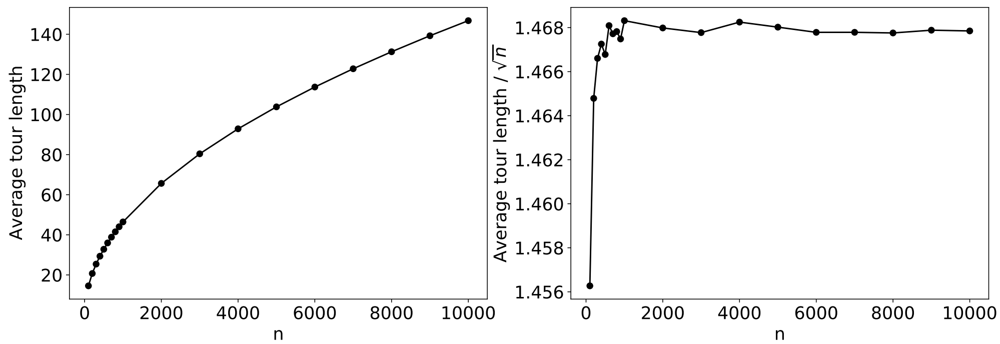

Figure 3: Numerical evaluation of the average-case performance of X-opt.

suggests an approximation ratio of $\Omega(\sqrt{n})$, in Figure 3 we cannot distinguish the approximation ratio from constant to the precision we obtained.

One explanation for this discrepancy is that we compare the optimal solution on any instance to local optima specifically constructed to be bad. This is a standard approach, and it is not surprising that it gives pessimistic results. However, the results in this case are especially pessimistic, considering that we can show a tight lower bound for the expected tour length in the average case. We consider this to be an indication that this approach is incapable of explaining the practical approximation performance of local search heuristics. In order to more closely model the true behavior of heuristics, it seems one must analyze the landscape of local optima, and the probability of reaching different local optima.

References

[1] E. Aarts and J. K. Lenstra, editors. *Local Search in Combinatorial Optimization*. 2003.

[2] A. Biniaz, A. Maheshwari, and M. Smid. Flip Distance to some Plane Configurations, May 2019.

[3] U. A. Brodowsky, S. Hougardy, and X. Zhong. The Approximation Ratio of the k-Opt Heuristic for the Euclidean Traveling Salesman Problem, Aug. 2021.

[4] B. Chandra, H. Karloff, and C. Tovey. New Results on the Old k-opt Algorithm for the Traveling Salesman Problem. *SIAM J. Comput.*, 28(6):1998–2029, Jan. 1999.

[5] G. D. da Fonseca, Y. Gerard, and B. Rivier. On the Longest Flip Sequence to Untangle Segments in the Plane, Dec. 2022.

[6] M. Englert, H. Röglin, and B. Vöcking. Smoothed Analysis of the 2-Opt Algorithm for the General TSP. *TALG*, 13(1):10:1–10:15, Sept. 2016.

[7] B. Korte and J. Vygen. *Combinatorial Optimization: Theory and Algorithms*. 2000.

[8] C. H. Papadimitriou. The Euclidean travelling salesman problem is NP-complete. *Theoretical Computer Science*, 4(3):237–244, June 1977.

[9] J. Van Leeuwen and A. Schoone. Untangling a traveling salesman tour in the plane. *Proc. 7th Internat. Workshop Graph-Theoret. Concepts Comput. Sci.*, Jan. 1980.