On the Smoothed Approximation Ratio of the 2-Opt Heuristic for the TSP

Marvin Künnemann[1] and Bodo Manthey[2]

[1]Max Planck Institute for Informatics, Saarbrücken, Germany, `marvin@mpi-inf.mpg.de`
[2]University of Twente, Enschede, The Netherlands. `b.manthey@utwente.nl`

The 2-Opt heuristic is a simple, easy-to-implement local search heuristic for the traveling salesman problem. While it usually provides good approximations to the optimal tour in experiments, its worst-case performance is poor.

In an attempt to explain the approximation performance of 2-Opt, we prove an upper bound of $\exp(O(\sqrt{\log(1/\sigma)}))$ for the smoothed approximation ratio of 2-Opt. As a lower bound, we prove that the worst-case lower bound of $\Omega(\frac{\log n}{\log \log n})$ for the approximation ratio holds for $\sigma = O(1/\sqrt{n})$.

Our main technical novelty is that, different from existing smoothed analyses, we do not separately analyze objective values of the global and the local optimum on all inputs, but simultaneously bound them on the same input.

1 2-Opt and Smoothed Analysis

The traveling salesman problem (TSP) is one of the best-studied combinatorial optimization problems. Euclidean TSP is the following variant: given points $X \subseteq [0,1]^d$, find the shortest Hamiltonian cycle that visits all points in X (also called a tour).

A very simple and popular heuristic for finding near-optimal TSP tours quickly is the 2-Opt heuristic: starting from an initial tour, we iteratively replace two edges by two other edges to obtain a shorter tour until we have found a local optimum. Experiments indicate that 2-Opt converges quickly and produces solutions that are within a few percent of the optimal solution [3]. In contrast to its success on practical instances, 2-Opt performs poorly in the worst case: the worst-case running-time is exponential even for $d = 2$ [2] and its worst-case approximation ratio of $O(\log n)$ has an almost matching lower bound of $\Omega(\log n / \log \log n)$ [1].

In order to explain the performance of algorithms whose worst-case performance guarantee does not reflect the observed performance, smoothed analysis has been introduced [6], which is a hybrid of worst-case analysis (which is often too pessimistic) and average-case analysis (which is often dominated by completely random instances that have special properties not shared by typical instances). In smoothed analysis, an adversary specifies an instance, and then this instance is slightly randomly perturbed. The smoothed performance is the expected performance, where the expected value is taken over the random perturbation. The motivating assumption of smoothed analysis is that practical instances are often subjected to a small

amount of random noise that can, e.g., come from measurement errors or numerical imprecision. Smoothed analysis often allows more realistic conclusions about the performance of an algorithm than mere worst-case or average-case analysis.

Smoothed analysis has been applied successfully to explain the running-time of the 2-Opt heuristic [2, 5] as well as other algorithms [4]. Much less is known about the smoothed approximation performance of algorithms.

Englert et al. [2] proved an upper bound of $O(\phi^{1/d})$ for the smoothed approximation ratio of 2-Opt in the so-called one-step model [2]. Here, ϕ is an upper bound on the density functions according to which the points are drawn. Translated to Gaussian perturbation, we would obtain an upper bound of $O(1/\sigma)$ if we truncate the Gaussian distribution such that all points lie in a hypercube of constant sidelength.

In order to explain the practical approximation performance of 2-Opt, we provide an improved smoothed analysis of its approximation ratio. More precisely, we provide a bound on the quality of the worst local optimum, when the n data points from $[0, 1]^d$ are perturbed by Gaussian distributions of standard deviation σ (Section 3). Our improved bound is better than the worst case bound of $O(\log n)$ in particular if $1/\sigma \leq (\log n)^c$ for some constant c.

To complement our upper bound, we show that the lower bound by Chandra et al. [1] remains robust for $\sigma = O(1/\sqrt{n})$ (Section 4).

2 Preliminaries

TSP and 2-Opt. We consider input in the unit hypercube $[0, 1]^d$ and assume the dimension d to be a fixed constant. Given a sequence of points $X = (X_1, \ldots, X_n)$ in $\mathbb{R}^d$, we call a collection T of edges connecting these points a *tour* if T is connected and every point has a degree of exactly 2 in T. Given any collection of edges T, its length is denoted by $L(T) = \sum_{(u,v) \in T} d(u, v)$, where $d(u, v)$ denotes the Euclidean distance between points X_u and X_v. We call a tour T *2-optimal*, if $d(u, v) + d(w, z) \leq d(u, w) + d(v, z)$ for all edge pairs $(u, v), (w, z) \in T$, where u, v, w, z appear in this order in T. Equivalently, it is not possible to obtain a shorter tour by replacing (u, v) and (w, z) in a 2-optimal tour T by two new edges. The 2-Opt heuristic replaces a pair of edges (u, v) and (w, z) by (u, w) and (v, z) if this decreases the tour length while this is possible. Thus, it terminates with a 2-optimal tour.

Two functionals are of particular interest: the TSP functional $\mathsf{TSP}(X) = \min_{\text{tour } T} L(T)$ and the functional $\mathsf{2OPT}(X) = \max_{\text{2-optimal tour } T} L(T)$ that maps X to the length of the longest 2-optimal tour through X. In the following, we are interested in the ratio $\mathsf{TSP}(X)/\mathsf{2OPT}(X)$.

Perturbation Models. In the Gaussian perturbation model (also called *two-step model*) for smoothed analysis, an adversary specifies points $x_1, \ldots, x_n$ in $[0, 1]^d$ that serve as unperturbed *origins*. Each such point x_i is perturbed independently by adding a normally distributed variable of mean 0 and standard deviation σ independently to each coordinate. Equivalently, we draw n random noise vectors according to independent multivariante Gaussian distributions with mean 0 and variance σ^2 to obtain the *perturbed input* $X_1 = x_1 + Z_1, \ldots, X_n = x_n + Z_n$. For compactness, we denote the set of unperturbed points by $\overline{X} = \{x_1, \ldots, x_n\}$ and the set of perturbed points by $X = \{X_1, \ldots, X_n\}$.

We assume $\sigma \leq 1$ without loss of generality. If $\sigma > 1$, we can rescale the instance to be contained in $[0, 1/\sigma]^d$ and perturb the points by Gaussians with standard deviation 1 instead,

which gives an equivalent instance. Thus, every upper bound for $\sigma = 1$ carries over to larger values of σ.

The *one-step model* lets the adversary directly specify (not necessarily identical) distributions by choosing probability density functions $f_1, \ldots, f_n : [0,1]^d \to [0, \phi]$. The perturbed input is then generated by independently sampling $X_1 \sim f_1, \ldots, X_n \sim f_n$. Note that the resulting input is always contained in $[0,1]^d$ and with higher ϕ, the adversary can concentrate points to smaller regions of the input space. Roughly speaking, when translating Gaussian perturbations to the one-step model, ϕ is proportional to σ^{-d} for fixed d.

3 Upper Bound on the Approximation Performance

We only give a rough sketch of the idea behind the proof of our upper bound. The crucial difference to existing smoothed analyses of approximation ratios is to take the origins of the points into account. Although this information is not available to the algorithm, it can be exploited in the analysis. The smoothed analyses of approximation ratios so far essentially ignored this information. While this simplifies the analysis, being oblivious to the origins seems too pessimistic. Taking the positions of the unperturbed points carefully into account yields our bound of $\exp(O(\sqrt{\log(1/\sigma)}))$.

To analyze the approximation ratio, we classify edges into *long* and *short* edges and bound the length of long edges by a global argument and the short edges locally against the partial optimal tour on subinstances. The local arguments for short edges will exploit how many unperturbed origins lie in the vicinity of a given region. Since the approximation ratio of 2-Opt is bounded by $O(\log n)$ in the worst-case, we may assume that $1/\sigma = O(n^\varepsilon)$ for all $\varepsilon > 0$, since otherwise our smoothed result is superseded by the $O(\log n)$ worst-case upper bound.

Theorem 1. *Let $\overline{X}$ be n points in $[0,1]^d$, and let X be obtained by perturbation of $\overline{X}$ with standard deviation $\sigma \leq 1$. With probability $1 - \exp(-\Omega(n^{1/2-\varepsilon}))$ for any constant $\varepsilon > 0$, we have $2\mathsf{OPT}(X) \leq \exp(O(\sqrt{\log(1/\sigma)})) \cdot \mathsf{TSP}(X)$. Furthermore,*

$$\mathbb{E}\left[\frac{2\mathsf{OPT}(X)}{\mathsf{TSP}(X)}\right] = \exp\big(O\big(\sqrt{\log(1/\sigma)}\big)\big).$$

4 Lower Bound on the Approximation Ratio

A central claim behind smoothed analysis is that worst-case analysis might be too pessimistic when the worst case is attained at pathological, specifically constructed instances. Since 2-Opt rarely performs as bad on real-word Euclidean instances as its worst-case analysis might suggest, it is interesting to investigate whether perturbation indeed destroys such bad examples. For this, we face the technical difficulty that in general, a single outlier might destroy the 2-optimality of a desired long tour, potentially cascading into a series of 2-Opt iterations that result in a substantially different or even optimal tour. We can show that an adaptation of the worst-case lower bound example by Chandra et al. [1] remains robust for $\sigma = O(1/\sqrt{n})$.

Theorem 2. *Let $\sigma = O(1/\sqrt{n})$. For infinitely many n, there is an instance $\overline{X}$ of points in $[0,1]^2$ such that, with a probability of $1 - O(n^{-s})$ for any constant $s > 0$, we have $2\mathsf{OPT}(X) = \Omega(\log n/\log\log n) \cdot \mathsf{TSP}(X)$. This also yields*

$$\mathbb{E}\left[\frac{2\mathsf{OPT}(X)}{\mathsf{TSP}(X)}\right] = \Omega\left(\frac{\log n}{\log\log n}\right).$$

Often, the one-step model is recognized as more general than Gaussian perturbations, since for sufficiently large ϕ (depending on σ) this model can closely resemble the Gaussian distribution with standard deviation σ, except for being bounded to the unit cube. Indeed, our result naturally transfers to the one-step model and interestingly, holds with a probability of 1 over the random perturbations.

Theorem 3. *Let* $\phi = \Omega(n)$*. For infinitely many* n*, there are probability density functions* $f_1, \ldots, f_n : [0,1]^2 \to [0, \phi]$ *such that* $\mathsf{2OPT}(X) = \Omega(\log n / \log \log n) \cdot \mathsf{TSP}(X)$ *for all instances* $X = (X_1, \ldots, X_n)$*, where each* X_i *is drawn according to* f_i*.*

5 Discussions and Open Problems

While our bound of $\exp(O(\sqrt{\log(1/\sigma)}))$ significantly improves the previously known bound for the smoothed approximation ratio of 2-Opt, we readily admit that it still does not explain the performance observed in practice. This seems to be mainly because of the following two reasons.

First, the initial tour is usually not picked by an adversary but comes from some construction heuristic. such as an insertion heuristic. In practice, tours are initialized using, e.g., the nearest neighbor heuristic (which constitutes a $O(\log n)$-approximation in itself) or insertion heuristics (which for some variants yield 2-approximations). We mention in passing that even when we initialize the tour using the nearest neighbor heuristic, 2-Opt might, with constant probability, return a 2-optimal tour of length $\Omega(\log n / \log \log n) \cdot \mathsf{TSP}(X)$ on perturbed inputs.

Second, we have analyzed the worst local optimum, which might not be the local optimum that we find, even without choosing the initial tour carefully.

We leave as an open problem to analyze the smoothed approximation ratio of 2-Opt with such initializations.

References

[1] Barun Chandra, Howard Karloff, and Craig Tovey. New results on the old k-opt algorithm for the traveling salesman problem. *SIAM J. Comput.*, 28(6):1998–2029, 1999.

[2] Matthias Englert, Heiko Röglin, and Berthold Vöcking. Worst case and probabilistic analysis of the 2-Opt algorithm for the TSP. *Algorithmica*, 68(1):190–264, 2014.

[3] David S. Johnson and Lyle A. McGeoch. Experimental analysis of heuristics for the STSP. In Gregory Gutin and Abraham P. Punnen, editors, *The Traveling Salesman Problem and its Variations*, chapter 9. Kluwer Academic Publishers, 2002.

[4] Bodo Manthey and Heiko Röglin. Smoothed analysis: Analysis of algorithms beyond worst case. *it – Information Technology*, 53(6):280–286, 2011.

[5] Bodo Manthey and Rianne Veenstra. Smoothed Analysis of the 2-Opt Heuristic for the TSP: Polynomial Bounds for Gaussian Noise. In *Proc. 24th Int. Symp. on Algorithms and Computation (ISAAC)*, vol. 8283 of *Lecture Notes in Comput. Sci.*, pp. 579–589. Springer, 2013.

[6] Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *J. ACM*, 51(3):385–463, 2004.