Smoothed Analysis of the 2-Opt Heuristic for the TSP: Polynomial Bounds for Gaussian Noise

Bodo Manthey and Rianne Veenstra

University of Twente, Department of Applied Mathematics, Enschede, The Netherlands
`b.manthey@utwente.nl`, `h.j.veenstra@student.utwente.nl`

Abstract. The 2-opt heuristic is a very simple local search heuristic for the traveling salesman problem. While it usually converges quickly in practice, its running-time can be exponential in the worst case.

In order to explain the performance of 2-opt, Englert, Röglin, and Vöcking (*Algorithmica*, to appear) provided a smoothed analysis in the so-called one-step model on d-dimensional Euclidean instances. However, translating their results to the classical model of smoothed analysis, where points are perturbed by Gaussian distributions with standard deviation σ, yields a bound that is only polynomial in n and $1/\sigma^d$.

We prove bounds that are polynomial in n and $1/\sigma$ for the smoothed running-time with Gaussian perturbations. In particular our analysis for Euclidean distances is much simpler than the existing smoothed analysis.

1 2-Opt and Smoothed Analysis

The traveling salesman problem (TSP) is one of the classical 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*). Even this restricted variant is NP-hard for $d \geq 2$ [16]. We consider Euclidean TSP with Manhattan and Euclidean distances as well as squared Euclidean distances to measure the distances between points. For the former two, there exist polynomial-time approximation schemes (PTAS) [1, 14]. The latter, which has applications in power assignment problems for wireless networks [8], admits a PTAS for $d = 2$ and is APX-hard for $d \geq 3$ [15].

As it is unlikely that there are efficient algorithms for solving Euclidean TSP optimally, heuristics have been developed in order to find near-optimal solutions quickly. One very simple and popular heuristic is 2-opt: 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 to near-optimal solutions quite quickly [9, 10], but its worst-case performance is bad: the worst-case running-time is exponential even for $d = 2$ [7] and the approximation ratio can be $\Omega(\log n/ \log \log n)$ for Euclidean instances [5].

An alternative to worst-case analysis is average-case analysis, where the expected performance with respect to some probability distribution is measured. The average-case running-time for Euclidean instances and the average-case approximation ratio for non-metric instances of 2-opt were analyzed [4–6, 11]. However, while worst-case analysis is often too pessimistic because it is dominated by

L. Cai, S.-W. Cheng, and T.-W. Lam (Eds.): ISAAC2013, LNCS 8283, pp. 579–589, 2013.
© Springer-Verlag Berlin Heidelberg 2013

artificial instances that are rarely encountered in practice, average-case analysis is dominated by random instances, which have often very special properties with high probability that they do not share with typical instances.

In order to overcome the drawbacks of both worst-case and average-case analysis and to explain the performance of the simplex method, Spielman and Teng invented smoothed analysis [17]: 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 underlying assumption is that real-world instances are often subjected to a small amount of random noise, which can, e.g., come from measurement or rounding errors. Smoothed analysis often allows more realistic conclusions about the performance than worst-case or average-case analysis. Since its invention, it has been applied successfully to explain the performance of a variety of algorithms [12, 18].

Englert, Röglin, and Vöcking [7] provided a smoothed analysis of 2-opt in order to explain its performance. They used the *one-step model*: an adversary specifies n density functions $f_1, \ldots, f_n : [0,1]^d \to [0, \phi]$. Then the n points $x_1, \ldots, x_n$ are drawn independently according to the densities $f_1, \ldots, f_n$, respectively. Here, ϕ is the perturbation parameter. If $\phi = 1$, then the only possibility is the uniform distribution on $[0,1]^d$, and we obtain an average-case analysis. The larger ϕ, the more powerful the adversary. Englert et al. [7] proved that the expected running-time of 2-opt is $O(n^4 \phi \log n)$ and $O(n^{4+\frac{1}{3}} \phi^{\frac{8}{3}} \log^2(n\phi))$ for Manhattan and Euclidean distances, respectively. These bounds can be improved slightly by choosing the initial tour with an insertion heuristic. However, if we transfer these bounds to the classical model of points perturbed by Gaussian distributions of standard deviation σ, we obtain bounds that are polynomial in n and $1/\sigma^d$. This is because the maximum density of a d-dimensional Gaussian with standard deviation σ is $\Theta(\sigma^{-d})$. While this is polynomial for any fixed d, it is unsatisfactory that the degree of the polynomial depends on d.

Our Contribution. We provide a smoothed analysis of the running-time of 2-opt in the classical model, where points in $[0,1]^d$ are perturbed by independent Gaussian distributions of standard deviation σ. The bounds that we prove for Gaussian perturbations are polynomial in n and $1/\sigma$, and the degree of the polynomial is independent of d. As distance measures, we consider Manhattan (Section 3), Euclidean (Section 5), and squared Euclidean distances (Section 4).

The analysis for Manhattan distances is a straightforward adaptation of the existing analysis by Englert et al. However, while the degree of the polynomial in n is independent of d in our bound, we still have a factor in the bound that is exponential in d.

Our analysis for Euclidean distances is considerably simpler than the one by Englert et al., which is rather technical and takes more than 20 pages [7].

The analysis for squared Euclidean distances is, to our knowledge, not preceded by a smoothed analysis in the one-step model. Because of the nice properties of squared Euclidean distances and Gaussian perturbations, this smoothed analysis is relatively compact and elegant: the only concept needed for

Theorem 4.3 is pairs of linked 2-changes (Section 2.1), and we can even get rid of this at the price of a slightly worse bound (Remark 4.4). This might be of independent interest, as smoothed analysis of local search heuristics is often rather technical [2,3,7,13].

We did not try to optimize our bounds, but rather tried to keep the analysis simple. We believe that much stronger bounds hold for Euclidean and squared Euclidean distances (see also Section 6).

2 Notation, Preliminaries and Outline

Throughout the rest of this paper, X denotes a set of n points in $\mathbb{R}^d$, where each point is drawn according to an independent d-dimensional Gaussian distribution with mean in $[0,1]^d$ and standard deviation σ. The dimension d is considered to be constant. We discuss the dependence of our bounds on d in Section 6.

We assume that $\sigma \leq 1$. This is justified by two reasons. First, small σ are actually the interesting case, i.e., when the order of magnitude of the perturbation is relatively small. Second, the smoothed number of iterations that 2-opt needs is a monotonically decreasing function of σ: if we have $\sigma > 1$, then this is equivalent to adversarial instances in $[0, 1/\sigma]^d$ that are perturbed with standard deviation 1. This in turn is dominated by adversarial instances in $[0,1]^d$ that are perturbed with standard deviation 1, as $[0,1/\sigma]^d \subseteq [0,1]^d$. Thus, any bound for $\sigma = 1$ holds also for larger σ. Sometimes we even assume $\sigma = O(1/\sqrt{n \log n})$ to simplify the analysis.

2.1 2-Opt State Graph and Linked 2-Changes

Given a tour H that visits all points in X, a *2-change* replaces two edges $\{x_1, x_2\}$ and $\{x_3, x_4\}$ of H by two new edges $\{x_1, x_3\}$ and $\{x_2, x_4\}$, provided that this yields again a tour (this is the case if x_1, x_2, x_3, x_4 appear in this order in the tour) and that this decreases the length of the tour, i.e., $d(x_1, x_2) + d(x_3, x_4) - d(x_1, x_3) - d(x_2, x_4) > 0$, where $d(a, b) = \|a - b\|_2$ (Euclidean distances), $d(a, b) = \|a - b\|_1$ (Manhattan distances), or $d(a, b) = \|a - b\|_2^2$ (squared Euclidean distances). The 2-opt heuristic iteratively improves an initial tour by applying 2-changes until it reaches a local optimum.

The number of iterations that 2-opt needs depends of course heavily on the initial tour and on which 2-change is chosen in each iteration. We do not make any assumptions about the initial tour and about which 2-change is chosen. Following Englert et al. [7], we consider the *2-opt state graph*: we have a node for every tour and a directed edge from tour H to tour H' if H' can be obtained by one 2-change. The 2-opt state graph is a directed acyclic graph, and the length of the longest path in the 2-opt state graph is an upper bound for the number of iterations that 2-opt needs.

In order to improve the bounds (for Manhattan distances) or to allow bounds on the expected number of iterations in the first place (for Euclidean and squared Euclidean distances), we also consider *pairs of linked 2-changes* [7]. Two 2-changes form a pair of linked 2-changes if there is one edge added in one 2-change and removed in the other 2-change. Formally, one 2-change replaces $\{x_1, x_2\}$ and $\{x_3, x_4\}$ by $\{x_1, x_3\}$ and $\{x_2, x_4\}$ and the other 2-change replaces $\{x_1, x_3\}$ and $\{x_5, x_6\}$ by $\{x_1, x_5\}$ and $\{x_2, x_6\}$. It can be that $\{x_2, x_4\}$ and $\{x_5, x_6\}$ intersect. Englert et al. [7] called a pair of linked 2-changes a *type i pair* if $|\{x_2, x_4\} \cap \{x_5, x_6\}| = i$. As type 2 pairs, which involve in fact only four nodes, are difficult to analyze because of dependencies, we ignore them. Fortunately, the following lemma states that we will find enough disjoint pairs of linked 2-changes of type 0 and 1 in any sufficiently long sequence of 2-changes.

Lemma 2.1 (Englert et al. [7]). *Every sequence of t consecutive 2-changes contains at least $t/6 - 7n(n-1)/24$ disjoint pairs of linked 2-changes of type 0 or type 1.*

2.2 Technical Lemmas

In order to get an upper bound for the length of the initial tour, we need an upper bound for the diameter of the point set X. Such an upper bound is also necessary for the analysis of 2-changes with Euclidean distances (Section 5). We choose $D_{\max}$ such that $X \subseteq [-D_{\max}, D_{\max}]^d$ with a probability of at least $1 - 1/n!$. For fixed d and $\sigma \leq 1$, we can choose $D_{\max} = \Theta(1 + \sigma\sqrt{n \log n})$ according to the following lemma. For $\sigma = O(1/\sqrt{n \log n})$, we have $D_{\max} = \Theta(1)$.

Lemma 2.2. *Let $c > 0$ be a sufficiently large constant, and let $D_{\max} = c \cdot (\sigma\sqrt{n \log n} + 1)$. Then $\mathbb{P}(X \not\subseteq [-D_{\max}, D_{\max}]^d) \leq 1/n!$.*

We need the following simple fact a few times.

Lemma 2.3 (Arthur et al. [2, Fact 2.1]). *Let $p \in [0, 1]$ be a probability, and let a, c, b, d, and e be non-negative real numbers with $c \geq 1$ and $e \geq d$. If $p \leq a + c \cdot b^e$, then $p \leq a + c \cdot b^d$.*

For $x, y \in \mathbb{R}^d$ with $x \neq y$, let $L(x, y) = \{\xi \cdot (y - x) + x \mid \xi \in \mathbb{R}\}$ denote the straight line through x and y.

Lemma 2.4. *Let $a, b \in \mathbb{R}^d$ be arbitrary with $a \neq b$. Let $c \in \mathbb{R}^d$ be drawn according to a d-dimensional Gaussian distribution with standard deviation σ. Then the probability that c is ε-close to $L(a, b)$, i.e., $\min_{c^\star \in L(a,b)} \|c - c^\star\|_2 \leq \varepsilon$, is bounded from above by $(\varepsilon/\sigma)^{d-1}$.*

Let $\delta_{\text{close}} = \min_{a,b \in X, a \neq b} \|a - b\|_2$ be the minimum distance of points in X.

Lemma 2.5. *For any $\varepsilon > 0$, we have $\mathbb{P}(\delta_{\text{close}} \leq \varepsilon) \leq n^2 \cdot (\varepsilon/\sigma)^d$.*

We need the following lemma in Section 5.

Lemma 2.6. *Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function whose derivative is bounded from above by B, let c be distributed according to Gaussian distribution with standard deviation σ. Let I be an interval of size ε, and let $f(I) = \{f(x) \mid x \in I\}$ be the image of I. Then $\mathbb{P}(c \in f(I)) = O(B\varepsilon/\sigma)$.*

2.3 Outline

The main idea in the proofs by Englert et al. [7] and in our proofs is to bound the minimal improvement of any 2-change or, for pairs of linked 2-changes, the minimal improvement of any pair of linked 2-changes. We denote the smallest improvement of any linked 2-change by $\Delta_{\min}$ and the smallest improvement of any pair of linked 2-changes by $\Delta_{\min}^{\text{link}}$. It will be clear from the context which distance measure is used for $\Delta_{\min}$ and $\Delta_{\min}^{\text{link}}$. Suppose that the initial tour has a length of at most L, then 2-opt cannot run for more than $L/\Delta_{\min}$ iterations and not for more than $\Theta(L/\Delta_{\min}^{\text{link}})$ iterations. The following lemma formalizes this.

Lemma 2.7. *Suppose that, with a probability of at least $1 - 1/n!$, any tour has a length of at most L. Let $\gamma > 1$. Then*

1. *If $\mathbb{P}(\Delta_{\min} \leq \varepsilon) = O(P\varepsilon)$, then the expected length of the longest path in the 2-opt state graph is bounded from above by $O(PLn \log n)$.*
2. *If $\mathbb{P}(\Delta_{\min} \leq \varepsilon) = O(P\varepsilon^{\gamma})$, then the expected length of the longest path in the 2-opt state graph is bounded from above by $O(P^{1/\gamma}L)$.*
3. *The same bounds hold if we replace $\Delta_{\min}$ by $\Delta_{\min}^{\text{link}}$, provided that $PL = \Omega(n^2)$ for Case 1 and $P^{1/\gamma}L = \Omega(n^2)$ for Case 2.*

For Euclidean and squared Euclidean distances, it turns out to be useful to study $\Delta_{a,b}(c) = d(c,a) - d(c,b)$ for points $a, b, c \in X$. By abusing notation, we sometimes write $\Delta_{i,j}(k)$ instead of $\Delta_{x_i,x_j}(x_k)$ for short. A 2-change that replaces $\{x_1, x_2\}$ and $\{x_3, x_4\}$ by $\{x_1, x_3\}$ and $\{x_2, x_4\}$ improves the tour length by $\Delta_{1,4}(2) - \Delta_{1,4}(3) = \Delta_{2,3}(1) - \Delta_{2,3}(4)$.

3 Manhattan Distances

The analysis for Manhattan distances is a straightforward adaptation of the analysis in the one-step model. We obtain a bound of $O(n^4 D_{\max}/\sigma)$. The term $D_{\max}$ in the bound comes from the bound of the initial tour.

Lemma 3.1. $\mathbb{P}(\Delta_{\min}^{\text{link}} \leq \varepsilon) = O(n^6\varepsilon^2/\sigma^2)$.

Theorem 3.2. *The expected length of the longest path in the 2-opt state graph corresponding to d-dimensional instances with Manhattan distances is at most $O(n^4 D_{\max}/\sigma)$.*

4 Squared Euclidean Distances

In this section, we have $\Delta_{a,b}(c) = \|c - a\|_2^2 - \|c - b\|_2^2$ for $a, b, c \in \mathbb{R}^d$.

Lemma 4.1. *Let $a, b \in \mathbb{R}^d$, $a \neq b$, and let c be drawn according to a Gaussian distribution with standard deviation σ. Let $I \subseteq \mathbb{R}$ be an interval of length ε. Then $\mathbb{P}\big(\Delta_{a,b}(c) \in I\big) \leq \frac{\varepsilon}{4\sigma \cdot \|a-b\|_2}$.*

Proof. Since Gaussian distributions are rotationally symmetric, we can assume without loss of generality that $a = (0, \dots, 0)$ and $b = (\delta, 0, \dots, 0)$. We have $\delta = \|a - b\|_2$. Let $c = (c_1, \dots, c_d)$. Then $\Delta_{a,b}(c) = c_1^2 - (c_1 - \delta)^2 = 2c_1\delta + \delta^2$. Thus, $\Delta_{a,b}(c)$ falls into I if and only if c_1 falls into an interval of length $\frac{\varepsilon}{2\delta}$. Since c_1 is a 1-dimensional Gaussian random variable with a standard deviation of σ, the probability for this is bounded from above by $\frac{\varepsilon}{4\delta\sigma}$. □

We analyze $\Delta_{\min}^{\text{link}}$ since it seems to be difficult to obtain bounds for the expected value using $\Delta_{\min}$.

Lemma 4.2. *For $d \geq 2$, we have $\mathbb{P}(\Delta_{\min}^{\text{link}} \leq \varepsilon) = O(n^6 \varepsilon \sigma^{-2})$.*

Proof. Consider a pair of linked 2-changes where $\{x_1, x_2\}$ and $\{x_3, x_4\}$ are replaced by $\{x_1, x_3\}$ and $\{x_2, x_4\}$ and then $\{x_1, x_3\}$ and $\{x_5, x_6\}$ by $\{x_1, x_5\}$ and $\{x_3, x_6\}$. We assume that $x_2 \notin \{x_5, x_6\}$ and $x_5 \notin \{x_2, x_4\}$. The other cases are identical.

If the pair yields an improvement of at most ε then $\Delta_{1,3}(2)$ falls into some interval of length at most ε and $\Delta_{1,6}(5)$ falls into some interval of length at most ε. We have $\|x_1 - x_3\|_2 \leq \sqrt{\varepsilon}$ or $\|x_1 - x_6\|_2 \leq \sqrt{\varepsilon}$ only if $\delta_{\text{close}} \leq \sqrt{\varepsilon}$, which happens with a probability of at most $n^2(\sqrt{\varepsilon}/\sigma)^d \leq n^2 \varepsilon \sigma^{-2}$ by Lemmas 2.5 and 2.3 since $d \geq 2$. From now on, we assume that $\|x_1 - x_3\|_2, \|x_1 - x_6\|_2 \geq \sqrt{\varepsilon}$. By independence of x_2 and x_5 and Lemma 4.1, the probability that both $\Delta_{1,3}(2)$ and $\Delta_{1,6}(5)$ fall into their "bad" interval of length ε is thus bounded from above by $\left(\frac{\sqrt{\varepsilon}}{4\sigma}\right)^2 = O(\varepsilon\sigma^{-2})$.

The lemma follows now by a union bound over all $O(n^6)$ pairs of linked 2-changes and the fact that we do not have to apply the union bound to the probability that δ_{close} is small. □

Theorem 4.3. *For $d \geq 2$, the expected length of the longest path in the 2-opt state graph corresponding to d-dimensional instances with squared Euclidean distances is at most $O(n^8 \log n \cdot D_{\max}^2/\sigma^2)$.*

Proof. The theorem follows by using Lemma 2.7 with Lemma 4.2 and the observation that the initial tour has a length of at most $O(D_{\max}^2 n)$ with a probability of at least $1 - 1/n!$ by Lemma 2.2. □

Remark 4.4. The proof of Theorem 4.3 can be simplified by getting rid of the pairs of linked 2-changes (Lemma 2.1) and slightly worsening the bound to $O(n^{10} \log n \cdot D_{\max}^2/\sigma^2)$: we observe that two consecutive 2-changes involve between five and eight nodes as they cannot involve the same four points. Thus, there is always one node that is only involved in the first of the two and one node that is only involved in the second of the two. This is sufficient but worsens the bound of Lemma 4.2 as we have to take a union bound over $O(n^8)$ choices for the two 2-changes instead of $O(n^6)$ choices for pairs of linked 2-changes.

5 Euclidean Distances

In this section, we have $\Delta_{a,b}(c) = \|c - a\|_2 - \|c - b\|_2$ for $a, b, c \in \mathbb{R}^d$. Analyzing $\|c-a\|_2 - \|c-b\|_2$ turns out to be more difficult than analyzing $\|c-a\|_2^2 - \|c-b\|_2^2$

in the previous section. In particular the case when $\|c-a\|_2 - \|c-b\|_2$ is close to the maximal possible value of $\|a-b\|_2$ requires special attention.

5.1 Difference of Euclidean Distances

As for squared Euclidean distances, we analyze the probability that a pair of linked 2-changes yields a small improvement. Assume that a, b, and c are already drawn. Then the 2-change that replaces $\{z,a\}$ and $\{b,c\}$ by $\{z,b\}$ and $\{a,c\}$ yields an improvement of at most ε only if $\eta = \|z-a\|_2 - \|z-b\|_2 = \Delta_{a,b}(z)$ falls in a particular interval of length ε. For this analysis, it does not matter which of the four points involved in the 2-change is chosen as z.

We observe that η is essentially 2-dimensional: it depends only on the distance of z from $L(a,b)$ (this is x in the following lemma) and on the position of the projection z to $L(a,b)$ (this is y in the following lemma). Furthermore, it depends on the distance $\|a-b\|_2$ between a and b (this is δ in the following lemma). The following lemma makes the connection between x and y explicit for a given η.

Lemma 5.1. *Let $z=(x,y)\in\mathbb{R}^2$, $x\geq 0$, $y \geq 0$. Let $a=(0,-\delta/2)$ and $b=(0,\delta/2)$ be two points at a distance of δ. Let $\eta = \|z-a\|_2 - \|z-b\|_2$. Then we have*

$$y^2 = \frac{\eta^2\delta^2 + 4\eta^2x^2 - \eta^4}{4\delta^2 - 4\eta^2} = \frac{\eta^2}{4} + \frac{\eta^2x^2}{\delta^2-\eta^2} \tag{1}$$

for $0 \leq \eta < \delta$ and

$$x^2 = \frac{y^2\cdot\left(4\delta^2-4\eta^2\right) + \eta^4 - \eta^2\delta^2}{4\eta^2} = \frac{y^2\cdot\left(\delta^2-\eta^2\right)}{\eta^2} - \frac{\delta^2-\eta^2}{4}. \tag{2}$$

for $\delta \geq \eta > 0$. Furthermore, $\eta > \delta$ is impossible.

In order to apply Lemma 2.6, we need the following upper bound on the derivative of y with respect to η, given that x is fixed.

Lemma 5.2. *For $x, y \geq 0$, let $y = \sqrt{\frac{\eta^2}{4} + \frac{\eta^2x^2}{\delta^2-\eta^2}}$. Assume that $0 \leq \eta \leq \delta - \kappa$ and $\kappa > 0$. Then the derivative of y with respect to η is bounded by $O\big(\frac{\delta^2+x^2}{\kappa^2}\big)$.*

If δ and x are bounded by $O(D_{\max})$, then the derivative of y with respect to η is bounded by $O(D_{\max}^2/\kappa^2)$.

We stress that Lemma 5.2 provides a rather bad upper bound on the derivative: We use an upper bound of $D_{\max}$ for x in the numerator, while $x \approx D_{\max}$ would lead to a much larger denominator and, thus, to a better bound. However, we try to keep the analysis simple, and it seems difficult to get a better compact upper bound for the derivative without case distinctions.

Using Lemmas 5.2 and 2.6, we can bound the probability that $\Delta_{a,b}(z)$ assumes a value in an interval of size ε.

Lemma 5.3. *Let $a, b \in [-D_{\max}, D_{\max}]^d$ be arbitrary, $a \neq b$, and let z be drawn according to a Gaussian distribution with standard deviation σ. Let $\delta = \|a-b\|_2 = O(D_{\max})$. Let I be an interval of length ε with $I \subseteq [0, \delta-\kappa]$. Then*

$$\mathbb{P}\big(\Delta_{a,b}(z) \in I \text{ and } z \in [-D_{\max}, D_{\max}]^d\big) = O(\varepsilon D_{\max}^2 \kappa^{-2}\sigma^{-1}).$$

5.2 Bad Events

Lemma 5.2 and, thus, Lemma 5.3 become quite weak if κ is small. This is the case if $\Delta_{a,b}(z)$ is close to its maximal possible value of $\|a-b\|_2$. In this case, z must be very close to $L(a,b)$. The following lemma states that this is unlikely.

Lemma 5.4. *For $d \geq 2$ and $0 < \alpha < \beta < 1$, let $E_{\varepsilon,\alpha,\beta}$ be the event that at least one of the following bad events occur:*

1. $X \not\subseteq [-D_{\max}, D_{\max}]^d$.
2. $\delta_{\text{close}} \leq \varepsilon^\alpha$.
3. *There exist four different points $a, b, c, c' \in X$ with $|\Delta_{a,b}(c)| \geq \|a-b\|_2 - \varepsilon^\beta$ and $|\Delta_{a,b}(c')| \geq \|a-b\|_2 - 2\varepsilon^\beta$.*

Then, for all $\varepsilon \leq \varepsilon_0$ for some ε_0 that depends on α and β, we have

$$\mathbb{P}(E_{\varepsilon,\alpha,\beta}) \leq \frac{1}{n!} + n^2 \cdot \left(\frac{\varepsilon^\alpha}{\sigma}\right)^d + n^4 \cdot \left(\frac{8\varepsilon^{\beta-\alpha} D_{\max}^2}{\sigma^2}\right)^{d-1}$$

Proof (sketch). The three terms of the bound correspond to the three parts of the bad events. The first two are immediate consequences of Lemmas 2.2 and 2.5.

For the last term and Item 3, we observe that, because $X \subseteq [-D_{\max}, D_{\max}]^d$, the event $\Delta_{a,b}(c) \geq \|a-b\|_2 - \varepsilon^\beta$ can only occur if c is within a distance of at most $2\sqrt{\varepsilon^{\alpha-\beta}} D_{\max}$ of $L(a,b)$. The probability that this happens can be bounded using Lemma 2.4. In the same way, the probability of the event $\Delta_{a,b}(c') \geq \|a-b\|_2 - 2\varepsilon^\beta$ can be bounded from above. □

5.3 Smallest Improvement of a Pair of Linked 2-Changes

In this section, we analyze the probability that there exists a pair of linked 2-changes that yields an improvement of at most ε. Simple 2-changes do not seem sufficient to yield a bound on the expected number of iterations.

Lemma 5.5. *Fix α and β, and let $\varepsilon > 0$ be sufficiently small as in Lemma 5.4. Then*

$$\mathbb{P}\big(\Delta_{\min}^{\text{link}} < \varepsilon \text{ and not } E_{\varepsilon,\alpha,\beta}\big) = O\left(n^6 \varepsilon^{2-4\beta} D_{\max}^4 \sigma^{-2}\right).$$

Proof. We analyze a fixed pair of linked 2-changes as described in Section 2.1. Then the lemma follows by a union bound over the $O(n^6)$ possible pairs. We assume that $\delta_{\text{close}} \geq \varepsilon^\alpha$. Otherwise, we would have event $E_{\varepsilon,\alpha,\beta}$ (Lemma 5.4, Item 2).

Suppose that $|\Delta_{1,4}(3)| \geq \|x_1 - x_3\| - \varepsilon^\beta$. Then, because we do not have $E_{\varepsilon,\alpha,\beta}$ (Lemma 5.4, Item 3), we have $|\Delta_{1,4}(2)| \leq \|x_1 - x_3\| - 2\varepsilon^\beta$. The improvement of the first 2-change of the linked pair is $|\Delta_{1,4}(2) - \Delta_{1,4}(3)| \geq \varepsilon^\beta \geq \varepsilon$ or it is not a 2-change as there is no improvement. In the same way, if $|\Delta_{1,4}(2)| \geq \|x_1 - x_3\| - \varepsilon^\beta$ or $\Delta_{1,6}(3) \geq \|x_1 - x_6\| - \varepsilon^\beta$ or $\Delta_{1,6}(5) \geq \|x_1 - x_6\| - \varepsilon^\beta$, at least one of the two 2-changes yields an improvement of at least $\varepsilon^\beta \geq \varepsilon$. Thus, we can ignore these cases from now on and apply Lemma 5.3 with $\kappa = \varepsilon^\beta$.

We first draw all but two of the five or six points (depending on which type of linked pair we have) such that one of the two remaining points (x_i with $i \in \{2,4\}$) is only involved in the first 2-change and the other point (x_j with $j \in \{5,6\}$) is only involved in the second 2-change. We only consider the case $i = 2$ and $j = 5$, the other cases are identical.

The first 2-change yields an improvement of at most ε only if $\Delta_{1,4}(2)$ falls into an interval of size at most ε. According to Lemma 5.3, the probability for this is at most $O(\varepsilon^{1-2\beta} D_{\max}^2 \sigma^{-1})$, as we have already ruled out the case that $X \not\subseteq [-D_{\max}, D_{\max}]^d$. Analogously, the probability that $\Delta_{1,6}(5)$ falls into an interval of length at most ε is at most $O(\varepsilon^{1-2\beta} D_{\max}^2 \sigma^{-1})$, and this is necessary for the second 2-change to yield an improvement of at most ε. By independence of x_2 and x_5, the probability that none of the two 2-changes yields an improvement of at least ε and that we do not have event $E_{\varepsilon,\alpha,\beta}$ is bounded from above by $O(\varepsilon^{2-4\beta} D_{\max}^4 \sigma^{-2})$. □

The following lemma is an immediate consequence of Lemmas 5.4 and 5.5.

Lemma 5.6. *For any* $0 < \alpha < \beta < 1$, *we have*

$$\mathbb{P}\big(\Delta_{\min}^{\text{link}} \leq \varepsilon \text{ and } X \subseteq [-D_{\max}, D_{\max}]^d\big) = O\left(n^6 \cdot \frac{\varepsilon^{2-4\beta} D_{\max}^4}{\sigma^2} + n^2 \cdot \left(\frac{\varepsilon^\alpha}{\sigma}\right)^d + n^4 \cdot \left(\frac{\varepsilon^{\beta-\alpha} D_{\max}^2}{\sigma^2}\right)^{d-1}\right).$$

Now we choose $\beta = 0.247$ and $\alpha = 0.12$. Then, for $d \geq 9$, this yields $2 - 4\beta > 1.01$, $\alpha d > 1.08$, and $(\beta - \alpha) \cdot (d - 1) > 1.01$. Using Lemma 2.3, this allows us to remove the d from the exponent, and we obtain the following simplified version of Lemma 5.6. We assume that $\sigma = O(1/\sqrt{n \log n})$ for simplicity. Thus, $D_{\max} = O(1)$.

Lemma 5.7. *For* $d \geq 9$ *and* $\sigma = O(1/\sqrt{n \log n})$, *we have* $\mathbb{P}\big(\Delta_{\min}^{\text{link}} \leq \varepsilon \text{ and } X \subseteq [-D_{\max}, D_{\max}]^d\big) = O(\varepsilon^{1.01} n^4 \sigma^{-16})$.

Using this lemma, we can prove the main result of this section.

Theorem 5.8. *For* $d \geq 9$ *and* $\sigma = O(1/\sqrt{n \log n})$, *the expected length of the longest path in the 2-opt state graph corresponding to d-dimensional instances with Euclidean distances is at most* $O(n^5/\sigma^{16})$.

6 Concluding Remarks

Improving the bounds. Our smoothed analysis for Euclidean instances works only for $d \geq 9$ and the dependence of the bound on σ is bad. With the same analysis, we can get a better bound – in particular with respect to σ – for larger values of d by adjusting Lemma 5.7. While our goal was to keep the analysis simple, we believe that a much better bound holds, also for smaller d, by exploiting techniques of Englert et al. [7] for Euclidean distances.

Similarly, we can obtain an improved bound for squared Euclidean distances by considering $d \geq 3$ and adapting Lemma 4.2.

Polynomial bound for Euclidean distances for all d. For $d \leq 8$, the bound proved by Englert et al. [7] for Euclidean distances is $O(n^{4+\frac{1}{3}} \log(n/\sigma)\sigma^{-21.4})$. By combining this with our bound, we obtain a smoothed polynomial number of iterations for all d and without d in the exponent.

Initial tour. One reason that we obtain worse bounds is that our upper bound for the length of the initial tour is worse because we do not truncate the Gaussian distributions. This effect is even stronger for Euclidean distances, where the maximum distance between points plays a role also in the analysis of the 2-changes (Lemmas 5.3 and 5.4). Only for $\sigma = O(1/\sqrt{n \log n})$, this effect is negligible, as then $D_{\max} = O(1)$.

In the same way as Englert et al. [7], we can slightly improve the smoothed number of iterations by using an insertion heuristic to choose the initial tour. We save a factor of $n^{1/d}$ for Manhattan and Euclidean distances and a factor of $n^{2/d}$ for squared Euclidean distances. The reason is that there always exist tours of length $O(D_{\max} n^{1-\frac{1}{d}})$ for n points in $[-D_{\max}, D_{\max}]^d$ for Euclidean and Manhattan distances and of length $O(D_{\max}^2 n^{1-\frac{2}{d}})$ for squared Euclidean distances for $d \geq 2$ [19].

Dependence on d. For Manhattan distances, the term hidden in the O depends exponentially on d. For Euclidean distances, the dependence is polynomially on d. For squared Euclidean distances, the term depends only linearly on d.

We conjecture that also for Manhattan distances, a bound that avoids exponential dependence on d can be proved.

Approximation ratio. Using the fact that any local optimum of 2-opt yields a tour of length at most $O(D_{\max} n^{1-\frac{1}{d}})$ [5] and that the optimal tour has a length of $\Omega(n^{1-\frac{1}{d}}\sigma)$ [7], we obtain a smoothed approximation ratio of $O(D_{\max}/\sigma)$. This, however, is worse than the worst-case ratio of $O(\log n)$ [5] as $D_{\max}/\sigma = \Omega(\sqrt{n/\log n})$. The reason for this bound is that the upper bound for the local optimum involves $D_{\max}$.

We conjecture an approximation ratio of $O(1/\sigma)$, which is what we would obtain if plugging $\sigma = \Theta(\phi^{-d})$ into the bound of Englert et al. [7] were allowed.

References

1. Arora, S.: Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems. Journal of the ACM 45(5), 753–782 (1998)
2. Arthur, D., Manthey, B., Röglin, H.: Smoothed analysis of the k-means method. Journal of the ACM 58(5) (2011)
3. Arthur, D., Vassilvitskii, S.: Worst-case and smoothed analysis of the ICP algorithm, with an application to the k-means method. SIAM Journal on Computing 39(2), 766–782 (2009)
4. Bringmann, K., Engels, C., Manthey, B., Rao, B.V.R.: Random shortest paths: Non-euclidean instances for metric optimization problems. In: Chatterjee, K., Sgall, J. (eds.) MFCS 2013. LNCS, vol. 8087, pp. 219–230. Springer, Heidelberg (2013)

5. Chandra, B., Karloff, H., Tovey, C.: New results on the old k-opt algorithm for the traveling salesman problem. SIAM Journal on Computing 28(6), 1998–2029 (1999)
6. Engels, C., Manthey, B.: Average-case approximation ratio of the 2-opt algorithm for the TSP. Operations Research Letters 37(2), 83–84 (2009)
7. Englert, M., Röglin, H., Vöcking, B.: Worst case and probabilistic analysis of the 2-Opt algorithm for the TSP. Algorithmica (to appear)
8. Funke, S., Laue, S., Lotker, Z., Naujoks, R.: Power assignment problems in wireless communication: Covering points by disks, reaching few receivers quickly, and energy-efficient travelling salesman tours. Ad Hoc Networks 9(6), 1028–1035 (2011)
9. Johnson, D.S., McGeoch, L.A.: The traveling salesman problem: A case study. In: Aarts, E., Lenstra, J.K. (eds.) Local Search in Combinatorial Optimization, ch. 8. John Wiley & Sons (1997)
10. Johnson, D.S., McGeoch, L.A.: Experimental analysis of heuristics for the STSP. In: Gutin, G., Punnen, A.P. (eds.) The Traveling Salesman Problem and its Variations, ch. 9. Kluwer Academic Publishers (2002)
11. Kern, W.: A probabilistic analysis of the switching algorithm for the TSP. Mathematical Programming 44(2), 213–219 (1989)
12. Manthey, B., Röglin, H.: Smoothed analysis: Analysis of algorithms beyond worst case. it – Information Technology 53(6), 280–286 (2011)
13. Manthey, B., Röglin, H.: Worst-case and smoothed analysis of k-means clustering with Bregman divergences. Journal of Computational Geometry 4(1), 94–132 (2013)
14. Mitchell, J.S.B.: Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric tsp, k-mst, and related problems. SIAM Journal on Computing 28(4), 1298–1309 (1999)
15. van Nijnatten, F., Sitters, R., Woeginger, G.J., Wolff, A., de Berg, M.: The traveling salesman problem under squared Euclidean distances. In: Proc. of the 27th Int. Symp. on Theoretical Aspects of Computer Science (STACS). LIPIcs, vol. 5, pp. 239–250. Schloss Dagstuhl (2010)
16. Papadimitriou, C.H.: The Euclidean traveling salesman problem is NP-complete. Theoretical Computer Science 4(3), 237–244 (1977)
17. Spielman, D.A., Teng, S.H.: Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. Journal of the ACM 51(3), 385–463 (2004)
18. Spielman, D.A., Teng, S.H.: Smoothed analysis: An attempt to explain the behavior of algorithms in practice. Communications of the ACM 52(10), 76–84 (2009)
19. Yukich, J.E.: Probability Theory of Classical Euclidean Optimization Problems. Lecture Notes in Mathematics, vol. 1675. Springer (1998)