A Fixed Budget Analysis of Randomized Search Heuristics for the Traveling Salesperson Problem

Samadhi Nallaperuma
Optimisation and Logistics
School of Computer Science
The University of Adelaide
Adelaide, Australia

Frank Neumann
Optimisation and Logistics
School of Computer Science
The University of Adelaide
Adelaide, Australia

Dirk Sudholt
Dept. of Computer Science
The University of Sheffield
Sheffield, United Kingdom

ABSTRACT

Randomized Search heuristics are frequently applied to NP-hard combinatorial optimization problems. The runtime analysis of randomized search heuristics has contributed tremendously to their theoretical understanding. Recently, randomized search heuristics have been examined regarding their achievable progress within a fixed time budget. We follow this approach and present a first fixed budget runtime analysis for a NP-hard combinatorial optimization problem. We consider the well-known Traveling Salesperson problem (TSP) and analyze the fitness increase that randomized search heuristics are able to achieve within a given fixed budget.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity

Keywords

Traveling Salesperson Problem; fitness gain; theory

1. INTRODUCTION

Randomized Search heuristics (RSH) such as randomized local search, evolutionary algorithms and ant colony optimization have become very popular in recent years to solve a wide range of hard combinatorial optimization problems. Regarding them as classical randomized algorithms [11], a lot of progress has been made in recent years on their theoretical understanding [1, 6]. Initially, most of the studies were focused on simple example functions [4]. Gradually, the analysis on combinatorial optimization [10, 12, 14] problems was also established. We refer the textbook of Neumann and Witt [13] for a comprehensive presentation on the runtime analysis of randomized search heuristics for problems from combinatorial optimization.

All these studies on analyzing of time complexity were based on a single perspective, the expected optimization time. There were slight variations on this considering the number of iterations/generations or fitness evaluations. Jansen and Zarges [7] and Zhou et al. [18] pointed out that there is a gap between the empirical results and the theoretical results obtained on the optimization time. Theoretical research most often yields asymptotic results on finding the global optimum while practitioners concern more about achieving some good result within a reasonable time budget. Furthermore, it is beneficial to know how much progress an algorithm can make given some additional time budget. Experimental studies on this topic have been carried out in the domain of algorithm engineering [15] and the term fixed budget runtime analysis has been introduced by Jansen and Zarges [8].

So far, the fixed budget analysis has been conducted for very simple test functions, on which the considered randomized search heuristics such as randomized local search and the (1+1) EA follow a typical search trajectory [3, 7, 8] with high probability. This implies that on these functions the development of the best fitness over time forms an almost deterministic curve that describes the algorithm's typical behavior. Given such a strong relation between fitness and time, and the availability of upper and lower tail bounds, it is then possible to derive tight upper and lower bounds on the expected fitness increase over any given period of time.

The goal of obtaining tight upper and lower bounds is only feasible for functions where randomized search heuristics show a typical search trajectory and tail bounds are available to bound deviations from this trajectory. This usually does not apply to hard combinatorial problems like the Traveling Salesperson Problem, and currently no fixed-budget analysis is available for such problems. We argue that for these problems fixed budget results can be obtained by relaxing the above goal towards only considering *lower bounds* on the expected fitness gain. Lower bounds can be determined based on the expected minimum improvement made in an iteration. In this manner there is no requirement for obtaining tail bounds, which drastically widens the scope of problems that can be tackled with this approach. Even though lower bounds on the expected fitness gain may not be tight, they provide proven guarantees on the progress made by a RSH. The aim of this approach is to establish guarantees on the expected fitness gain for various kinds of RSH, hence providing guidance for choosing, designing, and tuning RSH such that they find high-fitness solutions in short time.

This study provides a starting point for fixed-budget analysis of randomized search heuristics for combinatorial optimization problems. In particular, we consider random-

GECCO'14, July 12-16, 2014, Vancouver, BC, Canada.

http://dx.doi.org/10.1145/2576768.2598302

ized local search (RLS) and (1+1) Evolutionary Algorithm ((1+1) EA) on the famous Traveling Salesperson Problem (TSP). We analyze TSP instances on Manhattan and Euclidean instances in the setting of smoothed complexity [16]. Smoothed analysis provides a generic framework to analyze algorithms like 2-Opt for TSP with the capability to interpolate between average and worst case analysis. This analysis was first proposed by Spielman and Teng [16] focusing on the simplex algorithm to explain the discrepancy between its exponential worst case runtime and the fast performance in practice. The probabilistic model proposed by Englert et al. [5] is a reminiscence of the original smoothed analysis model. Later, these results were refined by Manthey and Veestra [17]. Here, we will adhere to the initial analysis by Englert et al. [5] as our major focus is on transferring these results to a fixed budget analysis of RSH.

We build on the analysis of Englert et al. [5] for 2-Opt which allows to get bounds on the expected progress of a 2-Opt operation in the smoothed setting. First, we obtain fixed budget results based on the minimum improvement that RLS and (1+1) EA can make in one iteration. We further improve these results, following [5], by analyzing a sequence of consecutive 2-Opt steps together to identify linked pairs. Interestingly, considering only single improving steps gives a constant lower bound on the progress achievable in each of the t iterations whereas the analysis of a sequence of consecutive 2-Opt steps gives a larger expected progress per step if t is large.

The organization of the paper is as follows. Section 2 describes problem context and the considered algorithms. Section 3 and section 4 contain the analysis for Manhattan and Euclidean instances respectively. Finally, Section 5 concludes with highlights and future directions.

2. PRELIMINARIES

The Traveling Salesperson problem (TSP) is one of the most famous NP-hard combinatorial optimization problems. Given a set of n cities $\{1, \ldots, n\}$ and a distance matrix $d = (d_{ij})$, $1 \leq i, j \leq n$, the goal is to compute a tour of minimal length that visits each city exactly once and returns to the origin.

A TSP instance is considered to be metric if its distance function is in metric space. Metric space satisfies reflexivity, symmetry and triangle inequality conditions. A pair (V, d) of a nonempty set V and a function $d : V \times V \rightarrow R^+$ is called a metric space if for all $x, y, z \in V$ the following properties are satisfied:

- $d(x, y) = 0$ if and only if $x = y$,
- $d(x, y) = d(y, x)$
- $d(x, z) \leq d(x, y) + d(y, z)$.

We consider the case where the n cities are given by points (x_i, y_i), $1 \leq i \leq n$, in the plane and distances are given according to the L_1 or L_2 metric. For a distance metric L_p the distance of two points $p_i = (x_i, y_i)$ and $p_j = (x_j, y_j)$ is

$$d_p(p_i, p_j) = (|x_i - x_j|^p + |y_i - y_j|^p)^{1/p}\,.$$

L_1 and L_2 are called Manhattan and Euclidean metric, respectively.

2.1 RLS and Simple Evolutionary Algorithm

We consider simple randomized search heuristics and analyze them with respect to the progress that they make within a given time budget.

We investigate Randomized Local Search (RLS) (Algorithm 1) and variants of the (1+1) EA (Algorithm 2). All algorithms work with a population size of 1 and produce 1 offspring in each iteration. A basic mutation is given by the well-known 2-Opt operator. The 2-Opt operator selects two edges $\{u_1, u_2\}$ and $\{v_1, v_2\}$ from the tour such that u_1, u_2, v_1, v_2 are distinct and appear in this order in the tour, and it replaces these edges by the edges $\{u_1, v_1\}$ and $\{u_2, v_2\}$.

RLS performs one 2-Opt step in each iteration to produce an offspring. (1+1) EA chooses an integer variable s drawn from the Poisson distribution with expectation 1 in each mutation step and performs sequentially $s+1$ 2-Opt operations. In case $s + 1 = 1$, we speak of a singular mutation, or a singular generation.

The reason we are studying the (1+1) EA is that it can simulate a 2-Opt step in singular generations, which occur with probability $1/e$, $e = \exp(1)$. Moreover, it has a positive probability of generating a global optimum in every generation through executing the right number and sequence of 2-Opt steps. So it is guaranteed to find a global optimum in finite time, though this time may be exponential in n.

Note that in our algorithms we consider the notion of fitness with regard to the minimization of the tour-length. As evolutionary algorithms often maximize fitness, we use the term *fitness gain* to describe fitness improvements, that is, the decrease of the tour length.

Algorithm 1 RLS

```
x ← a random permutation of [n]
repeat
    y := x
    y ← apply a 2-Opt step chosen uniformly to y.
    if f(y) ≤ f(x) then
        x ← y
    end if
until forever
```

Algorithm 2 (1+1) EA

```
x ← a random permutation of [n]
repeat
    s from a Poisson distribution with unit expectation
    y := x
    for s + 1 times do
        y ← apply a 2-Opt step chosen uniformly to y.
    end for
    if f(y) ≤ f(x) then
        x ← y
    end if
until forever
```

We study our algorithms regarding the expected progress that they make within a given number of t iterations. We consider the algorithms on random instances in the setting of smoothed analysis [5, 16].

In this model, n points are placed in a d-dimensional unit hypercube $[0, 1]^d$ for $d \geq 2$. Each point v_i, $i = 1, 2, \ldots, n$, is chosen independently according to its own probability density function $f_i\colon [0, 1]^d \rightarrow [0, \phi]$ for some parameter $\phi > 1$. To model worst-case instances, it is assumed that these densities are chosen by an adversary who is trying to create the most difficult random instances possible. By adjusting the parameter ϕ, one can tune the power of this adversary and hence interpolate between worst and average case. For example, for $\phi = 1$ there is only one valid choice for the densities, and every point is chosen uniformly at random from the unit hypercube. The larger ϕ, the more concentrated the probability mass can be, hence the adversary becomes more powerful in creating a worst case, and the analysis gets closer to a worst-case analysis.

This model covers a smoothed analysis also with a slight modification. There the adversary determines the initial distribution of points and then a slight perturbation is applied to each position, adding a Gaussian random variable with small standard deviation σ. There ϕ has to be set as $1/\sigma^2$. The two types of instances are called ϕ-perturbed Manhattan instances and ϕ-perturbed Euclidean instances.

Analyzing our algorithms in this setting, we may assume that any two different tours have different function value. Hence, both algorithms always accept strict improvements.

2.2 Minimum Improvement of a 2-Opt Step

We now summarize results by Englert et al. [5] on the minimal improvement of a 2-Opt step. These results will later on be used in our analysis of the randomized search heuristics. We denote the random variables Δ that describes the fitness gain obtained in one iteration and Δ_t the fitness gain in t iterations. Based on the smallest improvement of any 2-Opt step we can find the expected improvement made in t iterations of (1+1) EA. The interval $(0, \varepsilon]$ is determined by an adversary. Let us first consider a fixed 2-Opt step in which the edges e_1 and e_2 are exchanged with the edges e_3 and e_4. This 2-Opt step decreases the length of the tour by $\Delta(e_1, e_2, e_3, e_4) = d(e_1) + d(e_2) - d(e_3) - d(e_4)$.

Let $\Delta_{\min}$ denote the smallest possible improvement made by any improving 2-Opt step:

$$\Delta_{\min} = \min\{\Delta \mid \Delta > 0\}.$$

Inspired by the original ideas of Kern [9], Chandra et al. [2] bounded the probability that this smallest improvement lies within the interval $(0, \varepsilon]$ with a high probability for the uniform distribution.

We will make use of the following theorem by Englert et al. [5] which gives an upper bound on the probability that an improving 2-Opt step gives an improvement of at most ε for the Manhattan metric.

THEOREM 1 (MANHATTAN METRIC [5, THEOREM 7]). *For the Manhattan metric and any $\varepsilon > 0$, it holds*

$$\Pr(\Delta_{\min} \leq \varepsilon) \leq 4!^2 \varepsilon n^4 \phi.$$

Based on this result we get a lower bound on the probability that the smallest improvement is greater than any given ε.

Similar to the Manhattan instances, for the Euclidean instances also, the minimum improvement per a 2-Opt step is inspired by the original ideas of Kern [9]. Based on this, the expected runtime was proved polynomial for uniform distribution by Chandra et al. [2]. This was later extended for a more generalized setting having any probability distribution by Englert et al. [5].

LEMMA 2 (EUCLIDEAN METRIC [5, LEMMA 18]). *For the Euclidean metric and any $\varepsilon > 0$, it holds*

$$\Pr(\Delta_{\min} \leq \varepsilon) \leq n^4 \cdot \varepsilon \cdot \log(1/\varepsilon) \cdot \phi^3.$$

In case the considered algorithms reach a local optimum, we cannot guarantee a steady fitness gain. So instead we use the fact that local optima have a good approximation ratio. The approximation ratio for the worst local optimum with regard to 2-Opt was proven originally in Chandra et al. [2] for the uniform distribution. This was later generalized by Englert et al. [5] for any probability distribution with a given density function ϕ.

THEOREM 3 ([5, THEOREM 4]). *Let $p \in \mathbb{N} \cup \{\infty\}$. For ϕ perturbed L_p instances the expected approximation ratio of the worst tour that is locally optimal for 2-Opt is $O(\sqrt[d]{\phi})$, where d represents the dimension*

As a consequence, for Manhattan and Euclidean instances the expected approximation ratio is at most $O(\sqrt{\phi})$ as we consider instances in a 2 dimensional unit hypercube $[0, 1]^2$.

3. ANALYSIS FOR MANHATTAN INSTANCES

In this section, we first present the analysis for RLS and (1+1) EA based on the minimum possible improvement for a single 2-Opt step. We later extend the analysis for the improvement in a sequence of consecutive 2-Opt steps.

3.1 Analysis of a Single 2-Opt Step

We start by showing a lower bound on the fitness gain achievable by RLS.

THEOREM 4. *In t iterations, RLS achieves an expected fitness gain of $\Omega(t/(n^6\phi))$ or reaches a local optimal solution with expected approximation ratio $O(\sqrt{\phi})$.*

Proof. Based on Theorem 1, we get

$$\Pr(\Delta_{\min} > \varepsilon) \geq 1 - 4!^2 \varepsilon n^4 \phi = 1 - 576 \varepsilon n^4 \phi$$

as a lower bound on the probability that the minimum improvement is at least ε.

Let Δ_{imp} denote the random variable describing the fitness gain obtained in a single improving 2-Opt step. This is obviously no less than the minimum possible improvement $\Delta_{\min}$. For any fixed $\varepsilon > 0$, the expected fitness gain per one improving 2-Opt step can be lower bounded as follows:

$$\begin{aligned} E(\Delta_{\text{imp}}) &= \int_{\Delta_{\text{imp}}} \Pr(\Delta_{\text{imp}}) \cdot \Delta_{\text{imp}} \\ &\geq \Pr(\Delta_{\text{imp}} \geq \varepsilon) \cdot \varepsilon \\ &\geq \Pr(\Delta_{\min} \geq \varepsilon) \cdot \varepsilon \\ &= (1 - 576 \varepsilon n^4 \phi) \cdot \varepsilon \end{aligned}$$

Setting $\varepsilon = 1/(2 \cdot 576 n^4 \phi)$ we get

$$\Pr(\Delta_{\min} \geq \varepsilon) \geq 1/2$$

and accordingly

$$E(\Delta_{\text{imp}}) \geq 1/(2304 n^4 \phi).$$

The number of mutations occurring in one iteration is 1 and the probability for an improving 2-Opt step is at least $1/\binom{n}{2} \geq 2/n^2$ if the current solution is not locally optimal. Therefore, the expected value for the fitness gain Δ in any 2-Opt step can be lower bounded as

$$E(\Delta) \geq 1/(1152 n^6 \phi).$$

Hence, the expected value for the fitness gain in t iterations if no locally optimal solution has been obtained in between can be derived as $E(\Delta_t) \geq t/(1152n^6\phi)$.

By Theorem 3, a locally optimal solution has expected approximation ratio $O(\sqrt{\phi})$ which completes the proof. $\square$

THEOREM 5. *In t iterations (1+1) EA achieves an expected fitness gain of $\Omega(t/(n^6\phi))$ or reaches a local optimal solution with expected approximation ratio $O(\sqrt{\phi})$.*

Proof. The expectation for the fitness gain in an improving 2-Opt step can be derived similar to the above proof in Theorem 4

$$E(\Delta_{\text{imp}}) \geq 1/(2304n^4\phi).$$

The probability of singular mutation occur in a generation is $1/e$ due to the Poisson distribution with unit expectation. The probability of an improving 2-Opt step is therefore at least $2/(en^2)$. Thus the expected value for the fitness gain Δ in any 2-Opt step is

$$E(\Delta) \geq 1/(1152en^6\phi).$$

Hence, the expected value for the fitness gain in t iterations can be derived as $E(\Delta_t) \geq t/(1152en^6\phi)$.

By Theorem 3, a locally optimal solution has expected approximation ratio $O(\sqrt{\phi})$ which completes the proof. $\square$

3.2 Analysis of Linked Steps for RLS

The lower bound for the expected fitness gain presented in previous section is based on the minimum improvement a single 2-Opt step can make. This bound can be improved further by considering the improvement made in a sequence of consecutive 2-Opt steps.

The analysis of consecutive steps in Englert et al. [5] is based on the number of disjoint pairs of 2-Opt steps linked by an edge, such that in one step an edge is added and in the other it is removed. Different types of linked pairs of 2-Opt steps are considered as follows. Let $\{v_1, v_2\}$ and $\{v_3, v_4\}$ be the edges that are replaced by $\{v_1, v_3\}$ and $\{v_2, v_4\}$ in the first 2-Opt step, and $\{v_1, v_3\}$ and $\{v_5, v_6\}$ be replaced by $\{v_1, v_5\}$ and $\{v_3, v_6\}$ in the second 2-Opt step.

Following [5], we consider three different types of steps:

type 0: $|\{v_2, v_4\} \cap \{v_5, v_6\}| = 0$.

type 1: $|\{v_2, v_4\} \cap \{v_5, v_6\}| = 1$.

type 2: $|\{v_2, v_4\} \cap \{v_5, v_6\}| = 2$.

As explained in [5], it is important to limit the number of occurrences of type 2 as no guarantee on the fitness gain made by type 2 steps is available. We need to show that there is a sufficient number of linked pairs of type 0 and 1 as for linked pairs of type 0 and 1 a good progress can be guaranteed.

Due to [5, Lemma 9] there are at least $t/6 - 7n(n-1)/24$ such pairs in a sequence of t consecutive 2-Opt steps. The analysis in [5] considers all 2-Opt steps $S_1, \ldots, S_t$ in sequence and constructs disjoint linked pairs (of any type) in a greedy fashion. When processing some step S_i, we search for steps S_j and S'_j, where the two edges inserted by S_i are being removed again, if such steps exist. If either S_j or S'_j exist, the respective pair (S_i, S_j) or (S_i, S'_j) is being added to a list of disjoint linked 2-Opt steps, and both S_j and S'_j are being removed from the list to ensure disjointness of pairs. The proof of [5, Lemma 9] then shows that when removing all pairs of type 2 from this list, at least $t/6 - 7n(n-1)/24$ pairs of type 0 or 1 remain.

We further improve this bound, considering the fact that the possible number of pairs excluded is constrained by the number of edges in the final tour.

LEMMA 6. *Let u be the total number of linked pairs in an improving 2-Opt sequence. Then the number of linked pairs v of type 0 or 1 in that sequence is at least $u/2 - n/4$.*

Proof. Following the argument in the proof of [5, Lemma 9] there cannot be a type 2 linked pair that associates with another type 2 linked pair. Therefore, each of the type 2 pairs (S_i, S_j) can be associated with at most two different pairs (S_j, S_ℓ) and $(S_j, S_{\ell'})$ of type 0 or 1, unless the steps S_ℓ or $S_{\ell'}$ are undefined. This happens if the edges added to the tour in S_j are never removed. Since the final tour contains n edges, at most $n/2$ pairs are excluded due to this. If we consider the number of type 2 pairs as x then the total number of pairs of type 0 or 1 must be at least $x - n/2$. This implies $u \geq x + (x - n/2)$ and $x \leq u/2 + n/4$. The number of good pairs is therefore $u - x \geq u/2 - n/4$. $\square$

Using the same argument, we can improve [5, Lemma 8] on the total number of disjoint linked pairs. There, the authors show that for each processed 2-Opt step S_i at most two other steps S_j S'_j are excluded from being processed if j or j' is defined. Hence, for a sequence of t steps there are at least $t/3$ pairs, from which we have to subtract the number of steps S_i where neither j nor j' are defined. Englert et al. [5, Lemma 8] argue that we need to subtract a number of $n(n-1)/4$ steps. However, this number can be improved from $n(n-1)/4$ to $n/2$ considering that the number of edges in the final tour is exactly n, and S_i can only be excluded if both edges inserted in the tour are never removed again. Therefore, the total number of disjoint pairs is at least $u = t/3 - n/2$. Combining the result of above Lemma 6 to this we obtain the number of disjoint pairs of type 0 and 1.

LEMMA 7. *In every sequence of t consecutive 2-Opt steps, the number of disjoint pairs of 2-Opt steps of type 0 or 1 is at least $t/6 - n/2$.*

Due to above Lemma 7 there are at least $t/6 - n/2$ such pairs in a sequence of t consecutive 2-Opt steps. Here we consider the probability of both 2-Opt steps in a linked pair having improvement at least ε.

LEMMA 8 ([5, LEMMA 10]). *In a ϕ perturbed L1 instance with n vertices, the probability that there exists a pair of type 0 or 1 in which both 2-Opt steps are improvements by at most ε is bounded by $O(n^6 \cdot \varepsilon^2 \cdot \phi^2)$.*

Based on the above Lemmas (7 and 8), we can bound the fitness gain for a given number of t iterations. Note that the Theorem requires a lower bound on the number of iterations as only for large enough t we can guarantee that linked 2-Opt steps of type 0 or 1 do exist.

THEOREM 9. *In $t \geq cn^3$ iterations, $c > 3/2$ constant, RLS obtains an expected fitness gain of $\Omega(t/(n^5\phi))$ unless it reaches a local optimum. In that case, expected approximation ratio of the solution is $O(\sqrt{\phi})$.*

Proof. Let $\Delta_{\min}$ denote the minimum possible improvement made by any pair of type 0 or 1. Using above result in Lemma 8,

$$\Pr(\Delta_{\min} > \varepsilon) \geq 1 - n^6 \cdot \varepsilon^2 \cdot \phi^2.$$

Let Δ be the random variable describes the fitness gain obtained in a pair of linked 2-Opt steps of type 0 or 1. For any $\varepsilon > 0$, the expected fitness gain $E(\Delta)$ can be bounded as follows.

$$\begin{aligned} E(\Delta) &= \int_{\Delta} \Pr(\Delta) \cdot \Delta \\ &\geq \Pr(\Delta \geq \varepsilon) \cdot \varepsilon \\ &\geq (1 - n^6 \cdot \varepsilon^2 \cdot \phi^2) \cdot \varepsilon. \end{aligned}$$

Setting $\varepsilon = 1/(2\sqrt{(n^6 \cdot \phi^2)})$ we get $\Pr(\Delta > \varepsilon) \geq 1/2$ and as a consequence

$$E(\Delta) \geq 1/(4\sqrt{(n^6 \cdot \phi^2)}) \geq 1/(4n^3\phi).$$

The expected number of improving 2-Opt steps made in t iterations is at least $2t/n^2$. Let t^* be the number of improving steps. By Lemma 7 we know there are at least $t^*/6 - n/2$ type 0 or 1 pairs in a sequence of t^* improving steps. As $t \geq cn^3$ for $c > 3/2$, we have $2t = (2-3/c)\cdot t + 3t/c \geq (2-3/c)\cdot t + 3n^3$ and

$$E(t^*/6 - n/2) \geq \frac{2t}{6n^2} - \frac{n}{2} \geq \frac{(2-3/c)\cdot t}{6n^2} = \Omega(t/n^2).$$

A lower bound for the expected fitness gain for t iterations is therefore

$$E(\Delta_t) = E(\Delta) \cdot \Omega(t/n^2) \geq \Omega(t/(n^5\phi)).$$

By Theorem 3, a locally optimal solution has expected approximation ratio $O(\sqrt{\phi})$ which completes the proof. □

3.3 Analysis of Linked Steps for (1+1) EA

The challenge for analyzing the (1+1) EA instead of RLS lies in the fact that the (1+1) EA can execute multiple 2-Opt steps in one generation. For RLS Englert et al. [5] showed that certain pairs of improving 2-Opt steps yield a large fitness increase on perturbed instances, with high probability. Executing multiple 2-Opt steps in one generation complicates this argument, as some of these mutations may not be improving. As such, they might interfere with the mentioned pairs, and prohibit a large fitness increase. In the following, we show that there are sufficiently many linked 2-Opt operations that take place in generations where only one 2-Opt step is executed.

To this end, we consider a slightly modified variant of the (1+1) EA, which we call (1+1) EA* (see Algorithm 3). The (1+1) EA* will exclude generations containing multiple 2-Opt steps where an edge e is being inserted in one of these 2-Opt steps and being removed in a later 2-Opt step of the same generation.

The purpose of this modification is to enable a theoretical analysis as some of the excluded steps are difficult to handle. (1+1) EA and (1+1) EA* show identical behavior most of the time; it is easy to show that the probability of removing an edge that was inserted in the same generation is at most $O(1/n)$. So (1+1) EA and (1+1) EA* are identical most of the time, apart from a vanishingly small fraction of steps.

Algorithm 3 (1+1) EA*

```
x ← a random permutation of [n]
repeat
    Choose s from a Poison distribution with unit expectation
    for s + 1 times do
        y ← Mutate(x)
    end for
    if f(y) ≤ f(x) then
        check whether one of the above mutations has removed
an edge that was inserted in the same generation
        if so, reject y. Otherwise, x ← y
    end if
until forever
```

Note that, when the (1+1) EA* does behave differently from the (1+1) EA, it rejects a new offspring that would otherwise improve the current tour. We therefore believe that we are being pessimistic by considering the progress of the (1+1) EA* instead of that of the (1+1) EA.

LEMMA 10. *In every sequence of t generations of the (1+1) EA*, the expected number of disjoint pairs of 2-Opt steps, both of which are singular, is at least*

$$\frac{t}{3e^2n^2} - n/2,$$

unless a local optimum is reached beforehand.

Proof. We call a 2-Opt step *improving* if it does not decrease the current fitness. A 2-Opt step is called *singular* if it is the only 2-Opt step executed in that generation.

We adapt the proof of Lemma 8 in [5] to take into account steps that are rejected by the (1+1) EA*, and the fact that the (1+1) EA* can accept non-improving 2-Opt steps in generations with multiple 2-Opt steps.

Let $\mathcal{S} = S_1, S_2, \ldots$ be a list of all 2-Opt steps executed in t generations. Then we process this list to create a list $\mathcal{L}$ of linked 2-Opt steps, both of which are singular.

The probability of the (1+1) EA* making an improving 2-Opt step is at least $1/\binom{n}{2} \geq 2/n^2$, so long as no local optimum has been reached. The probability that an improving step takes place in a singular generation is $1/e$ due to the Poisson distribution. So, the probability of having an improving and singular step is at least $2/(en^2)$.

Let S_i be such an improving and singular 2-Opt step, and assume w. l. o. g. that edges e_1, e_2 are exchanged with the edges e_3, e_4. Then we process S_i to try to find a linked operation, which is both improving and singular. More precisely, let S_j be the next 2-Opt step where e_3 is being removed and the outcome of that generation is accepted, if such a step exists. Let S'_j be the next 2-Opt step where e_4 is being removed and the outcome of that generation is accepted, if such a step exists. If either S_j or S'_j exists and if one of these steps is singular, we add the corresponding pair (S_i, S_j) or (S_i, S'_j) to $\mathcal{L}$ and remove both S_j and S'_j from $\mathcal{S}$ to ensure disjointness of pairs. Otherwise, we proceed with the next improving and singular 2-Opt step following S_i.

We estimate the probability of a step S_j occurring and being a singular step. Let $A(e_3)$ denote the event that an accepted generation contains an improving 2-Opt step where e_3 is being removed from the tour.

Let $R(e_3)$ denote the set of all edges e such that a 2-Opt move removing e_3 and e results in a strict fitness improvement. Let x_ℓ denote the search point of the (1+1) EA* at time ℓ, and let us regard x_ℓ as a set of edges in the tour. Note that then $|R(e_3) \cap x_\ell|$ describes the number of improving 2-Opt moves where e_3 is being removed from the tour.

Let ℓ be the index of the first 2-Opt step in a new generation, and let $S+1$ be the random number of 2-Opt steps being executed in that generation. If $S = 0$, that is, only one 2-Opt step is executed, the conditional probability of $A(e_3)$ is given by

$$\Pr(A(e_3) \mid S = 0) = \frac{|R(e_3) \cap x_\ell|}{\binom{n}{2}} := p.$$

If $S+1 > 1$ operations are being executed in that generation, the probability of $A(e_3)$ is bounded by the union bound:

$$\Pr(A(e_3) \mid S = s) \leq \sum_{k=0}^{s} \frac{|R(e_3) \cap x_{\ell+k}|}{\binom{n}{2}}.$$

Note that $|R(e_3) \cap x_\ell|$ might increase if edges from $R(e_3)$ are being inserted into the tour. However, the additional selection criterion on the (1+1) EA* implies that, if a following step removes e_3 and one of the edges inserted previously, in the same generation, this sequence of 2-Opt steps will be rejected. Thus, only $|R(e_3) \cap x_\ell|$ edges can cause $A(e_3)$ and

$$\Pr(A(e_3) \mid S = s) \leq \sum_{k=0}^{s} \frac{|R(e_3) \cap x_\ell|}{\binom{n}{2}} = (s+1) \cdot p.$$

Note that, using the union bound for $S+1$ trials,

$$\begin{aligned}\Pr(A(e_3)) &= \sum_{s=0}^{\infty} \frac{1}{es!} \cdot \Pr(A(e_3) \mid S = s) \\ &\leq \sum_{s=0}^{\infty} \frac{1}{es!} \cdot (s+1)p = 2p.\end{aligned}$$

Combining this with Bayes' Theorem, we get

$$\begin{aligned}\Pr(S = 0 \mid A(e_3)) &= \frac{\Pr(A(e_3) \mid S = 0) \cdot \Pr(S = 0)}{\Pr(A(e_3))} \\ &\leq \frac{p \cdot 1/e}{2p} = \frac{1}{2e}.\end{aligned}$$

It follows that the probability of finding a linked pair (S_i, S_j) or (S_i, S'_j) is at least $1/(2e)$, if one of the steps S_j or S'_j exists.

Recall that the expected number of singular and improving steps S_i in t generations is at least $2t/(en^2)$. Each processed element S_i excludes at most 2 other elements of $\mathcal{S}$. This leaves an expected number of $2t/(3en^2)$ processed elements S_i, each of which has a probability of $1/(2e)$ for pairing with some S_j or S'_j, if one of them exists. A processed element S_i is excluded if neither S_j nor S'_j exist. This only happens if both edges are never removed from the tour. Since the final tour contains n edges, at most $n/2$ steps S_i are excluded. Hence, the resulting expected number of pairs is at least

$$\frac{t}{3e^2n^2} - n/2.$$

□

The following Theorem now gives a lower bound on the expected fitness gain of the (1+1) EA*.

THEOREM 11. *In $t \geq cn^3$ generations, $c > 3e^2$ constant, (1+1) EA* obtains an expected fitness gain of $\Omega(t/(n^5\phi))$ unless it reaches a local optimum. In that case, expected approximation ratio of the solution is $O(\sqrt{\phi})$.*

Proof. As in the proof of Theorem 9, we have

$$E(\Delta) \geq 1/(4n^3\phi).$$

From Lemma 10 we know that the expected number of disjoint pairs of 2-Opt steps, both of which are singular, is at least

$$u := \frac{t}{3e^2n^2} - \frac{n}{2}.$$

Lemma 7 implies that among these there are at least

$$\frac{u}{2} - \frac{n}{4} = \frac{t}{6e^2n^2} - \frac{n}{2} = \frac{(1 - 3e^2/c)t}{6e^2n^2} = \Omega(t/n^2)$$

type 0 or 1 pairs. The expected fitness gain in t generations, $E(\Delta_t)$, is therefore at least

$$E(\Delta) \cdot \Omega(t) \geq \Omega(t/(n^5\phi)).$$

The expected approximation ratio is proved in Theorem 3. □

4. ANALYSIS FOR EUCLIDEAN INSTANCES

We now turn our attention to the Euclidean instances. First we obtain the expected progress based on a single 2-Opt step for RLS and (1+1) EA, later improve these results by analyzing a sequence of consecutive 2-Opt steps.

4.1 Analysis of a Single 2-Opt Step

THEOREM 12. *In t iterations RLS achieves an expected fitness gain of $\Omega(t \log(n\phi)/(n^6\phi^3))$ unless it reaches a local optimum. In that case, expected approximation ratio of the solution is $O(\sqrt{\phi})$.*

Proof. Due to Theorem 2, we have

$$\Pr(\Delta_{\min} < \varepsilon) \leq n^4 \cdot \varepsilon \cdot \log(1/\varepsilon) \cdot \phi^3.$$

Let Δ_{imp} denote the random variable that describes the fitness gain in an improving 2-Opt step. Then similar to the proof for the Manhattan instances (Theorem 4), we get

$$E(\Delta_{\text{imp}}) \geq (1 - (n^4 \cdot \varepsilon \cdot \log(1/\varepsilon) \cdot \phi^3)) \cdot \varepsilon.$$

Setting $\varepsilon = c\log(n^4\phi^3)/(n^4\phi^3)$, $c > 0$ a constant such that $\Pr(\Delta_{\text{imp}} > \varepsilon) \geq 1/2$, we get

$$E(\Delta_{\text{imp}}) \geq c\log(n^4\phi^3)/(2n^4\phi^3).$$

The number of mutations occurring in one iteration is 1 and the probability for an improving 2-Opt step is at least $1/\binom{n}{2} \geq 2/n^2$. Therefore, the expected value for the fitness gain Δ in any 2-Opt step is

$$E(\Delta) \geq c\log(n^4\phi^3)/(n^6\phi^3).$$

The expected value for the fitness gain in t iterations is

$$\begin{aligned}E(\Delta_t) &\geq \Omega(t\log(n^4\phi^3)/(n^6\phi^3)) \\ &= \Omega(t\log(n\phi)/(n^6\phi^3)).\end{aligned}$$

Having obtained a locally optimal solution during the run implies an expected approximation ratio of $O(\sqrt{\phi})$ according to Theorem 3. □

THEOREM 13. *In t generations (1+1) EA achieves an expected fitness gain of $\Omega(t\log(n\phi)/(n^6\phi^3))$ unless it reaches a local optimum. In that case, the expected approximation ratio of the solution is $O(\sqrt{\phi})$.*

Proof. Similar to the above proof on RLS in Theorem 12 the expected fitness gain for an improving singular generation can be derived as

$$E(\Delta_{\mathrm{imp}}) \geq c\log(n^4\phi^3)/(2n^4\phi^3).$$

The probability of a single step mutation occur in a generation is $1/e$. This minimum fitness gain is due to any accepted singular steps. Therefore, we consider only singular steps. And the waiting time for the correct mutation is at most $2/n^2$. Hence, the expected fitness gain $E(\Delta)$ for any generation is

$$E(\Delta) \geq c\log(n^4\phi^3)/(en^6\phi^3).$$

The expected value for the fitness gain in t generations is derived accordingly:

$$\begin{aligned} E(\Delta_t) &\geq (2/(en^2))\varepsilon t = \Omega(t\log(n^4\phi^3)/(n^6\phi^3)) \\ &= \Omega(t\log(n\phi)/(n^6\phi^3)). \end{aligned}$$

Having obtained a locally optimal solution during the run implies an expected approximation ratio of $O(\sqrt{\phi})$ according to Theorem 3. □

4.2 Analysis of Linked Steps for RLS

The above lower bounds are based on the minimum possible improvement a single 2-Opt step can make. We can further improve this bound considering the improvement made in a sequence of consecutive steps. Similar to the analysis on the consecutive steps for Manhattan instances in section 3 here also we consider the set of linked pairs of type 0 and 1. In a sequence of t iterations, there are at least $t/6 - n/2$ such pairs due to Lemma 7. Here we consider Lemma 14 in Englert et al. [5] related to the probability of existence of each of the two types of linked pairs in a sequence of consecutive steps for Euclidean instances. Based on these we can bound the expected fitness gain made in t iterations.

LEMMA 14 ([5, LEMMA 14]). *For ϕ perturbed L_2 instances, the probability that there exists a pair of type 0 and 1 in which both 2-Opt steps are improvements by at most $\varepsilon \leq 1/2$ is bounded by $O(n^6\cdot\phi^5\cdot\varepsilon^2\cdot\log^2(1/\varepsilon)) + O(n^5\cdot\phi^4\cdot\varepsilon^{3/2}\log(1/\varepsilon))$.*

THEOREM 15. *In $t \geq cn^3$ iterations, $c > 3/2$ constant, RLS achieves an expected fitness gain of $\Omega(t\sqrt{\log(n\phi)}/(n^5\phi^{5/2}))$ unless it reaches a local optimum. In that case, the expected approximation ratio is $O(\sqrt{\phi})$.*

Proof. Using Lemma 14, the probability that the improvement $\Delta_{\min}$ in a linked 2-Opt step of type 0 or 1 is less than ε is at most

$$\Pr(\varepsilon) = O(n^6\cdot\phi^5\cdot\varepsilon^2\cdot\log^2(1/\varepsilon)) + O(n^5\cdot\phi^4\cdot\varepsilon^{3/2}\cdot\log(1/\varepsilon)).$$

Following the proof ideas of Theorem 9 on the consecutive 2-Opt steps for Manhattan instances, the expected fitness gain $E(\Delta)$ for a pair of linked 2-Opt steps of type 0 or 1 can be bounded from below as

$$E(\Delta) \geq (1-\Pr(\varepsilon))\cdot\varepsilon.$$

We set $\varepsilon = c'\sqrt{\log(n^6\phi^5)}/\sqrt{n^6\phi^5}$ for a constant $c' > 0$ such that $\Pr(\Delta > \varepsilon) \geq 1/2$. This implies

$$E(\Delta) \geq c'\sqrt{\log(n^6\phi^5)}/(2n^3\phi^{5/2}).$$

The expected number of improving 2-Opt steps made in t iterations is at least $2t/n^2$. Let t^* be the number of improving steps. By Lemma 7 we know there are at least $t^*/6 - n/2$ type 0 or 1 pairs in a sequence of t^* improving steps. As $t \geq cn^3$ for $c > 3/2$, we get

$$E(t^*/6 - n/2) = \frac{2t}{6n^2} - \frac{n}{2} = \frac{(2-3/c)t}{6n^2} = \Omega(t/n^2).$$

A lower bound for the expected fitness gain for t iterations is therefore

$$\begin{aligned} E(\Delta_t) = E(\Delta)\cdot\Omega(t/n^2) &\geq \Omega(t\sqrt{\log(n^6\phi^5)}/(n^5\phi^{5/2})) \\ &= \Omega(t\sqrt{\log(n\phi)}/(n^5\phi^{5/2})). \end{aligned}$$

The bound on the expected approximation ratio in the case that a locally optimal solution has been obtained holds according to Theorem 3. □

4.3 Analysis of Linked Steps for (1+1) EA

We improve the current results for (1+1) EA with the analysis for consecutive 2-Opt steps in a similar way to the analysis presented in the previous section. Again, we consider the (1+1) EA* but conjecture that the expected fitness gain in the (1+1) EA is no smaller than that for the (1+1) EA*. Based on our arguments on the number of type 0 or 1 linked pairs from Lemmas 10 and 7 and the stated Lemma 14 of [5] on the probability of the existence of a pairs of both improving steps we can bound the expected fitness gain in t generations.

THEOREM 16. *In $t \geq cn^3$ generations, $c > 3e^2$ constant, (1+1) EA* achieves an expected fitness gain of $\Omega(t\sqrt{\log(n\phi)}/(n^3\phi^{5/2}))$ unless it reaches a local optimum. In that case, the expected approximation ratio is $O(\sqrt{\phi})$.*

Proof. Following the proof ideas in above Theorem 15 we get for a $c' > 0$

$$E(\Delta) \geq c'\sqrt{\log(n^6\phi^5)}/(2n^3\phi^{5/2}).$$

From Lemma 10 we know that the number of disjoint pairs of 2-Opt steps, both of which are singular, is at least

$$u := \frac{t}{3e^2n^2} - \frac{n}{2}.$$

Lemma 7 implies that among these there are at least

$$\frac{u}{2} - \frac{n}{4} = \frac{t}{6e^2n^2} - \frac{n}{2} = \frac{(1-3e^2/c)t}{6e^2n^2} = \Omega(t/n^2)$$

type 0 or 1 pairs.

The expected fitness gain in t generations, $E(\Delta_t)$, is therefore at least

$$\begin{aligned} E(\Delta)\cdot\Omega(t/n^2) &\geq \Omega(t\sqrt{\log(n^6\phi^5)}/(n^5\phi^{5/2}) \\ &= \Omega(t\sqrt{\log(n\phi)}/(n^5\phi^{5/2}). \end{aligned}$$

Having obtained a locally optimal solution with respect to the 2-Opt neighbourhood implies the bound on the expected approximation ratio according to Theorem 3. □

Metric	RLS		(1+1) EA	
	Single Step (any t)	Consecutive Steps ($t \geq 3/2 \cdot n^3$)	Single Step (any t)	Consecutive Steps ($t \geq 3e^2/2 \cdot n^3$)
Manhattan	$\Omega(t/(n^6\phi))$	$\Omega(t/(n^5\phi))$	$\Omega(t/(n^6\phi))$	$\Omega(t/(n^5\phi))$
Euclidean	$\Omega(t\log(n\phi)/(n^6\phi^3))$	$\Omega(t\sqrt{\log(n\phi)}/(n^5\phi^{5/2}))$	$\Omega(t\log(n\phi)/(n^6\phi^3))$	$\Omega(t\sqrt{\log(n\phi)}/(n^5\phi^{5/2}))$

Table 1: Expected fitness gain in t iterations for RLS and (1+1) EA for Manhattan and Euclidean instances due to single-step and consecutive-steps analysis. The former applies for any time span t; the latter requires $t = \Omega(n^3)$. The consecutive-steps analysis was formally proven for the (1+1) EA* and transfers to the (1+1) EA if, as conjectured, the latter does not perform worse. All fitness gains assume that no local optimum is reached. otherwise the expected approximation ratio is $O(\sqrt{\phi})$.

5. CONCLUSIONS

We have carried out a fixed budget analysis of randomized local search (RLS) and variants of the (1+1) EA on the well-known Traveling Salesperson Problem (TSP). Our analysis allows to estimate the progress, or fitness gain, that these algorithms make within a given number of t iterations. This is, in particular, useful as it gives a guarantee to practitioners on the progress that such algorithms can make when deciding between stopping the algorithm or giving it additional running time.

We analyzed the algorithms in the setting of smoothed complexity for the Manhattan and Euclidean metric. We provided lower bounds on the expected fitness gain based on the minimum improvement the algorithms (RLS and (1+1) EA) can make in an iteration. The results show that for any number of iterations both algorithms gain a fair improvement based on single 2-Opt steps. We further improved these results by analyzing a sequence of consecutive 2-Opt steps together to identify linked pairs. Table 5 summarizes these results. It is observed that a larger improvement can be obtained considering the consecutive steps, for this however, the number of iterations t needs to be at least $\Omega(n^3)$.

The variant of the (1+1) EA ((1+1) EA*) analyzed for the linked steps accepts fewer solutions than the classical (1+1) EA and therefore we expect (1+1) EA* to be slower than (1+1) EA. Proving this is an interesting technical open problem and would give additional insights into the advantages of mutations that make multiple changes at the same time.

Acknowledgments: This research has been supported by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 618091 (SAGE) and by the Australian Research Council through grants DP130104395 and DP140103400. We thank Joseph Kempka and Anh Quang Nguyen for early discussions, and the participants and organizers of Dagstuhl seminar 13271 "Theory of Evolutionary Algorithms", where this research started.

References

[1] A. Auger and B. Doerr. *Theory of Randomized Search Heuristics: Foundations and Recent Developments*. World Scientific Publishing Co., Inc., 2011.

[2] B. Chandra, H. Karloff, and C. Tovey. New results on the old k-Opt algorithm for the TSP. In *Proc. of SODA'94*, pages 150–159, 1994.

[3] B. Doerr, T. Jansen, C. Witt, and C. Zarges. A method to derive fixed budget results from expected optimisation times. In *Proc. of GECCO'13*, pages 1581–1588. ACM, 2013.

[4] S. Droste, T. Jansen, and I. Wegener. On the analysis of the (1+1) evolutionary algorithm. *Theoretical Computer Science*, 276(1–2):51–81, 2002.

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

[6] T. Jansen. *Analyzing Evolutionary Algorithms - The Computer Science Perspective*. Natural Computing Series. Springer, 2013.

[7] T. Jansen and C. Zarges. Analysis of evolutionary algorithms: From computational complexity analysis to algorithm engineering. In *Proc. of FOGA '11*, pages 1–14. ACM, 2011.

[8] T. Jansen and C. Zarges. Fixed budget computations: a different perspective on run time analysis. In *Proc. of GECCO'12*, pages 1325–1332, 2012.

[9] W. Kern. A probabilistic analysis of the switching algorithm for the euclidean TSP. *Mathematical programming: Series A*, 44(1-3):213–219, 1989.

[10] J. Lässig and D. Sudholt. Analysis of speedups in parallel evolutionary algorithms for combinatorial optimization. In *Proc. of ISAAC'11*, pages 405–414. Springer, 2011.

[11] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[12] F. Neumann and I. Wegener. Randomized local search, evolutionary algorithms, and the minimum spanning tree problem. In *Proc. of GECCO'04*, pages 713–724, 2004.

[13] F. Neumann and C. Witt. *Bioinspired Computation in Combinatorial Optimization: Algorithms and Their Computational Complexity*. Springer, 2010.

[14] P. S. Oliveto, J. He, and X. Yao. Time complexity of evolutionary algorithms for combinatorial optimization: A decade of results. *International Journal of Automation and Computing*, 4(3):281–293, 2007.

[15] P. Sanders and D. Wagner. Algorithm engineering. *Informatik-Spektrum*, 36(2):187–190, 2013.

[16] D. A. Spielman and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *J. ACM*, 51(3):385–463, May 2004.

[17] B. M. R. Veenstra. Smoothed analysis of the 2-Opt heuristic for the TSP: Polynomial bounds for gaussian noise. In *Proc. of ISAAC'13*, 2013.

[18] D. Zhou, D. Luo, R. Lu, and Z. Han. The use of tail inequalities on the probable computational time of randomized search heuristics. *Theoretical Computer Science*, 436(0):106 – 117, 2012.