Swarm Intell (2012) 6:1–21
DOI 10.1007/s11721-011-0059-7

Theoretical analysis of two ACO approaches for the traveling salesman problem

Timo Kötzing · Frank Neumann · Heiko Röglin · Carsten Witt

Received: 26 October 2010 / Accepted: 8 September 2011 / Published online: 15 October 2011

Abstract Bioinspired algorithms, such as evolutionary algorithms and ant colony optimization, are widely used for different combinatorial optimization problems. These algorithms rely heavily on the use of randomness and are hard to understand from a theoretical point of view. This paper contributes to the theoretical analysis of ant colony optimization and studies this type of algorithm on one of the most prominent combinatorial optimization problems, namely the traveling salesperson problem (TSP). We present a new construction graph and show that it has a stronger local property than one commonly used for constructing solutions of the TSP. The rigorous runtime analysis for two ant colony optimization algorithms, based on these two construction procedures, shows that they lead to good approximation in expected polynomial time on random instances. Furthermore, we point out in which situations our algorithms get trapped in local optima and show where the use of the right amount of heuristic information is provably beneficial.

Keywords Ant colony optimization · Traveling salesperson problem · Run time analysis · Approximation

1 Introduction

Bioinspired algorithms, such as evolutionary algorithms (EAs) (Eiben and Smith 2007) and ant colony optimization (ACO) (Dorigo and Stützle 2004), are robust problem solvers that

T. Kötzing (✉)
Algorithms and Complexity, Max-Planck-Institut für Informatik, 66123 Saarbrücken, Germany
e-mail: koetzing@mpi-inf.mpg.de

F. Neumann
School of Computer Science, University of Adelaide, Adelaide, SA 5005, Australia

H. Röglin
Department of Computer Science, University of Bonn, Bonn, Germany

C. Witt
DTU Informatics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

have found a wide range of applications in various problem domains. There are many successful applications of this type of algorithms, but the theoretical understanding lags far behind their practical success. Therefore, it is highly desirable to increase the theoretical understanding of these algorithms in order to gain new insights into their working principles, which may lead to the design of even more successful approaches.

With this paper, we contribute to the theoretical understanding of bioinspired computing by rigorous runtime analyses. There have been many successful investigations of the runtime behavior of evolutionary algorithms, which have highly increased their theoretical foundation (see Neumann and Witt 2010 for a wide range of results on classical combinatorial optimization problems). Regarding ACO algorithms, the theoretical analysis of their runtime behavior seems to be much more difficult and has been started in 2006. First rigorous insights have been provided by runtime analysis on classical pseudo-Boolean functions (Gutjahr 2007; Gutjahr and Sebastiani 2008; Neumann et al. 2008, 2009; Neumann and Witt 2009). Furthermore, there are results on classical combinatorial optimization problems such as the computation of minimum spanning trees (Neumann and Witt 2010) and shortest paths (Horoba and Sudholt 2009). Moreover, initial results have been obtained for the well-known traveling salesperson problem (TSP) (Zhou 2009). Regarding ACO algorithms, the TSP is the first problem to which this kind of algorithm has been applied. This makes it a natural choice for studying the behavior of ACO algorithms for the TSP from a theoretical point of view. We contribute to this line of research and increase the theoretical understanding of ACO algorithms by increasing the knowledge on the runtime behavior of ACO algorithms for the TSP problem.

The design of ACO algorithms is inspired by the ability of ants to find a shortest path between their nest and a common source of food. It is well known that ants find such a path very quickly by using indirect communication via pheromones. This ants behavior inspired researchers to build an algorithmic framework, that uses artificial ants to solve optimization problems. In an ACO algorithm, solutions for a given problem are constructed by artificial ants that carry out random walks on a so-called construction graph. The random walk (and the resulting solution) depends on pheromone values that are values on the edges of the construction graph. The probability of traversing a certain edge depends on its pheromone value. One widely used construction procedure for tackling the TSP has already been analyzed in Zhou (2009). This paper carries out the first runtime analysis of ant colony optimization for the TSP problem. A tour is constructed in an *ordered* manner, where the iteratively chosen edges form a path at all times. In this paper, we give new runtime bounds for ACO algorithms using this construction procedure. In addition, we propose a new construction procedure, where, in each iteration, an *arbitrary* edge not creating a cycle or a vertex of degree three may be added to extend the partial tour. We analyze both construction methods and point out their differences.

We examine the two mentioned ACO variants in a rigorous way. Our first goal is to get insights into the local search behavior of the two approaches when the pheromone update is high. We examine the locality of changes made, i.e., how many edges of the current-best solution are also in the newly sampled tour, and how many are *exchanged* for other edges. Afterwards, we consider the expected time until certain desired local changes have been made to derive upper bounds on the optimization time for a particular TSP instance.

We show that the ordered edge insertion algorithm exchanges an expected number of $\Omega(\log(n))$ many edges while the arbitrary edge insertion exchanges only an expected constant number of edges. Arbitrary edge insertion has a probability of $\Theta(1/n^2)$ for any specific exchange of two edges, while ordered edge insertion has one of $\Theta(1/n^3)$ (Zhou 2009). Investigating the simple TSP-instance analyzed in Zhou (2009) for arbitrary edge insertions, we show an upper bound of $O(n^3 \log(n))$ on the expected number of steps to reach an optimal solution, while the best known bound for ordered edge insertion is $O(n^6)$ (Zhou 2009). Afterwards, we consider random graphs and show that both construction graphs lead in expected polynomial time to a good approximation on random instances. The paper extends its conference version published as Kötzing et al. (2010) by presenting worst-case instances for both algorithms. On these instances, the expected optimization time of both approaches is exponential. Furthermore, we investigate the use of heuristic information and point out its impact on the runtime behavior in a rigorous way.

Our theoretical results show that arbitrary edge insertion allows for better runtime bounds thanks to its locality when compared to ordered edge insertion and gives a deeper understanding on the impact of various parameter settings of ACO algorithms for the TSP problem.

The rest of the paper is organized as follows. In Sect. 2 we introduce the problem and the algorithms that we investigate. We analyze the number of edge exchanges for large pheromone updates in Sect. 3 and prove runtime bounds for certain classes of instances in Sect. 4. Section 5 shows instances where the behavior of both algorithms is exponential, and Sect. 6 investigates the use of heuristic information in our algorithms. Finally, we finish with some concluding remarks and topics for future work.

2 Problem and algorithms

We consider the symmetric traveling salesperson problem (TSP). Given is a complete undirected graph $G = (V, E)$ and a weight function $w : E \to \mathbb{R}_+$ that assigns positive weights to the edges. The TSP problem asks for a tour of minimum weight that visits every vertex exactly once and then returns to the start vertex. We investigate the behavior of ACO algorithms on this problem. Our goal is to study the computational complexity of such algorithms for the TSP. These studies provide new insights based on rigorous proofs regarding situations in which ACO algorithms are successful.

In Stützle and Hoos (2000) a search heuristic was proposed for solving this problem, based on ACO, called MMAS* (Min–Max Ant System). We reproduce a simpler version thereof in Algorithm 1. This algorithm (or variations) was already used in different theoretical studies (Neumann et al. 2009; Zhou 2009). MMAS* works as follows. The best solution that was found so far is always stored. The algorithm tries to find better solutions by iteratively generating candidate solutions that are discarded if they are at most as good as the best-so-far solution.

The construction of a new candidate solution for a target graph G works as follows. We imagine an artificial ant performing a random walk on an underlying graph, called the *construction graph*, step by step choosing components of a new candidate solution. Our components are the edges of the graph we are trying to find a minimal tour for.

Formally, for a given graph G, the construction graph $C(G)$ is defined on the nodes $\{s\} \cup E$ (where s is our distinguished start node), and an ant is allowed to travel to a node y iff y is in the *neighbor set* N of the sequence of nodes of $C(G)$ visited so far. We present two possible choices for the neighbor sets for ACO on TSP, discussed in Sects. 2.1 and 2.2. The G-edges thus visited by an ant in the random walk on $C(G)$ are the components chosen by the ant for the new tour.

In each step of its random walk on the construction graph, we want the ant to choose an edge e in G with a probability based on the *pheromone value* $\tau(e)$ and on the *heuristic value* $\eta(e)$ of that edge. Pheromone values represent the memory of the ACO algorithm about the quality of previously sampled tours and direct the search toward promising areas; we will say more about pheromones later. The heuristic information on an edge e is the inverse of the weight of e, i.e., $\eta(e) = 1/w(e)$; we use the heuristic value to bias the algorithm to favor light edges over heavy ones (this choice of η is the most frequent choice for TSP in the previous literature). Our ACO algorithm takes two parameters α and β; when choosing a new edge, we sample proportionally to the pheromone value to the power of α times the heuristic value to the power of β.

We use a procedure `construct` based on the pheromones τ as given in Algorithm 2. In this paper, we consider two different approaches of constructing new solutions by specifying the neighborhood function N of Algorithm 2 in Sects. 2.1 and 2.2.

After each iteration, the pheromones will be updated; edges that are part of the best-so-far solution are rewarded, and their pheromone is increased; all other edges will have their pheromone decreased. Formally, we use a procedure `update`. This procedure depends on the *evaporation factor* ρ ($0 \le \rho \le 1$, a parameter of the ACO algorithm); small values of ρ (close to 0) indicate low evaporation and small changes to pheromones per iterations; large values (close to 1) indicate fast changes. Max–Min Ant System derives its name from maximal and minimal pheromone values that can be attained. We use $\tau_{\max}$ and $\tau_{\min}$ as parameters to the ACO algorithm to indicate the upper and lower bound, respectively. Formally, the `update` procedure works as follows. For a tour x, let $E(x)$ be the set of edges used in x; for each edge e, the pheromone values are updated so that the new pheromone values

Algorithm 1: The algorithm MMAS*.

function *MMAS* on* $G = (V, E)$ **is**
  $\tau(e) \leftarrow 1/|V|$, for all $e \in E$;
  $x^* \leftarrow$ `construct`(τ);
  `update`(τ, x^*);
  while *true* **do**
    $x \leftarrow$ `construct`(τ);
    if $f(x) < f(x^*)$ **then**
      $x^* \leftarrow x$;
    $\tau \leftarrow$ `update`(τ, x^*);

Algorithm 2: The algorithm `construct`.

function `construct` *based on* $\tau, \eta, \alpha, \beta$ **is**
  for $k = 0$ **to** $n - 2$ **do**
    $R \leftarrow \sum_{y \in N(e_1,\ldots,e_k)} \tau(y)^\alpha \cdot \eta(y)^\beta$;
    Choose one neighbor $z \in N(e_1, \ldots, e_k)$ where the probability of selection of any fixed $z \in N(e_1, \ldots, e_k)$ is $\frac{\tau(z)^\alpha \cdot \eta(z)^\beta}{R}$;
  Let e_n be the (unique) edge completing the tour;
  return $(e_1, \ldots, e_n)$;

$\tau' = \texttt{update}(\tau, x)$ are such that

$$\tau'(e) = \begin{cases} \min\{(1-\rho)\cdot\tau(e)+\rho, \tau_{\max}\} & \text{if } e \in E(x), \\ \max\{(1-\rho)\cdot\tau(e), \tau_{\min}\} & \text{otherwise.} \end{cases}$$

The evaporation factor ρ determines the strength of an update. For several of our investigations, we assume a strong pheromone update, i.e., $\rho = 1$, such that only the pheromone values $\tau_{\min}$ and $\tau_{\max}$ are attained. Our goal is to get insights into the construction of solutions implied by two different construction graphs. We will see that this is already difficult as the probabilities of choosing certain components of a solution depend significantly on the components chosen so far during the construction run. We are optimistic that the analyses carried out in this paper will set the basis for further results regarding smaller pheromone updates.

Following Zhou (2009), we use $\tau_{\min} = 1/n^2$ and $\tau_{\max} = 1 - 1/n$ throughout this paper, where n is the number of nodes of the input graph; we initialize the values for pheromones with $1/n$. We say that the pheromones in an iteration of MMAS* are *saturated* if exactly the edges of the best-so-far tour are at $\tau_{\max}$ and all others are at $\tau_{\min}$. Our choice of $\tau_{\min}$ and $\tau_{\max}$ ensures that, for saturated pheromones, resampling the previous tour has constant probability for both constructions graphs below. Note that other choices for $\tau_{\min}$ and $\tau_{\max}$ are also possible but not considered in this paper.

To measure the runtime of MMAS*, it is common to consider the number of constructed solutions. Often we investigate the expected number of constructed solutions until an optimal tour or a good approximation of an optimal tour is obtained.

2.1 The input graph as construction graph

The construction graph is implicitly defined by introducing the neighborhood function N (see Algorithm 2). The most common way of constructing a tour for the TSP is to start in one vertex and to visit all the vertices one by one, thus growing a path containing all the vertices; finally the path is closed and becomes a tour (see, e.g., Dorigo and Gambardella 1997). This is captured by defining a neighbor set function as follows. For each sequence σ of previously chosen edges, we let $U(\sigma)$ be the set of unvisited nodes and $l(\sigma)$ be the most recently visited node (or, if σ is empty, some fixed node of the input graph); let

$$N_{\mathrm{Ord}}(\sigma) = \bigl\{\{l(\sigma), u\} \bigm| u \in U(\sigma)\bigr\}.$$

This neighborhood function is easy to compute and has linear size in the number of edges needed to complete the tour. There are drawbacks of this neighborhood set that we will discuss later. With $\mathrm{MMAS}^*_{\mathrm{Ord}}$ we denote the MMAS* that uses this neighborhood ("Ord" is mnemonic for the "ordered" way in which edges are inserted into the new tour).

2.2 An edge-based construction graph

There is no intrinsic reason for choosing the edges of the tour in an ordered way; we can let the ant choose to add any edge to the set of edges chosen so far, as long as no cycle and no vertex with degree at least three are created. This can be expressed by a neighbor set function as follows. For each sequence σ of chosen edges, let $P(\sigma)$ be the set of previously chosen edges, and let

$$\begin{aligned} N_{\mathrm{Arb}}(\sigma) = {} & \bigl(E \setminus P(\sigma)\bigr) \\ & \setminus \bigl\{e' \in E \bigm| \bigl(V, P(\sigma) \cup \{e'\}\bigr) \text{ contains a cycle or a vertex of degree} \geq 3\bigr\}. \end{aligned}$$

A drawback of this neighborhood function is that the resulting set has a size quadratic in the number of edges required to complete the tour (since all edges between the components of the partial tour constructed so far are choosable). With $\text{MMAS}^*_{\text{Arb}}$ we denote the MMAS^* that uses this neighborhood ("Arb" is mnemonic for the "arbitrary" way in which edges are inserted into the new tour).

3 Number of edge exchanges

Now we turn to properties of locality of the two algorithms. In particular, we consider the case where all pheromone values are saturated; in this case we can bound the expected number of edges by which a newly constructed solution x differs from the best-so-far solution x^*, if we do not use heuristic information. The values we use for $\tau_{\min}$ and $\tau_{\max}$ ensure that the solution x^* is reproduced with constant probability. We show that $\text{MMAS}^*_{\text{Ord}}$ and $\text{MMAS}^*_{\text{Arb}}$ differ significantly in the expected number of edges that they exchange in this situation. The expected number of edge exchanges is closely related to the local search ability of the algorithms. Local search procedures such as 2-opt are very common for the TSP problem, and our goal is to determine whether $\text{MMAS}^*_{\text{Ord}}$ and $\text{MMAS}^*_{\text{Arb}}$ have a strong local search behavior such that they are able to work without additional local search procedures.

In the following, we distinguish two algorithms with respect to their local search ability. For a tour t, we are in particular interested in tours t' such that t and t' differ by exchanging 2 or 3 edges, called a 2-*Opt* or a 3-*Opt neighbor*, respectively. For the remainder of this section, we let $\alpha = 1$ and $\beta = 0$.

3.1 The behavior of $\text{MMAS}^*_{\text{Ord}}$

We examine a particular iteration of $\text{MMAS}^*_{\text{Ord}}$ with best-so-far solution x^* and saturated pheromone values. We show that the expected number of edges where x^* and the newly constructed solution x differ is $\Omega(\log n)$. Thus, $\text{MMAS}^*_{\text{Ord}}$ does *not* have a strong local property.

The reason for this large number of exchange operations is as follows. With constant probability, the ant will first rediscover a constant fraction of the tour T corresponding to the currently best solution and then choose an edge not belonging to T. Again with constant probability, from now on, the ant will always choose edges from T, if possible. With high probability, after choosing an edge *not* from T, there are two edges from T available to be chosen next. The ant will follow one of them until no more such edges are available. A subpath of T, most likely a constant fraction of n long, remains to be traversed, starting from a random vertex. Again, the ant will have two edges to choose from and follow this subpath to the end. Thus, intuitively, a logarithmic number of new edges will be introduced in the new solution.

To turn this reasoning into a proof, we give a lemma in which we consider the following random process, capturing the situation after an ant has left the high pheromone path for the first time. Let W_t be the random variable of a walk on the sequence $(1,\dots,t)$ of t vertices. W_t starts at a random vertex and will go to the just previous or following vertex in the sequence with equal probability, if both are available and unvisited. If only one is available and unvisited, W_t will go to this one. If none are available and unvisited, the walk will *jump* uniformly at random to an unvisited vertex. The walk ends as soon as all vertices are visited.

Lemma 1 *For each t, let X_t be the random variable denoting the* number of jumps *made by the walk W_t on the sequence $(1, \ldots, t)$ of t vertices. Then we have*

$$\forall t \geq 3: \quad E(X_t) \geq \frac{1}{6}\ln(t).$$

Proof We start by giving a recursive definition of X_t. Clearly, $X_1 = 0$ and $X_2 = 0$. Let $t \geq 3$. The walk can start with uniform probability in any vertex and will never jump if the first or last vertex has been chosen. Otherwise, with equal probability, the walk will start in either of the two possible directions. After visiting all nodes in the chosen direction, the walk will jump once and then perform a walk according to X_j, where j is the number of unvisited nodes just before the jump. Thus, we get, for all $t \geq 3$,

$$\begin{aligned}
E(X_t) &= \frac{1}{t}\sum_{i=2}^{t-1}\left(\frac{1}{2}\big(1 + E(X_{i-1})\big) + \frac{1}{2}\big(1 + E(X_{t-i})\big)\right)\\
&= \frac{t-2}{t} + \frac{1}{t}\left(\frac{1}{2}\sum_{i=2}^{t-1} E(X_{i-1}) + \frac{1}{2}\sum_{i=2}^{t-1} E(X_{t-1})\right)\\
&= \frac{t-2}{t} + \frac{1}{t}\sum_{i=1}^{t-2} E(X_i) = \frac{t-2}{t} + \frac{1}{t}\sum_{i=3}^{t-2} E(X_i).
\end{aligned}$$

The claim of the lemma is true for $t = 3$. We show the remainder of the claim by induction on t. Let $t \geq 4$ and for all i, $3 \leq i < t$, $E(X_i) \geq \frac{1}{6}\ln(i)$.

We have

$$E(X_t) = \frac{t-2}{t} + \frac{1}{t}\sum_{i=3}^{t-2} E(X_i).$$

Using $t \geq 3$, we have $(t-2)/t \geq 1/3$. Thus, also using the induction hypothesis,

$$\begin{aligned}
E(X_t) &\geq \frac{1}{3} + \frac{1}{t}\sum_{i=3}^{t-2}\frac{1}{6}\ln(i)\\
&= \frac{1}{3} + \frac{1}{t}\frac{1}{6}\ln\left(\prod_{i=3}^{t-2} i\right) = \frac{1}{3} + \frac{1}{t}\frac{1}{6}\ln\big((t-2)!/2\big)\\
&\geq \frac{1}{3} + \frac{1}{t}\frac{1}{6}\big((t-2)\ln\big((t-2)/e\big) + 1 - \ln(2)\big)\\
&\geq \frac{1}{3} + \frac{1}{t}\frac{1}{6}\big((t-2)\ln\big((t-2)/e\big)\big)\\
&\geq \frac{1}{6} - \frac{t-2}{6t} + \frac{1}{6}\left(1 + \frac{t-2}{t}\ln(t-2)\right)\\
&\geq \frac{1}{6}\left(\ln(t-2) + 1 - \frac{2}{t}\ln(t-2)\right)\\
&\geq \frac{1}{6}\ln(t).
\end{aligned}$$

□

We can now use this lemma to give one of the main theorems in this section, bounding the expected number of edge exchanges which $\text{MMAS}^*_{\text{Ord}}$ will make when saturated.

Theorem 2 *If in an iteration of* $MMAS^*_{\text{Ord}}$ *the pheromone values are saturated, then, in the next iteration of* $MMAS^*_{\text{Ord}}$*, the newly constructed tour will exchange an expected number of* $\Omega(\log(n))$ *of edges.*

Proof Let E be the event that an ant leaves the tour T corresponding to the currently best solution x^* before having visited at most $n/2$ vertices and at least 2. In each of the $n/2$ first iterations $\Omega(n)$ edges *not* from T are available with pheromone $\geq 1/n^2$ each; only one (or two) edges from T with pheromone ≤ 1 is available. Thus, in each of these iterations, an edge *not* from T is chosen with probability $1/n$. We have that $P(E)$ is $\Omega(1)$, as $(1 - 1/n)^{n/2} \leq e^{-1/2}$.

Let E' be the event that, after the ant left the path for the first time, the ant will never choose an edge not from T. Using Bernoulli's inequality, we have that $P(E)$ is $\Omega(1)$. Suppose that r vertices are left to be chosen, and let X be the random variable of the number of edges not from T chosen after the first choice not from T, not counting the closing of the tour at the very end. Then we have that the random variable of X given E' equals W_r.

Now we get the desired result from Lemma 1. □

However, constructing new solutions with few exchanged edges is still somewhat likely. In Zhou (2009) it is shown that the probability for a particular 2-Opt step is $\Omega(1/n^3)$.

Theorem 3 (Zhou 2009) *Let* t *be a tour found by* $MMAS^*_{\text{Ord}}$*, and let* t' *be a* 2*-Opt neighbor of* t*. Suppose that the pheromone values are saturated. Then* $MMAS^*_{\text{Ord}}$ *constructs* t' *in the next iteration with probability* $\Omega(1/n^3)$.

Adding some simple observations, we also get a matching upper bound.

Theorem 4 *Let* t *be a tour found by* $MMAS^*_{\text{Ord}}$*. Suppose that the pheromone values are saturated. Then there is a tour* t' *which is a* 2*-Opt neighbor of* t*, and* $MMAS^*_{\text{Ord}}$ *constructs* t' *in the next iteration with probability* $\Theta(1/n^3)$.

Proof The lower bound is from Zhou (2009). Let t' be a 2-Opt neighbor of t as follows. t' follows t from the start vertex for $\lfloor n/3 \rfloor$ cities, then skips $\lfloor n/3 \rfloor$ to some city v. Then t' visits all skipped cities in reverse order before jumping to the city after v (with respect to the order in which t visits cities). Then t' follows t to the end.

We now analyze the likelihood for $\text{MMAS}^*_{\text{Ord}}$ to sampling t'. Following t in original or reversed order has a constant probability, even when accumulated over all the tour. Thus, we focus on the times when t' does not follow t. At the time of the first change, the ant does not choose the edge included in t, which has pheromone $\tau_{\max}$, but instead an edge with pheromone value $\tau_{\min}$. According to the edge selection mechanism of $\text{MMAS}^*_{\text{Ord}}$, this has a probability of $\Theta(1/n^2)$. The second time t' does not follow t, and $\text{MMAS}^*_{\text{Ord}}$ still has a constant fraction of the cities left to visit. Each of them will be the next with equal probability (as all of them have an edge with pheromone value $\tau_{\min}$ leading to them), which gives this particular choice of t' a probability of $\Theta(1/n)$. The result follows from the independence of all the choices. □

3.2 The behavior of $\text{MMAS}^*_{\text{Arb}}$

In this section we examine the expected number of edge exchanges of $\text{MMAS}^*_{\text{Arb}}$. In Theorem 7 we show that the expected number of edges where x^* and x differ is $\Theta(1)$. Thus, $\text{MMAS}^*_{\text{Arb}}$ *does* make only local changes.

We start this section with the following lemma regarding the probability of choosing edges with low pheromones.

Lemma 5 *Let $k \leq \sqrt{n}$. Consider an iteration of $MMAS^*_{\text{Arb}}$ where the pheromone values are saturated. Suppose that, in the next iteration of $MMAS^*_{\text{Arb}}$, at most k edges are chosen that are not from the best-so-far tour. Then the ant will choose edges of the best-so-far tour with probability $1 - O(1/n)$ as long as any is admissible.*

Proof We call an edge with pheromone level $\tau_{\max}$ a "high" edge; the others are "low" edges. We consider an iteration of $\text{MMAS}^*_{\text{Arb}}$. We analyze the situation where, out of the n edges to be chosen to create a new tour, there are still i edges left to be chosen, and at most k of these edges already chosen are low edges. Let a_i be the number of high edges left to be chosen in this case, and let b_i be the number of low edges still choosable. The edges chosen so far partition the graph into exactly i components. For each two components, there are between 1 and 4 edges to connect them (each component is a path with at most 2 endpoints, only the endpoints can be chosen for connecting with another component); thus, there are between $\binom{i}{2}$ and $\min\left(4\binom{i}{2}, \binom{n}{2}\right)$ edges left to be chosen. As there are i edges left to be chosen for the tour, at most k of which are low edges, and each low edge can block at most 3 high edges from being chosen (one at each endpoint, plus one which would complete a premature cycle), we get

$$i - 3k \leq a_i \leq i + k,$$
$$\binom{i}{2} - (i + k) \leq b_i \leq \min\left(4\binom{i}{2}, \binom{n}{2}\right).$$

Thus, the probability to choose a low edge is at most

$$\frac{b_i}{a_i} \cdot \frac{\tau_{\min}}{\tau_{\max}} \leq \frac{\min\left(4\binom{i}{2}, \binom{n}{2}\right)}{i - 3k} \frac{\tau_{\min}}{\tau_{\max}} \leq \frac{4i^2}{i - 3k} \cdot \frac{1}{n^2(1 - 1/n)}.$$

For $i \geq 3\sqrt{n}$, this term is $O(1/n)$. For $i \leq 3\sqrt{n}$, this is also $O(1/n)$ as long as there is at least one high edge to choose. □

Next we show that, for a fixed k, making k edge exchanges is likely.

Theorem 6 *Let $k = O(1)$. If in an iteration of $MMAS^*_{\text{Arb}}$ the pheromone values are saturated, then, in the next iteration of $MMAS^*_{\text{Arb}}$ with probability $\Theta(1)$, the newly constructed tour will choose k new edges and otherwise rechoose edges of the best-so-far tour as long as any are admissible.*

Proof We use the terminology from Lemma 5, as well as the bounds on a_i and b_i.

The ant iteratively chooses n edges for the new tour. Let M be a set of k iterations (a k-elementary subset of $\{1, \ldots, n\}$). We would like to bound the probability that the ant, in all iterations *not* belonging to M, will choose high edges as long as there are any admissible (regardless of what happens in the iterations belonging to M). Using the union bound and Lemma 5, this probability is $\Theta(1)$. Let $d > 0$ be a constant lower bound on this probability.

For each k-element subset M of $\{1, \ldots, n\}$, the probability of choosing a low edge on all positions of M, and choosing a high edge on all other positions is lower bounded by

$$\begin{aligned} d \prod_{i \in M} \left(\binom{i}{2} - (i+k) \right) \tau_{\min} \Big/ \left((i+k)\tau_{\max} + n^2 \tau_{\min} \right) \\ \geq d\tau_{\min}^k \prod_{i \in M} \left(\frac{i^2 - i}{2} - (i+k) \right) \Big/ (i+k+1) \\ \geq d\tau_{\min}^k \prod_{i \in M} \left(\frac{i^2}{2(i+k+1)} - 2 \right) \\ \geq d\tau_{\min}^k \prod_{i \in M} \left(\frac{i}{2k+4} - 2 \right). \end{aligned}$$

Let $c_{i,k} = i/(2k+4) - 2$. Note that, for any set M with $|M| \leq k$, we have $\sum_{i=1, i \notin M}^{n} c_{i,k} = \Theta(n^2)$. Now we have that the probability of choosing low edges on *any* k positions is lower bounded by

$$\begin{aligned} & d\tau_{\min}^k \sum_{\substack{M \subseteq \{1,\ldots,n\} \\ |M|=k}} \prod_{i \in M} c_{i,k} \\ & = \frac{d}{k! n^{2k}} \sum_{i_1=1}^{n} \left(\sum_{i_2=1, i_2 \notin \{i_1\}}^{n} \left(\cdots \left(\sum_{i_k=1, i_k \notin \{i_1,\ldots,i_{k-1}\}}^{n} \prod_{j=1}^{k} c_{i_j,k} \right) \right) \right) \\ & = \frac{d}{k! n^{2k}} \left(\sum_{i_1=1}^{n} c_{i_1,k} \right) \left(\sum_{i_2=1, i_2 \notin \{i_1\}}^{n} c_{i_2,k} \right) \cdots \left(\sum_{i_k=1, i_k \notin \{i_1,\ldots,i_{k-1}\}}^{n} c_{i_k,k} \right) \\ & \geq \frac{d}{k! n^{2k}} \left(\sum_{i=1}^{n} c_{i,k} \right)^k \\ & = \Theta(1). \end{aligned}$$

□

For the expected number of edge exchanges, we get the following.

Theorem 7 *If in an iteration of MMAS*$^*_{\text{Arb}}$ *the pheromone values are saturated, then, in the next iteration of MMAS*$^*_{\text{Arb}}$, *the newly constructed tour will exchange an expected number of* $O(1)$ *edges.*

Proof We use the terminology from Lemma 5, as well as the bounds on a_i and b_i.

Our goal is to bound the expected number of low edges that are introduced into a new solution, so let the random variable X_i denote the number of chosen new edges after i edges have been chosen. In the beginning of the construction, $X_0 = 0$ holds and afterwards, we may choose (with small probability) a low edge in a given step of the construction. We want to study the construction process and show that we choose only a small number of low edges during the construction process.

Let $(X_i)_{i\in\mathbb{N}}$ be the Markov-chain where X_i denotes the number of chosen low edges after i iterations. Then we have $X_0 = 0$ and, for all i, $X_{i+1} \in \{X_i, X_i + 1\}$. We are interested in $E(X_n)$ (and we want this to be $O(1)$).

We call a run *good* if the number of low edges is at most $\sqrt{n}$, i.e., $k \le \sqrt{n}$ holds. We call a run *bad* if $k > \sqrt{n}$ holds. In the following, we show that the expected number of edge exchanges of all *good* runs is $O(1)$. Later on, we show that the probability of having a *bad* run is exponentially small, which implies that this has almost no effect on the expected number of edge exchanges.

Let us consider the *good* runs, i.e., the case $k \le \sqrt{n}$. Let T be the random variable denoting the first time when no high edge is choosable. We have

$$E(X_n) = E(X_T + n - T) = E(X_T) + n - E(T),$$

as, after T iterations, all chosen edges will be low edges. The number of choosable high edges goes down by one if we choose a high edge and by at most three if we choose a low edge. In a good run, we will choose a low edge with probability $O(1/n)$ according to Lemma 5. Thus, we can apply an additive drift theorem (He and Yao 2004) to get $E(T) = n - O(1)$. Another application gives $E(X_T) = O(1)$, as $T \le n$ and we have an expected increase of $O(1/n)$ per iteration. This implies that the expected number of low edges in a *good* run is $O(1)$.

It remains to show that a *bad* run happens only with small probability. Let C_i be the condition that we have chosen at most $\sqrt{n}$ low edges after i steps, i.e., the run has not been bad until step i. Clearly, $\mathrm{Prob}(C_0) = 1$ since the number of chosen low edges is 0 at the beginning. Let us assume that C_{i-1} has happened. Then the probability of choosing a low edge in step j, $1 \le j \le i$, is at most c/n, c an appropriate constant, and the probability to choose within i steps at least $\sqrt{n}$ low edges is at most

$$\binom{i}{\sqrt{n}}(c/n)^{\sqrt{n}} = e^{-\Omega(\sqrt{n})}.$$

Hence, $\mathrm{Prob}(\neg C_i \mid C_{i-1}) = e^{-\Omega(\sqrt{n})}$, and $\mathrm{Prob}(\neg C_n) = \mathrm{Prob}(\bigcup_{i=1}^{n}(\neg C_i \mid C_{i-1})) = e^{-\Omega(\sqrt{n})}$ by the union bound. Hence, a *bad* run happens with probability most $e^{-\Omega(\sqrt{n})}$. Pessimistically, we assume that each *bad* run contributes n low edges. The contribution of the bad runs to the expected number of low edges is therefore at most

$$e^{-\Omega(\sqrt{n})} \cdot n,$$

i.e., exponentially small, which completes the proof. □

As a corollary to Theorem 6, we get the following:

Corollary 8 *Let t be a tour found by MMAS$^*_{\mathrm{Arb}}$, and let t' be a tour which is a 2-Opt neighbor of t. Suppose that the pheromone values are saturated. Then MMAS$^*_{\mathrm{Arb}}$ constructs t' in the next iteration with probability $\Theta(1/n^2)$.*

Proof The tour t has $\Theta(n^2)$ many 2-Opt neighbors. By Theorem 6, MMAS$^*_{\mathrm{Arb}}$ will construct, with constant probability, a tour that exchanges one edge and otherwise rechooses edges of t as long as possible. This new tour is a 2-Opt neighbor of t. As all 2-Opt neighbors of t are constructed equiprobably (thanks to the symmetry of the construction procedure), we obtain the desired result. □

4 Polynomial upper runtime bounds

In this section we give polynomial upper bounds for the ACO algorithms. First we analyze the optimization behavior on a simple TSP instance already studied by Zhou (2009). Then we look at random instances. For this section, we let $\alpha = 1$ and $\beta = 0$ again.

4.1 A simple instance

Zhou (2009) carried out the first runtime analysis of ACO algorithms for the TSP. He analyzed the time it takes ACO algorithms to obtain optimal solutions for some simple example instances. His analyses are based on the observation that ACO algorithms are able to imitate 2-Opt and 3-Opt operations. One particularly simple instance that is considered in Zhou (2009) consists of a single Hamiltonian cycle in which all edges have cost 1 (called light edges), while all remaining edges have a large weight of n (called heavy edges). This instance is called G_1, and it is clear that the Hamiltonian cycle of light edges is the unique optimum. The author shows that $\text{MMAS}^*_{\text{Ord}}$ obtains an optimal solution for G_1 in expected time $O(n^6 + (1/\rho)n \log n)$ for arbitrary ρ with $0 < \rho \leq 1$. The proof is based on the following idea: As long as the current solution is not optimal, there is always a 2-Opt or 3-Opt operation that leads to a better tour. Zhou shows that the probability of performing this operation is $\Omega(1/n^5)$. From this the result follows because at most n improvements are possible and $O(\log n/\rho)$ is the so-called freezing time, i.e., the time to bring all pheromone values to upper or lower bounds.

In this section, we prove that the expected optimization time of $\text{MMAS}^*_{\text{Arb}}$ for the instance G_1 is $O(n^3 \log n + (n \log n)/\rho)$. This is considerably better than the bound of $O(n^6)$ proved before in Zhou (2009) for $\text{MMAS}^*_{\text{Ord}}$. Note that we do not provide lower bounds, so it is open whether $\text{MMAS}^*_{\text{Ord}}$ has a running time as low as the one of $\text{MMAS}^*_{\text{Arb}}$. We strongly suspect that it has *not*. Furthermore, our analysis is much simpler than the one from Zhou (2009) and saves an unnecessary case analysis, which could also be avoided in the analysis of $\text{MMAS}^*_{\text{Ord}}$.

In the following lemma we consider a single improvement. Following the notation in Zhou (2009), let A_k, $k \leq n$, denote the set of all tours consisting of exactly $n - k$ light and k heavy edges. These are exactly the tours with a total weight of $n - k + kn$.

Lemma 9 *Let $\tau_{\min} = 1/n^2$ and $\tau_{\max} = 1 - 1/n$. Denote by X^t the best-so-far tour sequence produced by $MMAS^*_{\text{Arb}}$ on TSP instance G_1 until iteration $t > 0$ and assume that X_t is saturated. Then the probability of an improvement, given $1 \leq k \leq n$ heavy edges in X_t, satisfies $s_k = P(X^{t+1} \in A_{k-1} \cup \cdots \cup A_0 \mid X^t \in A_k) = \Omega(k/n^3)$.*

Proof We consider an arbitrary light edge $e = \{u, v\} \notin T$ that is not contained in the best-so-far tour T. Since each vertex of G_1 is incident to two light edges, both u and v are incident to exactly one light edge different from e. Since edge e does not belong to T, there must exist two different heavy edges $e_0, e_1 \in T$ on the tour such that e_0 is incident on u and e_1 incident on v. Now we denote by $e'_0 \in T$ and $e'_1 \in T$ with $e'_0 \neq e_0$ and $e'_1 \neq e_1$ the other two edges on T that are incident to u and v, respectively. We would like to form a new tour that contains e and still e'_0 and e'_1 but no longer e_0 and e_1. The set of edges $(T \cup \{e\}) \setminus \{e_0, e_1\}$ has cardinality $n - 1$ but might contain a cycle. If that is the case, there must be a heavy edge $e_2 \in T$ from the old tour on that cycle (since there is a unique cycle of light edges in G_1). In this case, we additionally demand that the new tour does not contain e_2. Since the undesired edges e_0, e_1 and possibly e_2 are heavy and e is a light edge outside the previous tour, any tour that is a superset of $(T \cup \{e\}) \setminus \{e_0, e_1, e_2\}$ is better than T.

For $1 \leq j \leq n/4$, we denote by $M_e(j)$ the intersection of the following three events, and we prove that $\text{Prob}(M_e(j)) = \Omega(1/n^4)$; later, we take a union over different j and e to get an improved bound.

1. The first $j-1$ steps of the construction procedure choose edges from $T^* := T \setminus \{e_0, e_1, e_2\}$, and the jth step chooses e.
2. e_0' is chosen before e_0 and e_1' before e_1.
3. All steps except the first one choose from T^* as long as this set contains applicable edges.

Note that e_0 and e_1 are no longer applicable once $\{e, e_0', e_1'\}$ is a subset of the new tour.

For the first subevent, assume that all of the first $i < j$ edges have been chosen from the set T^*. Then there are $n-i$ edges from T and $n-i-3$ edges from T^* left. Finally, there are at most $n^2/2$ edges outside of T. Using that X_t is saturated, the probability of choosing another edge from T^* can be seen to be at least

$$\frac{(n-i-3)\tau_{\max}}{(n-i)\tau_{\max} + n^2\tau_{\min}/2} \geq \frac{n-i-3}{n-i+1}$$

(assuming $n \geq 2$). Altogether, the probability of only choosing edges from T^* in the first $j-1$ steps is at least

$$\prod_{i=0}^{j-2} \frac{n-i-3}{n-i+1} \geq \left(\frac{3n/4-1}{3n/4+3}\right)^{n/4-1} = \Omega(1)$$

because $j \leq n/4$. The probability of choosing e in the jth step is at least $\tau_{\min}/n = 1/n^3$ since the total amount of pheromone in the system is at most n. Altogether, this shows that the first subevent has a probability of occurring of $\Omega(1/n^4)$.

The second subevent has a probability of occurring of at least $(1/2)^2 = 1/4$ because all applicable edges in T are chosen with the same probability (using that X_t is saturated).

For the third subevent, we study a step of the construction procedure in which there are i applicable edges from T left, and all edges chosen so far are from the set $T \cup \{e\}$. We need a more precise bound on the number of applicable edges outside of T. Removing $k \geq 1$ edges from T breaks the tour into k connected components, each of which has at most two vertices of degree less than 2. Since $e \notin T$ has been chosen, at most two edges from T are excluded from our consideration. Therefore, the number of connected components in the considered step of the construction procedure is at most $i+2$. Hence there are at most $\binom{2(i+2)}{2} \leq 2(i+2)^2 \leq 18i^2$ edges outside of T applicable. The probability of choosing neither e_2 nor an edge outside T in this situation is at least

$$\frac{i\tau_{\max}}{(i+1)\tau_{\max} + 18i^2\tau_{\min}}.$$

Hence, given the second subevent, the probability of the third subevent is at least

$$\prod_{i=1}^{n-1} \frac{i \cdot \tau_{\max}}{(i+1)\tau_{\max} + 18i^2\tau_{\min}}$$
$$= \prod_{i=1}^{n-1} \left(\frac{i}{i+1} \cdot \frac{(i+1) \cdot \tau_{\max}}{(i+1)\tau_{\max} + 18i^2\tau_{\min}}\right)$$

$$\geq \frac{1}{n}\prod_{i=1}^{n-1}\frac{i+1}{(i+1)+18(i+1)^2/(\tau_{\max}\cdot n^2)} \geq \frac{1}{n}\left(\prod_{i=1}^{n} 1+18i/(\tau_{\max}\cdot n^2)\right)^{-1}$$

$$\geq \frac{1}{n}\left(1+\frac{18}{n-1}\right)^{-n} = \Omega(1/n).$$

Altogether, this shows that the intersection $M_e(j)$ of the three subevents happens with a probability of $\Omega(1/n^4)$.

Finally, we consider the union $M_e := \bigcup_{j\leq n/4} M_e(j)$, which refers to including e in any of the first $n/4$ steps. Since the events $M_e(j)$ are disjoint for different j, we obtain $\text{Prob}(M_e) = (n/4)\cdot\Omega(1/n^4) = \Omega(1/n^3)$. Similarly, for all light edges $e \notin T$ (of which there are k), the events M_e are disjoint (as a different new edge is picked in the first step). Thus, the probability of an improvement is $\Omega(k/n^3)$ as desired. □

Theorem 10 *Let $\tau_{\min} = 1/n^2$ and $\tau_{\max} = 1 - 1/n$. Then the expected optimization time of MMAS$^*_{\text{Arb}}$ on G_1 is $O(n^3 \log n + n(\log n)/\rho)$.*

Proof Using Lemma 9 and the bound $O(\log n/\rho)$ on the freezing time, the waiting time until a best-so-far solution with k heavy edges is improved is bounded by $O((\log n)/\rho) + 1/s_k) = O((\log n)/\rho + n^3/k)$. Summing up, we obtain a total expected optimization time of $O(n(\log n)/\rho) + \sum_{k=1}^{n}(1/s_k) = O(n^3 \log n + n(\log n)/\rho)$. □

4.2 Random instances

The 2-Opt heuristic is a simple local search heuristic for the TSP, which starts with an arbitrary tour and performs 2-Opt steps until a local optimum is found. This heuristic performs well in practice both in terms of running time and approximation ratio (Johnson and McGeoch 1997). On the other hand, it has been shown to have exponential running time in the worst case (Englert et al. 2007), and there are instances with local optima whose approximation ratio is $\Omega(\log n/\log\log n)$ (Chandra et al. 1999). To explain this discrepancy between theory and practice, 2-Opt has been analyzed in a model of random instances reminiscent of smoothed analysis (Spielman and Teng 2004). In this model, n points are placed at random in the d-dimensional Euclidean space. Each point v_i $(i = 1, 2, \ldots, n)$ is chosen independently according to its own probability density $f_i\colon [0,1]^d \to [0,\phi]$ for some parameter $\phi \geq 1$. It is assumed that these densities are chosen by an adversary, and hence, by adjusting the parameter ϕ, one can interpolate between worst and average case: If $\phi = 1$, 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 can the probability mass be, and the closer is the analysis to a worst-case analysis.

It has been shown that the expected number of steps of 2-Opt is $O(n^{4+1/3}\cdot\log(n\phi)\cdot\phi^{8/3})$ and that the expected approximation ratio is $O(\sqrt[d]{\phi})$, partially explaining the success of 2-Opt in practice (Englert et al. 2007). We analyze the expected running time and approximation ratio of MMAS$^*_{\text{Arb}}$ and MMAS$^*_{\text{Ord}}$ on random instances. For this, we have to take a closer look into the results from Englert et al. (2007), which bound the expected number of 2-Opt steps until a good approximation has been achieved.

The upper bound on the expected number of steps is based on the observation that with high probability every 2-Opt step leads to a significant decrease of the tour length. Hence, the argument is not affected if between the 2-Opt steps other changes are made to the tour as long as these changes do not increase the length of the tour. We show the following theorem.

Theorem 11 *For $\rho = 1$, MMAS$^*_{\text{Arb}}$ finds in time $O(n^{6+2/3} \cdot \phi^3)$ with probability $1 - o(1)$ a solution with approximation ratio $O(\sqrt[d]{\phi})$.*

Proof As we have argued in Corollary 8, if all edges are saturated and there is an improving 2-Opt step possible, then this step is performed with probability at least $\Omega(1/n^2)$. From Englert et al. (2007) we know that from any state, the expected number of 2-Opt steps until a tour is reached that is locally optimal for 2-Opt is at most $O(n^{4+1/3} \cdot \log(n\phi) \cdot \phi^{8/3})$ even if in between other changes are made to the tour that do not increase its length. Hence, using Markov's inequality, we can conclude that MMAS$^*_{\text{Arb}}$ has reached a local optimum after $O(n^{6+2/3} \cdot \phi^3)$ steps with probability $1 - o(1)$.

From Englert et al. (2007) we also know that every locally optimal tour has an expected approximation ratio of $O(\sqrt[d]{\phi})$. Implicitly, the proof of this result also contains a tail bound showing that with probability $1-o(1)$ every local optimum achieves an approximation ratio of $O(\sqrt[d]{\phi})$. The theorem follows by combining the previous observations and taking into account that for our choice of ρ, all edges are saturated after the first iteration of MMAS$^*_{\text{Arb}}$. □

Taking into account that a specific 2-Opt operation in MMAS$^*_{\text{Ord}}$ happens with probability of $\Omega(1/n^3)$ in the next step, we get the following results.

Theorem 12 *For $\rho = 1$, MMAS$^*_{\text{Ord}}$ finds in time $O(n^{7+2/3} \cdot \phi^3)$ with probability $1 - o(1)$ a solution with approximation ratio $O(\sqrt[d]{\phi})$.*

5 Exponentially hard instances

Our previous investigations show that MMAS$^*_{\text{Ord}}$ and MMAS$^*_{\text{Arb}}$ achieve good approximations for a large class of problems. In this section we present instances where MMAS$^*_{\text{Ord}}$ and MMAS$^*_{\text{Arb}}$ have an exponential optimization time. This way we hope to understand better what kind of structures in graphs present problems for ACO for the TSP.

In the following, we assume $\tau_{\min}$ and $\tau_{\max}$ to be given bounds on the pheromones ($\in \mathbb{R}_{>0}$), and $\rho = 1$.

We consider the family $(G_n, w_n)_{n\in\mathbb{N}}$ of *metric* graphs on $\{1, \ldots, n\} \times \{0, 1\}$ such that

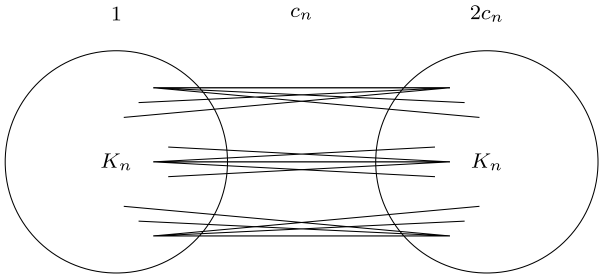
$$\forall n, \forall i, i', j, j': \quad w_n\big((i, j), (i', j')\big) = \begin{cases} 1 & \text{if } i = 0 = i', \\ 2c_n & \text{if } i = 1 = i', \\ c_n & \text{otherwise.} \end{cases} \tag{1}$$

We call the edges of weight c_n the *bridge edges*, the edges of weight 1 the *trap edges* (as they trap weight-sensitive MMAS), and the others *far edges*. One can depict the graph (G_n, w_n) as shown in Fig. 1. Within the left K_n, all edge weights are 1, within the right all are $2c_n$. All edges between the two are of weight c_n.

The graph has the following property which we will use for our analysis.

Property 13 *Let p be a tour in G_n. If p consists of k* trap edges, *k* far edges, *and $2n - 2k$* bridge edges, *the total weight of p is $w(p) = k + 2kc_n + (2n - 2k)c_n = k + 2nc_n$.*

Fig. 1 The Graph (G_n, w_n)



Note that for each $n \in \mathbb{N}$, the optimal tour of (G_n, w_n) only consists of *bridge edges* and the total cost of an optimal tour is $\mathrm{OPT} = 2nc_n$.

In our calculations, we will make use of the following inequalities. By Bernoulli's inequality, we have

$$\forall n \in \mathbb{N}, \forall x \in [0,1]: \quad (1-x)^n \geq 1 - nx. \tag{2}$$

In particular,

$$\forall n \in \mathbb{N}, \forall x \in [0,n]: \quad \left(1 - \frac{x}{n}\right)^n \geq 1 - x. \tag{3}$$

For all $n \in \mathbb{N}$, we set

$$c_n = 2^{2n+2} n^3 \frac{\tau_{\max}}{\tau_{\min}}. \tag{4}$$

Theorem 14 *With probability* $1 - 2^{-\Omega(n)}$, *the optimization time of* $\mathrm{MMAS}^*_{\mathrm{Ord}}$ *and* $\mathrm{MMAS}^*_{\mathrm{Arb}}$ *on* G_n *using* $\alpha = 1$ *and* $\beta = 1$ *is at least* 2^n.

Proof Let $\mathcal{E}$ be the event that $\mathrm{MMAS}^*_{\mathrm{Ord}}$ or $\mathrm{MMAS}^*_{\mathrm{Arb}}$ with $\alpha = \beta = 1$ on G_n in the first 2^n iterations always chooses an edge of weight $\leq n$ over and edge of weight $\geq c_n$, when possible. We show that

$$P(\mathcal{E}) \geq 1 - 2^{-n}.$$

Let p be the probability to choose in a single choice one out of $2n-2$ weight-c_n edges over a single edge of weight $\leq n$. From the construction rule of the algorithm we get

$$\begin{aligned} p &\leq \frac{\sum_{j=1}^{2n-2} \tau_{\max}/c_n}{\tau_{\min}/n + \sum_{j=1}^{2n-2} \tau_{\min}/c_n} \\ &\leq \frac{2n\tau_{\max}/c_n}{\tau_{\min}/n} \\ &\leq \frac{2n^2}{c_n} \cdot \frac{\tau_{\max}}{\tau_{\min}} \\ &= \frac{1}{2^{2n+1}n}. \end{aligned} \tag{5}$$

In the worst case, the ant is presented with such a choice $2n$ times in each of the 2^n iterations. Thus,

$$P(\mathcal{E}) \geq \prod_{i=1}^{2^n \cdot 2n} (1-p) \tag{6}$$

$$= (1-p)^{2^{n+1} \cdot n} \tag{7}$$

$$\underset{(2)}{\geq} 1 - \left(2^{n+1} \cdot n\right) p \tag{8}$$

$$\underset{(5)}{\geq} 1 - 2^{-n}. \tag{9}$$

□

As a corollary to the proof of Theorem 14, we have the following:

Corollary 15 *Let $f : \mathbb{N} \to \mathbb{N}$. There is a family $(G_n)_{n \in \mathbb{N}}$ of graphs such that, with probability $1 - 2^{-\Omega(n)}$, after 2^n steps of MMAS$^*_{\text{Ord}}$ and MMAS$^*_{\text{Arb}}$, the currently best solution has value at least $f(n) \cdot \text{OPT}$.*

Proof This can be proven by modifying (G_n, w_n) such that all far edges have weight $(f(n) + 1) \cdot (2nc_n)$. The optimal path remains the same (with the same cost), while the tour that starts with $n - 1$ trap edges has a weight of $f(n) \cdot \text{OPT}$. □

In contrast to the previous negative results, we show that $\alpha = 1$ and $\beta = 0$ lead to a polynomial expected optimization time. Together with Theorem 14, this shows that the use of heuristic information is destructive for the optimization of G_n.

Theorem 16 *With probability $1 - 2^{-\Omega(n)}$, the optimization times of MMAS$^*_{\text{Ord}}$ and MMAS$^*_{\text{Arb}}$ on G_n using $\alpha = 1$ and $\beta = 0$ are $O(n^4)$ and $O(n^3)$, respectively.*

Proof Let $n \in \mathbb{N}$. We give a *fitness based partition* of the possible tours through (G_n, w_n). We let, for all k with $0 \leq k < n$,

$$A_k = \left\{ p \text{ tour} \mid w(p) = k + 2c_n \right\}.$$

According to Property 13, $\{A_k \mid 0 \leq k < n\}$ is a partitioning of all possible tours of G_n. Note that the worst path cost is given by first traversing only trap edges, for a total cost of $2nc_n + n - 1$. This path traverses exactly two bridge edges.

Let $k \in \mathbb{N}$ with $0 < k < n$, and let a tour in A_k be given. Pick any trap edge $\{u_1, u_2\}$ and any far edge $\{v_1, v_2\}$ in the tour. The probability that MMAS will, as the only change, change the edges $\{u_1, u_2\}$ and $\{v_1, v_2\}$ for $\{u_1, v_1\}$ and $\{v_2, u_2\}$ (2-change) is lower bounded by either $1/n^3$ (Theorem 3) or $1/n^2$ (Corollary 8), and will yield a tour from A_{k-1}. Thus, we get the claimed runtime bounds. □

6 Benefit of heuristic information

In the previous section, we have shown how the use of heuristic information can harm the optimization of ACO algorithms. We now describe an example where a sufficiently large

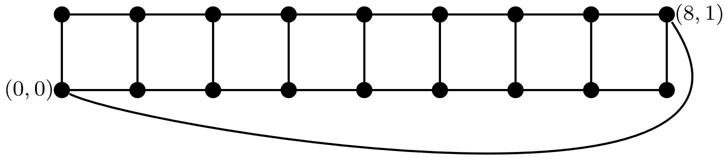


Fig. 2 The graph $G_\square$

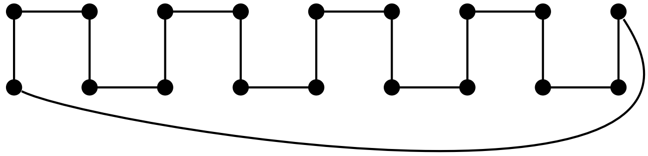


Fig. 3 The tour T_{zigzag}

value of β is essential for efficient optimization. Let k be even. Take a graph on $n = 2k + 2$ nodes consisting of $k + 1$ columns and two rows, hence k adjacent cells. More precisely, the vertex set is $V = \{0, \dots, k\} \times \{0, 1\}$, and the edge set is

$$\underbrace{\big\{\{(i,0),(i,1)\}\,\big|\,0 \le i \le k\big\} \cup \big\{(0,0),(k,1)\big\}}_{=:E_|}$$

$$\cup \underbrace{\big\{\{(i,0),(i+1,0)\}\,\big|\,0 \le i \le k-1\big\} \cup \big\{\{(i,1),(i+1,1)\}\,\big|\,0 \le i \le k-1\big\}}_{=:E_-}.$$

The edges in E_- are called *horizontal* and receive weight 1, and the edges in $E_|$ are called *vertical* and receive weight 2. Finally, we make the graph complete by giving all edges that are not contained in $\{E_- \cup E_|\}$ the large weight $c_n := 2^n$; we call these edges heavy and the other ones light. Let the resulting instance be called $G_\square$ (see Fig. 2 for an illustration). Obviously, the shortest tour T_{opt} in $G_\square$ travels along the border of the structure by taking all edges in E_- plus $\{(0,0),(0,1)\}$ and $\{(k,0),(k,1)\}$. There is only one more tour that does not use heavy edges, namely the following tour called T_{zigzag}: It contains all edges from $E_|$ and the edges $\{(i,0),(i+1,0)\}$ for odd i as well as $\{(i,1),(i+1,1)\}$ for even i. Finally, the tour is closed by the edge $\{(0,0),(k,1)\}$. See Fig. 3 for an illustration.

Lemma 17 *All tours on $G_\square$ except for T_{opt} and T_{zigzag} use at least one heavy edge.*

Proof We consider a tour not using heavy edges and show that is must equal either T_{opt} or T_{zigzag}. Each node on the tour must be connected by two light edges. Since node $(0,1)$ is incident on only two light edges, these two edges $\{(0,0),(0,1)\}$ and $\{(0,1),(1,1)\}$ must be used. The other edge to connect $\{0,0\}$ must be either $\{(0,0),(1,0)\}$ or $\{(0,0),(k,1)\}$. In the first case, the vertical edge $\{(1,1),(1,1)\}$ is unavailable since it would close a cycle, hence the horizontal edges $\{(1,0),(2,0)\}$ and $\{(1,1),(2,1)\}$ must be taken in order to include the nodes in column 2. Iterating this argument, we obtain T_{opt} as the only remaining tour without heavy edges.

In the second case, only the two light edges $\{(1,0),(1,1)\}$ and $\{(1,0),(2,0)\}$ are available in order to connect $(1,0)$. Then only two light edges are left in order to connect $(2,1)$, and iteratively we obtain T_{zigzag} as the only remaining tour without heavy edges. $\square$

Using the previous lemma, we prove that the optimization time of our algorithms is exponential with probability exponentially close to 1.

Theorem 18 *Consider MMAS*$^*_{\mathrm{Ord}}$ *with* $\alpha = 1$ *and starting vertex* $(0, 0)$. *If* $\beta = 1$, *it needs an expected number* $n^{\Omega(n)}$ *of iterations to find the optimum tour for* $G_\square$, *while with* $\beta = n$ *it finds the optimum in* 1 *iteration with probability* $1 - 2^{-\Omega(n)}$.

Proof Let $\beta = 1$. We claim that the probability of creating T_{zigzag} as initial solution is $2^{-O(n)}$. Assume that the ant has already followed the first j, $0 \le j \le n-1$, edges of T_{zigzag}. Then there are either one or two admissible light edges in the current neighborhood. Since $\tau_{\min} = 1/n^2$ and the weight of a light edge is at most 2, the ant will continue on a heavy edge only with probability at most $|E| \cdot \frac{1/c_n}{\tau_{\min}/2} = 2^{-\Omega(n)}$. Since the weight of light edges is within a ratio of 2, the ant continues on $(j+1)$st edge of T_{zigzag} with probability at least $1/3 - 2^{-\Omega(n)}$; if $j = n-1$, the probability is even 1. Altogether, the claim follows.

In the following, we assume that T_{zigzag} is the best-so-far tour and claim that the probability of constructing a better tour is $n^{-\Omega(n)}$. Since the probability of having T_{zigzag} as best-so-far tour is at least $2^{-O(n)}$, this will imply an expected number of $2^{-O(n)} \cdot n^{\Omega(n)} = n^{\Omega(n)}$ iterations.

To show the claim, we use Lemma 17, which implies that the only tour improving T_{zigzag} is T_{opt}. We consider for $1 \le j \le k-1$ the event that the ant has already created a path consisting of j edges without a vertical edge between inner columns. Then the ant is either at vertex $(0, j)$ or $(1, j-1)$. If j is even, the ant has just taken an edge from E_-, and the only permissible edge in the neighborhood with pheromone value $\tau_{\max}$ is from $E_|$. The probability of not taking that edge is at most $\frac{(2+n^2/c_n)\cdot\tau_{\min}}{\tau_{\max}/2} \le 3/n$, where we have used that there are at most 2 light edges incident on any node. There are at least $\lfloor (k-1)/2 \rfloor = \Omega(n)$ occasions where j is even. The probability of never taking the $E_|$-edge in any of these occasions is $n^{-\Omega(n)}$.

Finally, we study the case $\beta = n$. Then the probability of taking a permissible weight-2 edge in the neighborhood is by a factor of 2^{-n} less likely than taking a weight-1 edge. The ratio is even smaller with respect to weight-2 and weight-c_n edges. Hence, if already j edges of T_{opt} have been included, the next edge taken is not from T_{opt} with probability at most $n^2 \cdot 2^{-n}$. The probability of making such a mistake at least once in n trials is at most $n \cdot n^2 \cdot 2^{-n} = 2^{-\Omega(n)}$. □

7 Conclusions

Our theoretical results show that ACO with the usual construction procedure for the TSP (with our ranges of parameters) samples solutions that are in expectation far away from the currently best one in terms of edge exchanges, even if the pheromone values have touched their corresponding bounds. We have examined a new construction graph, which we showed to have stronger locality. This provably stronger locality has led to better provable runtime bounds; in particular, we showed both algorithms to perform well on random instances if the pheromone update is high. Finally, we gave example graphs on which both variants of ACO will require exponential optimization time with various parameter settings and investigated the impact of heuristic information.

It remains open whether there are TSP instances for which the nonlocality of ordered edge insertion provably gives better runtime bounds than the more local arbitrary edge insertion. Furthermore, it remains to be seen whether the more local construction procedure can outperform the classical procedure in experimental benchmarks. Practical implementations of ACO on TSP usually perform additional local searches interleaved with the searches of the ACO, which might make up for the lower locality of the classical construction procedure.

Future work should consider the behavior of the algorithms for small pheromone updates. It is not clear how to transfer the proofs for the random instances given in this paper to the ACO algorithms with small choices of ρ, as there may be many small improvements preventing us to prove freezing of the pheromones (in polynomial time). Furthermore, it is desirable to investigate the new construction procedure by experimental studies. Initial experiments show an advantage of the new procedure, but we need to carry out further investigations to confirm such a claim.

Acknowledgements Timo Kötzing, Frank Neumann, and Carsten Witt were supported by the Deutsche Forschungsgemeinschaft (DFG) under grants NE 1182/5-1 and WI 3552/1-1, respectively. Heiko Röglin was supported by a Veni grant from the Netherlands Organisation for Scientific Research (NWO).

References

Chandra, B., Karloff, H. J., & Tovey, C. A. (1999). New results on the old k-Opt algorithm for the traveling salesman problem. *SIAM Journal on Computing*, *28*(6), 1998–2029.

Dorigo, M., & Gambardella, L. M. (1997). Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, *1*(1), 53–66.

Dorigo, M., & Stützle, T. (2004). *Ant colony optimization*. Cambridge: MIT Press.

Eiben, A., & Smith, J. (2007). *Introduction to evolutionary computing* (2nd ed.). Berlin: Springer.

Englert, M., Röglin, H., & Vöcking, B. (2007). Worst case and probabilistic analysis of the 2-opt algorithm for the TSP: extended abstract. In N. Bansal, K. Pruhs, & C. Stein (Eds.), *SODA'07: Proceedings of the eighteenth annual ACM–SIAM symposium on discrete algorithms* (pp. 1295–1304). Philadelphia: Society for Industrial and Applied Mathematics.

Gutjahr, W. J. (2007). Mathematical runtime analysis of ACO algorithms: survey on an emerging issue. *Swarm Intelligence*, *1*(1), 59–79.

Gutjahr, W. J., & Sebastiani, G. (2008). Runtime analysis of ant colony optimization with best-so-far reinforcement. *Methodology and Computing in Applied Probability*, *10*(3), 409–433.

He, J., & Yao, X. (2004). A study of drift analysis for estimating computation time of evolutionary algorithms. *Natural Computing*, *3*(1), 21–35.

Horoba, C., & Sudholt, D. (2009). Running time analysis of ACO systems for shortest path problems. In T. Stützle, M. Birattari, & H. H. Hoos (Eds.), *Lecture notes in computer science: Vol. 5752. Engineering stochastic local search algorithms. Designing, implementing and analyzing effective heuristics, second international workshop, SLS*, 2009 (pp. 76–91). Berlin: Springer.

Johnson, D. S., & McGeoch, L. A. (1997). The traveling salesman problem: a case study in local optimization. In E. H. L. Aarts & J. K. Lenstra (Eds.), *Local search in combinatorial optimization*. Somerset: Wiley.

Kötzing, T., Neumann, F., Röglin, H., & Witt, C. (2010). Theoretical properties of two ACO approaches for the traveling salesman problem. In M. Dorigo, M. Birattari, G. A. D. Caro, R. Doursat, A. P. Engelbrecht, D. Floreano, L. M. Gambardella, R. Groß, E. Sahin, H. Sayama, & T. Stützle (Eds.), *Lecture notes in computer science: Vol. 6234. Swarm intelligence, 7th international conference, ANTS*, 2010 (pp. 324–335). Berlin: Springer.

Neumann, F., Sudholt, D., & Witt, C. (2008). Rigorous analyses for the combination of ant colony optimization and local search. In M. Dorigo, M. Birattari, C. Blum, M. Clerc, T. Stützle, & A. F. T. Winfield (Eds.), *Lecture notes in computer science: Vol. 5217. Ant colony optimization and swarm intelligence, 6th international conference, ANTS*, 2008 (pp. 132–143). Berlin: Springer.

Neumann, F., Sudholt, D., & Witt, C. (2009). Analysis of different MMAS ACO algorithms on unimodal functions and plateaus. *Swarm Intelligence*, *3*(1), 35–68.

Neumann, F., & Witt, C. (2009). Runtime analysis of a simple ant colony optimization algorithm. *Algorithmica*, *54*(2), 243–255.

Neumann, F., & Witt, C. (2010). Ant colony optimization and the minimum spanning tree problem. *Theoretical Computer Science*, *411*(25), 2406–2413.

Neumann, F., & Witt, C. (2010). *Bioinspired computation in combinatorial optimization—algorithms and their computational complexity*. Berlin: Springer.

Spielman, D. A., & Teng, S.-H. (2004). Smoothed analysis of algorithms: why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, *51*(3), 385–463.

Stützle, T., & Hoos, H. H. (2000). $\mathcal{MAX}$–$\mathcal{MIN}$ ant system. *Future Generations Computer Systems*, *16*(8), 889–914.

Zhou, Y. (2009). Runtime analysis of an ant colony optimization algorithm for TSP instances. *IEEE Transactions on Evolutionary Computation*, *13*(5), 1083–1092.