

A Converging ACO Algorithm for Stochastic Combinatorial Optimization

Walter J. Gutjahr

Dept. of Statistics and Decision Support Systems, University of Vienna
walter.gutjahr@univie.ac.at,
<http://mailbox.univie.ac.at/walter.gutjahr/>

Abstract. The paper presents a general-purpose algorithm for solving stochastic combinatorial optimization problems with the expected value of a random variable as objective and deterministic constraints. The algorithm follows the Ant Colony Optimization (ACO) approach and uses Monte-Carlo sampling for estimating the objective. It is shown that on rather mild conditions, including that of linear increment of the sample size, the algorithm converges with probability one to the globally optimal solution of the stochastic combinatorial optimization problem. Contrary to most convergence results for metaheuristics in the deterministic case, the algorithm can usually be recommended for practical application in an unchanged form, i.e., with the “theoretical” parameter schedule.

Keywords. Ant colony optimization, combinatorial optimization, convergence results, metaheuristics, Monte-Carlo simulation, stochastic optimization.

1 Introduction

In many practical applications of combinatorial optimization, a smaller or larger extent of uncertainty on the outcome, given a special choice of the decision maker, must be taken account of. A well-established way to represent uncertainty is by using a stochastic model. If this approach is chosen, the objective function of the optimization problem under consideration gets dependent not only on the decision, but on a random influence as well; in other word, it gets a random variable. The aim is then to optimize a specific functional of this random variable. Usually, this functional is the *expected value*; e.g., if the objective function represents cost, then the quantity to be minimized can be the expected cost. (Particular applications of risk theory, especially in financial engineering, also consider other functionals, e.g. the variance of the objective. We do not deal with this situation here.)

In some formally relative simple stochastic models, the expected value of the objective function can either be represented explicitly as a mathematical expression, or at least be easily computed numerically to any desired degree of accuracy. Then, the solution of the stochastic optimization problem is not essentially different from that of a deterministic optimization problem; the stochastic

structure is hidden in the representation of the expected objective function, and exact or heuristic techniques of combinatorial optimization can be used. The situation changes if it is only possible to determine *estimates* of the expected objective function by means of sampling or simulation. For example, consider the *single machine total tardiness problem*, an NP-hard sequencing problem (see Du and Leung [8]): n jobs have to be scheduled on a single machine, each job has a processing time and a due date, and the objective is to minimize the sum of the tardiness values, where tardiness is defined as the positive part of the difference between completion time and due date. Although the formula for the objective in the deterministic case is simple, no closed-form expression for the expected total tardiness in the case where the processing times are random variables (following a given joint distribution) is known, and its numerical computation would be very complicated and time-consuming. However, a relatively straightforward approach for approximating the expected total tardiness is to draw a sample of random scenarios and to take the average total tardiness over these scenarios as an estimate. Examples for other problems where the same approach seems promising are stochastic vehicle routing problems (see, e.g., Bertsimas and Simchi-Levi [3]), emergency planning based on simulation (Bakuli and Smith [2]), facility location problems involving queuing models (Marianov and Serra [18]), project financing with uncertain costs and incomes (Norkin, Ermoliev and Ruszczyński [19]), manpower planning under uncertainty (Futschik and Pflug [9]), or activity crashing using PERT (Gutjahr, Strauss and Wagner [16]).

For the approximate solution of hard deterministic combinatorial optimization problems, several *metaheuristics* have been developed. One of these metaheuristics with a currently rapidly growing number of applications is *Ant Colony Optimization* (ACO), rooted in work by Dorigo, Maniezzo and Colnari [7] at the beginning of the nineties and formulated more recently as a metaheuristic by Dorigo and DiCaro [6]. Like some other metaheuristics, ACO derives its basic idea from a biological analogy; in the case of ACO, the food-searching behavior of biological ant colonies is considered as an optimization process, and from this metaphor, strategies for solving a given combinatorial optimization problem by simulating walks of “artificial ants” are derived. It has been shown that certain ACO variants have the favorable property that the intermediate solutions found by the system converge to the globally optimal solution of the problem (see [12]–[14]).

The aim of the present investigation is to develop an ACO algorithm that is able to treat the more general case of a *stochastic* combinatorial optimization problem, using the generally applicable sampling approach described above. As in the deterministic case, guarantees on the convergence to the optimal solution are highly desirable. It turns out that such a convergence result is indeed possible for the algorithm presented here. It will be argued that, contrary to most convergence results for metaheuristics for deterministic problems, our algorithm can be recommended for practical use in an unchanged form, i.e., with the same parameter schedule as assumed for obtaining the convergence result.

Whereas for other metaheuristic approaches, extensions to stochastic problems have already been studied intensely (see, e.g., Arnold [1] for Evolutionary Strategies or Gutjahr and Pflug [15] for Simulated Annealing), the research on ACO algorithms for stochastic problems is currently only at the very beginning. An interesting first paper has been published by Bianchi, Gambardella and Dorigo [4], it concerns the solution of the probabilistic travelling salesman problem. Nevertheless, the approach chosen in [4] is tailored to the specific problem under consideration, and it depends on the availability of a closed-form expression of the expected objective function value. The algorithm presented here has a considerably broader range of application.

We think that ACO is especially promising for problems of the considered type for three reasons: First, it works with a “memory” (the pheromone trails, see below) which effects a certain robustness against noise; this is a common feature with Evolutionary Strategies and Genetic Algorithms, but different from Simulated Annealing or Tabu Search. Secondly, also problems with a highly constrained solution space (e.g., permutation problems) can be encoded in a natural way. Third, problem-specific heuristics can be incorporated to improve the performance. The two last issues seem to give the ACO approach a specific advantage in the field of highly constrained combinatorial optimization.

2 Stochastic Combinatorial Optimization Problems

We deal with stochastic combinatorial optimization problems of the following general form:

$$\begin{aligned} \text{Minimize } F(x) &= \mathbb{E}(f(x, \omega)) \\ \text{s.t. } x &\in S. \end{aligned} \quad (1)$$

Therein, x is the decision variable, f is the (deterministic) objective function, ω denotes the influence of randomness (formally: $\omega \in \Omega$, where (Ω, Σ, P) is the probability space specifying the chosen stochastic model), \mathbb{E} denotes the mathematical expectation, and S is a finite set of feasible decisions.

We need not to assume that $\mathbb{E}(f(x, \omega))$ is numerically computable, since it can be estimated by sampling: Draw N random *scenarios* $\omega_1, \dots, \omega_N$ independently from each other. A *sample estimate* is given by

$$\tilde{F}(x) = \frac{1}{N} \sum_{\nu=1}^N f(x, \omega_\nu) \approx \mathbb{E}(f(x, \omega)). \quad (2)$$

Obviously, $\tilde{F}(x)$ is an unbiased estimator for $F(x)$. For example, in the single machine total tardiness problem mentioned in Section 1, N arrays, each consisting of n random processing times for job 1 to n according to the given distribution(s), can be generated from independent random numbers. For each of these arrays, the total tardiness of the considered schedule (permutation) x can be computed. The average over the N total tardiness values is the sample estimate $\tilde{F}(x)$ of $F(x)$.

Let us emphasize that, contrary to its deterministic counterpart, problem (1) can be nontrivial already for a *very* small number $|S|$ of feasible solutions: Even for $|S| = 2$, we obtain, except if $F(x)$ can be computed directly, a nontrivial statistical hypothesis testing problem (see [19]).

3 Ant Colony Optimization

For the sake of a clearer understanding of the algorithm given in the next section, we recapitulate the main ideas of ACO by presenting one particular ACO algorithm, GBAS (see [12]), designed for *deterministic* problems. GBAS has been chosen since it is also the kernel of the algorithm S-ACO in Section 4.

Essential general features of ACO are the following:

- Solutions are constructed randomly and step-by-step.
- Construction steps that have turned out as part of good solutions are favored.
- Construction steps that can be expected to be part of good solutions are favored.

In GBAS (Graph-Based Ant System), the given problem instance is encoded by a *construction graph* \mathcal{C} , a directed graph with a distinguished start node. For sequencing problems as the TSP or the single-machine total tardiness problem mentioned above, the construction graph is essentially a complete graph with the items to be scheduled as nodes. For problems with other combinatorial structures (e.g., subset problems), suitable other graphs are used.

The stepwise construction of a solution is represented by a random walk in the construction graph. In this walk, each node is visited at most once, already visited nodes are “tabu” (infeasible). There may also be additional rules defining particular nodes as infeasible after a certain partial walk has been traversed. When there is no feasible unvisited successor node anymore, the walk stops and is decoded as a complete solution for the problem.

The encoding must be such that each walk that is feasible in the sense above corresponds to exactly one feasible solution. (Usually, also the reverse holds, but we do not make this to an explicit condition.) Since, if the indicated condition is satisfied, the objective function value is uniquely determined by a feasible walk, we may denote a walk by the same symbol x as a decision or solution and consider S as the set of feasible walks.

When constructing a walk in the algorithm, the probability p_{kl} to go from a node k to a feasible successor node l is chosen as proportional to $\tau_{kl} \cdot \eta_{kl}(u)$, where τ_{kl} is the so-called *pheromone* or *trail level*, a memory value storing how good step (k, l) has been in previous runs, and $\eta_{kl}(u)$ is the so-called *attractiveness* or *visibility*, a pre-evaluation of how good step (k, l) will presumably be (e.g., the reciprocal of the distance from k to l in a TSP). $\eta_{kl}(u)$ is allowed to depend on the given partial walk u up to now. This pre-evaluation is done in a problem-specific manner. Pheromone initialization and update is performed as follows:

Pheromone initialization: Set $\tau_{kl} = 1/m$, where m is the number of arcs of the construction graph.

Pheromone update: First, set, for each arc, $\tau_{kl} = (1 - \rho)\tau_{kl}$, where ρ is a so-called *evaporation factor* between 0 and 1. This step is called *evaporation*. Then, on each arc of the best walk \hat{x} found up to now, increase τ_{kl} by $\rho / L(\hat{x})$, where $L(x)$ denotes the length of walk x , defined as the number of arcs on x . Thus, the overall amount of pheromone remains equal to unity. This step reinforces the arcs (partial construction steps) of already found good solutions.

Random walk construction and pheromone update are iterated. Instead of a single walk (“one ant”), s walks ($s > 1$) are usually constructed sequentially or, in parallel implementations, simultaneously (“ s ants”).

Note that for being able to do the pheromone update as described above, the best found walk up to now has to be stored in a global variable \hat{x} . Each time a new random walk x is completed, the objective function value of the corresponding feasible solution is computed and compared with the objective function value of \hat{x} . If x turns out to be better than \hat{w} , the walk stored in \hat{x} is replaced by x .

4 Extension of the algorithm to stochastic problems

We present now an extension S-ACO of the algorithm GBAS of the last section to the case of the stochastic optimization problem (1). S-ACO leaves the basic procedure largely unchanged, but modifies the pheromone-update subprocedure by introducing a stochastic test whether the solution stored as the current best one should still be considered as optimal.

In the pseudo-code formulation below, we write $\tau_{kl}(n)$ instead of τ_{kl} in order to denote the dependence on round number n ; the same for $p_{kl}(n)$. For $\tau_{min}(n)$, see the comments after the procedure.

Feasibility of a continuation (k, l) of a partial walk u ending with node k is defined as in Section 3: The continuation (k, l) is feasible if node l is not yet contained in u , and none of the (eventual) additional rules specifies l as infeasible after u has been traversed.

```

procedure S-ACO {
  do pheromone-initialization;
  for round  $n = 1, 2, \dots$  {
    for ant  $\sigma = 1, \dots, s$  {
      set  $k$ , the current position of the ant, equal to the start node of  $\mathcal{C}$ ;
      set  $u$ , the current walk of the ant, equal to the empty list;
      while (a feasible continuation  $(k, l)$  of the walk  $u$  of the ant exists) {
        select successor node  $l$  with probability  $p_{kl}(n)$ , where

          
$$p_{kl}(n) = \begin{cases} 0, & \text{if } (k, l) \text{ is infeasible,} \\ \tau_{kl}(n) \eta_{kl}(u) / \left( \sum_{(k,r)} \tau_{kr}(n) \eta_{kr}(u) \right), & \text{otherwise,} \end{cases}$$


          the sum being over all feasible  $(k, r)$ ;
        set  $k = l$ , and append  $l$  to  $u$ ;
      }
    }
  }

```

```

    set  $x_\sigma = u$ ;
  }
  from  $\{x_1, \dots, x_s\}$ , select a walk  $x$ ;
  do S-pheromone-update( $x, n$ ); // see below
}
}

```

procedure S-pheromone-update(x, n) {
 compute estimate

$$\tilde{F}(n) = \frac{1}{N_n} \sum_{\nu=1}^{N_n} f(x, \omega_\nu)$$

by applying N_n independent sample scenarios ω_ν to x ;

```

if ( $n = 1$ ) {
  set  $\hat{x} = x$ ;
  set  $\hat{F}(n) = \tilde{F}(n)$ ;
}
else {
  compute estimate

```

$$\hat{F}(n) = \frac{1}{N_n} \sum_{\nu=1}^{N_n} f(\hat{x}, \omega'_\nu)$$

by applying N_n independent sample scenarios ω'_ν to \hat{x} ;

```

if ( $\hat{F}(n) < \tilde{F}(n)$ )
  set  $\hat{x} = x$ ;
}

```

set, with $L(x)$ denoting the length of walk x ,

$$\tau_{kl}(n+1) := \begin{cases} \max((1-\rho)\tau_{kl}(n) + \rho/L(\hat{x}), \tau_{min}(n)), & \text{if } (k, l) \in \hat{x}, \\ \max((1-\rho)\tau_{kl}(n), \tau_{min}(n)), & \text{otherwise;} \end{cases} \quad (3)$$

```

}

```

Comments:

The essential difference to the deterministic case is that in the stochastic case, it is not possible anymore to decide with certainty whether a current solution x is better than the solution currently considered as the best found, \hat{x} , or not. This can only be tested by statistical sampling, which happens in the specific pheromone update subprocedure used here, S-pheromone-update. Even the result of this test can be erroneous, due to the stochastic nature of all objective function evaluations, i.e., the test yields the correct comparison result only with a certain probability.

For the same reason, it is not even possible to decide which ant has, in the current round, produced the best walk. The procedure above prescribes that one

of the s produced walks is selected, according to whatever rule. A promising way to do that would be to evaluate each x_σ at a random scenario drawn specifically for this round and to take x as the walk with best objective value.

The first part of the subprocedure S-pheromone-update compares the solution x selected in the present round with the solution considered currently as the best, \hat{x} . This is done by determining sample estimates for both solutions (practically speaking: by estimating the expected costs of both solutions by means of Monte-Carlo simulation with N_n runs each). Scenarios ω'_ν have to be drawn independently from scenarios ω_ν (i.e., the simulation runs have to be executed with two independent series of random numbers). The winner of the comparison is stored as the new \hat{x} . The question which sample size N_n should be chosen in round n will be dealt with in the next section.

In the second part of the subprocedure, pheromone update is performed essentially in the same way as described in Section 3, but with an additional feature: If the computed pheromone value $\tau_{kl}(n)$ would fall below some predefined lower bound $\tau_{min}(n)$, we set $\tau_{kl}(n) = \tau_{min}(n)$. (The idea of using lower pheromone bounds in ACO is due to Stützle and Hoos [20], [21]). Again, the question how $\tau_{min}(n)$ should be chosen in dependence of n will be treated in the following section.

The computation of the attractiveness values $\eta_{kl}(u)$ needs some explanation. As mentioned, these values are obtained from a suitable problem-specific heuristic. Although, in principle, one could work with “zero-information” attractiveness values, all set equal to a constant, the choice of a good attractiveness heuristic will improve the performance of the algorithm considerably. In the stochastic case, there is the difficulty that certain variables possibly used by such a heuristic are not known with certainty, because they depend on the random influence ω . This difficulty can be solved either by taking the *expected values* (with respect to the distribution of ω) of the required variables as the base of the attractiveness computation (in most stochastic models, these expected values are directly given as model parameters), or by taking those variable values that result from a random scenario ω drawn for the current round. Presumably, both will perform much better than applying zero-information attractiveness.

5 Convergence

For the validity of the algorithm S-ACO presented in Section 4, we are able to give a strong theoretical justification: It is possible to prove that, under rather mild conditions, the current solutions produced by the algorithm converge with probability one to the globally optimal solution. In the sequel, we first present and then discuss these conditions.

- (i) The optimal walk x^* is unique.
- (ii) The function $f(x, \omega_\nu)$ observed at random scenario ω_ν can be decomposed in expected value and error term as follows:

$$f(x, \omega_\nu) = f(x) + \epsilon_{x\nu},$$

where $f(x) = \mathbb{E}(f(x, \omega))$, $\epsilon_{x\nu}$ is normally distributed with mean 0 and variance $(\sigma(x))^2$, and all error variables $\epsilon_{x\nu}$ are stochastically independent.

(iii) The attractiveness values $\eta_{kl}(u)$ satisfy

$$\eta_{kl}(u) > 0 \quad \text{for all prefixes } u \text{ of } x^* \text{ and for all } (k, l) \text{ on } x^*.$$

(iv) The lower pheromone bound is chosen as

$$\tau_{min}(n) = \frac{c_n}{\log(n+1)}$$

with $\lim_{n \rightarrow \infty} c_n > 0$. (E.g., $\tau_{min}(n) = c/\log n$ for some $c > 0$ fulfills the condition.)

(v) The sample size N_n grows at least linearly in n , i.e., $N_n \geq C \cdot n$ for some constant $C > 0$.

Condition (i) is in some sense the strongest of the four conditions, but it can probably be removed along the same lines as the corresponding condition in [14] for the deterministic special case. Also if this should not be the case, a negligible change of the objective function (e.g., adding $\epsilon \cdot i(x)$, where $i(x)$ is the index of solution x according to some order, and ϵ is sufficiently small) makes (i) satisfied.

As to Condition (ii), it should be observed that it is always possible to decompose a random variable $f(x, \omega)$ with existing expected value into this expected value and an error term. That the error terms are normally distributed is not a very restrictive assumption, since in S-ACO, the observations $f(x, \omega_\nu)$ are always used as independent terms in the sample estimate (2), where they produce (after suitable normalization and for large sample size N_n) an approximately normally distributed random variable by the central limit theorem, so it does not make an essential difference if they are assumed as normally distributed from the beginning. If one wants to get rid of the assumption of normally distributed error terms, one can also apply stochastic dominance arguments, as, e.g., in [15] for the convergence of stochastic Simulated Annealing. By independent simulation runs each time a value $f(x, \omega_\nu)$ is required, the condition on stochastic independence of the error terms can easily be made satisfied.

Condition (iii) is very weak, since it is only violated if the problem-specific attractiveness values have been chosen in such an inappropriate way that the optimal walk x^* is blocked by them *a priori*.

Condition (iv) is easy to satisfy.

Also Condition (v) makes, in general, no problems (cf. Remark 1 after the theorem below).

Theorem 1. *If conditions (i) – (v) are satisfied, then for the currently best found walk $\hat{x}(n)$ in round n , for the pheromone values $\tau_{kl}(n)$ in round n and for the probability $P_\sigma(n)$ that some fixed considered ant σ traverses the optimal walk x^* in round n , the following assertions hold:*

$$\lim_{n \rightarrow \infty} \hat{x}(n) = x^* \quad \text{with probability 1,} \quad (4)$$

$$\lim_{n \rightarrow \infty} \tau_{kl}(n) = \begin{cases} 1/L(x^*), & \text{if } (k, l) \text{ on } x^*, \\ 0, & \text{otherwise} \end{cases} \quad \text{with probability 1,} \quad (5)$$

$$\lim_{n \rightarrow \infty} P_\sigma(n) = 1. \quad (6)$$

In informal terms: On the indicated conditions, the solutions subsequently produced by S-ACO tend to the (globally) optimal solution, pheromone concentrates on the optimal walk and vanishes outside, and the current walks of the ants concentrate on the optimal walk.

We prove Theorem 1 with the help of five lemmas. In the proofs, we use the following notational conventions:

- $x(n)$ is the walk selected in round n , i.e., the first parameter given to the procedure S-pheromone-update when it is called in round n ,
- $\hat{x}(n)$ is the current value of \hat{x} before the update of \hat{x} in the else-branch of S-pheromone-update in round n . In particular: $\hat{x}(1) = x(1)$.

In all five lemmas, we always assume implicitly that conditions (i) – (v) are satisfied.

Lemma 1. *For each fixed positive integer n_1 , there exists with probability one an integer $n = n(\omega) \geq n_1$, such that x^* is traversed by all ants in round n .*

Proof. Because of the lower pheromone bound as given by condition (iv),

$$\tau_{kl}(n) \geq \tau_{min}(n) = \frac{c_n}{\log(n+1)} \geq \frac{c}{2 \log(n+1)} \quad (7)$$

for some $c > 0$ and for $n \geq n_0$ with some $n_0 \in \mathbb{N}$. By condition (iii),

$$\eta_{kl}(u) \geq \gamma > 0 \quad \text{for all prefixes } u \text{ of } x^* \text{ and for all } (k, l) \text{ on } x^*,$$

since the optimal walk x^* contains only a finite number of arcs. Moreover, $\eta_{kl}(u) \leq \Gamma$ for some $\Gamma \in \mathbb{R}$, since there is only a finite number of feasible paths. Therefore, for the probability that the optimal walk x^* is traversed by a fixed ant in round n , the estimate below is obtained, where $u_k(x^*)$ denotes the prefix of walk x^* ending with node k (note that the sum of the pheromone values is unity):

$$\begin{aligned} \prod_{(k,l) \in x^*} p_{kl}(n, u_k(x^*)) &= \prod_{(k,l) \in x^*} \frac{\tau_{kl}(n) \eta_{kl}(u_k(x^*))}{\sum_{(k,r)} \tau_{kr}(n) \eta_{kr}(u_k(x^*))} \\ &\geq \prod_{(k,l) \in x^*} \frac{\gamma}{\Gamma} \frac{\tau_{kl}(n)}{\sum_{(k,r)} \tau_{k,r}(n)} \geq \prod_{(k,l) \in x^*} \frac{\gamma}{\Gamma} \tau_{kl}(n) \\ &\geq \prod_{(k,l) \in x^*} \frac{\gamma c}{2\Gamma \log(n+1)} = \left(\frac{\gamma c}{2\Gamma \log(n+1)} \right)^{L(x^*)}. \end{aligned} \quad (8)$$

Obviously, estimation (8) holds as well, if the l.h.s. refers to the probability of traversing x^* conditional on any event in round 1 to $n - 1$.

Now, let B_n denote the event that x^* is traversed in round n by all ants. Evidently,

$$\neg B_{n_1} \wedge \neg B_{n_1+1} \wedge \dots$$

is equivalent to the statement that no round $n \geq n_1$ exists such that x^* is traversed in round n by all ants. We show that

$$\text{Prob}(\neg B_{n_1} \wedge \neg B_{n_1+1} \wedge \dots) = 0. \quad (9)$$

This is seen as follows. With $n' = \max(n_0, n_1)$, the last probability is equal to

$$\text{Prob}(\neg B_{n_1}) \cdot \text{Prob}(\neg B_{n_1+1} \mid \neg B_{n_1}) \cdot \text{Prob}(\neg B_{n_1+2} \mid \neg B_{n_1} \wedge \neg B_{n_1+1}) \cdot \dots$$

$$\begin{aligned} &\leq \prod_{n=n'}^{\infty} \text{Prob}(\neg B_n \mid \neg B_{n_1} \wedge \neg B_{n_1+1} \wedge \dots \wedge \neg B_{n-1}) \\ &\leq \prod_{n=n'}^{\infty} \left[1 - \left(\frac{\gamma c}{2\Gamma \log(n+1)} \right)^{L(x^*) \cdot s} \right] \end{aligned}$$

because of (8) and the remark thereafter. The logarithm of the last expression is

$$\begin{aligned} &\sum_{n=n'}^{\infty} \log \left(1 - \left(\frac{\gamma c}{2\Gamma \log(n+1)} \right)^{L(x^*) \cdot s} \right) \\ &\leq - \sum_{n=n'}^{\infty} \left(\frac{\gamma c}{2\Gamma \log(n+1)} \right)^{L(x^*) \cdot s} = -\infty, \end{aligned}$$

since $\sum_n (\log n)^{-\lambda}$ diverges for $\lambda > 0$. It follows that (9) holds, which proves the lemma. \square

Lemma 2. *Conditionally on the event that $\hat{x}(n) = x^*$ and $x(n) \neq x^*$,*

$$\text{Prob}(\tilde{F}(n) < \hat{F}(n)) \leq g(n),$$

and conversely, conditionally on the event that $\hat{x}(n) \neq x^$ and $x(n) = x^*$,*

$$\text{Prob}(\hat{F}(n) < \tilde{F}(n)) \leq g(n),$$

where $g(n) = \phi(-C\sqrt{n})$ with ϕ denoting the distribution function of the standard normal distribution, and C is a constant only depending on

$$\sigma = \min_{x \in S} \sigma(x) \quad (10)$$

(cf. condition (ii)) and on

$$\delta = \min\{F(x) - F(x^*) \mid x \neq x^*\} > 0 \quad (11)$$

(cf. condition (i)). In other words: The probability that x^* loses a comparison against a suboptimal solution is always smaller or equal to $g(n)$.

Proof. Because of condition (ii) and by definition of $\tilde{F}(n)$,

$$\tilde{F}(n) = \frac{1}{N_n} \sum_{\nu=1}^{N_n} f(x, \omega_\nu)$$

is normally distributed with mean $F(x)$ and

$$\text{var}(\tilde{F}(n)) = \frac{1}{N_n^2} \cdot N_n \cdot (\sigma(x))^2 = \frac{(\sigma(x))^2}{N_n}.$$

For the same reason, $\hat{F}(n)$ is normally distributed with mean $F(\hat{x})$ and

$$\text{var}(\hat{F}(n)) = \frac{(\sigma(\hat{x}))^2}{N_n},$$

and $\tilde{F}(n)$ and $\hat{F}(n)$ are stochastically independent. Hence $\tilde{F}(n) - \hat{F}(n)$ is normally distributed with mean $F(x) - F(\hat{x})$ and variance

$$\frac{(\sigma(x))^2}{N_n} + \frac{(\sigma(\hat{x}))^2}{N_n} \leq 2 \frac{\sigma^2}{N_n} \leq \frac{2\sigma^2}{an}$$

with $a > 0$ given by condition (v), and σ given by (10).

For $\hat{x}(n) = x^*$, this yields:

$$\begin{aligned} \text{Prob}(\tilde{F}(n) - \hat{F}(n) < 0) &= \phi\left(-\frac{F(x) - F(x^*)}{\sqrt{(\sigma(x))^2/N_n + (\sigma(x^*))^2/N_n}}\right) \\ &\leq \phi\left(-\frac{\delta}{\sqrt{2\sigma^2/an}}\right) = \phi(-C\sqrt{n}) \end{aligned}$$

with $C = \delta\sqrt{a} / (\sqrt{2}\sigma) > 0$.

The second part of the assertion follows immediately because of the symmetry in the computation of $\hat{F}(n)$ and $\tilde{F}(n)$. \square

Lemma 3. For the function $g(n)$ defined in Lemma 2,

$$\lim_{n_1 \rightarrow \infty} \prod_{n=n_1}^{\infty} (1 - g(n)) = 1 \tag{12}$$

holds.

Proof. Because of $C > 0$, we have $0 < g(n) < 1$. Taking logarithm, we obtain that (12) is equivalent to

$$\lim_{n_1 \rightarrow \infty} \sum_{n=n_1}^{\infty} (-\log(1 - g(n))) = 0, \tag{13}$$

where each term in the sum is positive. Because of $-\log(1-x) \leq x$ for all x , a sufficient condition for (13) being satisfied is

$$\lim_{n_1 \rightarrow \infty} \sum_{n_1=1}^{\infty} g(n) = 0. \quad (14)$$

Let $\varphi(x) = \phi'(x)$ denote the density function of a standard normal distribution. By elementary calculations, it is seen that $\phi(x) \leq \varphi(x)/(-x)$ for $x < 0$. Therefore one obtains

$$\begin{aligned} \sum_{n=1}^{\infty} g(n) &= \sum_{n=1}^{\infty} \phi(-C\sqrt{n}) \leq \sum_{n=1}^{\infty} \frac{\varphi(-C\sqrt{n})}{C\sqrt{n}} \\ &\leq \frac{1}{C} \sum_{n=1}^{\infty} \varphi(-C\sqrt{n}) = \frac{1}{C\sqrt{2\pi}} \sum_{n=1}^{\infty} \exp\left(-\frac{C^2 n}{2}\right). \end{aligned}$$

The function $\exp\left(-\frac{C^2 n}{2}\right)$ is decreasing in n , so

$$\sum_{n=1}^{\infty} \exp\left(-\frac{C^2 n}{2}\right) \leq \int_0^{\infty} \exp\left(-\frac{C^2 x}{2}\right) dx < \infty.$$

Thus, $\sum_{n=1}^{\infty} g(n) < \infty$. Since $g(n) > 0$, this proves (14) and therefore also the lemma. \square

Lemma 4. *With probability one, there is an $n_2 = n_2(\omega)$ such that $\hat{x}(n) = x^*$ for all $n \geq n_2$.*

Proof. Let n_0 be the index introduced in the proof of Lemma 1, such that (7) holds for all $n \geq n_0$. We choose $n_1 \geq n_0$ in such a way that

$$\prod_{n=n_1}^{\infty} (1 - g(n)) \geq 1 - \epsilon,$$

which is possible by Lemma 3. Let G_n denote the event that in round n , the optimal solution x^* is taken for the comparison in S-pheromone-update, either as the currently selected solution $x(n)$, or as the current best-solution candidate $\hat{x}(n)$, or both, and that x^* wins the comparison, such that $\hat{x}(n+1) = x^*$. Event G_n occurs in two possible situations:

- (a) $x(n) = \hat{x}(n) = x^*$. Then automatically (i.e., with probability 1) $\hat{x}(n+1) = x^*$.
- (b) $x(n) \neq \hat{x}(n)$, and either $x(n) = x^*$ or $\hat{x}(n) = x^*$. In this situation, by Lemma 2, x^* wins the comparison with a probability of at least $1 - g(n)$, with the effect that $\hat{x}(n+1) = x^*$.

Furthermore, let D_n denote the event that round n is the first round with $n \geq n_1$ where x^* is traversed by all ants. With the notation in the proof of Lemma 1,

$$D_n = \neg B_{n_1} \wedge \neg B_{n_1+1} \wedge \dots \wedge \neg B_{n-1} \wedge B_n.$$

Consider two arbitrary fixed rounds n_2 and n with $n_1 \leq n_2 \leq n$. For $n > n_2$, the event $D_{n_2} \wedge G_{n_2} \wedge G_{n_2+1} \wedge \dots \wedge G_{n-1}$ implies that $\hat{x}(n) = x^*$, hence

$$\text{Prob}(G_n | D_{n_2} \wedge G_{n_2} \wedge G_{n_2+1} \wedge \dots \wedge G_{n-1}) \geq 1 - g(n) \quad (15)$$

by the consideration above. For $n = n_2$, on the other hand, the event $D_{n_2} \wedge G_{n_2} \wedge \dots \wedge G_{n-1}$ reduces to D_{n_2} , which implies $x(n_2) = x^*$, such that also in this case, (15) holds by the consideration above. Therefore,

$$\begin{aligned} \text{Prob}(G_{n_2} \wedge G_{n_2+1} \wedge \dots | D_{n_2}) &= \prod_{n=n_2}^{\infty} \text{Prob}(G_n | D_{n_2} \wedge G_{n_2} \wedge G_{n_2+1} \wedge \dots \wedge G_{n-1}) \\ &\geq \prod_{n=n_2}^{\infty} (1 - g(n)) \geq \prod_{n=n_1}^{\infty} (1 - g(n)) \geq 1 - \epsilon. \end{aligned}$$

The events $D_{n_1}, D_{n_1+1}, \dots$ are mutually exclusive, and by Lemma 1,

$$\text{Prob}(D_{n_1}) + \text{Prob}(D_{n_1+1}) + \dots = 1.$$

Using this, we obtain: The probability that there is a round $n_2 \geq n_1$, such that round n_2 is the first round after round n_1 where x^* is traversed by all ants and $\hat{x}(n) = x^*$ for all $n \geq n_2$, is given by

$$\begin{aligned} &\sum_{n_2=n_1}^{\infty} \text{Prob}(D_{n_2} \wedge G_{n_2} \wedge G_{n_2+1} \wedge \dots) \\ &= \sum_{n_2=n_1}^{\infty} \text{Prob}(G_{n_2} \wedge G_{n_2+1} \wedge \dots | D_{n_2}) \cdot \text{Prob}(D_{n_2}) \\ &\geq (1 - \epsilon) \sum_{n_2=n_1}^{\infty} \text{Prob}(D_{n_2}) = 1 - \epsilon. \end{aligned} \quad (16)$$

Since the l.h.s. of (16) does not depend on ϵ and $\epsilon > 0$ is arbitrary, the considered probability must be exactly 1, which proves the assertion. \square

Lemma 5. *With probability one, $\tau_{kl}(n) \rightarrow 1/L(x^*)$ for $(k, l) \in x^*$ and $\tau_{kl}(n) \rightarrow 0$ for $(k, l) \notin x^*$, as $n \rightarrow \infty$.*

Proof. By Lemma 4, there is with probability one an integer n_2 such that $\hat{x}(n) = x^*$ for all $n \geq n_2$.

(i) Let $(k, l) \in x^*$. In round n_2 and all subsequent rounds, (k, l) is always reinforced. Set $L = L(x^*)$ for abbreviation. A lower bound for $\tau_{kl}(n)$ is obtained by omitting the rule that $\tau_{kl}(n+1)$ is set equal to $\tau_{min}(n)$ if it would otherwise decrease below $\tau_{min}(n)$ by evaporation. Based on this lower bound estimation, we get by induction w.r.t. $t = 1, 2, \dots$ that

$$\tau_{kl}(n_2 + t) \geq (1 - \rho)^t \tau_{kl}(n_2) + \frac{\rho}{L} \sum_{i=0}^{t-1} (1 - \rho)^i. \quad (17)$$

As $t \rightarrow \infty$, the expression on the r.h.s. of (17) tends to

$$\frac{\rho}{L} \sum_{i=0}^{\infty} (1-\rho)^i = \frac{1}{L}.$$

Therefore, for sufficiently large t , $\tau_{kl}(n_2 + t) > 1/(2L)$. On the other hand, $\tau_{min}(n) \rightarrow 0$, hence $\tau_{min}(n_2 + t) < 1/(2L)$ for sufficiently large t , which means that updates by setting $\tau_{kl}(n+1)$ equal to $\tau_{min}(n)$ do not happen anymore for large values of t . Thus, for some t_0 and integers $t' \geq 1$, we find in analogy to (17) (but now with equality instead of inequality) that

$$\tau_{kl}(n_2 + t_0 + t') = (1-\rho)^{t'} \tau_{kl}(n_2 + t_0) + \frac{\rho}{L} \sum_{i=0}^{t'-1} (1-\rho)^i,$$

and the expression on the r.h.s. tends to $1/L$ as $t' \rightarrow \infty$.

(ii) Let $(k, l) \notin x^*$. Then (k, l) is never reinforced anymore in round n_2 and any subsequent round. Thus the pheromone on (k, l) decreases geometrically until the lower bound τ_{min} is reached. Since $\tau_{min} \rightarrow 0$ as well, we have $\tau_{kl}(n) \rightarrow 0$ as $n \rightarrow \infty$. \square

Proof of Theorem 1. The first two assertions of the theorem, eqs. (4) and (5), are the assertions of Lemma 4 and Lemma 5, respectively. The third assertion, eq. (6), is seen as follows: From (5), we obtain for $(k, l) \in x^*$ and prefix u of x^* that, with $\delta_{kr} = 1$ if $k = r$ and $\delta_{kr} = 0$ otherwise,

$$\lim_{n \rightarrow \infty} p_{kl}(n, u) = \frac{1 \cdot \eta_{kl}(u)}{\sum_{(k,r)} \delta_{kr} \cdot \eta_{kr}(u)} = 1.$$

Therefore also the probability that a fixed ant σ traverses x^* , which is given by

$$P_{\sigma}(n) = \prod_{(k,l) \in x^*} p_{kl}(n, u),$$

tends to unity as $n \rightarrow \infty$. \square

Remark 1. In Gutjahr and Pflug [15], a similar convergence result has been shown for a modification of the Simulated Annealing metaheuristic designed for the application to stochastic optimization problems. There, however, a growth of the sample size N_n of order $\Omega(n^{2\gamma})$ with $\gamma > 1$ was required for obtaining the convergence property. The growth of order $\Omega(n)$ required in Theorem 1 is much more favorable. While runtime limits are reached soon when the sample size is increased faster than with quadratic order, a linear increment usually does not impose severe practical restrictions.

Remark 2. For the solution of *deterministic* combinatorial optimization problems by metaheuristic approaches, some convergence results exist. For Simulated Annealing, e.g., it has been demonstrated by Gelfand and Mitter [11] and by Hajek [17] that by applying a suitable “cooling schedule”, one can achieve that the

distribution of the current solution tends to the uniform distribution on the set of globally optimal solutions. A related result for two particular ACO algorithms (both of the GBAS-type outlined above) has been obtained in [13].

Nevertheless, it is clear that when applied to NP-hard problems, these algorithms cannot overcome the general limitations demonstrated by NP-completeness theory: If an algorithm is designed in such a way that it is guaranteed to find the optimal solution of any (or even: some) NP-hard problem, a price must be paid: runtime will get prohibitive for larger problem instances. For example, the theoretical cooling schedule assumed in the convergence results for Simulated Annealing is too slow to be well-suited for practical applications; it has to be modified towards faster cooling, which, on the other hand, introduces the risk of premature convergence to suboptimal solutions. (This dilemma has sometimes been formulated under the term of “No-Free-Lunch Theorems”.) For this reason, algorithms with theoretical guarantee of convergence to optimality are sometimes considered as not practicable.

It is interesting to see that this restriction needs not to hold for the algorithm S-ACO presented here: Of course, when applied to large instances of problems that are NP-hard even in the deterministic boundary case, S-ACO is subject to the same limitations as the deterministic-problem algorithms converging to optimality. Very large problem instances, however, are not typical for *stochastic* combinatorial problems in current practice. As argued at the end of Section 2, such problems are already nontrivial in the case of small feasible sets, say, with a few hundred elements or even less. For such problem instances, the algorithm S-ACO can be implemented *without any modification*; also the linear increase of the sample size will not lead to prohibitive runtime behavior.

6 Modifications

The algorithm S-ACO can be modified in several different ways. Let us only indicate one possible line of extension:

Our procedure S-pheromone-update follows a “global-best” reinforcement strategy (see Gambardella and Dorigo [10]): the arcs on that walk that is considered as the best found up to now (in any of the previous rounds) are reinforced. An alternative strategy is the classical pheromone update of Ant System [7], where the amount of reinforcement is chosen proportional to the “fitness” of the solution, or the *rank-based* pheromone update, introduced by Bullnheimer, Hartl and Strauss [5]: the arcs on the k best walks found *in the current round* are reinforced by a pheromone increment proportional to $(k - j + 1)/k$ for the walk with rank j ($j = 1, \dots, k$). We shortly outline the rank-based case; the classical case can be treated analogously. In the stochastic context, one cannot determine the *absolute* ranks of the walks, but, as indicated in the Comments in Section 4, one can evaluate the walks at a random scenario or at a small sample of random scenarios drawn specifically for this round. In this way, ranks relative to the current scenarios(s) can be computed. Now, one can choose between two alternatives:

- (i) Perform S-ACO in *two phases*: In the first phase, replace in S-pheromone-update the global-best update rule by rank-based pheromone update w.r.t. the the currently drawn scenario(s). For this kind of update, sampling for getting the estimates \tilde{F} and \hat{F} is *not* required. In the second phase, start with the pheromone values obtained in the first phase, and perform, from now on, in S-pheromone-update the (global-best) update rule described in Section 4.
- (ii) Instead of working in two phases, perform pheromone update in each round by a *weighted mix* between the global-best update described in Section 4 and the rank-based update w.r.t. the current scenario(s).

It is likely that the convergence result of Section 5 can be generalized to alternative (i) above. A generalization to alternative (ii) is much more difficult; presumably, convergence to the optimal solution can only be obtained if the weight for the application of the rank-based update scheme is gradually reduced.

Both alternatives may be advantageous in practice compared with the basic algorithm, since they allow a broad initial exploration of the solution space (the results of this “learning” rounds are stored in the pheromone), which can possibly speed up convergence by guiding the search in later rounds.

7 Conclusion

We have presented a general-purpose algorithm S-ACO applicable to all problems of one of the most frequent problem type in stochastic combinatorial optimization, namely expected-value optimization under deterministic constraints, and shown that on specific, rather mild conditions, S-ACO converges with probability one to the globally optimal solution of the given stochastic optimization problem. Since the algorithm can usually be applied without the necessity of tuning parameters from “theoretical” to “practical” schemes and still keeps the property of convergence to optimality, it might be a promising candidate for computational experiments in diverse areas of application of stochastic combinatorial optimization. Of course, experimental comparisons with other meta-heuristic algorithms for this problem field, either ACO-based or derived from other concepts, would be very interesting and could be of considerable practical value.

References

1. Arnold, D.V., “Evolution strategies in noisy environments - a survey of existing work”, *Theoretical Aspects of Evolutionary Computing*, Kallel, L., Nauts, B., Rogers, A. (eds.), Springer (2001), pp. 239–250.
2. Bakuli, D.L., MacGregor Smith, J., “Resource allocation in state-dependent emergency evacuation networks”, *European J. of Op. Res.* 89 (1996), pp. 543–555.
3. Bertsimas, D., Simchi-Levi, D., “A new generation of vehicle routing research: robust algorithms, addressing uncertainty”, *Operations Research* 44 (1996), pp. 286–304.

4. Bianchi, L., Gambardella, L.M., Dorigo, M., "Solving the homogeneous probabilistic travelling salesman problem by the ACO metaheuristic", *Proc. ANTS '02*, 3rd Int. Workshop on Ant Algorithms (2002), pp. 177–187.
5. Bullnheimer, B., Hartl, R. F., Strauss, C., "A new rank-based version of the Ant System: A computational study", *Central European Journal for Operations Research* 7 (1) (1999), pp. 25–38.
6. Dorigo, M., Di Caro, G., "The Ant Colony Optimization metaheuristic", in: *New Ideas in Optimization*, D. Corne, M. Dorigo, F. Glover (eds.), pp. 11-32, McGraw-Hill (1999).
7. Dorigo, M., Maniezzo, V., Colorni, A., "The Ant System: An autocatalytic optimization process", Technical Report 91–016, Dept. of Electronics, Politecnico di Milano, Italy (1991).
8. Du, J., Leung, J.Y.T., "Minimizing total tardiness on one machine is NP-hard", *Mathematics of Operations Research* 15 (1990), 483–495.
9. Futschik, A., Pflug, Ch., "Confidence sets for discrete stochastic optimization", *Annals of Operations Research* 56 (1995), pp. 95–108.
10. Gambardella, L.M., Dorigo, M., "Ant-Q: A Reinforcement Learning approach to the traveling salesman problem", *Proc. of ML-95*, Twelfth Intern. Conf. on Machine Learning (1995), pp. 252–260.
11. Gelfand, S.B., Mitter, S.K., "Analysis of Simulated Annealing for Optimization", *Proc. 24th IEEE Conf. on Decision and Control* (1985), pp. 779–786.
12. Gutjahr, W.J., "A graph-based Ant System and its convergence", *Future Generation Computer Systems* 16 (2000), pp. 873–888.
13. Gutjahr, W.J., "ACO algorithms with guaranteed convergence to the optimal solution", *Information Processing Letters* 82 (2002), pp. 145–153.
14. Gutjahr, W.J., "A generalized convergence result for the Graph-based Ant System", accepted for publication in: *Probability in the Engineering and Informational Sciences*.
15. Gutjahr, W.J., Pflug, G., "Simulated annealing for noisy cost functions", *J. of Global Optimization*, 8 (1996), pp. 1–13.
16. Gutjahr, W.J., Strauss, Ch, Wagner, E., "A stochastic branch-and-bound approach to activity crashing in project management", *INFORMS J. on Computing*, 12 (2000), pp. 125-135.
17. Hajek, B., "Cooling schedules for optimal annealing", *Mathematics of OR*, 13 (1988), pp. 311–329.
18. Marianov, V., Serra, D., "Probabilistic maximal covering location-allocation models for congested systems", *J. of Regional Science*, 38 (1998), 401–424.
19. Norkin, V.I., Ermoliev, Y.M., Ruszczyński, A., "On optimal allocation of indivisibles under uncertainty", *Operations Research* 46 (1998), pp. 381–395.
20. Stützle, T., Hoos, H.H., "The MAX-MIN Ant system and local search for the travelling salesman problem", in: T. Baeck, Z. Michalewicz and X. Yao (eds.), *Proc. ICEC '97* (Int. Conf. on Evolutionary Computation) (1997), pp. 309–314.
21. Stützle, T., Hoos, H.H., "MAX-MIN Ant System", *Future Generation Computer Systems*, 16 (2000), pp. 889–914.