The Dynamics of Evolution Strategies in the Optimization of Traveling Salesman Problems

Hans-Thomas Nürnberg *
Fraunhofer Institut
Materialfluß und Logistik
D-44227 Dortmund
nuernberg@iml.fhg.de

Hans-Georg Beyer
Universität Dortmund
Fachbereich Informatik, LS XI
D-44221 Dortmund
beyer@ls11.informatik.uni-dortmund.de

Abstract

An analysis of the dynamic behavior of Evolution Strategies applied to Traveling Salesman Problems is presented. For a special class of Traveling Salesman Problems a stochastic model of the optimization process is introduced. Based on this model different features determining the optimization process of Evolution Strategies are analyzed. In addition the stochastic model is extended to explain some aspects of Simulated Annealing.

1 Introduction

The Evolution Strategy (ES) is a probabilistic search algorithm based on the model of biological evolution. Although the Evolution Strategy was primarily designed for parameter optimization [20] [22], it is known to be a powerful strategy for the optimization of NP-complete combinatorial optimization problems as well [12]. As a representative for NP-complete optimization problems we chose the Traveling Salesman Problem (TSP). This is the problem of finding a tour with minimal length through n cities.

While there is a comprehensive theory explaining the ES for parameter optimization [3] [6] [5], a theoretical framework for combinatorial optimization is still lacking [4]. Most research concerning evolutionary algorithms for the TSP has been done in the fields of Genetic Algorithms and Simulated Annealing [7] [1].

Convergence proofs have been obtained for simulated annealing algorithms using logarithmic cooling schedules. However, these results hold for infinite running time t ($t \to \infty$) only. Therefore it remains the questions on the quality of the approximate solutions, if the algorithm is terminated after a finite time t. Recently, this question has been addressed by a 'fitness landscape theory' [25] [24]. This approach is aiming at an explanation of the structure of combinatorial optimization problems and its influence on the quality of the approximate solutions.

In contrast to the fitness landscape approach, which is based on random walks ex-

*Principle author – this student submission is an extract of the author's diploma thesis.

ploring the search space, we consider the *actual* optimization process itself. In our opinion it is not sufficient to restrict the analysis of evolutionary algorithms to the quality of approximate solutions when we want to learn something about the way how evolution performs the optimum seeking. We should rather concentrate on the whole optimization *process* to obtain information about the problem structure and the capabilities of evolutionary algorithms.

In this work we look at the dynamic behavior of the Evolution Strategy during the optimization process. That is, the time evolution of the tour length $L(t)$ over the generations t. Comparing the dynamics of various strategies and different kinds of the TSP, we try to identify the determining factors of the optimization process. To explain the observed dynamics we introduce a stochastic model for a special variant of the Traveling Salesman Problem.

The various 'schools' of evolutionary algorithms reach quite different conclusions about the influence of the genetic operators selection, mutation and recombination. Therefore, we would like to have a model that allows quantitative predictions of the influence of these operators [4]. Such a model should be possible with respect to certain problem instances. In this work we concentrate on the $(1 \stackrel{+}{,} \lambda)$-ES with different mutation operators and without recombination. Results for these simple cases may serve as a base for more complex models later on.

A brief survey of the examined TSP instances and strategy variants follows in the next section. Then we will present empirically obtained curves from typical optimization processes. After that we introduce a stochastic model for a special class of the TSP that will be mathematically tractable. Finally we will discuss the scope of our model and its consequences for other classes of the TSP.

2 The test-field

The optimization process of Evolution Strategies for Traveling Salesman Problems depends on characteristics of the problem instance and on the strategy variant as well. We looked especially at the following TSP properties:

- Dimension,
- Symmetry,
- Distribution of edge weights,
- Dependencies between edge weights (e.g. triangle inequality).

To analyze the influence of these problem properties we generated symmetric and asymmetric TSP between 50 and 10000 cities in our experiments. Since our main interest was the influence of the distribution of edge weights in the distance matrices, we defined the following TSP-classes [18]:

- TSPs with independently generated uniformly distributed edge weights,
- Euclidean TSPs with uniformly generated coordinates in the plane,
- TSPs where the independently generated edge weights obey a distance distribution [13] which matches the distribution of the considered problems in 2-dimensional Euclidean space,
- (1,2)-TSPs where all edges have the weight one or two and are randomly generated according to different 2-point distributions [14].

The Evolution Strategy applied can be described by the following scheme:

1. INITIALIZATION:
   Generate a population $P(t = 0)$ with μ randomly chosen individuals and calculate their fitness.

2. REPRODUCTION / RECOMBINATION:
   Generate λ descendants from the actual population.

3. MUTATION:
   Mutate the offspring.

4. SELECTION:
   Calculate the fitness of the offspring and select a new population $P(t+1)$ from the offspring (e.g. (μ, λ)-ES) or from both the offspring and the parent population (e.g. $(\mu + \lambda)$-ES).

5. ITERATION:
   Repeat steps 2 - 4 until the termination criterion is fulfilled.

In this work we concentrate on the mutation operators and different selection schemes. Thus we use populations with only one individual and we do not use recombination. Typical elements of the Evolution Strategy for parameter optimization such as the self-adaptation of strategy parameters are not considered yet.

2.1 Mutation operators

In our experiments we worked with the following mutation operators:

- Inversion,
- Translocation,
- Transposition,
- 3-Point-Exchange.

Inversion cuts out a segment of a tour and reinserts it in opposite direction. Thus inversion changes two edges on symmetric problems and $O(\frac{n}{2})$ edges on the average on asymmetric problems. The inversion operator is also known as 2-opt move (cf. LIN & KERNIGHAN [16]). Translocation (3-opt) cuts a tour into three segments and builds a new tour by changing position and direction of those segments. We restricted ourselves to a translocation operator which conserves the direction of the segments. Thus translocation changes three edges in a tour for both symmetric and asymmetric problems. Transposition simply exchanges two randomly chosen points in a tour and therefore changes up to four edges per mutation. 3-Point-Exchange swaps three randomly chosen points in a tour and changes up to 6 edges per mutation.

3 Some empirical observations

In this section typical results for the optimization of Traveling Salesman Problems with the Evolution Strategy are discussed.

Figure 1 shows fitness values over time produced by an (1+1)-ES on symmetric problems with uniformly distributed edge weights. The *log*/*log*-plot supports the assumption that the dynamics can be approximated over a certain period of time by a power law of the form

$$L \quad \propto \quad t^{-\alpha} \qquad (\alpha > 0) \; , \qquad (1)$$

where L is the tour length and t the generation (cf. BEYER [2]). An analysis for TSP with independently generated edge weights has been performed in [18]. It supports the power law assumption of equation 1. However, due to space limitations it cannot presented here.

A further observation is that the optimization process can be separated in two (overlapping) phases. First there is a 'convergence phase' and afterwards a 'stagnation phase'.

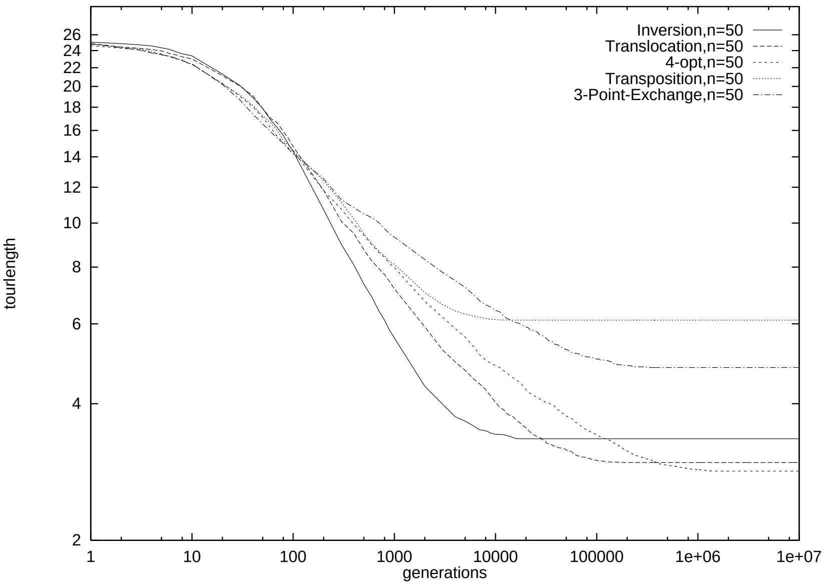

Figure 1: *Dynamics for the (1+1)-ES with different mutation operators, symmetric TSP with* $n = 50$ *and uniformly distributed edge weights between* 0 *and* 1.

In the convergence phase the convergence rate mainly depends on the number of exchanged edges per mutation and on the edge weight distribution [18]. In the stagnation phase the optimization process runs into a local sub-optimum. In this phase other influence factors become important. For example, the neighborhood size of the mutation operator influences the quality of reachable suboptimal solutions. The neighborhood size is defined as the number of states which can be reached within one mutation.

In the convergence phase we can state the following thesis for the observed TSP classes. 'The optimization process converges faster, if fewer edges are exchanged per mutation'. As a consequence the inversion operator is a good operator for symmetric problems only.

The assumption that the number of exchanged edges per mutation is a principal factor in the convergence phase is supported by a comparison between symmetric and asymmetric problem instances. Under the condition that symmetric and asymmetric TSP instances belong to the same class (e.g. they are randomly generated according to the same edge weight distribution), then there is no significant difference between the processes, provided that the applied mutation operator changes the same number of edges for symmetric and asymmetric problems (this holds for the described translocation and point-exchange operators).

Although the number of exchanged edges and the edge weight distributions are very important influence factors on the optimization process, other factors can be detected [18]. A comparison between Euclidean TSP in the plane and problems with independently generated edge weights which possess the same distribution function yields differences in the convergence behavior. Thus we have a dependency between the mutation success rate and a structural property of the problem instance. The surprising observation is that the ES for Euclidean TSP instances exhibits a larger convergence velocity than for problem instances with independent edge weights. This means that in this case the optimization process benefits from the property that the edge weights are Euclidean distances although the strategy does not use that knowledge in any way.

4 A stochastic model

The $(1,2)$-TSP is a good candidate for modeling the dynamics of Evolution Strategies. Since our main interest lies in the distribution of the edge weights in a given distance matrix and in the tour at a given time, all calculations are extremely simplified. The considered mutation operators are modeled with respect to the number of exchanged edges only. Just looking at the distribution functions and the number of exchanged edges, we abstract from the TSP-typical restriction of correct tours. Thus we expect that the dynamic behavior of the optimization process essentially depends on the edge distributions and the number of exchanged edges.

For the $(1,2)$-TSP the edge weights X of the distance matrix have the following distribution:

$$\Pr[X=1] = a_0 \qquad \Pr[X=2] = 1-a_0$$

with a given a_0 $(0 < a_0 < 1)$.

The distribution of the edges in a tour at time t has the same form:

$$\Pr_T^{(t)}[X=1] = a_t \qquad \Pr_T^{(t)}[X=2] = 1-a_t$$

The average tour length at time t calculates to

$$\begin{aligned} L_{a_0}(t) &= n \cdot a_t + 2n \cdot (1-a_t) \\ &= n(2-a_t)\ , \end{aligned} \tag{2}$$

where n is the number of cities.

4.1 $(1+1)$-ES

First we consider the $(1+1)$-ES. This is a simple probabilistic descent algorithm. It can also be interpreted as Simulated Annealing at zero temperature.

4.1.1 Inversion

The inversion operator changes two edges in a tour for symmetric TSP. For this case we can calculate the success probabilities in dependence on the tour and matrix distributions. The probability for cutting 'good' edges out of the tour calculates to

$$\begin{aligned} \Pr_T^{(t)}[N=2] &= a_t^2 \\ \Pr_T^{(t)}[N=1] &= -2a_t^2 + 2a_t \\ \Pr_T^{(t)}[N=0] &= (1-a_t)^2, \end{aligned}$$

where N is the number of edges with weight 1.

In our model we make the assumption that the probability for inserting 'good' edges $\Pr_M$ is time invariant and only depends on the matrix distribution. This assumption is justified for sufficiently large n:

$$\begin{aligned} \Pr_M[N=2] &= a_0^2 \\ \Pr_M[N=1] &= -2a_0^2 + 2a_0 \\ \Pr_M[N=0] &= (1-a_0)^2 \end{aligned}$$

Since the $(1+1)$-ES only accepts successful mutations, we now can calculate the success probabilities. Let N be the number of the 'good' edges which are finally inserted into in the tour. This may be the pair of old edges from the tour or a pair of new edges. The success probabilities read:

$$\begin{aligned}
\Pr_N[N=2] &= \Pr_T[N=2] \\
&\quad + \Pr_T[N=1]\cdot\Pr_M[N=2] \\
&\quad + \Pr_T[N=0]\cdot\Pr_M[N=2] \\
\Pr_N[N=1] &= \Pr_T[N=1]\cdot\Pr_M[N=1] \\
&\quad + \Pr_T[N=1]\cdot\Pr_M[N=0] \\
&\quad + \Pr_T[N=0]\cdot\Pr_M[N=1] \\
\Pr_N[N=0] &= \Pr_T[N=0]\cdot\Pr_M[N=0]
\end{aligned}$$

Thus, we get a new edge distribution in the tour at time $t+1$:

$$\begin{aligned}
a_{t+1} &= \frac{n-2}{n}a_t + \frac{2}{n}\Pr_N[N=2] \\
&\quad + \frac{1}{n}\Pr_N[N=1] \\
&= \frac{1}{n}\Big[2a_0(1-a_0)\cdot a_t^2 \\
&\quad + \big(n+2a_0(a_0-2)\big)\cdot a_t + 2a_0\Big] \qquad (3)
\end{aligned}$$

Equation 3 establishes an iteration scheme that can be used to calculate the a_t dynamics. However, it is also possible to derive an asymptotic exact analytical expression for a_t. Using equation 3 one obtains the difference equation:

$$a_{t+1} - a_t = \frac{2a_0}{n}\left[(1-a_0)a_t^2 - (2-a_0)a_t + 1\right] \qquad (4)$$

For $\frac{a_0}{n} \to 0$ this equation can be approximated by the differential equation:

$$\frac{da}{dt} = \frac{2a_0}{n}\left[(1-a_0)a^2 - (2-a_0)a + 1\right] \qquad (5)$$

Solving this equation by the separation method one yields:

$$a(t) = \frac{1 - a_0 - (a_0^2 - a_0 + 1)\cdot\exp\frac{2a_0^2}{n}t}{(1-a_0)^2 - (a_0^2 - a_0 + 1)\cdot\exp\frac{2a_0^2}{n}t} \qquad (6)$$

Figure 2 shows the dynamics of the (1+1)-ES for (1,2)-TSP with $n = 1000$ and different edge distributions. The dotted lines show plots of equation 6.

For $a_0 = 0.01$ we can see a deviation between the model and the empirical results of the (1+1)-ES. This deviation seems to be the result of stagnation processes in the vicinity of local suboptimal solutions and depends on the applied mutation operators. This supports our expectations, that the model is far too simple in order to explain the whole optimization process.

However, the model fits very well over a long period of time. Therefore we hope to learn more about the optimization process in the area that is covered by the model. Besides we hope that the differences between the model and the observed optimization processes can give us hints in which direction the model should be extended.

4.1.2 Translocation

In analogy to the modeling of the (1+1)-ES with inversion operator, we can look at the translocation operator. Due to the binomial distribution we can define the probabilities for cutting N 'good' edges out of a tour at time t:

$$\begin{aligned}
\Pr_T^{(t)}[N=3] &= a_t^3 \\
\Pr_T^{(t)}[N=2] &= 3a_t^2(1-a_t) \\
\Pr_T^{(t)}[N=1] &= 3a_t(1-a_t)^2 \\
\Pr_T^{(t)}[N=0] &= (1-a_t)^3
\end{aligned}$$

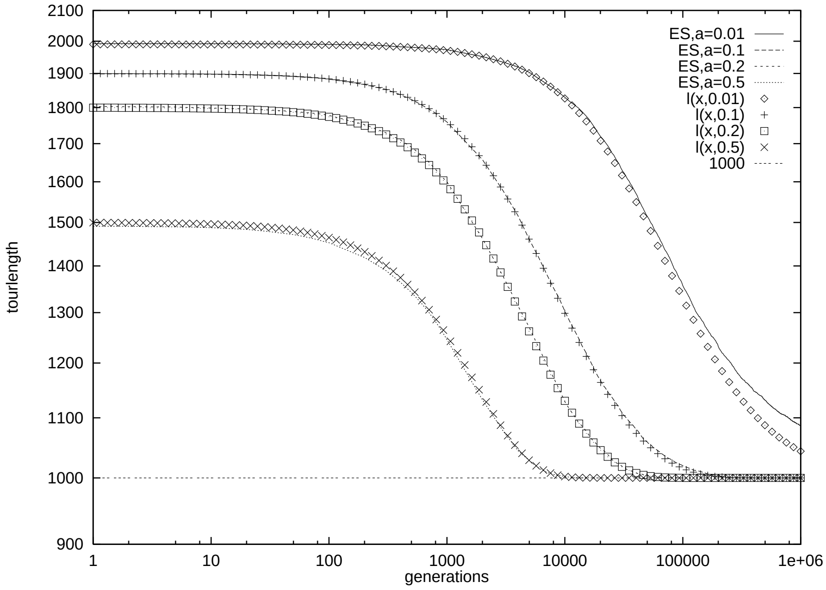

Figure 2: *(1+1)-ES with inversion, TSP with* $n = 1000$ *and different distributions, numerical results of equation 6.*

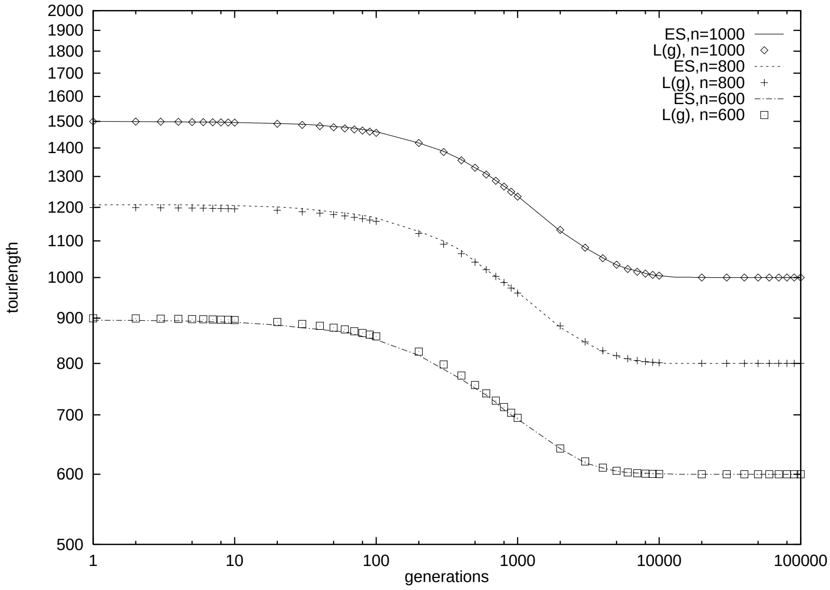

Figure 3: *(1+1)-ES with translocation, TSP with* $n = 600$, $n = 800$ *and* $n = 1000$, *numerical results of equation 7.*

With a calculation similar to the one for the inversion operator one yields the new edge distribution in the tour:

$$\begin{aligned}\Pr_T^{(t+1)}[X=1] &= \frac{1}{n}\Big[(-6a_0^3+9a_0^2-3a_0)a_t^3 \\ &\quad + (9a_0^3-18a_0^2+9a_0)a_t^2 \\ &\quad + (n-3a_0^3+9a_0^2-9a_0)a_t \\ &\quad + 3a_0\Big] \qquad (7)\end{aligned}$$

Figure 3 shows the dynamics of the $(1+1)$-ES with the translocation operator applied to symmetric $(1,2)$-TSP of different dimensions. The dotted lines are calculated by means of equation 7.

4.2 $(1+\lambda)$-ES

Now we will extend our model to the $(1+\lambda)$-ES applied to $(1,2)$-TSP. Here we restrict ourselves to the inversion operator, since other edge-exchange operators can be modeled in the same way.

First we define the following probabilities:

$$\begin{aligned}\Pr[N=+2] &= (1-a_t)^2a_0^2 \\ \Pr[N\geq+1] &= (1-a_t)^2\cdot(2a_0-a_0^2) \\ &\quad + 2a_t(1-a_t)a_0^2 \\ \Pr[N=0] &= (1-a_t)^2(1-a_0)^2 \\ &\quad + 2a_t(1-a_t)\cdot(1-a_0^2)+a_t^2\end{aligned}$$

These probabilities describe the number of 'good' edges that can be obtained by one mutation. Since the $(1+\lambda)$-ES performs λ mutations per generation, we can obtain 'good' edges in one generation with the following probabilities:

$$\begin{aligned}\Pr_\lambda[N=+2] &= 1-(1-\Pr[N=+2])^\lambda \\ \Pr_\lambda[N\geq+1] &= 1-(1-\Pr[N\geq+1])^\lambda \\ \Pr_\lambda[N=0] &= \Pr[N=0]^\lambda\end{aligned}$$

Now we can easily calculate

$$\begin{aligned}\Pr_\lambda[N=+1] &= (1-\Pr[N=+2])^\lambda \\ &\quad -(1-\Pr[N\geq+1])^\lambda \qquad (8)\end{aligned}$$

Therefore we get the following equation for the number of edges with weight 1 in a tour at time t:

$$\begin{aligned}a_{t+1} &= a_t+\frac{2}{n}\Pr_\lambda[N=+2] \\ &\quad +\frac{1}{n}\Pr_\lambda[N=+1] \\ &= \frac{1}{n}\Big[na_t+2-\big(1-(1-a_t)^2a_0^2\big)^\lambda \\ &\quad -\big(1-(1-a_t)^2\cdot(2a_0-a_0^2) \\ &\quad + 2a_t(1-a_t)a_0^2\big)^\lambda\Big] \qquad (9)\end{aligned}$$

Figure 4 shows empirical results of an $(1+\lambda)$-ES with the inversion operator. In addition the results of equation 2 and equation 9 are shown.

4.3 $(1,\lambda)$-ES

The $(1,\lambda)$-ES is a strategy variant where we allow that the fitness can deteriorate. As above, we restrict ourselves to the inversion operator. Since this operator has no way of influencing the mutation success probabilities, the success probabilities mainly depend on the fitness of the current parent individual.

Therefore the optimization process will reach an equilibrium, where the probability for a successful mutation in λ trials equals the probability for not finding a successful mutation in λ trials.

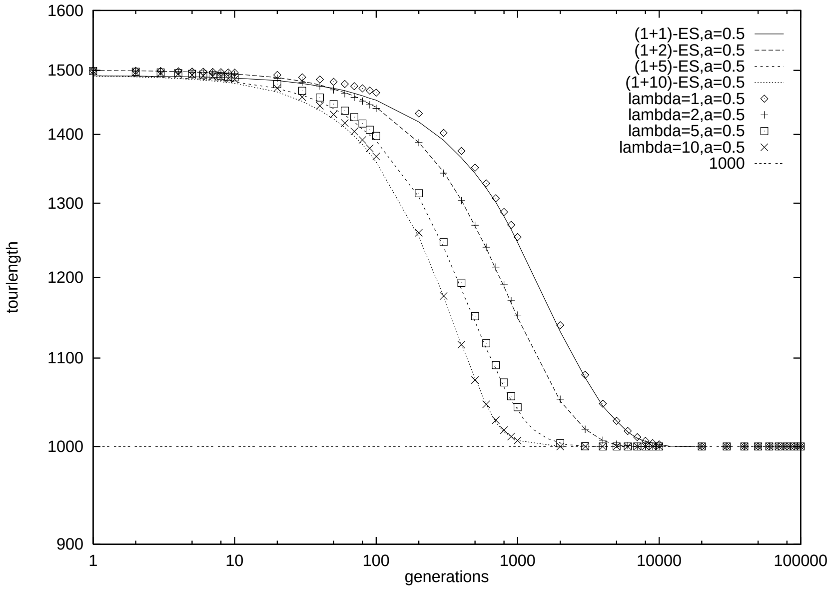

Figure 4: $(1 + \lambda)$-*ES with inversion,* $(1, 2)$-*TSP with* $a = 0.5$ *and* $n = 1000$*, numerical results of equation 2 and equation 9.*

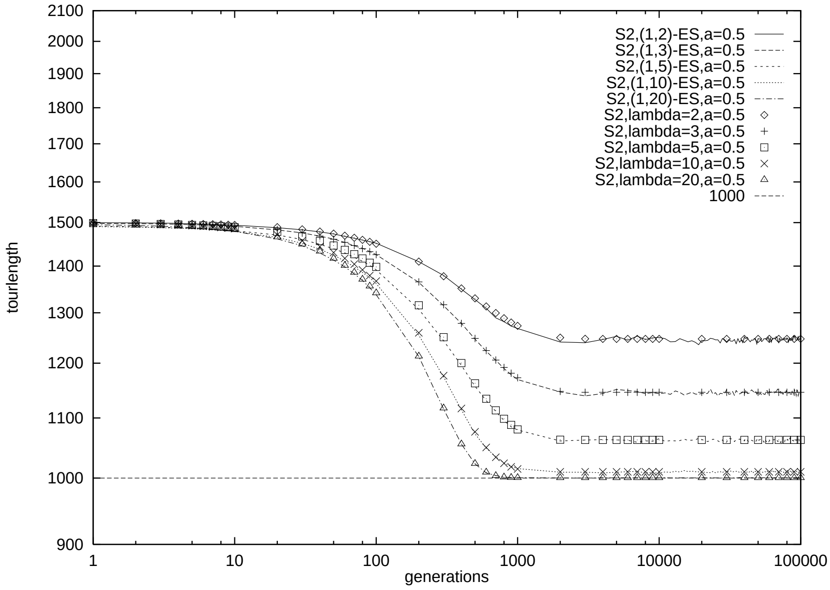

Figure 5: $(1, \lambda)$-*ES with inversion,* $(1, 2)$-*TSP with* $a = 0.5$ *and* $n = 1000$*, numerical results of equation 2 and 10.*

For our model we will define the following probabilities where N is the number of 'good' edges that we will win or loose in a mutation.

$$
\begin{aligned}
\Pr[N=+2] &= (1-a_t)^2 a_0^2 \\
\Pr[N\geq+1] &= (1-a_t)^2\cdot(2a_0-a_0^2) \\
&\quad + 2a_t(1-a_t)a_0^2 \\
\Pr[N\geq 0] &= (1-a_t)^2 \\
&\quad + 2a_t(1-a_t)\cdot(2a_0-a_0^2) \\
&\quad + a_t^2a_0^2 \\
\Pr[N\geq -1] &= 1-a_t^2 \\
&\quad + a_t^2(2a_0-a_0^2) \\
\Pr[N=-2] &= a_t^2(1-a_0)^2
\end{aligned}
$$

Now we can define the probabilities for winning or loosing N edges with weight 1 in λ trials or in one generation.

$$
\begin{aligned}
\Pr_\lambda[N=+2] &= 1-(1-\Pr[N=+2])^\lambda \\
\Pr_\lambda[N\geq+1] &= 1-(1-\Pr[N\geq+1])^\lambda \\
\Pr_\lambda[N\geq 0] &= 1-(1-\Pr[N\geq 0])^\lambda \\
\Pr_\lambda[N\geq -1] &= 1-(1-\Pr[N\geq -1])^\lambda \\
\Pr_\lambda[N=-2] &= \Pr[N=-2]^\lambda
\end{aligned}
$$

With these probabilities we can calculate the desired probabilities:

$$
\begin{aligned}
\Pr_\lambda[N=+2] &= 1-(1-\Pr[N=+2])^\lambda \\
\Pr_\lambda[N=+1] &= (1-\Pr[N=+2])^\lambda \\
&\quad -\ (1-\Pr[N\geq+1])^\lambda \\
\Pr_\lambda[N=0] &= (1-\Pr[N\geq+1])^\lambda \\
&\quad -\ (1-\Pr[N\geq 0])^\lambda \\
\Pr_\lambda[N=-1] &= (1-\Pr[N\geq 0])^\lambda \\
&\quad -\ (1-\Pr[N\geq -1])^\lambda \\
\Pr_\lambda[N=-2] &= \Pr[N=-2]^\lambda
\end{aligned}
$$

One obtains the following iteration instruction for the portion of 'good' edges in a tour at time t:

$$
\begin{aligned}
a_{t+1} &= a_t + \frac{2}{n}\Pr_\lambda[N=+2] \\
&\quad + \frac{1}{n}\Pr_\lambda[N=+1] \\
&\quad - \frac{1}{n}\Pr_\lambda[N=-1] \\
&\quad - \frac{2}{n}\Pr_\lambda[N=-2] \qquad (10)
\end{aligned}
$$

Figure 5 shows some empirical results from the $(1,\lambda)$-ES for $(1,2)$-TSP and the results of our stochastic model.

4.4 Simulated Annealing

The dynamics of Simulated Annealing for the $(1,2)$-TSP can be modeled in the same way as the dynamics of the Evolution Strategy. Despite the fact that the $(1,2)$-TSP in connection with the inversion operator is not an adequate test field for Simulated Annealing, we can show that the dynamics can be described in the terms of our stochastic model. One of the main ideas in Simulated Annealing is that a small deterioration is more easily accepted than a large deterioration. With only two possible deteriorations when using the inversion operator with the $(1,2)$-TSP, this aspect of Simulated Annealing can not be covered by our model. As above we define the probabilities for winning or loosing 'good' edges in one mutation:

$$
\begin{aligned}
\Pr[N=+2] &= (1-a_t)^2\cdot a_0^2 \\
\Pr[N=+1] &= (1-a_t)^2\cdot 2a_0(1-a_0) \\
&\quad + 2a_t(1-a_t)\cdot a_0^2 \\
\Pr[N=0] &= 1-\Pr[N=+2]-\Pr[N=+1] \\
&\quad -\ \Pr[N=-1]-\Pr[N=-2] \\
\Pr[N=-1] &= \left[a_t^2\cdot 2a_0(1-a_0)\right. \\
&\quad \left.+\ 2a_t(1-a_t)(1-a_0)^2\right]\exp^{-\frac{1}{T_t}} \\
\Pr[N=-2] &= \left[a_t^2\cdot(1-a_0)^2\right]\cdot\exp^{-\frac{2}{T_t}}
\end{aligned}
$$

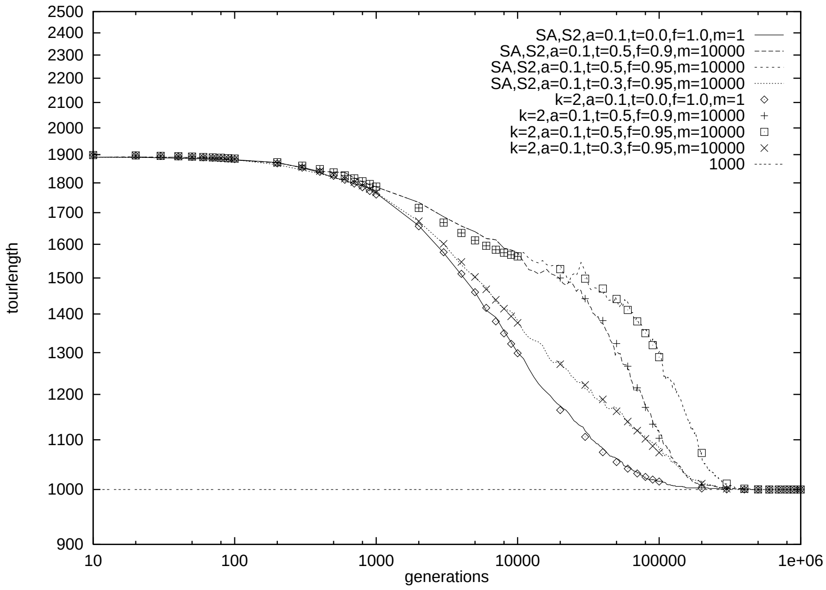

Figure 6: *Simulated Annealing with inversion,* $(1,2)$*-TSP with* $a = 0.1$ *and* $n = 1000$*, numerical results of equation 2 and 12.*

The temperature T_t depends on the cooling schedule. We worked with the following cooling schedule: T_t is initialized with a start temperature T_0 and after each m steps of the Metropolis algorithm, T_t is multiplied with a cooling factor $f < 1$. Therefore T_t calculates to

$$T_t = T_0 \cdot f^{\lfloor \frac{t}{m} \rfloor} . \qquad (11)$$

The dynamics can now be described with the following iteration instruction:

$$\begin{aligned} a_{t+1} = \; & a_t + \frac{2}{n}\,\Pr[N=+2] \\ & + \frac{1}{n}\,\Pr[N=+1] \\ & - \frac{1}{n}\,\Pr[N=-1] \\ & - \frac{2}{n}\,\Pr[N=-2] \qquad (12) \end{aligned}$$

Figure 6 shows empirical results from Simulated Annealing and numerical results of our model for different cooling schedules.

5 Conclusions

In this work we outlined an approach to obtain theoretical results which explain how Evolution Strategies work on combinatorial optimization problems. Although the stochastic model presented is only valid for a very restricted TSP, the principles can be transfered to more general problems. The essential steps are the analysis of the whole optimization process, the identification of problem and strategy characteristics and the formulation of a stochastic model based on these characteristics.

The model for the (1,2)-TSP covers the convergence phase of the optimization process only. Problem specific or mutation operator specific stagnation processes were not modeled. Nevertheless the model is able to predict the convergence behavior of different ES-variants and Simulated Annealing in a very accurate way. For the special case of the (1+1)-ES with inversion operator it was possible to derive a closed expression for the expected tourlength over the time. Iteration formulas were obtained for the $(1 \overset{+}{,} \lambda)$-ES and for Simulated Annealing

The ideas for the (1,2)-TSP are also applied to TSP with independently generated edge weights with different distribution functions in [18]. But the iterated convolution of distribution functions appeared to be a non-trivial task.

The empirical analysis of optimization processes showed that there can be a trade-off between good convergence rates and reachable solutions depending on the mutation operators used. This suggests strategies with multiple mutation operators. The choice of appropriate mutation operators could be done by some kind of empirical data analysis drawn from the actual evolutionary dynamics or even by a self-adaptive process. Developing such strategies remains as a future task.

References

[1] E. Aarts and J. Korst. *Simulated Annealing and Boltzmann Machines.* Wiley, Chichester, 1989.

[2] H.-G. Beyer. Some Aspects of the 'Evolution Strategy' for Solving TSP-Like Optimization Problems Appearing at the Design Studies of a $0.5TeV$ e^+e^--Linear Collider. In R. Männer and B.Manderick, editors, *Parallel Problem Solving from Nature, 2*, pages 361–370, Amsterdam, 1992. Elsevier Science Publishers B.V.

[3] H.-G. Beyer. Toward a Theory of Evolution Strategies: Some Asymptotical Results from the $(1,^+ \lambda)$-Theory. *Evolutionary Computation*, 1(2):165–188, 1993.

[4] H.-G. Beyer. An Alternative Explanation for the Manner in which Genetic Algorithms Operate. *BioSystems*, 1995. in print.

[5] H.-G. Beyer. Toward a Theory of Evolution Strategies: On the Benefit of Sex - the $(\mu/\mu, \lambda)$-Theory. *Evolutionary Computation*, 3(1):81–111, 1995.

[6] H.-G. Beyer. Toward a Theory of Evolution Strategies: The (μ, λ)-Theory. *Evolutionary Computation*, 2(4):381–407, 1995.

[7] L. Davis and M. Steenstrup. *Genetic algorithms and simulated annealing.* Pitman, London, 1987.

[8] Frank-Michael Dittes. Optimization on rugged landscapes: A democratic approach. Technical report, Forschungszentrum Rossendorf e.V., Dresden, 1995.

[9] G. Dueck and T. Scheuer. Treshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing. Technical report, IBM Scientific Center, Heidelberg, 1988.

[10] John Dzubera and Darrell Whitley. Advanced correlation analysis of operators for the traveling salesman problem. In Y. Davidor, R. Männer, and H.P. Schwefel, editors, *Parallel Problem Solving from Nature, 3*, Heidelberg, 1994. Springer Verlag.

[11] David E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison Wesley, Reading Massachusetts, 1989.

[12] M. Herdy. Application of the 'Evolutionsstrategie' to Discrete Optimization Problems. In H.-P. Schwefel and R. Männer, editors, *Parallel Problem Solving from Nature, 1*, pages 188–192. Springer-Verlag, Berlin, 1990.

[13] N.L. Johnson and S. Kotz. *continuos univariate distributions-2*. Houghton Mifflin Company, Boston, 1970.

[14] Viggo Kann. *On the Approximability of NP-complete Optimization Problems*. PhD thesis, Royal Institute of Technology and Computing Science, Stockholm, 1992.

[15] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by Simulated Annealing. *Science*, 220:671–680, 1983.

[16] S. Lin and B. Kernighan. An efficient heuristic procedure for the traveling salesman problem. *Operations Research*, 21:498–516, 1973.

[17] K. Mathias and D. Whitley. Genetic operators, the fitness landscape and the traveling salesman problem. In R. Männer and B. Manderick, editors, *Parallel Problem Solving from Nature, 2*. Elsevier Science Publishers B.V., 1992.

[18] Hans-Thomas Nürnberg. Das Konvergenzverhalten der Evolutionsstrategie für Traveling-Salesman-Probleme. Master's thesis, Universität Dortmund, 1996.

[19] C.H. Papadimitriou and K. Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Prentice Hall, Englewood Cliffs, New Jersey, 1982.

[20] I. Rechenberg. *Evolutionsstrategie '94*. Frommann–Holzboog Verlag, Stuttgart, 1994.

[21] Gerhard Reinelt. *TSPLIB 95*. Institut für Angewandte Mathematik, Universität Heidelberg, 1995.

[22] Hans-Paul Schwefel. *Evolution and Optimum Seeking*. Wiley, New York, 1995.

[23] P. Sibani, J.M. Pedersen, K.H. Hoffmann, and P.Salamon. Monte carlo dynamics of optimization problems. *Phys. Rev. A 42*, 1990.

[24] G.B. Sorkin. Efficient simulated annealing on fractal energy landscapes. *Algorithmica*, 1991.

[25] P.F. Stadler. Towards a theory of landscapes. Technical report, Santa Fe Institute, 1995.

[26] P.F. Stadler and W.Schnabl. The landscape of the traveling salesman problem. *Physics Letters A*, 1992.

[27] J. Michael Steele and Timothy Law Snyder. Worst-case growth rates of some classical problems of combinatorial optimization. *SIAM Journal on Computing*, 18:278–287, 1989.

[28] R. Tafelmayer and K.H. Hoffmann. Adaptive schedules for ensemble based treshold accepting. Technical report, Bereich Physik, Technische Universität Chemnitz-Zwickau, 1995.