J Stat Phys (2010) 141: 767–784
DOI 10.1007/s10955-010-0080-z

Traveling Salesman Problem with Clustering

Johannes J. Schneider · Thomas Bukur · Antje Krause

Received: 27 April 2010 / Accepted: 21 October 2010 / Published online: 3 November 2010

Abstract In the original traveling salesman problem, the traveling salesman has the task to find the shortest closed tour through a proposed set of nodes, touching each node exactly once and returning to the initial node at the end. For the sake of the tour length to be minimized, nodes close to each other might not be visited one after the other but separated in the tour. However, for some practical applications, it is useful to group nodes to clusters, such that all nodes of a cluster are visited contiguously. Here we present an approach which leads to an automatic clustering with a clustering parameter governing the sizes of the clusters.

Keywords Traveling salesman problem · TSP · Clustering · Simulated annealing

1 Introduction

The traveling salesman problem (TSP) [1–4], which was originally defined in 1832 [5], is one of the most thoroughly investigated optimization problems. Scientists from applied mathematics, computer science, statistical physics, computational biology, and operations research applied their various exact algorithms and heuristic methods to this NP-complete problem. There is no exact algorithm solving this problem in polynomial time, but so far, instances up to 33,810 nodes have already been solved to optimality [6]. In order to better compare the qualities of various heuristics applied to the TSP, libraries of benchmark instances have been developed, e.g., the TSPLIB95, which is maintained by Reinelt's group [7] and from which two instances are shown in Fig. 1.

The traveling salesman problem is defined as follows: given N nodes with distances $D(i, j)$ between each pair (i, j) of nodes, the traveling salesman has the task to find the

J.J. Schneider (✉)
Center for Computational Research Methods in Natural Sciences, Department of Physics, Mathematics, and Computer Science, Johannes Gutenberg University of Mainz, Staudinger Weg 7, 55099 Mainz, Germany
e-mail: schneidj@uni-mainz.de

T. Bukur · A. Krause
Fachhochschule Bingen, University of Applied Sciences, Berlinstr. 109, 55411 Bingen, Germany

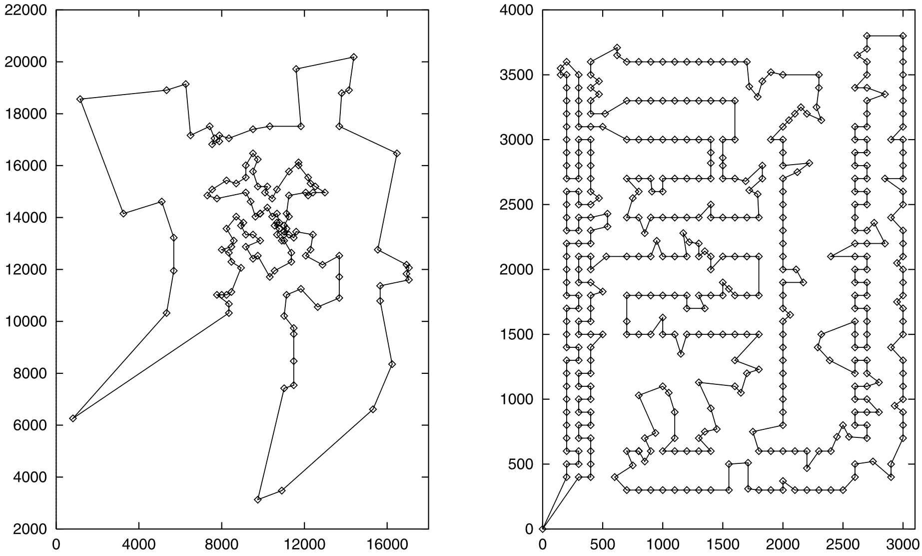

Fig. 1 Optimum solutions of two instances of the traveling salesman problem: optimum round trips through 127 beer gardens in Augsburg (BEER127, *left*) and through 442 drilling holes on a computer circuit (PCB442, *right*)

shortest closed tour, touching each node exactly once and returning to the initial node at the end. If denoting a feasible configuration as a permutation σ of the numbers $1, 2, \ldots, N$, the Hamiltonian of this problem can be written as

$$\mathcal{H}_{\text{TSP}}(\sigma) = \sum_{i=1}^{N} D(\sigma(i), \sigma(i+1)) \tag{1}$$

with $\sigma(N+1) \equiv \sigma(1)$. Although this problem seems to be academic at first sight (Hardly anybody will ever perform a round trip through the 127 beer gardens of Augsburg, whose locations are shown in the left graphic of Fig. 1), it has various practical applications, like the problem of minimizing the length of the way a drilling head has to go for drilling 442 holes in a computer circuit, as exemplarily shown in the right graphic of Fig. 1. For the case that a real traveling salesman or saleslady has to find the shortest closed tour, the original problem definition has to be extended: often enough, one is not interested in the round trip which has the shortest length but that round trip for which the smallest amount of time is needed. Then the distances are not measured in units of length, but of time. Additionally, often one has to consider the case that these distances depend on the time of day and perhaps also on the day of the week [8, 9]. Furthermore, the customers might have specific time windows in which goods can be delivered to them or fetched from them [10, 11]. This more realistic scenario can be easily extended to a vehicle routing problem, in which several traveling salesmen drive trucks with finite capacities [12], often also having to consider the time windows of their customers and depots [13]. Also other extensions of the traveling salesman problem exist, e.g., for solving sequencing problems in production processes of car producers [14].

Besides these extensions of the TSP with temporal or capacity constraints, there are also variants of the original TSP considering a related problem only on the basis of the underlying graph. There is e.g. the Chinese postman problem, in which both a set of nodes and a set of edges connecting these nodes is given. The Chinese postman has to find the shortest closed round trip, in which each of the given edges is traversed at least once [15]. For the bottleneck TSP, that closed round trip has to be found for which the longest edge has the shortest length [16]. In the traveling salesman problem with a center, the traveling salesman has to find a short round trip with the constraint to stay close to a center point if possible [17]. In the generalized TSP, clusters of nodes are given. The traveling salesman has to find the shortest closed round trip, containing one node of each cluster. In the clustered TSP (CTSP), again clusters of nodes are given. The traveling salesman has to find the shortest closed round trip through all nodes, visiting all nodes of each cluster contiguously [18, 19]. Hereby the order in which the nodes of a cluster have to be visited is not specified. Thus, one has to simultaneously find the optimum order within each cluster and the optimum order among the clusters, which are not independent problems.

2 Traveling Salesman Problem with Clustering

In practical applications, sometimes these cluster definitions as in the CTSP are a priori provided, such that nodes within each proposed cluster have to be visited contiguously within the tour. More often, however, one faces the problem, that no such clusters are a priori given, only the nodes to be visited and the distances between them are known. When having a look at the optimum round trip of a TSP instance, however, one often sees that nodes close to each other are sometimes not visited one after the other, but separated in the tour, for the sake of an overall minimum tour length. Therefore, sometimes one feels the urge to find a slightly worse tour in which nodes close to each other are visited contiguously. On the other hand, the deterioration of the tour length should remain small. Thus, the clustering of the nodes should be performed simultaneously to the minimization of the overall tour length. For this problem, we want to coin the term Traveling Salesman Problem with Clustering (TSPC), for which we introduce a Hamiltonian to be minimized as

$$\mathcal{H}(\sigma) = \sum_{i<j} D(\sigma(i), \sigma(j)) c(i, j) \tag{2}$$

with clustering factors

$$c(i, j) = \begin{cases} \alpha^{|j-i|-1} & \text{if } |j-i| \leq N/2, \\ \alpha^{N-|j-i|-1} & \text{otherwise} \end{cases} \tag{3}$$

depending on a clustering constant $0 < \alpha < 1$. Please note that the calculation of the energy of a configuration requires a computing time of order $\mathcal{O}(N^2)$, whereas it is of linear order for the original TSP. Furthermore note that the addends of the Hamiltonian of the original TSP in (1) are also addends of the Hamiltonian in (2), such that we can write

$$\mathcal{H} = \mathcal{H}_{\text{TSP}} + \mathcal{H}_{\text{C}}. \tag{4}$$

In the first addend, we consider the overall tour length, i.e., the distances to the preceding and succeeding nodes in the tour with a weight of 1. In the second addend $\mathcal{H}_{\text{C}}$, we consider the distances to the second predecessor and successor with a weight of α, to the third predecessor and successor with a weight of α^2, and so on. We can split $\mathcal{H}_{\text{C}}$ in a sum of addends by defining a TSP related Hamiltonian as

$$\mathcal{H}_\tau(\sigma) = \sum_{i=1}^{N} D(\sigma(i), \sigma(i+\tau)), \tag{5}$$

in which the distances to the τ-th neighbors are summed up. For $\tau = 1$, we have $\mathcal{H}_1 = \mathcal{H}_{\text{TSP}}$. We can write

$$\mathcal{H}_{\text{C}} = \alpha\mathcal{H}_2 + \alpha^2\mathcal{H}_3 + \cdots = \sum_{\tau=2}^{[N/2]} \alpha^{\tau-1}\mathcal{H}_\tau. \tag{6}$$

Depending on the size of the clustering constant α, more or less emphasis is put on a clustering of nodes, which are located close to each other, in the tour.

Of course, we could also have chosen any other formula with which we introduce weight values decreasing in an other way with an increasing distance of two node positions in the tour. One might think e.g. of a linear decrease

$$c(i, j) = \begin{cases} (N - |j-i| + 1)/N & \text{if } |j-i| \le N/2, \\ (|j-i| + 1)/N & \text{otherwise,} \end{cases} \tag{7}$$

but according to our opinion, this linear formula puts too much weight on nodes farther away from each other in the tour. On the other hand, an exponential decrease of the form

$$c(i, j) = \begin{cases} \exp(-\alpha(|j-i| - 1)) & \text{if } |j-i| \le N/2, \\ \exp(-\alpha(N - |j-i| - 1)) & \text{otherwise} \end{cases} \tag{8}$$

leads to a too strong decrease of the weight values. Thus, throughout this paper, we stay with the power law decrease as defined in (3).

Please note that the traveling salesman problem with clustering, as defined above, is quite different from the clustered traveling salesman problem, which was already defined 35 years ago in [18]. There, the various clusters are proposed a priori. Furthermore, visiting all nodes of a cluster in a contiguous way is a hard constraint in the CTSP. In our problem definition, clusters of nodes should arise naturally during the optimization process. Visiting all nodes of a cluster contiguously is only a weak constraint and may be violated for the sake of improving the overall tour length. Thus, there is no mapping transferring one problem directly to the other.

3 Optimization of the Traveling Salesman Problem with Clustering with Simulated Annealing

As there is no exact algorithm solving this problem, we apply the heuristic method simulated annealing [20], with which we have already found excellent solutions for the standard TSP [21–23] and with which we have recently been able to find densest packings of multidisperse systems of hard disks [24, 25].

Simulated annealing interprets the problem to be optimized as physical system, whose energy is gradually reduced in a cooling process, finally ending up in the ground state of the system with minimum energy or at least in a local minimum of the energy landscape with a comparatively small energy. Thus, the cost function of the optimization problem is interpreted as the Hamiltonian of the physical system, a feasible solution of the optimization problem as configuration of the physical system. The application of simulated annealing is rather straightforward: starting at a randomly generated initial configuration σ_0 at high temperatures, the system is subjected to a cooling schedule, until it freezes at very small temperatures in a hopefully very good solution. We use an exponential cooling schedule of the form $T_{\text{new}} = fT_{\text{old}}$ with a cooling factor f slightly smaller than 1, e.g. $f = 0.95$. In

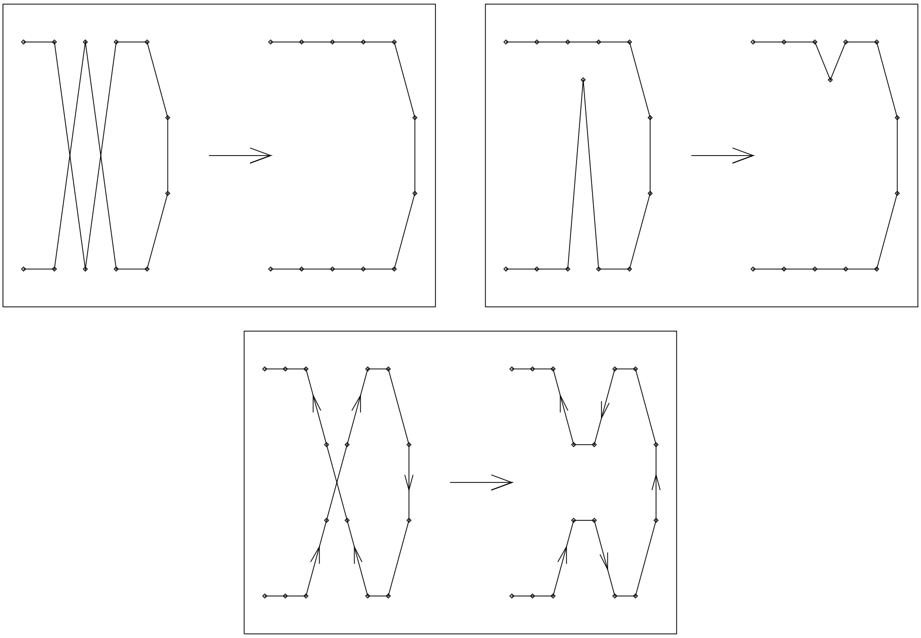

Fig. 2 Three local search moves for the TSP: Exchange (*top left*), Node Insertion Move (*top right*), and Lin-2-Opt (*bottom*)

each temperature step, a series of moves $\sigma_i \to \sigma_{i+1}$ is performed. Each move is accepted according to the Metropolis acceptance probability [26]

$$p(\sigma_i \to \sigma_{i+1}) = \begin{cases} 1 & \text{if } \Delta\mathcal{H} \le 0, \\ \exp(-\Delta\mathcal{H}/T) & \text{otherwise} \end{cases} \tag{9}$$

with T being the temperature and $\Delta\mathcal{H} = \mathcal{H}(\sigma_{i+1}) - \mathcal{H}(\sigma_i)$ being the energy difference between the current configuration σ_i and the tentative new configuration σ_{i+1}. Thus, a move is always accepted if it leads to an improvement. If a move leads to a deterioration, it is only accepted with some probability, which is the smaller, the larger the deterioration and the smaller the temperature is. In case of rejection, one sets $\sigma_{i+1} := \sigma_i$.

There are various ways of how to change the current configuration of a TSP in a new one. The simplest moves, which are shown in Fig. 2, are the Exchange, which exchanges two randomly selected nodes in the tour, the Node Insertion Move, which shifts a randomly selected node to a randomly chosen new position, and the Lin-2-Opt, which removes two randomly chosen edges from the tour, thus creating two partial sequences, turns the direction of one of these two partial sequences around, and reconnects the two partial sequences to a new feasible configuration, with two new edges. Thus. the Lin-2-Opt always removes or introduces a crossing in the tour. For these moves, the neighborhood size, i.e., the number of configurations which can be created from the current configuration, is of order $\mathcal{O}(N^2)$ [27]. The move leading to the best results is the Lin-2-Opt [21, 27].

However, it is useful also to include moves of the next higher order. There are four variants of the Lin-3-Opt [28], which remove three randomly selected edges from the tour and reconnect the three thus created partial sequences to a new feasible tour with three new edges, as shown in Fig. 3. The inclusion of even higher moves, such as the Lin-4-Opt, for

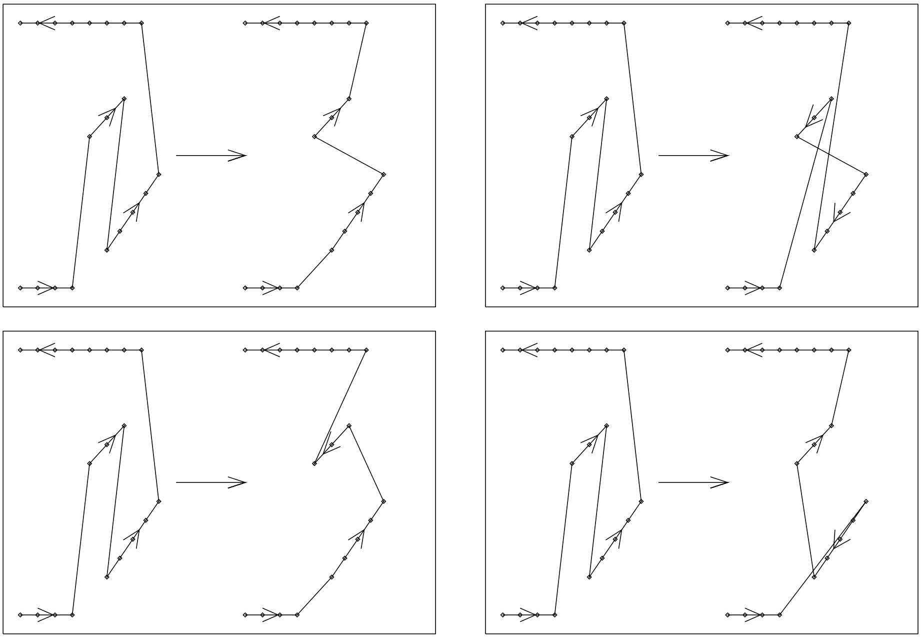

Fig. 3 The four variants of the Lin-3-Opt

which 25 variants exist, does not lead to any further improvements in the quality of the results for the TSP [28]. Thus, we also use for our traveling salesman problem with clustering only the seven above mentioned moves (Exchange, Node Insertion Move, Lin-2-Opt, and the four variants of the Lin-3-Opt) and call them with equal probability.

Please note that the calculation of the energy difference between the current and the tentative new configuration takes much more computing time in the case of the TSPC than in the case of the original TSP. While for the original TSP, only the lengths of the new edges have to be added and the lengths of the removed edges subtracted, the Hamiltonian for the traveling salesman problem with clustering contains many more addends, which have to be considered. In the case of the Exchange, a computing time of order $\mathcal{O}(N)$ has to be spent for the calculation of the energy difference. But specifically for the four variants of the Lin-3-Opt, it is necessary to calculate the overall energy of the tentative new configuration, which takes a computing time of order $\mathcal{O}(N^2)$. Also the calculation time in number of tried moves has to be increased at least linearly with the system size, such that our overall computing time scales at least with $\mathcal{O}(N^3)$. Thus, in our investigations, we are restricted to system sizes with only a few hundreds of nodes in order to perform thorough investigations of these instances.

4 Computational Results for Optimum and Quasi Optimum Solutions

In order to demonstrate both the potential and the limitations of our definition of a traveling salesman problem with clustering, we do not use TSP instances, in which the nodes can trivially be grouped to clusters, separated from each other, but use TSP instances, for which such a clustering is nontrivial. For this purpose, we choose the two instances shown in Fig. 1:

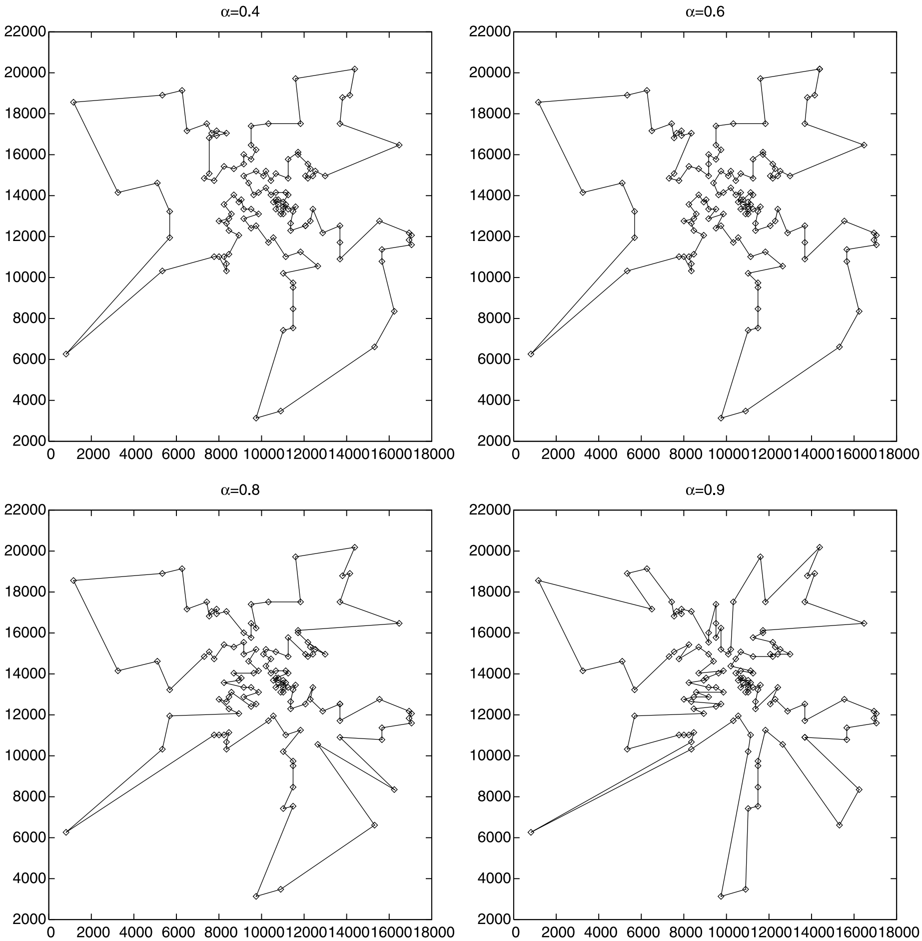

Fig. 4 Optimum solutions for the TSPC instance based on the BEER127 TSP instance using various clustering values $\alpha = 0.4, 0.6, 0.8$, and 0.9

the BEER127 instance contains a large cluster of nodes in the city center of Augsburg and some smaller clusters and isolated nodes in the villages around Augsburg. The wish to drive contiguously through all nodes in the city center of Augsburg strongly contradicts the wish to get an overall short tour length. The cluster in the city center has to be broken up in order to best insert the beer gardens in the suburbs of Augsburg and thus to decrease the tour length. An other example, in which the additional frustration by the clustering constraint will have different interesting effects, is the PCB442 instance. Although many nodes of this instance are placed on a regular grid, the PCB442 instance exhibits properties also known from random instances [27]. Furthermore, it has a highly degenerate ground state [12].

As shown for these two instances in Figs. 4 and 5, the optimum solutions of the TSPC instances change with changing clustering factor α. For small values of α, the TSP part dominates, i.e., the minimization of the overall tour length is most strongly emphasized. Increasing α to values larger than 0.5, the distances to the next nearest neighbors and the second next nearest neighbors in the tour are added with a considerable weight, thus putting

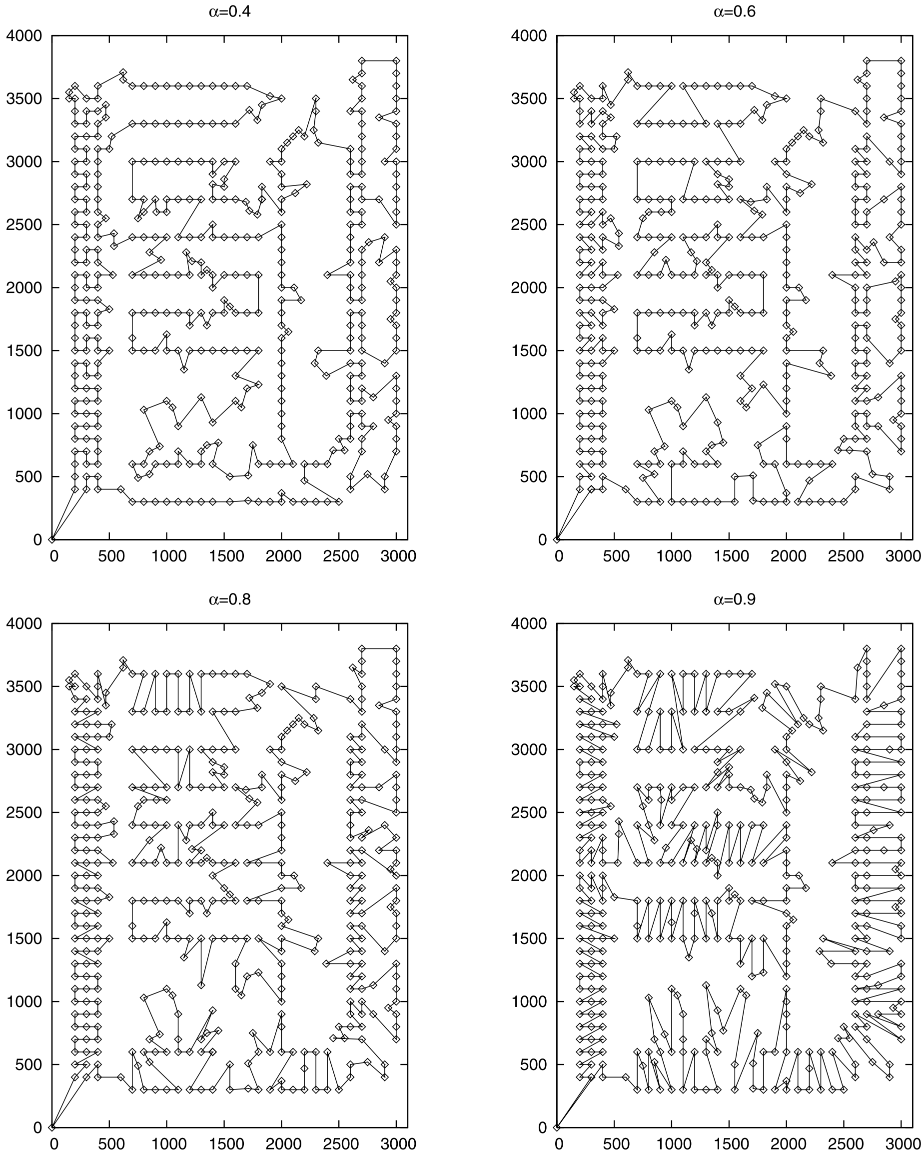

Fig. 5 Exemplary solutions for the TSPC instance based on the PCB442 TSP instance using various clustering values $\alpha = 0.4, 0.6, 0.8$, and 0.9

more emphasis on the clustering aspect. The larger α, the more the clustering part dominates, as can be seen in Fig. 5.

In order to thoroughly investigate the transition from a TSP-like system to a clustered system, we performed 100 optimization runs each for values of α from $\alpha = 0.01$ to $\alpha = 0.99$ in steps of 0.01 and for $\alpha = 0.995, 0.999, 0.9995, 0.9999, 0.99995$, and 0.99999 for the TSPC instance based on the BEER127 TSP instance. In each optimization run, we decreased the temperature T from 10^5 to 10^{-5} by a factor of 0.95. In each temperature step, we performed

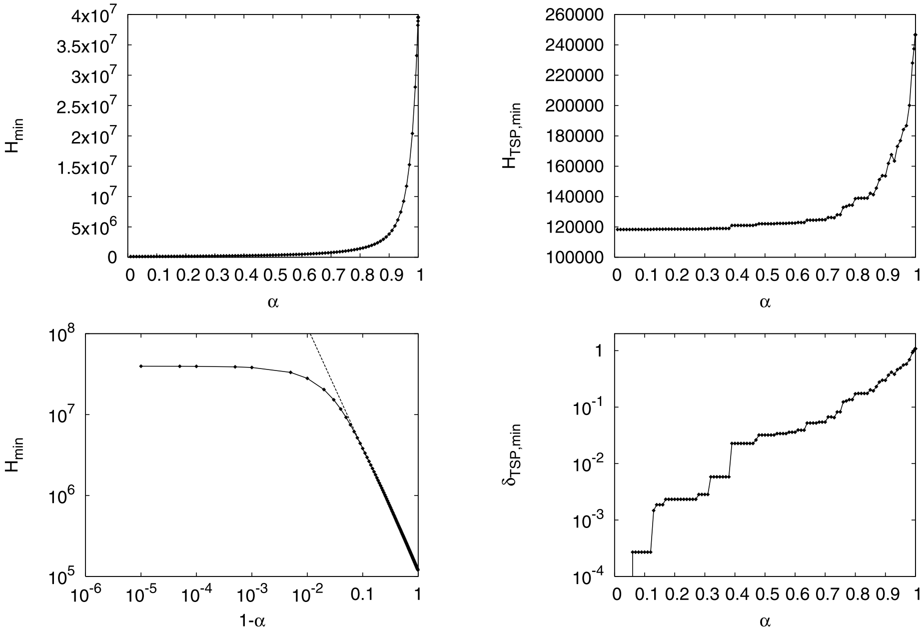

Fig. 6 Results for various values of α for the TSPC instance based on the BEER127 TSP instance: *top left*: increase of the ground state energy with increasing α, *bottom left*: ground state energy redrawn vs. $1-\alpha$ on a logarithmic scale with the fit function $1.2 \times 10^5(1-\alpha)^{-1.5}$, *top right*: development of the length of the ground state, *bottom right*: deviation of the length of the ground state from the ground state length of the corresponding TSP instance

10^4 move trials. For each value of α, we got the best result found several times, such that we are quite sure that this best result is also the optimum result of the TSPC instance.

The left graphics in Fig. 6 show these best results for the TSPC instance based on the BEER127 TSP instance. We find that the overall ground state energy increases strongly with an increasing value of α. Close to $\alpha = 0$, the ground state energy increases like

$$\mathcal{H}_{\min} \sim \frac{1}{(1-\alpha)^{1.5}}. \tag{10}$$

This power law increase holds roughly until $\alpha = 0.9$. For values between $\alpha = 0.9$ and $\alpha = 1.0$, a transition occurs: in the limit $\alpha \to 1$, the ground state energy converges towards a constant value $\approx 3.96 \times 10^7$. Please note that for the limiting case $\alpha = 1$, all edges of the round trip of the traveling salesman problem are considered with equal weight, the cluster is thus extended uniformly over the whole system. This problem with a Hamiltonian

$$\mathcal{H}(\sigma) = \sum_{i<j} D(\sigma(i), \sigma(j)) \tag{11}$$

was already studied for the determination of the most important industrial sectors by Fred Ramsey and his coworkers. In their problem, the entry $D(i, j)$ in the matrix D denotes the sum of the values industrial sector i buys from industrial sector j [29, 30]. This problem exhibits the property that there is no optimization potential at all in the case of a symmetric matrix with $D(i, j) = D(j, i)$ for all pairs (i, j). Therefore, all solutions have the same energy value in the case $\alpha = 1$. This does of course not mean that the configurations have an equal TSP length. As one gets all configurations with equal probability, the mean value of the lengths of the configurations is

$$\langle \mathcal{H}_{\text{TSP}} \rangle = N\bar{D} \tag{12}$$

with the mean distance

$$\bar{D} = \frac{1}{N(N-1)} \sum_{i,j=1}^{N} D(i,j). \tag{13}$$

For the BEER127 instance, we have $\bar{D} = 4952.476\ldots$ and thus an average of $628964.46\ldots$. Analogously, we get $\bar{D} = 1747.997\ldots$ and an average of $772614.92\ldots$ for the PCB442 instance.

The top right graphic in Fig. 6 displays the corresponding lengths of the ground state configurations for the various values of α. We find that the lengths mostly increase monotonously with an increasing value of α. In order to better visualize this increase, we additionally plot the deviation of the ground state lengths of these TSPC instances from the ground state length of the corresponding TSP instance,

$$\delta_{\text{TSP,min}} = \frac{\mathcal{H}_{\text{TSP,min}}(\alpha) - \mathcal{H}_{\text{TSP,min}}(\alpha = 0)}{\mathcal{H}_{\text{TSP,min}}(\alpha = 0)}. \tag{14}$$

Looking at the bottom right graphic in Fig. 6, we find that the ground state length stays constant for some α intervals: for $0 \le \alpha \le 0.05$, the ground state of the original TSP instance remains optimal. For $0.06 \le \alpha \le 0.12$, a configuration with the slightly larger length $118325.39\ldots$ is optimal. Further intervals are listed in Table 1. Mostly, we get shorter intervals for larger values of α.

Interestingly, we also find an interval with a constant ground state in the range $0.999 \le \alpha \le 0.99999$, exhibiting a length of $246671.17\ldots$. This length is of course much smaller than the average length $628964.46\ldots$ for $\alpha = 1$, but its corresponding configuration is also one of the degenerate ground states for $\alpha = 1$. It is the ground state with minimum length.

In order to measure the clustering effect, we have a look at the values of the Hamiltonians $\mathcal{H}_\tau$ as defined in (5) for the ground states, whose lengths are given in Table 1. There the values for $\tau = 2$ and $\tau = 3$ are listed. Up to $\alpha = 0.53$, we get monotonically decreasing values for both $\mathcal{H}_2$ and $\mathcal{H}_3$, displaying that more emphasis is put on the distances to the next nearest neighbors and third nearest neighbors. For large values of α, no further monotonic decrease can be expected, as the clusters become larger and more emphasis is also put on distances between nodes not so close to each other in the round trip.

5 Results for the Cooling Process

Secondly, we have a look at the dynamics of the cooling process in the application of simulated annealing to our traveling salesman problem with clustering. For this purpose, we performed long-running optimization runs for various values of α for the TSPC instances based on the BEER127 TSP instance and on the PCB442 TSP instance. We decreased the temperature from an initial value of $T = 10^5$ to a final value of $T = 10^{-5}$ exponentially by a cooling factor of 0.95. In each temperature step, 10^4 measurements were taken. Between two measurements each, 10 sweeps were performed. Due to the computing time for a single

Table 1 Intervals of α-values with constant length of the ground state configuration of the TSPC instance based on the BEER127 instance

$\mathcal{H}_{\text{TSP,min}}$	$\mathcal{H}_2$	$\mathcal{H}_3$	interval
118293.52...	190172.36...	245680.49...	$0.00 \le \alpha \le 0.05$
118325.39...	189600.62...	244520.84...	$0.06 \le \alpha \le 0.12$
118513.97...	188377.49...	242552.95...	$0.14 \le \alpha \le 0.16$
118568.74...	188068.05...	242366.24...	$0.17 \le \alpha \le 0.27$
118628.78...	187848.93...	242340.36...	$0.28 \le \alpha \le 0.31$
118981.00...	187452.97...	241593.14...	$0.32 \le \alpha \le 0.38$
120980.31...	186001.88...	236560.87...	$0.39 \le \alpha \le 0.46$
122086.16...	184992.98...	235488.24...	$0.48 \le \alpha \le 0.53$
122311.91...	184776.92...	235528.79...	$0.54 \le \alpha \le 0.57$
122543.14...	184637.54...	235447.42...	$0.58 \le \alpha \le 0.60$
122915.38...	184004.30...	235519.40...	$0.62 \le \alpha \le 0.63$
124482.30...	182489.70...	233363.73...	$0.64 \le \alpha \le 0.67$
124714.30...	182301.55...	233157.65...	$0.68 \le \alpha \le 0.70$
126226.78...	181172.84...	232683.79...	$0.71 \le \alpha \le 0.72$
134345.44...	193210.75...	239488.86...	$0.78 \le \alpha \le 0.79$
138969.71...	188956.09...	242937.73...	$0.81 \le \alpha \le 0.84$
246671.17...	265843.09...	312001.05...	$0.999 \le \alpha \le 0.99999$

move scaling with the order $\mathcal{O}(N^2)$, we were unable to perform even more measurements, such that there are still some larger fluctuations in the curves, especially for the specific heat.

First, we consider the decrease of the mean energy $\langle \mathcal{H} \rangle$ with decreasing temperature. At very high temperatures, the system accepts almost every deterioration and is therefore in a quasi random walk mode. Therefore, each distance value can occur at each position with equal probability. In the random walk regime, the expectation value of Hamiltonian (2) reduces to

$$\langle \mathcal{H} \rangle = \bar{D}\gamma \tag{15}$$

with the mean distance $\bar{D}$ as defined in (13) specific for each TSP instance and the clustering parameter

$$\gamma = \sum_{i<j} c(i, j) \tag{16}$$

with the clustering factors $c(i, j)$ as defined in (3), thus depending only on the clustering parameter α and the number N of nodes. For small values of α, the clustering parameter γ can be approximated by

$$\gamma \approx \frac{N}{1-\alpha}. \tag{17}$$

This approximation holds up to $\alpha \approx 0.9$. In the limit $\alpha \to 1$, γ approaches the limiting value

$$\lim_{\alpha \to 1} \gamma = N(N-1)/2. \tag{18}$$

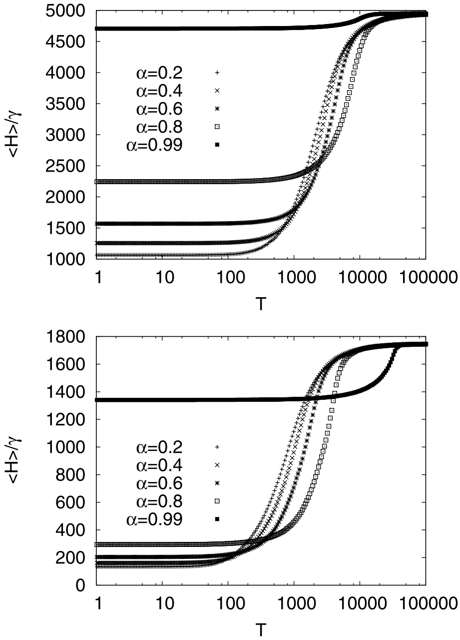

Fig. 7 Decrease of the normalized mean energy with decreasing temperature for various values of the clustering parameter α and for the two TSPC instances based on the BEER127 (*top*) and on the PCB442 (*bottom*) TSP instance

Figure 7 shows the decrease of the mean energy with decreasing temperature for various values of the clustering parameter α and for the TSPC instances based on the BEER127 TSP instance and on the PCB442 TSP instance. We normalized the mean energy by the factor γ. For infinitely high temperatures, the normalized mean value is equal to the mean distance $\bar{D}$. The mean energy decreases sigmoidally, until it freezes at small temperatures. The larger α, the higher is the transition temperature and the smaller is the difference between the resulting normalized mean value at small temperatures and $\bar{D}$. As already mentioned above, for $\alpha = 1$, there would be no optimization potential at all, $\langle\mathcal{H}\rangle$ would be constant $\bar{D}N(N-1)/2$.

In order to have a better look at the transition between the unordered high-energy regime at high temperatures and the ordered low-energy regime at small temperatures, we have a look at the specific heat

$$C = \frac{\mathrm{d}\langle\mathcal{H}\rangle}{\mathrm{d}T}, \tag{19}$$

which is given in thermal equilibrium as

$$C = \frac{\mathrm{Var}(\mathcal{H})}{T^2}. \tag{20}$$

Looking at Fig. 8, we can clearly distinguish two different scenarios for the specific heat: for small and medium-sized values of α, the specific heat looks similar to the specific heat of the corresponding TSP instance [21], which is rather symmetric around its peak when plotted vs. the logarithm of the temperature. The location of the peak, which denotes the freezing temperature of the system, is shifted towards higher temperatures for increasing values of α. For small values of α, the freezing temperature T_f scales with $(1-\alpha)^{-1.5\pm0.2}$. For the values of α we considered, we find the maximum freezing temperature for $\alpha = 0.95$ in the case of the TSPC instance based on the BEER127 TSP instance and for $\alpha = 0.99$ in the case of the TSPC instance based on the PCB442 TSP instance. For α-values approaching 1,

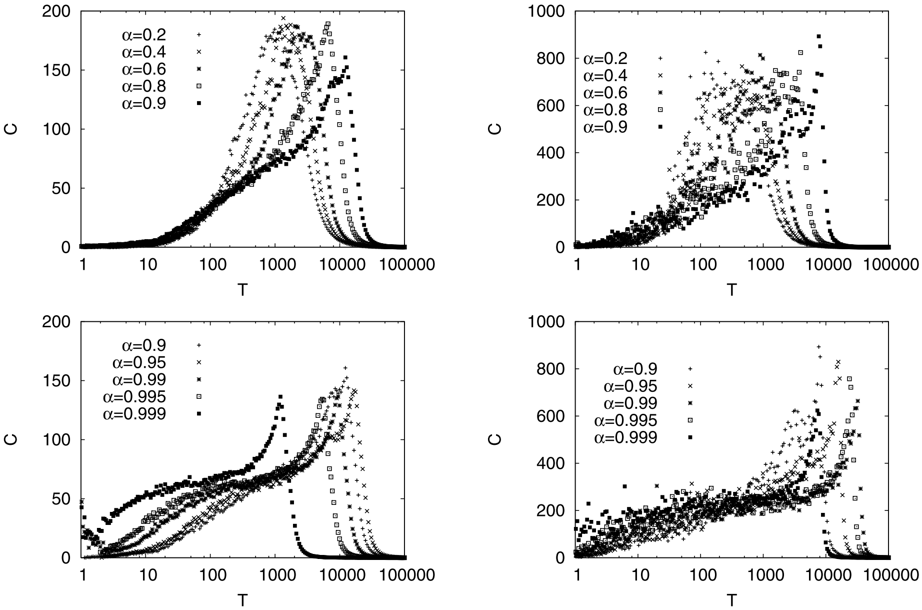

Fig. 8 Development of the specific heat for various values of α and for the TSPC instances based on the BEER127 TSP instance (*left*) and on the PCB442 TSP instance (*right*): we can clearly distinguish two scenarios for small and medium-sized values of α (*top*) and values close to the limit $\alpha \to 1$ (*bottom*)

T_f decreases again. In this regime, the specific heat is strongly asymmetric: while cooling the system, it rises sharply to its maximum and then decreases slowly over several orders of magnitude of the temperature.

As the minimization of the overall tour length is still part of the traveling salesman problem with clustering, we have also a look at the decrease of the mean length $\langle \mathcal{H}_{\text{TSP}} \rangle$ with decreasing temperature. Figure 9 shows a sigmoidal decrease for various values of α. The curves decrease from a value of $\langle \mathcal{H}_{\text{TSP}} \rangle = N\bar{D}$, which is the average value of the tour lengths occurring in the random walk. The resulting values, at which the various optimization runs freeze at small temperatures, are mostly the larger, the larger α, a result, which we already found in the last section. The decrease of $\langle \mathcal{H}_{\text{TSP}} \rangle$ is the steeper, the larger the value of α. For small and medium-sized values of α, the critical temperature T_c, at which the system orders itself according to the length constraint, increases with increasing α.

For this addend to the Hamiltonian, we can define a susceptibility. We introduce this susceptibility via an analogy to the Hamiltonian of the Ising system: there the Hamiltonian is given as

$$\mathcal{H} = -\sum_{\langle i,j \rangle} J_{ij} S_i S_j - H \sum_{i=1}^{N} S_i = \mathcal{H}_0 - HNM \tag{21}$$

with N Ising spins $S_i = \pm 1$, the interactions J_{ij} between neighboring spins, the external magnetic field H, and the magnetization M. The susceptibility is defined as

$$\chi = \frac{\partial \langle M \rangle}{\partial H}. \tag{22}$$

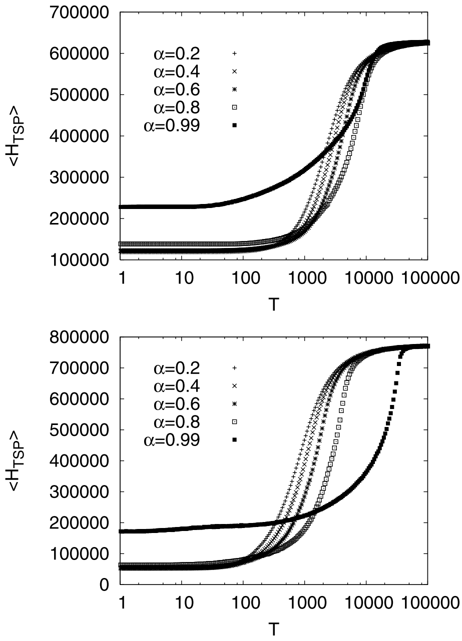

Fig. 9 Decrease of the mean length with decreasing temperature for various values of the clustering parameter α and for the two TSPC instances based on the BEER127 (*top*) and on the PCB442 (*bottom*) TSP instance

Analogously, we define for our Hamiltonian $\mathcal{H} = \lambda\mathcal{H}_{\text{TSP}} + \mathcal{H}_{\text{c}}$ with $\lambda = 1$ the susceptibility

$$\chi = \frac{\partial \langle \mathcal{H}_{\text{TSP}} \rangle}{\partial \lambda}, \tag{23}$$

which can be rewritten to

$$\chi = \frac{\text{Var}(\mathcal{H}_{\text{TSP}})}{T} \tag{24}$$

in thermal equilibrium.

Figure 10 shows that there are also two α-regimes for the susceptibility: for small and medium-sized values of α, the peak of the susceptibility and thus the critical temperature T_c is shifted towards higher temperatures. The displacement scales with $(1-\alpha)^{-1\pm 0.1}$. The height of the peak is the smaller, the larger the value of α. Increasing α, a second peak of the susceptibility at smaller temperatures emerges, which becomes larger than the right peak for α approaching the limiting value 1. The smaller the deviation of α from 1, the larger the height of this second peak and the smaller the corresponding critical temperature. For the TSPC instance based on the BEER127 TSP instance, we can even see a third peak, which we cannot distinguish for the TSPC instance based on the PCB442 TSP instance.

6 Summary and Outlook

As a variation of the original traveling salesman problem, in which a traveling salesman has the task to find the overall shortest round trip through a proposed set of nodes, we consider the problem that he also takes the distances to further succeeding nodes in the tour into account in order to visit nodes lying geographically close to each other contiguously in the tour. The weights which are put on the distances to the next nearest neighbors in the tour

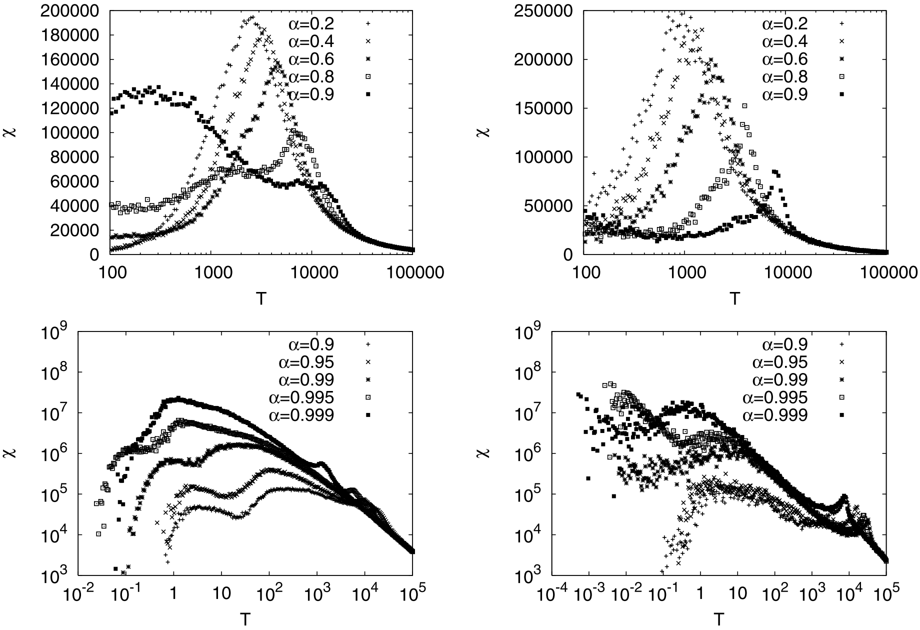

Fig. 10 Development of the susceptibility (as defined in the text) for various values of α and for the TSPC instances based on the BEER127 TSP instance (*left*) and on the PCB442 TSP instance (*right*): we can clearly distinguish two scenarios for small and medium-sized values of α (*top*) and values close to the limit $\alpha \to 1$ (*bottom*)

decrease with increasing differences between the tour position numbers of the corresponding nodes. Depending on these weights, the clustering aspect of contiguously visiting nodes lying close to each other is more or less emphasized compared to the minimization of the overall tour length. Due to a simultaneous optimization, clusters of nodes lying close to each other and being contiguously visited arise naturally while the overall tour length is minimized. The more the clustering aspect dominates, the larger the optimum overall tour length becomes. For intermediate values of the clustering constant, the deterioration of the overall tour length remains small while some small clusters can already develop.

We apply the general purpose optimization algorithm simulated annealing to this problem, finding that the dynamics of the cooling process is rather similar to the process for the original traveling salesman problem for small and intermediate values of the clustering constant. For large values of the clustering constant, the specific heat becomes strongly asymmetric, the susceptibility exhibits more than one peak, and the difference between the mean value of random configurations and the optimum value decreases.

Not always is such a clustering extension suitable for a real-world traveling salesman problem. For example, if a real traveling salesman has to deal with customers having time windows, the fulfillment of time windows often already contradicts the aim to get a minimum overall tour length. The introduction of an additional clustering term would be highly disadvantageous. On the other hand, if no such constraints exist, it would sometimes provide a positive psychologic effect to visit customers close to each other contiguously. But there are also other variants of the traveling salesman problem, for which an additional clustering approach would be highly useful: for example, the traveling salesman approach was applied to determine the temporal sequence between tombs in prehistoric graveyards. Here the distances are measured by differences in the properties of grave goods and in the positions of the skeletons. Often one finds that such a graveyard is strongly used during some prehistoric time interval, then the usage was interrupted for some time, until it was used again, and so on. For detecting such time intervals, our clustering approach might be highly useful.

We have already started applying a variant of our Traveling Salesman Problem with Clustering to various problems. Just as in the well-known Traveling Salesman Problem with Open End Points, for which the Hamiltonian is given as

$$\mathcal{H}_{\text{TSPOEP}}(\sigma) = \sum_{i=1}^{N-1} D(\sigma(i), \sigma(i+1)) \tag{25}$$

and in which the shortest tour with arbitrary initial and final nodes, touching each node exactly once has to be found, we have a look at the Traveling Salesman Problem with Clustering and Open End Points, for which the Hamiltonian can be written as

$$\mathcal{H}_{\text{TSPCOEP}}(\sigma) = \sum_{i<j} D(\sigma(i), \sigma(j)) \alpha^{j-i-1}. \tag{26}$$

Using this Hamiltonian [25], we first of all intend to visualize the Parisi block structure found for overlaps $q_{\sigma,\tau}$ between (quasi) optimum configurations σ and τ for Sherrington-Kirkpatrick spin glasses [31]. The distances between the configurations are given as $D(\sigma,\tau) = 1 - q_{\sigma,\tau}$, if the overlaps are normalized. We expect to see a structure like

$$(q_{\phi(\sigma),\phi(\tau)}) = \begin{pmatrix} \boxed{\begin{matrix} \boxed{\begin{matrix} 1 & q_1 \\ q_1 & 1 \end{matrix}} & q_2 \\ q_2 & \boxed{\begin{matrix} 1 & q_1 \\ q_1 & 1 \end{matrix}} \end{matrix}} & q_3 \\ q_3 & \boxed{\begin{matrix} \boxed{\begin{matrix} 1 & q_1 \\ q_1 & 1 \end{matrix}} & q_2 \\ q_2 & \boxed{\begin{matrix} 1 & q_1 \\ q_1 & 1 \end{matrix}} \end{matrix}} \end{pmatrix}, \tag{27}$$

if the (quasi) optimum solutions are ordered according to an optimum sequence ϕ. We expect to see this Parisi block structure not only for the Sherrington-Kirkpatrick model but also for other optimization problems, for which ultrametric properties were found, such as the traveling salesman problem [33] and a multi-disperse packing problem [24, 25], which was recently studied by us.

Secondly, we intend to use the Hamiltonian (26) for generating phylogenetic trees. Hereby the distances $D(i, j)$ can be defined in various ways for measuring the differences between species i and j. Again we want to get a picture with block structures denoting families of species and sub-blocks within these blocks showing genera of species, and so on.

Of course, the results of these attempts will strongly depend on the size of the clustering constant α, just as the results for the Traveling Salesman Problem with Clustering studied in this paper. The results shown here provide an insight in our clustering approach, displaying both the gains which could be possibly achieved by it as well as the dangers of misinterpreting results by choosing wrong values for the clustering constant.

Acknowledgements J.J.S. would like to thank the John von Neumann Institute for Computing at the Research Center Julich for a large grant of computing time under project number HMZ15. Furthermore, J.J.S. would like to thank the Center for Computational Research Methods in Natural Sciences of the Johannes Gutenberg University of Mainz for financial support.

References

1. Reinelt, G.: The Traveling Salesman. Springer, Berlin (1994)
2. Lawler, E.L., Lenstra, J.K., Rinnooy-Kan, A.H.G., Shmoys, D.B. (eds.): The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization. Wiley Interscience Series in Discrete Mathematics. Wiley, New York (1985)
3. Applegate, D.L., Bixby, R.E., Chvátal, V., Cook, W.J.: The Traveling Salesman Problem: A Computational Study. Princeton Series in Applied Mathematics. Princeton University Press, Princeton (2007)
4. Gutin, G., Punnen, A.P. (eds.): The Traveling Salesman Problem and its Variations. Springer, Berlin (2007)
5. Der Handlungsreisende – wie er seyn soll und was er zu thun hat, um Aufträge zu erhalten und eines glücklichen Erfolgs in seinen Geschäften gewiß zu seyn – von einem alten Commis-Voyageur (1832)
6. Cook, W., Espinoza, D., Goycoolea, M.: Computing with domino-parity inequalities for the TSP. Preprint (2005)
7. www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95
8. Bentner, J., Bauer, G., Obermair, G.M., Morgenstern, I., Schneider, J.: Optimization of the time-dependent traveling salesman problem with Monte Carlo methods. Phys. Rev. E **64**, 036701 (2001)
9. Schneider, J.: The time-dependent traveling salesman problem. Physica A **314**, 151–155 (2002)
10. Dumas, Y., Desrosiers, J., Gelinas, E., Solomon, M.M.: An optimal algorithm for the traveling salesman problem with time windows. Oper. Res. **43**, 367–371 (1995)
11. Gendreau, M., Hertz, A., Laporte, G., Stan, M.: A generalized insertion heuristic for the traveling salesman problem with time windows. Oper. Res. **43**, 330–335 (1998)
12. Schneider, J.: Searching for backbones—a high-performance parallel algorithm for solving combinatorial optimization problems. Future Gener. Comput. Syst. **19**, 121–131 (2003)
13. Schrimpf, G., Schneider, J., Stamm-Wilbrandt, H., Dueck, G.: Record breaking optimization results—using the ruin & recreate principle. J. Comput. Phys. **159**, 139–171 (2000)
14. Schneider, J., Britze, J., Ebersbach, A., Morgenstern, I., Puchta, M.: Optimization of production planning problems—a case study for assembly lines. Int. J. Mod. Phys. C **11**, 949–972 (2000)
15. Edmonds, J., Johnson, E.L.: Matching, Euler tours, and the Chinese postman. Math. Program. **5**, 88–124 (1973)
16. Parker, R.G., Rardin, R.L.: On the bottleneck traveling salesman problem. Oper. Res. Lett. **2**, 269–272 (1982)
17. Lipowski, A., Lipowska, D.: Traveling salesman problem with a center. Phys. Rev. E **71**, 067701 (2005)
18. Chisman, J.A.: The clustered traveling salesman problem. Comput. Oper. Res. **2**, 115–119 (1975)
19. Laporte, G., Palekar, U.: Some applications of the clustered travelling salesman problem. J. Oper. Res. Soc. **53**, 972–976 (2002)
20. Kirkpatrick, S., Gelatt, C.D. Jr., Vecchi, M.P.: Optimization by simulated annealing. Science **220**, 671–680 (1983)
21. Schneider, J., Kirkpatrick, S.: Stochastic Optimization. Springer, Berlin (2006)
22. Schneider, J., Froschhammer, Ch., Morgenstern, I., Husslein, Th., Singer, J.M.: Searching for backbones—an efficient parallel algorithm for the traveling salesman problem. Comput. Phys. Commun. **96**, 173–188 (1996)
23. Schneider, J., Morgenstern, I., Singer, J.M.: Bouncing towards the optimum: improving the results of Monte Carlo optimization algorithms. Phys. Rev. E **58**, 5085–5095 (1998)
24. Müller, A., Schneider, J.J., Schömer, E.: Packing a multidisperse system of hard disks in a circular environment. Phys. Rev. E **79**, 021102 (2009)
25. Schneider, J.J., Müller, A., Schömer, E.: Ultrametricity property of energy landscapes of multidisperse packing problems. Phys. Rev. E **79**, 031122 (2009)
26. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E.: Equation of state calculations by fast computing machines. J. Chem. Phys. **21**, 1087–1092 (1953)
27. Stadler, P.F., Schnabl, W.: The landscape of the traveling salesman problem. Phys. Lett. A **161**, 337–344 (1992)
28. Stattenberger, G., Dankesreiter, M., Baumgartner, F., Schneider, J.J.: On the neighborhood structure of the traveling salesman problem generated by local search moves. J. Stat. Phys. **129**, 623–648 (2007)

29. Ramsey, F.L., Pierce, D.A., Bowman, V.J.: Triangularization of input-output matrices. Technical Report No. 16, Department of Statistics, Oregon State University (1969)
30. Ramsey, F.L., Pierce, D.A., Bowman, V.J.: Triangularization of input-output matrices. Proc. Bus. Econ. Stat. Sect. Am. Stat. Assoc. 676 (1969)
31. Sherrington, D., Kirkpatrick, S.: Solvable model of a spin glass. Phys. Rev. Lett. **35**, 1792–1796 (1975)
32. Parisi, G.: A sequence of approximated solutions to the S-K model for spin glasses. J. Phys. A, Math. Gen. **13**, L115–L121 (1980)
33. Kirkpatrick, S., Toulouse, G.: Configuration space analysis of travelling salesman problems. J. Phys. **46**, 1277–1292 (1985)