Replica symmetry and replica symmetry breaking for the traveling salesperson problem

Hendrik Schawe [1,*] Jitesh Kumar Jha,[1,2,†] and Alexander K. Hartmann[1,‡]

[1]*Institut für Physik, Universität Oldenburg, D-26111 Oldenburg, Germany*
[2]*Manipal Institute of Technology, 576104 Karnataka, India*

We study the energy landscape of the traveling salesperson problem (TSP) using exact ground states and a novel linear programming approach to generate excited states with closely defined properties. We look at four different ensembles, notably the classic finite dimensional Euclidean TSP and the mean-field-like (1,2)-TSP, which has its origin directly in the mapping of the Hamiltonian circuit problem on the TSP. Our data supports previous conjectures that the Euclidean TSP does not show signatures of replica symmetry breaking neither in two nor in higher dimension. On the other hand the (1,2)-TSP exhibits some signature which does not exclude broken replica symmetry, making it a candidate for further studies in the future.

I. INTRODUCTION

The concept of *replica symmetry breaking* (RSB) was introduced in the context of spin glasses [1,2], where it has a long history of debate to which models it applies [3]. RSB is an assumption about the structure of the phase space (or "energy landscape"), which leads to the correct results for the Sherrington-Kirkpatrick (SK) spin glass [4]. RSB basically means that the phase space is hierarchically structured such that two configurations of very similar energy may be far away from each other in the configuration space. The phase space becomes *complex*.

The physics-inspired analysis of the phase-space structure has also been applied to combinatorial optimization problems, namely problems belonging to the class of nondeterministic polynomial (NP)-hard [5–7] problems (or the corresponding decision problems belonging to the class of NP-complete problems). For NP-hard problems currently only algorithms are known which exhibit a worst-case running time which grows exponentially with system size. Examples of NP-hard problems are satisfiability [8] and vertex cover [9]. Here, ensembles are known where replica symmetry (RS) breaks at some value of a control parameter [10–12]. This appears not to be surprising to many researchers because intuitively a hard optimization problem may correspond to a nontrivial energy landscape. This prompted many attempts to distinguish easy from hard instances or explore the energy landscape of such problems [13–19].

One of the best-known NP-hard combinatorial optimization problems is the traveling salesperson problem (TSP) [20]. Somewhat surprisingly, in contrast to the aforementioned problems, only indications for RS have been found within studies of some TSP ensembles so far [21–24]. Nevertheless, for these analytical and numerical studies various approxi-

mations had to be used, somehow questioning the previous claims for RS.

In this work, by performing computer simulations [25] via calculating numerically exact ground states [26] and excitations, we confirm the previous results for these specific ensembles. But on the other hand we show that there is at least one ensemble also for the TSP where RSB cannot be excluded, namely the (1,2)-TSP ensemble [27]. In particular, in contrast to previous numerical studies, which used heuristics to generate tours near the optimum [23,24], we use an exact algorithm to find the true optimum and very specific excitations. This approach is facilitated by the combination of flexibility and high performance (compared to other exact algorithms for the TSP) of linear programming (LP) with branch and cut. Combined with the general increase in computing power and the improvement of algorithms for TSP optimization, it enables us to simulate comparatively large instances.

II. MODELS

The traveling salesperson problem [28,29] is defined on a complete weighted graph, where the vertices are usually called *cities* and the symmetric edge weights $c_{ij} = c_{ji}$ *distances* or *costs*. On this graph one searches for the shortest cyclic path through all N cities, which is called *tour* and can be represented by a set of edges T. An equivalent representation is through a symmetric adjacency matrix $\{x_{ij}\}$ where $x_{ij} = 1$ if city i is followed by city j on the tour and $x_{ij} = 0$ otherwise. The *length* of the tour, which we will also call *energy*, is thus

$$L = \sum_{\{i,j\} \in T} c_{ij} = \sum_i \sum_{j<i} c_{ij} x_{ij}.$$

Note that an instance of the problem is completely encoded in the distance matrix c_{ij}.

To compare two tours T_1 and T_2, their *distance* or *difference* d is defined as the number of edges, which are in T_1 but not in T_2 [13],

$$d = \sum_{\{i,j\} \in T_1} 1 - x_{ij}^{(2)},$$

_____
*hendrik.schawe@uni-oldenburg.de
†jiteshjha96@gmail.com
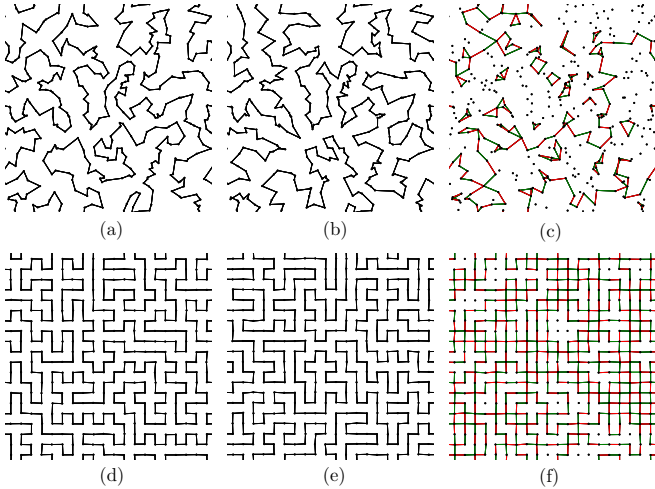‡a.hartmann@uni-oldenburg.de

032135-1

FIG. 1. (a) and (b) A configuration with $N = 400$ of the ETSP. Left is the optimal tour, right the MaxDiff excitation with $d = 129$ difference to the optimum. (d) and (e) The optimal and excited tour for an STSP realization. (c) and (f) The difference between the optimum and the excitation of the respective instances; red edges are removed, green are added for the excitation.

where $x_{ij}^{(2)}$ is the adjacency matrix corresponding to T_2. Like the link overlap for spin glasses is robust against the flipping of *compact* clusters with a low domain-wall energy, this observable is robust against partial reversals of the tour. If one considered instead the order of the cities in the tour, roughly analogous to the spin overlap used for spin glasses, this could introduce a difference in the order of N by just changing two links.

Here, we study four TSP ensembles to evaluate the influence of the quenched randomness on the complexity of the solutions.

(1) First, the most intuitive and probably the most scrutinized [15,20,30–33] ensemble is the *Euclidean TSP* (ETSP). Here a Poisson point process in a square determines the locations of the cities and the distance matrix is filled with their Euclidean distances. We use periodic boundary conditions. An example for an optimal tour in such a configuration is shown in Fig. 1(a). It is straight forward to generalize this in higher dimensions using a Poisson point process in a hypercube and the corresponding Euclidean distances.

(2) The *random link model* (RLTSP) [13,34] is an approximation of the ETSP, which disregards any correlations of the entries in the distance matrix c_{ij} and therefore does not obey, e.g., the triangle inequality. For this approximation in the statistical physics literature solutions were obtained under the premise that replica symmetry holds based on the replica method [21] and cavity method [24,34,35]. In this work we study the ETSP and RLTSP ensembles in which the density of the cities is constant, such that the average optimal tour length $L^o \sim N$ [30], i.e., the energy is extensive. The random link length is therefore drawn uniformly from $[0, N]$.

(3) The $(1, 2)$-TSP is the result of the classical mapping of the Hamilton circuit problem (HCP) onto the TSP [6]. The HCP is whether a cycle visiting every vertex exactly once exists on a given graph G. The mapping from HCP to TSP

is simply assigning the distance matrix as

$$c_{ij} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are adjacent in } G, \\ 2, & \text{otherwise.} \end{cases}$$

A Hamiltonian cycle exists, iff the length of the optimal tour is equal N. For simplicity sake, the ensemble we are looking at is derived from an Erdős-Rényi graph (ER) [36] where edges occur with probability $p = 1/N$, which results in an average degree of 1. Note that both limiting cases $p = 0$ and $p = 1$ are trivial since every tour will be optimal with length $2N$, respectively, N. $p = 1/N$ was chosen since it is the percolation threshold for the underlying graph ensemble, i.e., G exhibits a forestlike structure and to form a cycle in the corresponding TSP realization almost surely edges nonexisting in G, i.e., distance 2 in the TSP, need to be used.

(4) An additional ensemble that we look at is an Euclidean TSP, where the cities are arranged on a square lattice with lattice constant 1 (STSP). Every city is displaced by at most $1/N$ in a random direction to avoid degeneracy. An optimal tour in such a configuration is shown in Fig. 1(d). While this ensemble may appear arbitrary and trivial at first, because it is very similar to a grid which is easy to solve, it is actually rather nicely motivated. First, the constructions to map the *exact cover* problem onto the ETSP [20,33] result in instances where most cities lie on the sites of a square lattice, though not every site is occupied. This mapping is the usual way to show that even the ETSP is NP-hard. Second, historically the "ts225" instance of the TSPLIB [37] with 225 cities was solved only in 1994—three years after the record of the largest optimally solved nontrivial instance was set to 2392 cities and 10 years after its inclusion in the TSPLIB [38]. The empirically hard ts225 instance consists of cities on square lattice sites and equidistant cities on straight lines between nearest neighbor sites. Since we want to look at an easy to define ensemble, we propose the slightly disturbed square lattice, which we suspect could show typical properties of these square-lattice-like configurations. It turns out that open boundary conditions lead to strong finite-size effects, since overlaps between two arbitrary tours are coerced at the boundary. Therefore, like for the ETSP, we use periodic boundary conditions for the STSP ensemble.

The STSP is obviously a very specific subset of the ETSP. The justification to expect a different behavior is that the typical ETSP instance might be diluted by entropically favored instances with trivial solution space structure, but the solution space structure of the STSP subset might look complex. Subspaces in the problem domain which behave dramatically different are quite common. For example, a subspace of the spin glass configuration space are ferromagnets, which have a trivial solution space structure, while general spin glasses—at least in high dimensions—have complex ones.

III. METHODS

Like other studies on the solution space structure of different optimization problems, we look at excitations [39–41]. To test whether RSB is a possibility, we test a necessary criterion introduced in the context of TSP by Mézard and Parisi in Ref. [22]. A configuration is called *quasioptimal* if the relative difference of its energy L^* to the optimal energy L^o behaves

as

$$\frac{L^* - L^o}{L^o} = \mathcal{O}\left(\frac{1}{N}\right). \tag{1}$$

According to Ref. [22], in order for replica symmetry to be broken, it is necessary that there exist quasioptimal configurations, whose differences to the optimum behave as

$$d(T^o, T^*) = \mathcal{O}(N). \tag{2}$$

This does not say anything about other configurations which will have other distances to the optimum; there will be always a distribution of distances to the optimum. This distribution depends on the instances and on the system size, similar to the distribution of overlaps in spin glasses [2]. Thus, for the present analysis, it is not relevant whether this distribution converges, or whether the mean converges or whether in case of convergence they are self-averaging. Intuitively Eq. (2) means that a finite, i.e., $\mathcal{O}(1)$, energy is sufficient to find *some* change of a finite fraction, i.e., $\mathcal{O}(N)$, of the system [39]. If this criterion is not fulfilled, we will conclude that RS holds.

But this criterion alone does not suffice. Furthermore, we have to ensure that some kind of order exists in the ground state. Consider, for example, a system, where every edge has equal length. Its solution space structure is like a paramagnet, i.e., trivial, since every tour is of identical length. On the other hand, this system also fulfills the criterion Eq. (2).

While a random tour and the optimal tour in this degenerate ensemble behave the same in every aspect, this is not true for the (1, 2)-TSP, where a random tour has $\mathcal{O}(1)$ edges of length *one* but an optimal tour has $\mathcal{O}(N)$ edges of length *one*. This distinguishes order from disorder. In more detail, our measurements show the actual number of length *one* edges for the (1, 2)-TSP is $0.4240(8)N$, thus, corresponding to an ordered ground state. We obtained this constant by using the Beardwood-Halton-Hammersley constant β, which will be scrutinized in the beginning of Sec. IV. For the (1, 2)-TSP β is the mean length of the edges constituting the tour and since all edges in the (1, 2)-TSP ensemble are of length 1 or 2, $\beta - 1$ is the fraction of length 2 edges in the optimal tour. The ETSP shows a very similar behavior [42]. We will further show that the STSP on the other hand, while fulfilling the criterion Eq. (2), behaves still trivial and the fulfillment of the criterion is caused by a high degeneracy.

Note that degeneracy alone does not mean that a solution space structure is trivial, since the degenerate solutions may be contained in one big cluster, at least in the thermodynamic limit. Famous examples, where this is the case include the two-dimensional Ising spin glass with ± 1 couplings [43] and the satisfiability problem in the range of few constraints [44].

Anyway, for the cases where we cannot rule out RSB, we cannot reach a definitive conclusion since RSB is a more complex phenomenon not only caught by one quantity of interest. However, we can identify cases which might be worthwhile to study in more detail to determine whether they are RS or RSB, or exhibit a complex behavior in another way.

Going from measurable quantities to algorithms, to solve numerically any instance of the TSP, the following integer program, i.e., an LP with additional integer constraints Eq. (6), can be used [45]:

$$\text{minimize} \qquad \sum_i \sum_{j<i} c_{ij} x_{ij}, \tag{3}$$

$$\text{subject to} \qquad \sum_j x_{ij} = 2 \quad i = 1, 2, \ldots, N \tag{4}$$

$$\sum_{i \in S, j \notin S} x_{ij} \geqslant 2 \quad \forall S \subset V, \tag{5}$$

$$x_{ij} \in \{0, 1\}, \tag{6}$$

where x_{ij} is the searched for adjacency matrix defining the tour, V is the set of all cities, and S a proper, nonempty subset of V. Equation (3) minimizes the tour length, Eq. (4) ensures that the number of incident edges into every city is two, such that the salesperson enters every city once and leaves it again. Equation (5) is the *subtour elimination constraint* (SEC), which prevents the tour to fragment into multiple not-connected subtours.

As a technical detail, we use fixed point data types for the distances. This discretization means effectively that the entries of the cost matrix are rounded and can therefore lead to different results than exact Euclidean distances, however, this is a fundamental problem of any computer simulation. Tests with different precisions did not show any systematic and notable influence on the mean values, such that we are confident that no systematic error is introduced by this choice. We use CONCORDE [46] to generate optimal tours, which implements the LP from Eqs. (3) to (6) at its core but also extends it with additional constraints and heuristics to speed up the solution process.

Note that optima found with this method are not necessarily drawn uniformly from all existing optima and we do not perform unbiased ground-state sampling. However, most of our ensembles are not degenerate anyway. And in the case of the (1, 2)-TSP, the only ensemble with many optima studied by us, we tested whether this possible bias has influence on our results. Therefore we applied random perturbations on the edge lengths to lift the degeneracy, which was tested on other models to result in uniform, unbiased sampling of the optima [47,48]. This procedure yielded within error bars the same results as the degenerate ensemble, such that we are confident that any possible bias of the optimum selection does not have considerable influence on our results.

To construct the excitations T^*, we modify the linear program formulation using the obtained optimal tour T^o. This allows us to construct excitations with very specific properties. Since we want to check the criterion Eqs. (1) and (2), we construct a very specific integer program which fixes Eq. (1) to be fulfilled and maximizes Eq. (2). If the problem is RS, the result should show the criterion to be violated.

So we fix the allowed energy difference $L^* - L^o = \epsilon$ to a constant, which will lead to the desired relative energy difference Eq. (1) if the energy is extensive. For this reason our definitions of the ensembles are formulated in a way that leads to extensive energy, i.e., $\langle L^o \rangle \sim N$. Within this excitation energy window ϵ, the number of common edges with the optimal tour T^o needs to be minimized to maximize the distance of the configurations. Thus replacing the objective

with

$$\text{minimize} \quad \sum_{\{i,j\} \in T^o} x_{ij}, \tag{7}$$

and adding the additional constraint,

$$\sum_i \sum_{j<i} c_{ij} x_{ij} \leqslant L^o + \epsilon, \tag{8}$$

results in a suitable LP. We will call this LP *MaxDiff*. Technically, we used a custom implementation of the LP. We used CPLEX [49] as the LP solver and for branch and cut. Two exemplary solutions of this LP are visualized in Fig. 1 in comparison to the optimal tours.

IV. RESULTS

We performed the calculation of optimum and excited tours for the four ensembles ETSP, RLTSP, (1,2)-TSP, and STSP, for various system sizes ranging from $N = 64$ to $N = 1448$ cities. All results are averaged over a few hundred realizations of the disorder.

Due to the hardness nature of the TSP and our exact solution approach, some realizations of the largest system sizes take far more computational resources than most realization and could not be solved in reasonable time, respectively, memory. If we just omitted the unsolved instances from our results, the results would be subject to selection bias since the hardest realizations (for the used algorithm) are excluded from the results, which can lead to systematical errors—especially since we are interested in properties linked to hardness. To ensure that the results are not tainted by such systematic errors, we perform a very conservative error estimation. The basic idea is that we determine intervals of possible values for the means, where unsolved instances enter with their minimum and maximum possible values, thus taking care of these systematic errors. As we will see in the results section, these intervals are very small, showing that the unsolved instances have no significant effect, which is visible also when comparing fits which used lower and upper possible values. Nevertheless, in detail, there are two ways in which the optimization might fail, which we treated differently. In the first case, already the optimal tour, i.e., the ground state, cannot be found for a given realization. System sizes for which this happened for at least one realization are omitted completely from our analysis. In other words, our data contains only system sizes where we always found the optimum tour for all instances of this system size and ensemble. In the second failure case, we found the optimal tour Eq. (3) for an instance, but were not able to determine the excitation Eq. (7). In this case we can often use intermediate results of the branch-and-cut procedure to estimate upper and lower bounds. The upper bound of d is always available as the solution of the LP relaxation, i.e., the solution of the LP defined by Eqs. (7), (4), (5), and (8) without the integer constraints. In fact, the bound is even tighter, as we can round down the relaxation solution to the next integer. The lower bound of d is available if the branching procedure produces an integer solution [50]; otherwise it is assumed as the minimum possible value, i.e., $d = 0$. Similarly, we can estimate bounds on the relative energy difference as it is bounded by 0 and ϵ/L^o.
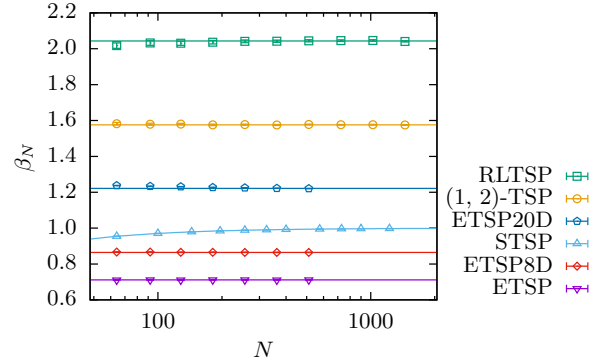

FIG. 2. Finite-size Beardwood-Halton-Hammersley constant β_N measured at different system sizes N. Except for the STSP, all lines show the mean value calculated from the results for $N \geqslant 200$. For STSP the line is a fit to $\beta_N = \beta + aN^b$ to extrapolate the asymptotic β, which yields $a = -2.62(5)$ and $b = -0.971(4)$. The extrapolated values for β are tabulated in Table I.

The range of sizes N we study for each TSP ensemble is determined by the largest size for which we could calculate always the optimum and obtain results with sensibly small variations due to the inclusion of instances where the excitation could not be obtained. We used as a criterion that not more than one instance of the second failure kind occurs which comes *without* an estimate for the lower bound. On the other hand, failures of the second kind, coming *with* upper and lower bounds are quite benign, since the bounds are typically reasonably tight, such that data points are shown, for which, for the largest sizes, up to 10% of the 100 (or 200 depending on the ensemble) samples belong to this failure category. All fits performed in the remainder of this study are done twice, once using the upper bounds and once using the lower bounds. The results are always compatible within the statistical errors. All fit results shown in the following are obtained from using the upper bound, since it is usually tighter than the lower bound.

We start the presentation of the results with the behavior of the optimum tour length. Here only data for system sizes N is included where an optimum was found for all instances of this size N. For the ETSP it is well known that the mean optimal length $\langle L^o \rangle$ through N cities placed on a unit square by a Poisson point process approaches a limit value for large N, if scaled appropriately:

$$\lim_{N \to \infty} \langle L^o \rangle / \sqrt{N} = \beta. \tag{9}$$

This constant β is the Beardwood-Halton-Hammersley constant [30] and some estimates for its value exist [31,51,52]. Similarly, such a constant should exist for the random link model. For the pseudo-one-dimensional case, it is even known exactly [53]. For the STSP and (1,2)-TSP the authors are not aware of previous work, but it is easy to recognize that the optimal tour in the STSP traverses N horizontal or vertical edges, which should have each a length of 1 for large N. Thus, we expect the corresponding constant to be $\lim_{N \to \infty} \langle L^o \rangle / N = 1$.

Comparing these expectations to our data serves as a crosscheck to establish some level of confidence in our data. For the (1,2)-TSP case these are novel results. In Fig. 2 the

TABLE I. Beardwood-Halton-Hammersley constants β for different ensembles of the TSP determined from our data and the current best estimates for their actual values. Note that we only need to find the ground state and not the excitation for this analysis such that we can show larger system sizes than in later results. The literature values marked by an asterisk * are values according to a high dimensional limit conjecture for the RLTSP, which should coincide in this limit with the ETSP [34]. Therefore, a perfect agreement for the finite dimensions simulated is not expected.

	β (measured)	β (literature)
ETSP, 2 dim.	0.7112(6)	0.712403(7) [52]
ETSP, 8 dim.	0.8645(3)	0.8531* [34]
ETSP, 20 dim.	1.2218(2)	1.2093* [34]
RLTSP, 1 dim.	2.044(3)	2.0415.. [53]
STSP	1.0005(2)	1
(1, 2)-TSP (ER, $p = \frac{1}{N}$)	1.5760(8)	–

rescaled mean optimal tour lengths β_N are plotted. Generally the finite-size effects are small, such that we determine estimates for β simply as the average of all data points $N \geqslant 200$. Except for the STSP this works reasonably well. Since the STSP shows the largest finite-size effect, we use an offsetted power law $\beta_N = \beta + aN^b$ to extrapolate the measurements. The results of this analysis are shown in Table I together with the currently best known values for this constant. Since we do not have a good model to extrapolate the values, the given error bars are only statistical and do not account for errors in the extrapolation. Considering this, our estimates are reasonably close to the expectations.

Next, the results for the MaxDiff excitation simulations for the two-dimensional ETSP are shown in Fig. 3. We chose $\epsilon = 3.16$ for ETSP in two dimensions, $\epsilon = 1$ in higher dimensions and $\epsilon = 30$ for the RLTSP. Note that these choices are motivated by the different typical optimal lengths of the tours and chosen to minimize finite-size effects. That means ϵ should be small enough to not allow every single edge to be different, which would clamp the data points for every size to $d = 1$. On the other hand ϵ should be large enough to allow for all simulated system sizes some edges to be different.

However, the choice of ϵ should have no influence on the principal result, and will lead mainly to a shift up or down in the logarithmical scale. A more detailed analysis of the influence of the value of ϵ is shown later for the (1, 2)-TSP.

We found a $1/N$ behavior of the relative energy difference (inset) as required by Eq. (1). Nevertheless the difference d of the tours also vanishes in the large N limit as a power law, thus the behavior of Eq. (2) is not found. Therefore, to change a finite fraction of an infinite system, a finite energy ϵ does not suffice for this ensemble. Thus, the results do not show the signature of replica symmetry breaking, hinting at a trivial solution space structure. This is consistent with previous studies [21,23,24,34,35] using the RLTSP as an approximation for the ETSP. They used, e.g., the cavity method to estimate some properties and compared them to tours obtained by heuristics (for smaller system sizes N) leading to the claim that the ETSP is replica symmetric. Also our results actually for the RLTSP, shown in Fig. 3, lead to the same conclusion and confirm the previous results.
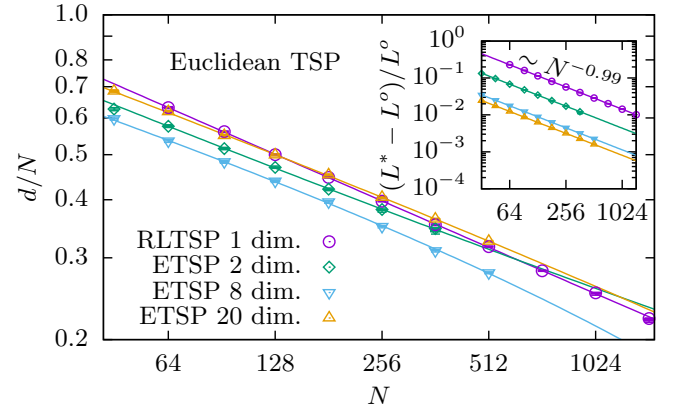

FIG. 3. The relative difference d/N of the optimum and the MaxDiff excitation decreases as a power law with the system size N. Its exponent depends on the ensemble. For large N the difference d/N vanishes which is a hint for replica symmetry and a trivial solution space structure. The inset shows that the premise Eq. (1) is fulfilled. The bounds of the value are visualized as filled boxes and the statistical errors as error bars; note that both are always smaller than the symbols. The smallness (hardly visible) of the boxes indicate that the few instances where only bounds for the distance d could be obtained have basically no influence as the relative difference between upper and lower bound is always less than 2%. The exponent and offset, which is always compatible with 0, are obtained by fits to $\frac{d}{N} = aN^b + D^\infty$ and shown in Table II.

For spin glasses, the energy landscape becomes complex and exhibits many features of RSB in high enough dimensions. Above the upper critical dimension the system is believed to behave [54–57] like the mean-field SK model [1,2], corresponding to RSB. This motivated us to investigate the ETSP for high dimensions as well. Our results for the eight-dimensional and 20-dimensional ETPS are also shown in Fig. 3. Evidently, even for higher dimensions the same behavior indicates that RS is present. Thus a simple increase in dimensionality does apparently not change the behavior regarding replica symmetry much. This is in strong contrast to spin glasses.

Next, we will look at an ensemble which is closer to a direct mapping from the Hamilton circuit, which is usually used to prove the TSP NP-complete. The mapping creates an instance of the (1, 2)-TSP. For three tested values of the finite excitation energy $\epsilon \in \{10, 20, 30\}$, we calculated the difference between the optimal and excited tours d, shown in Fig. 4.

First (see inset), the relative energy difference decreases as $1/N$ as required by Eq. (1). The measured difference d does not follow a pure power law, but seems to converge to a nonzero offset. Extrapolating the difference for large N with $\frac{d}{N} = aN^b + D^\infty$ (cf. Ref. [39]) leads to offsets for each value of ϵ, which are reasonably close to the most accurate value we obtained $D^\infty = 0.645(2)$ and exponents close to $b = -1$. All values are shown in Table II. Note that for small system sizes N finite-size effects are visible, where ϵ is of the order of the optimal length and the excitation can differ in every single edge. Therefore, the difference is clamped at $d/N = 1$. For larger system sizes N this does not seem to play a role
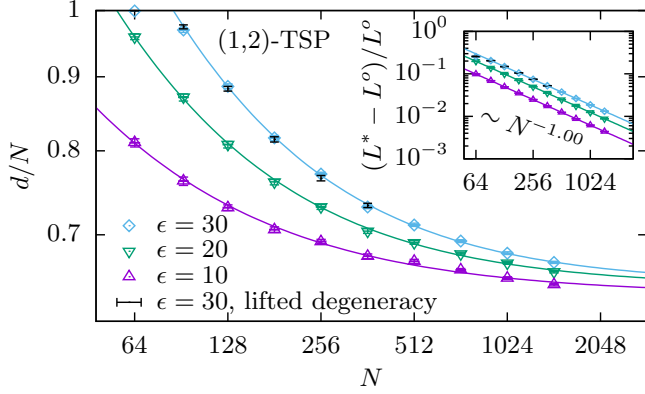
FIG. 4. Statistics of the $(1, 2)$-TSP for a connectivity of $Np = 1$. The MaxDiff constraints with the finite excitation energy $\epsilon \in \{10, 20, 30\}$ are used for the three curves, respectively. The distance of the excitation to the optimal tour is extrapolated with an offsetted power law ansatz $\frac{d}{N} = aN^b + D^\infty$. The fit parameters are shown in Table II. All three result in a convergence to a finite D^∞ for large N, i.e., a finite fraction. The possibility of RSB can therefore not be excluded. The inset shows the relative energy difference of the optimum and the excitation, showing nearly a perfect $1/N$ form, as required by the RSB criterion. The bounds of the value are visualized as filled boxes and the statistical errors as error bars; note that both are always smaller than the symbols. The best estimate, i.e., the upper bound, is used for fits.

anymore. To reduce the influence of this finite-size effect, the fits for larger ϵ disregard the small system sizes $N < 128$ for $\epsilon = 30$ and $N < 64$ for $\epsilon = 20$. In particular, different values of ϵ lead to consistent results. According to the criterion Eq. (2) our results cannot exclude the possibility that replica symmetry is actually *broken* for this ensemble.

To further test these results, we conducted simulations above the percolation threshold, for $p = 3/N$, and below the threshold for $p = 1/2N$. The results exhibit qualitatively the same behavior (not shown), but with different values of the asymptotic D^∞. Apart from the limits $p \to 0$ and $p \to 1$, where every tour is optimal, the precise structure of the graph does not seem to have a critical influence on this result.

To exclude that the degeneracy has a special influence on our results, we lift the degeneracy by adding a slight perturbation on each edge. Therefore we scale the edge weights and

TABLE II. Values of the fit parameters extrapolating the behavior of d/N. Interestingly all ensembles converging to a finite value of D^∞ show an exponent close to $b = -1$.

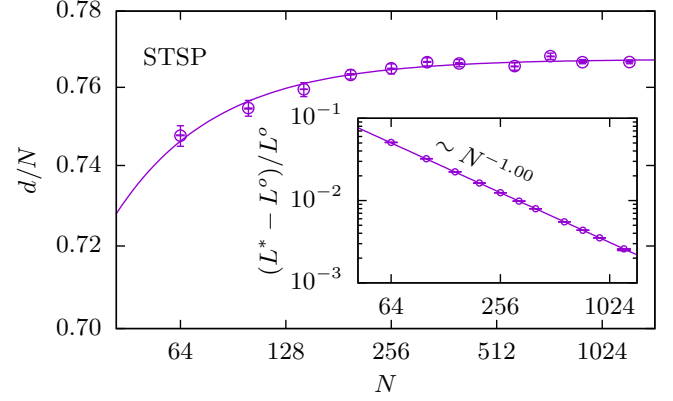|	b	D^∞	RS
ETSP, 2 dim.	$-0.32(8)$	$0.04(10)$	✓
ETSP, 8 dim.	$-0.21(4)$	$-0.18(9)$	✓
ETSP, 20 dim.	$-0.27(3)$	$-0.05(5)$	✓
RLTSP, 1 dim.	$-0.336(12)$	$-0.004(13)$	✓
STSP,	$-1.5(5)$	$0.767(1)$	Degenerate
$(1, 2)$-TSP, $\epsilon = 10$	$-0.82(2)$	$0.636(2)$	RSB possible
$(1, 2)$-TSP, $\epsilon = 20$	$-0.93(3)$	$0.645(2)$	RSB possible
$(1, 2)$-TSP, $\epsilon = 30$	$-0.97(4)$	$0.648(3)$	RSB possible
$c_{ij} = 1$		1	Degenerate

FIG. 5. For the STSP the relative difference d/N converges to a finite value, which means that finite energy is sufficient to change a macroscopic part of the system. The value it converges to is estimated by an offsetted power-law ansatz $\frac{d}{N} = aN^b + D^\infty$ and fulfills the criterion Eq. (2). The exponent and offset are shown in Table II. The inset shows that Eq. (1) is fulfilled. The bounds of the value are visualized as filled boxes and the statistical errors as error bars; note that the bounds are always smaller than the symbols. The best estimate, i.e., the upper bound, is used for fits.

ϵ by 10^5 and add a random disturbance $U(-10, 10)$ to each edge. Except for a vanishing degeneracy, this procedure also does not change the results beyond statistical errors, which are indicated as additional black error bars for a selection of data points for $\epsilon = 30$ in Fig. 4.

The last ensemble we study is the STSP, which is a very special subspace of the ETSP configuration space. The STSP, where cities are placed on a square lattice and are displaced by a distance proportional to $1/N$, does show a qualitatively very different behavior to the ETSP. In contrast to the ETSP case, the difference does not follow a pure power law, but seems to converge to a nonzero offset. But different than the $(1, 2)$-TSP case, it approaches the limiting value from below. In Fig. 5 this behavior is fitted with an offsetted power law $\frac{d}{N} = aN^b + D^\infty$. While according to criterion Eq. (2) this is not compatible with the trivial behavior of RS, it is rather easy to see that this is an effect caused by high degeneracy. In fact, large realizations basically look like a square lattice, where many tours which do not use diagonals have almost the same lengths. This is compatible with the value of β which is apparently 1 (cf. Table I). The slight displacements avoid perfect degeneration, but are not strong enough to destroy this effect and thus no ordered phase can be observed. Thus, the system behaves like a paramagnet, where many solutions are indistinguishably close in energy. Note that a displacement by a fixed amount, e.g., 5% of the lattice constant, does lead to the same trivial behavior as the ETSP before (not shown). The same is true for a diluted square lattice, where a fixed fraction of sites is removed (also not shown). We therefore conclude that the energy landscape is most likely trivial.

V. CONCLUSION

To summarize, we studied multiple ensembles of the TSP by applying sophisticated exact combinatorial optimization algorithms in extensive simulations. As suspected before,

we find evidence for the replica symmetry of the Euclidean TSP and the related random link model. Interestingly, we find these results also in high space dimensions, in contrast to spin glasses where RSB is believed to appear above the upper critical dimension $d_u = 6$. Our results strengthen the conjecture that replica symmetry holds for these ensembles, which is often used to tackle this problem from a statistical mechanics point of view.

On the other hand, for the (1,2)-TSP, inspired by the classical mapping of the Hamilton circuit to the TSP, we cannot exclude replica symmetry breaking. Thus, we provide the first evidence for a complex phase-space behavior of this classical NP-hard optimization problem. This should motivate further studies to find out whether the solution space is clustered and whether replica symmetry breaking might actually be present.

For future work, especially for the degenerate case of the (1,2)-TSP, it would be interesting to study the solution space structure with a focus on clustering. One could define a neighborhood relationship in the configuration space, e.g., k-opt moves [58], and search for clusters of configurations which can be reached from each other by paths traversing only neighboring instances [8,59–61].

The linear programming approach we use is very general and can be applied to a large range of problems. Since for many problems mappings to integer programs are already known and it is quite straight forward to formulate additional constraints enforcing some specific excitations, this technique could be quite generally used to explore a very specific range of the energy landscape of many problems.

[1] G. Parisi, Phys. Rev. Lett. **43**, 1754 (1979).

[2] G. Parisi, Phys. Rev. Lett. **50**, 1946 (1983).

[3] D. L. Stein, in *Decoherence and Entropy in Complex Systems: Selected Lectures from DICE 2002*, edited by H.-T. Elze (Springer, Berlin/Heidelberg, 2004), pp. 349–361.

[4] M. Talagrand, Annal. Math. **163**, 221 (2006).

[5] S. A. Cook, in *Proceedings of the Third Annual ACM Symposium on Theory of Computing* (ACM Press, New York, 1971), pp. 151–158.

[6] R. M. Karp, *Reducibility among Combinatorial Problems* (Springer, Berlin, 1972).

[7] S. Mertens, Comput. Sci. Eng. **4**, 31 (2002).

[8] A. Montanari, F. Ricci-Tersenghi, and G. Semerjian, J. Stat. Mech. (2008) P04004.

[9] M. Weigt and A. K. Hartmann, Phys. Rev. E **63**, 056127 (2001).

[10] A. K. Hartmann and M. Weigt, *Phase Transitions in Combinatorial Optimization Problems: Basics, Algorithms and Statistical Mechanics* (John Wiley & Sons, New York, 2006).

[11] C. Moore and S. Mertens, *The Nature of Computation* (Oxford University Press, Oxford, 2011).

[12] M. Mézard and A. Montanari, *Information, Physics and Computation* (Oxford University Press, Oxford, 2009).

[13] S. Kirkpatrick and G. Toulouse, J. Phys. **46**, 1277 (1985).

[14] P. Cheeseman, B. Kanefsky, and W. M. Taylor, in *Proceedings of the 12th International Joint Conference on Artificial Intelligence* (Morgan Kaufmann, San Francisco, 1991), Vol. 1, pp. 331–337.

[15] I. P. Gent and T. Walsh, Artif. Intell. **88**, 349 (1996).

[16] A. K. Hartmann and M. Weigt, J. Phys. A: Math. Gen. **36**, 11069 (2003).

[17] K. Smith-Miles, J. van Hemert, and X. Y. Lim, in *International Conference on Learning and Intelligent Optimization* (Springer, Berlin, 2010), pp. 266–280.

[18] T. Dewenter and A. K. Hartmann, Phys. Rev. E **86**, 041128 (2012).

[19] H. Schawe and A. K. Hartmann, Europhys. Lett. **113**, 30004 (2016).

[20] C. H. Papadimitriou, Theor. Comput. Sci. **4**, 237 (1977).

[21] M. Mézard and G. Parisi, J. Phys. **47**, 1285 (1986).

[22] M. Mézard and G. Parisi, Europhys. Lett. **2**, 913 (1986).

[23] N. Sourlas, Europhys. Lett. **2**, 919 (1986).

[24] W. Krauth and M. Mézard, Europhys. Lett. **8**, 213 (1989).

[25] A. K. Hartmann, *Big Practical Guide to Computer Simulations* (World Scientific, Singapore, 2015).

[26] A. K. Hartmann and H. Rieger, *Optimization Algorithms in Physics* (Wiley-VCH, Weinheim, 2001).

[27] C. H. Papadimitriou and M. Yannakakis, Math. Oper. Res. **18**, 1 (1993).

[28] K. Menger, Monatsh. Math. Physik **38**, 17 (1931).

[29] W. Cook, *In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation* (Princeton University Press, Princeton, 2012).

[30] J. Beardwood, J. H. Halton, and J. M. Hammersley, in *Mathematical Proceedings of the Cambridge Philosophical Society* (Cambridge University Press, Cambridge, 1959), Vol. 55, pp. 299–327.

[31] A. G. Percus and O. C. Martin, Phys. Rev. Lett. **76**, 1188 (1996).

[32] S. Arora, J. ACM (JACM) **45**, 753 (1998).

[33] M. R. Garey, R. L. Graham, and D. S. Johnson, in *Proceedings of the Eighth Annual ACM Symposium on Theory of Computing*, STOC '76 (ACM Press, New York, 1976), pp. 10–22.

[34] N. J. Cerf, J. B. de Monvel, O. Bohigas, O. C. Martin, and A. G. Percus, J. Phys. I France **7**, 117 (1997).

[35] A. G. Percus and O. C. Martin, J. Stat. Phys. **94**, 739 (1999).

[36] P. Erdős and A. Rényi, Publ. Math. Inst. Hungar. Acad. Sci. **5**, 17 (1960).

[37] G. Reinelt, ORSA J. Comput. **3**, 376 (1991).

[38] D. Applegate, R. Bixby, W. Cook, and V. Chvátal, Documenta Mathematica **Extra Volume ICM III**, 645 (1998).

[39] M. Palassini and A. P. Young, Phys. Rev. Lett. **85**, 3017 (2000).

[40] M. Zumsande and A. K. Hartmann, Eur. Phys. J. B **72**, 619 (2009).

[41] M. Zumsande, M. J. Alava, and A. K. Hartmann, J. Stat. Mech. (2008) P02012.

[42] J. Vannimenus and M. Mézard, J. Phys. Lett. **45**, 1145 (1984).

[43] G. Hed, A. K. Hartmann, D. Stauffer, and E. Domany, Phys. Rev. Lett. **86**, 3148 (2001).

[44] R. Monasson and R. Zecchina, Phys. Rev. E **56**, 1357 (1997).

[45] G. Dantzig, R. Fulkerson, and S. Johnson, J. Oper. Res. Soc. Am. **2**, 393 (1954).

[46] D. Applegate, R. Bixby, V. Chvátal, and W. Cook, Math. Program. **97**, 91 (2003).

[47] C. Amoruso and A. K. Hartmann, Phys. Rev. B **70**, 134425 (2004).

[48] S. von Ohr and A. K. Hartmann, Phys. Rev. E **98**, 012108 (2018).

[49] IBM, IBM ILOG CPLEX Optimization Studio (IBM, Armonk, 2013).

[50] To find an integer solution during the branching, a strategy with an emphasis on the integrality constraints instead of optimality should be used. Integer programming libraries often allow one to choose such a strategy, e.g., in CPLEX with the *MIPEmphasis* parameter or in Gurobi with the *MIPFocus* parameter. We also observed that this focus does lead typically not only to nonoptimal integer solutions, which can be used as bounds, but also to faster termination of the algorithm with an optimal integer solution.

[51] J. L. Jacobsen, N. Read, and H. Saleur, Phys. Rev. Lett. **93**, 038701 (2004).

[52] D. Applegate, W. Cook, D. Johnson, and N. Sloane, Using large-scale computation to estimate the Beardwood-Halton-Hammersley TSP constant, conference presentation, XLII Brazilian Symposium on Operations Research (2010), http://w3.ufsm.br/42sbpo/material/sbpo_bbh.ppt

[53] J. Wästlund, Acta Math. **204**, 91 (2010).

[54] A. B. Harris, T. C. Lubensky, and J.-H. Chen, Phys. Rev. Lett. **36**, 415 (1976).

[55] H. G. Katzgraber and A. P. Young, Phys. Rev. B **72**, 184416 (2005).

[56] H. G. Katzgraber, D. Larson, and A. P. Young, Phys. Rev. Lett. **102**, 177205 (2009).

[57] M. A. Moore and A. J. Bray, Phys. Rev. B **83**, 224408 (2011).

[58] S. Lin, Bell Syst. Tech. J. **44**, 2245 (1965).

[59] A. K. Hartmann, Phys. Rev. E **63**, 016106 (2000).

[60] W. Barthel and A. K. Hartmann, Phys. Rev. E **70**, 066120 (2004).

[61] A. K. Hartmann, A. Mann, and W. Radenbach, J. Phys.: Conf. Ser. **95**, 012011 (2008).