Replica symmetry breaking for Ulam's problem

Phil Krabbe,[1] Hendrik Schawe,[2] and Alexander K. Hartmann[1]
[1]*Institut für Physik, Universität Oldenburg, D-26111 Oldenburg, Germany*
[2]*Laboratoire de Physique Théorique et Modélisation, UMR-8089 CNRS-Université Cergy Pontoise, Cergy-Pontoise, France*

(Received 20 September 2022; revised 7 February 2023; accepted 13 February 2023; published 24 February 2023)

We study increasing subsequences (ISs) for an ensemble of sequences given by a permutation of numbers $\{1, 2, \ldots, n\}$. We consider a Boltzmann ensemble at temperature T. Thus each IS appears with the corresponding Boltzmann probability where the energy is the negative length $-l$ of the IS. For $T \to 0$, only ground states, i.e., the longest IS (LIS), contribute, also called Ulam's problem. We introduce an algorithm which allows us to directly sample ISs in perfect equilibrium in polynomial time, for any given sequence and any temperature. Thus, we can study very large sizes. We obtain averages for the first and second moments of the number of ISs as functions of n and confirm analytical predictions. Furthermore, we analyze for low temperature T the sampled ISs by computing the distribution of overlaps and performing hierarchical cluster analyses. In the thermodynamic limit $n \to \infty$ the distribution of overlaps stays broad and the configuration landscape remains complex. Thus, Ulam's problem exhibits replica symmetry breaking (RSB). This means it constitutes a model with complex equilibrium behavior which can be studied numerically exactly in a highly efficient way. This is in contrast to other models, where RSB becomes exponentially irrelevant for the equilibrium behavior in the thermodynamic limit, such as in a random exclusive-or satisfiability (XORSAT) problem, or models where RSB remains relevant, such as spin glasses or NP-hard optimization problems, but where no fast exact algorithms are known.

DOI: 10.1103/PhysRevB.107.064208

I. INTRODUCTION

The mathematician Stanisław Ulam was also a well-known pioneer in computer simulations. One of the problems he studied numerically, back in the 1950s, was [1] the scaling of the length L of the longest increasing subsequence (LIS) [2] of random permutations of n numbers. Based on the knowledge [3] that the average length increases at least as the square root of n, he proposed that the average length scales as $\langle L \rangle = c\sqrt{n}$ with $c \approx 1.7$. In the meantime, $\langle L \rangle = 2\sqrt{n}$ for $n \to \infty$ has been proven [4]. Also, the distribution $P(l)$ of maximum lengths has been studied analytically [5–7] and it was found [8] that the central part is given by the Tracy-Widom distribution. This was confirmed numerically by large-deviation simulations [9], and also considered for other sequence ensembles [10]. Furthermore, the expectation values of the number of increasing subsequences (ISs) of a certain length [11] and of all ISs [12] have been obtained analytically. The actual distribution of the number of LISs was obtained numerically over a large range of the support again by applying large-deviation algorithms [13].

As a tool, the calculation of LIS finds also applications outside mathematics, such as in data analysis [14], financial fraud detection [15], or sequence alignment in bioinformatics [16].

In spite of these connections to many fields, to our knowledge, the behavior of IS and LIS was studied so far surprisingly only with respect to the length and to the exponentially growing number of increasing subsequences. Thus, the actual structure of the exponentially large IS configuration space still needs to be studied.

Such questions about the phase-space organization lie at the heart of the statistical mechanics of complex systems such as glasses, spin glasses, and machine learning or optimization problems [17–25]. One is interested whether the configuration space is rather simple, as for a ferromagnet, often referred to as *replica symmetric*, or whether it is complex with a hierarchical organization of phase space, referred to as *replica symmetry breaking* (RSB), as it appears for mean-field spin glasses [26].

In most cases, analytical solutions cannot be obtained, so one has to use computer simulations [27]. Unfortunately, all standard models which exhibit a complex RSB-like behavior, such as spin glasses, are numerically very hard to treat. Hence, only rather limited system sizes could be considered when performing equilibrium sampling, even when using special parallel computers such as JANUS [28]. Note that for combinatorial optimization problems such as the satisfiability problem, in some cases efficient algorithms exist [29]. But they allow only to find *some* solution, i.e., this sampling is not controlled. Thus, these algorithms do not allow us to sample the configuration space in equilibrium which is necessary to study the configuration-space structure. For this purpose one has to use Monte Carlo Markov-chain sampling, which requires equilibration and is slow therefore. Note, the exclusive-or satisfiability (XORSAT) problem can be exactly sampled in polynomial time [30–32], but only when the energy is strictly zero. The model exhibits one-step RSB, but due to exponentially many clusters, it leads to a trivial

equilibrium distribution of overlaps in the thermodynamic limit. Still, RSB becomes visible in the nonequilibrium behavior of XORSAT. For directed polymers in random media, a fast polynomial sampling is possible everywhere [33–35]. Here, a complex RSB equilibrium behavior was found [36], but only for ensembles with correlations in the disorder.

Here, we introduce an algorithm which allows one to count the number of ISs for any given length l as well as sampling ISs exactly for any given distribution which depends only on the IS length l, in particular for any given length the sampling is uniform. Both the calculation of the numbers and the sampling can be performed in polynomial time, which allows us to treat large systems exactly. We study the sequence ensemble of random permutations, which does not exhibit correlations and is the classical and most-studied ensemble for IS and LIS. Our results indicate that the structure of the configuration space exhibits properties of replica symmetry breaking, i.e., a broad distribution of overlaps and a hierarchical clustering of configurations, even in the thermodynamic limit $n \to \infty$.

Next, we present all necessary definitions and introduce the algorithms. Then we show our results and finish by a summary and discussion.

II. MODEL AND ALGORITHMS

Let $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ be a sequence of n distinct numbers. A subsequence $\lambda = (\sigma_{i_1}, \sigma_{i_2}, \ldots, \sigma_{i_l})$ of length $l = l(\lambda)$ fulfills $1 \leqslant i_1 < i_2 < \cdots < i_l \leqslant n$ and is called *increasing* if $\sigma_{i_j} < \sigma_{i_{j+1}}$ for all $j = 1, \ldots, l-1$. To calculate the longest among all possible ISs, the *patience sort algorithm* [4] is a popular choice which runs in polynomial time. Recently, an extension was proposed [13], which allows one to calculate the number of LISs. Here, we introduce a further extension and variant of the algorithm, which enables one to count all ISs and sample them efficiently and exactly for any desired probability distribution which depends on the IS length l.

Let H be a *precedence matrix*, which encodes possible joint occurrences of entries σ_i and σ_j in an IS λ, i.e.,

$$H_{ij} = \begin{cases} 1 & \text{if } i < j \text{ and } \sigma_i < \sigma_j, \\ 0 & \text{else.} \end{cases} \tag{1}$$

Note that the matrix H can be efficiently stored as a graph with neighbor lists. To set up H, we run an extended variant of patience sort, which gives also the length L of the LIS and allows us to restrict the number of candidates i, j which have to be checked whether one has to assign $H_{ij} = 1$. Still, this requires $O(n^2)$ steps.

To count IS and LIS, we denote by Ψ_j^l the number of ISs of length l which end at position j. Clearly, each single entry of σ represents an IS of length $l = 1$, i.e., we have $\Psi_i^1 = 1$ for $i = 1, \ldots, n$. Now, ISs of length $l > 1$ can be constructed by selecting a final entry σ_j precedent by an IS of length $l-1$ where all entries are smaller than σ_j and appear before position j. For the number of ISs this turns into

$$\Psi_j^l = \sum_{i<j} H_{ij} \Psi_i^{l-1} \quad \text{for } l = 2, \ldots, L, \tag{2}$$

which can be computed in a convenient way recursively, i.e., by *dynamic programming* in $O(n^2 L)$. The total number of ISs of length l is given by $\Psi^l = \sum_j \Psi_j^l$, where we also include the empty subsequence $\Psi^0 = 1$. The total number of ISs is given by $\Psi = \sum_{l=0}^{L} \Psi^l$.

To sample an IS for given length l, one starts by sampling the final entry j which appears with probability Ψ_j^l / Ψ^l. Next, the preceding entry i is sampled among all possible predecessors $i < j$, i.e., where $H_{ij} = 1$. Each possible entry i is selected with probability Ψ_i^{l-1} / Ψ_j^l. This is continued iteratively for length $l-2, l-3$, etc., always given the just sampled entry, until length 0 is reached. This algorithm takes $O(n\,l)$ steps.

The sampling can be easily extended to include any probability which depends on the length. Here, we take a physical viewpoint by considering $E = -l$ as energy within the canonical ensemble at temperature T, i.e., by using probabilities $\sim \exp(l/T)$. Thus, for an IS λ we have the probability given by

$$p(\lambda) = \exp[l(\lambda)/T]/Z, \quad Z = \sum_l \Psi^l \exp(l/T). \tag{3}$$

This includes in particular all LISs for $T \to 0$. Sampling an IS now consists of first drawing a length l according the probabilities $\Psi^l \exp(l/T)/Z$, and then uniformly sampling an IS of length l as explained before. Note that this approach is exact and direct, i.e., for each run of the algorithm an independently sampled configuration is returned. Our approach runs in polynomial time, such that we can treat rather large systems in perfect equilibrium.

III. RESULTS

We performed simulations [27] for ensembles of permutations of n numbers in the range $n = 128$ to $n = 8192$. We studied for all sizes 10 000 realizations of the disorder, i.e., independent permutations. For comparison, we also considered in some cases the ordered sequence $\sigma^o = (1, 2, \ldots, n)$.

We start by considering the number Ψ of IS. The asymptotic behavior of the expectation value is analytically given by [12]

$$E(\Psi) = \frac{1}{2\sqrt{\pi e}} n^{-1/4} \exp(2\sqrt{n}). \tag{4}$$

In Fig. 1 we compare the numerical average $\langle \Psi \rangle$ with the analytical result and find very good agreement, even for rather small system sizes. Note that the average is "annealed" in the sense that is represents an exponentially growing quantity, such that sequences with exceptionally large values of Ψ will dominate. This means we need a rather large number of samples to observe agreement, as we do. This also indicates the correctness of our approach. We have also evaluated the second moment $\langle \Psi^2 \rangle$ (not shown). Here, the agreement with the analytical result [12] is fair, i.e., a bit lower, due to the even stronger dominance of exponentially large but exponentially rare sequences. To find a good agreement here, one would have to obtain the distribution $P(\Psi)$ down to the tails. This should be possible by using a large-deviation approach, as it has been used to obtain the distribution of the number of LISs [13], but lies outside the scope of the present study.

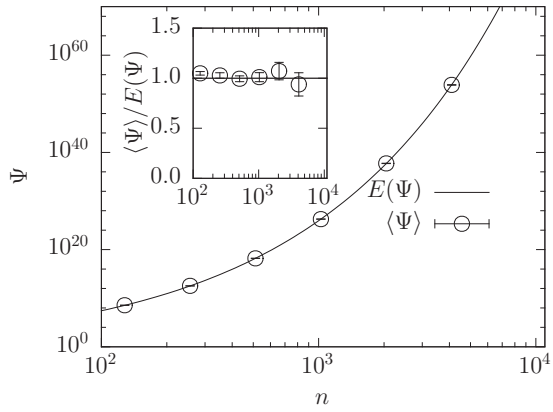

FIG. 1. The number Ψ of ISs as function of sequence length n, for the analytical expectation value $E(\Psi)$ and the numerical average $\langle\Psi\rangle$. The inset shows the ratio $\langle\Psi\rangle/E(\Psi)$, the horizontal line is at value 1.

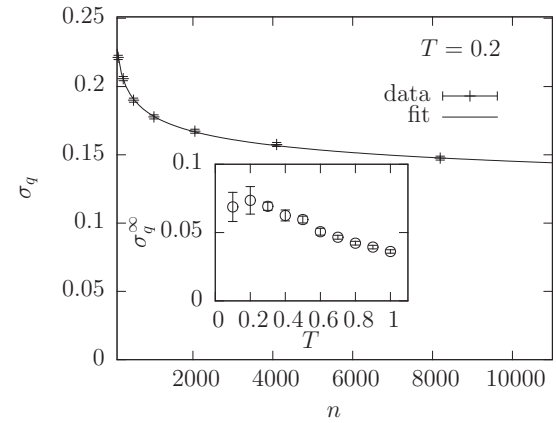

FIG. 3. Width $\sigma_q(n)$ of $P(q)$ at $T = 0.2$. The line shows a fit of the data to a power law with offset $\sigma_q(n) = \sigma_q^\infty + an^{-b}$. The inset shows the extrapolated value $\sigma_q^\infty = \lim_{n\to\infty}\sigma_q(n)$ as a function of the temperature.

Next, we analyze ISs sampled in equilibrium according to Eq. (3) at a low temperature $T = 0.2$ for several sequence lengths n. For independently sampled pairs $\lambda^{(1)}$, $\lambda^{(2)}$ of IS, we calculate the similarity of the two ISs via the *overlap* q. Here, considering the ISs as sets of the contained numbers, we use the Jaccard index [37] as given by

$$q = \frac{|\lambda^{(1)} \cap \lambda^{(2)}|}{|\lambda^{(1)} \cup \lambda^{(2)}|}. \tag{5}$$

This definition takes care of the fact that ISs can have varying length. Also, since most entries of the sequence σ are not part of any IS λ, using an overlap based on the Hamming distance would give rise to all ISs being very similar to each other, irrespective of the underlying configuration space.

The distributions $P(q)$ of overlaps, where for each of the 10 000 realizations 3000 independent pairs of configurations were sampled, are shown for $T = 0.2$ in Fig. 2 for three sequence sizes n. Apparently the distribution is broad, even for large systems, indicating a complex configuration landscape.

Although the results exhibit broad distributions of overlaps for rather large system sizes, the relevant question is what happens in the thermodynamic limit $n \to \infty$. As it is standard for numerical simulation, we perform a finite-size scaling extrapolation. Therefore, to investigate whether this signature of RSB is visible in the thermodynamic limit, we have evaluated the width σ_q of the distribution as function of system size n. The result for $T = 0.2$ is shown in Fig. 3. We fitted a power law $\sigma_q(n) = \sigma_q^\infty + an^{-b}$ and obtained the limiting value $\sigma_q^\infty = 0.073(10)$, which is significantly different from zero and shows that the distribution remains broad in the thermodynamic limit. This is in contrast to replica symmetric systems, e.g., ferromagnets, where the distribution of overlaps converges to one delta peak, or two peaks due to a trivial symmetry.

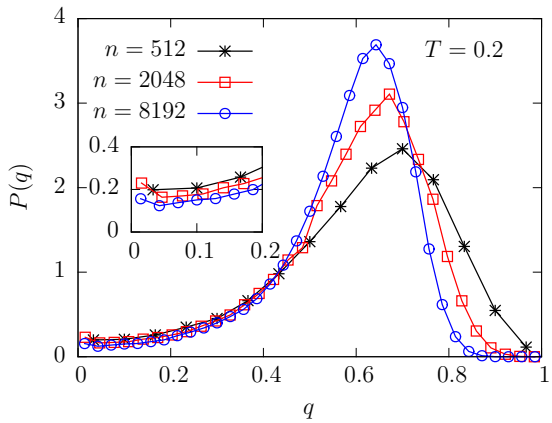

FIG. 2. Distribution of overlaps at $T = 0.2$, for three system sizes n. The inset shows a magnification of the region near $q = 0$. The lines are guides to the eyes only.

We also analyzed the integrated weight $W_{q_1,q_2} = \int_{q_1}^{q_2} P(d)dq$ in intervals $[q_1, q_2]$. While near $q = 0$ the weight seems to decrease (see the inset of Fig. 2), in the intervals [0.4,0.5], [0.5,0.6], and [0.6,0.7] we observe an increase of the weights with increasing sequence length n, which indicates a convergence to a nonzero value. This also speaks in favor of a broad distribution $P(q)$ for $n \to \infty$. For the random XORSAT problem, which exhibits one-step RSB, the limiting equilibrium $P(q)$ has zero width, i.e., appears trivial, due to the exponentially growing number of clusters [30–32].

When evaluating the *limiting*σ_q^∞ as function of T (see the inset of Fig. 3), a small peak near $T = 0.2$ and a rather smooth decrease beyond are visible. Thus, there is no sign of a transition, where one would expect a power-law decrease $\sim L^{-\eta}$, i.e., $\sigma_q^\infty = 0$ at and beyond the transition. Since we measured up to $T = 1$, i.e., five times the temperature of the peak, we expect that the crossoverlike behavior continues even for larger temperatures. We also observe this behavior for the average overlap (not shown here).

We also obtained the specific heat via the variance of the length by calculating $C = (\langle l^2\rangle - \langle l\rangle^2)/(nT^2)$. Interestingly, the disorder-averaged $C(T)$, see Fig. 4, exhibits peaks near $T = 0.4$ for all system sizes, but the peak height decreases with growing system size n. Thus, this behavior provides also no sign for a phase transition. Note that a nongrowing peak

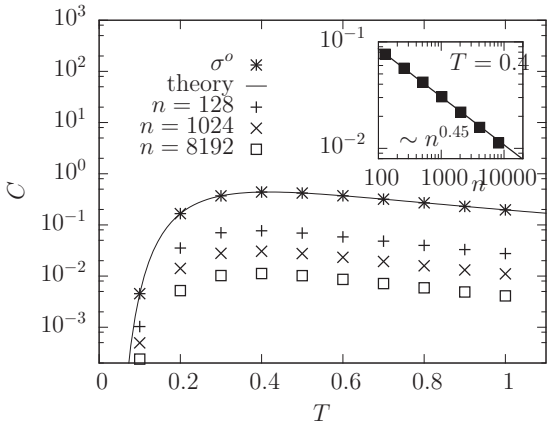

FIG. 4. The specific heat C as function of temperature T, for the ordered sequence σ^0, where the line is the analytic result, and averaged for the permutations. Shown are some system sizes n. The inset shows the scaling of the value near the peak for the permutations as function of system size, following a power law n^{-y}, shown as a straight line, with a fit value $y = 0.451(5)$.

is obtained also when one considers just the single ordered sequence σ^o, where each number can independently be part of an IS with probability $p(T) = e^{1/T}/(e^{1/T} + 1)$. Thus, σ^o represents n independent paramagnets in a field, where the variance of the length is just the sum of the single-number variances $p(T)[1 - p(T)]$ and therefore $C(T) = p(T)/[1 - p(T)]/T^2$ is readily available [38]. This $C(T)$ exhibits also a peak at the same temperature $T \approx 0.4$. Thus, from the energetic point of view no phase transition is visible. This may be similar to spin glasses, where the transition to the RSB phase is also not visible when studying the specific heat [19].

The configuration-space structure was further analyzed by applying the *agglomerative clustering approach* of Ward [39]. The hierarchical structure obtained by the clustering can be visualized by a tree, usually called *dendrogram*, where each branching corresponds to a subspace of configurations (see Fig. 5). The sequence of configurations as located in the leaves defines a partial order. This order can be used to display the matrix of the pairwise overlaps or distances where the order of

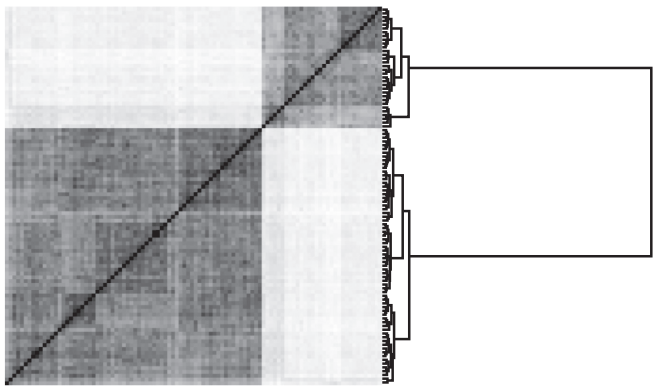

FIG. 5. Clustered overlap matrix and dendrogram obtained by clustering 200 equilibrium ISs sampled at $T = 0.2$ for one random permutation with length $n = 8192$.

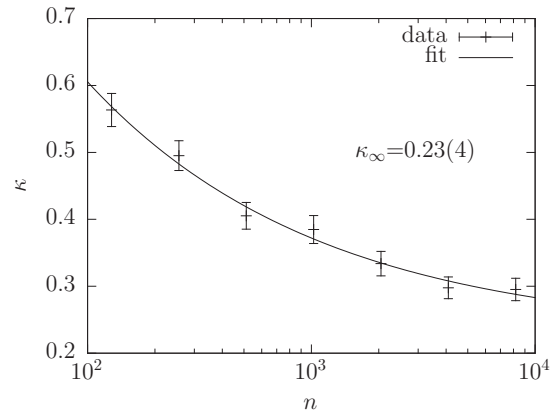

FIG. 6. Average distance-dendrogram correlation κ as function of sequence length n, for random permutations at $T = 0.2$. The line shows a result when fitting to a power law plus constant (see text).

the rows and columns is exactly given by the leaf order. The resulting matrix for 200 samples IS ($T = 0.2$) of one random permutation of length $n = 8192$ is displayed in Fig. 5. One observes a hierarchical structure given by two major clusters, visible by dark squares, i.e., similar configurations, which are subdivided into subclusters, with relatively smaller similarities on the off diagonals, respectively. This could indicate a multilevel of RSB, also since the matrices look similar to ones obtained for mean-field spin glasses [40]. On the contrary, for random XORSAT, which is known to exhibit one-step RSB, the matrices look trivial, i.e., plain gray, for $n \to \infty$ due to the exponentially diverging number of clusters [30–32].

The extent of the hierarchical structure can be made quantitative by calculating the *cophenetic correlation*$\kappa \equiv [d \cdot d_c]_P - [d][d_c]_P$, where $d = 1 - q$ is the distance corresponding to overlap value q. The *cophenetic distance* d_c between two states is measured on the dendrogram as the distance of the two largest clusters that contain only one of the states, respectively. $[\cdots]_P$ denotes the combined average over the sampled IS and the disorder ensemble. Thus, this κ measures the correlation between the original distance d of two states and the distance d_c imposed by the clustering, i.e., the degree of hierarchical structure. In Fig. 6 κ is shown as function of n. By fitting a power law $\kappa(n) = \kappa_\infty + an^{-b}$ we obtained $\kappa_\infty = 0.23(4)$. Thus, the IS landscape of permutations exhibits also for $n \to \infty$ a nested hierarchical structure, such as for problems exhibiting RSB as mean-field spin glasses or some hard combinatorial optimization problems [41].

IV. SUMMARY AND DISCUSSION

The original problem of Ulam is to find the longest increasing subsequence for random permutations. With so-far known algorithms it was possible to generate one LIS, but in a statistically uncontrolled way. Here, we have introduced an algorithm which allows for exact and direct sampling of increasing subsequences in polynomial time. For the uncorrelated and most natural ensemble of permutations, we find for the annealed mean and second moment of the number of ISs a good agreement with the analytical calculations. By using our algorithm, we numerically find a broad distribution

of overlaps and a hierarchical structure of configuration space. We show that this persists in the thermodynamic limit, thus the model exhibits for the equilibrium behavior thermodynamically relevant replica symmetry breaking. Simultaneously, this model exhibits RSB that is thermodynamically relevant for the equilibrium behavior and the model can be addressed by a polynomial exact sampling algorithm. This is in contrast to computationally hard problems such as mean-field spin glasses and NP-hard optimization problems. Thus, Ulam's problem provides an ideal test bed to study other phenomena or variants of interest. In particular one can address the nonequilibrium behavior, the scaling of excitations, the coupling of replicas, or an extended random model obtained by assigning individual local lengths for the numbers. It could also be of interest to consider ensembles with correlation or structure, in the spirit of a recent work on directed polymers in random media [36]. Furthermore, this study might motivate or help to identify other models with complex RSB behavior which can also be treated by polynomial algorithms. Finally, as our results are numerical and rely on an extrapolation for $n \to \infty$, it would be of great interest, if our study motivates analytical work on the configuration-space structure for Ulam's problem.

ACKNOWLEDGMENTS

We are grateful to A. Peter Young for critically reading the manuscript. We thank Federico Ricci-Tersenghi and Fred Hucht for useful discussions. The simulations were performed at the the HPC cluster CARL, located at the University of Oldenburg (Germany) and funded by the DFG through its Major Research Instrumentation Program (INST 184/157-1 FUGG) and the Ministry of Science and Culture (MWK) of the Lower Saxony State.

[1] S. M. Ulam, in *Modern Mathematics for the Engineer: Second Series*, edited by E. Beckenbach and M. Hestenes, Dover Books on Engineering Series (Dover, New York, 2013), Chap. 11, pp. 261–281.
[2] D. Romik, *The Surprising Mathematics of Longest Increasing Subsequences* (Cambridge University Press, New York, 2015).
[3] P. Erdös and G. Szekeres, Compos. Math. **2**, 463 (1935).
[4] D. Aldous and P. Diaconis, Bull. Am. Math. Soc. **36**, 413 (1999).
[5] T. Seppäläinen, Probab. Theory Relat. Fields **112**, 221 (1998).
[6] B. F. Logan and L. A. Shepp, Adv. Math. **26**, 206 (1977).
[7] J.-D. Deuschel and O. Zeitouni, Comb. Probab. Comput. **8**, 247 (1999).
[8] J. Baik, P. Deift, and K. Johansson, J. Am. Math. Soc. **12**, 1119 (1999).
[9] J. Börjes, H. Schawe, and A. K. Hartmann, Phys. Rev. E **99**, 042104 (2019).
[10] J. R. G. Mendonça, H. Schawe, and A. K. Hartmann, Phys. Rev. E **101**, 032102 (2020).
[11] J. M. Hammersley, in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Theory of Statistics* (University of California Press, Berkeley, CA, 1972), pp. 345–394.
[12] V. Lifschitz and B. Pittel, J. Combin. Theory, Ser. A **31**, 1 (1981).
[13] P. Krabbe, H. Schawe, and A. K. Hartmann, Phys. Rev. E **101**, 062109 (2020).
[14] P. Gopalan, T. S. Jayram, R. Krauthgamer, and R. Kumar, in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '07* (Society for Industrial and Applied Mathematics, Philadelphia, 2007), pp. 318–327.
[15] L. Bonomi and L. Xiong, Trans. Data Priv. **9**, 73 (2016).
[16] H. Zhang, Bioinformatics **19**, 1391 (2003).
[17] K. Binder and A. Young, Rev. Mod. Phys. **58**, 801 (1986).
[18] M. Mézard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[19] K. H. Fischer and J. A. Hertz, *Spin Glasses* (Cambridge University Press, Cambridge, UK, 1991).
[20] *Spin Glasses and Random Fields*, edited by A. P. Young (World Scientific, Singapore, 1998).
[21] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Oxford University Press, Oxford, UK, 2001).
[22] A. K. Hartmann and M. Weigt, *Phase Transitions in Combinatorial Optimization Problems* (Wiley-VCH, Weinheim, 2005).
[23] M. Mézard and A. Montanari, *Information, Physics and Computation* (Oxford University Press, Oxford, UK, 2009).
[24] C. Moore and S. Mertens, *The Nature of Computation* (Oxford University Press, Oxford, UK, 2011).
[25] N. Kawashima and H. Rieger, in *Frustrated Spin Systems*, 2nd ed., edited by H. T. Diep (World Scientific, Singapore, 2013), pp. 509–614.
[26] G. Parisi, Phys. Rev. Lett. **43**, 1754 (1979).
[27] A. K. Hartmann, *Big Practical Guide to Computer Simulations* (World Scientific, Singapore, 2015).
[28] F. Belletti, M. Cotallo, A. Cruz, L. A. Fernandez, A. Gordillo-Guerrero, M. Guidetti, A. Maiorano, F. Mantovani, E. Marinari, V. Martin-Mayor *et al.*, Comput. Sci. Eng. **11**, 48 (2009).
[29] M. Mézard, G. Parisi, and R. Zecchina, Science **297**, 812 (2002).
[30] F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, Phys. Rev. E **63**, 026702 (2001).
[31] M. Mézard, F. Ricci-Tersenghi, and R. Zecchina, J. Stat. Phys. **111**, 505 (2003).
[32] F. Ricci-Tersenghi, Science **330**, 1639 (2010).
[33] D. A. Huse and C. L. Henley, Phys. Rev. Lett. **54**, 2708 (1985).
[34] M. Kardar, Phys. Rev. Lett. **55**, 2235 (1985).
[35] M. Kardar and Y.-C. Zhang, Phys. Rev. Lett. **58**, 2087 (1987).
[36] A. K. Hartmann, Europhys. Lett. **137**, 41002 (2022).
[37] P. Jaccard, New Phytol. **11**, 37 (1912).
[38] As another check of our algorithm, we sampled ISs for also σ^o and obtained the same $C(T)$ curve.
[39] The clustering approach [42,43] operates on a set of M sampled configurations by initializing a set of M clusters each containing one configuration. One maintains pairwise distances between all clusters, which are initially the distances $d = 1 - q$ between the configurations, obtained from the overlaps q. Then iter-

atively two clusters exhibiting the currently shortest distance between them are selected and merged to one single cluster, thereby reducing the cluster number by one. For this new merged cluster, an updated distance to all other still existing clusters have to be obtained. Here, the update is done with the approach of Ward [42], which has been used previously for the analysis of disordered systems [40,41,44] (for more details, see there). The merging process is iterated until only one cluster is left.

[40] H. G. Katzgraber and A. K. Hartmann, Phys. Rev. Lett. **102**, 037207 (2009).
[41] W. Barthel and A. K. Hartmann, Phys. Rev. E **70**, 066120 (2004).
[42] J. Ward, J. Am. Stat. Assoc. **58**, 236 (1963).
[43] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data* (Prentice-Hall, Englewood Cliffs, NJ, 1988).
[44] A. Mann and A. K. Hartmann, Phys. Rev. E **82**, 056702 (2010).