

Evidence of many thermodynamic states of the three-dimensional Ising spin glass

Wenlong Wang ^{1,2,*}, Mats Wallin ¹ and Jack Lidmar ¹

¹*Department of Physics, Royal Institute of Technology, Stockholm, SE-106 91, Sweden*

²*College of Physics, Sichuan University, Chengdu 610065, China*

 (Received 25 May 2020; revised 4 September 2020; accepted 15 October 2020; published 16 November 2020)

We present a large-scale simulation of the three-dimensional Ising spin glass with Gaussian disorder to low temperatures and large sizes using optimized population annealing Monte Carlo. Our primary focus is investigating the number of pure states regarding a controversial statistic, characterizing the fraction of centrally peaked disorder instances, of the overlap function order parameter. We observe that this statistic is subtly and sensitively influenced by the slight fluctuations of the integrated central weight of the disorder-averaged overlap function, making the asymptotic growth behavior very difficult to identify. Modified statistics effectively reducing this correlation are studied, and essentially monotonic growth trends are obtained. The effect of temperature is also studied, finding a larger growth rate at a higher temperature. Our state-of-the-art simulation and variance reduction data analysis suggest that the many pure states picture is most likely and coherent.

DOI: [10.1103/PhysRevResearch.2.043241](https://doi.org/10.1103/PhysRevResearch.2.043241)

I. INTRODUCTION

Spin glasses are fascinating disordered and frustrated magnets with a wide array of applications in diverse fields such as biology, computer science, and optimization problems [1,2]. The mean-field Sherrington-Kirkpatrick (SK) model [3] has an unusual low-temperature phase of many pure states described by the replica symmetry breaking (RSB) [4–6]. Here a pure state refers to a self-sustained thermodynamic state characterized by a time-averaged spin orientational pattern. Despite several decades of efforts, it is still an outstanding problem whether the more realistic Edwards-Anderson (EA) spin glass [7] in three dimensions has a single pair or many pairs of pure states. The RSB picture [8,9] assumes that the mean-field theory is qualitatively correct for the EA model. On the other hand, the droplet picture [10–14] based on the domain-wall renormalization group method predicts only a single pair of pure states much like a ferromagnet. The two pictures also differ on the geometrical aspect of excitations (fractals or space-filling) [15] and the existence of a spin-glass phase in a weak external magnetic field [16]. There are other pictures as well [2]. The applicability of RSB is of broad interest and is related to, e.g., the Gardner transition in structural glasses [17,18].

Despite much research aiming at discriminating which picture is suitable in three (and also four) dimensions, the problem has not been definitely settled. In this work we

solely discuss the number of pure states, as a solid answer on one individual property can put stringent constraint on possible theories. There is mounting evidence that the disorder-averaged overlap function is nontrivial (corresponding to many pure states) for the sizes available, which have been steadily growing over time. One exception is the works focusing on the ground states at zero temperature [19,20]. However, we conjecture that a single pair of ground states is a strong support for neither a two-state nor many-state picture. It seems likely what is going on in this case is that there are nonzero energy gaps between the lowest pure states. In this way, even $O(1)$ large-scale droplet excitations are forbidden at $T = 0$. This is motivated by the observation that the disorder-averaged central weight [see Eq. (2)] decreases approximately linearly with decreasing temperature [21]. Next, we turn to the finite temperatures, which is the focus of this work.

We find the computational controversies regarding the number of pure states are essentially from investigating new statistics, new boundary conditions, or a combination of the two. To our best knowledge, all numerical simulations find nontrivial disorder-averaged overlap functions under periodic boundary conditions at a typical low temperature. Therefore, to argue against many states, it is necessary that one or several of the conditions have to be altered. The free boundary condition was thought to potentially support a two-state picture as $I(0.2)$ [see Eq. (2)] is found to rapidly decrease for small sizes and remarkably agrees with the $1/L^{\theta_{\text{DW}}}$ scaling, where θ_{DW} is the interface free energy exponent [22]. However, this was later found to be a finite-size effect from the surfaces [23] and $I(0.2)$ of the free and periodic boundary conditions run together for larger sizes, supporting many pure states. This also suggests that the many pure states are genuinely due to droplet excitations rather than topologically protected domain walls. By contrast, various statistics have been proposed other than $I(0.2)$, but most of these are not very successful; see, e.g., Refs. [24] and the references therein for a collection of

*wenlongcmp@scu.edu.cn

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by [Bibsam](https://www.bibsam.com/).

examples. One controversial but stimulating statistic is the fraction of centrally peaked instances [25–27], which we discuss in detail below. For new statistics and boundary conditions, a work on sample stiffness in thermal boundary conditions argued against many states [28]. This is also partially addressed [29], and to fully resolve this problem a contrast experiment of the SK model should be conducted, which shall be discussed elsewhere.

In this work we focus on the controversial statistic Δ of the fraction of centrally peaked instances [25–27] and find again that there is no violation of many pure states. This is significant since Δ appears to do the best job among the new statistics supporting the two-state picture [24]. In Ref. [25] it was found that Δ at $T = 0.42$ grows with system size up to about $L = 10$, then it levels off or stops growing appreciably. By contrast, Δ of the SK model at a similar T/T_C grows prominently for the similar range of sizes. A criticism in Ref. [26] suggested that comparing $T = 0.4T_C$ for different models has no theoretical basis, and the difference is from the effective lower temperature of the EA model, i.e., a smaller central weight $I(0.2)$. Increasing the temperature of the EA model such that it has a similar $I(0.2)$ as the SK model, it was found that Δ also grows prominently in the EA model. However, the problem was not fully addressed in spite of the profound insight. An explanation of the irregular low temperature data is missing, and slightly different models were compared. The former group used Gaussian disorder and a low temperature [25], while the latter group used $\pm J$ disorder and a relatively high temperature [26].

The main purpose of this work is to resolve this problem systematically by carrying out a large-scale simulation of the three-dimensional EA model at the same parameters using the same model as the original work [25] but including larger sizes. In light of Ref. [26], data at a higher temperature are also collected for comparison. Using massively parallel population annealing Monte Carlo [30–34] and scalable MPI parallel computing, and taking further advantage of the recent optimizations of the algorithm [34–36], we have managed with considerable efforts to increase the largest size from 12^3 spins [25] to 16^3 spins. We refer to Ref. [35] for a discussion of how notoriously the spin glass computational complexity grows at low temperatures with the number of spins. Our larger range of sizes crucially enables us to identify a *subtle* correlation between $\Delta(q_0, \kappa)$ and $I(q_0)$, showing that even small $I(q_0)$ fluctuations can significantly influence the behavior of Δ . Motivated by our observation, we define a slightly modified Δ compensating effectively for this correlation effect. The modified Δ essentially grows monotonically, providing a coherent many-state picture. Our results also confirm that the different results of earlier works originate from the use of different effective temperatures or central weights [26].

II. MODEL, METHOD, AND OBSERVABLES

We study the three-dimensional Edwards-Anderson Ising spin glass [7] defined by the Hamiltonian $H = -\sum_{\langle ij \rangle} J_{ij} S_i S_j$, where $S_i = \pm 1$ are Ising spins and the sum is over nearest neighbors on a simple cubic lattice under periodic boundary conditions. For a linear size L , there are $N = L^3$

TABLE I. Preliminary parameters of the population annealing simulation. L is the linear system size, R is the number of replicas, T_0 is the lowest temperature simulated, N_T is the number of temperatures used in the annealing schedule, N_S is the number of sweeps applied to each replica after resampling, and M is the number of instances studied. Note that unequilibrated instances were rerun with (much) larger simulation parameters; see the text for details.

L	R	T_0	N_T	N_S	M
4	8×10^4	0.42	101	10	5000
6	1.6×10^5	0.42	101	10	5000
8	4×10^5	0.42	201	10	5000
10	9.6×10^5	0.42	301	10	5000
12	9.6×10^5	0.42	301	10	5000
14	9.6×10^5	0.42	401	20	3500
16	9.6×10^5	0.42	401	40	3424

spins. The random couplings J_{ij} are chosen independently from the standard Gaussian distribution $n(0, 1)$. A set of quenched couplings $\mathcal{J} = \{J_{ij}\}$ defines a disorder realization sample or an instance. The model has a spin-glass phase transition at $T_C \approx 1$ [37].

Population annealing cools gradually a population of R random replicas starting from $\beta = 0$ with alternating resampling and Metropolis sweeps until reaching the lowest temperature. In a resampling step, when the inverse temperature is increased from β to β' , a replica i is copied according to its energy E_i with the expectation number $n_i = \exp[-(\beta' - \beta)E_i]/Q$. Here $Q = (1/R) \sum_i \exp[-(\beta' - \beta)E_i]$ is a normalization factor to keep the population size approximately the same as R . In our simulation, the number of copies is randomly chosen as either the floor or the ceiling of n_i with the proper probability to minimize fluctuations. After the resampling step, Monte Carlo sweeps are applied to each replica at the new temperature. Population annealing is used because it is both efficient and massively parallel [33,36,38]. Our equilibration criterion is based on the replica family entropy $S_f = -\sum_i f_i \log(f_i)$, where f_i is the fraction of replicas descended from replica i of the initial population. We require $S_f \geq \log(100)$ at the lowest temperature for each individual instance [28,33]. The preliminary simulation parameters are summarized in Table I, and unequilibrated instances were rerun with larger parameters until equilibration is reached. It should be noted that the hard instances may require substantially more work than a typical instance. For example, our typical top 5% hard instances at $L = 16$ require approximately $R \sim 10^7$ replicas, $N_T = 501$ temperatures, and as large as $N_S = 200$ sweeps on each replica per temperature; cf., the preliminary parameters in the table. Finally, our data readily pass the disorder-average based equilibration test of [39] at the lowest and a higher temperature ($T = 2/3$) for all sizes.

Our primary observable is the spin overlap distribution function $P_{\mathcal{J}}(q)$ where the spin overlap q is defined as

$$q = \frac{1}{N} \sum_i S_i^{(1)} S_i^{(2)}, \quad (1)$$

where spin configurations “(1)” and “(2)” are chosen independently from the Boltzmann distribution. We have collected

the distributions at two typical low temperatures $T = 2/3$ and $T = 0.42$. Other regular observables like energy, free energy, and the replica family entropy are collected at all temperatures.

We further introduce two statistics of the individual overlap function: the central weight and central peakedness [25]:

$$I_{\mathcal{J}}(q_0) = \int_{-q_0}^{q_0} P_{\mathcal{J}}(q) dq, \quad (2)$$

$$\Delta_{\mathcal{J}}(q_0, \kappa) = \theta \left(\max_{|q| \leq q_0} ([P_{\mathcal{J}}(q) + P_{\mathcal{J}}(-q)]/2) - \kappa \right). \quad (3)$$

Here q_0 characterizes the half length of a chosen interval around $q = 0$, κ is a chosen threshold to determine whether or not an instance is centrally peaked, and the Heaviside function $\theta(x) = 1$ if $x \geq 0$ and 0 otherwise. The statistic $\Delta_{\mathcal{J}}$ takes either 0 or 1. When the subscript is dropped, we refer to the disorder-averaged quantity. Hence, Δ refers to the fraction of centrally peaked instances. Δ should decrease to 0 for a two-state picture, while it should increase to 1 for a many-state picture in the thermodynamic limit.

III. RESULTS

The disorder-averaged overlap function and the central weight $I(0.2)$ at both $T = 2/3$ and 0.42 are shown in Fig. 1. The central weight is essentially flat up to fluctuations as a function of the system size, in agreement with numerous previous results [25,33,40,41]. This result is well known, except that we here further extend it to two larger system sizes at low temperatures. The result is in excellent agreement with RSB, but is strikingly different from the $1/L^{\theta_{\text{DW}}}$ ($\theta_{\text{DW}} \approx 0.2$) droplet scaling. To our best knowledge, there is no straightforward way to explain this as a finite-size effect, because the interface free energy scales well with this exponent for the same range of sizes. Finally, the weights of $T = 0.42$ are smaller than those of $T = 2/3$ as expected, as the effective number of “active” pure states is suppressed at lower temperatures.

Next, we discuss the statistic Δ in detail. The results of $\Delta(0.2, 1)$ and $\Delta(0.4, 1)$ are shown in the top left panel of Fig. 2. At first sight, the data look quite irregular as observed in Ref. [25]. There appears to be a growing trend in general, and the trend is much clearer at $T = 2/3$ than at the lower temperature $T = 0.42$. We shall discuss the effect of the temperature below, and first focus on the origin of the irregular low temperature data. We take the $\Delta(0.2, 1)$ as an example without loss of generality. After an increasing trend at small sizes, the grow from $L = 10$ to 12 is very marginal (marginal or slightly negative depending on disorder realizations; here we call them collectively marginal), in agreement with Ref. [25]. Then we observe a noticeable increase from $L = 12$ to 14, a somewhat promising result for many pure states. But the statistic subsequently grows rather marginally again from $L = 14$ to 16, resulting in a rather confusing situation. This puzzle is finally understood when we recognized a subtle correlation between Δ and the central weight. As illustrated in Fig. 1, $I(0.2)$ drops slightly from $L = 10$ to 12; this is where the corresponding Δ has a marginal increase. Then $I(0.2)$ increases slightly from $L = 12$ to 14 and then drops slightly at $L = 16$, explaining the correlated trend in Δ . This together with other similar observations throughout our

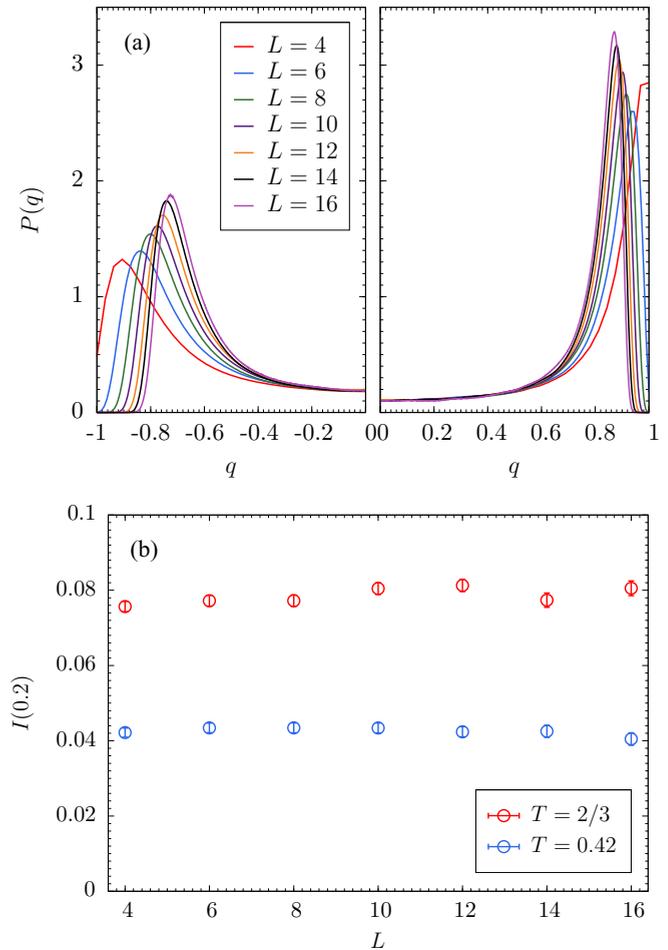


FIG. 1. (a) The disorder-averaged overlap function for different system sizes L at typical low temperatures $T = 2/3$ (left) and $T = 0.42$ (right), respectively. (b) The central weight $I(0.2)$ is approximately independent of the system size L for both temperatures, in agreement with the many-state picture. Note that the maximum size is $L = 16$, compared with $L = 12$ of [25].

data collection process strongly suggest that the two statistics are correlated. Then Δ is presumably a growing function with increasing system size, but the growth is very sensitively affected by the fluctuations of the central weights, producing an irregular growth trend. This also partially explains why the data at a higher temperature have a clearer growth trend, as the relative fluctuation of the central weights with respect to the size-averaged mean is smaller at higher temperatures. An additional reason is that Δ has a larger growth rate at higher temperatures; note the crossings of the data at the two temperatures.

The correlation between Δ and I is reasonable from the following argument. The central weight is like a supply of peaks, and higher weights supply more peaks, which tend to statistically produce more peaked instances. Take two extreme examples: if $I(0.2) = 0$, it is clear that $\Delta(0.2, 1)$ is bounded to be zero. On the other hand, if $I_{\mathcal{J}}(0.2)$ for an instance is large, there is almost certainly a peak present, at least when the size is large. The more detailed sample-wise correlation is further illustrated in the Appendix. Moreover, it is found therein the

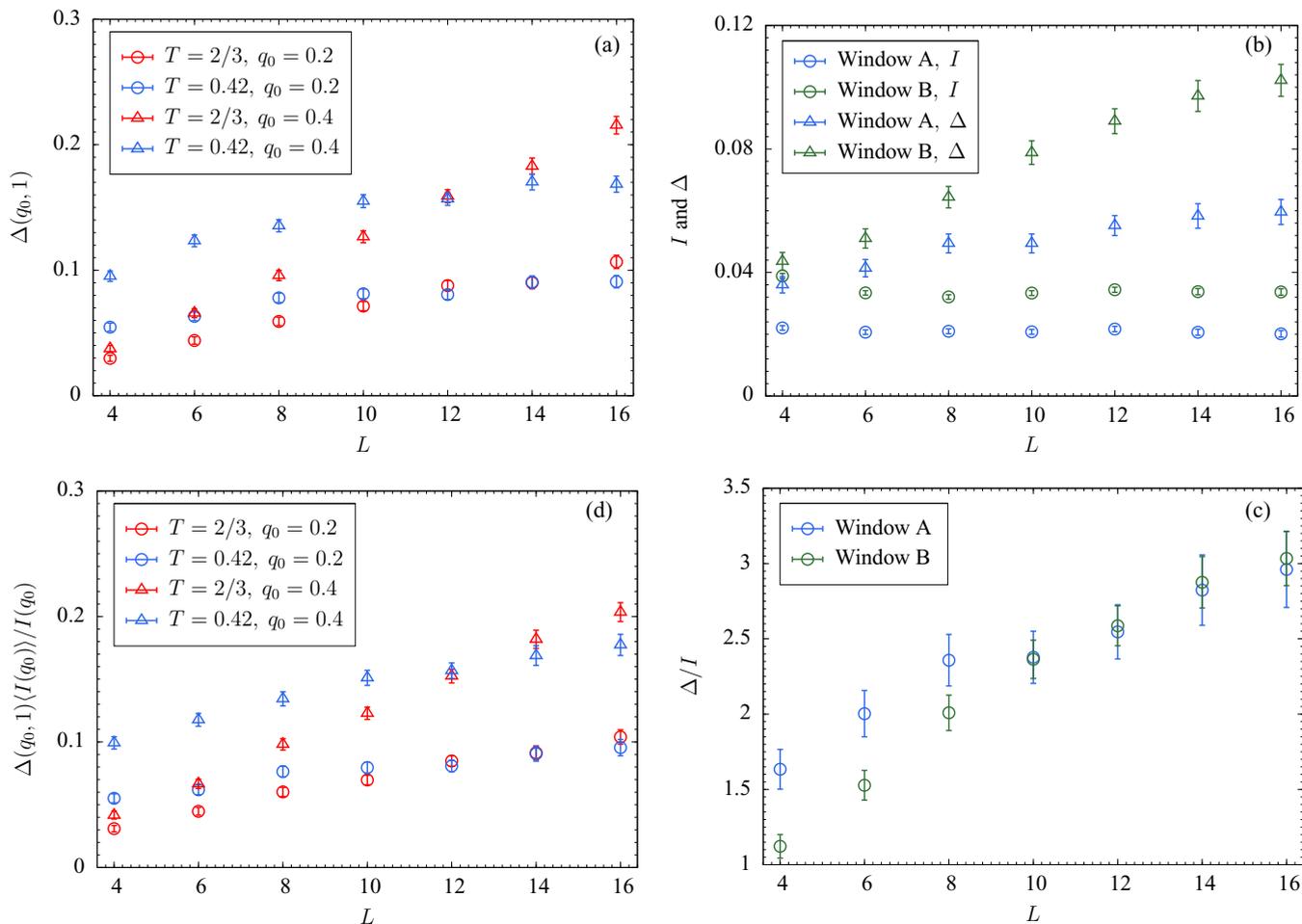


FIG. 2. The statistics $\Delta(0.2, 1)$ and $\Delta(0.4, 1)$ as a function of L at two temperatures $T = 2/3$ and $T = 0.42$ (a). While Δ appears to have a growing trend, it is not fully regular. We find that Δ is subtly influenced by the central weight; cf., Fig. 1. To illustrate this correlation, two windows of the same size but of very different weights are studied at $T = 0.42$, and the Window B with a larger weight also has consistently larger Δ (b). This correlation can be effectively reduced by looking at the statistic Δ/I as shown in panel (c). Similarly, the modified Δ as shown in panel (d) has a much cleaner growth trend and is in fact remarkably monotonic. Here $\langle I \rangle$ is the average of I over the sizes. See the discussion for more details.

correlation becomes stronger at larger sizes and lower temperatures, as the system moves closer to the δ -peaks regime. This is reasonable as in this regime, the absence or presence of the central weight would directly correspond to $\Delta_{\mathcal{J}} = 0$ or 1, respectively. Since the fluctuation of I has a profound influence on Δ , we next look for modified statistics to compensate this correlation effect by a variance reduction method.

We define a slightly modified statistic of weighted Δ as $\Delta(q_0, \kappa)[\langle I(q_0) \rangle / I(q_0)]$, where the angular bracket is an average over the system size. While this definition assumes in the first place I is approximately constant, it has strong numerical supports as mentioned previously and it only slightly modifies the Δ data. Nevertheless, we shall present below another statistic which is similar in spirit but does not have this “problem.” First, we look at an example to *motivate* the weighted Δ using two windows that have very different weights, showing that it is effective for comparing Δ with different weights. From the overlap functions, it is clear that at the lower temperature there are wide q ranges where the major q_{EA} peaks have little influence. In addition, the weight density is higher at larger q than at the neighborhood of $q = 0$. Therefore, we

select the following two windows and study the behaviors of I and Δ at $T = 0.42$: Window A as $q \in [-0.1, 0.1]$ with a small weight and Window B as $|q| \in [0.4, 0.5]$ with the same length but a noticeably larger weight. Here I and Δ are measured in the respective supports. The two statistics as a function of the system size for these two windows are shown in the top right panel of Fig. 2. Since Window B has consistently larger weights I , it also has consistently larger Δ as expected. The ratio Δ/I is shown in the bottom right panel, and this simple statistic brings the two Δ data sets much closer particularly for the pertinent large sizes. Similar behavior is also found for the higher temperature, despite that Window B is slightly affected by the q_{EA} peaks. These demonstrate that Δ/I is an effective statistic to reduce the correlation effect from the weights, despite that it may not fully remove it.

The modified Δ is shown in the bottom left panel of Fig. 2. It is remarkable that this simple statistic has a clean growth trend, i.e., the growth trend is not only improved but also monotonic. We have carried out a quantitative growing trend test to leading linear order using a linear fit, as nearby data variations can be within error bars. The computed slopes are

0.0034(5) (low T , $q_0 = 0.2$), 0.0062(4) (high T , $q_0 = 0.2$), 0.0067(6) (low T , $q_0 = 0.4$), 0.0138(5) (high T , $q_0 = 0.4$), respectively. These values, especially the high-temperature data, are all significantly larger than 0, suggesting a collective growing trend of this statistic. We conclude therefore that the seemingly nonmonotonic growth of Δ is a result of its sensitivity to the fluctuations of the central weight. From this perspective, the statistic Δ is not as good as the central weight in discriminating the number of pure states due to its discrete nature as it amplifies fluctuations.

We now comment on the effect of the temperature on the growth rate of Δ . The modified Δ suggests that the growth rate at the higher temperature is noticeably larger than at the lower temperature. There is also an interesting crossing in Δ for each interval despite that I has no crossing, showing the complex quantitative relations of I and Δ in general. Nevertheless, the crossings can be qualitatively understood by the two peak-sharpening mechanisms mentioned above either by increasing the system size or by lowering the temperature. When the system size is sufficiently large or the correlation is sufficiently strong, i.e., typical peaks regardless of the temperature are sharp and tall with respect to κ , the order of I should sufficiently determine the order of Δ as the temperature is tuned. Then the Δ at the higher temperature should be larger as the weight is larger. This is found numerically in Fig. 2, as the red points exceed the blue points when the size is increased. On the other hand, the impact of the temperature on Δ for small sizes is, however, more complex. In this regime, correlation may strongly increase with decreasing temperature, i.e., typical peaks are wide and short at high temperatures and sharp and tall at low temperatures. Meanwhile, the weight decreases with decreasing temperature. Therefore, as the temperature decreases, these two effects *compete* with each other. This provides a possibility that Δ may increase in a temperature interval as the temperature is decreased, and Fig. 2 suggests that this is actually realized, e.g., at $L = 4$. Fortunately, it is not computationally hard to study the small size in detail in a wide range of temperatures. To this end, we have conducted an additional set of simulation to study the evolution of $\Delta(T)$ and $I(T)$ at $L = 4$ down to $T = 0.1$. Our picture is numerically confirmed, and the results are shown in Fig. 3. In a broad range of temperatures, here approximately $0.3 \lesssim T < 0.8$, Δ grows with decreasing temperature. When typical peaks are sufficiently narrow and tall at low temperatures $T \lesssim 0.3$, similar to the large size regime, the weight effect takes over and Δ decreases along with decreasing I , converging to 0 as $T \rightarrow 0$. The *opposite* trends in the small- and large-size regimes strongly suggest that the growth rate of Δ is an increasing function of temperature in this wide temperature range $0.3 \lesssim T < 0.8$, if we reasonably assume the growth is approximately linear particularly for the weighted Δ . This is also in line with Refs. [25,26], and our argument indicates that Δ should grow even slower and become more irregular at a slightly lower temperature, e.g., $T = 0.3$.

Next we look at an alternative method for variance reduction [42] by considering $\Delta - I$. The modified Δ above assumes I is approximately constant, but this statistic does not need such an assumption. The results are shown in the top panels of Fig. 4. Similar to the modified Δ , this statistic also has a clean growth trend, in agreement with many pure states.

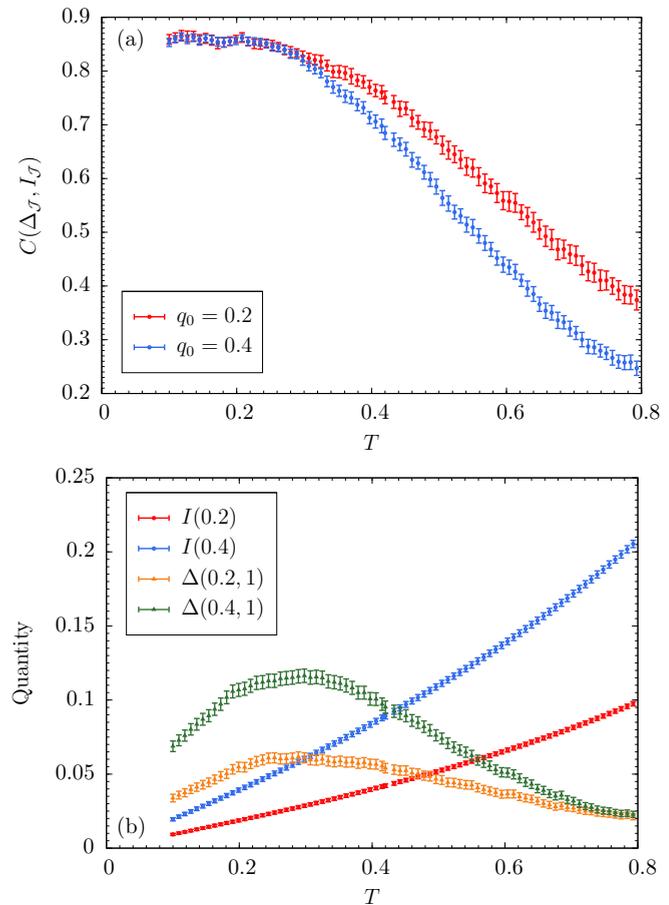


FIG. 3. Different from the large-size regime, the effect of the temperature on Δ for small sizes is complex because the rapidly increasing correlation and decreasing weight I compete with each other as the temperature is decreased. The top panel (a) shows the Pearson correlation coefficient, and the bottom panel (b) shows Δ and I for $L = 4$. Here a local maximum of Δ is observed. As T decreases, Δ first increases due to the rapidly growing correlation or peak sharpening between $0.3 \lesssim T < 0.8$. When typical peaks are sufficiently narrow and tall $T \lesssim 0.3$, the weight effect takes over and Δ decreases with I , converging to 0 as $T \rightarrow 0$. The figure also illustrates clearly the fluctuation of Δ is systematically larger than that of the central weight due to its discrete nature and thus amplifies fluctuation.

By contrast, this statistic should converge to 0 for a two-state picture, as the two terms should separately converge to 0. The data are clearly running above 0 and are still growing, and should reach finite limits if there are many pure states. In addition, this statistic can be further refined using the control variate analysis [43,44]. This method utilizes a parametrized control variate estimator $\Delta - \lambda I$ where λ is a free parameter. In order to minimize the variance of this statistic, it is straightforward to show that $\lambda = (\sigma_{\Delta}/\sigma_I)C(\Delta_{\mathcal{J}}, I_{\mathcal{J}})$. The optimum $\lambda(T, q_0)$ depends on the temperature and the specific Δ window and varies reasonably slowly with the system size. The size-averaged estimates are approximately $\lambda(2/3, 0.2) = 1.51$, $\lambda(0.42, 0.2) = 2.34$, $\lambda(2/3, 0.4) = 0.97$, and $\lambda(0.42, 0.4) = 1.85$, respectively. These refined statistics are shown in the bottom

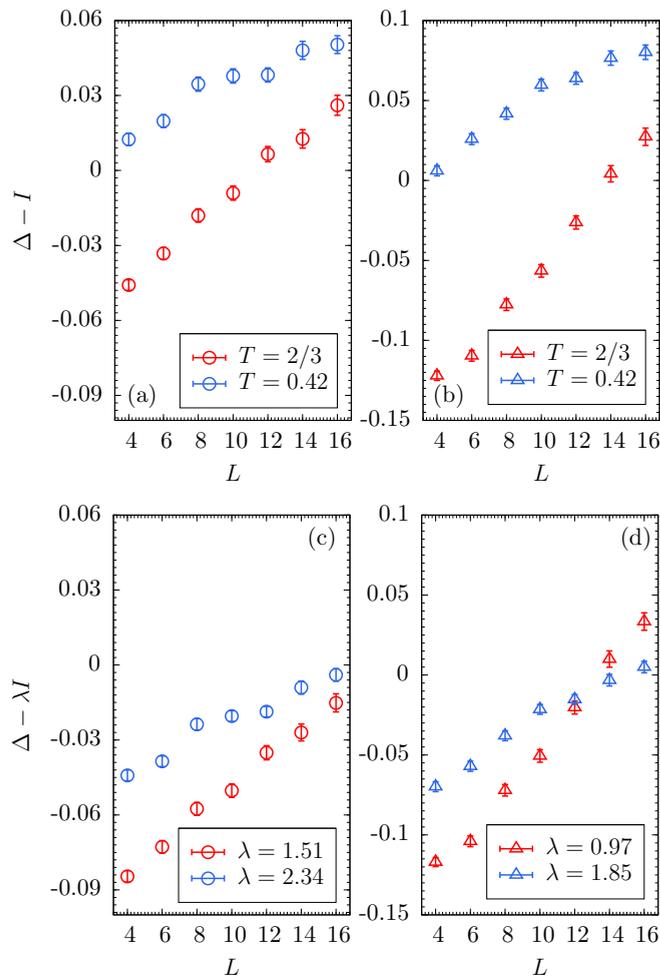


FIG. 4. (a, b) A more traditional statistic $\Delta - I$ to decorrelate Δ and I ; the left (a) and right (b) panels are for $q_0 = 0.2$ and 0.4 , respectively. The statistic shows a growing behavior as the modified Δ , in agreement with the many-state picture. By contrast, the statistic should converge to 0 for a two-state picture. (c, d) The same as above but for a refined statistic $\Delta - \lambda I$ using the control variate analysis with a size-averaged parameter λ . The data appear more regular as the error bar is further suppressed, showing again a growing behavior.

panels of Fig. 4, showing similar behaviors, and it seems that the data are even more regular, as the error bar is further suppressed particularly at the low temperature. Note that $\Delta - I$ remains important, despite its larger error bar, as it clearly runs above 0. Therefore, the statistics I , modified Δ , $\Delta - I$, and $\Delta - \lambda I$ are all coherently in agreement with the many-state picture.

Our results are in full qualitative agreement with Refs. [25,26]. The former group found a small growth rate of Δ at a low temperature, while the latter group found instead a much larger growth rate by operating at a higher temperature. The size of the central weight again has a significant impact on the growth behaviors of Δ [25]. Our large range of sizes is crucial in identifying this subtle correlation. Therefore, we interpret the low-temperature irregular Δ as fluctuations because of a small growth rate, rather than a genuine onset of the two-state picture. We conclude that the statistic Δ has

no evidence of violating many pure states, and instead it is consistent with a coherent many-state picture [26]. We cannot rule out the possibility that the two-state picture behavior is realized at yet much larger sizes currently not accessible, but we cannot glimpse such a trend, and it appears an unlikely scenario.

Finally, we briefly discuss the difficulties of the two-state picture with current numerical results. (1) In the same range of sizes, we clearly get a finite domain-wall exponent θ_{DW} yet a flat $I(0.2)$. It is unclear what finite-size effect is responsible for a flat central weight in the droplet framework. (2) The θ_{DW} exponent is a growing function of dimensionality [45], and even the SK model has a positive exponent [46] which is clearly described by RSB. It seems likely that $\theta_{\text{DW}} > 0$ is not capable of excluding large-scale $O(1)$ droplet excitations as suggested by the droplet picture, which assumes that droplet and domain-wall excitations are similar; see, e.g., Ref. [29] for an interesting possibility of how a positive domain-wall exponent and many pure states can coexist. Otherwise, an explanation should be provided on why this argument does not apply to, e.g., the SK model. These difficulties in our opinion must be addressed for a two-state picture to be consistent.

IV. CONCLUSIONS

In this work we carried out a state-of-the-art simulation of the Gaussian Ising spin glass in three dimensions and examined the statistic Δ in detail. Our results reveal that the nonmonotonic growth of the statistic with system size is a result of its sensitivity to the fluctuations of the central weight I . By looking at a modified Δ and also $\Delta - I$ compensating for this correlation effect, we find essentially monotonic growth of the statistics. Combining with the relatively flat central weight, we conclude that the statistic Δ is in full agreement with the many-state picture but not with the two-state picture.

The spin-glass literature is overall currently far from conclusive, our investigation of the number of pure states benefits from running state-of-the-art simulations and using variance reduction data analysis. Since simulating much larger system sizes is currently not an option, it is highly motivated to carry out careful statistical analysis of the data to reduce the fluctuations in the observed quantities, as we have demonstrated here. To ultimately decide between different pictures might require considerably larger simulations that are presently out of reach, but our results suggest that the many pure states picture is most likely.

ACKNOWLEDGMENTS

We thank J. Machta, M. Weigel, and B. Yucesoy for helpful discussions. W.W. acknowledges support from the Swedish Research Council Grant No. 642-2013-7837, and the Goran Gustafsson Foundation for Research in Natural Sciences and Medicine, and the Fundamental Research Funds for the Central Universities, China. M.W. acknowledges support from the Swedish Research Council Grant No. 621-2012-3984. The computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at the National Supercomputer Centre (NSC), the High Performance Computing Center North (HPC2N), and the Emei cluster at Sichuan University.

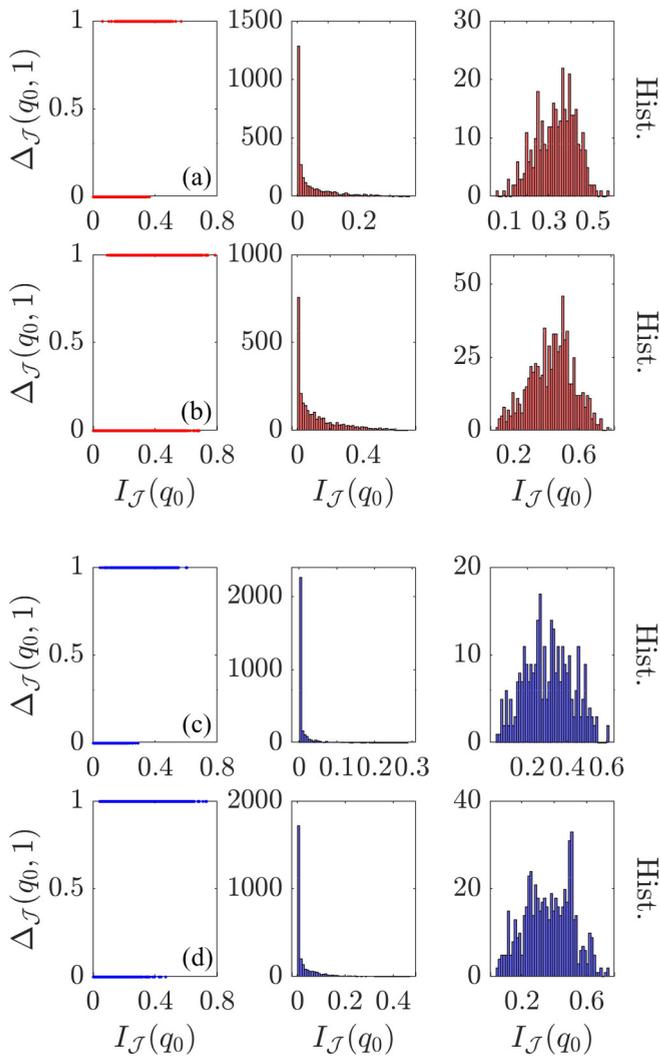


FIG. 5. The left column of top panels (a) shows the scatter plot of $\Delta_{\mathcal{J}}$ vs $I_{\mathcal{J}}$ for $q_0 = 0.2$ at $T = 2/3$. Note the empty top left and bottom right regions in the scatter plot, suggesting sample-wise correlation. To further illustrate the scatter plot, the histograms of $I_{\mathcal{J}}$ for the $\Delta_{\mathcal{J}} = 0$ and 1 classes are shown in the second and third columns, respectively. Similar correlations are found for $q_0 = 0.4$ (b), as well as the lower temperature $T = 0.42$ in the corresponding panels (c, d). Other system sizes show very similar features.

APPENDIX: CORRELATION BETWEEN $\Delta_{\mathcal{J}}$ AND $I_{\mathcal{J}}$

In this Appendix, we characterize the correlation between $\Delta_{\mathcal{J}}$ and $I_{\mathcal{J}}$ in detail at the sample-wise level. The scatter

TABLE II. Pearson correlation coefficient $C(\Delta_{\mathcal{J}}, I_{\mathcal{J}})$ for different system sizes L , temperatures T , and overlap intervals $[-q_0, q_0]$. Here the pair represents (T, q_0) . The correlation is stronger at the lower temperature and larger sizes, where typical peaks are narrow and tall.

L	(2/3, 0.2)	(2/3, 0.4)	(0.42, 0.2)	(0.42, 0.4)
4	0.485(17)	0.351(14)	0.751(13)	0.685(14)
6	0.556(16)	0.413(14)	0.787(12)	0.740(11)
8	0.629(15)	0.508(13)	0.808(10)	0.772(09)
10	0.677(13)	0.548(11)	0.833(09)	0.800(08)
12	0.717(11)	0.616(10)	0.838(09)	0.805(08)
14	0.733(14)	0.646(11)	0.840(09)	0.829(08)
16	0.751(12)	0.670(10)	0.861(08)	0.832(08)

plot of these two quantities of $L = 16$ at both temperatures and intervals is shown in Fig. 5. Since Δ is a discrete binary statistic, the conditional distributions of I for $\Delta_{\mathcal{J}} = 0$ and 1 are illustrated. These plots, and also those of other system sizes, have very similar features.

In the scatter plot, there is an interesting shift of $I_{\mathcal{J}}$ to larger values when $\Delta_{\mathcal{J}}$ steps from 0 to 1. This means that when the central weight is close to 0, it is almost surely $\Delta_{\mathcal{J}} = 0$. At the opposite end, samples of the largest central weights have almost surely $\Delta_{\mathcal{J}} = 1$. In the histograms, the central weight conditional distribution has a very biased distribution towards 0 for the class $\Delta_{\mathcal{J}} = 0$, while it is approximately Gaussian distributed for the class $\Delta_{\mathcal{J}} = 1$. Note that a sizable fraction of samples have $I_{\mathcal{J}}$ close to 0, indicating the correlation is prominent. The fraction of small weighted samples increases as the temperature is decreased, suggesting that the correlation should be stronger at lower temperatures.

In order to characterize the correlation more quantitatively, we have calculated the Pearson correlation coefficient, and the results are summarized in Table II. First, all coefficients are significantly larger than 0, confirming the strong positive correlation between the two statistics. Moreover, the correlation becomes stronger at larger sizes and the lower temperature. A more detailed study of the size $L = 4$ down to $T = 0.1$ (see Fig. 3) suggests that the correlation is an increasing function of the decreasing temperature. It should be noted that in both cases the system is moving closer to the δ -peaks regime. This trend is reasonable as in this regime, the absence or presence of the central weight would directly correspond to $\Delta_{\mathcal{J}} = 0$ or 1, respectively.

[1] K. Binder and A. P. Young, Spin glasses: Experimental facts, theoretical concepts and open questions, *Rev. Mod. Phys.* **58**, 801 (1986).
 [2] D. Stein and C. Newman, *Spin Glasses and Complexity, Primers in Complex Systems* (Princeton University Press, Princeton, 2013).
 [3] D. Sherrington and S. Kirkpatrick, Solvable Model of a Spin Glass, *Phys. Rev. Lett.* **35**, 1792 (1975).
 [4] G. Parisi, Infinite Number of Order Parameters for Spin-Glasses, *Phys. Rev. Lett.* **43**, 1754 (1979).

[5] G. Parisi, The order parameter for spin glasses: A function on the interval 0–1, *J. Phys. A* **13**, 1101 (1980).
 [6] G. Parisi, Order Parameter for Spin-Glasses, *Phys. Rev. Lett.* **50**, 1946 (1983).
 [7] S. F. Edwards and P. W. Anderson, Theory of spin glasses, *J. Phys. F: Met. Phys.* **5**, 965 (1975).
 [8] G. Parisi, Some considerations of finite dimensional spin glasses, *J. Phys. A* **41**, 324002 (2008).
 [9] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).

- [10] W. L. McMillan, Domain-wall renormalization-group study of the two-dimensional random Ising model, *Phys. Rev. B* **29**, 4026 (1984).
- [11] A. J. Bray and M. A. Moore, Scaling theory of the ordered phase of spin glasses, in *Heidelberg Colloquium on Glassy Dynamics and Optimization*, edited by L. Van Hemmen and I. Morgenstern (Springer, New York, 1986), p. 121.
- [12] D. S. Fisher and D. A. Huse, Ordered Phase of Short-Range Ising Spin-Glasses, *Phys. Rev. Lett.* **56**, 1601 (1986).
- [13] D. S. Fisher and D. A. Huse, Absence of many states in realistic spin glasses, *J. Phys. A* **20**, L1005 (1987).
- [14] D. S. Fisher and D. A. Huse, Equilibrium behavior of the spin-glass ordered phase, *Phys. Rev. B* **38**, 386 (1988).
- [15] W. Wang, M. A. Moore, and H. G. Katzgraber, Fractal Dimension of Interfaces in Edwards-Anderson and Long-Range Ising Spin Glasses: Determining the Applicability of Different Theoretical Descriptions, *Phys. Rev. Lett.* **119**, 100602 (2017).
- [16] J. R. L. de Almeida and D. J. Thouless, Stability of the Sherrington-Kirkpatrick solution of a spin glass model, *J. Phys. A* **11**, 983 (1978).
- [17] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, Fractal free energy landscapes in structural glasses, *Nat. Comm.* **5**, 3725 (2014).
- [18] C. L. Hicks, M. J. Wheatley, M. J. Godfrey, and M. A. Moore, Gardner Transition in Physical Dimensions, *Phys. Rev. Lett.* **120**, 225501 (2018).
- [19] M. Palassini and A. P. Young, Triviality of the Ground State Structure in Ising Spin Glasses, *Phys. Rev. Lett.* **83**, 5126 (1999).
- [20] N. Hatano and J. E. Gubernatis, Evidence for the double degeneracy of the ground state in the three-dimensional $\pm J$ spin glass, *Phys. Rev. B* **66**, 054437 (2002).
- [21] R. Alvarez Baños, A. Cruz, L. A. Fernandez, J. M. Gil-Narvion, A. Gordillo-Guerrero, M. Guidetti, A. Maiorano, F. Mantovani, E. Marinari, V. Martin-Mayor *et al.*, Nature of the spin-glass phase at experimental length scales, *J. Stat. Mech.: Theory Exp.* (2010) P06026.
- [22] H. G. Katzgraber and A. P. Young, Monte Carlo simulations of spin-glasses at low temperatures: Effects of free boundary conditions, *Phys. Rev. B* **65**, 214402 (2002).
- [23] W. Wang, Numerical simulations of Ising spin glasses with free boundary conditions: The role of droplet excitations and domain walls, *Phys. Rev. E* **95**, 032143 (2017).
- [24] M. Wittmann, B. Yucesoy, H. G. Katzgraber, J. Machta, and A. P. Young, Low-temperature behavior of the statistics of the overlap distribution in Ising spin-glass models, *Phys. Rev. B* **90**, 134419 (2014).
- [25] B. Yucesoy, H. G. Katzgraber, and J. Machta, Evidence of Non-Mean-Field-Like Low-Temperature Behavior in the Edwards-Anderson Spin-Glass Model, *Phys. Rev. Lett.* **109**, 177204 (2012).
- [26] A. Billoire, L. A. Fernandez, A. Maiorano, E. Marinari, V. Martin-Mayor, G. Parisi, F. Ricci-Tersenghi, J. J. Ruiz-Lorenzo, and D. Yllanes, Comment on “Evidence of Non-Mean-Field-Like Low-Temperature Behavior in the Edwards-Anderson Spin-Glass Model,” *Phys. Rev. Lett.* **110**, 219701 (2013).
- [27] B. Yucesoy, H. G. Katzgraber, and J. Machta, Yucesoy, Katzgraber, and Machta Reply, *Phys. Rev. Lett.* **110**, 219702 (2013).
- [28] W. Wang, J. Machta, and H. G. Katzgraber, Evidence against a mean-field description of short-range spin glasses revealed through thermal boundary conditions, *Phys. Rev. B* **90**, 184412 (2014).
- [29] W. Wang, M. Wallin, and J. Lidmar, Chaotic temperature and bond dependence of four-dimensional Gaussian spin glasses with partial thermal boundary conditions, *Phys. Rev. E* **98**, 062122 (2018).
- [30] K. Hukushima and Y. Iba, in *The Monte Carlo Method in the Physical Sciences: Celebrating the 50th Anniversary of the Metropolis Algorithm*, edited by J. E. Gubernatis, AIP Conf. Proc. No. 690 (AIP, New York, 2003), p. 200.
- [31] E. Zhou and X. Chen, in *Proceedings of the 2010 Winter Simulation Conference (WSC), Baltimore, MD* (IEEE, Piscataway, NJ, 2010), pp. 1211–1222.
- [32] J. Machta, Population annealing with weighted averages: A Monte Carlo method for rough free-energy landscapes, *Phys. Rev. E* **82**, 026704 (2010).
- [33] W. Wang, J. Machta, and H. G. Katzgraber, Population annealing: Theory and application in spin glasses, *Phys. Rev. E* **92**, 063307 (2015).
- [34] L. Y. Barash, M. Weigel, M. Borovský, W. Janke, and L. N. Shchur, GPU accelerated population annealing algorithm, *Comput. Phys. Commun.* **220**, 341 (2017).
- [35] C. Amey and J. Machta, Analysis and optimization of population annealing, *Phys. Rev. E* **97**, 033301 (2018).
- [36] A. Barzegar, C. Pattison, W. Wang, and H. G. Katzgraber, Optimization of population annealing Monte Carlo for large-scale spin-glass simulations, *Phys. Rev. E* **98**, 053308 (2018).
- [37] H. G. Katzgraber, M. Körner, and A. P. Young, Universality in three-dimensional Ising spin glasses: A Monte Carlo study, *Phys. Rev. B* **73**, 224432 (2006).
- [38] W. Wang, J. Machta, and H. G. Katzgraber, Comparing Monte Carlo methods for finding ground states of Ising spin glasses: Population annealing, simulated annealing, and parallel tempering, *Phys. Rev. E* **92**, 013303 (2015).
- [39] H. G. Katzgraber, M. Palassini, and A. P. Young, Monte Carlo simulations of spin glasses at low temperatures, *Phys. Rev. B* **63**, 184422 (2001).
- [40] E. Marinari, G. Parisi, F. Ricci-Tersenghi, J. J. Ruiz-Lorenzo, and F. Zuliani, Replica symmetry breaking in short-range spin glasses: Theoretical foundations and numerical evidences, *J. Stat. Phys.* **98**, 973 (2000).
- [41] M. Palassini and A. P. Young, Nature of the Spin Glass State, *Phys. Rev. Lett.* **85**, 3017 (2000).
- [42] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*, Monographs on Statistics and Applied Probability (Springer, Amsterdam, 1964).
- [43] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer Series in Statistics (Springer, New York, 2004).
- [44] L. A. Fernandez and V. Martin-Mayor, Mean-value identities as an opportunity for Monte Carlo error reduction, *Phys. Rev. E* **79**, 051109 (2009).
- [45] S. Boettcher, Stiffness of the Edwards-Anderson Model in all Dimensions, *Phys. Rev. Lett.* **95**, 197205 (2005).
- [46] T. Aspmeier, W. Wang, M. A. Moore, and H. G. Katzgraber, Interface free-energy exponent in the one-dimensional Ising spin glass with long-range interactions in both the droplet and broken replica symmetry regions, *Phys. Rev. E* **94**, 022116 (2016).