Evidence of Many Thermodynamic States of the Three-dimensional Ising Spin Glass

Wenlong Wang,1,2, * Mats Wallin,1 and Jack Lidmar1

1Department of Physics, Royal Institute of Technology, Stockholm, SE-106 91, Sweden
2College of Physics, Sichuan University, Chengdu 610065, China

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 behaviour 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 detailed examination and state-of-the-art simulation provide a coherent picture of many pure states, explain the previous findings, and the controversy is solved. The pertinent status of the number of pure states beyond this statistic is also discussed, and we find the spin glass balance is overall tilting towards many pure states studied by simulations.

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 orientation 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. Our approach is to focus on individual spin glass properties, and a solid answer on one property can put stringent constraint on possible theories. Therefore, in this work we solely discuss computationally the number of pure states. 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 state 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 (finite if there are many pure states) energy gaps between the lowest pure states. In this way, even $O(1)$ large-scale droplet excitations are forbidden at $T = 0$. This conjecture 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. As mentioned above, the overlap function is found to be nontrivial for the sizes available at finite temperatures. It is helpful to rephrase this important statement more precisely as: “The disorder-averaged overlap distribution function is nontrivial under periodic boundary conditions at a typical low temperature”. To our best knowledge, there is no numerical work that contradicts this result. Therefore, to argue against this statement, it is necessary that one or several of the conditions have to be altered. At finite temperatures, these efforts come broadly in three types: a new statistic, a new boundary condition, or a combination of the two. It should be mentioned that not all of these pioneering works concluded strongly against many pure states, some only argued their results may have violated many pure states, and many of these have already been solved. 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 $1/L^{\theta_{DW}}$ scaling, where $\theta_{DW}$ is the interface free energy exponent [22]. However, this was later found to be a finite-size effect from the surfaces [23]. The scaling no longer holds for larger sizes and $I(0.2)$ of the free boundary condition and periodic boundary condition run together for larger sizes, supporting many pure states. By contrast, various statistics have been proposed other than $I(0.2)$, but most of these are not very successful; see, e.g., [24] and the references therein for a collection of 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 pure 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 $\Delta$ 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 $\Delta$ appears to do the best job among the new statistics supporting the two state picture [24]. In [25], it was found that $\Delta(0.2, 1)$ [see Eq. (3)] at $T = 0.42$ grows with system size up to about $L = 10$, then it levels off or stops growing appreciably. By contrast, $\Delta$ of the SK model at a similar $T/T_C$ grows prominently for the similar range of sizes. A criticism in [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 $\Delta$ 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 investigated. 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 Letter 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 system sizes. In light of [26], data at a higher temperature are also collected for comparison. Using massively parallel population annealing Monte Carlo [30–34] and 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 [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 behaviour of $\Delta$. Motivated by our observation, we define a slightly modified $\Delta$ compensating effectively for this correlation effect. The modified $\Delta$ essentially grows monotonically, providing a coherent and coherent state picture. Our results also reveal that the notably different results of earlier works originate from the use of the different effective temperatures or central weights, as suggested by [26].

**Model, method, and observables–** We study the three-dimensional Edwards-Anderson Ising spin glass [7] defined by the Hamiltonian $H = -\sum_{ij} J_{ij} S_i S_j$, where $S_i = \pm 1$ are Ising spins and the sum is over nearest neighbours on a simple cubic lattice under periodic boundary conditions. For a linear size $L$, there are $N = L^3$ 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 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 $T = 0.42$. Population annealing is used because it is both efficient and massively parallel [33, 38]. Our criterion for equilibration is based on the replica family entropy $S_f \geq \log(100)$ and we ensure equilibration for each individual instance [28, 33]. The preliminary simulation parameters are summarized in Table I, and the 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.

**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 the unequilibrated instances were rerun with (much) larger simulation parameters; see the text for details.

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

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

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

where spin configurations “(1)” and “(2)” are chosen independently from the Boltzmann distribution or the population. Measuring the overlap distribution is a costly process in parallel computing, and 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]:

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

$$\Delta_{\mathcal{J}}(q_0, \kappa) = \max_{\|\eta\| \leq q_0} \left[ \langle P_{\mathcal{J}}(\eta) + P_{\mathcal{J}}(-\eta) \rangle / 2 > \kappa \right].$$  

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

Results– The disorder-averaged spin overlap function and the central weight $I(0.2)$ at both $T = 2/3$ and 0.42 are shown in Fig. 1. It is clear that the central weight is essentially flat up to fluctuations as a function of the system size, in agreement with numerous previous results [25, 33, 39, 40]. This result is well known, except that we have further extend it to two larger system sizes at low temperatures. The result is in excellent agreement with RSB. By contrast, this is strikingly different from the $1/L^\beta_{Dw}$ ($\theta_{Dw} \approx 0.2$) scaling predicted by the droplet picture. 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 $\Delta$ 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 [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 growth from $L = 10$ to $L = 12$ is very marginal (marginal or slightly negative depending on disorder realizations, here we call it collectively as marginal), in agreement with [25]. Then we observe a noticeable increase from $L = 12$ to $L = 14$, a somewhat promising result for many pure states. But the statistic subsequently grows rather marginally again from $L = 14$ to $L = 16$, resulting a rather confusing situation. This puzzle is finally understood after checking back and forth when we realize a subtle correlation between the $\Delta$ data and the central weight data. As illustrated in Fig. 1, $I(0.2)$ drops slightly from $L = 10$ to $L = 12$, this is where the corresponding $\Delta$ has a marginal increase. Then $I(0.2)$ increases slightly from $L = 12$ to $L = 14$ and then drops slightly at $L = 16$, explaining the correlated trend in $\Delta$. This together with other similar observations throughout our data collection process strongly suggest that the two statistics are indeed correlated. Then $\Delta$ 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 $\Delta$ has a larger growth rate at higher temperatures, note the crossings of the data at the two temperatures.

The correlation between $\Delta$ 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 as well by definition. On the other hand, if $I(0.2)$ for an instance is large, there is almost certainly a peak present, at least when the size is large. Since the fluctuation of $I$ has a profound influence on the statistic $\Delta$, we next look for modified statistics to compensate this correlation effect by a variance reduction method.

We define a slightly modified statistic of weighted $\Delta$ as $\Delta(q_0, \kappa) = \langle I(q_0) / I(q_0) \rangle$, where the angular bracket is an average over the system size. This definition takes advantage of the approximately constant nature of $I(q_0)$ and only slightly modifies the $\Delta$ data. Before we show that this statistic works well, we look at another example to first motivate this statistic. We look at two extreme windows that have very different weights. From the overlap distribution functions, it is clear that at the lower temperature there are wide $q$ ranges where the major $q_{FEA}$ peaks have little influence. In addition, the weight density is higher at larger $q$ than at the neighbourhood of $q = 0$. Therefore, we select the following two different
windows and study the behaviours of \( I \) and \( \Delta \) at \( T = 0.42 \): (1) Window A defined as the interval \( q \in [-0.1, 0.1] \) with a small weight and (2) Window B defined as \( |q| \in [0.4, 0.5] \) with the same length but a noticeably larger weight. Here, \( I \) and \( \Delta \) are measured in the respective supports. The two statistics as a function of 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 \( \Delta \) as expected. The ratio \( \Delta / I \) is shown in the bottom right panel, and this simple statistic brings the two \( \Delta \) data sets much closer, particularly for the pertinent large sizes. Similar behaviour is also found for the higher temperature, despite that Window B is slightly affected by the \( q_{\text{EA}} \) peaks. These demonstrate that \( \Delta / I \) is an effective statistic to reduce the correlation effect from the weights.

The modified \( \Delta \) is shown in the bottom left panel of Fig. 2. It is remarkable that this simple modified \( \Delta \) has a clean growth trend with increasing system size, i.e., the growing trend is not only improved but also the data are monotonic. To be more quantitative, we have carried out a growing trend test to leading linear order using a linear fit. 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. Note that all of these values, especially the high temperature data, are significantly larger than 0, suggesting a collective growth trend of this statistic. We conclude therefore that the seemingly nonmonotonic growth of \( \Delta \) is a result of its sensitivity to the fluctuations of the central weight. From this perspective, the statistic \( \Delta \) is not as good as the central weight in discriminating the number of pure states due to the discrete nature of its support.

The modified \( \Delta \) also shows clearly the effect of temperature on \( \Delta \). The growth rate at the higher temperature (red symbols) is noticeably larger than at the lower temperature (blue symbols). There is also an interesting crossing in \( \Delta \) at each temperature despite that there is no crossing in \( I \) at the two
temperatures, showing the complex relations of $I$ and $\Delta$ in general. Nevertheless, the crossings could be qualitatively understood by the two mechanisms of the sharpening of peaks either by increasing the system size or lowering the temperature. When the system size is sufficiently large, all peaks regardless of the temperature tend to be sharp and tall. In this case $I$ will sufficiently determine the order of $\Delta$, the $\Delta$ at higher temperatures will be larger as the weight is larger. On the other hand, peaks for small sizes tend to be wide and short at high temperatures and sharp and tall at low temperatures, with respect to the chosen cutoff $\kappa$. This provides a possibility that one could register more peaked instances at lower temperatures due to the sharper and taller peaks, despite that the central weight is smaller. This explains the crossings and leads to a conclusion that the growth rate of $\Delta$ is higher at higher temperatures than at lower temperatures.

Next we look at an alternative method for variance reduction [41] by considering $\Delta-I$. The results are shown in Fig. 3. Similar to the modified $\Delta$, this statistic also has a clean growth trend. By contrast, this statistic should converge to 0 for a two state picture. The data are clearly running above 0 and are still growing, and should reach finite limits if there are many pure states. Therefore, the statistics $I$, modified $\Delta$, and $\Delta-I$ are all coherently in agreement with the many state picture.

Our results are in full qualitative agreement with [25, 26]. The former group found a small growth rate of $\Delta$ 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 behaviours of $\Delta$ [25]. Our large range of sizes is crucial in identifying this subtle correlation. We conclude that the statistic $\Delta$ has no evidence of violating the many state picture, and instead it is consistent with a coherent picture of many pure states [26]. The controversy of $\Delta$ is solved.

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 free energy exponent $\theta_{\text{DW}}$ yet a flat $I(0.2)$. It is unclear what finite-size effect is responsible for a flat central weight, should one predict it would eventually decay as $1/L^{\theta_{\text{DW}}}$. (2) The $\theta_{\text{DW}}$ exponent is a growing function of dimensionality [42] and even the SK model [43] has a positive exponent which is clearly described by RSB [44]. It seems likely that $\theta_{\text{DW}} > 0$ is not capable of excluding large-scale $O(1)$ droplet excitations; see, e.g., [29] for an interesting possibility of how a positive domain-wall exponent and nontrivial overlap distribution functions can coexist. These difficulties in our opinion must be addressed for a two state picture to be acceptable and meaningful.

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 $\Delta$ in detail. Our results reveal that the nonmonotonic growth of the statistic as a function of the system size is a result of its sensitivity to the fluctuations of the central weight $I$. By looking at a modified $\Delta$ 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 $\Delta$ is in full agreement with the many state picture but not with the two state picture. With one more controversy solved and supporting (again) many pure states and the difficulties of the two state picture mentioned above, it appears that a coherent picture of many pure states is emerging and that the spin glass balance is significantly tiling towards many pure states by recent simulations. Further explorations such as the SK spin glass in thermal boundary conditions [28] are currently in progress and will be published in the future.

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), and the High Performance Computing Center North (HPC2N), and the Emei cluster at Sichuan university.

[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, 2013).
[3] D. Sherrington and S. Kirkpatrick, *Solvable model of a spin glass*, Phys. Rev. Lett. **35**, 1792 (1975).
