REPLICA EXCHANGE MONTE CARLO SIMULATIONS OF THE ISING SPIN GLASS:
STATIC AND DYNAMIC PROPERTIES

A Dissertation Presented

by

BURCU YUCESOY

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2013

Department of Physics

REPLICA EXCHANGE MONTE CARLO SIMULATIONS OF THE ISING SPIN GLASS: STATIC AND DYNAMIC PROPERTIES

A Dissertation Presented

by

BURCU YUCESOY

Approved as to style and content by:

Jonathan Machta, Chair

Helmut G. Katzgraber, Member

Murugappan Muthukumar, Member

Boris Svistunov, Member

Donald Candela, Department Chair
Department of Physics

ACKNOWLEDGMENTS

First and foremost, I want to thank my thesis advisor Jonathan Machta. When we first met, I told him in no uncertain terms that I wanted to work exclusively on spin glasses, and he was gracious enough to not see it as arrogance but, I hope, as dedication, and he accepted me as his student. Thanks to him I learned to be precise, to pay attention to details, to question and double check everything, and to always, always put error bars on plots. His calm demeanor and sound advice helped me to overcome many stressful situations, and I hope I have been a student worth his time.

Next, our main collaborator in this work, Helmut Katzgraber, has my gratitude. Thanks to his helpful insights, strong work ethic and great humor, it was a joy and an inspiration to work with him.

There are many other people who helped me along the way. I want to thank Ruben Andrist not only for teaching me the ropes in scripting but also for his company and to Don Blair for inspiring, in one case even life-changing, discussions. Katie Scilla made life in Amherst, or in the US in general, much more fun. She is the greatest friend one can hope for, thanks to her, Chris Scilla and Derek and Erin Breid I always felt at home, that I was included, and they gave me the courage to become a bigger geek than I already was. When they moved away, Katie Williams made me look forward to our weekly coffee, thanks to her, Amherst was still fun to come to.

Thanks to my parents and my sister for their encouragement and support, and not fleeing the room when I start talking about spin glasses yet again, or not making fun of me when I finally got it tattooed on my forearm.

I will always be grateful to Nihat Berker, for awakening the love of physics in me, for believing in me, and for being there with support and encouragement whenever I needed it the most. I will always strive to become the physicist he sees in me, and hope to make him proud.

A final thanks to the love of my life, Arto Suren. I can live the life as it should be lived, only because you are in it...

ABSTRACT

REPLICA EXCHANGE MONTE CARLO SIMULATIONS OF THE ISING SPIN GLASS: STATIC AND DYNAMIC PROPERTIES

SEPTEMBER 2013

BURCU YUCESOY
B.Sc., ISTANBUL TECHNICAL UNIVERSITY
M.Sc., ISTANBUL TECHNICAL UNIVERSITY
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Jonathan Machta

Spin glasses have been the subject of intense study and considerable controversy for decades, and the low-temperature phase of short-range spin glasses is still poorly understood. Our main goal is to improve our understanding in this area and find an answer to the following question: Are there only a single pair or a countable infinity of pure states in the low temperature phase of the EA spin glass? To that aim we first start by introducing spin glasses and provide a brief history of their research, then proceed to describe our method of simulation, the parallel tempering Monte Carlo algorithm. Next, we present the results of a large-scale numerical study of the equilibrium three-dimensional Edwards-Anderson Ising spin glass with Gaussian disorder. In order to understand how the parallel tempering algorithm works, we measure various static, as well as dynamical quantities, such as the autocorrelation times and round-trip times for the parallel tempering Monte Carlo method. We examine the correlation between static and dynamic observables for $\sim$ 5000 disorder realizations and up to 1000 spins down to temperatures at 20% of the critical temperature, and our results show that autocorrelation times are directly correlated with the roughness of the free-energy landscape. In the following chapters, the three- and four-dimensional Edwards-Anderson and mean-field Sherrington-Kirkpatrick Ising spin glasses are studied again via large-scale Monte Carlo simulations at low temperatures, deep within the spin-glass phase. Performing a careful statistical analysis of several thousand independent disorder realizations and using an observable that detects peaks in the overlap distribution, we show that the Sherrington-Kirkpatrick and Edwards-Anderson models have a distinctly different low-temperature behavior. We arrive to the following conclusion: The structure of the spin-glass overlap distribution for the Edwards-Anderson model suggests that its low-temperature phase has only a single pair of pure states. Finally we present results for several new observables, along with a few preliminary studies and suggestions for future research.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

Table **Page**

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

In 1988, Philip Anderson wrote in the columns of Physics Today [4]:

> The history of spin glass may be the best example I know of the dictum that a real scientific mystery is worth pursuing to the ends of the Earth for it's own sake, independently of any obvious practical importance or intellectual glamor.

As one does not judge a book by solely on its cover, a topic of research can or should not be judged solely on the apparent practical applications, or the perceived fame it would bring to the researcher. Every open question is worth pursuing, and what seems to be a dire prospect at first glance, can end up being a major discovery and a source of inspiration. It was certainly the case for spin glasses. Let us start from the beginning: A disordered magnetic material such as the spin glass does not seem very exciting at first glance. If it is metallic, it cannot conduct well, and as insulators, they are practically useless [66]. Then why have they attracted not only many condensed matter and statistical physics, but also complexity scientists and even mathematicians? Let us again quote from Anderson:

> The pursuit of the spin glass mystery led, *inter alia* and aside from all the good solid-state physics that resulted, to new algorithms for computer optimization, a new statistical mechanics, a new view of protein structure, a new view of evolution and new ideas in neurosciences.

And why that happened has many answers. First, these are seemingly simple systems, described by a one line equation, not all that different from the simplest of many body statistical models, the Ising model. Therefore, they are easy to picture conceptually and easy to start working on. The 'simplicity' ends there, however. Spin glasses are complex systems. They exhibit both frustration and quenched disorder, their behavior is very hard to predict. Together, these properties make spin glasses one of the simplest of complex systems, and consequently an ideal candidate for all complexity studies. Second, finding their ground state is a hard combinatorial optimization problem, therefore mathematical tools developed for those studies are invaluable for analyzing many real world problems. Finally, on the pure physics side, they represent a gap in our understanding of the solid state: The structural glass in general is a poorly understood material [66]. Even the basic question why a window glass is rigid when it does not have a crystalline structure has not been answered to everyones satisfaction. Here, the magnetic counterpart the spin glass, which can easily be described and modeled mathematically, can be a great starting point, before moving on to much complicated ordinary glasses.

Despite all the interest spin glasses generated over the years, and despite all the resources allocated to their research resulting in countless published work, these systems are still not fully understood. Today, nearly half a century after their first discovery, even for the simplest mathematical model there are many completing theories, each with considerable number of supporters. And therein lies our main motivation as well. The main purpose of this document is to understand the behavior of the spin glass at low temperatures, by studying the mathematical model constructed by Edwards and Anderson back in 1975, and to be able decide which of the proposed explanations is the correct one. To that end, we now present a brief historical description of spin glasses and their research so far. For a more detailed outlook, one can consult the many books and review articles on the topic, some of which we are referencing here [10, 43, 22, 48].

1.2 Spin Glasses

1.2.1 ‘Real’ Spin Glasses

The ‘real’ spin glasses are alloys, in which a non-magnetic host metal is mixed with trace amounts of magnetic transition metals. These ‘impurities’ then occupy random sites, and their magnetic dipoles or *spins* magnetically polarize the conduction electrons of the host metal. This polarization can be ferromagnetic at some distances, and antiferromagnetic in others, and as other impurity spins try to align themselves according to the thusly created local magnetic fields, interactions between impurities that are randomly ferromagnetic or antiferromagnetic are created. Historically, the first types of spin glasses that have been studied were hosts of copper (Cu) or gold (Au) with impurities of manganese (Mn) or iron (Fe), hence, we can consider the the alloys $\mathrm{Cu}_{1-x}\mathrm{Mn}_x$ or $\mathrm{Au}_{1-x}\mathrm{Fe}_x$ the 'classical' spin glasses. Here x denotes the concentration of impurities and is usually < 0.1 for spin glasses. And the classical spin-spin interaction $J(r)$ with respect to the distance r between spins is of RKKY-type (named for Ruderman, Kittel, Kasuya and Yoshida) and defined as

$$J(r) \propto \frac{\cos(Ar)}{Br^3}, \tag{1.1}$$

where A and B are constants related to the materials used. In summary, the couplings oscillate between ferro- and antiferromagnetic nature and their strengths diminish as $1/r^3$.

The research on spin glasses first started gain momentum in the 1970s following some surprising results observed in experiments: When subjected to a small oscillating magnetic field, the magnetic susceptibility χ (defined as $\chi = \delta M/\delta h$ where M is the magnetization and h is the external magnetic field) of CuMn and AuFe shows a cusp at a specific temperature [13]. This cusp that would increase in magnitude and in temperature as the concentration of impurities increases was surprising and needed to be understood, especially since the specific heat is not diverging at the same temperatures, showing that the energy fluctuations themselves are not divergent. This physically interesting phenomenon has been only observed at low temperatures; the high temperature behavior is paramagnetic, i.e. the spins rotate randomly, independent from each other. The spin glass transition has been found to occur at a well defined critical temperature T_c. Below T_c, the average magnetization is still zero, but unlike the paramagnetic phase where each spin is constantly changing directions, now the spins are frozen in their places, in random orientations. This creates local spontaneous magnetizations which are the cause for the decrease in χ at these low temperatures, even though no conventional structural long range order of ferro- or antiferromagnetic nature can be detected.

In summary, the essential features of spin glasses seem to boil down to two important ingredients: Randomness, and competing interactions. The couplings between spins have random signs, i.e. they are randomly of ferromagnetic or antiferromagnetic nature, and therefore, no spin configuration can satisfy all couplings simultaneously. This effect is called *frustration*. And as noted, no conventional long range order can be established, instead a freezing transition occurs to a state where the spins are frozen in random orientations. This freezing behavior actually gives the system its name; it is reminiscent of the translational disorder in the atomic arrangement of an ordinary glass.

1.2.2 The Edwards-Anderson Model

Edwards and Anderson are considered to have initiated the modern theory of spin glasses as we regard them now, by noting that the competition between quenched ferromagnetic and antiferromagnetic couplings define the essential physics, and not other microscopic details [17]. They proposed the prominent theoretical model for spin glasses, which is known as the Edwards-Anderson (EA) model. Here, spins are located on sites of an hypercubic lattice of D dimensions. The couplings J_{ij} are between nearest neighbor spins, and chosen according to a Gaussian distribution

$$P(J_{ij}) = \frac{1}{\sqrt{2\pi\delta}} \exp\Big(\frac{-(J_{ij}-\mu)^2)}{2\delta^2}\Big), \tag{1.2}$$

where the variance δ is usually unity and the mean $\mu = 0$. Alternatively, there is a binary version where J_{ij} is chosen to be $\pm J$ randomly. Since the couplings are chosen once and then keep their value, the disorder is called to be quenched. The standard Hamiltonian

$$\mathcal{H}_{\mathcal{J}} = -\sum_{\langle i,j \rangle} J_{ij} s_i s_j - h \sum_i s_i \tag{1.3}$$

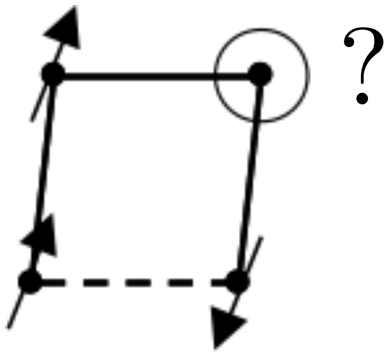

Figure 1.1. Example of a frustrated square, where solid lines denote ferromagnetic and dashed lines denote antiferromagnetic interactions.

is then sufficient to explore the nature of the state. Here, i is a site on the hypercubic lattice, s_i and s_j are the Ising spins on those sites that can be up or down, hence $s_i \in \{\pm 1\}$. The sum $\langle i,j \rangle$ is over the nearest neighbor pairs only, and h stands for the external magnetic field. We will, from now on, only consider the case $h = 0$ where the model has global spin-flip symmetry, provided that μ defined above is also zero. A positive J_{ij} indicates a ferromagnetic and a negative one an antiferromagnetic coupling, and since these are randomly distributed, competitions arise. An example of the frustration caused by this competition can be seen in Fig. 1.1. In this square, one coupling (dashed line) is negative and the remaining three (solid lines) are positive, therefore, aligning the top right spin correctly to satisfy both neighboring couplings is not possible.

For this model, under the freezing temperature T_c, the magnetization per spin M would always be zero, and consequently a new order parameter is needed to define the spin glass order. Edwards and Anderson suggested that

$$q_{EA} = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \langle s_i \rangle^2, \tag{1.4}$$

would be sufficient. Here N is the total number of spins and $\langle\ \rangle$ denotes a thermal average. This way one can identify the two properties of the spin glass phase as $M = 0$ and $q_{EA} > 0$.

Even after more than three decades of research, this simple model is not fully understood. It is now known that for $D = 2$, there is no spin glass phase for $T > 0$ [64]. For the higher dimensional models, however, finite transition temperatures exist [8], but how these models would behave deep within the spin glass phase is still a topic of high interest. There are competing theories, first of which was developed based on the mean field approximation of the model.

1.2.3 The Mean Field Theory and the Replica Symmetry Breaking Picture

One important step in understanding spin glasses was a model developed by Sherrington and Kirkpatrick in 1975 [63]. They proposed that the correct mean field approach to the EA model would be a model where all spins are coupled to each other regardless of their distance. This model is called the Sherrington-Kirkpatrick (SK) model, and even though it is physically unrealistic to explain real spin glasses, solving it should help to establish a theory to build upon by adding corrections due to short range interactions later.

To solve this model, one needs first to determine the average free energy, and since this is a system with randomness, an average over disorder realizations has to be taken in addition to the usual thermal averaging. To that end, we start with the partition function

$$Z = \sum_J \exp\left(\frac{-\mathcal{H}_J}{k_B T}\right), \tag{1.5}$$

where k_B is the Boltzmann factor and the sum is over all possible disorder realizations, and then take the logarithm to obtain the free energy

$$F = -k_B T \ln Z. \tag{1.6}$$

To calculate a disorder average of this logarithm is extremely difficult, so instead Sherrington and Kirkpatrick came up with an approximation called the *replica ansatz*. Here, one studies the n-th power of the partition function, which can be averaged easily and then use the following equality

$$\ln Z = \lim_{n \to 0} \left[\frac{1}{n} (Z^n - 1) \right]. \tag{1.7}$$

The idea behind is that the n-th power would represent n identical replicas of the original system, and after the averaging process, the $n \to 0$ limit can be taken. This is called the *replica symmetric solution*, since all the replicas are treated as equal.

This solution captured the essence of the experiments by showing the cusp in susceptibility at the critical temperature, however, at very low temperatures, the entropy would become unphysically negative. Four years later, Georgio Parisi finally came up with a correct solution, which is now called the *replica symmetry breaking* (RSB) solution [56, 58, 57, 59], since now the replicas are considered not to be identical. Instead, they correspond to the many pure states that exist at low temperatures. A pure state means an equilibrium distribution in the thermodynamic limit that cannot be expressed as a linear combination of other distributions and according to the RSB solution, there has to be an infinite number of those pure states resembling or differing from each other in varying degrees. One can then define an order parameter by comparing those states to each other by measuring their spin overlap. If we consider two such states α and β, their spin overlap will then be

$$q_{\alpha,\beta} = \frac{1}{N} \sum_i \langle s_i \rangle_\alpha \langle s_i \rangle_\beta, \tag{1.8}$$

where the $\langle \; \rangle$ denotes a thermal average. One can see that for $\alpha = \beta$, one recovers the Edwards-Anderson order parameter q_{EA}. Since the spin overlap can only compare two states at a time, it is more informative to consider its probability distribution. RSB states that at finite but low temperatures, for a given disorder realization, only a few among the infinite number of states will have weights large enough to be noticeable. Therefore, even in the thermodynamic limit, the positions and weights of all overlaps between $\pm q_{EA}$ will depend on the disorder realization J and the overlap probability distribution $P_J(q)$ will not have a unique shape (Fig. 1.2, left), i.e. it will be non-self-averaging. The disorder average $P(q)$ will have a continuous part and a non-zero value near $q = 0$, which becomes a constant independent of system size for large L (Fig. 1.2, right). The RSB solution fully explains all aspects of the SK model and is now known to be the correct mean field solution. It is from this exact solution that the RSB picture for the EA model is constructed [10, 43]. There, we have a system with short range interactions, and the RSB picture states that there must be again a countable infinity of pure states with varying overlaps. Consequently, the overlap probability distribution will be non-self-averaging and fluctuations between disorder realization samples will always be significant, independent of system size. Even though these features are proven to hold for the SK model, they are only conjectured for the short range models. In fact, in the following section we present a competing theory with very different implications.

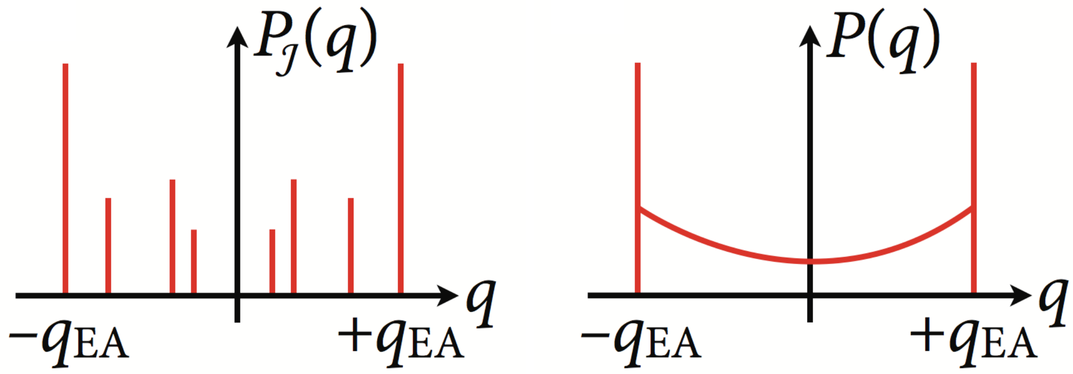

Figure 1.2. The left-hand side picture is a sketch of the overlap probability distribution as described in the RSB picture $P_J(q)$ for a sample disorder realization J. For different realizations, the states will also be different, thus changing the location of the inside peaks. On the right, a sketch of $P(q)$ as a disorder average over all $P_J(q)$ is shown.

1.2.4 The Droplet Picture

This scenario, called the 'droplet' or 'scaling' picture, proposes a phenomenological scaling approach and has been independently developed by McMillan [42], Bray and

Moore [12] and Fisher and Huse [19, 20, 21]. It is based on an analysis of how the energy cost of flipping a large region of spins would depend on the size of that region. Here, the lowest energy excitation spanning the length L around a particular spin site is called a *droplet*. To flip one such droplet one needs an energy of order L^{θ}, where θ is the droplet exponent. Additionally, these droplet excitations are not space-filling since their surface is a fractal of dimension d_s, where d_s is smaller than the space dimension. In comparison, in the RSB picture the energy of flipping a cluster of length scale L would remain of order $\mathcal{O}(1)$ even if L grows infinitely big ($\theta_{RBS} = 0$), and the surface dimension d_s of the excitations would be the same as the space dimension.

The main implication of the droplet picture is that in the spin glass phase, there can only be one pair of pure states which are related by the spin flip symmetry, like in a ferromagnet. Consequently, the resulting $P(q)$ will be self-averaging and in the thermodynamic limit it will solely consist of two δ peaks at $q = \pm q_{EA}$ (Fig. 1.3). For finite systems, the droplet

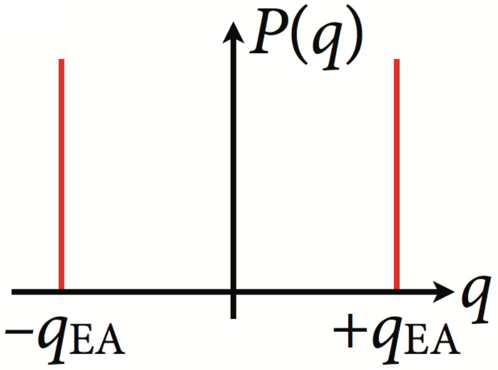

Figure 1.3. Sketch of the overlap probability distribution as described in both the droplet and the chaotic pairs pictures.

exponent θ relates the domain wall excitation energy to the system size, and from this relation the droplet picture predicts that the temperature and size dependence of $P(q = 0)$ should scale like $TL^{-\theta}$. θ can be measured using different excitations generated by changing the boundary conditions (BC), e.g. from periodic to antiperiodic. According to the droplet picture, such a change would have no effect on spins far away from the boundary.

1.2.5 Intermediate Theories

So far we have seen two theories about the low temperature of the EA spin glass that are firmly in disagreement with each other. In this section we consider intermediate theories attempting the reconcile those differences.

One of the most prominent of these is the 'chaotic pairs' picture, which was developed by Newman and Stein based on rigorous mathematical calculations [49, 50, 51, 52, 53]. Their *metastates* idea can be seen as the unifying theory, which is intended to move the focus away from the traditional thermodynamic states to probabilistic distributions of those, i.e. ensembles of thermodynamic states. Here, each finite volume sample can have only a single pair of spin-flip related states, but unlike the droplet picture, now that pair will vary chaotically with system size L, giving this picture a very different thermodynamic structure. It could actually be considered a many-state picture, but unlike the RSB, only a single pair of states appear in a large volume and the chaotic size dependence of that pair results in not a countable, but an uncountable infinity of states. Here the overlap probability distribution $P(q)$ cannot distinguish between the droplet and the chaotic pairs pictures as can be seen from Fig. 1.3.

Another such intermediate theory is the 'trivial-nontrivial' picture, developed by Krzkala and Martin [37], and Palassini and Young [55]. Here, the energy of the low-lying excitations does not increase with the size of the excitation like in the droplet picture, but unlike the RSB picture, those excitations have fractal surfaces, i.e. they are not space-filling. Additionally, a new scaling exponent θ' is proposed. In this scenario, unlike the original θ which measures the change in the energy of a cluster when it is excited by changing the boundary conditions, θ' measures the change in the energy of a cluster when the excitations are happening within the system with fixed boundaries. For the droplet picture, these two exponents θ and θ' should be the same [47].

1.3 Our Approach

Now that we have summarized the theoretical framework and predictions about the nature of the low temperature phase of spin glasses, we can concentrate on our contribution which is presented in the following chapters. Our main focus is to decide which of the proposed theories correctly describes the number of pure states, and to specifically answer the following question: Are there only a single pair or a countable infinity of pure states in the low temperature phase of the EA spin glass?

Our research is based on numerical simulations done using the parallel tempering Monte Carlo algorithm [23, 40, 25]. Therefore, in Ch. 2 we explain this algorithm in detail as well as the way we implement it to simulate spin glasses. We go into the details of how the algorithm was developed and how it works as it is applied to spin glasses, the advantages of using this method as opposed to single spin flip Monte Carlo schemes, and also explain how we implemented it.

The following chapter, Ch. 3, explains the dynamics at work when we are simulating spin glasses. The main concern there, in general, is to ensure equilibrium for all disorder realizations simulated, since unequilibrated samples may lead to erroneous conclusions by having significant effects on the averages and fractions calculated. We therefore investigate several time scales such as the round trip and autocorrelation times and their validity as markers of equilibration, as well as possible factors contributing to the long equilibration times known to be expected when simulating systems with complex free energy landscapes such as spin glasses. We lay out a rigorously tested, effective method for ensuring full equilibration, which we then proceed to use in our subsequent research. The contents of this chapter has been published as a paper in 2013 [74].

In Ch. 4 we finally approach our main question, the single or many pairs problem. To that end, we introduce a new statistic that clearly distinguishes between the two scenarios and then present results from simulations of the 3D EA as well as the SK model. For the SK model, our main purpose is to observe the properties described by the RSB picture, since it is known to be the exact solution of the model. We then compare those properties to the observations made on the 3D EA model at same reduced temperatures to see if the two models behave in a similar fashion, i.e. if the mean field approach can fully explain the low temperature phase of the EA model as well. The contents of this chapter has been published as a paper in 2012 [73] and 2013 [72].

The next chapter, Ch. 5, contains complementary material to Ch. 4 based on our most recent research done after our paper [73] was published, as well as some preliminary explorations to motivate future projects. First of all we show results from another simulation, this time the 4D EA model, which we then compare to previously simulated models. Second, we investigate several new observables that have been suggested by Middleton [45] and Monthus [46], using our data. We then consider the behavior of observables within a small window embedded in a larger system and how its behavior will be affected by size, to address a concern first raised by Ref. [49], which states that for small systems, boundaries will introduce domain walls to the system causing droplet-like excitations, thus affecting the results. Finally we present a preliminary look at the effects of changing the boundary conditions on systems with fixed bond configurations.

Our final chapter contains a summary of our conclusions and several future project ideas worth exploring, motivated by those conclusions. You can find additional details such as proof of detailed balance for the parallel tempering Monte Carlo and information about how we create, store and analyze the data in the appendices.

CHAPTER 2

NUMERICAL METHODS

2.1 Introduction

Our method of simulation is the parallel tempering algorithm. This method is also known as the replica exchange Monte Carlo and was independently introduced by Geyer [23], Marinari and Parisi [40], as well as Hukushima and Nemoto [25]. Furthermore, an earlier algorithm with many of the same essential features was introduced by Swendsen and Wang [67]. In parallel tempering, many copies of the system at different temperatures are simulated in parallel. Each copy (henceforth referred to as 'replica') is simulated using a standard, single-temperature Monte Carlo scheme such as the Metropolis or heat bath algorithms [36].

In addition to the single-temperature Monte Carlo moves, there are also *replica exchange* moves in which replicas at neighboring temperatures swap their temperatures. These moves permit replicas to diffuse in temperature space [32], from the lowest temperature to the highest temperature and back again. These *round trips* enhance mixing and greatly reduce equilibration times for systems with rough energy landscapes. Another advantage of this method is that it allows one to check the temperature dependence of observables since it provides data for various temperatures in a single run.

In this chapter we first explain the parallel tempering algorithm in detail, and then show how we implement this algorithm to simulate the Ising spin glass, using the C++ programming language.

2.2 Parallel Tempering Algorithm

Parallel tempering is a powerful tool for simulating systems with rough free-energy landscapes [16, 32, 38], with applications across many disciplines. To date, it is the most efficient method for simulating spin glasses and other disordered systems in more than two dimensions at low temperatures where dynamics are slow.

The algorithm works as described in Fig. 2.1: Several replicas of the system with the same disorder are simulated in parallel, each at a different temperature. To do this, a standard Markov chain Monte Carlo (MCMC) technique such as the Metropolis or heat bath algorithm [54, 36, 29] is used. In this work we employ the heat bath algorithm because of its better performance at low temperature. The set of temperatures is chosen to span both the low temperatures of interest, where equilibration is not feasible using single-temperature MCMC methods, and high temperatures where equilibration is fast. In addition to the MCMC sweeps at each temperature, there are also *replica exchange* moves between replicas. A proposed replica exchange move involves swapping the temperatures of two replicas at neighboring temperatures. In order to satisfy detailed balance [54] and ensure that the entire set of replicas equilibrates, the probability for accepting such a swap needs to be

$$p = \min[1, e^{(\beta-\beta')(E-E')}]. \tag{2.1}$$

Here $\beta = 1/T$ is the inverse temperature of one replica, β' is the inverse temperature of the neighboring replica, and E and E' are the corresponding energies of the two aforementioned replicas. These moves permit replicas to diffuse in temperature space [32] and this diffusion process permits an enormous speed-up relative to single-temperature Monte Carlo. Instead of directly surmounting the high free-energy barriers present at low temperatures, barriers are indirectly surmounted via the following process: A replica in one free-energy well at the lowest temperature diffuses to the highest temperature where it easily moves to another well and then diffuses back to the lowest temperature. This process is called a “round trip”. The whole diffusion process allows parallel tempering to reduce exponential time scales for surmounting free-energy barriers to power law time scales [38], thus reducing the equilibration time tremendously, even though that time is still growing exponentially with system size. As we explain in our dynamics chapter (Ch. 3), a complete round trip is not always necessary to ensure equilibration.

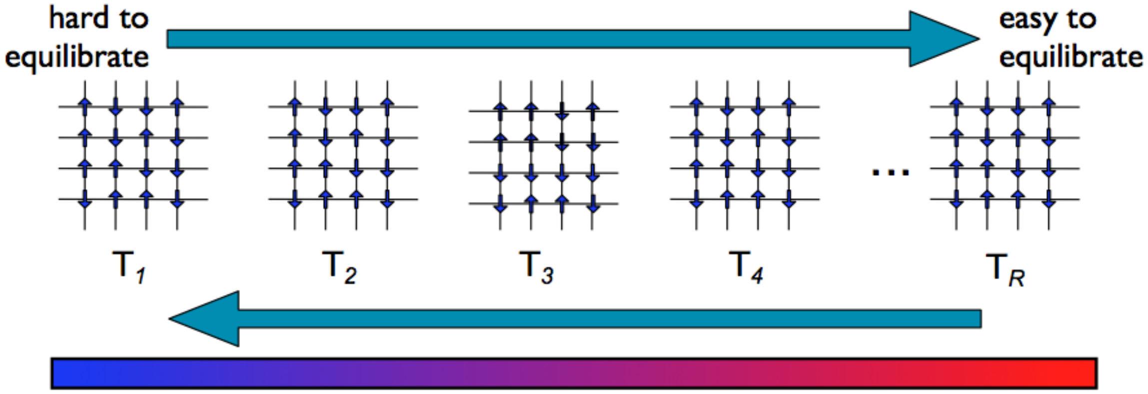

Figure 2.1. A sketch of how parallel tempering algorithm works: R copies of the system at different temperatures are simulated in parallel. The set of temperatures includes both high temperatures (where rapid equilibration is possible) and low temperatures of interest (where equilibration is hard). The replica exchange moves (swapping of two adjacent copies) permit replicas to diffuse in temperature space, as shown by the large arrows.

As noted, Equation (2.1) ensures that the entire Markov chain, including both single-temperature MCMC sweeps and replica exchange moves, converges to equilibrium and satisfies detailed balance. A proof of the latter can be found in Appendix A.

The parallel tempering algorithm has free parameters that include the set of temperatures and the ratio of single-temperature sweeps to replica-exchange sweeps. Optimizing parallel tempering therefore requires an appropriate choice of these parameters. The choice of the number of replicas N_T and their temperatures involves the following trade-off: If there are many closely-spaced temperatures, the energy distribution between adjacent temperatures overlaps strongly and the probability that a proposed replica exchange is accepted, see Eq. (2.1), is large. This suggests that the more replicas one uses, the better apart from the extra computational work involved in carrying out single-temperature sweeps of each replica. However, the motion of replicas in temperature space is diffusive such that the time scale for a round-trip scales approximately as the square of the number of temperatures. At least for simple model systems, parallel tempering is optimized if the number of temperatures scales as the square root of the number of spins [38], and a geometric progression of temperatures is used [29]. The average time to complete one round trip is often used to characterize the performance of parallel tempering, therefore choosing a set of temperatures that minimizes the round trip time is also one of the ways proposed to optimize parallel tempering [32, 70, 11]. However, in Ch. 3 we find that using the round trip time for optimization purposes can be misleading. Instead, the acceptance ratio of proposed swaps as a function of temperature can be considered a good measure for efficiency. Ideally, this ratio should be between $\sim 20-80\%$ and roughly independent of temperature, and large fluctuations should be avoided not to break up the random walk into segments in the temperature space. As a result, the best method for deciding on a temperature set is to run a preliminary, short simulation to test these acceptance ratios, starting with an initial, geometrically progressing temperature set. One can then use the information for further adjustments, such as adding intermediate temperatures where the ratios are low.

2.3 Implementation

In this section we explain the way we implemented the parallel tempering algorithm, using the C++ programming language.

Our program starts with an initialization function. In this function, all couplings J_{ij} are randomly selected from a Gaussian distribution of mean zero and standard deviation unity and all spins in every replica are chosen randomly as ±1. We store the random number generator seed for the couplings to identify the different disorder realizations, so we can reproduce those runs if necessary. Then, the initial energies are calculated for all R replicas in both sets.

A single time-step of the simulation consists of nested loops with two characteristic moves: First, we have the *sweep* move for the single-temperature MCMC updates, where we update each spin in a replica sequentially using the heat bath algorithm. Second, we repeat the *swap* move corresponding to the replica exchange between temperatures $R-1$ times, choosing one random neighboring pair to swap each time. This can also be done sequentially, without disturbing detailed balance. In our simulation, one sweep move for every replica is then followed by $R-1$ swap moves, and together they make up one time-step, which is shown in the code below:

```
for (s=0; s<2; s++) {
    for (t=0; t< R; t++)  sweep(t, s)
    for (i=0; i<(R-1); i++) swap(s);
}
```

The variable s here distinguishes between the two sets of replicas of the system we are simulating, since we need two separate sets to calculate our order parameter. The variable t accounts for which of the replicas we are updating, and the variable i accounts for the number of swap moves.

Before going into the details of the two moves, let us first explain how we assign temperatures to replicas in order to implement the parallel tempering algorithm. We have a two dimensional array called $\mathrm{map[t][s]}$, where the second index is just the set number. The first index points to the temperature we are working at, referring to the temperature array $\mathrm{T[t]}$. The value of the mapping array points to the replica 'r' we are using, such that $\mathrm{r{=}map[t][s]}$. As such, the $\mathrm{swap(\ )}$ function simply updates this mapping array in accordance with the probability given in Eq. 2.1. Both $\mathrm{sweep(\ )}$ and $\mathrm{swap(\ )}$ functions are given below.

```
void swap(int s)
{     int r, pair, temp;
      double P, rand;

      // select a neighboring replica pair at random:
      pair=random_range(0, R-2);

      //compute the probability of a swap:
      P=(1./T[pair]-1./T[pair+1])*(E[map[pair][s]][s]-E[map[pair+1][s]][s])

      //carry out swap:
      if (P>=0) {
         temp=map[pair][s]; map[pair][s]=map[pair+1][s]; map[pair+1][s]=temp;
      }
        else { rand=random( );
               if (rand<exp(P)) {
```

```
17                  temp=map[pair][s];  map[pair][s]=map[pair+1][s];  map[pair+1][s]=temp;
18              }
19      }
20  }
```

```
1   void sweep (int t, int s)
2   {
3       int r;
4       double delta=0.0;
5       r=map[t][s];
6
7       for(int i = 0; i < N; i++){
8
9        //compute the energy difference if spin is flipped:
10       delta = 2.*(double)s[i][r][s]*h[i][r][s];
11
12       //use delta to compute probability and carry out the flip:
13       if (((1. + exp((1./T[t])*delta))*random()) < 1.0){
14             s[i][r][s]=-s[i][r][s];
15             for(int j = 0; j < N; j++){
16                 h[j][r][s] += s[i][r][s]*J[j][i];  //update h accordingly
17           }
18             E[r][s]+= delta;   //update energy accordingly
19         }
20
21      }
22  }
```

Here $\mathrm{h[i][r][s]} = (1/2)\sum_j s_j^{s,r} J_{ij}$, where the sum is over all neighbors of s_i; and $\mathrm{E[r][s]}$ stands for energy, both for the replica r of set s. Please note that the way h is updated in the above code is intended for the fully connected SK model, where the sum in line 15 is over all spins N. For finite dimensional models such as the EA it would be more efficient to sum over the number of neighbors N_b, and compute h using only the neighboring spins.

In Markov chain Monte Carlo simulations such as ours, right after the initializing part there needs to be an initial thermalization where the code is run for many time-steps without taking any measurements. This is done to ensure independence from the initial conditions. In our case, we employ equal amounts of thermalizing and measuring time-steps. Also, we typically measure observables every 100^{th} time-step to minimize correlations. Additional details about how we generate, analyze and present the data can be found in Appendix B.

CHAPTER 3

DYNAMICS

3.1 Introduction

Systems in statistical physics with rough free-energy landscapes, such as spin glasses, are notoriously difficult to equilibrate using Markov chain Monte Carlo methods. The basic reason for this difficulty is that extremely long times are required to surmount free-energy barriers and thus to efficiently and fully explore the minima in the free-energy landscape. The parallel tempering algorithm (also known as "replica exchange Monte Carlo") [23, 40, 25], as well as other multicanonical techniques [28, 7] have partially overcome this problem. However there is little understanding of the relevant time scales for these methods. The purpose of the work presented in this chapter is to shed light on the equilibration time of the parallel tempering algorithm and to understand how it is related to equilibrium properties such as the roughness of the free-energy landscape. To this end, we apply parallel tempering to a large ensemble of disorder realizations of the three-dimensional Edwards-Anderson (EA) Ising spin glass [17, 10]. Our main results are that equilibration times vary significantly from one disorder realization to another and that this variation is correlated with the structure of the free-energy landscape. The fact that equilibration times are broadly distributed has been previously noted in a different context (see, for example, Refs. [1] and [2]) but the extent of variation was not fully appreciated because different measures of equilibration time were used in past studies. The correlation of dynamic time scales for parallel tempering and equilibrium properties has not been previously recognized, again because of the use of insensitive measures of dynamic time scales. In this study we compute autocorrelation functions of observables to measure equilibration times whereas previous studies have primarily relied on the round-trip time and related quantities. Although our study is focused on parallel tempering and the Edwards-Anderson Ising spin-glass model, we believe that our qualitative results are likely to be relevant both to other disordered systems with rough free-energy landscapes, as well as other multicanonical simulation techniques such as umbrella sampling methods [69], the Wang-Landau algorithm [71], and simulated tempering [40].

In the context of simple models with rough free-energy landscapes it has recently been shown [38] that the equilibration time of parallel tempering depends strongly on the structure of the free-energy landscape. If there is a single dominant minimum in the landscape and no first-order transition within the range of temperatures, then the equilibration time is short. On the other hand, the presence of several nearly degenerate minima or the existence of a first-order transition leads to much longer equilibration times. These results in simple model landscapes prompted us to look for similar phenomena in the Edwards-Anderson model where different disorder realizations may have quite different free-energy landscapes. In this study, rather than measuring the free-energy landscape directly, we used the spin overlap (order parameter) distribution as a proxy. The overlap distribution is closely related to the free-energy landscape but is substantially easier to measure in simulations. When the free-energy landscape has a simple shape consisting of only one minima or 'well', the system will have only one pair of preferred spin configurations related by spin-flip symmetry, and correspondingly the overlap probability distribution will only have two peaks at $\pm q_{EA}$ (Fig.3.2 (a)). If there are many minima in the free-energy landscape, however, systems will spend considerable time in different wells. Consequently, their spin configurations will differ or overlap in varying degrees, resulting in many peaks of various heights and shapes in the overlap probability distribution (Fig.3.2 (b)). Based on the results from simple model free-energy landscapes, we hypothesize that when the overlap distribution has a complex structure implying many free-energy minima the time scale for parallel tempering will tend to be longer than when the overlap distribution is simple. Our results confirm this hypothesis.

Furthermore, we have measured both equilibration times and overlap distributions for all disorder realizations and correlated these quantities. Finally, we also demonstrate that the conservative use of the equilibration criterion introduced in Ref. [31] is sufficient to ensure that nearly all disorder samples are equilibrated.

Our results suggest both a method for improving parallel tempering and a warning when using it in spin-glass simulations. On the one hand, we find that many disorder realizations have quite short equilibration times. Thus it might be useful to implement an adaptive scheme where some disorder realization are simulated for shorter times than other realizations. State-of-the-art parallel tempering simulations of equilibrium spin glasses require huge amounts of CPU time because of the difficulty of reaching equilibrium and the need for a large ensemble of disorder realizations. Thus, such an adaptive scheme has the potential for large savings in CPU time. On the other hand, the fact that the overlap distribution (a quantity directly related to the controversy surrounding the low-temperature phase of spin glasses) is significantly correlated with the dynamics of the algorithm serves as a warning that one must be extremely careful to ensure that essentially all samples are well equilibrated in order to avoid systematic errors in measuring disordered-averaged equilibrium properties.

3.2 Model and Methods

3.2.1 The Edwards-Anderson Ising spin glass

We study the three-dimensional Edwards-Anderson (EA) Ising spin-glass model, defined by the energy function

$$H = -\sum_{\langle i,j \rangle} J_{ij} s_i s_j. \tag{3.1}$$

The sum is over the nearest neighbors on a simple cubic lattice with periodic boundary conditions and side length L. The Ising spins s_i take values ± 1 and the interactions J_{ij}

are quenched random couplings chosen from a Gaussian distribution with zero mean and variance one.

3.2.2 Simulation parameters

For the parallel tempering algorithm (described in detail in Ch. 2) we use $N_T = 16$ temperatures between $T = 0.2$ to 2.0 [1]. This set of temperatures was chosen heuristically in Refs. [31] and [33] to perform well for $L = 6$ and 8. The corresponding average acceptance fractions for replica exchange moves are plotted in Fig. 3.1. It is likely that the simulations would be more efficient if more replicas were used near and above the critical region ($T_c \approx 0.96$ [30]) where the acceptance fractions are small. However, we believe that as long as the swapping probability is nonzero, our results will not change qualitatively. A bottleneck merely slows down the diffusion process but does not prevent it. For $L = 10$, the small acceptance fractions suggest that a larger number of temperatures might be more efficient. However, our goal is to understand correlations between static and dynamic quantities in parallel tempering for a given system size and so we have not sought to optimize the algorithm for each system size. We believe that the qualitative results obtained using our parallel tempering parameters would also hold for other reasonable values of the parameters.

By placing the hottest system temperature at $T \approx 2T_c$, we ensure that equilibration happens quickly, and we use one exchange sweep for each heat bath sweep ($N = L^3$ attempted spin updates). To calculate the spin overlap [Eq. (3.2)] we use two copies of the system at each temperature, therefore for each sample we simulate two sets of replicas independently from each other.

As stated before, we employ one replica exchange sweep after one heat bath sweep [32]. A heat bath sweep corresponds to sequentially attempting to update each spin at each replica once using the heat bath algorithm as shown in Ch. 2. An exchange sweep corre-

---

[1] $T \in \{0.20, 0.27, 0.34, 0.42, 0.50, 0.60, 0.70, 0.82, 0.94, 1.00, 1.20, 1.34, 1.50, 1.66, 1.82, 2.00\}$

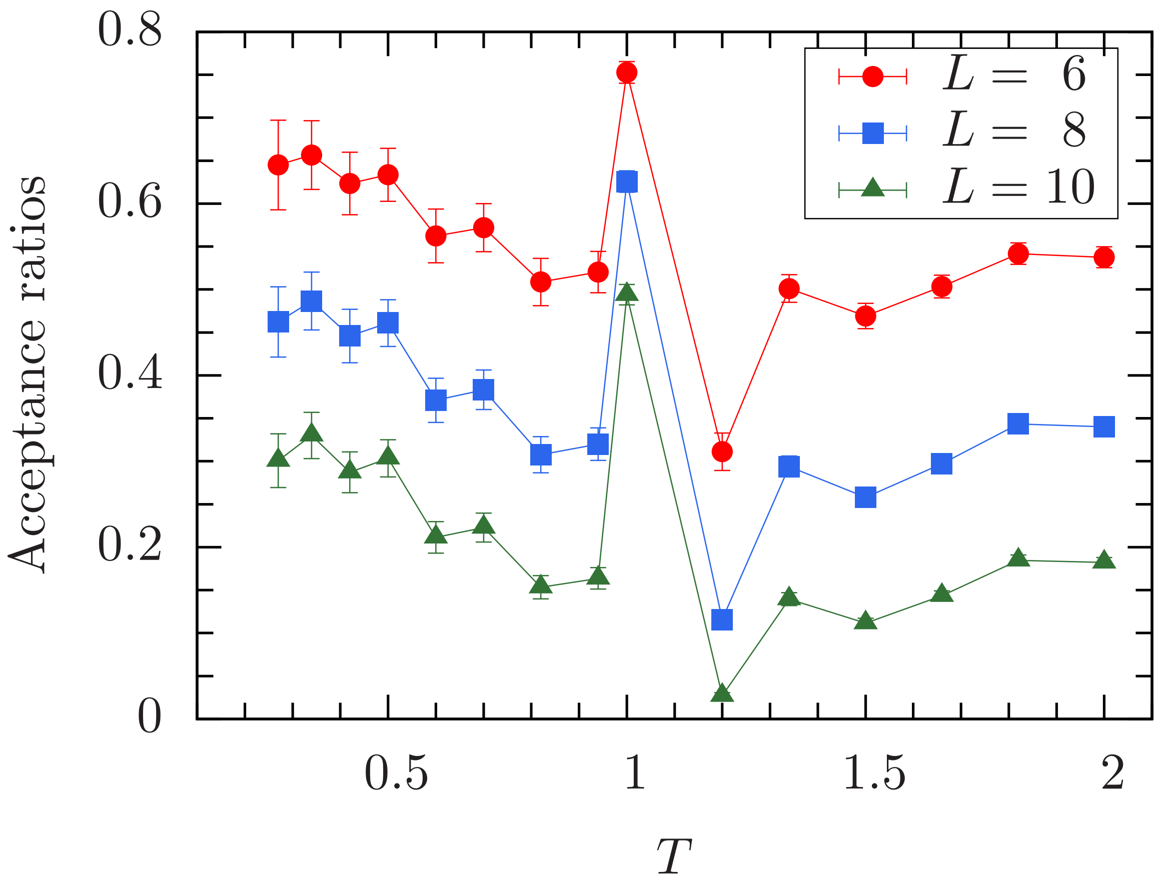

Figure 3.1. Average acceptance fraction as a function of temperature T for replica exchange moves. Note that $T_c \approx 0.96$.

sponds to randomly choosing $N_T - 1$ pairs of adjacent replicas and proposing exchanges for each. Both types of sweeps together make up one parallel tempering sweep. When no confusion arises we will call this unit of time simply a *sweep*.

For each system size we simulate approximately 5000 disorder realizations or *samples*. The simulations are equilibrated for at least 2^{24} sweeps for $L = 6$ and 2^{27} sweeps for $L = 8$ and 10, then measurements are done for the same number of sweeps, see Table 3.1 for details. A fraction of samples required longer runs to meet the equilibration criteria discussed in Sec. 3.2.3.2. For samples with very short equilibration times we performed runs to obtain fine-grained autocorrelation information, see Sec. 3.2.3.2. In total, the data collection took approximately 140 CPU years on 12-core AMD Opteron 6174 CPUs.

Table 3.1. For each system size L we equilibrate and then measure for at least 2^b Monte Carlo sweeps. $T_{\min}$ ($T_{\max}$) is the lowest (highest) temperature used, N_T is the number of temperatures, and N_{sa} is the number of disorder realizations. For some $L = 10$ samples longer runs had to be performed to ensure equilibration.

L	b	$T_{\min}$	$T_{\max}$	N_T	N_{sa}
6	24	0.2	2.0	16	4961
8	27	0.2	2.0	16	5126
8	28	0.2	2.0	16	4
10	27	0.2	2.0	16	4360
10	28	0.2	2.0	16	353
10	29	0.2	2.0	16	241
10	30	0.2	2.0	16	73

3.2.3 Observables

3.2.3.1 Overlap distributions

The focus of this study is to relate static equilibrium properties of individual spin-glass samples to the dynamics of the parallel tempering algorithm acting on that sample. The primary quantity that we measure to study both equilibrium properties and the dynamics of parallel tempering is the spin overlap q,

$$q = \frac{1}{N} \sum_{i=0}^{N} s_i^{\alpha} s_i^{\beta} \tag{3.2}$$

where $N = L^3$ is the number of spins and the superscripts α and β indicate two independent copies of the system with the same disorder. The thermal and disorder average of the overlap is an order parameter for the system and the probability distribution of the thermally averaged overlap $P_J(q)$ for a given sample J reveals aspects of the free-energy landscape for that particular sample. The behavior of $P_J(q)$ for large systems and at low temperatures is one of the major open questions in the theory of spin glasses. Overlap distributions vary widely from sample to sample, as can be seen in the three examples shown in Fig. 3.2. Although there is no direct mapping between the free-energy landscape and $P_J(q)$, it is clear that numerous peaks in $P_J(q)$ imply numerous minima in the free-energy landscape.

For example, samples corresponding to Figs. 3.2(b) and 3.2(c) have a more complicated free-energy landscapes than Fig. 3.2(a). Our central hypothesis is that samples with a more complicated free-energy landscape tend to have longer dynamic time scales.

From $P_J(q)$ we define $I_J(q_0)$ for different q_0,

$$I_J(q_0) = \int_{-q_0}^{q_0} P_J(q). \tag{3.3}$$

$I_J(q_0)$ is the weight of $P_J(q)$ in an interval near the origin and it serves as an approximate measure of the complexity of $P_J(q)$. For example, $I_J(0.2) \approx 0$ for the samples shown in Figs. 3.2(a) and 3.2(b), while $I_J(0.2)$ is large for the sample shown in Fig. 3.2(c). We compute $I_J(q_0)$ for eight values of q_0 from $q_0 = 0.2$ to 0.9. It is worth noting that $I_J(q_0)$–especially with $q_0 = 0.2$–has been extensively used in studies of the low-temperature equilibrium properties of the EA model [31, 3].

3.2.3.2 Characteristic time scales

We measure two time scales for parallel tempering from the autocorrelation function of the overlap at the lowest temperatures. For an arbitrary observable $\mathcal{A}$, the autocorrelation function $\Gamma_{\mathcal{A}}(t)$ is defined via

$$\Gamma_{\mathcal{A}}(t) = \frac{\langle \mathcal{A}(0)\mathcal{A}(t)\rangle - \langle \mathcal{A}\rangle^2}{\langle \mathcal{A}^2\rangle - \langle \mathcal{A}\rangle^2}. \tag{3.4}$$

Here, $\mathcal{A}(t)$ is the observable measured at time t and $\langle \cdots \rangle$ represents a thermal average (To implement this definition of the autocorrelation function in the simulations, the observable $\mathcal{A}$ is measured at regular intervals and correlated with its value at a time lag t earlier. See, for example, Ref. [54]). Integrated and exponential autocorrelation times are computed from the autocorrelation function. The integrated autocorrelation time $\tau_{\text{int}}^{\mathcal{A}}$ is the integral of the autocorrelation function or, for discrete time, the sum of the autocorrelation function,

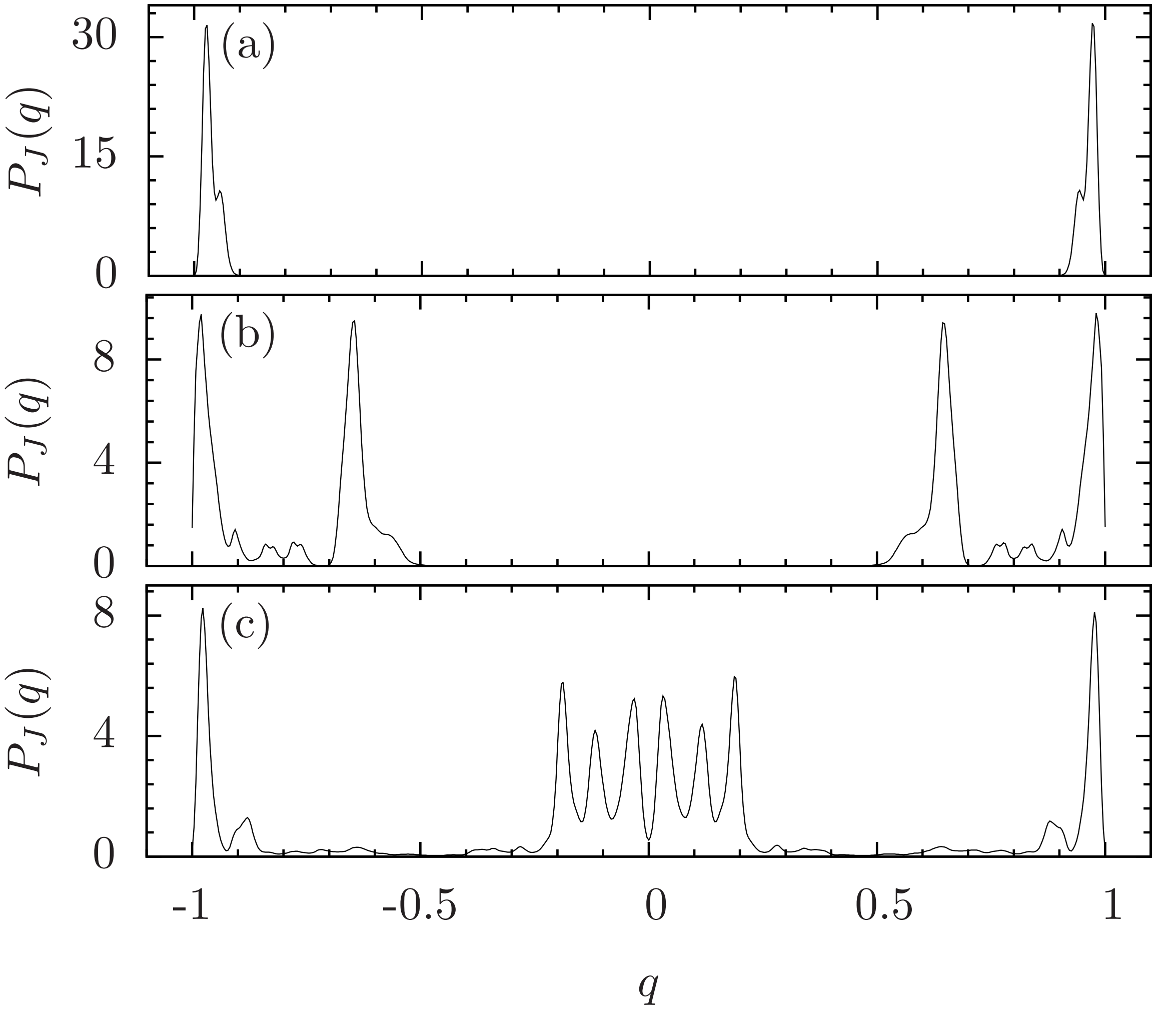

Figure 3.2. Examples of overlap distributions $P_J(q)$ for three different disorder realizations J for $L = 8$. While the distribution in panel (a) has only two peaks, the distribution in panel (c) implies a very complex free-energy landscape. All panels have the same horizontal scale.

$$\tau_{\text{int}}^{\mathcal{A}} = \frac{1}{2} + \sum_{t=1}^{\infty} \Gamma_{\mathcal{A}}(t). \tag{3.5}$$

The integrated autocorrelation time is the time needed for two subsequent measurements of $\mathcal{A}$ to decorrelate. It is a *lower bound* on the time needed to reach equilibrium from an arbitrary initial condition. To calculate $\tau_{\text{int}}^{\mathcal{A}}$ care must be taken when truncating the sum in Eq. (3.5) to avoid large statistical errors (explained below).

Figure 3.3 shows two typical examples of the autocorrelation function for the absolute value of the spin overlap $|q|$. For large times, the autocorrelation function is eventually dominated by noise. The noise floor is indicated in Figs. 3.3 by the horizontal (red) solid line. Because the noise floor is determined primarily by the number of data points used to compute the autocorrelation function, it can be chosen to be the same for all simulations of the same length. Our truncation procedure is to sum the autocorrelation function until it first hits the noise floor. We obtain the noise floor for each run length by visual inspection of $\Gamma_{|q|}(t)$ for many samples. We note that this approach introduces some slight systematic bias in our estimates of the integrated autocorrelation times. However, the main emphasis of the study presented in this chapter is to understand the very large sample-to-sample variations in time scales and, compared to these variations, the errors introduced by our truncation protocol are small.

We have computed autocorrelation functions for both $|q|$ and q and from these we have obtained $\tau_{\text{int}}^{|q|}$ and τ_{int}^{q}, respectively. To do this, we first compute a Fourier transform of the data and then invert the function to obtain the autocorrelation function. This method is considerably faster than a direct calculation. For $L = 6$, the simulations are run for 2^{24} sweeps, recording data for the computation of the autocorrelation functions every 10 sweeps. The noise floor for these runs is set to $\Gamma = 0.005$. Some samples have $\tau_{\text{int}}^{|q|}$ values of order 10 or less. For these, shorter data collection runs of 2^{18} sweeps have been done starting from an equilibrated spin configuration stored at the end of the longer run. For these short runs data are collected every sweep to accurately measure the autocorrelation

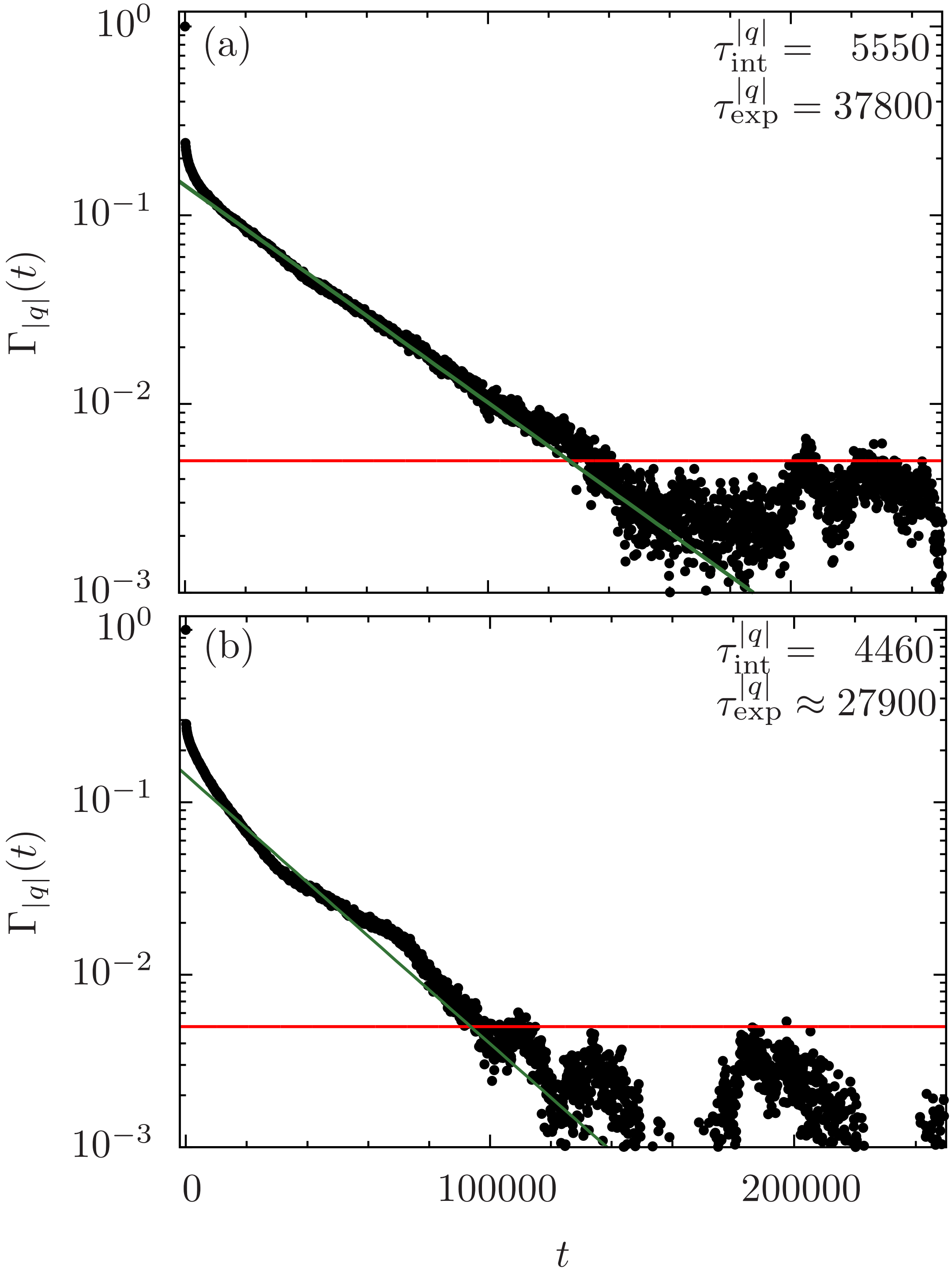

Figure 3.3. Autocorrelation function for the absolute value of the order parameter $\Gamma_{|q|}(t)$ as a function of simulation time for two different disorder realizations and system size $L = 8$. The horizontal (red) solid line at $\Gamma_{|q|} = 0.005$ represents the noise floor. The diagonal solid (green) lines represent exponential fits to the data. Both panels have the same horizontal scale.

function up to time 10 and the full autocorrelation function is patched together from the short and long runs. The noise floor for these shorter runs is set to $\Gamma = 0.01$. If the noise floor is reached before $t = 10$, only the autocorrelation function generated by the short run is used. For both $L = 8$ and $L = 10$ the simulations are run for at least 2^{27} sweeps, recording data every 100 sweeps with shorter runs also needed for samples with small $\tau_{\text{int}}^{|q|}$ values. The noise floor used is the same as for $L = 6$.

The autocorrelation function is always calculated up to a maximum time lag of 1% of the total run length. Therefore, for some samples and, in particular for $L = 10$, the noise floor was not reached by this time and longer runs were necessary to obtain good statistics. For these $L = 10$ samples we simulate up to 2^{30} sweeps. 32 samples (approximately 0.6%) stayed above the noise floor even for the longest runs. To prevent a biasing of the results, these were not included in the analysis. However, it would be interesting to study these samples in detail in the future to see why their time scales are this long.

There is another method for truncating the sum in Eq. (3.5) introduced by Madras and Sokal [39] where the upper limit of the sum is determined recursively. An initial upper limit is chosen and $\tau_{\text{int}}^{\mathcal{A}}$ is computed for that value, the new upper limit is determined as some factor (6 is suggested in Ref. [39]) times the current $\tau_{\text{int}}^{\mathcal{A}}$. The estimated value of $\tau_{\text{int}}^{\mathcal{A}}$ is obtained when this process has converged. We experimented with this approach but did not find that it converged in a consistent way across all samples with very different autocorrelation functions. However, we note that our approach should have systematic errors no larger than those resulting from the method of Madras and Sokal with the upper limit set to $6\tau_{\text{int}}^{|q|}$ because the noise floor is not reached until at least $10\tau_{\text{int}}^{|q|}$ except for the very few samples with extremely small $\tau_{\text{int}}^{|q|}$ values.

The second time scale that we consider is the exponential autocorrelation time $\tau_{\text{exp}}^{\mathcal{A}}$, defined by the asymptotic exponential decay of the autocorrelation function,

$$\Gamma_{\mathcal{A}}(t) \sim A e^{-t/\tau_{\text{exp}}^{\mathcal{A}}}. \tag{3.6}$$

Except for observables orthogonal to the slowest mode of the Markov chain, $\tau_{\rm exp}^{\mathcal{A}}$ is the characteristic time of the slowest mode and is thus the exponential time scale to reach equilibrium from an arbitrary initial state [65].

In principle, $\tau_{\rm exp}^{\mathcal{A}}$ is the most important time scale for studying disordered systems because the main difficulty is reaching thermal equilibrium (controlled by $\tau_{\rm exp}^{\mathcal{A}}$) rather than collecting enough uncorrelated data at equilibrium (controlled by $\tau_{\rm int}^{\mathcal{A}}$). Unfortunately, calculating $\tau_{\rm exp}^{\mathcal{A}}$ for a large number of samples is difficult because it requires an automatized fitting process. For some samples, such as the one shown in Fig. 3.3(a), the autocorrelation function decays nearly exponentially over most of the observable time range, allowing for a precise fitting and extraction of $\tau_{\rm int}^{|q|}$. However, for many samples the autocorrelation function is *not* exponential for the measured times, as can be seen in Fig. 3.3(b) and, for these samples, an automatic fitting procedure would not be reliable. For that reason, the time scale that we correlate with the static properties of the EA spin glass is the integrated autocorrelation time. We have studied $\tau_{\rm exp}^{|q|}$ by hand for a small subset of the data set and find that for those samples where a reasonable fit is possible, $\tau_{\rm exp}^{|q|}$ is in the range of 1 to 15 times $\tau_{\rm int}^{|q|}$. Most, but not all, samples share the behavior seen in the examples in Figs. 3.3 that there is an initial sharp decline before the (approximately) exponential decay sets in. This sharp decline explains why $\tau_{\rm exp}^{|q|}$ is typically larger than $\tau_{\rm int}^{|q|}$.

Finally, we also measured the round-trip time $\tau_{\rm RT}$ for each sample, defined as the average time after equilibration is achieved, for a replica to diffuse from the lowest to the highest temperature and then back to the lowest temperature again [32].

3.3 Results

Figure 3.4 shows the probability distribution for $\tau_{\rm int}^{|q|}$ for $L = 8$. In agreement with previous studies [1, 30, 2], the figure reveals that the equilibration times for spin glasses are very broadly distributed. The distributions for $L = 6$ and 10 are similar although with different means and standard deviations, as shown in Tables 3.2 and 3.3, respectively. Ta-

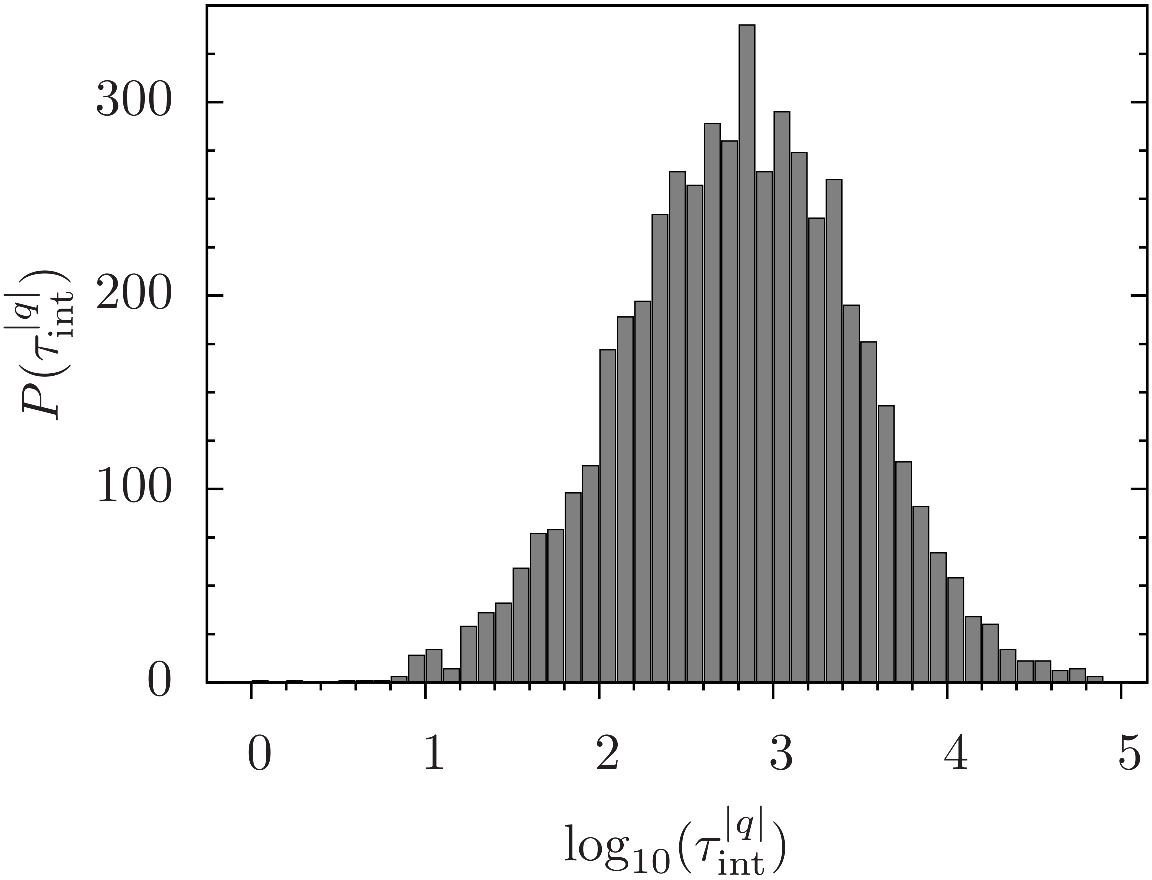

Figure 3.4. Logarithmic histogram of the integrated autocorrelation times $\tau_{\mathrm{int}}^{|q|}$ for for $L = 8$. The data are normally distributed and show that a small fraction of samples is extremely difficult to equilibrate.

Table 3.2. Mean values of the logarithms of τ_{int}^{q}, $\tau_{\text{int}}^{|q|}$ and τ_{RT} for system sizes $L = 6 – 10$.

L	6	8	10
$\log_{10}(\tau_{\text{int}}^{q})$	$1.7164(58)$	$2.8366(75)$	$3.8711(84)$
$\log_{10}(\tau_{\text{int}}^{|q|})$	$1.6035(90)$	$2.8065(94)$	$3.8881(97)$
$\log_{10}(\tau_{\text{RT}})$	$3.6217(28)$	$4.4571(34)$	$5.2421(42)$

Table 3.3. Standard deviations of the logarithms of τ_{int}^{q}, $\tau_{\text{int}}^{|q|}$ and τ_{RT} for system sizes $L = 6$ – 10.

L	6	8	10
$\log_{10}(\tau_{\text{int}}^{q})$	0.408	0.534	0.592
$\log_{10}(\tau_{\text{int}}^{|q|})$	0.631	0.672	0.685
$\log_{10}(\tau_{\text{RT}})$	0.194	0.242	0.293

bles 3.2 and 3.3 also show the mean and standard deviation of τ_{int}^{q}. The values for τ_{int}^{q} are similar to those for $\tau_{\text{int}}^{|q|}$, as can be seen more clearly in Fig. 3.5, however the standard deviations are somewhat smaller. The mean time scales increase with system size, as expected, and the round trip time τ_{RT} is, on average, longer than the integrated autocorelation times. Figure 3.6 shows a scatter plot with each point representing one sample for $L = 8$. The x and y coordinates are determined by τ_{int}^{q} and $\tau_{\text{int}}^{|q|}$, respectively. The data illustrate that τ_{int}^{q} and $\tau_{\text{int}}^{|q|}$ are strongly correlated, especially for those samples with the longest integrated autocorrelation times. The observation that $\tau_{\text{int}}^{|q|}$ can take smaller values than τ_{int}^{q} is presumably a consequence of the fact that the decorrelation of q requires global spin flips while the decorrelation of $|q|$ does not.

Figures 3.7(a) and 3.7(b) show scatter plots of $\tau_{\text{int}}^{|q|}$ vs $I_J(q_0)$ for $L = 8$ and $q_0 = 0.2$ and $q_0 = 0.8$, respectively. Each point represents one sample. These figures show that $\tau_{\text{int}}^{|q|}$ is correlated with $I_J(q_0)$ and the correlation is stronger for larger values of q_0. It is also striking that large values of $I_J(q_0)$ are never associated with very small integrated autocorrelation times (similar results are obtained for the other system sizes studied). This

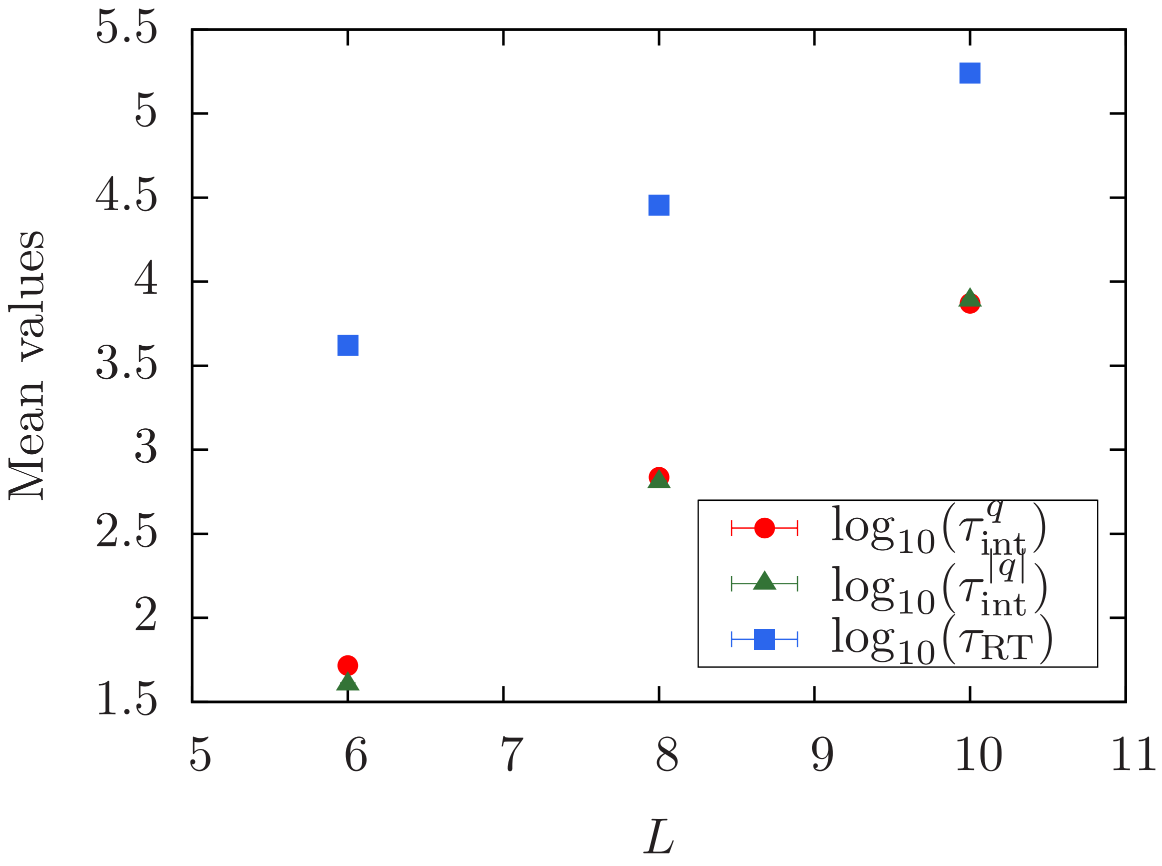

Figure 3.5. The change in mean values of the logarithms of τ_{int}^{q}, $\tau_{\text{int}}^{|q|}$ and τ_{RT} with respect to system size L, from Table 3.2.

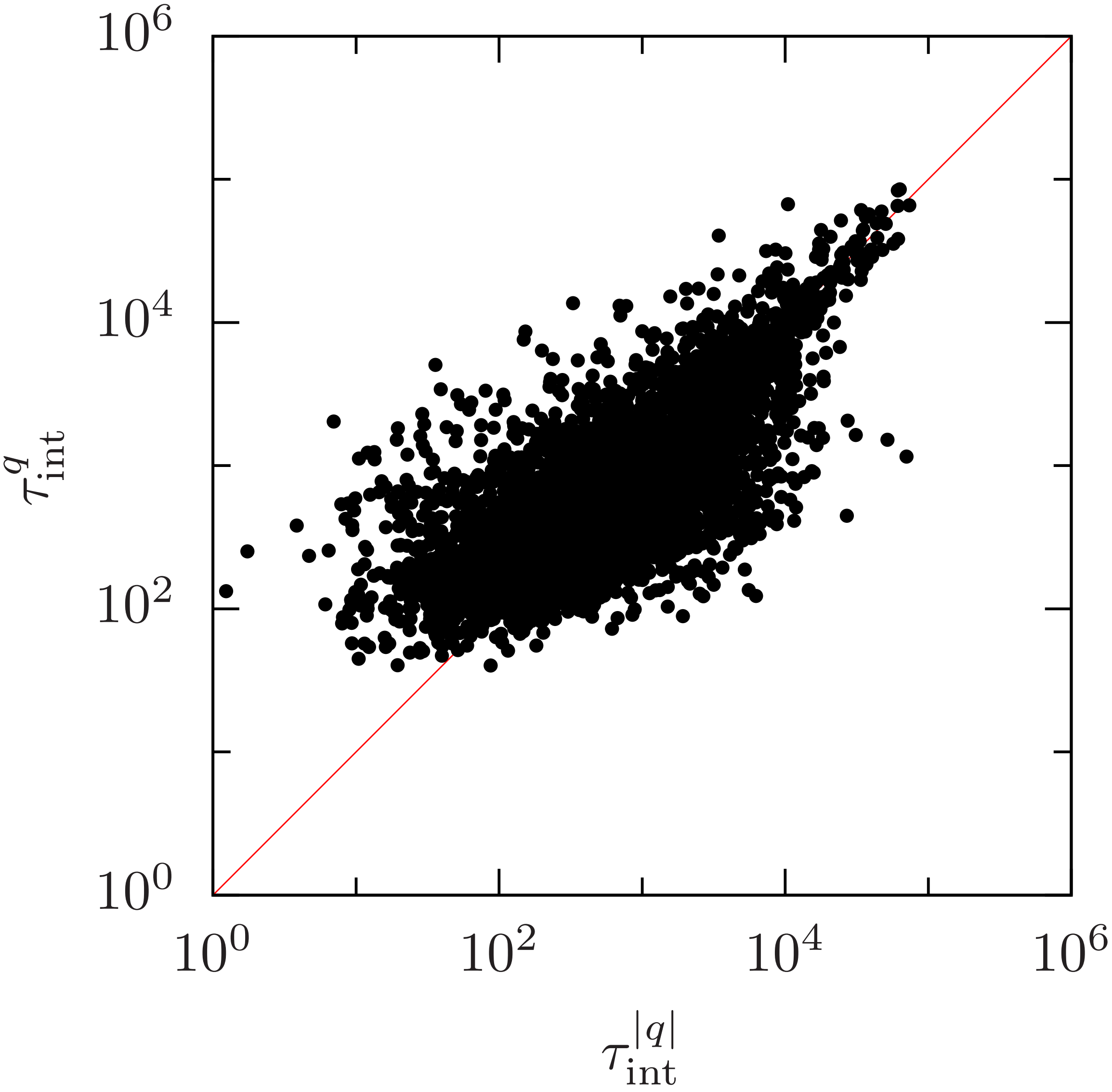

Figure 3.6. Scatter plot of τ_{int}^{q} vs. $\tau_{\mathrm{int}}^{|q|}$ for $L = 8$. The diagonal (red) solid line corresponds to $\tau_{\mathrm{int}}^{q} = \tau_{\mathrm{int}}^{|q|}$. Note that $\tau_{\mathrm{int}}^{|q|}$ is often smaller than τ_{int}^{q} because the decorrelation of q requires global spin flips while the decorrelation of $|q|$ does not.

suggest that when the free-energy landscape is complex, time-scales are large, as one would naively expect.

To quantify the correlations between $\tau_{\text{int}}^{|q|}$ and $I_J(q_0)$, we use the Pearson correlation coefficient r [61], whose estimator for a finite data set $\{(x_i, y_i)\}$ is

$$r = \frac{\sum_{i=0}^{M}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=0}^{M}(x_i - \bar{x})^2 \sum_{i=0}^{M}(y_i - \bar{y})^2}}, \tag{3.7}$$

where M is the size of the data set and the overbar indicates the average over the data set. If $r = 1$ there is an exact linear relationship between x and y while a small value of r indicates the absence of any linear dependence. Pearson r values for $\log_{10}(\tau_{\text{int}}^{|q|})$ vs $\log_{10}[I_J(q_0)]$ for our data are shown in Table 3.4 and Fig. 3.8 for the system sizes studied and different values of q_0. These results clearly demonstrate that longer parallel tempering time scales are correlated with more complicated overlap distributions. The strength of the correlation is slightly weaker for larger systems. The fact that r increases with q_0 suggests that the presence of additional peaks in $P_J(q)$ increases $\tau_{\text{int}}^{|q|}$ independent of whether those peaks are near the origin or not. Here it is worth mentioning that to observe correlations, q_0 needs to be chosen small enough to exclude the two outer peaks at $\pm q_{EA}$, since those peaks are always present and not indicative of the complexity of the overlap distribution. Consequently, including them will only increase $I_J(q_0)$ but not the equilibration time, as can be seen by the decreasing r value for $q_0 > 0.8$ in Fig. 3.8.

On the other hand, we find that the round trip time is not correlated with $\tau_{\text{int}}^{|q|}$ or, by extension, $I_J(q_0)$. Figure 3.9 is a scatter plot of τ_{RT} vs $\tau_{\text{int}}^{|q|}$ for $L = 8$ (compare with Fig. 3.6). Each point represents one sample and the diagonal line is $\tau_{\text{RT}} = \tau_{\text{int}}^{|q|}$. It is clear from this figure that τ_{RT} has a much smaller variance than $\tau_{\text{int}}^{|q|}$ and that the correlation between the two is minimal. The last row of Table 3.4 shows the Pearson r value between $\log_{10}(\tau_{\text{RT}})$ and $\log_{10}(\tau_{\text{int}}^{|q|})$ and quantifies the lack of correlation between τ_{RT} and $\tau_{\text{int}}^{|q|}$. Note

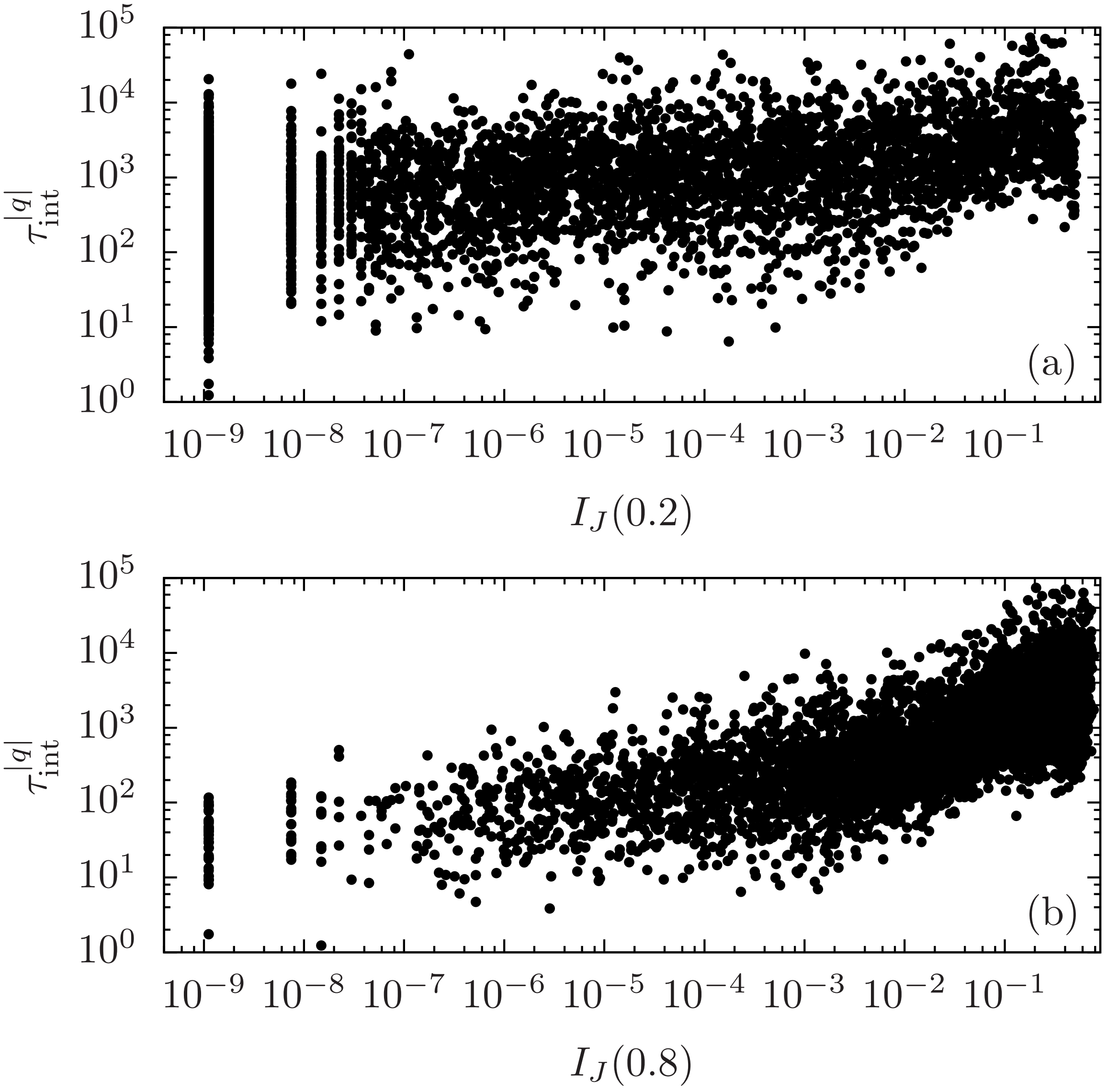

Figure 3.7. Scatter plot for all $L = 8$ samples of $\tau_{\text{int}}^{|q|}$ vs $I_J(q_0)$ for (a) $q_0 = 0.2$ and (b) $q_0 = 0.8$. The empty area underneath the scatter for large $I_J(q_0)$ clearly shows that realizations with large values of $I_J(q_0)$ can never have small values of $\tau_{\text{int}}^{|q|}$.

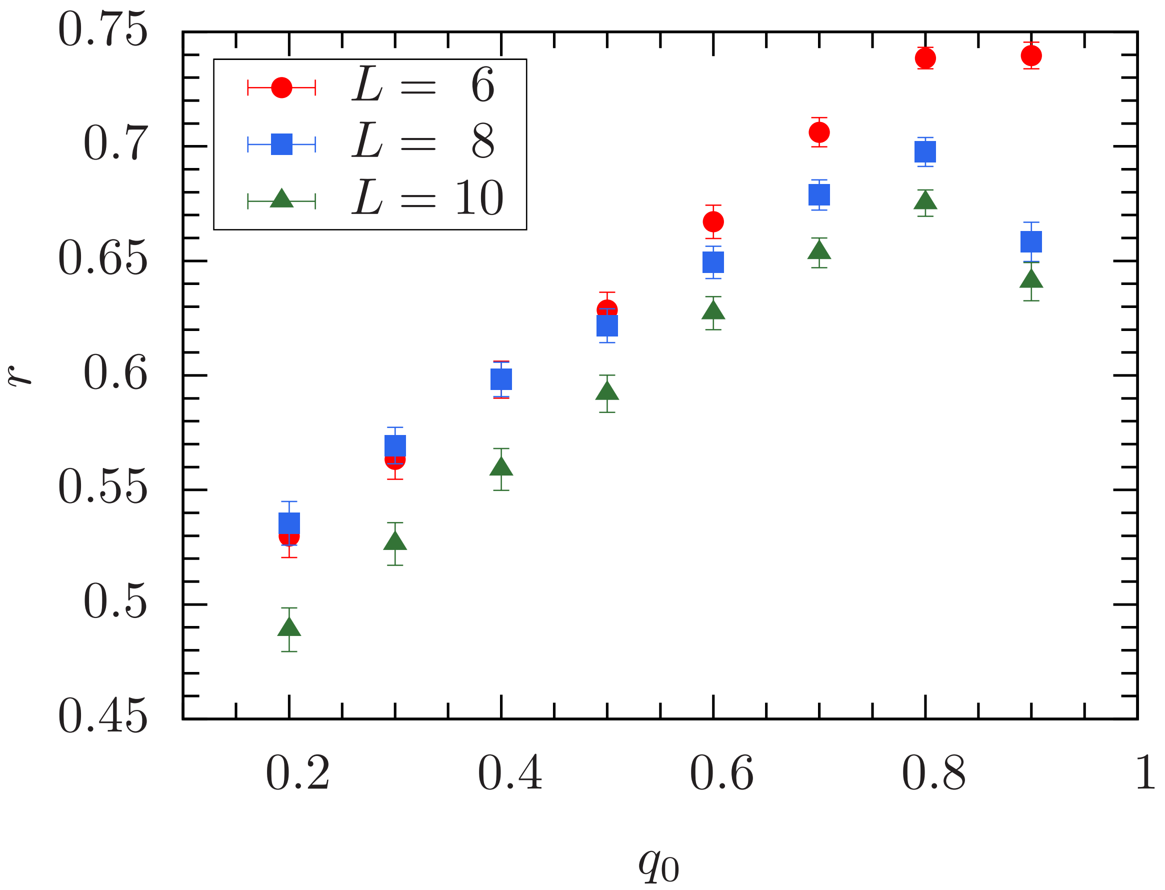

Figure 3.8. Pearson correlation r between $\log_{10}[I_J(q_0)]$ and $\log_{10}(\tau_{\text{int}}^{|q|})$ vs q_0. The data suggest that longer parallel tempering time scales are correlated with more complicated overlap distributions.

that the authors of Ref. [3] consider a time scale related to the motion of replicas in parallel tempering and also find no correlation with $I_J(q_0)$.

If the highest temperature is chosen such that its mixing time is one heat bath sweep, then $\tau_{\text{RT}} \geq \tau_{\text{int}}^{|q|}$ because there will be no memory of the spin state of the replica between successive visits to the lowest temperature, as can be seen in Tables 3.2 and 3.3, as well as Fig. 3.9. Figure 3.9 shows that τ_{RT} is much larger than $\tau_{\text{int}}^{|q|}$ for most samples. One explanation for this is that (for most samples) the spin state of the coldest replica can be decorrelated by diffusion from the lowest temperature to an intermediate temperature and back; full round trips are not needed for decorrelation.

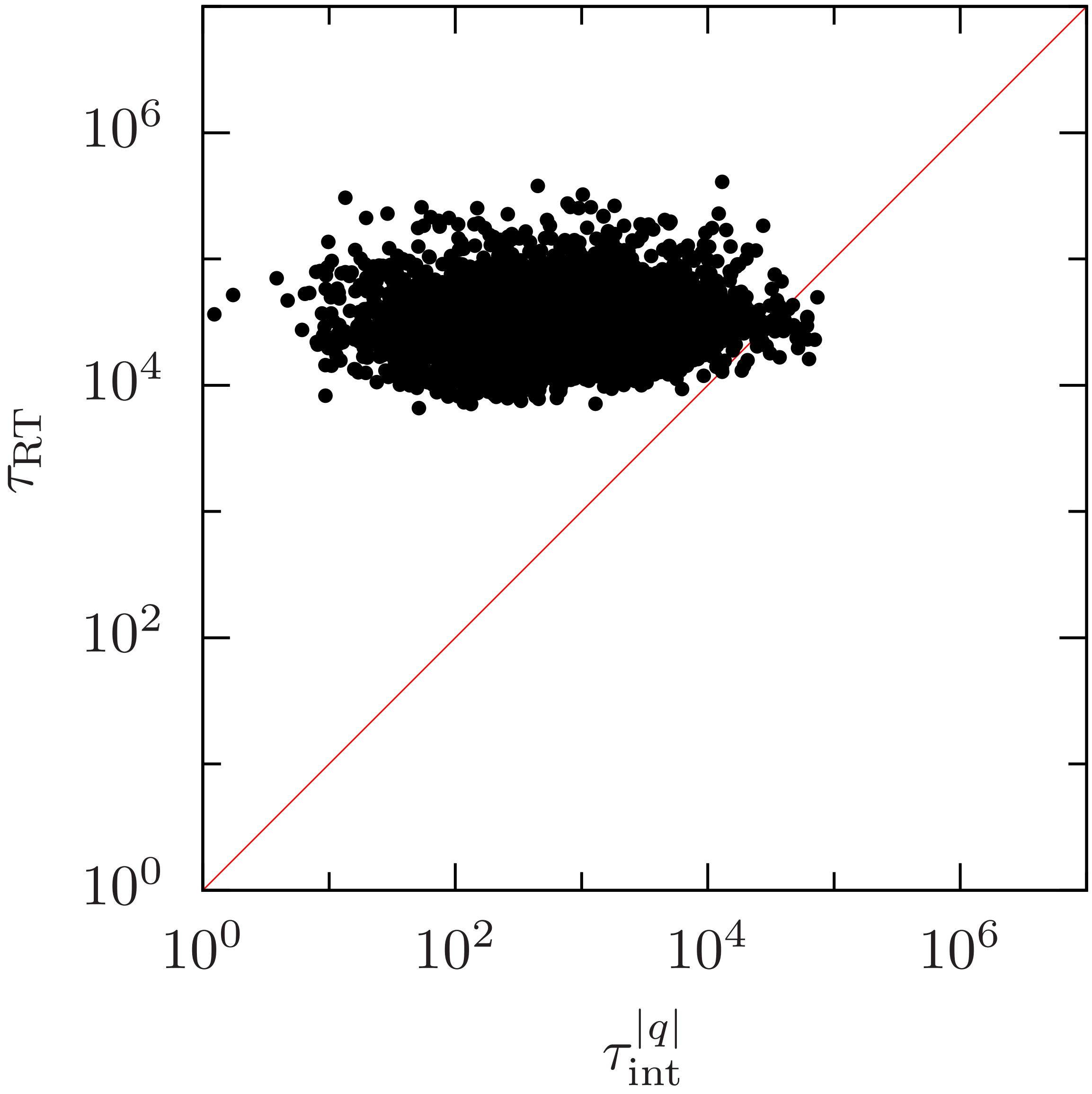

Figure 3.9. Scatter plot of the average round trip time τ_{RT} vs $\tau_{\mathrm{int}}^{|q|}$ for $L = 8$. The diagonal (red) line is $\tau_{\mathrm{RT}} = \tau_{\mathrm{int}}^{|q|}$. The data show almost no correlation between these quantities.

Table 3.4. Pearson correlation r between the overlap weight $\log_{10}[I_J(q_0)]$ for various q_0 and the logarithm of the integrated autocorrelation time for the overlap $\log_{10}(\tau_{\rm int}^{|q|})$ for the three system sizes, L. The bottom row shows the correlation coefficient between the logarithm of the round trip time $\log_{10}(\tau_{\rm RT})$ and $\log_{10}(\tau_{\rm int}^{|q|})$.

q_0	$L = 6$	$L = 8$	$L = 10$
0.2	$0.5299(94)$	$0.5355(95)$	$0.4889(95)$
0.3	$0.5634(86)$	$0.5693(79)$	$0.5264(93)$
0.4	$0.5982(80)$	$0.5983(75)$	$0.5590(91)$
0.5	$0.6286(77)$	$0.6217(73)$	$0.5920(80)$
0.6	$0.6671(72)$	$0.6494(70)$	$0.6272(72)$
0.7	$0.7062(63)$	$0.6788(66)$	$0.6535(65)$
0.8	$0.7385(46)$	$0.6975(63)$	$0.6752(57)$
0.9	$0.7397(58)$	$0.6583(85)$	$0.6410(83)$
$\log_{10}(\tau_{\rm RT})$	$0.014(19)$	$0.114(17)$	$0.126(20)$

Note that if we repeat the study for $\tau_{\rm int}^{q}$ instead of $\tau_{\rm int}^{|q|}$, the correlation values would show the same qualitative trend as for $\tau_{\rm int}^{|q|}$, but they would on average be slightly smaller.

Finally, we also test equilibration with the method introduced by Katzgraber *et al.* in Ref. [31] for short-range spin glasses with Gaussian disorder between the spins. In this method, the average energy per spin

$$u = -\frac{1}{N}\sum_{\langle i,j\rangle}[\, J_{ij}\langle s_i s_j\rangle\,]_{\rm av}, \tag{3.8}$$

where $[\cdots]_{\rm av}$ represents a disorder average and, as before, $\langle\cdots\rangle$ denotes the Monte Carlo average for a given set of bonds, can be related to the link overlap

$$q_l = \frac{1}{N_b}\sum_{\langle i,j\rangle} s_i^\alpha s_j^\alpha s_i^\beta s_j^\beta \tag{3.9}$$

via an integration by parts over the interactions J_{ij} between the spins. In Eq. (3.9) $N_b = dN$ with $d = 3$ the space dimension represents the number of bonds in the system. One obtains:

$$[\langle q_l\rangle]_{\rm av} = 1 - \frac{T|u|}{d}, \tag{3.10}$$

where T is the temperature.

Figure 3.10 shows representative results for $L = 8$ and $T = 0.2$, the lowest temperature simulated. At approximately 10^5 Monte Carlo sweeps both data computed directly from the energy per spin (blue squares) and data computed from the link overlap (red circles) agree within error bars. Note that 99% of the $\tau_{\rm int}^{|q|}$ values for this system size are smaller than the point marked with the solid vertical green line. Thus, for the simulated time of over 10^8 Monte Carlos sweeps, the data are well equilibrated. This suggests that a conservative use of the equilibration criterion developed in Ref. [31] will guarantee that almost all of the samples are equilibrated.

3.4 Conclusions

We have found one factor that explains the broad distribution of time scales in parallel tempering for spin glasses: The roughness of the free-energy landscape of each individual sample directly affects its equilibration time scales. In accordance with previous observations on the parallel tempering method [38], if there are many minima in the free-energy landscape, a sample requires more time to equilibrate. The equilibrium distribution of the spin overlap serves as proxy for the complexity of the free-energy landscape. Our results show that individual samples need to be tested individually to ensure proper equilibration, as previously suggested in Ref. [2]. However, we show that the equilibration test developed in Ref. [31] is a viable alternative for system sizes currently accessible in simulations if applied conservatively. Furthermore, the parallel tempering round-trip times seem to be unaffected by the complexity of the free-energy landscape and should therefore not be used as a measure of how well a system is equilibrated.

The clear correlations between algorithmic (nonequilibrium) time scales and the equilibrium complexity of the free-energy landscape opens the door for alternate studies of the nature of the spin-glass state, a problem that we consider in the following chapters.

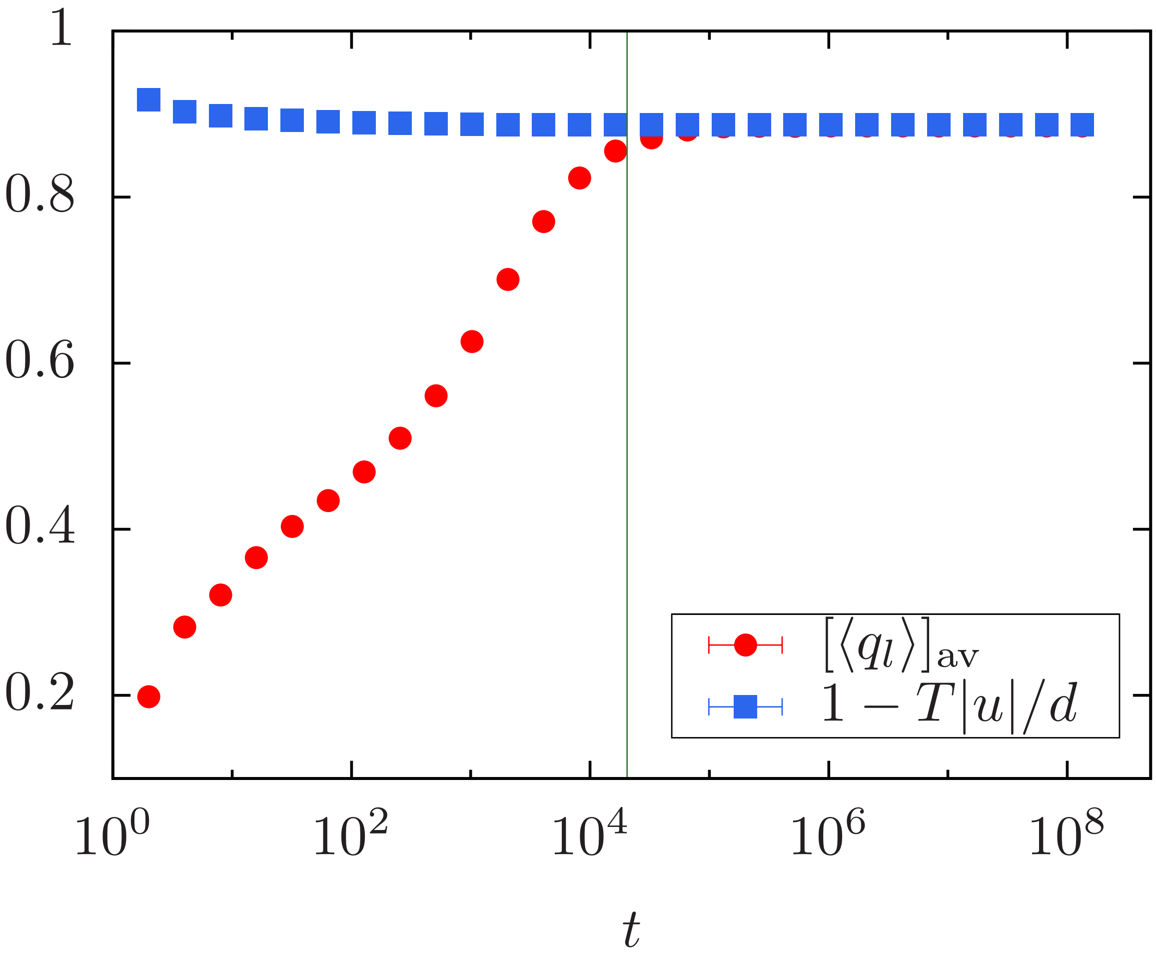

Figure 3.10. Average link overlap $[\langle q_l \rangle]_{\rm av}$ (red circles) and $1 - T|u|/d$ (see Eq. (3.10), blue squares) as a function of time t for $L = 8$ and $T = 0.20$. For $t \gtrsim 10^5$ sweeps both data sets agree suggesting that the system is in thermal equilibrium. The vertical (green) line in the figure represents the time for which 99% of the samples have values of $\tau_{\rm int}^{|q|}$ less than this time. Error bars are smaller than the symbols.

CHAPTER 4

EQUILIBRIUM PROPERTIES

4.1 Introduction

Spin glasses [10] have been the subject of intense study and considerable controversy for decades. These models are perhaps the simplest, physically-motivated examples of frustrated systems in classical statistical mechanics. Given their wide applicability across disciplines, it is important that their behavior is understood. Despite four decades of research, the low-temperature phase of short-range spin glasses is hotly debated and poorly understood. Here we study both the three-dimensional (3D) Edwards-Anderson (EA) Ising spin glass [17] and the Ising spin glass on a complete graph—known as the Sherrington-Kirkpatrick (SK) model [63]—in an effort to gain a deeper understanding of the low-temperature spin-glass state, as well as the applicability of different theoretical pictures. Our results suggest that these models are qualitatively different at low temperatures.

Parisi's solution of the SK model [56, 59] involves an unusual form of symmetry breaking among replicas. These were originally introduced as a tool to carry out the disorder average of the logarithm of the partition function. The low-temperature phase of the model within the replica symmetry breaking (RSB) solution [56, 59] has several unusual features such as the breakdown of self-averaging and the co-existence of a countable infinity of pure states in the thermodynamic limit.

There is no analytic theory for the EA model but it is now well-accepted on the basis of numerical simulations [35] that the EA model undergoes a continuous phase transition. However, the low-temperature broken-symmetry phase is not understood, even qualitatively. Different mutually-exclusive scenarios have been proposed: The replica symmetry

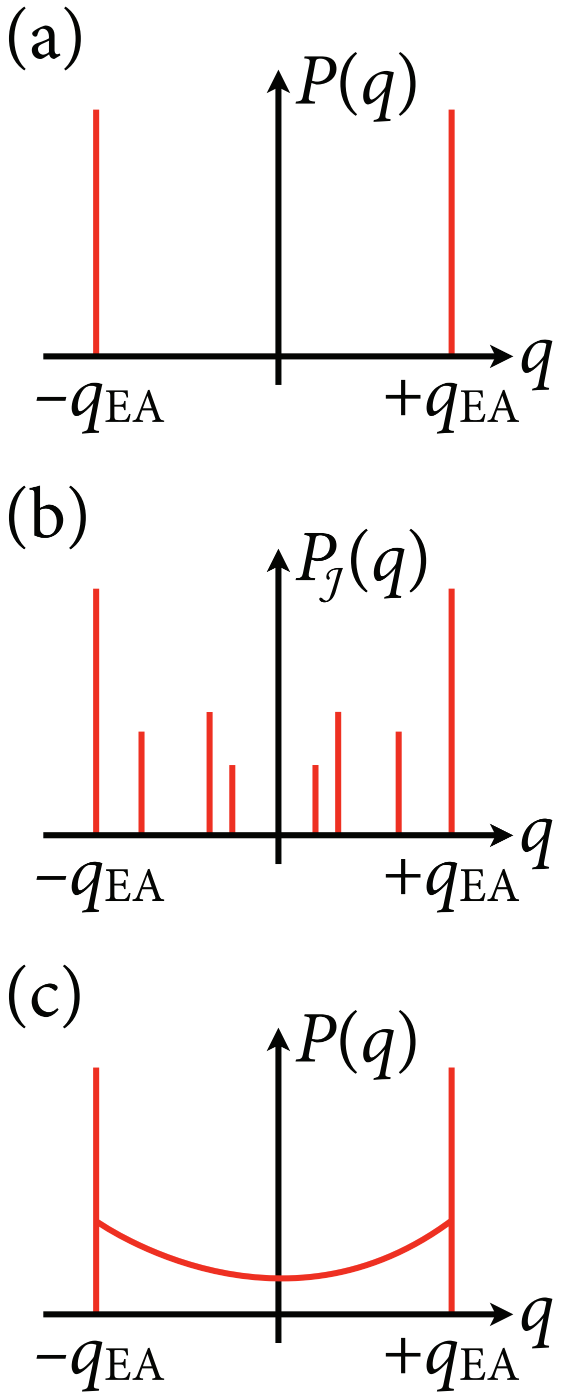

Figure 4.1. (a) In the droplet (and chaotic pairs picture for fixed volume) $P(q)$ is trivial (two δ functions at $\pm q_{\mathrm{EA}}$, i.e., one pair of pure states). (b) In the RSB picture individual samples have many pairs of pure states (many δ functions in $P_{\mathcal{J}}(q)$). (c) In the RSB picture $P(q)$ is nontrivial (continuous support for $|q| < q_{\mathrm{EA}}$).

breaking (RSB) picture is based on an analogy with the solution of the SK model. It assumes that self-averaging breaks down and that there are a countable infinity of pure states in the thermodynamic limit. A qualitatively different and simpler picture was proposed to describe the EA model by McMillan, Fisher and Huse, as well as Bray and Moore [42, 19, 20, 21, 12]. In the "droplet scaling" picture, the low-temperature phase is described by one pair of pure states related by a spin flip with low-lying excitations that are isolated, compact droplets of the opposite phase. In the droplet picture the low-temperature phase of the EA model is qualitatively similar to the Ising ferromagnet where there are also two phases related by a global spin flip and where low-lying excitations are compact regions of the minority phase. A central difference between RSB and the droplet picture for the EA model is whether there is a single pair of pure states or a large number of pairs of pure states for large systems, see Fig. 4.1.

Newman and Stein [49, 50, 51] explained that the usual way of constructing the thermodynamic limit cannot be applied to finite-dimensional spin glasses because of the possibility of a chaotic system-size dependence in which different thermodynamic states may appear for different system sizes. They showed that for translationally-invariant lattices and disorder that is statistically translationally invariant, the key ideas of RSB—non-self-averaging and a countable infinity of pure states—cannot hold within the naïve way that they were first proposed. However, their results do not completely rule out a nonstandard interpretation of RSB with an uncountable infinity of pure states but in every fixed volume, only a countable infinity of these states appearing. They also proposed a more plausible many-states picture, known as the "chaotic pairs" picture, in which for a fixed choice of couplings, there are uncountably many pure states but that in a single finite volume only one pair is manifest.

In this chapter we report the results of large-scale Monte Carlo simulations of both the SK and EA models. Our objective is to shed light on the qualitative nature of the low-temperature phase of the EA model by comparing and contrasting to the SK model using comparable system sizes. Previous numerical studies, e.g., [34] using the average spin overlap distribution suggested that both the SK and EA models are well described by the RSB picture, see Fig. 4.1. However, for the numerically-accessible system sizes the two main peaks are still converging to $\pm q_{\mathrm{EA}}$ (see Figs. 4.1 and 4.4) and therefore results might be plagued by finite-size effects. On the other hand, studies of the average link overlap [34] distribution suggest agreement with the droplet picture. The "trivial nontrivial" scenario [55, 37, 34] reconciles these numerical results by postulating that excitations are compact, as in the droplet picture, but their energy cost is independent of system size, as in the RSB picture. In an effort to resolve these discrepancies, in this work we introduce a statistic obtained from the spin overlap distribution that detects sharp peaks in *individual* samples, inspired by a recent study on the SK model [5]. This statistic clearly differentiates the RSB and droplet pictures: it converges to zero in the large-volume limit if there is a single pair of pure states and to unity if there are countably many. Our results for this quantity shows clear differences between the EA and SK models.

4.2 Models and Numerical Details

The SK and EA models are defined by the Hamiltonian

$$\mathcal{H} = -\sum_{i,j=1}^{N} J_{ij} s_i s_j, \tag{4.1}$$

with $s_i \in \{\pm 1\}$ Ising spins. For the EA model (Fig. 4.2(a)), the sum is over nearest neighbors on a cubic lattice of size $N = L^3$ with periodic boundaries. The couplings J_{ij} are chosen from a Gaussian distribution with zero mean and standard deviation unity. A set of couplings $\mathcal{J} = \{J_{ij}\}$ defines a disorder realization or, simply, a "sample." For the SK model (Fig. 4.2(b)), the sum is over all pairs of spins and the J_{ij} are chosen from a Gaussian distribution with zero mean and standard deviation $1/\sqrt{N-1}$.

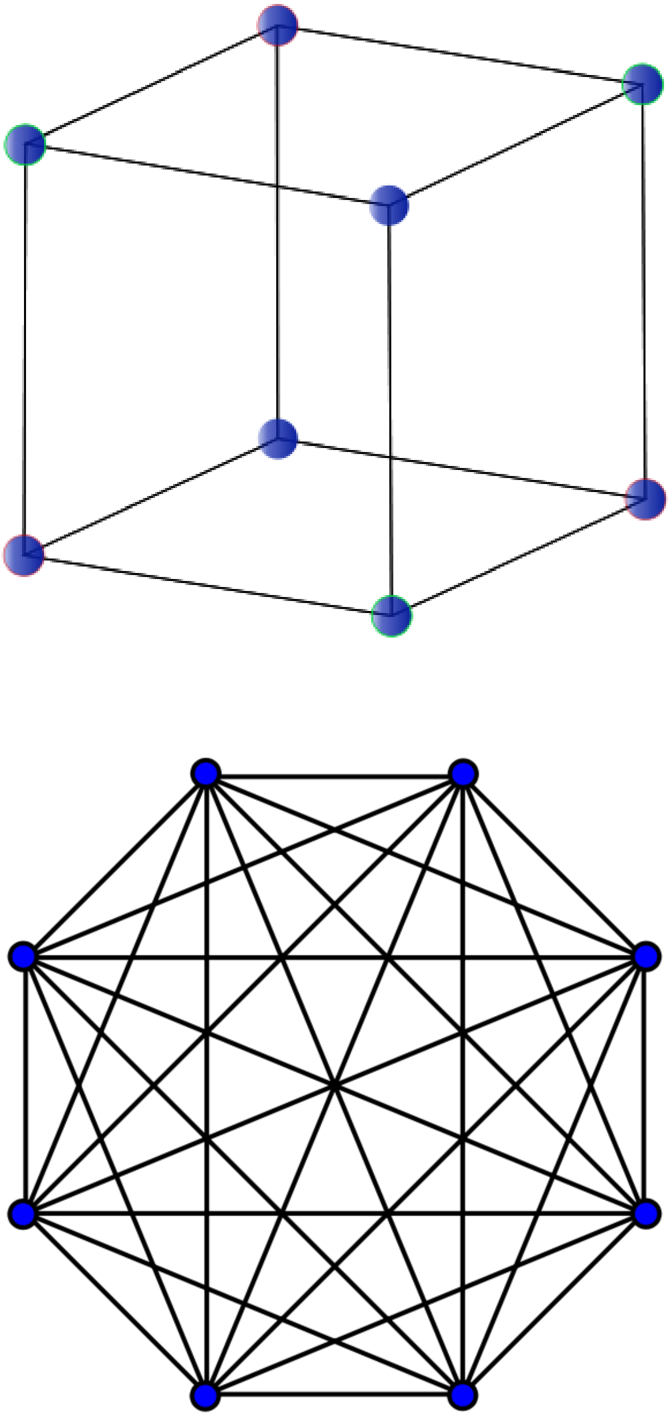

Figure 4.2. (a) Cubic lattice for the EA model for a system of size $L = 2$ with $N = 8$ spins. (b) Complete graph for the SK model with $N = 8$ spins. In both cases, the Ising spins are located at the corners marked by the blue dots and the edges/lines correspond to couplings.

Ordering in spin glasses is detected by measuring the spin overlap

$$q = \frac{1}{N}\sum_i s_i^\alpha s_i^\beta, \tag{4.2}$$

where "α" and "β" indicate two independent spin configurations for the same sample $\mathcal{J}$. The primary observable we consider for fixed disorder $\mathcal{J}$ and system size N is the overlap probability density, $P_\mathcal{J}(q)$. In the high-temperature phase there is a well-defined thermodynamic limit and $P_\mathcal{J}(q) \to \delta(q)$ for $N \to \infty$ for almost every $\mathcal{J}$. The behavior of $P_\mathcal{J}(q)$ for large N and $T < T_c$, T_c the critical temperature, distinguishes the RSB picture from other theories. If there is only a single pair of states for each system size, $P_\mathcal{J}(q)$ consists for large N of a symmetric pair of δ functions at the Edwards-Anderson order parameter $q = \pm q_{\mathrm{EA}}$, see Fig. 4.1(a). In the RSB picture there are many sharp peaks symmetrically distributed in the range $-q_{\mathrm{EA}} < q < q_{\mathrm{EA}}$ as shown in Fig. 4.1(b), corresponding to multiple pairs of pure states. In the RSB picture, the distribution of peaks depends on $\mathcal{J}$ but the disorder average of the overlap distribution $P(q)$ exists, and for large N is expected take the form shown in Fig. 4.1(c).

We have carried out parallel tempering Monte Carlo [25] simulations of both models. Parameters are shown in Tables 4.1 and 4.2. For each sample we equilibrate two independent sets of replicas to compute the overlap distribution. Equilibration in the EA model is tested with the method introduced in Ref. [34], in accordance with our findings explained in the previous chapter. For the SK model we use the approach from Ref. [26]. The number of equilibration and data collection sweeps are chosen to be long enough to ensure that samples are well equilibrated and that $P_\mathcal{J}(q)$ is accurately measured for each sample. We report results for $T = 0.42$ [$T = 0.4231$] for the EA [SK] model. For the EA model, $T_c \approx 0.96$ [35], while for the SK model $T_c = 1$, such that for both models our simulations are at $\sim 0.4T_c$, i.e., deep within the spin-glass phase where critical fluctuations are unimportant. It is worth noting here that tests at low T give qualitatively similar results.

Table 4.1. Simulation parameters for the EA spin glass. For each number of spins $N = L^3$ we equilibrate and measure for 2^b Monte Carlo sweeps. $T_{\rm min}$ [$T_{\rm max}$] is the lowest [highest] temperature used and N_T is the number of temperatures. $N_{\rm sa}$ is the number of disorder samples. For $T \geq 0.42$ all system sizes are in thermal equilibrium.

N	L	b	$T_{\rm min}$	$T_{\rm max}$	N_T	$N_{\rm sa}$
64	4	18	0.2000	2.0000	16	4891
216	6	24	0.2000	2.0000	16	4961
512	8	27	0.2000	2.0000	16	5130
1000	10	27	0.2000	2.0000	16	5027
1728	12	25	0.4200	1.8000	26	3257

Table 4.2. Simulation parameters for the SK spin glass. For details see the caption of Table 4.1. For $T \geq 0.4231$ all data are equilibrated.

N	b	$T_{\rm min}$	$T_{\rm max}$	N_T	$N_{\rm sa}$
64	22	0.2000	1.5000	48	5068
128	22	0.2000	1.5000	48	5302
256	22	0.2000	1.5000	48	5085
512	18	0.2000	1.5000	48	4989
1024	18	0.2000	1.5000	48	3054
2048	16	0.4231	1.5000	34	3020

4.3 Results

Figure 4.3 shows $P_{\mathcal{J}}(q)$ for three different EA samples ($N = 512 = 8^3$, $T = 0.42$). It is noteworthy that $P_{\mathcal{J}}(q)$ varies considerably between samples. Qualitatively similar overlap distributions are seen for the SK model. Figure 4.4, left panel [right panel], shows the disorder averaged overlap distribution $P(q)$ for the EA [SK] model for different system sizes at $T = 0.42$ [$T = 0.4231$]. Because $P(q)$ is an even function, we show one side of $P(q)$ for each model. At this low temperature, $P(q)$ consists of large peaks at the finite-size value of the EA order parameter, $\pm q_{\rm EA}(N)$. $P(q)$ is reasonably flat, non-zero, and nearly independent of N in the approximate range $-0.4 \lesssim q \lesssim 0.4$ for the sizes studied here. We can quantify this observation by considering the integrated overlap,

$$I(q_0) = \int_{|q|<q_0} P(q) dq. \tag{4.3}$$

Figure 4.5 shows $I(0.2)$ as a function of N for both the EA and SK models at $T \approx 0.4 T_c$ (Tests at low T give qualitatively similar results). It is noteworthy that $I(0.2)$ is nearly independent of N. We found qualitatively similar results for other values of q_0 up to $q_0 \approx$

0.5 and temperatures down to $0.2T_c$ for smaller systems. The constancy of $I(0.2)$ has been observed in a number of studies (see, for example, Refs. [34] and [3]) and is among the strongest evidence in favor of the validity of the RSB picture for short-range systems.

Although $I(q_0)$ in Fig. 4.5 is nearly constant over the range of sizes simulated in this and other studies of the EA model, it is also clear that, for these same sizes there are strong finite-size effects. These corrections can be seen, for example, by looking at the size dependence of $q_{\mathrm{EA}}(N)$, the location of the main peak in $P(q)$ for different system sizes N. The peak moves to smaller values of q_{EA} as N increases, similar to recent results [3] for larger N. The presence of these strong finite-size corrections makes the absence of any significant N dependence of $P(q)$ for small q somewhat surprising. In the droplet picture, $I(q_0)$ is expected to decay with a small power of L, $I(q_0) \sim TL^{-\theta}$ ($\theta \approx 0.2$ in 3D [24]) and this slow asymptotic behavior may not set in until large sizes where the droplet energy is larger than the temperature are reached. Thus the behavior of $I(q_0)$ shown in Fig. 4.5 may not be a sensitive indicator of the nature of the low-temperature phase for system sizes currently accessible to simulation.

To better understand the size dependence of the overlap distributions, we must go beyond disorder averages and consider other statistics obtained from $P_{\mathcal{J}}(q)$. In particular, we want to identify the emergence, or not, of δ functions in the range $-q_{\mathrm{EA}} < q < q_{\mathrm{EA}}$ as N increases, which would signal more than one pair of pure states. A finite-size broadened δ function at q is characterized by a large value of $P_{\mathcal{J}}(q)$. To detect a δ-function-like behavior for finite N we consider the statistic

$$\Delta(q_0, \kappa) = \mathrm{Prob}\left[\max_{|q|<q_0}\left\{\frac{1}{2}\big(P_{\mathcal{J}}(q) + P_{\mathcal{J}}(-q)\big)\right\} > \kappa\right]. \tag{4.4}$$

The probability is defined with respect to the distribution of couplings, so that $\Delta(q_0, \kappa)$ is the fraction of samples with at least one peak greater than κ in $P_{\mathcal{J}}(q)$ in the range $|q| < q_0$. Since $P_{\mathcal{J}}(q)$ is an even function, we average over q and $-q$ to reduce statistical errors. κ is chosen to be large enough to exclude some but not all samples. We refer to samples

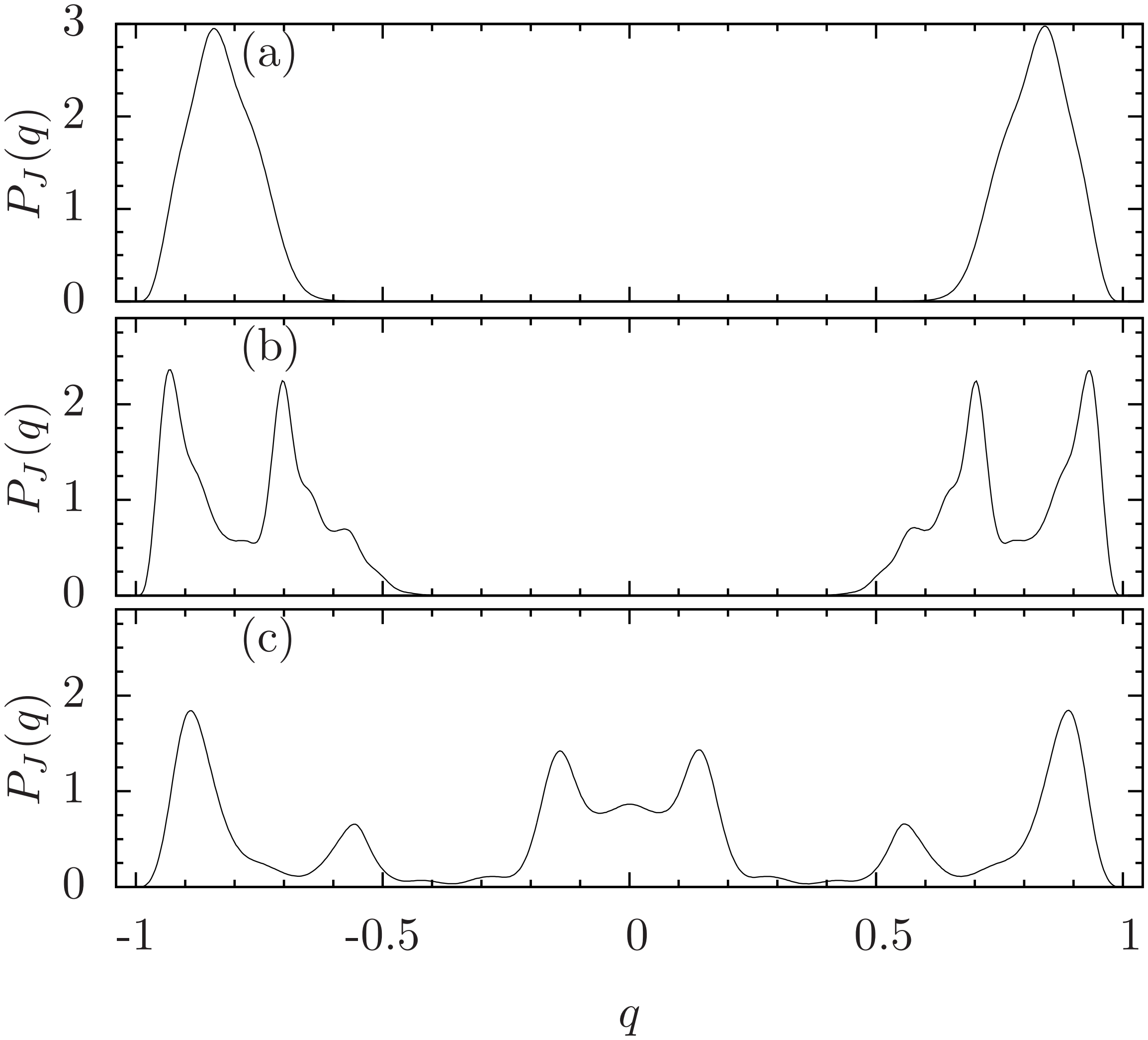

Figure 4.3. Typical overlap distributions $P_{\mathcal{J}}(q)$ for three disorder realizations for the EA model with $N = 8^3$ and $T = 0.42$. $P_{\mathcal{J}}(q)$ varies considerably between samples, and for a choice of $\kappa = 1$, only the sample with the central peaks (c) is peaked for $q_0 \gtrsim 0.1$, whereas the two other samples (a) and (b) are not for $q_0 \lesssim 0.5$.

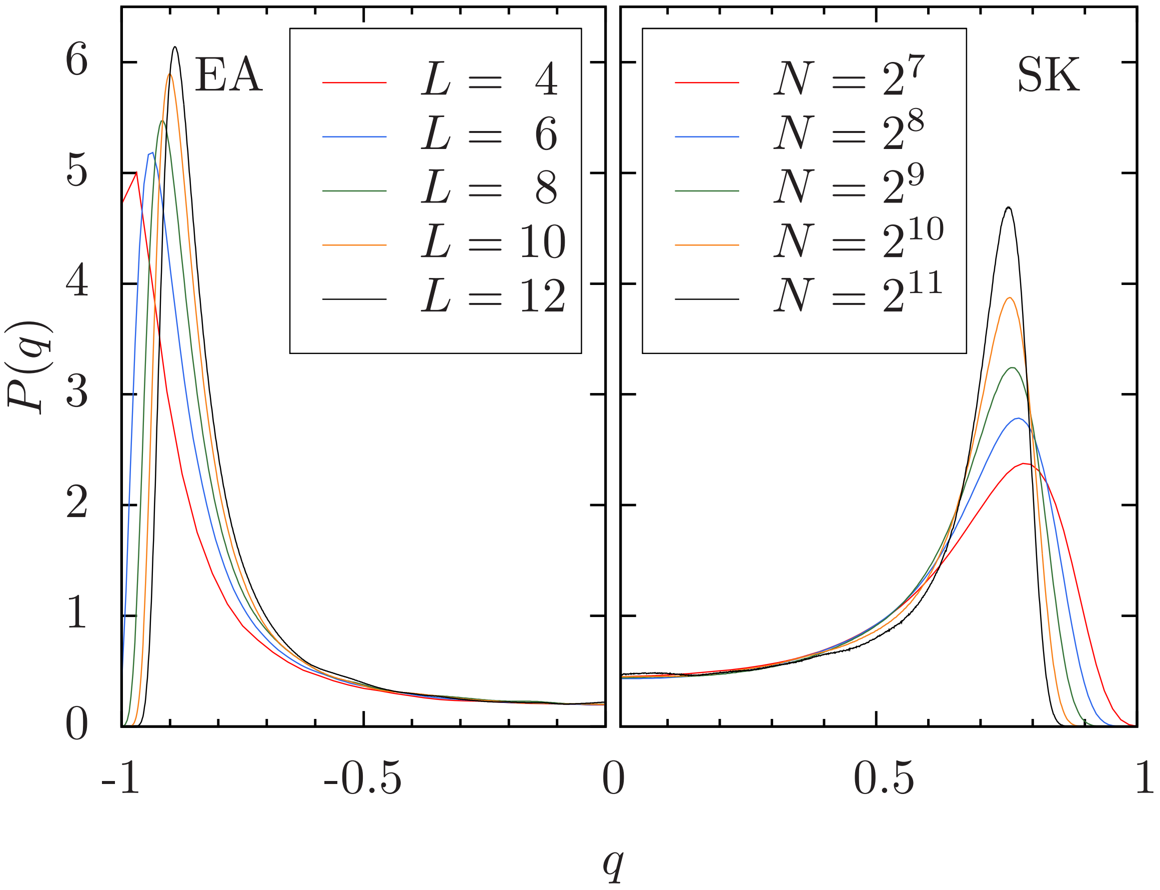

Figure 4.4. Disorder-averaged overlap probability distribution $P(q)$ for different system sizes at $T = 0.42$ and $T = 0.4231$ for the EA model (left) and SK model (right), respectively. Strong finite-size effects are present, which can be noted by looking at the height and the location of the main peaks for different system sizes. As the system size increases, the peaks move to smaller values of q_{EA} and their height increases.

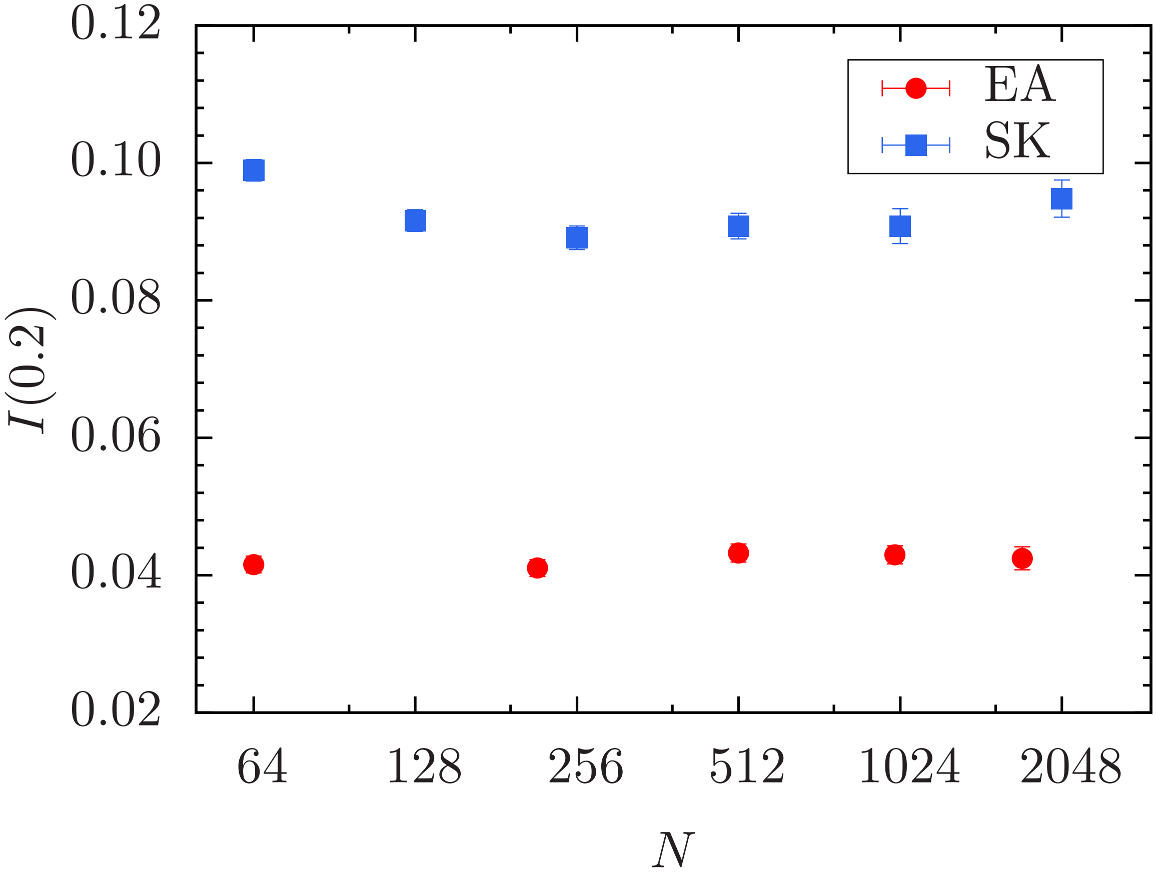

Figure 4.5. Disorder average of the weight of the overlap distribution $I(0.2)$ as a function of N for $T \approx 0.4T_c$ for both the EA and SK models. The data are nearly independent of N.

counted in $\Delta(q_0, \kappa)$ as "*peaked*." For example, with $\kappa = 1$ the sample with the central peaks (c) in Fig. 4.3 is peaked for $q_0 \gtrsim 0.1$, whereas the two other samples (a) and (b) are not for $q_0 \lesssim 0.5$.

The droplet and chaotic pairs pictures make dramatically different predictions from the RSB picture for $\Delta(q_0, \kappa)$. For the droplet or chaotic pairs picture there is only a single pair of states for any large volume so that $\Delta(q_0, \kappa) \to 0$ for any $\kappa > 0$ and any $q_0 < q_{\mathrm{EA}}$ when $N \to \infty$. However, for the RSB picture one expects δ functions in $P_{\mathcal{J}}(q)$ for any range of q, i.e., $\Delta(q_0, \kappa) \to 1$ as $N \to \infty$ for any q_0 and $\kappa > 0$.

Figure 4.6 shows $\Delta(q_0, \kappa)$ as a function of system size for $q_0 = 0.2$ and 0.4, as well as $\kappa = 1$. The error bars there are one standard deviation statistical errors due to the finite number of samples. There are also errors due to the finite length of the data collection. To estimate these errors we measured $\Delta^+(q_0, \kappa)$ and $\Delta^-(q_0, \kappa)$, defined as in Eq. (4.4) but from the $q > 0$ and $q < 0$ components of $P_{\mathcal{J}}(q)$, respectively. These are expected to be reasonably independent and their differences provide an estimate of the error due to finite run lengths. For all sizes, the average absolute difference between these quantities, $(|\Delta^+(q_0, \kappa) - \Delta(q_0, \kappa)| + |\Delta^-(q_0, \kappa) - \Delta(q_0, \kappa)|)/2$ is less than the statistical error.

We found qualitatively similar results for other values of q_0 and κ, as well as for lower temperatures. Our most important observation is that the fraction of peaked samples $\Delta(q_0, \kappa)$ is nearly constant and small for the EA model while $\Delta(q_0, \kappa)$ is increasing over the same range of N for the SK model. Similar results for the SK model were obtained in Ref. [5] by counting individual peaks. The result for the SK model is expected from Parisi's RSB solution. The contrasting result for the EA model suggests that the number of pure states does not grow with the system size for low T; a result consistent with the droplet and chaotic pairs pictures.

The difference in the behavior of Δ for the SK model in comparison to the EA model might be explained by the fact—visible in Fig. 4.4 and discussed more quantitatively in Ref. [18]—that peaks sharpen more quickly with N for the SK than for the EA model. To

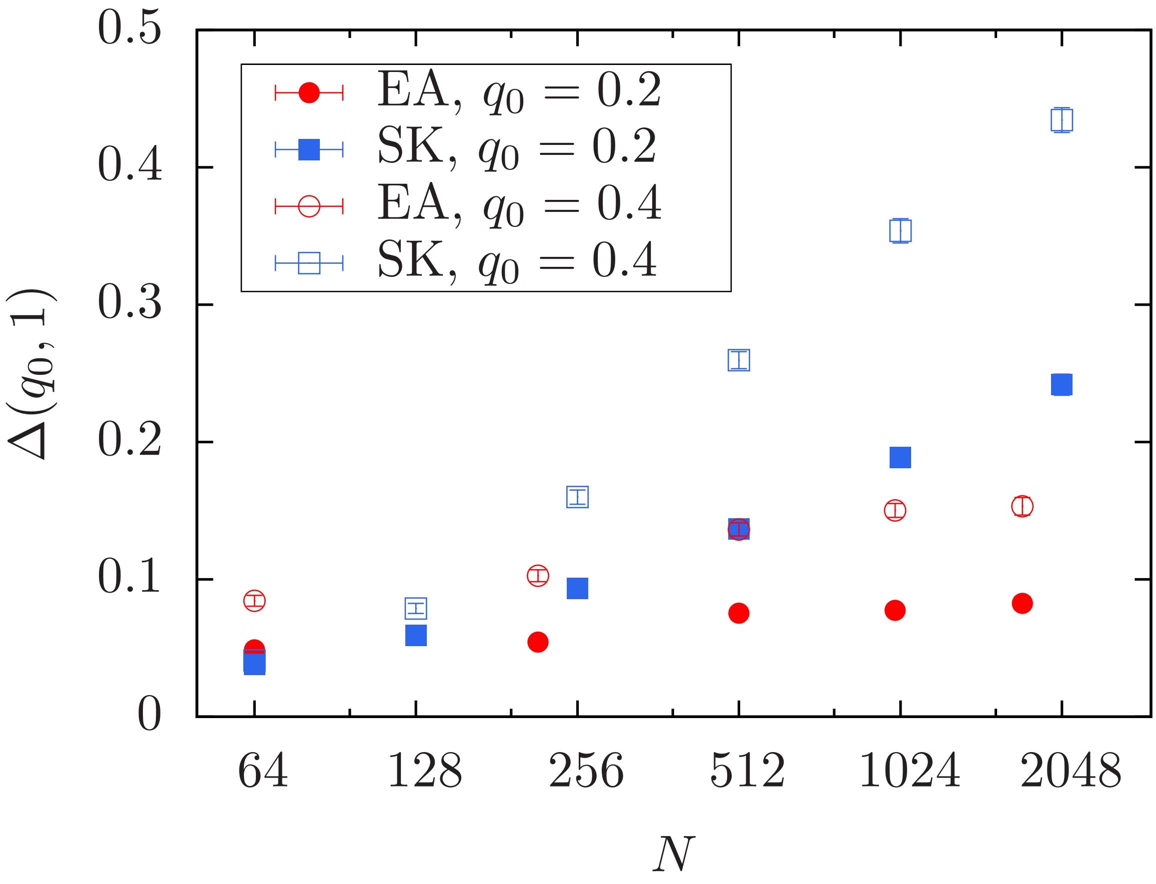

Figure 4.6. Fraction of peaked samples $\Delta(q_0, \kappa)$ at $T \approx 0.4T_c$ as a function of N for $\kappa = 1$, $q_0 = 0.2$ and 0.4. While data for the EA model are nearly constant as a function of N (consistent with the droplet/chaotic pairs pictures) the data for the SK model increase with N, in agreement with the RSB picture.

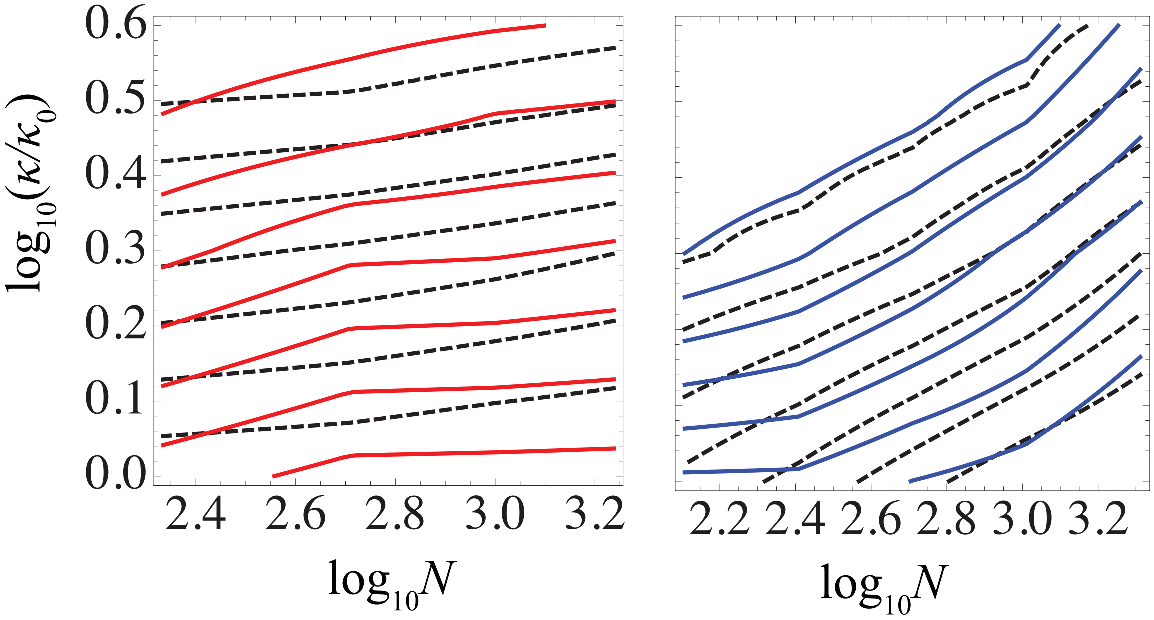

Figure 4.7. Contours of constant Δ for the EA model (left) and the SK model (right) as a function of $\log_{10}(N)$ and $\log_{10}(\kappa/\kappa_0)$ with $\kappa_0 = 0.5$ and 1.5 for $q_0 = 0.2$ and 1.0, respectively. The solid [dashed] lines are contours of constant Δ for $q_0 = 0.2$ [$q_0 = 1.0$]. The contours are equally spaced in Δ.

study the role of this effect, we compare Δ for the two values, $q_0 = 0.2$ and $q_0 = 1$, for each model separately. For $q_0 = 1$, Δ is controlled by the peaks at $\pm q_{\rm EA}$ and must converge to unity for both models because for $N \to \infty$ the $q_{\rm EA}$ peaks become δ functions. Figure 4.7, left [right] panel, shows contour plots of constant Δ for the EA [SK] model. The horizontal axis is the logarithm of the number of spins and the vertical axis is the logarithm of κ/κ_0 with $\kappa_0 = 0.5$ for $q_0 = 0.2$ and $\kappa_0 = 1.5$ for $q_0 = 1$. The curves are lines of constant Δ obtained from a linear interpolation of the data. Each set of curves are equally spaced in Δ with Δ decreasing as κ increases. For the EA model the ranges are $(0.2, 0.8)$ for $q_0 = 1$ and $(0.02, 0.14)$ for $q_0 = 0.2$. For the SK model the ranges are $(0.05, 0.77)$ for $q_0 = 1$ and $(0.06, 0.288)$ for $q_0 = 0.2$. The dashed contours are for $q_0 = 1$ and thus include the $q_{\rm EA}$ peaks. As expected, the dashed contours are clearly increasing functions for *both* models although they rise more rapidly for the SK model than for the EA model. The solid curves are contours of constant Δ for $q_0 = 0.2$. Close inspection of the data reveals a *qualitative* difference between both models. For large N and large Δ, the SK $q_0 = 0.2$ contours rise more steeply than the corresponding $q_0 = 1$ contours, suggesting that not only are peaks sharpening, but the number of peaks is also increasing. In fact, Ref. [5] shows that the number of peaks in $P_J(q)$ should scale as $N^{1/6}$ for the SK model. On the other hand, for large N and large Δ, the EA contours for $q_0 = 0.2$ are nearly flat, rising less steeply than for $q_0 = 1$, suggesting that the number of peaks is either decreasing or staying constant.

To further illustrate how various choices of κ would affect information gained from the quantity $\Delta(q_0, \kappa)$, let us consider two additional cases with both smaller ($\kappa = 0.5$) and larger ($\kappa = 2.5$) values. We mentioned that κ has to be chosen large enough to exclude some but not all samples. This is actually a very large window, since one would have to look at extremely small κ values to get nearly all of the samples to have at least one peak in the range of $-q_0 < q < q_0$ that is higher than κ. This shows the robustness of our conclusions with regards to $\Delta(q_0, \kappa)$, as can be seen in Fig. 4.8. There we have $\kappa = 0.5$, but we can still observe the qualitative difference between the EA and the SK models, especially for large system sizes. The only difference is that, for both models, for small system sizes only little change in behavior can be observed with increasing system size, thus making it harder to reach clear conclusions. Hence, using a larger κ illustrates our point more clearly. The next figure, Fig. 4.9, points to the opposite case, where κ is too large. Here, only the outer-most peaks located at q_{EA} would be high enough to exceed κ for most of our samples. But the q_0 values we are considering here are much smaller than q_{EA}, therefore, especially for smaller systems sizes, we get insignificantly small values. This is, in itself not necessarily misleading, but makes any observation about the behavior of the models difficult to interpret. In conclusion, observations made by looking at $\Delta(q_0, \kappa)$ are quite robust without any significant dependence on the value of κ, and the main reason we selected $\kappa = 1$ is to be able to make the most from our whole data set, including the smallest system sizes.

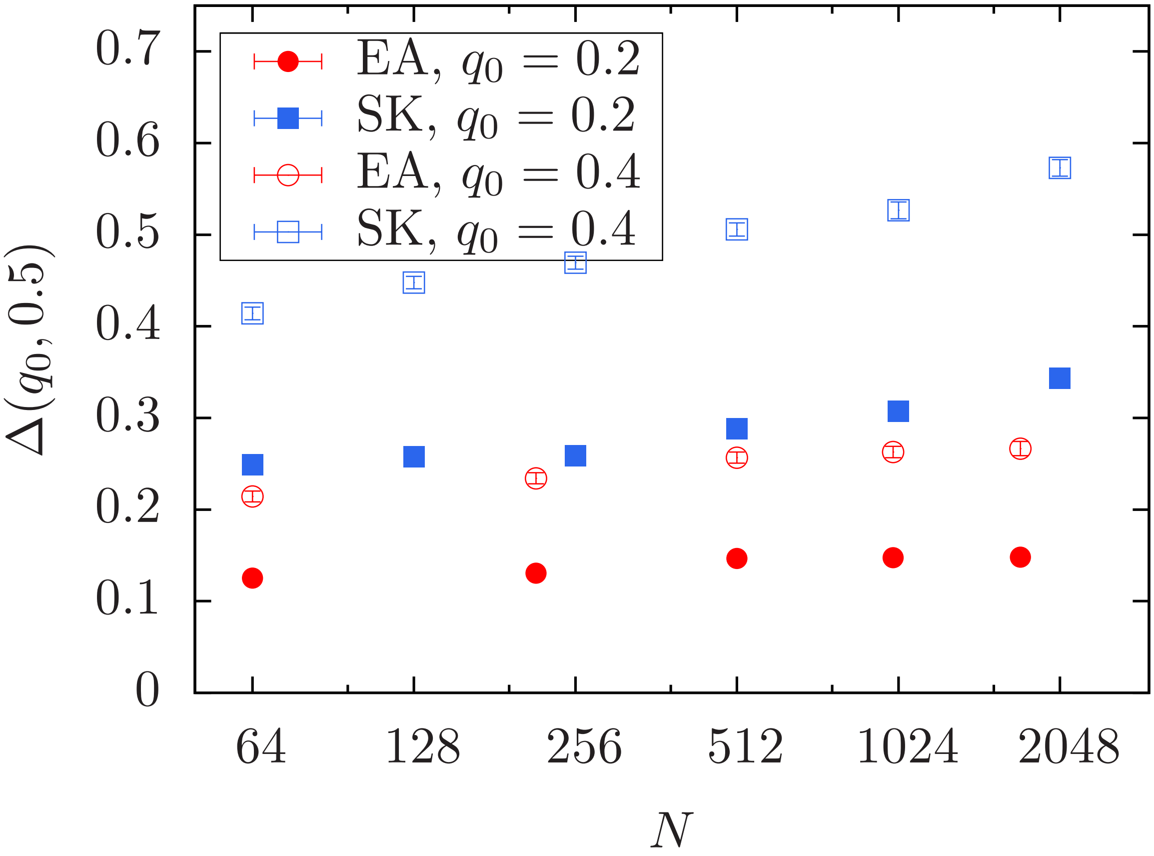

Figure 4.8. Fraction of peaked samples $\Delta(q_0, \kappa)$ at $T \approx 0.4T_c$ as a function of N for $\kappa = 0.5$, $q_0 = 0.2$ and 0.4. For such a low value of κ, $\Delta(q_0, \kappa)$ is at first not significantly affected by the increasing system size, only for larger system sizes we start to see the qualitative difference between the two models.

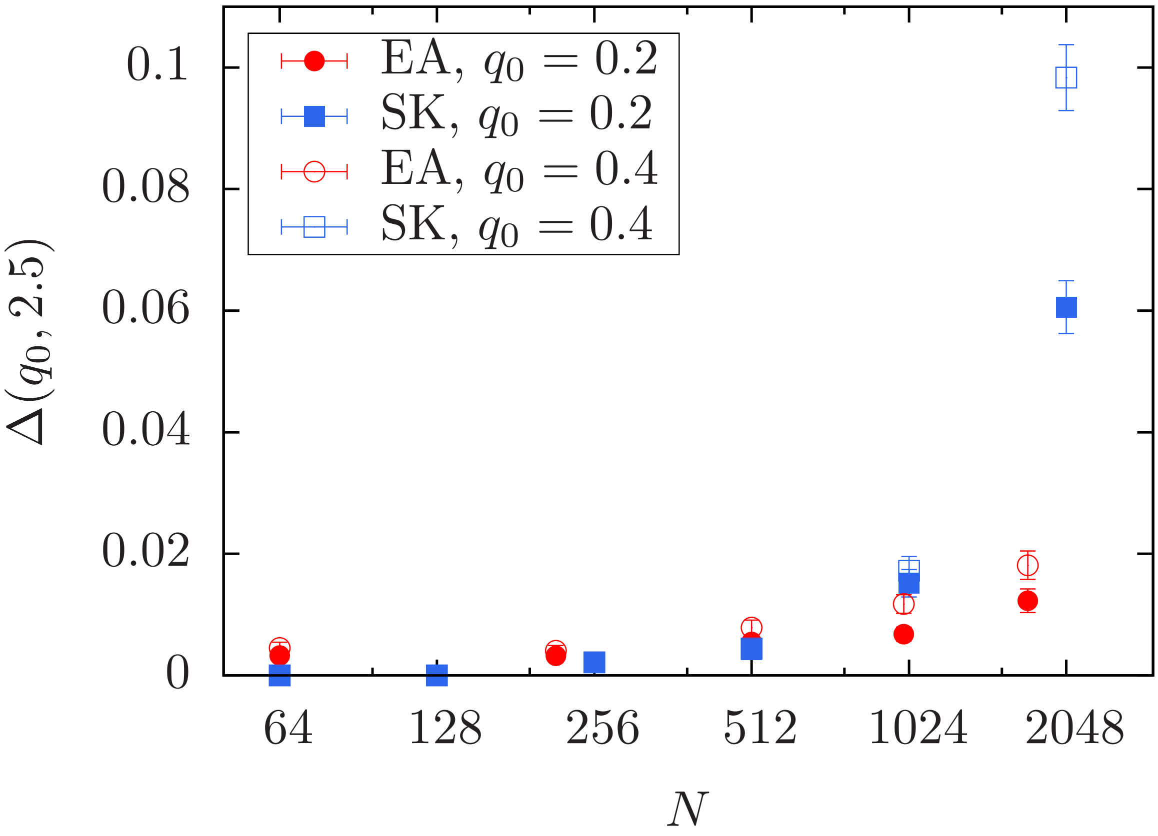

Figure 4.9. Fraction of peaked samples $\Delta(q_0, \kappa)$ at $T \approx 0.4T_c$ as a function of N for $\kappa = 2.5$, $q_0 = 0.2$ and 0.4. For such a high value of κ, very few samples have at least one peak higher than κ in the $\pm q_0$ range, thus, for most of the systems, $\Delta(q_0, \kappa)$ is too small to make any meaningful observations.

4.4 Conclusions

In this chapter we introduced a statistic Δ that detects the fraction of samples with δ function behavior in $P_{\mathcal{J}}(q)$ near the origin and sharply distinguishes the RSB picture from scenarios with only a single pair of states such as the droplet picture. While our results for the SK model are consistent with RSB, as expected, the EA model does not display a trend towards many pairs of pure states. These results lend weight to the droplet picture, the chaotic pairs picture and other scenarios with only a single pair of pure states in any large volume. It is also possible that for the EA model, Δ increases very slowly in N and ultimately converges to unity in agreement with the RSB picture. However, our data show no indication of such trend. Therefore, it would be interesting to perform a similar analysis with extremely large data sets computed with special-purpose computers, such as the Janus machine [6].

4.5 A scaling theory

Shortly after the paper [73] on which this chapter is based was published, Billoire *et al.* published a comment [9] with an interesting theory, based on RSB, to explain our 3D EA data. We responded [72], and the following is a summary of the exchange.

Their theory attributes the plateau in Δ for our EA data to a combination of a small value of $I(q_0)$ and a small value of the exponent describing the growth of $P(q_{EA})$ as a function of L. It predicts that Δ for the EA model will grow from its small values at $L = 12$ to unity extremely slowly in L. But our Fig. 4.7 shows that, even after factoring out the slower growth in $P(q_{EA})$ for the EA model relative to the SK model, we still find a qualitative difference between the EA and SK models. The scaling theory in their comment asserts that

$$\Delta(q_0, \kappa) \sim [P(q_{EA})/\kappa]^{I(q_0)}. \tag{4.5}$$

This can be simplified in the regime that it is applicable, i.e., when $I(q_0)$ is small. Noting that $I(q_0) \approx q_0 P(0)$ one obtains

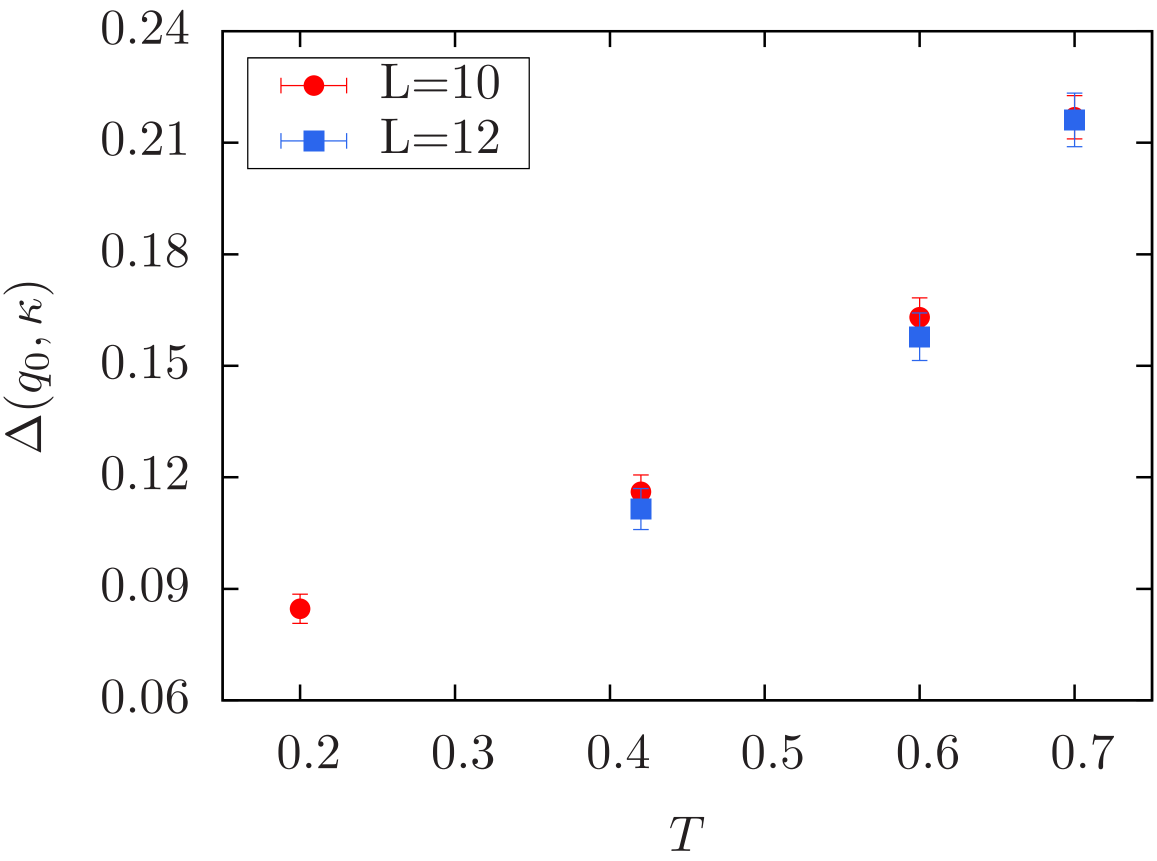

Figure 4.10. $\Delta(q_0, \kappa)$ vs T for $L = 10$ and 12 (EA model). q_0 and κ are chosen such that $I(q_0) \approx 0.067$ and $P(q_{EA})/\kappa = 3$.

$$\Delta(q_0, \kappa) \approx q_0 P(0) \log[P(q_{EA})/\kappa]. \tag{4.6}$$

The predicted linear dependence of Δ on q_0 is consistent with our data and is to be expected if the peaks in $P_J(q)$ are distributed roughly uniformly in q for $q < q_{\mathrm{EA}}$. To test the validity of the scaling theory of Ref. [9], we compared our data for the EA model at several temperatures holding the scaling quantities $(P(q_{EA})/\kappa)$ and $I(q_0)$ fixed across all T. The theory predicts that if these scaling variables are held fixed, then Δ should remain constant. Figure 4.10 shows $\Delta(q_0, \kappa)$ for a range of temperatures and our two largest sizes $L = 10$ and $L = 12$ (except for $T = 0.2$ where we do not have data for $L = 12$). For each T and L, both q_0 and κ are adjusted so that $I(q_0) \approx 0.067$ and $P(q_{EA})/\kappa = 3$, which requires that q_0 ranges from 0.16 to 0.56 and κ from 0.5 to 2.6 as T decreases from 0.7 to 0.2. It is clear from Fig. 4.10 that Δ is not constant as predicted by the theory in Ref. [9]. Therefore, the theory presented in their comment does not explain our EA data.

CHAPTER 5

THE 4D EA MODEL AND ADDITIONAL OBSERVATIONS

5.1 Introduction

In this chapter we expand on some of our findings in the previous chapter. First we consider the 4D Edwards-Anderson (EA) model [8, 62, 14, 60], in addition to the 3D EA and the Sherrington-Kirkpatrick (SK) models we looked at before, to have a more complete understanding of the Ising spin-glass in general. We then present supplementary measurements that were inspired by recent papers on the topic [45, 46], further investigating the low temperature behavior of all of our models. We finish up with two sections of preliminary results which seem promising to motivate further investigation. The first of those is to consider smaller windows inside larger systems [49, 44, 41], and how the observables change depending on the size of the window or the system simulated. The second is a quick look at the effects of boundary conditions (BC) [27], and there we explain what observations can be made by considering not only the fully periodic BC, but also antiperiodic ones in several, or all directions.

5.2 The 4D EA Model

In this section we consider an additional model, the 4D EA model [8, 62, 14, 60], to see whether the observations we made in the previous chapter is generic to the low temperature phase of short-range spin glasses, or is specific to three dimensions. There is a consideration about an upper critical dimension D_u, beyond which corrections to the mean field approach would be insignificant ($D_u \geq 6$) [68]. By considering $D = 4$ we are staying below that limit, as well.

Table 5.1. Simulation parameters for the 4D EA spin glass. For each number of spins $N = L^4$ we equilibrate and measure for 2^b Monte Carlo sweeps. $T_{\rm min}$ [$T_{\rm max}$] is the lowest [highest] temperature used and N_T is the number of temperatures. $N_{\rm sa}$ is the number of disorder samples.

N	L	b	$T_{\rm min}$	$T_{\rm max}$	N_T	$N_{\rm sa}$
256	4	23	0.7200	2.3800	52	3252
625	5	23	0.9101	2.3800	42	4086
1296	6	23	0.7200	2.3800	52	3282
2401	7	23	0.9101	2.3800	42	4274
4096	8	23	0.7200	2.3800	52	3074
6561	9	24	0.7200	2.3800	52	3010

The EA model, as explained in the previous chapter, is defined by the Hamiltonian

$$\mathcal{H} = -\sum_{\langle i,j\rangle}^{N} J_{ij} s_i s_j, \tag{5.1}$$

with $s_i \in \{\pm 1\}$ Ising spins. For the 4D version, the sum is over nearest neighbors on an hypercubic lattice of size $N = L^4$ with periodic boundaries. The couplings J_{ij} are again chosen from a Gaussian distribution with zero mean and standard deviation unity. The parameters of the replica exchange Monte Carlo simulations are given in Table 5.1. We are reporting results for $T = 0.9101$, which corresponds to $T \approx 0.5T_c$ for this model ($T_c = 1.78$ [60]) . This allows us to compare with the previously investigated models, since at this temperature our system is deep within the spin-glass phase. Additionally, we use the SK model data at a similar reduced temperature for our one-to-one comparisons, and we added one more system size, $N = 3072$, to the SK model data as well.

The largest system size we have simulated for this model is $L = 9$. We believe it is important to start by questioning the equilibration of these simulations, since we are now dealing with much larger systems compared to the previous ones we simulated. Figure 5.1 shows representative results for the method introduced in [31] and used in Section 3.3. Here, we can see that at the end of simulation of 2^{24} time steps, the model can be considered well equilibrated, since at least the last four of the measurements fully overlap. However, we closely examined the time-evolution of the link overlap q_l and the overlap square q^2 values for individual samples, and discovered cases where those values are not converging

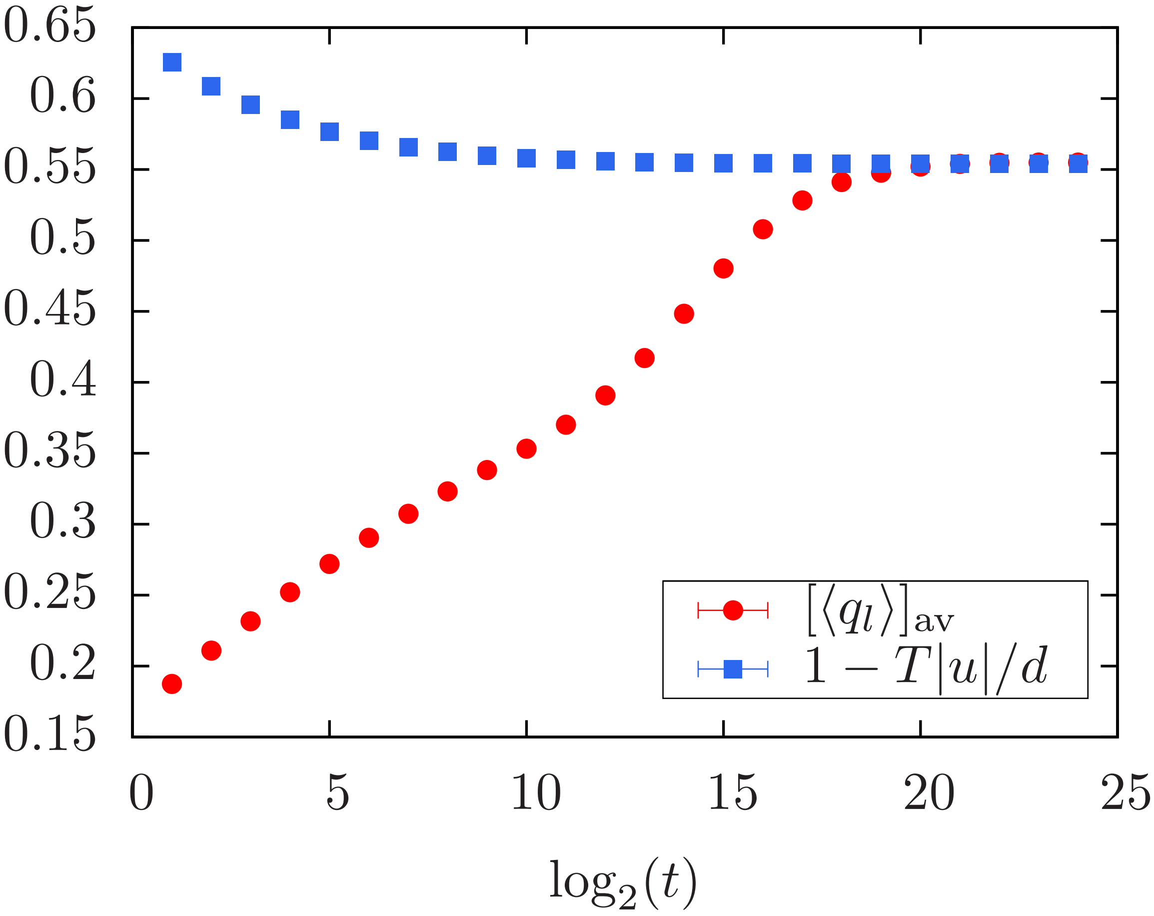

Figure 5.1. Average link overlap $[\langle q_l \rangle]_{\mathrm{av}}$ (red circles) and $1 - T|u|/d$ (blue squares), where u is the energy (see Eq. (3.10)), as a function of time t for $L = 9$ and $T = 0.9101$. Error bars are smaller than the symbols.

fast enough. Hence, we heed our own results from Ch. 3 and take a more conservative approach by discounting any conclusions made solely based on the largest system sizes. As we will see, some of the $L = 9$ of the 4D EA model and $N = 3072$ of the SK model results do seem anomalous.

The first step towards understanding this model is to look at the disorder-averaged overlap probability distribution $P(q)$ for all simulated system sizes. In Fig. 5.2 we see that significant finite size effects are present, as can be noted from the change in positions of the q_{EA} peaks with L. The distributions seem to be overlapping in the middle section, around $q = 0$ with the possible exception of the largest system $L = 9$, which is lower. This effect

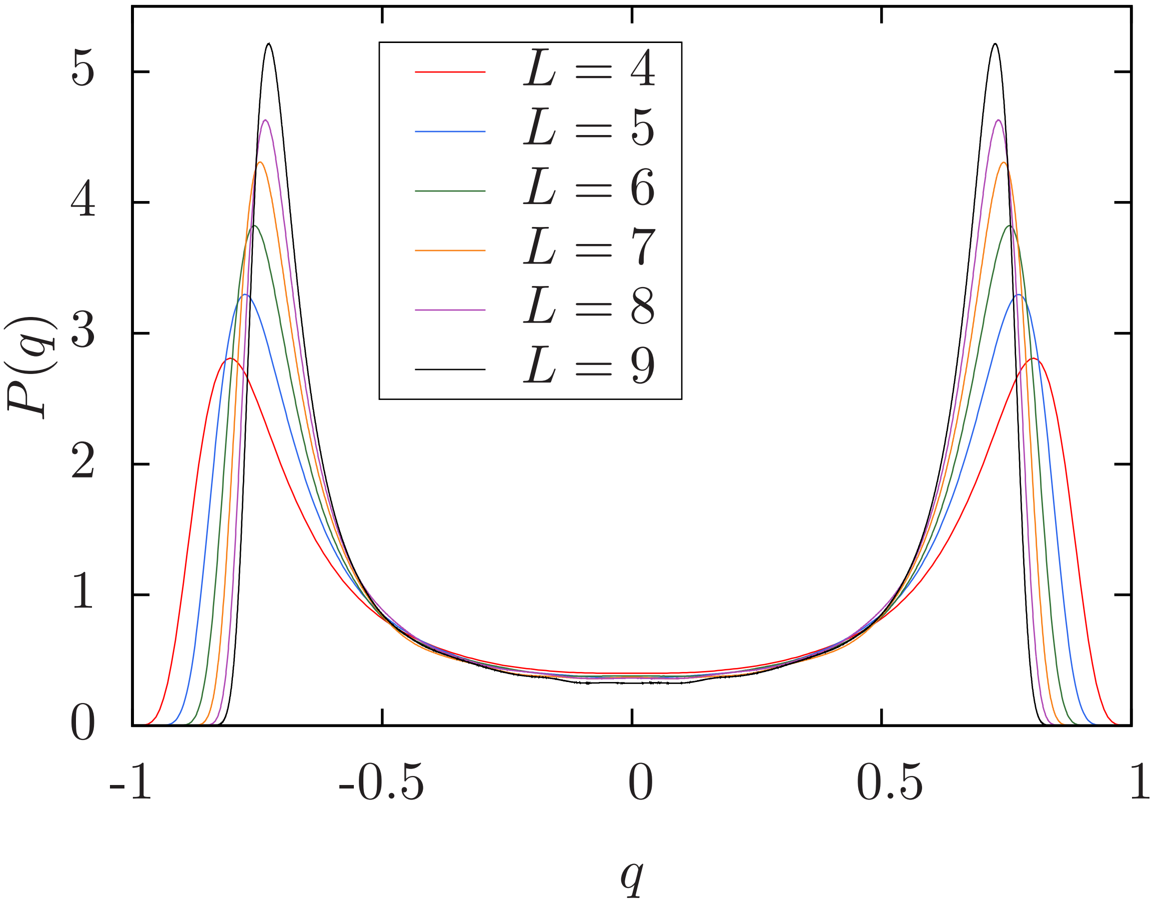

Figure 5.2. Disorder-averaged overlap probability distribution $P(q)$ for different system sizes at $T = 0.9101$ for the 4D EA model.

can be better observed in Fig. 5.3, where we plotted $I(0.2)$, the disorder average of the weight of the overlap distribution falling in that middle region (see Eq. 3.3). Here, $L = 9$ systems clearly have smaller $I_J(0.2)$, but as we mentioned earlier in this section, we will not draw any conclusion from the smaller value of $I(0.2)$ for $L = 9$ only. All other values are within each others error bars for the 4D EA model, similar to what has been observed for the 3D EA model.

Following the same path as in our previous chapter, we now consider not only the averages, but other statistics of individual $P_J(q)$'s as well. We are interested in the δ functions in individual probability distributions, and if they occur more and more frequently as we go to larger and larger systems. To do that, we use $\Delta(q_0, \kappa)$, defined in Eq. 4.4. Figure 5.4

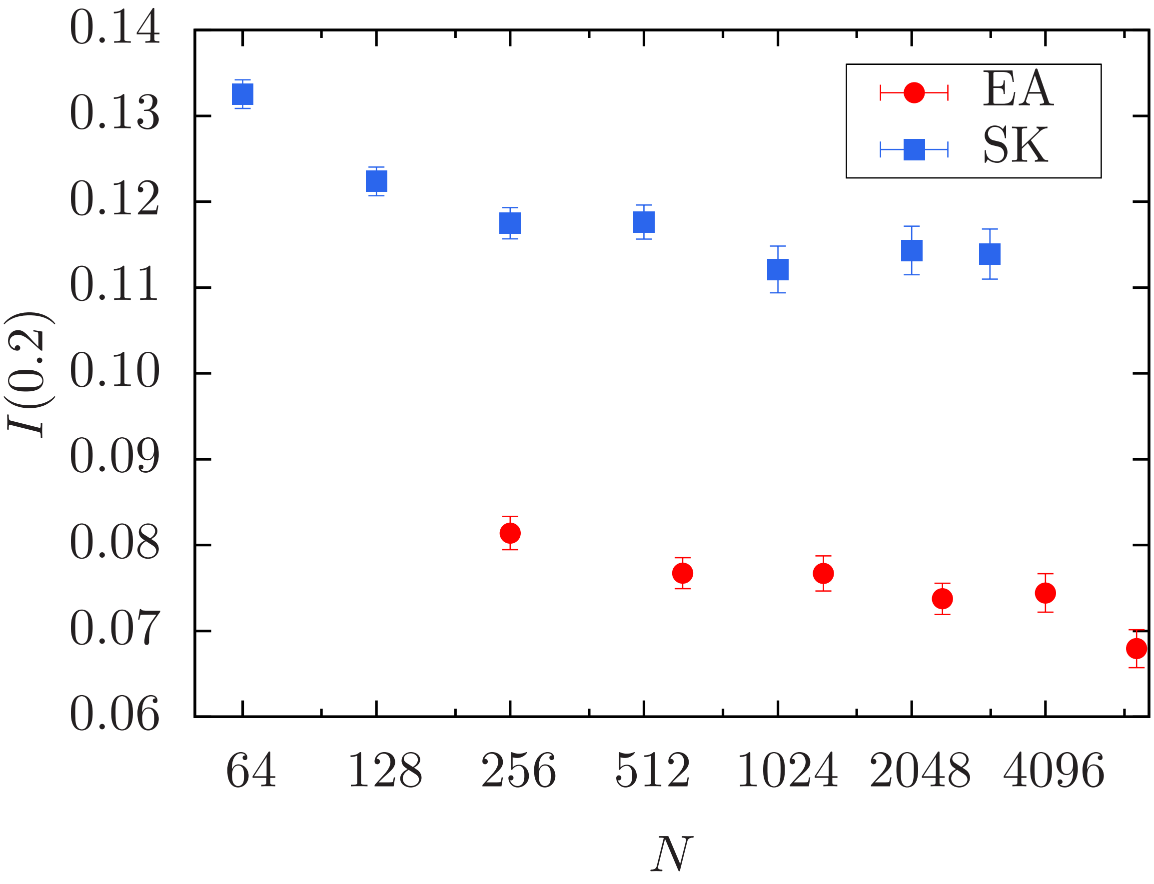

Figure 5.3. Disorder average of the weight of the overlap distribution $I(0.2)$ as a function of N for $T \approx 0.5T_c$ for both the EA and SK models. The data are nearly independent of N with the exception of the two smallest sizes for the SK model, and the largest size for the EA model. See text for an explanation of the latter.

shows a plot of $\Delta(q_0, \kappa)$ vs N for $\kappa = 1$ and $q_0 = 0.2$ and 0.4 for both the 4D EA and SK models. Here we can easily make the same observation as we did when comparing the 3D EA model with the SK model: $\Delta(q_0, \kappa)$ for the SK model increases much faster compared to the $\Delta(q_0, \kappa)$ of the 4D EA model. Once again in the plot we see a slight decrease for the largest system size of $L = 9$, and if true, this would be the first evidence we observe for the onset of the scenario predicted by the droplet picture, namely $\Delta(q_0, \kappa)$ going to zero. However, we feel we need at least one more, larger system size in comparison, and proof that all samples are fully equilibrated, to be able to reach a definite conclusion. Finally, Figures 5.5 and 5.6 are examples of different, less ideal choices of κ. Both plots only become informative for the largest few systems, therefore making it hard to deduce general trends from.

In conclusion, the observations made on the 4D EA model agree qualitatively with the results we obtained from the 3D counterpart. At low temperature, it behaves differently compared to the SK model, suggesting a different picture from the RSB hypothesis.

5.3 Additional statistics from $P(q)$

A recent paper by Alan Middleton suggests various ways to extract useful information from the overlap probability distribution [45]. He considers the quantity $I_J(q_0) = \int_{-q_0}^{+q_0} P_J(q)dq$, and defines its sample median as $\tilde{I}(q_0)$, in addition to the mean of the same function, $I(q_0)$. We decided to compare the mean $I(q_0)$ and the median $\tilde{I}(q_0)$ for our data, and Figures 5.7, 5.8 and 5.9 show both quantities on log-linear plots for the 3D EA, 4D EA and SK models as a function of q_0, respectively. For small values of q_0, $\tilde{I}(q_0)$ is clearly smaller than $I(q_0)$ for all models, indicating that the average is actually determined by a few samples with very large $I_J(q_0)$ values. Additionally, we see that for the small q_0 values $\tilde{I}(q_0)$ decreases with system size, whereas it would go the opposite way for large q_0 values. We observe a cross-over roughly around $q_0 = 0.6$ for all models. In conclusion, we noticed that for the SK model, the $I(q_0)$ and $\tilde{I}(q_0)$ values are closer to each other then they are for

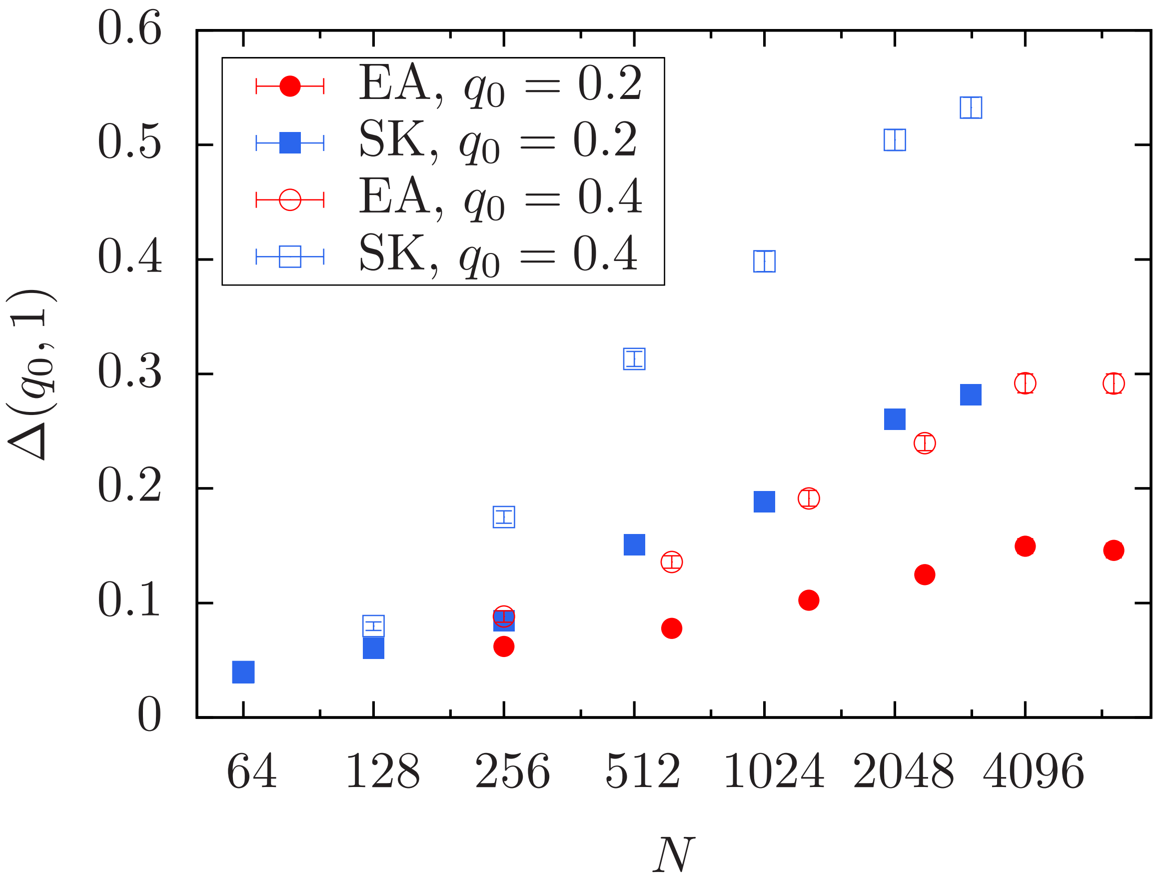

Figure 5.4. Fraction of peaked samples $\Delta(q_0, \kappa)$ at $T \approx 0.5T_c$ as a function of N for $\kappa = 1$, $q_0 = 0.2$ and 0.4.

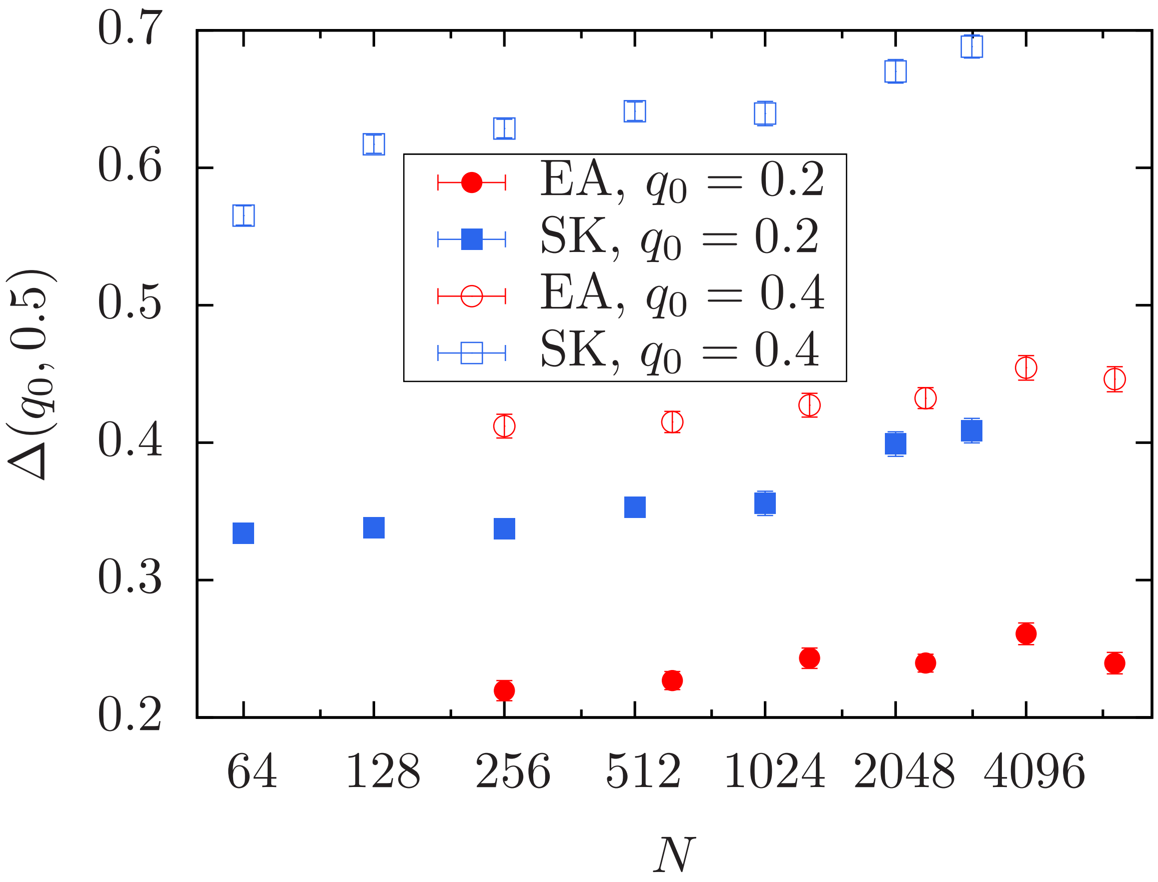

Figure 5.5. Fraction of peaked samples $\Delta(q_0, \kappa)$ at $T \approx 0.5T_c$ as a function of N for $\kappa = 0.5$, $q_0 = 0.2$ and 0.4.

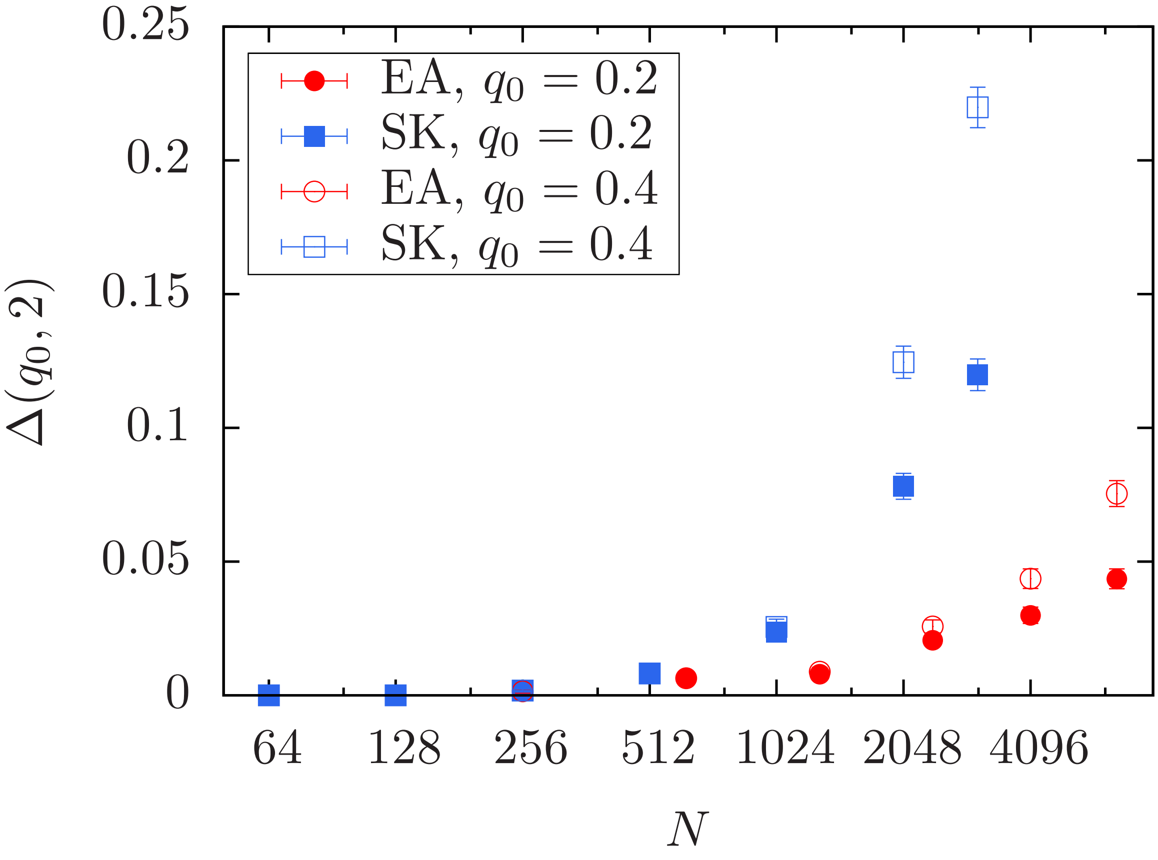

Figure 5.6. Fraction of peaked samples $\Delta(q_0, \kappa)$ at $T \approx 0.5T_c$ as a function of N for $\kappa = 2$, $q_0 = 0.2$ and 0.4.

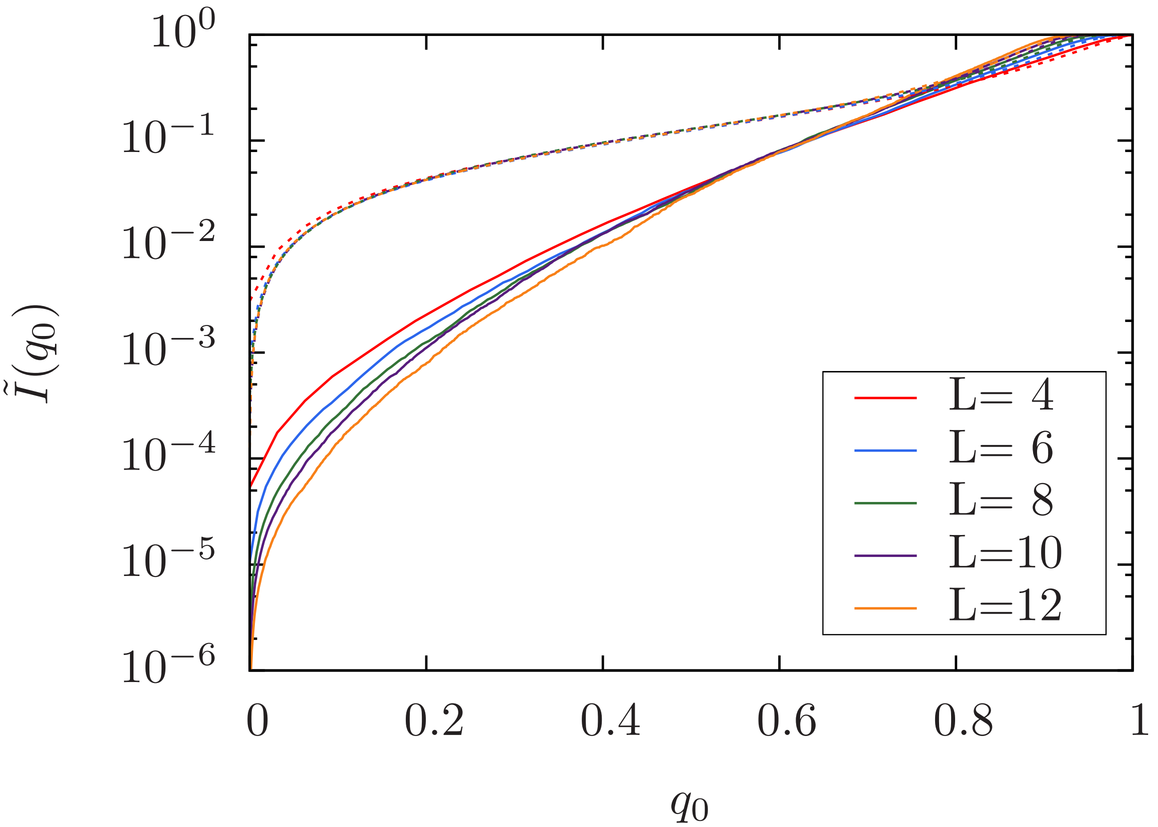

Figure 5.7. Log-Linear $\tilde{I}(q_0)$ (median) plot for the 3d EA model at $T \sim 0.42$. The dashed lines indicate $I(q_0)$ (mean).

both EA models, but we did not see a clear qualitative difference between the EA and the SK models otherwise.

Shortly after Middleton's paper, another observable based on the overlap probability distribution was introduced by Monthus and Garel [46], called the "typical value" $P^{typ}(q)$. It is constructed in the following way

$$P^{typ}(q) = \exp(\overline{\ln P_J(q)}), \tag{5.2}$$

where the bar above indicates an average over disorder realizations. The difference from the usual average probability distribution $P(q)$ is that here we take the logarithm of individual $P_J(q)$s, then average those before finally taking the exponential of that average. The purpose of this procedure is to avoid the disproportionate contribution to the average from

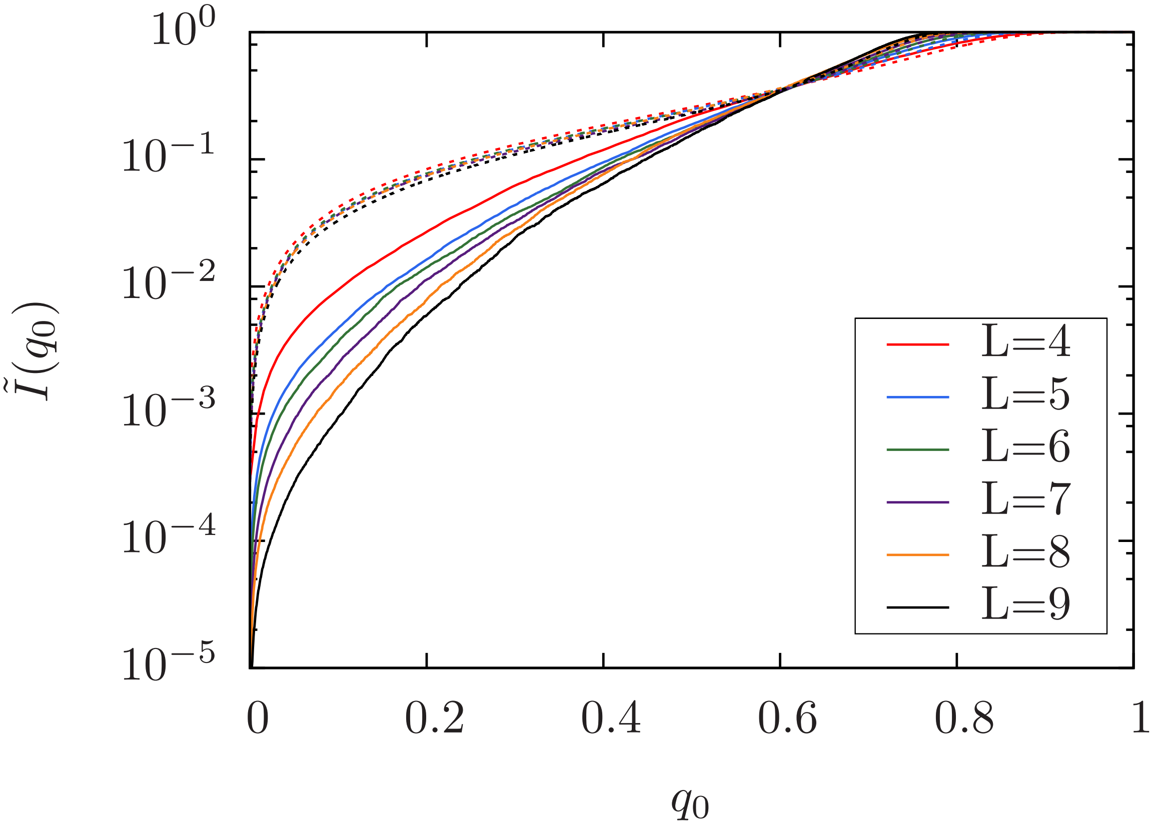

Figure 5.8. Log-Linear $\tilde{I}(q_0)$ (median) plot for the 4d EA model at $T \sim 0.9$. The dashed lines indicate $I(q_0)$ (mean).

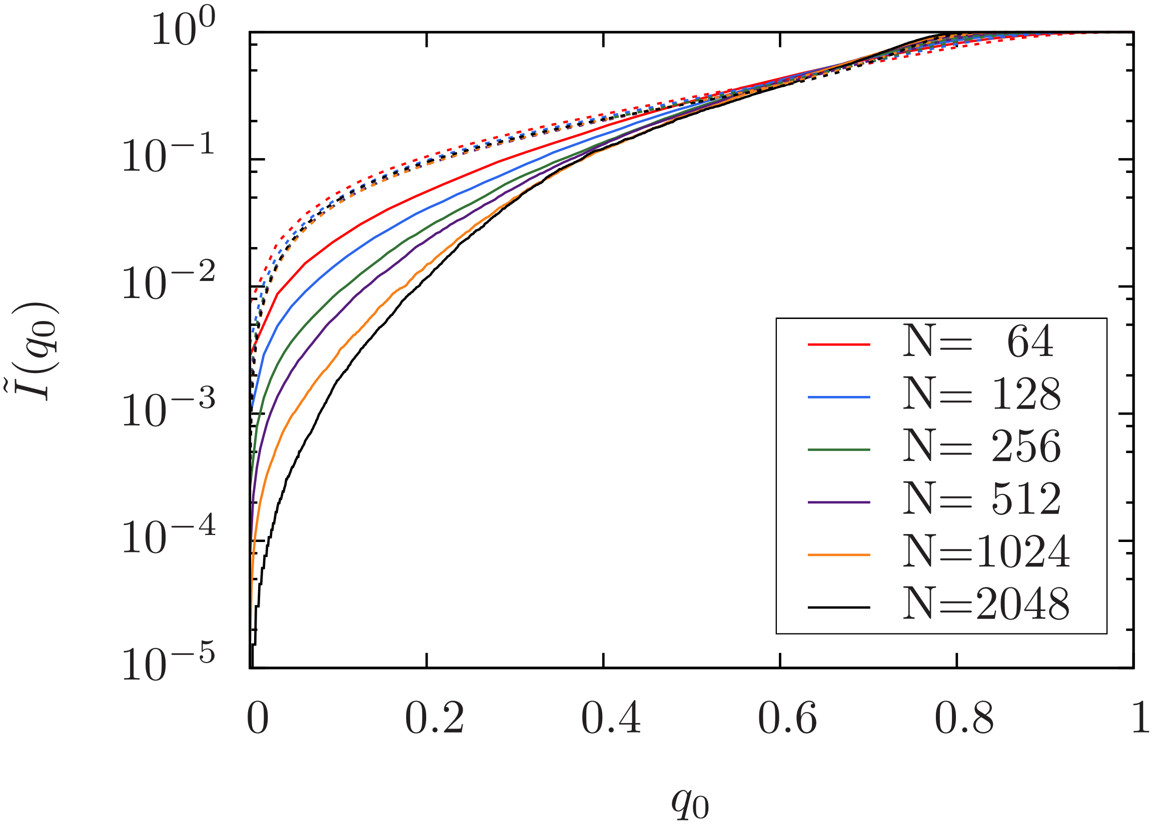

Figure 5.9. Log-Linear $\tilde{I}(q_0)$ (median) plot for the SK model at $T \sim 0.42$. The dashed lines indicate $I(q_0)$ (mean).

rare samples with large values of $P_J(q)$ and to observe the true typical behavior. The same procedure can be done for $I_J(q_0)$, thus obtaining an $I^{typ}(q_0)$ value.

$$I^{typ}(q_0) = \exp\Big(\overline{\int_{-q_0}^{+q_0} \ln P_J(q)dq}\Big). \tag{5.3}$$

These quantities should be easy enough to calculate, but because of the limitations of our data, we run into one important problem: Since we are simulating finite systems, $P_J(q)$ should actually be greater than zero for all q values, but many of our samples have zero weight at smaller q values because of our finite run length in measuring $P_J(q)$. This makes it impossible to take the natural logarithm directly. To overcome this, we decided to do the following: The smallest value possible to be generated by a simulation, ϵ, is equal to the inverse of the total running time i.e. if a simulation is run for 2^{27} time steps, $\epsilon = 2^{-27}$. Thus, to estimate the effect of this lower limit, we replace all zero weights by a fixed fraction of ϵ to calculate the $P^{typ}(q)$ and $I^{typ}(q_0)$ with the hope that the results would be independent of the choice of said fraction. Figures 5.10 to 5.15 show the log-linear plots of these quantities, measured using $\epsilon/50$ in place of zero, for all of our models.

We found out, however, that replacing the zero weights by a fixed fraction of ϵ creates a strong dependence on that fraction, an example of which can be seen in Fig. 5.16. This is the case because a considerable percentage of our samples have $P_J(q) = 0$ for smaller q values. To give an example, 20% of the samples of the size $L = 8$ 3D EA model have $P_J(0) = 0$. Therefore, conclusions drawn from the quantities $P^{typ}(q)$ and $I^{typ}(q_0)$ can be misleading, since there is no clear way of accurately estimating typical values due to artifacts created by our finite run length. For completeness sake, we still saw fit to include these calculations, so that they may constitute a starting point for further research, but while the typical values $P^{typ}(q)$ and $I^{typ}(q_0)$ may have theoretical value, their accurate measurement in MC simulations does not appear to be feasible.

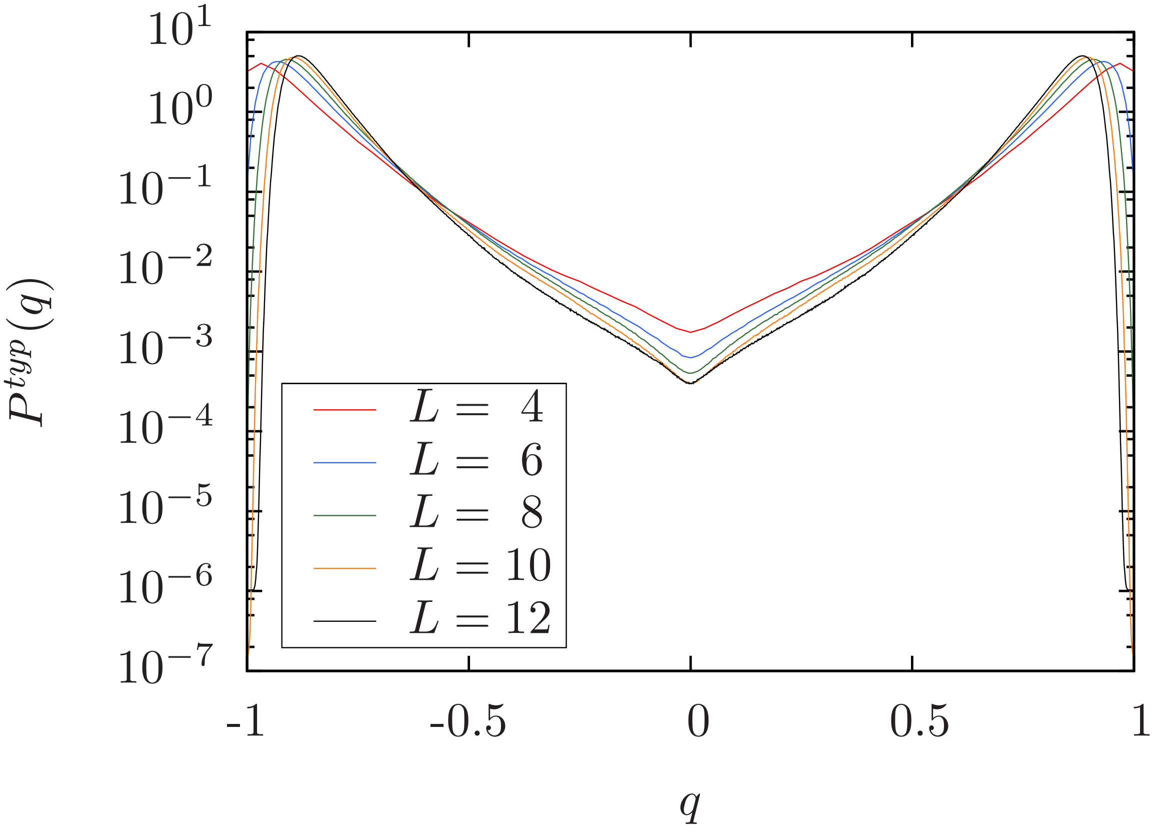

Figure 5.10. Log-Linear $P^{typ}(q)$ plot for the 3d EA model at $T \sim 0.42$. All zeros were replaced by $\epsilon/50$.

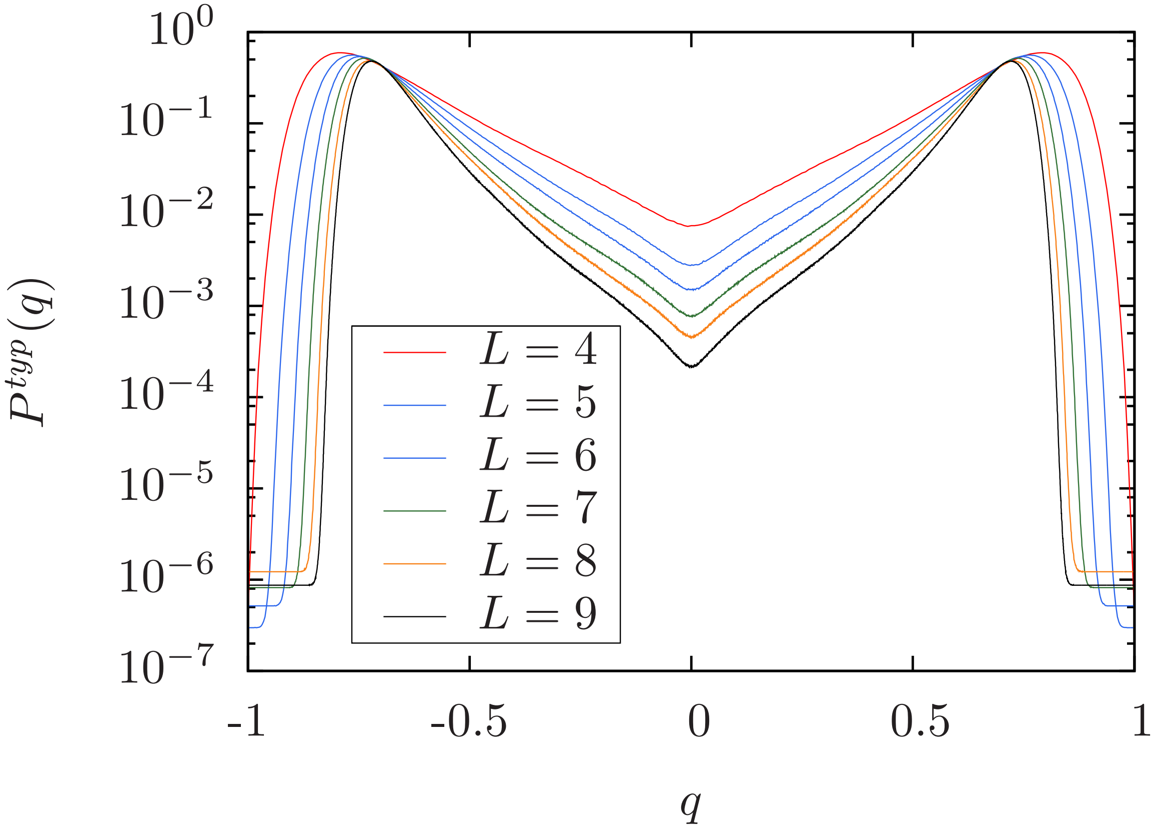

Figure 5.11. Log-Linear $P^{typ}(q)$ plot for the 4d EA model at $T \sim 0.9$. All zeros were replaced by $\epsilon/50$.

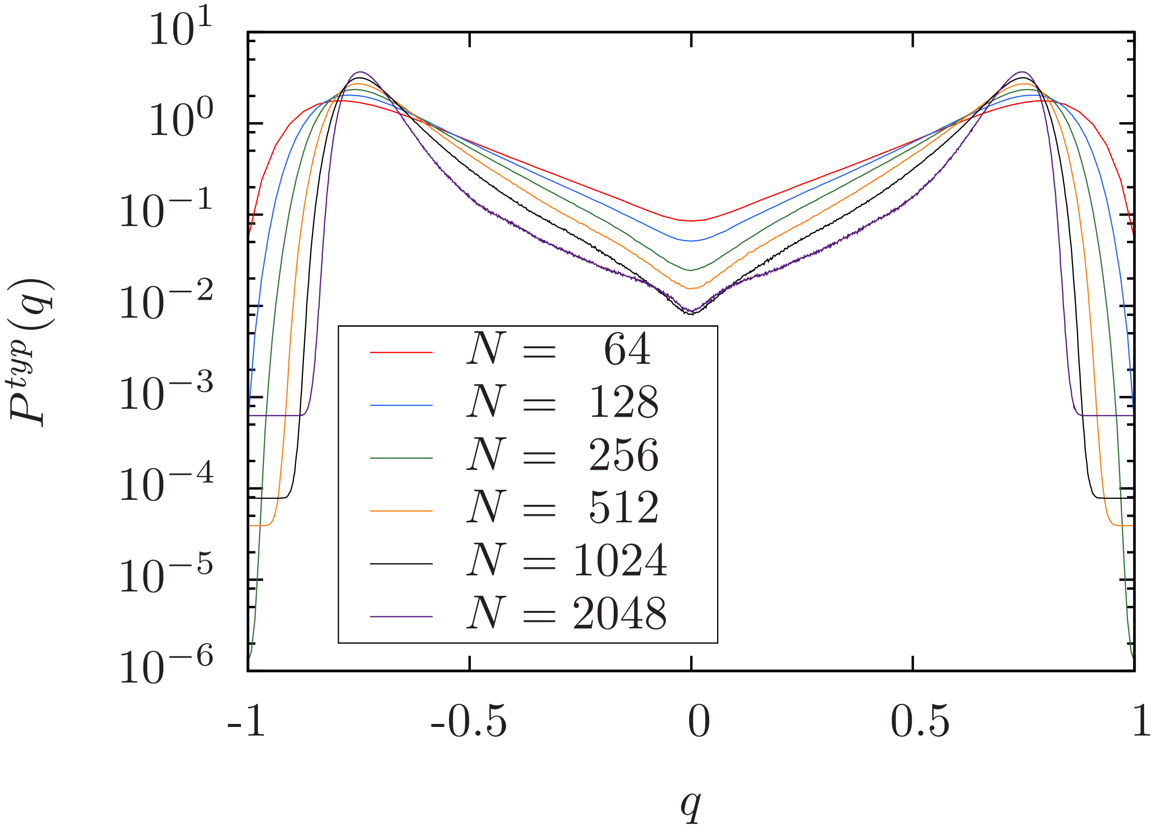

Figure 5.12. Log-Linear $P^{typ}(q)$ plot for the SK model at $T \sim 0.42$. All zeros were replaced by $\epsilon/50$.

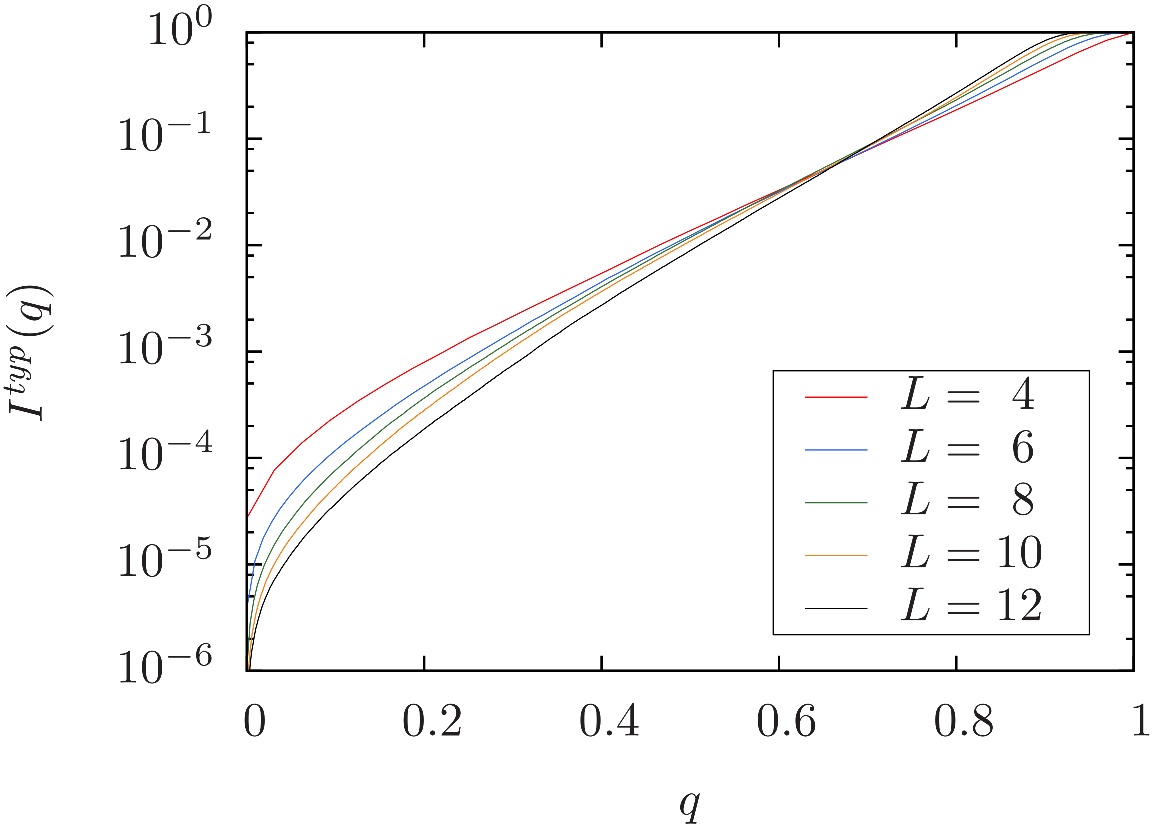

Figure 5.13. Log-Linear $I^{typ}(q)$ plot for the 3d EA model at $T \sim 0.42$. All zeros were replaced by $\epsilon/50$.

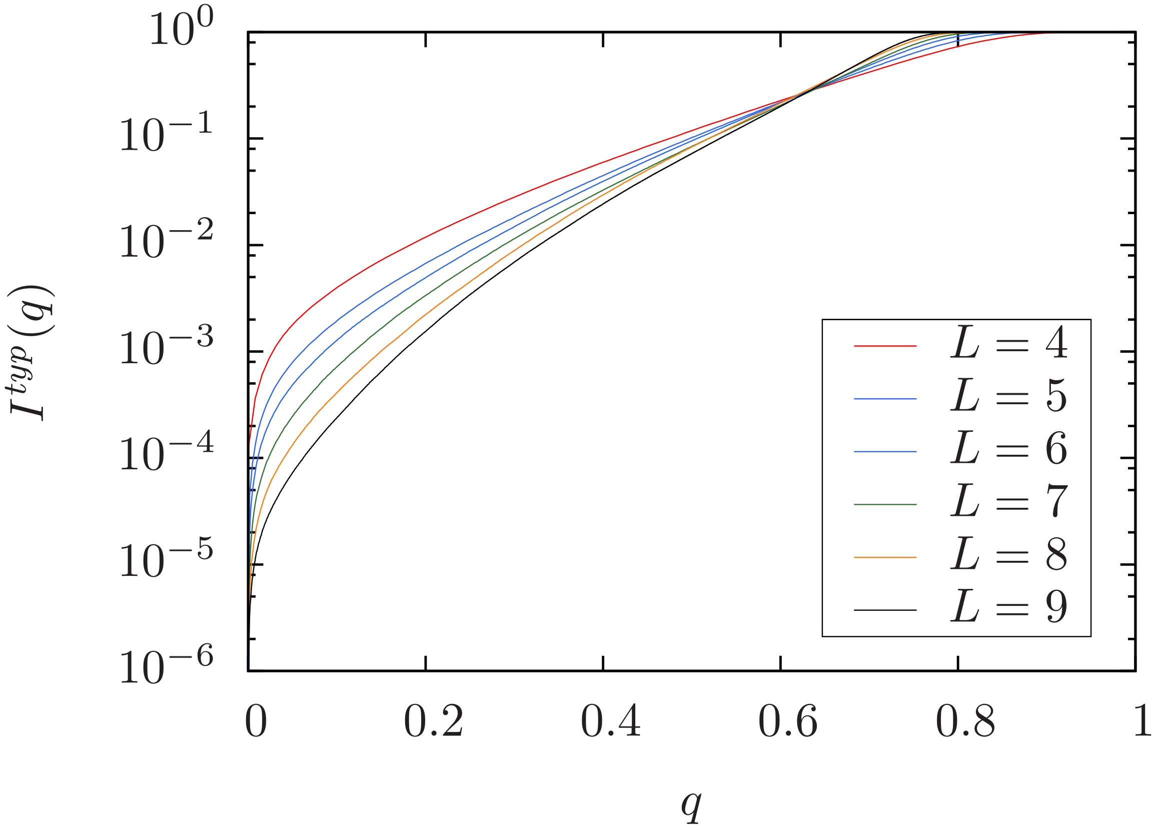

Figure 5.14. Log-Linear $I^{typ}(q)$ plot for the 4d EA model at $T \sim 0.9$. All zeros were replaced by $\epsilon/50$.

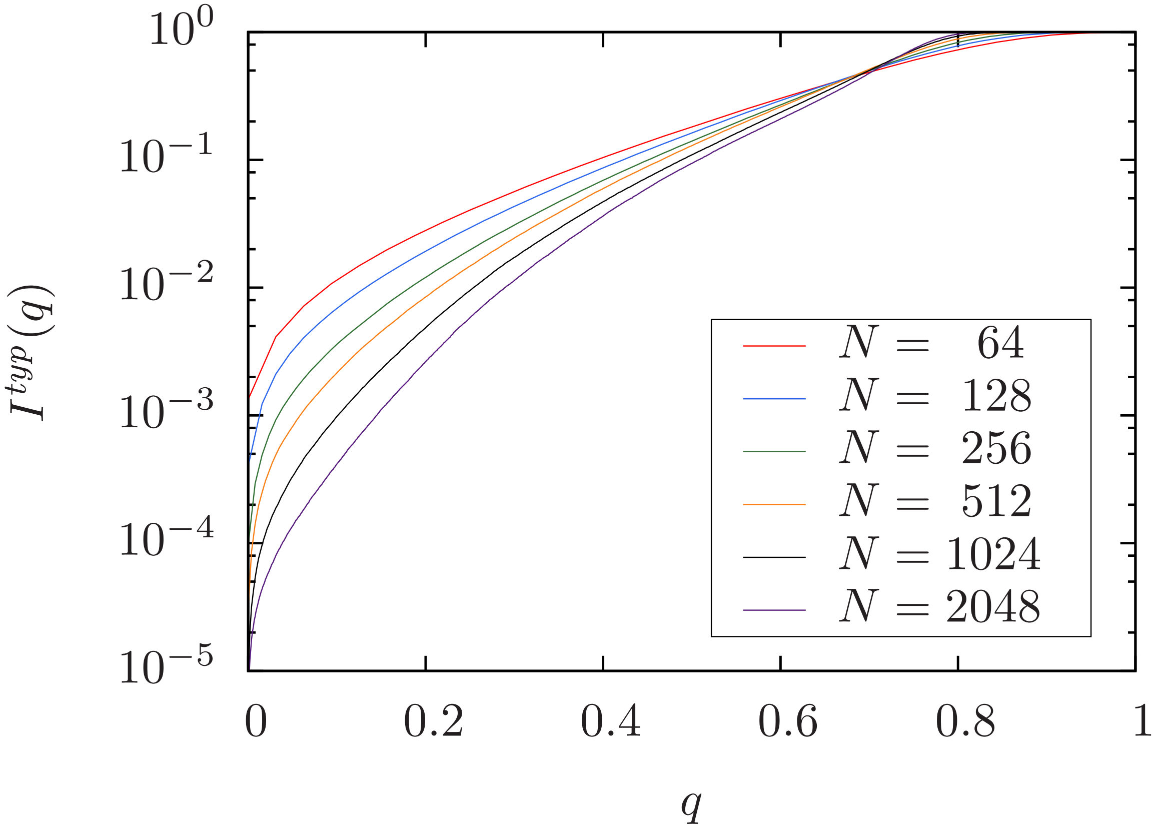

Figure 5.15. Log-Linear $I^{typ}(q)$ plot for the SK model at $T \sim 0.42$. All zeros were replaced by $\epsilon/50$.

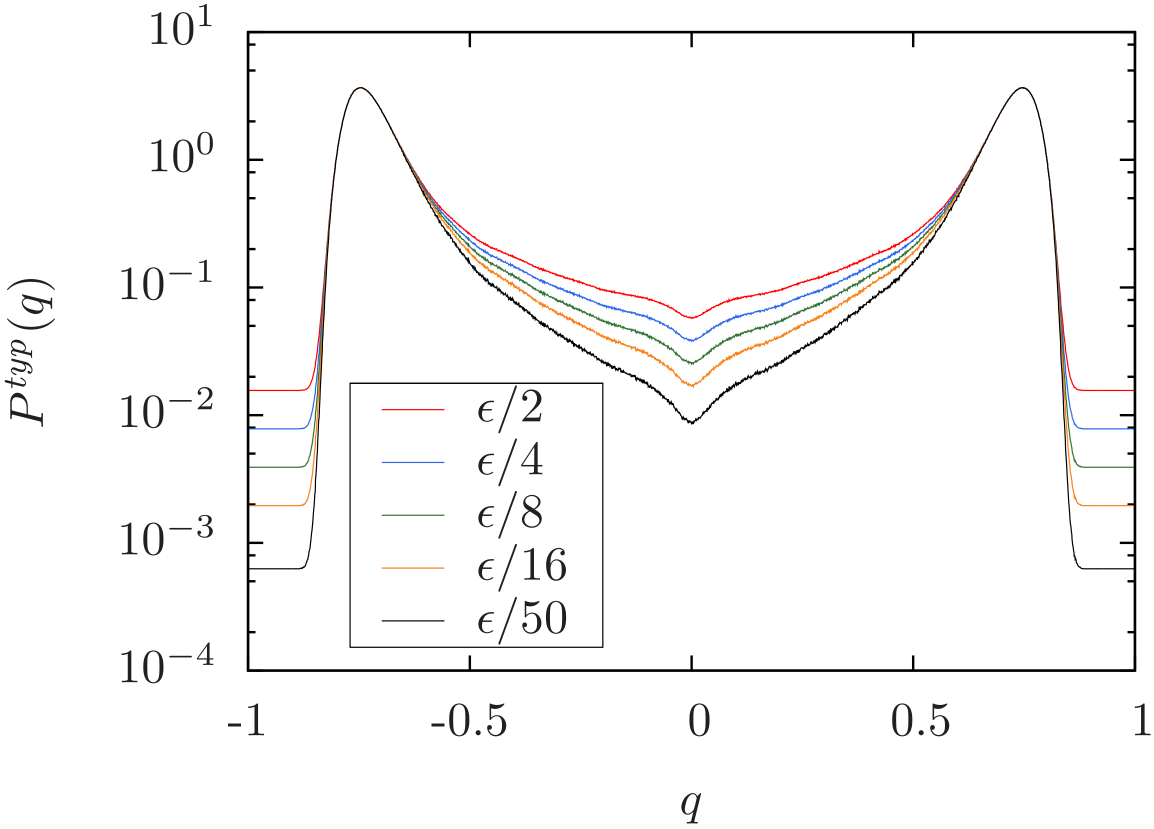

Figure 5.16. Log-Linear $P^{typ}(q)$ plot for the SK model for N=2048 showing the dependency on the zero-replacement value.

5.4 Windows

A method developed to define the infinite-volume limit for disordered systems theoretically is to consider fixed volumes, or 'windows', embedded in systems of successively larger sizes [49, 44, 41]. This method originates from the concern that for small system sizes, the chosen boundary conditions would introduce domain walls to the system, causing droplet-like excitations. Consequently, if only a small window far away from the boundary is considered, where such a domain wall would be unlikely to go through, more accurate observations can be made.

We wanted to test this idea numerically by selecting a window inside the system and comparing the observables obtained from that window only with the measurements on larger windows, the whole system, or even another system with the same size as the selected window. To do that, we took measurements on all but the largest system sizes of the 3D EA model for such windows. For example for a size $L = 10$ system we collected data for overlaps for cubic windows of size $l = 4, 6$ and 8. Here we present what we have observed in two ways: First, we compare the effects of the size l of a window in a large system of $L = 10$ of the 3D EA Model, at $T = 0.2$ (Scenario A). Next, we address the main concern described above by investigating how a fixed window of $l = 4$ behaves when simulated as a part of different system sizes L, again at $T = 0.2$ (Scenario B).

As with previous experiments, we start by looking at the disorder-averaged overlap probability distribution $P(q)$. $P(q)$ is symmetrical, and since the main divergence occurs at the q_{EA} peaks, in Figures 5.17(a) and 5.17(b) we are showing the large-q behavior, since in the region around $q = 0$, the lines fall on top of each other. This fact can be better seen in Fig. 5.18, where we plotted $I(q_0)$ for a small region of $q_0 = 0.2$. The fluctuations in both cases are quite small. The finite size effects, however, are still significant as can be seen by the changing heights and positions of the peaks in Fig. 5.17(a). The right version (b) does not indicate to such change, with the exception of our smallest system size $L = 4$, which is the same size as our chosen window.

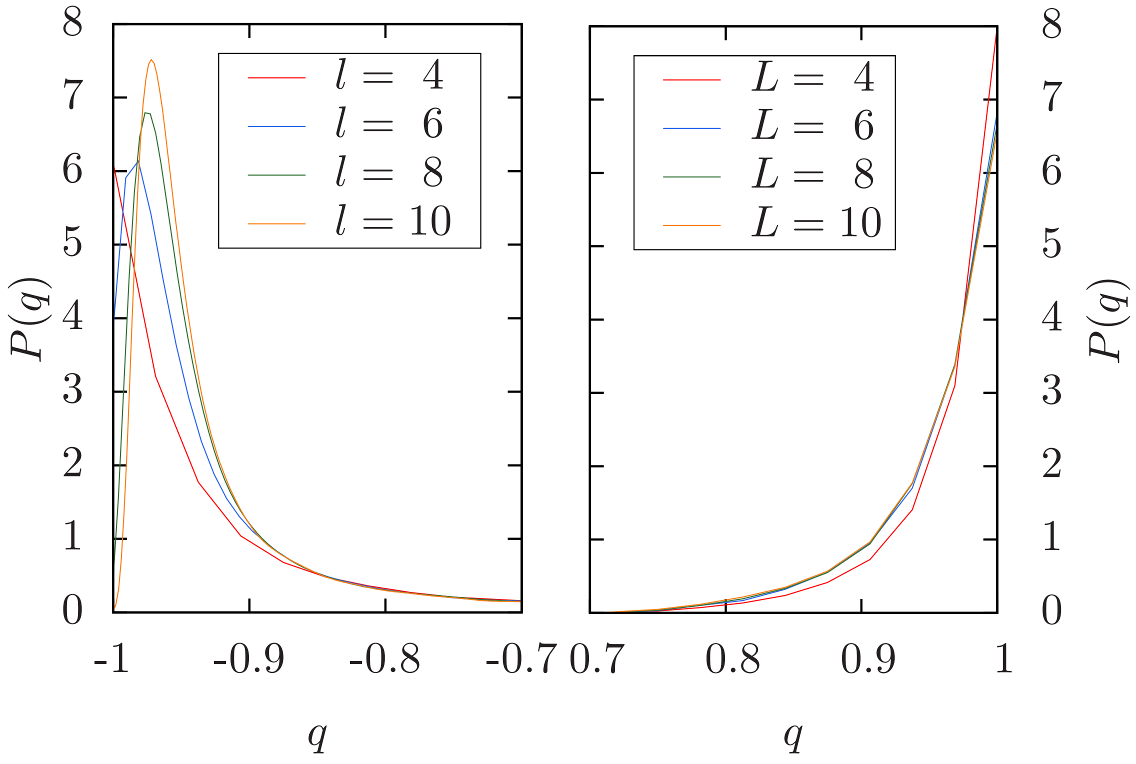

Figure 5.17. Large q behavior of the disorder-averaged overlap probability distribution $P(q)$ (a) for different window sizes l for the system size $L = 10$, (b) for a window of $l = 4$ in different system sizes L, both at $T = 0.2$.

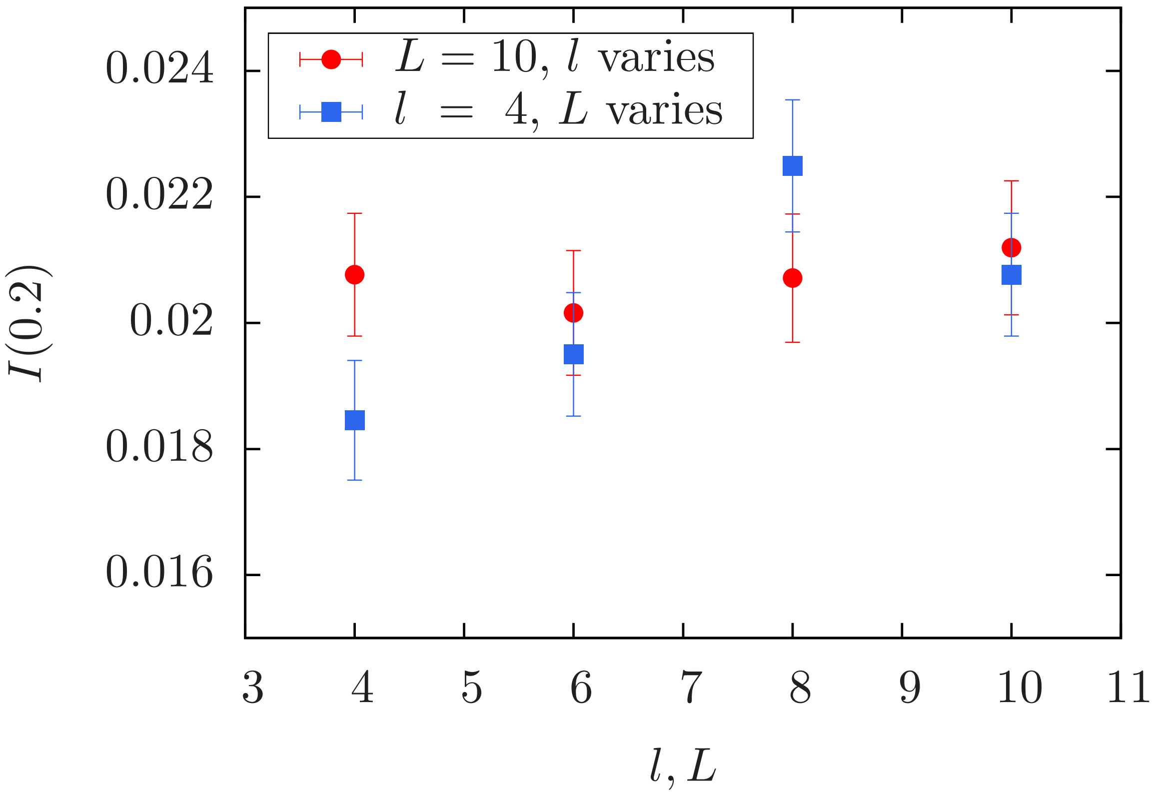

Figure 5.18. Disorder average of the weight of the overlap distribution $I(0.2)$ as a function of system/window sizes for $T = 0.2$.

Now we can look for the emergence of δ functions in individual $P_J(q)$'s for scenario A, by increasing the window size within the same system, as well as for scenario B, simulating the same size window as part of larger and larger systems. Fig. 5.19 shows $\Delta(q_0, \kappa)$ defined in Eq. 4.4 for scenario A, and Fig. 5.20 for scenario B. We first see that increasing the window size affects $\Delta(q_0, 1)$, and more and more δ functions emerge at first. But this increase is rather slow, showing the possibility of reaching a plateau, as we have seen the system behave in Fig. 4.6 of the previous chapter. In the following figure corresponding to Scenario B, however, we see no effect on $\Delta(q_0, \kappa)$ from the increase in system size, as long as we keep the window size the same. This, combined with what we observe in Fig. 5.17(b) is actually a striking result, since it indicates that looking at a small window of size l embedded in a larger system of size L does not give any different results than considering the window alone, $l = L$.

Both figures show that we are not able to gain significant additional insights about the nature of the low temperature behavior of the 3D EA model by looking at the windows in the way we did. Larger systems will have to be simulated, with varying windows inside them measured separately, in order to gain further information, but for now, our original way of considering whole systems seems to be a better way of analyzing this model. The main conclusion that we can draw from this experiment is that results obtained from windows are not significantly different than those obtained from full systems.

5.5 Changing the boundary conditions

In this section we briefly mention the effects of boundary conditions (BC) [27] by presenting a preliminary study we did on ~ 1000 samples of size $L = 8$ of the 3D EA model. In all of our previous studies, we only considered periodic BC in all directions with the hope of minimizing finite size corrections. An equally valid choice would be to use antiperiodic BC in one, a few, or all directions, yielding eight possible choices of BC in total. Here we wanted to test if such a choice would have any effect on the samples simulated, es-

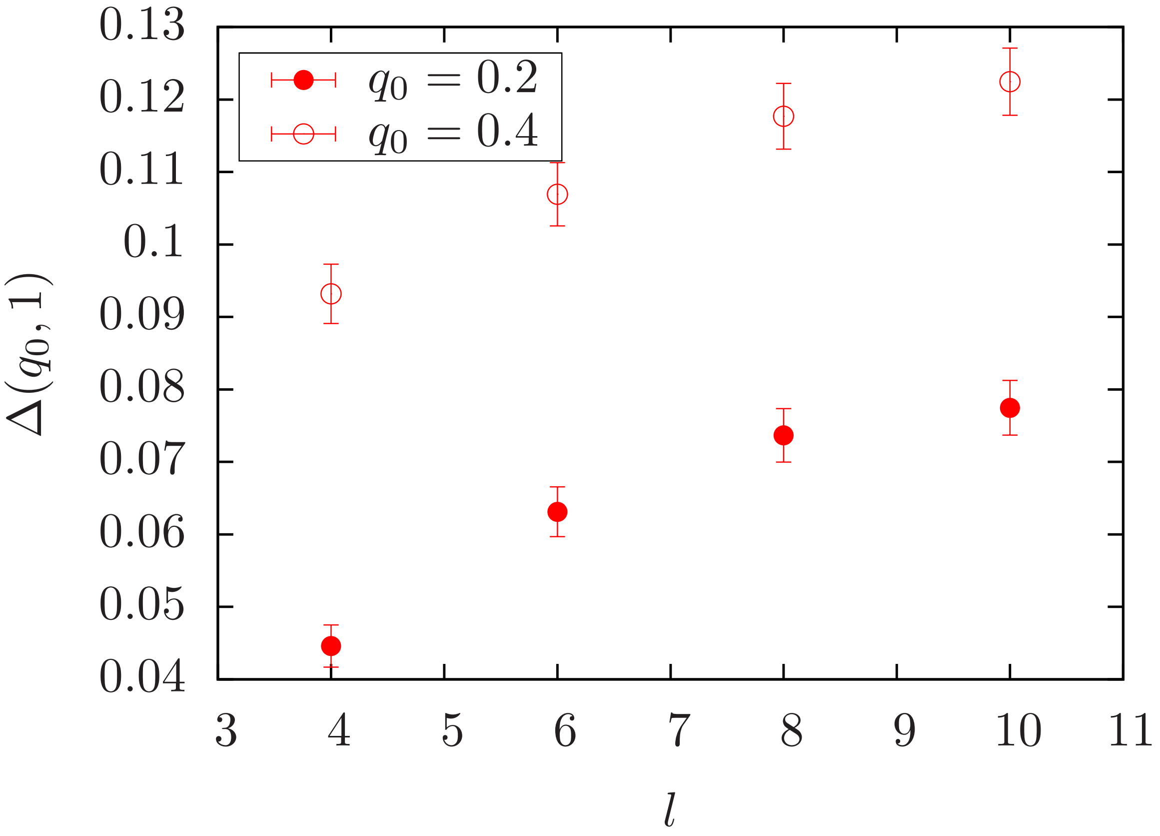

Figure 5.19. Fraction of peaked samples $\Delta(q_0, \kappa)$ at $T = 0.2$ for different windows l in the system size $L = 10$. Here $\kappa = 1$ and $q_0 = 0.2$ and 0.4. We observe a slow increase in Δ as the window size increases.

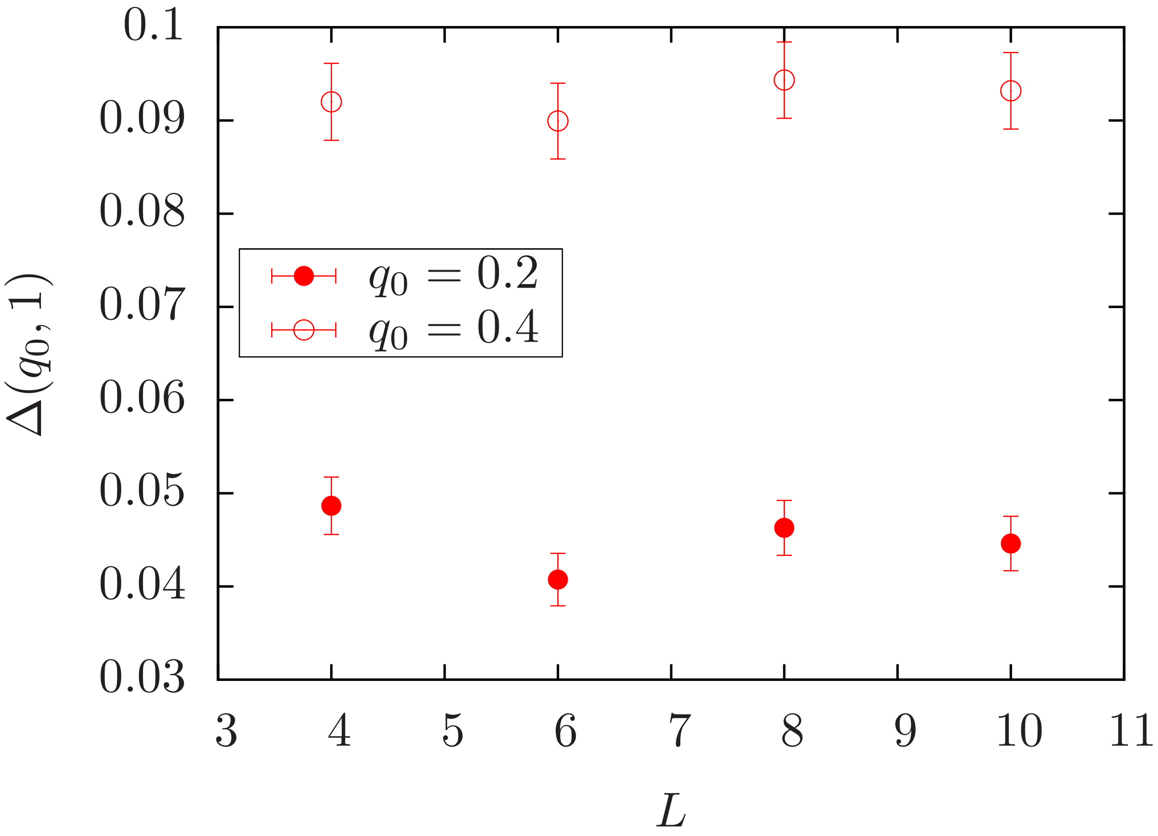

Figure 5.20. Fraction of peaked samples $\Delta(q_0, \kappa)$ at $T = 0.2$ for a window of size $l = 4$ in varying system sizes L. Here $\kappa = 1$ and $q_0 = 0.2$ and 0.4. Changing the system size simulated does not seem to have any effect on Δ, if the size of the window considered remains the same.

pecially on the value of $I_J(0.2)$, given that the bond configuration $\{J_{ij}\}$ remained the same except on the boundary. We chose such samples that have a large $I_J(0.2)$ value (large compared to ϵ defined in Sec. 5.3), and we stopped trying different BC once we encountered one that would result in $I_J(0.2) < \epsilon$. This test provided us with one interesting observation: For every single one of our disorder realizations, at least one choice of BC resulted in $I_J(0.2)$ being smaller than ϵ. We believe this is a line of research worth pursuing, since it indicates that some observables, especially $I_J(0.2)$ may be sensitive to BC. Therefore, for future work, simulating an ensemble with all eight BC and then averaging over those can be considered.

5.6 Conclusions

In this chapter we wanted to fill in some of the details left out in the previous chapter, as well as make use of the newest developments in the field that came after the paper [73] on which Ch. 4 is based was published. We showed results of another simulation, this time of the 4D EA model, looked at some interesting quantities such as $\tilde{I}(q_0)$ and $P^{typ}(q)$, and did some preliminary work to motivate future projects. The observations made in this chapter in general seem to lend further support to our conclusion in the previous chapter, namely that the EA model, both in 3D and 4D, does not seem to behave mean-field-like, i.e. it does not display a trend towards many pairs of pure states as opposed to only a single pair. In the following chapter we try to elaborate on what can be done in terms of future research to investigate this point further.

CHAPTER 6

SUMMARY AND FUTURE RESEARCH

Our main purpose in this study was to understand the low temperature behavior of spin glasses, and to answer the question at the heart of a controversy which has been occupying many physicists' minds for decades: Are there only a single pair or a countable infinity of pure states in the low temperature phase of the Edwards-Anderson spin glass?

To that end, we started first by giving a short description of spin glasses and their history, then briefly explained what the controversy is by describing the competing theories and subsequent attempts at uniting those.

For our research we have undertaken a massive numerical study, using the parallel tempering Monte Carlo algorithm, and in our second chapter we introduced this method of simulation. We discussed how this algorithm works and why it is well suited to simulate spin glasses and explained how we implemented it using the C++ programming language. In the following chapter, we proceeded to explain how and how effective this algorithm works, by laying out the observations we made by carefully studying the simulation dynamics. From these dynamics studies we found out that the degree of roughness of the free-energy landscape of each individual disorder realization sample was directly affecting the sample's equilibration time. That the equilibration time scales are broadly distributed across samples has been known and identified as a potential problem before [1, 2], and we came to the conclusion that this influence of the free energy landscape is the reason why different samples would equilibrate in different time frames. By using the overlap probability distribution as a proxy for the free energy landscape, we concluded that ideally, individual realizations need to be individually tested to ensure complete equilibration. As a more viable method, we have shown that a conservative application of the Katzgraber-Young equilibration test [31] would be sufficient for further studies.

In Ch. 4 we then proceeded to present our observations made on systems at equilibrium. Of those observations, arguably the most striking concerns a new statistic we introduced. $\Delta(q_0, \kappa)$, as it is called, can distinguish between the many-pairs and the single-pair scenarios by detecting the fraction of samples with δ function behavior in $P_J(q)$ near the origin. An increase of this fraction with increased system size would indicate a trend towards the many-pairs picture, and a decrease would point towards a single-pair scenario. We first tested $\Delta(q_0, \kappa)$ on the fully understood Sherrington-Kirkpatrick model and confirmed that it does indeed behave as expected from the RSB picture and show many-pairs behavior. For the EA model, however, we have seen that it behaved differently, increasing slowly at first and then leveling off, which showed that the mean field solution (and the many-pairs hypothesis) was not able to explain this model's properties. At the end of that chapter we addressed a scaling theory attempting to reconcile our data with the RSB picture [9], and showed that the theory presented in their comment does not explain our EA data.

The results obtained in Ch. 4, when published [73], generated considerable interest, and many studies came out suggesting additional statistics that would show the same effect [45, 46]. We tried to answer these in our final chapter, and we presented results of an additional simulation, the 4D EA model, as well. We finished up with some preliminary studies to motivate further research.

Our main conclusion from this entire study is that the low temperature phase of the EA model does not seem to be mean-field like. We believe this to be a major contribution to the discussion in the field, and some of the observations we have made can be used as a starting point for further exploration. First of those future projects should be to expand on the 4D EA model. Due to equilibration concerns on our largest system size $L = 9$, we could not make a firm statement about the Δ data. It would be interesting to run longer simulations of that system size, as well as obtain data for one or ideally two larger system sizes to see if the observed trends hold, and if Δ indeed starts to decrease with increasing system size. This would be a striking result, since it would mean the onset of a single-pair scenario. Additionally, simulating larger systems for all of our models would be desirable. In the end, our aim is to be able to draw conclusions about the thermodynamic limit, and larger systems would therefore make stronger statements.

Another interesting project to undertake would be to carefully consider all possible boundary conditions (BC), and the effects of changing those on various observables. From our preliminary results in Sec. 5.5 we have seen that the weight of the overlap probability distribution around zero, $I_J(0.2)$, is already susceptible to the changes in BC, and therefore an average over all boundary conditions described there may give a more accurate reading of that quantity. If other statistics are also susceptible, conclusions drawn based on such an average instead of only the periodic BC would carry more weight, in our opinion.

Furthermore, for our equilibrium studies, we have only considered statistics from the spin overlap q. But it has been shown that the link overlap

$$q_l = \frac{1}{N_b} \sum_{\langle i,j \rangle} s_i^\alpha s_j^\alpha s_i^\beta s_j^\beta \tag{6.1}$$

is equally interesting [15]. It would make sense to carry out a similar analysis to what we did with q on the q_l data, to look for other characteristic features there. Finally, we choose our couplings from a Gaussian distribution. An alternative would be to consider the bimodal case, where the couplings can be $\pm J$ with $1/2$ probability.

APPENDIX A

DETAILED BALANCE FOR PARALLEL TEMPERING

In a set of N replicas, the probability of a swap move between two replicas X and X' at inverse temperatures β and β' can be obtained using the detailed balance condition [54]. To do that, we start with the partition function for the whole system:

$$Z_{EXT} = \prod_{i=1}^{N} Z(\beta_i), \tag{A.1}$$

where

$$Z(\beta_i) = \sum_{\{X_i\}} \exp[-\beta_i \mathcal{H}(X_i)]. \tag{A.2}$$

For a given set of β's, $(\beta_1, ..., \beta_N)$, the probability of picking a configuration $X = (X_1, ..., X_N)$ is

$$P(X; \beta_1, ..., \beta_N) = \frac{1}{Z_{EXT}} \exp\left[- \sum_{i=1}^{N} \beta_i \mathcal{H}(X_i) \right]. \tag{A.3}$$

Now we can define a Markov process for this extended system. To do this we first need to define a transition probability matrix $W(X, \beta; X', \beta')$. Here we can use the detailed balance condition for this system, given as

$$\begin{aligned} P(..., X, ..., X', ...; ...\beta, ..., \beta', ...) W(X, \beta; X', \beta') = \\ P(..., X', ..., X, ...; ...\beta, ..., \beta', ...) W(X', \beta; X, \beta'). \end{aligned} \tag{A.4}$$

Inserting Eq. (A.3) to the previous equation we obtain

$$\frac{W(X, \beta; X', \beta')}{W(X', \beta; X, \beta')} = \exp(-\Delta), \tag{A.5}$$

where $\Delta = (\beta' - \beta)(\mathcal{H}(X) - \mathcal{H}(X'))$. The Metropolis solution is

$$W(X,\beta;X',\beta') = \begin{cases} 1, & \Delta < 0. \\ exp(-\Delta), & \Delta > 0. \end{cases} \tag{A.6}$$

Hence we arrive to the swap probability given in Eq. (2.1)

$$p = \min[1, e^{(\beta-\beta')(E-E')}], \tag{A.7}$$

where $E = \mathcal{H}(X)$ and $E' = \mathcal{H}(X')$.

APPENDIX B

HANDLING THE DATA

In this appendix we want to briefly explain our process of generating, analyzing and presenting the data, as well as the computational tools we use.

In the beginning we have our main C++ code with a check-pointing routine that, at a given point of the simulation, stores not only the observables but also the exact coordinates of the random number generator. Therefore, if need arises, we can continue the simulation from that point on any time, as if the process has never been interrupted. This is useful for many reasons. First of all, since our calculations require considerable running times, an accidental interruption is always a possibility. And second, check-pointing gives us the opportunity to do a longer run, or measure other observables once the equilibration is reached, even at a later time, just by initiating the code from the stored file.

Once the main code is run, data files with a lot of information are generated. Since we store many observables and their evolutions in time, these files consist of many sections and can get up to 1GB in size for each individual realization. Considering there are thousands of realizations for each system size of all models we investigate, a careful selection and extraction of the needed portions is necessary before analyzing the data. To facilitate this process, we make sure that every section of the data file is labeled correctly at the beginning, and those label lines always start with the symbol $\#$ for easier identification. Additionally, we start each data file with a basic information section where we state the simulation parameters, such as the system size, duration of the run in terms of time steps, the seed for the random number generator and the temperature set. For the selection and extraction of necessary portions of the data, we use unix bash scripts, and a sample script with various key command examples can be seen below.

```
#!/bin/bash

# Find all files in format S8∗. txt , erase the common beginning and end,
# then write the remaining part of the file names on a . txt file :
for FILE in S8∗. txt ; do echo $FILE | sed −e”s/S8//” −e”s/. txt //”; done > List . txt

#Copy all files in that List into a new folder :
cat List . txt | while read NAME; do cp S8$NAME’.txt’ NewFolder/ ; done

for FILE in S8∗     #Find all files starting with S8,
do
#open all files and find the second number in the top 13 lines
cat $FILE | awk ’{if(NR<14) {print $2}}’
done > NewFile.txt     #Print them out in a new file .

#OR print the numbers into individual files , using the original file name as identifier :
for FILE in S8∗. txt
do
NAME=‘echo $FILE | sed −e”s/S8//” −e”s/.txt //”‘
cat S8_$NAME’_0.txt’ | awk ’{if(NR<14) {print $2}}’ > Numbers_$NAME.txt
done

#Find a particular phrase in all matching files using ’grep’ command,
#and print the numbers following that phrase into a new file :
for file in S8∗. txt
do
cat $file |grep ”# Number of round trips :” | sed −e”s/# Number of round trips :// ”
done > NewFile.txt
```

The final analysis is done using a combination of such scripts, as well as simple C++ codes, depending on what quantities or characteristics we are looking for. Last but not least, we plot our results and do any necessary fitting on those using gnuplot.

BIBLIOGRAPHY

[1] Alder, S., Trebst, S., Hartmann, A. K., and Troyer, M. Dynamics of the Wang Landau algorithm and complexity of rare events for the three-dimensional bimodal Ising spin glass. *J. Stat. Mech.* P07008 (2004).

[2] Alvarez Baños, R., Cruz, A., Fernandez, L. A., Gil-Narvion, J. M., Gordillo-Guerrero, A., Guidetti, M., Maiorano, A., Mantovani, F., Marinari, E., Martin-Mayor, V., Monforte-Garcia, J., Muñoz Sudupe, A., Navarro, D., Parisi, G., Perez-Gaviro, S., Ruiz-Lorenzo, J. J., Schifano, S. F., Seoane, B., Tarancon, A., Tripiccione, R., and Yllanes, D. Nature of the spin-glass phase at experimental length scales. *J. Stat. Mech.* P06026 (2010).

[3] Alvarez Baños, R., Cruz, A., Fernandez, L. A., Gordillo-Guerrero, A., Gil-Narvion, J. M., Guidetti, M., Maiorano, A., Mantovani, F., Marinari, E., Martin-Mayor, V., Monforte-Garcia, J., Muñoz Sudupe, A., Navarro, D., Parisi, G., Perez-Gaviro, S., Ruiz-Lorenzo, J. J., Seoane, B., Schifano, S. F., Tarancon, A., Tripiccione, R., and Yllanes, D. Critical Behavior of Three-Dimensional Disordered Potts Models with Many States. *J. Stat. Mech.* P05002 (2010).

[4] Anderson, P. W. Spin glass i: A scaling law rescued. *Physics Today 41* (1988), 9.

[5] Aspelmeier *et al.*, T. Finite-size corrections in the Sherrington Kirkpatrick model. *J. Phys. A: Math. Theor. 41* (2008), 324008.

[6] Belletti *et al.*, F. Simulating spin systems on IANUS, an FPGA-based computer. *Comp. Phys. Comm. 178* (2008), 208.

[7] Berg, B. Introduction to multicanonical Monte Carlo simulations. *Fields Inst. Commun. 26* (1999), 1. (cond-mat/9909236).

[8] Bhatt, R. N., and Young, A. P. Numerical studies of Ising spin glasses in two, three and four dimensions. *Phys. Rev. B 37* (1988), 5606.

[9] Billoire *et al.*, A. Comment on "Evidence of Non-Mean-Field-Like Low-Temperature Behavior in the Edwards-Anderson Spin-Glass Model". *Phys. Rev. Lett. 110* (2013), 219701.

[10] Binder, K., and Young, A. P. Spin glasses: Experimental facts, theoretical concepts and open questions. *Rev. Mod. Phys. 58* (1986), 801.

[11] Bittner, E., Nußbaumer, A., and Janke, W. Make life simple: Unleash the full power of the parallel tempering algorithm. *Phys. Rev. Lett. 101* (2008), 130603.

[12] Bray, A. J., and Moore, M. A. Scaling theory of the ordered phase of spin glasses. In *Heidelberg Colloquium on Glassy Dynamics and Optimization*, L. Van Hemmen and I. Morgenstern, Eds. Springer, New York, 1986, p. 121.

[13] Cannella, Y., and Mydosh, J. A. Magnetic ordering in gold-iron alloys (susceptibility and thermopower studies). *Phys. Rev. B 6* (1972), 4220.

[14] Ciria, J. C., Parisi, G., and Ritort, F. Four-dimensional Ising spin glass: scaling within the spin-glass phase. *J. Phys. A 26* (1993), 6731.

[15] Drossel, B., Bokil, H., Moore, M. A., and Bray, A. J. The link overlap and finite size effects for the 3d Ising spin glass. *Euro. Phys. J. 13* (2000), 369.

[16] Earl, D. J., and Deem, M. W. Parallel Tempering: Theory, Applications, and New Perspectives. *Phys. Chem. Chem. Phys. 7* (2005), 3910.

[17] Edwards, S. F., and Anderson, P. W. Theory of spin glasses. *J. Phys. F: Met. Phys. 5* (1975), 965.

[18] Fernández, J. F., and Alonso, J. J. Pair correlation function for spin glasses. (arXiv:cond-mat/1207.4008), 2012.

[19] Fisher, D. S., and Huse, D. A. Ordered phase of short-range Ising spin-glasses. *Phys. Rev. Lett. 56* (1986), 1601.

[20] Fisher, D. S., and Huse, D. A. Absence of many states in realistic spin glasses. *J. Phys. A 20* (1987), L1005.

[21] Fisher, D. S., and Huse, D. A. Equilibrium behavior of the spin-glass ordered phase. *Phys. Rev. B 38* (1988), 386.

[22] Fisher, K. H., and Hertz, J. A. *Spin Glasses*. Cambridge University Press, Cambridge, 1991.

[23] Geyer, C. Monte Carlo Maximum Likelihood for Dependent Data. In *23rd Symposium on the Interface* (Fairfax Station, 1991), E. M. Keramidas, Ed., Interface Foundation, p. 156.

[24] Hartmann, A. K. Scaling of stiffness energy for three-dimensional $\pm J$ Ising spin glasses. *Phys. Rev. E 59* (1999), 84.

[25] Hukushima, K., and Nemoto, K. Exchange Monte Carlo method and application to spin glass simulations. *J. Phys. Soc. Jpn. 65* (1996), 1604.

[26] Hukushima *et al.*, K. Exchange Monte Carlo dynamics in the sk model. *J. Phys. Soc. Jpn. 67* (1998), 12.

[27] Huse, D. A., and Fisher, D. S. Pure states in spin glasses. *J. Phys. A 20* (1987), L997.

[28] Janke, W. Multicanonical Monte Carlo simulations. *Physica A 254* (1998), 164.

[29] Katzgraber, H. G. Introduction to Monte Carlo Methods. (arXiv:0905.1629), 2009.

[30] Katzgraber, H. G., Körner, M., and Young, A. P. Universality in three-dimensional Ising spin glasses: A Monte Carlo study. *Phys. Rev. B 73* (2006), 224432.

[31] Katzgraber, H. G., Palassini, M., and Young, A. P. Monte Carlo simulations of spin glasses at low temperatures. *Phys. Rev. B 63* (2001), 184422.

[32] Katzgraber, H. G., Trebst, S., Huse, D. A., and Troyer, M. Feedback-optimized parallel tempering Monte Carlo. *J. Stat. Mech.* P03018 (2006).

[33] Katzgraber, H. G., and Young, A. P. Monte Carlo simulations of spin-glasses at low temperatures: Effects of free boundary conditions. *Phys. Rev. B 65* (2002), 214402.

[34] Katzgraber *et al.*, H. G. Monte Carlo simulations of spin glasses at low temperatures. *Phys. Rev. B 63* (2001), 184422.

[35] Katzgraber *et al.*, H. G. Universality in three-dimensional Ising spin glasses: A Monte Carlo study. *Phys. Rev. B 73* (2006), 224432.

[36] Krauth, W. *Algorithms and Computations*. Oxford University Press, New York, 2006.

[37] Krzakala, F., and Martin, O. C. Spin and link overlaps in 3-dimensional spin glasses. *Phys. Rev. Lett. 85* (2000), 3013.

[38] Machta, J. Strengths and weaknesses of parallel tempering. *Phys. Rev. E 80* (2009), 056706.

[39] Madras, N., and Sokal, A. D. The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk. *J. Stat. Phys. 50* (1988), 109.

[40] Marinari, E., and Parisi, G. Simulated tempering: A new Monte Carlo scheme. *Europhys. Lett. 19* (1992), 451.

[41] Marinari, E., Parisi, G., Ricci-Tersenghi, F., and Riuz-Lorenzo, J. J. Small window overlaps are effective probes of replica symmetry breaking in three-dimensional spin glasses. *J. Phys. A: Math. Gen. 31* (1998).

[42] McMillan, W. L. Domain-wall renormalization-group study of the two-dimensional random Ising model. *Phys. Rev. B 29* (1984), 4026.

[43] Mézard, M., Parisi, G., and Virasoro, M. A. *Spin Glass Theory and Beyond*. World Scientific, Singapore, 1987.

[44] Middleton, A. A. Numerical investigation of the thermodynamic limit for ground states in models with quenched disorder. *Phys. Rev. Lett. 83* (1999), 1672.

[45] Middleton, A. A. Extracting thermodynamic behavior of spin glasses from the overlap function. *Phys. Rev. B 87* (2013), 220201.

[46] Monthus, C., and Garel, T. Typical versus averaged overlap distribution in Spin-Glasses : Evidence for the droplet scaling theory. (arXiv:cond-mat/1306.0423).

[47] Moore, M. A. Corrections to scaling in the droplet picture of spin glasses. (cond-mat/0203469).

[48] Mydosh, J. A. *Spin glasses: an experimental introduction*. Taylor & Francis London, Washington, DC, 1993.

[49] Newman, C. M., and Stein, D. L. Multiple states and thermodynamic limits in short-ranged Ising spin-glass models. *Phys. Rev. B 46* (1992), 973.

[50] Newman, C. M., and Stein, D. L. Non-mean-field behavior of realistic spin glasses. *Phys. Rev. Lett. 76* (1996), 515.

[51] Newman, C. M., and Stein, D. L. Simplicity of state and overlap structure in finite-volume realistic spin glasses. *Phys. Rev. E 57* (1998), 1356.

[52] Newman, C. M., and Stein, D. L. TOPICAL REVIEW: Ordering and broken symmetry in short-ranged spin glasses. *J. Phys.: Condensed Matter 15* (2003), 1319.

[53] Newman, C. M., and Stein, D. L. Short-range spin glasses: Results and speculations. In *Lecture Notes in Mathematics 1900*. Springer-Verlag, Berlin, 2007, p. 159. (cond-mat/0503345).

[54] Newman, M. E. J., and Barkema, G. T. *Monte Carlo Methods in Statistical Physics*. Oxford University Press Inc., New York, USA, 1999.

[55] Palassini, M., and Young, A. P. Triviality of the ground state structure in Ising spin glasses. *Phys. Rev. Lett. 83* (1999), 5126.

[56] Parisi, G. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett. 43* (1979), 1754.

[57] Parisi, G. A sequence of approximated solutions to the S-K model for spin glasses. *J. Phys. A 13* (1980), L115.

[58] Parisi, G. The order parameter for spin glasses: a function on the interval 0–1. *J. Phys. A 13* (1980), 1101.

[59] Parisi, G. Order parameter for spin-glasses. *Phys. Rev. Lett. 50* (1983), 1946.

[60] Parisi, G., Ricci-Tersenghi, F., and Ruiz-Lorenzo, J. J. Equilibrium and off-equilibrium simulations of the 4d Gaussian spin glass. *J. Phys. A 29* (1996), 7943.

[61] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. *Numerical Recipes in C*. Cambridge University Press, Cambridge, 1995.

[62] Reger, J. D., Bhatt, R. N., and Young, A. P. A Monte Carlo study of the order parameter distribution in the four-dimensional Ising spin glass. *Phys. Rev. Lett. 64* (1990), 1859.

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

[64] Singh, R. R. P., and Chakravarty, S. Critical behavior of an Ising spin-glass. *Phys. Rev. Lett. 57* (1986), 245.

[65] Sokal, A. D. *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms (Lecture notes, Cours de Troisieme Cycle de la Physique en Suisse Romande).* Lausanne, 1989.

[66] Stein, D.L., and Newman, C.M. *Spin Glasses and Complexity*. Primers in Complex Systems. Princeton University Press, 2013.

[67] Swendsen, R. H., and Wang, J. Replica Monte Carlo simulation of spin-glasses. *Phys. Rev. Lett. 57* (1986), 2607.

[68] Tasaki, H. On the upper critical dimensions of random spin systems. *J. Stat. Phys. 54* (1989), 163.

[69] Torrie, G. M., and Valleau, J. P. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys. 23* (1977), 187.

[70] Trebst, S., Hansmann, U. H. E., and Troyer, M. Optimized parallel tempering simulations of proteins. *J. Chem. Phys. 124* (2006), 174903.

[71] Wang, F., and Landau, D. P. An efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett. 86* (2001), 2050.

[72] Yucesoy, B., Katzgraber, H. G., and Machta, J. Yucesoy, katzgraber, and machta reply:. *Phys. Rev. Lett. 110* (2013), 219702.

[73] Yucesoy, B., Katzgraber, Helmut G., and Machta, J. Evidence of Non-Mean-Field-Like Low-Temperature Behavior in the Edwards-Anderson Spin-Glass Model. *Phys. Rev. Lett. 109* (2012), 177204.

[74] Yucesoy, B., Machta, J., and Katzgraber, Helmut G. Correlations between the dynamics of parallel tempering and the free-energy landscape in spin glasses. *Phys. Rev. E 87* (2013), 012104.