Nature of the spin-glass phase at experimental length scales

J. Stat. Mech. (2010) P06026

(http://iopscience.iop.org/1742-5468/2010/06/P06026)

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Nature of the spin-glass phase at experimental length scales

R Alvarez Baños[1,2], A Cruz[1,2], L A Fernandez[1,3], J M Gil-Narvion[1], A Gordillo-Guerrero[1,4], M Guidetti[5], A Maiorano[1,6], F Mantovani[5], E Marinari[6], V Martin-Mayor[1,3], J Monforte-Garcia[1,2], A Muñoz Sudupe[3], D Navarro[7], G Parisi[6], S Perez-Gaviro[1,6], J J Ruiz-Lorenzo[1,8], S F Schifano[5], B Seoane[1,3], A Tarancon[1,2], R Tripiccione[5] and D Yllanes[1,3]

[1] Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Zaragoza, Spain
[2] Departamento de Física Teórica, Universidad de Zaragoza, 50009 Zaragoza, Spain
[3] Departamento de Física Teórica I, Universidad Complutense, 28040 Madrid, Spain
[4] Departamento de Ingeniería Eléctrica, Electrónica y Automática, Universidad de Extremadura, Avda. de la Universidad s/n, 10071 Cáceres, Spain
[5] Dipartimento di Fisica Università di Ferrara and INFN—Sezione di Ferrara, Ferrara, Italy
[6] Dipartimento di Fisica, SMC of INFM-CNR and INFN, Università di Roma La Sapienza, 00185 Roma, Italy
[7] Departamento de Ingeniería, Electrónica y Comunicaciones and Instituto de Investigación en Ingeniería de Aragón (I3A), Universidad de Zaragoza, 50018 Zaragoza, Spain
[8] Departamento de Física, Universidad de Extremadura, 06071 Badajoz, Spain
E-mail: yllanes@lattice.fis.ucm.es

Received 23 March 2010
Accepted 2 June 2010
Published 28 June 2010

Online at stacks.iop.org/JSTAT/2010/P06026
doi:10.1088/1742-5468/2010/06/P06026

Abstract. We present a massive equilibrium simulation of the three-dimensional Ising spin glass at low temperatures. The Janus special-purpose computer has allowed us to equilibrate, using parallel tempering, $L = 32$ lattices down to $T \approx 0.64T_c$. We demonstrate the relevance of equilibrium finite size simulations to understanding experimental non-equilibrium spin glasses in the

J. Stat. Mech. (2010) P06026

thermodynamical limit by establishing a time-length dictionary. We conclude that non-equilibrium experiments performed on a timescale of 1 h can be matched with equilibrium results on $L \approx 110$ lattices. A detailed investigation of the probability distribution functions of the spin and link overlap, as well as of their correlation functions, shows that Replica Symmetry Breaking is the appropriate theoretical framework for the physically relevant length scales. Besides, we improve over existing methodologies in ensuring equilibration in parallel tempering simulations.

Keywords: classical Monte Carlo simulations, disordered systems (theory), spin glasses (theory)

Contents

---

1. Introduction

Spin glasses (SG) are disordered magnetic alloys that are generally regarded as particularly convenient model systems for the study of glassy behaviour [1, 2]. Indeed, ideas originating in the SG context have been fruitful in the study of structural glasses, optimization in computer science, quantum information, econophysics, etc.

A distinctive feature of SG is that, below their glass temperature, they remain out of equilibrium even if they are left to relax under constant experimental conditions for days or weeks. In spite of this, the *equilibrium* properties of their low temperature phase are believed to control their non-equilibrium behaviour. Indeed, both theory [3, 4] and experiment [5] agree in that the sluggish dynamics is due to a *thermodynamic* phase transition at a critical temperature, T_{c}, that separates the paramagnetic phase from a low temperature one where the spins freeze according to extremely complex, essentially unpredictable, ordering patterns. Furthermore, it has been now established that an accurate knowledge of the thermodynamic equilibrium properties would allow us to predict in detail many relevant features of their non-equilibrium relaxation [6, 7].

There is an already 30-year-old theoretical controversy regarding the defining properties of the SG phase. On the one hand, the replica symmetry breaking (RSB) theory that stems from Parisi's solution of the SG in the mean-field approximation [8, 9]. A system well described by the RSB is in a critical state for all $T < T_{\mathrm{c}}$, where the surfaces of the magnetic domains are space filling. On the other hand, the droplet theory [10]–[13] views the SG phase as a disguised ferromagnet. It provides the solution of SG models as computed in the Migdal–Kadanoff approximation [14]. We refer the reader to section 2 for the detailed predictions of the RSB and droplet theories for the different physical observables in the SG phase. The predictions of the somewhat intermediate TNT theory [15, 16] are discussed also in section 2.

Numerical simulations are the main tool that theoretical physicists have to make progress in the understanding of the SG phase in $D = 3$ systems. Basically without exceptions, numerical work in $D = 3$ is best described by RSB theory (see [9] for a review, [17]–[19] for recent work and [15, 16, 20] for some somewhat dissenting views). Yet, numerical investigations have also received severe criticism. It has been claimed that basically all simulations doable to date are contaminated by critical effects [21]. One would need to simulate still larger systems at still lower temperatures, in order to observe the asymptotic behaviour corresponding to large enough systems.

Here we present the results of a large-scale simulation performed on Janus [22, 23], a special-purpose computer designed for the simulation of SG. For this particular task, Janus outperforms standard computers by several orders of magnitude. We have devoted (the equivalent of) 200 days of the full Janus computer to an equilibrium, parallel tempering simulation of the Ising SG in $D = 3$. We have been able to thermalize lattices of size $L = 32$ down to temperatures $T \approx 0.64 T_{\mathrm{c}}$. This is not only a world record, but provides as well an unprecedented glimpse at the low temperature SG phase.

Our main objectives here have been (see section 2 for definitions):

- To perform a precision comparison of equilibrium and non-equilibrium spatial correlation functions. It turns out that a time-length dictionary exists, which relates with amazing accuracy our previous results at finite times [24, 25] (on non-equilibrium infinite systems) with equilibrium *finite* lattice *sizes*. The unavoidable conclusion is that experimental SG are in the dynamical non-equilibrium regimes that correspond to equilibrium results on lattices $L \sim 110$. There is no doubt that at these length scales, the appropriate effective theory is RSB, irrespectively of which of the competing theories is correct for much larger L.

- To perform a study of the probability density function (pdf) of the spin overlap, and to extrapolate important quantities to the thermodynamic limit. So doing, we will gather important information about the correlation length in the spin-glass phase.
- To provide a detailed study of the link overlap.
- Last, but not least, to obtain a large set of configurations, fireproof thermalized, which will serve as a starting point for more sophisticated studies (such as investigation of ultrametricity, or temperature chaos). In particular, a detailed study of the spatial correlation functions will appear elsewhere [26].

The layout of the rest of this paper is as follows. In section 2 we briefly recall the definition of the Edwards–Anderson model. In particular, in section 2.2 we describe the observables considered and discuss the scaling behaviour predicted for them by the different theoretical scenarios. In section 3, we describe our simulations and address the crucial problem of ensuring thermal equilibrium. We have found it most useful to study the random walk in temperature space performed in our parallel tempering simulations (section 3.3). In particular, our thermalization checks significantly expand the methodology introduced in [27]. At this point, we are ready to study in section 4 the pdf of the spin overlap. In particular, in section 4.3 we determine through finite size effects a correlation length in the spin-glass phase. We focus on the spatial correlation functions in section 5, finding (section 6) crystal-clear indications of the relevance of our *equilibrium* investigations to the *non-equilibrium* experimental work. The properties of the link overlap are addressed in section 7. Our conclusions are presented in section 8. Technical details are provided in two appendices.

2. Model, observables, theoretical expectations

We divide this section into five parts. In section 2.1 we describe our model. The spin overlap and related quantities are defined in section 2.2. We discuss spatial correlation functions in section 2.3. Their non-equilibrium counterparts are recalled in section 2.4. We address the link overlap in section 2.5. Even though most of this section consists of results and definitions well known in the spin-glass community, we consider it convenient as a quick reference. We also introduce some specific (and sometimes new or seldom used) physical quantities for this paper.

2.1. The model

We consider the $D = 3$ Edwards–Anderson model [28, 29]. Our dynamical variables are Ising spins $s_{\boldsymbol{x}} = \pm 1$, which are placed on the nodes, $\boldsymbol{x}$, of a cubic lattice of linear size L, containing $V = L^3$ sites, and with periodic boundary conditions. Their interaction is restricted to lattice nearest neighbours and is given by the Hamiltonian:

$$\mathcal{H} = -\sum_{\langle \boldsymbol{x}\boldsymbol{y} \rangle} J_{\boldsymbol{x},\boldsymbol{y}}\, s_{\boldsymbol{x}}\, s_{\boldsymbol{y}}. \tag{1}$$

Note that the couplings $J_{\boldsymbol{x},\boldsymbol{y}}$ in the Hamiltonian are themselves stochastic variables: they take the values ±1 with 50% probability. The coupling constants attached to different lattice links are statistically independent. The physical motivation for working with a random Hamiltonian is modelling the effects of impurities in a magnetic alloy.

We shall consider the *quenched* approximation: in the timescale relevant to the spin dynamics, the impurities can be regarded as static. Hence, we will not allow for any back-reaction of the spins over the coupling constants. A given realization of the $\{J_{\boldsymbol{x},\boldsymbol{y}}\}$ (a sample, from now on), will be fixed from the start and considered non-dynamical [1].

A random Hamiltonian implies a double averaging procedure. For any observable O (an arbitrary function of the spins and the coupling constants), we shall *first* compute the thermal average $\langle O\rangle$ using the Boltzmann weight at temperature T for the Hamiltonian (1). The average over the coupling constants distribution, $\overline{\langle O\rangle}$, is only taken afterwards. We will refer sometimes to the second averaging, $\overline{(\cdots)}$, as disorder average.

The reader will notice that the disorder average induces a non-dynamical gauge symmetry [30]. Let us choose a random sign per site $\epsilon_{\boldsymbol{x}} = \pm 1$. Hence, the energy (1) is invariant under the transformation

$$s_{\boldsymbol{x}} \longrightarrow \epsilon_{\boldsymbol{x}} s_{\boldsymbol{x}}, \qquad J_{\boldsymbol{x},\boldsymbol{y}} \longrightarrow \epsilon_{\boldsymbol{x}}\epsilon_{\boldsymbol{y}} J_{\boldsymbol{x},\boldsymbol{y}}. \tag{2}$$

Since the gauge-transformed couplings $\epsilon_{\boldsymbol{x}}\epsilon_{\boldsymbol{y}} J_{\boldsymbol{x},\boldsymbol{y}}$ are just as probable as the original ones, the quenched mean value of $\overline{\langle O(\{s_{\boldsymbol{x}}\})\rangle}$ is identical to that of its gauge average $\sum_{\{\epsilon_{\boldsymbol{x}}=\pm1\}} \overline{\langle O(\{\epsilon_{\boldsymbol{x}} s_{\boldsymbol{x}}\})\rangle}/2^{L^D}$, which typically is an uninteresting constant value. We show in section 2.2 how to overcome this problem.

We remark as well that the Hamiltonian (1) also has a global $\mathbf{Z}_2$ symmetry (if all spins are simultaneously reversed $s_{\boldsymbol{x}} \to -s_{\boldsymbol{x}}$ the energy is unchanged), corresponding to time-reversal symmetry. This symmetry gets spontaneously broken in three dimensions upon lowering the temperature at the SG transition at $T_{\mathrm{c}} = 1.109(10)$ [31, 32].

2.2. The spin overlap

We need observables that remain invariant under the transformation (2). The Hamiltonian (1) provides, of course, a first example. To make further progress we consider *real* replicas $\{s_{\boldsymbol{x}}^{(1)}\}, \{s_{\boldsymbol{x}}^{(2)}\}$, copies of the system that evolve under the same set of couplings $\{J_{\boldsymbol{x},\boldsymbol{y}}\}$ but are otherwise statistically uncorrelated[9].

Using them we form the *overlap field*:

$$q_{\boldsymbol{x}} = s_{\boldsymbol{x}}^{(1)} s_{\boldsymbol{x}}^{(2)}, \tag{3}$$

which is obviously invariant under (2).

The Edwards–Anderson order parameter, the *spin overlap*, is the spatial average of the overlap field:

$$q = \frac{1}{V}\sum_{\boldsymbol{x}} q_{\boldsymbol{x}}\,. \tag{4}$$

---

[9] For the thermal average of any observable depending on a *single* spin configuration, $O(\{s_{\boldsymbol{x}}^{(1)}\})$, we have $\langle O(\{s_{\boldsymbol{x}}^{(1)}\})\rangle^2 = \langle O(\{s_{\boldsymbol{x}}^{(1)}\})\, O(\{s_{\boldsymbol{x}}^{(2)}\})\rangle$.

In particular, it yields the (non-connected) spin-glass susceptibility

$$\chi_{\mathrm{NC}}(T) = V\overline{\langle q^2 \rangle}, \tag{5}$$

that diverges at T_{c} with the critical exponent γ. For all $T < T_{\mathrm{c}}$, one expects $\chi_{\mathrm{NC}} = \mathcal{O}(V)$. We shall also consider the Binder ratio

$$B(T) = \frac{\overline{\langle q^4 \rangle}}{\overline{\langle q^2 \rangle}^2}. \tag{6}$$

In particular, for all $T > T_c$, the fluctuations of q are expected to be Gaussian in the large-L limit, hence $\lim_{L\to\infty} B = 3$, $(T > T_{\mathrm{c}})$. Its behaviour in the low temperature phase is controversial. For a *disguised ferromagnet* picture one expects B to approach 1 in the limit of large lattices. On the other hand, for an RSB system one expects $1 < B < 3$ in the SG phase $(T < T_{\mathrm{c}})$. We recall also that one may consider as well the overlap computed in small boxes, in order to avoid the effect of the interphases (physical results are equivalent to those obtained with the standard overlap [33]).

A great deal of attention will be devoted to the probability density function (pdf) of the overlap

$$\tilde{P}(q) = \overline{\left\langle \delta\left(q - \frac{1}{V}\sum_{\boldsymbol{x}} q_{\boldsymbol{x}}\right)\right\rangle}. \tag{7}$$

Note that, in a finite system, the pdf is not smooth, but composed of $N+1$ Dirac deltas at $q = -1, -((N-2)/N), \ldots, ((N-2)/N), 1$. Here, we have solved this problem by a convolution of the comb-like pdf (7) with a Gaussian of width $1/\sqrt{V}$, $\mathcal{G}_V(x) = \sqrt{V/2\pi}\,\exp[-Vx^2/2]$:

$$P(q=c) = \int_{-\infty}^{\infty} \mathrm{d}q'\, \tilde{P}(q')\, \mathcal{G}_V(c-q') = \overline{\left\langle \mathcal{G}_V\left(c - \frac{1}{V}\sum_{\boldsymbol{x}} q_{\boldsymbol{x}}\right)\right\rangle}. \tag{8}$$

In this way, we basically add the contribution of $\mathcal{O}(\sqrt{V})$ microscopic values of q, belonging to an interval of width $\sim 1/\sqrt{V}$ [34]. Note, however, that equations (5) and (6) are computed out of moments of $\tilde{P}(q)$, rather than of $P(q)$.

The Edwards–Anderson order parameter q_{EA} vanishes for all $T \geq T_{\mathrm{c}}$. Below T_{c}, in a droplet system, $P(q)$ collapses in the large-L limit in a pair of Dirac delta functions of equal weight, centred at $q = \pm q_{\mathrm{EA}}$. In an RSB system, $P(q)$ contains as well a pair of delta functions at q_{EA}, but it also has a continuous piece, non-vanishing for every q such that $-q_{\mathrm{EA}} < q < q_{\mathrm{EA}}$. This is the origin of the differences in the predictions that both theories make for B in the low temperature phase.

We will find it useful to consider as well *conditional* expectation values at fixed q. Let O be an arbitrary function of the spins. We define its conditional expectation

$$\mathrm{E}(O|q=c) = \overline{\left\langle O\mathcal{G}_V\left(c - \frac{1}{V}\sum_{\boldsymbol{x}} q_{\boldsymbol{x}}\right)\right\rangle} \Big/ \overline{\left\langle \mathcal{G}_V\left(c - \frac{1}{V}\sum_{\boldsymbol{x}} q_{\boldsymbol{x}}\right)\right\rangle}. \tag{9}$$

Of course, one may easily recover standard expectation values from $\mathrm{E}(O|q)$:

$$\overline{\langle O \rangle} = \int_{-\infty}^{\infty} \mathrm{d}q\, P(q)\, \mathrm{E}(O|q). \tag{10}$$

Strictly speaking, the integration limits should be $\pm\infty$. However, truncating the integral to $-1 < q < 1$, the error is exponentially small in $L^{D/2}$ (yet, for $L = 8$ and 12 we had to extend the limits beyond ± 1).

We can also define the conditional variances as

$$\mathrm{Var}(O|q = c) = \mathrm{E}(O^2|q = c) - \mathrm{E}(O|q = c)^2, \tag{11}$$

where we have the identity

$$\overline{\langle O^2\rangle} - \overline{\langle O\rangle}^2 = \int_{-\infty}^{\infty} \mathrm{d}q\, P(q)[\mathrm{Var}(O|q) + (\mathrm{E}(O|q) - \overline{\langle O\rangle})^2]. \tag{12}$$

2.3. Spatial correlation functions

The overlap correlation function is

$$C_4(\boldsymbol{r}) = \frac{1}{V}\sum_{\boldsymbol{x}} \overline{\langle q_{\boldsymbol{x}}\, q_{\boldsymbol{x}+\boldsymbol{r}}\rangle}. \tag{13}$$

$C_4(\boldsymbol{r})$ decays to zero for large $\boldsymbol{r}$ only for $T > T_\mathrm{c}$. Thus we have considered as well conditional correlation functions, recall equation (9):

$$C_4(\boldsymbol{r}|q) = \mathrm{E}\left(\frac{1}{V}\sum_{\boldsymbol{x}} q_{\boldsymbol{x}} q_{\boldsymbol{x}+\boldsymbol{r}}\,\middle|\, q\right). \tag{14}$$

Equation (10) allows us to recover $C_4(\boldsymbol{r})$ from $C_4(\boldsymbol{r}|q)$.

The two main theoretical pictures for the SG phase, the droplet and RSB pictures, dramatically differ on their predictions for $C_4(\boldsymbol{r}|q)$. Let us discuss them in detail:

- In the RSB picture, the *connected* correlation functions tend to zero at large $\boldsymbol{r}$. For all $q \in [-q_\mathrm{EA}, q_\mathrm{EA}]$ we expect the asymptotic behaviour

  $$C_4(\boldsymbol{r}|q) \sim q^2 + \frac{A_q}{r^{\theta(q)}} + \cdots, \tag{15}$$

  where the dots stand for scaling corrections, subleading in the limit of large r. The exponent $\theta(q)$ in equation (15) has been computed for D larger than the upper critical dimension $D_\mathrm{u} = 6$: [35, 36]

  $$\theta(q = 0) = D - 4, \tag{16}$$

  $$\theta(0 < |q| < q_\mathrm{EA}) = D - 3, \tag{17}$$

  $$\theta(|q| = q_\mathrm{EA}) = D - 2. \tag{18}$$

  These mean-field results for $\theta(q)$ become inconsistent for $D < 4$ (the correlations should *decrease* for large r, implying $\theta(q) > 0$, recall equation (15)). An expansion in $\epsilon = 6 - D$ suggests that $\theta(q)$ will renormalize [37]. Note as well that, at least for large D, $\theta(q)$ is discontinuous at $q = 0$. However, we remark that there are no compelling theoretical arguments supporting the discontinuity of $\theta(q)$ in $D = 3$. Indeed, recent numerical studies found no evidence for it [19, 26]. We finally recall a non-equilibrium computation [25] yielding in $D = 3$:[10]

$$\theta(q = 0) = 0.38(2). \tag{19}$$

- Quite the opposite to the RSB case, in a system well described by a droplet model and for $|q| < q_{\mathrm{EA}}$, $C_4(\boldsymbol{r}|q)$ does not tend to q^2 for large r (we are referring, of course, to the regime $1 \ll r \ll L$). In fact, spin configurations with $|q| < q_{\mathrm{EA}}$ are spatially heterogeneous mixtures of the two pure phases. One should find *bubbles* or slabs of linear size $\sim L$ of one of the two phases, say $q = +q_{\mathrm{EA}}$, surrounded by a matrix of the complementary state (see e.g. [38, 39]). It follows that

  $$C_4(\boldsymbol{r}|q) = q_{\mathrm{EA}}^2 f_{\boldsymbol{r}/r}(r/L), \qquad \text{if } |q| < q_{\mathrm{EA}} \qquad \text{and} \qquad 1 \ll r \ll L, \tag{20}$$

  ($f_{\boldsymbol{r}/r}(x)$ is a direction-dependent scaling function with $f_{\boldsymbol{r}/r}(0) = 1$). Indeed, the probability that two spins at fixed distance r belong to domains of opposite orientation is proportional to r/L in the large-L limit. On the other hand, precisely at $|q| = q_{\mathrm{EA}}$, but only there, droplet theory predicts that the connected correlation function vanishes for asymptotically large r. The same behaviour of equation (15) was predicted [11]. The exponent $\theta(q_{\mathrm{EA}})$ is identical to the scaling exponent of the coupling strength, denoted as θ or y in the literature, and has a value of $\theta(q_{\mathrm{EA}}) \sim 0.2$ [11].

2.4. Non-equilibrium correlation functions

Let us recall that non-equilibrium counterparts exist of q and $C_4(\boldsymbol{r}|q)$. We shall not be computing them here, but we *will* compare previous computations with our equilibrium results. Hence, we briefly recall the definitions [25]. One considers pairs of times t_{w} and $t + t_{\mathrm{w}}$, with $t, t_{\mathrm{w}} > 0$, after a sudden quench from a fully disordered state to the working temperature T. The analogue of the spin overlap is

$$C(t, t_{\mathrm{w}}) = \frac{1}{V} \sum_{\boldsymbol{x}} \overline{\langle s_{\boldsymbol{x}}(t_{\mathrm{w}}) s_{\boldsymbol{x}}(t + t_{\mathrm{w}}) \rangle}. \tag{21}$$

The non-equilibrium spatial correlation function is

$$C_{2+2}(\boldsymbol{r}; t, t_{\mathrm{w}}) = \frac{1}{V} \sum_{\boldsymbol{x}} \overline{\langle s_{\boldsymbol{x}}(t_{\mathrm{w}}) s_{\boldsymbol{x}}(t + t_{\mathrm{w}}) s_{\boldsymbol{x}+\boldsymbol{r}}(t_{\mathrm{w}}) s_{\boldsymbol{x}+\boldsymbol{r}}(t + t_{\mathrm{w}}) \rangle}. \tag{22}$$

At fixed t_{w}, $C(t, t_{\mathrm{w}})$ monotonically decreases from $C = 1$ at $t = 0$, to $C = 0$ at $t \to \infty$. Hence, one may consider C, rather than t, as an independent variable. We will compare the non-equilibrium $C_{2+2}(\boldsymbol{r}; t, t_{\mathrm{w}})$, computed in very large lattices [24, 25], with our equilibrium results for $C_4(\boldsymbol{r}|q = C(t, t_{\mathrm{w}}))$. To do so, we shall need to relate the finite *time* t_{w} (on very large lattices) with the finite *size* L. As we shall see in section 6, the correspondence between the non-equilibrium and the equilibrium correlation functions is amazingly accurate.

---

[10] We may mention as well three conjectures: $\theta(0) = (D - 2 + \eta)/2$ [37] (that from the results in [32], yields $\theta(0) = 0.313(5)$), $\theta(0) = 1/\hat{\nu}$ ($\hat{\nu}$ is the exponent that rules finite size effects at q_{EA}) and $\theta(0) + 1/\hat{\nu} = \theta(q_{\mathrm{EA}})$. There is also an exact scaling relation $\theta(q_{\mathrm{EA}}) = 2/\hat{\nu}$ [26].

2.5. The link overlap

The link overlap is defined as[11]

$$Q_{\mathrm{link}} = \frac{1}{DV} \sum_{\|\boldsymbol{x}-\boldsymbol{y}\|=1} q_{\boldsymbol{x}} q_{\boldsymbol{y}}. \tag{23}$$

It is a more sensitive quantity than the spin overlap to the differences between a system described by droplet theory or an RSB system [40]. Since it is invariant under time-reversal symmetry (the global reversal of every spin in either of our two real replicas $s_{\boldsymbol{x}}^{(i)} \longrightarrow -s_{\boldsymbol{x}}^{(i)}$) its expectation value is non-vanishing, even in a finite system at high temperatures. Its pdf can be defined as we did with the spin overlap, recall equations (7), (8). In fact, it has been proposed that the link overlap (rather than the spin overlap) should be considered as the fundamental quantity to describe the spin-glass phase below the upper critical dimension [41, 17]. There are both physical and mathematical reasons for this:

- On the physical side, Q_{link} provides an estimate of the volume of the domains' surfaces. Indeed, consider two configurations of the overlap field (3) differing only in that a *domain* of size $\sim L$ has flipped. This will result in a large change of the spin overlap, q. Yet, the only changing contribution to Q_{link} is that of the lattice links crossed by the domain's surface. In a droplet theory, where the surface-to-volume ratio of the domains vanishes in the large-L limit, one does not expect any q variation of the conditional expectation $\mathrm{E}(Q_{\mathrm{link}}|q)$, not even in the $|q| < q_{\mathrm{EA}}$ region. Hence, the pdf for Q_{link} is expected to collapse to a single-valued delta function in the large-L limit. The intermediate TNT picture coincides with the droplet theory in this respect. For an RSB system, the domains' surfaces are space filling. Hence, when q suffers a variation of order one, the variation of Q_{link} will be of order one, too. Accordingly, a non-trivial pdf is expected for Q_{link}, in the limit of large systems.
- On the mathematical side, theorems have been proven for the link overlap [42]–[44], valid for three-dimensional systems, which are the exact correlate of mean-field results for the spin overlap[12]. Specifically, the replica equivalence property holds for the link overlap in three-dimensional systems. Replica equivalence [45, 46] is a property of the Parisi matrix which yields an infinite hierarchy of identities relating linear combinations of moments of Q_{link} in the large-L limit. A specific example that we shall be using here is

$$\lim_{L\to\infty} \overline{\langle Q_{\mathrm{link}}\rangle^2} = \lim_{L\to\infty} \left[\tfrac{2}{3}\,\overline{\langle Q_{\mathrm{link}}\rangle}^2 + \tfrac{1}{3}\,\overline{\langle Q_{\mathrm{link}}^2\rangle}\right], \tag{24}$$

(at finite L, the equality is not expected to hold). This is just a particular case of the family of identities valid for all $k, s = 0, 1, 2, \ldots$

$$\lim_{L\to\infty} \overline{\langle Q_{\mathrm{link}}^k\rangle\langle Q_{\mathrm{link}}^s\rangle} = \lim_{L\to\infty} \left[\tfrac{2}{3}\,\overline{\langle Q_{\mathrm{link}}^k\rangle}\;\overline{\langle Q_{\mathrm{link}}^s\rangle} + \tfrac{1}{3}\,\overline{\langle Q_{\mathrm{link}}^{k+s}\rangle}\right], \tag{25}$$

(replica equivalence implies infinitely many relations such as this). It is amusing that the mathematical proof for the three-dimensional theorem does *not* use Parisi

---

[11] Clearly, $\overline{\langle Q_{\mathrm{link}}\rangle} = C_4(1,0,0)$.

[12] The mathematical proof known so far is valid only for Gaussian-distributed couplings in equation (1). However, physical intuition strongly suggests that the theorems are valid in more general cases such as our bimodal couplings.

matrices, relying instead on stochastic stability. Let us stress that ultrametricity implies replica equivalence, but the converse statement (i.e. replica equivalence implies ultrametricity) does not hold, in general[13].

The distinction between *spin* overlap and *link* overlap seems somewhat artificial from the point of view of the mean-field approximation. In fact, in the Sherrington–Kirkpatrick model one easily shows that $Q_{\mathrm{link}} = q^2$. For finite-connectivity mean-field models, non-equilibrium numerical computations yield $Q_{\mathrm{link}} = aq^2 + b$ [48] (a and b are numerical constants). In $D = 3$ there are also clear indications that fixing the spin overlap fixes as well the link overlap: the conditional variance $\mathrm{Var}(Q_{\mathrm{link}}|q)$, equation (11), tends to zero for large lattices, see [17] and figure 16, below. Furthermore, in a TNT or droplet system, the derivative $\mathrm{d}E(Q_{\mathrm{link}}|q)/\mathrm{d}q^2$ should vanish in the large-L limit for all $|q| < q_{\mathrm{EA}}$ (since there is a single valid value for Q_{link}, there can be no q^2 dependency left). Numerical simulations, both in equilibrium [17, 18] and out of equilibrium [24, 49], find so far a non-vanishing derivative that nevertheless decreases for larger L. The extrapolation to $L = \infty$ is still an open issue, see section 7.1.

We wish to emphasize that Q_{link} unveils that the spin-glass phase is a critical state where minimal perturbations can produce enormous changes. In fact, let us couple two otherwise independent copies of the system through Q_{link},

$$\mathcal{H} = -\sum_{\langle \boldsymbol{x}\boldsymbol{y}\rangle} J_{\boldsymbol{x},\boldsymbol{y}}\,(s^{(1)}_{\boldsymbol{x}}\, s^{(1)}_{\boldsymbol{y}} + s^{(2)}_{\boldsymbol{x}}\, s^{(2)}_{\boldsymbol{y}}) - T\epsilon V Q_{\mathrm{link}}. \tag{26}$$

In a system described by droplet theory, one expects the link susceptibility

$$\chi_{\mathrm{link}} \equiv \left.\frac{\partial \overline{\langle Q_{\mathrm{link}}\rangle}}{\partial \epsilon}\right|_{\epsilon=0} = V\left[\overline{\langle Q^2_{\mathrm{link}}\rangle - \langle Q_{\mathrm{link}}\rangle^2}\right], \tag{27}$$

to remain finite in the large-L limit, for all $T < T_{\mathrm{c}}$ (precisely at T_{c}, a critical divergence might arise). Hence $\overline{\langle Q_{\mathrm{link}}\rangle}_{\epsilon} = \overline{\langle Q_{\mathrm{link}}\rangle}_{\epsilon=0} + \epsilon\chi_{\mathrm{link}} + \ldots$ in a droplet or TNT system.

On the other hand, in the mean-field approximation, one finds for RSB systems a discontinuity with ϵ [50]:

$$\overline{\langle Q_{\mathrm{link}}\rangle}_{\epsilon>0} = \mathrm{E}(Q_{\mathrm{link}}|q = q_{\mathrm{EA}}) + a_{+}\sqrt{\epsilon} + \cdots, \tag{28}$$

$$\overline{\langle Q_{\mathrm{link}}\rangle}_{\epsilon<0} = \mathrm{E}(Q_{\mathrm{link}}|q = 0) - a_{-}\sqrt{-\epsilon} + \cdots. \tag{29}$$

Actually, the mean-field computation was carried out for the *spin* overlap, yet, in mean-field models, Q_{link} is essentially q^2, hence we can borrow their result. We should emphasize that the situation is even more critical than for standard first-order phase transitions: $\chi_{\mathrm{link}}(\epsilon)$ diverges when $\epsilon \to 0$ (just as if the specific heat of liquid water approaching its boiling temperature showed a divergence!).

Below the upper critical dimension, there has been very little investigation of χ_{link} (see, however, [40]). In fact, equation (24) has interesting implications in this respect.

---

[13] For the sake of completeness, let us recall that replica and overlap equivalence, combined, imply ultrametricity [46]. In addition, replica equivalence and the Ansatz of a generic ultrametricity implies ultrametricity just as in the SK model [47]. Finally, we point out that replica equivalence is tantamount to stochastic stability and a self-averageness property.

Let us rewrite it in the equivalent form

$$\lim_{L\to\infty}\left[\overline{\langle Q^2_{\mathrm{link}}\rangle}-\overline{\langle Q_{\mathrm{link}}\rangle^2}\right]=\tfrac{2}{3}\lim_{L\to\infty}\left[\overline{\langle Q^2_{\mathrm{link}}\rangle}-\overline{\langle Q_{\mathrm{link}}\rangle}^2\right]. \tag{30}$$

In an RSB system, the right-hand side of equation (30) is positive (since Q_{link} may take values on a finite interval). Yet, equation (27), the lhs of (30) is nothing but the large-L limit of χ_{link}/L^D. Hence, RSB implies $\chi_{\mathrm{link}}\sim L^D$, as expected for first-order phase transitions (see e.g. [51]).

We note that for droplet, or TNT systems, equation (30) is merely an empty $0=\frac{2}{3}\times 0$ statement, just as for RSB systems in their paramagnetic phase. Hence we have found it of interest to study the dimensionless ratio

$$R_{\mathrm{link}}=\frac{\overline{\langle Q^2_{\mathrm{link}}\rangle-\langle Q_{\mathrm{link}}\rangle^2}}{\overline{\langle Q^2_{\mathrm{link}}\rangle}-\overline{\langle Q_{\mathrm{link}}\rangle}^2}. \tag{31}$$

Equation (30) implies that, for an RSB system on its large-L limit, $R_{\mathrm{link}}=\frac{2}{3}$ for all $T<T_{\mathrm{c}}$. For a droplet or TNT system any value $0\le R_{\mathrm{link}}\le 1$ is acceptable. In fact, the high temperature expansion for the $D=3$ EA model tells us that, in the large-L limit, $R_{\mathrm{link}}=1-\mathcal{O}(T^{-2})$.

We finally recall that the Chayes *et al* bound [52, 53] may seem to imply that χ_{link} can diverge at most as $L^{D/2}$, rather than as L^D as required by RSB. The way out of the paradox is a little technical[14].

3. Numerical methods

We describe here our numerical simulations. We describe the simulation organization on Janus in section 3.1. We explain our choice of parameters for the parallel tempering simulation in section 3.2. An absolutely crucial issue is that of thermalization criteria, section 3.3. We largely extend here the methods first introduced in [27], which allows us to distribute on a rational basis the computational resources according to the difficulty in thermalizing each particular sample. At variance with [27], which was restricted to the critical region, we are here probing the deep spin-glass phase, hence more demanding criteria need to be met. The statistical data analysis is described in section 3.4. Finally, in section 3.5 we describe some more traditional thermalization tests.

3.1. The Janus computer

Our Monte Carlo simulations have been carried out on the Janus special-purpose machine. Information about Janus' hardware as well as some details of low-level programming can be found in [22, 23, 54]. Janus is built out of 256 computing cores (Virtex-4 LX200 FPGAs) arranged on 16 boards. With the code used for this paper, each core updates

---

[14] Imagine generalizing model (1) in the following sense: the coupling is $J_{xy}=+1$ with probability p (and $J_{xy}=-1$ with probability $1-p$), so that our model is just the particular instance $p=0.5$. One may follow [52] to show that $\partial\overline{\langle Q_{\mathrm{link}}\rangle}/\partial p$ diverges at most as $L^{D/2}$. However, the critical value of ϵ would still be $\epsilon=0$ for p in a finite range around $p=0.5$ (this is the crucial point: in the standard argument [52, 53] one would require that the critical value of ϵ vary when p moves away from $p=0.5$). Hence, the rate of divergence of p-derivatives does not convey information on the rate of divergence of ϵ-derivatives.

doi:10.1088/1742-5468/2010/06/P06026

3×10^{10} spins s^{-1} with a heat-bath algorithm. The 16 FPGAs on a board communicate with a host PC via a 17th on-board control FPGA.

The controlling PC generates the couplings $\{J_{\boldsymbol{xy}}\}$, initializes the Janus random number generators, and provides as well the starting spin configurations. All the required data is transmitted to the FPGAs (one FPGA per *real* replica) that carry out both the heat-bath (HB) updating and the Parallel Tempering (PT) temperature exchange. Due to the special architecture of Janus, the PT step is not costless, as we previously need to compute the total energy for each temperature. We thus equilibrate the computational cost of both updates by performing several HB sweeps before a PT temperature swap is attempted. Fortunately, selecting a modest number of HB sweeps per PT update hardly affects the efficiency. After a suitable number of PT cycles, spin configurations of all replicas are copied to PC memory to take measurements. The measurement process on the PC is easily parallelized with the next simulation block in Janus so that the PC is always ready for the next reading.

During the simulation, we store on disk information about the PT dynamics (temperature random walk and acceptance rates), configuration energies, and measurements related to the overlap and link overlap fields. We also store full spin configurations every several measurement steps (usually a hundred) to be later used for *offline* measurements (see section 3.4) or as a checkpoint for continuing the simulation if needed.

In a few specific cases (namely one $L = 24$ sample and four $L = 32$ samples) the time required to fulfil our thermalization criteria was exceedingly long, more than six months. For these samples we have accelerated the simulation by increasing the level of parallelism. We have used a special low-level code that transfers the PT procedure to the control FPGA. This has allowed us to distribute the set of temperatures along several FPGAs on a board, speeding up the simulation accordingly.

For the smaller lattices ($L \leq 12$) we substitute the communication with Janus by a call to a simulation routine in the PC. Although these simulations are much less demanding, we go down to very small temperatures. As a consequence, the total cost is not negligible and we have used a PC cluster to complete the simulations.

3.2. Choosing parameters for parallel tempering

The key point in a parallel tempering [55, 56] simulation consists in ensuring that each configuration spends enough time at high temperatures so that its memory can be erased. Since we intend to study the physics of the Edwards–Anderson spin glass at very low temperatures, our simulations are necessarily very long. Because of this, we do not need to reach temperatures as high as those used in critical point studies. We can perform a quantitative analysis using the known behaviour of the heat-bath dynamics above the critical point.

Following [57], the equilibrium autocorrelation time in the thermodynamic limit is taken from a power law to a critical divergence

$$\tau_{\mathrm{HB}}(T) \sim (T - T_{\mathrm{c}})^{-z\nu}. \tag{32}$$

For instance, for the maximum temperature used in our largest lattice ($L = 32$) Ogielski found $\tau_{\mathrm{HB}}(T) \sim 10^5$ [57]. This is several orders of magnitude shorter than our shortest simulations (see table 1).

Table 1. Parameters of our parallel tempering simulations. In all cases we have simulated four independent real replicas per sample. The N_T temperatures are uniformly distributed between $T_{\rm min}$ and $T_{\rm max}$ (except for the runs of the first row, which have all the temperatures of the second one plus $T = 0.150$ and 0.340). In this table $N_{\rm mes}$ is the number of Monte Carlo Steps between measurements (one MCS consists of 10 heat-bath updates and 1 parallel tempering update). The simulation length was adapted to the thermalization time of each sample (see section 3.3). The table shows the minimum, maximum and medium simulation times ($N_{\rm HB}$) for each lattice, in heat-bath steps. Lattice sizes $L = 8, 12$ were simulated on conventional PCs, while sizes $L = 16, 24, 32$ were simulated on Janus. Whenever we have two runs with different $T_{\rm min}$ for the same L the sets of simulated samples are the same for both. The total spin updates for all lattice sizes sum 1.1×10^{20}.

L	$T_{\rm min}$	$T_{\rm max}$	N_T	$N_{\rm mes}$	$N_{\rm HB}^{\rm min}$	$N_{\rm HB}^{\rm max}$	$N_{\rm HB}^{\rm med}$	$N_{\rm s}$	System
8	0.150	1.575	10	10^3	5.0×10^6	8.30×10^8	7.82×10^6	4000	PC
8	0.245	1.575	8	10^3	1.0×10^6	6.48×10^8	2.30×10^6	4000	PC
12	0.414	1.575	12	5×10^3	1.0×10^7	1.53×10^{10}	3.13×10^7	4000	PC
16	0.479	1.575	16	10^5	4.0×10^8	2.79×10^{11}	9.71×10^8	4000	Janus
24	0.625	1.600	28	10^5	1.0×10^9	1.81×10^{12}	4.02×10^9	4000	Janus
32	0.703	1.549	34	2×10^5	4.0×10^9	7.68×10^{11}	1.90×10^{10}	1000	Janus
32	0.985	1.574	24	2×10^5	1.0×10^8	4.40×10^9	1.16×10^8	1000	Janus

The choice of the minimum temperature was taken so that the whole simulation campaign took about 200 days of the whole Janus machine and so that $T_{\rm c} - T_{\rm min} \sim L^{-1/\nu}$. With 4000 samples for $L = 16, 24$ and 1000 for $L = 32$, this resulted in $T_{\rm min} = 0.479, 0.625$ and 0.703, respectively. Smaller lattices, $L = 8, 12$, were simulated on conventional computers. In all cases, we simulated four independent real replicas per sample.

As to the other parallel tempering parameters, namely the number and distribution of intermediate temperatures and the frequency of the parallel tempering updates, the choice is more arbitrary. We dedicated several weeks of the machine to test several combinations trying, if not to optimize our decision, at least to avoid clearly bad choices.

Specifically, we varied the number N_T of temperatures keeping the acceptance of the parallel tempering update between 7% and 36%. This corresponds to an increase of roughly a factor of two in N_T. Noticing that the computational effort is proportional to N_T, we found that the efficiency hardly changed, even for such a wide acceptance range. Eventually, we chose a compromise value of about 20% in the acceptance, resulting in the parameters quoted in table 1. This both avoided unconventionally low acceptances and saved disk space.

In contrast to conventional computers, Janus needs about as much time to do a parallel tempering update as a heat-bath one. Therefore, while it is customary to perform both updates with the same frequency, after testing frequencies in the range 1–100 we have chosen to do a parallel tempering update each 10 heat-bath ones. In fact, even if the time to do a parallel tempering step were negligible, we have checked that doing a single heat-bath between parallel temperings would produce a practically immeasurable gain. We note, finally, that this issue was investigated as well in [58] (in that work clear conclusions were not reached, as far as the $D = 3$ Edwards–Anderson model at low temperatures and large L is concerned).

3.3. Thermalization criteria

In order to optimize the amount of information one can obtain given a computational budget, the length of the simulations must be carefully selected. It is well known that sample-to-sample fluctuation is the main source of statistical error. Thus, we want to simulate each sample for the shortest time that ensures thermalization.

The most common robust thermalization check consists in the determination of the autocorrelation times for physical observables [59]. However, in order for this determination to be precise one needs a much longer simulation than needed to thermalize the system (e.g., while ten exponential autocorrelation times can be enough to thermalize the system, we need an at least ten times longer simulation to determine this autocorrelation time). Notice that this is not an issue in ordered systems, where one employs very long simulations in order to reduce statistical errors.

The typical practical recipe to assess thermalization for disordered systems consists in studying the time evolution of the disorder-averaged physical observables. In particular, the so-called $\log_2$-binning procedure uses the evolution of the time averages along the intervals $I_n = (2^{-(n+1)}N_{\mathrm{HB}}, 2^{-n}N_{\mathrm{HB}}]$. The system is considered to be thermalized if the first few intervals are compatible.

This procedure is not optimal, because the thermalization time is wildly dependent on the sample. Thus, a simulation time long enough to thermalize the slowest samples will be excessive for most of the rest. Perhaps even more frightening, the average over samples may well hide that a few samples, the very worst ones, are still quite far from equilibrium.

Fortunately, the use of parallel tempering presents us with the possibility to use the dynamics of the temperature random walk to learn about the thermalization scale for each sample. In fact, in order to ensure thermalization, each of the participating configurations must cover the whole temperature range. Here, expanding on a method first used in [27], we have promoted this idea to a fully quantitative and physically meaningful level.

Let us consider the ordered set of N_T temperatures $\{T_1, \ldots, T_{N_T}\}$ and let us suppose that $T_{i_{\mathrm{c}}-1} < T_{\mathrm{c}} \leq T_{i_{\mathrm{c}}}$. In figure 1—left we show an instance of the random walk of the temperature index, $i(t) \in \{1, 2, \ldots, N_T\}$, performed by one of the N_T copies of the system considered in the parallel tempering. The random walk is clearly not Markovian, as the system remembers for a long time that it belongs to the high (low) temperature phase. This effect is also demonstrated in figure 1—right, where we plot the time spent over T_{c} as a function of the simulation time (note the long plateaux).

To make these arguments quantitative, we shall use the standard tools of correlated time series [51, 59]. We need a mapping defined on the $1, \ldots, N_T$ range of temperature indices so that

$$f(i) \geq 0, \qquad \forall i \geq i_{\mathrm{c}}, \tag{33}$$

$$f(i) < 0, \qquad \forall i < i_{\mathrm{c}}, \tag{34}$$

$$\sum_{i=1}^{N_T} f(i) = 0. \tag{35}$$

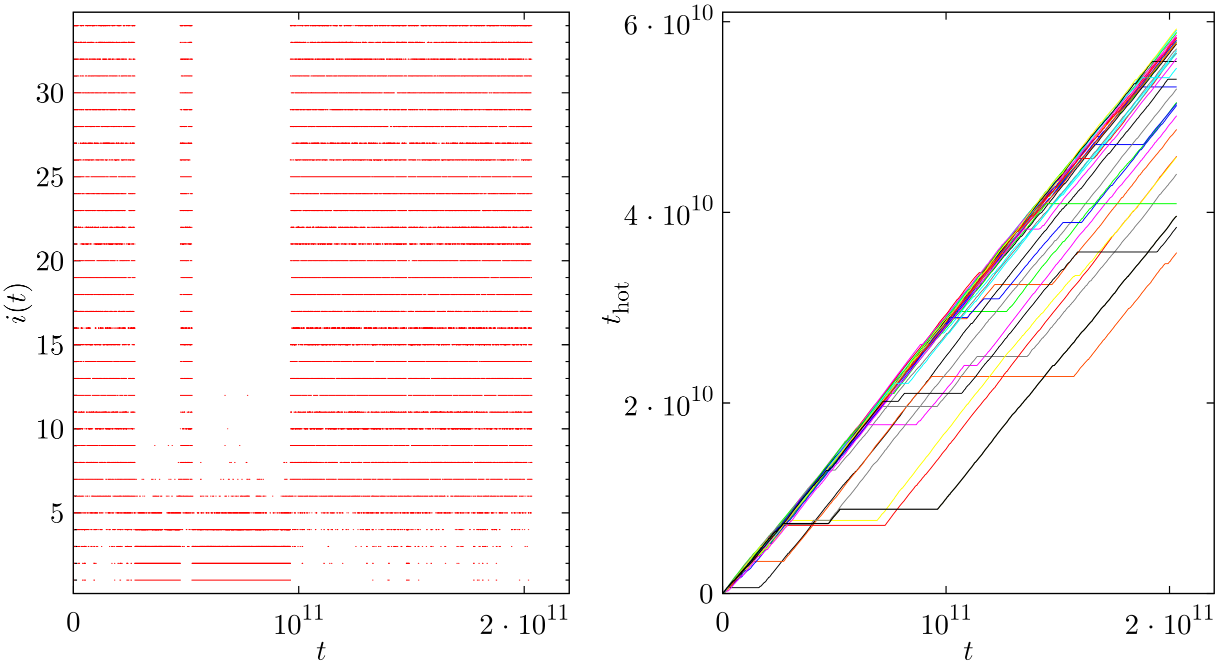

Figure 1. We plot (left panel) the temperature index of a fixed configuration of an $L = 32$ sample as a function of the number of HB sweeps. We plot one point every 5 million HB sweeps. The critical temperature corresponds to $i_{\mathrm{c}} = 17$. This specific sample has $\tau_{\mathrm{exp}} = 1.75 \times 10^{10}$ HB sweeps. In the right panel, we show the time that all configurations of the same replica spend in the paramagnetic phase.

It is also convenient that f be monotonic. Because we have chosen the same number of temperatures above and below T_{c}, a simple linear f is suitable, but the method works with any function fulfilling the above conditions.

For each of the participating configurations we consider the time evolution i_t of the temperature index. We define the equilibrium autocorrelation function as

$$C(t) = \frac{1}{N_{\mathrm{HB}} - t_0 - t} \sum_{t'=t_0}^{N_{\mathrm{HB}}-t} f(i_{t'}) f(i_{t'+t}), \tag{36}$$

where t_0 is long enough to ensure that the temperature random walk has reached a steady regime. Due to condition (35), we avoid subtracting the squared mean value of f in this definition. From the normalized $\hat{C}(t) = C(t)/C(0)$, see e.g. figure 2, we can define the integrated correlation times:

$$\tau_{\mathrm{int}} = \tfrac{1}{2} + \sum_{t=0}^{W} \hat{C}(t), \tag{37}$$

where W is a self-consistent window that avoids the divergence in the variance of τ_{int}.

The great advantage of these functions over the physical observables is that we can average over the N_T configurations in the parallel tempering[15].

---

[15] Even if these are not completely statistically independent, the averaged autocorrelation has a much smaller variance. In addition, the need to simulate several replicas provides independent determinations of $C(t)$, which permits a further error reduction and an estimate of the statistical errors.

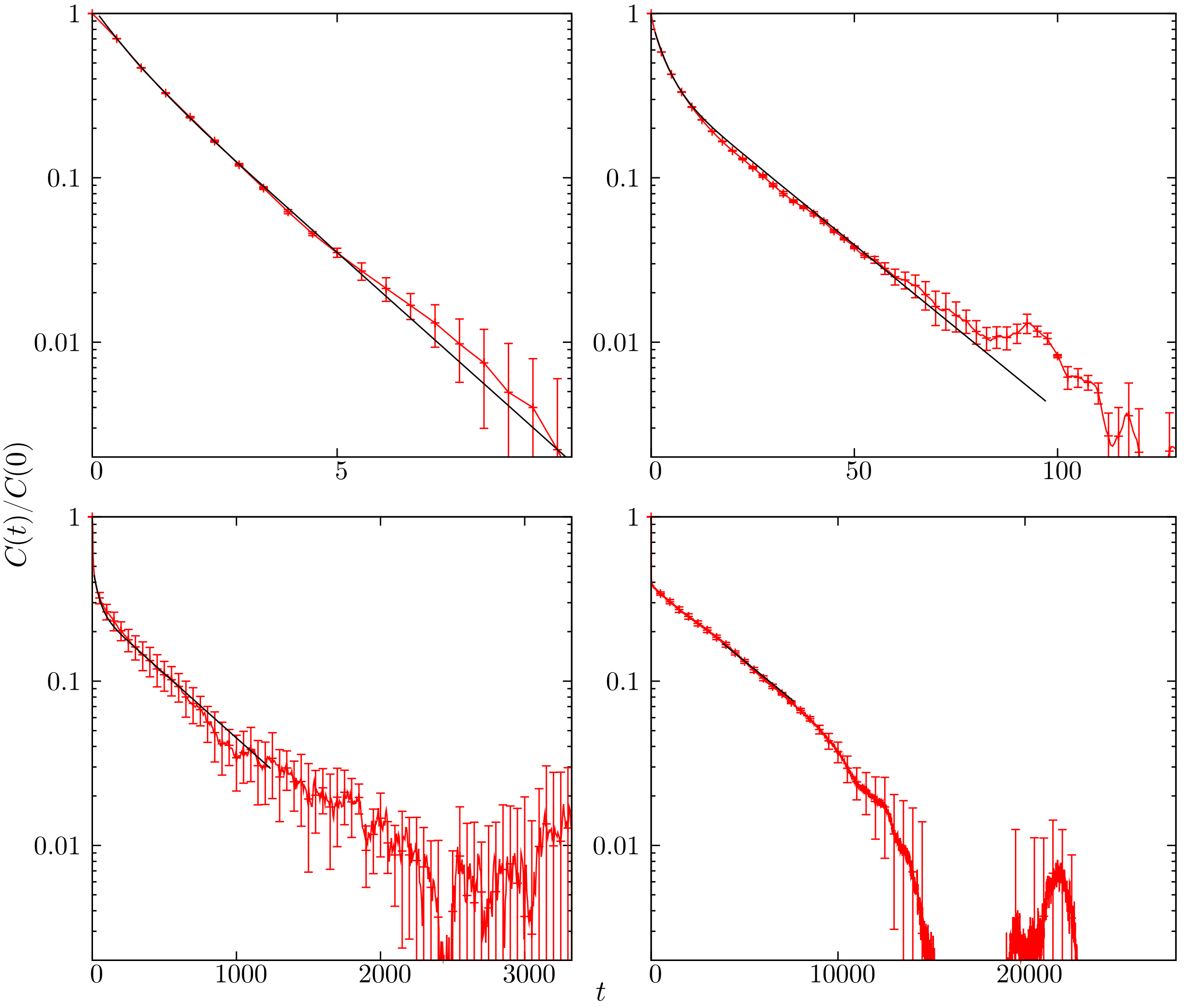

Figure 2. Autocorrelation functions for samples with $\tau_{\rm exp}$ of different orders of magnitude. We have plotted the range $[0, 6\tau_{\rm exp}]$. We include the automatic double exponential fit, see appendix A. In the last panel the fit fails due to the strong downwards fluctuation and our programme has chosen a restricted interval for a fit to a single exponential. In order to avoid cluttering the graphs, we have only plotted a few times (the actual correlation functions have many more points). The horizontal axis is in units of 10^6 heat-bath updates.

This procedure works surprisingly well, not only giving reliable estimates of the integrated time but even providing the, more physical but notoriously difficult to measure, exponential autocorrelation time. Indeed, the correlation function admits an expansion on exponentially decaying modes

$$\hat{C}(t) = \sum_i A_i \, \mathrm{e}^{-t/\tau_{\mathrm{exp},i}}, \qquad \sum_i A_i = 1. \tag{38}$$

In this representation, the exponential time $\tau_{\rm exp}$ is the largest of the $\tau_{{\rm exp},i}$.[16] Barring symmetry considerations, this exponential time should be the same for all random variables in the simulation, including the physical observables.

The relative sizes of the A_i, and hence $\tau_{\rm int}$, depend to a certain extent on the particular choice of f. Notice, however, that criteria (33) and (34) select a family of functions that hopefully reduce the amplitude of the irrelevant fast modes. In any case, $\tau_{\rm exp}$ has a physical meaning independently of these somewhat arbitrary considerations.

In practice, the simulations are too long (up to $N_{\rm HB} \sim 10^{12}$) to consider all the $f(i_t)$ individually and we have to introduce some data binning, averaging over a large number of consecutive measurements. As it turns out, this is not a very limiting issue for two reasons. On the one hand, as long as these bins are much shorter than τ, there is no real information loss. On the other hand, one can reconstruct any polynomial f up to degree k—in particular our linear f—by saving the sums of the first k powers of the i_t.

Even after this binning, we have worked with time series with a length of up to several million, so in order to compute the autocorrelation we have used a Fast Fourier Transform algorithm [60].

The details of the chosen thermalization protocol can be found in appendix A. We summarize by saying that our main thermalization criterion is ensuring that $N_{\rm HB} > 12\tau_{\rm exp}$ ($2\tau_{\rm exp}$ are discarded and the remaining $10\tau_{\rm exp}$ are used to measure and study $\hat{C}(t)$).

In figure 2 we plot several autocorrelation functions showing how the data quality allows for an exponential fit. We have chosen randomly four samples with very different exponential autocorrelation times: 6.5×10^6, 8.8×10^7, 1.5×10^9 and 1.8×10^{10}.

To summarize the distribution of the exponential autocorrelation times we have computed a histogram. Due to the large dispersion of these quantities we have chosen $\log_2 \tau_{\rm exp}$ as a variable. In figure 3 we show the results for the two runs performed in $L = 32$ (see table 1). Notice the dramatic increase of the $\tau_{\rm exp}$ when decreasing the minimum temperature of the simulation. The smooth shape of the curves defined by the histogram is a further test of our procedure for determining autocorrelation times.

In figure 4 we plot the logarithm of the histogram in the $L = 24$ case to show the exponential behaviour of the long-times tail. This result gives confidence that rare events, with very large (logarithms of) autocorrelation times, are at least exponentially suppressed. We have not made efforts to measure with precision the small autocorrelation times as they are immaterial regarding thermalization, which is ensured by the minimum number of iterations performed for all samples.

3.4. Monte Carlo evaluation of observables

We present now some technical details about our evaluation of mean values, functions of mean values and error estimation.

Some of the observables considered in this work were obtained by means of an online analysis: the internal energy, the link overlap, powers of the spin overlap (q, q^2, q^4), and Fourier transforms of the correlation function $C_4(\boldsymbol{r})$ for selected momenta. These quantities were computed as Monte Carlo time averages along the simulation. Note that the length of the simulation is sample dependent, something that would be a nuisance

---

[16] The number of modes equals the dimension of the dynamical matrix of the Monte Carlo Markov process, which in our case is $(N_T!) \times 2^{N_T V}$.

J. Stat. Mech. (2010) P06026

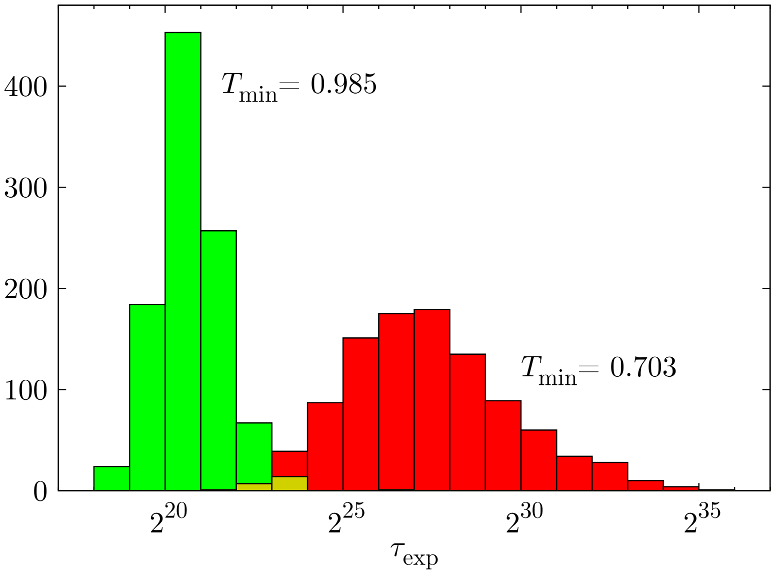

Figure 3. Histogram of exponential autocorrelation times for our simulations of the $L = 32$ lattice (1000 samples).

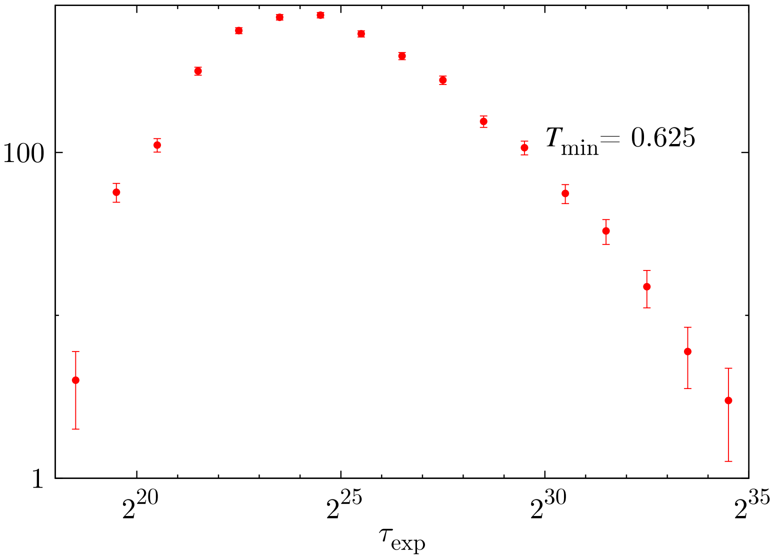

Figure 4. Logarithm of the histogram of exponential autocorrelation times for our simulations of the $L = 24$ lattice (4000 samples). Note the behaviour of the long-times tail.

in a multispin coding simulation, but not in Janus where each sample is simulated independently. The disorder averaging followed the Monte Carlo one. Statistical errors were computed using a jackknife method over the samples, see for instance [51].

However, when designing the simulation, one cannot anticipate all quantities that would be interesting, or these can be too expensive to be computed in runtime. In particular, we did not compute the conditional correlation functions $C_4(\boldsymbol{r}|q)$. Fortunately,

J. Stat. Mech. (2010) P06026

an offline analysis of the stored configurations has allowed us to estimate them. We had to overcome a difficulty, though, namely the scarcity of stored configurations. In fact, for the samples that were simulated only for the minimum simulation time, we had only $N_{\mathrm{conf}} \sim 100$ configurations stored on disk (ranging from $N_{\mathrm{conf}} = 10$ for $L = 12$ to $N_{\mathrm{conf}} = 200$ in the case $L = 32$). We regard the second half (in a Monte Carlo time sense) of these configurations as fireproof thermalized. Yet, when forming the overlap field, equation (3), one needs only that the two spin configurations, $\{s_{\boldsymbol{x}}^{(1)}\}$ and $\{s_{\boldsymbol{x}}^{(2)}\}$, be thermalized and independent. Clearly enough, as long as the two configurations belong to different real replicas and belong to the second half of the Monte Carlo history they will be suitable. There is no need that the two configurations were obtained at the same Monte Carlo time (as it is done for the online analyses). Furthermore, the four real replicas offer us 6 pair combinations. Hence, we had at least $6 \times (N_{\mathrm{conf}}/2)^2 \sim 10\,000$ (60 000 for $L = 32$) measurements to estimate the overlaps and the correlation functions. We used the Fast Fourier Transform to speed up the computation of the spatial correlations. For those samples that had more configurations (because their total simulation time exceeded $N_{\mathrm{min}}^{\mathrm{HB}}$), we considered nevertheless $N_{\mathrm{conf}}/2$ configurations evenly spaced along the full second half of the simulation. When some quantity, for instance the $P(q)$, could be computed in either way, online or offline, we have compared them. The two ways turn out to be not only compatible, but also equivalent from the point of view of the statistical errors. As an example of this let us compute the following quantity:

$$\sigma_{\mathrm{link}}^2 = \overline{\langle Q_{\mathrm{link}}^2 \rangle} - \overline{\langle Q_{\mathrm{link}} \rangle}^2 . \tag{39}$$

For $L = 32$, $T = 0.703$, the value of σ_{link}^2 computed from online measurements of Q_{link} and Q_{link}^2 is

$$V\sigma_{\mathrm{link,online}}^2 = 50.88(90). \tag{40}$$

We could now recompute this value from offline measurements of Q_{link} and Q_{link}^2. Instead, we are going to use equation (12), which involves the intermediate step of computing conditional expectation values and variances at fixed q and then integrating with $P(q)$. This will serve as a test both of the offline measurements' precision and of our Gaussian convolution method for the definition of clustering quantities. The result is

$$V\sigma_{\mathrm{link,conf}}^2 = 50.81(90). \tag{41}$$

The precision of $\sigma_{\mathrm{link,online}}^2$ and $\sigma_{\mathrm{link,conf}}^2$ is the same and the difference less than 10% of the error bar, even though we only analysed 100 configurations per sample for the second one. Of course, both determinations are very highly correlated, so the uncertainty in their difference is actually much smaller than their individual errors. Computing the difference for each jackknife block we see that

$$V[\sigma_{\mathrm{link,conf}}^2 - \sigma_{\mathrm{link,online}}^2] = -0.065(79), \tag{42}$$

which is indeed compatible with zero.

A subtle point regards non-linear functions of thermal mean values that are later on averaged over the disorder. In this work, the only instance is χ_{link}, see equation (27). Care is needed to estimate such non-linear functions because a naive evaluation would be biased, and the bias might be sizeable compared to the statistical errors [61]. This problem does not arise in non-linear functions such as equation (9), which are computed

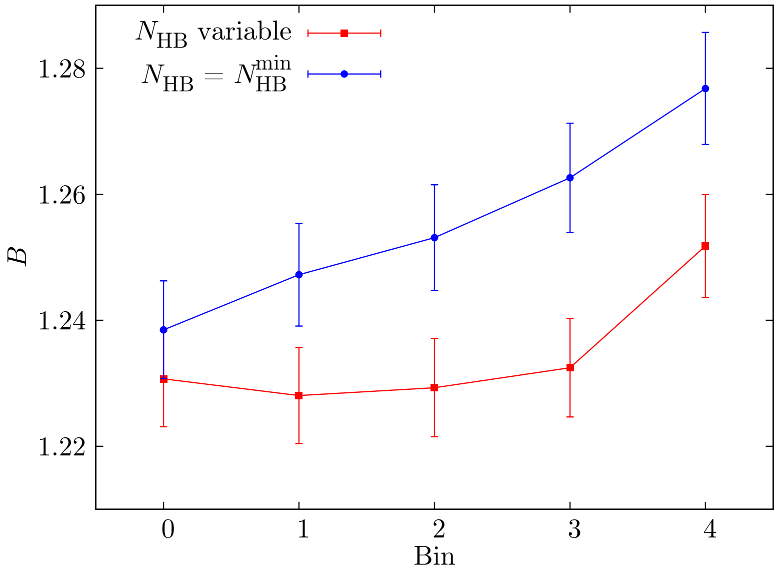

Figure 5. Evolution of the Binder parameter for $L = 32$, $T = 0.703$ using $\log_2$ binning (0 = second half, 1 = second quarter, $\ldots$). The blue curve (circles) is the result of stopping at step 1 of our thermalization protocol (i.e., all samples simulated for a fixed time of 4×10^9 heat-bath updates). The red curve (squares) is the result of completing all the steps, which implies an increase of roughly 150% in simulation time.

on observables only *after* the double averaging process over the thermal noise and over the samples. The problem and several solutions are discussed in appendix B (see also [32]).

A final issue is the comparison of data computed in different system sizes at the *same* temperatures. Unfortunately the grids of temperatures that we used for the different L differ. Hence we have interpolated our data by means of a cubic spline.

3.5. Thermalization tests

We will consider in this subsection thermalization tests directly based on physically interesting quantities.

We start with the traditional $\log_2$-binning procedure. We choose the Binder parameter for the overlap, see equation (6), which is specially sensitive to rare events. In figure 5 we show the results for $B(T_{\rm min})$ for $L = 32$, considering only the first 4×10^9 heat-bath steps of each of our 1000 samples, as if all the simulations were $N_{\rm min}^{\rm HB}$ heat-bath steps long (blue line). We could not affirm that even the last two bins were stable within errors. Things change dramatically if we consider Monte Carlo histories of a length proportional to the exponential autocorrelation time. Note that, thanks to our choice of $N_{\rm min}^{\rm HB}$ in table 1, the simulation time for most samples has not increased. If we first rescale data according to the total simulation length (itself proportional to the autocorrelation time) and average for equal *rescaled* time, the $\log_2$-binning procedure gives four steps of stability within errors. That is to say: we obtain the Binder parameter without thermalization bias just discarding 1/16 of the history (and taking up to 1/8). Regarding the Binder parameter our requirement of $12\tau_{\rm exp}$ is excessive.

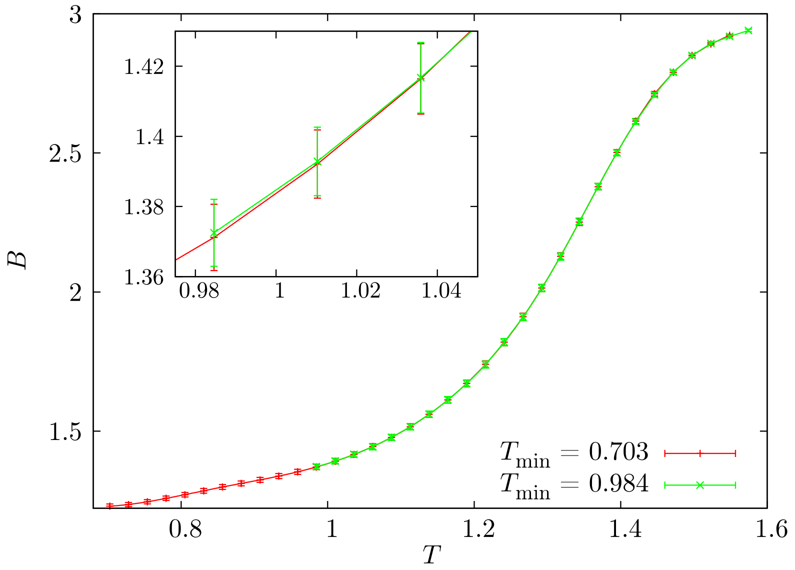

Figure 6. Binder ratio as a function of the temperature for $L = 32$. The good overlap between two different simulations (one of them in the much easier critical region) is a further thermalization check. We use the same set of 1000 samples.

In retrospect (see figure 5), shorter simulations would have produced indistinguishable physical results for most observables. We do not regret our choices, however, as we plan to use these thermalized configurations in the future [62] for very delicate analyses (such as temperature chaos), which are much more sensitive to thermalization effects.

A different test can be performed by comparing the difficult low temperature simulations of our largest lattice with simulations in the critical region of the *same samples*. A faulty thermalization (for instance, a configuration remains trapped at low temperatures) could be observable as inconsistencies in the values of quantities in common temperatures. In figure 6 we show the Binder parameter as a function of temperature for the two simulations with $L = 32$ (see table 1). The agreement between both simulations is excellent.

A very different test on the statistical quality of our data is the comparison of the values of $\chi_{\rm link}$ obtained using the different possible estimators for $\langle Q_{\rm link}\rangle^2$. We have an unbiased estimator if we use $Q^{(2)}_{\rm link,4R}$, see equation (B.6), the linearly bias-corrected estimator $Q^{(2)}_{\rm link,linear}$ in equation (B.4), and the quadratically bias-corrected estimator $Q^{(2)}_{\rm link,quadratic}$ in equation (B.5). The different determinations are equal only if the total simulation time (in each sample) is much longer than the integrated autocorrelation time for $Q_{\rm link}$. As we see in figure 7—left, only computing $\chi_{\rm link}$ from the biased estimator $[Q_{\rm link}]^2_{2/2}$ results in a measurable bias. Once bias correction is taken into account, differences are only a fraction of the statistical error for each estimator. Nevertheless, the different statistical estimators are dramatically correlated. Hence, their difference might be significant. In figure 7—right we plot these differences for $L = 24$ and 32 as a function of temperature, in units of the statistical error for that difference. As we see, at the lowest temperatures for $L = 32$, the bias for the estimate of $\chi_{\rm link}$ obtained from $Q^{(2)}_{\rm link,linear}$ is still measurable. Only the estimate from $Q^{(2)}_{\rm link,quadratic}$ is statistically compatible with

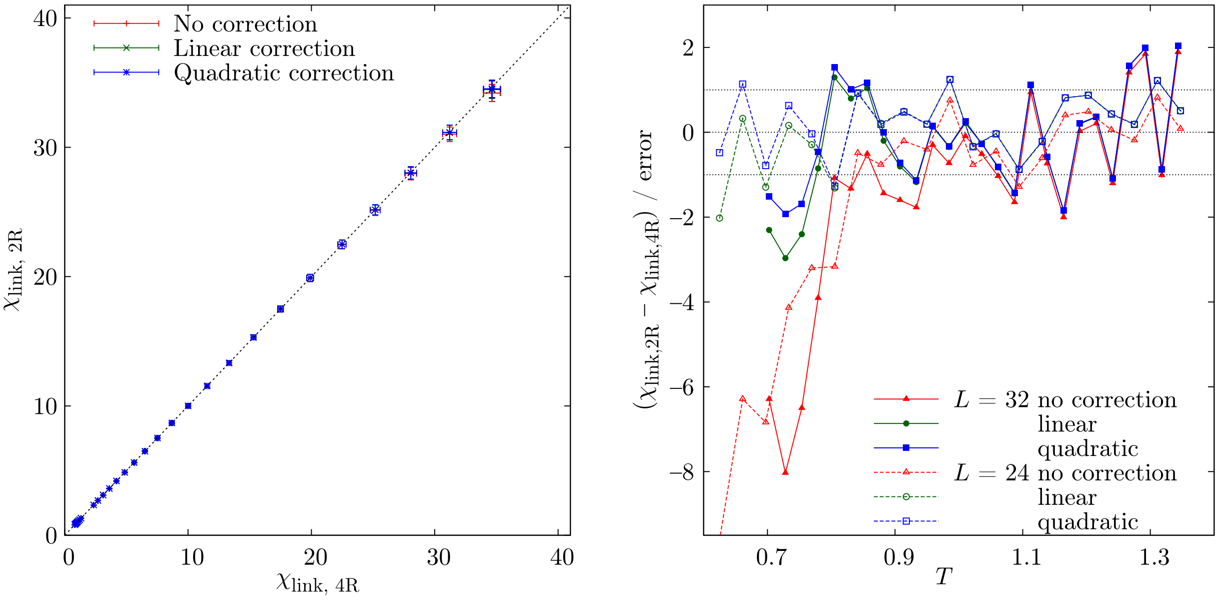

Figure 7. Bias correction in the computation of χ_{link}, equation (27). On the left panel we plot the two-replica estimators $\chi_{\text{link,2R}}$ as a function of the unbiased four-replica estimator $\chi_{\text{link,4R}}$, equation (B.10), for all our temperatures in the $L = 32$ lattice. The two-replica estimators $\chi_{\text{link,2R}}$ are computed with no bias correction, equation (B.7), with linear corrections, equation (B.8), and with quadratic corrections, equation (B.9). The right panel displays, for the three two-replica estimators, their difference with the four-replica estimator *in units of the statistical error for that difference*, as a function of temperature. We show our data for $L = 32$ and 24. Note that the statistical error in the *difference* between two estimators is largely reduced (as compared to individual errors) due to dramatic data correlation.

the unbiased estimator. Since our data fully complies with our expectations, we consider the above analysis as a confirmation of our expectation $N \gg \tau_{\text{int},Q_{\text{link}}}, \tau_{\text{exp}}$.

We carefully avoided making decisions during thermalization based on the values of physical quantities. However, one could worry about the possibility of important statistical correlations between the temperature random walk and interesting quantities. Such correlation could produce some small biases that would be difficult to eliminate. Fortunately, we have not found any correlation of this type. In figure 8 we show the correlation between τ_{exp} and two important quantities: probability of the overlap being small and the energy.

4. The overlap probability density

In this section we study the pdf of the spin overlap. This is a particularly interesting quantity because, as we saw in section 2, it has a qualitatively different behaviour in the droplet and RSB pictures of the spin-glass phase.

We have plotted $P(q)$ for $T = 0.703$ (the lowest for $L = 32$) and $T = 0.625$ (the lowest for $L = 24$) in figure 9. Notice that the convolution of the comb-like $\tilde{P}(q)$, equation (7),

J. Stat. Mech. (2010) P06026

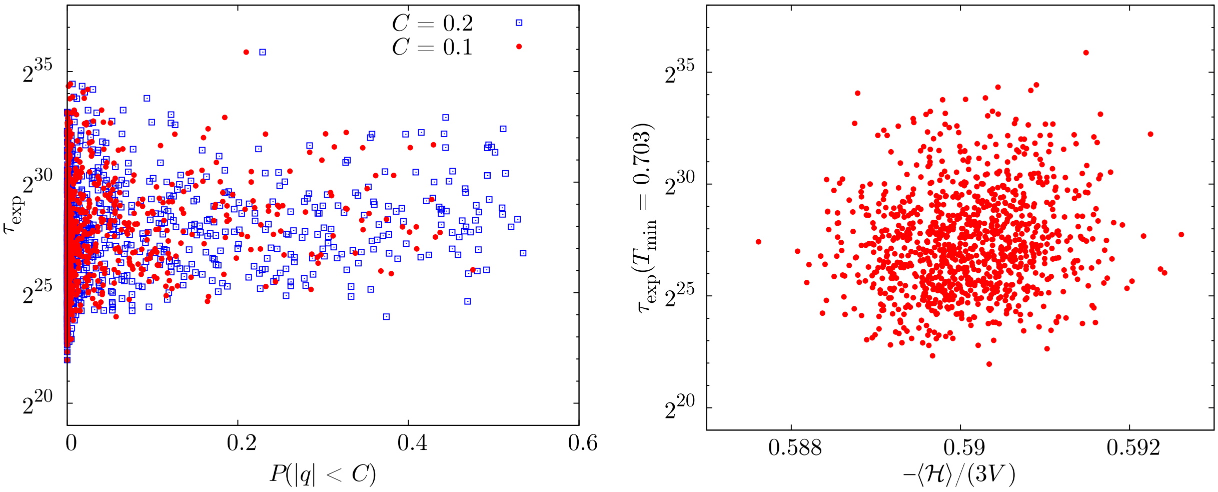

Figure 8. Scatter plot of the exponential autocorrelation time ($L = 32$) versus the probability of the overlap being less than a small quantity (left) and the energy (right). We do not observe correlation between the thermalization times and these physically relevant quantities.

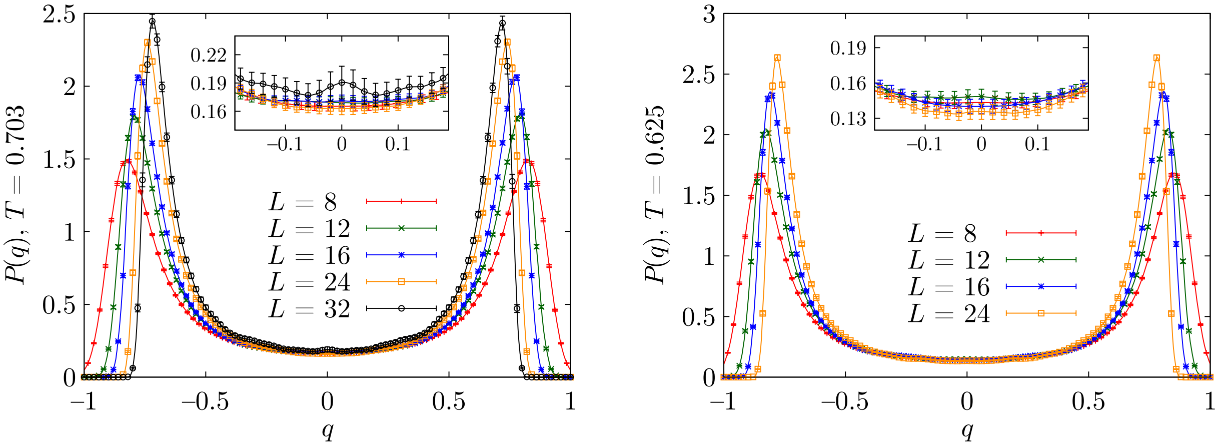

Figure 9. Overlap probability density function $P(q)$, equation (8), at $T = 0.625$ and 0.703. Notice that for the central sector of $q \sim 0$ the curves for the different system sizes quickly reach a plateau with $P(q) > 0$.

with the Gaussian function, equation (8), has yielded a very smooth $P(q)$. Initially, one would expect the peaks of this pdf to grow narrower and closer together as L increases, eventually becoming two Dirac deltas at $\pm q_{\mathrm{EA}}$. The shift in position is clearly visible in the figures, but a more careful analysis is needed to confirm that the peaks are indeed getting sharper (section 4.3). In addition, the probability in the $q = 0$ sector should either go to zero (droplet) or reach a stable non-zero value (RSB). Even if a visual inspection of figure 9 seems to favour the second scenario, we shall need a more quantitative analysis to draw conclusions.

In the remainder of this section we undertake such a quantitative characterization of $P(q)$ and, in particular, its thermodynamical limit. To this end, we will study the

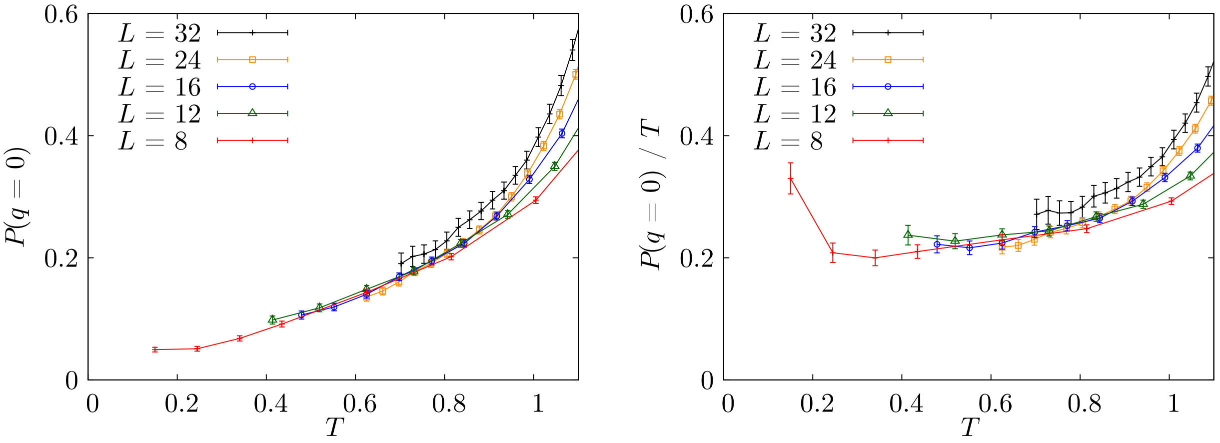

Figure 10. Overlap density distribution function at zero overlap as a function of temperature. We observe an enveloping curve with a linear behaviour, as expected in an RSB setting.

evolution of $P(q = 0)$ with T and L (section 4.1); the extrapolation to infinite volume of the Binder cumulant (section 4.2) and finally the evolution of the shape and position of the peaks with the system's size (section 4.3).

4.1. The $q = 0$ sector

We have plotted in figure 10—left the probability density at $q = 0$ as a function of T for all our lattices. There clearly is an enveloping curve in the region $T < T_\mathrm{c}$ with a decreasing, but positive, value of $P(0)$. In a mean-field setting [8] we expect this probability density to go to zero linearly in T. In order to check this, we have plotted $P(0)/T$ against T in figure 10—right. As we can see, this expectation is fulfilled. For a similar study see [63]. We remark that the seemingly out of control value of $P(0)$ for our lowest temperature in $L = 8$ is an artefact of the binary nature of the couplings (a finite system always has a finite energy gap). Indeed, in [64], the finite size behaviour of $P(0)$ for the Edwards–Anderson model with binary couplings was studied as a function of temperature. Finite size effects on $P(0)$ turned out to be stronger close to $T = 0$ than at finite temperature.

From a droplet model point of view, Moore *et al* [21] have argued that the apparent lack of a vanishing limit for $P(0)$ in numerical work in the 1990s was an artefact of critical fluctuations. In fact, at T_c, $P(0)$ diverges as $L^{\beta/\nu}$ while droplet theory predicts that, for very large lattices, it vanishes as $L^{-\zeta}$, with $\zeta \sim 0.2$, for all $T < T_\mathrm{c}$. These authors rationalize the numerical findings as a crossover between these two limiting behaviours. However, a numerical study at very low temperatures (so the critical regime is avoided) found for moderate system sizes a non-vanishing $P(0)$ [63]. Furthermore, we compute in section 4.3 a characteristic length for finite size effects in the spin-glass phase, which turns out to be small at $T = 0.703$.

4.2. The Binder cumulant

We have plotted the Binder cumulant (6) for $T = 0.625, 0.703$ as a function of the system size in figure 11. As discussed in section 2.2, the evolution (and thermodynamical limit)

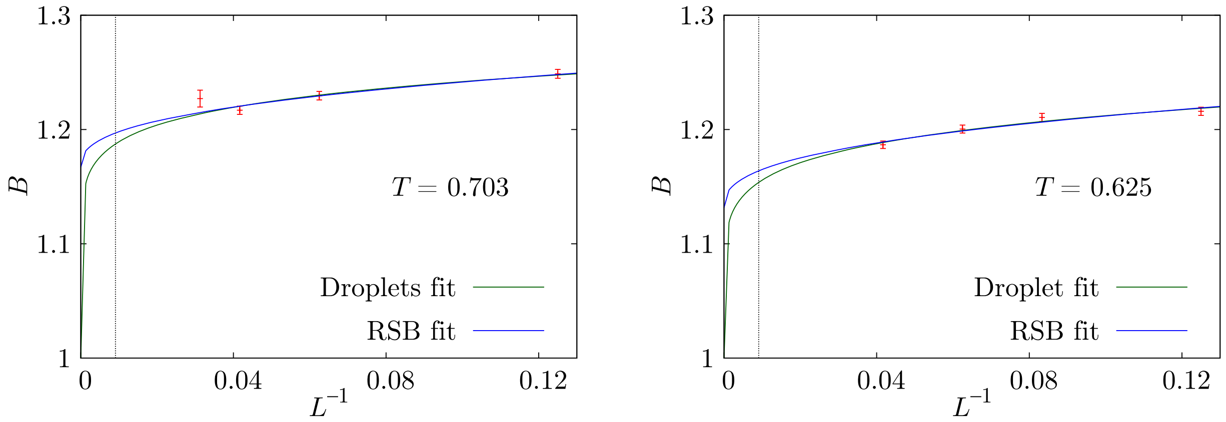

Figure 11. Infinite volume extrapolation of the Binder parameter at $T = 0.703$ and 0.625 and fits to the behaviour expected in the RSB, equation (44), and droplet, equation (43), pictures. See table 2. For the experimentally relevant scale of $L = 110$ (dotted vertical line, see section 6) both fits are well above the $B = 1$ value of a coarsening system.

Table 2. Scaling of the Binder parameter and fit to the behaviour expected in the droplet, equation (43), and RSB pictures, equation (44).

	Droplet fit			RSB fit		
T	χ^2/d.o.f.	a	ζ	χ^2/d.o.f.	c	d
0.703	3.78/3	0.312(17)	0.110(17)	3.44/3	1.165(12)[34]	0.186(34)[03]
0.625	2.00/2	0.289(16)	0.134(21)	2.73/2	1.128(11)[33]	0.193(28)[03]

of this observable is different in the droplet and RSB pictures:

$$\text{Droplet:} \qquad B(T;L) = 1 + aL^{-\zeta}, \tag{43}$$

$$\text{RSB:} \qquad B(T;L) = c + dL^{-1/\hat{\nu}}, \tag{44}$$

where $1/\hat{\nu} = 0.39(5)$ [26]. Since it is compatible with our best estimate for the replicon exponent, $\theta(0) = 0.38(2)$, we prefer to use the second, more accurate value (there is some analytical ground for this identification [26]). We will attempt to distinguish between these two behaviours by fitting our data to (43) and (44).

These two-parameter fits are plotted in figure 11 and the resulting parameters are gathered in table 2. In the case of the RSB fit, equation (44), we have included two error bars: the number enclosed in parentheses $(\cdot)$ comes from the statistical error in a fit fixing $1/\hat{\nu}$ to $\theta(0)$ and the one inside square brackets $[\cdot]$ is the systematic error due to the uncertainty in $\theta(0)$.

As it turns out, both fits have acceptable values of χ^2 per degree of freedom (d.o.f.). However, the evolution of B with L is very slow, so in order to accommodate the limit value of $B(L \to \infty) = 1$ consistent with the droplet picture, we have needed a very small exponent ($\zeta \sim 0.12$, smaller than the droplet prediction of $\zeta \approx 0.2$ [11]). On the other hand, according to droplet theory [11], the connected spatial correlation function at $q = q_{\rm EA}$ decays as $1/r^{\zeta}$. A direct study [26], however, indicates that these correlations decay as $1/r^{0.6}$.

The reader may find it disputable, from an RSB point of view, that a single power law should govern finite size effects. It would be rather more natural that corrections were of order

$$\frac{1}{L^{\theta_{\rm eff}(L)}} = \int_0^1 {\rm d}q\, \frac{P(q)}{L^{\theta(q)}}. \tag{45}$$

It turns out, however, that $\theta(q)$ hardly depends on q (except on the neighbourhood of $q_{\rm EA}$), see [26] and section 5. The neighbourhood of $q_{\rm EA}$ would produce a subleading correction of order $1/L^{0.6}$.

In any case, see section 6, we remark that the relevant regime for comparison with experimental work is $L \approx 110$, where both the RSB and the droplet fits predict that $B(T, L)$ is well above 1 (see figure 11).

4.3. The peaks of $P(q)$, $q_{\rm EA}$, and finite size effects

One of the features of the $P(q)$ about which droplet and RSB agree is the fate of its two symmetric peaks as we approach the thermodynamical limit. These should grow increasingly narrow and shift their position until they eventually become two Dirac deltas at $q = \pm q_{\rm EA}$. The actual value of $q_{\rm EA}$ is notoriously difficult to compute [25, 65, 66], see, however, [26].

Characterizing the evolution of these peaks as we increase the system size is the goal of this section. We start by defining $q_{\rm EA}(L)$ as the position of the maximum of $P(q; L)$ (since the pdf is symmetric, we shall consider all overlaps to be positive in the remainder of this section). Thanks to the Gaussian smoothing procedure described in equation (8), this maximum is very well defined. We compute its position by fitting the peak to a third-order polynomial (notice that the peaks are very asymmetric).

In order to further describe the peaks, we will also employ the half-widths $\sigma^{(\pm)}$ at half height $[P(q^{(\pm)}) = P(q_{\rm EA}(L))/2]$:

$$\sigma^{(\pm)} = |q^{(\pm)} - q_{\rm EA}(L)| \tag{46}$$

where $q^{(-)} < q_{\rm EA}(L) < q^{(+)}$.

We have plotted these parameters as a function of temperature in figure 12. In table 3 we can see that the width of the peaks does decrease with a power law in L, although very slowly. The product $\sigma P(q_{\rm EA}(L))$ has a small dependence on L.

We can now extrapolate $q_{\rm EA}(L)$ to find the order parameter in the thermodynamical limit. A finite size scaling study [26] shows that

$$q_{\rm EA}(L, T) = q_{\rm EA}^{\infty}(T)\left[1 + \frac{A(T)}{L^{1/\hat{\nu}}}\right], \qquad A(T) = [L_{\rm c}(T)]^{1/\hat{\nu}}, \tag{47}$$

where $1/\hat{\nu} = 0.39(5)$. Yet, as discussed after equation (44), we prefer to identify $1/\hat{\nu}$ with the replicon exponent, $\theta(0) = 0.38(2)$. A disagreeing reader merely needs to double the error estimate in the extrapolation of $q_{\rm EA}$. Note that one should not attempt a three-parameter fit to equation (47), as there are too few degrees of freedom. An independent estimate of $1/\hat{\nu}$ is required. Similar extrapolations were attempted in [47], with smaller system sizes ($L \leq 16$) and a lesser control over $1/\hat{\nu}$.

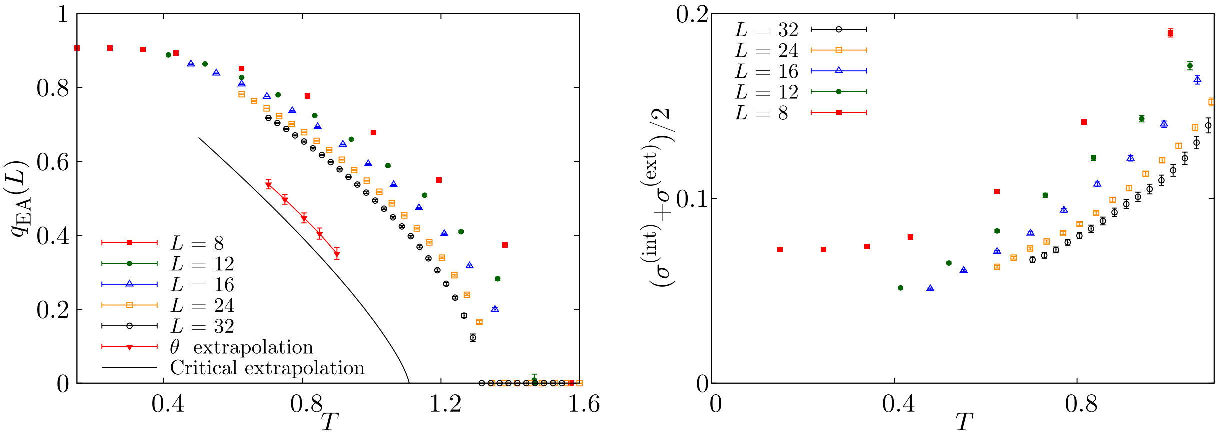

Figure 12. Left: $q_{\rm EA}(L)$ as a function of the temperature. We include two different infinite volume extrapolations: using the replicon exponent, equation (47) and table 4, and the one obtained from finite size scaling arguments in the critical region, equations (49) and (52). Right: width of the peaks of $P(q)$, equation (46), as a function of T for all our lattice sizes.

Table 3. Width $\sigma = (\sigma^{(+)} + \sigma^{(-)})/2$ of the peaks in $P(q)$ and fit to a power-law $\sigma(L) = AL^B$ in the range $[L_{\rm min}, 32]$. We also include the product $\sigma P(q_{\rm EA}(L))$.

	$T = 0.703$		$T = 0.805$	
L	σ	$\sigma P(q_{\rm EA}(L))$	σ	$\sigma P(q_{\rm EA}(L))$
8	0.1177(20)	0.1784(10)	0.1391(25)	0.1833(10)
12	0.0963(21)	0.1740(12)	0.1165(25)	0.1809(12)
16	0.0817(16)	0.1696(11)	0.1001(22)	0.1756(11)
24	0.0735(16)	0.1690(12)	0.0860(19)	0.1728(12)
32	0.0668(29)	0.1631(23)	0.0798(34)	0.1669(22)
$L_{\rm min}$	16		16	
χ^2/d.o.f.	0.43/1		1.13/1	
B	−0.278(28)		−0.346(30)	

We present the values of $q_{\rm EA}(L)$ and the result of a fit to equation (47) in table 4. As we can see, the errors due to the uncertainty in the exponent, denoted by $[\cdot]$, are greater than those caused by the statistical error in the individual points, $(\cdot)$. In fact, our data admit good fits for a very wide range of values in $1/\hat{\nu}$. For instance, if we try to input the value of the exponent obtained in the droplet-like extrapolation of the Binder parameter, $\zeta \sim 0.12$ (see equation (43) and table 2), we still obtain a good fit, even though the extrapolated value for $q_{\rm EA}$ is almost zero at $T = 0.703$ and negative at $T = 0.805$. Therefore, using the droplet exponent ζ the spin-glass phase would be non-existent.

Also included in table 4 is the confidence interval for this observable computed from non-equilibrium considerations in [25]. Notice that the equilibrium values are much more precise, but consistent. The extrapolations included in this table (and analogous ones for other values of T) are plotted on figure 12.

We remark that the estimate of $q_{\rm EA}$ from equation (47) is fully compatible with the results of a finite size scaling analysis of the conditional correlation functions [26].

Table 4. Extrapolation to infinite volume of $q_{\rm EA}(L,T)$ using the replicon exponent, equation (47). We also include the confidence interval previously obtained in a non-equilibrium study [25].

L	$T = 0.703$	$T = 0.805$
8	0.82461(83)	0.7818(11)
12	0.79333(85)	0.7412(11)
16	0.77300(75)	0.71681(95)
24	0.74027(71)	0.67905(83)
32	0.7174(14)	0.6535(16)
$L_{\rm min}$	16	16
χ^2/d.o.f.	1.83/1	0.98/1
$q_{\rm EA}$	0.538[11](6)	0.447[12](6)
Bounds from [25]	$0.474 \leq q_{\rm EA} \leq 0.637$	$0.368 \leq q_{\rm EA} \leq 0.556$

Table 5. Determination of $L_{\rm c}$ in equation (47) for several temperatures below $T_{\rm c}$. Errors are given as in table 4. The characteristic length $L_{\rm c}(T)$ scales as a correlation length when T approaches $T_{\rm c}$ ($\nu \approx 2.45$ from [32]). We warn the reader that the χ^2/d.o.f. for the fits at $T = 0.85$ and 0.90 are, respectively, 2.6/1 and 2.7/1.

T	$L_{\rm c}^{1/\hat{\nu}}$	$L_{\rm c}$	$L_{\rm c}(T_{\rm c}-T)^{\nu}$
0.703	1.253[10](32)	1.78[4](11)	0.197[4](13)
0.75	1.448[12](34)	2.58[6](16)	0.210[4](13)
0.805	1.731[14](44)	4.08[9](27)	0.221[5](15)
0.85	2.023[16](54)	6.09[13](42)	0.222[5](15)
0.90	2.514[21](66)	10.63[22](71)	0.230[5](15)

Interestingly enough the estimate of $q_{\rm EA}$ provides a determination of the correlation length in the spin-glass phase. The reader might be surprised that a correlation length can be defined in a phase where correlations decay algebraically. Actually, finite size effects are ruled by a crossover length $L_{\rm c}(T)$ [67], that scales as a correlation length (i.e. $L_{\rm c}(T) \propto (T_{\rm c}-T)^{-\nu}$). In fact, one would expect $q_{\rm EA}(T,L)/q_{\rm EA}(T) = 1 + h[L/L_{\rm c}(T)]$. The only thing we know about the crossover function is that it behaves for large x as $h(x) \sim x^{-1/\hat{\nu}}$. Making the simplest ansatz $h(x) = x^{-1/\hat{\nu}}$, the amplitude for the finite size corrections in equation (47) can be interpreted as a power of the crossover length $L_{\rm c}(T)$, table 5. We note that our determination of $L_{\rm c}(T)$ really scales as a bulk correlation length, with $T_{\rm c}$ and ν taken from [32]. It turns out to be remarkably small at $T = 0.703$.

The above argument tells us that good determinations of $q_{\rm EA}(T)$ are possible, provided that $L \gg L_{\rm c}(T)$. Yet, finite size scaling can be used as well to extrapolate $q_{\rm EA}(T,L)$ to the large volume limit, even closer to $T_{\rm c}$ where L becomes *smaller* than $L_{\rm c}$. This somehow unconventional use of finite size scaling was started in [68]–[71], and has also been used in the spin-glass context [4, 72]. Most of the time, these ideas are used in the paramagnetic phase, but we show below how to implement them in the low temperature phase.

Close to $T_{\rm c}$, we know that

$$q_{\rm EA}^{\infty}(T) = \lambda(T_{\rm c}-T)^{\beta}[1 + \mu(T_{\rm c}-T)^{\omega\nu} + \cdots]. \tag{48}$$

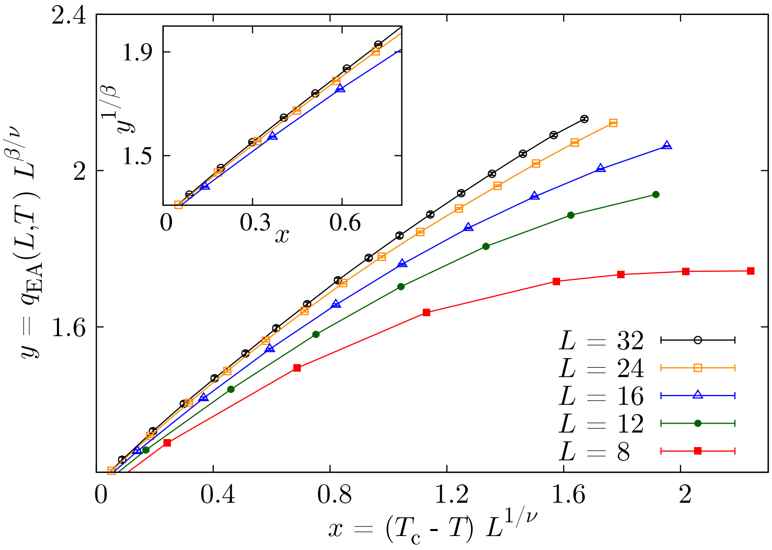

Figure 13. Scaling plot of $y = q_{\mathrm{EA}}(L,T)L^{\beta/\nu}$ in the critical region below T_{c}, following equation (49) and using the values given in [32] for the critical exponents and T_{c}. Inset: close-up of the region near T_{c} in the representation of equation (52), showing a linear behaviour for large L.

We have excellent determinations of T_{c} and β from the work in [32], so we need only to estimate the amplitude λ. In fact, Wegner's confluent corrections $(T_{\mathrm{c}} - T)^{\omega\nu}$ are small close to T_{c}. To proceed, we note that finite size scaling tells us that

$$q_{\mathrm{EA}}(L,T) = L^{-\beta/\nu}F(x)[1 + L^{-\omega}G(x) + \cdots], \qquad x = L^{1/\nu}(T_{\mathrm{c}} - T), \tag{49}$$

where the critical exponents are (from [32]),

$$\nu = 2.45(15), \qquad \beta = 0.77(5), \qquad \omega = 1.0(1). \tag{50}$$

In order to connect equation (49) with the infinite volume limit in equation (48) the asymptotic behaviour of the scaling functions $F(x)$ and $G(x)$ must be for large x

$$F(x) \sim x^{\beta}, \qquad G(x) \sim x^{\omega\nu}. \tag{51}$$

The resulting scaling plot is represented on figure 13. Varying the values of T_{c} and the critical exponents inside their error margins does not make significant changes in the plot. Notice how the curves collapse for small values of the scaling variable x and large L, but how for our lowest temperatures scaling corrections become important. In fact, equation (49) implies that when the temperature is lowered away from T_{c} the amplitude for scaling corrections grows as $x^{\omega\nu} \approx x^{2.45}$.

In order to estimate the amplitude λ we shall concentrate on the small-x region where finite size scaling corrections are smallest. Disregarding scaling corrections in equation (49),

$$(q_{\mathrm{EA}}(L,T)L^{\beta/\nu})^{1/\beta} = F(x)^{1/\beta} \xrightarrow[x\to\infty]{} x. \tag{52}$$

The inset of figure 13 shows that we reach this asymptotic behaviour for $L \geq 24$. Then, using the simplest parameterization, $F(x) = (\lambda^{1/\beta} x + B)^\beta$,

$$q_{\mathrm{EA}}(L,T) = \lambda (T_{\mathrm{c}} - T)^\beta \left[1 + \frac{\beta B}{\lambda^{1/\beta} (T_{\mathrm{c}} - T) L^{1/\nu}} + \cdots \right]. \tag{53}$$

We can fit our $L = 32$ data for $x < 0.4$ (where the curves for $L = 24$ and 32 are compatible) and use the resulting value of λ to extrapolate in equation (53) to infinite volume. This extrapolation is represented as a function of T on figure 12. It is clear that this critical extrapolation differs with the extrapolation from (47) at most by two standard deviations. The difference, if any, could be explained as Wegner's confluent corrections. However, to make any strong claim on confluent corrections, one would need to estimate the error in the critical extrapolation. Unfortunately, we have found that this error estimate is quite sensitive to the statistical correlation between T_{c}, ν, and β (as far as we know, the corresponding covariance matrix has not been published).

One could be tempted to compare equation (53) with (47) and conclude $\hat{\nu} = \nu$. We observe that, at the numerical level, $\nu = 2.45(15)$ [32] and $\hat{\nu} = 2.6(3)$ [26]. However, we do not regard this as fireproof. Indeed, it is a consequence of our somewhat arbitrary parameterization $F(x) = (\lambda^{1/\beta} x + B)^\beta$. To investigate this issue further, the small-x region is not enough. One is interested in the asymptotic behaviour of $F(x)$ for large x, where unfortunately corrections to scaling are crucial. A careful study of the crossover region can be done only by considering corrections to scaling both at the critical temperature (at $q = 0$) and below the critical temperature (at $q = q_{\mathrm{EA}}$).

Finally, the reader could worry about the applicability of (48) well below T_{c}. The issue has been considered recently within the framework of droplet theory [73]. It was found that (48) is adequate for all $T < T_{\mathrm{c}}$ (actually, no Wegner's scaling corrections were discussed in [73]). Thus, the fact that our data are describable as scaling behaviour with leading Wegner's correction does not imply that they are not representative of the low temperature phase.

5. Conditional correlation functions

Let us consider the conditional spatial correlation function $C_4(r|q)$, equation (14). A thorough study in the Fourier space is performed in [26]. Here, we provide some complementary information, concentrating on real space and considering as well the statistical fluctuations on the correlators.

We first concentrate on $q = 0$, the region where the droplet and RSB theory most differ. In figure 14—left we show $C_4(r|q = 0)$ for $T = 0.703$, which is seen to tend to zero for large r. Furthermore, if we use the droplet scaling of equation (20), we see that we need to rescale the correlation function by a factor $L^{\theta(0)}$, with $\theta(0) = 0.38(2)$ the replicon exponent, in order to collapse the curves.

As for other values of q, we may consider the differences

$$C_4(r = L/4|q) - C_4(r = L/2|q) \sim \frac{1}{L^{\theta(q)}}, \tag{54}$$

where the subtraction takes care of the large-r background in $C_4(r|q)$. As we show in figure 15, the subtracted correlation function scales in the range $q^2 < 0.2$ as $L^{-\theta(0)}$. This

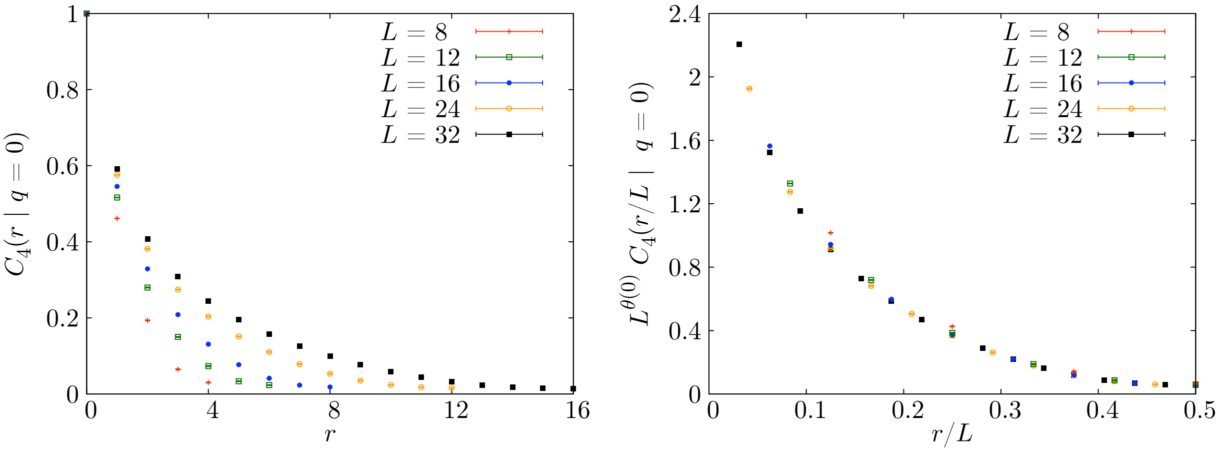

Figure 14. Spatial correlation function $C_4((r,0,0)|q=0)$ at $T = 0.703$. We show on the right panel a rescaled version using the replicon exponent $\theta(0) = 0.38(2)$ and the scaling variable r/L.

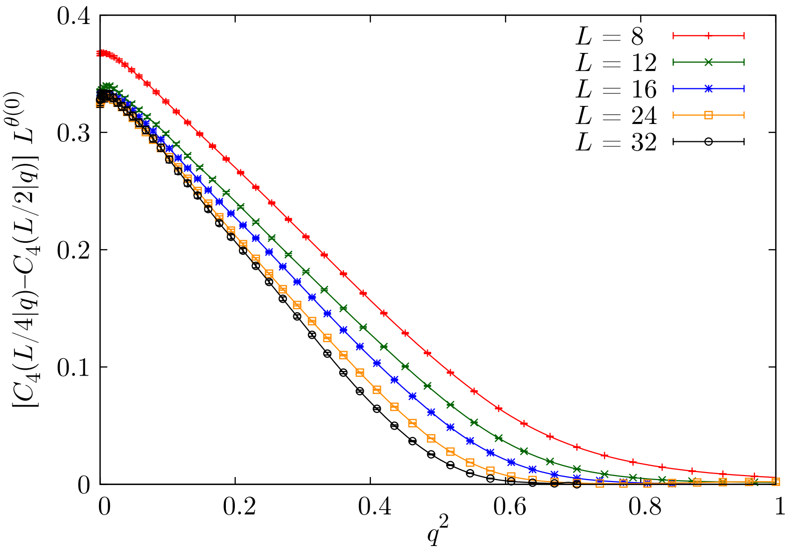

Figure 15. Subtracted correlation function, equation (54), in units of $1/L^{\theta(0)}$ as function of q^2. We took the non-equilibrium determination of the replicon exponent, $\theta(0) = 0.38(2)$ [25].

implies that the connected correlation functions $C_4(r|q) - q^2$ decay algebraically for large r (a similar conclusion was reached in [19]). On the other hand, for $q^2 = q_{\mathrm{EA}}^2 \approx 0.3$, the exponent $\theta(q)$ is definitively larger than $\theta(0)$ (a detailed analysis indicates $\theta(q_{\mathrm{EA}}) \sim 0.6$ [26]). The crossover from the scaling $C_4(r = L/4|q) - C_4(r = L/2|q) \sim 1/L^{\theta(0)}$ to $C_4(r = L/4|q) - C_4(r = L/2|q) \sim 1/L^{\theta(q_{\mathrm{EA}})}$ can be described by means of finite size scaling [26].

Recalling that $\overline{\langle Q_{\mathrm{link}}\rangle} = C_4(r = 1)$, we can consider the spatial correlation as a sort of generalization of the link overlap. In this sense it is worth recalling that in a mean-field setting fixing q^2 also fixes Q_{link}. In a three-dimensional RSB system one would,

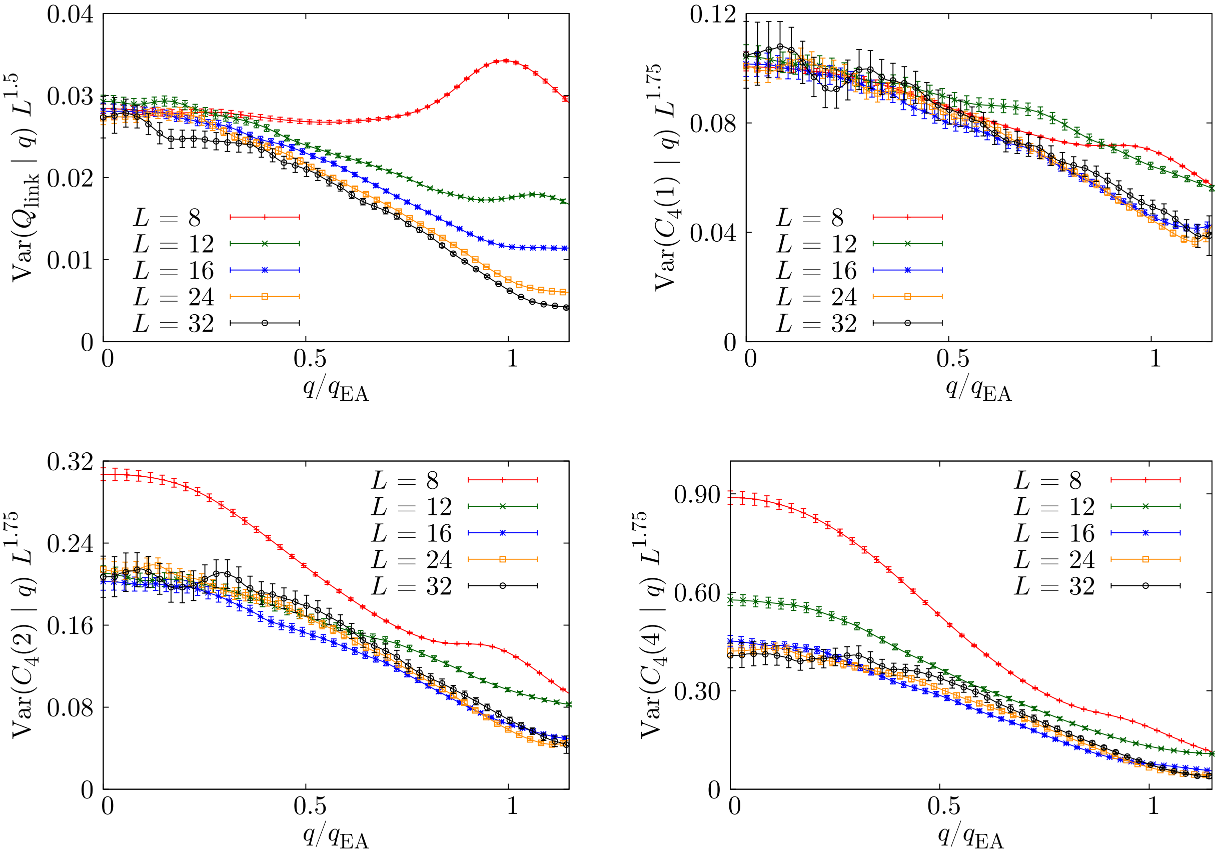

Figure 16. Plots of the conditional variance at fixed q of Q_{link} and $C_4(r)$ at $T = 0.703$, rescaled by appropriate powers of L (we chose exponents that provided a good scaling at $q = 0$). The abscissas correspond to q in units of $q_{\mathrm{EA}}(L, T = 0.703)$.

therefore, expect the conditional variance $\mathrm{Var}(Q_{\mathrm{link}}|q)$, equation (11), to tend to zero for large lattices [17]. The first panel of figure 16 demonstrates that this is the case in our simulations, where we find that $\mathrm{Var}(Q_{\mathrm{link}}|q) \sim L^{-D/2}$. We can extend this result to $r > 1$ by considering the conditional variances of C_4. Notice that, unlike Q_{link}, C_4 is already defined as an averaged quantity in equation (13) and not as a stochastic variable, so speaking of its variance is either trivial or an abuse of language. However, to avoid clutter, we have maintained the notation $\mathrm{Var}(C_4(r)|q)$, as its intended meaning is clear. These are plotted in figure 16, where we see that they decrease even faster than $\mathrm{Var}(Q_{\mathrm{link}}|q)$, with a power of L that does not seem to depend on r.

6. Non-equilibrium versus equilibrium

In [24], we suggested the existence of a time-length dictionary, relating results in the thermodynamical limit at finite time t_{w} with equilibrium results for finite size L. The matching for $T = 0.7$ was $L \approx 3.7\xi(t_{\mathrm{w}})$, where $\xi(t_{\mathrm{w}})$ is the coherence length at time t_{w}. The comparison there was restricted to $L \leq 20$. The expectation value $\mathrm{E}(Q_{\mathrm{link}}|q)$ was confronted with the correlation function $C_{2+2}(r = 1)$, recall the definitions in section 2.4.

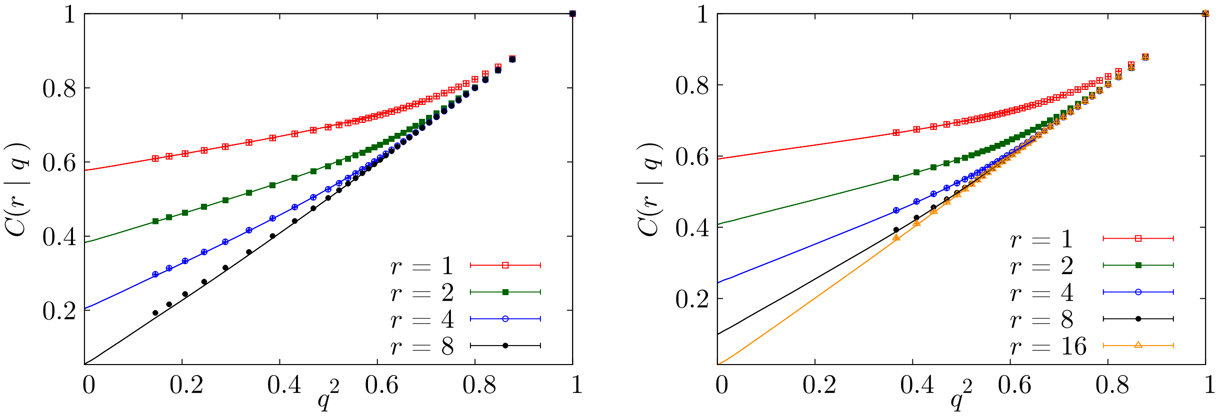

Figure 17. Equilibrium $C_4(r|q)$ as a function of q^2 (lines) for $L = 24$ (left) and $L = 32$ (right) lattices at $T = 0.703$. We compare with non-equilibrium data from [25] (points) of $C_{2+2}(r, t, t_\mathrm{w})$ as a function of $C^2(t, t_\mathrm{w})$ for $t_\mathrm{w} = 2^{26}$ (left) and $t_\mathrm{w} = 2^{31}$ (right), (see section 2.4 for definitions). The errors in both sets of data are comparable, and smaller than the point size.

We also predicted that the equilibrium data for $L = 33$ would match our non-equilibrium results for $t_\mathrm{w} = 2^{32}$. Using the same time-length dictionary our $L = 32$ simulations would correspond to $t_\mathrm{w} \approx 2^{31}$ and those for $L = 24$ would correspond to $t_\mathrm{w} \approx 2^{26}$.

Now, recalling that $\overline{\langle Q_\mathrm{link}\rangle}$ is merely $C_4(r = 1)$, it is natural to extend this correspondence between $C_4(r)$ and $C_{2+2}(r)$ to $r > 1$. Of course, care must be exercised because C_4 in a finite lattice cannot be computed beyond $r = L/2$, while C_{2+2} is defined for arbitrary r. However, the matching is very accurate, even for r dangerously close to $L/2$, see figure 17. It is interesting to point out that the off-equilibrium results of [24] and our equilibrium simulations have similar precision, even though the latter required about twenty times more computation time on Janus, not to mention a much more complicated simulation protocol. In this sense we arrive at the conclusion that simulating the dynamics may be the best way to obtain certain equilibrium quantities. On the other hand, only the equilibrium simulations give access to the crucial $C(t, t_\mathrm{w}) = 0$ physics.

We may now wonder about the experimentally relevant scale of 1 h ($t_\mathrm{w} \sim 3.6 \times 10^{15}$, taking one MC step as one picosecond [1]). Assuming a power-law behaviour, $\xi(t_\mathrm{w}) = At_\mathrm{w}^{1/z(T)}$, with $z(0.64T_\mathrm{c}) = 11.64(15)$ [25], we conclude that the correspondence is 1 h $\longleftrightarrow$ $L \approx 110$. Note, see for instance figure 11, that $L = 110$ is close enough to $L = 32$ to allow a safe extrapolation.

Let us finally stress that the modified droplet scaling for $\xi(t_\mathrm{w})$ [74] would predict that 1 h of physical time would correspond to equilibrium data on L even smaller than 110. Indeed, according to these authors the time needed to reach some coherence length $\xi(t_w)$ grows as

$$t_\mathrm{w} \sim \tau_0 \xi^{z_\mathrm{c}} \exp\left(\frac{Y(T)\xi^\psi}{T}\right), \tag{55}$$

where τ_0 is the microscopical time associated to the dynamics; z_c is the dynamical critical exponent computed at the critical point; ψ is the exponent that takes the free energy barriers into account (from the dynamical point of view) and $Y(T) = Y_0(1 - T/T_\mathrm{c})^{\psi\nu}$, with the ν exponent being the static critical exponent linked to the coherence length. Near the critical point $Y(T) \to 0$ and the power-law critical dynamics is recovered. On the other hand, if we stay below T_c, equation (55) predicts an algebraic grow of t_w with $\xi(t_\mathrm{w})$ only for very small coherence lengths. However, as the coherence length grows, the time needed to reach it diverges exponentially on $\xi(t_\mathrm{w})$.

7. The link overlap

We shall address here three separated problems: overlap equivalence (section 7.1), replica equivalence (section 7.2), and the scaling of the link susceptibility (section 7.3).

7.1. Overlap equivalence

As we have discussed previously, it has been proposed [41, 17] that attention should be shifted from the spin overlap (the primary object for mean-field systems) to the link overlap (the would-be primary object below the upper critical dimension). Two requirements should be met for this change of variable to be feasible:

(i) The conditional variance $\mathrm{Var}(Q_\mathrm{link}|q)$ must vanish in the large L limit.
(ii) The conditional expectation $\mathrm{E}(Q_\mathrm{link}|q)$ should be a strictly increasing function of q^2.

The scaling with L of $\mathrm{Var}(Q_\mathrm{link}|q)$, section 5, does suggest that the first requirement holds. We shall investigate here the second requirement. We remark that the RSB theory expects it to hold, while droplet expects it not to. Furthermore, this point is actually the only disagreement between the RSB and the TNT picture. In fact, RSB expects the derivative $\mathrm{dE}(Q_\mathrm{link}|q)/\mathrm{d}q^2$ never to vanish. On the other hand, TNT supporters expect this derivative to scale as L^{D_s-D}, where D_s represents the (would be) fractal dimension of the surface of the spin-glass domains. In $D=3$, $D-D_s \approx 0.44$ [16].

To estimate the derivative $\mathrm{dE}(Q_\mathrm{link}|q)/\mathrm{d}q^2$, we observe that $E(Q_\mathrm{link}|q)$ is an extremely smooth function of q^2 (see the $r = 1$ curves in figure 17). Hence we can attempt a polynomial fit:

$$\mathrm{E}(Q_\mathrm{link}|q) - \mathrm{E}(Q_\mathrm{link}|q=0) = \sum_{k=1}^{m} c_{2k}^{(2m)} q^{2k}. \tag{56}$$

In particular, the coefficient $c_2^{(2m)}$ provides an estimate of $\mathrm{d}E(Q_\mathrm{link}|q)/\mathrm{d}q^2$ at $q^2 = 0$. Playing with the order $2m$ of the polynomials, one can control systematic errors. Note that it is very important to fit the *difference* $\mathrm{E}(Q_\mathrm{link}|q) - \mathrm{E}(Q_\mathrm{link}|q=0)$, which, due to statistical correlations, has much reduced statistical errors. On the other hand, data for different q are so strongly correlated that standard fitting techniques are inappropriate. We thus used the approach explained in [25]. The results, see table 6, indicate that $c_2^{(4)}$ offers a reasonable compromise between systematic and statistical errors.

Once we have the derivatives in our hands, we may try to extrapolate them to large L by means of an RSB fit ($A/L+b$, middle part of table 6) or using a TNT fit ($B/L^{0.44}+d$, bottom part of table 6). The two functional forms produce a reasonable fit. As expected, the $1/L$ extrapolation to $L=\infty$ yields a non-vanishing derivative, while the $1/L^{0.44}$ extrapolation suggests that, for large L, $\mathrm{E}(Q_\mathrm{link}|q)$ is constant as q^2 varies. We remark

Table 6. Coefficients $c_2^{(2m)}$ in the fit to equation (56), for various orders of the fitting polynomial, and $T = 0.703$ and 0.625. This coefficient is interpreted as $[\mathrm{dE}(Q_{\mathrm{link}}|q)/\mathrm{d}q^2]_{q^2=0}$. We report as well the results for fits of the form $c_2^{(4)} = A/L + c$ (centre) and $c_2^{(4)} = B/L^{0.44} + d$ (bottom). For both fits, we also provide the extrapolation to $L = 110$ which, according to the time-length dictionary, corresponds to the experimentally relevant length scale.

	$T = 0.703$			$T = 0.625$		
L	$c_2^{(2)}$	$c_2^{(4)}$	$c_2^{(6)}$	$c_2^{(2)}$	$c_2^{(4)}$	$c_2^{(6)}$
8	0.403(5)	0.405(16)	0.43(3)	0.414(7)	0.423(19)	0.45(4)
12	0.317(5)	0.321(14)	0.35(3)	0.331(6)	0.335(18)	0.36(3)
16	0.271(4)	0.262(11)	0.26(2)	0.282(6)	0.275(16)	0.28(3)
24	0.224(5)	0.222(15)	0.22(3)	0.231(5)	0.220(14)	0.22(3)
32	0.199(6)	0.201(18)	0.20(4)	—	—	—
χ^2/d.o.f.		0.57/3			0.46/2	
A		2.23(21)			2.46(27)	
c		0.129(16)			0.121(21)	
$L = 110$		0.149(14)			0.143(19)	
χ^2/d.o.f.		2.39/3			0.18/2	
B		1.45(11)			1.32(15)	
d		−0.06(3)			−0.11(5)	
$L = 110$		0.082(20)			0.058(28)	

as well that the very same conclusion was reached in the analysis of the non-equilibrium temporal correlation functions [25].

However, we have far more accurate data at our disposal than the derivative $\mathrm{d}E(Q_{\mathrm{link}}|q)/\mathrm{d}q^2$ at $q^2 = 0$, namely the correlation functions themselves. In table 7 we give our estimates for $C(r = 1|q = 0)$ and $C(r = 1|q = 0.523 \approx q_{\mathrm{EA}})$. According to a TNT picture of the SG phase, the two correlation functions should be equal. As the reader can check, an infinite volume extrapolation as $L^{-0.44}$ is unbearable for both correlation functions (even if we discard the two smallest sizes). The same conclusions hold substituting L by

$$\ell = \pi/\sin(\pi/L), \tag{57}$$

which is more natural for lattice systems. Yet, it could be argued that our data are preasymptotic. Hence, we may try a TNT extrapolation including scaling corrections.

$$C(r = 1|q) = C_\infty + A_q L^{-0.44}(1 + B_q L^{-y}). \tag{58}$$

We have performed a joint fit of the data in table 7 to equation (58). The fitting parameters were the four amplitudes $A_0, B_0, A_{0.523}$ and $B_{0.523}$, the common scaling corrections exponent y and the common large-L extrapolation C_∞. We take into account the (almost negligible) correlation in data for the same L by computing χ^2 with the covariance matrix, which can be reconstructed from the data in table 7. The result is (notice the highly asymmetric errors)

$$C_\infty = 0.677^{+0.012}_{-0.005}, \qquad y = 0.57^{+0.26}_{-0.08}, \qquad \chi^2/\mathrm{d.o.f.} = 9.1/4. \tag{59}$$

Table 7. $C(r=1|q)$ for $q=0$ and q_{EA} for all our system sizes at $T=0.703$. For each L, we include the correlation coefficient between both values of q. Specifically, for two quantities A and B, $\mathcal{R}_{AB} = \overline{(\langle A\rangle - \overline{\langle A\rangle})(\langle B\rangle - \overline{\langle B\rangle})}/\sqrt{\overline{(\langle A\rangle - \overline{\langle A\rangle})^2(\langle B\rangle - \overline{\langle B\rangle})^2}}$.

L	$C(1\|0)$	$C(1\|q_{\mathrm{EA}})$	$\mathcal{R}$
8	0.461 38(82)	0.572 53(33)	0.134
12	0.516 49(71)	0.603 90(28)	0.051
16	0.545 52(60)	0.620 89(22)	0.060
24	0.575 73(77)	0.637 42(17)	−0.119
32	0.591 31(94)	0.645 79(24)	0.063

Were the functional form in equation (58) correct, the probability of χ^2 being even larger than we found would be only 6%.

On the other hand, in an RSB setting, one would expect $C(r=1|q)$ to scale as $1/L$, with a q-dependent infinite volume value $C_\infty(q)$. Indeed, if we fit the data in table 7 to $C(1|q) = C_\infty(q) + A/\ell$ we obtain

$$C_\infty(q=0) = 0.6349(8), \qquad \chi^2/\mathrm{d.o.f.} = 3.63/3, \tag{60}$$

$$C_\infty(q=q_{\mathrm{EA}}) = 0.6711(2), \qquad \chi^2/\mathrm{d.o.f.} = 2.86/3. \tag{61}$$

We note as well that $[C_\infty(q=q_{\mathrm{EA}}) - C_\infty(q=0)]/q_{\mathrm{EA}}^2 \approx 0.132$, in fair agreement with the $1/L$ extrapolation for the derivative in table 6.

However, more important than the extrapolation to $L=\infty$ is the extrapolation to $L=110$, the length scale that, for $T=0.7$, matches the experimental timescales. For $T=0.625$, $L=110$ is surely larger than the relevant length scale but, unfortunately, the time-length dictionary at such a low temperature still needs to be tuned. As it can be seen in the middle and bottom parts of table 6, the two extrapolations yield a non-vanishing derivative.

Thus, whichever the standpoint adopted, the conclusion is identical for RSB and TNT theories: at the experimentally relevant length scales, overlap equivalence can be assumed.

7.2. Replica equivalence

We consider now the ratio

$$R_{\mathrm{link}} = \frac{\overline{\langle Q_{\mathrm{link}}^2\rangle - \langle Q_{\mathrm{link}}\rangle^2}}{\overline{\langle Q_{\mathrm{link}}^2\rangle} - \overline{\langle Q_{\mathrm{link}}\rangle}^2}, \tag{62}$$

defined in section 2.5. As was explained there, the RSB theory expects it to reach a constant value 2/3 below T_c, whereas the droplet and TNT theories lack a definite prediction. Our numerical data fit very well the RSB expectation (see figure 18—left).

Besides, we can also study a similar ratio, in which the mean-field substitution $Q_{\mathrm{link}} \rightarrow q^2$ is performed:

$$R_{q^2} = \frac{\overline{\langle q^4\rangle - \langle q^2\rangle^2}}{\overline{\langle q^4\rangle} - \overline{\langle q^2\rangle}^2}. \tag{63}$$

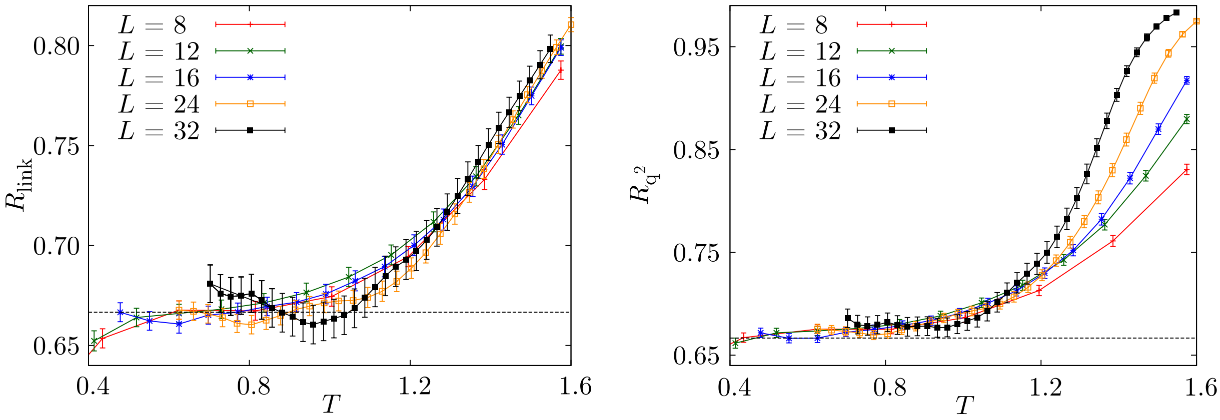

Figure 18. The ratios $R_{\rm link}$, equation (31), (left panel) and $R_{\rm q^2}$, equation (63), (right panel) versus T for the different system sizes. The replica equivalence property implies that, in an RSB system below $T_{\rm c}$, $R^{\rm link} = 2/3$ in the large-L limit. Recall that $T_{\rm c} \approx 1.1$.

Overlap equivalence suggests that R_{q^2} approaches 2/3 in the large L limit (again neither the droplet nor the TNT theories have a definite prediction). Our data at low temperatures seem compatible with the 2/3 expectation, see figure 18—right. On the other hand, the convergence to the thermodynamic limit seems fairly slower close to $T_{\rm c}$. We recall that a previous computation also concluded that violations of $R_{q^2} = 2/3$ are due to critical fluctuations [9].

7.3. Link susceptibility

We show in figure 19—left the link susceptibility $\chi_{\rm link}$, equation (27), as a function of temperature for different lattice sizes. It is clear enough that this susceptibility is divergent in the spin-glass phase and that the lower the temperature, the more violent the divergence. Hence, it is clear that this particular effect is not due to critical fluctuations.

We perform a more quantitative study in figure 19—right. As discussed in section 2.5, according to RSB theory, one would expect $\chi_{\rm link} \sim L^D$ in the SG phase.

We find evidence of a critical divergence. At and above $T_{\rm c}$, our data grow very softly with L (at $T = 1.3 \approx 1.17 T_{\rm c}$, data seem to reach a limiting value). However, below $T_{\rm c}$, we observe an effective exponent that grows when we lower the temperature. We observe that the effective exponent, for our lattice sizes and temperatures, has already grown beyond the Chayes bound of $D/2$ but still has not reached the RSB expectation of D. Note that no existing theory of the spin-glass phase can accommodate a temperature-dependent exponent. Therefore, the most economic scenario is that our lattice sizes are not large enough, so we are still in a preasymptotic regime for this quantity.

Let us take a slightly different point of view. Rigorous theorems discussed in section 2.5 tell us that, equation (30), if

$$\lim_{L\to\infty} \chi_{\rm link}/L^D > 0, \tag{64}$$

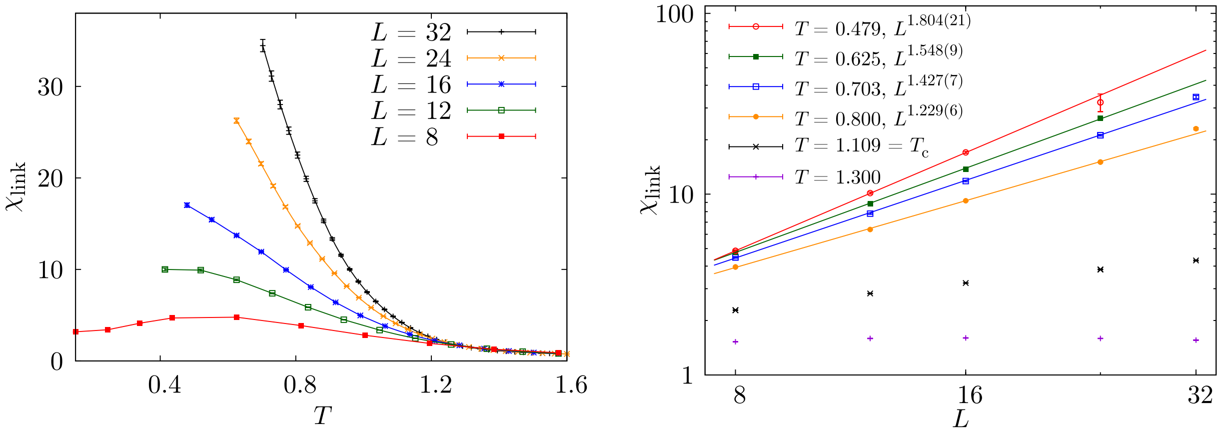

Figure 19. (Left) Susceptibility χ_{link} versus temperature for the different system sizes. (Right) Behaviour of χ_{link} with L for different temperatures. Lines are power-law fits. The effective exponents found in each of these fits is reported in the legends.

also the width $\sigma_{Q_{\text{link}}}$ of the probability density function for Q_{link},

$$\sigma^2_{Q_{\text{link}}} = \overline{\langle Q^2_{\text{link}} \rangle} - \overline{\langle Q_{\text{link}} \rangle}^2, \tag{65}$$

will be non-vanishing in the thermodynamic limit (see also section 7.2). It is very important that the converse statement also holds.

Now, using the identity (12), we can split this variance into two different contributions:

$$\sigma^2_{Q_{\text{link}}} = \int_{-\infty}^{\infty} \mathrm{d}q\, P(q) \left(\mathrm{Var}(Q_{\text{link}}|q) + \left[\mathrm{E}\left(Q_{\text{link}}|q\right) - \overline{\langle Q_{\text{link}} \rangle} \right]^2 \right). \tag{66}$$

Since $\mathrm{Var}(Q_{\text{link}}|q)$ scales as $L^{-D/2}$ (see [17] and figure 16), only the second term may survive the large L limit.

This suggests the definition of a modified link susceptibility:

$$\hat{\chi}_{\text{link}} = \frac{L^D \int_{-\infty}^{\infty} \mathrm{d}q\, P(q) \left[\mathrm{E}\left(Q_{\text{link}}|q\right) - \overline{\langle Q_{\text{link}} \rangle} \right]^2}{\overline{\langle q^4 \rangle} - \overline{\langle q^2 \rangle}^2}. \tag{67}$$

According to RSB theory, $\hat{\chi}_{\text{link}}$ should scale as L^D whereas it would not diverge as violently in a droplet or TNT scenario. The rationale for dividing out the $\overline{\langle q^4 \rangle} - \overline{\langle q^2 \rangle}^2$ can be found in equation (56). Assuming that the lowest-order polynomial is adequate, one finds that (of course, the particular value of the index m should be immaterial)

$$\hat{\chi}_{\text{link}} \approx L^D [c_2^{(2m)}]^2. \tag{68}$$

Hence, the TNT theory would expect χ_{link}/L^D to tend to zero, just because it predicts that in the large-L limit $c_2^{(2m)} = 0$. Note that the droplet theory would predict a vanishing χ_{link}/L^D for a different reason, namely because they expect that $\overline{\langle q^4 \rangle} - \overline{\langle q^2 \rangle}^2$ should vanish.

Table 8. Ratio $S^{(2m)}_{\mathrm{link}}$, equation (69), for all our lattice sizes at $T = 0.625, 0.703$, using the coefficients from table 6.

	$T = 0.703$		$T = 0.625$	
L	$S^{(2)}_{\mathrm{link}}$	$S^{(4)}_{\mathrm{link}}$	$S^{(2)}_{\mathrm{link}}$	$S^{(4)}_{\mathrm{link}}$
8	0.838(21)	0.846(67)	0.859(29)	0.897(81)
12	0.777(25)	0.797(70)	0.801(29)	0.821(88)
16	0.755(22)	0.706(59)	0.766(33)	0.729(85)
24	0.776(35)	0.76(10)	0.745(32)	0.675(86)
32	0.816(49)	0.83(15)	—	—

Table 9. χ_{link} in the 2D Ising model ($T_{\mathrm{c}} = 2/\log(1+\sqrt{2}) \approx 2.269\,185\,31\cdots$).

L	$T = 0.992T_{\mathrm{c}}$	$T = 0.986T_{\mathrm{c}}$
8	7.16(1)	7.091(8)
12	8.776(6)	8.614(15)
16	10.247(12)	9.868(9)
24	11.47(2)	10.51(2)
32	12.058(15)	10.639(12)

Let us check to what extent the estimate (68) is accurate. We show in table 8 the ratios

$$S^{(2)}_{\mathrm{link}} = \frac{L^D[c_2^{(2)}]^2}{\hat{\chi}_{\mathrm{link}}}, \qquad S^{(4)}_{\mathrm{link}} = \frac{L^D[c_2^{(4)}]^2}{\hat{\chi}_{\mathrm{link}}}. \tag{69}$$

Referring again to equation (56), it is clear that the contribution linear in q^2 explains a large fraction of $\hat{\chi}_{\mathrm{link}}$, and that this fraction is not likely to vanish in the large-L limit.

Hence the question of whether χ_{link} diverges as L^D or not, turns out to be strictly equivalent to that of overlap equivalence that we discussed at length in section 7.1. Our interpretation is that the effective scaling in figure 19—right is mostly due to strong finite size effects in $c_2^{(2m)}$. In this light, the effective exponents reported in figure 19—right are preasymptotic. In fact, the ratio A/c is large ($c_2^{(2m)}(L) = c + A/L$, see table 6), which tells us that for χ_{link} and related quantities, finite volume corrections are particularly large and naive power-law fits may give wrong results.

Let us conclude this section by checking how these quantities behave in a 2D Ising ferromagnet (i.e. with no disorder built in). Although this model is clearly too simple, it is also true that, to the best of our knowledge, the quantities investigated here have not been looked at before. Hence, it is interesting to see what happens even in this simple case. We use two replicas to compute χ_{link}. Results for χ_{link} are presented in table 9 for two different temperatures below the critical temperature T_{c}. There we can see that χ_{link} approaches a limiting $\mathcal{O}(L^0)$ value when L grows. Furthermore, the limiting value decreases when lowering the temperature away from T_{c}. Hence, a divergent link susceptibility below T_{c} is something that should *not* be taken for granted.

8. Conclusions

We have obtained equilibrium configurations of the Ising spin glass ($D = 3$, ± 1 Edwards–Anderson model) on large lattices at low temperatures ($T = 0.64T_c$ for $L = 32$, $T = 0.56T_c$ for $L = 24$, and even lower temperatures for smaller systems, see table 1). This unprecedented computation has been made possible by the Janus computer. However, the parallel tempering had never before been put to such stress, and we have devoted a large effort to convince ourselves that thermalization was achieved. New thermalization tests were devised. Furthermore, a new simulation strategy had to be employed: the simulation time needs to be tailored sample by sample (for one cannot afford adopting worst-case parameters).

The main conclusion we draw is that the correspondence between equilibrium results and non-equilibrium dynamics (much easier to compare with experimental work), is deeper than anticipated. In fact, one can construct a time-length dictionary, such that equilibrium correlation functions on finite systems match non-equilibrium correlators at finite time (but infinite system size). The evidence for this correspondence consists of: (i) quantitative comparison of the spatial correlation functions and (ii) the analysis of overlap equivalence on equilibrium (this work) and non-equilibrium settings [25]. In addition, there is a remarkable coincidence between the replicon exponent obtained from equilibrium methods [26], and from non-equilibrium dynamics [24, 25].

The unavoidable consequence of this time-length correspondence is that the system size that is relevant for the experimental work (timescales of 1 h, say) at $T = 0.64T_c$ is not infinite, but $L = 110$. Note that this correspondence was obtained assuming a power-law growth with time of the spin-glass coherence length in experimental samples. Should the modified droplet scaling for $\xi(t_w)$ hold [74], the relevant equilibrium system size would be even smaller. It is obvious that extrapolating numerical data from $L = 32$ to 110 is far less demanding than extrapolating them to infinite size. All such extrapolations in this work (even those assuming droplet scaling) were conclusive. The only effective theory that is relevant at experimental timescales is replica symmetry breaking.

However, the question of whether RSB is only an effective theory in $D = 3$ or a fundamental one does not lack theoretical interest. We have attempted several extrapolations to infinite system size in this work, finding that droplet theory is ruled out, unless a change of regime arises for system sizes much larger than our reached $L = 32$. We remark that in section 4.3 we have numerically determined a crossover length that rules finite size effects. As expected for a large enough system, it scales with temperature as a *bulk* correlation length. However, on the basis of numerical data alone, one can never discard that new behaviour might appear for much larger system sizes, irrelevant for current experimental work.

We found three contradictions with droplet theory. First, in order to have a trivial Binder cumulant, finite size corrections had to be of order $\sim L^{-0.11}$. Such finite size corrections would imply a vanishing, or even negative, spin-glass order parameter q_{EA}. Second, according to droplet theory (see [11], page 139) finite size corrections $\sim L^{-0.11}$ imply that the connected spatial correlation function at $q = q_{\mathrm{EA}}$ decays as $1/r^{0.11}$. A direct estimate indicates that, at $q = q_{\mathrm{EA}}$, correlations decay as $1/r^{0.6}$ [26]. Third, the probability density function $P(q = 0)$ does not decrease with increasing system size (a similar conclusion was reached in [63, 75]).

J. Stat. Mech. (2010) P06026

Our analysis of overlap equivalence is compatible with the RSB picture, without invoking sophisticated finite size effects. On the other hand, the statistical likelihood for TNT theory, as formulated in [16], has been quantified to be 6%. In any case, TNT scaling predicts that for $L = 110$ the surface-to-volume ratio of the magnetic domains is still of order one (in agreement with RSB). In addition, we find that replica equivalence is consistent with the RSB picture (while TNT lacks a definite prediction). Furthermore, the link susceptibility, χ_{link}, is definitively divergent in the spin-glass phase (since the divergence is stronger the lower the temperature, its origin is obviously non-critical). We are aware of no argument in TNT theory implying the divergence of the link susceptibility. On the other hand, RSB theory does require a divergent χ_{link}. However, RSB demands a scaling $\chi_{\text{link}} \sim L^D$. Such a growth regime has still not been reached for our system sizes, although we have identified the origin of this preasymptotic behaviour.

A final lesson from the present numerical study is that careful non-equilibrium simulations [25] are almost as rewarding as the equilibrium work. Indeed, our previous non-equilibrium study [24, 25] reached a timescale that corresponds to the present equilibrium $L = 32$ simulation. Yet, the numerical effort to obtain the data in figure 17 has been larger by, roughly, a factor of 20 in the case of the equilibrium work. It is true that the equilibrium approach allows one to investigate directly the crucial $q = 0$ region, where in the non-equilibrium case one would need to rely on difficult extrapolations to infinite time. However, we do not think that there is much road ahead for equilibrium studies, due to the failure of the parallel tempering algorithm. Indeed, see table 1, it takes about 3.5 times more numerical work to equilibrate 1000 samples of $L = 32$ at $T = 0.64T_{\text{c}}$ than 4000 samples of $L = 24$ down to $T = 0.56T_{\text{c}}$. Clearly enough, the temperature window accessible with the parallel tempering algorithm decreases very fast as the system size grows. We believe this failure to be due to a genuine temperature-chaos effect. However, in order to analyse quantitatively the effect one needs to correlate the (sample-dependent) temperature bottlenecks, see figure 1—left, with the spin overlap at different temperatures. This analysis is left for future work [62].

Acknowledgments

We acknowledge support from MICINN, Spain, through research contracts No. TEC2007-64188, FIS2006-08533-C03, FIS2007-60977, FIS2009-12648-C03 and from UCM-Banco de Santander. BS and DY are FPU fellows (Spain) and RAB and JM-G are DGA fellows. SP-G was supported by FECYT (Spain). The authors would like to thank the Arénaire team, especially J Detrey and F de Dinechin for the VHDL code of the logarithm function [76]. M Moore posed interesting questions that helped us sharpen the discussion in section 4.3.

Appendix A. Our thermalization protocol

We have followed a three-step procedure to thermalize each sample:

(i) We simulate for a fixed minimum length of $N_{\text{HB}}^{\text{min}}$ MCS, chosen to be enough to thermalize most of the samples. Notice that most published parallel tempering simulations stop here, assessing the thermalization only through the time evolution of disorder-averaged observables.

(ii) We discard the first sixth of the measurements and compute the integrated autocorrelation time, choosing the self-consistent window W of (37) so that $W > 6\tau_{\mathrm{int}}$. Using this first estimate of the integrated time, we enlarge the simulation until $N_{\mathrm{HB}} > 22\tau_{\mathrm{int}}$, always discarding its first sixth. A criterion based on τ_{int} was first used in [27].

(iii) Now that we have a reasonably dimensioned simulation, we can compute the exponential autocorrelation time (which is typically bigger than, but of the same order of magnitude as, the integrated time). We demand that N_{HB} be larger than $12\tau_{\mathrm{exp}}$.

This last step is the main innovation of these simulations. Notice that we have to perform non-linear fits in some 10^4 autocorrelation functions, with a sample-dependent fitting range. This is a somewhat delicate procedure, so we have taken great care to ensure it is failsafe.

We start by assuming that the correlation function (38) can be approximated by the sum of two exponentials:

$$\hat{C}(t) \simeq A_1 \mathrm{e}^{-t/\tau_1} + A_2 \mathrm{e}^{-t/\tau_2}, \qquad \tau_1 = \tau_{\mathrm{exp}} > \tau_2. \tag{A.1}$$

Usually, one chooses the range for such a fit manually but this is not practical here, due to the sheer number of correlation functions we have to study. Hence, assuming that the exponential time is not much larger than the integrated one, we have used the latter in order to define our fitting range (notice that if $A_2 = 0$, $A_1 = 1$ and $\tau_1 = \tau_{\mathrm{int}}$). Our fitting procedure has three steps

(a) We perform a first fit to a single exponential in the range $[2\tau_{\mathrm{int}}, 3\tau_{\mathrm{int}}]$, from which we obtain an amplitude A and a time τ.

(b) Using $\tau_1 = \tau$, $\tau_2 = \tau/10$, $A_1 = A$ and $A_2 = 1 - A$ as a starting point we perform the non-linear fit to (A.1) with a Levenberg–Marquardt scheme [77]. The fitting range is chosen as $[\tau_{\mathrm{int}}/10, 10\tau_{\mathrm{int}}]$.

(c) Sometimes τ_2 is very small and $\hat{C}(t)$ is indistinguishable from a single exponential in $[\tau_{\mathrm{int}}/10, 10\tau_{\mathrm{int}}]$. In these occasions the fit in step 2 fails, which can be detected in a number of ways (very large or even negative values for one of the A_i, absurdly large values for τ_1 or even a complete breakdown of the iterative method). For these samples, a third fit to a single exponential is performed in the range $[5\tau_{\mathrm{int}}, 10\tau_{\mathrm{int}}]$. There is one exception: when one of the A_i is negative, there appears a very pronounced downwards fluctuation in $\hat{C}(t)$ for large times, which can lead to an underestimation of τ_{exp}. In these occasions, the third fit is performed in $[2.5\tau_{\mathrm{int}}, 5\tau_{\mathrm{int}}]$.

This automatic and fully quantitative procedure works for most samples, but there are some potential pitfalls which may lead to our underestimating τ_{exp}. Sometimes, the exponential time is much larger than τ_{int}. This can result in a failure of the automatic method for two reasons: (1) as $\tau_{\mathrm{exp}} \gg \tau_{\mathrm{int}}$, the fitting ranges are no longer well adjusted (2) a very large $\tau_{\mathrm{exp}}/\tau_{\mathrm{int}}$ implies a very low value for A_1. We address this problem by enlarging the measurement bins by a factor of 10 in case $\tau_{\mathrm{exp}} > 10\tau_{\mathrm{int}}$. This way, both τ_{int} and A_1 grow, and the fit works much better.

The possibility also exists that $\hat{C}(t)$ may be misleading, because the simulation is so much shorter than the exponential time that some of the configurations have not yet explored the relevant minima of the free energy (i.e., the $\hat{C}(t)$ we are measuring is not yet the *equilibrium* one). This happens when some of the $4N_T$ configurations have not crossed the critical temperature in the parallel tempering dynamics. The assumption here, key to the parallel tempering method, is that once a configuration spends a few MCS at high temperatures it becomes completely decorrelated (remember that, due to Janus' special characteristics, the interval between measurements is very large). To prevent this from happening, we measure the time t_{hot} that each configuration spends at temperatures greater than T_{c}. In case any of the $4N_T$ configurations has a value of t_{hot} smaller than one third of the median, we consider that the simulation is far too short for us to measure τ_{exp} and we simply double N_{HB}. Notice that this last criterion is unlike the others in that it is not completely quantitative. It simply detects that our starting point is very badly dimensioned.

As a final test, we have increased N_{min} by a factor of 10 for the first 1% of the samples in all lattices. None of the τ_{exp} estimates changed within errors.

Appendix B. Unbiased estimators of non-linear functions

Non-linear functions of thermal mean values, that are computed sample by sample and afterwards averaged over disorder, are prone to suffer systematic errors larger than the statistical ones. General cures for this problem are known [61, 32]. In our case, the only such quantity is χ_{link}, defined in equation (27). Since we have four replicas, the bias problem could be avoided for χ_{link}, see equation (B.6) below. However, if one decides instead to face it, a nice test for the statistical quality of the data is obtained.

Indeed, consider equation (27), and let $[Q_{\text{link}}]$ be our Monte Carlo estimate of $\langle Q_{\text{link}}\rangle$ as computed from N measurements for a given sample. The expectation value $\langle [Q_{\text{link}}]^2\rangle$ is *not*$\langle Q_{\text{link}}\rangle^2$. To quantify the effect, we need some notation [51, 59]. The normalized equilibrium autocorrelation function for Q_{link}, at a given temperature and for a given sample, is

$$\hat{C}_{Q_{\text{link}}}(t) = \frac{\langle (Q_{\text{link}}^{(s+t)} - \langle Q_{\text{link}}\rangle)(Q_{\text{link}}^{(s)} - \langle Q_{\text{link}}\rangle)\rangle}{\langle Q_{\text{link}}^2\rangle - \langle Q_{\text{link}}\rangle^2}, \tag{B.1}$$

where $Q_{\text{link}}^{(s)}$ stands for the value taken by Q_{link} at time s. Two characteristic timescales are relevant to us:

$$\tau_{\text{int},Q_{\text{link}}} = \frac{1}{2}\sum_{t=-\infty}^{t=+\infty} \hat{C}_{Q_{\text{link}}}(t), \qquad \tau_{\text{avg},Q_{\text{link}}} = \frac{\sum_{t=-\infty}^{t=+\infty} |t| \hat{C}_{Q_{\text{link}}}(t)}{\sum_{t=-\infty}^{t=+\infty} \hat{C}_{Q_{\text{link}}}(t)}. \tag{B.2}$$

Then a straightforward computation shows that

$$\langle [Q_{\text{link}}]^2\rangle = \langle Q_{\text{link}}\rangle^2 + \frac{2\tau_{\text{int},Q_{\text{link}}}[\langle Q_{\text{link}}^2\rangle - \langle Q_{\text{link}}\rangle^2]}{N}\left(1 - \frac{\tau_{\text{avg},Q_{\text{link}}}}{N}\right), \tag{B.3}$$

up to corrections of order $\mathcal{O}(\mathrm{e}^{-N/\tau_{\text{exp}}})$ (the exponential autocorrelation time τ_{exp} was discussed in section 3.3). This computation is performed in textbooks [51, 59] only to order $1/N$, and with a rather different aim: it provides an estimate of the (squared) statistical error in the Monte Carlo estimation of $\langle Q_{\text{link}}\rangle$. Our interest in equation (B.3) is different. It tells us that, when taking $[Q_{\text{link}}]^2$ as $\langle Q_{\text{link}}\rangle^2$ we are incurring in bias not only of order $1/N$, but also of order $1/N^2$.

It is easy to obtain bias-corrected estimators [61]: one divides the Monte Carlo history in two halves, four quarters, and eight eighths. Recall that we will be dropping in the analysis the full first half of the Monte Carlo history. Then, one computes $[Q_{\mathrm{link}}]^2_{2/2}$ from the last half of the data as well as $[Q_{\mathrm{link}}]^2_{3/4}$ and $[Q_{\mathrm{link}}]^2_{4/4}$ from the third and fourth quarters respectively. Similarly, we compute $[Q_{\mathrm{link}}]^2_{5/8}$, $[Q_{\mathrm{link}}]^2_{6/8}$, $[Q_{\mathrm{link}}]^2_{7/8}$ and $[Q_{\mathrm{link}}]^2_{8/8}$. Then, equation (B.3), the thermal expectation value of

$$Q^{(2)}_{\mathrm{link,linear}} = 2[Q_{\mathrm{link}}]^2_{2/2} - \frac{[Q_{\mathrm{link}}]^2_{3/4} + [Q_{\mathrm{link}}]^2_{4/4}}{2}, \tag{B.4}$$

is $\langle Q_{\mathrm{link}}\rangle^2$, up to a bias of order $\tau_{\mathrm{int},Q_{\mathrm{link}}}\tau_{\mathrm{avg},Q_{\mathrm{link}}}/N^2$. We can do it even better:

$$\begin{aligned} Q^{(2)}_{\mathrm{link,quadratic}} = {} & \frac{8}{3}[Q_{\mathrm{link}}]^2_{2/2} - 2\frac{[Q_{\mathrm{link}}]^2_{3/4} + [Q_{\mathrm{link}}]^2_{4/4}}{2} \\ & + \frac{1}{3}\frac{[Q_{\mathrm{link}}]^2_{5/8} + [Q_{\mathrm{link}}]^2_{6/8} + [Q_{\mathrm{link}}]^2_{7/8} + [Q_{\mathrm{link}}]^2_{8/8}}{4}, \end{aligned} \tag{B.5}$$

has thermal expectation value $\langle Q_{\mathrm{link}}\rangle^2$, up to corrections of order $\mathcal{O}(\mathrm{e}^{-N/\tau_{\mathrm{exp}}})$.

The fact that we have four real replicas offers us an alternative way of overcoming this problem. In fact, denoting by $Q^{(ij)}_{\mathrm{link}}$ the link overlap computed from replicas i and j, we have

$$Q^{(2)}_{\mathrm{link,4R}} = \frac{[Q^{(12)}_{\mathrm{link}}Q^{(34)}_{\mathrm{link}} + Q^{(13)}_{\mathrm{link}}Q^{(24)}_{\mathrm{link}} + Q^{(14)}_{\mathrm{link}}Q^{(23)}_{\mathrm{link}}]}{3}, \tag{B.6}$$

with $\langle Q^{(2)}_{\mathrm{link,4R}}\rangle = \langle Q_{\mathrm{link}}\rangle^2$ (we average over the three equivalent replica pairings to reduce statistical errors). The comparison of the two procedures offers an interesting test on the statistical quality of our data, because equation (B.3) holds only for $N \gg \tau_{\mathrm{int},Q_{\mathrm{link}}}, \tau_{\mathrm{exp}}$.

From the different estimators for the $\langle Q_{\mathrm{link}}\rangle^2$ we finally obtain four estimators of χ_{link}:

$$\chi^{\mathrm{biased}}_{\mathrm{link,2R}} = V\overline{([Q^2_{\mathrm{link}}] - [Q_{\mathrm{link}}]^2)}, \tag{B.7}$$

$$\chi^{\mathrm{linear}}_{\mathrm{link,2R}} = V\overline{([Q^2_{\mathrm{link}}] - Q^{(2)}_{\mathrm{link,linear}})}, \tag{B.8}$$

$$\chi^{\mathrm{quadratic}}_{\mathrm{link,2R}} = V\overline{([Q^2_{\mathrm{link}}] - Q^{(2)}_{\mathrm{link,quadratic}})}, \tag{B.9}$$

$$\chi_{\mathrm{link,4R}} = V\overline{([Q^2_{\mathrm{link}}] - Q^{(2)}_{\mathrm{link,4R}})}. \tag{B.10}$$

References

[1] Mydosh J A, 1993 *Spin Glasses: An Experimental Introduction* (London: Taylor and Francis)
[2] Fisher K H and Hertz J A, 1993 *Spin Glasses* (Cambridge: Cambridge University Press)
[3] Ballesteros H G, Cruz A, Fernandez L A, Martin-Mayor V, Pech J, Ruiz-Lorenzo J J, Tarancon A, Tellez P, Ullod C L and Ungil C, 2000 *Phys. Rev.* B **62** 14237
[4] Palassini M and Caracciolo S, 1999 *Phys. Rev. Lett.* **82** 5128
[5] Gunnarsson K, Svendlindh P, Nordblad P, Lundgren L, Aruga H and Ito A, 1991 *Phys. Rev.* B **43** 8199
[6] Franz S, Mézard M, Parisi G and Peliti L, 1998 *Phys. Rev. Lett.* **81** 1758
[7] Franz S, Mézard M, Parisi G and Peliti L, 1999 *J. Stat. Phys.* **97** 459
[8] Mézard M, Parisi G and Virasoro M, 1987 *Spin-Glass Theory and Beyond* (Singapore: World Scientific)
[9] Marinari E, Parisi G, Ricci-Tersenghi F, Ruiz-Lorenzo J J and Zuliani F, 2000 *J. Stat. Phys.* **98** 973

[10] McMillan W L, 1984 *J. Phys. C: Solid State Phys.* **17** 3179
[11] Bray A J and Moore M A, 1987 *Heidelberg Colloquium on Glassy Dynamics* (*Springer Lecture Notes in Physics* vol 275) ed J L van Hemmen and I Morgenstern (Berlin: Springer)
[12] Fisher D S and Huse D A, 1986 *Phys. Rev. Lett.* **56** 1601
[13] Fisher D S and Huse D A, 1988 *Phys. Rev.* B **38** 373
[14] Gardner E, 1984 *J. Physique* **45** 1755
[15] Krzakala F and Martin O C, 2000 *Phys. Rev. Lett.* **85** 3013
[16] Palassini M and Young A P, 2000 *Phys. Rev. Lett.* **85** 3017
[17] Contucci P, Giardinà C, Giberti C and Vernia C, 2006 *Phys. Rev. Lett.* **96** 217204
[18] Contucci P, Giardinà C, Giberti C, Parisi G and Vernia C, 2007 *Phys. Rev. Lett.* **99** 057206
[19] Contucci P, Giardinà C, Giberti C, Parisi G and Vernia C, 2009 *Phys. Rev. Lett.* **103** 017201
[20] Jörg T and Katzgraber H G, 2008 *Phys. Rev. Lett.* **101** 197205
[21] Moore M, Bokil H and Drossel B, 1998 *Phys. Rev. Lett.* **81** 4252
[22] Belletti F *et al* (Janus Collaboration), 2006 *Comput. Sci. Eng.* **8** 41
[23] Belletti F *et al* (Janus Collaboration), 2008 *Comput. Phys. Commun.* **178** 208
[24] Belletti F *et al* (Janus Collaboration), 2008 *Phys. Rev. Lett.* **101** 157201
[25] Belletti F *et al* (Janus Collaboration), 2009 *J. Stat. Phys.* **135** 1121
[26] Álvarez Baños R *et al* (Janus Collaboration), 2010 arXiv:1003:2943
[27] Fernandez L A, Martin-Mayor V, Perez-Gaviro S, Tarancon A and Young A P, 2009 *Phys. Rev.* B **80** 024422
[28] Edwards S F and Anderson P W, 1975 *J. Phys. F: Met. Phys.* F **5** 975
[29] Edwards S F and Anderson P W, 1976 *J. Phys. F: Met. Phys.* F **6** 1927
[30] Toulousse G, 1977 *Commun. Phys.* **2** 115
[31] Hasenbusch M, Pelissetto A and Vicari E, 2008 *J. Stat. Mech.* L02001
[32] Hasenbusch M, Pelissetto A and Vicari E, 2008 *Phys. Rev.* B **78** 214205
[33] Marinari E, Parisi G, Ricci-Tersenghi F and Ruiz-Lorenzo J J, 1998 *J. Phys. A: Math. Gen.* **31** L481
[34] Fernandez L A, Martin-Mayor V and Yllanes D, 2009 *Nucl. Phys.* B **807** 424
[35] de Dominicis C, Kondor I and Temesvári T, 1998 *Spin Glasses and Random Fields* ed A P Young (Singapore: World Scientific)
[36] de Dominicis C, Kondor I and Temesvári T, 1999 *Eur. Phys. J.* B **11** 629
[37] de Dominicis C and Giardina I, 2006 *Random Fields and Spin Glasses* (Cambridge: Cambridge University Press)
[38] Martin-Mayor V, 2007 *Phys. Rev. Lett.* **98** 137207
[39] Macdowell L G, Shen V and Errington J R, 2006 *J. Chem. Phys.* **125** 034705
[40] Marinari E, Parisi G, Ruiz-Lorenzo J J and Zuliani F, 1999 *Phys. Rev. Lett.* **82** 5176
[41] Contucci P and Giardinà C, 2005 *Phys. Rev.* B **72** 014456
[42] Contucci P, 2003 *J. Phys. A: Math. Gen.* **36** 10961
[43] Contucci P and Giardinà C, 2005 *Ann. Henri Poincaré* **6** 915
[44] Contucci P and Giardinà C, 2007 *J. Stat. Phys.* **126** 917
[45] Parisi G, 1998 arXiv:cond-mat/9801081
[46] Parisi G and Ricci-Tersenghi F, 2000 *J. Phys. A: Math. Gen.* **33** 113
[47] Iñiguez D, Parisi G and Ruiz-Lorenzo J J, 1996 *J. Phys. A: Math. Gen.* **29** 4337
[48] Fernandez L A, Martin-Mayor V, Parisi G and Seoane B, 2010 *Phys. Rev.* B **81** 134403
[49] Jimenez S, Martin-Mayor V, Parisi G and Tarancon A, 2003 *J. Phys. A: Math. Gen.* **36** 10755
[50] Franz S, Parisi G and Virasoro M, 1992 *J. Physique* **2** 1869
[51] Amit D J and Martin-Mayor V, 2005 *Field Theory, the Renormalization Group and Critical Phenomena* 3rd edn (Singapore: World Scientific)
[52] Chayes J, Chayes L, Fisher D S and Spencer T, 1986 *Phys. Rev. Lett.* **57** 2999
[53] Maiorano A, Martin-Mayor V, Ruiz-Lorenzo J J and Tarancón A, 2007 *Phys. Rev.* B **76** 064435
[54] Belletti F *et al* (Janus Collaboration), 2009 *Comput. Sci. Eng.* **11** 48
[55] Hukushima K and Nemoto K, 1996 *J. Phys. Soc. Japan* **65** 1604
[56] Marinari E, 1998 *Advances in Computer Simulation* ed J Kerstész and I Kondor (Berlin: Springer)
[57] Ogielski A, 1985 *Phys. Rev.* B **32** 7384
[58] Bittner E, Nubaumer A and Janke W, 2008 *Phys. Rev. Lett.* **101** 130603
[59] Sokal A D, 1997 *Functional Integration: Basics and Applications (1996 Cargèse School)* ed C DeWitt-Morette, P Cartier and A Folacci (New York: Plenum)
[60] Frigo M and Johnson S G, 2005 *Proc. IEEE* **93** 216
[61] Ballesteros H G, Fernandez L A, Martin-Mayor V and Muñoz Sudupe A, 1997 *Nucl. Phys.* B **483** 707

[62] Janus Collaboration, in preparation
[63] Katzgraber H, Palassini M and Young A, 2001 *Phys. Rev.* B **63** 184422
[64] Palassini M and Young A P, 2001 *Phys. Rev.* B **63** 140408(R)
[65] Perez-Gaviro S, Ruiz-Lorenzo J J and Tarancón A, 2006 *J. Phys. A: Math. Gen.* **39** 8567
[66] Iñiguez D, Marinari E, Parisi G and Ruiz-Lorenzo J J, 1997 *J. Phys. A: Math. Gen.* **30** 7337
[67] Josephson B D, 1966 *Phys. Lett.* **21** 608
[68] Luescher M, Weisz P and Wolff U, 1991 *Nucl. Phys.* B **359** 221
[69] Kim J K, 1993 *Phys. Rev. Lett.* **70** 1735
[70] Caracciolo S, Edwards R G, Ferreira S J, Pelissetto A and Sokal A D, 1995 *Phys. Rev. Lett.* **74** 2969
[71] Caracciolo S, Edwards R G, Pelissetto A and Sokal A D, 1995 *Phys. Rev. Lett.* **75** 1891
[72] Jörg T, 2006 *Phys. Rev.* B **73** 224431
[73] Moore M, 2010 arXiv:1005.0561
[74] Bouchaud J P, Dupuis V, Hammann J and Vincent E, 2001 *Phys. Rev.* B **65** 024439
[75] Katzgraber H and Young A, 2003 *Phys. Rev.* B **67** 134410
[76] Detrey J and de Dinechin F, 2007 *Microprocess. Microsyst.* **31** 537
[77] Press W H, Teukolsky S A, Vetterling W T and Flannery B P, 1992 *Numerical Recipes in C* 2nd edn (Cambridge: Cambridge University Press)