# Correct extrapolation of overlap distribution in spin glasses

Guy Hed[1], Alexander K. Hartmann[2] and Eytan Domany[1]

[1] *Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot 76100, Israel*
[2] *Institut für Theoretische Physik, Universität Göttingen, Bunsenstr. 9, 37073 Göttingen, Germany*

(February 23, 2001)

We study in $d = 3$ dimensions the short range Ising spin glass with $J_{ij} = \pm 1$ couplings at $T = 0$. We show that the overlap distribution is non-trivial in the limit of large system size.



## I. INTRODUCTION

There has been an upsurge of interest in the Edwards-Anderson model [1] of the short-range Ising spin glass with binary couplings,

$$\mathcal{H} = \sum_{i<j} J_{ij} S_i S_j \qquad\qquad J_{ij} = \pm 1 \tag{1}$$

in $d = 3$ dimensions. Several papers addressed the following well-posed question:

> *is the overlap distribution $P(q)$ of this model, measured at $T = 0$, trivial or non-trivial in the thermodynamic limit?*

A trivial $P(q)$ consists of a single delta function, $P(q) = \delta(q - q_{\mathrm{EA}})$; a non-trivial $P(q)$ has non-vanishing support also for $0 < q < q_{\mathrm{EA}}$. Claims were made to the effect that the nature of $P(q)$ bears on the validity of the droplet picture (DP) [2] versus the scenario [3] based on mean-field theory (MF) [4]. Our understanding is that a non-trivial global[1] $P(q)$ is consistent with both [5]. In this communication we do not take sides in the DP vs MF controversy; rather, we address the well-defined technical question posed above.

> *Our conclusion is that $P(q)$ is non-trivial at $T = 0$.*

Berg et al [6] addressed the issue directly by generating ground states $\mathbf{S}^\mu = (S_1^\mu, S_2^\mu, \ldots S_N^\mu)$, for 512 realizations $\{J\}$ of systems with sizes $L = 4, 6, 8$ (and for 7 realizations of $L = 12$). For each $\{J\}$ they computed the overlap distribution function $P_J(q)$, where the overlap $q^{\mu\nu} = (1/N)\mathbf{S}^\mu \cdot \mathbf{S}^\nu$ is calculated between all pairs of ground states $\mu, \nu$. They studied the function obtained by averaging over all realizations, $P(q) = [P_J(q)]_J$. In particular, they evaluated $P(0)$; the second moment of the distribution $\sigma^2(q)$, and the quantity $x_{1/2}$, where $x_a$ is defined by

$$x_a = 2 \int_0^a P(q)dq \tag{2}$$

If $P(q) \to \delta(|q| - q_{\mathrm{EA}})/2$ for large $L$, all these quantities should extrapolate to zero (provided one uses $a < q_{\mathrm{EA}}$). Berg et al found that all three quantities decrease as $L$ grows; they could, however, fit the data to $L^{-y}$, with $y = 0.72 \pm .12$, as well as to $A + BL^{-3}$, indicating consistency with extrapolation to both vanishing and non-vanishing limiting values.

Hartmann [7] also studied the size-dependence of $x_{1/2}$ and found that it behaves as $L^{-y}$ with $y = 1.25 \pm .05$, indicating a trivial $P(q)$; the same conclusion was reached by Hatano and Gubernatis [8] who studied $P(0)$ at finite temperatures. Krzakala and Martin [9] presented arguments that also support a trivial $P(q)$. Finally, very recently Palassini and Young [10] evaluated $P(q)$ for a sequence of temperatures and sizes $L = 4, 6, 8, 10$. They evaluated $x_{1/2}$ as function of $L$ and $T$ and demonstrated that the data are consistent with a scaling form. According to their scaling, for fixed $T > 0$ and sizes $L \gg L_c(T)$, $x_{1/2}$ goes to a constant, $x^\infty(T) \propto T$; hence they find that $P(q)$ is non-trivial at $T > 0$ and trivial at $T = 0$.

---

[1] We assume that $P(q)$ is measured for an entire system, whose size is then extrapolated to $L \to \infty$.

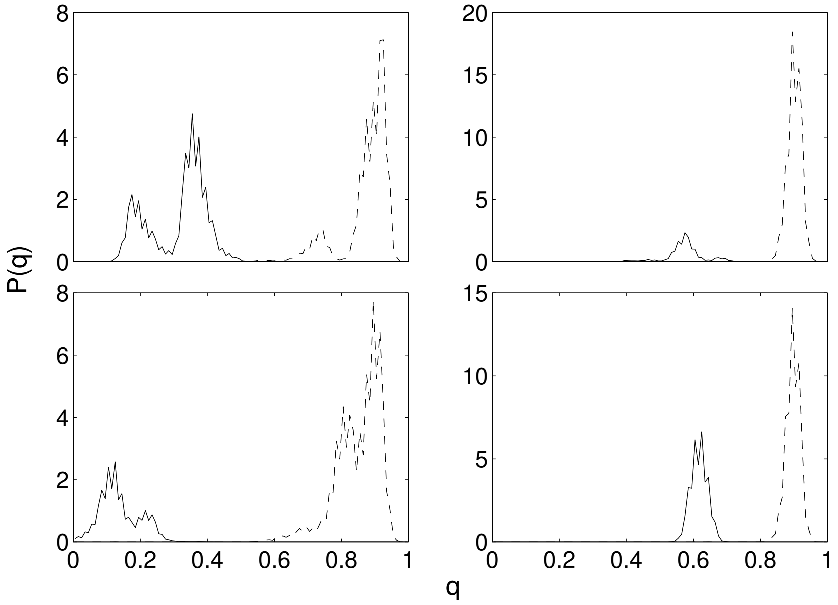


FIG. 1. The state overlap distribution $P(q)$ for four different realizations $\{J\}$ for system size $L=8$. Each distribution is divided to its components. The partial distribution $\tilde{P}^o_J(q)$ (see text), is represented by a solid line. The rest of the distribution, $P_J(q)-\tilde{P}^o_J(q)$, which includes $P^i_J(q)$, is represented by a dashed line.

## II. OUTLINE OF STRATEGY

We will argue now that all the studies mentioned measured a compound quantity, $x_{1/2}$, which is the sum of two parts; one which is relevant to the question asked, and another which is irrelevant. Furthermore, for some of the sizes studied, the irrelevant part is as large as the relevant one. We will show how can one isolate the relevant part, and present the results obtained when this is done. These results indicate that $P(q)$ is non-trivial at $T=0$.

For a particular realization $P_J(q)$ has, at $T=0$, the typical form presented in Fig 1. It has a large peak centered at some $q_0 \approx q_{\rm EA}$, and one or more smaller peaks. The largest peak is due to the overlap of pairs of states that belong to the *same* pure state. Denote the overlap distribution of such pairs by $P^i_J(q)$. The other peaks, at lower $q$, are due to the overlap between states that belong to two *different* pure states. The corresponding overlap distribution is $P^o_J(q)$ and we have

$$P_J(q) = P^i_J(q) + P^o_J(q) \qquad \text{and} \qquad P(q) = P^i(q) + P^o(q) \tag{3}$$

where the second equation is the average of the first over all realizations. Hence we can write

$$x_a = x^i_a + x^o_a = 2\int_0^a P^i(q) + 2\int_0^a P^o(q) \tag{4}$$

Irrespectively of whether $P(q)$ is trivial or non-trivial, one expects that the width of $P^i(q)$ decreases with increasing size, since $P^i(q) \to \delta(|q|-q_{\rm EA})/2$ as $L\to\infty$. Therefore as $L$ increases, the integral $x^i_a$ decreases towards 0;

$$x^i_a \approx L^{-y_i} \tag{5}$$

On the other hand, the behavior of $P^o(q)$ (and $x^o_a$) *does* distinguish a trivial $P(q)$ from a non-trivial one; in the first case $P^o(q)\to 0$, while in the non-trivial case $P^o(q)$ and $x^o_a$ do not vanish as $L\to\infty$. We believe that previous analysis was hindered by the lack of ability to decompose $P(q)$ and $x_{1/2}$ into its two constituent parts; a method that we developed recently enables us to perform this task. We describe below how we can use a very recently developed method [11] to identify unabiguously, for a large majority of the realizations, a partial distribution $\tilde{P}^o_J(q)$, which is a lower bound to $P^o_J(q)$. For realizations in which this identification is ambiguous we set $\tilde{P}^o_J(q)$ to zero; averageing yields $\tilde{P}^o(q) = [\tilde{P}^o_J(q)]_J$, which is a lower bound on $P^o(q)$. The corresponding lower bound on $x^o_a$ is given by

$$\tilde{x}^o_a = 2\int_0^a \tilde{P}^o(q)dq \ , \tag{6}$$

We found that the rate of convergence of $\tilde{P}^o(q)$ to its limiting large-$L$ form is non-uniform; for the sizes studied, convergence (with increasing $L$) is much slower, and statistical errors are much larger in the interval $0 \leq q \leq 0.5$ than in $0.4 \leq q \leq 0.7$. Hence we base our analysis on the latter interval, calculate

$$x^* = 2 \int_{0.4}^{0.7} P(q) dq \tag{7}$$

and show that it approaches a non-vanishing limit as $L \to \infty$.

## III. DECOMPOSING $P_J(q)$

Our method has been presented in [11], together with results obtained for the model (1). Full details of the method are given in [12]; here we give a brief summary of the main ingredients.

We have shown that an unbiased sample [11] of $M$ ground states breaks naturally into two large groups, $\mathcal{C}$ and $\bar{\mathcal{C}}$. The states of the two sets are related by spin reversal. For a large majority of realizations the set $\mathcal{C}$ also breaks into two natural subsets, $\mathcal{C}_1$ and $\mathcal{C}_2$. By natural we mean that the overlap between two states that belong to the same group (say $\mathcal{C}_1$) is significantly larger than between two that belong to two different groups. This suggest that the states in these clusters belong to different pure states, separated by free energy barriers. These barriers consist of correlated spin domains $\mathcal{G}_1$ and $\mathcal{G}_2$, which flip collectively when we move from a state in one cluster to a state in another cluster. The spins that belong to $\mathcal{G}_1$ are reversed in at least 95% of the pairs of states $\mu \in \mathcal{C}$ and $\nu \in \bar{\mathcal{C}}$. Similarly, the second largest domain $\mathcal{G}_2$ contains those spins that flip in 95% (or more) of the times we pass between pairs states, with one member in $\mathcal{C}_1$ and the other in $\mathcal{C}_2$.

The domains $\mathcal{G}_1$ and $\mathcal{G}_2$ play the role of the cores of macroscopic "zero energy excitations" [13] that flip as we go from one pure state to another. $\mathcal{G}_1$ separates state space into $\mathcal{C}$ and $\bar{\mathcal{C}}$. Within $\mathcal{C}$, $\mathcal{G}_2$ induces a further non-trivial separation of the states, into clusters $\mathcal{C}_1$ and $\mathcal{C}_2$. Each cluster $\mathcal{C}_\alpha$ contains one or more pure states. When $\mathcal{G}_2$ is large ("macroscopic"), a pair of states $\mu \in \mathcal{C}_1$ and $\nu \in \mathcal{C}_2$ will belong to different pure states, and their overlap $q^{\mu\nu}$ will contribute to $P^o_J(q)$. Hence, we define a new distribution $\tilde{P}^o_J(q)$, to which only pairs of states $\mu \in \mathcal{C}_1$ and $\nu \in \mathcal{C}_2$ contribute. This function is a lower bound to $P^o_J(q)$, since we might have for some realizations a third macroscopic cluster, in which case $\mathcal{C}_1$ contains states from more than one pure state. When this happens, some pairs of states, both taken from $\mathcal{C}_1$, contribute to $P^o_J(q)$, and we do not include them in $\tilde{P}^o_J(q)$. In order to assure that $\mathcal{C}_1$ and $\mathcal{C}_2$ indeed do not belong to one pure state, we consider only those realizations for which $|\mathcal{G}_2| > 0.05N$. Otherwise, we set $\tilde{P}^o_J(q) = 0$.

The method we used to partition the states was based on a clustering procedure. It is important to stress the fact that the main result of the present study, that there are states whose overlap contribution should be separated from the self-overlap peak and does not vanish in the thermodynamic limit, does *not* depend qualitatively on the way the state clusters are determined. In fact, any method, which projectes out a particular contribution to $P(q)$ and has a nonvanishing weight in the $L \to \infty$ limit, will lead to the same conclusion. The only requirements are that the method is applied for all system sizes in the same way and the contribution is measured in absolute weights with respect to the total $P(q)$.

We determined [11] for each $L$ the size distributions $|\mathcal{G}_2|/N$ and found that they are nearly the same for $4 \leq L \leq 8$, indicating convergence. $|\mathcal{G}_2|$ scales as $N = L^3$ ; for $L = 6$ the average value of $|\mathcal{G}_2|/N$ is 0.07 and its standard deviation 0.09; for $L = 8$ the numbers are 0.08 and 0.10, respectively.

Since the limiting size distribution of $\mathcal{G}_2$ is non-trivial, we expect a non-trivial $\tilde{P}^o_J(q)$ as long as $\mathcal{C}_2$ does not vanish. The size $|\mathcal{C}_2|$ of this state cluster is determined by the correlation between the spins of $\mathcal{G}_1$ and $\mathcal{G}_2$. If this correlation approaches 1, this means that $\mathcal{G}_2$ has a low probability to flip without $\mathcal{G}_1$, resulting in $|\mathcal{C}_1| \gg |\mathcal{C}_2|$. The average correlation between these domains is

$$\bar{c}_{12} = \frac{1}{|\mathcal{G}_1||\mathcal{G}_2|} \sum_{i \in \mathcal{G}_1} \sum_{j \in \mathcal{G}_2} {c_{ij}}^2 , \tag{8}$$

where $c_{ij} = \langle S_i S_j \rangle$ is the correlation between spins $i$ and $j$. The weight of the contribution to $P(q)$ by pairs of states in which $\mathcal{G}_2$ is flipped without $\mathcal{G}_1$ or vice versa can be evaluated [11] by $(1 - \bar{c}_{12})/2$. We found [11] that $\bar{c}_{12}$ does not extrapolate to 1 as $L \to \infty$. For a realization $\{J\}$, in which $\bar{c}_{12} < 1$ and $|\mathcal{G}_2| > 0$, the function $P^o_J(q)$ will be non-trivial, i.e. it will have a finite support for $-1 < q < 1$.

To show explicitly that this indeed is the case, we studied $\tilde{P}^o(q)$ and $P(q)$. The function $\tilde{P}^o(q) = [\tilde{P}^o_J(q)]_J$, presented in Fig. 2, is a conservative estimate (and a lower bound) for $P^o(q)$. It also has a clear physical meaning. $\tilde{P}^o(q)$ is

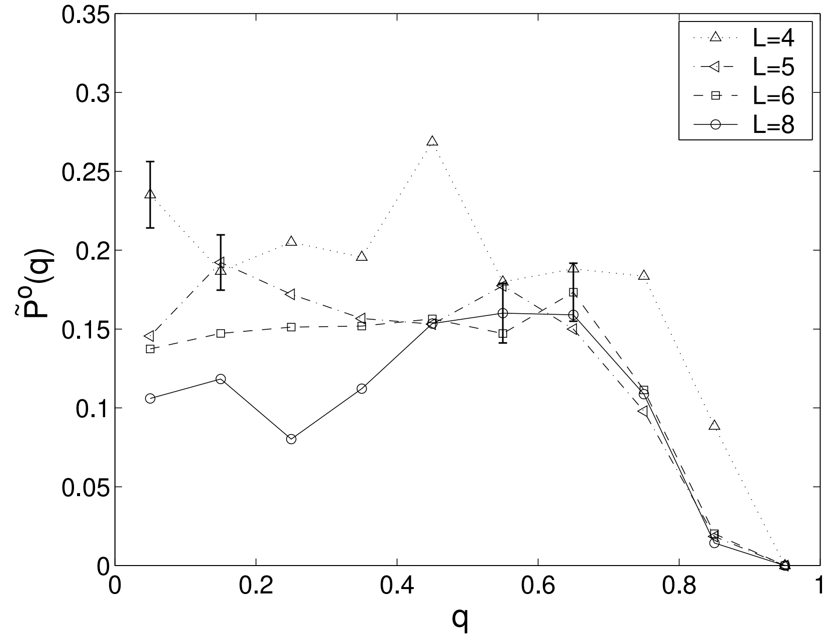


FIG. 2. The partial distribution $\tilde{P}^o(q)$ for $L = 4, 5, 6, 8$. It is normalized so that $2\int_0^1 \tilde{P}^o(q)dq$ is its weight in the total $P(q)$. For each $L$ the largest error bar is shown.

the distribution of overlaps between pairs of states on the two sides of the second largest free energy barrier in the system. For comparison, we also present the full $P(q)$ in Fig 3.

$\tilde{P}^o(q)$ has low values and large relative errors for $q < 0.4$. In this range its values decrease with increasing $L$. On the other hand, in the interval $0.4 \le q \le 0.7$ it seems to have converged. Therefore we chose this range for our analysis, and calculated the integrals $x^*$ (see eq. (7)) and

$$\tilde{x}^{*o} = 2\int_{0.4}^{0.7} \tilde{P}^o(q)dq \; . \tag{9}$$

The values obtained for $L = 4, 5, 6, 8$ are presented in Table I. Perhaps the most direct evidence for our claim is the manner in which the values of $\tilde{x}^{*o}$ level off as the size increases, at 0.047. On the other hand, those of $x^*$ decrease with size. We performed a fit of the latter to the form

$$x^* = A + BL^{-y} \tag{10}$$

The results of several attempts to fit the data to this form are summarized in Table II. The best fit (with $\chi^2 = 1.0 \times 10^{-5}$) was obtained for $y = 2.06(49)$ and $A = 0.042(15)$, which is close to 0.047. Imposing this value, i.e. setting $A = 0.047$ and fitting $B$ and $y$, we had a somewhat larger $\chi^2 = 1.1 \times 10^{-5}$; imposing $A = 0$ yields a worse fit, with $\chi^2 = 4.2 \times 10^{-5}$. We believe that these results clearly show that $P(q)$ is non-trivial.

| $L$ | $x^*$ | $\tilde{x}^{*o}$ | $x_{1/2}$ | $\tilde{x}^o_{1/2}$ |
|---|---|---|---|---|
| 4 | 0.161(5) | 0.064(10) | 0.157(7) | 0.109(14) |
| 5 | 0.115(5) | 0.048(11) | 0.105(6) | 0.082(15) |
| 6 | 0.096(5) | 0.048(10) | 0.095(5) | 0.074(14) |
| 8 | 0.070(4) | 0.047(12) | 0.062(4) | 0.057(14) |

TABLE I. Values of the observables (defined in eq. (2), (6), (7) and (9)) for different system sizes.

To make contact with previous analysis we also calculated $x_{1/2}$ and performed similar fits, the results of which are also presented in Table II. As discussed above, in this range of $q$ the function $\tilde{P}^o(q)$ has larger statistical fluctuations, and is decreasing with size (to a limiting value that is expected to be small, albeit non-zero). Indeed the best fit for

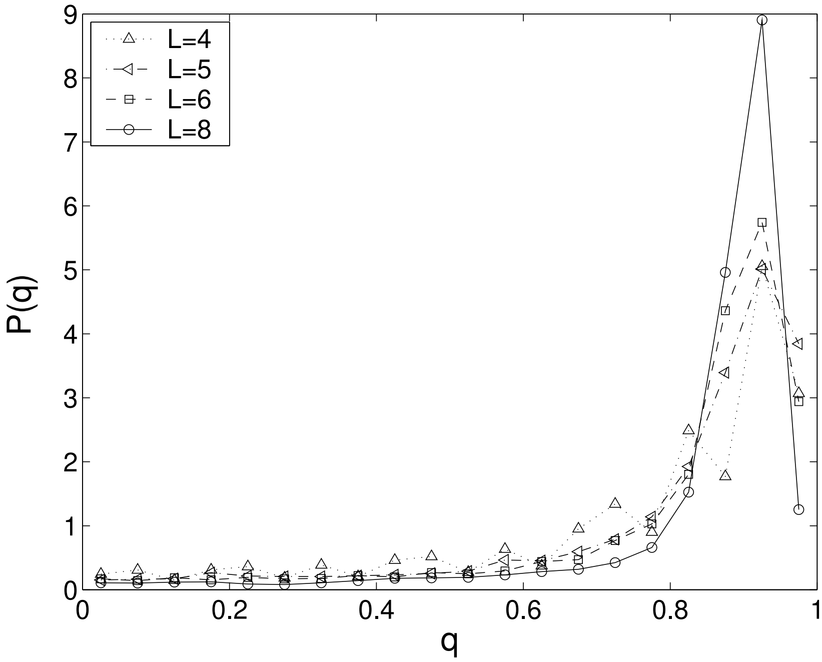


FIG. 3. The distribution $P(q)$ for $L = 4, 5, 6, 8$. The size of the error bars is of the order or less than the size of the symbols.

$x_{1/2}$ is attained for $y = 2.07(1.51)$ and $A = 0.036(47)$, with $\chi^2 = 1.0 \times 10^{-4}$. Note that this $\chi^2$ is 10 times the value obtained when fitting $x^*$. Since our estimated value of $A$, as well as the estimates of others [6,7] is much smaller then the values of $x_{1/2}$ used to perform the fit, it is hard to distinguish, by means of this extrapolation, between $A = 0$ and a small positive $A$. Indeed, when we impose $A = 0$ or both $A = 0, B = 1$ we get fits of comparable quality, with an exponent which is consistent with Hartmann's estimate.

Finally, we attempted to fit the data for $\tilde{x}^o_{1/2}$. The values of $\tilde{x}^o_{1/2}$ for the system sizes used are smaller and noisier the the results for $x^*$. Nevertheless, using the same fit for $\tilde{x}^o_{1/2}$ yields minimum of $\chi^2 = 1.5 \times 10^{-5}$ for $y = 1.94(1.02)$ and $A = 0.040(20)$, quite consistent with the results obtained for $x^*$.

The authors thank A.P. Young and M. Palassini for most helpful correspondence. AKH was supported by the Graduiertenkolleg "Modellierung und Wissenschaftliches Rechnen in Mathematik und Naturwissenschaften" at the *Interdisziplinäres Zentrum für Wissenschaftliches Rechnen* in Heidelberg and the *Paderborn Center for Parallel Computing* by the allocation of computer time. AKH obtained financial support from the DFG (*Deutsche Forschungs Gemeinschaft*) under grant Zi209/6-1. ED and GH were supported by the Germany-Israel Science Foundation (GIF).

| fit | $A$ | $B$ | $y$ | $\chi^2$ |
|---|---|---|---|---|
| $x^*$, best | 0.042(15) | 2.06(1.16) | 2.06(49) | $1.0 \times 10^{-5}$ |
| $x^*$, imposed $A$ | 0.047 | 2.45(36) | 2.21(10) | $1.1 \times 10^{-5}$ |
| $x^*$, imposed $A$ | 0 | 0.88(13) | 1.24(9) | $4.2 \times 10^{-5}$ |
| $x_{1/2}$, best | 0.036(47) | 2.10(3.68) | 2.07(1.51) | $1.0 \times 10^{-4}$ |
| $x_{1/2}$, imposed $A$ | 0 | 0.97(26) | 1.33(17) | $1.3 \times 10^{-4}$ |
| $x_{1/2}$, imposed $A, B$ | 0 | 1 | 1.35(2) | $1.3 \times 10^{-4}$ |
| $\tilde{x}^o_{1/2}$, best | 0.040(20) | 1.00(1.16) | 1.94(1.02) | $1.5 \times 10^{-5}$ |

TABLE II. Best fit parametrs for $x(L) = A + BL^{-y}$.

[1] S. F. Edwards and P. W. Anderson, J. Phys. **F5**, 965 (1975).
[2] D. S. Fisher and D. A. Huse, Phys. Rev. B **38**, 386 (1988).
[3] M. Mézard, G. Parisi, N. Sourlas, G. Toulouse and M. A. Virasoro, J. Phys. **45**, 843 (1984).
[4] D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).
[5] D. A. Huse and D. S. Fisher, J. Phys A **20**, L997 (1987).
[6] B. A. Berg, U. E. Hansmann and T. Celik, Phys. Rev. B **50**, 16444 (1994).
[7] A. K. Hartmann, Eur. Phys. J. B **13**, 539 (2000).
[8] N. Hatano and J. E. Gubernatis, cond-mat/0008115; see, however, E. Marinari, G. Parisi, F. Ricci-Tersenghi and F. Zuliani, cond-mat/0011039.
[9] F. Krzakala and O. C. Martin, cond-mat/0010010.
[10] M. Palassini and A. P. Young, cond-mat/0012161.
[11] G. Hed, A. K. Hartmann, D. Stauffer and E. Domany, cond-mat/0007356.
[12] To be published.
[13] F. Krzakala and O. C. Martin, Phys. Rev. Lett. **85**, 3013 (2000).