Graph bipartitioning and statistical mechanics

1987 J. Phys. A: Math. Gen. 20 L1

(http://iopscience.iop.org/0305-4470/20/1/001)

J. Phys. A: Math. Gen. **20** (1987) L1-L8. Printed in the UK

LETTER TO THE EDITOR

Graph bipartitioning and statistical mechanics

Jayanth R Banavar†, David Sherrington†‡ and Nicolas Sourlas†§

† Schlumberger-Doll Research, Old Quarry Road, Ridgefield, CT 06877-4108, USA
‡ Physics Department, Imperial College, London SW7 2BZ, UK‖
§ Laboratoire de Physique Théorique de l'Ecole Normale Supérieure, 24 rue Lhomond, 75005 Paris, France‖

Received 29 August 1986

Abstract. The problem of bipartitioning a random graph of fixed finite local valence (connectivity) so as to minimise the number of cross edges is studied by the application of Monte Carlo simulation of thermodynamics and annealing. This problem is *NP*-complete and is relevant as an idealisation of several practical organisational problems. Evidence is presented for the existence of a multiplicity of metastable states and for non-self-averaging and ultrametricity in the space of overlaps, albeit with a possible critical valence. An empirical formula is presented for the optimal cost, in excellent accord with values obtained by simulation and those known exactly.

Thanks to progress in our understanding of the physics of strongly disordered and frustrated systems, the methods of statistical mechanics are now being applied fruitfully to a much wider variety of problems. One of the most promising is combinatorial optimisation. This letter is concerned with the application of methods and concepts from the statistical physics of spin glasses [1] to the optimal bipartitioning of graphs of complex random connectivity, a problem known to be *NP*-complete [2] and relevant to several interesting practical applications [3].

It is now widely believed that the characteristic feature of *NP* systems, which makes them particularly difficult to deal with, is the existence of a large number of local optima, not simply related by any symmetry, growing in number with the size of the system and separated in phase space by highly non-optimal barriers. In certain spin-glass models a further hierarchical organisation of these optimal states manifests itself in non-self-averaging [4, 5] and ultrametricity [4] in the space of overlaps between states. From a simulational study, we find similar behaviour in graph partitioning, but with interesting dependence on coordination number. We also obtain a new empirical formula for the (self-averaging) optimal cost.

Let us turn to a more precise statement of the problem. A graph is specified by a set of vertices $V=\{v_1, v_2, \ldots, v_N\}$ and a set of edges $E=\{(v_i, v_j)\}$. We shall be concerned with N even. We define a connectivity matrix $\boldsymbol{a}$ to have elements $a_{ij}=1$ if there is an edge connecting vertices v_i and v_j, and $a_{ij}=0$ otherwise. The bipartitioning problem consists of partitioning V into two subsets V_1, V_2 of equal size in such a way as to minimise the number of edges N_{ct} connecting V_1 and V_2. N_{ct} is thus our cost function. For general graphs this problem is known to be *NP*-complete [2].

‖ Permanent address.

0305-4470/87/010001+08$02.50

The bipartitioning problem is easily mapped into that of finding the ground state of the ferromagnetic Ising Hamiltonian

$$H_1 = -\sum_{i<j} a_{ij}\sigma_i\sigma_j \tag{1}$$

where the $\sigma_i = \pm 1$ are Ising spins, subject to the constraint

$$\sum_i \sigma_i = 0. \tag{2}$$

V_1 is the set of $\sigma_i = +1$ and V_2 the set of $\sigma_i = -1$. The constraint of zero magnetisation ensures that the sizes of V_1 and V_2 are equal and introduces frustration [6] into the problem, which is the origin of its interesting properties. The cost function N_{ct} is simply related to the ground-state energy E_0 of H_1 by

$$E_0 = -N_{ed} + 2N_{ct} \tag{3}$$

where $N_{ed} = \Sigma_{i<j}\, a_{ij}$ is the total number of edges.

Clearly, there are special cases of this problem which are trivial, such as if the vertices (spins) can be considered to lie on the sites of a finite-dimensional lattice with the edges corresponding to short-range lattice-translationally invariant interactions. We are not concerned with such trivial situations here. Rather, to emphasise and typify the *NP* character, we concern outselves with the case of random (but quenched) connectivity, investigating statistical measures of its consequence.

Fu and Anderson [7] (hereafter referred to as FA) have shown that in the case where the a_{ij} are independent random variables, present with probability p, independent of N, absent with probability $(1-p)$, the graph bipartitioning problem is equivalent in the thermodynamic ($N \to \infty$) limit to finding the ground-state energy of the infinite-range Sherrington-Kirkpatrick (SK) model spin glass [8]†. Using the Parisi solution [9] within a replica treatment [8, 10] they demonstrate the result for the average cost function:

$$\lim_{N\to\infty} \frac{N_{ct}}{N_{ed}} = \frac{1}{2}\left[1 - c\left(\frac{1-p}{pN}\right)^{1/2}\right] \tag{4}$$

where $c = 1.5266 \pm 0.0002$.

In the present letter we study the case where the a_{ij} are again distributed randomly but every vertex is now connected to exactly α other vertices, with α independent of N, i.e. $a_{ij} = a_{ji}$ are random parameters, equal to 1 or 0, subject only to the constraint

$$\sum_j a_{ij} = \alpha \qquad \text{for all } i. \tag{5}$$

In the above-mentioned Fu-Anderson model the individual vertex connectivities (valences) are not restricted and the average connectivity, $\bar{\alpha} = pN$, is extensive. Clearly, finite connectivity is more realistic for practical applications, such as partitioning electronic circuit elements between microchips [3]. It is also well known from studies of spin glasses that results of the high connectivity infinite-ranged mean-field-like models cannot be assumed to carry over to more realistic short-range models, although remarkable practical similarities have been found (if not completely explained) [1].

† In common with the SK model a scaling of the effective interaction a_{ij} as $N^{-1/2}$ is needed to make the problem thermodynamically sensible. This also results in a simplifying truncation in the effective Hamiltonian of a replica treatment [7, 8].

In fact, the model we study lies in an intermediate class, having an effectively infinite spatial dimensionality in spite of its intensive connectivity. A third canonical random-connection bipartitioning problem is that with the a_{ij} independently randomly present or absent but with α independent of N, i.e. with the less restrictive constraint

$$N^{-1}\sum_{ij} a_{ij} = \bar{\alpha}. \tag{6}$$

We do not consider this problem, with only average but finite valence (equation (6)), here, but note that (i) it is likely to be easier to study analytically than the model with fixed finite valence (equation (5)) [11, 12], and (ii) the two problems are not obviously equivalent, as a consideration of the different structures permitted in the two cases rapidly demonstrates.

Spin-glass studies have been concerned not only with the ground-state energy but also, and in fact more particularly, with the questions of the existence of phase transitions (of both conventional and unconventional character), the extent to which properties are self-averaging and the occurrence of ultrametricity. A property is said to be self-averaging if its relative variance over many systems constructed with the same stochastic preparation rules (here different actual choices of $\{a_{ij}\}$ with the same α) vanishes as the number of samples tends to infinity. Ultrametricity [4] characterises hierarchical clustering of macrostates, even for a single manifestation of the random preparation. For an Ising model it may be examined within the space of magnetisation overlaps†

$$q_{ss'} = N^{-1}\sum_{i} m_i^s m_i^{s'} \tag{7}$$

where s, s' label states and m_i^s is the magnetisation at site i in state s. A space of overlaps is ultrametric if, given any three states s, s', s'', the two smallest $q_{RR'}$; $R, R' = s, s', s''$ are equal. Alternatively, one may define a distance in overlap space by $d_{ss'} = (q_{ss} - q_{s's'} - 2q_{ss'})^{1/2}$. The distance space is ultrametric if all triangles are isosceles, with the two largest sizes equal, or are equilateral.

We have studied numerically the problem of bipartitioning of graphs with random connections but fixed finite valence, using Monte Carlo simulation with importance sampling. Our results can be summarised as follows.

(a) The cost function is self-averaging and is given by the following empirical formula:

$$\frac{N_{ct}}{N_{ed}} = \frac{1}{2}\left(1 - \frac{c}{(\alpha + c^2 - 2)^{1/2}}\right) \tag{8a}$$

where, as before, $c = 1.5266$. This result quite accurately reproduces our numerical results for $3 \leqslant \alpha \leqslant 20$ and $450 \leqslant N \leqslant 4000$, it correctly gives the $N \to \infty$ result for $\alpha = 2$, namely zero‡, and furthermore it reproduces the result of Fu and Anderson for the case $\alpha = pN$, $1 \gg p \gg N^{-1}$. We can extend the formula to reproduce our results and the FA results for $\alpha = pN$, any finite p, by

$$\frac{N_{ct}}{N_{ed}} = \frac{1}{2}\left[1 - c\left(\frac{(1-\alpha/N)}{[\alpha + (c^2-2)(1-\alpha/N)]}\right)^{1/2}\right]. \tag{8b}$$

Unfortunately, we have so far been unable to derive it analytically.

† It is also reflected in other overlaps [13].

‡ For $\alpha = 2$, our graph is a collection of rings with N_{ct} at most 2, while $N_{ed} \to O(N)$.

(*b*) At the temperatures studied, there exists a large number of (macro) states. We find a non-trivial overlap function

$$P(q) = \sum_{s,s'} p^s p^{s'} \delta(q - q_{ss'}) \tag{9}$$

where p^s is the probability of state s. Not only is $P(q)$ not delta function-like, also it is non-self-averaging [4], at least for the sizes studied. This has been verified for $\alpha = 4$, $\alpha = 6$ and $\alpha = 12$ and for sizes $400 \leqslant N \leqslant 2400$.

(*c*) Regarding ultrametricity, our observations suggest the existence of a critical value α_c of α such that for $\alpha < \alpha_c$ there is no ultrametricity, while for $\alpha > \alpha_c$ the space of state overlaps is ultrametric. More specifically, we found strong evidence for ultrametricity for $\alpha = 12$, while the indications for $\alpha = 4$ and $\alpha = 6$ are poor.

(*d*) We were unable to resolve the question of the existence of a phase transition. Because of the lack of self-averaging of $P(q)$, the size scaling necessary to resolve this issue requires much more computer time than we have been able to devote.

All our results are numerical. In principle, one could analyse this problem mathematically using the replica method, as was done with other optimisation problems [11], but in the case of finite connectivity, α or $\bar{\alpha}$ independent of N, the effective replicated Hamiltonian contains an infinity of relevant terms. Furthermore, fixed connectivity also imposes additional constraints making the analysis particularly hard.

More specifically, we proceeded as follows. Instead of imposing the rigid constraint $M \equiv \Sigma\, \sigma_i = 0$, we added a restraining penalty term λM^2 to the Hamiltonian. This term ensures that $M/N \sim 1/\sqrt{N}$. In practice we took $\lambda = 0.4$ and found, for all of the ground states, that $-4 \leqslant M \leqslant 4$, even for N as large as 4000. The simulation methods we used were very similar to those of [14]. Connectivity matrices were constructed randomly with fixed valence. For any given connectivity matrix $\{a_{ij}\}$, we considered a large number N_S of copies differing only in the initial conditions, which were chosen at random; typically we took $N_S = 32$. We used the heat bath Monte Carlo algorithm [15], starting from high temperatures and slowly cooling down, each copy treated independently. Once at the desired temperature we started making measurements *after* the internal energy was stabilised for *all* of the copies. In addition to the energy and the magnetisation, we measured the distribution of overlaps $P(q)$ and the joint distribution function $P(q_1, q_2, q_3)$, defined below, proceeding as follows.

For every pair (α, β) of copies ($\alpha \neq \beta$; $\alpha, \beta = 1, \ldots, NS$) we measured the overlap $q_{\alpha\beta} = N^{-1} \Sigma \sigma_i^\alpha \sigma_i^\beta$. $P(q)$ is the resultant (probability) distribution of $q_{\alpha\beta}$†. Similarly, for any three *different* copies α, β, γ we measured $q_1 = q_{\alpha\beta}$, $q_2 = q_{\alpha\gamma}$, $q_3 = q_{\beta\gamma}$ and $P(q_1, q_2, q_3)$. In order to get a reliable estimate of $P(q)$, our measuring time t (=number of Monte Carlo steps, during which we performed measurements of the $q_{\alpha\beta}$) should be long enough to permit a representative part of the space of states to be visited. To assess if this is satisfied for a given measuring time t we define the self-overlap

$$q_{EA} \equiv (q_{10}/q_8)^{1/2} \tag{10}$$

q_n being the nth moment of $P(q)$, and, for every copy α,

$$Q_\alpha(t) = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{t} \sum_{t'=1}^{t} \sigma_i(t') \right)^2. \tag{11}$$

† It may be shown that the $P(q)$ given by this procedure is equivalent to that defined in equation (9) [16].

We considered t to be long enough if $Q_\alpha(t)/q_{EA} < 0.1$ for every copy, $\alpha = 1, \ldots, N_S$. (Typically, for $\alpha = 6$, $N = 1200$, $\beta = 0.64$†, we use $t \sim 150\,000$ Monte Carlo steps per spin and find $q_{EA} \sim 0.45$.) While the fulfilment of this condition is important for $P(q)$, it is not important for the ultrametricity tests (even if we take a small subspace of the space of states it should be ultrametric) nor the internal energy (as we found that this is state independent).

Figure 1 shows $P(q)$ for two typical realisations of the connectivity matrix $\{\alpha_{ij}\}$, for $N = 1200$, $\alpha = 6$ and $\beta = 0.62$. These two curves indicate both the existence of a large number of states and the lack of self-averaging of $P(q)$ at $N = 1200$. In order to decide whether this is a finite volume artefact or a genuine effect, we tried to study the variation with N of the probability distributions of the nth moment $q_n(\beta, N)$ of $P(q)$. Unfortunately, however, it turned out that much more computer time than we had available would be needed in order to reach any conclusion on this issue.

In order to measure the ground-state energy and assess its self-averaging we slowly and independently cooled down to zero temperature NS (typically ten) copies of the system for several realisations of the connectivity matrix. Table 1 shows for $N = 1800$

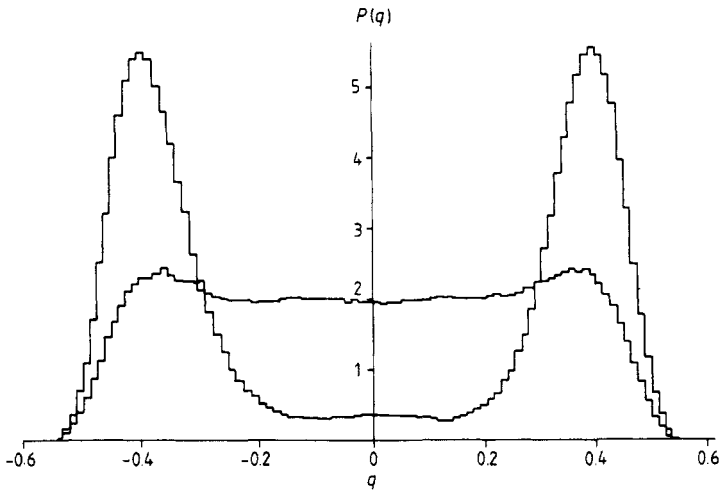

Figure 1. Probability distribution of the overlap $P(q)$ for two typical realisations of the connectivity matrix, for $N = 1200$, $\alpha = 6$ and $\beta = 0.62$.

Table 1. Ground-state results for $\alpha = 6$, $N = 1800$, for 5 different realisations of the disorder. E_m is the minimum, E_{av} the average and σ_E the variance found in each case for ten separate annealing searches.

E_m	E_{av}	σ_E
1.8112	1.8082	3.0×10^{-3}
1.8155	1.8130	4.5×10^{-3}
1.8135	1.8094	3.0×10^{-3}
1.8154	1.8088	7.2×10^{-3}
1.8124	1.8111	2.1×10^{-3}

† β is the usual inverse temperature used in the Boltzmann weighting $e^{-\beta H}$.

and $\alpha = 6$ the minimum value E_m, the average E_{av} and the variance σ_E of the ground-state energy found for the different copies, for five realisations of the connectivity matrix. σ_E can be considered as an estimate of our error on the ground-state energy. Similar results were obtained for different values of N and α. We concluded that, within the accuracy we were able to reach, the ground-state energy is self-averaging.

Table 2 shows the ratio of the ground-state energy E to the ground-state energy E_f of the corresponding ferromagnetic system (i.e. without the constraint $M = 0$: in our notation $E_f = N_{ed}/2$) for different values of the connectivity α. It also shows the prediction for this ratio, as given by our empirical formula $R_e = c/(\alpha + c^2 - 2)^{1/2}$. The agreement is very good for all the values of α we studied. (We give in the table a conservative estimate of E, i.e. the average of the lowest value found in each copy for a particular α.) We also found that, within our accuracy, E/E_f does not depend on N, for $450 \leqslant N \leqslant 4000$. In order to understand this, we may consider the graph equipartitioning problem on a symmetric hyperplanar section of a regular hypercubic lattice, with nearest-neighbour connections in D dimensions. The optimal equipartition is obviously through a hyperplane, perpendicular to one of the principal axes. This means that $E/E_f \sim L^{D-1}/L^D = L^{-1} = N^{-1/D}$ where L is the linear dimension of the hypercube. Thus our result that $E/E_f \sim$ constant can be rephrased by saying that, in our case, the dimensionality of the lattice is infinite. Another way of seeing this is to try to build the connectivity matrix a_{ij}. One starts from one site, i_0, and connects it to α sites j_γ, $\gamma = 1, \ldots, \alpha$. Then every newly connected site has to be connected to $(\alpha - 1)$ other sites, k_δ, every site being equally probable. In the $N \to \infty$ limit, with probability one, these 'k' will be different from the 'j' sites. The same argument can be repeated several times, until a substantial portion of the total number of sites has been connected. This means that the probability of small loops is $1/N$ and that, looked at on a fixed scale, as $N \to \infty$, our graph has a tree-like structure, which, as any tree, is infinitely dimensional. Another demonstration of this idea follows from a computation of the critical temperature and the spontaneous magnetisation of the corresponding ferromagnetic model (i.e. without the constraint $M = 0$). In such a study we found the known results for the Bethe lattice, with coordination number α.

Finally, we discuss ultrametricity. Our analysis follows closely that of [14], using the following definition of the distance between two microstates: $d^2_{ss'} = q_{EA} - q_{ss'}$. Figures 2 and 3 show the density of triangles in the space of states as a function of

Table 2. Ratio of the ground-state energies of the restricted ($M = 0$; graph partitioning) and unrestricted ($M = N$; ferromagnetic) problems for various α, compared with the values R_e predicted by the empirical formula, equation (8). Each quoted value of E/E_f is the average over several different realisations of $\{a_{ij}\}$.

α	E/E_f	R_e
3	0.840	0.836
4	0.732	0.734
5	0.663	0.661
6	0.605	0.607
8	0.528	0.529
9	0.499	0.500
10	0.470	0.475
15	0.386	0.390
20	0.335	0.339

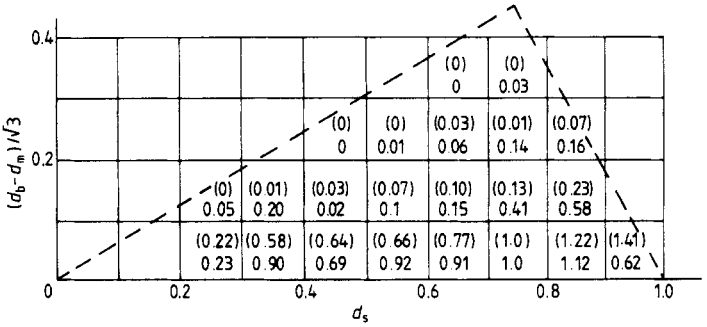

Figure 2. Probability distribution of the density of triangles in the space of states for $\alpha = 12$ and $\beta = 0.64$. d_s is the length of the smallest side, while $d_b - d_m$ is the difference between the lengths of the other two sides. Two different sizes are shown, $N = 600$ and $N = 1200$. In each case $\alpha = 12$ and $\beta = 0.64$, and the results are for a single realisation of the connectivity matrix. The two sets of figures refer to the two different sizes of the system: lower set $N = 600$ and upper set (bracketed) $N = 1200$. The overall normalisation is arbitrary.

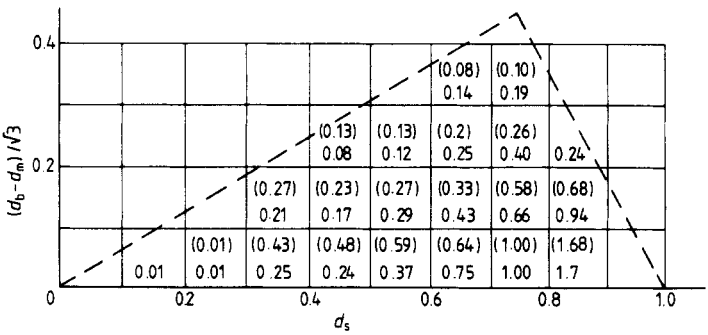

Figure 3. The same as in figure 2, but for $\alpha = 4$ and $\beta = 4.0$ for two sizes of the system: $N = 1800$ (unbracketed) and $N = 2400$ (bracketed).

d_s, the smallest distance between the three states, and of $d_b - d_m$, where d_b is the largest distance and d_m the middle one. The distances have been rescaled such that $d_b + d_m + d_s = 3$ for every triangle. The broken lines indicate the bounds from the triangular inequality and from $d_b \geqslant d_m \geqslant d_s$. Perfect ultrametricity means zero density except for $d_b = d_m$, i.e. for every d_s the weight in $(d_b - d_m)$ should be concentrated at the $(d_b - d_m) = 0$ axis. Equilateral triangles are mapped onto the point $d_s = 1$, $(d_b - d_m) = 0$. As perfect ultrametricity is expected only in the limit $N \to \infty$ (in the case of the spin glasses, at least), we have studied the density of triangles as a function of N. The lower set of numbers in figure 2 shows our results for $\alpha = 12$, $\beta = 0.64$ and $N = 600$, and the upper set those for the same values of α and β but for $N = 1200$. There is clear evidence for clustering near $(d_b - d_m) = 0$, increasing with increasing N. However, while for $\alpha = 12$ the evidence for ultrametricity is very strong, this is not the case for $\alpha = 4$, as shown in figure 3, although further (and costly) examination would be needed to make a definitive statement as to whether ultrametricity is actually absent. Note that each of figures 2 and 3 corresponds to a *single* realisation of the connectivity matrix.

In conclusion, we have found empirically that several features of the Parisi solution to the SK spin-glass model, demonstrated to be applicable to optimal bipartitioning of

graphs of random connectivity with extensive valence, appear to be applicable also to random graphs of fixed finite connectivity. These include the existence of a multiplicity of metastable states of non-trivial overlap, the lack of self-averaging of the overlap distributions and the existence of ultrametricity, at least for high enough valence. We have obtained an empirical formula for the average cost function which agrees strikingly well with our simulations, with the exact result for $\alpha = 2$, and which yields the Parisi formula in the limit of extensive valence. However, there remain open questions such as the existence of a phase transition and a critical non-trivial minimum valence for ultrametricity, as well as an analytic basis for our empirical results.

Fruitful discussions with P W Anderson, B Derrida and M Mézard are gratefully acknowledged. Two of us (DS and NS) would like to thank Schlumberger-Doll Research, where the first part of this work was done. We would also like to thank the GRECO 'Expérimentation Numérique' at ENS, Paris, for further computer time support.

References

[1] Binder K and Young A P 1986 *Rev. Mod. Phys.* to be published
Van Hemmen L and Morgenstern I 1983 (ed) *Proc. Heidelberg Colloq. on Spin Glasses* (Berlin: Springer)
[2] Garey M R and Johnson D S 1979 *Computers and Intractability* (San Francisco: Freeman)
Papamidimitriou C H and Steglitz K 1982 *Combinatorial Optimization* (Englewood Cliffs, NJ: Prentice-Hall)
[3] Kirkpatrick S 1981 *Lecture Notes in Physics* **149** (Berlin: Springer) p 280
Kirkpatrick S, Gelatt C D and Vecchi M P 1983 *Science* **220** 671
[4] Mézard M, Parisi G, Sourlas N, Toulouse G and Virasoro M A 1984 *Phys. Rev. Lett.* **52** 1156; 1984 *J. Physique* **45** 843
Rammal R, Toulouse G and Virasoro M A 1986 *Rev. Mod. Phys.* **58** 765
[5] Bray A J, Moore M A and Young A P 1984 *J. Phys. C: Solid State Phys.* **17** L155
[6] Toulouse G 1977 *Commun. Phys.* **2** 115
[7] Fu Y and Anderson P W 1986 *J. Phys. A: Math. Gen.* **19** 1605
[8] Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **35** 1792
[9] Parisi G 1979 *Phys. Rev. Lett.* **43** 1754; 1980 *J. Phys. A: Math. Gen.* **13** L115, 1101, 1887
[10] Edwards S F and Anderson P W 1975 *J. Phys. F: Met. Phys.* **5** 965
[11] Mézard M and Parisi G 1985 *J. Physique Lett.* **46** L771; 1986 to be published
[12] Viana L and Bray A J 1985 *J. Phys. C: Solid State Phys.* **18** 3037
[13] Thomsen M, Thorpe M F, Choy T C, Sherrington D and Sommers H-J 1986 *Phys. Rev.* B **33** 1931
Athanasiu G G, Bachas C P and Wolff W F 1986 to be published
[14] Sourlas N 1984 *J. Physique Lett.* **45** L969
[15] Binder K 1979 *Monte Carlo Simulation in Statistical Physics* (Berlin: Springer)
[16] Parisi G 1983 *Phys. Rev. Lett.* **50** 1946