ELSEVIER

Physica A 232 (1996) 61–73

PHYSICA A

Averaging and finite-size analysis for disorder: The Hopfield model

Thomas Stiefvater[a], Klaus-Robert Müller[b,*], Reimer Kühn[c]

[a] *MPW Lasec, Anklamerstr. 32, 10115 Berlin, Germany*

[b] *GMD FIRST, Rudower Chaussee 5, 12489 Berlin, Germany*

[c] *Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 19, 69120 Heidelberg, Germany*

Received 4 March 1996

Abstract

A finite-size scaling study of the capacity problem for the Hopfield model is presented. Questions of identifying the correct shape of the scaling function, of corrections to finite-size scaling and, in particular, the problem of properly dealing with disorder are carefully addressed. At first-order phase transitions, like the one considered here, relevant physical quantities typically scale exponentially with system size, and it is argued that in disordered systems reliable information about the phase transition can therefore be obtained only by averaging their logarithm rather than by considering the logarithm of their average – an issue reminiscent of the difference between quenched and annealed disorder, but previously ignored in the problem at hand. Our data for the Hopfield model yield $\alpha_c = 0.141 \pm 0.0015$. They are thus closer to the results of a recent one- and two-step replica symmetry breaking (RSB) analysis, and disagree with that of an earlier one-step RSB study, with those of previous simulations, and with that of a recent paper using an infinite-step RSB scheme.

Keywords: Hopfield model; Disordered systems; First-order transitions; Finite-size scaling

The present contribution is concerned with a seemingly old question – that of the storage capacity α_c of the Hopfield model [1]. There was, indeed, a time when this question appeared to have been settled. Applying the so-called replica method borrowed from spin-glass theory, Amit et al. [2] obtained $\alpha_c^{RS} \simeq 0.1379$ in the replica-symmetric (RS) approximation to the full theory. Knowing that the RS results were not really thermodynamically acceptable, they also performed numerical simulations, which yielded [2] $\alpha_c^{sim} = 0.144 \pm 0.009$, and they conjectured that the discrepancy

* Corresponding author. E-mail: klaus@first.gmd.de.

PII S0378-4371(96)00134-3

between these two results might be put down to effects of replica-symmetry breaking (RSB). A subsequent 1-step replica-symmetry breaking (1RSB) analysis performed by Crisanti et al. [3] did, indeed, yield $\alpha_c^{\mathrm{1RSB}} \simeq 0.144$, so that the conjecture of Amit et al. [2] appeared to be warranted, and further steps in Parisi's hierarchical replica-symmetry breaking scheme [4] were not expected to significantly change the results. Ensuing numerical investigations of Horner et al. [5] in fact gave $\alpha_c^{\mathrm{sim}} = 0.1455 \pm 0.001$, thereby confirming the earlier simulation results, but narrowing down error bars. A large-scale simulation of Kohring [6] was interpreted to yield $\alpha_c^{\mathrm{sim}} = 0.143 \pm 0.001$, so that the overall agreement between theory [2, 3] and numerical experiment [2, 5, 6] appeared to be rather good.

The 1RSB calculation was recently redone in a wider context and supplemented by a 2RSB analysis [7]. For the Hopfield model, these authors found $\alpha_c^{\mathrm{1RSB}} \simeq 0.138186$ and $\alpha_c^{\mathrm{2RSB}} \simeq 0.138187$, and they were forced to conclude that the 1RSB result obtained earlier [3] is very likely in error. Both, the 1RSB and the 2RSB result are found to lie within the reentrance bound [7, 8] $\alpha_{\mathrm{max}}^{\mathrm{RS}}(T) \simeq 0.138189$ which denotes the capacity maximally attainable at non-zero temperatures in the RS approximation. On the basis of the Parisi–Toulouse hypothesis [9] and very strong analogies with the SK model it was conjectured [7] that a complete RSB solution, presumably providing the exact answer, would yield a $T = 0$ storage capacity not higher than $\alpha_{\mathrm{max}}^{\mathrm{RS}}(T)$ – and probably very close to that value.

At this point we should like to mention another dissenting vote, viz. that of Tokita [10], who claims $\alpha_c^{\infty\mathrm{RSB}} \simeq 0.159$, working in the RSB scheme of De Dominicis et al. [11].

The current state of affairs thus finds theory and numerical simulations, once more, in conflicting positions, rendering the old question completely open again.

In what follows, we reinvestigate the numerical side of the problem. Our main hypothesis, which we shall try to substantiate below, is that Monte-Carlo data on the capacity problem have up to now been evaluated in questionable ways. We shall explain our views as to how the finite-size scaling (FSS) analysis of the simulation data should be performed, and produce new Monte-Carlo data which, when evaluated accordingly, yield $\alpha_c \simeq 0.141 \pm 0.0015$ and are thus closer to the result of Ref. [7], whereas they clearly disagree with those of Tokita [10].

Two main issues will be dealt with: (i) the question of identifying the correct shape of the scaling function that describes the FSS-signature of the phase transition and (ii) the problem of properly dealing with the effects of disorder in the FSS analysis. While the first may be regarded as specific for the model investigated, the second is of a more fundamental nature and of *general* relevance to FSS analyses of first-order phase transitions in disordered systems. Nevertheless, for reasons to be explained below, we find that the first issue has a stronger effect on the evaluation of our simulation data than the second.

The Hopfield model of a neural network [1] is designed to store and retrieve a set of unbiased binary random patterns $\{\xi_i^\mu;\ i = 1, \ldots, N;\ \mu = 1, \ldots, p\}$ with $\xi_i^\mu \in \{\pm 1\}$, where i enumerates the neurons of the net and μ is the pattern index. The patterns $\boldsymbol{\xi}^\mu$

are encoded in couplings of the form

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu , \qquad J_{ii} = 0, \tag{1}$$

and the dynamics of the neuron activities $s_i \in \{\pm 1\}$ is given by a simple threshold rule

$$s_i(t+1) = \mathrm{sgn}(h_i(t)), \tag{2}$$

with h_i denoting the local field potential

$$h_i(t) = \sum_{j=1}^{N} J_{ij} s_j(t). \tag{3}$$

Updating in (2) may be either synchronous or asynchronous.

The Hopfield model exhibits a *discontinuous* phase transition from a state of good retrieval, where the system has stable fixed points exhibiting $\mathcal{O}(1)$ overlaps

$$m^\mu = \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu s_i \tag{4}$$

with the stored patterns, satisfying $m^\mu > 0.9$, to a spin-glass phase with fixed points showing no (or rather small remanent) overlaps with the stored patterns, as $\alpha = p/N$ is increased through α_c.

The finite-size signature of this phase transition, as it is observed in numerical simulations, can be described as follows. One follows the evolution of the system dynamics (2), taking pure pattern states as initial conditions. Each initial state evolves to a target state characterized by its final overlap m_f with the initial pure pattern state. For sufficiently large system sizes N, the distribution of final overlaps has a double peak structure, one peak concentrated at $m_f \simeq 0.95$, the other near $m_f \simeq 0.3$, with a rather clear gap between the two (see Fig. 1). As N is increased, one observes that the mass of the m_f distribution goes into the high-m_f peak, if $\alpha < \alpha_c$, whereas it goes into the low-m_f peak, if $\alpha > \alpha_c$. Since the transition is discontinuous (or first order), one expects on the basis of heuristic arguments [12] that the scaling of these trends with system size N should be exponential in N. To be specific, denoting by f the fraction of pure pattern states of one given pattern set that evolves towards the high-m_f peak, one expects the scaling

$$f \simeq \begin{cases} \exp\{a_{(>)} - N\, b_{(>)}(\alpha - \alpha_c)\} & \text{if } \alpha > \alpha_c \\ 1 - \exp\{a_{(<)} - N\, b_{(<)}(\alpha_c - \alpha)\} & \text{if } \alpha < \alpha_c. \end{cases} \tag{5}$$

Eq. (5) is believed to describe the *dominant* trends in the scaling limit $|\alpha - \alpha_c| \ll 1$, and $N|\alpha - \alpha_c| \gg 1$.

An alternative representation that lumps the $\alpha < \alpha_c$ and the $\alpha > \alpha_c$ case into a single scaling-function [13] assumes a scaling of the form

$$g = \frac{f}{1-f} \simeq \exp\{a - N\, b\,(\alpha - \alpha_c)\}. \tag{6}$$

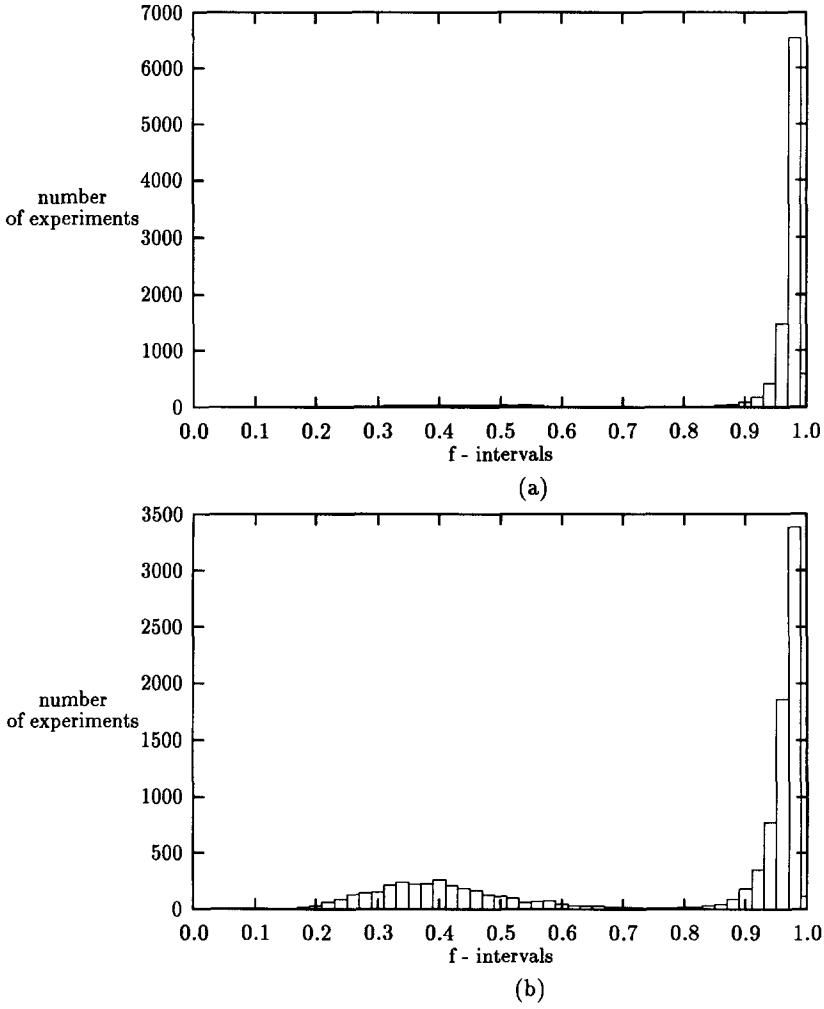
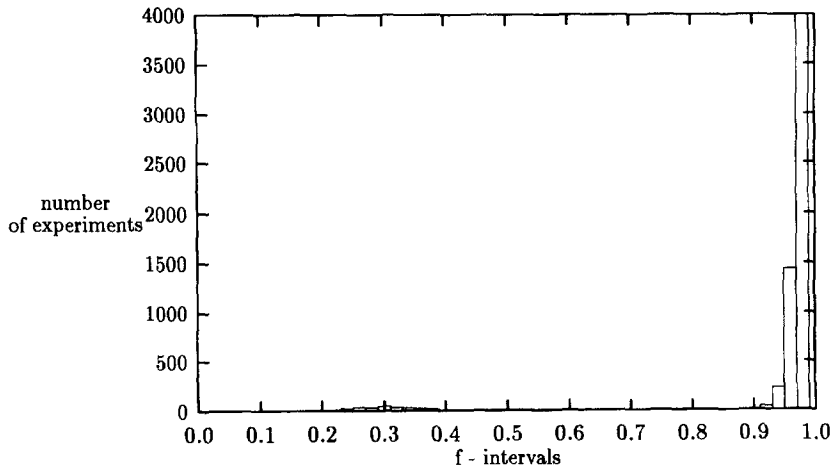


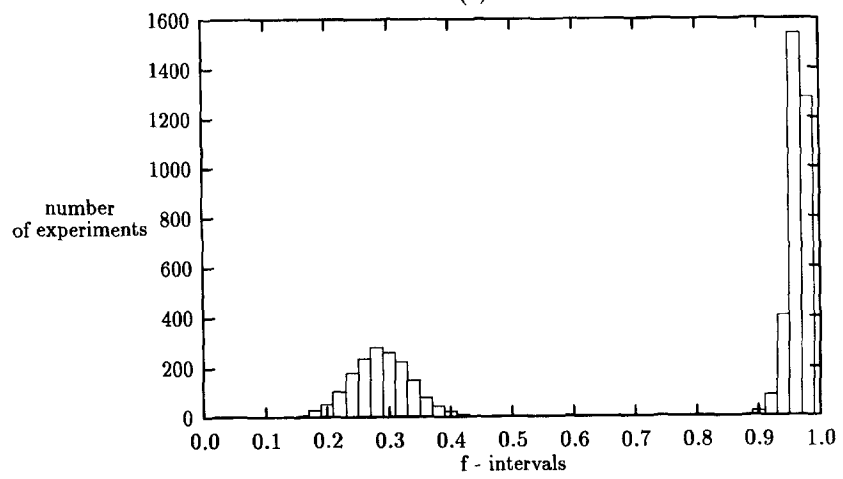
Fig. 1. Histogram of the final overlap distribution for $N = 1024$ at (a) $\alpha = 0.136$, (b) $\alpha = 0.154$ and for $N = 5120$ at (c) $\alpha = 0.136$ and (d) $\alpha = 0.146$.

If desired, different sets of constants might be used for the cases $\alpha > \alpha_c$ and $\alpha < \alpha_c$ in (6) as well. In the scaling limit $N|\alpha_c - \alpha| \gg 1$ this representation is equivalent to (5), but small corrections will be generated on transforming one into the other. More importantly, if corrections to FSS are expected to affect the data to begin with, they will be treated differently in the two approaches, and the question as to which scaling assumption provides a better modelling of the data is clearly meaningful, and it should be addressed rather than that an answer to it hypothesized.

Although the analogy with conventional temperature- or field-driven first-order phase transitions as considered in Ref. [12] is not complete, let us – like previous authors



(c)



(d)

Fig. 1. Continued.

[2, 5, 6, 13–15] – accept (5) or (6) as working hypotheses, leaving as yet open the question of which ansatz provides a better model of the data.

At a more fundamental level, our *disagreement* with previous numerical analyses is related to the question of dealing with the influence of disorder on (5) or (6). Clearly, since the Hopfield couplings are determined from a set of random patterns, one must expect f in (5) and (6) to vary randomly from pattern set to pattern set. For this reason, the question of *averaging* arises, and it is here that we would like to make our second, and main point – explaining it for the sake of definiteness for the ansatz (5) at $\alpha > \alpha_c$.

To make ourselves understood as clearly as possible, let us restate in detail and with a bit more care what the assumption (5) of FSS (at $\alpha > \alpha_c$) precisely means for the

random system at hand. FSS for this system really asserts *three* things, namely (i) that f scales exponentially with system size, i.e., accounting for $\mathcal{O}(1)$ prefactors,

$$f = f_N = \exp\{A_N - NB_N\}, \tag{7}$$

with A_N and B_N both $\mathcal{O}(1)$ and randomly varying with the pattern set actually embedded in the net – $A_N = A_N(\{\xi^\mu\})$ and $B_N = B_N(\{\xi^\mu\})$, further (ii) that both A_N and B_N, as random variables, converge to non-random quantities in the thermodynamic limit in the sense that their distributions become sharp about non-random averages: $C_{AA} = \langle A_N^2\rangle - \langle A_N\rangle^2 \to 0$, $C_{BB} = \langle B_N^2\rangle - \langle B_N\rangle^2 \to 0$ (hence $C_{AB} = \langle A_N B_N\rangle - \langle A_N\rangle\langle B_N\rangle \to 0$ by virtue of the Schwarz inequality $|C_{AB}| \leqslant \sqrt{C_{AA}C_{BB}}$), finally (iii) that the α-dependent averages behave according to $a_N(\alpha) := \langle A_N\rangle \to a_N$ and $b_N(\alpha) := \langle B_N\rangle \sim b_N(\alpha - \alpha_c)$, as $\alpha \searrow \alpha_c$, with constants a_N and $b_N > 0$ which tend to finite limits a and b as N becomes large. In particular, $b_N(\alpha)$ approaches zero from above as $\alpha \searrow \alpha_c$, and it is this fact which is used to locate α_c.

Keeping these three points in mind, we are now able to discuss the averaging issue.

It appears that previous authors have in their FSS analyses of the capacity problem invariably averaged f itself[1] and *assumed*

$$\lim_{N\to\infty} N^{-1}\log\langle f\rangle = -\tilde{b}(\alpha - \alpha_c) \tag{8}$$

with $\tilde{b}$ constant in the sense that it does not depend on N or α. In what follows, we argue that this assumption is generally *wrong*, even if A_N and B_N in (7) have *all* properties assumed in points (i)–(iii). To see this, note that precisely *because* of the exponential scaling of f with N, one should average $\log f$ instead of f itself, in order to get *typical* results. Indeed, for sufficiently large N, the average of f itself is likely to be dominated by rare events for which $B_N = B_N(\{\xi^\mu\})$ is unusually small, and it cannot be excluded that these untypical events *invalidate* the assumption expressed in (8) that $\lim_{N\to\infty} N^{-1}\log\langle f\rangle$ approaches zero at the true limit of capacity, and thereby invalidate all results based on that assumption. This holds true even if A_N and B_N are self-averaging in the thermodynamic limit and satisfy point (iii) above.

Averaging $\log f$ instead, one obtains

$$\lim_{N\to\infty} N^{-1}\langle\log f\rangle = -b(\alpha - \alpha_c) \tag{9}$$

by virtue of (i)–(iii) – a quantity which does, indeed, approach zero as $\alpha \to \alpha_c$.

It is perhaps worth pointing out that this issue is very well known in the statistical mechanics of disordered systems, where it is the partition function which scales exponentially with system size, $Z_N = \exp\{-N\psi_N\}$, with a (dimensionless) free energy ψ_N which is random but self-averaging in the thermodynamic limit. That is, ψ_N as a random variable converges, as $N \to \infty$, to its configuration average, called the *quenched* free energy: $\psi_N \to \psi_q = \lim_{N\to\infty}\langle\psi_N\rangle = -\lim_{N\to\infty} N^{-1}\langle\log Z_N\rangle$. Note that averaging

[1] This cannot in all cases be reconstructed with certainty from the texts of published papers. We would like to thank H. Horner, B. Forrest, G. Kohring and D.J. Amit for helpful correspondence regarding this point.

the *logarithm* of a quantity scaling exponentially with system size N is required to get the typical result for the free energy, viz., ψ_q. Had one, on the other hand, interchanged averaging and computation of the logarithm, one would have computed the *annealed* free energy $\psi_a = -\lim_{N\to\infty} N^{-1}\log\langle Z_N\rangle$, which is generally *smaller* than the quenched free energy, $\psi_a < \psi_q$.

In the capacity problem, the quantity f of (7) is the analog of the partition function Z of a disordered system, and the quantity B_N introduced in (i) above is, for large N, the analog of the (dimensionless) free energy. Assertion (iii) of FSS implies that the proper analog of the quenched free energy, $b_N(\alpha)$ approaches zero from above, as $\alpha \searrow \alpha_c$. Since – by Jensens inequality – the analog of the *annealed* free energy obtained by computing $\log\langle f\rangle$ is generally *smaller* than its quenched counterpart, it may be expected to cross zero already at larger values of α, thus leading to an *overestimation* of α_c! So much can be said in general, since we do not a priori know the precise form of the A_N- and B_N-distribution in (7).

With one additional assumption, which should however produce something typical for the dominant effect in general, a quick calculation can illustrate our point *quantitatively*. Assume that A_N and B_N in (7) are correlated *Gaussians* with $C_{AA} = \tilde{C}_{AA}/N$, $C_{BB} = \tilde{C}_{BB}/N$, and $C_{AB} = \tilde{C}_{AB}/N$, so that the distributions of A_N and B_N become sharp in the thermodynamic limit. The scaled elements $\tilde{C}_{AA}$, etc. of the covariance matrix may, of course, still depend on α and (weakly) on N. Then the result obtained by computing $\langle\log f\rangle$ is given by (9), whereas points (i)–(iii) together with our assumption about the distribution of the A_N and the B_N imply that

$$\lim_{N\to\infty} N^{-1}\log\langle f\rangle = -b(\alpha-\alpha_c) + \tfrac{1}{2}\tilde{C}_{BB}, \tag{10}$$

which is, indeed, different from (8), confirming our assertion that the assumption expressed in that equation would be drastically wrong in this case, *despite* the fact that we had assumed A_N and B_N to be self-averaging in the thermodynamic limit. Eq. (10) expresses the said influence of untypical events quantitatively for the Gaussian case. In particular, it also shows that the proper analog of the annealed free energy for the capacity problem approaches zero from above at

$$\alpha_0 = \alpha_c + \frac{1}{2b}\tilde{C}_{BB}. \tag{11}$$

This value would erroneously be taken to denote the storage capacity of the model, if the evaluation is based on an analysis of $\log\langle f\rangle$ data, as in Refs. [2, 5, 6] and [13–15], but it is obviously *greater* than the true α_c, even in the thermodynamic limit. Note therefore that simulating very large systems would not help to avoid wrong results for the capacity α_c, if the evaluation is based on (8).

The same arguments apply to the other cases in (5) and (6): Taking the logarithm of the average instead of averaging the logarithm in quantities scaling exponentially in system size N will produce spurious results. Moreover, on transforming (5) into (6) or vice versa, higher order corrections in $\alpha - \alpha_c$ or $1/N$ to b are generated as well (because a_N, b_N, and the scaled elements covariance matrix will generally depend on α,

hence $\alpha - \alpha_c$). This holds also, if, starting from (6), averaging and function evaluation are exchanged through a *second* level by evaluating $\log[\langle f\rangle/(1-\langle f\rangle)]$ instead of $\langle\log[f/(1-f)]\rangle$, as has been done in Refs. 13–15. Incidentally, since this implies an exchange of averaging and function evaluation through two levels, of which the first involves a concave function and the second a convex function, the errors made at the first level will be partially compensated by those made at the second level.

We believe that the failure of previous investigators [2, 5, 6, 13–15] to notice these effects has lead to erroneous numerical estimates for the storage capacity α_c, and that a proper evaluation might reconcile theoretical and numerical results on this problem. In what follows we report results of a MC study of networks containing up to $N = 5120$ neurons, which yields $\alpha_c = 0.141 \pm 0.0015$.

Let us mention a few points related to achieving a high efficiency in our simulation. First, we used parallel dynamics, because some of the points mentioned below will give a speed-up only for this case. Asynchronous and parallel dynamics have the same fixed points; in parallel dynamics there are also two-cycles. Where they occur, we find that they involve only a tiny fraction of all neurons, so they never switch between the high-m_f and the low-m_f peak. Moreover, if one turns to asynchronous dynamics when a two-cycle has been reached, there is always a fixed-point of the asynchronous dynamics "nearby". Next, we do not store couplings and write h_i in terms of overlaps

$$h_i = \sum_{j\neq i} J_{ij}s_j = \sum_{j\neq i}\frac{1}{N}\sum_{\mu}\xi_i^{\mu}\xi_j^{\mu}s_j = \frac{1}{N}\sum_{\mu}\xi_i^{\mu}(Nm^{\mu} - \xi_i^{\mu}s_i). \tag{12}$$

Bit-coding techniques are used to store the ξ_i^{μ} and the s_i. Since we are interested only in the sign of h_i and $\mathrm{sgn}(h_i) = \mathrm{sgn}(Nh_i)$, integer or logical operations can be used throughout the simulation. Up to this point, our strategy is similar to that proposed by Penna and Oliveira [16]. An improvement over their approach is obtained by noting that the states s_i change drastically only within the first few steps. Hence updates of the overlaps after each time step are efficiently performed once the set I of nodes at which changes do have occurred has been identified,

$$\begin{aligned} Nm^{\mu}(t) &= \sum_{i=1}^{N}\xi_i^{\mu}s_i(t) \\ &= Nm^{\mu}(t-1) + \sum_{i\in I}\xi_i^{\mu}(s_i(t) - s_i(t-1)). \end{aligned} \tag{13}$$

As everything is bitcoded, the search for changes in the state vector s can be done word-wise. Within a word in which changes are detected, changed bits (set bits resulting from an XOR operation) are read out sequentially. Still a significant speed-up is obtained – in particular, as the size of the set I of nodes in which changes do occur *decreases* as the dynamics proceeds and converges to a fixed point. So the numerical effort to update the overlaps according to (13) becomes progressively smaller as the dynamics is homing on a stable configuration. This feature is not available in the approach of Ref. [16]. However, it is particularly useful in our case, since we are simulating above

α_c and the dynamics is approaching fixed points *slowly*. Quantitatively, the gain in efficiency is found to depend on the loading level, increasing with the distance from α_c. To give an order of magnitude, for the system sizes and loading levels simulated in the present study (see below) there are typically $\mathcal{O}(10)$ parallel updates – making up for the larger fraction of the total number of updates needed to find a stable configuration– in which only a tiny fraction of nodes, viz. 3% or less, do actually change.

We have simulated systems with sizes ranging from $N = 1024$ to $N = 5120$, for a range of α-values close to the anticipated α_c, always using 100 patterns from each pattern set to measure f. Results are averaged over 200 pattern sets for $N = 1024$, over 120 pattern sets for $N = 2048$, and over 60 pattern sets for $N = 3072$, $N = 4096$, and $N = 5120$.

From the $\langle \log f \rangle$ and $\log\langle f \rangle$ data at $\alpha > \alpha_c$, it follows that the interchange of averaging and log-evaluation has a smaller effect than anticipated, basically because the scaling variable $N(\alpha - \alpha_c)$ is still rather small for the loading levels and system sizes simulated; differences are barely visible in graphical representations, so in Fig. 2 we only reproduce the $\langle \log f \rangle$ data. The difference is more clearly visible, albeit small, in the $\langle \log[f/(1-f)] \rangle$ and the $\log\langle f/(1-f) \rangle$ data of Fig. 3, because the $(1-f)^{-1}$ for not too small f provides an amplifying mechanism. That is, the effect *exists*.

The $\langle \log f \rangle$ and the $\log\langle f \rangle$ data together imply that higher powers of $\alpha - \alpha_c$ would be needed in the scaling ansatz (5) to begin with (see Fig. 2). That is, Eq. (5) ought to be replaced by

$$f \simeq \begin{cases} \exp\{a_{(>)} - N\, b_{(>)}(\alpha - \alpha_c) - N\, c_{(>)}(\alpha - \alpha_c)^2\} & \text{if } \alpha > \alpha_c, \\ 1 - \exp\{a_{(<)} - N\, b_{(<)}(\alpha_c - \alpha) - N\, c_{(<)}(\alpha_c - \alpha)^2\} & \text{if } \alpha < \alpha_c, \end{cases} \tag{14}$$

where it is understood that averaging is being dealt with in the correct way. The $\langle \log[f/(1-f)] \rangle$ data on the other hand are well fitted by (6). Both, the $\langle \log f \rangle$ and the $\langle \log[f/(1-f)] \rangle$ curves, do not intersect exactly at one point, but rather in a sequence of points at α values $\alpha_{N,N'}$ between $\alpha \simeq 0.14$ and $\alpha \simeq 0.142$. For large N and N' the $\alpha_{N,N'}$ should converge to α_c which, in principle, provides a first method to locate α_c. However, since convergence is nonmonotonic for the smaller system sizes, we use this fact only for a first orientation. From the fact that not all curves intersect in the same point, it can immediately be concluded that there is indeed an N dependence in the average a_N, i.e., there are corrections to FSS. Similar dependences are found in b_N (and c_N in (14)), although they cannot be read off directly from the figures. The data are consistent with $1/N$ corrections of the form $a_N = \langle A_N \rangle = a + a_{(1)}/N$ (and similarly for b_N and c_N). Given that (6) does provide a good description of the data, the appearance of higher order $(\alpha - \alpha_c)$ corrections, both, in the $\langle \log f \rangle$ and the $\log\langle f \rangle$ data, is not unexpected and is, indeed, observed.

The unknown constants characterizing the properly averaged scaling functions are extracted from a set of simulations at various α and N. In practice, *two* values of α, each simulated for a range of system sizes N which are large enough to allow a safe extrapolation of $\phi_N(\alpha) := -\frac{1}{N}\langle \log[f/(1-f)] \rangle$ to $N = \infty$, the result of which is

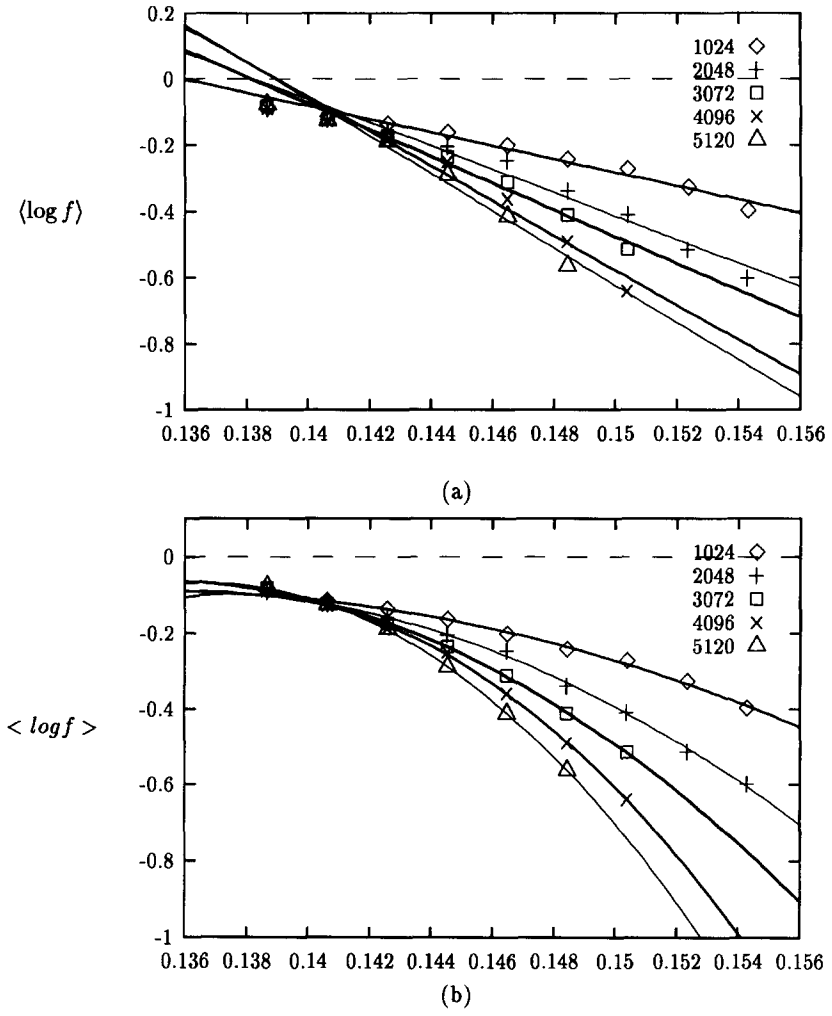
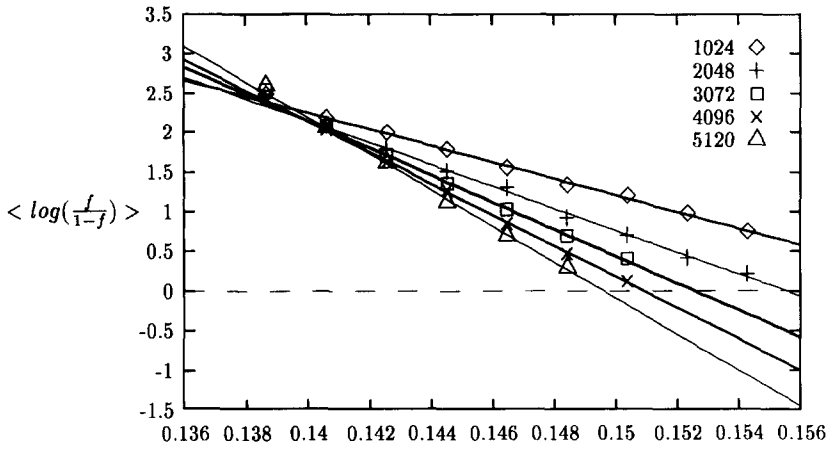


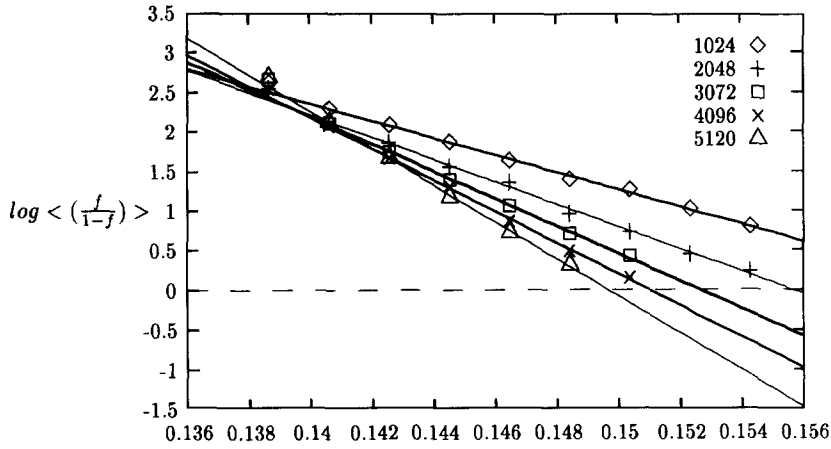
Fig. 2. (a) straight line and (b) quadratic interpolation of $\langle \log f \rangle$ versus memory loading $\alpha = p/N$. The extrapolation for $N \to \infty$ yields (a) $\alpha_c = 0.139 \pm 0.0015$ and (b) $\alpha_c = 0.141 \pm 0.0015$.

$b(\alpha - \alpha_c)$ – with *two* unknown constants, b and α_c – suffice to determine α_c (see Fig. 4). The $\phi_N(\alpha)$ are read off from the interpolated curves in Fig. 3, in order to reduce the (statistical) error. If one uses the $\langle \log f \rangle$ data and the scaling ansatz (14) instead, one needs three different values of α to fix all three unknown constants, b, c, and α_c. *Both* procedures yield $\alpha_c = 0.141 \pm 0.0015$.

If one – erroneously – uses an exponential scaling fit to the $\log\langle f/(1-f)\rangle$ data, the value for α_c increases slightly to 0.142 ± 0.0015. Failure to include the $\mathcal{O}((\alpha_c - \alpha)^2)$ contribution in the exponential fit to the $\langle \log f \rangle$ data, on the other hand, yields an α_c value of 0.139 ± 0.0015, which we believe to be too low – despite its better agreement with the results of Ref. [7].



(a)



(b)

Fig. 3. (a) straight line interpolation for (a) $\left\langle \log \frac{f}{1-f} \right\rangle$ and (b) $\log \left\langle \frac{f}{1-f} \right\rangle$ versus memory loading $\alpha = p/N$. The extrapolation yields (a) $\alpha_c = 0.141 \pm 0.0015$ and (b) $\alpha_c = 0.142 \pm 0.0015$.

In summary, a FSS study of the capacity problem for the Hopfield model has been presented, in which the questions of corrections to FSS, of identifying the correct shape of the scaling function, and of properly dealing with the disorder averaging problem are carefully discussed. The latter, in particular, is of general relevance for FSS analyses of first-order phase transitions in disordered systems in which quantities scaling exponentially with system size *typically* occur [12]. While the system-sizes simulated in the present paper are certainly to be termed "moderate" at best, we believe our statistics to be good enough for a reliable analysis. Indeed, our key point is that even extraordinarily large systems would not have helped to avoid wrong conclusions, had they been evaluated "the old way".

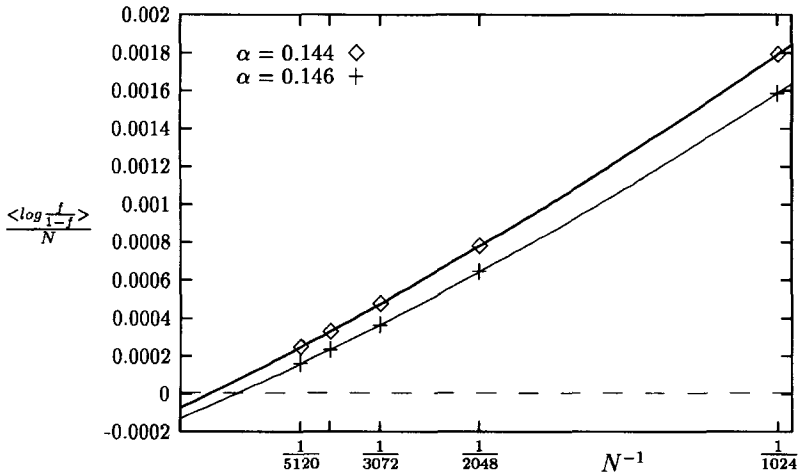


Fig. 4. Extrapolation for $N \to \infty$. Two intercepts with the vertical axes allow the determination of the remaining unknown constants α_c, b (see also Eq. (6)). We obtain $\alpha_c = 0.141 \pm 0.0015$.

Our results significantly reduce the discrepancy between earlier simulation results [2, 5, 6] and those of a recent one- and two-step RSB study [7], and they clearly disagree with those of an earlier one-step RSB analysis of Crisanti et al. [3] and the infinite-step RSB result of Tokita [10]. Whether the remaining small discrepancy between our results and those of Ref. [7] is significant or not, cannot be assessed with absolute certainty, but we are inclined to believe that it is. Possible mechanisms to explain the gap that remains are the following. Our $T = 0$ simulation is not the same as the $T \to 0$ limit of a thermodynamic analysis to which the replica theory refers. The $T = 0$ Monte-Carlo dynamics might after all get trapped by finite energy barriers which are not seen in thermodynamics. On the other hand, for first-order phase transitions in spin-glass-like systems, there may also exist true dynamic freezing transitions (even at nonzero temperatures) at values different from those predicted by replica theory (and we are *not* talking of spinodals here), such as have recently been found in dynamical analyses of the capacity problem for the binary perceptron [17] and of the spherical p-spin interaction spin-glass [18], and we may well have seen the effects of such a true freezing transition in our simulations. One way to clarify this question might be to investigate the capacity problem for low activity systems, for which the discrepancy between existing simulation results [13] – evaluated "the old way" – and predictions from replica analyses [19, 7] are much stronger.

Acknowledgements

It is a pleasure to thank Helmut Steffan and Heinz Horner for stimulating discussions. RK is supported by a Heisenberg fellowship. Moreover, he would like to thank the

Instituut voor Theoretische Fysica at K.U. Leuven and GMD FIRST for hospitality while parts of the paper were being written.

References

[1] J.J. Hopfield, Proc. Natl. Acad. Sci. USA 79 (1982) 2554 .
[2] D.J. Amit, H. Gutfreund and H. Sompolinsky, Phys. Rev. Lett. 55 (1985) 1530; Ann. Phys. NY 173 (1987) 30.
[3] A. Crisanti, D.J. Amit and H. Gutfreund, Europhys. Lett. 2 (1986) 337.
[4] G. Parisi, J. Phys. A 13 (1980) L115; J. Phys. A. (1980) 1101.
[5] H. Horner, D. Bormann, M. Frick, H. Kinzelbach and A. Schmidt, Z. Phys. B 76 (1989) 381.
[6] G.A. Kohring, J. Stat. Phys. 59 (1990) 1077.
[7] H. Steffan, PhD thesis, Heidelberg, unpublished (1993); H. Steffan and R. Kühn, Z. Phys. B 95 (1994) 249.
[8] J.P. Naef and A. Canning, J. Physique 2 (1992) 247.
[9] G. Parisi and G. Toulouse, J. Physique Lett. 41 (1980) L361.
[10] K. Tokita, J. Phys. A 27 (1994) 4413.
[11] C. De Dominicis, M. Gabay and H. Orland, J. Physique Lett. 42 (1981) L523.
[12] M.E. Fisher and V.J. Privman, Stat. Phys. 33 (1983) 385; K. Binder and D.P. Landau, Phys. Rev. B 30 (1984) 1477.
[13] B.M. Forrest and A. Loettgers, Z. Phys. B 86 (1992) 309.
[14] K.-R. Müller, PhD Thesis, Karlsruhe, (1992), published as GMD Bericht 218, R. Oldenbourg Verlag, München/Wien (1994).
[15] T. Stiefvater and K.-R. Müller, J. Phys. A 25 (1992) 5919.
[16] T.J.P. Penna and P.M.C. Oliveira, J. Phys. A 22 (1989) L719.
[17] H. Horner, Z. Phys. B 86 (1992) 291.
[18] A. Crisanti and H.J. Sommers, Z. Phys. B 87 (1992) 341; A. Crisanti, H. Horner and H.J. Sommers, Z. Phys. B 92 (1993) 257.
[19] H. Horner, Z. Phys. B 75 (1989) 133.