New Method for Studying the Dynamics of Disordered Spin Systems without Finite-Size Effects

H. Eissfeller and M. Opper
Institut für Theoretische Physik III, Justus-Liebig-Universität Giessen, Heinrich-Buff-Ring 16, W-6300 Giessen, Federal Republic of Germany
(Received 14 October 1991)

A new method is presented which allows the determination of dynamical properties of disordered spin systems avoiding finite-size effects. The method is based on exact dynamical mean-field equations for the infinite large system. The resulting single-spin dynamics is solved by Monte Carlo simulations. We outline the formalism for the parallel dynamics of a fully connected model with random couplings. The decay of remanent magnetization of the model is studied. We find a power-law decay: $m(t) - m(\infty) \propto t^{-a}$, with $a = 0.474$ and $m(\infty) = 0.184$ for the infinite system.

PACS numbers: 87.10.+e

Networks composed of spinlike elements which interact via long-range random interactions are of growing interest. Such models served as approximations for disordered materials as, for example, the Sherrington-Kirkpatrick (SK) model for spin glasses [1]. More recently similar systems have become popular as models for neural networks [2].

The equilibrium behavior of such random spin systems is now well understood by applying the replica approach to determine thermal averages for disordered systems [3].

Because of the lack of ergodicity these models show many interesting features that are essentially of *nonequilibrium* nature. Usually the dynamics at low temperatures will reach a state which will strongly depend on its initial conditions. This property made networks of spins useful for the models of associative memories [2].

To study the dynamics of fully connected spin models, there are mainly two different approaches. The first one is to calculate the time evolution of dynamic quantities analytically using the so-called *dynamical functional method* [4]. This method has been successfully applied to the behavior of disordered spin systems at long time scales being able to recover and understand the equilibrium results mentioned above from a purely dynamical viewpoint [5–7].

Unfortunately the situation is less satisfactory for the nonequilibrium, transient behavior. Apart from approximate treatments [8,9] exact calculations are only possible for very few time steps, e.g., up to four time steps for the SK model and two time steps for the Hopfield model [10].

So most of the studies on dynamics of disordered spin systems rest on numerical simulations, which is the other possible approach. However, there are strong finite-size effects, which for example do not even allow a decision if there is a finite remanent magnetization for the infinite spin system. In this Letter we propose a new method to avoid these finite-size problems. Our method combines the dynamical functional method, which allows us to perform the limit $N \to \infty$ exactly, and a Monte Carlo simulation of the resulting stochastic one-particle equations. The method is demonstrated for a disordered spin system with synchronous dynamics.

This model consists of N Ising spins $S_i = \pm 1$, where every spin S_i is connected to all other spins S_j by couplings J_{ij}, which are Gaussian random variables with distribution $P(J_{ij}) = \sqrt{N/2\pi}\exp(-\frac{1}{2}NJ_{ij}^2)$. Additionally the symmetry of the couplings is given by the symmetry parameter η:

$$[J_{ij}J_{ji}] = \eta/N . \tag{1}$$

This means the couplings are fully antisymmetric ($J_{ij} = -J_{ji}$) for $\eta = -1$ and fully symmetric ($J_{ij} = J_{ji}$) for $\eta = 1$. In the latter case the model corresponds to the Sherrington-Kirkpatrick model.

These couplings can be constructed from

$$J_{ij} = \sqrt{(1+\eta)/2}\,J_{ij}^s + \sqrt{(1-\eta)/2}\,J_{ij}^a \tag{2}$$

with $J_{ij}^s = J_{ji}^s$ and $J_{ij}^a = -J_{ji}^a$, where J_{ji}^s and J_{ji}^a are now independent random Gaussian variables for all i and j. We shall restrict ourselves to the simplest type of noise free dynamics where all spins are updated in parallel:

$$S_i(t+1) = \mathrm{sgn}[h_i(t)], \quad i = 1, \ldots, N . \tag{3}$$

The internal field of the spin S_i is given by

$$h_i(t) = \sum_{j \neq i} J_{ij} S_j(t) . \tag{4}$$

Instead of simulating the system (3), we follow the dynamic functional approach mentioned above. It uses the fact that average values of dynamical quantities can be obtained from a *generating function*:

$$[Z(t)]_J = \left[\mathrm{Tr}_{S_i(t)} \int \prod_{i,t} \left\{ dh_i(t)\Theta(S_i(t+1)h_i(t))\delta\left(h_i(t) - \sum_{j\neq i}[\sqrt{(1+\eta)/2}\,J_{ij}^s + \sqrt{(1-\eta)/2}\,J_{ij}^a]S_j(t)\right)\right\} \times \exp\left(i\sum_{i,t} l_i(t)h_i(t)\right)\right]_J , \tag{5}$$

where $[\cdots]_J$ denotes an average over the random couplings. The normalization $[Z(0)]_J=1$ holds for the generating function.

Following the usual approach of dynamical mean-field theory [7] we perform the average over the J_{ij} and transform the remaining expression using saddle-point methods, which are exact in the limit $N\to\infty$. The result can be written in the following form:

$$[Z(I)]_J=\left[\mathrm{Tr}_{S_i(t)}\int\prod_{i,t}\left\{dh_i(t)\Theta(S_i(t+1)h_i(t))\delta\left(h_i(t)-\phi_i(t)-\eta\sum_s K_{ts}S_i(s)\right)\right\}\exp\left(i\sum_{i,t}l_i(t)h_i(t)\right)\right]_\phi . \tag{6}$$

Here $[\cdots]_\phi$ denotes an average over the time-dependent Gaussian random variables $\phi_i(t)$. The K_{ts} represent time-dependent order parameters. Generating functions can be used to derive equations of motion for various averaged quantities. In equilibrium these could be calculated, e.g., by means of perturbation expansions or, if a fluctuation dissipation theorem [7] is valid, even exactly.

To treat the nonequilibrium case, we directly use the fact that (6) describes a new system of *stochastic* dynamical equations which reads

$$\begin{aligned} S_i(t+1)&=\mathrm{sgn}[h_i(t)]\,,\\ h_i(t)&=\phi_i(t)+\eta\sum_s K_{ts}S_i(s)\,. \end{aligned} \tag{7}$$

The time correlations of the Gaussian noise variables are prescribed by saddle-point equations as

$$[\phi_i(s)\phi_j(\tau)]_\phi=\delta_{ij}C_{s\tau}=\delta_{ij}[S_i(s)S_i(\tau)]_\phi . \tag{8}$$

In contrast to the original system (3) the Gaussian random variables are *uncorrelated* on different sites i. Thus the dynamical equations (7) are no longer coupled to each other.

The coefficients K_{ts} are obtained self-consistently from

$$K_{ts}=\left[\frac{\partial S(t)}{\partial\phi(s)}\right]_\phi , \tag{9}$$

where the index "i" was omitted since the average does not depend on the site.

We are left with a "one-particle" equation for each spin $S(t)$. The internal field of S at time t consists of the Gaussian noise variable at the same time and a part which describes a *deterministic* coupling to the spin at the same *site* but at all previous time steps [$S(s)$ for $s<t$]. The time-dependent order parameters K_{ts} determine the strength of these couplings. Obviously $K_{ts}=0$ for $s\geq t$ because of causality.

The values of all spins [$S(s)$ for $s\leq t$] at a given site can be replaced by functions of their internal fields at one time step before. Doing this recursively, we will end up with an explicit function for $S(t+1)$, depending only on the Gaussian noise variables $\phi(s)$ for $s\leq t$. Given the values of K_{ts} for $s<t$, and correlations $C_{\tau,s}$ for $\tau,s\leq t$, averages for all dynamical quantities (spins, internal fields, etc.) defined at an arbitrary site i can be found for the next time $t+1$ from one *single* equation of (7) without any finite-size effects. Since $K_{t+1,s}$ and $C_{t+1,s}$ themselves are averages of this type, the entire procedure is *self-consistent*.

Dynamical mean-field equations for disordered systems could be solved analytically only in rather limiting cases. These include models where the equations are essentially linear (see, for example, [11]), or the case of completely asymmetric interactions ($\eta=0$), where the complicated memory terms vanish [12–15].

In this Letter we treat the nonlinear case (7) for general η by performing the averages via direct Monte Carlo simulations of the stochastic single spin dynamics (7). We generate a large number, say N_T, of trajectories for the spin variables at only *one single site i*. N_T should *not be confused with N, the number of spins in the network*, which equals infinity.

Expression (9) in its present form is not suitable for simulations. Since the $\phi(s)$ are Gaussian variables and $S(t)$ is only a function of these, we can apply the following identity (see, for example, van Kampen, p. 26 of [16]):

$$[\phi_\tau S_t]_\phi=\sum_s K_{ts}[\phi_s\phi_\tau]_\phi . \tag{10}$$

For a given time t this is a system of linear equations which determines the coefficients K_{ts} for $s=0,\ldots,t-1$. Now we are able to describe the algorithm for the Monte Carlo simulation of the single spin equation.

(1) Start at $t=0$.—(a) Set $S^k(0)=1$ for all $k=1,\ldots,N_T$, where N_T denotes the number of spin trajectories. (b) Set $h^k(0)=\phi^k(0)$, where the $\phi^k(0)$ are drawn independently from the distribution $P(\phi^k(0))=1/\sqrt{2\pi}\exp\{-[\phi^k(0)]^2/2\}$.

(2) An arbitrary time step t.—(a) Evaluate the spins at time t from the dynamical equations

$$S^k(t)=\mathrm{sgn}[h^k(t-1)]\ \text{for}\ k=1,\ldots,N_T . \tag{11}$$

Note that the internal field $h^k(t-1)$ was already calculated at time step $t-1$. (b) Calculate the sample averages over all trajectories,

$$\overline{S_tS_\tau}\ \text{for}\ \tau=0,\ldots,t-1\,, \tag{12}$$

which give the correlation matrix for the Gaussian noise variables:

$$\hat{\Phi}_t=\begin{pmatrix}\overline{\phi_0\phi_0} & \overline{\phi_1\phi_0} & \cdots & \overline{\phi_t\phi_0}\\ \overline{\phi_1\phi_0} & \overline{\phi_1\phi_1} & & \\ \vdots & & \ddots & \vdots\\ \overline{\phi_t\phi_0} & \cdots & & \overline{\phi_t\phi_t}\end{pmatrix}=\begin{pmatrix}1 & \overline{S_1S_0} & \cdots & \overline{S_tS_0}\\ \overline{S_1S_0} & 1 & & \\ \vdots & & \ddots & \vdots\\ \overline{S_tS_0} & \cdots & & 1\end{pmatrix} . \tag{13}$$

(c) Now perform the decomposition $\hat{\Phi}_t = \hat{A}_t \cdot \hat{A}_t^T$, where $\hat{A}_t$ is a triangular matrix. (d) Draw the components of the vector $\mathbf{N}_t = (n_0, n_1, \ldots, n_t)^T$ independently from a normal distribution. Transform this vector according to

$$\boldsymbol{\phi}_t = \hat{A}_t \cdot \mathbf{N}_t , \tag{14}$$

which gives the right correlations for the components of $\boldsymbol{\phi}_t$. (e) Compute the sample averages $\overline{\phi_\tau S_t}$ for $\tau = 0, \ldots, t$. (f) Obtain the coefficients $K_{t0}, \ldots, K_{tt-1}$ by solving the system of linear equations

$$\mathbf{x}_t = \mathbf{K}_t \cdot \hat{\Phi}_t , \tag{15}$$

with

$$\mathbf{x}_t = (\overline{\phi_0 S_t}, \ldots, \overline{\phi_t S_t})^T, \qquad \mathbf{K}_t = (K_{t0}, K_{t1}, \ldots, K_{tt-1}, 0)^T . \tag{16}$$

(g) Determine the internal fields from

$$h^k(t) = \phi^k(t) + \eta \sum_{\tau < t} K_{t\tau} S^k(\tau) . \tag{17}$$

(3) Iterate the steps 2(*a*)-2(*g*).

With the algorithm described above, we are immediately able to calculate all averages of dynamical quantities of the model, which depend on a single spin. As an example, we estimated the temporal behavior of the magnetization $m(t) = \overline{S(t)S(0)}$, for $\eta = 1$, which measures the memory of the system to its initial conditions.

For the magnetization after two and four time steps we get

$$m(2) = 0.574 \pm 0.001, \quad m(4) = 0.469 \pm 0.001 , \tag{18}$$

which is good agreement with the values obtained by Gardner, Derrida, and Mottishaw [10] from an analytical calculation:

$$m(2) = 0.575, \quad m(4) = 0.468 . \tag{19}$$

Of special interest in the limiting value of $m(t)$ as $t \to \infty$, the remanent magnetization m_r. Especially for this quantity finite-size effects were found to be very strong [17-20] and it is still unclear what is the correct finite-size scaling function to be used. So the extrapolated value $m_r(\infty)$ obtained from finite-size scaling changed over the years by going to larger maximum system sizes.

We performed simulations of the one-particle equations up to $t = 100$ time steps and $N_T = 1\,000\,000$ spin trajectories. The data are plotted in Fig. 1 (open squares). It turned out that the decay of magnetization can be fitted very well by a power law:

$$m(t) - m_r(\infty) \propto t^{-a} , \tag{20}$$

with

$$m_r(\infty) = 0.184 \pm 0.002, \quad a = 0.474 \pm 0.005 . \tag{21}$$

This corresponds to the curve in Fig. 1. The resulting value of $m_r(\infty)$ in the thermodynamic limit is in very good agreement with the extrapolated value for the up to now largest systems investigated by Kohring and Schreckenberg [20]. Again note that these results do not contain any uncontrolled dependencies on the system size N as all previous results. We only have statistical errors, which originate in the finite number of trajectories N_T. Since these trajectories are statistically independent the errors are expected to be of order $1/N_T^{1/2}$. In fact, a comparison of our numerical estimates for $K_{t,s}$ for all odd time differences $t - s$ with their exact values, which were found to be zero in Ref. [10], confirms this behavior.

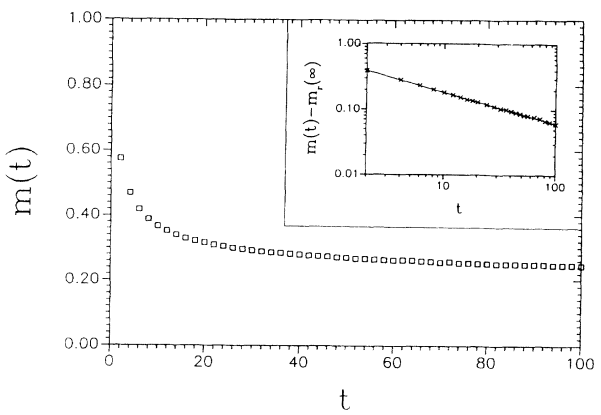

FIG. 1. Decay of the magnetization for the neural network model with random couplings and $\eta = 1$, which corresponds to the SK model. The data are shown (open squares) for the first 100 time steps. Each point is an average over $N_T = 1\,000\,000$ spin trajectories. The statistical errors are much smaller than the symbol sizes. Inset: A log-log plot of these data. The straight line corresponds to the fitted power law.

Because of the uncontrolled errors in finite system simulations we cannot compare the CPU times used by simulations for finite systems and the algorithm presented here. Nevertheless we would like to mention that our simulations took only 6 h of CPU time on a Cray Y-MP, which we think indicates that the method can be implemented efficiently in simulation practice. So we think the new method passed the first test.

Further results, like the dependence of the remanent magnetization on the symmetry η will be published in a forthcoming paper together with a more detailed analytical derivation of the one-particle equations. Clearly our approach is not limited to the deterministic dynamics (3). We can easily include a fast external noise to mimic the effects of temperature. Presently it is not clear whether, by an investigation of the dynamical functional developed by Sompolinsky, Crisanti, and Sommers [13,21], one can derive a similar single-spin equation for random sequential update of the spins.

In general our method will be applicable to all systems with an infinite range of interactions such that a mean-field treatment becomes exact. This includes most of the models of attractor neural networks. Other interesting

candidates for applications will be the dynamics of learning algorithms for feedforward networks [22] and other optimization problems.

The authors like to thank R. D. Henkel, M. Biehl, and W. Kinzel for many stimulating discussions. The numerical simulations were carried out on the Cray Y-MP of the HLRZ Jülich. The work was supported by grants of the Deutsche Forschungsgemeinschaft.

[1] S. Kirkpatrick and D. Sherrington, Phys. Rev. B **17**, 4384 (1978).
[2] J. J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).
[3] G. Parisi, J. Phys. A **13**, 1101 (1980).
[4] C. de Dominicis, Phys. Rev. B **18**, 4913 (1978).
[5] H. Sompolinsky, Phys. Rev. Lett. **47**, 935 (1981).
[6] H. Horner, Z. Phys. B **57**, 29 (1984).
[7] H. Sompolinsky and A. Zippelius, Phys. Rev. Lett. **47**, 359 (1981); Phys. Rev. A **25**, 6860 (1982).
[8] H. Horner *et al.*, Z. Phys. B **76**, 381 (1989).
[9] R. D. Henkel and M. Opper, J. Phys. A **24**, 2201 (1991); R. D. Henkel, Ph.D. thesis, Justus-Liebig-Universität Giessen, 1991 (unpublished).
[10] E. Gardner, B. Derrida, and P. Mottishaw, J. Phys. (Paris) **48**, 741 (1987).
[11] A. Crisanti and H. Sompolinsky, Phys. Rev. A **36**, 4922 (1987).
[12] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Phys. Rev. Lett. **61**, 259 (1988).
[13] H. Rieger, M. Schreckenberg, and J. Zittartz, J. Phys. A **21**, L263 (1988).
[14] H. Rieger, J. Phys. A **22**, 3447 (1989).
[15] A. Crisanti and H. Sompolinsky, Phys. Rev. A **37**, 4865 (1988).
[16] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland Personal Library, Amsterdam, 1987).
[17] W. Kinzel, Phys. Rev. B **33**, 5086 (1986).
[18] R. D. Henkel and W. Kinzel, J. Phys. A **20**, L727 (1987).
[19] T. Pfenning, H. Rieger, and M. Schreckenberg, J. Phys. I (France) **1**, 323 (1991).
[20] G. A. Kohring and M. Schreckenberg (to be published).
[21] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Phys. Rev. Lett. **58**, 1268 (1987).
[22] First attempts for mean-field theories of learning algorithms can be found in J. A. Hertz, A. Krogh, and G. I. Thorbergson, J. Phys. A **22**, 2133 (1989), and W. Kinzel and M. Opper, in *Models of Neural Networks*, edited by E. Domany, J. L. Hemmen, and K. Schulten (Springer, New York, 1991).