Retrieval-time properties of the Little-Hopfield model and their physiological relevance

Sebastián Risau-Gusman* and Marco A. P. Idiart†
Instituto de Física, UFRGS, Caixa Postal 15051, 91501-970 Porto Alegre, RS, Brazil
(Received 31 December 2004; published 12 October 2005)

We perform an extensive numerical investigation on the retrieval dynamics of the synchronous Hopfield model, also known as Little-Hopfield model, up to sizes of 2^{18} neurons. Our results correct and extend much of the early simulations on the model. We find that the average convergence time has a power law behavior for a wide range of system sizes, whose exponent depends both on the network loading and the initial overlap with the memory to be retrieved. Surprisingly, we also find that the variance of the convergence time grows as fast as its average, making it a non-self-averaging quantity. Based on the simulation data we differentiate between two definitions for memory retrieval time, one that is mathematically strict, τ_c, the number of updates needed to reach the attractor whose properties we just described, and a second definition correspondent to the time τ_η when the network stabilizes within a tolerance threshold η such that the difference of two consecutive overlaps with a stored memory is smaller that η. We show that the scaling relationships between τ_c and τ_η and the typical network parameters as the memory load α or the size of the network N vary greatly, being τ_η relatively insensitive to system sizes and loading. We propose τ_η as the physiological realistic measure for the typical attractor network response.

DOI: 10.1103/PhysRevE.72.041913 PACS number(s): 87.19.La, 05.90.+m, 84.35.+i, 05.50.+q

I. INTRODUCTION

Attractor neural networks are paradigmatic in computational neuroscience research. They present among other interesting properties, the possibility that, by local adjustment of the synaptic weights [1], specific patterns of activity be stored as dynamic attractors. The basin of attraction of such attractors is usually finite, endowing the network with the property of content addressability [2]. These networks are robust with respect to synaptic as well as learning noise and degrade gracefully with connection removal (for an extensive discussion on the attractor model and its physiological relevance see Ref. [3]).

But to be definitely plausible an attractor network must generate an output in physiologically realistic times. Neurons are slow processors, an action potential lasting approximately 1 ms and the typical speed of action potential propagation along thin axons being in the range 0.5–5 mm/ms [4,5]. If an attractor neural network is to be a candidate model for short-term-memory retrieval, for instance, when scaled to physiological sizes (10^9-10^{10} neurons) it should give a meaningful output in a number of time steps that when translated to physical time should be as low as 40 ms and not larger than, say, 200 ms [6] (see Ref. [7] for a quick review on short-term memory). Therefore it is of utmost importance to evaluate how the typical response times of the attractor networks scale with parameters like the system size, memory storage loading and the degradation level of the input. The response time should be related to the convergence time, the time elapsed between the presentation of the initial state (input) and the convergence to a stable attractor (output).

Interestingly, the convergence time of artificial neural networks is one of the few quantities that have resisted theoretical analysis. In fact, even numerical studies have been few and restricted to the synchronous Hopfield model, or Little-Hopfield model [1,8]. Kanter [9] performed simulations for fully connected systems of $\approx 10^3$ neurons, and found that the average convergence time to a stable state was logarithmic on N, the size of the net. It also depends on the initial overlap with the pattern to be retrieved. Surprisingly, Kanter also found that the variance of the distribution of convergence times grows with N, with the same scaling as the mean.

Some time later Kohring [10], using a multispin coding technique developed by Penna and de Oliveira [11], was able to perform a similar study on much larger systems, $N \approx 10^5$. His results disagree with those of Kanter, because he finds that for initial overlaps near to the critical value, the convergence time deviates from a simple logarithmic growth already for $N \approx 10^4$. In any case his main conclusion was that much bigger systems are needed to draw sound conclusions. Interestingly, for the system sizes simulated by Kohring, the variance of the distribution is not monotonous, and tends to decrease for large values of N. He speculates that the mechanism responsible for this decrease is the disappearance of metastable states near the memory for increasing N.

The loading of the network is measured by $\alpha = M/N$, the relative number of stored memories. Up to a certain critical value, $\alpha_c(m_0)$, there is always (in the infinite size limit) a value of the initial overlap m_0 above which the network always converges to the desired memory. The behavior of the convergence time as the critical value of α is approached was studied by Ghosh *et al.* [12]. Using the fact that the relaxation time for spin glasses is thought to follow a power law or a Vogel-Fulcher law ($\tau \approx \exp[A/(T-T_c)]$, where T is the temperature), they tried to determine whether the convergence time of neural networks belongs to one of these re-

*Current address: Departamento de Física, CNEA – Centro Atómico Bariloche, San Carlos de Bariloche, Rio Negro, Argentina. Electronic address: srisau@cab.cnea.gov.ar
†Electronic address: idiart@if.ufrgs.br

gimes, using system sizes as large as $N=16\,000$. Instead, they found that it follows a stretched exponential law, $\tau \approx \exp[-A(N)(\alpha_c-\alpha)^{\beta}]$, with $\beta \approx 0.6$. But they warn that much bigger systems are needed to confirm these results.

More recently, Frolov and Husek [13] have performed simulations from which they draw the conclusion that the convergence time grows as N^{γ}, where $\gamma<1$ is a number that grows with α and *does not depend* on the initial overlap with the pattern to be retrieved.

The only exact results available for convergence times are those of Komlos and Paturi [14]. Unfortunately, they concern only the case of low loading of the network, for a number of patterns $M<N/\ln N$ they find that the convergence time is of order $\ln \ln N$. For large loading, $M=\alpha N$ it is only possible to show that, for small α, all initial states at a distance ρN from the memory to retrieve will end up, in constant time, at a distance $\exp(-1/4\alpha)$ (for the exact definition of the increasing function $\rho(\alpha)$ see Ref. [14]). Some bounds have also been obtained, but they turn out to be so loose that they have no practical value.

The aim of the present work is to revisit the convergence time problem for the Little-Hopfield model. With the newly available computer power we are able to perform extensive numerical simulations up to 2^{18} neurons.

II. METHODS

The Little-Hopfield model consists of a network composed of N binary neurons interconnected, where the weight of each connection is given by Hebb's rule,

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

The set of patterns $\xi^{\mu}\in\{-1,1\}^N$ with $\mu=1,\ldots,M$ is called the set of stored patterns and it is randomly chosen without bias. We have studied the parallel dynamics of this model, at $T=0$, where the state of all the neurons is updated at each time step by

$$\sigma_i^t=\mathrm{sgn}\left(\sum_j J_{ij}\sigma_j^{t-1}\right). \tag{2}$$

Another possibility is to study the serial dynamics of the model, where at each time step only the state of a randomly chosen neuron is updated. We have preferred to study the parallel dynamics because in this case, for a given set of stored patterns, the final state of the networks is solely determined by the initial state. One of its drawbacks is the fact that the final state of the dynamics is not always a fixed point (as happens with asynchronous dynamics). There can also appear 2-cycles, where the state of the network oscillates between two different configurations. Nevertheless, we show below that the difference between the two configurations in a cycle is very small for large networks.

We have performed simulations for systems of up to 2^{18} neurons in the regime of parallel dynamics. We have used the multineuron coding of Penna and de Oliveira [11], where only one bit is used to encode the state of each neuron. As we are interested in the properties of the network in the regime of recall of one of the memories, the configurations that we have selected as initial states have a given overlap with the memory to be retrieved. In the following the overlap of the state of the network with the memory will be called magnetization. For large values of α, after the first few time steps the number of neurons that change their state before the fixed point is reached is only a very small fraction of the total. The improvement proposed by Stiefvater *et al.* [15] takes this into account and allows us to reduce significantly the time of computation. We have recorded the time taken by the network to reach the fixed point. In the case of 2-cycles, we have defined as convergence time the number of time steps needed to reach the first configuration of the cycle. In both cases only attractors with final overlaps larger than 0.9 are considered. We have taken the mean of the convergence time found for several realizations of the stored patterns. We remark here that for every realization we have used only one starting point, so as to maximize the independence of our runs.

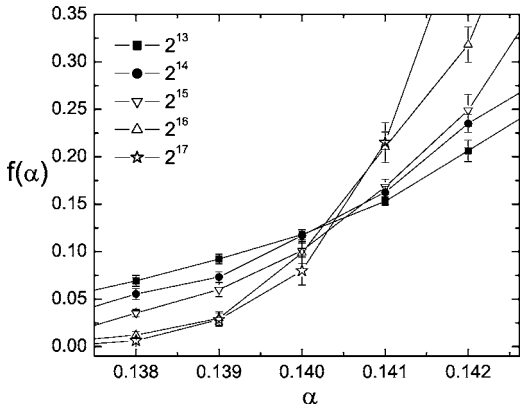

FIG. 1. Fraction of initial conditions that did not converge in our simulations. The initial overlap is $m_0=0.8$. The lines are only guides to the eye.

Naturally, for finite network sizes not all the runs converge to the memory attractor. Besides the minimal overlap criterion we have also set a limit time after which nonconvergence is assumed. This limit time is of the order of 5–10 times the average convergence time. The correctness of this criterion is somewhat confirmed by the fact that the largest convergence times recorded were usually much smaller than the limit time. In Fig. 1 the fraction of nonconverging initial conditions is displayed as a function of α. The critical value of α, $\alpha_c(m_0)$, is the value below which all initial conditions (with a definite overlap) converge, and above which none converge [2], for an infinite size system. For the initial overlap used in this paper $m_0=0.8$, the critical α could be determined by finite size scaling. According to Fig. 1, our estimation of it gives $0.14<\alpha_c(0.8)<0.141$ which agrees with the results obtained in Ref. [15].

III. RESULTS

A. The attractors

Attractors in the Hopfield-Little model are either fixed points or cycles of period 2. We find that the fraction of

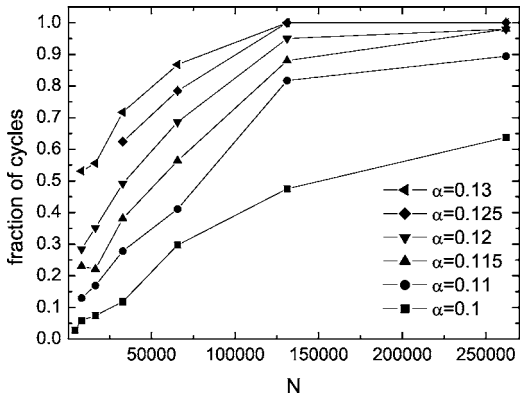

FIG. 2. Average fraction of final states that are cycles. The initial overlap is $m_0=0.8$. The lines are only guides to the eye.

retrieval states that are cycles increases with N (see Fig. 2). But, interestingly, we also find that the difference between the two configurations in each cycle tends to zero. We have measured this by taking the difference between the magnetization of both states (see Fig. 3). Notice that from this difference only a lower bound to the number of differing bits can be derived; for example, two states having the same magnetization need not be equal.

In the statistical mechanics calculations [16] it is obtained that, for values of α smaller than the critical, the final states are only fixed points of the dynamics. But the statistical mechanics results are only valid in the infinite size limit. Thus, for this to be the correct result, for finite values of N one of two things must happen: either the fraction of cycles vanishes with increasing N, or the fraction of cycles does not vanish but it is the overlap between its configurations that does. Our simulations show that the second possibility is the correct one. Thus, even though the number of cycles increases with N, this should not be a hindrance, as the two configurations in the cycle turn out to be almost identical.

B. The convergence time

For the average convergence time, our results are displayed in Fig. 4, as a function of the loading of the network.

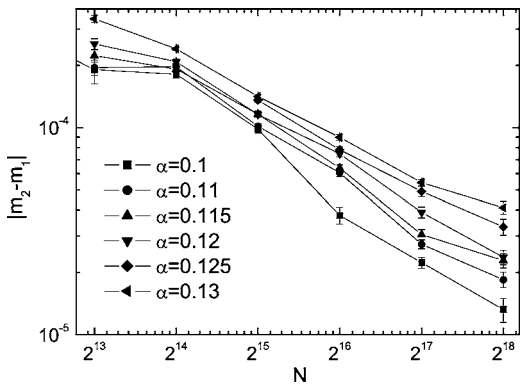

FIG. 3. Average difference of magnetization between the two states of a cycle. The initial overlap is $m_0=0.8$. The lines are only guides to the eye.

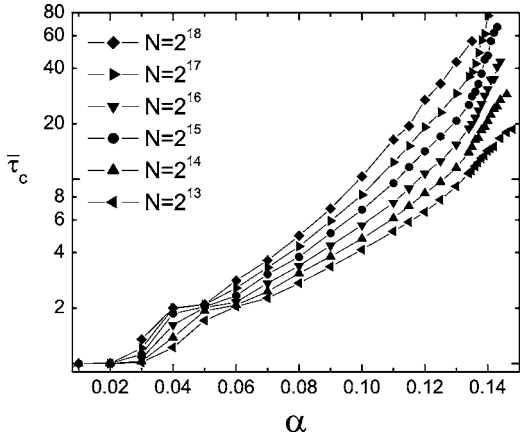

FIG. 4. Convergence times as a function of memory storage. The initial overlap is $m_0=0.8$. The lines are only guides to the eye. Error bars are smaller than the symbols.

For loadings smaller than 0.06 finite size effects are present, therefore in the rest of the paper we restrict ourselves to consider only larger loadings. Our results (see Fig. 5) differ from those found in Refs. [9,10] but are similar to those found in Ref. [13], the average convergence time seems to follow a power law for large N. The exponent of this power law grows with α [see Fig. 5(b)]. At variance with what was reported in Ref. [13], we find that the exponent depends also

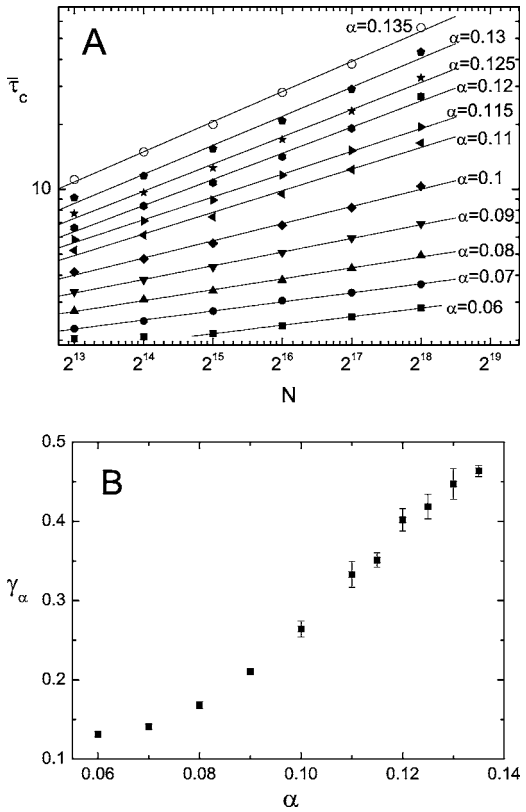

FIG. 5. (A) Convergence time for several values of $\alpha=M/N$ (symbols). The initial overlap is $m_0=0.8$. Error bars are smaller than the symbols. The lines shown are best fits. (B) The corresponding exponents of the best fits.

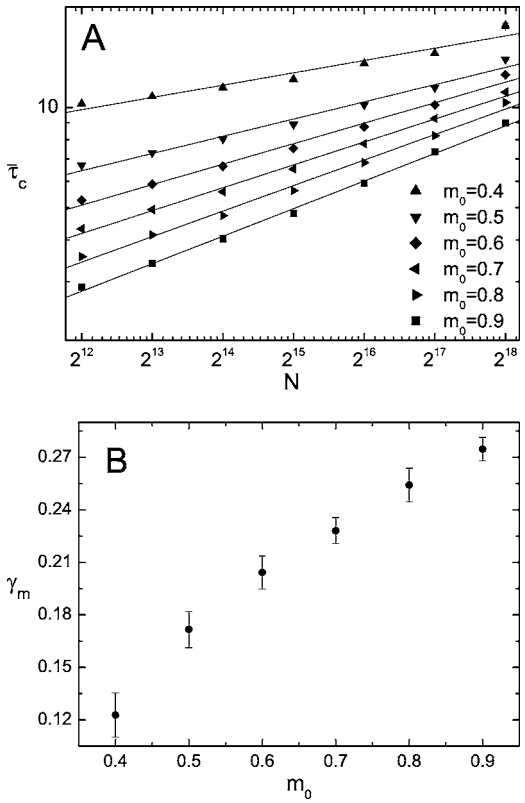

FIG. 6. (A) Convergence time for several values of m_0, the overlap of the initial state (symbols). The loading is $\alpha=0.1$. Error bars are smaller than the symbols. The lines shown are best fits. (B) The corresponding exponents of the best fits.

on the initial state of the network, at least for the values of N shown (see Fig. 6). It is obvious however that these exponents must change as N grows; otherwise it would imply that there exists a value of N where the curves intersect and eventually change their "ordering." This is not possible because for the same value of N the convergence times must necessarily be decreasing functions of m_0. Nevertheless, our results seem to suggest that for very high values of N the convergence times for different initial states should be very close.

C. The convergence time variance

For the variance of the convergence time, the behavior we find is a very interesting one. As mentioned in the introduction, from his simulations, Kanter concluded that, as a function of N, the variance grows at approximately the same rate as the average value. Kohring, studying larger systems, showed that this growth seemed only to be a transient after which the variance decreased. But he shows only the case of initial overlap $m_0=0.4$. We find in our simulations that the situation described by Kohring also seems to be a transient (see Fig. 7). We find that even though the variance increases for small values of N and then decreases for intermediate values, in the asymptotic limit it increases again, this time with no bound (as far as we can see). This asymptotic behavior is common to all the values of the initial overlap that

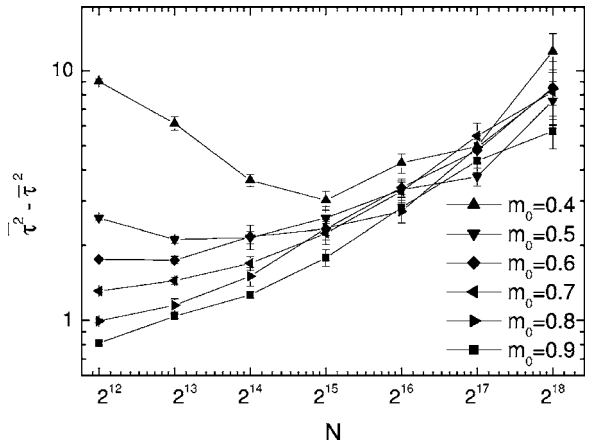

FIG. 7. Variance of the convergence time for several values of m_0, the overlap of the initial state (symbols). The lines shown are only guides to the eye.

we have analyzed. On the other hand, the initial "bump" of the variance seems to be highly dependent on the initial overlap, being biggest for the smallest overlaps and seems to disappear for $m_0>0.6$.

Some of the features of the variance can be understood by looking at the trajectories followed by the network in the dynamics (see Fig. 8). It is easy to see that there are two different effects responsible for the fact that there are many different routes leading to the same fixed point. The fact that, for small values of N, the beginning of the trajectories can be very different for different initial states seems to show that the basin of attraction is not very "deep," or at least seems to be very irregular, far from the bottom [notice how different the trajectories are in Fig. 8(a)]. On the other hand, for large values of N the basin seems to be much more regular, or more isotropic, the trajectories almost coincide in their beginning [see Fig. 8(b)]. But now it is the bottom of the basin that seems to be very irregular, in such a way that for some initial states the network wanders around the bottom of the basin for some time steps before finding the true minimum.

The convergence displayed in Figs. 8(c) and 8(d) make us question the practical value of finding the convergence time to the true mathematical attractor of the network. After a close inspection of Fig. 8(c), it is evident that the network has already converged around 10 time steps with very low variance among different runs, however the attractor stabilizes at $t=15$, 23, and 30 for three different runs. A similar thing happens in Fig. 8(d).

We thus decided to redo the graph of Fig. 5(a) using as a convergence criterion a finite precision on the overlap, in other words, we considered that a pattern has settled at time τ_η if $|m(\tau_\eta)-m(\tau_\eta-1)|<\eta$. With this new definition we obtain the surprising result displayed in Fig. 9. It shows that this convergence time is actually independent of the size of the system for low memory loading while it surprisingly decreases as we approach the critical loading value.

Figures 10(a) and 10(b) show the typical behavior of the convergence times for moderate ($\alpha=0.110$) and high ($\alpha=0.135$) memory loading as we change the precision η. We observe that the time retrieval properties seem to be ro-

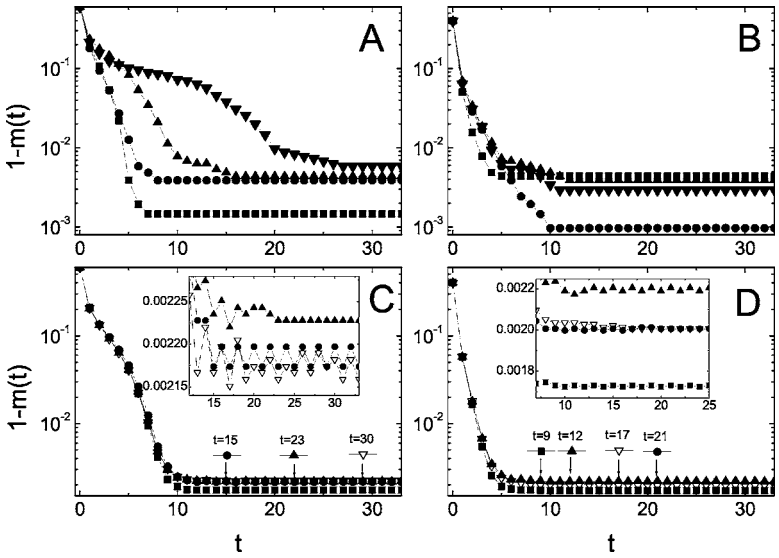

FIG. 8. Overlap as a function of time, for several different initial states (symbols). The lines shown are only guides to the eye. (A) $N=4096$, $m_0=0.4$; (B) $N=4096$, $m_0=0.6$; (C) $N=262\,144$, $m_0=0.4$; (D) $N=262\,144$, $m_0=0.6$. The number of stored memories is $M=0.1N$. The arrows with the symbols show where the corresponding line reaches its final state.

bust to the choice of precision. For low precision the convergence times are relatively small and insensitive to system size, that seems to hold true for moderate storage levels even as we decreased η. When α is close the critical storage capacity the network presents a pronounced decrease of retrieval time. Figure 11 displays the same data as a function of memory load α for two different precision values.

IV. DISCUSSION AND CONCLUSIONS

The results presented in this paper show that the details of the retrieval dynamics of the Little-Hopfield model are still a matter of scientific debate. We obtain that for a fairly large region of systems parameters, $m_0>0.4$ and $0.08<\alpha<\alpha_c$, the average retrieval time is a power law of the form $\bar{\tau}_c \propto N^{\gamma(\alpha,m_0)}$ where γ range from 0.1 to 1, and seems to be slightly increasing with N. That contradicts the early results [9,10] and agree somewhat with Ref. [13]. We find no conclusive evidence in our data for the exponential behavior proposed in Ref. [12] near saturation. We also observe that the behavior of the convergence time variance is much more unusual than what was reported in Ref. [10]. After a transient nonmonotonous behavior, the variance seems to increase regularly and apparently without bound with the system size. Its growth with N would not be a problem if it was smaller than the corresponding growth of the mean value. However, our simulations show clearly that this is not the case (see Fig.

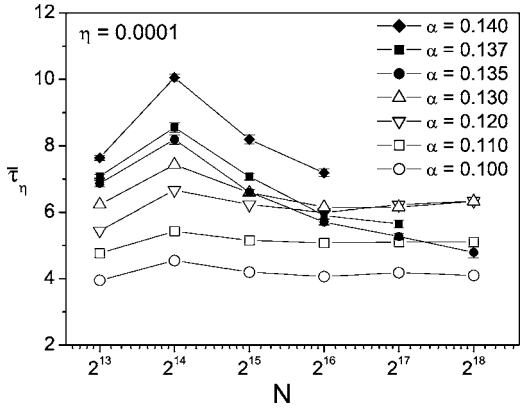

FIG. 9. Convergence times for a precision of $\eta=0.0001$ as a function of the size of the network for several values of $\alpha=M/N$, for initial state $m_0=0.8$. The lines shown are only guides to the eye and the error bars are smaller than the symbols.

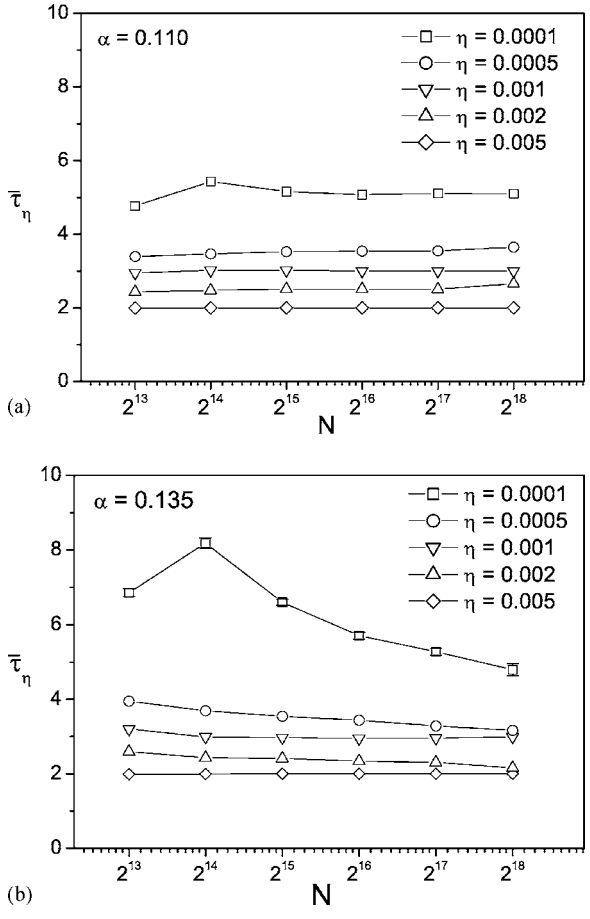

FIG. 10. Convergence times as a function of the system size for several values of the precision η for $\alpha=0.110$ and 0.135. The initial state is again $m_0=0.8$.

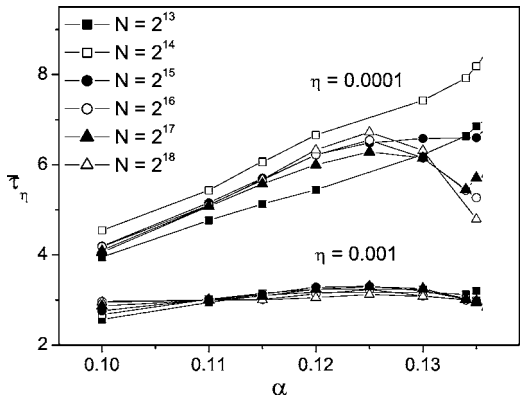

FIG. 11. Convergence times as a function of the memory load α, for η=0.0001 and 0.001 for several values of N, for initial state m_0=0.8.

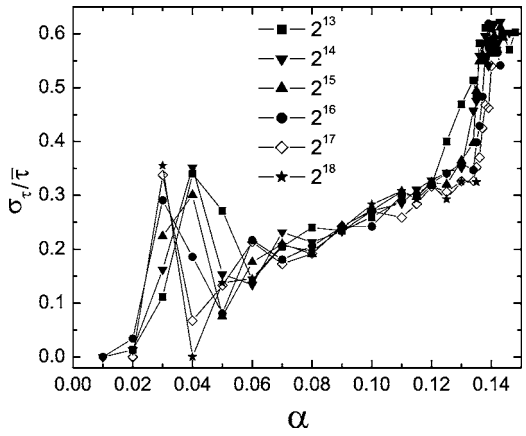

FIG. 12. Quotient between the statistical deviation and the mean value of the convergence time, for several values of N. The initial state has m_0=0.8. The lines shown are only guides to the eye.

12), the quotient between both quantities is of order 1 throughout the range of values of N that we have explored. If this was true for all values of N, it would mean that the average time (divided by some suitable function of N) is not a self-averaging quantity, i.e., its value would differ significantly for different initial states, even in the infinite N limit. It is interesting to notice that many other quantities are self-averaging. For example, the final overlap and, most interestingly, also the overlap at each time step. We have checked that the distributions for the convergence times τ_c are well described by Gaussians. We have also analyzed our results using the median of the convergence time distribution, as well as averages with subsamples with the 50% fast times. In both cases the results agree with the average results using the full samples. This rules out the possibility that our findings are artifacts coming from large time values in the tail of the distribution.

The correct theoretical interpretation of the properties of the convergence time τ_c can provide important insights to the general understanding of frustrated systems approaching equilibrium. As for physiological adequacy of the model the results for τ_c would be discouraging. For moderate storage levels $\alpha \simeq 0.10-0.12$, and relatively close initial states m_0=0.8, we obtain $\tau_c \simeq N^{1/3}$ what would indicate a multiplicative factor of order 20 when we go from system sizes of 2^{18} to more realistic sizes of 2^{31}. Even if we take the optimistic view that a network update can be accomplished in 2 ms (1 ms action potential and 1 ms axonal delay and integration time) it implies a total time of 400 ms for convergence, that can be considered physiologically too slow.

However, as was pointed out earlier, the model's retrieval dynamics seems to have two well-defined regimes. That gets more evident with the increase of size of the network (see Fig. 8). The first regime corresponds to a fast overlap increase up to more than 99.99% of its final value. The times involved in this first phase have very small variance among realizations. The second regime seems to correspond to a search of the true attractor in a very irregular energy landscape. The second phase results in a correction smaller than 0.01% in the state of the network, but takes usually much longer than the first phase and it is responsible for most of the scaling properties described in this paper, such as the power law dependence with size and the odd variance behavior.

If we take τ_η for a reasonable value of η, as the true relevant response time of the network the results become much more promising. We observe that up to the sizes simulated in this work the times τ_η are insensitive to the system size and fairly insensitive to the memory storage load. In fact, it decreases with α as it approaches the critical value α_c. But this result is to be taken with caution since the basins of attraction are too narrow [2] and Fig. 9 displays only the trials that did converge.

We therefore conclude that the Little-Hopfield network have a response time in terms of number of steps that could be compatible with the characteristic times in biology, both physiological and behavioral, if we consider as the relevant time parameter the fixed precision convergence time τ_η. Since τ_η is usually small this result highlights the importance of the theoretical calculations for the first few time steps of the retrieval dynamics started by Gardner *et al.* [17], and developed into the statistical neurodynamics [18–20] and recently redone with a generating functional approach [21].

ACKNOWLEDGMENTS

This work was supported in part by the Brazilian agencies CNPq and FAPERGS. One of the authors (S.R.G.) acknowledges support from the Centro Latinoamericano de Física.

[1] J. J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).
[2] B. M. Forrest, J. Phys. A **21**, 245 (1988).
[3] E. Rolls and A. Treves, *Neural Networks and Brain Function* (Oxford University Press, Oxford, 1998).
[4] R. D. Traub, A. Bibbig, F. E. N. LeBeau, E. H. Buhl, and M. A. Whittington, Annu. Rev. Neurosci. **27**, 247 (2004).
[5] *Fundamental Neuroscience*, edited by M. J. Zigmond, F. E. Bloom, S. C. Landis, J. L. Roberts, and L. R. Squire (Aca-

demic, New York, 1999).
[6] To calculate exactly this equivalence one must take into account the processing response of single neurons, integration times, as well as axon delays.
[7] J. E. Lisman and M. A. P. Idiart, Science **267**, 1512 (1995).
[8] W. A. Little, Math. Biosci. **19**, 101 (1974).
[9] I. Kanter, Phys. Rev. A **40**, 2611 (1989).
[10] G. A. Kohring, J. Phys. A **23**, 2237 (1990).
[11] T. J. P. Penna and P. M. C. de Oliveira, J. Phys. A **22**, L719 (1989).
[12] M. Ghosh, A. K. Sen, B. K. Chakrabarti, and G. A. Kohring, J. Stat. Phys. **61**, 501 (1990).
[13] A. A. Frolov and D. Husek, International Joint Conference on Neural Networks 5 (IJCNN'00), 2000, p. 622.
[14] J. Komlos and R. Paturi, Neural Networks **1**, 239 (1988).
[15] T. Stiefvater, K.-R. Müller, and R. Kühn, Physica A **232**, 61 (1996).
[16] J. F. Fontanari and R. Koeberle, J. Phys. (France) **49**, 13 (1988).
[17] E. Gardner, B. Derrida, and P. Mottishaw, J. Phys. (Paris) **48**, 741 (1987).
[18] S. Amari and K. Maginu, Neural Networks **1**, 63 (1988).
[19] M. Okada, Neural Networks **8**, 833 (1995).
[20] H. Koyama, N. Fujie, and H. Seyama, Neural Networks **12**, 247 (1999).
[21] D. Bolle, J. B. Blanco, and T. Verbeiren, J. Phys. A **37**, 1951 (2004).