An Application of the Saturated Attractor Analysis to Three Typical Models

J. FENG B. TIROZZI

Mathematical Department
University of Rome "La Sapienza", P.le A. Moro, 00185 Rome

Abstract

The saturated attractor analysis, an approach proposed first in [FP] for a comprehensive study of the dynamics of the Linsker model and then successfully applied to the dynamic link model[FT1], is further developed here. By a unified approach to the Hopfield model, the Linsker model and the dynamic link model, three typical models in the field of the neural networks, we show a way to choose the parameters of these dynamics in order to obtain any chosen saturated attractor which is general enough in most applications. We generalize our previous results for the Linsker model and the dynamic link model with the clipping function to the case of the sigmoid like function. Our results allow us for the first time to understand the underlying mechanism among these models and thus to furnish a useful guidance in the further possible applications.

§1 INTRODUCTION

The past decade has seen an explosive growth in the studies of neural networks, the theory underlying learning and computing in networks has developed into a mature subfield existing somewhere between mathematics, physics, computer science and neurobiology. In part this was the result of many deep and interesting theoretical exposition in physics and mathematics, for example, the application of the spin glass theory to the Hopfield model allows us to understand clearly the phase transition from the retrieval to non retrieval state. Another major impulse was provided by the successful explanation of some biological phenomena, at least in a primitive level, for example, the Linsker model mimics the ontogenesis development of the primary visual system[Lin]. Of course, the most important impulse comes from the learning techniques successfully applied to some practical problems which were traditionally thought of as some of the hardest problems in the AI. One of the recent examples of such an application is the face recognition using the dynamic link model, a model proposed by von der Malsburg first in 1981[KMM].

However, at this moment, the theoretical treatment of these models is obviously far away from being satisfactory, mainly due to the lack of theoretical tools to deal with the nonlinearity exploited in most of the models reported today. In the present paper, in terms of our previous work on the Linsker model and the dynamic link model we develop a unified theoretical framework for tackling the Hopfield model, the Linsker model and the dynamic link model.

Our approach allows us to reformulate many problems studied in the Hopfield model. A concrete criterion to check whether a stored pattern is an attractor of the network is given. The capacity, a quantity which plays a central role in the spin glass approach to the Hopfield model, is naturally introduced here. One advantage of the present approach is that we do not impose the restriction of the symmetry of the connection matrix. Our results also reveal the role of different parameters in the Hopfield model. We consider the Linsker model with the sigmoid like function in the updating dynamics of its synaptic connections(a definition of the sigmoid like function is in section 2). All conclusions in [FP][FPR] are reobtained, where the clipping function, a special case of the sigmoid like function and so a special case of the present paper, is used for the development of the synaptic connections. The present paper tells that the appearance of the structured receptive fields is independent of the choice of the clipping function, which is thought of as one of the drawback of the Linsker model. Furthermore, we also take into account on the reason for the appearance of the oriented receptive field in the further layers of the Linsker network. For the dynamic link model, a principle to choose all five parameter employed in the model is furnished, which confirms our previous claim that all results contained in [FT1] for the clipping function are true for a more general class of function, i.e. for the sigmoid like function.

Although here we confine ourselves to the models on which we worked before [ATYD] [FP] [FPR] [FQ] [FT1], the essential part of our approach is to analyze the dynamics with the sigmoid like function, and it is possible to adopt our method here to analyze other models in the field of the neural networks such as the B.P. and the recurrent network.

The general idea behind the saturated attractor analysis is quite straightforward. Consider a dynamics defined on the space $[-1,1]^N$, where N is either the number of neurons (the Hopfield model and the dynamic link model) or the number of connections (the Linsker model). It is reasonable to confine ourselves to a subset of all the fixed points of the dynamics, i.e. to all saturated fixed points in $\{-1,1\}^N$ since in the Hopfield model all the stored patterns take values on $\{-1,1\}^N$, while in the Linsker model and the dynamic link model this confinement has been confirmed by the numerical simulations[Lin][KMM]. In particular, the fast dynamic link model is proposed in terms of this observation[KMM]. As we all know, it is relatively easy to determine the *whole* region of the dynamic parameters, say $\Gamma(w)$, in which a given pattern w of particular interest (in the Hopfield it is one of the stored patterns, in the Linsker model it is the structured receptive field, in the dynamic link model it is the on-center configuration) is a fixed point. If we are further able to prove the stability of the fixed point w, we assert that if and only if as the dynamic parameters are in the region $\Gamma(w)$, w is an attractor of the dynamics. Fortunately, due to the special form of the sigmoid like function and we restrict ourselves to all saturated fixed points, the idea above can be carried out as in [FP][FPR][FT1], but for a more general and more significant class of functions, the sigmoid like functions. We call such an approach *the saturated attractor analysis.*

Due to the limitation of the space, we are only able to briefly report our results. For a full exposition and detailed proofs, we refer the reader to our whole paper[FT2].

§2 General Model and Notation

For a given positive integer N, an $N \times N$ matrix $Q = (q_{ij}, i,j = 1,\cdots,N)$ and an N dimensional vector $r = (r_i, i = 1,\cdots,N)$, consider the following dynamics

$$w_i(\tau+1) = f(w_i(\tau) + k_1 + \sum_{j=1}^{N}[(q_{ij}+k_2)r_j w_j(\tau)]) \qquad (1)$$

where $\tau = 1,2,\cdots$ is the discrete time, $w(\tau) = (w_i(\tau), i = 1,\cdots,N) \in \mathbb{R}^N$, (k_1,k_2) are two parameters of the dynamics, and f is a continuous function defined on $\mathbb{R}^1$ satisfying

(f1). $f(x) = 1$, if $x \geq 1$, $f(x) = -1$, if $x \leq -1$,

(f2). $f(x)$ is a strictly increasing and continuous function for $x \in [-1,1]$, $f(x) \geq x$ if $x \in (0,1]$ and $f(x) \leq x$ if $x \in [-1,0)$.

We call a function with the properties (f1) and (f2) a *sigmoid like function.*

Note that for the sigmoid function σ_β with range between -1 and 1, $\sigma_\beta(x) = \frac{2}{1+\exp(-\beta x)} - 1$, both conditions (f1) and (f2) are approximately satisfied when β is large. It is reasonable to expect that in the numerical simulation, both (f1) and (f2) are true for the sigmoid function σ_β with large β. Due to this reason, we believe that our results on the dynamics (1) with the sigmoid like function below reflect the exact properties of dynamics (1) with $f = \sigma_\beta$ (β large) which are mostly observed by numerical simulation. The function termed as *the clipping function* and used in the dynamics of the development of the synaptic connection of the Linsker network is defined by $f_c(x) = x$ if $|x| < 1$, and $f_c(x) = 1$ if $x > 1$, $f_c(x) = -1$ if $x < -1$, which of course fulfills both (f1) and (f2) [Lin][FP][FPR]. In the dynamic link model, fast dynamic link model, or the discrete version of it, the function f adopted for the dynamics is either the clipping function or the sigmoid function [KMM][FT1].

It is easily seen that the conditions on the range of the function (f1) is not essential and can be relaxed.

Let us now introduce three functions which will play a crucial role in our later development. Let $w \in \{-1,1\}^N$ be a given configuration then $J^+(w) = \{i, w_i = 1\}$, $J^-(w) = \{i, w_i = -1\}$ are (respectively) the set of all sites with $w_i = 1$ and all sites with $w_i = -1$.

First we introduce the *slope function* $c(w)$ on $\{-1,1\}^N$ defined by

$$c(w) = \sum_{i\in J^+(w)} r_j - \sum_{i \in J^-(w)} r_j. \qquad (2)$$

Then we consider the *intercept functions* $d_1(w)$ and $d_2(w)$:

$$d_1(w) = \max_{i\in J^+(w)}[\sum_{j\in J^-(w)} q_{ij}r_j - \sum_{j\in J^+(w)} q_{ij}r_j] \qquad (3)$$

and

$$d_2(w) = \min_{i \in J^-(w)} [\sum_{j \in J^-(w)} q_{ij} r_j - \sum_{j \in J^+(w)} q_{ij} r_j]. \tag{4}$$

The reason why we call them slope function and intercept functions will be clear after the Theorem 1 below.

§3 The Set of All Saturated Attractors

The set of all fixed points of the dynamics (1) is $FP = \{w; w_i = f(w_i + \sum_{j=1}^N (q_{ij} + k_2) r_j w_j + k_1),\ i = 1, \cdots, N\}$. From the compactness of the range of the function f and the continuity of f, we get that the set FP is nonempty by the Brouwer's fixed point theorem. A fixed point is called an attractor if it is a stable fixed point. We will confine ourselves to a subset of all attractors in $\{-1,1\}^N$ which is general enough in most of applications.

Definition 1. *A configuration in the set $\Omega := \{w \in \{-1,1\}^N$; there exists a nonempty neighborhood $B(w)$ of w in $\{-1,1\}^N$ such that $\lim_{\tau \to \infty} w(\tau) = w$ if $w(0) \in B(w)$ and $k_1 + \sum_{j=1}^N (q_{ij} + k_2) w_j \neq 0, \forall i = 1, \cdots, N\}$ is called a saturated attractor of the dynamics (1).*

The following theorem establishes that, for the case of the dynamics (1), the condition (5) below is strong enough to ensure that w is an attractor of the dynamics.

Theorem 1. *If w is a saturated attractor of the dynamics (1), $\lim_{\tau \to \infty} w(\tau) = w$, then there exists a $T > 0$ such that $w = w(T + \tau), \forall \tau \geq 0$. Furthermore w is a saturated attractor of the dynamics (1) if and only if*

$$d_1(w) < k_1 + c(w) k_2 < d_2(w). \tag{5}$$

For a given configuration w, Theorem 1 tells that w is a saturated attractor of the dynamics (1) if and only if (k_1, k_2) lies in between the two parallel lines(see Fig. 1) $k_1 + k_2 c(w) = d_1(w)$ and $k_1 + k_2 c(w) = d_2(w)$. Hence $c(w)$ is the slope function of the lines above, and d_1, d_2 are the two intercept functions. If $d_2(w) > d_1(w)$, the parameter region $\Gamma(w) := \{(k_1, k_2)$ such that w is a saturated attractor of the dynamics (1)$\}$ is a nonempty set. If $d_2(w) < d_1(w)$ $\Gamma(w)$ is an empty set. So in this sense the larger is the difference between $d_2(w)$ and $d_1(w)$, the more stable is the attractor w. From Theorem 1 we can derive some interesting consequences which are shown in the following corollaries:

Corollary 1. *(Fig. 1)*
1) The parameter region of (k_1, k_2) in which $(1, \cdots, 1)$ is a saturated attractor of the dynamics (1) is

$$k_1 + \sum_j r_j k_2 > d(+) := -\min_{i=1,\cdots,N} \sum_{j=1}^N q_{ij} r_j \tag{6}$$

2) The parameter region of (k_1, k_2) in which $(-1, \cdots, -1)$ is a saturated attractor of the dynamics (1) is

$$k_1 - \sum_j r_j k_2 < d(-) := \min_{i=1,\cdots,N} \sum_{j=1}^N q_{ij} r_j \tag{7}$$

Corollary 2. *(Fig. 1)*
1) If q_{ij} depends only on j, then only the configuration $(1, \cdots, 1)$ and $(-1, \cdots, -1)$ are saturated attractors of the dynamics (1).
2) If $q_{ij} = \delta_{ij}$, and $\min\{r_j, j = 1, \cdots, N\} > 0$, then any configuration $w \in \{-1,1\}^N$ is a saturated attractor of the dynamics (1).

§4 Application to the Hopfield Model, the Linsker model and the dynamic link model

The Hopfield model, to which most of the theoretical investigations in the field of the neural network have been devoted so far, is defined by $q_{ij} = T_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu,\ i, j = 1, \cdots, N$ and by the equalities $k_1 = \theta$, the threshold, $k_2 = h$, the external field and $r_i = 1, i = 1, \cdots, N$. $w_i(\tau)$ is the neural activity at time τ of the i-th neuron, and $\xi^\mu = (\xi_i^\mu, i = 1, \cdots, N)$ is the μ-th pattern to be stored in the network. The dynamics (1) now reads

$$w_i(\tau+1) = f(w_i(\tau) + \sum_{j=1}^{N}(T_{ij} + h)w_j(\tau) + \theta), i = 1, \cdots, N. \quad (8)$$

In most of the theoretical investigations, in particular in the statistical physics approach, ξ_i^μ is assumed to be i.i.d. and $p(\xi_i^\mu = 1) = p(\xi_i^\mu = -1) = \frac{1}{2}, \forall i, \mu$.

The dynamics (8) is a discrete time version of the continuous Hopfield model. Next we apply our results of section 3 to the Hopfield model. Since the stored patterns take values +1 and −1, it is enough general for us to restrict ourselves to the space $\{-1, 1\}^N$. $d_2(w)$ and $d_1(w)$ may be expressed using the overlap parameters $m(w, \xi^\mu) := \frac{1}{N}\sum_{i=1}^{N} w_i \xi_i^\mu$

$$d_1(w) = -\min_{i\in J^+(w)} \sum_{\mu=1}^{p} \xi_i^\mu m(w, \xi^\mu) \quad (9)$$

and similarly,

$$d_2(w) = -\max_{i\in J^-(w)} \sum_{\mu=1}^{p} \xi_i^\mu m(w, \xi^\mu). \quad (10)$$

Combining (9), (10) and Theorem 1, we see that the criterion for the existence of a saturated attractor of the Hopfield model is that

Theorem 2. *For the dynamics (8), a configuration $w \in \{-1, 1\}^N$ is a saturated attractor of the Hopfield model if and only if*

$$-\min_{i\in J^+(w)} \sum_{\mu=1}^{p} \xi_i^\mu m(w, \xi^\mu) < \theta + c(w)h < -\max_{i\in J^-(w)} \sum_{\mu=1}^{p} \xi_i^\mu m(w, \xi^\mu). \quad (11)$$

In the practical applications, we are mainly interested to establish if $w = \xi^\mu, \mu = 1, \cdots, p$ is a saturated attractor of the dynamics (8), a fact which can be easily checked by using Theorem 2. Note that this criterion is not based on the independence of the patterns ξ^μ. Many interesting examples can be constructed using this approach[FT2].

In spite of the extensive investigation of the Hopfield model, a little attention was paid to the parameter (θ, h) until now. Our theorem allows us for the first time to have a clear understanding of the role played by the two parameters in the dynamics (8) as explained below. The Hopfield model is described by a picture of the type of Fig. 1 , which is redrawn in Fig. 2. It is easily seen from the Fig. 2 that the number of stored patterns, i.e, of saturated attractors, of the Hopfield model depends on the parameters (θ, h). There is one region in which many saturated attractors coexist(see Fig. 2). In this region, the network will have the highest capacity, a quantity studied extensively in the literature. Outside this region, the capacity will become lower and lower. When h, the external field, is negative, there will be only one saturated attractor corresponding to one of the stored patterns if $c(\xi^\mu) \neq c(\xi^\nu)$ for $\mu \neq \nu$. Thus the capacity for the network is only $1/N$ in this case. However this region is good for retrieving a specific memory w if it is a saturated attractor. This remark suggests a way to recall an information avoiding the spurious states[FQ]. The above example suggests us the following definition of the critical capacity of the Hopfield model which can be applied to more general models also with dependent patterns:

Definition 2. *The critical capacity α_c of the dynamics (8) is*

$$\alpha_c := \inf\{\alpha = p/N, \langle d_2(\xi^\mu)\rangle - \langle d_1(\xi^\mu)\rangle = 0 \; for \; any \; \mu = 1, \cdots, p\}, \quad (12)$$

where $\langle\cdot\rangle$ represents the expectation with respect to the distribution P of ξ^μ.

Further discussion on the relation of α_c defined above and the critical capacity founded in the spin glass approach is contained in [FT2] and the numerical simulation of α_c is shown in [FT2] also.

Now we consider the Linsker model.

The Linsker's model [Lin][FP][FPR] resembles the visual system, with an input feeding onto a number of layers corresponding to the layers of the visual cortex. The units of the network are linear and are organized into two-dimensional layers indexed L_0(input), $L_1, \cdots$ and so on. We suppose that each layer has N neurons and periodic boundary condition(wrapped up). There are feed-forward connections between adjacent layers, with each unit receiving inputs decreasing monotonically with the distance from the neurons belonging to the underlying layer. Let $w_{ki}^{(n)}(\tau)$ be the synaptic connection between the neuron i of the $(n-1)$-th layer and the neuron k of n-th layer, $r_{ki}^{(n)}$ is the synaptic density function between the $(n-1)$-th layer and the n-th layer, $(\sum_{i=1}^{N} r_{ki}^{(n)} = 1, \ \forall\ n, k)$. Making the averages with respect to the neuron activities we get

$$w_{ki}^{(n)}(\tau+1) = f(w_{ki}^{(n)}(\tau) + k_1 + \sum_{j=1}^{N}(q_{ij}^{(n)} + k_2)r_{kj}^{(n)}w_{kj}^{(n)}(\tau)). \tag{13}$$

(see the above bibliography for more details). The dynamics (13) is the updating process of the synaptic connections and characterizes the Linsker network. The index k can be dropped from the equation (13) since the appearance of a structured receptive field does not depend on it. We change our notation a little bit in order to apply theorem 1 of the section 3 to the dynamics (13). Let

$$d_1(w,n) = \max_{i\in J^+(w)}[\sum_{j\in J^-(w)} q_{ij}^{(n)}r_j^{(n)} - \sum_{j\in J^+(w)} q_{ij}^{(n)}r_j^{(n)}],$$

$$d_2(w,n) = \min_{j\in J^-(w)}[\sum_{i\in J^-(w)} q_{ij}^{(n)}r_j^{(n)} - \sum_{j\in J^+(w)} q_{ij}^{(n)}r_j^{(n)}].$$

Theorem 3. *If w is a saturated attractor of the Linsker model, then there exists $T > 0$ such that $w = w(T+\tau), \forall\tau \geq 0$. Furthermore w is a saturated attractor of the dynamics (13) if and only if*

$$d_1(w,n) < k_1 + c(w)k_2 < d_2(w,n), \ \ n = 1, \cdots. \tag{14}$$

In the Linsker model a structured (an on-center or an oriented) receptive field is of particular interest. We know from the simulations that these receptive fields appear between the L_1 and L_2 layers if all the synaptic connections between the L_0 layer and the L_1 layer are kept positive. Recently this kind of structured receptive field has been founded important in the application of similar networks to the image recognition (see next section). Theorem 3 gives results which agree with what has been discovered by Linsker in its numerical simulations for the third layer (L_2) [FPR]. These results are shown in Fig. 3. The application of this theorem becomes more important as we encounter the necessity, in the practical application, to control the size of the on-center receptive field configuration by selecting the parameter of the dynamics (13). All results in [FP] [FPR] are true for the dynamics (13), we will not repeat them here and refer the reader to them for more details.

An interesting problem for us to do in the future is to check which model is the most optimal one in the sense to have the biggest $d_2(w) - d_1(w)$ among the models proposed to describe the ontogenesis of the visual system[Mal] here w is a structured receptive field. Our approach here makes this comparison possible.

Next we are going to consider the dynamic link model.

The power of the dynamic link network, a model proposed by von der Malsburg first in 1981, is demonstrated and developed in recent years in different applications, see for example [KMM]. In [FT1], a discrete version of the dynamic link model is proposed and a principle for choosing parameters used in the dynamic link model is given for the clipping function defined as in section 2. Here, by our results of section 3, we are able to reobtain all results in [FT1]. The dynamic link network is essentially a two layers network, say layer X and layer Y with both inter-layer connections and intra-layer connections. There are N neurons in the two layers and they are arranged in a two dimensional torus (i.e with periodic boundary conditions). For the details of the model the reader can refer to the paper [KMM]. In this model there are two time scales, one (τ) varied rapidly corresponding to the neural dynamics and the other one (ν) associated to the characteristic time scale of the synaptic dynamics. Making analogous transformations as those in [FT1] the neuron dynamics can be written in a form which is similar for the two layers

$$\begin{cases} \eta_i(\tau+1,\nu) & = f(\sum_{j=1}^N k_{ij}\eta_j(\tau,\nu) + I_i(\tau,\nu)) \\ \eta_i(0,\nu) & = 0 \end{cases} \tag{15}$$

and which is the dynamics we will focus on. The input $I_i(\tau,\nu)$ is different for the two layers and contains all the information connected with the image recognition problem. We suppose here that $I_i(\tau,\nu)$ is independent of i and τ denoting it as $I(\nu)$. The case of the dependence of I on i and τ is considered in [FT1]. The matrix k_{ij} is defined by $k_{ij} = \gamma e^{-\|i-j\|^2/s} - \mu = \gamma p_{ij} - \mu$ where p_{ij} is the weight interaction function inside the layer X or Y, $\gamma > 0$ and $\mu > 0$ are the intensities of excitatory or inhibitory connections respectively. Let $c(w)$ be defined as in Section 3 and $e_1(w)$ and $e_2(w)$ be two functions of the configuration w defined similar to $d_1(w)$ and $d_2(w)$

$$e_1(w) = \max_{i\in J^+(w)}[\sum_{j\in J^-(w)} p_{ij} - \sum_{j\in J^+(w)} p_{ij}], \quad e_2(w) = \min_{i\in J^-(w)}[\sum_{j\in J^-(w)} p_{ij} - \sum_{j\in J^+(w)} p_{ij}]. \tag{16}$$

Then we can apply Theorem 1 of section 3 to the dynamic link model. The proofs are similar to that in [FT1].

Theorem 4.
(1). For $\forall w \in \{-1,1\}^N$, $w \neq (1\cdots,1),(-1,\cdots,-1)$, *a necessary and sufficient condition ensuring that there exists a nonempty set of* (μ,γ,s,I) *in which* w *is a saturated attractor of the dynamic (15) is* $e_2(w) > e_1(w)$ *and* $\gamma > \gamma_0 := 2/[e_2(w) - e_1(w)]$. *Furthermore the larger the* γ, *the bigger the parameter region ensuring that* w *is a saturated attractor of the dynamics (15).*
(2). In the circumstances of (1), there exists a positive number μ_0 *such that when* μ *is in the set* $\{\mu,\mu \geq \mu_0\} \cap \{\mu, \gamma e_1(w) + 1 < I(\nu) - c(w)\mu < \gamma e_2(w) - 1\}$ *then* w *is a saturated attractor of the dynamics (15) and* $(1,\cdots,1)$, $(-1,\cdots,-1)$ *will no longer be attractors of the dynamics (15).*
(3). If s *is large enough so that* p_{ij}, $i,j = 1,\cdots,N$ *are constants independent of* i,j, *then only* $(1,\cdots,1)$ *and* $(-1,\cdots,-1)$ *are the only possible saturated attractors of the dynamics (15).*
(4). If s *is small enough so that* $p_{ij} = \delta_{ij}$ *with* $\gamma > 1$, *then any state* $w \in \{-1,1\}^N$ *is an attractor of the dynamics (15).*
(5). w *is a saturated attractor of the dynamics (15) if and only if* $I(\nu) \in [\gamma e_1(w) + c(w)\mu + 1, \gamma e_2(w) + c(w)\mu - 1]$.

For an explanation of Theorem 4, we refer the reader to Fig. 4. Theorem 4 shows that the effect of the lateral inhibition is to have some non trivial pattern as an attractor of the dynamics (15) and to avoid the region in which the trivial configuration $(1,\ldots,1)$ or $(-1,\ldots,-1)$ is the global minima. Theorem 4, (5) establishes the good fluctuation region of the input signal. If $I(\nu)$ is in the region of $[\gamma e_1(w) + c(w)\mu + 1, \gamma e_2(w) + c(w)\mu - 1]$, w will remain as an attractor of the dynamics (15) (Fig. 4). The interval in which $I(\nu)$ changes can be taken as an estimate for an effective interval in the case when the input signal is not translation invariant and depends on the time τ of the neural dynamics. From all the above arguments one derives a useful way for choosing the four parameters μ,γ,s,I^X in order to have an on-center configuration which is an attractor of the dynamics (15).

From Theorem 4 one can show that if the input $I(\tau)$ increases, the size of the on-center of the configuration which is an attractor of the dynamics (15) will increase also(see Fig. 4) if $\mu > 0$, the opposite will happen if μ is negative. One important open question here is how to ensure the convergence of the algorithm. A simple way to achieve it, as one may suggest similar to that in the Kohonen network , is to shrink the size of the on-center field. This can be done by decreasing the input $I(\tau)$ as well(Fig. 4). However from our Theorem 4 it follows after some estimates that in general it is impossible to shrink arbitrarily the size of the on-center pattern if s is fixed. This is a main difference between the Kohonen network and the dynamic link model[KMM].

§5 Conclusions

This paper unifies the approach to the Hopfield model, the Linsker model and the dynamic link model, three typical networks arising from three typical areas in the study of the neural network. Since most of models proposed so far in the field of the neural networks use the sigmoid function in their dynamics of learning or retrieving procedure, as discussed in section 2, we are able to characterize the attractors of these models in terms of the present approach. So the power of the present method is not restricted to these three models reported here.

In the Hopfield model, we give a sufficient and necessary condition in order to check if a given pattern is an attractor of the network. The capacity of the network is reconsidered from a point of view different from the usual statistical physics approach. It is also obvious that we could apply our method here to analyze generalizations of the Hopfield model.

The present approach becomes more efficient if we are mainly interested in one or a few kind of patterns. This is the case in the Linsker network and in the dynamic link network. For the former network, the appearance of the on-center and oriented receptive field is the core of its dynamics. For the latter, the on-center structured pattern is an important one in its dynamics. The present paper asserts that the appearance of the structured receptive field in the Linsker model is universal in the sense that the appearance of the structured receptive field is independent of the specific choice of the clipping function used in the numerical simulation in the Linsker's network. For the dynamic link network, we propose a principle for the selection of these parameters employed in the model.

The significance of this unification approach is obvious. It helps us to understand the mechanism underlying each model more deeply. It furnishs a useful guidance in the practical application of these models, in particular in choosing the parameters of the dynamics. For example, the results in the Linsker network suggest a way to shrink the size of the on-center receptive field, which may help the convergence of the algorithm used in the pattern recognition based upon the dynamic link network; essentially the self-organization in the Linsker model and the dynamic link model is a procedure of retrieving 'memory', the structured receptive field ('memory') is stored in the model already; and so on. Besides these findings we report here and in [FT2], there is still a lot of work to be done further.

Acknowledgement. This paper is partially supported by the CNR of Italy.

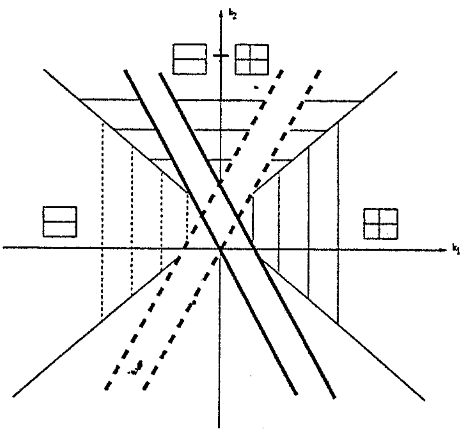


Fig. 1

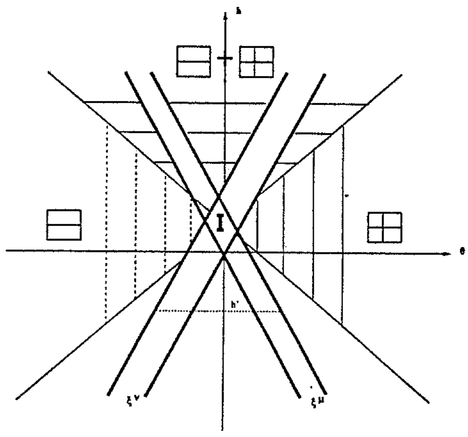
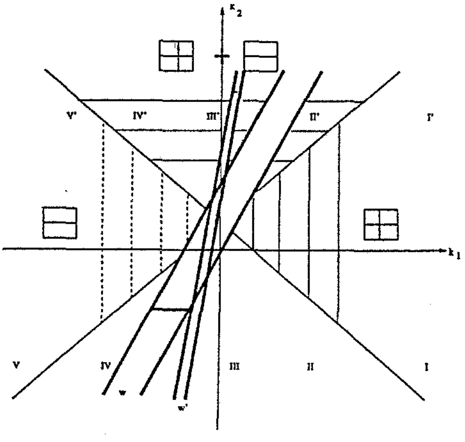


Fig. 2

Fig. 1. The parameter region of different saturated attractors of the dynamics (1). ⊞ is the parameter region of the all positive attractors $w_i = 1, i = 1, \cdots, N$ (Corollary 1). ⊟ is the parameter region of the all negative attractors $w_i = -1, i = 1, \cdots, N$ (Corollary 1) and ⊞ + ⊟ is the parameter region of the all positive and the all negative attractor. The region of (k_1, k_2) between two dark lines is the parameter region in which $w \neq (1, \cdots, 1), (-1, \cdots, -1)$ is an attractor of the dynamics (1). The region of (k_1, k_2) between two dash dark lines is the parameter region in which $-w$ is an attractor of the dynamics (1)(see also [FT2]).

Fig. 2. The parameter region of (θ, h) in which w is a saturated attractor of the Hopfield model (see Fig. 1 also). In the region I enclosed by dark lines, the Hopfield model has the highest capacity. In this region, for example, ξ^μ, ξ^ν are both attractors of the Hopfield model. When $h = h'$(horizontal dash line), the capacity of the model becomes lower.



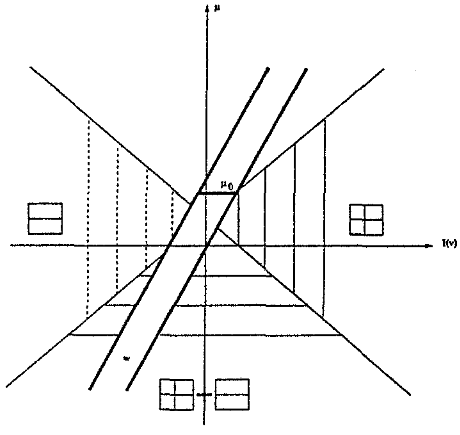


Fig. 3. The smaller is the size of the on-center, the narrower is the band in which that the on-center receptive field is an attractor([FPR]). The same conclusion is true for the off-center receptive field. When k_1 decreases($k_2 < 0$), the system passes through region I(all-excitatory), II(on-center), III(several attractors coexist), IV(off-center), V(all-inhibitory). If k_1 decreases($k_2 > 0$), we go from region I'(all-excitatory), II'(off-center), III'(several attractors coexist), IV'(on-center), V'(all-inhibitory). If we fix $k_1 > 0$, decrease $k_2 < 0$ the size of the on-center of the configuration which is an attractor of the Linsker model becomes smaller and smaller.

Fig. 4. The parameter region of $(I(\nu), \mu)$ in the dynamic link model. The slope of the dark lines is $c(w)$. So for fixed $\mu = \mu_0$, the smaller is the $I(\nu)$, the smaller is the radius of the on-center of the configuration which is an attractor of the dynamics (15). As r is small enough, the configuration with on-center radius r will no long be an attractor of the dynamics. For fixed $I(\nu) > 0$, when $\mu > 0$ becomes small, the size of the on-center of the configuration which is a saturated attractor of the dynamics (15) will become large.

References

[ATYD] Antonucci, M. & Tirozzi, B. & Yarunin, N.D. & Dotsenko, U.S.(1994), *Numerical Simulation of Neural Networks with Translation and Rotation Invariant Pattern Recognition*, Inter. Jour. of Modern Physics B, vol. 8, nos. 11 & 12, pp. 1529-1541.

[FP] Feng, J. & Pan, H.(1993),*Analysis of Linsker-type Hebbian Learning: Rigorous Results*, Proc. 1993 IEEE Int. Conf. on Neural Networks - San Francisco Vol. III, (Piscataway, NJ: IEEE) pp. 1516-1521.

[FPR] Feng, J. & Pan, H. & Roychowdhury, V. P.(1995), *A Rigorous analysis of Linsker's Hebbian learning network*, to appear in Advances of Neural Information Processing Systems, Vol .6 (NIP*S 94).

[FQ] Feng, J. F. & Qian, M. P.(1993), *Two-Stage Annealing in Retrieving Memories I*, In: Badrikian, A.; Meyer P-A, and Yan, J-A (ed.), Probability and Statistics, Singapore: World Scientific, pp. 149-176.

[FT1] Feng, J. & Tirozzi, B.(1995), *A Discrete Version of the dynamic link model*, to appear in Neural Computation.

[FT2] Feng, J. & Tirozzi, B.(1994), *An analysis of the dynamics with the sigmoid like function*, preprint.

[KMM] Konen, W. & Maurer, T. & von der Malsburg, C.(1994), *A fast dynamic link matching algorithm for invariant pattern recognition*, Neural Network, Vol. 7, Nos, 6/7, pp.1019-1030.

[Lin] Linsker, R.(1986), *From Basic Network Principle to Neural Architecture (series)*, Proc. Natl. Acad. Sci. USA, 83, 7508-7512, 8390-8394, 8779-8783.

[Mal] von der Malsberg, C.(1995),*Network self-organization in the ontogenesis of the mammalian visual system*, An Introduction to Neural and Electronic Networks, Second Edition, S.F. Zornetzer, J. Davis and C. Lau, Eds. (Academic Press).