ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 147-154

What is Observable in a Class of Neurodynamics?

Jianfeng Feng David Brown
Biomathematics Laboratory, The Babraham Institute, Cambridge
CB2 4AT, UK

Abstract. Nearly all models in neural networks start from the assumption that the input-output characteristic is a sigmoidal function. We study a class of dynamical systems of neural network models with saturated sigmoidal functions as their input-output characteristics. A complete spectrum on the parameter space for all possible outcomes of the dynamics is obtained. Under a stated condition we show that the possible outcomes of the dynamics are all saturated or all-but-one saturated fixed point attractors. An exact parameter region is given for all saturated attractors and all-but-one saturated fixed points.

1. Introduction

Interestingly, to date, sigmoidal functions have been utilized in the vast majority of neural networks as the input-output characteristic [7], either on the state space or on the weight space. The sigmoidal function is nearly saturated outside a region, by suitably adjusting some parameters of the function. In the present paper we report our recent work on neurodynamics defined by

$$y_i(t+1) = f(y_i(t) + \sum_{j=1}^{N}(a_{ij} + k_2)r_j y_j(t) + k_1), \ i = 1, \cdots, N, \ t = 1, \cdots, \quad (1)$$

where $\mathbf{y}(t) = (y_i(t), i = 1, \cdots, N) \in I\!R^N$, $A = (a_{ij}, i, j = 1, \cdots, N)$ an $N \times N$ matrix representing interaction between units (either weights or states), (k_1, k_2) are two key parameters of the dynamics, N can be thought of either as the number of neurons or the number of synaptic efficacies connected to the i-th neuron, $R = (r_i\delta_{ij}, i, j = 1, \cdots, N)$ plays the role of normalization with $r_i \geq 0, i = 1, \cdots, N$, f is a *saturated sigmoidal function* which is continuous and defined by

$$f(x) = \begin{cases} y_{\max} & \text{if } x > y_{\max} \\ f(x) & f(x) \text{ is strictly increasing for } x \in [y_{\min}, y_{\max}] \\ y_{\min} & \text{if } x < y_{\min} \end{cases} \quad (2)$$

Dynamics (1) can be realized either by a synchronous dynamics or asynchronous dynamics[1]. A variety of methods (see [3, 6]) have been developed

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 147-154

in recent years for exploring different aspects of the properties of dynamics (1) with some further assumptions on f or A. In [3] and references therein the behavior of a neurodynamical system with a *limiter function* as its input-output characteristic is analyzed. Nevertheless there are two severe restrictions which prevent a wide application of the approach presented in [3].

1. Limiter functions have been applied in a few models and, because of this, linear analysis can be carried out and informative results obtained as in [6]. But nearly all models of neural networks start from the assumption that the firing rate is a *sigmoidal* function of the summed inputs rather than a *linear* function, although both allow single neurons to make linear discriminations in the space of input features.

2. All results obtained in [3] are based upon an assumption that we exclusively consider the set of all saturated attractors and tacitly suppose that this is a generic case. This assumption is partly confirmed by Linsker [5] for the limiter function, using both numerical simulation and theoretical proof. But he also points out the possibility of the emergence of all-but-one saturated attractors, even in the limiter function case. When we consider the (more general) sigmoidal function case (1), we are not certain, at least at a first glance, what form of attractor occurs in general.

With the aim of providing a systematic and feasible tool for grasping some informative properties of dynamics (1), here we generalize the *saturated* fixed point attractor analysis on parameter space with *limiter* functions developed in [3] to fixed point attractor analysis with saturated sigmoidal functions. By this we mean the following.

- We perform an analysis on parameter space for the whole set of fixed point attractors rather than only the set of saturated fixed point attractors. We provide an argument to claim that we are only likely to observe the set of all saturated or all-but-one saturated attractors for dynamics (1). In other words, the generic outcome of dynamics (1) is an all saturated or all-but-one saturated attractor.

- We derive a necessary and sufficient condition to test whether a given all saturated (all-but-one saturated) state is an attractor (fixed point) or not for any given set of system parameters. This result in turn enables us to study Linsker's model in a more general setting and consider the set of all saturated attractors, as well as the set of all-but-one fixed points. Using extreme value theory in statistics, we give an exact parameter region for the threshold of the Hopfield model within which a stored pattern is an attractor[2].

Miller and MacKay [6] have carried out a detailed discussion directed at elucidating the effect of different constraints–subtractive and divisive. They show

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 147-154

that divisive enforcement causes the weight pattern to tend to the principal eigenvector of the synaptic development operator(matrix AR), whereas subtractive enforcement causes almost all weights to reach either their minimum or maximum values. Our results on dynamics (1) partly serve as a complement of their results: under an exactly given condition we assert that only all saturated or all-but-one saturated attractors are possible outcomes of the dynamics, similar to the situation of subtractive enforcement.

Applications of our results to competitive learning[4], the Hopfield model, Linsker's model and a continuous time model can be found in our full paper [2].

2. Fixed Point Attractors Of Dynamics (1)

2.1. All Saturated Fixed Point Attractors

The following definition and theorem are keys for our further development of the present paper. Without loss of generality we suppose that $y_{\text{min}} = -1$ and $y_{\text{max}} = 1$. As we already pointed out in [3] saturated states in the space $\{-1,1\}^N$ represent the most common outcome of many learning and retrieval models of neural networks and so we address the following definition.

Definition 1 *A fixed point attractor* $\mathbf{y} \in \{-1,1\}^N$ *is called a saturated attractor if*

$$k_1 + \sum_{j=1}^{N} (a_{ij} + k_2) r_j y_j \neq 0 \quad \text{for all } i \tag{3}$$

Restriction (3) will be relaxed gradually in the following subsection. The case that one unit violates condition (3) is dealt with in the next subsection; the case of more than one unit not satisfying (3) is discussed in subsection 2.3..

Motivated by the Hopfield model(see [2]) we introduce the following definition.

Definition 2 *The quantity*

$$h_i(\mathbf{y}) := \sum_{j \in J^+(\mathbf{y})} a_{ij} r_j - \sum_{j \in J^-(\mathbf{y})} a_{ij} r_j \tag{4}$$

is called the local field of the i-th neuron where $J^+(\mathbf{y}) = \{i, y_i = 1\}$, $J^-(\mathbf{y}) = \{i, y_i = -1\}$. *We say that there is a local field gap between neurons in* $J^+(\mathbf{y})$ *and* $J^-(\mathbf{y})$ *if and only if*

$$\min_{i \in J^+(\mathbf{y})} h_i(\mathbf{y}) > \max_{i \in J^-(\mathbf{y})} h_i(\mathbf{y}) \tag{5}$$

In spite of the fact that dynamics (1) is a generalization of what we consider in [3], the proof of the following theorem is similar to that of Theorem 2 in [3].

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 147-154

Theorem 1 *(see also theorem 2 in [3]) y is a saturated attractor of dynamics (1) if and only if*

$$d_1(\mathbf{y}) < k_1 + k_2 c(\mathbf{y}) < d_2(\mathbf{y}) \tag{6}$$

where the slope function $c(\mathbf{y}) = [\sum_{j\in J^-(\mathbf{y})} r_j - \sum_{j\in J^+(\mathbf{y})} r_j]$ *and two intercept functions*

$$d_1(\mathbf{y}) = \begin{cases} \max_{i\in J^+(\mathbf{y})}[-h_i(\mathbf{y})] & \text{if } J^+(\mathbf{y}) \neq \phi \\ -\infty & \text{otherwise} \end{cases} \tag{7}$$

and

$$d_2(\mathbf{y}) = \begin{cases} \min_{i\in J^-(\mathbf{y})}[-h_i(\mathbf{y})] & \text{if } J^-(\mathbf{y}) \neq \phi \\ \infty & \text{otherwise} \end{cases} \tag{8}$$

In other words, a saturated state $\mathbf{y}$ *is a saturated fixed point attractor of dynamics (1) if and only if there is a local field gap between neurons in* $J^+(\mathbf{y})$ *and* $J^-(\mathbf{y})$.

These two functions d_2 and d_1 were introduced in 1993(see [3] and references therein) but their physical meaning, extremes of local fields, is clear only after we apply Theorem 1 to the Hopfield model [2].

2.2. All-But-One Saturated Attractors

Now we consider the set of all-but-one saturated attractors. Without loss of generality we assume that $y_1 \in (-1, 1)$ is the only unsaturated state and $y_i \in \{-1, 1\}$ with

$$k_1 + \sum_{j=1}^{N}(a_{ij} + k_2) r_j y_j \neq 0 \tag{9}$$

for $i \neq 1$.

Since y_i, $i \neq 1$, are saturated fulfilling condition (9) our arguments of the previous subsection hold which imply that

$$\bar{d}_1(\mathbf{y}) < k_1 + \bar{c}(\mathbf{y}) k_2 < \bar{d}_2(\mathbf{y}). \tag{10}$$

for

$$\begin{cases} \bar{c}(\mathbf{y}) &= \sum_{j\in J^-(\mathbf{y})} r_j - \sum_{j\in J^+(\mathbf{y})} r_j + y_1 r_1 \\ \bar{d}_2(\mathbf{y}) &= \min_{i\in J^-(\mathbf{y})} [\sum_{j\in J^-(\mathbf{y})} a_{ij} r_j - \sum_{j\in J^+(\mathbf{y})} a_{ij} r_j + a_{i1} y_1 r_1] \\ \bar{d}_1(\mathbf{y}) &= \max_{i\in J^+(\mathbf{y})} [\sum_{j\in J^-(\mathbf{y})} a_{ij} r_j - \sum_{j\in J^+(\mathbf{y})} a_{ij} r_j + a_{i1} y_1 r_1] \end{cases} \tag{11}$$

Note that there is a slight difference between the definition of $\bar{d}$'s and d's: the maximum and minimum for d_2 and d_1 is taken over a set of N elements, but

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 147-154

for $\bar{d}_2$ and $\bar{d}_1$ it is over a set of $N-1$ elements. For y_1 we have the following identity

$$y_1 = f(y_1 + \sum_{j=1}^{N}(a_{1j} + k_2)r_j y_j + k_1)$$

or equivalently

$$d^{(1)}(\mathbf{y}) = k_1 + \bar{c}(\mathbf{y})k_2 = f^{-1}(y_1) - y_1 - \sum_{j\in J^+(\mathbf{y})} a_{1j}r_j + \sum_{j\in J^-(\mathbf{y})} a_{1j}r_j - a_{11}y_1r_1 \tag{12}$$

Hence the parameter region of (k_1, k_2) in which $\mathbf{y}$, an all-but-one saturated state, is a fixed point of dynamics (1) is not empty if and only if

$$\bar{d}_1(\mathbf{y}) < d^{(1)}(\mathbf{y}) < \bar{d}_2(\mathbf{y}) \tag{13}$$

Under condition (13) the parameter region for $\mathbf{y}$ to be a fixed point of dynamics (1) is the line given by

$$\{(k_1, k_2) : k_1 + \bar{c}(\mathbf{y})k_2 = d^{(1)}(\mathbf{y})\}$$

Theorem 2 *Under condition (9) an all-but-one saturated configuration* $\mathbf{y}$ *is a fixed point of dynamics (1) if and only if* (k_1, k_2) *is in the set*

$$\{(k_1, k_2) : \bar{d}_1(\mathbf{y}) < k_1 + \bar{c}(\mathbf{y})k_2 = d^{(1)}(\mathbf{y}) < \bar{d}_2(\mathbf{y})\} \tag{14}$$

Remark 1 For an all saturated configuration $\mathbf{y}$ except one unit say y_1 which violates restriction (3)

$$k_1 + \sum_{j=1}^{N}(a_{1j} + k_1)r_j y_j = 0 \tag{15}$$

we have a similar conclusion as Theorem 2, namely the parameter region in which $\mathbf{y}$ is a fixed point is line (14) inside a band.

2.3. Other Forms Of Attractors

For concreteness of expression we assume that $y_1, y_2 \in (-1, 1)$, the only two unsaturated states, and $y_i \in \{-1, 1\}$ with the property

$$k_1 + \sum_{j}(a_{ij} + k_2)r_j y_j \neq 0, \tag{16}$$

for $i \neq 1, 2$. After proceeding similarly as done above for all-but-one saturated configuration, we readily see that a necessary and sufficient condition for $\mathbf{y}$ to be a fixed point of dynamics (1) is

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 147-154

$$\tilde{d}_2(\mathbf{y}) > k_1 + k_2\tilde{c}(\mathbf{y}) > \tilde{d}_1(\mathbf{y}) \tag{17}$$

and

$$\begin{cases} d_1^{(2)}(\mathbf{y}) := k_1 + k_2\tilde{c}(\mathbf{y}) &= \sum_{j\in J^-(\mathbf{y})} a_{1j}r_j - \sum_{j\in J^+(\mathbf{y})} a_{1j}r_j - a_{11}y_1r_1 \\ & \quad -a_{12}y_2r_2 + f^{-1}(y_1) - y_1 \\ d_2^{(2)}(\mathbf{y}) := k_1 + k_2\tilde{c}(\mathbf{y}) &= \sum_{j\in J^-(\mathbf{y})} a_{2j}r_j - \sum_{j\in J^+(\mathbf{y})} a_{1j}r_j - a_{21}y_1r_1 \\ & \quad -a_{22}y_2r_2 + f^{-1}(y_2) - y_2 \end{cases} \tag{18}$$

where

$$\begin{cases} \tilde{c}(\mathbf{y}) &= \sum_{j\in J^-(\mathbf{y})} r_j - \sum_{j\in J^+(\mathbf{y})} r_j + y_1r_1 + y_2r_2 \\ \tilde{d}_2(\mathbf{y}) &= \min_{i\in J^-(\mathbf{y})} [\sum_{j\in J^-(\mathbf{y})} a_{ij}r_j - \sum_{j\in J^+(\mathbf{y})} a_{ij}r_j + a_{i1}y_1r_1 + a_{i2}y_2r_2] \\ \tilde{d}_1(\mathbf{y}) &= \max_{i\in J^+(\mathbf{y})} [\sum_{j\in J^-(\mathbf{y})} a_{ij}r_j - \sum_{j\in J^+(\mathbf{y})} a_{ij}r_j + a_{i1}y_1r_1 + a_{i2}y_2r_2] \end{cases} \tag{19}$$

An interesting new phenomenon occurs: the two lines corresponding to the two unsaturated states defined by Eq. (18) are *parallel*, which indicates that as long as

$$d_1^{(2)}(\mathbf{y}) \neq d_2^{(2)}(\mathbf{y}) \tag{20}$$

then the parameter region in which $\mathbf{y}$ is a fixed point of dynamics (1) is *empty*. The fulfillment of Eq. (20) is a generic situation essentially depending on the property of the matrix A. When $f(x) = x,\ x \in I\!R$ this conclusion has been confirmed by Linsker[5] in his numerical simulation and theoretical proof.

Theorem 3 *If and only if* (k_1, k_2) *is in the following set*

$$\{(k_1,k_2); \tilde{d}_1(\mathbf{y}) < k_1 + \tilde{c}(\mathbf{y})k_2 = d_1^{(2)}(\mathbf{y}) = d_2^{(2)}(\mathbf{y}) < \tilde{d}_2(\mathbf{y})\} \tag{21}$$

an all-but-two saturated configuration $\mathbf{y}$ *is a fixed point of dynamics (1).*

Remark 2 If for an all saturated configuration there are two saturated units not satisfying Eq. (3) or an all-but-one saturated configuration with one saturated unit violating Eq. (9) we have a similar conclusion as in Theorem 3.

We are able to carry out a cascade study, continuing to consider three unsaturated units and so on. The situation to ensure the existence of a nonempty parameter region in which $\mathbf{y}$ is a fixed point of dynamics (1) becomes more and more difficult when the number of unsaturated units is larger and larger since it requires all parallel lines corresponding to unsaturated units to intersect. In general two parallel lines $k_1 + \tilde{c}(\mathbf{y})k_2 = d_2^{(2)}(\mathbf{y})$ and $k_1 + \tilde{c}(\mathbf{y})k_2 = d_1^{(2)}(\mathbf{y})$ do

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 147-154

not coincide. Hence we stop here and believe that the general outcomes of dynamics (1) are all saturated and all-but-one saturated attractors.

In conclusion for dynamics (1) the full spectrum of its outcomes is summarized in table 1.

ATTRACTOR TYPE	PARAMETER REGION	POSSIBILITY
$\mathbf{y}$: all saturated	A band determined by $d_2(\mathbf{y}) > k_1 + c(\mathbf{y})k_2 > d_1(\mathbf{y})$ (Theorem 1), independent of f	Most Possible
$\mathbf{y}$: all-but-one saturated	A line $k_1 + \bar{c}(\mathbf{y})k_2 = d^{(1)}(\mathbf{y})$ inside a band $\bar{d}_1(\mathbf{y}) < k_1 + \bar{c}(\mathbf{y})k_2 < \bar{d}_2(\mathbf{y})$ (Theorem 2), dependent on f	Less Possible
$\mathbf{y}$: all but two saturated	Intersection of two parallel lines inside a band (Theorem 3), dependent on f	Hardly possible
	⋮	

Table 1: The General Parameter Region

3. Conclusions

We have studied the dynamics of neural network models with saturated sigmoidal functions as their input-output characteristics. A complete spectrum on the parameter space for all possible outcomes of dynamics (1) is obtained. Under a stated condition we have shown that the possible outcomes of dynamics (1) are all saturated or all-but-one saturated fixed point attractors. An exact parameter region is given for all saturated attractors and all-but-one saturated fixed points. Our approach provides a systematic and feasible tool to deal with nonlinear dyanamics in many neural network models[2].

In a single theoretical framework we have managed to treat diverse models in neural networks[2]. The significance of this unified treatment lies in that, in additional to some novel discoveries after revisiting these models, we have exposed some common mechanisms behind them (for example we have understood the physical meaning of d_2 and d_1 from the study of the Hopfield model [2]) which will provide useful guidance in further designing and understanding new models, both for learning and retrieving.

Acknowledgement. The paper was partially supported by a BBSRC LINK program and an ESEP of the Royal Society.

References

[1] Feng, J. 1997. Lyapunov functions for neural nets with nondifferentiable input-output characteristics. *Neural Computation* **9**, 45-51.

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 147-154

[2] Feng, J., and Brown, D. 1997. Fixed point attractors for a class of neurodynamics. *Neural Computation* **10**(1), (in press).

[3] Feng, J., Pan, H., and Roychowdhury, V. P. 1996. On neurodynamics with limiter function and Linsker's developmental model. *Neural Computation* **8**, 1003-1019.

[4] Goodhill, G., and Barrow, H.G. 1994. The role of weight normalization in competitive learning. *Neural Computation* **6**, 255-269.

[5] Linsker, R. 1986. From basic network principle to neural architecture (series). *Proc. Natl. Acad. Sci. USA* **83**, 7508-7512, 8390-8394, 8779-8783.

[6] Miller, K., and MacKay, D. 1994. The role of constraints in Hebbian learning. *Neural Computation* **6**, 100-126.

[7] Sejnowski, T.J. 1995. Time for a new neural code? *Nature* **376**, 21-22.