# On Classification of Some Hopfield-Type Learning Rules via Stability Measures

Mohammad Reza Rajati, Mohammad Bagher Menhaj

**Summary.** This paper first reviews several learning methods for training Hopfield-type associative memories as well as a novel architecture with neurons of nonmonotonic stimulus functions. These learning rules are classified into three groups according to a measure of stability closely related to the storage capacity. This measure helps us better study the ability of a network to store patterns as stable states of its dynamics in case it is highly loaded. We then analyze the experimental data related to the stability measure and classify the previously studied learning methods according to the measure. We also show that the behavior of those learning rules converges to either the behavior of Hebbian learning or that of the pseudo-inverse method.

**Key words:** Hopfield-type neural networks, Learning rules, Stability measures, Storage capacity, Equilibrium points.

## 1 Introduction

Hopfield-type neural networks are shown to be amenable to thorough analysis. They have simple synthesis procedures and interesting aspects for scientific investigations including those related to the content addressability [3], storage capacity [8], robustness against noise and adaptability to the neurons' malfunction [10]. Fundamentally, networks of Hopfield-type suffer from poor capacity and performance. Many architectures are innovated to remedy the restrictions of the Hopfield associative memory, including those trying to modify the connections and updating schemes (for example [12]) and those employing new learning rules [2].

This paper tends to classify some architectures and learning rules according to a stability condition related to the storage capacity of the network. Of course, there are many other capacity and performance indices, which are open to examine for Hopfield-like architectures, and we leave them to the forthcoming articles. We first review the conventional Hopfield network model. Then we discuss the capacity issues in recurrent associative memories. Some learning methods for Hopfield networks are discussed in Sect. 3 and finally we classify them according to a stability measure by experimental analysis.

## 2 The Hopfield Model and Its Storage Capacity

The Hopfield network is a recurrent neural network governed by the difference equation:

$$\begin{aligned} & a_i(0) = p_i,\ n_i(t+1) = \sum_j^N w_{ij} a_j(t) + b_i,\ a_i(t+1) = f(n_i(t+1)) \\ & f(u) = \mathrm{sgn}(u) \end{aligned} \tag{1}$$

where $p_i$ is the unknown pattern to be recognized, $a_i$ is the output of the network, $w_{ij}$ is the connection weight between neurons $i, j$, $b_i$ is the threshold term and the stimulus function is $f(u)$. This is a discrete version of the Hopfield model. The connection weight matrix should be calculated to store the required prototype patterns as fixed states of the network dynamics so that patterns can be recalled from noisy or incomplete initial inputs.

### 2.1 Storage Capacity

The storage capacity of the Hopfield-like associative memories is of great consideration in the neurocomputing literature. It is formulated by either big O notation in terms of the number of neurons ($M_c$), or the relative capacity $\alpha_c$ defined as $\alpha_c = L/N$, where L is the number of patterns stored and N is the number of neurons.

Although one can place any load upon a neural system, there is obviously a value for $\alpha$ above which some of the vectors in the training set will not be stored as stable states. We refer to this as the maximum permissible loading (or just loading) and denote it by $\alpha_{\max}$. Both of the aforementioned formulations are deeply discussed and it has been shown that for randomly realized unbiased binary patterns, $\alpha_{\max} \approx 0.14$ and $M_c = N/\log N$ [8].

## 3 Some Learning Methods and Architectures

There are several methods to obtain a weight matrix with higher performance for a recurrent associative memory. We consider here, Hebbian, pseudo-inverse, Menhaj–Seifipour, and Li–Michel learning rules. Besides, we examine the architecture proposed by Yanai and Amari [12] which uses nonmonotonous stimulus functions for the neurons.

### 3.1 Hebbian Learning

This is the conventional learning rule for recurrent associative memories with:

$$w_{ij} = \sum_{l=1}^{L} p_i^l p_j^l \tag{2}$$

### 3.2 The Pseudo-Inverse Method

The Pseudo-inverse rule is introduced by Personnaz et al. [11] and studied deeply by Yen and Michel [13] to generate the weight matrix according to the rule:

$$W = PP^{+} \tag{3}$$

where $P$ is the matrix whose columns are the $p^l$and $P^{+}$ is its pseudo-inverse, the matrix with the property that: $P^{+}P = I$. It is notable that many modified methods such as Perceptron-style methods are approximate versions of the pseudo-inverse method [3–5].

### 3.3 Menhaj–Seifipour Algorithm

Menhaj and Seifipour propose a new algorithm, and state that it has a better storage capacity and a higher speed of convergence [9]. They build the memory matrix by:

$$r_{ij} = \frac{1}{2^L}\prod_{l=1}^{L}(p_i^l + p_j^l) \tag{4}$$

$$W = R^T R \tag{5}$$

This network is proved to minimize the energy function:

$$E(t) = -\sum_i \sum_j w_{ij}[a_i(t) + a_j(t)]^2 - 4\sum_i b_i(t)a_i(t) \tag{6}$$

Additionally, the matrix built by the above rule is sparse and results in a higher decrease of the energy function in each time-step than the classical Hopfield network with Hebbian learning. The architecture of sparsely connected networks is studied by Liu and Michel and the sparse nature of the networks is proved to be beneficial [7].

### 3.4 Li–Michel Learning Rule

Li–Michel synthesis method relies on a vigorous mathematical foundation, i.e., analysis of linear systems operating on a hypercube [6]. For the sake of clarity, we just present a brief algorithm, without thorough theoretical considerations.

To store L prototype patterns in a Hopfield-type memory as asymptotically stable equilibrium points, let:

$$X = [x_1, x_2, \ldots, x_{L-1}],\ x_i = p^i - p^L, i = 1, 2, 3, \ldots, L-1 \tag{7}$$

Then obtain a Singular Value Decomposition of $X$:

$$X = USV^T, U = [u_1, u_2, \ldots, u_N] \tag{8}$$

Suppose that k is the rank of S, then k is the dimension of the space spanned by $x_i$'s. Then the weight matrix and the bias vector are obtained by:

$$W^{+} = \sum_{j=1}^{k} u_j u_j^T \quad W^{-} = \sum_{i=1}^{N} u_i u_i^T \tag{9}$$

$$W^{+} = \alpha W^{+} - \beta W^{-} b = \alpha p^L - W p^L \tag{10}$$

in which $\alpha, \beta$ are properly selected constants which satisfy: $\alpha > 1, \beta < 1$.

This algorithm guarantees the system to have at most $3^N$equilibrium points. Additionally, at most $2^N$of the equilibrium points are asymptotically stable.

### 3.5 Yanai–Amari Architecture

Yanai and Amari [12] propose an associative memory with two stage nonlinear dynamics:

$$a(k+1) = \mathrm{sgn}[W(a - f(Wa))] = \mathrm{sgn}[Wa - Wf(Wa)] \tag{11}$$

They use a nonmonotonic function for their network:

$$f(u) = \begin{cases} a(u+h) - c & u < -h \\ 0 & -h \le u \le h \\ a(u-h) + c & u > h \end{cases} \tag{12}$$

where $h$ and $c$ are non-negative. We used $a = 0.4,\ h = 0.1,\ c = 0$ in this paper.

## 4 Empirical Analysis of Storage Capacity

To obtain the relative capacity of the models examined, we trained Hopfield-type networks of 100 neurons, with a set of random unbiased prototype patterns. Loading was increased and the response of the network to an erratic version of one of the stored patterns (with a Hamming distance of 10) is evaluated, and the normalized overlap $(1 - H_d/N)$ of the response and the stored pattern is illustrated in Fig. 1. The absolute capacity is usually defined as the maximum loading in which the network can recall a pattern more than 90% (and sometimes exactly 100%) perfectly.

It is obvious from Fig. 1 that $\alpha_{\max}$ is 0.15, 0.35, 0.25, and 0.05 for Hebbian, pseudo-inverse, Li–Michel and Menhaj–Seifipour learning rules, respectively. It is also notable that in high loadings, Hebbian learning is able to recall the patterns with a 70% overlap, but other rules recall the patterns with more than 94% overlap, although loaded highly. Yanai and Amari reported $\alpha = 0.3$ for their architecture [12].

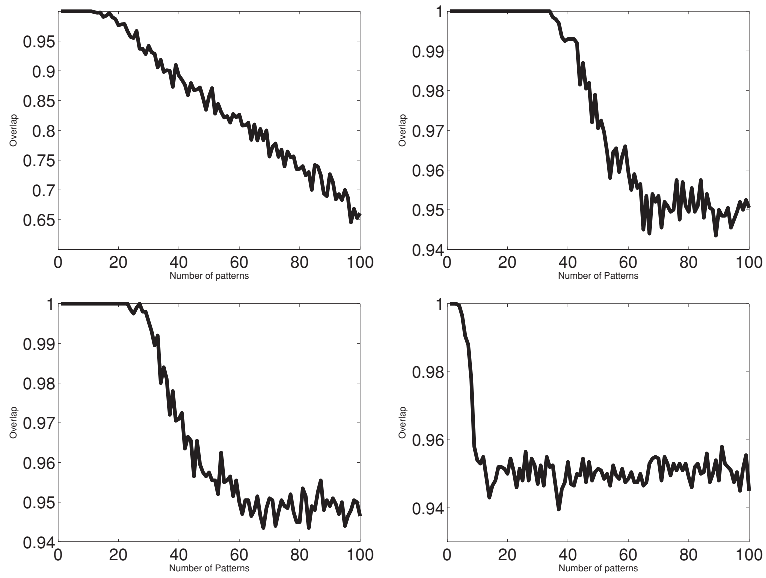


**Fig. 1.** Pattern recall by (**a**) Hebbian learning (**b**) pseudo–inverse rule (**c**) Li–Michel rule (**d**) Menhaj–Seifipour training method. Each network has 100 neurons and is trained by random patterns. Every point on the plots is generated by averaging over 20 runs. N = 100

## 5 Classification of Hopfield Memories via a Stability Measure

Abbott classified all Hopfield models into three groups. Any member of each group may have a different behavior when the loading upon it is not near $\alpha_{\max}$, but all members loaded near $\alpha_{\max}$ have the same behavior [1].

From the dynamic equations of the network, it can be seen that a state $a$ will be stable if $n_i$ has the same sign as $a_i$ for all $i$. So, the parameter $n_i a_i$ should be non-negative for all i in order for the network to have a pattern $p$ as its stable equilibrium point. Furthermore, assume a network with a set of stable states. The weight matrix could be scaled by any positive number, and thereby the synaptic signals will increase (and obviously the $n_i a_i$'s) but the domains of attraction of the stable states will not get wider. Thus, the following stability measure is defined to characterize the nature of stable states:

$$\gamma_i^l = \frac{n_i^l p_i^l}{\|W_i\|} \qquad \|W_i\| = \sqrt{\sum_{j=1}^{N} (w_{ij})^2} \tag{13}$$

Considering the worst case analysis, the minimum value of $\gamma_i^l$s is a parameter for identification of the network's basins of attraction [5].

The global groups of recurrent networks are different in the distribution of their $\gamma$ values. The first group, known as Hopfield group, has a normal distribution with a mean of $1/\sqrt{\alpha}, \alpha < 0.15$. In this group of models, negative values of $\gamma$ could be present in the network, and this is a sign of the existence of unstable patterns (Hopfield network with Hebbian learning is within this group of models).

The second group has matrices of pseudo-inverse type. The $\gamma$ values theoretically converge to the same value $\gamma_0 = \sqrt{(1-\alpha)/\alpha}$. So we suppose a notch distribution of $\gamma$ values in our numerical results. The third group has a clipped normal distribution, with positive $\gamma$ values [4].

## 6 Classification of the Models via Experimental Analysis

In this section, we analyze the $\gamma$ distributions of different algorithms in the paper. Figure 2 depicts different $\gamma$ distributions for the learning rules. The distributions are plotted by training networks consisting 1,000 neurons with a set of 500 unbiased random bipolar patterns.

It is easily observed that Hebbian learning causes a normal distribution of $\gamma$'s. It is notable that many values of $\gamma$ are negative, so in high loadings, some of the patterns will not be stored as stable states.

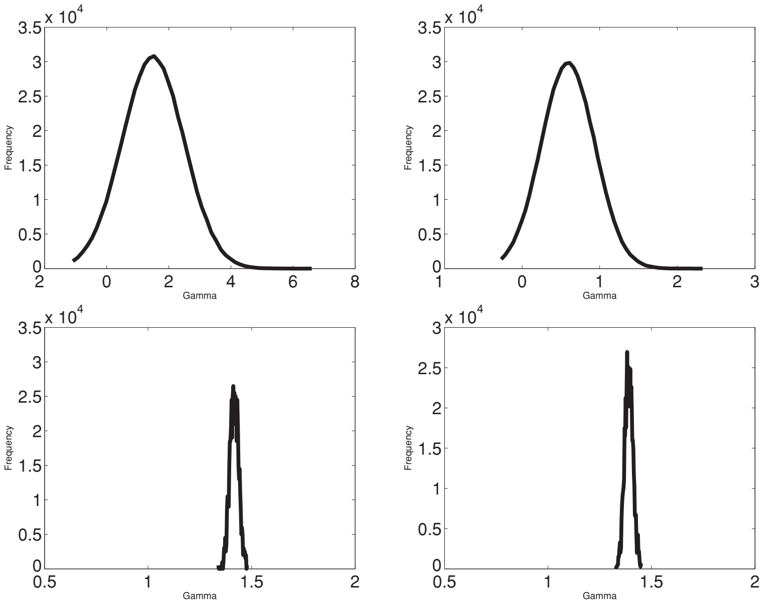


**Fig. 2.** Gamma distributions of (**a**) Hebbian learning (**b**) network with non-monotonic neurons (**c**) pseudo-inverse rule (**d**) Li–Michel learning rule. Each network is trained by 500 patterns and has 1,000 neurons

The pseudo-inverse rule, as the canonical model of the second group, is tested by $\gamma$ values as well and a notch distribution is resulted. Li–Michel learning rule results in a very notch distribution with no negative values. Thereby, it could be concluded that this rule falls into the pseudo-inverse class of models. It may be a cause of its high performance (this procedure was performed many times and $\gamma$ never became negative).

Yanai and Amari state that their proposed model is an approximation of the pseudo-inverse rule. The $\gamma$ distribution of this rule has negative values, so it could not be classified into the pseudo-inverse group, although the $\gamma$ values are not distributed very widely, and it is less likely that the $\gamma$'s be negative (in some cases, the authors encountered distributions without negative values). The architecture is therefore classified into the first group.

About Menhaj–Seifipour learning rule, we inspected the convergence properties of the network's $\gamma$ distribution (Fig. 3). We plotted the $\gamma$ distributions in different loadings. In high loadings the $\gamma$ values converge to the same amount of $\gamma_0 = 1$ and this fact helps us figure out that this rule could be classified in the pseudo-inverse group of models. When the number of stored patterns becomes large even though the loading is low, the matrix constructed by the

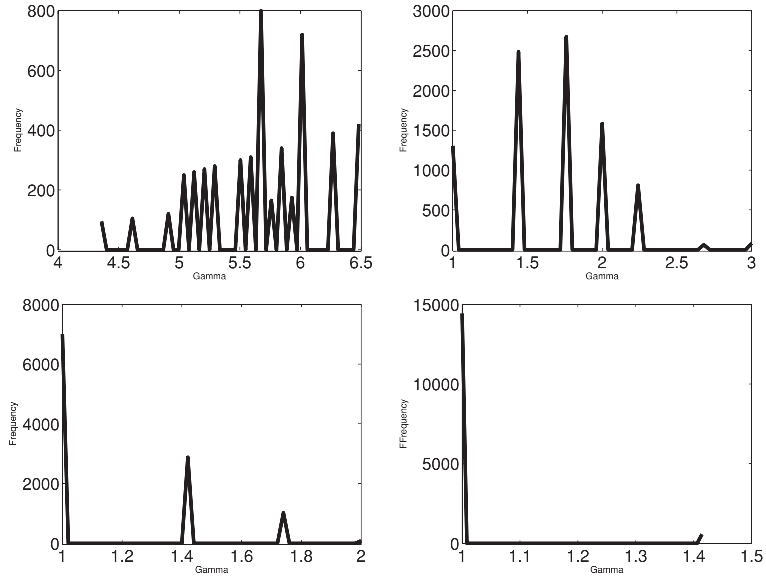


**Fig. 3.** Gamma distributions of Menhaj rule with (**a**) L = 5 (**b**) L = 9 the network is not loaded very highly and the distributions behavior is not similar to that in the saturation. With (**c**) L = 11 (**d**) L = 15 the network's distributions converge to a notch distribution. The notch distribution is reached when the number of patterns grows and it does not related to the number of neurons, because the resulting weight matrix converges to the identity matrix. For biased data this convergence will be slower. Here the number of neurons is N = 1,000

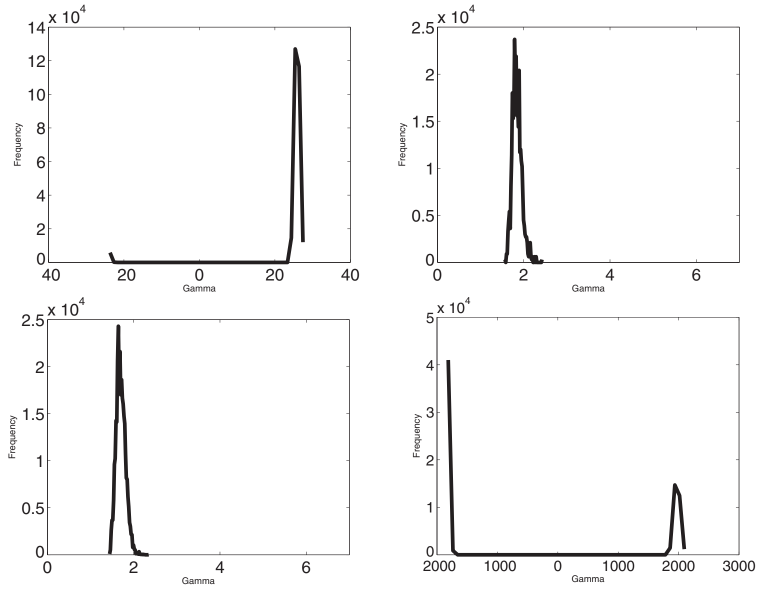


**Fig. 4.** Gamma distributions for biased patterns. (**a**) Hebbian (**b**) pseudo-inverse (**c**) Michel (**d**) Yanai–Amari N = 1000, L = 300

rule will converge to the identity matrix and all $\gamma$ values become the same, which is similar to the theoretically derived distribution for the pseudo-inverse group of models.

We also analyzed the effect of bias on patterns in $\gamma$ distributions. We trained the networks with a set of random patterns in which the probability of presence of 1 is 90%, in contrast to the case of unbiased patterns, in which this probability is 50% (Fig. 4). It is obvious that Hebbian learning and Yanai model are not tolerant to biased data, in contrast to Michel and pseudo-inverse rules which maintain their notch distribution.

## 7 Conclusion

In this article, we classified different algorithms of learning via stability measures proposed by Abbott. We showed that Li–Michel and Seifipour–Menhaj rules' behaviors converge to that of the pseudo-inverse method. The behavior of the Yanai–Amari network model converges to that of the Hopfield model, although because of its nature, it has a similarity with the pseudo-inverse group of models. We leave deeper discussions about the information theoretical aspects of Menhaj–Seifipour rule, finding the shape and size of the attraction basins of the discussed rules and the convergence properties of the networks trained by them to future contributions.

## References

1. Abbott, L.F., Kepler, T.B.: Universality in the Space of Interactions for Network Models, J. Phys. A: Math. Gen. 22 (1989) 2031–2038
2. Athitan G.: A Comparative Study of Two Learning Rules for Associative Memory, Pramana-J. Phys. 45(6) (1995) 569–582
3. Forrest, B.M: Content Addressability and Learning in Neural Networks, J. Phys. A 21 (1988) 245–255
4. Gardner, E.: The Space of Interactions in Neural Networks, J. Phys. A 21 (1988) 257–270
5. Krauth W., Mezard M.: Learning Algorithms with Optimal Stability for Neural Networks, J. Phys. A 20 (1987) L745–L752
6. Li J.H, Michel A.N., Porod W.: Analysis and Synthesis of a Class of Neural Networks: Linear Systems Operating on a Closed Hypercube, IEEE Trans. Circuits Syst. 36(11) (1989) 1405–1422
7. Liu D., Michel A.N.: Sparsely Interconnected Neural Networks for Associative Memories with Applications to Cellular Neural Networks, IEEE Trans. Circuits Syst.-II: Analog and Digital Signal Processing, 41(4) (1994) 205–307
8. McEliece, R.J., Posner, E.C., Rodemich, E.R., Venaktash, S.S.: The Capacity of the Hopfield Associative Memory. IEEE Trans. Inf. Theory, 33(4) (1987) 461–482
9. Menhaj M.B., Seifipour N.: A New Implementation of Discrete–Time Hopfield Network with Higher Capacity and Speed of Convergence, Proc. IJCNN '01, Vol. 1. (2001) 436–441
10. Nouri Shirazi, Mahdad, Nouri Shirazi, Mehdi, Mackawa, S.: The Capacity of Associative Memories with Malfunctioning Neurons, IEEE Trans. Neural Networks, 4(4) (1993) 628–635
11. Personnaz L., Guyon I., Dreyfus G.: Collective Computational Properties of Neural Networks: New Learning Mechanisms, Phys. Rev. A 34(5) (1986) 4217–4228
12. Yanai, H.F., Amari S.: Auto–Associative Memory with Two–Stage Dynamics of Nonmonotonic Neurons. IEEE Trans. Neural Networks, 7(4)(1996) 803–815
13. Yen, G., Michel, A.N.: A Learning and Forgetting Algorithm in Associative Memories: Results Involving Pseudo–Inverses, IEEE Trans. Circuits. Syst., 38(10) (1991) 1193–1205