01 Jan 1999

Frequency Domain Robustness Analysis of Hopfield and Modified Hopfield Neural Networks

Jie Shen

S. N. Balakrishnan
Missouri University of Science and Technology, bala@mst.edu

Proceedings of the American Control Conference
San Diego, California • June 1999

Frequency Domain Robustness Analysis of Hopfield and Modified Hopfield Neural Networks

Jie Shen S. N. Balakrishnan *
Department of Mechanical and Aerospace Engineering
and Engineering Mechanics
University of Missouri-Rolla
Rolla, MO 65401

Abstract

A variant of Hopfield neural network called modified Hopfield network is formulated in this study. This class of networks consists of parallel recurrent networks which have variable dimensions that can be changed to fit the problem under consideration. It has a structure to implement an inverse transformation that is essential for embedding optimal control gain sequences. Equilibrium solutions of this network are discussed. The robustness of this network and the classical Hopfield network are carried out in the frequency domain using describing functions.

1 INTRODUCTION

A neural network is a parallel, distributed information processing structure consisting of processing elements [1]. Artificial neural networks (ANN) represent an emerging technology rooted in many disciplines. Control theory is not only gaining benefits from ANN, but also contributing to the development of ANN [2].

In this paper, a variation of the Hopfield network (HN) called Modified Hopfield neural network (MHNN) is proposed [3]. Based on the equilibrium analysis, these networks can perform an inverse transformation on matrices and other auxiliary mathematical operations. This feature allows the networks to produce optimal control gain sequences. Unlike any other existing ANN [4, 5, 6], inputs to the networks are the parameters of system dynamics and control matrices, and the outputs are gain matrices [7, 8].

In control field, stability and robustness are two basic concerns for any controller design. The stability and robust property of MHNN will be investigated in this study. Due to the nonlinearity of the activation function in ANN, proper definition and measurement of robustness are needed. A linear equivalence of the nonlinear activation function will be obtained using the describing function technique and the Bode diagram of the whole system is drawn based on a sigmoidal input. The robustness properties of the recurrent ANN in the frequency domain can be established based on the Bode diagram.

*Associate Fellow, AIAA (to whom all correspondence should be sent)

2 MHNN AND ITS STABILITY

The Modified Hopfield Neural Network (MHNN) is a variant of the classical HN. Its dynamical model is shown in Fig. 1.

Since it is derived from HN, MHNN keeps the beneficial characteristic of the former; it is stable in the Lyapunov's sense.

2.1 Stability

The stability of MHNN will be demonstrated by analyzing its dynamics and using the energy function. The network has two clusters of neurons. The right part of the network is characterized by outputs $\phi_1, \phi_2, \cdots, \phi_m$ which are transformed by nonlinear functions f from their states $u_i, u_2, \ldots, u_m$; m is the number of outputs of neurons in the right part.

$$\phi_j = f_j(u_j) \tag{1}$$

with b_j the input current and v_i the output of the left cluster of amplifiers. The conductance w^r_{ij} connects the output of the jth neuron in the left part to the input of the ith neuron in the right part, which is indicated in Fig. 1 as ■. The superscript r indicates the location of weight w_{ij} in the right part of the network; n is the number of outputs of neurons in the left part.

Now, by Kirchhoff's law

$$-\sum_{i=1}^{m} w^l_{ji}\phi_i(t) - a_j = \frac{x_j(t)}{R_j} + c_j\frac{dx_j(t)}{dt} \qquad j = 1, \ldots, n. \tag{2}$$

Now define the following function as an energy function E for MHNN as in [6]

$$E(v) = \sum_{j=1}^{n} a_j v_j + \sum_{j=1}^{n} \frac{1}{R_j}\int_0^{v_j} f^{-1}(v)dv + \sum_{j=1}^{n}\int \left(\sum_{i=1}^{m} f(\sum_{k=1}^{n} v_k w^r_{ki} - b_i) w^l_{ij}\right) dv_j. \tag{3}$$

The time derivative of the energy function is

$$\frac{dE}{dt} = -\sum_{j=1}^{n} c_j g^{-1}(v_j)\left(\frac{dv_j}{dt}\right)^2. \quad (4)$$

after some algabra. By the same arguement in [6], the model is stable in accordance with Lyapunov's theorem.

This means that the evolution of dynamic system in state space always seeks the minima of the energy surface E.

2.2 Solution

In order to obtain the analytic estimation for the converged value of the networks, it is assumed that the network signals are small, and that they work in the linear region of the amplifiers. These assumptions are reasonable because ANN signals are usually normalized.For those big signals, log sigmoid and tangent sigmoid functions will limit their outputs magnitude, and their effects are localized and will not be propagated. Assuming tangent sigmoid activation functions have equal amplifer gains, we get

$$\begin{aligned} C\frac{dX}{dt} &= -A - GX - W^L k_f (W^R k_g X - B) \quad (5) \\ &= -(G + k_f k_g W^L W^R)X + k_f W^L B - A \end{aligned}$$

When the networks reach equilibrium, $dX/dt = 0$, and

$$V = \left(W^L W^R + \frac{G}{k_f k_g}\right)^{-1}\left(W^L B - \frac{A}{k_f}\right) \quad (6)$$

3 OPTIMAL CONTROL APPLICATION

3.1 Problem Formulation

Let the plant to be

$$x_{k+1} = A_k x_k + B_k u_k \quad (7)$$

with $x_k \in \mathbf{R^n}$ and $u_k \in \mathbf{R^m}$. The associated performance index is the quadratic function

$$J_i = \frac{1}{2} x_N^T S_N x_N + \frac{1}{2}\sum_{k=i}^{N-1}\left(x_k^T Q_k x_k + u_k^T R_k u_k\right) \quad (8)$$

defined over the time interval of interest $[i, N]$. We assume that Q_k, R_k and S_N are symmetric positive semidefinite matrices, and in addition that $|R_k| \neq 0$ for all k.

The objective is to find the control sequence $u_i, u_{i+1}, \cdots, u_{N-1}$ to minimize J_i.

The solutions [3] are given by

$$u_k = -K_k x_k, \qquad k < N$$

where the Kalman gain K_k is given by

$$K_k = \left(B_k^T S_{k+1} B_k + R_k\right)^{-1} B_k^T S_{k+1} A_k \quad (9)$$

In terms of the Ricatti variable S_k, now

$$S_k = A_k^T S_{k+1}\left(A_k - B_k K_k\right) + Q_k \quad (10)$$

In the application where the control interval is finite, S_N will be given. Alternatively use Eqns (9) and (10), we will get a series of K_k.

3.2 Network Solution/Implementation

The crucial operation here is the inverse to get the Kalman gain. The MHNN contain both invariant and variable parameters. Invariant parameters are fixed in the neuron-computing model, while variable parameters can be modified. By comparing Eqns. (9) and (10) with the stable output of the network Eqn. (6), the network will produce the Kalman sequence if set $W^L = B_k^T S_{k+1}$, $W^R = B_k$, $\frac{G}{k_f k_g} = R_k$, and $B = A_k$, $A = 0$.

4 HN & MHNN BLOCK DIAGRAMS

Since the potential of ANN come from their nonlinear transformation [7], how to deal with and analyze the nonlinearity in ANN becomes an important part of this study. In general, two philosophical approaches can be used in dealing with the nonlinearities in the networks. One is the "macro" analysis, which treats ANN collectively as black boxes and pays attention only to their input-output reactions regardless of how complex their inner structures are. The other one is the "micro" analysis, which studies mathematically or biologically the potential of every neuron so that the whole cluster of neurons will have a summation effect of every contributing element. In this study, a mixed approach is taken.

Since the whole mechanism of ANN derives from the functions of its basic unit — a neuron, the study of neurons will contribute to the research of the whole networks. After the understanding of the individual element, the overall system is synthesized or analyzed.

4.1 Open-loop Hopfield Network

HN is a feedback system in nature. In order to see the effect of the feedback mechanism, the open-loop characteristic is analyzed first. Referring to Fig. 3(a) with the loop broken at ■. which is indicated by "w", one can obtain Fig. 2(a). By Kirchhoff's Law, one can have

$$\begin{aligned} c\frac{du}{dt} &= a - ug \quad (11) \\ v &= f(u) \quad (12) \end{aligned}$$

Equation (11) can be regarded as a state equation while Eqn. (12) an output equation. The corresponding block diagram is shown in Fig. 2(b). The transfer function is obtained as

$$h(s) = \frac{1}{cs + g}\cdot f(s). \quad (13)$$

In order to maintain the simplicity of description, for the time being, it is assumed that $f(s)$ is a constant. (Later it can be seen, $f(s)$ is equivalent to a varying constant.) Then $h(s)$ can be written as

$$h(s) = \frac{f}{cs + g}. \tag{14}$$

4.2 Closed-loop Hopfield Network

Similar to the open-loop analysis, the Kirchhoff's Law produces

$$c\frac{du}{dt} = a + wv - ug \tag{15}$$

$$v = f(u). \tag{16}$$

The activation function f is again treated as a constant, and from the block diagram shown in Fig. 3, the whole system transfer function $h(s)$ from v to a can be expressed as

$$h(s) = \frac{1}{cs - wf + g} \cdot f(s) \tag{17}$$

4.3 Open-loop MHNN

While HN only has a linear weight in the feedback loop, MHNN has an exogenous input b and an extra nonlinear transformation f_2 in the feedback loop. When the feedback loop of MHNN is broken, MHNN degenerates to the open-loop HN.

4.4 Closed-loop MHNN

The analysis of MHNN is interesting. First, there are two sets of input-output relations will be available

$$c\frac{du_1}{dt} = a + w_1 v_2 - u_1 g \tag{18}$$

$$v_1 = f_1(u_1) \tag{19}$$

$$0 = b + w_2(v_1 - u_2) \tag{20}$$

$$v_2 = f_2(u_2). \tag{21}$$

Though this is a scalar system, it can have two inputs and two outputs. As done with HN, it is assumed f_1 and f_2 as constants weight.

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \frac{1}{cs - w_1 f_1 f_2 + g} \begin{bmatrix} f_1 & f_1 f_2 \frac{w_1}{w_2} \\ f_1 f_2 & \frac{cs+g}{w_2} f_2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \tag{22}$$

5 EQUIVALENCE OF THE ACTIVATION FUNCTION

Based on the information contained in Fig. 3(b) and 4(b), the following general nonlinear control block diagram as in Fig. 5 can be abstracted. In the figure, L_1, L_2 and L_3 are linear blocks and N is a nonlinear element which can represent an activation function in the ANN context. HN block digram falls into this category.

5.1 Linear Estimation of Nonlinearity

In this subsection, a "macro" approach is taken to approximate the given nonlinear element N by a linear time-invariant element L in the input-output sense.

The symbol Lr is used to indicate the output of system L when it is excited by input r [9]. A measure of how well the linear system L approximates the nonlinear system N is provided by the error criterion

$$E(L) = \lim_{T\to\infty} \frac{1}{T} \int_0^T [Nr(t) - Lr(t)]^2 dt \tag{23}$$

assuming the indicated limit exists. Thus the objective is to choose the linear system L in such a way that the error criterion $E(L)$ is minimized. The function L minimizes the error criterion E of (23), if and only if, their cross-correlation functions are equal, i.e.

$$\phi_{r,Lr}(\tau) = \phi_{r,Nr}(\tau), \quad \forall \tau > 0 \tag{24}$$

Consequently, the network nonlinearity can be replaced with a linear gain to produce similar responses of the nonlinearity and its approximation, in some sense, to the same sinusoidal input. In this vein it is observed that, if a nonlinearity $y(r, \dot{r})$ is excited by a sinusoidal input ($\psi = \omega t$), $r = A \sin \psi$, then the output can be expressed by a Fourier series and the corresponding nonlinear transfer function, denoted by $N(A, \omega)$, is

$$N(A, \omega) = \frac{A_1(A, \omega)}{A} e^{j\varphi_1(A,\omega)}. \tag{25}$$

5.2 Equivalence of Neural Hard-limit Nonlinearity

In the frequency domain, since only the input-output relations count, a replacement of nonlinear element with a linear element will not affect the rest of the system so that the overall system properties are preserved. In the following, it will be shown how this equivalence can be found for a typical ANN nonlinearity. A hard-limit describing function is selected for this purpose [9]. Let

$$f(x) = \begin{cases} mx, & 0 < \psi \le \psi_1 \\ m\delta, & \psi_1 < \psi \le \frac{\pi}{2} \end{cases} \tag{26}$$

L is found through equating $\phi_{r,y_l}(\tau)$ and $\phi_{r,y_n}(\tau)$. Since the characteristics of ANN nonlinearities are single-valued, memory-less, static and odd, using sinusoidal input in Eqn. (25), one can get

$$\begin{aligned} N(A,\omega) &= N(A) = \frac{4}{\pi A} \int_0^{\pi/2} y(A \sin \psi) \sin \psi d\psi. \quad (27) \\ &= \frac{2m}{\pi} \left[\sin^{-1} \frac{\delta}{A} + \frac{\delta}{A} \sqrt{1 - \left(\frac{\delta}{A}\right)^2} \right] \qquad A \ge \delta \\ &= m \qquad A < \delta \end{aligned}$$

As can be seen, the transfer function of a nonlinear element is also a function of the magnitude of the reference input. When the input signal is small, the nonlinear element acts like a linear element; but if the input signal is large, the gain of the nonlinear element decreases almost in proportion to the magnitude of input.

6 BODE DIAGRAMS OF NETWORKS

The objective of this section is to show how the theory in the last section can be applied to the analysis of a control loop with a nonlinear element and to reach conclusions about the robustness of recurrent ANN based on the stability margin concepts.

6.1 Frequency Response of Non-limit-cycling Nonlinear Systems

As shown in Eqn. (27), $N(A,\omega)$ is a function of both A and ω. Referring to Fig. 5, let $r(t)$ represent the simple harmonic excitation with a magnitude U If one assumes $x(t)$ to be sinusoid of amplitude A and frequency ω, the transfer function from $c(t)$ to $r(t)$ becomes

$$\frac{C}{R}(j\omega, A) = \frac{L_1(j\omega)N(A,\omega)L_2(j\omega)}{1 + L_1(j\omega)N(A,\omega)L_2(j\omega)L_3(j\omega)} \quad (28)$$

and consider A and ω as independent variables. By inspection, $|\frac{X}{R}|$ should be equal to $\frac{A}{U}$. So, this equation will yield the solution to A corresponding to some fixed ω. Let ω vary and the desired A's will be obtained. Substituting the A's into $|\frac{C}{R}|$ will yield a frequency response.

6.2 Bode Diagram of HN

Knowing the property of an individual element in the control loop, the system property can be obtained by putting them together. The block diagram for HN has $L_1(s) = \frac{1}{cs+g}$, $L_2(s) = 1$, $L_3(s) = w$, and $N(A,\omega) = N(A)$, is real. One can suppose that $U = 1$ for the worst situation because most of the signals in ANN are normalized. The other parameters are set at $w = 1$, $g = 10$, $c = 1$, $m = 1$, $\delta = 1$, for illustration.

By simple algebra, one can obtain

$$N(A) = \begin{cases} 1 & A \leq 1 \\ \frac{2}{\pi}\left[\sin^{-1}\frac{1}{A} + \frac{1}{A}\sqrt{1 - \left(\frac{1}{A}\right)^2}\right] & A > 1 \end{cases}$$

Therefore, the ratio of the output to the input is

$$\left|\frac{X}{R}(j\omega, A)\right| = \frac{A}{U} \quad (29)$$

The above equation can be solved to obtain A and be represented in a Bode diagrams as in Fig. 6.

From the Bode of HN, it is seen that HN is robust and has an infinite stability margin. Though HN contains a nonlinearity, it is beneficial. The nonlinearity delivers a smaller gain for the larger signal and a larger gain for the smaller signal so that the whole system does not become unstable. The maximum gain of nonlinearity is bounded and occurs around zero. For ANN nonlinearity, the tangent of the activation function around zero is very critical since the maximum gain of the system is near zero.

7 CONCLUSIONS

A new mutually recurrent ANN has been formulated. We have presented a frequency domain method to examine the ANN nonlinearity. This "micro" approach presented a fundamental concept about the functionality of nonlinear activation function in ANN. Form the approximate analytical results in this section, it is observed that the nonlinearity in an ANN is nothing but a "varying constant" changing with the magnitude of its input signal. The nature of this change is beneficial and favors the stability of the system.

REFERENCE

[1] P. K. Simpson. *Artificial Neural Systems.* Pergamon Press, Elmsford, NY, 1990.

[2] K. S. Narendra and K. Parthasarathy. Identification and Control of Dynamical Systems Using Neural Networks. *IEEE Transactions on Neural Networks*, 1(1):4-27, 1990.

[3] Jie Shen and S. N. Balakrishnan. A Class of Modified Hopfield Networks for Aircraft Identification and Control. In *Proceedings of AIAA Atmospheric Flight Mechanics Conference*, 1996.

[4] K. S. Narendra. *Adaptive Control using Neural Networks*, pages 115–142. In Miller et al. [5], 1990.

[5] W. Thomas Miller, III, Richard S. Sutton, and Paul J. Werbos, editors. *Neural Networks for Control.* The MIT Press, Cambridge, MA, 1990.

[6] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79:2554–2558, April 1982.

[7] Paolo Frasconi and Marco Gori. Computational Capabilities of Local-Feedback Recurrent Networks Acting as Finite-State Machines. *IEEE Transaction on Neural Networks*, 7(6):1521–1225, 1996.

[8] F. C. Chen and C. C. Liu A Daptively Controlling Non-Linear Continous-Time System Using Multilayer Neural Networks. *IEEE Transaction on Automatic Control*, 39(6):1306–1310, 1994

[9] Arthur Gelb and Wallace E. Van der Velde. *Multiple-input describing functions and nonlinear system design.* McGraw-Hill Book Company, New York, 1967.

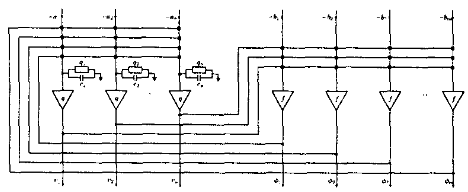


Figure 1: Modified Hopfield Neural Network

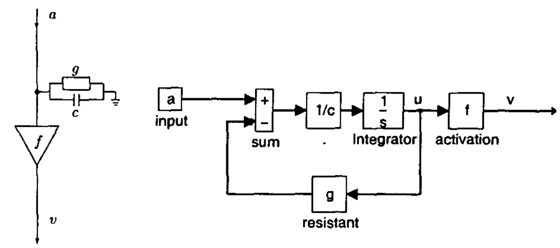
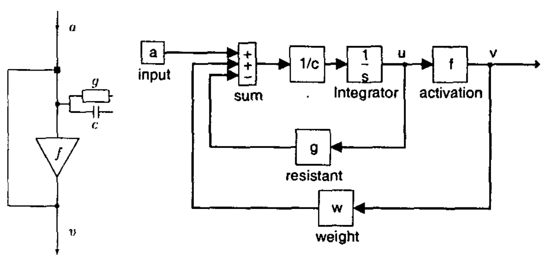
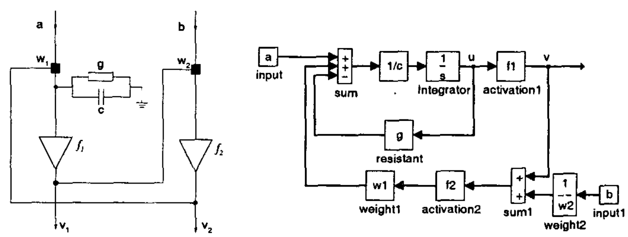


Figure 2: Open-loop Hopfield Network and Its Block Diagram



(a) (b)

Figure 3: Closed-loop Hopfield Network and Its Block Diagram



(a) (b)

Figure 4: Closed-loop MHNN and Its Block Diagram

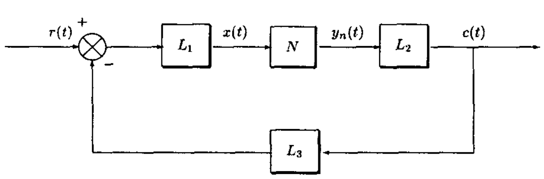


Figure 5: General Nonlinear Control Blocks

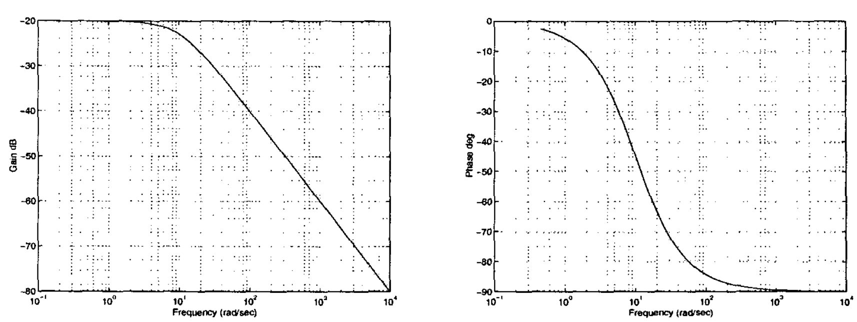


Figure 6: Nonlinear System Bode Diagram