



ON THE EFFECT OF NEURON ACTIVATION GAIN ON ROBUSTNESS OF COMPLETE STABILITY

M. DI MARCO, M. FORTI and P. NISTRI

*Dipartimento di Ingegneria dell'Informazione, Università di Siena,
v. Roma 56 – 53100 Siena, Italy*

A. TESI

*Dipartimento di Sistemi e Informatica, Università di Firenze,
v. S. Marta 3 – 50139 Firenze, Italy*

Received February 20, 2003; Revised March 17, 2003

The paper addresses robustness of complete stability with respect to perturbations of the interconnections of nominal symmetric neural networks. The influence of the maximum neuron activation gain on complete stability robustness is discussed for a class of third-order neural networks. It is shown that high values of the gain lead to an extremely small complete stability margin of all nominal symmetric neural networks, thus allowing to conclude that complete stability robustness cannot be, in general, guaranteed.

Keywords: Neural networks; complete stability; robustness; neuron activations.

1. Introduction

Hopfield neural networks (HNNs) and standard Cellular Neural Networks (CNNs) with a symmetric neuron interconnection matrix are known to enjoy the fundamental property of complete stability, i.e. each trajectory converges to some equilibrium point in the long run behavior [Hopfield, 1984; Chua & Yang, 1988]. In view of practical implementation it is also of key importance to address robustness of complete stability of nominal symmetric neural networks with respect to small perturbations of the neuron interconnections, since it is not possible to realize exactly symmetric interconnections [Wang & Michel, 1994; Vidyasagar, 1993].

Recent work has shown that complete stability of nominal symmetric neural networks may be not robust [Di Marco *et al.*, 2000; Di Marco *et al.*, 2002a; Di Marco *et al.*, 2002b]. In particular, for a class of third-order competitive CNNs it has been

shown that there is *some* nominal symmetric CNN which exhibits nonvanishing large-size oscillations even for arbitrarily small perturbations of the interconnections [Di Marco *et al.*, 2000].

A standard CNN such as that considered in [Di Marco *et al.*, 2000] possesses neuron activations with *unity* gain in the linear region [Chua & Yang, 1988]. In contrast, HNNs are characterized by neuron activations with a very *high* maximum gain, which are actually realized using a hard comparator. The goal of this paper is to address how high values of the activation gain influence the property of robustness of complete stability. To this end, the class of third-order neural networks introduced in [Di Marco *et al.*, 2000] is analyzed in the general case where the neuron activation gain can assume values larger than unity. It is shown that when the gain is increased, the complete stability margin of *all* nominal symmetric matrices decreases and tends to zero.

2. Main Results

Let us consider the inputless third-order neural network

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - \begin{pmatrix} 0 & \alpha & \beta \\ \beta & 0 & \alpha \\ \alpha & \beta & 0 \end{pmatrix} \times \begin{pmatrix} g_\lambda(x_1) \\ g_\lambda(x_2) \\ g_\lambda(x_3) \end{pmatrix} \quad (1)$$

where α and β are positive parameters, i.e. the neural network has negative (inhibitory) interconnections between distinct neurons and as such is said to be competitive. Note that the neuron interconnection matrix is symmetric if and only if $\alpha = \beta$.

The neuron activations are modeled by the piecewise linear function

$$g_\lambda(\rho) = \frac{1}{2} (|\lambda\rho + 1| - |\lambda\rho - 1|)$$

with a maximum gain $\lambda > 0$ in the linear region $|\rho| < 1/\lambda$, and ± 1 saturation levels.

When $\lambda = 1$, $g_1(\rho) = \frac{1}{2} (|\rho + 1| - |\rho - 1|)$ corresponds to the standard low gain CNN neuron activation introduced by Chua and Yang [1988]. When $\lambda \gg 1$, $g_\lambda(\rho)$ instead approaches a hard comparator function such as that typically employed to model the neuron activation of a HNN [Hopfield, 1984]. Note also that in accordance with the design procedure suggested by Hopfield [1984], the neuron self-connections in (1) are zero.

From standard results it follows that (1) is completely stable in the nominal symmetric case, i.e. $\alpha = \alpha_n$, $\beta = \beta_n$, and $\alpha_n = \beta_n$ [Chua & Yang, 1988]. Given some symmetric neural network defined by (α_n, β_n) with $\alpha_n = \beta_n$, our goal is to evaluate the l_2 margin of complete stability, which is defined as $\gamma(\alpha_n, \beta_n) = \sup\{\delta > 0\}$ such that (1) is completely stable for all (α, β) satisfying $\sqrt{(\alpha - \alpha_n)^2 + (\beta - \beta_n)^2} < \delta$. In particular, we are interested in studying how this margin is affected by the maximum neuron activation gain λ . To this end, we find it convenient to recall first the case $\lambda = 1$ in Sec. 2.1, which has been treated in detail in [Di Marco *et al.*, 2000]. Then, in Sec. 2.2 the case $\lambda \gg 1$ is investigated via a suitable coordinate transformation of (1) which allows us to reuse the results of the previous case.

2.1. The CNN case $\lambda = 1$

When $\lambda = 1$, (1) reduces to

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - \begin{pmatrix} 0 & \alpha & \beta \\ \beta & 0 & \alpha \\ \alpha & \beta & 0 \end{pmatrix} \times \begin{pmatrix} g_1(x_1) \\ g_1(x_2) \\ g_1(x_3) \end{pmatrix} \quad (2)$$

where $g_1(\rho) = \frac{1}{2} (|\rho + 1| - |\rho - 1|)$.

Let us consider the following curves in the parameter space (α, β)

$$C_{\text{sym}} = \{(\alpha, \beta) : \alpha - \beta = 0, \alpha \geq 0, \beta \geq 0\}$$

$$C_{\text{hop}} = \{(\alpha, \beta) : \alpha + \beta = 2, \alpha \geq 0, \beta \geq 0\}$$

$$C_{\text{het}} = \{(\alpha, \beta) : (\alpha\beta + \alpha - \beta^2 - 1) \times (\alpha\beta + \beta - \alpha^2 - 1) = 0, \alpha \geq 1, \beta \geq 1\}$$

which delimit the three open regions R_{gas} , R_{cs} and R_{osc} shown in Fig. 1. It is known that (2) is completely stable in the symmetric case $(\alpha, \beta) \in C_{\text{sym}}$. The following additional results have been proved in the nonsymmetric case $(\alpha, \beta) \notin C_{\text{sym}}$ [Di Marco *et al.*, 2000]. For $(\alpha, \beta) \in R_{\text{gas}}$, the origin is the unique equilibrium point of (2) and it is globally asymptotically stable. When $(\alpha, \beta) \in R_{\text{cs}}$, (2) possesses 13 hyperbolic equilibrium points, 6 of which are asymptotically stable. Moreover, within region R_{cs} , (2) is completely stable. For $(\alpha, \beta) \in R_{\text{osc}}$ the origin is the unique equilibrium point of (2), but now it turns out to be unstable. In region R_{osc} almost all trajectories of (2) display a large-size non-vanishing oscillation in the long run behavior. On C_{hop} a Hopf-like bifurcation occurs which causes the loss of complete stability when passing from R_{gas} to R_{osc} . Similarly, on C_{het} a heteroclinic bifurcation is present which again causes the loss of complete stability when passing from R_{cs} to R_{osc} . For the discussion that follows it is important to note that curve C_{het} tends to the asymptotes $\beta = \alpha \pm 1$, which are shown dashed in Fig. 1, as $\alpha \rightarrow \infty$. In particular, region R_{cs} is thoroughly confined within these asymptotes.

Given some nominal symmetric neural network (2) defined by $(\alpha_n, \beta_n) \in C_{\text{sym}}$, let us now consider its complete stability margin $\gamma(\alpha_n, \beta_n)$. Geometrically, for $\alpha_n = \beta_n \geq 1$, $\gamma(\alpha_n, \beta_n)$ corresponds to the minimum Euclidean distance between (α_n, β_n) and curve C_{het} , and it always results in $\gamma(\alpha_n, \beta_n) < 1$.

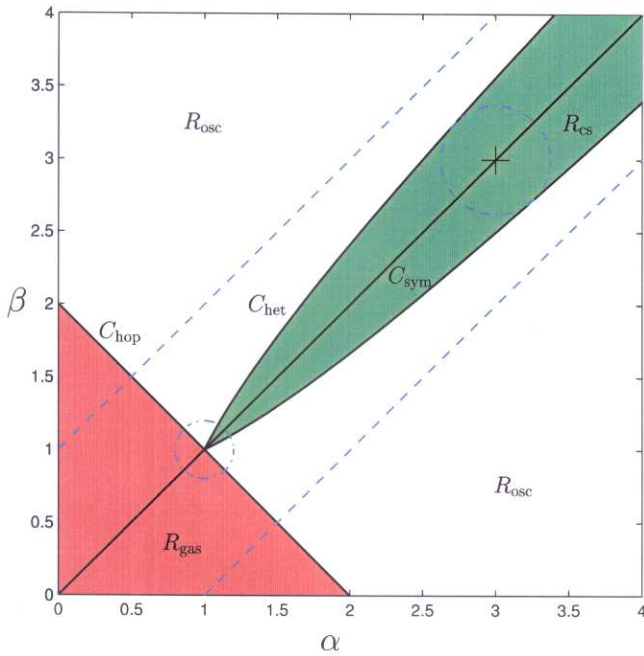


Fig. 1. Bifurcation diagram of neural network (2) ($\lambda = 1$).

One key observation is that complete stability is not robust for the CNN (2) described by the symmetric interconnection matrix obtained for $(\alpha_n, \beta_n) = (1, 1)$. Indeed from Fig. 1 it is seen that any (small) neighborhood of $(\alpha_n, \beta_n) = (1, 1)$ certainly intersects the oscillatory region R_{osc} , hence $\gamma(1, 1) = 0$.

It can also be observed that symmetric matrices sufficiently far from $(\alpha_n, \beta_n) = (1, 1)$ have instead some margin of complete stability. For example, it results in $\gamma(3, 3) \simeq 0.375$.

2.2. The HNN case $\lambda \gg 1$

Consider the change of variables $y_i = \lambda x_i, i = 1, 2, 3$, which transforms (1) into

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{pmatrix} = - \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} - \begin{pmatrix} 0 & \lambda\alpha & \lambda\beta \\ \lambda\beta & 0 & \lambda\alpha \\ \lambda\alpha & \lambda\alpha & 0 \end{pmatrix} \times \begin{pmatrix} g_\lambda \left(\frac{1}{\lambda} y_1 \right) \\ g_\lambda \left(\frac{1}{\lambda} y_2 \right) \\ g_\lambda \left(\frac{1}{\lambda} y_3 \right) \end{pmatrix}. \tag{3}$$

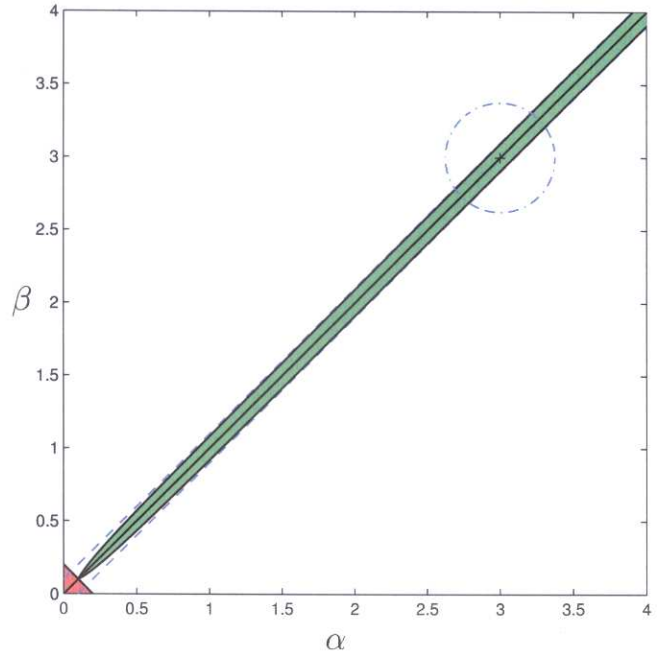


Fig. 2. Bifurcation diagram of neural network (3) ($\lambda = 10$).

Moreover, it is seen that it results in

$$g_\lambda \left(\frac{1}{\lambda} \rho \right) = \frac{1}{2} (|\rho + 1| - |\rho - 1|) = g_1(\rho) \tag{4}$$

where $g_1(\rho)$ is the nonlinearity with gain $\lambda = 1$ which characterizes the standard low gain CNN (2).

Comparison of (3) and (2), taking into account (4), allows one to conclude that the dynamical behavior of (3) coincides with that of (2), provided parameters α and β are re-scaled, i.e. α and β are replaced in (2) with $\lambda\alpha$ and $\lambda\beta$, respectively. In particular, the bifurcation diagram for (3) can be immediately obtained from that of (2) in Fig. 1 by means of the same rescaling of the parameters. As an example, Fig. 2 illustrates the situation in the case $\lambda = 10$.

The above considerations make it clear that the overall effect of increasing the gain λ is to reduce the region of global asymptotic stability, and the region of complete stability, while the oscillatory region tends to cover the whole parameter space. In particular, the region of complete stability is confined within the asymptotes $\beta = \alpha \pm 1/\lambda$, represented by the dashed lines in Fig. 2. As a consequence, the complete stability margin is such that $\gamma(\alpha_n, \beta_n) < 1/\lambda$ for all symmetric matrices $(\alpha_n, \beta_n) \in C_{sym}$ with $\alpha_n = \beta_n > 1$, and therefore no symmetric HNN (3) has an acceptable margin

of complete stability for large values of λ . In particular, in high-gain limit $\lambda \gg 1$, which is at the basis of the design procedure suggested by Hopfield [1984], it turns out that each nominal symmetric HNN (3) actually has an almost zero complete stability margin, so that its practical implementation cannot guarantee a completely stable behavior.

As an example, let us consider the nominal symmetric HNN (3) with $\lambda = 1000$ and $(\alpha_n, \beta_n) =$

$(3, 3)$. The HNN has 13 hyperbolic equilibrium points, six of which are asymptotically stable. Figure 3 shows the state space trajectory of this neural network starting at $(7, 3, 10)'$ (prime means transpose), which converges to the asymptotically stable equilibrium point $(0, -3, 3)'$. The complete stability margin $\gamma(3, 3)$ of the nominal HNN is less than 10^{-3} . Then, let us consider the perturbed non-symmetric HNN (3) defined by $\alpha = 3.002$ and

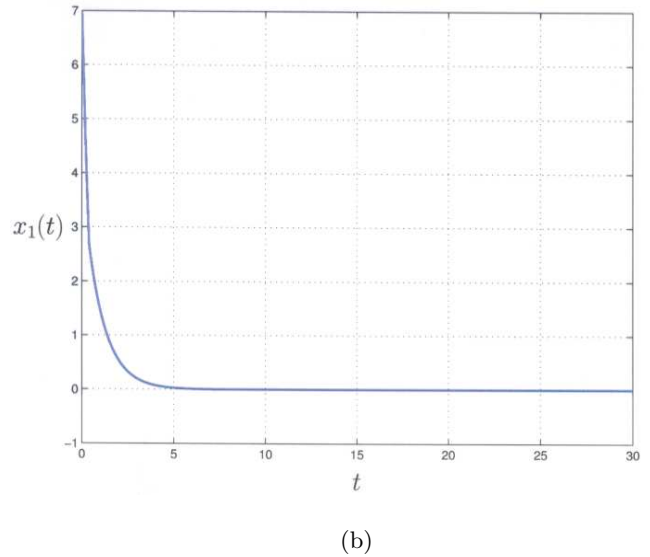
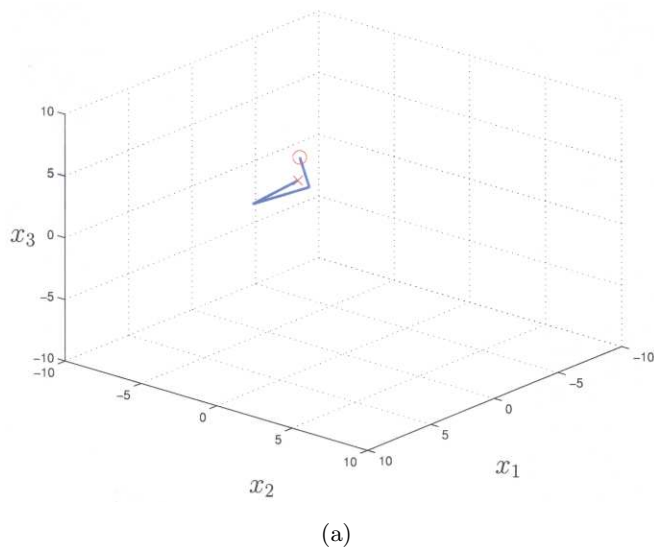


Fig. 3. (a) State-space trajectory for a symmetric neural network (3) with a high-gain activation function ($\alpha = 3, \beta = 3, \lambda = 1000$); \circ initial condition, \times stable equilibrium point. (b) Corresponding time evolution of $x_1(t)$.

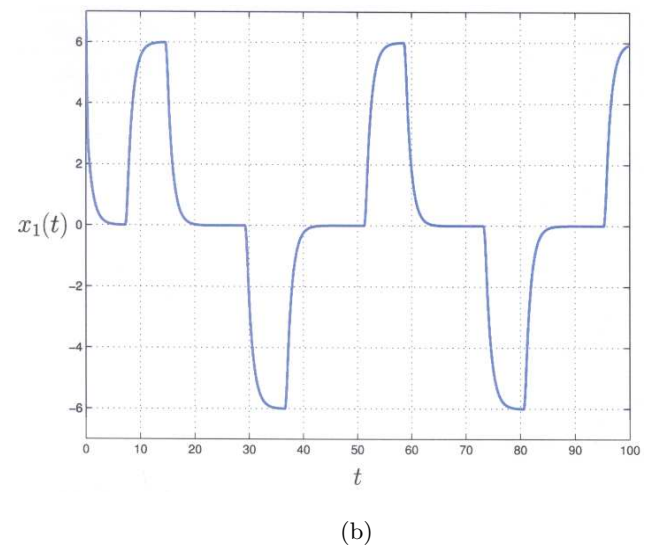
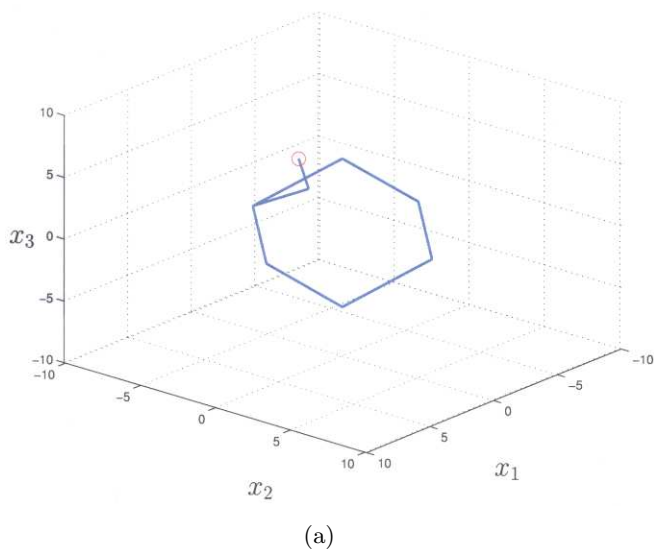


Fig. 4. (a) State-space trajectory for an almost symmetric neural network (3) with a high-gain activation function ($\alpha = 3.002, \beta = 2.998, \lambda = 1000$); \circ initial condition. (b) Corresponding time evolution of $x_1(t)$.

$\beta = 2.998$, and the same value of λ , which turns out to fall into the oscillatory region of Fig. 2. It is seen from Fig. 4 that the trajectory starting at the same initial condition now displays a large-size non-vanishing oscillation in the long run behavior. This is true also for almost all other trajectories, as it has been verified by simulations not reported here for brevity. Actually, the considered small perturbations, when combined with the high value of λ , are seen to quickly destabilize the configuration of equilibrium points of the nominal symmetric HNN, leading to a perturbed HNN that possesses a unique unstable equilibrium point at the origin.

3. Conclusion

The paper has addressed how the maximum gain of the neuron activations affects the property of robustness of complete stability with respect to small perturbations of nominal symmetric interconnection matrices. For a class of third-order neural networks it has been proved that high values of the neuron gain have a really negative effect on robustness of complete stability, since they lead to a situation where all nominal symmetric neural networks have an extremely small complete stability margin. In this respect, low gain neurons as those employed in the standard CNNs are preferable to high-gain neurons as those used in the HNNs. One main conclusion is that in the typical limiting case the HNN neurons are implemented via hard com-

parators, hence they have a very high gain, there is the risk that the stability margin be actually reduced to zero for all symmetric matrices, so that complete stability can be destroyed by arbitrarily small perturbations.

References

- Chua, L. O. & Yang, L. [1988] "Cellular neural networks: Theory," *IEEE Trans. Circuits Syst.* **35**, 1257–1272.
- Di Marco, M., Forti, M. & Tesi, A. [2000] "Bifurcations and oscillatory behavior in a class of competitive cellular neural networks," *Int. J. Bifurcation and Chaos* **10**, 1267–1293.
- Di Marco, M., Forti, M. & Tesi, A. [2002a] "Complex dynamics in nearly-symmetric three-cell cellular neural networks," *Int. J. Bifurcation and Chaos* **12**, 1357–1362.
- Di Marco, M., Forti, M. & Tesi, A. [2002b] "Existence and characterization of limit cycles in nearly symmetric neural networks," *IEEE Trans. Circuits Syst.-I* **49**, 979–992.
- Hopfield, J. J. [1984] "Neurons with graded response have collective computational properties like those of two-state neurons," *Proc. Nat. Acad. Sci.* **81**, 3088–3092.
- Vidyasagar, M. [1993] "Location and stability of the high-gain equilibria of nonlinear neural networks," *IEEE Trans. Neural Networks* **4**, 660–672.
- Wang, K. & Michel, A. N. [1994] "Robustness and perturbation analysis of a class of nonlinear systems with applications to neural networks," *IEEE Trans. Circuits Syst.-I* **41**, 24–32.