Stability, Robustness, and Approximation Properties of Gradient Recurrent High Order Neural Networks

E.B. Kosmatopoulos[†] and M.A. Christodoulou[†]

[†] Dept. of Electronic & Computer Engineering
Technical University of Crete
73100 Chania, Crete, GREECE

Extended Abstract

In this paper we examine recurrent high order neural network (RHONN) models whose dynamics are governed by a differential equation of the form

$$\dot{x}_i = -b_i(x_i, \gamma_i) - \sum_{k=1}^{L} w_{ik} \frac{d_i(k)}{S_i} \prod_{j \in I_k} S_j^{d_j(k)} = \mathcal{N}_i(x) \tag{1}$$

where $\{I_1, I_2, \ldots, I_L\}$ is a collection of L not-ordered subsets of $\{1, 2, \ldots, n\}$, w_{ik} are the (adjustable) synaptic weights of the neural network, and $d_j(k)$ are non-negative integers. The function $b_i(\cdot)$ is a smooth function of x_i that depends on a vector of parameters $\gamma_i \in \Re^{q_i}$, where q_i are positive integers. The function $b_i(\cdot)$ satisfies the follwoing assumptions.

(S.1) $\int_0^{x_i} b_i(\eta, \gamma_i) S'(\eta) d\eta \to +\infty$ as $|x_i| \to \infty$.

(S.2) $|b_i(x_i, \gamma_i)| \to \infty$ and $|b_i(x_i, \gamma_i) S'(x_i)| \to \infty$ as $|x_i| \to \infty$.

(S.3) For any positive constants a_i, d_i, there exist γ_i such $\frac{1}{2} b_i^2(x_i, \gamma_i) > |b_i'(x_i, \gamma_i)| a_i + |b_i(x_i, \gamma_i)| d_i$, where $b_i'(\cdot) = \frac{\partial b_i}{\partial x_i}(\cdot)$.

(S.4) There exists a positive number c_1 such that $c_1 \sum_{i=1}^{n} \int_0^{x_i} b_i(\eta, \gamma_i) S'(\eta) d\eta \geq \frac{1}{2} \sum_{i=1}^{n} S_i' b_i^2(x_i, \gamma_i)$.

(S.5) For any x_i and z_i, $sgn(x_i - z_i) = sgn(b_i(x_i,\gamma_i) - b_i(z_i,\gamma_i))$ and there exist positive constants a_i such that

$$|b_i(x_i,\gamma_i) - b_i(z_i,\gamma_i)| \geq a_i|x_i - z_i|$$

The synaptic weights are restricted to satisfy the following condition

$$w_{ik} = w_{jk},\ \ \forall i,j \in \{1,\ldots,n\} \tag{H1}$$

Such RHONNs are gradient dynamical systems and hence their stability properties can be easily detected. In order to distinguish these networks from the other RHONNs we will call them symmetric RHONNs or simply *s-RHONNs*. Also, in this chapter we will concentrate our attention in the case where the powers $d_j(k)$ are selected in such a way that the following relation always holds

$$d_i(k) = d_j(k) \geq 2,\ \forall i,j \in \{1,\ldots,n\}, \forall k \in \{1,\ldots,L\} \tag{H2}$$

The main results of the papaer are as follows.

Proposition 1 *Let the s-RHONN dynamics be perturbed as follows*

$$\dot{x}_i = -b_i(x_i,\gamma_i) - \sum_{k=1}^{L} w_{ik}\frac{d_i(k)}{S_i}\prod_{j\in I_k} S_j^{d_j(k)} + \xi_i(t) = \mathcal{N}_i(x) + \xi_i(t) \tag{2}$$

where $\xi \in \Re^n$ is a bounded signal. Let $\bar{\xi} = \sup_t |\xi(t)|^2$. Then, if $b_i(\cdot)$ satisfies assumptions (S.1), (S.2), *then*

(a) *The solution of the perturbed s-RHONN* (2) *are bounded for each t, and moreover the term $\mathcal{D}(x)$ converges to $\mathcal{B}(\bar{\xi})$, where*

$$\mathcal{D}(x) = \sum_{i=1}^{n} S_i' \left[b_i(x_i,\gamma_i) + \sum_{k=1}^{L} w_{ik}\frac{d_i(k)}{S_i}\prod_{j\in I_k} S_j^{d_j(k)} \right]^2$$

.

(b) *There exists a positive constant λ such that at each t*

$$\int_{t_0}^{t} \frac{1}{2}\mathcal{D}(\tau)d\tau \leq \lambda + \frac{1}{2}\int_{t_0}^{t} |\xi(\tau)|^2 d\tau$$

(c) *In the special case where $\xi \in \mathcal{L}_2$, the solutions of the s-RHONN converge to an equilibrium point asymptotically.*

□

We also investigate the following perturbed version of the s-RHONN (1):

$$dx_i = -b_i(x_i, \gamma_i)dt - \sum_{k=1}^{L} w_{ik} \frac{d_i(k)}{S_i} \prod_{j \in I_k} S_j^{d_j(k)} dt + \sum_{j=1}^{m} f_{ij}(x) d\zeta_j \tag{3}$$

where $f_{ij}(\cdot)$ are bounded and smooth functions, and ζ is a standard m-dimensional Wiener process with $E\{\dot{\zeta}(t)\} = 0$ and $E\{\dot{\zeta}(t)\dot{\zeta}(\tau)\} = \sqrt{2}I\delta(t-\tau)$, where I is the $L \times L$ identity matrix. As explained in the Appendix A, the stochastic differential equation (3) can be written into a more familiar to engineers form

$$\dot{x}_i = -b_i(x_i, \gamma_i) - \sum_{k=1}^{L} w_{ik} \frac{d_i(k)}{S_i} \prod_{j \in I_k} S_j^{d_j(k)} + \sum_{j=1}^{m} f_{ij}(x)\dot{\zeta}_j \tag{4}$$

where $\dot{\zeta}$ is an m-dimensional white noise process. The next proposition establishes stability of the perturbed s-RHONN (3)

Proposition 2 *Consider the perturbed s-RHONN* (3). *Assume that* $f_{ij}(\cdot)$ *are bounded functions, and that* $b_i(\cdot)$ *satisfy assumptions* (S.1)-(S.4). *Then, there exist positive numbers* c_i, $i = 1, 2$ *such that*

$$\sup_{0 \le t < \infty} \Pr\{\mathcal{V}(t) \ge C\} < \frac{\mathcal{V}(0) + \frac{c_2}{c_1}}{C}$$

□

Consider now the problem of approximating a general nonlinear dynamical system by a s-RHONN. The behavior of the system to be approximated is described by

$$\dot{\chi} = F(\chi) \tag{5}$$

where $\chi \in \mathcal{R}^n$ is the state of the system and $F : \mathcal{R}^n \to \mathcal{R}^n$ is a smooth vector field defined on a compact set $\mathcal{Y} \subset \mathcal{R}^n$.

The approximation problem consists of determining whether by allowing enough high order connections, there exist weights w_{ik} such that the s-RHONN model approximates the behavior of an arbitrary dynamical system of the form (5).

In order to have a well-posed problem, we assume that F is continuous and satisfies a local Lipschitz condition such that (5) has a unique solution and $\chi(t) \in \mathcal{Y}$ for all t in some time interval $J_T := \{t : 0 \le t \le T\}$, where $\mathcal{Y}$ is a compact subset of $\mathcal{R}^n$. The interval J_T represents the time period over which the approximation is to performed. Based on the above assumptions we obtain the following result.

Theorem 1 *Assume that the s-RHONN satisfies assumption* (H2). *Then*

(a) *For any $\varepsilon > 0$ there exists an integer L and a set of weights w_{ik}^* satisfying the symmetry property* (H1), *such that for any $\chi \in \mathcal{Y}$ the neural network vector field with L high-order connections and weight values $w_{ik} = w_{ik}^*$ satisfies*

$$\sup_{\chi \in \mathcal{Y}} |\mathcal{N}(\chi) - F(\chi)| < \varepsilon \tag{6}$$

(b) *Suppose that the system* (5) *and the s-RHONN* (1) *are initially at the same state $x(0) = \chi(0)$. Then for any $\varepsilon > 0$ and any finite $T > 0$, there exists an integer L and a set of weights w_{ik}^* satisfying the symmetry property* (H1), *such that the state $x(t)$ of the s-RHONN model* (1) *with L high-order connections and weight values $w_{ik} = w_{ik}^*$ satisfies*

$$\sup_{0 \le t \le T} |x(t) - \chi(t)| \le \varepsilon \tag{7}$$

□