Convergence and Stability of the FSR CNN Model

S. Espejo, A. Rodríguez-Vázquez, R. Domínguez-Castro, and R. Carmona
Centro Nacional de Microelectrónica-Universidad de Sevilla
Edificio CICA, C/Tarfia s/n, 41012-Sevilla, SPAIN
Phone: +34 - 5 - 423 99 23. Fax: +34 - 5 - 462 45 06. E-mail: espejo@cnm.us.es

***Abstract** - Stability and convergency results are reported for a modified continuous-time CNN model. The signal range of the state variables is equal to the unitary interval, independently of the application. Stability and convergency properties are similar to those of the original model and, for given templates and offset coefficients, the results are generally identical. In addition, robustness and area-efficiency of VLSI implementations are significantly advantageous.*

1. Introduction

From an implementation point of view, the signal-range of the state variables of CNNs is an important factor. In the Chua-Yang model [1],[2], this value (normalized to the signal range of the output variables) is usually in the range of 4 to 20, depending on the application. The difference between the two signal ranges results in increased error levels in CNN circuits. Furthermore, the dependency of the signal range of the state variables on the template coefficients represents an additional problem for the implementation of programmable CNN systems.

A modified CNN model which exhibits identical signal ranges for the state and output variables, independently of the template coefficients, was proposed by the authors in [3]. This contributions demonstrates some stability and convergency properties of this model. Among them, the stability results reported by Chua and Yang in their first CNN proposal [1],[2] are extended here to the modified model, from now on called the Full Signal Range (FSR) model.

For the purpose of mathematical demonstrations, we consider a parameterized family of CNN models, referred to as the Improved Signal Range (ISR) family, which includes the Chua-Yang and FSR models as particular cases.

2. The ISR model family

The dynamic behavior of an ISR cell c is described by the following differential equation, in which each symbol is used with its usual meaning (see Annex at the end of the paper),

$$\frac{dx^c}{dt} = -g(x^c) + \sum_{d \in N(c)} a_d^c y^d + \sum_{d \in N(c)} b_d^c u^d + i^c \quad (1)$$

and where

$$y^c = f(x^c) = \begin{cases} 1, \forall x^c \geq 1 \\ x^c, \forall |x^c| < 1 \\ -1, \forall x^c \leq -1 \end{cases} \quad \text{and} \quad g(x^c) = \begin{cases} m(x^c - 1) + 1 & \forall x^c > 1 \\ x^c & \forall |x^c| \leq 1 \\ m(x^c + 1) - 1 & \forall x^c < -1 \end{cases} \quad (2)$$

Note that the output and the state variables of each cell are related by the usual nonlinear relationship, while the loss term is modified introducing a nonlinear function $g(x^c)$ parameterized by $m \in [1, \infty)$. This function is depicted in Fig. 1. Note that if $m = 1$, then $g(x^c) = x^c$, thus resulting in the original CNN model. On the other hand, the FSR model is obtained when $m \to \infty$. Although, as compared to the original model, the implementation of an ISR cell according to Eq. 1, and the definitions in Eq. 2 requires a circuitry of increased complexity, the situation becomes is inverted when the FSR model is considered. In addition, the advantages of the ISR model become optimal when $m \to \infty$.

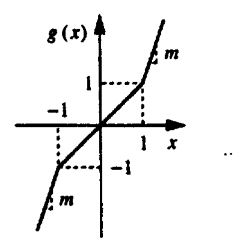

Figure 1 - Piece-wise linear loss term used in the ISR models.

3. Stability and convergency of the ISR model family

This section contains a sequence of statements and proofs. We make here two definitions,

$$k^c = \sum_{d \in N(c)} b^c_d u^d + d^c \qquad \text{and} \qquad M^c = \sum_{d \in N(c)} |a^c_d| + |k^c| \tag{3}$$

For simplicity, some statements below will be demonstrated assuming that $M^c \geq 1$, which is the most common case (for instance whenever $a^c_c \geq 1$). This assumption is not necessary. General demonstrations are available from the authors upon request [4].

Statement 1: "The absolute value of the state variable of every cell c in an ISR CNN is bounded for any set of finite feedback, control and offset coefficients (including nonuniform CNNs), and for any set of finite initial conditions, input to the cells and boundary conditions."

Proof: Using the definition of $g(x^c)$ into Eq. 1 we obtain a differential equation from which the following expressions can be obtained by traditional methods,

$$x^c(t) = x^c(0)e^{-mt} + \int_0^t \left[m - 1 + \sum_{d \in N(c)} a^c_d y^d(\tau) + k^c \right] e^{-m(t-\tau)} d\tau \qquad ; \forall x^c > 1 \tag{4}$$

$$x^c(t) = x^c(0)e^{-t} + \int_0^t \left[\sum_{d \in N(c)} a^c_d y^d(\tau) + k^c \right] e^{-(t-\tau)} d\tau \qquad ; \forall |x^c| \leq 1 \tag{5}$$

$$x^c(t) = x^c(0)e^{-mt} + \int_0^t \left[1 - m + \sum_{d \in N(c)} a^c_d y^d(\tau) + k^c \right] e^{-m(t-\tau)} d\tau \qquad ; \forall x^c < -1 \tag{6}$$

which correspond, respectively, to the upper-saturated region $(x^c > 1)$, the linear region $(|x^c| > 1)$, and the lower-saturated region $(x^c < -1)$. For these equations to be valid, the time variable must be shifted so as to make $t = 0$ at the instant in which the state variable entered a particular region, and the initial condition $x^c(0)$ must be defined accordingly. Since the state variable will always be in one of these regions, we only need to demonstrate that the above three equations are bounded. Using Schwartz's inequality several times on Eq. 4, taking into account that the output variables are $|y^c(t)| \leq 1$, and that $m \geq 1$, the following result is obtained

$$|x^c(t)| \leq \frac{m - 1 + M^c}{m} + \left[|x^c(0)| - \frac{m - 1 + M^c}{m} \right] e^{-mt} \tag{7}$$

which is bounded. The same result can be obtained from Eq. 6. Finally, within the linear region (Eq. 5) we have by definition that $|x^c(t)| \leq 1$. Thus, state variables are globally bounded.

Statement 2: "For any cell c, the maximum possible absolute value of an equilibrium point is an absolute bound for the state variable at any time instant, assuming that it is true at $t = 0$."

Proof: In order to determine the maximum possible absolute value of an equilibrium point, we must look at the dynamic route of ISR cells, obtained from Eq. 1,

$$\frac{dx^c}{dt} = \begin{cases} -m(x^c-1)-1+a_c^c y^c + i^c & ;\forall x^c > 1 \\ -x^c + a_c^c y^c + i^c & ;\forall |x^c| \le 1 \\ -m(x^c+1)+1+a_c^c y^c + i^c & ;\forall x^c < -1 \end{cases} \tag{8}$$

where i^c contains all the contributions on the right-hand side of Eq. 1 which do not directly depend on the state variable x^c,

$$i^c = \sum_{d \in N'(c)} a_d^c y^d + k^c \tag{9}$$

Note that the summation extends to the *reduced neighborhood* $N'(c) = N(c) - \{c\}$ of the cell. The value of the equilibrium points in each of the three regions in Eq. 8 can be obtained making the derivative equal to cero and using the definition of $y(x)$ given in Eq. 2,

$$x'^c_1 = \frac{m-1+a_c^c+i^c}{m} > 1; \quad x'^c_2 = \frac{-i^c}{a_c^c-1}, |x'^c_2| \le 1; \quad x'^c_3 = -\frac{m-1+a_c^c-i^c}{m} < -1 \tag{10}$$

where we include the condition required for the equilibrium points to be "real", otherwise being "virtual". Using $|y^c(t)| \le 1$ in Eq. 9 we obtain the maximum possible absolute value of i^c

$$|i^c| \le \sum_{d \in N'(c)} |a_d^c| + |k^c| \tag{11}$$

which used in Eq. 10 gives the maximum possible absolute value of a cell equilibrium point

$$|x'^c| \le \frac{m-1+\sum_{d \in N(c)} |a_d^c| + |k^c|}{m} = \frac{m-1+M^c}{m} \tag{12}$$

In order to calculate a global bound for the state variable, we must take into account that it can go in and out of the linear and saturated regions several times, defining intervals $[t_{Li}, t'_{Li})$, $i=1,\ldots,N$ and $[t_{Sj}, t'_{Sj})$, $j=1,\ldots,M$ within which the state variable is in the linear region or in some saturated region, respectively. The union of these intervals contains the complete transient evolution of the cell. We consider different initial conditions $x^c_{Li}(0) = x^c(t=t_{Li})$, $i=1,\ldots,N$, and $x^c_{Sj}(0) = x^c(t=t_{Sj})$, $j=1,\ldots,M$, and the corresponding time-shifts. This allows the use of one of Eq. 4 through Eq. 6 for each interval. Note that in general $|x^c_{Li}(0)| = |x^c_{Sj}(0)| = 1$ with the possible exception of the real (unshifted time) initial condition $x^c_0(0)$, which affects either to $x^c_{L1}(0)$ or to $x^c_{S1}(0)$. If $|x^c_0(0)| < 1$, then $|x^c_{L1}(0)| = |x^c_0(0)| < 1$, and if $|x^c_0(0)| > 1$ then $|x^c_{S1}(0)| = |x^c_0(0)| > 1$. In any case, $|x^c_{Li}(0)| \le 1$, $\forall\, i=1,2,\ldots,N$, and $|x^c_{Sj}(0)| \ge 1$, $\forall\, j=1,2,\ldots,M$.

From Eq. 7, which is valid for either saturated region, we can obtain a bound for the state variable in each interval $[t_{Sj}, t'_{Sj})$. We define function $B^c_{Sj}(t)$ as the right hand side of Eq. 7, defined within the interval, where it can be seen to be bounded by one of the following limits,

$$\text{if} \quad |x^c_{Sj}(0)| \le \frac{m-1+M^c}{m} \quad \text{then} \quad |x^c(t)| \le B^c_{Sj}(\infty) = \frac{m-1+M^c}{m} \tag{13}$$

$$\text{if} \quad |x^c_{Sj}(0)| > \frac{m-1+M^c}{m} \quad \text{then} \quad |x^c(t)| \le B^c_{Sj}(0) = |x^c_{Sj}(0)| \tag{14}$$

From where we can conclude that

$$|x^c(t)| \le \max\left\{\frac{m-1+M^c}{m}, |x^c_{Sj}(0)|\right\} \qquad \forall t \in [t_{Sj}, t'_{Sj}),\ j=1,\ldots,M \tag{15}$$

Taking into account the possible values of $|x^c{}_{Si}(0)|$, the following result is obtained

$$|x^c(t)| \le \max\left\{\frac{m-1+M^c}{m}, |x_0^c(0)|\right\} > 1 \qquad \forall t \in \bigcup_{j=1}^{M} [t_{Sj}, t'_{Sj}) \tag{16}$$

The same procedure can be used for the linear region, or we can simply take into account that $|x^c(t)| \le 1$, $\forall t \in [t_{Li}, t'_{Li})$ and for any $i = 1, \ldots, N$, which allows the extension of Eq. 16 to

$$|x^c(t)| \le \max\left\{\frac{m-1+M^c}{m}, |x_0^c(0)|\right\} \qquad \forall t \ge 0 \tag{17}$$

The right hand side of this equation coincides with that of Eq. 12 under the assumption made in the statement.

Statement 3: "There is a hyper-rectangular region around the origin of the state space that is a global attracting region, this is, if the state vector of the network is outside, it evolves towards the region, and if it is inside, it never scapes."

Proof: This hyper-rectangular region is defined by a set of intervals, one for each cell,

$$I^c = \left[-\frac{m-1+M^c}{m}, \frac{m-1+M^c}{m}\right] \qquad \forall c \in GD \tag{18}$$

From Eq. 13, if the state variable x^c of some cell belongs to I^c at some time instant (which can be redefined as $t = 0$), then it will belong to I^c at any later instant. On the other hand, from Eq. 7 we have that,

$$\lim_{t \to \infty} |x^c(t)| \le \frac{m-1+M^c}{m} \tag{19}$$

which shows that if the state variable of some cell is outside I^c at some time instant, the state variable will be either in I^c or arbitrarily close to its borders after a sufficient amount of time.

Statement 4: "The global attracting region approaches the $[-1,1]^n$ hypercube as $m \to \infty$ (FSR limit), regardless the particular template coefficients."

Proof: Simply note that

$$\lim_{m \to \infty} \frac{m-1+M^c}{m} = 1 \qquad \Rightarrow \qquad \lim_{m \to \infty} I^c = [-1, 1] \tag{20}$$

Statement 5: "ISR CNNs with central-feedback coefficients larger than unity $(a_c^c > 1)$ have no stable equilibrium points within the open hypercube $(-1,1)^n$, this is, if it converges, its output is binary."

Proof: The slope of the dynamic route in the linear region is given by $s = a_c^c - 1$, which is positive in the cases considered. Hence, any equilibrium point in the linear region is unstable.

Statement 6: "ISR CNNs with reciprocal feedback templates are convergent."

Proof: The demonstration is identical to that employed by Chua and Yang in [1]. A CNN is said to be reciprocal if

$$a_d^c = a_c^d \qquad , \forall c \in GD, \forall d \in GD \tag{21}$$

We will use the same function that was proposed for the original CNN model:

$$E(t) = \sum_{c \in GD} \left\{-\frac{1}{2} \sum_{d \in GD} a_d^c y^c(t) y^d(t) + \frac{1}{2}[y^c(t)]^2 - \sum_{d \in GD} b_d^c y^c(t) u^d - y^c(t) d^c\right\} \tag{22}$$

Since the output variables are restricted to the [-1,1] interval, this function is bounded for any set of finite CNN coefficients and inputs values. We will now demonstrate that this function is a continuously decreasing function of time. For this purpose, note that

$$\frac{dy^c}{dt} = \frac{dy^c}{dx^c}\frac{dx^c}{dt} \tag{23}$$

Then, taking the time derivative of Eq. 22, and after changing variables (c by d and vice-versa) in one of the resulting sums, and using Eq. 21, we obtain

$$\frac{dE(t)}{dt} = -\sum_{c \in GD} \{ \sum_{d \in GD} [a_d^c y^d + b_d^c u^d] + d^c - y^c \} \frac{dy^c}{dx^c} \frac{dx^c}{dt} \tag{24}$$

which, taking into account the following three facts

$$a_d^c = b_d^c = 0, \forall d \notin N(c); \qquad \frac{dy^c}{dx^c} = \begin{Bmatrix} 1, \forall |x^c| \le 1 \\ 0, \forall |x^c| > 1 \end{Bmatrix}; \qquad y^c = x^c, \forall |x^c| \le 1 \tag{25}$$

can be transformed into

$$\frac{dE(t)}{dt} = -\sum_{\substack{c \in GD \\ |x^c| \le 1}} \{ -x^c + \sum_{d \in N(c)} [a_d^c y^d + b_d^c u^d] + d^c \} \frac{dx^c}{dt} \tag{26}$$

Since only the cells whose state variables are in the linear region are included in the sum, we can use Eq. 1 with $g(x^c) = x^c$ in the above equation, yielding

$$\frac{dE(t)}{dt} = -\sum_{\substack{c \in GD \\ |x^c| \le 1}} \left[\frac{dx^c}{dt} \right]^2 \le 0 \tag{27}$$

which demonstrates that $E(t)$ is a monotonically decreasing function of time. Now, using Eq. 25 and Eq. 26, we can change Eq. 27 as follows,

$$\frac{dE(t)}{dt} = -\sum_{\substack{c \in GD \\ |x^c| \le 1}} \left[\frac{dx^c}{dt} \right]^2 = -\sum_{\substack{c \in GD \\ |x^c| \le 1}} \left[\frac{dy^c}{dt} \right]^2 = -\sum_{c \in GD} \left[\frac{dy^c}{dt} \right]^2 \le 0 \tag{28}$$

Since $E(t)$ is a bounded and monotonic decreasing function of time, its limit for $t \to \infty$ exists, from where the limit of its derivative must be cero,

$$\lim_{t \to \infty} \frac{dE(t)}{dt} = 0 \quad \Rightarrow \quad \lim_{t \to \infty} \sum_{c \in GD} \left[\frac{dy^c}{dt} \right]^2 = 0 \quad \Rightarrow \quad \lim_{t \to \infty} \frac{dy^c}{dt} = 0, \forall c \in GD \tag{29}$$

where we have used Eq. 28 and the fact that if a sum of positive summands is cero, every summand is cero. The final result in Eq. 29 demonstrates that all output variables converge.

We will show now that the state variables converge as well. Note that, after the output variables have converged, the state variables of the cells are confined in some region. We define

$$r^c(t) = \sum_{d \in N(c)} a_d^c y^d(t) + k^c \qquad \text{and} \qquad R^c = \lim_{t \to \infty} r^c(t) \tag{30}$$

The limit of $r^c(t)$ for $t \to \infty$ exists because the limits of all output variables exist.

If the state variable of a particular cell is in the linear region, we can obtain its limit using Eq. 5, rewritten using the above definitions, which yields,

$$\lim_{t \to \infty} x^c(t) = \lim_{t \to \infty} x^c(0)e^{-t} + \lim_{t \to \infty} e^{-t} \int_0^t r^c(\tau)e^{\tau} d\tau = \lim_{t \to \infty} \frac{\int_0^t r^c(\tau)e^{\tau} d\tau}{e^t} = \lim_{t \to \infty} \frac{r^c(t)e^t}{e^t} = R^c \tag{31}$$

where we have used the rule of L'Hopital. If the state variable is in the upper or lower saturated regions, we can use the same procedure starting from Eq. 4 or Eq. 6, respectively, which yields,

$$\lim_{t \to \infty} x^c(t) = \frac{m - 1 + R^c}{m} \qquad \text{or} \qquad \lim_{t \to \infty} x^c(t) = \frac{1 - m + R^c}{m} \qquad \text{also respectively} \tag{32}$$

4. The FSR model

The statements in the previous section are valid (except statement 4) in the range $m \in [1, \infty)$ which includes the Chua-Yang $(m = 0)$ and the FSR $(m \to \infty)$ models. In addition, in the FSR model, statement 4 is valid, and hence, the state vector never leaves the $[-1,1]^n$ hypercube, assuming that it is initially in it. Since the state vector is confined in $[-1,1]^n$, where $g(x^c) = x^c$ and $y^c = x^c$, the FSR model can be practically described by the following equation,

$$\frac{dx^c}{dt} = \left\{ \begin{array}{ll} \sum_{d \in N(c)} a'^c_d x^d + \sum_{d \in N(c)} b^c_d u^d + d^c & , \forall |x^c| < 1 \\ 0 & , \forall |x^c| = 1 \end{array} \right\}, \text{ with } \quad a'^c_d = \begin{cases} a^c_d, \forall c \neq d \\ a^c_d - 1, c = d \end{cases} \tag{33}$$

where the equilibrium points at $|x^c| = 1$ are stable or unstable depending on the sign of the derivative at the inner borders of interval $[-1,1]$, given by the upper expression in Eq. 33.

In this manner, the implementation of the nonlinear operator $y(x)$ in Eq. 2 is not required, neither that of the central segment of the dissipative term $g(x)$, which is grouped with the self feedback coefficient a^c_c. On the other hand, the state variable must be forced to remain in the unitary interval $[-1,1]$ using a hard limiter (external segments of $g(x)$).

The result is a CNN model with signal range equal to the unitary interval, independently of the application, which requires less circuitry for its implementation, and with stability and convergency properties similar to those of the original model. Numerical simulation shows that, for a given templates and offset coefficients, the results obtained using the Chua-Yang model and those obtained using ISR model are generally identical. It can also be shown that uniform (proportional) variations of the templates and offset coefficients affect only to the time constant of the network. This represent an increased robustness against global process parameter variations in VLSI implementations.

Annex: notation

c, d cell indexes
x^c state variable of cell c
y^c output variable of cell c
t normalized time. If t' is the physical time, $t = t'/\tau$, where τ is the time constant of the cells, assumed invariant through the network
a^c_d feedback coefficient. Weight of the contribution from the output y^d of cell d towards cell c.
b^c_d control coefficient. Weight of the contribution from the input u^d of cell d towards cell c.
d^c offset term. Constant contribution to cell c
$N(c)$ neighborhood of cell c
GD grid domain. Set of all inner cells in the network

References

[1] L.O. Chua and L. Yang: "Cellular Neural Networks: Theory". *IEEE Trans. Circuits and Systems*, Vol. 35, pp 1257-1272, October 1988.

[2] L.O. Chua and L. Yang: "Cellular Neural Networks: Applications". *IEEE Trans. Circuits and Systems*, Vol. 35, pp 1273-1290, October 1988.

[3] S. Espejo, A. Rodriguez-Vazquez, R. Dominguez-Castro, B. Linares and J.L. Huertas: "A Model for VLSI Implementation of CNN Image Processing Chips Using Current-Mode Techniques". *Proc. 1993 IEEE Int. Symp. on Circuits and Systems*, Chicago, May 1993.

[4] S. Espejo: "*VLSI Design and Modeling of CNNs*". Ph. Dissertation, University of Sevilla, March 1994.