Dynamical behaviour of neural networks iterated with memory

Paulin Melatagia Yonta, René Ndoundam, Maurice Tchuente

Department of Computer Science
University of Yaounde I
P.O. Box 812 Yaounde
CAMEROON
{pmelatag, ndoundam, tchuente}@uycdc.uninet.cm

RÉSUMÉ. Nous étudions l’itération avec mémoire où la mise à jour considère une suite d’états de chaque site et où l’ensemble des matrices d’interaction est palindromique. Nous analysons deux modes d’itération des réseaux : l’itération parallèle avec mémoire et l’itération série avec mémoire que nous introduisons dans ce papier. Nous définissons pour l’itération parallèle des fonctionnelles de Lyapunov qui nous permettent de caractériser les périodes et de borner les longueurs des transitoires des réseaux de neurones itérants avec mémoire. Pour l’itération série, nous utilisons un invariant algébrique pour caractériser les périodes du modèle d’évolution étudié.

ABSTRACT. We study memory iteration where the updating consider a longer history of each site and the set of interaction matrices is palindromic. We analyze two different ways of updating the networks: parallel iteration with memory and sequential iteration with memory that we introduce in this paper. For parallel iteration, we define Lyapunov functional which permits us to characterize the periods behaviour and explicitly bounds the transient lengths of neural networks iterated with memory. For sequential iteration, we use an algebraic invariant to characterize the periods behaviour of the studied model of neural computation.

MOTS-CLÉS : Réseau de neurones, Longueur de transitoire, Période, Itération avec mémoire, fonctionnelle de Lyapunov, Invariant Algébrique.

KEYWORDS : Neural network, Transient length, Period, Iteration with memory, Lyapunov functional, Algebraic invariant.
1 Introduction

Caianiello [1] has suggested that the dynamic behaviour of a neuron in a neural network with \( k \)-memory can be modeled by the following recurrence equation:

\[
x_i(t) = 1 \left( \sum_{j=1}^{n} \sum_{s=1}^{k} a_{ij}(s) x_j(t-s) - b_i \right), \quad t \geq k
\]

where

- \( i \) is the index of a neuron, \( i = 1, \ldots, n \).
- \( x_i(t) \in \{0, 1\} \) is a variable representing the state of the neuron \( i \) at time \( t \).
- \( k \) is the memory length, i.e., the state of a neuron \( i \) at time \( t \) depends on the states \( x_j(t-1), \ldots, x_j(t-k) \) assumed by all the neurons \( j = 1, \ldots, n \) at the previous steps \( t-1, \ldots, t-k \) \( (k \geq 1) \).
- \( a_{ij}(s) \) \( (1 \leq i, j \leq n \text{ and } 1 \leq s \leq k) \) are real numbers called the weighting coefficients. More precisely, \( a_{ij}(s) \) represents the influence of the state of the neuron \( j \) at time \( t-s \) on the state assumed by the neuron \( i \) at time \( t \).
- \( b_i \) is a real number called the threshold.
- \( 1 \) is the Heaviside function: \( 1(u) = 0 \) if \( u < 0 \), and \( 1(u) = 1 \) if \( u \geq 0 \).

Neural networks were introduced by McCulloch and Pitts [12], and are being investigated in many fields of artificial intelligence as a computational paradigm alternative to the conventional Von Neumann model. Neural networks are able to simulate any sequential machine or Turing machine if an infinite number of cells is provided. Recently, neural networks have been studied extensively as tools for solving various problems such as classification, speech recognition, and image processing [5]. The field of application of threshold functions is large [9, 11, 5].

Since neural network models have also been inspired by neurophysiological knowledge, the theoretical results may help to broaden understanding of the computational principles of mental processes.

The dynamics generated by Eq. (1) have been studied for some particular one-dimensional systems:

- when \( k = 1 \), one obtains a McCulloch and Pitts neural network [12]. Some results about the dynamical behaviour of these networks can be found in [12, 11].
- when \( n = 1 \), one obtains a single neuron (proposed by Caianiello and De Luca [2]) with memory that does not interact with other neurons. See [2, 3, 14] for some results.

There are few results in the study of evolution induced by Eq. (1). In [7] Goles established the following result:

**Theorem 1** [7]. If the class of interaction matrices \( A(s) : s = 1, \ldots, n \) is palindromic the periods \( T \) of parallel iteration with memory satisfies \( T | k + 1 \).

In [15] Tchuente generalized the preceding result by showing that the parallel iteration of a network of automata \( N \) can be sequentially simulated by another network \( N' \) whose local transition functions are the same as those of \( N \). By implementing a binary Borrow–Save counter, Ndoundam and Tchuente show that:

**Theorem 2** [13]. There exist a Caianiello automata network of size \( 2n + 2 \) and memory length \( k \) which describes a cycle of length \( k 2^{nk} \).

In this work, we show some dynamical results for parallel iteration with memory of neural network (Eq. (1)) when non-trivial regularities on coupling coefficients are satisfied. We also define the sequential iteration with memory of neural network and characterize its periodic behaviour. Our approach consists in defining appropriate Lyapunov functional [11] or algebraic invariant [10].
The remainder of the paper is organized as follows: in Section 2, some basic definitions and preliminary results are presented. In Section 3 we define two Lyapunov functional for parallel iteration with memory of neural network, characterize its periodic behaviour and bound its transient length. We also compare this bound with another obtained using sequential simulation of parallel iteration. In Section 4, we introduce sequential iteration with memory of neural network and study its periodic behaviour using an algebraic invariant.

2 Definitions and preliminary results

A neural network $N$ iterated with a memory of length $k$ is defined by $N = (I, A(1), ..., A(k), b)$, where $I = \{1, ..., n\}$ is the set of neurons indexes, $A(1), ..., A(k)$ are matrices of interactions and $b = (b_i : i \in \{1, ..., n\})$ is the threshold vector. Let $\{x(t) \in \{0,1\}^n : t \geq 0\}$ be the trajectory starting from $x(0), ..., x(k-1)$; since $\{0,1\}^n$ is finite, this trajectory must sooner or later encounter a state that occurred previously - it has entered an attractor cycle. The trajectory leading to the attractor is a transient. The period $(T)$ of the attractor is the number of states in its cycle, which may be just one - a fixed point. The transient length of the trajectory is noted $\tau(x(0), ..., x(k-1))$. The transient length of the neural network is defined as the greatest of transient lengths of trajectories, that is:

$$\tau(A(1), ..., A(k), b) = \max\{\tau(x(0), ..., x(k-1)) : x(t) \in \{0,1\}^n, 0 \leq t \leq k-1\}$$

The period and the transient length of sequences generated are good measures of the complexity of the neural network.

The updates of the state values of each neuron depends on the type of iteration associated to the model. The sequential iteration consists of one by one updating the neurons in a pre-established periodic order $(i_1, i_2, ..., i_n)$, where $I = \{i_1, i_2, ..., i_n\}$. The parallel iteration consists of updating the value of all the neurons at the same time.

3 Parallel iteration with memory

Let us consider the parallel iteration with memory of a finite neural network $N = (I, A(1), ..., A(s), b)$ give by Eq. (1); we can assume that:

$$\sum_{j=1}^{n} \sum_{x=1}^{k} a_{ij}(s) u_j \neq b_i, \forall i \in I, \forall u = (u_1, ..., u_n) \in \{0,1\}^n$$

(2)

We also assume that the set of interaction matrices $(A(s) : s = 1, ..., k)$ satisfy the palindrome condition:

$$A(k+1-s) = A(s)^T \text{ for } s = 1, ..., k$$

(3)

i.e.:

$$a_{ij}(k+1-s) = a_{ji}(s) \forall i, j \in \{1, ..., n\}, \forall s \in \{1, ..., k\}$$

(4)

Let $\{x(t) : t \geq 0\}$ be a trajectory of the parallel iteration, we define the following functional for $t \geq k$:

$$E(x(t)) = - \sum_{i=1}^{n} \left( \sum_{s=0}^{k-1} x_i(t-s) \left( \sum_{j=1}^{n} \sum_{s'=1}^{k-s} a_{ij}(s') x_j(t-s-s') \right) - b_i \sum_{s=0}^{k} x_i(t-s) \right)$$

(5)

**Proposition 1** If the class of interaction matrices $(A(s) : s = 1, ..., k)$ is palindrome, then the functional $E(x(t))$ is a strictly decreasing Lyapunov functional for the parallel iteration with memory of neural network.
We now give another proof of the Theorem 1 using the preceding Lyapunov functional.

**Proof of Theorem 1.** Let \(X = (x(0), ..., x(T - 1))\) a cycle of period \(T\). From the proof of Proposition 1 we found that, \(E(x(0)) = ... = E(x(T - 1))\) if \(\forall t = 0, ..., T - 1, x_i(t) = x_i(t + k + 1)\) for all \(i = 1, ..., n\), which implies that \(\tau (X_i) |k + 1\) \(\forall i \in I\). Then \(T \mid k + 1\). ■

To study the transient phase, we will work with another Lyapunov functional derived from \(E(x(t))\). Define:

\[
E^*(x(t)) = -\sum_{i=1}^{n} \left( \sum_{s=0}^{k-1} (2x_i(t - s) - 1) \sum_{j=1}^{n} \sum_{s'=1}^{k-1} a_{ij}(s') (2x_j(t - s' - 1) - 1) \right) \\
+ \sum_{i=1}^{n} \left( 2b_i - \sum_{j=1}^{n} \sum_{s=1}^{k} a_{ij}(s) \sum_{s=0}^{k-1} (2x_i(t - s) - 1) \right) 
\]

(6)

**Proposition 2.** If the class of interaction matrices \((A(s) : s = 1, ..., k)\) is palindromic, then the functional \(E^*(x(t))\) is a strictly decreasing Lyapunov functional for the parallel iteration with memory of neural network.

Denotes by \(\hat{X}\) the set of all initial conditions which do not belong to a period of length \(k + 1\):

\(\hat{X} = \{x(0) \in \{0, 1\}^n \text{ such that } x(0) \neq x(k + 1)\}\)

Recall that \(\hat{X}\) is empty iff the transient length of the neural network is null. If \(\hat{X} \neq \emptyset\) define:

\[e = \min \{-(E(x(k + 1)) - E(x(k))) : x(0) \in \hat{X}\}\]

(7)

We note \(e = 0\) if \(\hat{X} = \emptyset\).

**Proposition 3.** Let \(\{x(t) : t \geq 0\}\) be a trajectory; \(E^*(x(t))\) is bounded by:

\[
E^*(x(t)) \geq -(k + 1) \left\| 2b - \sum_{s=1}^{k} A(s) \cdot \mathbf{1} \right\| - \sum_{s=1}^{k} (k - s + 1) \| A(s) \| \]

(8)

and

\[
E^*(x(t)) \leq \left\| 2b - \sum_{s=1}^{k} A(s) \cdot \mathbf{1} \right\| - 2k \sum_{s=1}^{n} e_i + \sum_{s=1}^{k} (s - 1) \| A(s) \| \]

(9)

where:

\[e_i = \min \left\{ \sum_{j=1}^{n} \sum_{s=1}^{k} a_{ij}(s) u_j(s) - b_i : u(s) \in \{0, 1\}^n, s = 1, ..., k \right\}\]

(10)

and \(\|A(s)\| = \sum_{i=1}^{n} \sum_{j=1}^{k} |a_{ij}(s)|, \left\| u \right\|_1 = \sum_{i=1}^{n} |u_i| \) for any vector \(u \in \mathbb{R}^n\), \(\mathbf{1} = (1, ..., 1)^t\) is the \(1\)-constant vector.

**Theorem 3.** If the class of interaction matrices \((A(s) : s = 1, ..., k)\) is palindromic, then the transient length \(\tau (A(1), ..., A(k), b)\) of parallel iteration with memory of neural network is bounded by:

\[
\tau (A(1), ..., A(k), b) \leq \frac{1}{2e} \left( (k + 2) \left\| 2b - \sum_{s=1}^{k} A(s) \cdot \mathbf{1} \right\| + k \sum_{s=1}^{\bar{k}} \| A(s) \| - 2k \sum_{i=1}^{n} e_i \right) \text{ if } e > 0 \\
\tau (A(1), ..., A(k), b) = 0 \text{ if } e = 0 
\]

(11)
For \( k = 1 \), this bound is obtained in [8] and a family of neural network which attains it is given.

**Remark 1.** By using construction suggested in [15] to simulated sequentially a network of automata, we obtain for \( N = (I, A(1), \ldots, A(k), b) \) the sequential iterating network \( N' = (I', A', b') \) where:

\[
I' = \bigcup_{p=1}^{k+1} I_p \text{ with } I_p = \{ (p-1)n+1, \ldots, (p-1)n+n \}
\]

\[
A' = \begin{bmatrix}
\mathbf{0} & A(k) & A(k-1) & \cdots & A(1) \\
A(1) & \mathbf{0} & A(k) & \cdots & A(2) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
A(k-1) & A(k-2) & A(k-3) & \cdots & A(k) \\
A(k) & A(k-1) & A(k-2) & \cdots & \mathbf{0}
\end{bmatrix}
\]

\( \mathbf{0} \) is the 0-constant matrix

\( b' = (b_1, b_2, \ldots, b_{k+1}) \), \( b_p = b \forall p \in \{1, \ldots, k+1\} \)

This simulation can be used to bound the transient length of parallel iteration with memory. Indeed let \( t \geq k+1 \) such that \( t = (p-1) \mod (k+1) \) (\( 1 \leq p \leq k+1 \)). The updating of \( N' \) can be written:

\[
\begin{cases}
x_i(t) = x_i(t-1) & \text{if } i \in I_q \text{ and } q \neq p \\
x_i(t) = 1 \left( \sum_{q=1}^{k+1} \sum_{j \in I_q} a_{ij'}((p-q) \mod (k+1))x_{j'}(t-1) - b_i \right) & \text{if } i \in I_p
\end{cases}
\]

where \( i' = i - (p-1)n \), \( j' = j - (q-1)n \) and \( a(0) = \mathbf{0} \).

It is a block sequential iteration on neural network (sequential with respect to the order of blocks \( I_p \) and parallel within each block).

### 3.1 Transient length of block sequential iteration on neural network

Let \( N = (I, A, b) \) a neural network and \( I_1, \ldots, I_l \) an ordered partition (with respect the order of \( \mathbb{Z} \)) of \( I = \{1, \ldots, m\} \); i.e \( i < j \) if \( i \in I_r, j \in I_{r'} \) with \( r < r' \). The block sequential updating of \( N \) is: at time \( t > 0 \), \( t = (r-1) \mod l \) (\( 1 \leq r \leq l \)):

\[
\begin{cases}
x_i(t) = x_i(t-1) & \text{if } i \in I_r \text{ and } r' \neq r \\
x_i(t) = 1 \left( \sum_{r'=1}^{l} \sum_{j \in I_{r'}} a_{ij}x_j(t-1) - b_i \right) & \text{if } i \in I_r
\end{cases}
\]

We now show the following proposition:

**Proposition 4** If \( A \) is a real symmetric matrix with \( a_{ij} = 0 \) for \( i, j \in I_r \) (\( 1 \leq r \leq l \)), then the functional \( E_\beta \):

\[
E_\beta(x(t)) = -\frac{1}{2} \sum_{i=1}^{m} x_i(t) \sum_{j=1}^{m} a_{ij}x_j(t) + \sum_{i=1}^{m} b_i x_i(t)
\]

is a strictly decreasing Lyapunov functional for the block sequential iteration of the neural network.
Corollary 1 If A is a real symmetric matrix with \( a_{ij} = 0 \) for \( i, j \in I_r \) (\( 1 \leq r \leq l \)), then the periods \( T \) of the block sequential iteration on the neural network satisfies \( T = l \).

Let \( E^*_b(x(t)) \) be the functional [4] defined by:

\[
E^*_b(x(t)) = -\frac{1}{2} \sum_{i=1}^{m} (2x_i(t) - 1) \sum_{j=1}^{m} a_{ij} (2x_j(t) - 1) + \sum_{i=1}^{m} \left( 2b_i - \sum_{j=1}^{m} a_{ij} \right) (2x_i(t) - 1) \quad (15)
\]

Proposition 5 Let \( A \) be a real symmetric matrix with \( a_{ij} = 0 \) for \( i, j \in I_r \) (\( 1 \leq r \leq l \)). The difference \( \Delta_t E^*_b = E^*_b(x(t)) - E^*_b(x(t-1)) = 4\Delta_i E_b \) and then \( E^*_b(x(t)) \) is a strictly decreasing Lyapunov functional for the block sequential iteration on neural network.

The functional \( E^*_b(x(t)) \) is a more appropriate Lyapunov functional to study the transient length of \( N \). Indeed, it is easy to show that \( |E^*_b(x(t))| \leq \frac{1}{2} \|A\| + \|2b - AT\|_1 \). Denote by \( \tilde{X}^* \) the set of all initial conditions which do not belong to a period of length \( l \):

\[
\tilde{X}^* = \{ x(0) \in \{0,1\}^n \text{ such that } x(0) \neq x(l) \}
\]

If \( \tilde{X}^* \neq \emptyset \) define:

\[
e^* = \min \{ -(E_b(x(l)) - E_b(x(l - 1))) : x(0) \in \tilde{X}^* \}
\] (16)

We note \( e^* = 0 \) if \( \tilde{X}^* = \emptyset \).

Proposition 6 Let \( A \) be a real symmetric matrix with \( a_{ij} = 0 \) for \( i, j \in I_r \) (\( 1 \leq r \leq l \)). The transient length of block sequential iteration \( \tau_b(A, b) \) is bounded by:

\[
\tau_b(A, b) \leq \frac{1}{4e^*} \left( \|A\| + 2 \|2b - AT\|_1 \right)
\] (17)

Remark 2. Since \( N' \) is a block sequential iterating neural network with \( a'_{ij} = 0 \) when \( i, j \in I_p \) (\( 1 \leq p \leq k + 1 \)), its transient length \( \tau_b(A', b') \) is bounded by:

\[
\tau_b(A', b') \leq \frac{1}{4e} \left( \|A'\| + 2 \|2b' - A'T\|_1 \right)
\] (18)

To compare this bound with the ones obtained in Theorem 3, Eq. (18) must be rewritten as follows:

\[
\tau_b(A', b') \leq \frac{1}{4e} \left( 2(k + 1) \left\| 2b - \sum_{s=1}^{k} A(s)T \right\|_1 + (k + 1) \sum_{s=1}^{k} \|A(s)\| \right)
\] (19)

Let

\[
\tau = \frac{1}{4e} \left( (k + 2) \left\| 2b - \sum_{s=1}^{k} A(s)T \right\|_1 + k \sum_{s=1}^{k} \|A(s)\| - 2k \sum_{i=1}^{n} e_i \right)
\] (20)

and

\[
\tau' = \frac{1}{4e} \left( 2(k + 1) \left\| 2b - \sum_{s=1}^{k} A(s)T \right\|_1 + (k + 1) \sum_{s=1}^{k} \|A(s)\| \right)
\] (21)

we find

\[
\tau' - \tau = \frac{1}{4e} \left( k \left\| 2b - \sum_{s=1}^{k} A(s)T \right\|_1 + \sum_{s=1}^{k} \|A(s)\| + 2k \sum_{i=1}^{n} e_i \right) > 0
\] (22)

Hence, the first bound (\( \tau' \)) is better than the ones obtained by sequential simulation of the parallel iteration of a neural network with memory.
4 Sequential iteration with memory

We define the sequential iteration with memory as follows: the update of the neurons when the network evolves from \( t = 1 \) to \( t \) occurs hierarchically according to a pre-established periodic order on \( I \) (we shall assume, without loss of generality, that the order on \( I \) is the same order as \( I \) possesses as a subset of \( \mathbb{Z} \)). Thus, when the neuron \( i \) changes from \( x_i(t - 1) \) to \( x_i(t) \), all the vertices \( j < i \) have already evolved. The states considered for the iteration are \( x_j(t + 1 - s) \) for \( j < i \) and \( x_j(t - s) \) for \( j \geq i \); \( s = 1, \ldots, k \).

Thus the configuration of the system are \( x(t) \in \{0,1\}^n \), the set of interaction matrices is \( \{ A(s) = (a_{ij}(s) : i, j \in \{1, \ldots, n\} : s = 1, \ldots, k \} \) and the threshold vector is \( b = (b_i : i \in \{1, \ldots, n\}) \). Since a sum over an empty set of indexes is null \( ( \sum_{j=1}^{n} = 0 \) if \( i = 1 \)), the sequential updating with memory of the neural network is written:

\[
x_i(t) = 1 \left( \sum_{j=i}^{n-1} \sum_{s=1}^{k} a_{ij}(s) x_j(t + 1 - s) + \sum_{j=1}^{n} \sum_{s=1}^{k} a_{ij}(s) x_j(t - s) - b_i \right)
\]  \hspace{1cm} (23)

When \( k = 1 \), we obtain a Mc Culloch and Pitts neural network iterating sequentially [12, 6].

Let \( T \) be the period of the neural network. Let \( X = (x(0), \ldots, x(T - 1)) \) be a \( T \)-cycle. For any couple of local cycles \((X_i, X_j)\) we define the sequential functional (algebraic invariant) by:

\[
L(X_i, X_j) = \begin{cases} 
\sum_{s=1}^{k} a_{ij}(s) \Delta V^{k-s+1,s-1}(X_i, X_j) & \text{if } j < i \\
\sum_{s=1}^{k-1} a_{ij}(s) \Delta V^{k-s,s-1}(X_i, X_j) & \text{if } j = i \\
\sum_{s=1}^{k} a_{ij}(s) \Delta V^{k-s+1,s}(X_i, X_j) & \text{if } j > i 
\end{cases}
\]  \hspace{1cm} (24)

From Lemma results obtained by Goles et al. in [9] we find:

\[
\text{if } \tau(X_i) | k \text{ then } L(X_i, X_j) = 0 \text{ for any } j \in I
\]  \hspace{1cm} (25)

Now for evolution Eq. (23) we establish the following lemma:

**Lemma 1** For any family of interaction matrices \( \{ A(s) : s = 1, \ldots, k \} \) such that \( a_{ii}(k) \geq 0 \) for any \( i \in I \) we have:

\[
\sum_{j \in I} L(X_i, X_j) \leq 0 \text{ for any } i \in I
\]  \hspace{1cm} (26)

\[
L(X_i, X_j) = 0 \text{ for any } j \in I \text{ iff } \sum_{j \in I} L(X_i, X_j) = 0 \text{ iff } \tau(X_i) | k
\]  \hspace{1cm} (27)

\[
\sum_{j \in I} L(X_i, X_j) < 0 \text{ iff } \tau(X_i) \text{ does not divide } k
\]  \hspace{1cm} (28)

\[
\sum_{j \in I} \sum_{j \in I} L(X_i, X_j) = 0 \text{ iff } \tau(X_i) | k \text{ for any } i \in I
\]  \hspace{1cm} (29)

Now assume that the set of interaction matrices \( \{ A(s) : s = 1, \ldots, k \} \) satisfy:

\[
\text{diag}(A(s)) = \text{diag}(A(s + 1)) \quad \forall s = 1, \ldots, k - 1
\]  \hspace{1cm} (30)

where \( \text{diag}(A(s)) = (a_{ii}(s) : i \in I) \)

**Theorem 4** If the class of interaction matrices \( \{ A(s) : s = 1, \ldots, k \} \) is palindromic (i.e. satisfy (3)) and satisfy (30) then the period \( T \) of the neural network iterated sequentially with \( k \)-memory satisfies \( T | k \).
5 Conclusion

We study neural networks of Caianiello under some assumptions on interaction matrices. For parallel iteration, using Lyapunov functional, we characterize the periods and bounds explicitly the transient lengths of neural networks. The bound is compare with the ones obtained by sequential simulation of the parallel iteration of a neural network with memory and proves more better. We introduce sequential iteration with memory of neural networks and, using an algebraic invariant, characterize its period behaviour.

Acknowledgements

This work was done within the framework of the Associateship Scheme of the Abdus Salam International Centre of Theoretical Physics (ICTP), Trieste, Italy. Financial support from the Swedish International Development Cooperation Agency (SIDA) is acknowledged.

Références