Fragmented Basins of Attraction of Recursive Processing Elements in Associative Neural Networks and its Impact on Pattern Recovery Performance

Emilio Del Moral Hernandez[1]

Abstract -This paper addresses recurrent neural architectures based on coupled bifurcating nodes that exhibit chaotic dynamics. The nodes are composed of logistic recursive maps, which interact through parametric coupling, i.e., through dynamic modulation of the bifurcation parameters. These networks are used to implement associative memories in which the coding of binary strings is done through spatio-temporal attractors with period-2 cycles. The associative performance of such arrangements is measured under several levels of analog noise in the prompting pattern (initial conditions of the coupled recursions). The phenomena of unbalanced power of attractors is detected. The paper also identifies and analyzes the issue of fragmented (non-convex) regions associated to the attractors representing binary zeros and binary ones of the stored strings. This subject is approached in the context of associative networks operating under analog noise and the related degradation of performance. A simple pre-processing technique aiming to overcome the mentioned fragmentation of basins of attraction resulted in marked improvement of network performance, particularly in highly noisy situations.

I. INTRODUCTION

Artificial Neural Networks employing nodes with complex dynamics is a growing subject in both arenas, theory and application [1-5]. We have worked in the past years in the development of associative architectures based on RPEs – Recursive Processing Elements, producing several design methodologies and interesting results such as the superior immunity to prompting noise, by factors of *1.5* to *2.0* with respect to Hopfield associative networks (see Table I in Section II for a comparison) [6-9]. These RPEs nodes are parametric recursions, such as the logistic map, which interact through parametric coupling, self organize, and make the network evolve to spatio-temporal attractors that encode stored patterns. This strategy is used to implement associative memories in which the coding of binary strings is done through period-2 cycles [6-9].

In the following section of the paper (Section II), we describe the main elements for the understanding of the operation and design of associative RPEs networks. In Section III, we present results on the characterization of pattern recovery performance under noisy prompting conditions and the characterization of the network's landscape of attractors. This "landscape characterization" allows us to detect the unbalanced power of attractors under high levels of prompting noise. Section IV discusses the phenomenon of fragmented basins of the attractors in logistic RPEs nodes and relates it to the degradation of performance of associative RPEs networks. Then, Section V describes a successful method to cope with such a degradation of performance, and Section VI concludes the paper.

II. SUMMARY OF PRINCIPLES OF NEURAL NETWORKS OF RPEs NODES - BIFURCATING RECURSIVE PROCESSING ELEMENTS (RPEs)

The local dynamics of the logistic RPEs nodes is defined by Equation 1, and its parametric coupling is expressed generically in (2).

$$x_{i,n+1} = p_{i,n} \cdot x_{i,n} \cdot (1-x_{i,n}) \qquad (1)$$

$$p_{i,n} = f(w_{ij}; x_{j,n}) \qquad (2)$$

In (1), n represents the discrete time, $x_{i,n}$ and $x_{i,n+1}$ represent consecutive values of the state variable for node i, and the $p_{i,n}$ is the bifurcation parameter of the logistic recursion. Figure 1 shows illustrative return maps for a generic parametric recursion R_p relating $x_{i,n}$ and $x_{i,n+1}$.
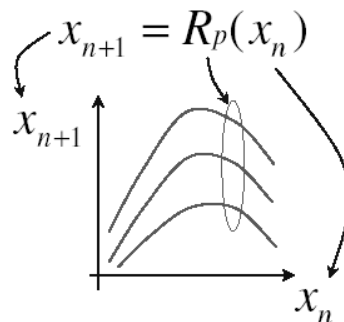

Fig. 1. Illustrative plot showing a family of parametric recursions, with three return maps relating x_n and x_{n+1}. According to the value of the numeric parameter p, the recursion R_p maps the values x_n into different values of x_{n+1}.

[1] E. Del-Moral-Hernandez is with the Department of Electronic Systems Engineering, Polytechnic School of the University of São Paulo
Av. Prof. Luciano Gualberto, trav.3, n.158 - 05508 900 – São Paulo – SP – Brazil (e-mail: emilio_del_moral@ieee.org).

In the Equation (2) above, f represents some function or functional, adequate for the formation of collective behavior, and its form for our studies is specified ahead. The w_{ij} are the synaptic connections from neighboring neurons j, and $x_{j,n}$ their outputs at time step n. Figure 2 represents the structure of an associative network of logistic RPEs with full connection among them. Each circle "L1 to L4" represents a bifurcating recursion, i.e., a logistic map in these studies.
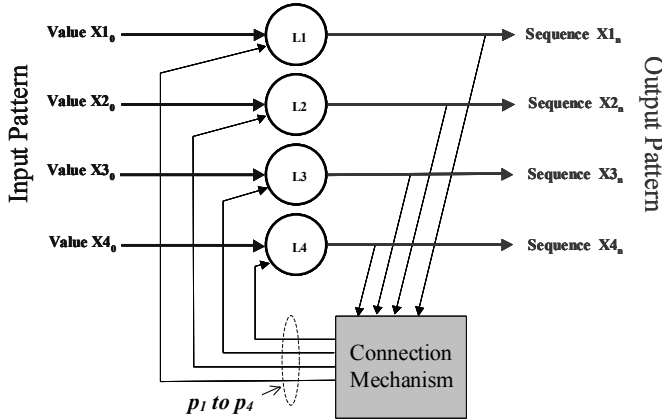

Fig. 2. Illustrative associative network composed of four logistic RPEs with full connection among them. Each circle "L1 to L4" represents a logistic recursion.

The parametric connection among nodes is provided through appropriate changes in the bifurcation parameters of the logistic nodes (parameters p_1 to p_4 in the illustration of Figure 2). These changes in the p_i are driven by the internal states which represent the network activity ($x_{1,n}$ to $x_{4,n}$), according to the generic Equation (2) above, which in our studies assumes the specific form of Equation (3) [8,9]:

$$\Delta p_i = - scale \cdot (x_i - avg) \cdot (\Sigma_j w_{ij} \cdot (x_j - avg)) - decay \quad (3)$$

In (3), Δp_i represents gradual increments in the bifurcation parameter, the constant avg is the average value of the state variables x, and the constants $scale$ and $decay$ have to be chosen to promote adequate balance between the chaotic behavior during the search for stored patterns and the ordered behavior of periodic attractors, once this search is achieved.

We will not have the space here for a complete explanation of this process, but instead we will refer to the papers presented in two previous IJCNN conferences, in which the design of associative RPEs networks was developed in detail [8,9]. We will show though the bifurcation diagram of the logistic RPE node in Figure 3 and mention that for stabilized networks which already have recovered a given stored pattern, the node's x_i trajectory is a period-2 cycle such as the one indicated in the diagram, in which the oscillation happens among two values A and B, and the p_i bifurcation parameter is settled at a given value p, related to the cycle A and B, according to the bifurcation diagram.
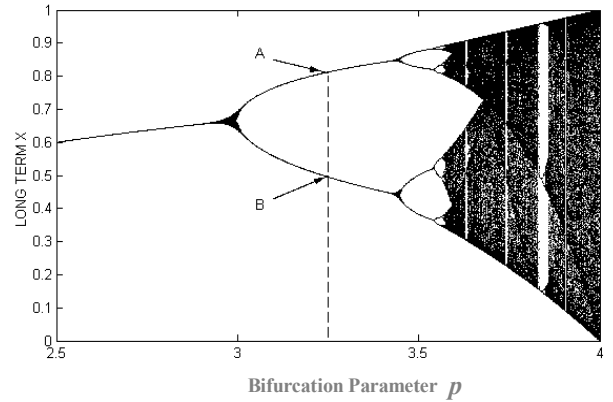

Fig 3. The bifurcation diagram for the logistic recursion. The horizontal axis represents the value of the bifurcation parameter p, and the vertical axis represents the values visited in the long term by the state variable x. The marked points A ($x = 0.81$) and B ($x = 0.50$) exemplify the period-2 cycle that emerges when p is 3.25.

During the use of the associative network, the prompting pattern defines the initial state of the network (the values $x_{i,0}$), and the coupled logistic elements are then free to evolve in time. In their evolution, the RPE nodes possibly exercise different modalities of dynamical behavior, and they eventually reach accommodation, with all the p_i settling at a lower limit p_{min}, and all the state variables x_i perform period-2 cycles, defined by the pair of values A and B. This period-2 oscillation is sufficient for the coding of binary strings, since we have two possible phases of oscillation. In this way, when the network stabilizes, some of the nodes oscillate in a given phase, coding the *zeros* of a binary string previously stored in the network, and some of the nodes oscillate in the opposite phase, coding the *ones*.

As in Hopfield networks, the synaptic connections w_{ij} among nodes are calculated from the M binary patterns P_k (binary strings) to be stored in the network, according to Equation (4) [9-11]:

$$w_{ij} = \Sigma_{k=1:M} (P_{i,k} \cdot P_{j,k}) \ , \ i \neq j \quad (4)$$

The terms $P_{i,k}$ and $P_{j,k}$ in the summation in (4) are two bits of the stored pattern P_k which are represented, respectively, by the long term cyclic activities x of the nodes i and j. In order to have (4) expressing the correlation among nodes i and j, the terms $P_{i,k}$ and $P_{j,k}$ have to be expressed (when used in the equation above) in the bipolar representation, i.e., having values $+1$ or -1. As mentioned before, many additional details on the design of associative RPEs networks can also be found in [8,9].

It is interesting to show here some comparison between the performances of associative RPEs networks and Hopfield networks [8]. Table I presents results with both architectures being used to store binary strings which are later recovered from distorted versions of them.

Table I – Hamming Recovery Error for Associative
Memories of type Hopfield and RPEs

(100 nodes, with load M=3 patterns)	Average Hamming Error in Hopfield Architecture	Average Hamming Error in RPEs Network
Prompting Noise 10%	0.00%	0.00%
Prompting Noise 20%	0.33%	0.18%
Prompting Noise 30%	2.12%	1.26%

The table presents performance results with three different levels of input distortion, measured as a percentage of flipped bits in the prompting pattern received by the associative arrangements. The performance of each associative architecture is quantified based on the average Hamming error of the recovered pattern. We can see from Table I that the average recovery error is much smaller in RPEs networks.

III. CHARACTERIZATION EXPERIMENTS ON THE PATTERN RECOVERY PERFORMANCE IN ASSOCIATIVE RPEs NETWORKS AND ITS DEGRADATION WITH ANALOG PROMPTING NOISE

Figure 4 presents an illustrative Limit Set Diagram (LSD), a type of diagram used to represent the dynamical state of a given network under study [6,12]. In this particular LSD, we show the outcome of one of the experiments with RPEs networks with 100 coupled nodes programmed to store simultaneously 3 patterns. The recovery experiments presented in this and the following sections of the paper were carried out with prompting initial conditions distorted by analog noise. The solid line in the LSD represents the values $x_{i,0}$, i.e., the noisy initial condition imposed to the 100 nodes. The horizontal axis represents / identifies the 100 nodes that compose the network, and right above each i (node number), we have two dots representing a limit cycle of period-2, which is exercised by the node once the network reaches stabilization. The series of circles surrounding some of these dots representing the limit cycles, defines one sample in time of the state variable vector (the values of the 100 x_i), showing thus which nodes are oscillating in phase, and which nodes are oscillating in counter phase [6].

In the studies presented here, the level of analog noise in the initial condition is defined in a normalized form. Notice that when we are dealing with analog noise environments, the issue of the relative scales of the analog noise and the separation between the representatives of binary *one* and binary *zero* is an important issue. Therefore, for the characterizations that we perform, we define the level of analog noise by comparing its average magnitude to the distance $|A-B|$. This distance is the amplitude of period-2 oscillations of the final state that represents the binary information. The so named "analog noise factor" is thus defined as the ratio $<|noise|> / |A-B|$.
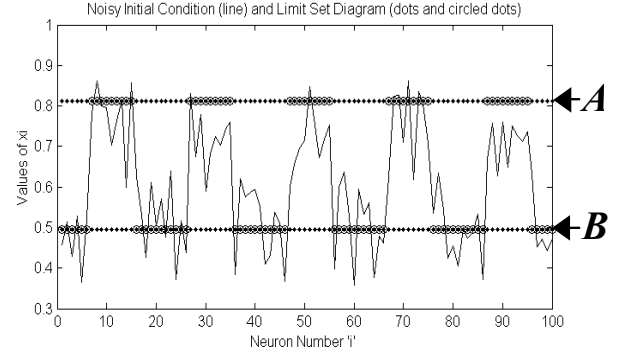

Fig. 4. Limit Set Diagram (LSD) showing both the prompting pattern (initial condition) of a network, through the solid line linking the values $x_{i,0}$, and the period-2 limit cycles exercised by the 100 nodes in the long term. Each node is represented in the horizontal axis (neuron number *i*), and the corresponding limit cycles are represented by the dot and circled dot right above each *i*. (Figure from [6]). The outcome corresponds to one of the stored patterns.

In Figure 5 we have a plot showing the evolution of the Hamming error in the recovered pattern at the network output (represented in the vertical axis), for arbitrary levels of analog noise in the initial conditions. The level of input noise is represented in the horizontal axis. The size of the network is 100 nodes and these particular experiments on degradation of performance with noise were conducted with networks storing 3 patterns, as mentioned before. We observe that virtually there are no errors in pattern recovery as long as the analog noise factor is bellow *0.35*.

As we can see in Figure 5, as the level of analog noise becomes significantly high, with values for the noise factor close to *0.8* for example, the average level of error in the recovered patterns goes to *75* %. Such a percentage is measured by the average Hamming distance between the clean version of the prompting pattern and the observed output given by the network (as in Table I).
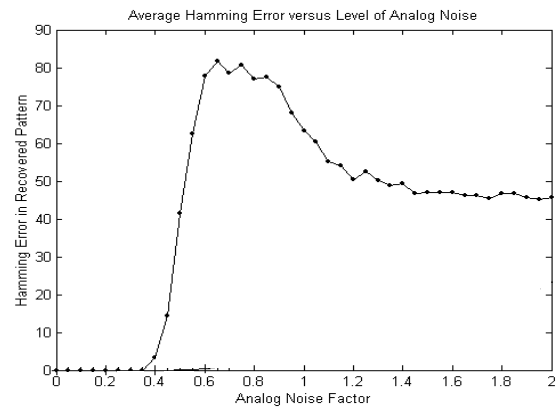

Fig. 5. Error plot for arbitrary levels of analog noise in the range of noise factors from 0 to 2. The horizontal axis shows the level of prompting noise (noise factor) and the vertical axis shows the average Hamming error in pattern recovery.

In order to have a better understanding of the process of performance degradation in the recovery of patterns by the associative structure, we developed additional tools that could help us in representing the outcomes of large series of experiments, as we have in the Error Plot (Figure 5), but still giving us some of the detailed information on the output patterns produced by the network, in a similar way as we have in the LSDs (Figure 4). A practical way that we used for that was by developing a histogram tool for the joint representation of the network outputs and the statistics on the final states.

Figure 6 shows some of these "Graphic Histograms", used in showing the statistical scenario in pattern recovery experiments dealing with different levels of prompting noise. We call this tool a Graphic Histogram because, in it, each frequency of occurrence of an output binary string is accompanied by a graphical representation of the string itself. We can use the Graphic Histograms for the analysis of the repertoire of final states that emerge during the experiments with high levels of noise, as well as for the analysis of experiments with random initial conditions, so that we can evaluate the relative frequencies of occurrence for each final state and detect unplanned preferential attractors in the network.
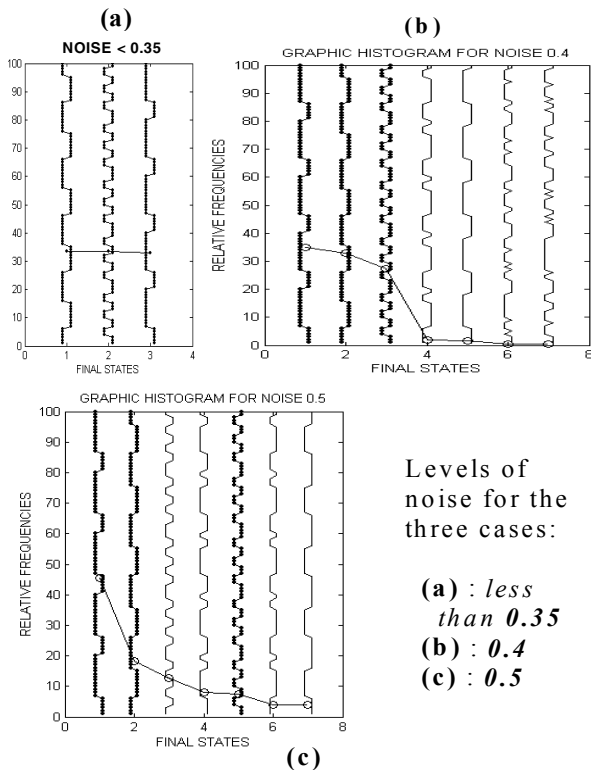
In the experiments represented in Figure 6, we have used as prompting patterns a balanced mixture of all the programmed patterns (all of them are given to the input of the network with the same frequency), corrupted by a certain level of analog noise. The first of these histograms (a) represents experiments with analog noise factors no larger than *0.35*. The second histogram (b) corresponds to a noise factor of *0.4* and the third one (c) corresponds to the noise factor *0.5*. In these graphic histograms, the binary patterns produced at the output of the network, coded in the final trajectories of the state variables of the network, appear ordered from the more frequent one to the less frequent one. The value of the relative frequency of each output pattern appears in the vertical axis represented in a scale from *0%* to *100%*. The line linking the circles shows the relative frequencies of the recovered patterns in the experiments with noisy promptings. The bolded outcomes represent planned attractors, i.e., stored patterns, and the other ones represent spurious attractors, including mirror patterns and other spurious patterns as well.

Using this technique, it is possible to easily verify the unbalanced power of the different network attractors, it is easy to identify the programmed patterns among all the patterns considered in the histogram, and it is easy to check similarities and differences between the programmed patterns and each spurious pattern appearing in the repertoire of network outcomes. We have thus a practical and easy to use tool, for the exploration of the "landscape of attractors".

We can see in Figure 6 that for reduced levels of noise, up to *0.35* (part (**a**) of the figure), only the stored patterns are produced at the output, all of them with the same relative frequency (*33%*). When we go to intermediate levels of noise, such as the experiments with noise factor *0.4*, as in part (**b**) of Figure 6, the stored patterns are still the predominant ones in the histograms, appearing as the three more frequent patterns (*35%*, *32%* and *28%*), although other patterns, in addition to the stored ones, start to appear at the output, as a result of the significant level of noise. We also observe the appearance of some small unbalance in the power of attraction of the three principal attractors.

Finally, for experiments with large noise factors, such as the ones represented in Figure (**c**), not only the stored patterns (which are at the positions *#1*, *#2* and *#5*) appear with significant frequencies in the initial part of the histogram, but also mirror patterns and other spurious patterns as well. We can see in this third histogram a dramatic unbalance in the power of attraction of the several stored patterns, with one of them being recovered in *45%* of the experiments. Graphical Histograms for higher levels of analog noise in the prompting initial conditions were also obtained, and the unbalance of the attractors became still more dramatic. Therefore, high levels of analog noise induce the appearance of strongly preferential attractors.

Although not included here due to limitation of space, another related series of experiments that we have done and

Fig. 6. Three Graphic Histograms, showing the relative frequency (vertical scale, from *0%* to *100%*) of each output pattern generated by the associative network, as a result of prompting inputs corrupted by analog noise. The noise factors are different and have increasing magnitudes: less than *0.35* in *(a)*, *0.4* in *(b)***,** and *0.5* in *(c)*.

that helped us in finding the reasons for the preferential attractors mentioned above is related to what we call "**digital noise**". In this form of experiment, the only type of distortion imposed to the initial conditions is the flipping of a certain percentage of bits in the input binary pattern, i.e., a certain percentage of the initial conditions that should be set $x_0 = A = 0.81$ are changed to $x_0 = B = 0.50$, and the same percentage of the initial conditions that should be set $x_0 = 0.50$ are changed to $x_0 = 0.81$.

We performed experiments with digital noise and obtained Error Plots and Graphic Histograms, as we did for analog noise (Figures 5 and 6 above). Although the Error Plots for digital noise have shown that the performance of recovery also decayed under situations of high digital noise, as expected, the Graphical Histograms for experiments with high levels of digital noise revealed a much better balance among the patterns generated at the output, as compared to the similar experiments with analog noise. The contrast between the Graphic Histograms for initial conditions with digital noise and the histograms for initial conditions with analog noise indicated that the attractors landscape for experiments with analog prompting noise favors certain patterns with respect to the others in a marked way. That didn't happened so heavily when the noise was purely digital, even when the number of flipped bits grew to high levels. This indicates that the phenomenon of preferential attractors is more related to the nature of the excursion allowed to the distorting signal than to its magnitude. Regarding this comment, notice that in experiments with digital noise with **50%** of the bits flipped, we have that half of the x_0 that should be A are changed to B and vice versa, this implying that the equivalent factor of analog noise ($<|noise|> / |A-B|$) is **0.5**. Nevertheless, the observed unbalance among recovered patterns is much less significant in such a situation of digital noise than it is in a purely random analog noise with the same factor **0.5**. This made us to look in more detail to the basins of attractions of the logistic map, our RPE node, in order to understand why is that similar levels of analog noise and digital noise can lead to so different scenarios of unbalance of attractors.

IV. STUDYING THE FRAGMENTED BASINS OF ATTRACTION IN LOGISTIC BIFURCATING RPES NODES

Our next steps in the understanding of the observed preferential attractors came from the analysis of the isolated logistic recursion in what respects the sets of values of initial conditions x_0 that result in each one of the two limit cycles used to represent binary **zeros** and binary **ones**. Experiments for the mapping of these two sets reveal that they have inconvenient features with regard to the setting of initial conditions which are distorted by analog noise.

In order to prepare the setting for the discussion that follows, it is important first to remember that the RPE nodes

represent information through period-2 cycles involving the values A and B described before (Sections II and III), and also to remember that the associative network relies on the attractor behavior of the RPEs in order to promote immunity to noise in the recovery of distorted versions of the stored patterns presented at the network input. At the single node level, this ingredient (attractor behavior) means that we have tolerance to limited deviations in the initial condition with respect to the ideal values $x_0 = A = 0.81$ and $x_0 = B = 0.50$, i.e., those values of x_0 which would instantaneously lead to the desired patterns of cycling that represent either a binary **one** or a binary **zero**. In particular, values of the initial conditions close to A and B will cause a limited transitory which will disappear in the long term, as illustrated in the parts (**a**) and (**b**) of Figure 7. In these two parts of the figure, we have experiments with the initial condition set to $x_0 = 0.85 = A+0.04$ and $x_0 = 0.45 = B-0.05$, i.e., initial conditions just a bit outside the segment AB.
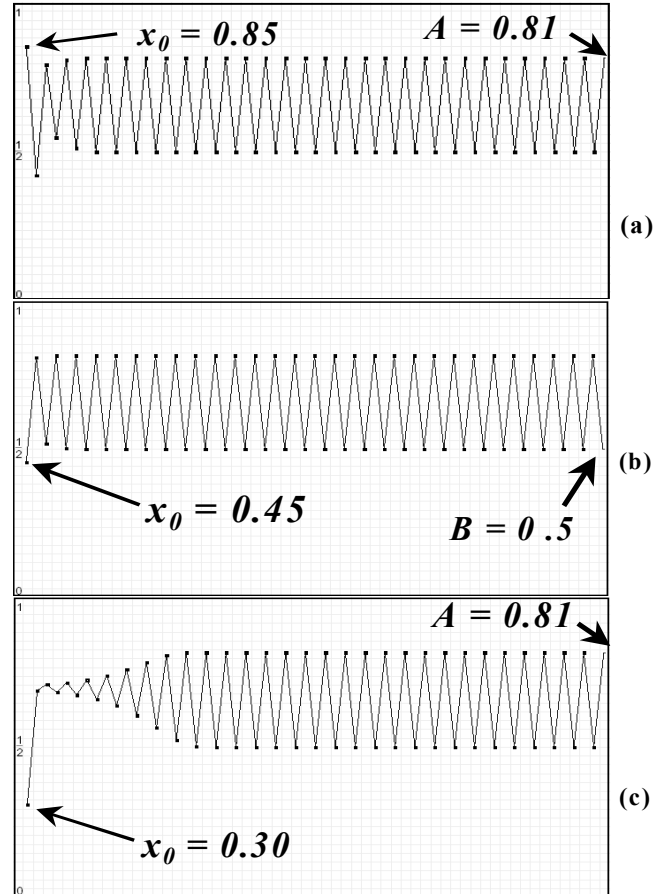

Fig 7. Long term phases of oscillation with different choices for the initial conditions. In (**a**), we have an x_0 value which is close to A and thus leads to the same long term behavior as $x_0 = A$ does. In (**b**), we have an x_0 value which is close to B and leads to the same long term behavior as $x_0 = B$ does. In (**c**), we have an x_0 value which is not so close to either A or B. Although its distance to B is smaller than its distance to A, it leads to the same long term behavior as $x_0 = A$ does.

If we perform similar experiments with x_0 values inside the *AB* segment, we observe the existence of a boundary at $x_0 = 0.69$ and a bipartition of the segment *AB*: an "upper part" of the segment, **[0.69,0.81=A]**, which defines initial conditions that promote a certain phase of oscillation in the long term, and a "lower part" of the segment, **[B=0.50,0.69]**, which defines initial conditions that promote the opposite phase of oscillation in the long term. This boundary corresponds to the fixed point defined by the condition $x_{n+1} = x_n = p \cdot x_n \cdot (1-x_n)$, which for *p = 3.25* is unstable [13]. Initial conditions right above this unstable fixed point result in period-2 long term behavior with a given phase of oscillation, and initial conditions right bellow it result in period-2 long term behavior with the opposite phase of oscillation.

Notice that the bipartition defined by the frontier at *0.69* is not valid for the whole range *[0,1]* of possible initial conditions, but only for those initial conditions inside the segment *AB* or in its neighborhood. If, for example, in the process of decreasing the value of x_0, we go lower and lower, beyond the value *B = 0.50* and more, we find that for values of x_0 around *0.31*, the basin of attraction for *B* ends, and after that, the basin of attraction for *A* restarts, even though we are working with values of x_0 that are closer to *B*. This "switching back" of the phase of cycling is illustrated through the part **(c)** of Figure 7, where we have an experiment with $x_0 = 0.30$ leading to the same phase of oscillation as an experiment with $x_0 = A = 0.81$ does, although $x_0 = 0.30$ is clearly much closer to *B = 0.50* than it is to *A = 0.81*. This secondary boundary, *0.31*, corresponds to a value of x_0 that after one iteration of the logistic map produces $x_1 = 0.69$, i.e., the value that corresponds to the main boundary discussed above. This explains the fact that initial conditions right above *0.31* and initial conditions right bellow *0.31* generate two opposite phases of period-2 oscillation in the long term, defining thus a new frontier between the two basins of attractions.

If we study the attraction basins in an exhaustive manner, i.e., exploring all the values of x_0 in the range *[0,1]* and observing what are the emerging long term patterns of oscillation, as we have done for the three cases of x_0 represented in Figure 7, we observe the following features regarding the sets of values of x_0 that compose the two attractors under study [13]:

- These two sets are not symmetric with respect to the coding values associated to binary *zeros* and binary *ones* (the values **A** and **B** discussed in Sections II and III). Notice for example that the main boundary at *0.69* is not exactly equal to *(A+B)/2*.
- In addition, they do not correspond to two convex sets, having a fragmented structure. Initial conditions closer to *A* can be part of the attraction region associated to *B* and vice versa, in an interleaved structure of both basins.

Figure 8 shows a diagram illustrating this fragmented nature of the basins of attraction that we observe in this process of mapping the attractors related to *A* and *B*. One of these basins is labeled here "basin for *A*", meaning that this basin is associated to the phase of oscillation which emerges for experiments with $x_0 = A$. Similarly, we have another basin associated to the other phase of period-2 oscillation, labeled "basin for *B*", and associated to the same phase of oscillation that emerges with $x_0 = B$.

In fact we can verify that we have a fractal structure in the fragmented sets that define the basins of attraction for *A* and *B*, with both sets repeatedly interleaving with each other, as we depart more and more from *A* upwards to *1*, as well as when we depart more and more from *B* downwards to *0*. This
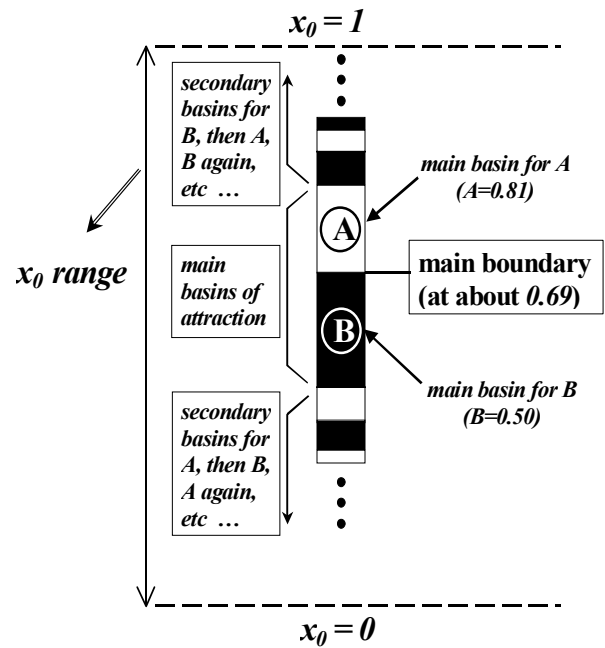

Fig. 8. Mapping the two basins of attraction associated to the long term cyclic trajectory for x_n in the logistic map. Depending on the value of the initial condition x_0, we have the state variable of the recursive map settling in one of two possible phases of oscillation. These two basins are identified in the above diagram through the alternating bands with wither and darker patterns.

fractal structure is confirmed by repeated experiments in which the step used for the sweep of the x_0 in the experiments is reduced more and more, as we approach the limit values $x_0 = 1$ and $x_0 = 0$, which are accumulation points for the repeated boundaries between the two basins. In particular, as we approach the $x_0 = 0$, one of these two accumulation points for the repeated boundaries among the alternating bands of the basin for *A* and the basin for *B*, we can verify that the values x_0 for consecutive boundaries between the bands is always scaled down by a factor of *3.25*. Notice that this value of scaling for the location of successive

boundaries is exactly the same as the value of the bifurcation parameter of the logistic map that we are using (i.e., $p = 3.25$).

The main conclusion of this analysis of the two basins of attraction in the logistic RPE is that allowing the initial conditions to go beyond a large distance from the segment **AB** will result in losing the control of the phase of cycling in the long term behavior of the period-2 attractor, what is not good for the operation of the associative RPEs network, since it relies on that phase of cycling for the coding of information. Of course, the phenomena of basins of attraction in a population of coupled RPE elements is much more complex than what we have in the isolated logistic map. Nevertheless, the fact that the basins of attraction of the isolated node are fragmented (not composed by convex sets) is a good indication for possible similar phenomena happening at the global scale. As a consequence of this reasoning, we decided to do experiments and to adopt design strategies in which the prompting the RPE nodes with values in the "fuzzy" parts of the basins of attraction is avoided.

V. APPLYING "SOFT CLAMPING" FOR IMPROVED PATTERN RECOVERY ABILITY

The analysis of the previous section indicates the convenience of using some kind of clamping pre-processing on the noisy initial conditions $x_{i,0}$, so to restrict them to a range of values in which the fractal forms of the basins of attraction discussed above would not be present. In this way, we can restrict the setting of initial conditions to convex sub-regions of attraction that are associated to the representation of **zeros** and **ones**. This clamping is in fact relatively simple to implement, since all the fragmented regions of the two sets are clearly apart from the segment **AB**.

With the clamping pre-processing, also named "soft clamping pre-processing" as justified ahead, we are able to improve the results for analog noise environments in a large extent, as it can be verified by producing a new Error Plot as the one previously presented in Figure 5, but this time for experiments involving networks operating with the soft clamping in their initial conditions. This important improvement result is presented in Figure 9, where we have the contrast between two Error Plots, showing that with the soft clamping technique (lower curve in the Figure), the Hamming error in the recovery of patterns under high levels of noise is reduced by factors of 10 or more, taking as reference the output errors obtained in the previous results presented in the Error Plot of Figure 5, where no clamping pre-processing was applied. Such a previous Error Plot is also reproduced here in Figure 9 for comparison, in its upper curve.

Notice that the proposed mechanism of clamping does not try to fix the values of the initial conditions exactly at A and

B through simple discretization of x_0 at one of these two values. This would in fact correspond to artificially remove a large part of the analog noise, what is something not feasible in a number of realistic implementations, such as in electronic hardware for example, where technological difficulties for the imposition of precise values of the $x_{i,0}$ are usually present [14,15]. On the contrary, the clamping of the initial conditions here discussed can be done in a relatively liberal way, what justifies thus the name that we have chosen for it: "soft clamping". This can be confirmed through a series of experiments with different lengths for the ranges of allowed x_0 values, i.e., experiments involving different limits for the clamping of the initial conditions. Table II shows the results of the output Hamming error for a fixed level of prompting analog noise (noise factor **0.5**), and 8 different sizes of clamping ranges. The extension of the range [**lower_clamp, upper_clamp**] in these experiments goes from values as small as **0.2** times |**A-B**| to values as large as **1.6** times |**A-B**|, as listed in the first row of Table II.
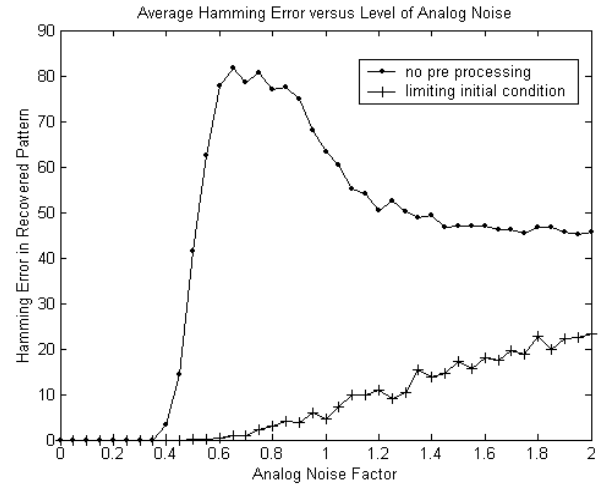

Fig. 9. Contrast involving two Error Plots, for arbitrary levels of noise in the prompting conditions. The horizontal axis shows the analog noise factor at the input (<|**noise**|> / |**A-B**|). The vertical axis shows the recovery error. The contrast of the two plots shows the performance improvement brought by the soft clamping of the initial conditions (lower Error Plot among the two) with respect to the performance obtained for associative networks without any control on the noisy initial conditions (upper Error Plot, already presented before in Figure 5).

Table II – Hamming Recovery Error (second row) *versus* normalized clamping range for analog noise factor **0.5**.

range / |A-B|	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6
Hamming error	35%	2%	2%	2%	1%	1%	1%	20%

In the first and last columns of Table II, where the output error goes up to *35%* and *20%*, we are clearly dealing with extreme cases, in which the clamping is not producing any help. The first of these extremes, i.e., *range = 0.2 |A-B|*, corresponds to experiments were we are excessively reducing the imprinting of the initial conditions by the prompting pattern. The other extreme, *range = 1.6 |A-B|*, corresponds to experiments in which we are not clamping enough the initial conditions so to make the coupled RPEs avoid the fragmented parts of the relevant basins of attraction.

On the other hand, if we exclude the extremes of the table and we only look at values of the ratio *(range / |A-B|)* from *0.4* to *1.4*, the results of the clamping pre-processing are excellent, all of them with performance numbers which are similar to the ones represented in the lower curve of the contrastive graph presented in the Figure 9. Therefore, the proposed mechanism of clamping, or perhaps more appropriately named "soft clamping", due to the wide flexibility on the definition of the clamping boundaries, only imposes that the values of $x_{i,0}$ are set inside a reasonable generous range of values, of length approximately *|A-B| ± 40%*, thus including the segment *AB* and a large part of its surroundings. Notice that this wide tolerance to the input noise certainly gives us the possibility of modeling with comfort a large variety of situations of imprecise definition of the initial conditions $x_{i,0}$.

VI. CONCLUSIONS

An important result of the analysis and experimental studies here presented was that we could define simple modifications in the mode of operation of associative RPEs networks under noisy initial conditions which resulted in a significant performance improvement. By "soft clamping" the noisy analog initial condition, having, for example, as the lower limit *0.44* and as the upper limit *0.86*, we can increase the average performance of the system when operating with high levels of analog noise by a factor of *12*, with the average Hamming error in the recovered pattern dropping from *63* % to *5* % – this is the case for analog noise factor *1.0*. The feature of the fragmented basins of attraction of the recursive nodes was identified as an important source of network degradation. The detection of the phenomenon and the definition of clamping regions so to create clean convex regions for the attractor for the binary *zero* limit cycle and the attractor for the binary *one* limit cycle was enough to obtain a very useful result, i.e., the improvement of the immunity to analog noise in large amounts. The final result of this very simple mechanism was a dramatic improvement in the associative network performance.

REFERENCES

[1] Nabil Hassan Farhat, S-Y Lin and Mostafa Eldelfrawy, "Complexity and chaotic dynamics in spiking neuron embodiment", SPIE Critical Review, vol. CR55, pp. 77-88, SPIE, Bellingham, Washington, 1994.
[2] DeLiang Wang, "A comparison of CNN and LEGION networks", IJCNN 2004, Budapest, 2004.
[3] R Kozma and Walter J. Freeman. "Encoding and recall of noisy data as chaotic spatio-temporal memory patterns in the style of the brains", IJCNN 2000, Como, Italy, 2000.
[4] Jose C. Principe, V.G. Tavares, J.G. Harris, W.J. Freeman, "Design and implementation of a biologically realistic olfactory cortex in analog VLSI", Proceedings of the IEEE, vol. 89, pp. 1030 – 1051, 2001.
[5] Kunihiko Kaneko and Ichiro Tsuda, "Complex Systems: Chaos and Beyond.", Springer-Verlag, 2001.
[6] E. D.-M.-Hernandez, "Studying neural networks of bifurcating recursive processing elements – quantitative methods for architecture design and performance analysis", Lecture Notes in Comp. Science, Springer, vol. 2084, p.546, 2001.
[7] E Del Moral Hernandez, L. A. da Silva, "A new hybrid neural architecture (MLP+RPE) for hetero association: multi layer perceptron and coupled recursive processing elements neural networks", IJCNN 2004, Budapest, 2004.
[8] E. Del-Moral-Hernandez, "Neural networks with chaotic recursive nodes: techniques for the design of associative memories, contrast with Hopfield architectures, and extensions for time-dependent inputs", Neural Networks, Elsevier, vol. 16, pp. 675–682, 2003.
[9] E. Del-Moral-Hernandez, "Non-Homogenous Neural Networks with Chaotic Recursive Nodes: Connectivity and Multi-assemblies Structures in Recursive Processing Elements Architectures", Neural Networks, Elsevier, vol. 18, pp. 532–540, 2005.
[10] Simon Haykin, "Neural networks: a comprehensive foundation", 2nd ed, Prentice Hall, Upper Saddle River, NJ, 1999.
[11] J. J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", in Proc. Nat. Acad. of Sciences, USA, Ap. 1982, vol. 79, p.2554-58.
[12] Nabil Farhat and E. Del-Moral-Hernandez, "Logistic Networks With DNA-Like Encoding and Interactions", Lecture Notes in Comp. Science, Springer-Verlag, vol. 930, pp.214-222, 1995.
[13] Robert L. Devaney, "An Introduction to Chaotic Dynamical Systems", 2nd ed, Addison-Wesley, Redwood City, CA, 1989.
[14] E. Del-Moral-Hernandez, Gee-Hyuk Lee and Nabil Farhat, "Analog Realization of Arbitrary One Dimensional Maps", IEEE Trans. on Circuits and Systems I: Fundamental Theory and Applications, vol. 50, pp. 1538 – 1547, 2003.
[15] A. Farfan-Pelaez, E Del-Moral-Hernandez, J. Navarro-Soares Jr., and W. Van Noije, "A CMOS Implementation of the Sine-circle Map", 48th IEEE Int'l Midwest Symposium on Circuits & Systems, Cincinnatti, 2005.