A mathematical approach on memory capacity of a simple synapses model
Pascal Helson

To cite this version:
Pascal Helson. A mathematical approach on memory capacity of a simple synapses model. International Conference on Mathematical NeuroScience (ICMNS), Jun 2018, Antibes, France. hal-01957292

HAL Id: hal-01957292
https://hal.archives-ouvertes.fr/hal-01957292
Submitted on 17 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Network models of Memory: Capacity of neural networks in memorizing external inputs is a complex problem which has given rise to numerous research. It is widely accepted that memory sits where communication between two neurons takes place, in synapses [1]. It involves a huge number of chemical reactions, cascades, ions flow, proteins states and even more mechanisms, which makes it really complex. Such a complexity stresses the need of simplifying models which is done in network models of memory.

Problem: Most of these models don’t take into account both synaptic plasticity and neural dynamic. Adding dynamics on the weights makes the analysis more difficult which explains why most models consider either a neural [2,3, 4] or a synaptic weight dynamic [5, 6, 7, 8]. We decided to study the binary synaptic model of [9], model we wish to complete with a neural network afterwards in order to get closer to biology.

Purpose: Propose a rigorous mathematical approach of the model of [9] as part of a more general framework which is to be have a general mathematical framework adapted to many models of memory.

3. Amit-Fusi Model [9]

Discrete time model with two coupled binary processes, stimuli $(q_{t}\in\{0,1\})$ and synaptic weight matrix $(W_{t+1}\in\{0,1\})^{N}\times N$.

Stimuli: $(q_{t}\in\{0,1\})$ i.i.d. random vector $\sim Bernoulli(\theta)$. $\theta$ is known.

Synaptic weights dynamics:

- $(W_{t+1,r}\neq 0 )$ and at each step a new stimulus $q_{t}$ is received by the network.
- The components $J_{r}^{+}$ and $J_{r}^{-}$ jump as follow:
  - If $(q_{t},J_{r}) = (1, J_{r}^{+})$ then $W_{t+1,r} = 0$ with probability $q_{t}^{1}$; 
  - If $(q_{t},J_{r}) = (0, J_{r}^{+})$ then $W_{t+1,r} = 1$ with probability $q_{t}^{0}$;
  - If $(q_{t},J_{r}) = (1, J_{r}^{+})$ then $W_{t+1,r} = 1$ with probability $q_{t}^{0}$;
  - If $(q_{t},J_{r}) = (0, J_{r}^{+})$ then $W_{t+1,r} = 0$.

Remark: The initial condition is not defined on synaptic weights but on the synaptic input into neurons defined as follows.

Synaptic input into neurons:

$$ h_{r}(t) = \sum_{i=0}^{N} W_{i,r} q_{i,t} $$

In the following, we are interested in the laws of $(h_{r}(t)\sim 0)$ and $(h_{r}(t)\sim 1)$, respectively called $p_{r,\theta}^{0}$ and $p_{r,\theta}^{1}$.

Remark: The state space of $h_{r}$ depends on the size of $r$, $K = \sum_{r} c_{r}^{j}$. Moreover, as neurons are simply use notation $h_{r,K}^{j}$. Finally, it is easy to show $p_{r,K}^{j}$ converges to a unique $p_{r,K}$.

Initial condition:

Initially, synaptic input follows the stationary distribution $\pi_{r,K}^{j} = p_{r,K}^{j}$.

5. First Results

Mathematical results:

As we have the general forms of $p_{r,K}$ and $p_{r}^{j}$ which is hard, we first studied the spectrum of the transition matrix $M_{K,r}$ of $(h_{r}(t))$ and got a first result.

Proposition 1 The spectrum of the transition matrix $M_{K} , h_{r}$ and the one of $(q_{t},J_{r})$ is the following:

$$ \text{M}_{K,r} = (1 - f) \sigma_{K}^{t} + \frac{1}{2} \left( (1-f) \sigma_{K}^{t} - \frac{1}{2} \right) \sigma_{K}^{t} , \quad 0 \leq K \leq \frac{N}{2} $$

$$ \text{M}_{K,r} = (1 - f) \sigma_{K}^{t} + \left( \frac{1}{2} \left( (1-f) \sigma_{K}^{t} - \frac{1}{2} \right) \sigma_{K}^{t} \right) , \quad \text{multiplicity} \quad \sigma_{K}^{t} $$

In fact, the spectrum is linked to the speed at which stimuli are forgotten. However, the slower this speed is, the less plastic the network is. It is a classical compromise in optimizing storage capacity.

Sketch of the proof for $\text{M}_{K,r}$(9):

We can write $M_{K,r} = [\Pi_{r, K, j}^{i} ]$ as a matrix by block with $p_{r,K}^{j} = \Pi_{r, K, j}$ and $M_{r}^{j}$ the probability matrix of $(q_{t},J_{r})^{j}$, knowing that $q_{t} = \xi_{t}$.

$$ M_{K,r} = \sum_{j=0}^{K} \Pi_{r, K, j}^{i} $$

It is not difficult to show

$$ \text{M}_{K,r} = \sum_{i} \Pi_{r, K, j}^{i} \times \Pi_{r, K, j}^{i} $$

In particular, if $q_{t}$ is an invariant measure of the process with matrix transition $M_{K,r}$, $S_{K,r} = \sigma_{K}^{t}$ then $p_{r,K}^{j} = \Pi_{r, K, j}^{i} \times \Pi_{r, K, j}^{i}$. One can then compute $M_{r}$ from the following 2 x 2 matrices:

$$ M_{0} = \Pi_{r, 0}^{0} = \Pi_{r, 0}^{1} = \Pi_{r, 1}^{0} = 0 , \quad M_{1} = 1 - q_{t}^{1} $$

Then, using Kneercke product properties, we have the lemma.

6. Conclusion

When previous studies have considered small coding level $f$, in order to get results on the storage capacity depending on $N$, such an assumption seems to make the model loose its initial interest in the correlations between synapses. Our approach aims at studying the synaptic input into a neuron taking correlations into account through a decision rule. It enabled us to get a first result on the speed of forgetting-stimulus thanks to a spectral analysis.

7. Perspectives

- Show $p_{r,K}^{j}$ converges to $\pi_{r,K}^{j}$ when $q_{t}$ are small
- Control the probability of error thanks to the study in the independent case $f_{s}, f_{g}$ small
- A study of the probability of error taking into account all the vector $h_{r}
- Add neural dynamics as a feedback to maintain weight structures longer and enhance storage capacity

References:


Many methods have been used to study the storage capacity of network models. The more intuitive is maybe to use stimulus to be learned as attractors of a neural dynamic [3]; the maximal number of attractors would then be the memory capacity of the model. Signal to Noise Ratio (SNR) analysis [8, 9] and mean first passage time to a threshold [7] have also been proposed. The underlying idea of those methods is that the neural dynamic is ruled by a threshold on the synaptic input: a linear decision rule. In our case, we don’t impose such a rule. Our retrieval criteria holds on the knowledge of the two distributions $p_{r,K}$ and $p_{r}$.

Decision rule [10]:

At fixed $K$, we aim at studying the minimal probability of error assuming the neuron $t$ knows $p_{r,t}$ and observes $h_{r}^{s}$. As $p_{r,\theta}^{0}$ and $p_{r,\theta}^{1}$ converges to $p_{r,\theta}$, the error increases with time as distributions get closer.