Synaptic value bounds for optimizing retrieval in recurrent neural networks

Daniel D. Ben Dayan Rubin

Department of Bioengineering, Polytechnic University of Milan,
P.zza Leonardo da Vinci 32, Milan, Italy
e-mail: **dbd@biomed.polimi.it**

Abstract. We present an analysis of the range of values of synaptical connections to enhance the storage proprieties of neural networks. Consider a random connected system to which random patterns are shown; these patterns impose specific activities over neuron pairs that might be connected evoking long term synapse modifications. Two approaches are given. The first one focuses on the noise context and the second one focuses on the learning rule. We find that increasing the variability among the synapse values within a given range, both the quality and the speed retrieval increase.

1 Introduction

In order to achieve experience, it is essential that each developing biological system, at the first stage of its development, should own the possibility of recognizing, indiscriminately, a very large number of patterns. Most neural circuits should be organized according to two mechanisms that bring to the formation of the neural maps: *anatomical*, in the formation of the *neural substrate*, and *synaptical*, in the formation of a *specified connectivity*. In certain stages of the development these two events are overlapped to create and to consolidate particular maps. In the neural substrate, different groups of cells should be able to perform the same functions better or worse. The specified connectivity is developed following to the modification of synaptical contacts, which are responsible for compensating the unfaithfulness of a distributed system. Yet, during the animal experience, the initial neural substrate may contain cellular groups that are functionally equivalent to or even more "efficient" than the ones already active in the modified maps [3]. We should evince from the stochastic nature of the neural substrate that the organization (that should reflect the initial condition of the developing brain) of a connected system gives the highest retrievable quality to a large number of patterns if the synapse values are taken within certain ranges. In order to prove this we will compute the ratio between the activity of a neuron which belongs to a pool of neurons that encode a specific pattern and any other neuron that does not participate to the encoding. This ratio is a measure of signal to noise ratio (SNR). Another measure for retrieval quality is also being given in terms of mutual information.

2 Context and initial conditions

In the following analysis we will consider sparse distributed representation in which a small proportion of neurons is active at any time. Suppose a pattern is shown to the system after the presentation of p consecutive different patterns; all these patterns belong to a set of statistically independent random patterns. In normal conditions of physiologically enabled synaptic dynamics, the obtained synaptic structure would be independent of the initial condition determined by the presentation of the very first (i.e. oldest) pattern. This condition, usually, depends on the speed of learning/forgetting implied in the synapse dynamics [4]. Let f be the fraction of the overall N neurons of a local module that encodes a specific stimulus (i.e. pattern); f measures the *sparseness* of the output firing pattern and will be referred to as *foreground* activity. For simplicity, we consider only the excitatory contribution since it is the only responsible for the emerging auto-associative characteristics of the network. The connection between a pair (i, j) within the network is a random variable J with distribution $P(J)$. For simplicity suppose that each neuron is excitatory, $J_{ij} \geq 0$. We will investigate the stochastic properties of the synaptic matrix $\mathbf{J}$ in order to evince analytically the conditions to express the SNR and the measure of information as a function of the synaptic value distribution.

Information quantification

Consider the information conveyed by an autoassociative memory that can store a number of patterns. In retrieving pattern μ-th, the network produces a distinct firing pattern ξ. The similarity between ξ^μ and ξ can be measured by the average mutual information

$$\langle I(\xi^\mu, \xi) \rangle = \sum_{\xi^\mu, \xi} P(\xi^\mu, \xi) \log_2 P(\xi^\mu, \xi) / P(\xi^\mu) P(\xi), \tag{1}$$

where $P(\xi^\mu, \xi)$ is the joint probability of receiving ξ^μ and emitting ξ. Substituting ξ in Eq.1 with $\xi_i = f(\sum_j \xi_j^\mu J_{ij})$ and in vectorial notation $\xi_i = f(\xi^\mu \mathbf{J}_i)$ for the single i component, and $\xi = \{\xi_i\} = \{f(\xi^\mu \mathbf{J}_i)\}$ for the entire ξ pattern, where $f(.)$ is the activation function and $\mathbf{J}_i$ is the i-th column of the synaptic matrix $\mathbf{J}$ leads to:

$$\langle I(\xi^\mu, \xi) \rangle = \sum_{\xi^\mu} P(\xi^\mu, \{f(\xi^\mu \mathbf{J}_i)\}) \log_2 P(\xi^\mu, \{f(\xi^\mu \mathbf{J}_i)\}) / P(\xi^\mu) P(\{f(\xi^\mu \mathbf{J}_i)\});$$

taking:
$P(\xi^\mu, \xi) = P(\xi^\mu, \{f(\xi^\mu \mathbf{J}_i)\}) \propto P(\mathbf{J}, \xi^\mu)$, and for Bayes: $P(\mathbf{J}, \xi^\mu) = P(\mathbf{J}|\xi^\mu) P(\xi^\mu)$, we can write:

$$\langle \hat{I}(\xi^\mu, \xi) \rangle = \sum_{\xi^\mu} P(\mathbf{J}|\xi^\mu) P(\xi^\mu) \log_2 P(\mathbf{J}|\xi^\mu) P(\xi). \tag{2}$$

Hence, $\langle I(\xi^\mu, \xi) \rangle \propto \langle \hat{I}(\xi^\mu, \xi) \rangle$ is a measure of the mutual information dependent of ξ^μ and $\mathbf{J}$. We can investigate the stochastic properties of $\mathbf{J}$ in order to assert the information content of a neural network with dependence of the different connection strengths.

3 Stochastic implications of the synaptic values

To simplify the analysis we assume that the set of stable internal synaptic states is discrete. The synapse dynamics can be described by a random walk confined into two reflecting barriers, where the reflecting barriers are the saturation values (i.e. the maximal and the minimal value) of the synapses. This random walk is induced by the presentation of the sequence of the p uncorrelated stimuli. The random walk can be formalized as a particular kind of Markov process. Let $\mathcal{M}_{KJ}$ be the probability that a synapse makes a transition from the internal stable state K to the stable state J given $p(\xi_i, \xi_j)$, which is the probability that a stimulus imposes the activities ξ_i, ξ_j to the pair on neurons i, j (see e.g. Fusi 2002 [4]). The conditional distribution function $p^p(J|(\xi_i^1, \xi_j^1))$ that a synapse is in state J following the presentation of p patterns, the first of which imposed ξ_i^1, ξ_j^1 on the synapse, satisfies the equation:

$$p^p(J|(\xi_i^1, \xi_j^1)) = \sum_{K=1}^{n_s} p^1(K|(\xi_i^1, \xi_j^1))\mathcal{M}_{KJ}^{(p-1)},$$ (3)

where the superscript $(p-1)$ indicates the rising power factor. K runs over all the stable n_s synaptic states. (3) states that after the presentation of $p-1$ patterns we can still recall the p-th pattern which is, actually, the first pattern shown to the system. Since this kind of dynamics is ergodic, for a large number of presentations, the system will be independent of the first stimulus, and hence, it will forget the first pattern [6, 5].

Simplifying the scenario
Reducing the random walk among the discrete states to the extreme situation in which only two stable synapses are present, facilitates the analytical analysis of the problem. Simulations [4] and experimental data [2, 7] have shown evidences for such situation in long time scenarios. Accordingly, we will analyse $\mathcal{M}_{KJ}$ (a 2x2 matrix), given by:

$$\mathcal{M}_{KJ} = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix},$$ (4)

where $\alpha = q_+ f^2$ is the probability of a transition to the upper state given the probability of being in the lower state when the pre- and post- synaptic neurons are both active (i.e. f^2), and $\beta = q_- f(1-f)$ is the probability of a transition to the lower state given the probability of being in the upper state when the pre- and the post- synaptic neuron are respectively active and inactive or viceversa (i.e. $f(1-f)$). Eq.(4) takes into account the distribution of the synapses given the interference of the other patterns. This interference should be further analyzed by introducing the correlation that can arise between random patterns that share the same encoding neurons. In this simplified approach we will focus only on Eq.(4) to investigate the synapse distribution, after having shown $p-1$ patterns, independently of the specific activated fraction of neurons.

4 Searching for the solutions

In this paragraph we shall use two different quantification measures to test how the distribution of the synaptic values modifies the retrievability of the stored patterns and the information content.

SNR approach

SNR is a direct measure of the quality of retrieval of patterns embedded in noise background. The immediate quantification regards the mean spiking rate of a fraction of neurons among the entire system's rate. If the spike rate arriving to each channel is low, but in an interval τ the number of arriving spikes is high, due to the large number of input channels C, then the source of the depolarization is Gaussian. To obtain an idea concerning the neuron's output rate, we will consider the simplified treatment of the spike emission process [1] which is a good approximation of the full Integrate&Fire neuron when the spike rates are much lower than $1/\tau$. In each of the integration time intervals τ the depolarization is equal to the sum of unitary inputs J_i arriving in that intervals. If the synaptic input to the neuron in the τ interval has mean μ and STD σ and the threshold is θ, the probability that a spike is being emitted in the time interval τ is: $P(\nu) = \int_\theta^{\inf} dI/(\sqrt{2\pi}\sigma(\nu))e^{-\frac{(I-\mu(\nu))^2}{2\sigma(\nu)^2}}$; if the output rate is given by $\nu_{out} = P(\nu)/\tau$, the mean μ and the STD σ of the afferent current is: $\mu = JC\nu\tau; \sigma = J\sqrt{C\nu\tau(1+\Delta^2)}$, being the value of the synapse STD equal to Δ times J . Since the $p.d.f.$ of the spike emission is Gaussian, the STD of the firing rate distribution can be taken equal to the STD of the afferent current for each neuron. The ratio between the σ_s and σ_n is a measure of SNR, where the subscripts indicate s-ignal (i.e. all those neurons activated by the stimulus presentation and whose activity is in foreground with respect to the system activity) or n-oise. Although Δ_n is being modified by the learning process, in order to obtain Δ_s for the fraction f of neurons belonging to the specific pattern, in this simplified approach we will assume that learning has been already carried out and Δ_s is proportional to Δ_n. A more complete treatment of this dependence upon $\mathcal{M}_{KJ}$ is currently studied and it will be reported elsewhere. We shall plot (see Fig.1a) the SNR in dB: $20\text{Log}\frac{\sigma_s}{\sigma_n}$, where $\frac{\sigma_s}{\sigma_n} = \frac{\sqrt{\nu_s(1+\Delta_s^2)}}{\sqrt{\nu_n(1+\Delta_n^2)}}$ for different values of the ratios $\frac{\nu_s}{\nu_n} = \alpha$ and of $\frac{1+\Delta_s^2}{1+\Delta_n^2} = \beta$.

Information theory approach

Let us recall the issues discussed in sections **2** and **3**. The distribution of the potentiated synapses may deviate from the extreme case (4) (depicted in section **2**), and be distributed around the potentiated (i.e. J_+) value and the depressed value (i.e. 0). This case is practically what we expected from the distribution of a sum of independent random variables[1] as are the synapse values. For this reason we may consider the distribution as a bi-Normal truncated distribution given by: $c_0 = I_{(0, \ J_+)}(x)(\phi_{0,\sigma^2}(x) + \phi_{J_+,\sigma^2}(x))/(2\Phi(J + /\sigma) - 1)$, where ϕ and Φ are

---

[1] i.e. law of large numbers and the central limit theorem
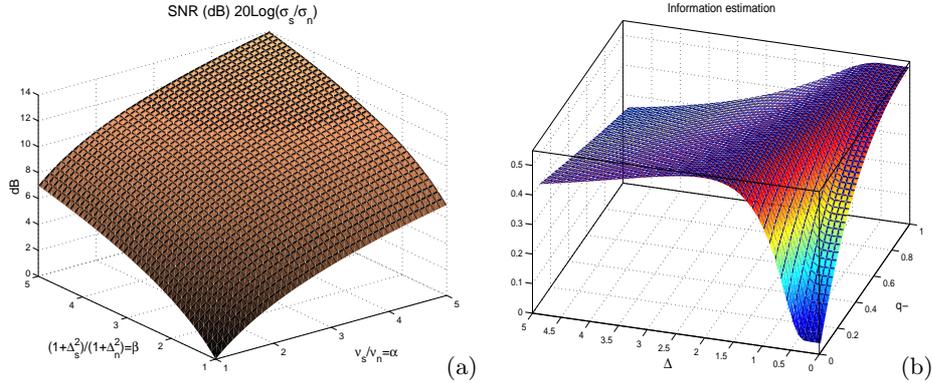
Fig. 1. (**a**) SNR (dB) over the z-axis; α the ratio between the signal spiking rate and the noise spiking rate over the x-axis; while over the y-axis, β a measure of the synapse variance ratio between the foreground neurons and the background neurons. The function is monotonically increasing for increasing values of the independent variables over the x, and y axes. (**b**) Average mutual information, which measures the quality of the retrieval varying the STD (Δ) of the synapses (x-axis) and the probability of depression q_- (y-axis) of synapses which bond background neurons not included in the stimulus retrieval. The surface is obtained for f=1% and p=100

respectively a Gaussian probability distribution function and its corresponding cumulative function. If the probability to have connectivity is given by C_0, the probability to have no connected synapses is $1 - C_0$.

We search for the probability to retrieve the first pattern given $p-1$ patterns already shown. Substituting $\mathrm{p}^1(K|(\xi_i^1, \xi_j^1))$ in (3) with $[1-C_0 \quad C_0]$ which is the vector of the initial probabilities, considering $\mathcal{M}_{KJ}$ in (4), being $\lambda = 1 - \alpha - \beta$ the smallest eigen-value of (4), and after some mathematics, we obtain:

$$\mathrm{p}(\mathbf{J}|\xi^1) = \pi_0 + \lambda^{p-1}(C_0 - \pi_0), \tag{5}$$

where $\pi_0 = \alpha/(\alpha + \beta) = q_+/(q_+ + q_-(\frac{1}{f} - 1))$ is the asymptotical distribution as λ^{p-1} fades to zero as p increases, and where λ is the smallest eigenvalue. Substituting (5) in (2) and considering that since ξ^1 is every possible pattern within the set of $p - 1$ patterns, its probability is $\frac{1}{p-1}$; in the same way we can take $\mathrm{P}(\xi)$=f. Since $\lambda \sim 1 - fq_-$, since $\mathrm{p}(\mathbf{J}|\xi^1)$ does not significantly depend on variations of q_+ in the range of [0 1], and if f=1% and p=100, in Fig.1b $\langle \hat{\mathrm{I}}(\xi^\mu, \xi) \rangle$ is plotted against Δ over the x-axis and q_- over the y-axis.

5   Discussion and Conclusion

The SNR analysis and the information theory approach may reveal different situations upon which the retrieval of a pattern can depend. The SNR quantification enables investigation of the influence of the background noise firing rate on the

retrieval capability of the system. It was shown that, even when the firing rates of the foreground and of the background neurons are equal, the pattern can however be retrieved if the variance of the synaptic connections, within the fraction of the neurons participating to retrieval of the pattern, is increased. Moreover, the information quantification approach revealed the dependency upon the biological compatible values of the probability to have long term modifications. However, this demonstrates that for increasing values of the variance of the synaptic connections within a specific fraction of pattern-stimulated neurons, information increases within a specific range. The mutual information function in Fig.1b clearly shows an upper bound to which corresponds the highest quality retrieval. It is interesting to note that for values of $\Delta \sim 0$, in order to have maximal information content it is important to have high probability of depressing the synapses that connect the background to the foreground neurons and viceversa. In this case only strong connections within the foreground neurons and weak connections between the background and the foreground neurons can elicit high quality retrieval of the specific pattern; in this case we fall back into the bistable synapse assumption.

References

[1] DJ Amit and N Brunel, *Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex*, Cerebral Cortex **7** (1997), 237–252.

[2] TVP Bliss and GL Collingridge, *A synaptic model of memory: long term potentiation in the hippocampus*, Nature **361** (1993), 31–39.

[3] GM Edelman, *Group selection as the basis for higher brain function, in The Organization of the Cerebral Cortex*, Schmitt F.O., Worden F.G., Adelman G., Dennis S.G. eds., Cambridge, Mass. MIT Press, London, 1981.

[4] S Fusi, *Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates*, Biological Cybernetics **17** (2002), 305–317.

[5] JP Nadal, G Toulouse, JP Changeux, and S Dehaene, *Network of formal neurons and memory palimpsests*, Europhys. Lett. **1** (1986), 535–542.

[6] G Parisi, *A memory which forgets*, J. Phys, A: Math Gen **19** (1986), 617–619.

[7] CCH Petersen, RC Malenka, RA Nicoll, and JJ Hopfield, *All-or-none potentiation at CA3-CA1 synapses*, Proc. Natl. Acad. Sci **95** (1998), 4732–4737.