Noise based memories

Daniel Ben Dayan Rubin[1], **Stefano Fusi**[1,2]

[1]Institute of Neuroinformatics, UNI ETH, Zurich, [2]Center for Theoretical Neuroscience, Columbia University, New York

Long term memories are widely believed to be stored in the pattern of synaptic couplings. In order to retrieve a memory the synapses are probed by pre-synaptic activity, which, weighted by the synapses, generates a current driving the post-synaptic neuron. By looking at the activity of post-synaptic neurons, it is possible to read out the information stored in the synaptic couplings. In classical neural network models (e.g. the Hopfield model), the mnemonic trace is assumed to be in the average value of the total synaptic currents. Usually neurons partition in two or more groups, each group being characterized by a different average synaptic current, and hence by a different mean activity. For example certain neurons will be considered ‘active’ because their activity is above some threshold, and others will be ‘inactive’ in response to the input which triggers the process of memory retrieval. The pattern of activities of all neurons represents a readout of the information stored in the synaptic efficacies. Here we investigate whether the variance of the synaptic currents across different neurons can also carry information about memories. Previous attempts [1] considered the efficiency of encoding information in the second order statistics of population responses. We study the case of on-line learning in binary synapses which is known [2] to be difficult because the mnemonic trace contained in the average synaptic currents decays exponentially fast. We show that there is a component of the variance which depends on the correlations between synapses which are on the same dendritic tree. Such a correlation is always present, even in the case of random uncorrelated pre- and post-synaptic patterns of activity. This component of the variance contains a mnemonic trace which also decays exponentially, at the same rate as the mnemonic trace contained in the average synaptic currents. Such a mnemonic trace can also be harnessed to retrieve information about past memories.

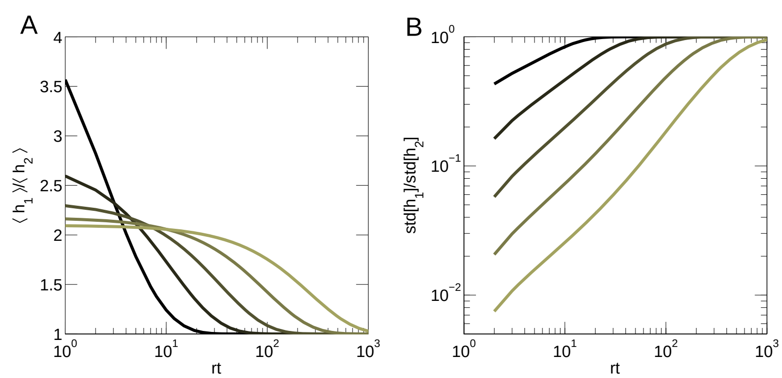

Figure: (**A**) Ratio between the mean synaptic currents (named h_1 and h_2) of two groups of neurons whose activity should be different when the specific memory that we are tracking is retrieved. As soon as $h_1 = h_2$, the memory is forgotten. h_1/h_2 is plotted against the number of new experiences which overwrite the old ones and disrupt the memory (we assume that there is constant rate r of new experiences, t is time). The tracked memory is stored at time zero. Different curves correspond to different statistics of the input patterns (sparse random uncorrelated patterns with an average fraction f of active neurons). (**B**) Dynamics of the ratio of the standard deviation of the two distributions h_1 and h_2. The memory is not forgotten as long as there are detectable differences between the two standard deviations. Lighter lines correspond to decreasing coding levels: $f = \{0.30, 0.19, 0.11, 0.07, 0.04\}$.

Acknowledgments We thank W. Bialek and L.F. Abbott for interesting discussions. Supported by the SNF grant PP00A-106556 and by NIH-2R01 MH58754.

Refs [1] M. Shamir, H. Sompolinsky, Neural Computation 2004, [2] D.J. Amit, S. Fusi, Neural Computation, 1994