Clique encoding in recurrent neural networks

Emanuele Varriale, Claudius Gros

Institute for Theoretical Physics, Goethe University, Frankfurt am Main, Germany

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Introduction

The internal brain activity is only modulated, not driven, by sensory input [1].
Therefore external sensory stimuli interact with an autonomously active network, in a way that allows semantic learning.
A network with competing neuronal assemblies can show a transient state dynamics, i.e. an infinite time series of meta-stable attractor states.
We propose a learning rule that correlates such attractor states with sensory inputs from the bars and stripes problem, expanding from [2].

Network architecture

With the right network architecture it can be easy to pinpoint attractor states.
We use networks as the shown in Fig. 1, which is formed by cliques, i.e. complete sub-graph, with excitatory connections within cliques and mostly inhibitory ones across cliques, so that cliques have a *competitive dynamics*.
Because of the network topology, the winning clique inhibits every other, so that this is a stable state, as long as it is active.

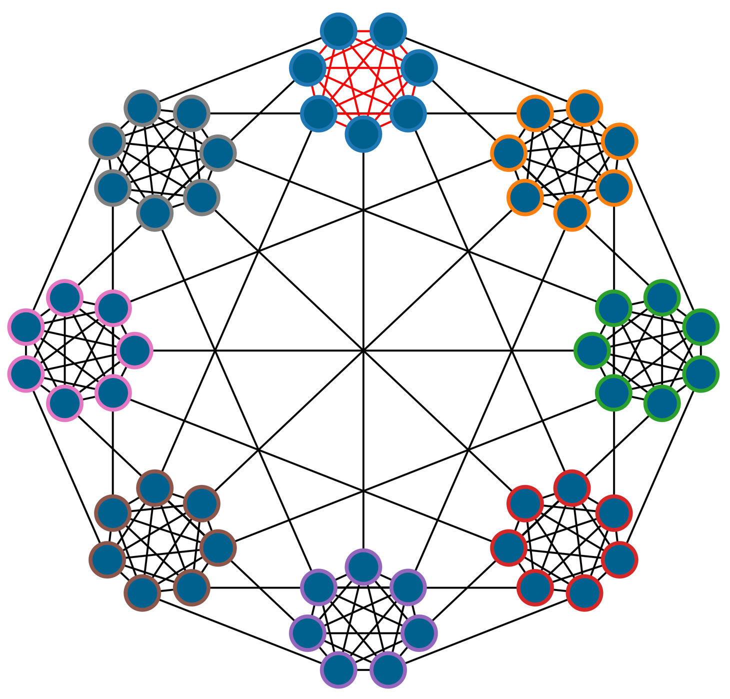


Figure: An example of our network architecture.

Neuronal dynamics

We use a rate-encoding model, in which each neuron has a membrane potential x_j, a sigmoidal activation function σ and receives excitatory input E_j and inhibitory I_j:

$$\tau_x \dot{x}_j = -x_j + E_j + I_j$$
$$y_j = \sigma(x_k) = \frac{1}{1+\exp(-a x_k)}$$
$$E_j = \sum_k w_{jk} y_k$$
$$I_j = \sum_k z_{jk} \tilde{y}_k$$

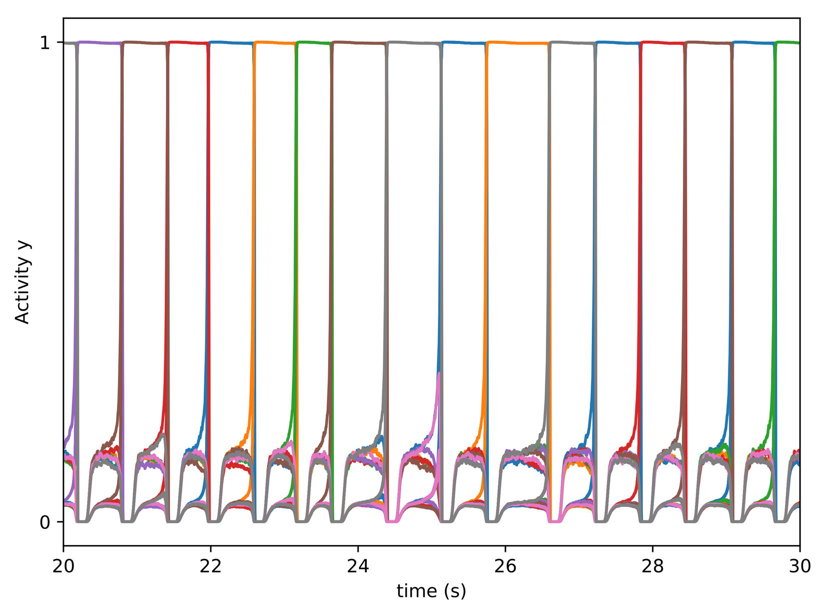


Figure: Internal activity in the network shown above.

Full depletion model

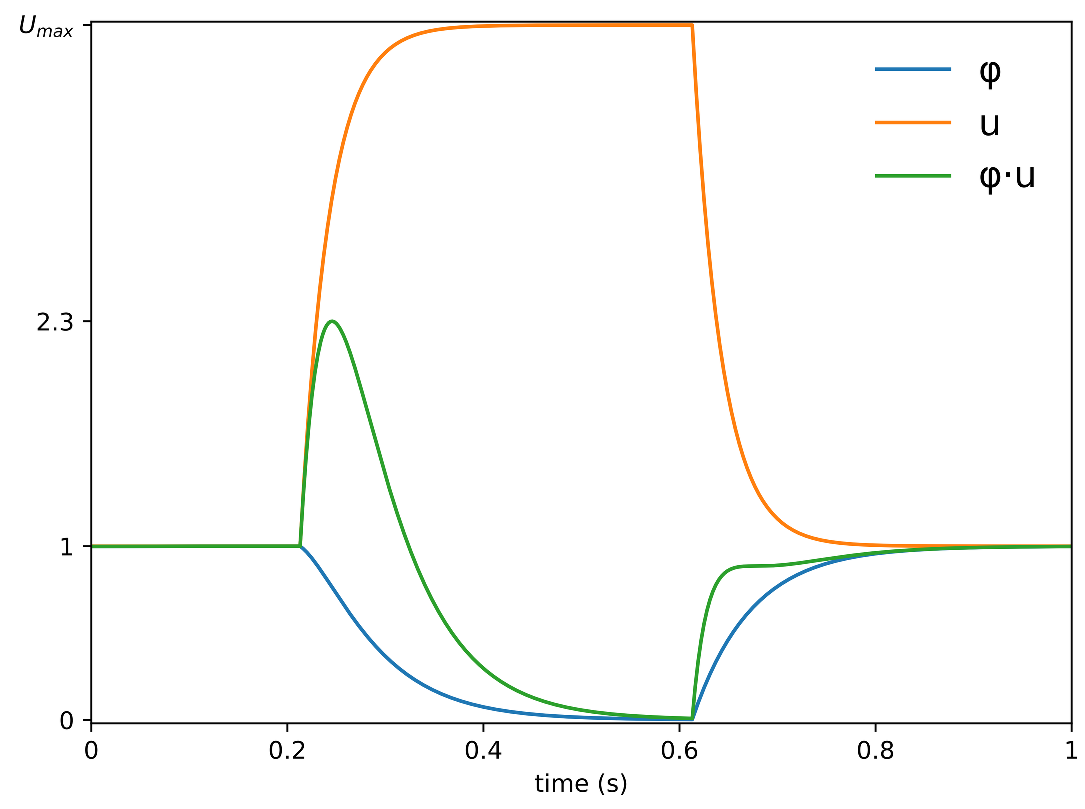
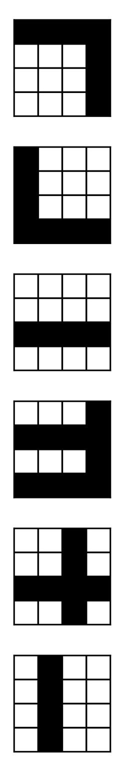


Figure: Dynamics of the full depletion model, with high presynaptic activity.

To enforce the transient state dynamics, the winning clique state has to become unstable after some time.
This is achieved by lowering inhibitory signals in case of sustained activity, through the *full depletion model*, that is inspired by the Tsodyks-Markram model [3].
The pre-synaptic activity $\tilde{y}_j$ is proportional to the somatic activity y_j, through two terms u_j and φ_j.
u_j represents the likelihood of neurotransmitter release, which increases with high inputs, while φ_j represents the concentration of available vesicles, that decreases while firing, according to the following dynamics:

$$\tilde{y}_j = y_j u_j \varphi_j$$
$$\dot{u}_j = \frac{U_y - u}{T_u}, \quad U_y = 1 + (U_{\text{max}} - 1)\, y_j$$
$$\dot{\varphi}_j = \frac{\Phi_u - \varphi}{T_\varphi}, \quad \Phi_u = 1 - \frac{u y_j}{U_{\text{max}}}$$

Sensory input

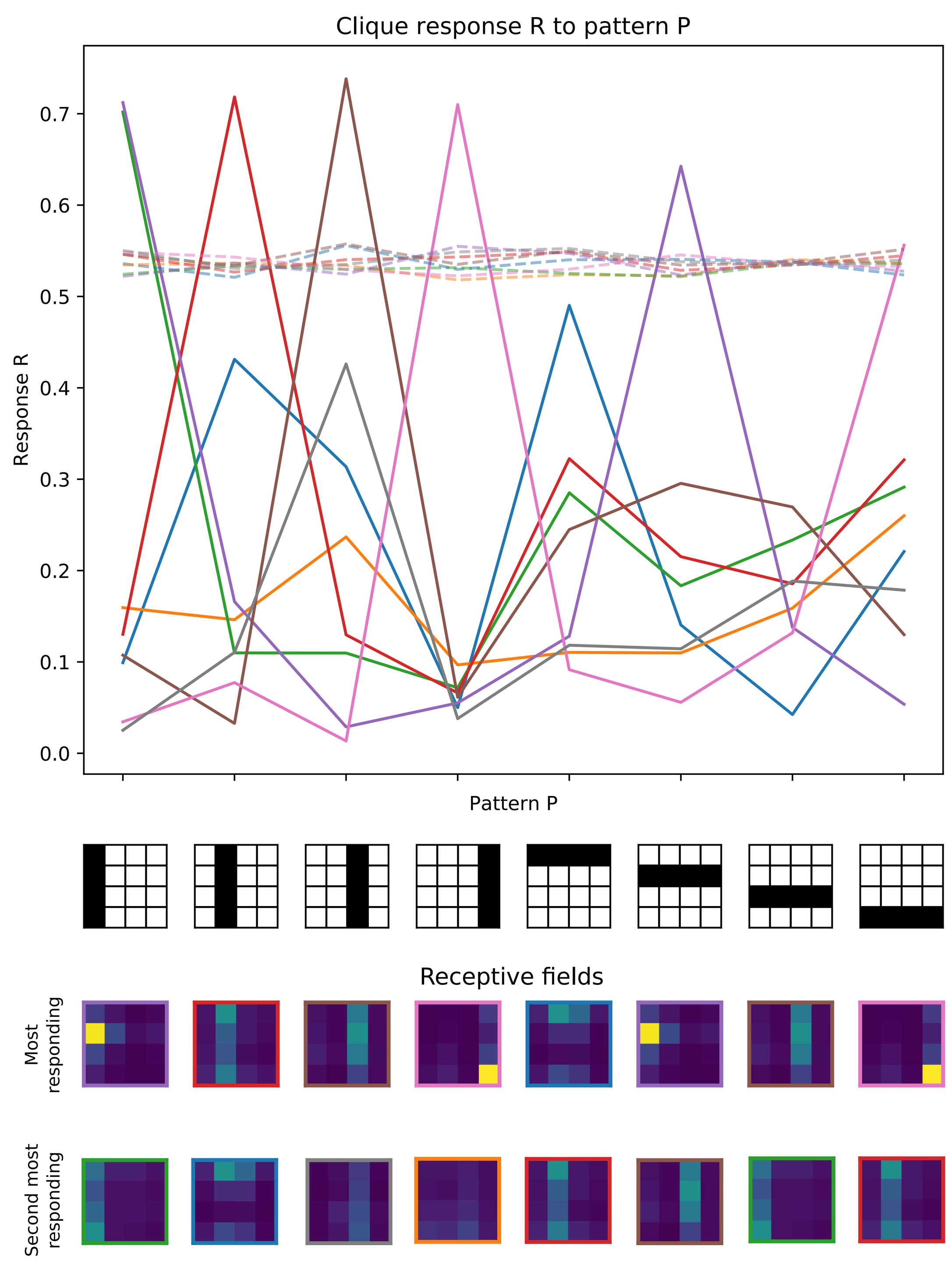


We want our network to extract information from an external environment via signal separation and feature extraction. An example would be to identify recurring patterns and to separate them, even if they occur simultaneously.
We use the bars problem, in which horizontal and vertical bars are presented on a retina of L x L pixels. Bars are independently drawn so that there is, on average, one bar per input. Active bar pixels have $y_l^{\text{ext}} = 1$, and inactive ones $y_l^{\text{ext}} = 0$, as shown in the Figure.
The bar intersection pixels also have value 1, so that a non linear independent component analysis has to be performed in order to separate bars.
This sensory input layer is connected to the network with excitatory connections v_{jl} so that an extra input term ΔE_j is added:

$$\Delta E_j = \sum_{l=1}^{L^2} v_{jl} y_l^{\text{ext}}$$

The initial weights are chosen so that, on average, $\Delta E_j \approx 0.5$. With these weights external input can, on average, change the internal dynamics only when the short-term plasticity tunes down inhibitory weights.

Learning rule



The learning procedure has to map different bars onto different cliques, by taking advantage of the fact that the sensory input can "steer" the competing dynamics, thus changing the order of winning cliques.
We employ the following learning rule:

$$\frac{d}{dt} v_{jl} = (\dot{y}_j c_j / \tau_v - 1/\tau_d)\, y_l^{\text{ext}} v_{jl}$$
$$c_j = \tanh\left[a\,(V_t - \Delta E_j)\right]$$
$$V_t = V^{\text{ina}} + y_j\left(V^{\text{act}} - V^{\text{ina}}\right)$$
$$\tau_v \ll \tau_d, \quad V^{\text{ina}} < V^{\text{act}}.$$

We note that:

- the factor v_{jl} ensures that the weights do not change sign,
- pre-synaptic activity y_l^{ext} is necessary to learning,
- $\dot{y}_j$ ensures that learning only takes place when the activity changes significantly,
- the term c_j prevents runaway growth or shrinking of weights, with the sliding threshold V_t,
- the slow decay term $-1/\tau_d$ shrinks non-useful synapses.

$$R(\alpha,\beta) = \frac{1}{S(C_\alpha)} \sum_{i \in C_\alpha, l} v_{il} y_l^{\text{ext}} \qquad (1)$$
$$F(\alpha,l) = \frac{1}{S(C_\alpha)} \sum_{i \in C_\alpha} v_{il} \qquad (2)$$

Conclusions

- That is really important to keep in mind.
- Furthermore ...

References

[1] Fiser et al., Small modulation of ongoing cortical dynamics by sensory input during natural vision. *Nature*, 2004.

[2] Gros and Kaczor, Semantic learning in autonomously active recurrent neural network. *Logic Journal of the IGPL*, 2017.

[3] Mongillo et al., Synaptic theory of working memory. *Science*, 2008

varriale@itp.uni-frankfurt.de