Probabilistic Inference and Learning in a Spiking Neural Network

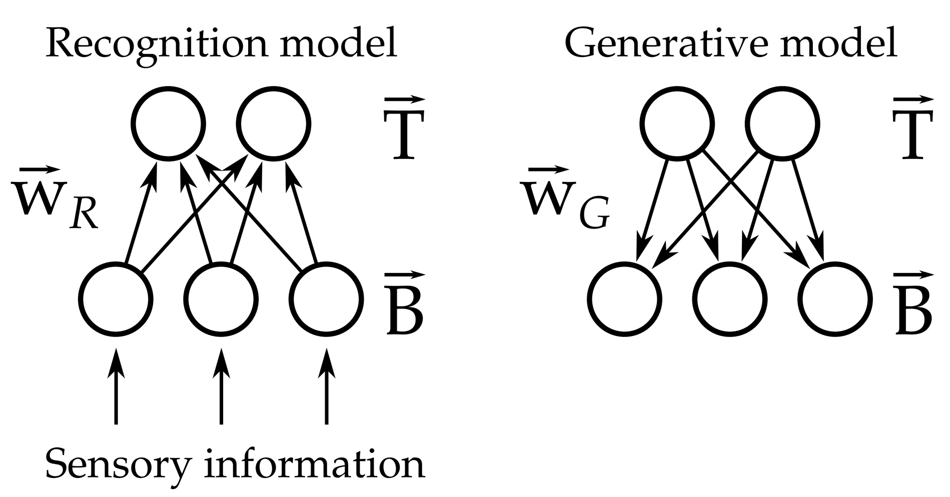
Pavel Sountsov, Paul Miller

sl157@brandeis.edu, pmiller@brandeis.edu

1. Introduction

Perceptual inference is performed in the brain by combining previously learned knowledge and recent sensory information. The mechanism that performs this computation is termed the recognition model. Theoretical work has shown that good inference is only possible if the recognition model does a good job at inverting the generative model of the world: its low and high order statistical structure.

The Helmholtz machine (Hinton et al., 1995) suggests how such a recognition model can be constructed and trained using local computations. The aim of this project is to implement the Helmholtz machine in a network of model conductance based spiking neurons and experimentally supported synaptic plasticity rules.



The simplest Helmholtz machine is composed of binary stochastic units arranged in a bottom layer $\vec{B}$ and a top layer $\vec{T}$. These two layers are interconnected with two separate sets of connections (with weights $\vec{w}_G$ and $\vec{w}_R$) creating two independently functioning networks. The bottom-up connections implement the recognition model used for inference. The top-down connections implement the generative model.

2. Learning using Wake-Sleep algorithm

$\vec{w}_G$ and $\vec{w}_R$ are learned in two phases:

- **Wake:** The bottom layer is activated using sensory data, and the recognition network is used to infer the activities in the top layer. The generative model is then used to reconstruct the sensory data given the inferred activity in the top layer. The error in this reconstruction is used to train the generative weights $\vec{w}_G$.
- **Sleep:** The top layer is activated spontaneously (or from activity in the higher areas), and the generative model is used to generate fictitious sensory activity in the bottom layer. The recognition network is then used to infer the activities in the top layer given this fictitious sensory activity. The error in this inference is used to train the recognition weights $\vec{w}_R$.

In general, during both phases some neuronal activity r_y is adjusted to match some target neuronal activity r_z by changing the connection strengths w while taking the input activity r_x into account. The simplest rule that can accomplish this is the delta rule:

$$\Delta w = \lambda r_x (r_z - r_y)$$

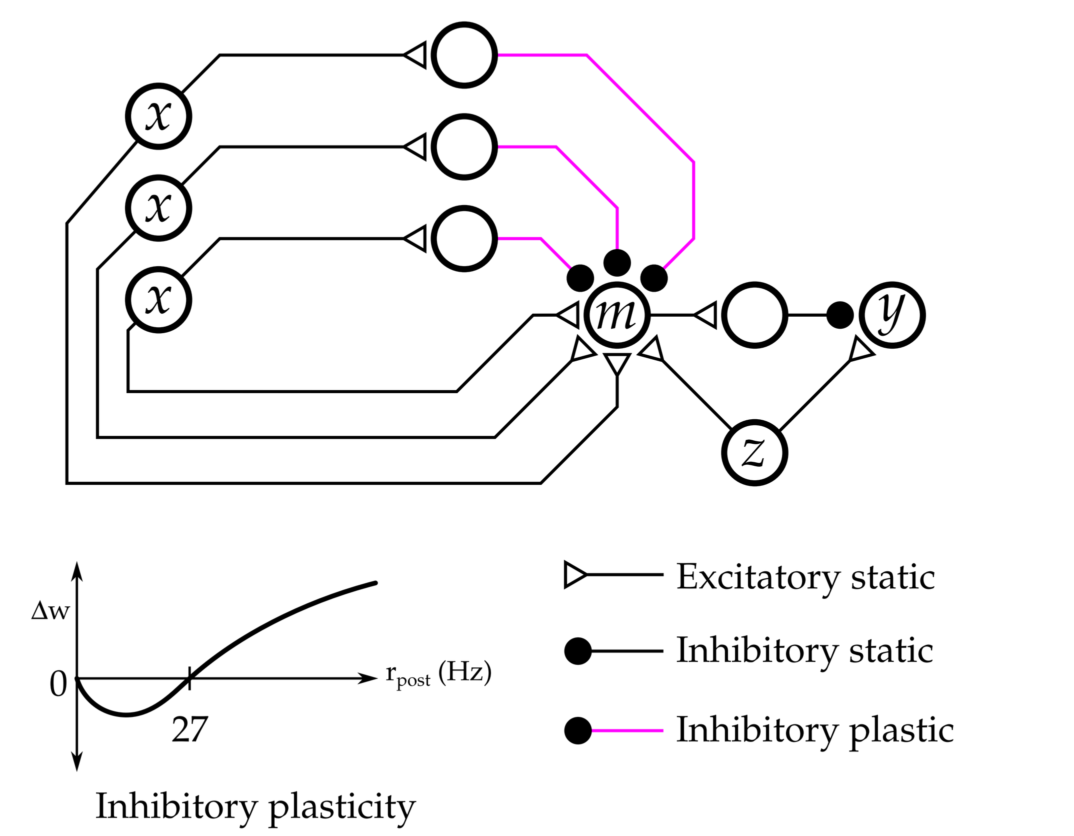
References

Haas, J. S., Nowotny, T., & Abarbanel, H. D. I. (2006). Spike-timing-dependent plasticity of inhibitory synapses in the entorhinal cortex. *Journal of neurophysiology*, 96(6), 3305-13.

Hinton, G., Dayan, P., Frey, B., & Neal, R. (1995). The "wake-sleep" algorithm for unsupervised neural networks. *Science*, 268(5214), 1158-1161.

3. Delta rule network

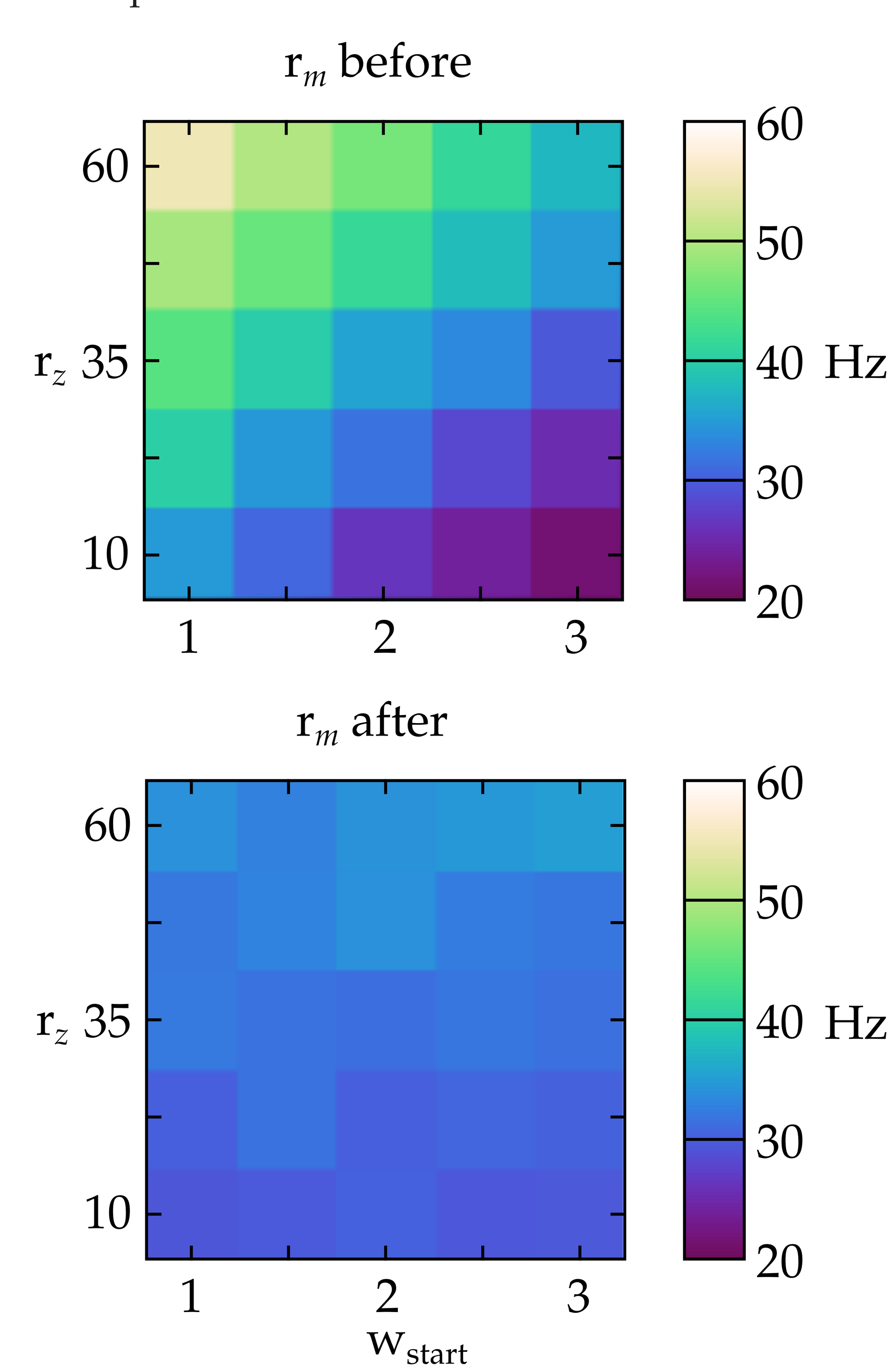
Previous efforts to model the delta rule required the use of discrete neuronal firing rates or relied on a special timing relationship between the target and output neuronal activities. The network below improves on those efforts and implements the delta rule with simultaneous, continuously varying firing rates.



The network is composed of interconnected pools of 10-20 spiking neurons each. Plastic inhibitory connections follow an anti-Hebbian synaptic plasticity rule similar to one characterized by Haas et al. (2006) in the entorhinal cortex.

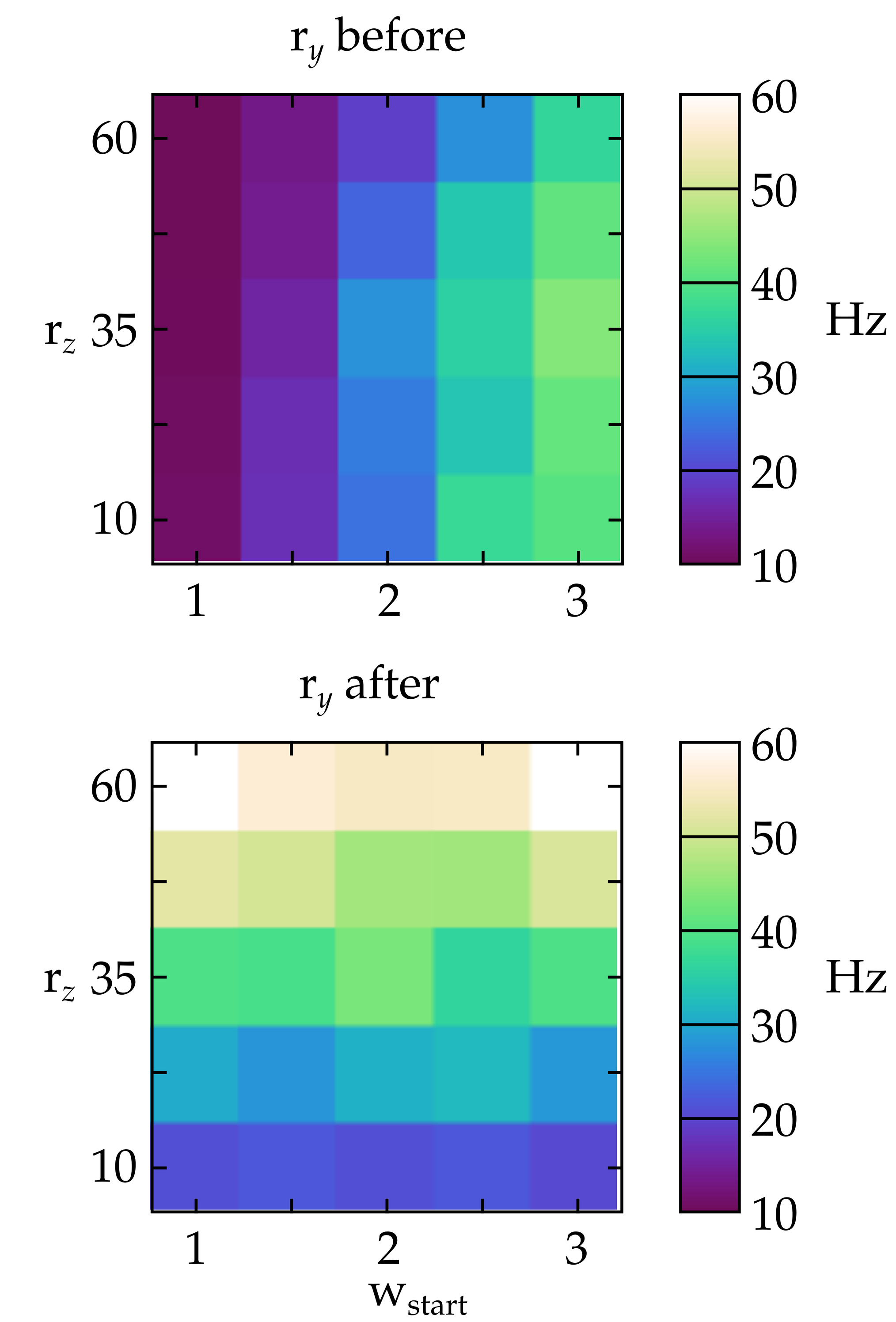
4a. Delta network results

25 separate networks were trained with a steady input, but different initial plastic inhibitory weights and target rates. r_m approaches the rate of ≈27 Hz in all networks after 40 seconds of stimulus presentation:

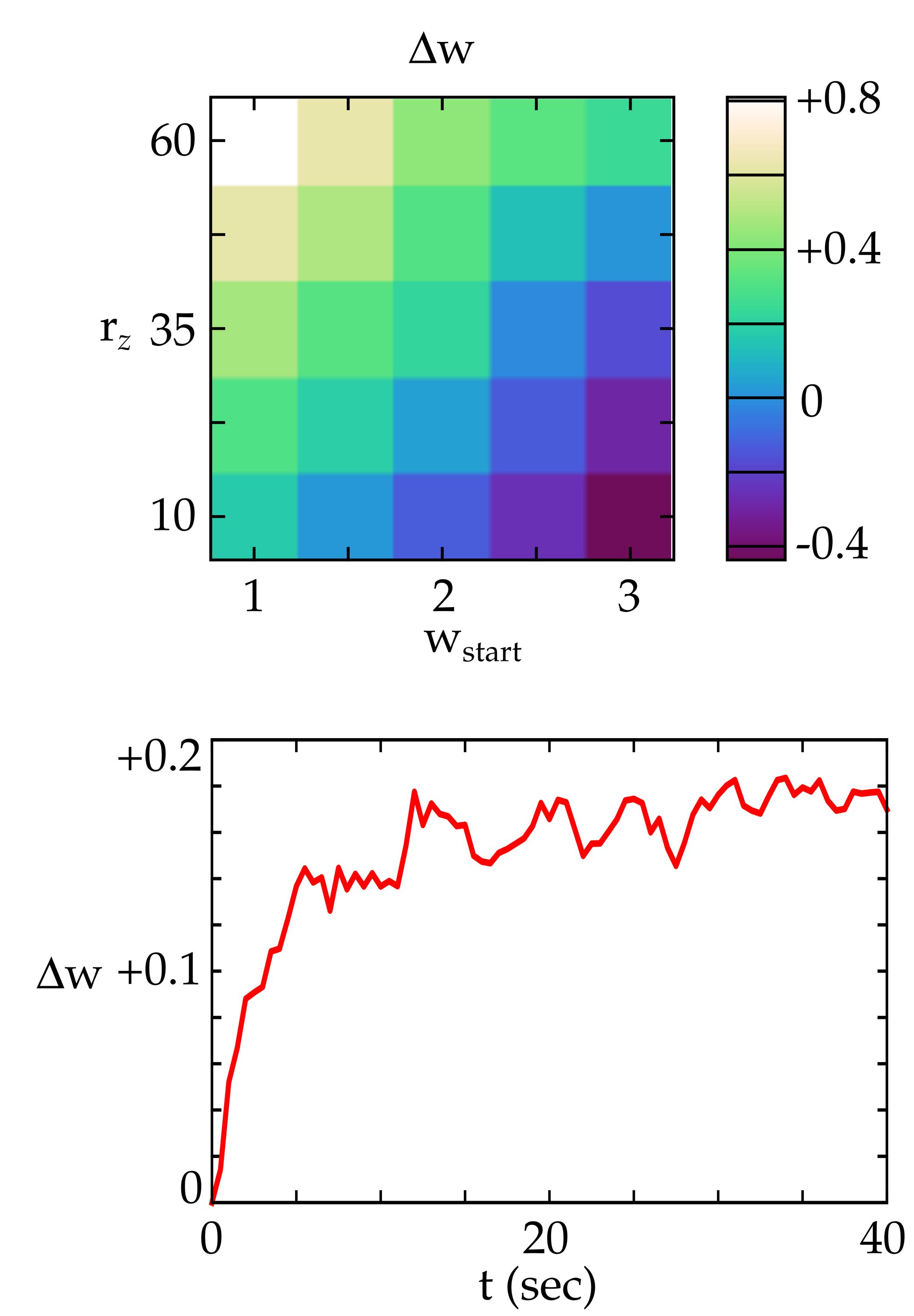


4b. Delta network results

Initially the r_y does not depend on r_z, but after training it matches it closely (given steady input rates):

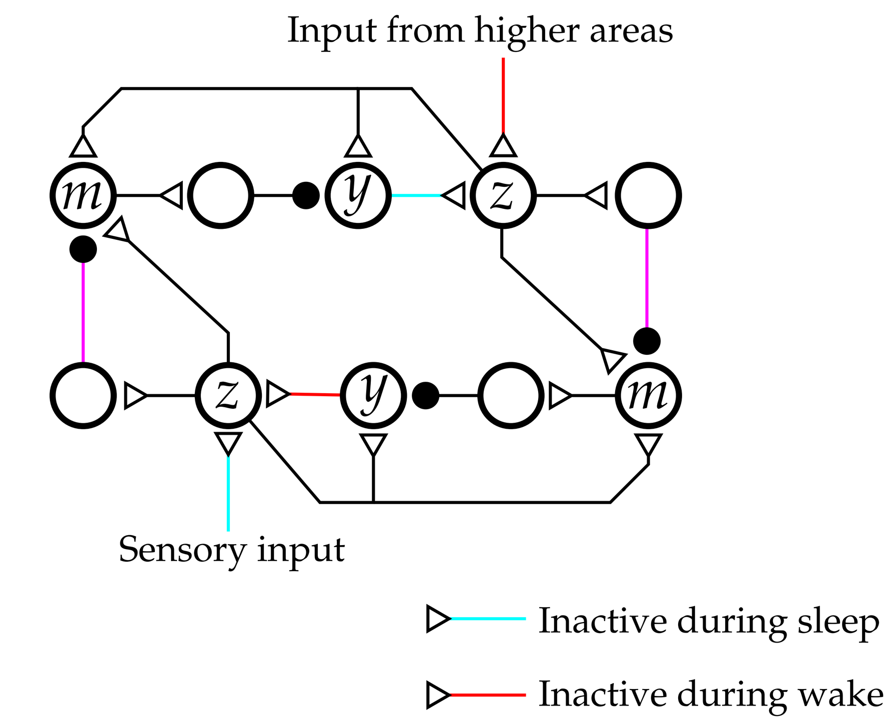


Weight changes follow the initial r_m values and converge to a steady state after ≈20 seconds of stimulus presentation:



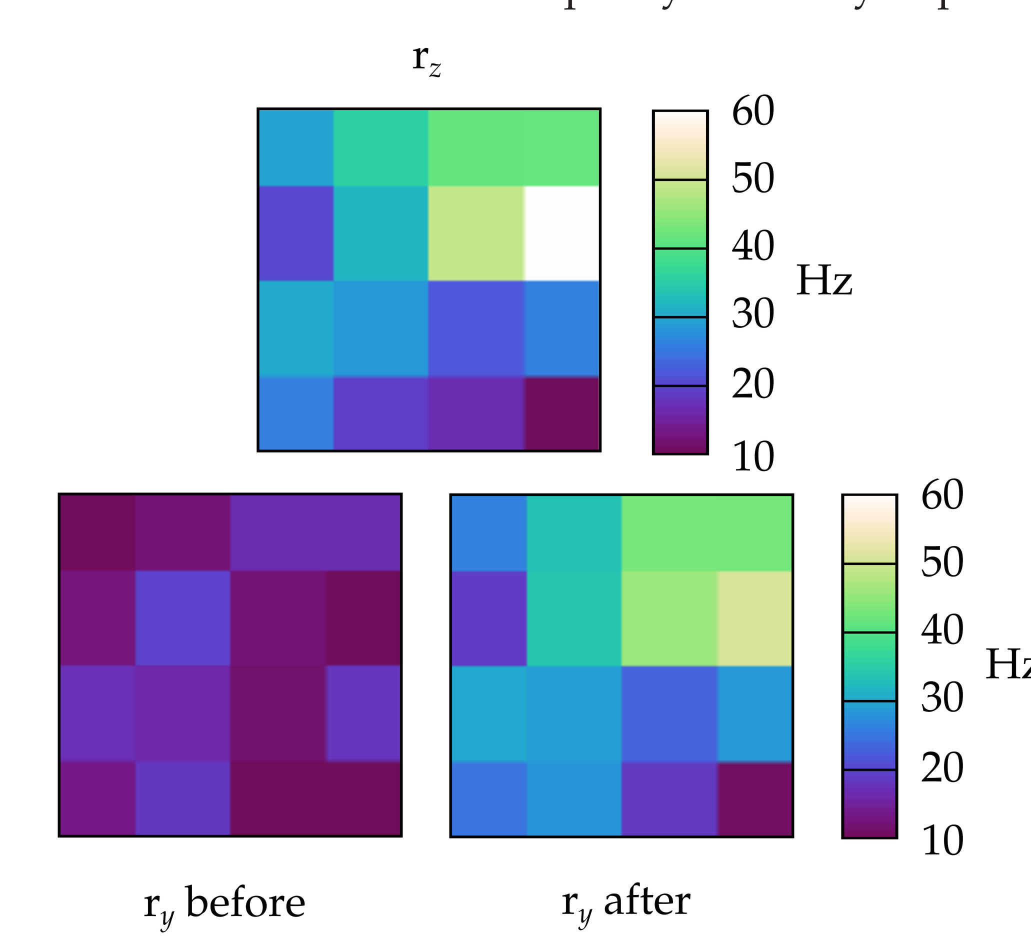
5. Helmholtz machine network

The full Helmholtz machine can be built by linking multiple delta-rule networks (each one functioning as a stochastic unit) together. Certain connections need to be inactivated during the different phases of the Wake-Sleep algorithm through either a neuro-modulatory signal, or silencing of the pre-synaptic neurons through inhibition.



6. Helmholtz machine network results

A network with 16 bottom units and 4 top units with all-to-all connectivity was run in the wake phase to test its ability to reconstruct the sensory input. Before training the reconstruction does not resemble the sensory input, but after 50 seconds of stimulus presentation the reconstruction quality markedly improves.



7. Conclusions

1. The proposed delta rule network seems very capable at implementing the computation in spite of the non-ideal behavior of the plasticity rule and spiking neurons.
2. Multiple delta rule networks combined into a Helmholtz machine seem capable of implementing the wake phase of the Wake-Sleep algorithm. Future work will examine if alternating sleep and wake configurations of the network produces correct learning of the recognition weights.

Acknowledgements

Thanks to József Fiser for suggesting to look into the Wake-Sleep algorithm and his continued assistance.

Funded by NIH CNTG T90 DA032432 and IGERT 0549390.