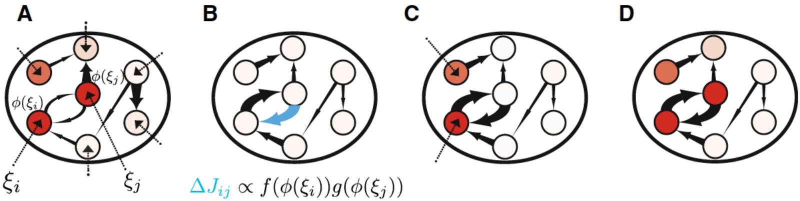
Information bounds and attractor dynamics of a Hebbian associative memory

Clayton Seitz

May 21, 2021

Introduction

RNNs trained with Hebbian learning rules



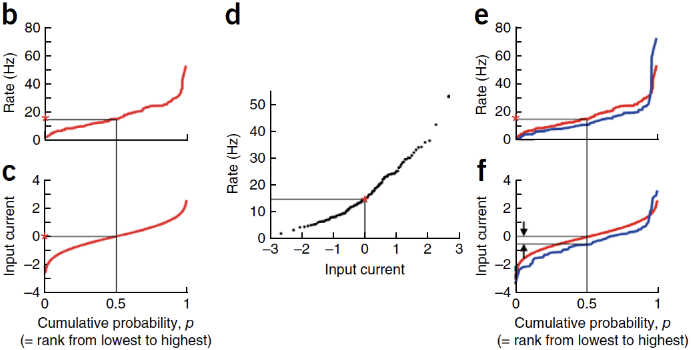
Let W_{ij} be a matrix of recurrent weights that evolves when stimulated by

$$\xi(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp -\frac{1}{2}(\boldsymbol{r} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{r} - \boldsymbol{\mu})$$

[1]

[1] [Peirera and Brunel, Neuron. 2018]

Inferring learning rules from firing rate distributions in ITC

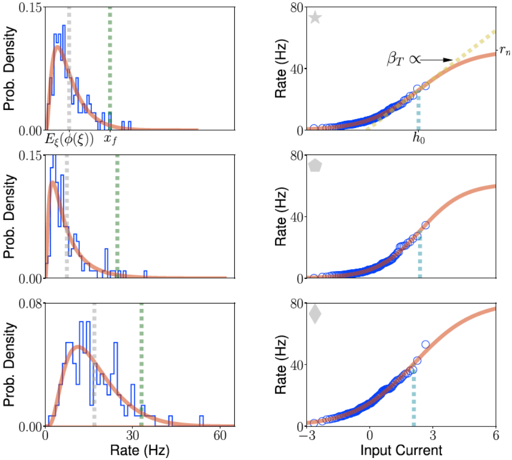


Inferring ΔW_{ij} from ITC neurons after presentation of novel and familiar images [2]

[2] [Lim et al., Nature Neuroscience. 2015]

Inferring the transfer function from ITC data

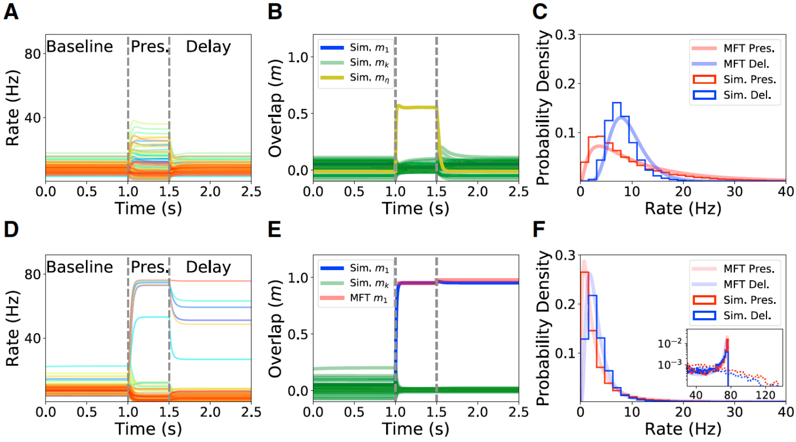
All you can really observe is the firing rate distribution. Assume the input currents are Gaussian



3

[3] [Peirera and Brunel, Neuron. 2018]

Presenting novel and familiar stimuli to the network



4

[4] [Peirera and Brunel, Neuron. 2018]

A Hebbian update for synaptic weights

Let the synaptic update be a separable function of the presynaptic (ξ_i) and postsynaptic (ξ_j) firing rates

$$\Delta W_{ij} = f(\phi(\xi_i))g(\phi(\xi_j)) \rightarrow W_{ij} = C_{ij} \sum f(\phi(\xi_i))g(\phi(\xi_j))$$

Evolution of the firing rate for neuron i is

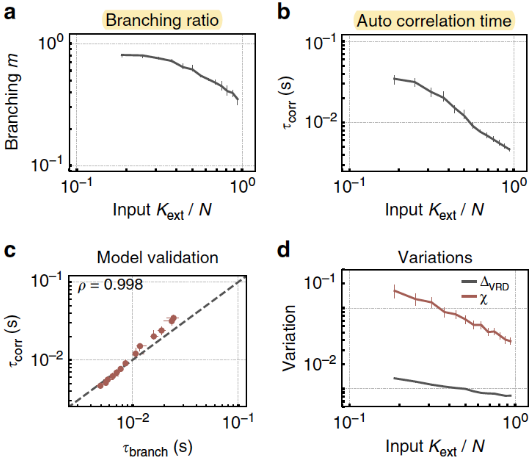
$$\tau \dot{r}_i = -r_i + \phi(I + \sum_{j \neq i} W_{ij} r_j)$$

5

[5] [J.J. Hopfield PNAS. 1982]

Do these networks optimize information transmission?

Are these networks functioning at a critical point? What about the balance between input and recurrence? (Cramer et al. 2020)



A coding theory perspective

How much information does the response R carry about the input pattern S i.e. $I(R;S)$ on novel and familiar stimuli?

What is the fundamental coding capacity of these networks?