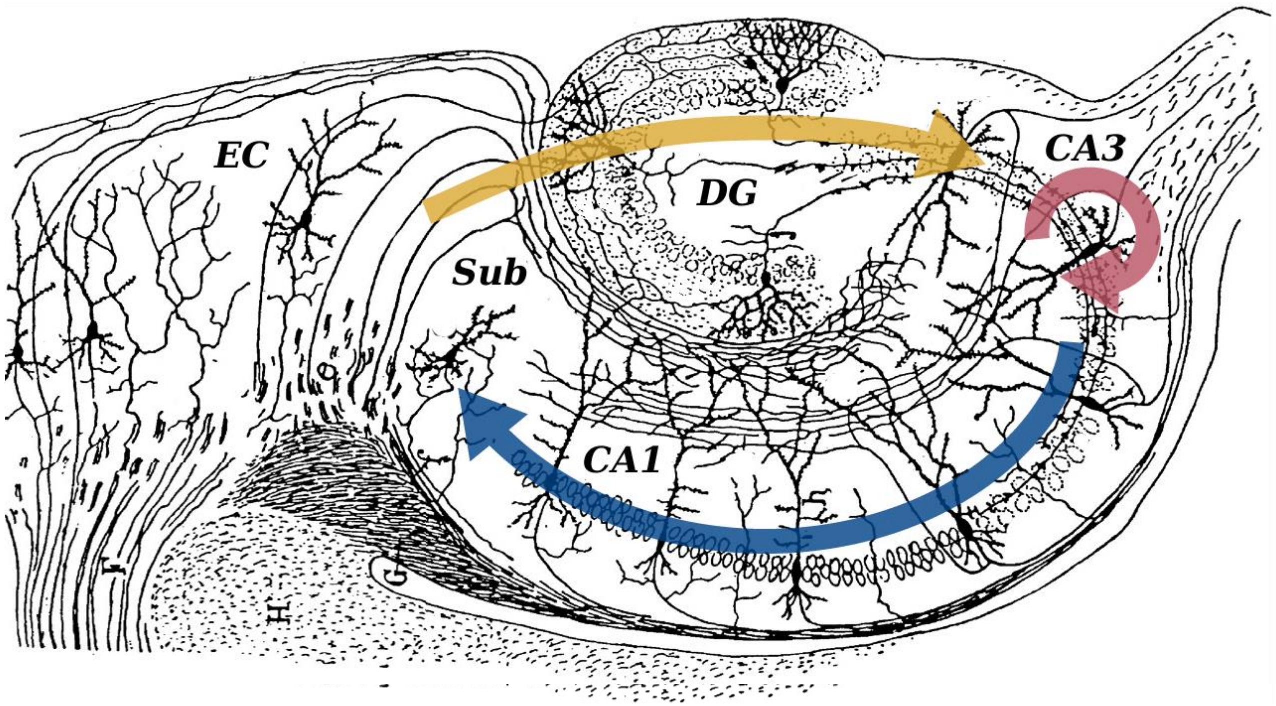
Predictive Recirculation: A model of encoding and replay of sequences in a recurrent neural network

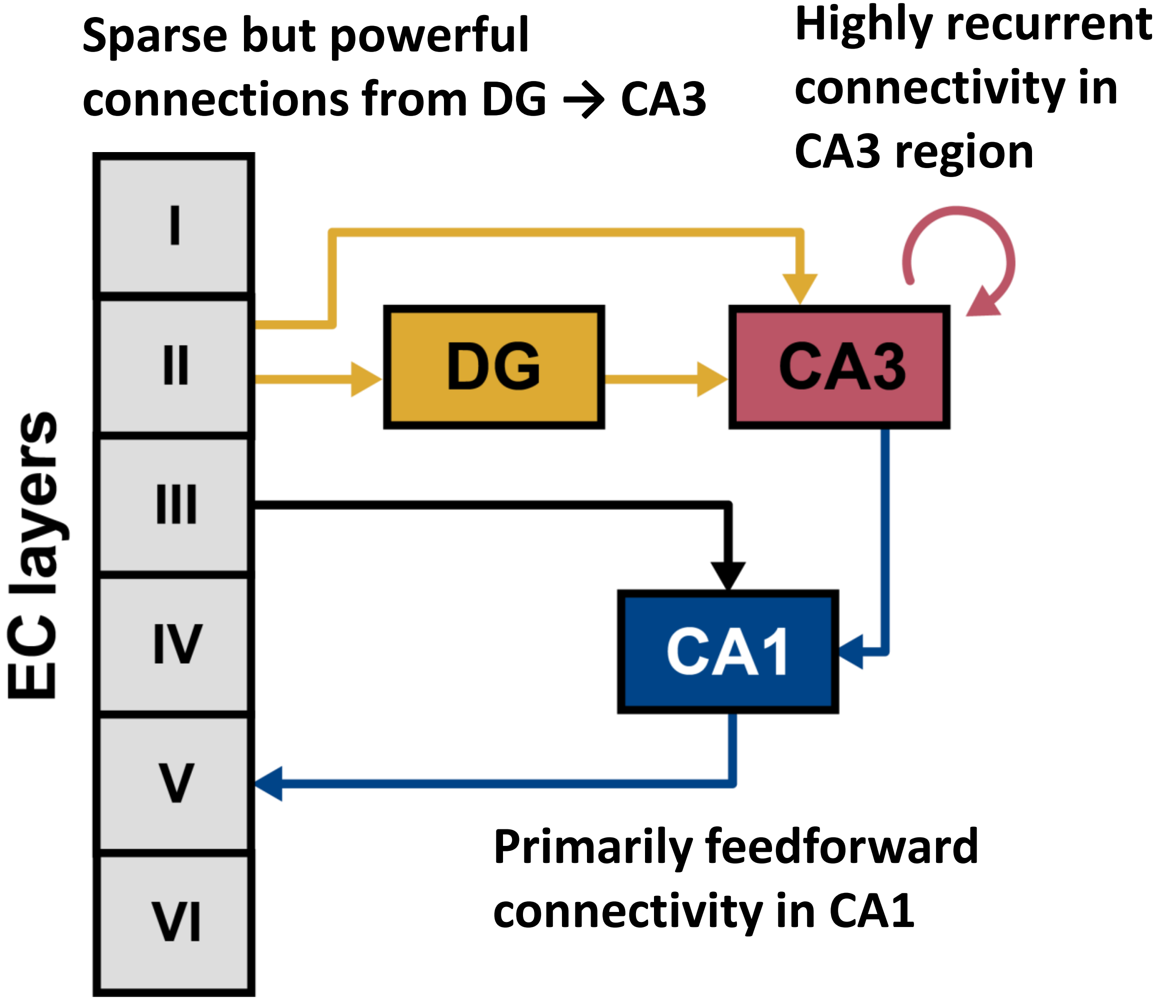
Mia Cameron*, Homero Esmeraldo*
Computational Neurobiology Laboratory
Salk Institute for Biological Studies
December 16^{th}, 2022

*co-first author

What are the computational processes underlying sequence storage and replay in the hippocampus?

Specialized connectivity patterns at different regions may facilitate different types of computation



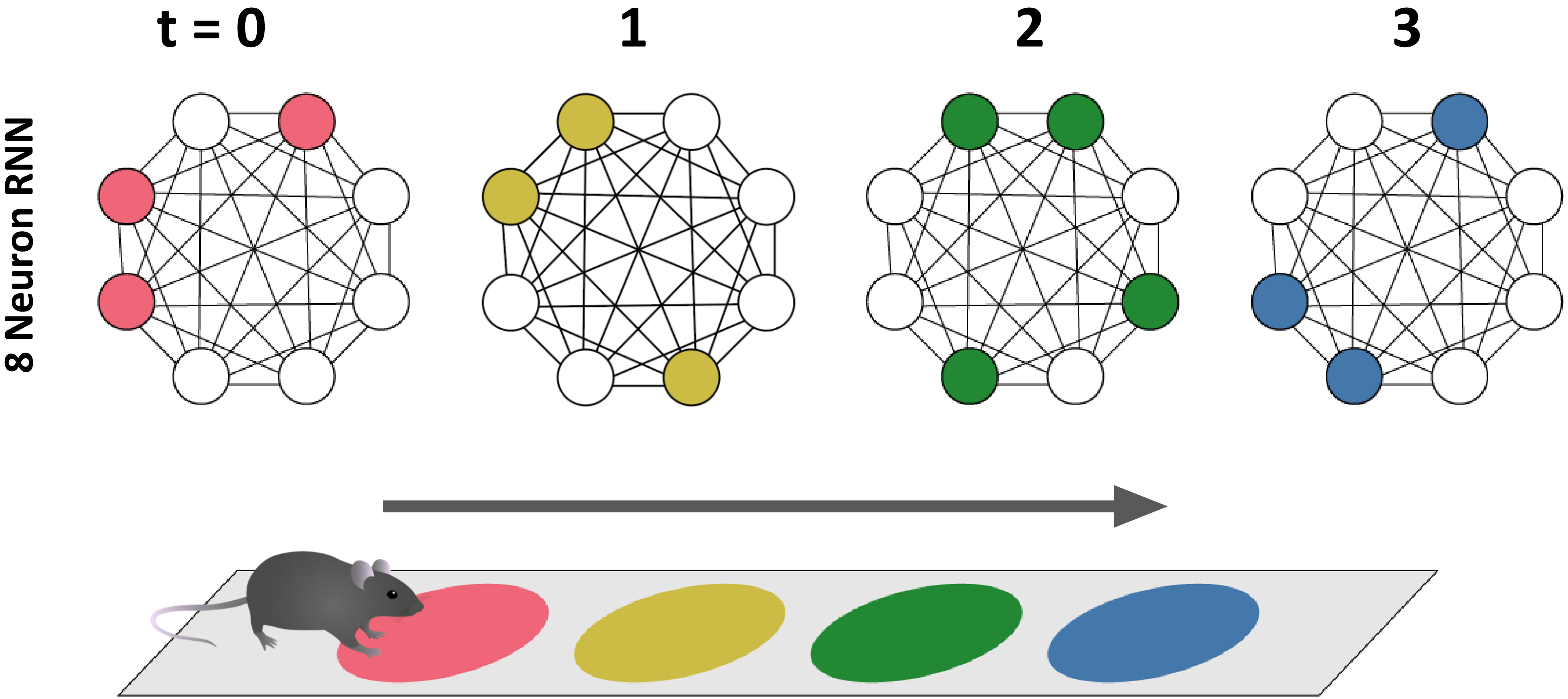


CA3 may store sequences in a recurrent network

Modelling CA3 as an artificial Recurrent Neural Network (RNN)

Each element of a sequence is represented as a static pattern in the RNN

During learning, each pattern is an external input to the network

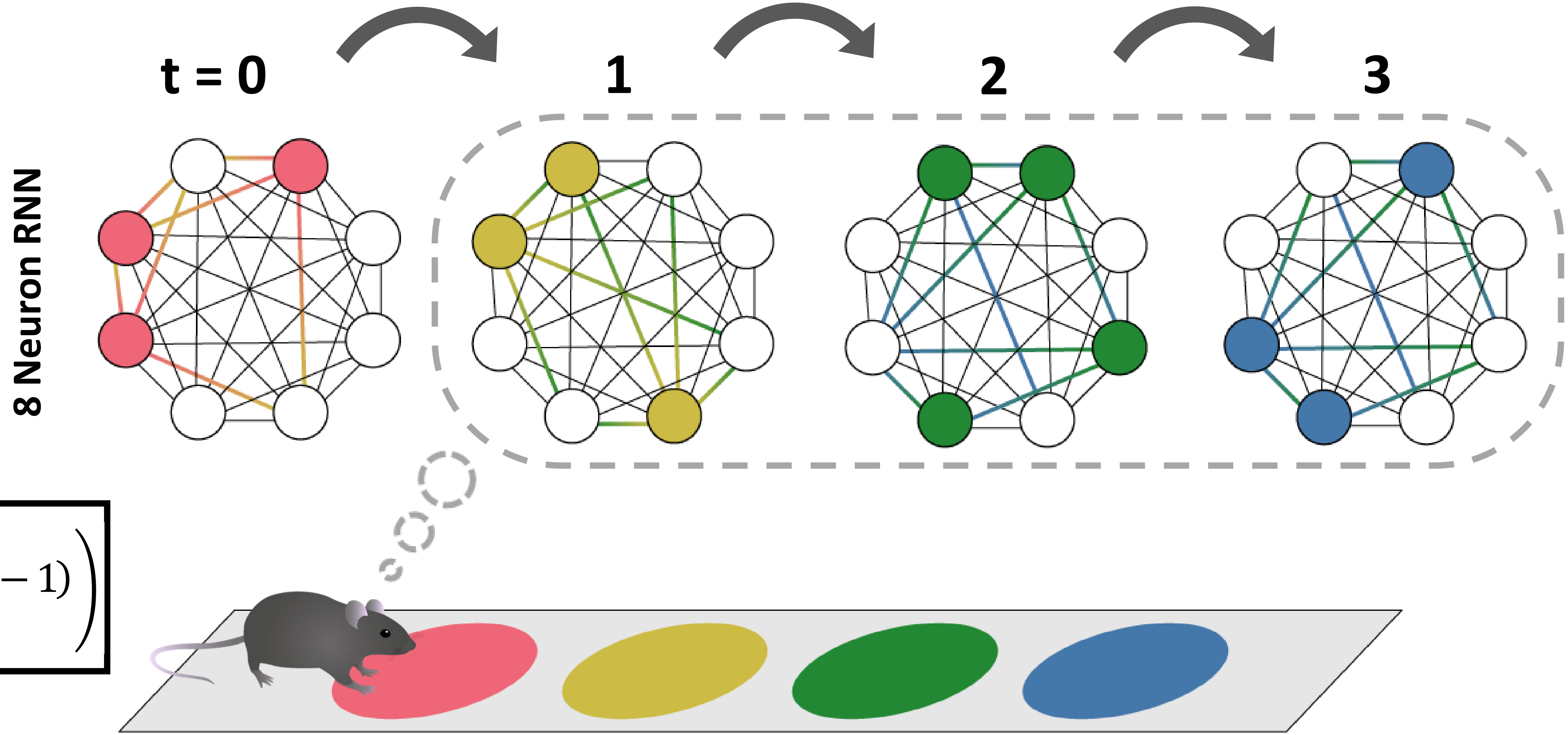


Asymmetric connectivity allows the network to recall the entire sequence, given the first element

During recall, patterns are sequentially activated by recurrent network weights

Network dynamics:

$$\hat{x}_i(t) = \sigma\left(\sum_j W_{ij}\, x_i(t-1)\right)$$



How can these synaptic weights be learned in a biologically plausible way?

Deriving a local recurrent learning rule:

Network dynamics:

$$\boxed{\hat{x}_i(t) = \sigma\left(\sum_j W_{ij}\, x_i(t-1)\right)}$$

$$\Delta W_{ij} = -\lambda \frac{\partial E_i(t)}{\partial W_{ij}}$$

Gradient descent at each time step

Error function:

$$\boxed{E_i(t) = \frac{1}{2}\left(x_i(t) - \hat{x}_i(t)\right)^2}$$

$$\frac{\partial E_i(t)}{\partial W_{ij}} = \left(x_i(t) - \hat{x}_i(t)\right)\frac{\partial \hat{x}_i(t)}{\partial W_{ij}}$$

$$= \left(x_i(t) - \hat{x}_i(t)\right)\sigma'\left(\hat{x}_i(t)\right)\left(x_j(t-1) + \boxed{\sum_{k\neq j} W_{ik}\frac{\partial x_k(t-1)}{\partial W_{ij}}}\right)$$

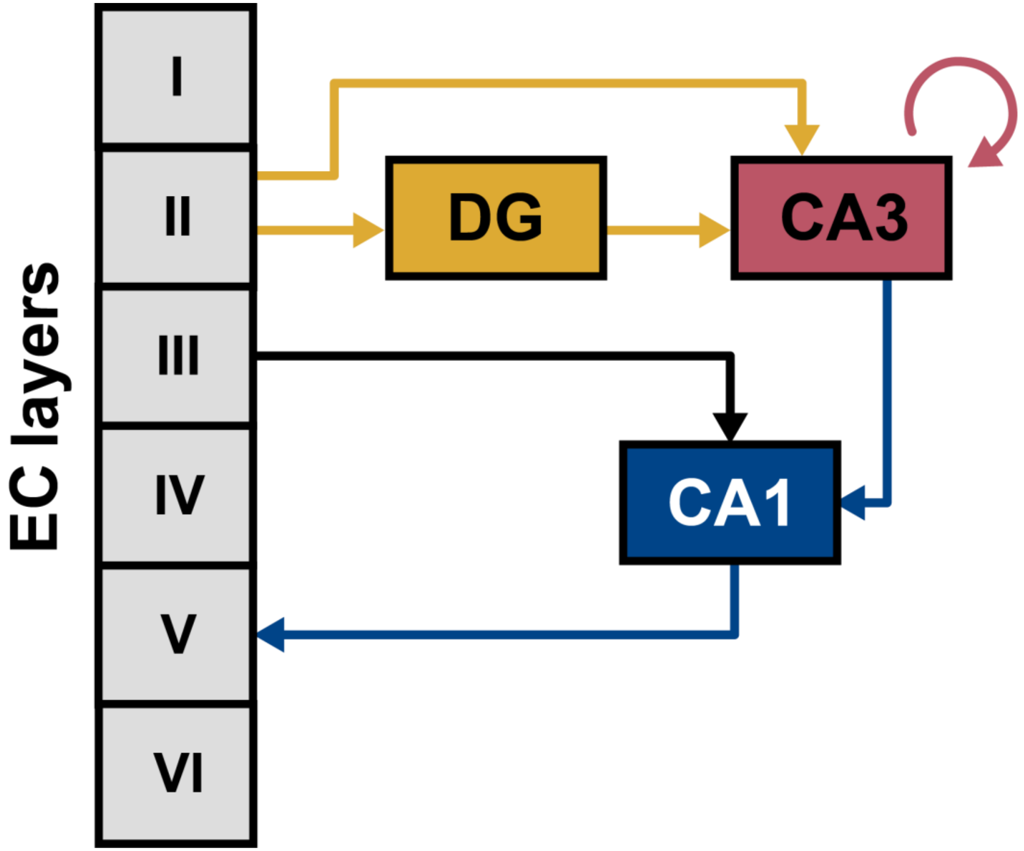
Restricted to spatially and temporally local information:

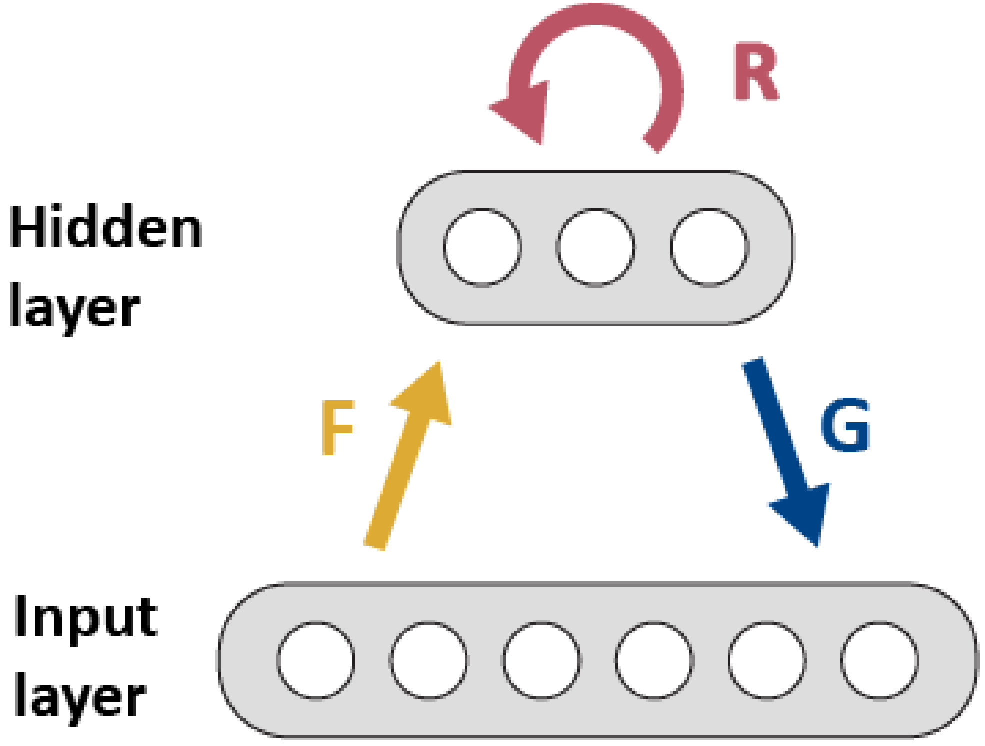
$$= \left(x_i(t) - \hat{x}_i(t)\right)\sigma'\left(\hat{x}_i(t)\right) x_j(t-1)$$

Recurrent Learning Rule:

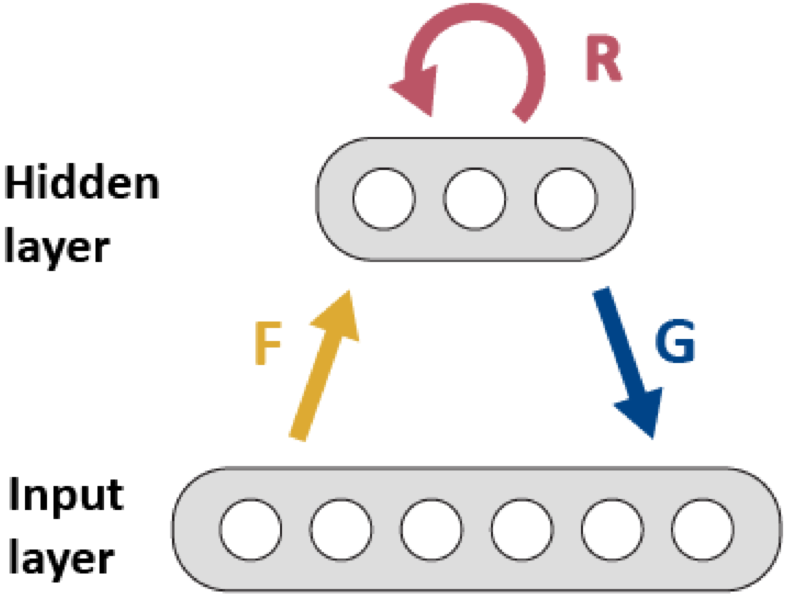
$$\boxed{\Delta W_{ij} = -\lambda\left(\mathrm{x_i(t)} - \hat{x}_i(t)\right)\sigma'\left(\hat{x}_i(t)\right) x_j(t-1)}$$

Sequences may be recoded to circumvent locally restricted recurrent learning rule

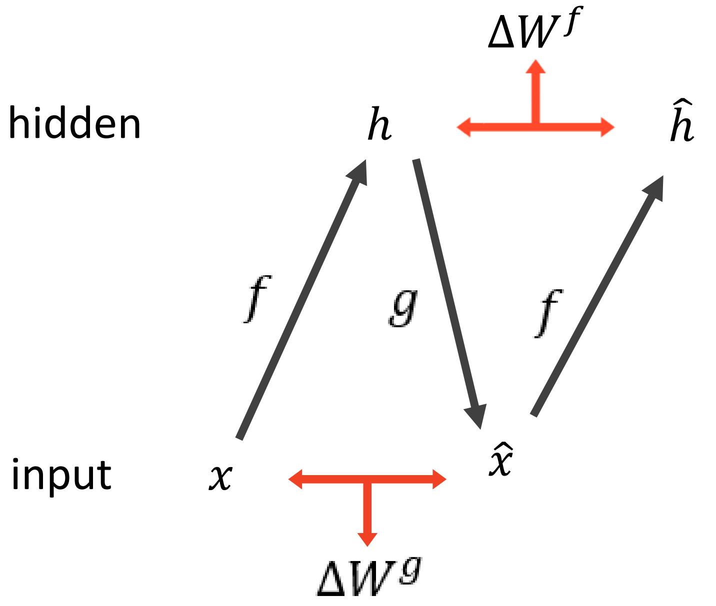




Predictive recirculation combines an autoencoder and RNN



Recirculation autoencoder learning procedure:



Network dynamics:

Synaptic input to unit i

$$s_i = \sum_j W_{ij} u_j$$

Logistic activation

$$u_i = \sigma(s_i) = \frac{1}{1 + e^{-s_i}}$$

Temporal regression

$$u_i^{t+1} = \lambda u_i^t + (1 - \lambda)\sigma(u_i^{t+1})$$

Recirculation Learning Rule

Visible-to-hidden

$$\Delta W_{ij}^f = \epsilon \hat{x}_j (h_i - \hat{h}_i)$$

Hidden-to-visible

$$\Delta W_{ij}^g = \epsilon h_j (x_i - \hat{x}_i)$$

Hinton, G. E., & McClelland, J. (1987). Learning representations by recirculation. In *Neural information processing systems*.

Combined Learning Procedure:

Recurrent Learning Rule:

$$\Delta W^r_{ij} \coloneqq \epsilon\left(h^t_i - p^t_i\right) h^{t-1}_j$$

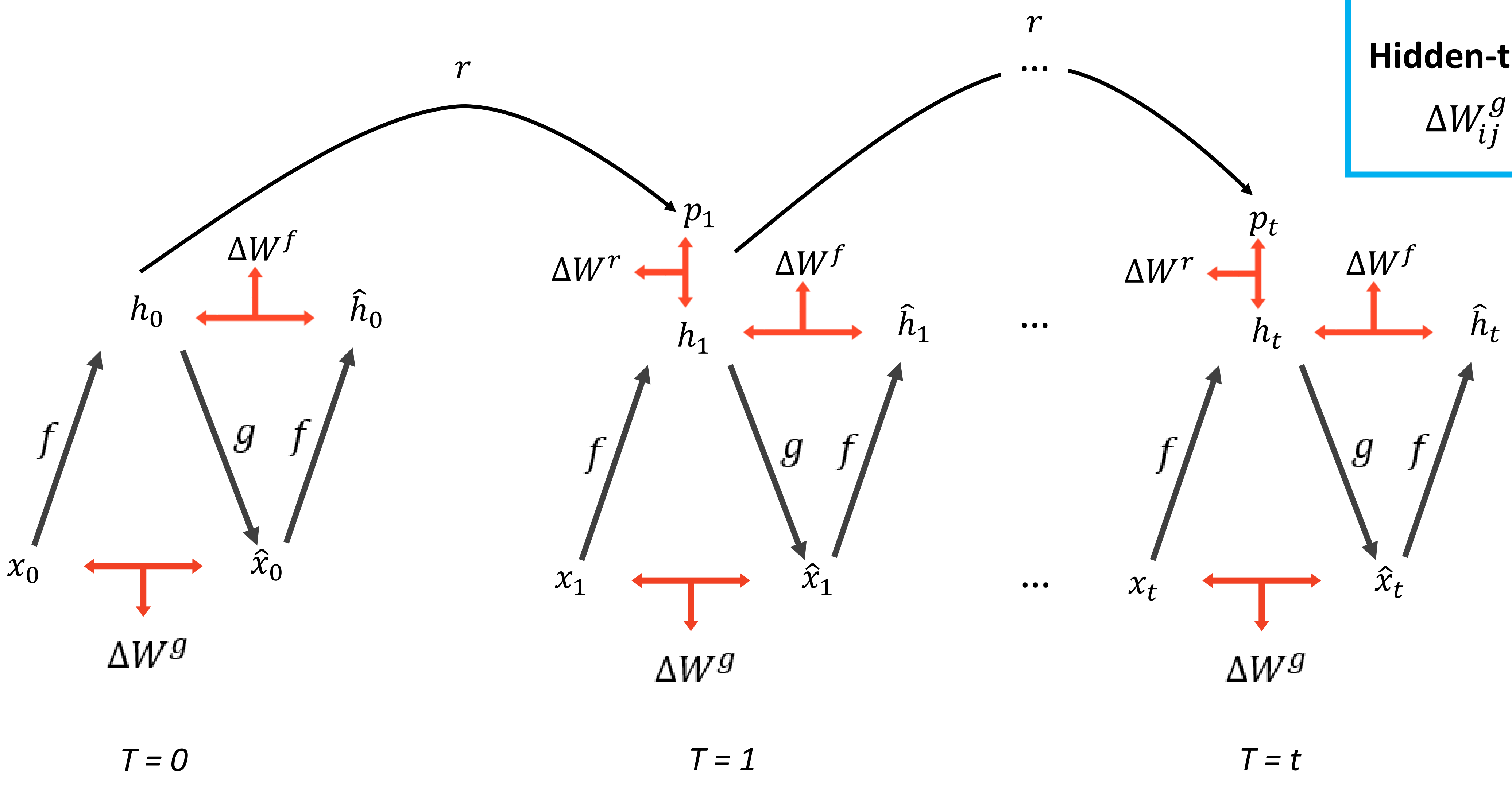
Recirculation Learning Rule

Visible-to-hidden

$$\Delta W^f_{ij} = \epsilon \hat{x}_j (h_i - \hat{h}_i)$$

Hidden-to-visible

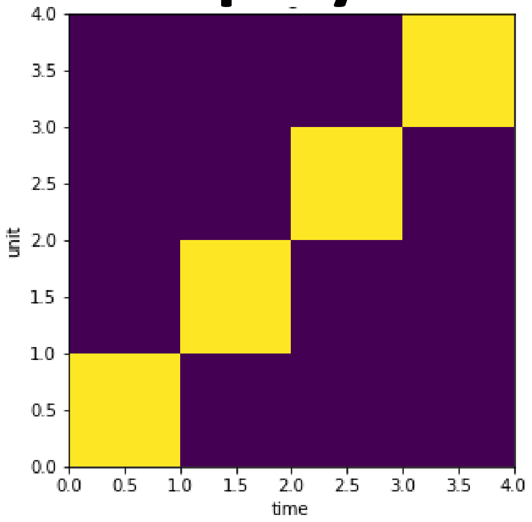
$$\Delta W^g_{ij} = \epsilon h_j (x_i - \hat{x}_i)$$



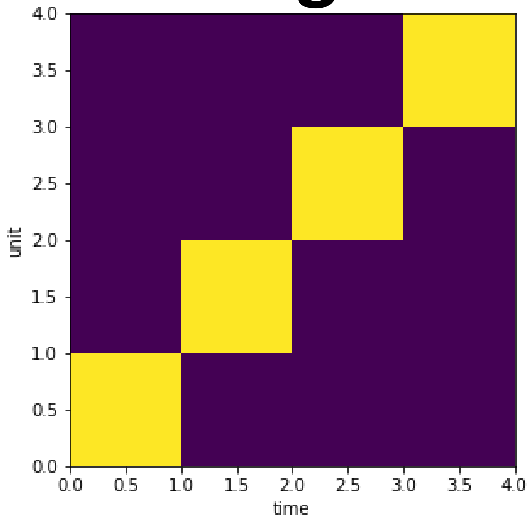
Binary sequences can be encoded and replayed

Sequential 4-2-4

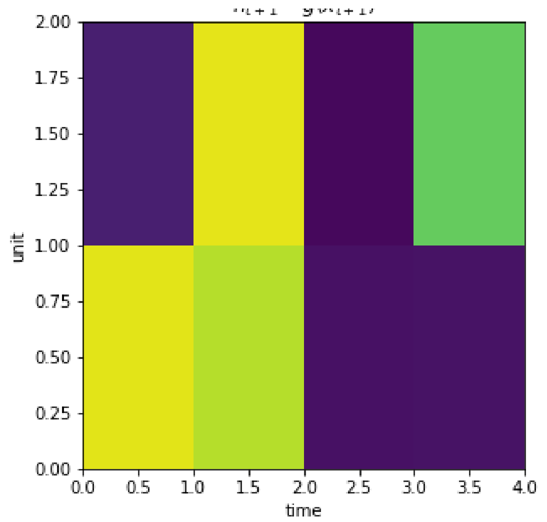
Replayed



Target

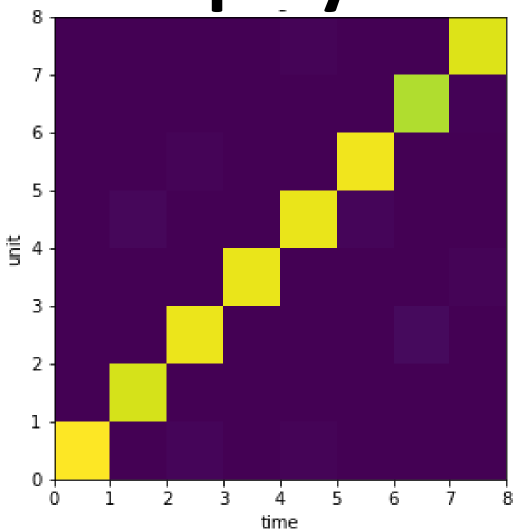


Learned H encoding:

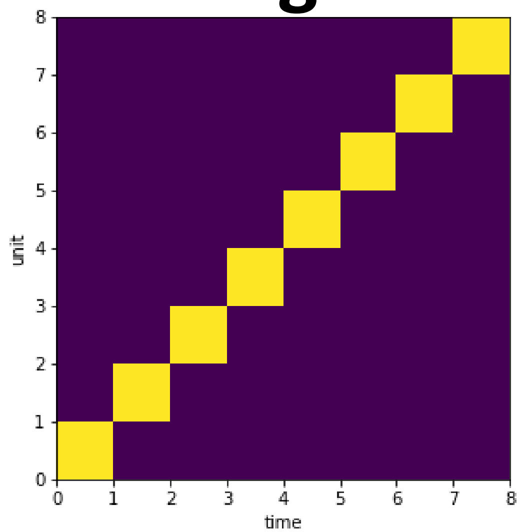


Sequential 8-4-8

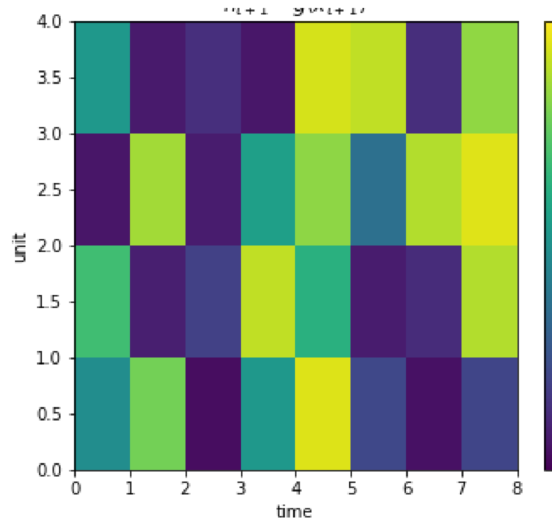
Replayed



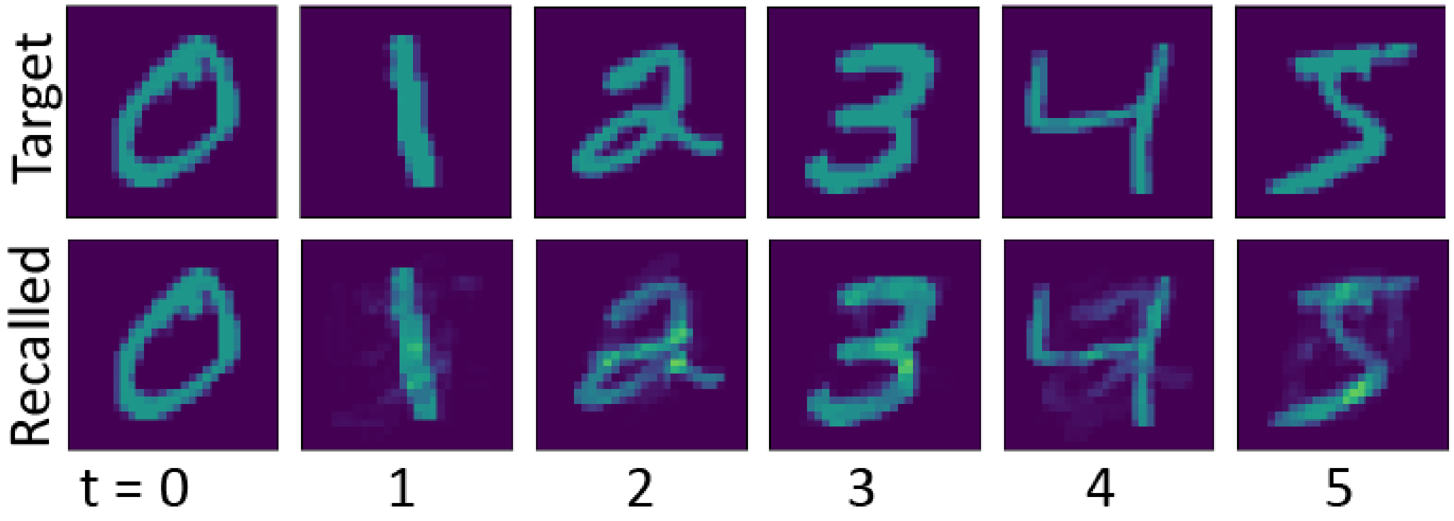
Target



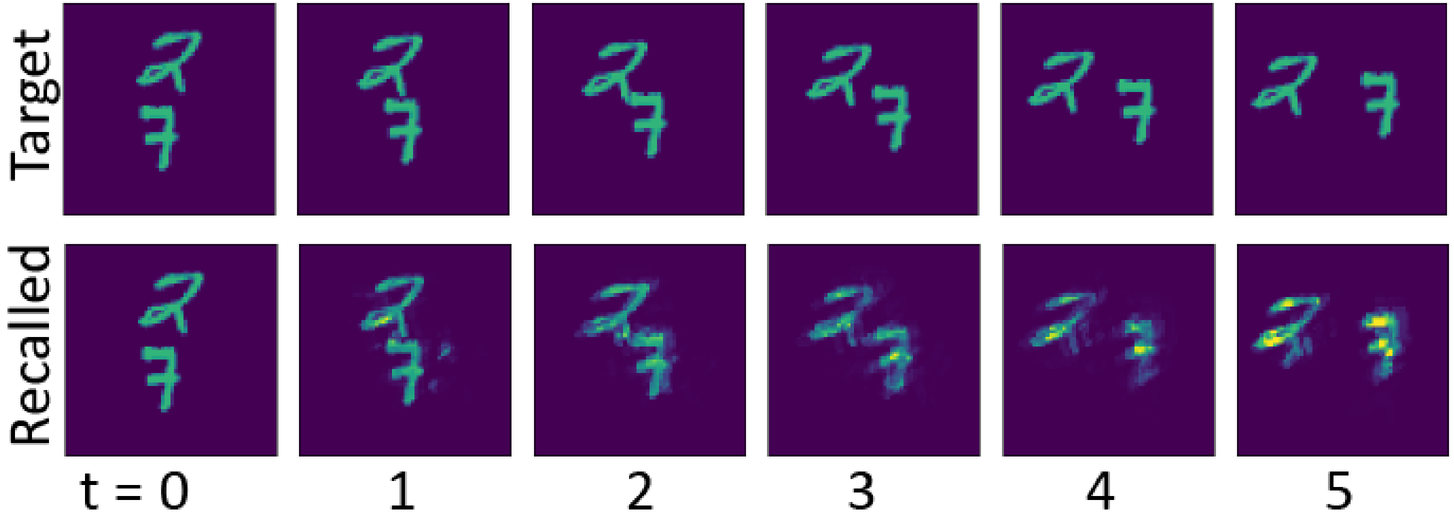
Learned H encoding:



Sequential MNIST



Moving MNIST



Sequences with high overlap between elements can also be learned

Acknowledgements

- Prof. Terrence Sejnowski
- Homero Esmeraldo
- Yusi Chen
- Vikrant Jaltare
- All members of CNL