A diverse range of factors affect the nature of neural representations underlying short-term memory

A. Emin Orhan and Wei Ji Ma

Published: February 2019

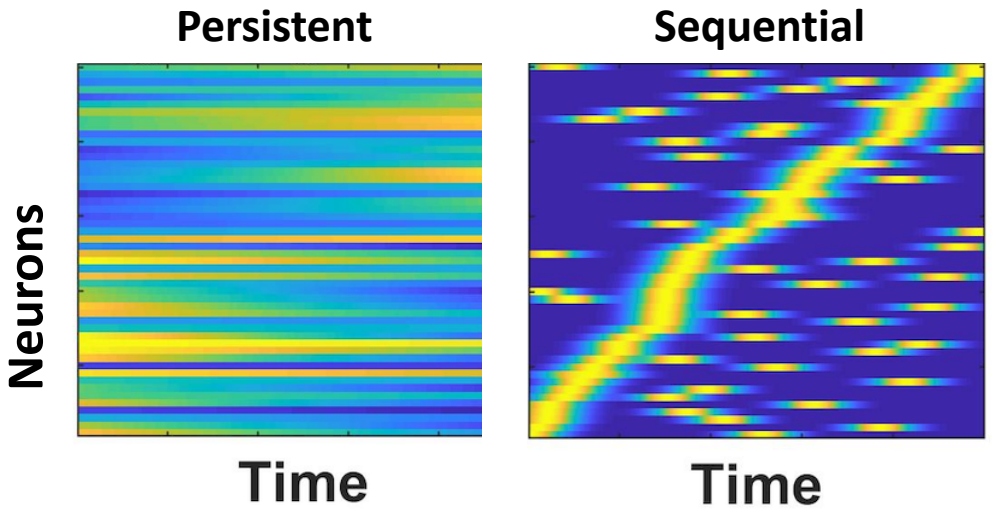
Rehan Chinoy

31 July 2020

Research Goal and Experimental Setup

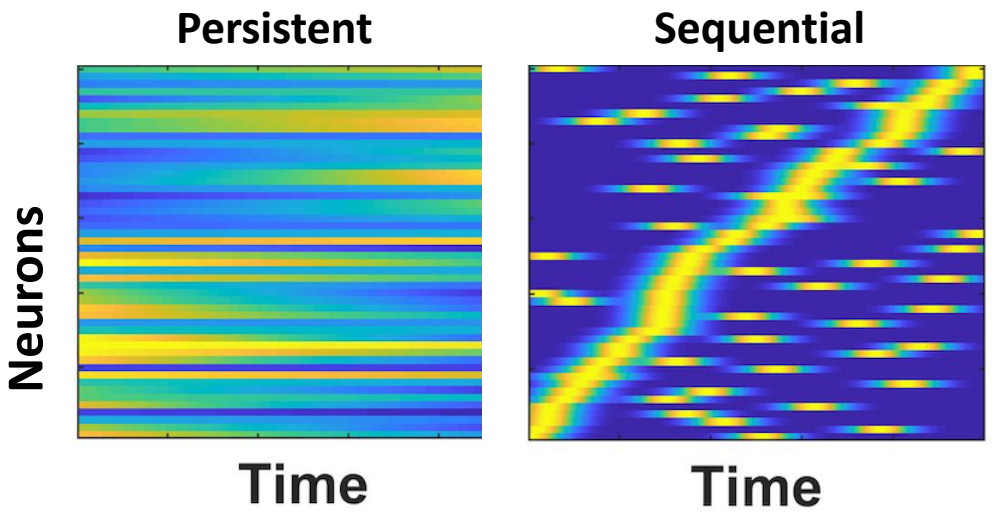
Research Goal

- Persistent and sequential activity models are two prominent models of short-term memory in neural circuits.



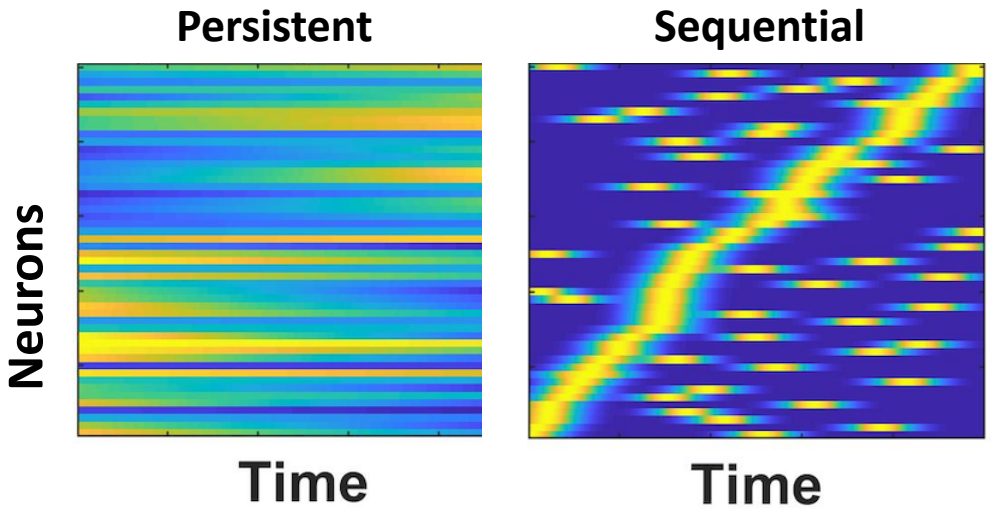
Research Goal

- Persistent and sequential activity models are two prominent models of short-term memory in neural circuits.
- *What is the underlying circuit mechanism that determines whether a persistent or sequential solution will emerge in the network?*



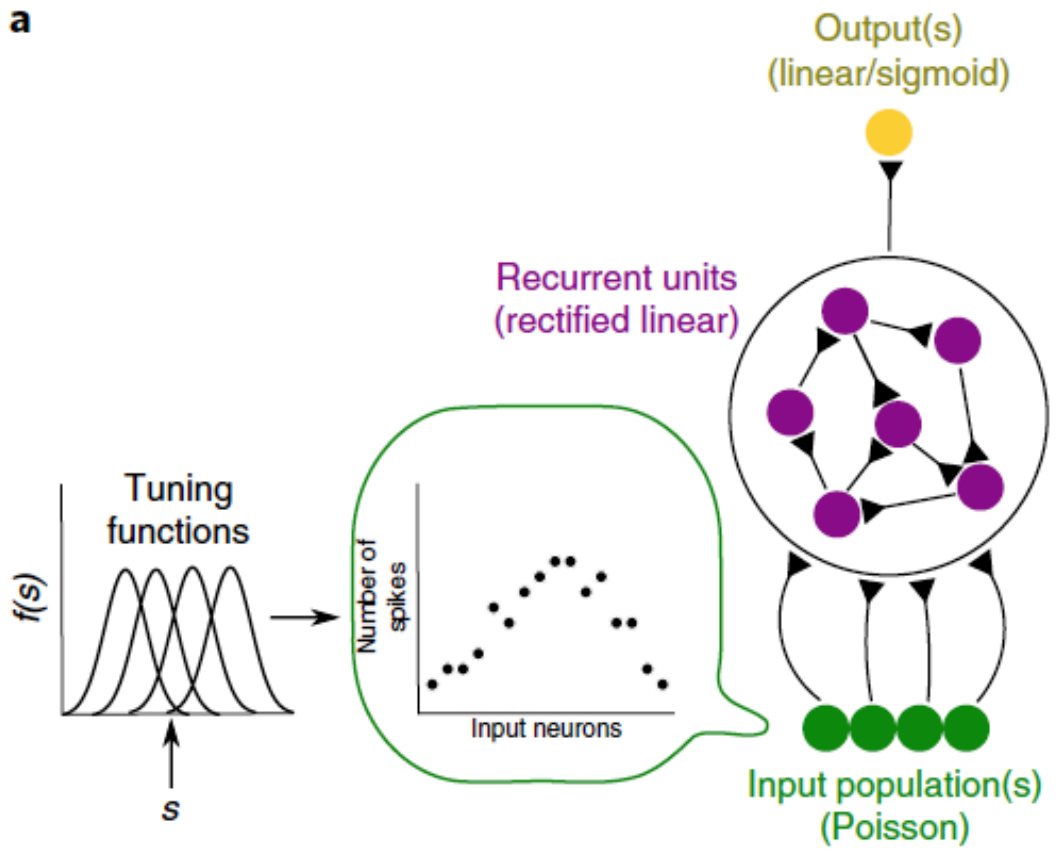
Research Goal

- Persistent and sequential activity models are two prominent models of short-term memory in neural circuits.
- *What is the underlying circuit mechanism that determines whether a persistent or sequential solution will emerge in the network?*
- Authors address this question by training RNNs on several short-term memory tasks under a wide range of circuit and task manipulations.

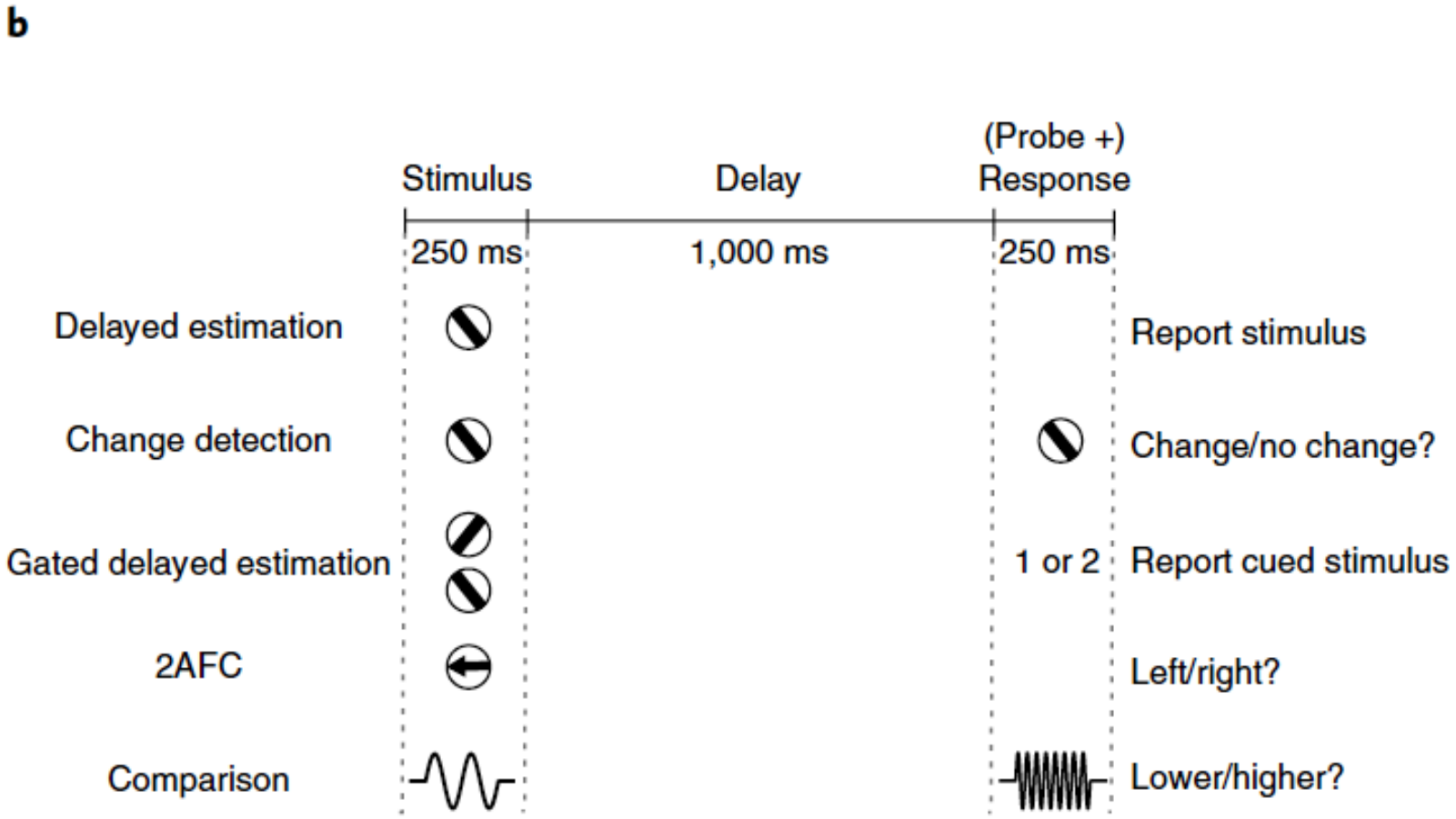


Networks

- Each network is trained to learn a short-term memory task
- Tuning functions map a stimulus to the firing rate of input neurons
- Poisson input neurons fire independently at each time step of the task
- ReLU units clip activity such that the lower bound is zero

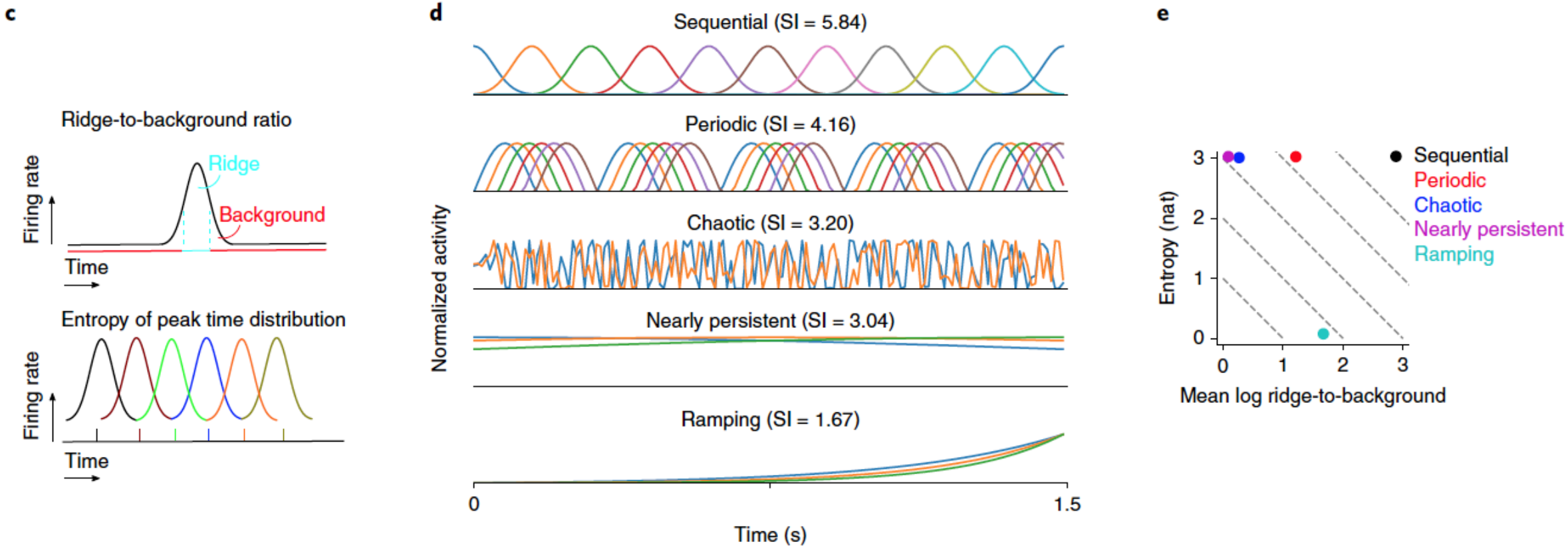


Tasks


b
Stimulus
Delay
(Probe +)
Response
250 ms
1,000 ms
250 ms
Delayed estimation
Report stimulus
Change detection
Change/no change?
Gated delayed estimation
1 or 2
Report cued stimulus
2AFC
Left/right?
Comparison
Lower/higher?

Sequentiality Index

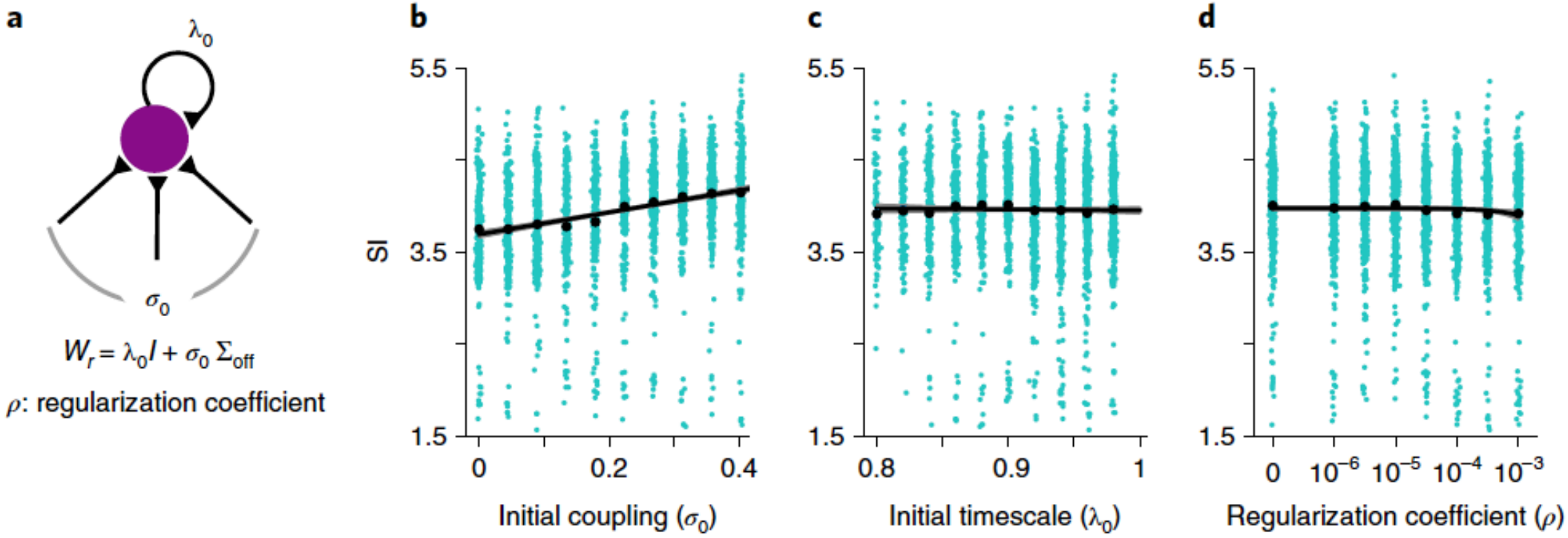
SI = (entropy of peak response time distribution of the recurrent neurons) + (mean log ridge-to-background ratio of the neurons)



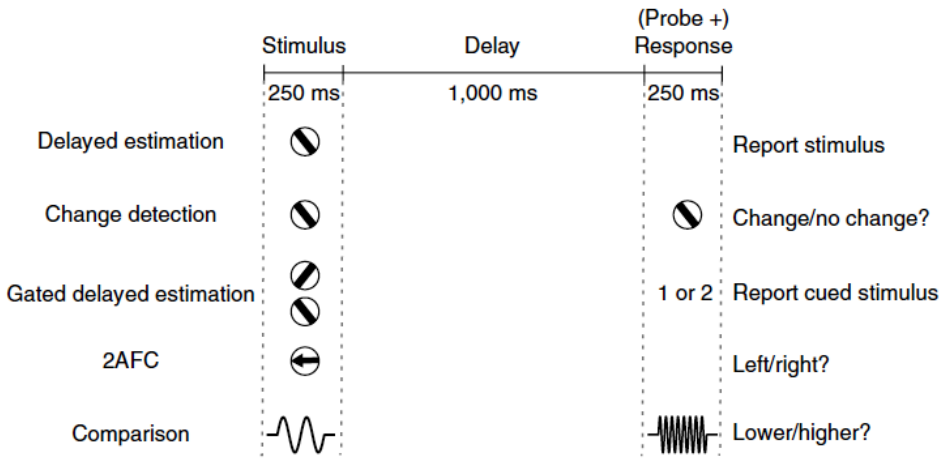
Factors Affecting SI of Network Solutions

- Intrinsic circuit properties
- Temporal complexity of tasks
- Hebbian STSP
- Delay duration variability

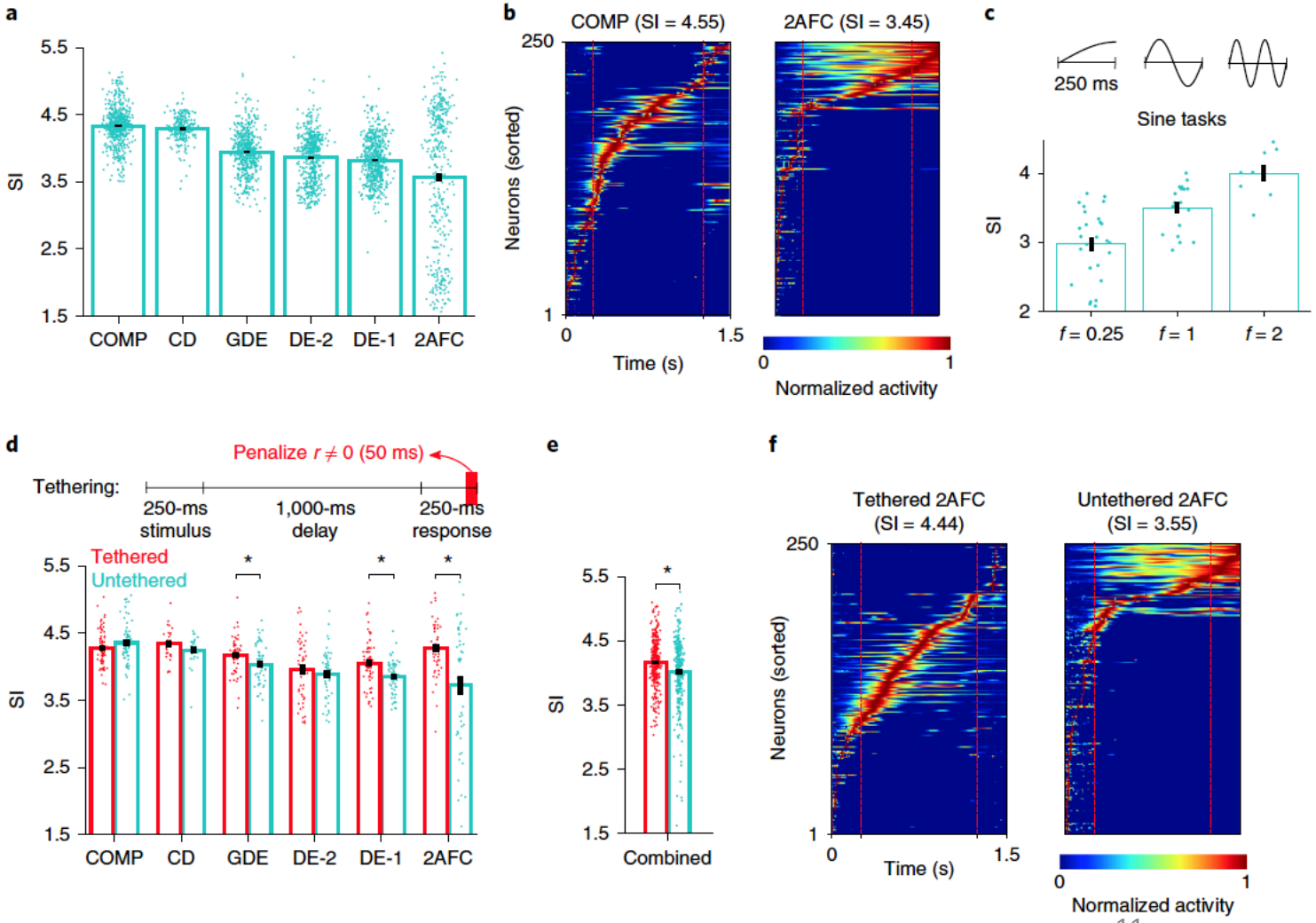
Intrinsic circuit properties



Temporal complexity of tasks increases SI



Hypothesis: increasing target function complexity (i.e. frequency) increases sequentiality of the network solution



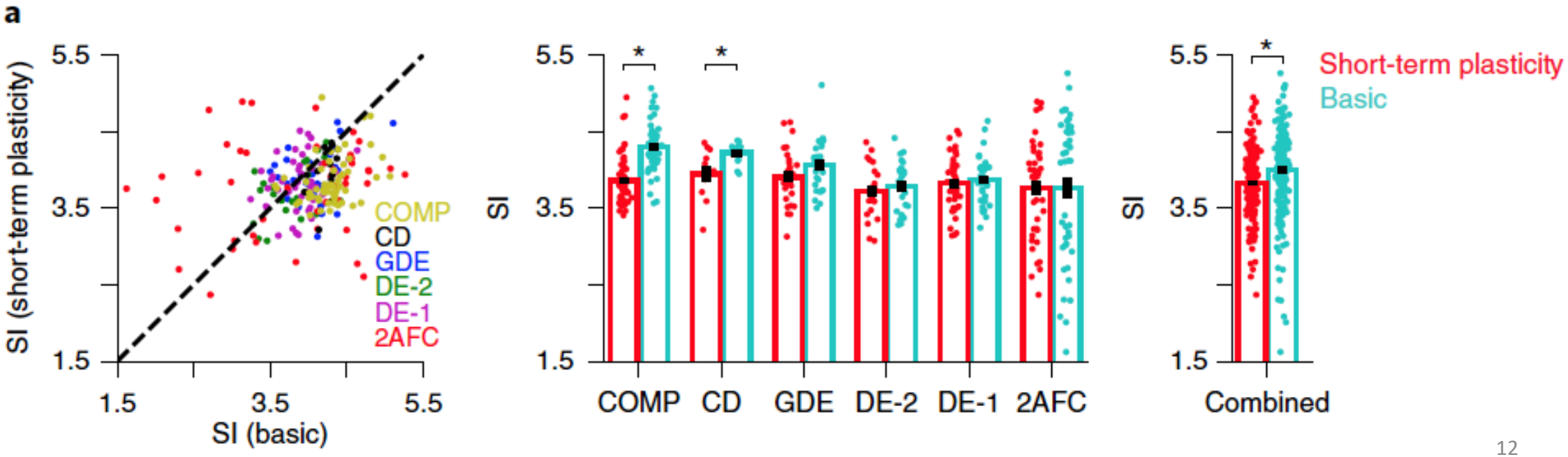
Symmetric Hebbian STSP decreases SI

Basic discrete-time formulation: $\mathbf{r}_t = f(W_r \mathbf{r}_{t-1} + W_h \mathbf{h}_t + \mathbf{b})$

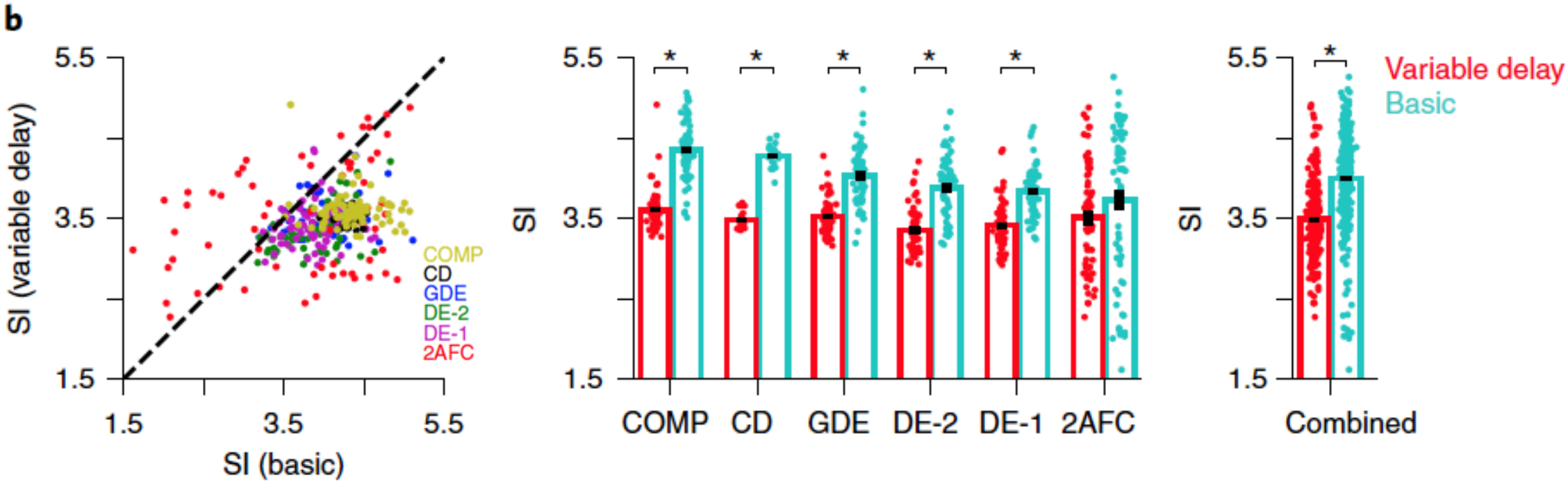
With Hebbian STSP, $\mathbf{r}_t = f\left(\left[W_r + \sum_{\tau=1}^{T} \gamma^{\tau} \mathbf{r}_{t-\tau-1}\, \mathbf{r}_{t-\tau-1}^{\top}\right] \mathbf{r}_{t-1} + W_h\, \mathbf{h}_t + \mathbf{b}\right)$

T>1 is unstable, so set T=1: $\mathbf{r}_t = f(W_r \mathbf{r}_{t-1} + \boxed{\gamma \kappa(\mathbf{r}_{t-2}^{\top} \mathbf{r}_{t-1})\, \mathbf{r}_{t-2}} + W_h \mathbf{h}_t + \mathbf{b})$

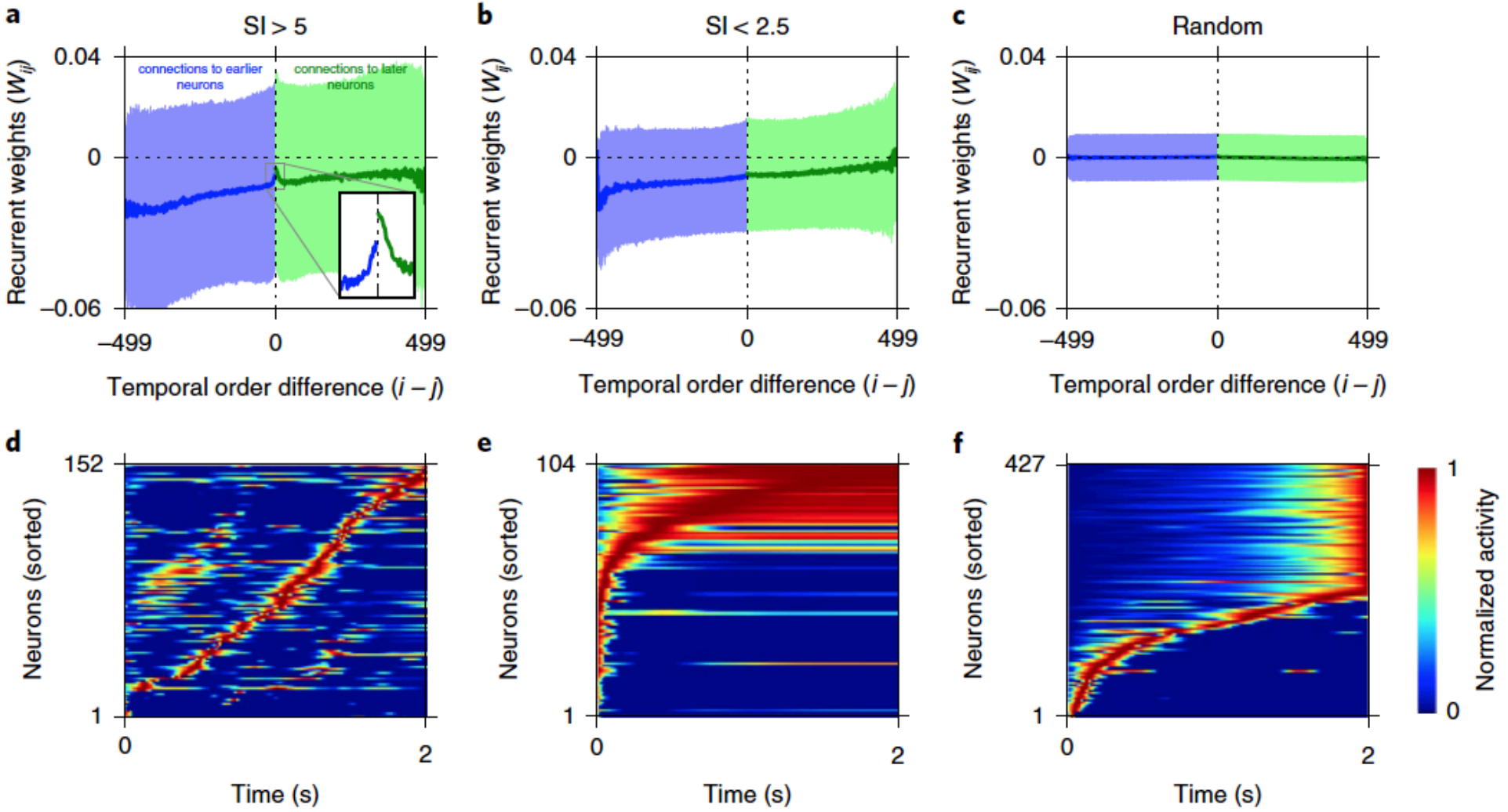
** Where f is ReLU, γ=0.0007 is Hebbian contribution, and $\kappa()$ is a clipping function for stability*



Delay duration variability decreases SI



Circuit mechanism that generates sequential versus persistent activity



Concluding Remarks

- This paper establishes a mechanism for the maintenance of STM and shows that persistent and sequential solutions are ends of a spectrum that emerges from training
- Why does the network's mechanism for maintaining short-term memory rely on non-normal dynamics?
- Is the model biologically realistic?