Attractor networks in systems with underlying random connectivity.

P.E. Latham
and
S. Nirenberg

University of California at Los Angeles

898.14
SFN 1999

Introduction

Most treatments of Hopfield networks (PNAS 1982) assume a weight matrix of the form

$$J_{ij} \propto \sum_{\mu} \epsilon_\mu \eta_i^\mu \eta_j^\mu$$

where ϵ_μ represents the strength of the μ^{th} memory and η^μ is a vector of 0s and 1s. Here we consider the more realistic case in which the weight matrix has additional components:

$$J_{ij} = W_{ij} + c_{ij} \sum_{\mu} N_\mu^{-1} \epsilon_\mu \eta_i^\mu (\eta_j^\mu - f_\mu)$$

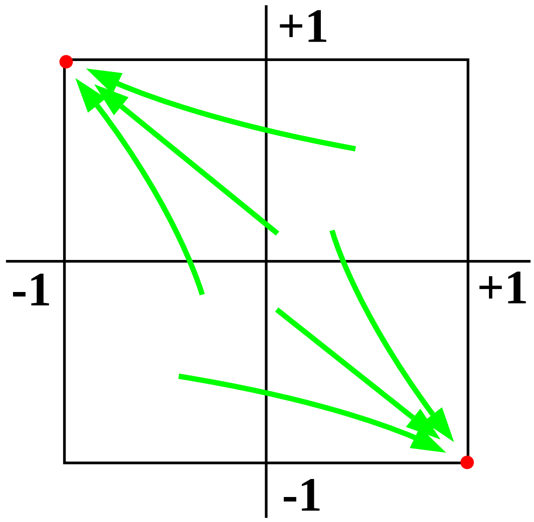
where W_{ij} is a random matrix that corresponds to the (sparse) connectivity in the absence of stored memories, c_{ij} is 1 if neuron j is connected to neuron i and 0 otherwise, a fraction f_μ of the components of η^μ are equal to 1, and N_μ neurons participate in the μ^{th} memory.

Randomly connected networks of excitatory and inhibitory neurons with no memories (all the ϵ_μ equal to zero) exhibit, over a broad range of parameters, a single stable state at low firing rate. We investigate, using both mean field analysis and simulations with spiking model neurons, the conditions for the formation of additional fixed points — new memories — as the ϵ_μ grow.

Supported by the Klingenstein foundation.

The problem

It is well known that idealized neurons can form attractor (Hopfield) networks:



2-neuron Hopfield network with fixed points at (+1, -1) and (-1, +1).

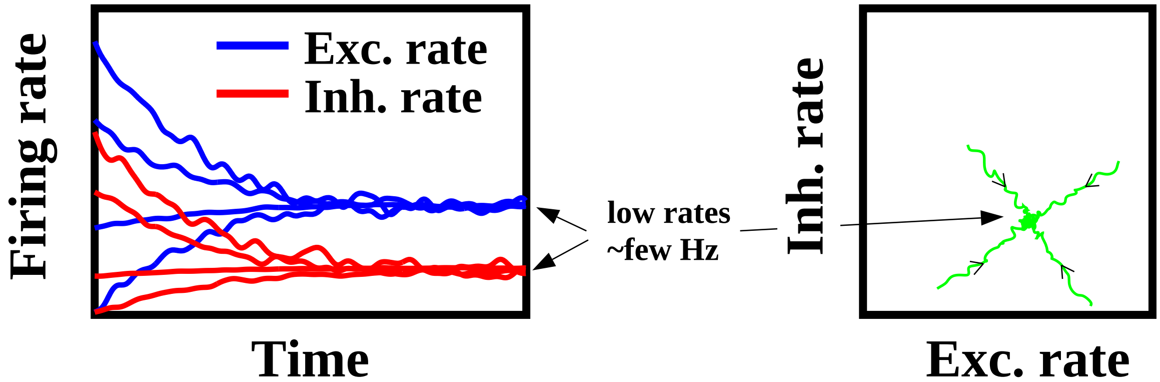
A beautiful model, but simplifications have been made:

- **Symmetric**
- **Units are +1 or -1**
- **All-all coupling**
- **Neurons are simple: no voltage gated channels ...**
- **Coupling is simple: no synapses or dendrites ...**

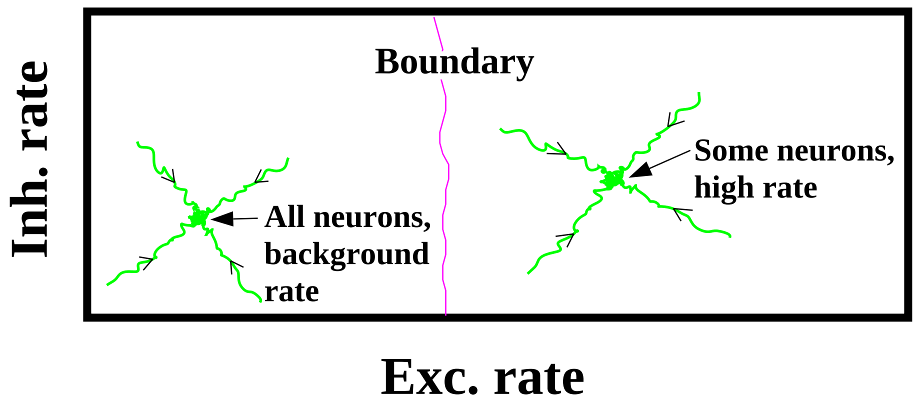
What about real, spiking, excitatory and inhibitory neurons with synaptic, non-symmetric coupling and sparse connectivity?

The Issues

Randomly connected excitatory and inhibitory neurons (often) have a globally attacting fixed point at low firing rate



Structured connectivity can embed memories



Constraints:

1. **If no memories are active, network fires at background rate.**
2. **At most, one memory can be active at a time.**

} **Important**

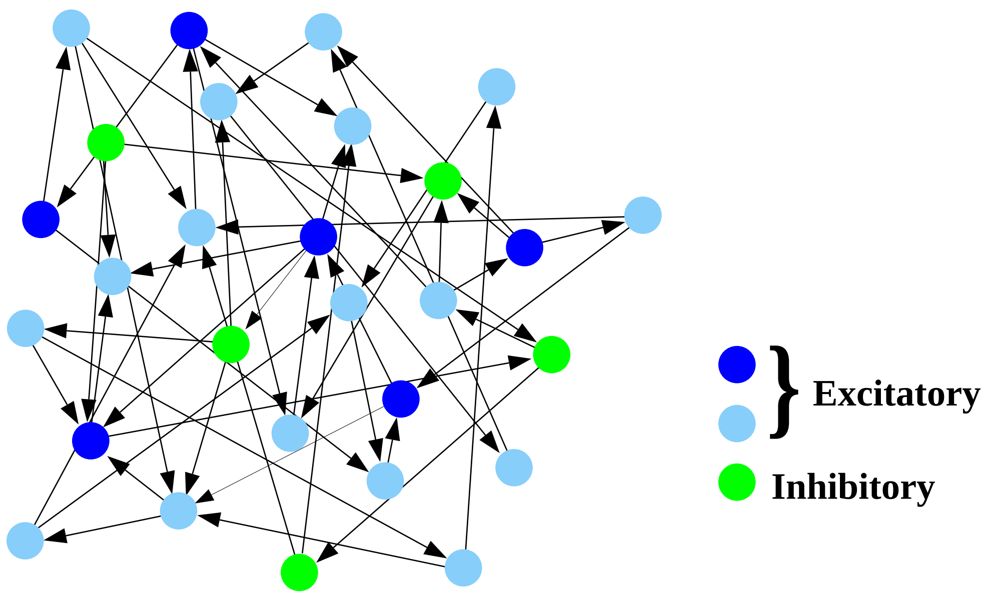
i. **Non-symmetric connectivity.**
ii. **Sparse connectivity.**

} **Not so important**

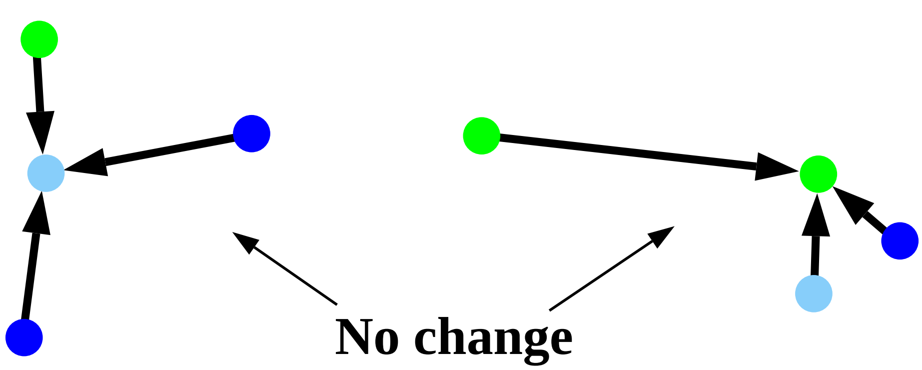
Can these constraints be satisfied?

The Prescription

Start with a randomly connected network:



Embed memories:

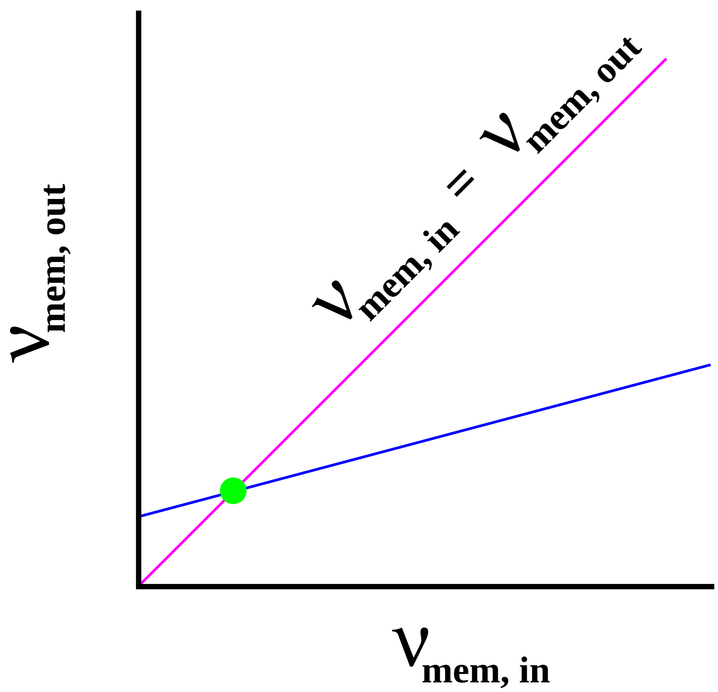


Analysis

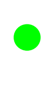
Take the limit *f* (fraction of neurons in a memory) $\rightarrow$ 0.

- **Each memory is all-excitatory network;**
- **Since *f* $\rightarrow$ 0, background firing rate is independent of firing rate of memory neurons.**

Can use (relatively) standard graphical techniques:



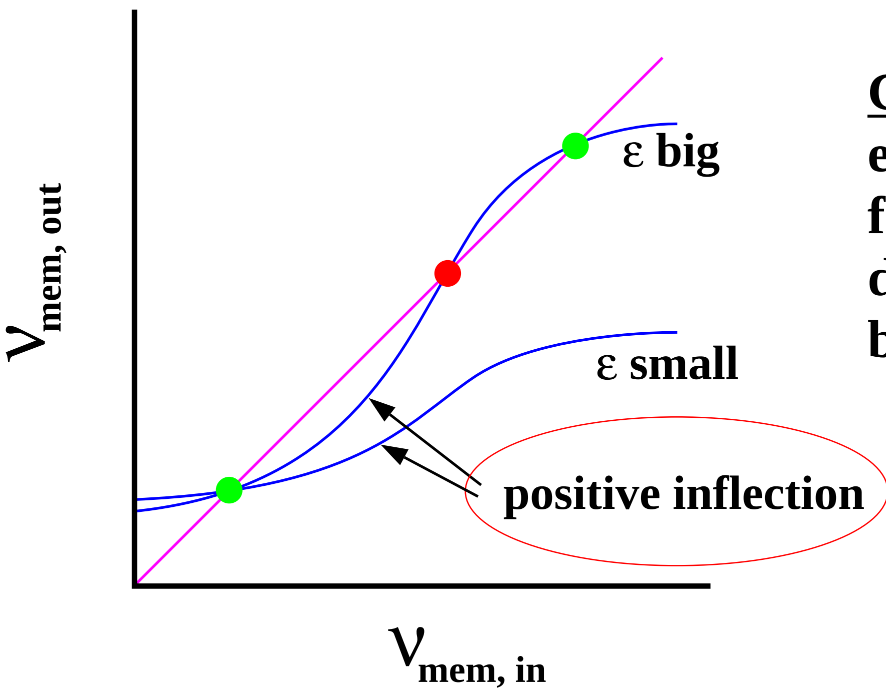
When strength of memory (i.e., the increase in connection strength among some subpopulation of neurons) is small, there is only one equilibrium and no memory is embedded.



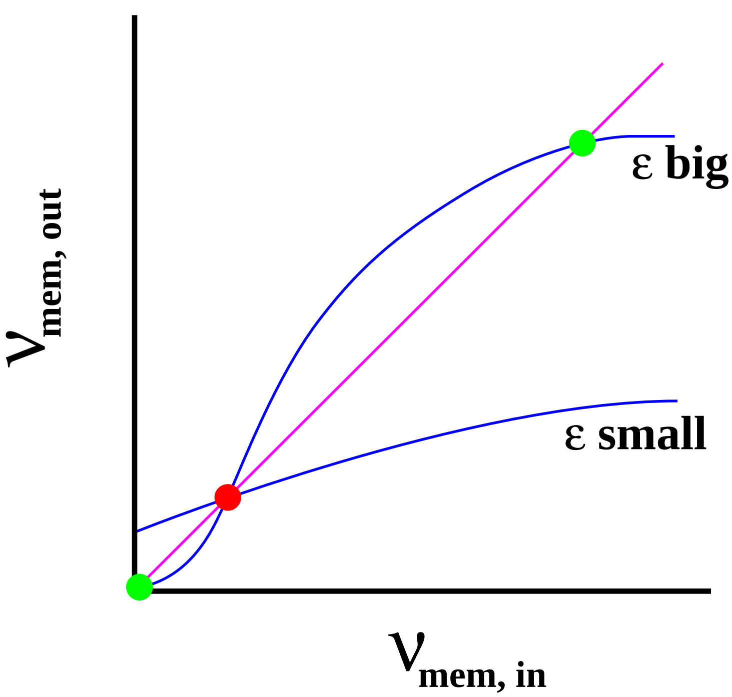
Stable equilibrium at background firing rate.

Gain functions: output firing rate of memory neurons as a function of input rate.

Two possibilities as ε increases:



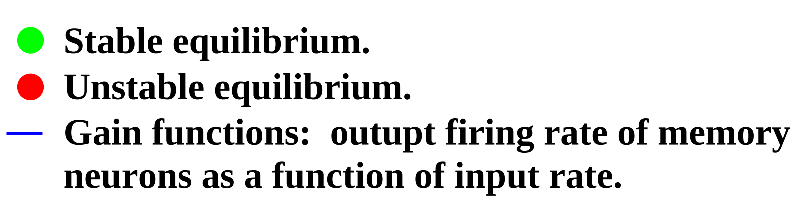
<u>Good:</u> Memory is embedded at high firing rate without disturbing the background.



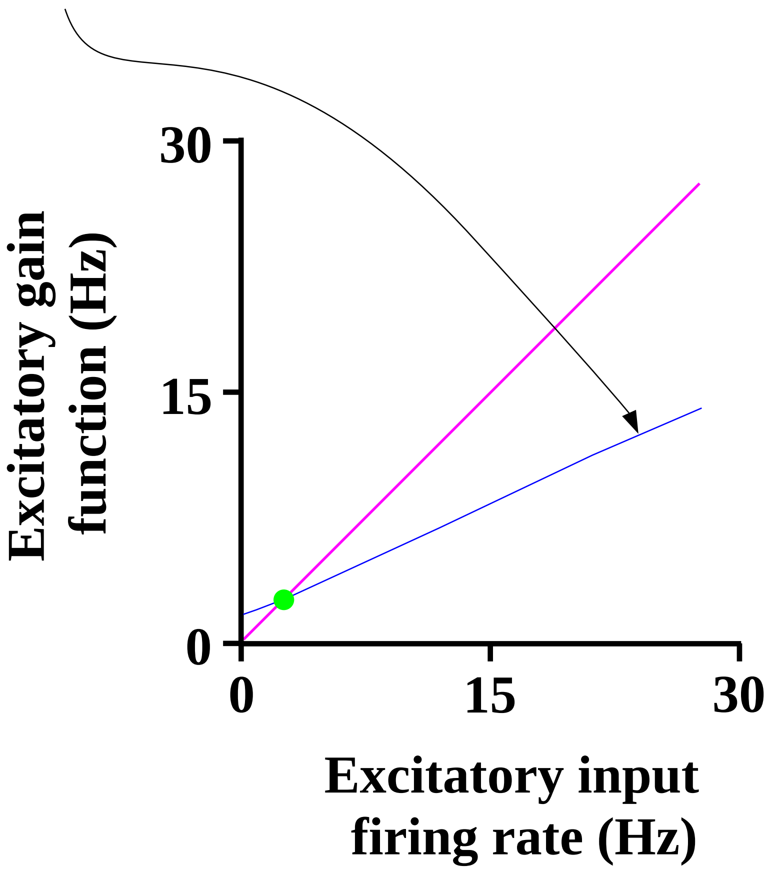
<u>Bad:</u>

1. Fluctuations typically destabilize new background.
2. Threshold is low -- this is a problem if you only want one memory to be active at a time.

2a. It’s also a problem if you want to avoid epilepsy ...



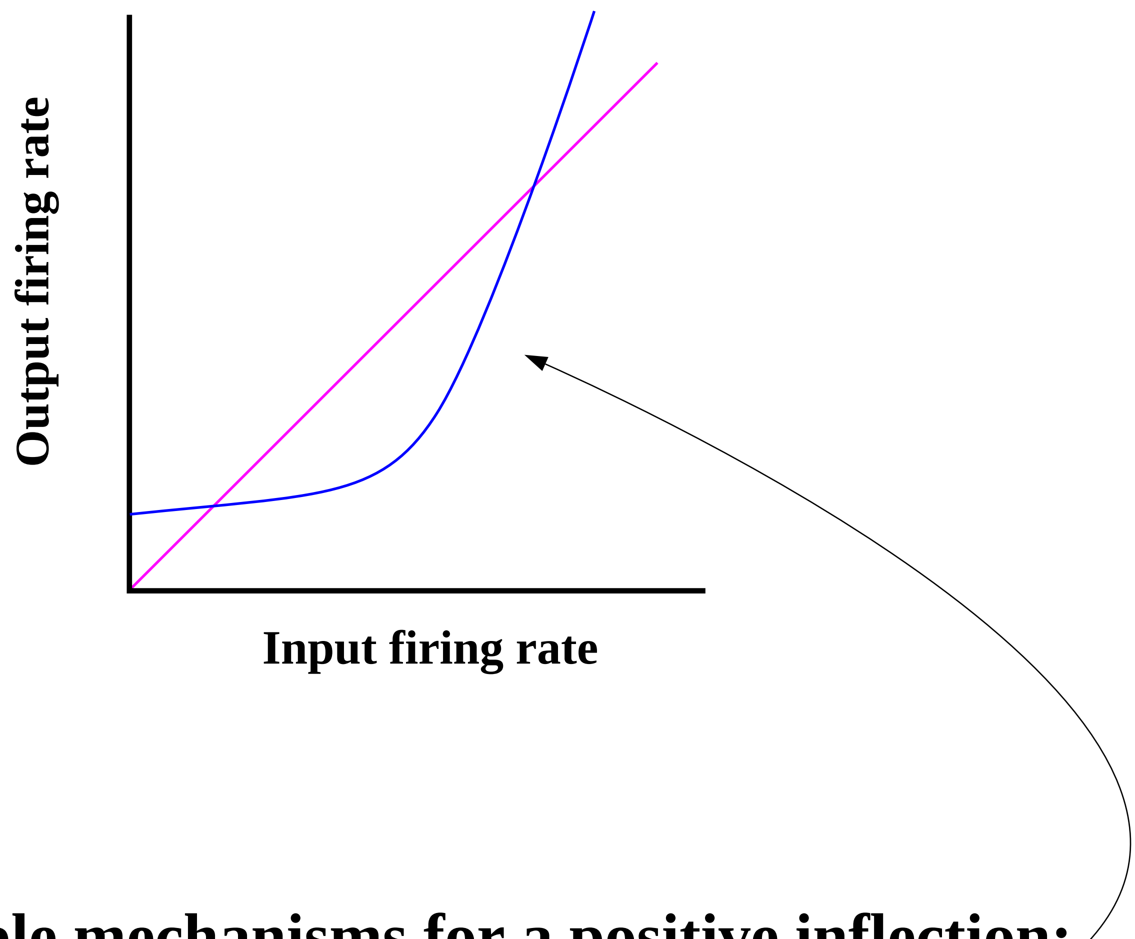
Gain curve from simulation with 10,000 θ-neurons



At equilibrium, no positive inflection

For details see: Latham et al, "Intrinsic dynamics in neuronal networks. I. Theory." Available at

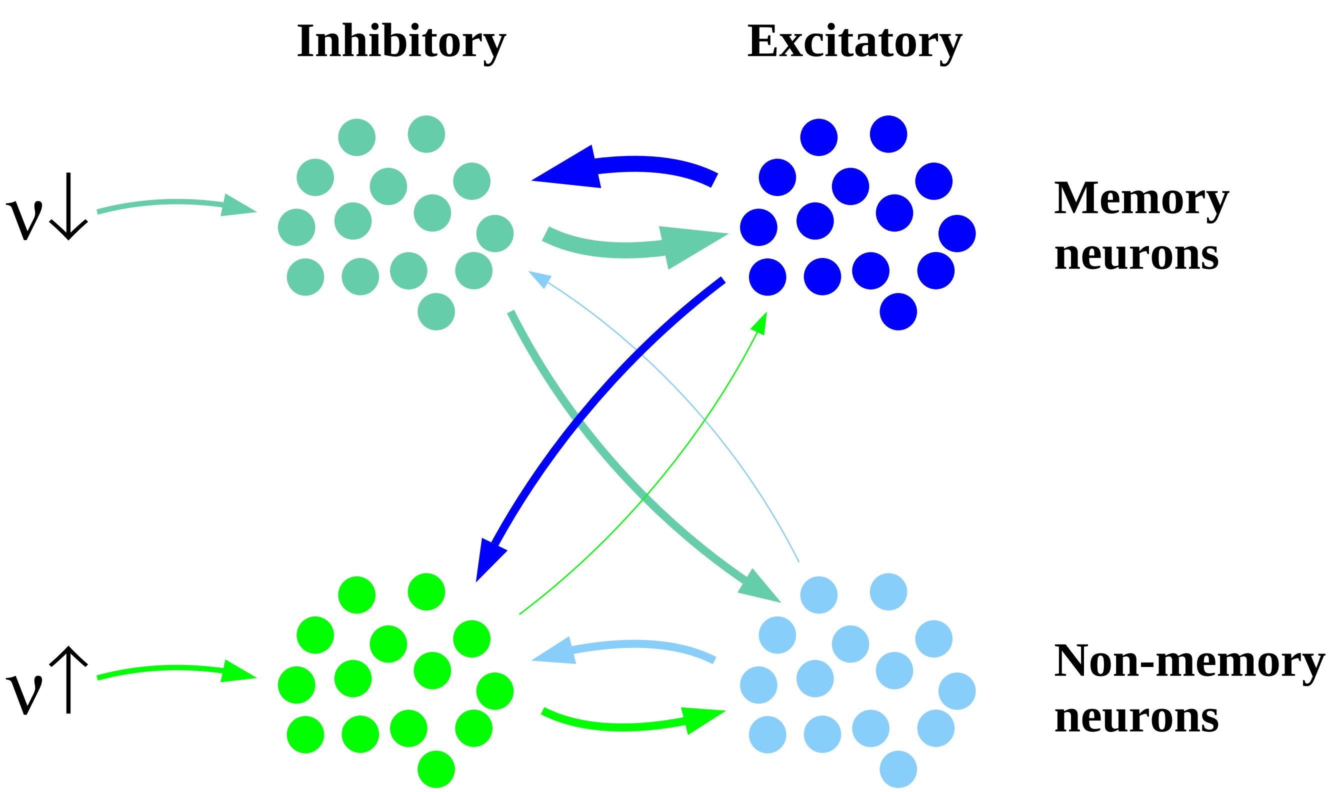
http://culture.neurobio.ucla.edu/~pel/



Possible mechanisms for a positive inflection:

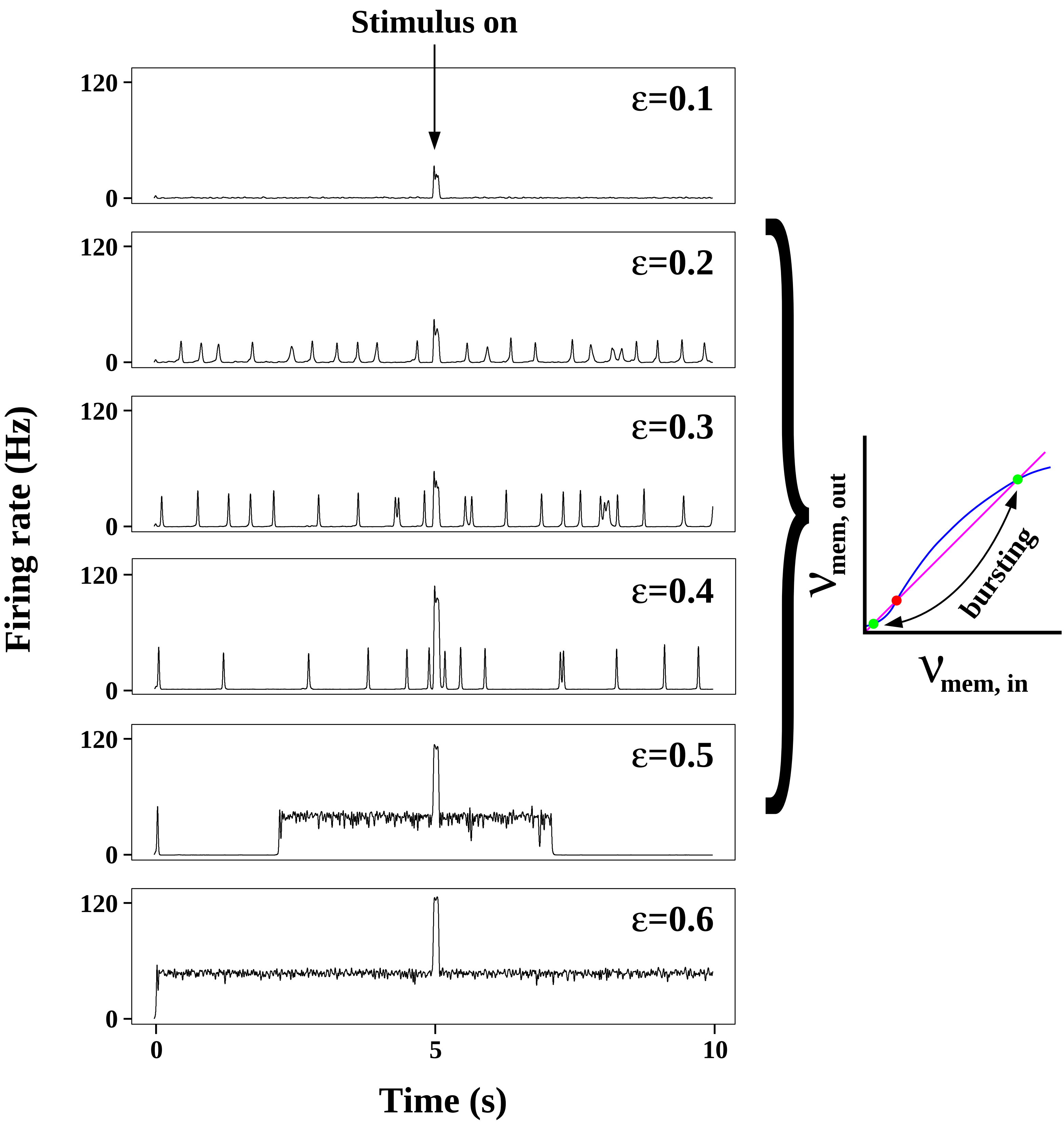
- **NMDA receptors,**
- **Paired-pulse facilitation.**

To enhance this effect, adjust connectivity so that the pool of inhibitory neurons that is firing at a relatively lower rate preferentially connects to the memory neurons:

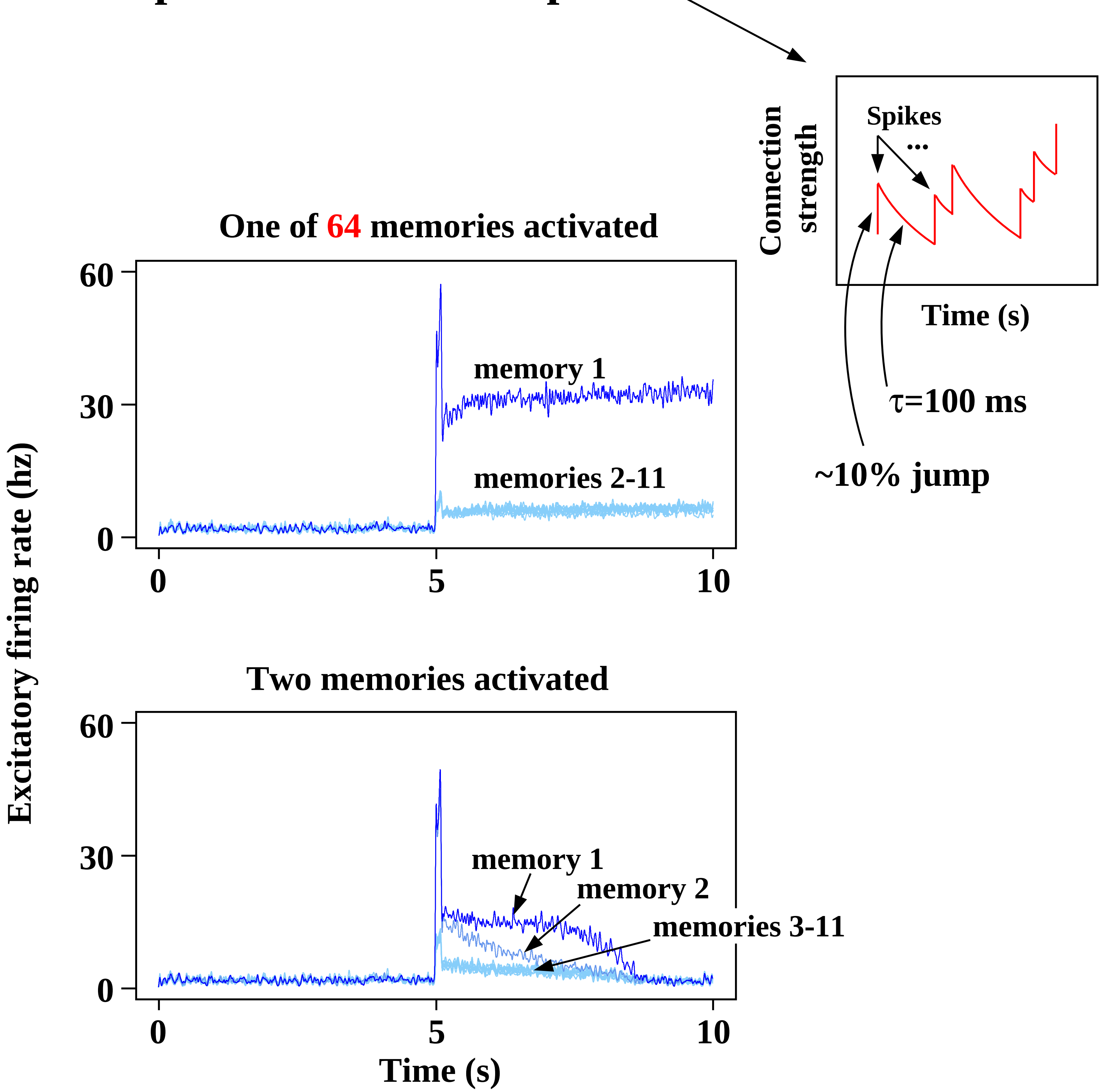


Simulations

10,000 spiking θ-neurons -- no NMDA channels



10,000 spiking θ-neurons with pseudo-NMDA receptors



For these parameters, network is sensitive to degredation of input. For a memory to last indefinitely:

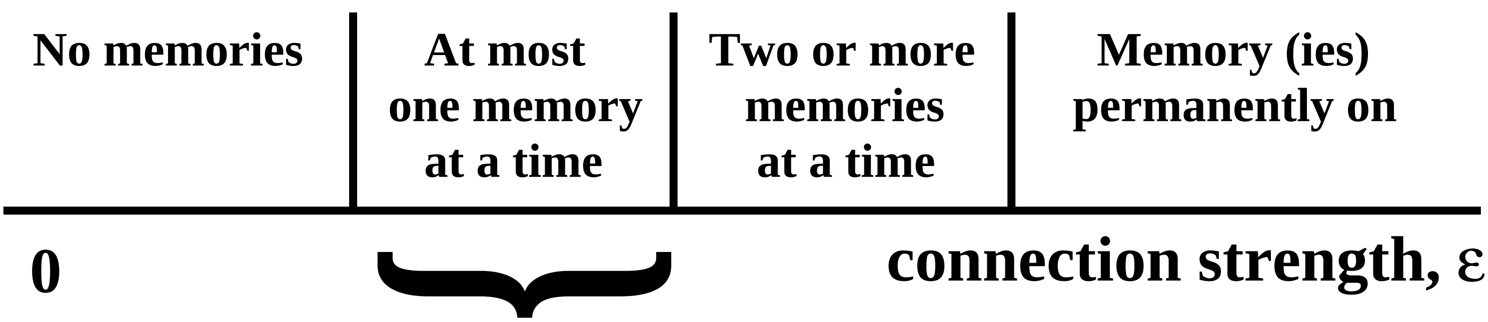
> 90% of the memory neurons must be activated
< 15% of non-memory neurons can be activated

Summary

- Can embed memories <u>if</u> the gain curve (input firing rate versus output firing rate) has a positive inflection at the backgound firing rate.

- **This will require something like NMDA channels or paired-pulse facilitation, for which the effective connection strength increases with post-synaptic voltage.**

<u>The picture:</u>



Desired regime.
In our simulations, this regime was small

Simulations with more realistic neurons are necessary!!

Mean Field Analysis

Equilibrium firing rate equations:

$$\nu_i = \Phi_i\left(\sum_j J_{ij}\,\nu_j\right)$$

Connectivity:

$$J_{ij} = W_{ij} + c_{ij}\sum_\mu N_\mu^{-1}\,\varepsilon_\mu\,\eta_i^\mu\left(\eta_j^\mu - f_\mu\right)$$

- W_{ij}: Sparse, random connectivity matrix
- c_{ij}: Sparse, random matrix of 0s and 1s.
- N_μ: Number of neurons in memory μ
- ε_μ: Strength of memory
- η_i^μ, η_j^μ: Determines which neurons are active during memory μ (vector of 1s and 0s)
- f_μ: Fraction of neurons in memory μ; post-synaptic normalization

Sources of randomness:

1. **W** and **c** -- random connectivity.
2. Non-active memories (see Chapter 10 of Hertz, Krogh and Palmer).

Define overlaps:

$$m^{\mu} = N_{\mu}^{-1} \sum_i \eta_i^{\mu} \nu_i$$

Perform suitable averaging, arrive at mean-field equations when 1 neuron is active:

$$m^{\mu} = \overline{\Phi}\left(C_0 \overline{\nu} + C_{\mu} m^{\mu}\right)$$

Mean firing rate (which has its own equilibrium equation)

Average is over random connectivity and non-active memories

Warning: the existence of excitatory neurons adds considerable algebra, but not much new conceptually.

Firing-rate-model simulations -- no memories