

Neural circuits for cognition

MIT 9.49/9.490/6.S076

Instructor: Professor Ila Fiete

TA: Gregg Heller

Senior Instruction Assistant: Adnan Rebei

Logistics/reminders

- HW # 2 posted; due Tuesday Oct 11 by midnight

Recall: Linear networks summary

- Every linear network has at most three fixed-point behaviors: single fixed point, no fixed points, infinitely many fixed points.
- The behavior is set by the eigenvalues of the system: eigenvalue < 1
- This is true for symmetric ($\mathbf{W}=\mathbf{W}^T$) and non-symmetric weights.
- However, dynamics can be quite different for symmetric and non-symmetric networks.

Recall: *Symmetric* linear networks summary

- Only the following are possible:
 - exponential decay dynamics to single fixed point, or exponential growth away, one behavior per eigendirection
 - If highly tuned along some eigendirection(s) → integrating/memory mode (line (plane or hyperplane) attractor along that direction(s))
- Oculomotor integrator: biological example of system which might operate as a line attractor.
- Nonlinear symmetric networks: permit finite multistability (e.g. bistable switch).

Today: Asymmetric networks

- Eigenvalues of matrix need not be real
- Complex eigenvalues → oscillatory/rotational dynamics
- Transient dynamics can also be different from simple exponential decay or growth

(non)Linear asymmetric networks

The balanced state

Network: Simple random connectivity

$$\tau \frac{d\mathbf{h}}{dt} = -\mathbf{h} + W\phi(\mathbf{h})$$

N neurons

$$\phi(\mathbf{h}) = \tanh(\mathbf{h})$$

$$W = \mathcal{N}(0, \frac{g^2}{N})$$

Not symmetric, since W_{ij} set independently of W_{ji}

Approximate E-I “balance” Strong weights ($\sim \frac{1}{\sqrt{N}}$)

Single parameter model: synaptic gain g

Fixed points

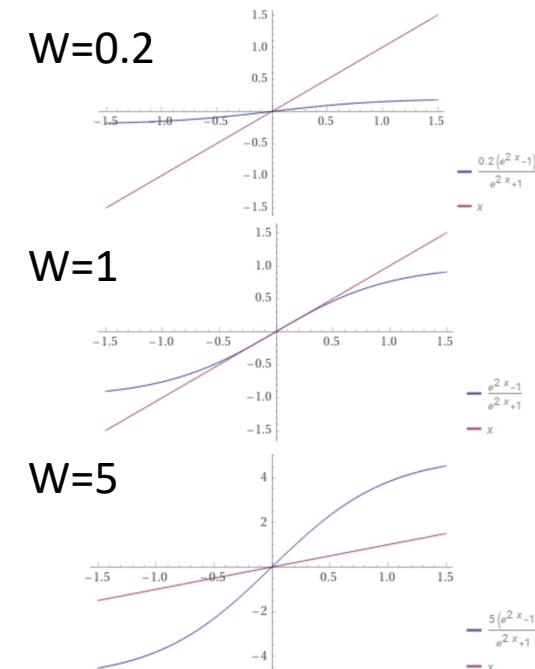
$$\tau \frac{d\mathbf{h}}{dt} = -\mathbf{h} + W\phi(\mathbf{h})$$

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} = \frac{e^{2z} - 1}{e^{2z} + 1}$$

$$\tanh(0) = 0$$

So $\mathbf{h} = 0$ is a fixed point of the dynamics

$$\phi(\mathbf{h}) = \tanh(\mathbf{h})$$



Stability of the $\mathbf{h} = 0$ fixed point? Linearize at 0

$$\tau \frac{d\delta\mathbf{h}}{dt} = -\delta\mathbf{h} + W\phi'(\bar{\mathbf{h}})\delta\mathbf{h}$$
$$W = \mathcal{N}(0, \frac{g^2}{N})$$

$$\left. \frac{d \tanh(z)}{dz} \right|_{z=0} = \left. \frac{4e^{2z}}{(e^{2z} + 1)^2} \right|_{z=0} = 1$$

$$\tau \frac{d\delta\mathbf{h}}{dt} = -\delta\mathbf{h} + W\delta\mathbf{h}$$

All eigenvalues of $W < 1 \rightarrow$ stable.
Any eigenvalue $> 1 \rightarrow$ unstable.

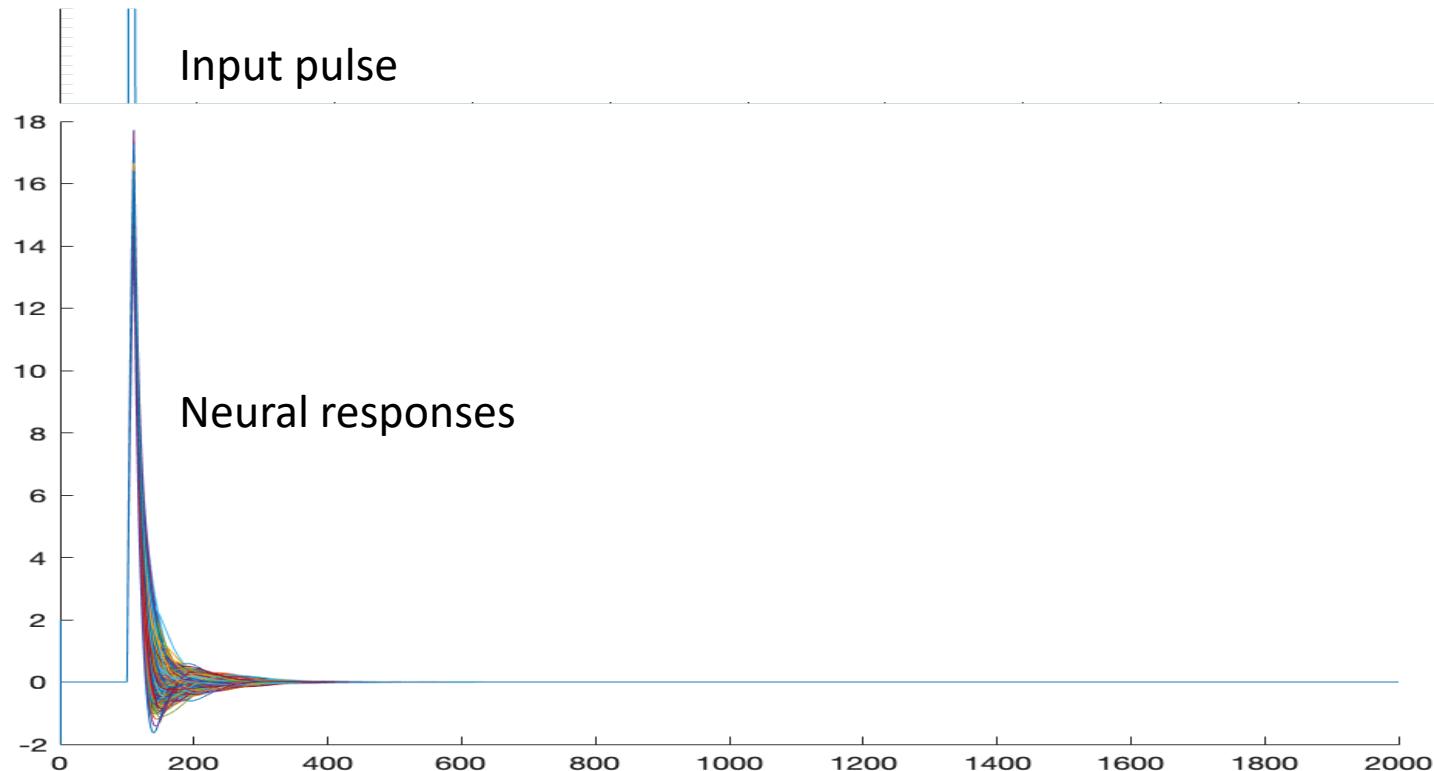
Code it up:

Explore dynamics

Dependence of dynamics on g

State-space plots

What we observed: Stability of fixed point at 0

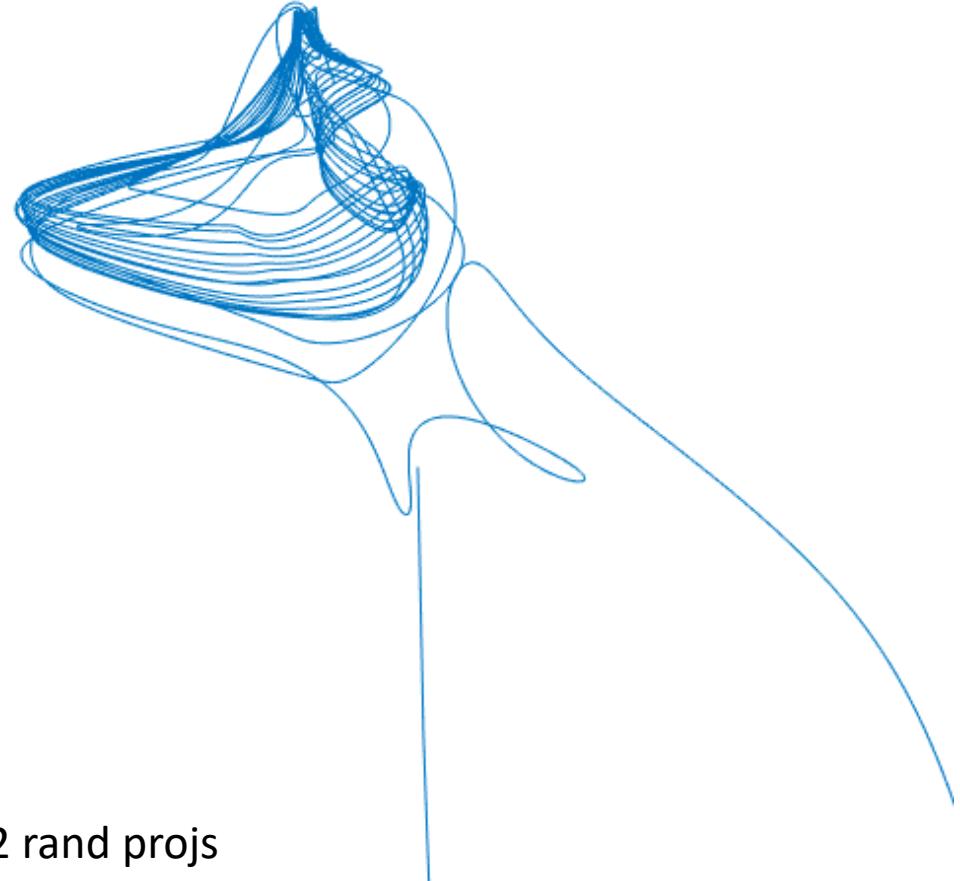


$g = 0.9$; $n = 512$

Visualizing dynamics of high-dim system: plot low-dimensional projections

Well-known method: PCA -- use the highest-variance directions (eigenvectors with largest eigenvalues)

Johnson-Lindenstrauss lemma says random low-dimensional projections can be sufficient



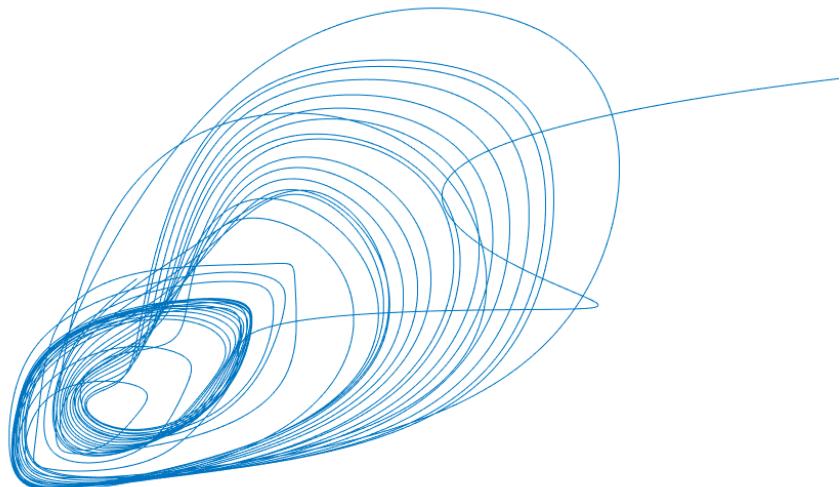
M randomly placed points in a space of dimension N can be embedded with preservation of relative distances (distortion $< \varepsilon$) *through random projections* into a space of dimension $> K = \log(M)/\varepsilon^2$.

This carries over to smooth manifolds of linear dimension M ($\leq M$ linearly independent directions along the manifold).

Closely related to compressive sensing.

Visualizing dynamics in state space even if system only partially observed: Takens's delay embedding theorem

Takens, 1981

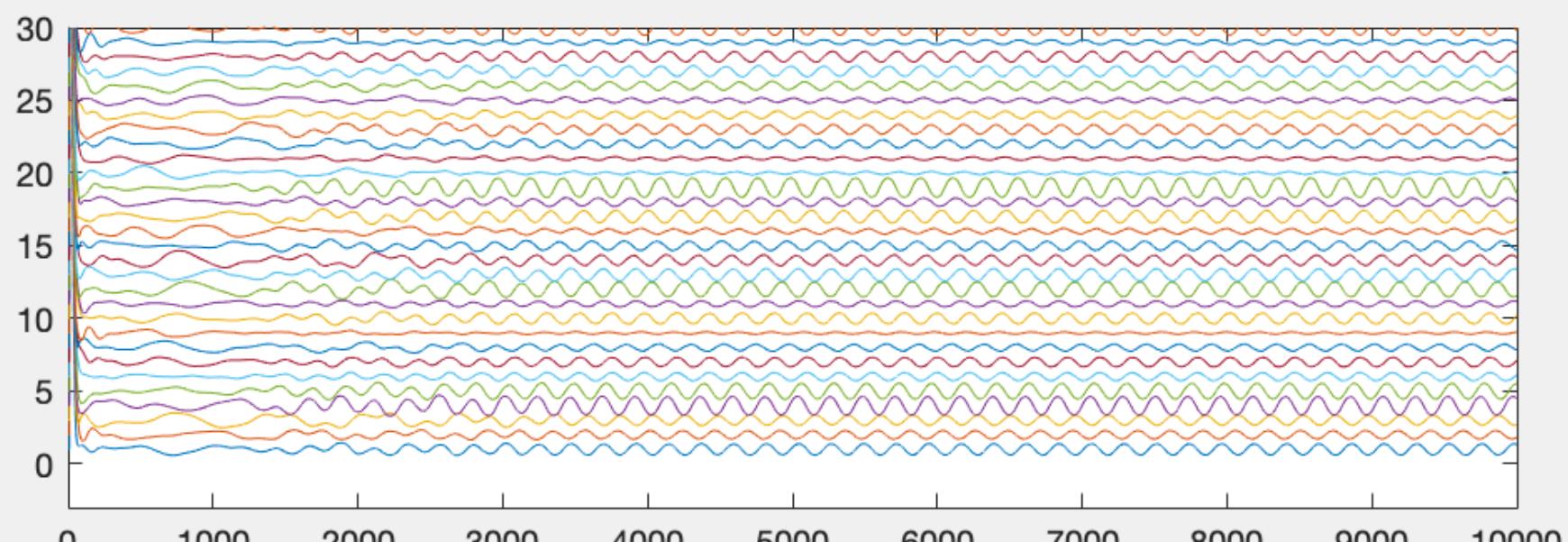


1 cell vs delayed version of itself
Same pop trajectory

Consider a d -dimensional state vector \mathbf{y} evolving deterministically over time. Now consider trajectories of the vector $\mathbf{y}(t)$ and trajectories of the vector $(y_i(t), y_i(t + \tau), \dots, y_i(t + n\tau))$ for some scalar projection j of the dynamics (e.g. one-dimensional observations). The two trajectories are the same, upto a diffeomorphism (smooth, invertible coordinate change), when $n \geq 2d + 1$ (and often less).

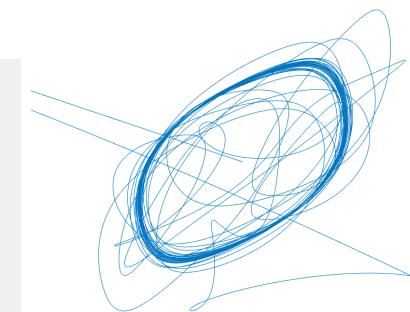
Almost magic! Under relatively generic conditions, can reconstruct all $d-1$ unobserved dims from enough delayed observations of single dim!

A periodic regime: limit cycles



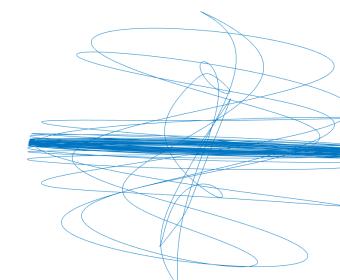
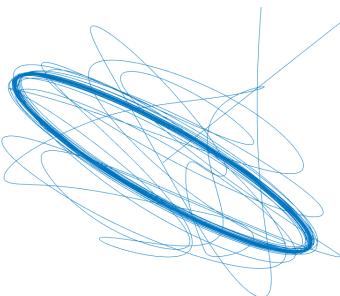
$N=512; g = 1.1$

Two random
cells



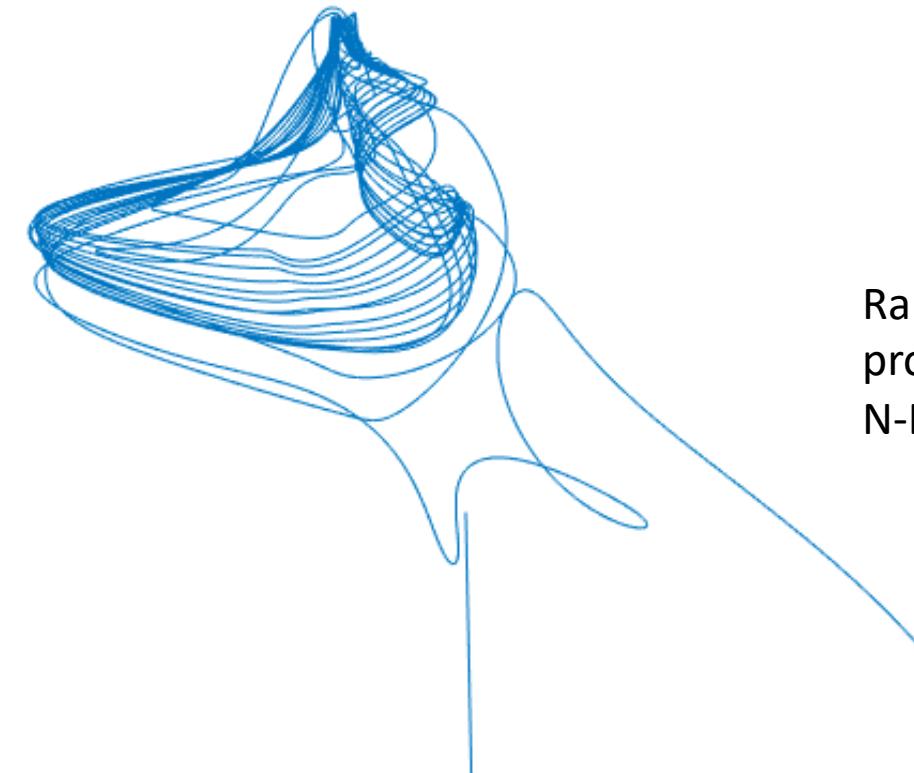
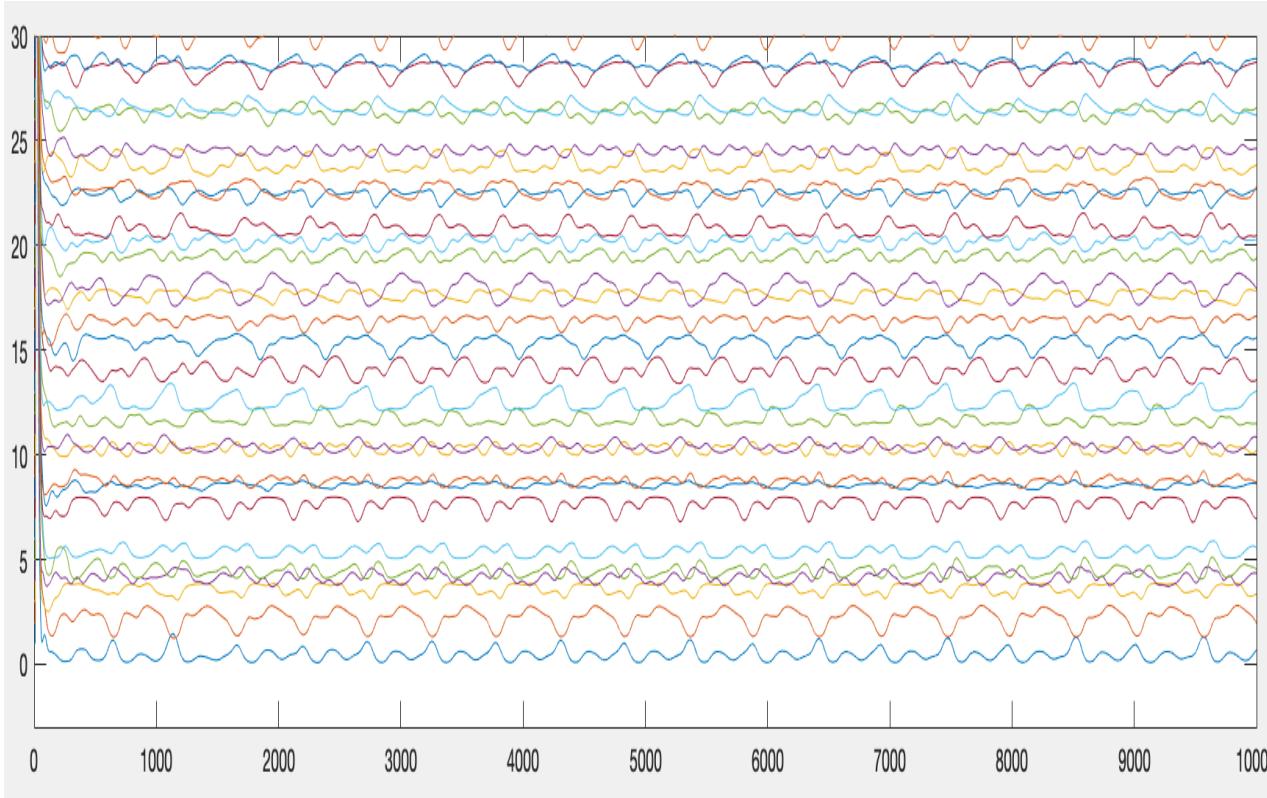
PCA proj 2

Random 2-D
projection of
N-D dynamics



PCA proj 3

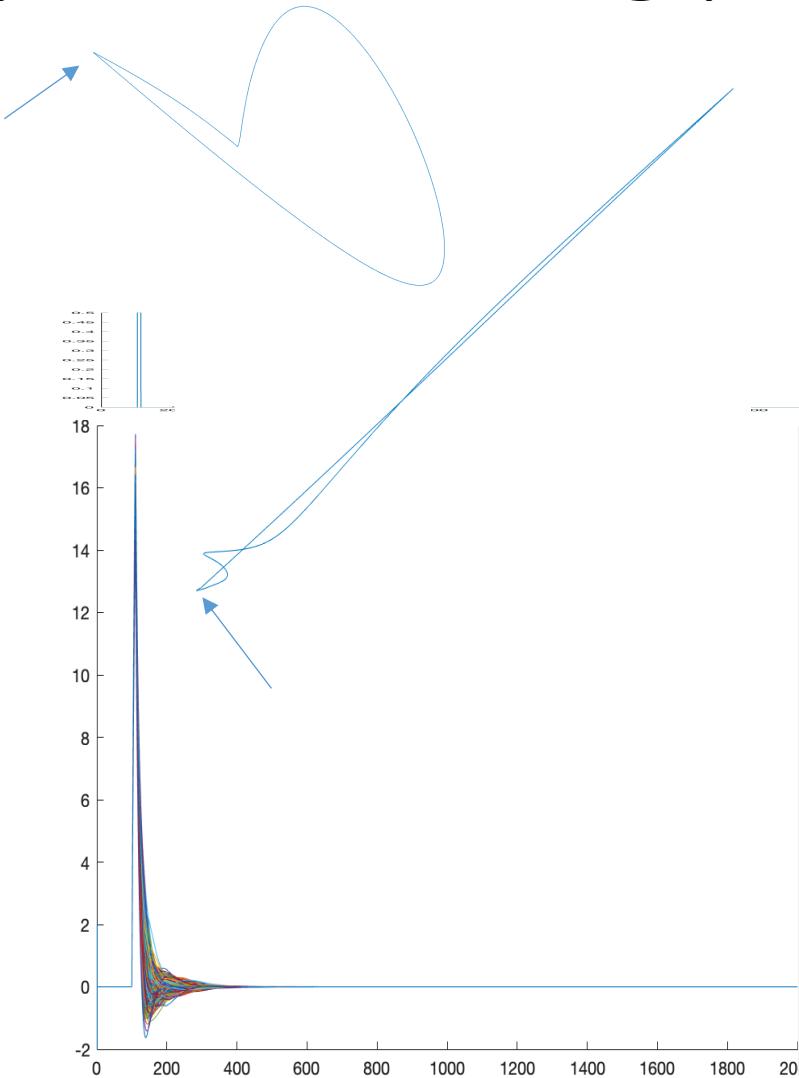
A non-periodic/chaotic regime



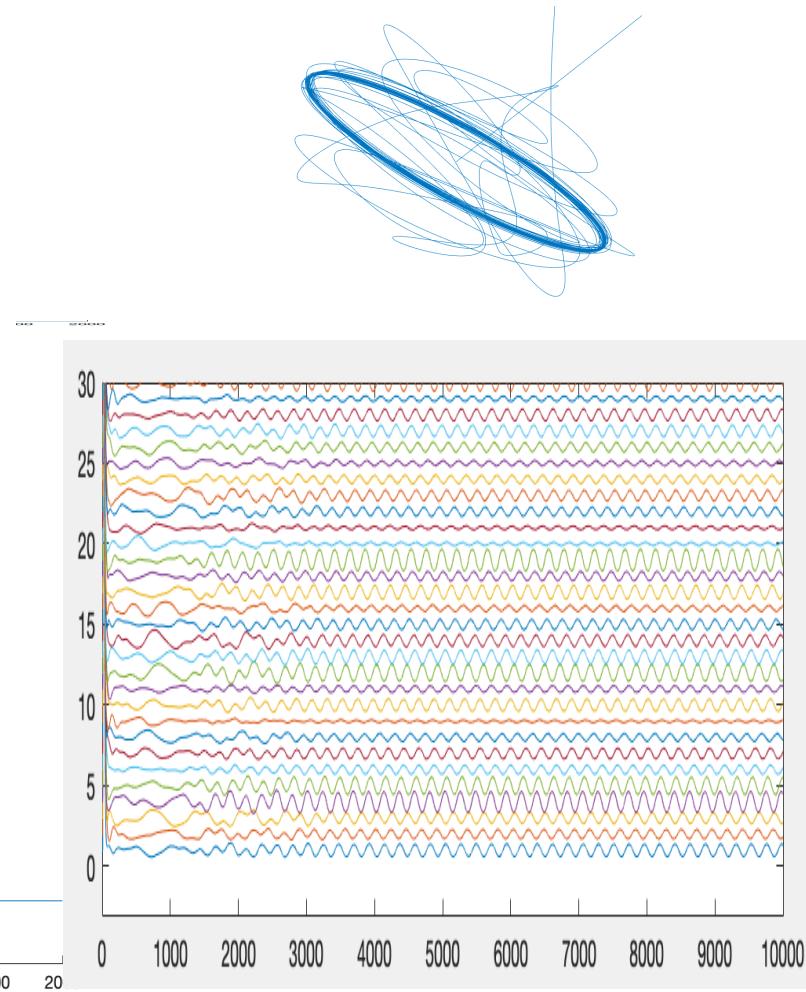
Random 2-D
projection of
N-D dynamics

$N=512; g=1.4$

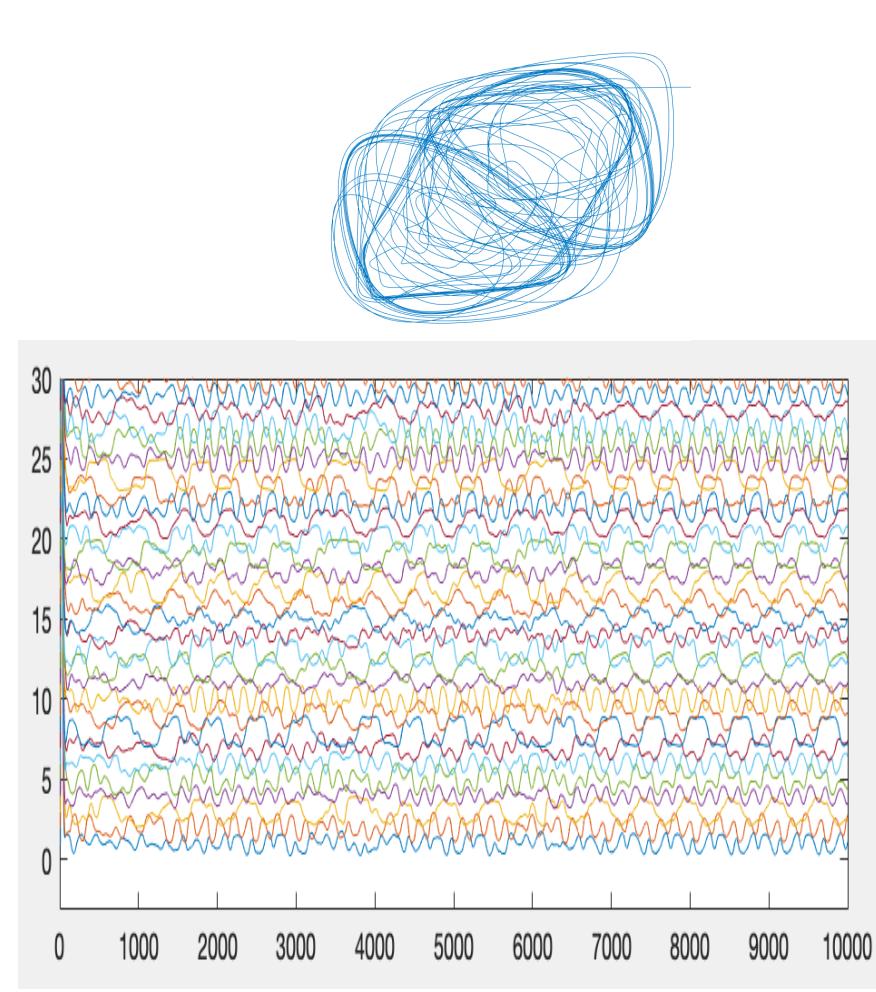
Synaptic scaling parameter g



$g = 0.9; n = 512$



$g = 1.2; n = 256$

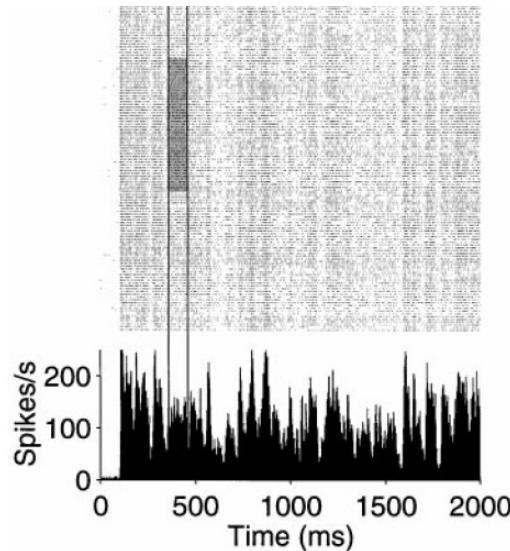


$g = 1.4; n = 512$

Could the brain be operating in a chaotic regime?

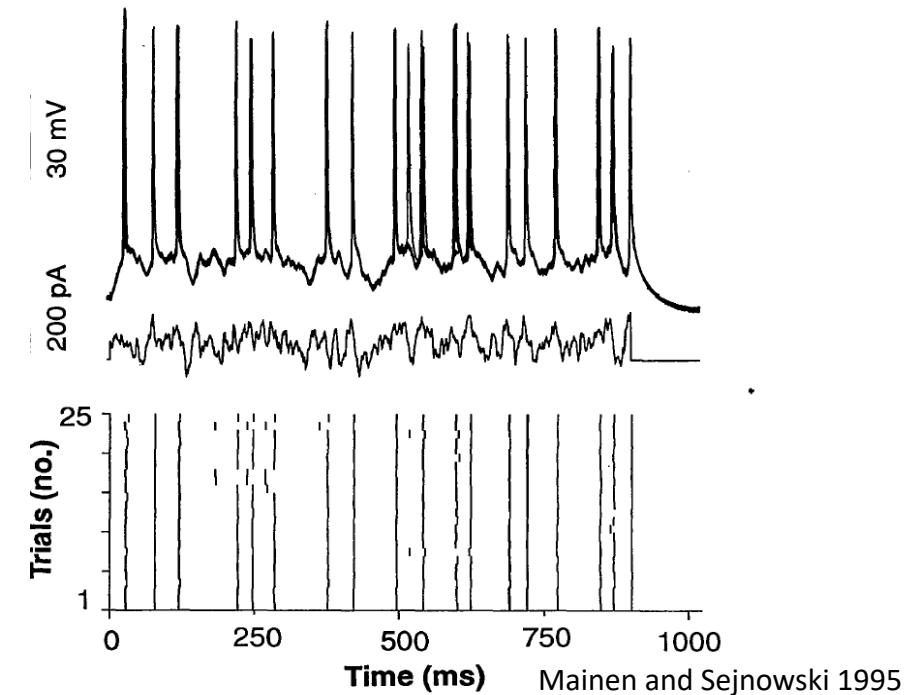
Observation: Neural spiking is highly variable

Variable spiking on repeated response to same stimulus



Softky and Koch, 1993
Shadlen and Newsome, 1998

..... But not because of spike-generation stochasticity in single cells



Mainen and Sejnowski 1995

This suggests that synaptic input fluctuations might drive fluctuations in spiking

Suppression of variance in large networks driven by non-zero mean input

The mean input to each neuron (connected to K neurons) in N-neuron network is:

$$g_i = \sum_j^K W_{ij} s_j$$

If $W_{ij} = \frac{W}{K}$ and $\langle s_i \rangle = m$ then the mean input to each neuron is:

$$\langle g_i \rangle = Wm$$

The variance of input into each neuron is:

$$\begin{aligned} \langle (g_i - \langle g \rangle)^2 \rangle &= \left(\frac{W}{K} \right)^2 \left\langle \sum_{j,k} (s_j - m)(s_k - m) \right\rangle = \left(\frac{W}{K} \right)^2 K(m - m^2) \\ &= \frac{\langle g \rangle^2}{K} \frac{(1 - m)}{m} \end{aligned}$$

iid Poisson
assumption
for s

Ratio of input standard deviation to mean tends to 0

Input mean: $\langle g_i \rangle = Wm$

Input variance: $\langle (g_i - \langle g \rangle)^2 \rangle = \frac{\langle g \rangle^2}{K} \frac{(1-m)}{m}$

Ratio: $\frac{var(g)}{\langle g \rangle^2} = \frac{1}{K} \frac{(1-m)}{m} \rightarrow 0 \text{ as } K \rightarrow \infty$

Cell is strongly mean-driven:
vanishing effect of fluctuations

For typical CNS neuron: $K \sim 10^2 - 10^5$

Weight scaling and balanced +, - inputs

$$g_i = \sum_j^K W_{ij}^E s_j^E + \sum_j^K W_{ij}^I s_j^I$$

One cell in network; E, I inputs
N neurons, connectivity $K < N$

If $W_{ij}^E = \frac{W^E}{\sqrt{K}}$, $W_{ij}^I = -\frac{W^I}{\sqrt{K}}$ and $\langle s_i \rangle = m$ then the mean input to each neuron is:

$$\langle g_i \rangle = \frac{W^E}{\sqrt{K}} K m^E - \frac{W^I}{\sqrt{K}} K m^I = \sqrt{K}(W^E m^E - W^I m^I)$$

The variance of input into each neuron is:

$$\langle g_i^2 \rangle = (W^E)^2 m^E (1 - m^E) + (W^I)^2 m^I (1 - m^I)$$

independent of K, remains finite

Single cell: Weight scaling and balanced +, - inputs for variance-driven regime

$$g_i = \sum_j^K W_{ij}^E s_j^E + \sum_j^K W_{ij}^I s_j^I$$

One cell in network; E, I inputs
 N neurons, connectivity $K < N$

If $W_{ij}^E = \frac{W^E}{\sqrt{K}}$, $W_{ij}^I = -\frac{W^I}{\sqrt{K}}$ and $\langle s_i \rangle = m$

Ratio of variance to squared mean is:

$$\frac{\text{var}(g)}{\langle g \rangle^2} = \frac{\text{const}}{(\sqrt{K}(W^E m^E - W^I m^I))^2}$$

If tune W's so that $(W^E m^E - W^I m^I) \sim \frac{1}{\sqrt{K}}$

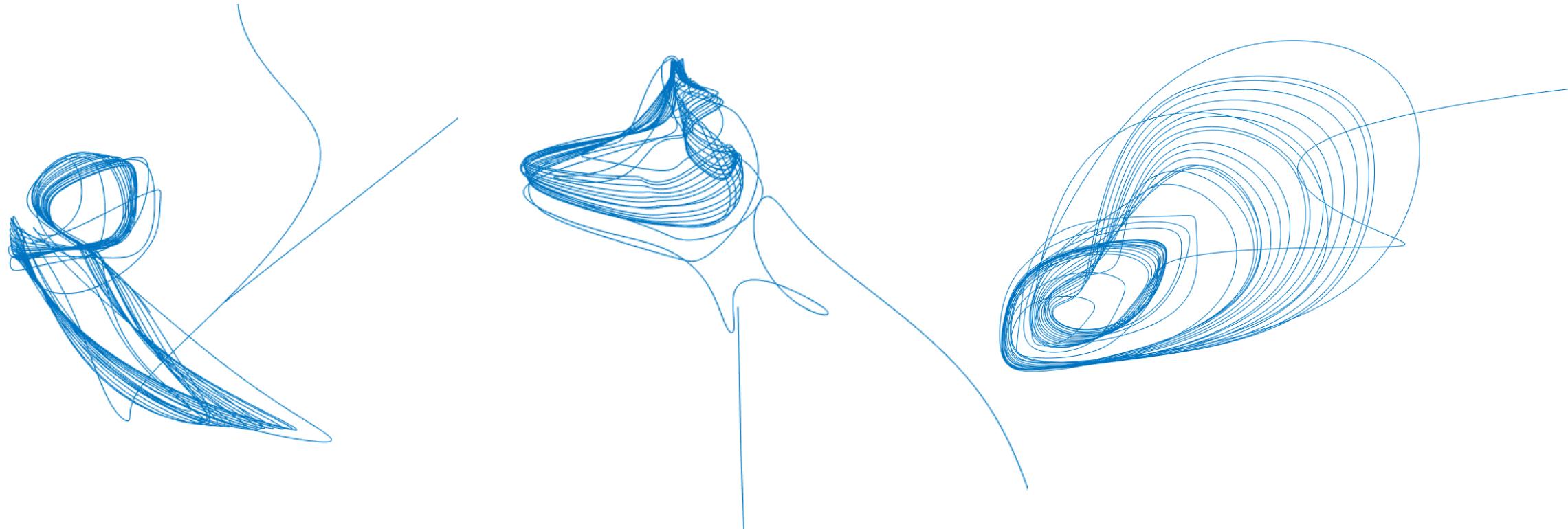
then mean term is $O(1)$, indep of K ,
and the ratio of variance to squared
mean is remains finite at all K

→ Balanced regime

Ingredients that permit variance in highly-connected neurons

- Strong weights: they scale as $\frac{1}{\sqrt{K}}$ rather than $\frac{1}{K}$
- Near-cancellation of strong positive ("excitatory") and strong negative ("inhibitory") inputs.
- So far, considered single neurons whose inputs we controlled by hand. Can these conditions remain true in a self-consistent way in a recurrent network?

The non-periodic/chaotic state



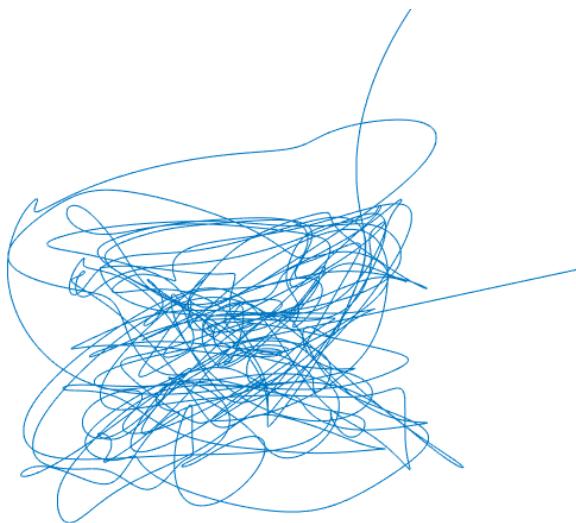
2 randomly chosen
cells, same pop traj

Random 2-D
projection of
N-D dynamics

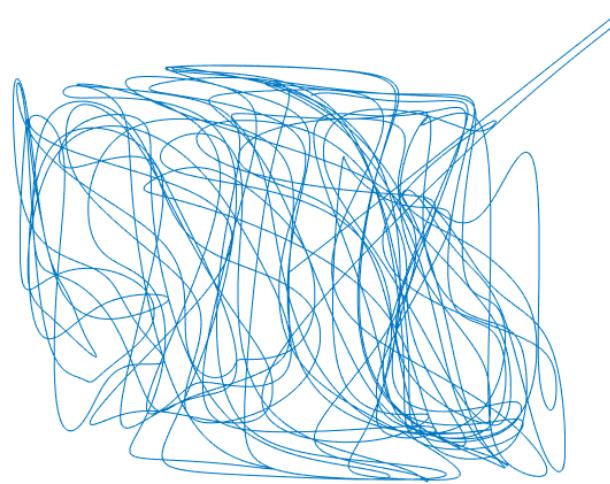
1 cell versus delayed version of itself

N=512; g=1.4

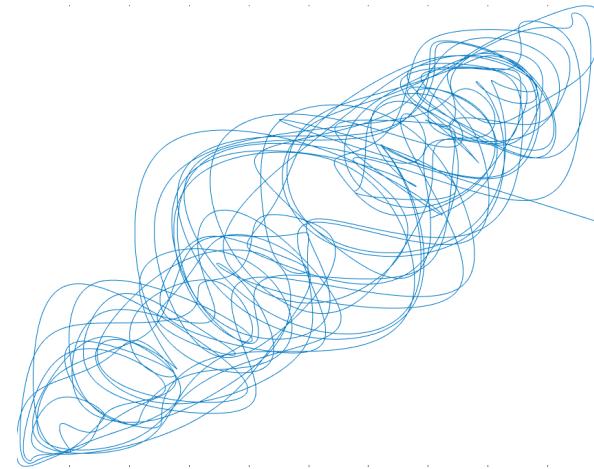
Another run of the non-periodic/chaotic state



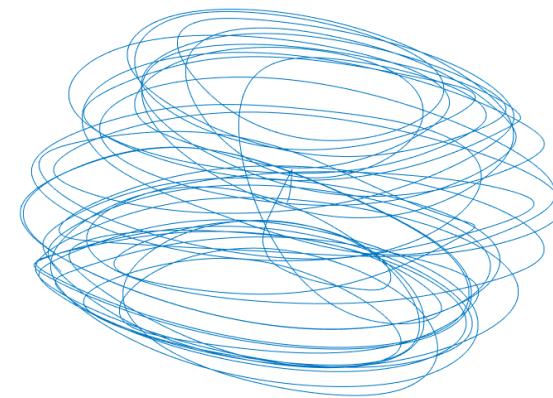
2 randomly chosen
cells, same pop traj



Random 2-D
projection of
N-D dynamics



1 cell activity against
delayed version of
itself, same pop traj



PC2, PC3

Stability of $h = 0$ fixed point with random Gaussian matrix

For a random Gaussian matrix $M \in \mathcal{R}^{N \times N}$, $M_{ij} = \mathcal{N}(0, \sigma^2)$

eigenvalues lie on a disc of radius $\sigma\sqrt{N}$ in the complex plane.

Ginibre 1965; Girko 1984;
Tao & Vu 2008

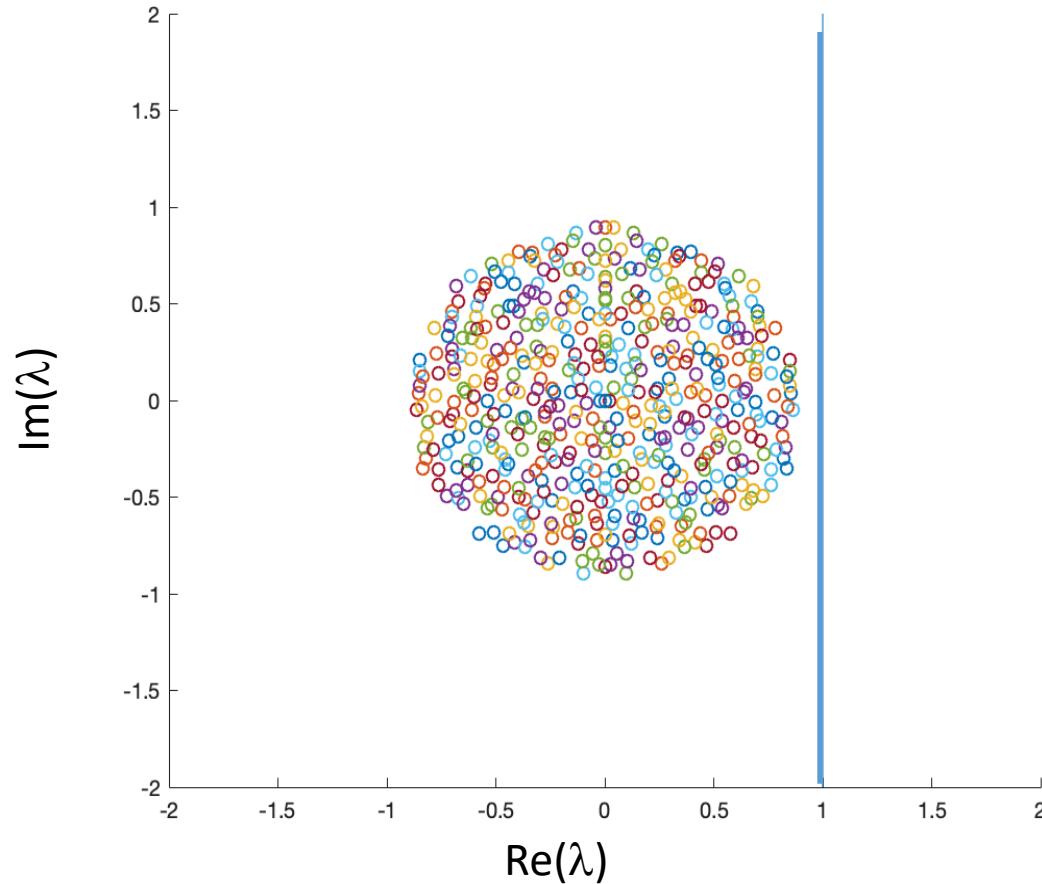
For \mathbf{W} with $\sigma = \frac{g}{\sqrt{N}}$: eigenvalue disc has radius g

Thus $h = 0$ fixed point becomes unstable for $g > 1$ (exact in large N limit).

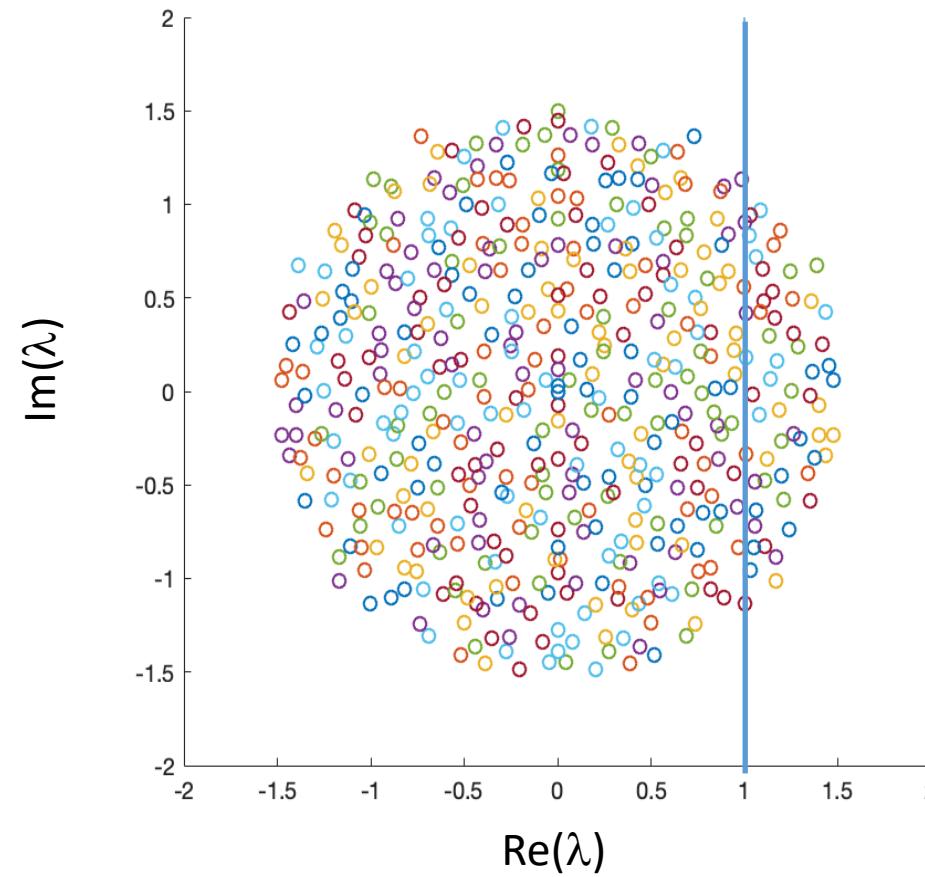
Related:

- Semicircular law: symmetric random gaussian matrices (1955 Wigner)
<https://terrytao.wordpress.com/2010/02/02/254a-notes-4-the-semi-circular-law/>
- Elliptical law: random matrices with correlated entries (1988; Somers, Crisanti, Sompolinsky)

Eigenvalues of the random matrix: complex-valued, in a disc of radius g

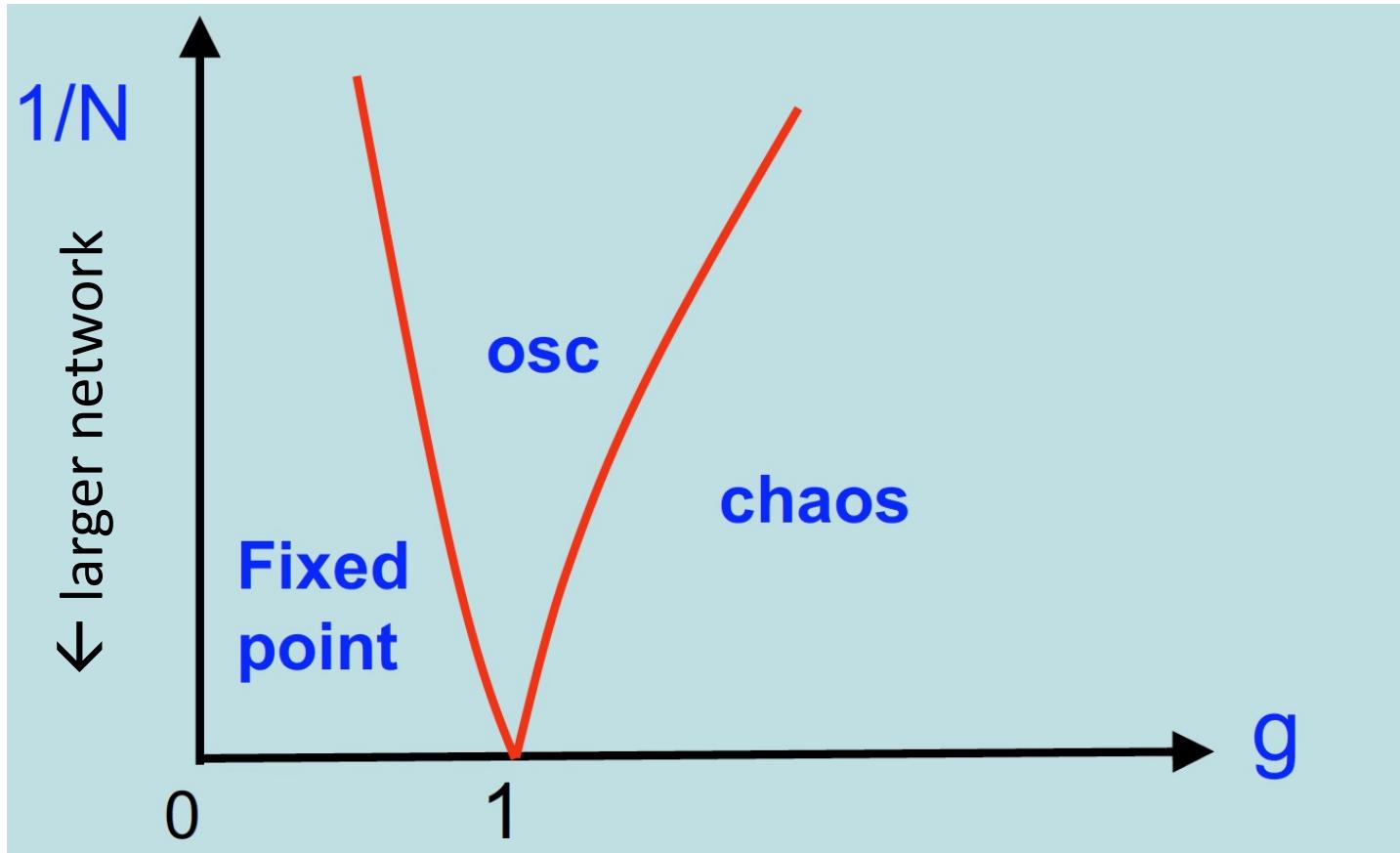


$g = 0.9; n = 512$



$g = 1.5; n = 512$

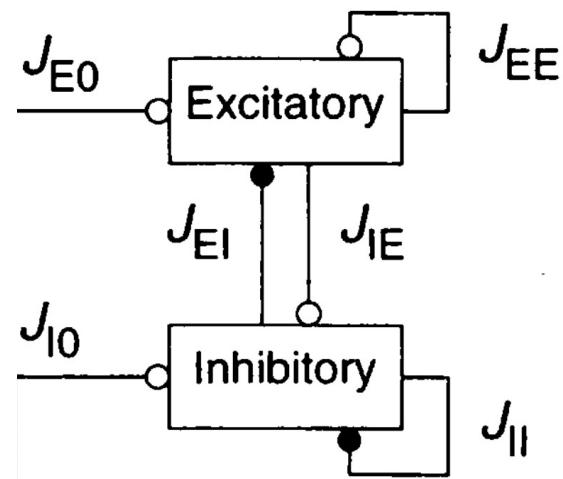
Phase diagram



- Weak coupling: single fixed point at 0.
- Small network, intermediate coupling: periodic.
- Strong coupling: chaos

Biologically-plausible balanced networks

- Dale's law: a neuron releases the same (set of) neurotransmitters at all its synapses.
- Separate E, I cells: Entries of weight matrix within a column should all have the same sign.
- These networks are far more non-normal than the random Gaussian balanced networks.



E-I balanced network equations

$$\tau \frac{d\mathbf{h}}{dt} = -\mathbf{h} + J\phi(\mathbf{h}) + \mathbf{h}^0 \quad \phi(x) = [x]^+$$

$$J_{ij} = \frac{J_{\alpha\beta}}{\sqrt{K}} C_{ij} + \eta_{ij}; \quad \alpha, \beta \in \{E, I\}$$

$$C_{ij} \in \{0, 1\} : \langle C_{ij} \rangle = \frac{K}{N}$$

$$C_{ij} = \frac{K}{N} + \delta C_{ij}$$

$$\implies J_{ij} = \underbrace{\frac{\sqrt{K}}{N} J^{\alpha\beta}}_{\text{"structured"}} + \underbrace{\frac{J^{\alpha\beta}}{\sqrt{K}} \delta C_{ij}}_{\text{random}} + \eta_{ij}$$

“structured”

random

$$\bar{J}^{\alpha\beta}$$

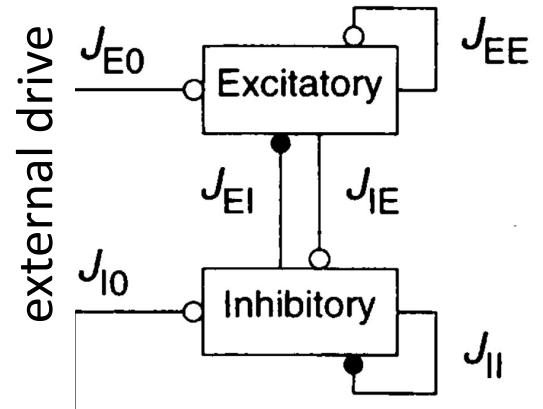
weak

$$\delta J_{ij}$$

strong

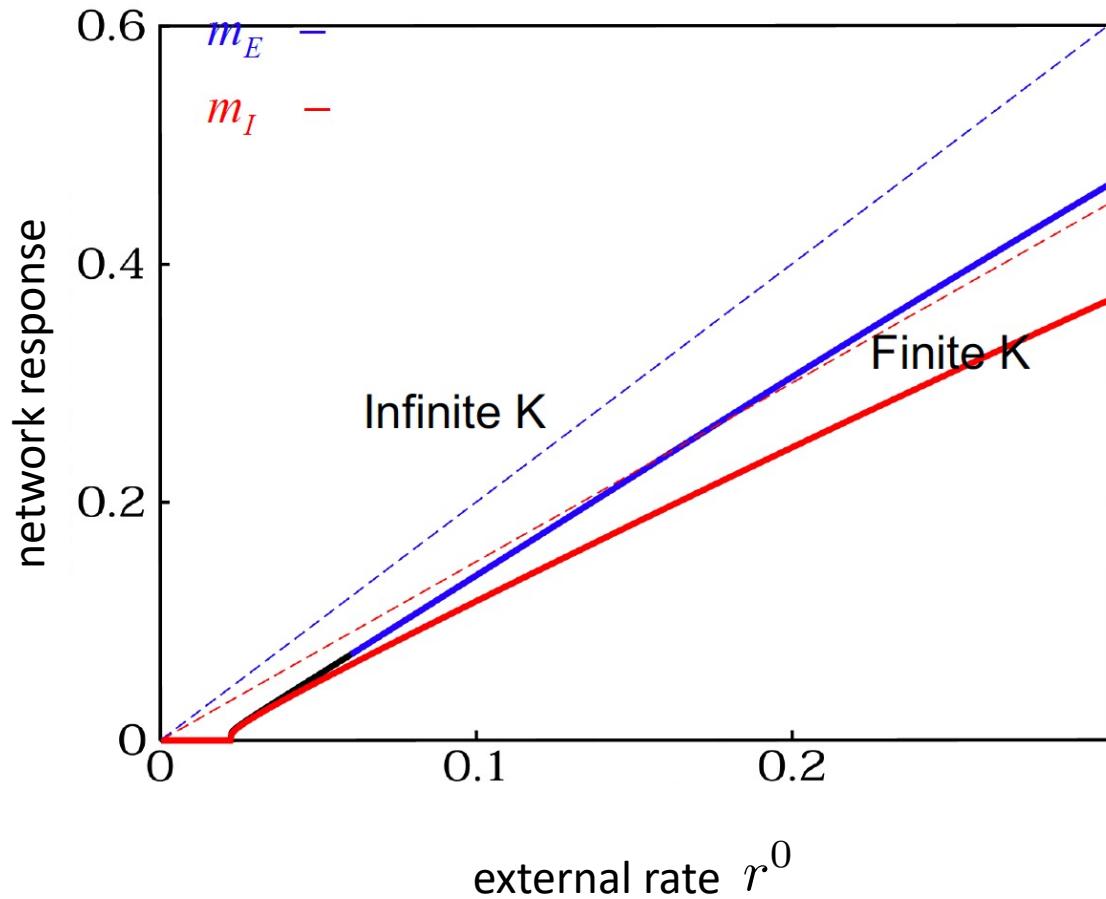
Dilute EI rate models (Kadmon & Sompolinsky):

- Same transition from fixed point to chaos through oscillation as in the single-population random Gaussian weight model.
- But if neural nonlinearity sharp enough near threshold, then no stable fixed point, just chaos. Non-normality?



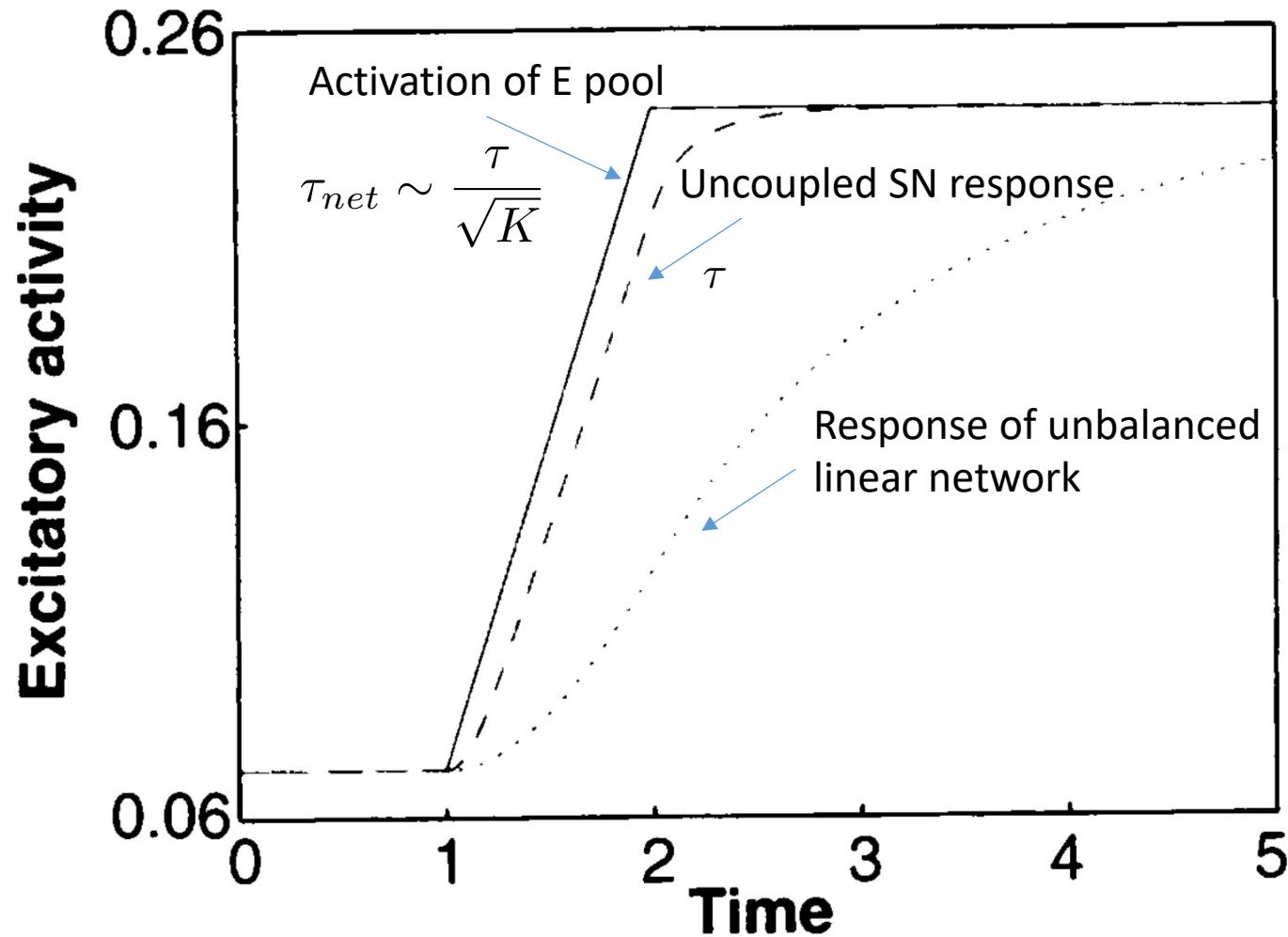
Properties of E-I balanced network:

1. Linear population response



Properties of EI balanced network:

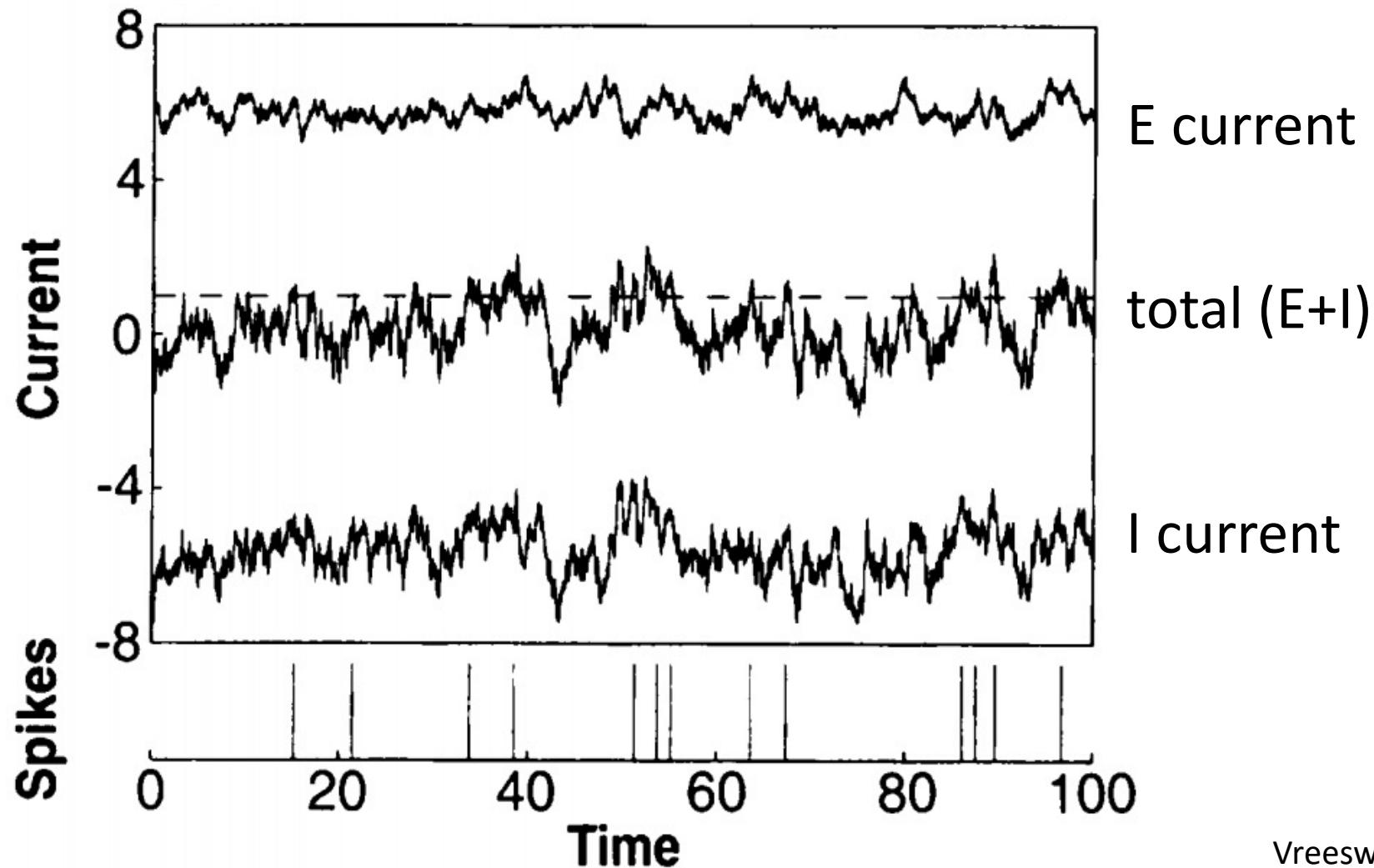
2. Rapid input response/tracking (faster than biophysical)



$$\tau_{net} \sim \frac{\tau}{\sqrt{K}}$$

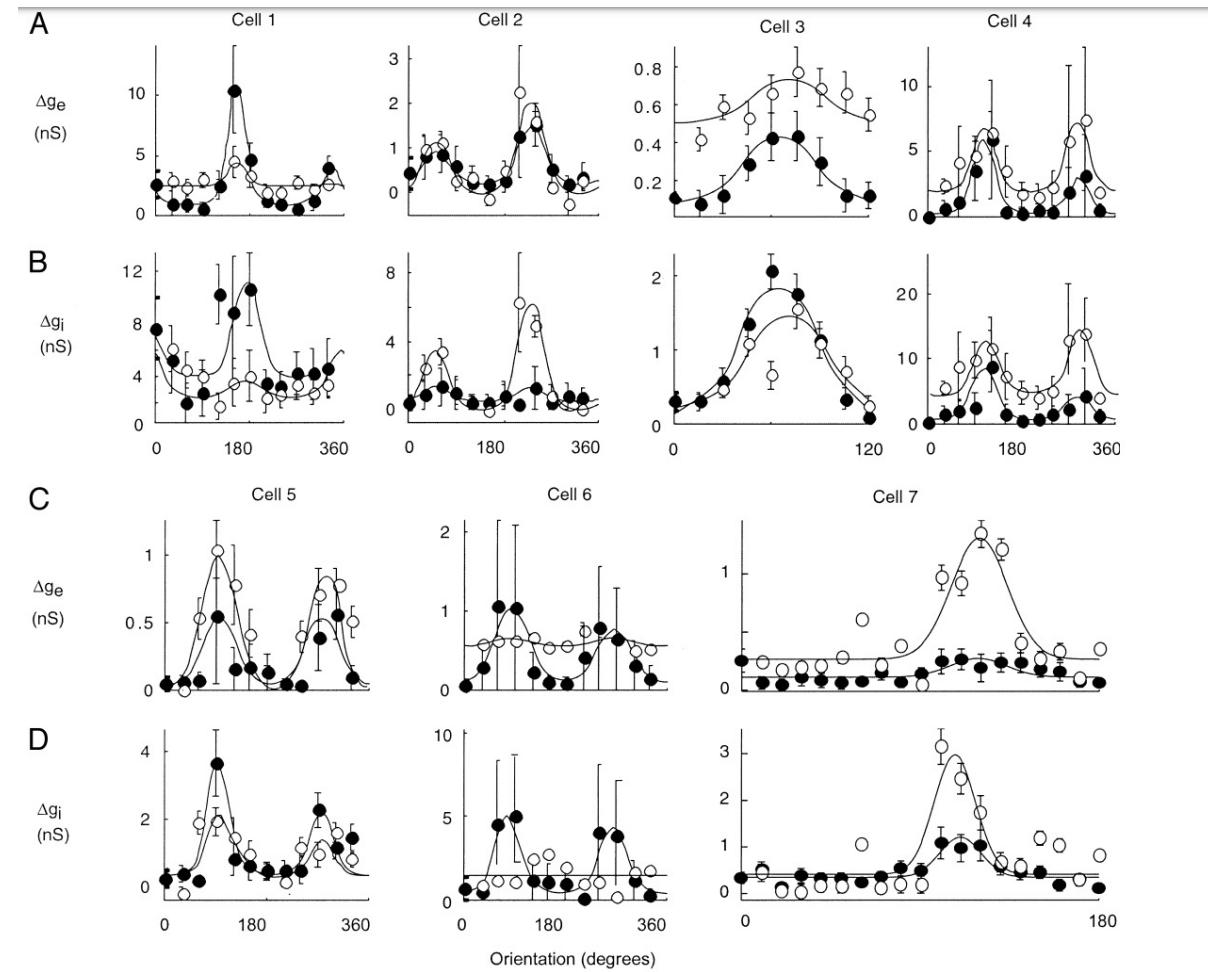
Properties of EI balanced network:

3. Co-varying (co-tuned) excitation and inhibition



Survey of some experimental results
consistent with balanced network predictions

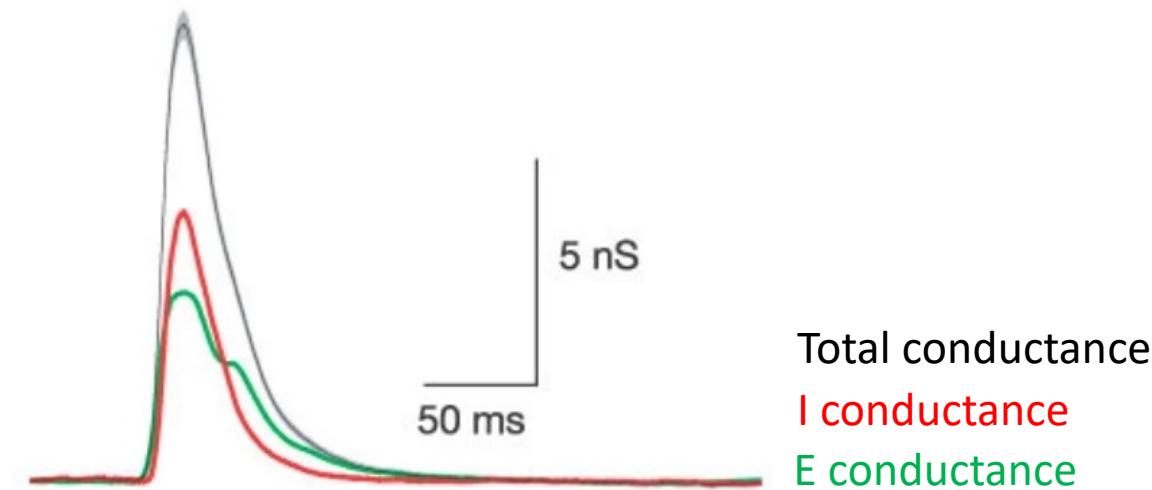
Cat V1 orientation tuning: co-tuned E, I inputs



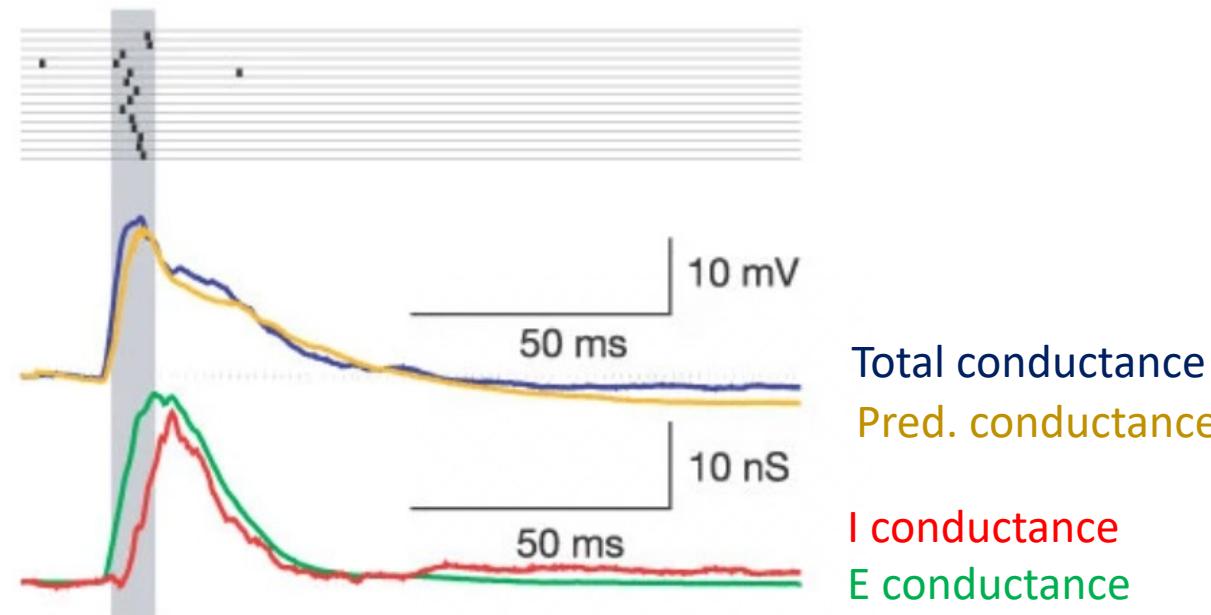
Anderson, Carandini, Ferster, 2000

Primary auditory cortex: co-tuned E, I inputs

Cell 1,
Spikes blocked



Cell 2,
Spikes not blocked



Effect: Slightly offset co-tuned I input truncates the broad E input, to generate temporally precise spiking responses.

Spontaneous activity in nearby neurons shows correlated E, I inputs

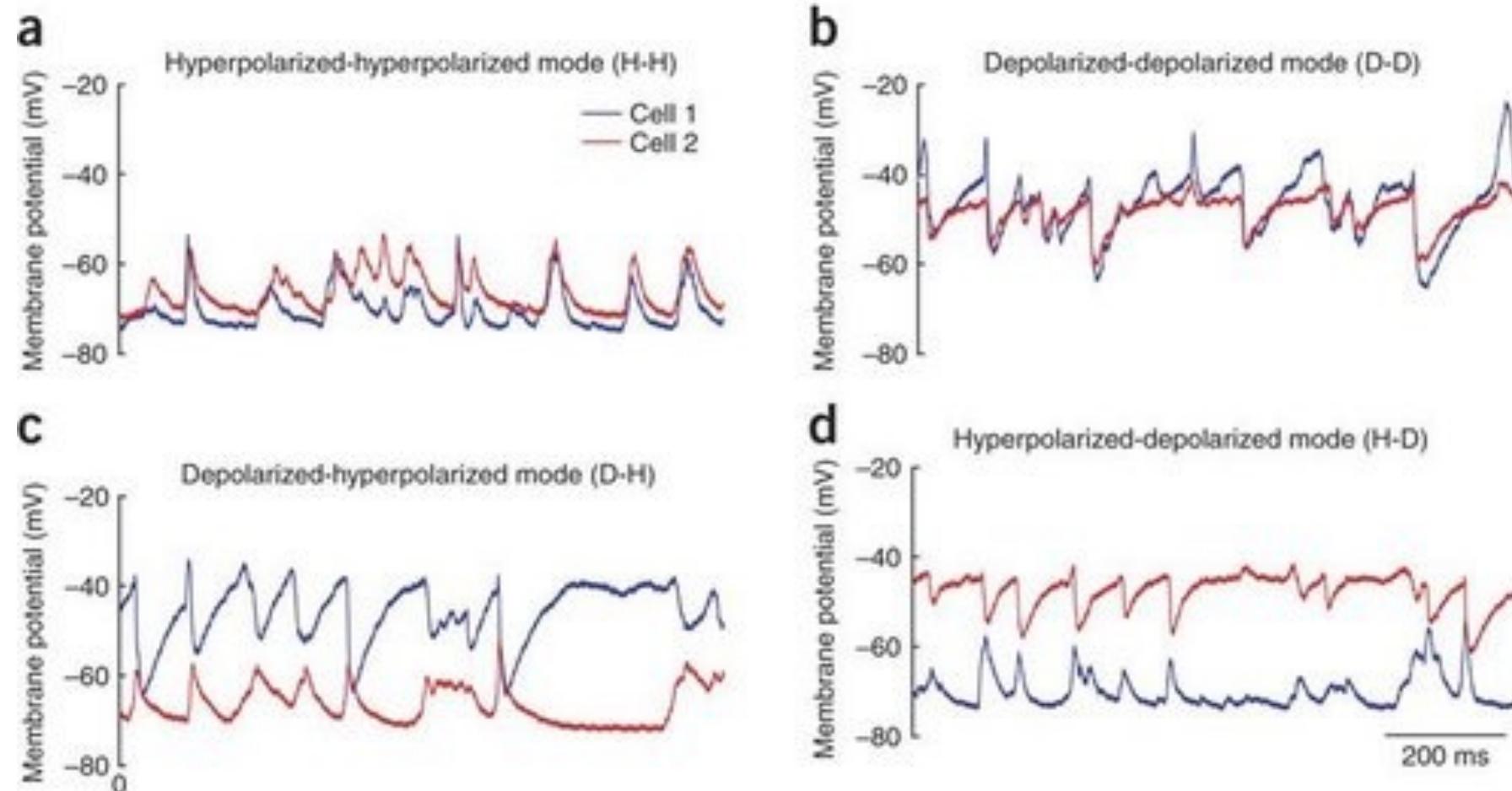
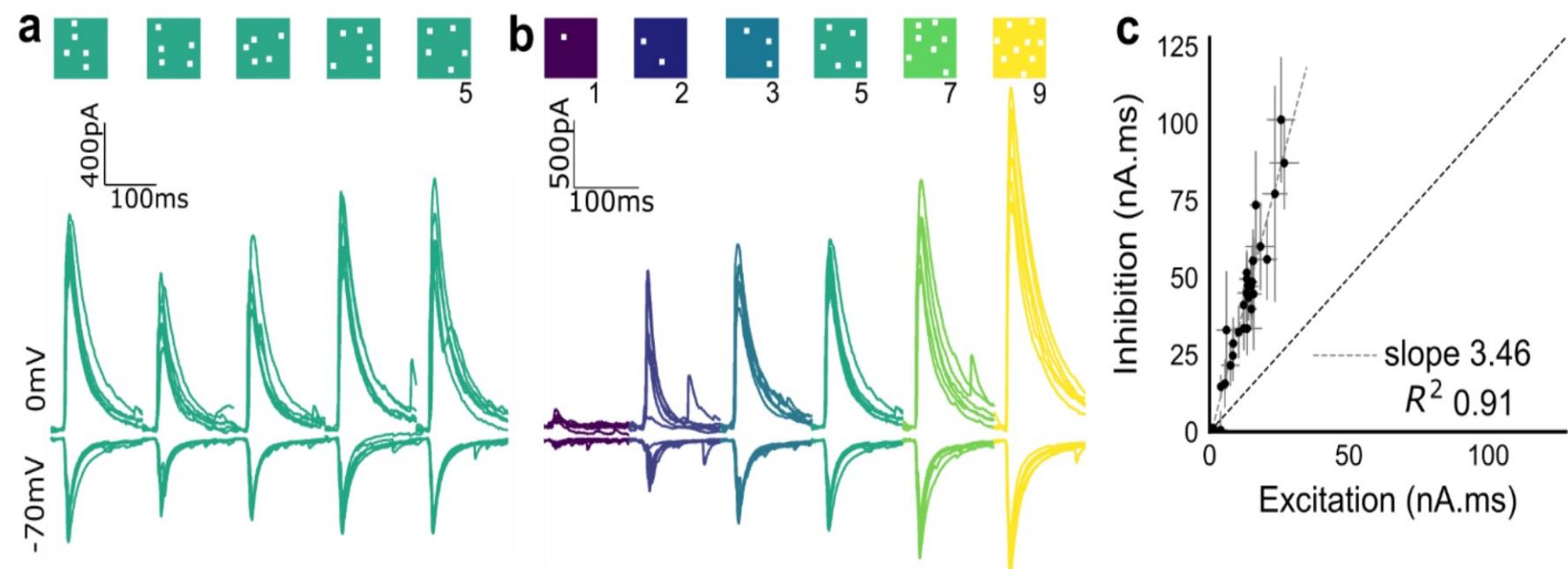
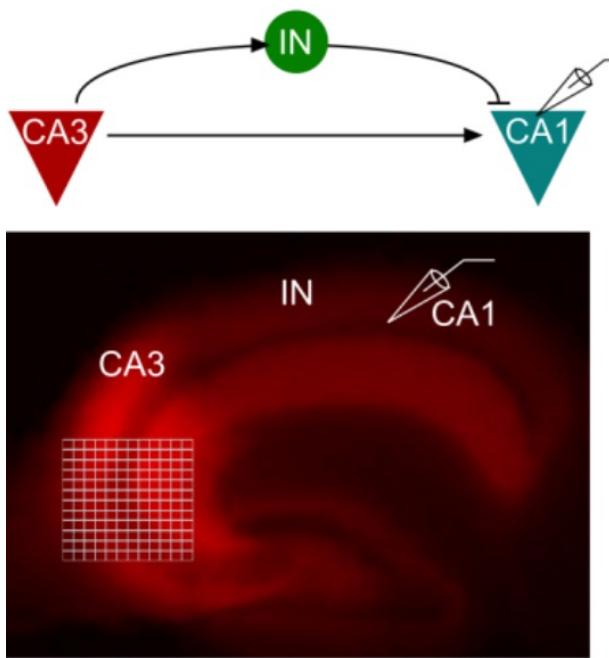


Figure 3: Excitatory and inhibitory inputs are synchronized during spontaneous activity. Two nearby neurons are simultaneously recorded when (a) both are at their resting potential, close to the reversal potential of inhibition (hyperpolarized-hyperpolarized mode); (b) both neurons are depolarized close to the reversal potential of excitation (depolarized-depolarized mode); (c-d) one of the neurons is in the hyperpolarized mode while the other is in the depolarized mode. In (a) the activity is dominated by excitatory inputs, which are seen to be highly synchronized between the neurons. Similarly, in (b) the activity is dominated by inhibitory inputs which are also highly synchronized. Finally, the mixed mode recordings (c-d) demonstrate that the excitatory and inhibitory inputs possess a high degree of synchrony.

Hippocampus brain slice



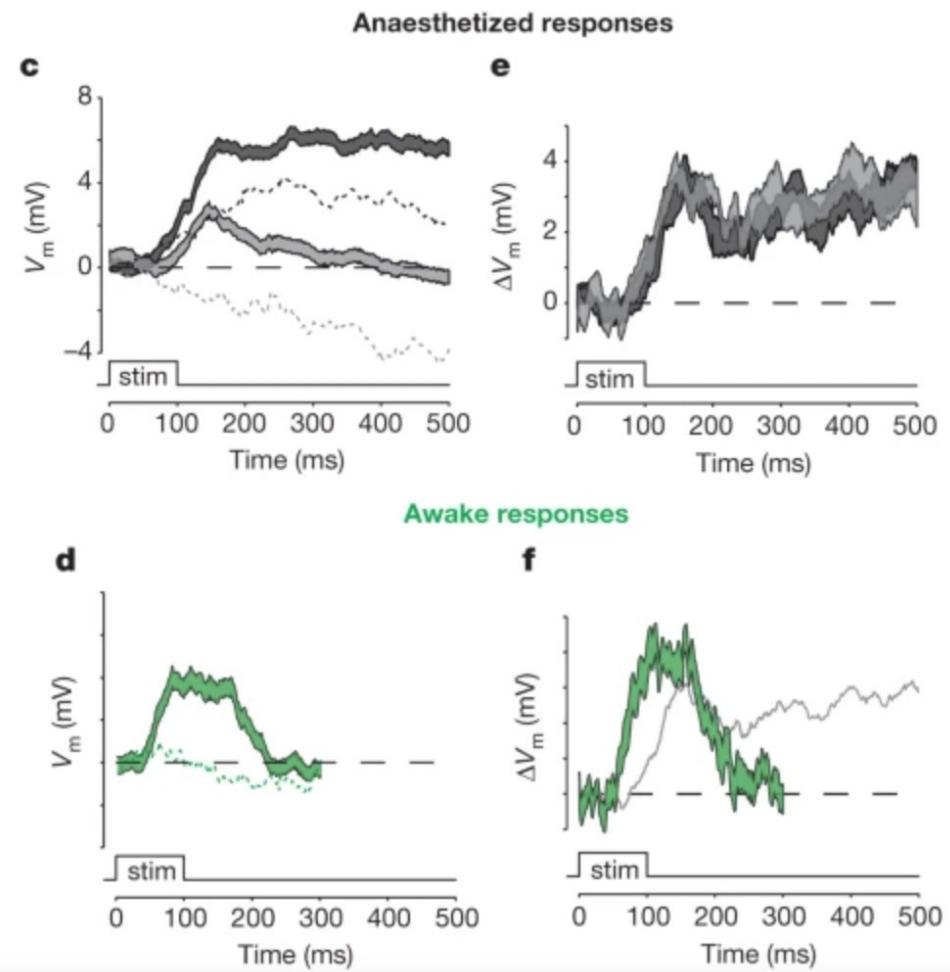
Bhatia, Moza, Bhalla 2017

Optical stimulation in CA3; patch recording in CA1.
Direct excitation, disynaptic inhibition → rapid inhibitory co-activation

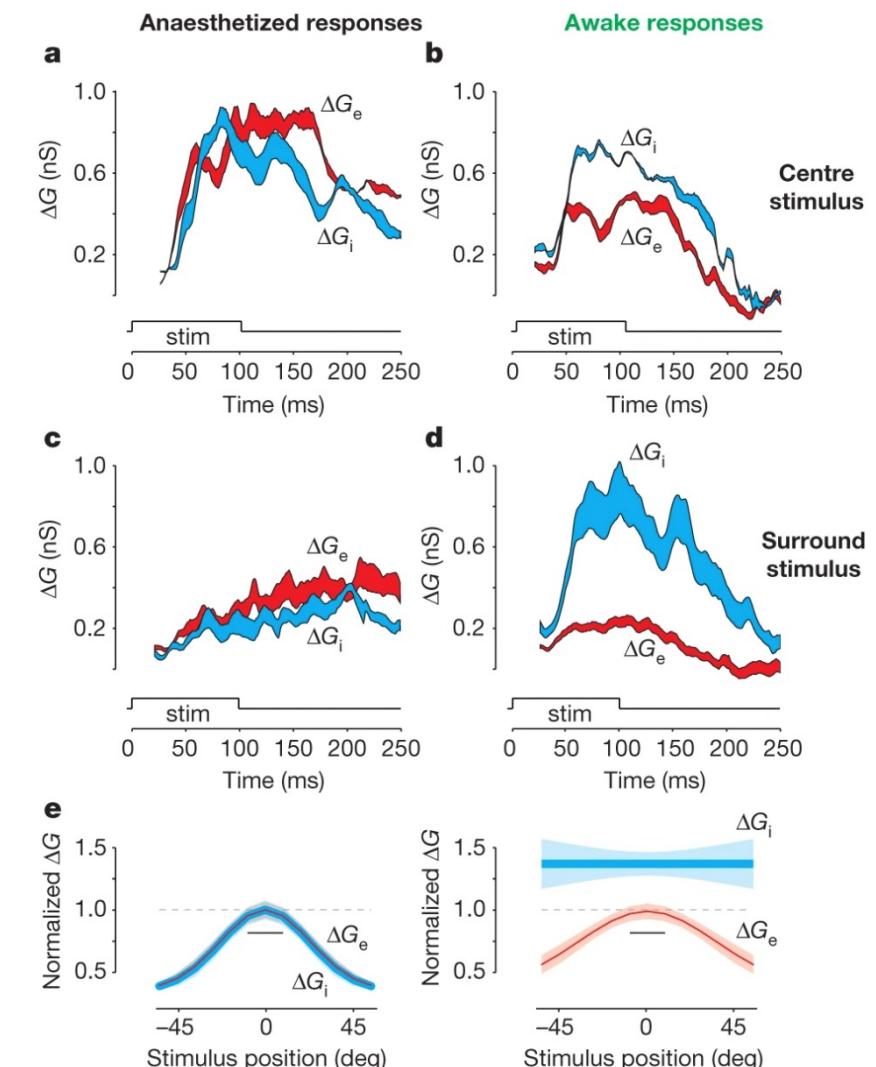
Finding: stimulus-invariant proportionality of excitation and inhibition for random input patterns of different size.
But effect is a sub-linear summation or divisive normalization rather than simple subtraction, and proportionality is not 1.

Cortical responses in anaesthetized vs. awake animals

Haider, Hausser, Carandini 2012



Anaesthetized responses slow (whether in up or down states)



Awake responses are inhibition-dominated and not spatially co-tuned (broad, untuned inhibition)

Potential uses/relevance...

- Variance found: solution to problem of explaining variance in large networks?
- Mean subtraction/dynamic range extension.
- Fast amplification in sensory systems. Murphy and Miller 2009; Hennequin, Vogels, Gerstner 2012
- Reservoir computing/Liquid State Machines: generating rich spatiotemporal dynamics for use as temporal basis functions for learning sequences for motor control, other tasks. Kirby 1991; Maass & Markram 2002; Sussillo and Abbott 2009
- Efficient coding by prediction. S. Deneve
- EI balance dysfunction and human disease: hypothesized in autism, mental retardation, epilepsy, schizophrenia, Alzheimer's disease. Eichler & Meier 2008
- Hard to harness for memory/persistent states.

Summary

- Large random networks with sufficiently strong coupling are chaotic.
- Transition from single-fixed point (at 0) to limit cycles to chaos as strength increases.
- E,I networks with strength of E, I set to cancel on average produce similar dynamics: chaos and fast time-scale response to input.
- Co-tuned inhibition and excitation: a common finding in cortex (at least in anaesthetized animals; some doubt about awake animals).