

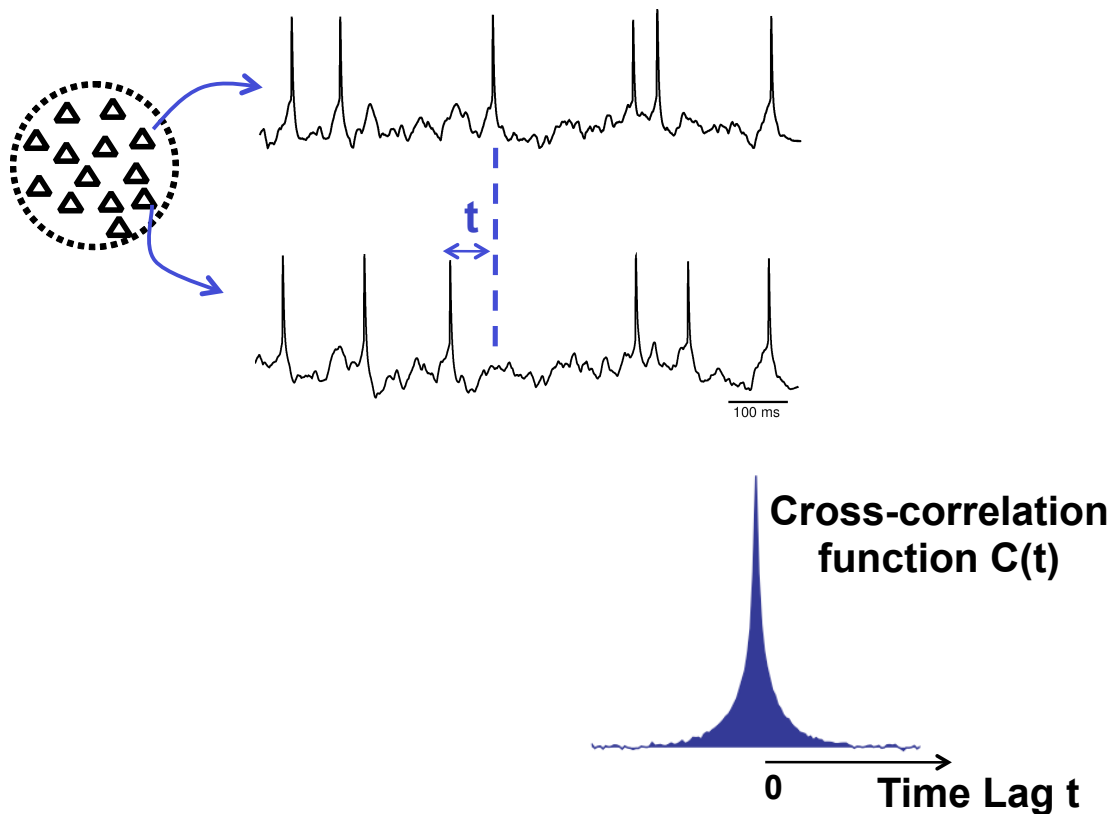
Introduction to *correlated spiking* in neural coding and dynamics

Eric Shea-Brown
U. Washington

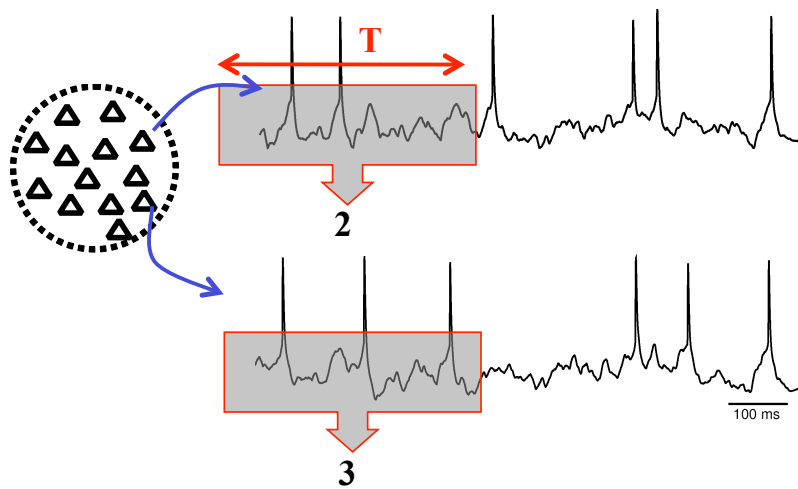
amath.washington.edu/~etsb

1

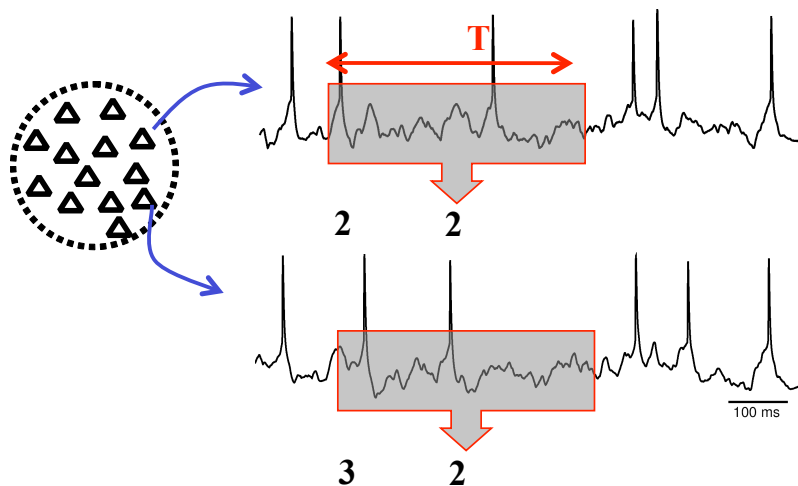
What do we mean by correlation?



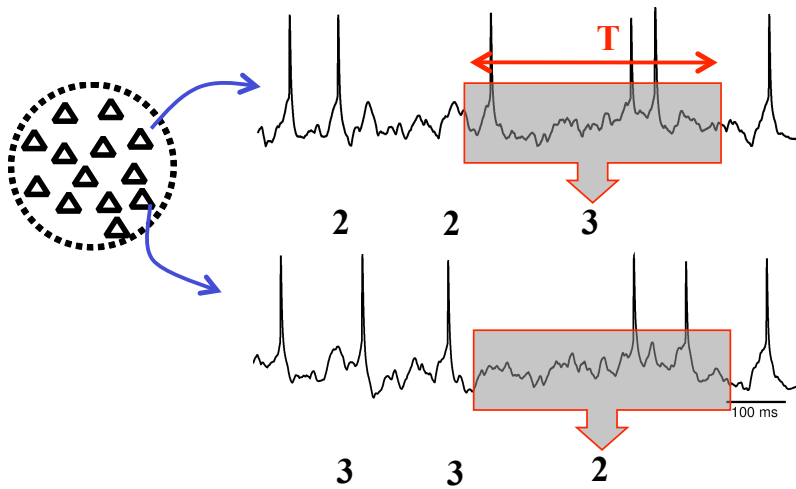
What do we mean by correlation?



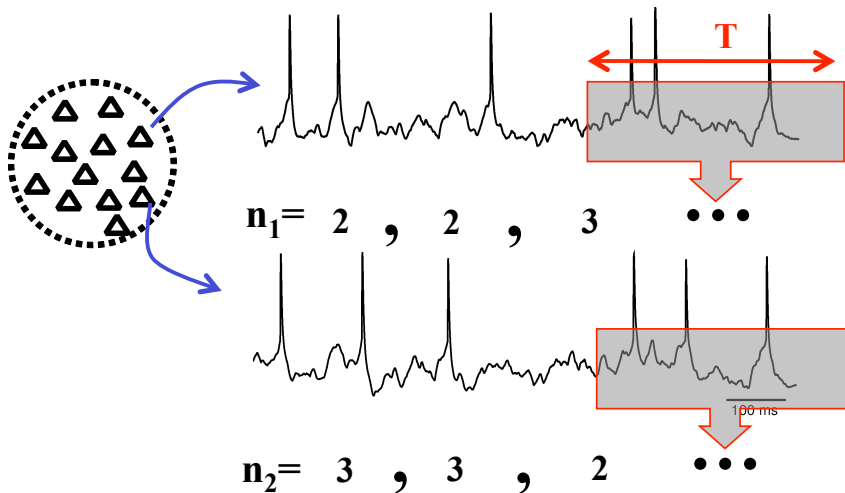
What do we mean by correlation?



What do we mean by correlation?

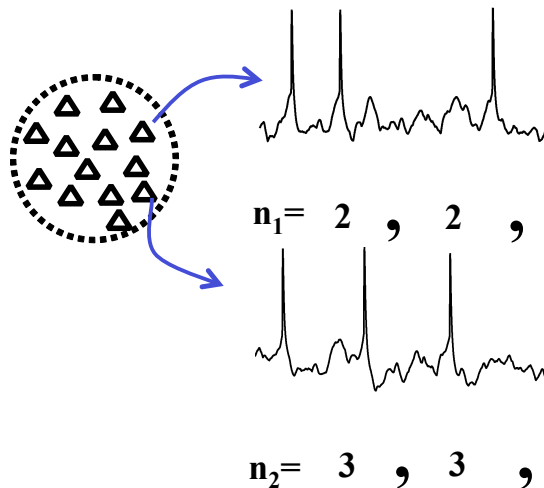
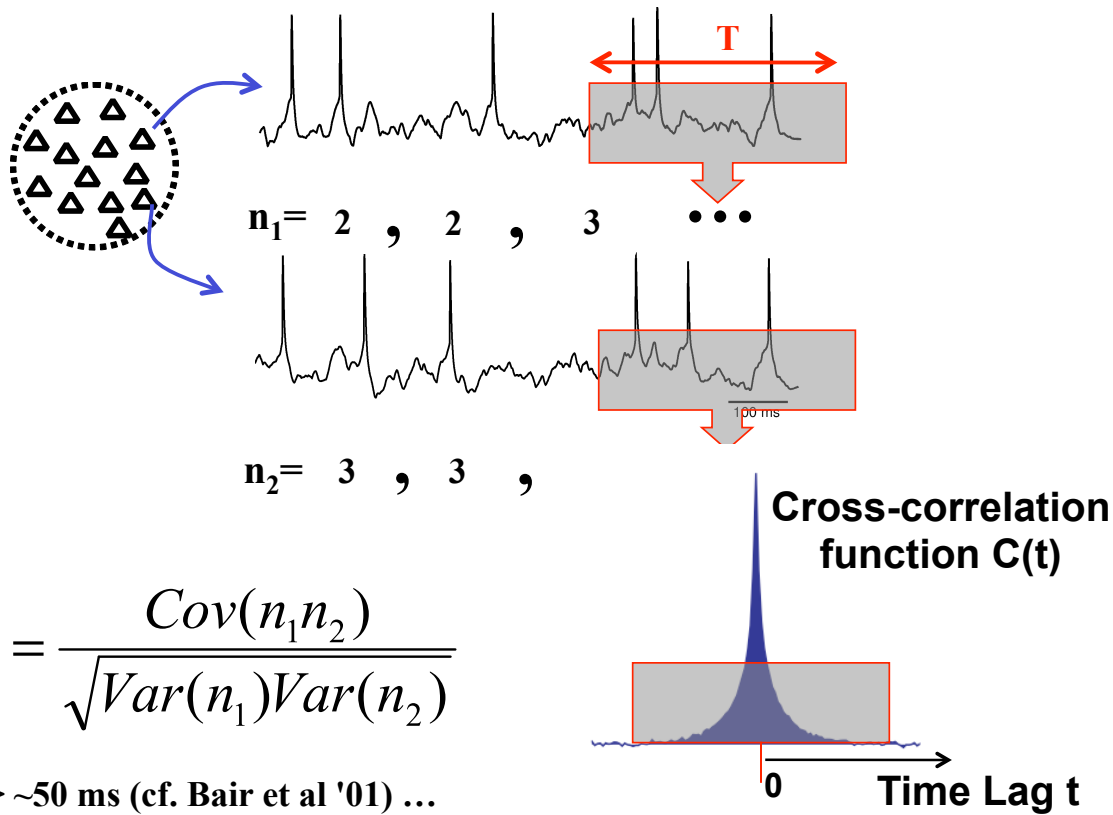


What do we mean by correlation?



$$\rho_T = \frac{\text{Cov}(n_1 n_2)}{\sqrt{\text{Var}(n_1) \text{Var}(n_2)}}$$

What do we mean by correlation?



**Correlation $\rho \neq 0$
ubiquitous *in vivo*:**

- Retina: *Mastronade 1983.*
- LGN: *Alonso et al 1996*
- V1: *Kohn and Smith 2005*
see also Ecker et al 2010
- IT: *Gawne & Richmond 1993*
- PF: *Constanidis & Goldman-Rakic 2002.*
- Motor cortex: *Vaadia et al 1995*
- Parietal Cortex: *Lee et al 1998*
- Somatosensory thal.: *Bruno & Sakmann 2006*
- A1: *deCharms & Merzenich 1996*
- SI: *Romo et al 2003. ...*

$$\rho = \frac{\text{Cov}(n_1 n_2)}{\sqrt{\text{Var}(n_1) \text{Var}(n_2)}}$$

Why the correlations? $p(n_1, n_2) \neq p(n_1)p(n_2)$

Common signal input → Common spike response
→ **SIGNAL CORRELATIONS**



**ADDITIONAL “NETWORK-DRIVEN” CORRELATIONS
ARE ...
NOISE CORRELATIONS**

$$p(n_1, n_2 | s(t)) \neq p(n_1 | s(t))p(n_2 | s(t))$$

We'll focus on noise correlations.

OUTLINE

CONSEQUENCES OF CORRELATED SPIKING

Impact on coding
(a) Homogeneous populations
...

Impact on signal propagation

BASIC MECHANISMS FOR CORRELATED SPIKING

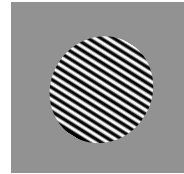
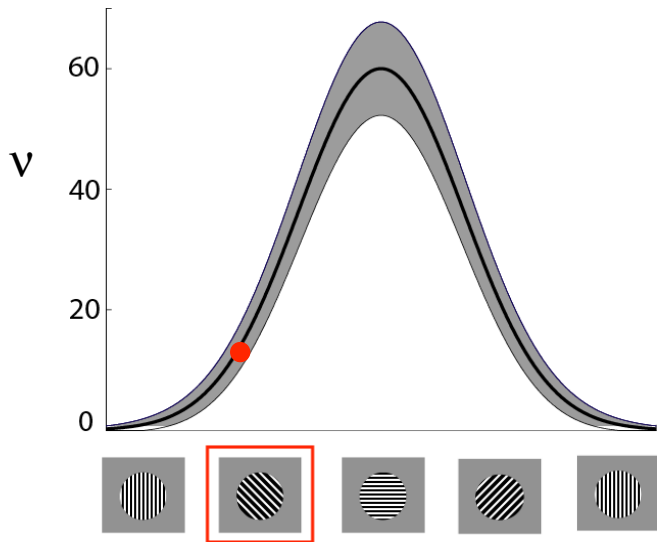
...

BEYOND CELL-PAIRS: HIGHER-ORDER CORRELATIONS

...

Rate coding

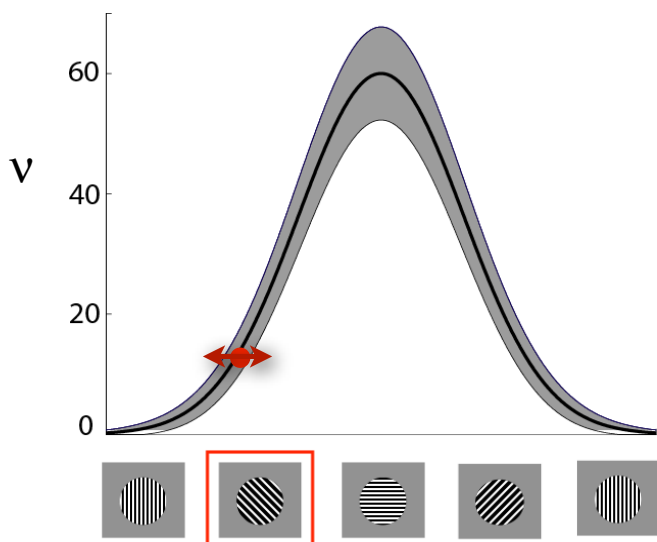
$$RATE : v = \frac{n}{T}$$



Hubel and Weisel. J. Physiol., 1962
Reddy, Kreiman, Koch, and Fried (20

Response Variability

$$RATE : v = \frac{n}{T}$$



Variable spike
count introduces
ambiguity.

| | | | | 5

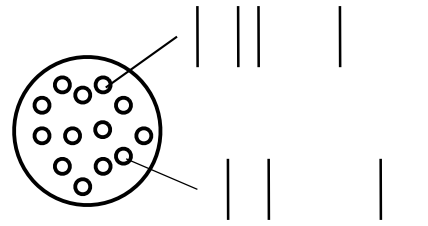
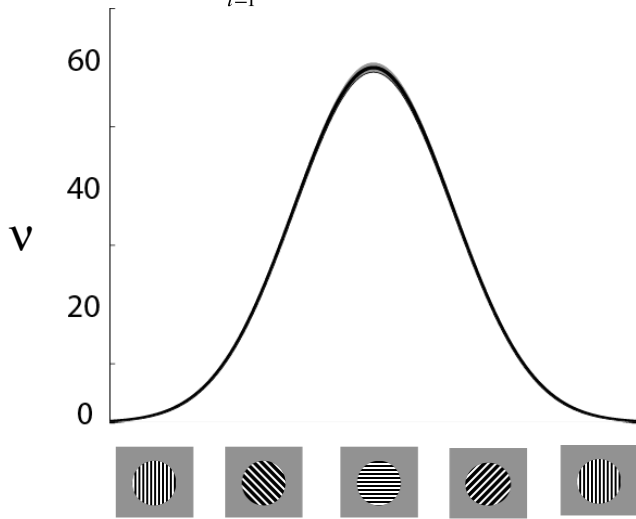
(Poisson process)

| | | | 3

| | | | | | 6

Population codes – average over M independent cells

$$\text{RATE } \nu = \frac{1}{TM} \sum_{i=1}^M n_i$$



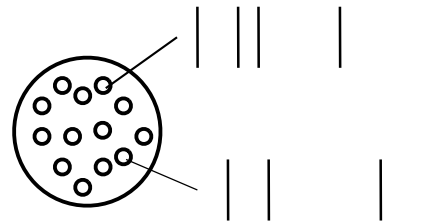
M cells
 n_i spikes each
in time window T

Population codes – average over M independent cells

$$\text{RATE } \nu = \frac{1}{TM} \sum_{i=1}^M n_i$$

$$\langle \nu \rangle = \frac{1}{TM} \sum_i^M \langle n_i \rangle$$

$$= \frac{1}{TM} MrT = r$$



M cells
 n_i spikes each
in time window T

$$\text{var}(\nu) = \frac{1}{T^2 M^2} \sum_i^M \text{var}(n_i)$$

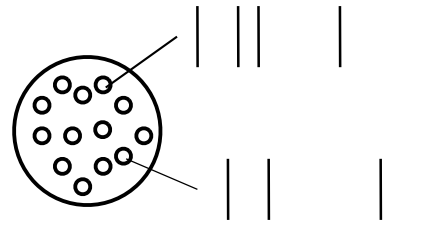
$$= \frac{1}{T^2 M^2} MrT \sim \frac{1}{M} r$$

Population codes – average over M independent cells

$$RATE \nu = \frac{1}{TM} \sum_{i=1}^M n_i$$

$$\langle \nu \rangle = \frac{1}{TM} \sum_i^M \langle n_i \rangle$$

$$= \frac{1}{TM} MrT = r$$



M cells
 n_i spikes each
in time window T

$$var(\nu) = \frac{1}{T^2 M^2} \sum_i^M var(n_i)$$

$$= \frac{1}{T^2 M^2} MrT \sim \frac{1}{M} r$$

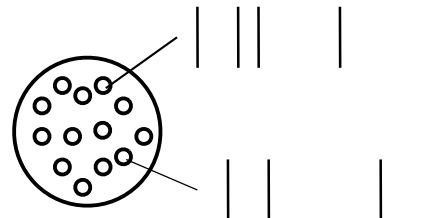
$$\frac{\langle \nu \rangle}{var(\nu)} = SNR(\nu) \sim M$$

Population codes – average over M independent cells

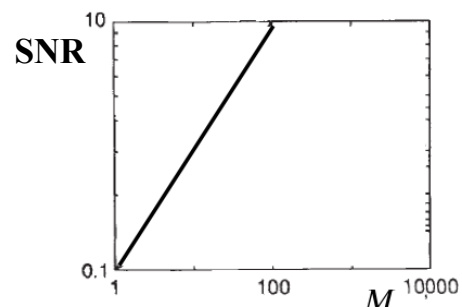
$$RATE \nu = \frac{1}{TM} \sum_{i=1}^M n_i$$

$$\langle \nu \rangle = \frac{1}{TM} \sum_i^M \langle n_i \rangle$$

$$= \frac{1}{TM} MrT = r$$



M cells
 n_i spikes each
in time window T



**Population averaging
improves SNR.**

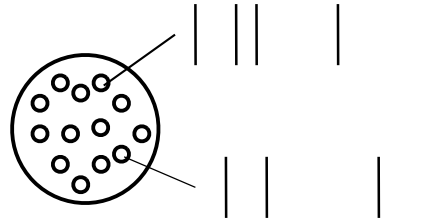
$$\frac{\langle \nu \rangle}{var(\nu)} = SNR(\nu) \sim M$$

Population codes – average over M correlated cells

$$RATE \nu = \frac{1}{TM} \sum_{i=1}^M n_i$$

$$\langle \nu \rangle = \frac{1}{TM} \sum_i^M \langle n_i \rangle$$

$$= \frac{1}{TM} MrT = r$$



M cells

n_i spikes each

in time window T

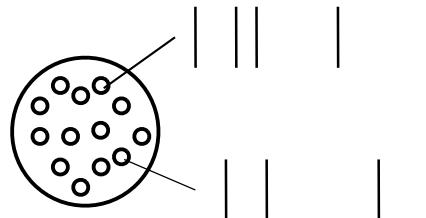
n_i have correlation coefficient ρ

Population codes – average over M correlated cells

$$RATE \nu = \frac{1}{TM} \sum_{i=1}^M n_i$$

$$\langle \nu \rangle = \frac{1}{TM} \sum_i^M \langle n_i \rangle$$

$$= \frac{1}{TM} MrT = r$$



M cells

n_i spikes each

in time window T

n_i have correlation coefficient ρ

$$\boxed{\frac{\langle \nu \rangle}{var(\nu)} = SNR(\nu) \sim \frac{1}{\rho}}$$

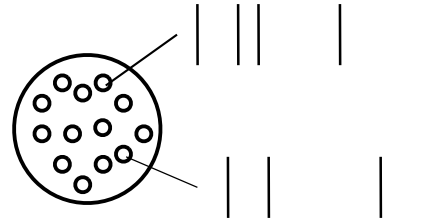
$$\begin{aligned} var(\nu) &= \\ &= \frac{1}{T^2 M^2} \left(\sum_i^M var(n_i) + \sum_{i \neq j} cov(n_i, n_j) \right) \\ &\sim \frac{1}{T^2 M^2} M^2 r T \rho \sim \rho \end{aligned}$$

Population codes – average over M correlated cells

$$RATE \nu = \frac{1}{TM} \sum_{i=1}^M n_i$$

$$\langle \nu \rangle = \frac{1}{TM} \sum_i^M \langle n_i \rangle$$

$$= \frac{1}{TM} M r T = r$$



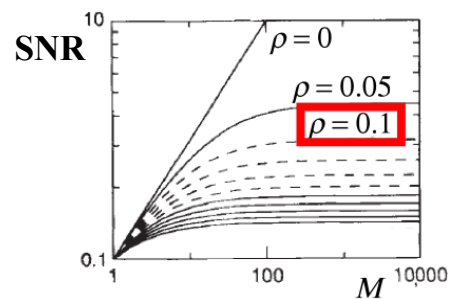
M cells

n_i spikes each

in time window T

n_i have correlation coefficient ρ

Zohary, Shadlen and Newsome (1994)



$$\frac{\langle \nu \rangle}{var(\nu)} = SNR(\nu) \sim \frac{1}{\rho}$$

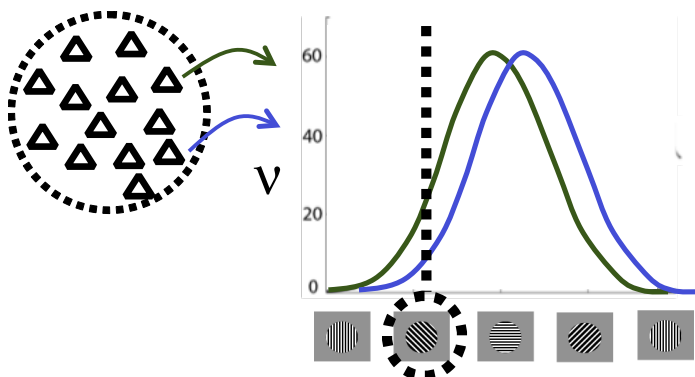
OUTLINE

CONSEQUENCES OF CORRELATED SPIKING

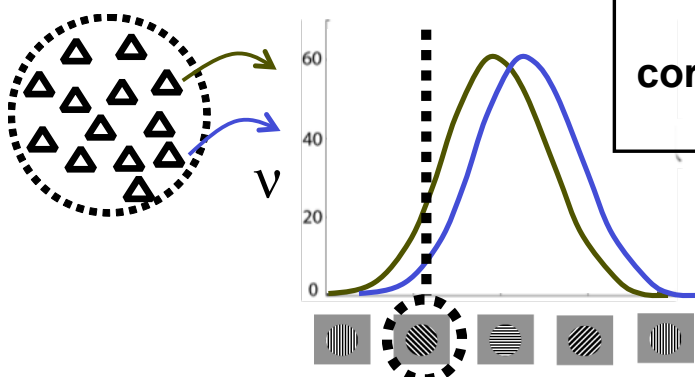
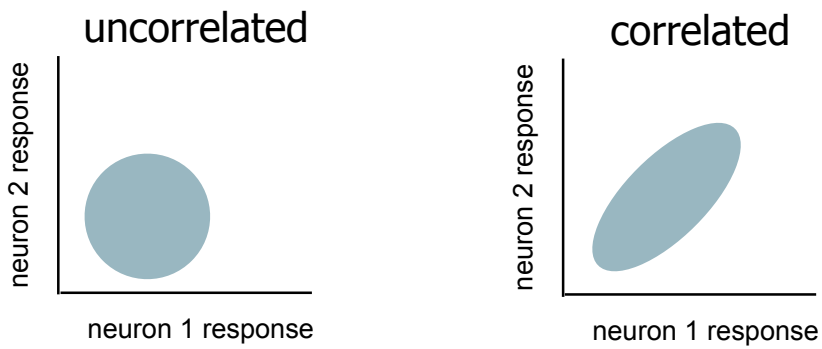
Impact on coding

(a) Homogeneous populations: limits population averaging / degrades info

(b) Heterogeneous cell pairs ...



21



What are effects of correlations on information content in cell pair?

22

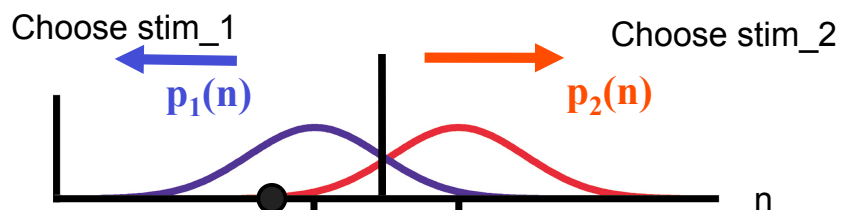
Consider discriminating two nearby stimuli

For one neuron:

$p_1(n)$: response n under stimulus 1

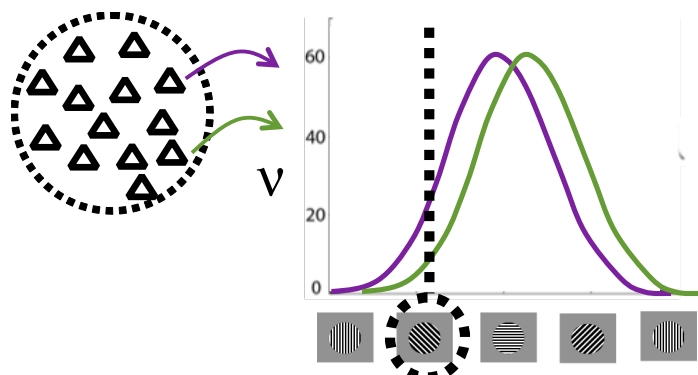
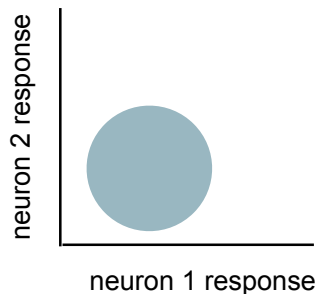
$p_2(n)$: response n under stimulus 2

Decode ... via (optimal) maximum likelihood discrimination

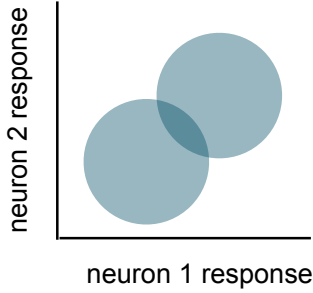


Abbott+Dayan, N. Comp. 99, Sompolinsky et al., PRE 01, Averbeck et al., Nat Rev Nsci 06

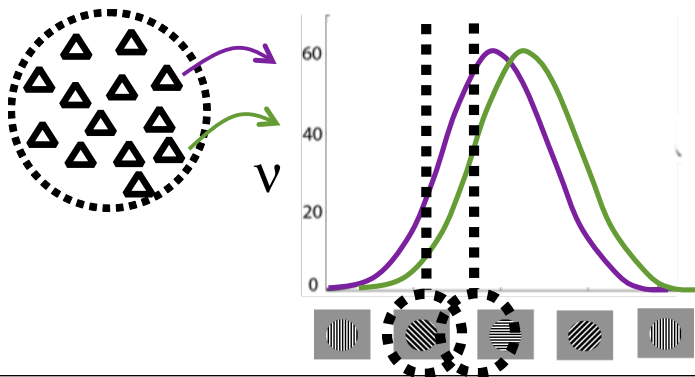
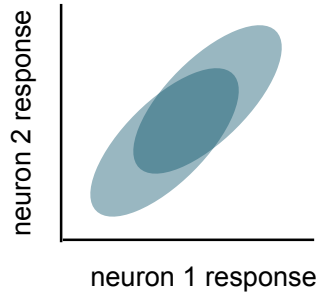
uncorrelated



uncorrelated



correlated



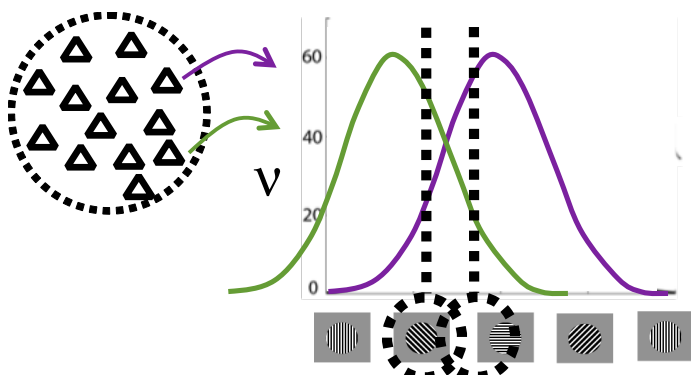
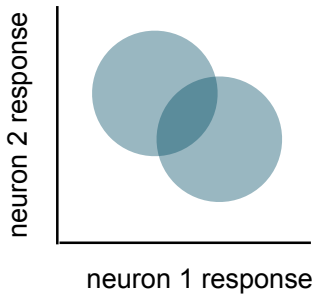
Positive **noise correlation** degrades signal encoding.

...

when also have positive signal correlation.

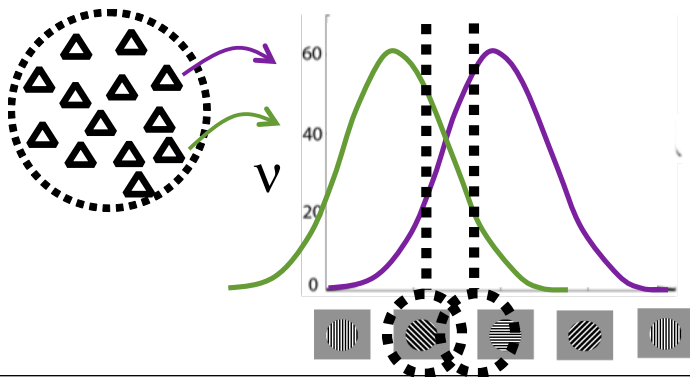
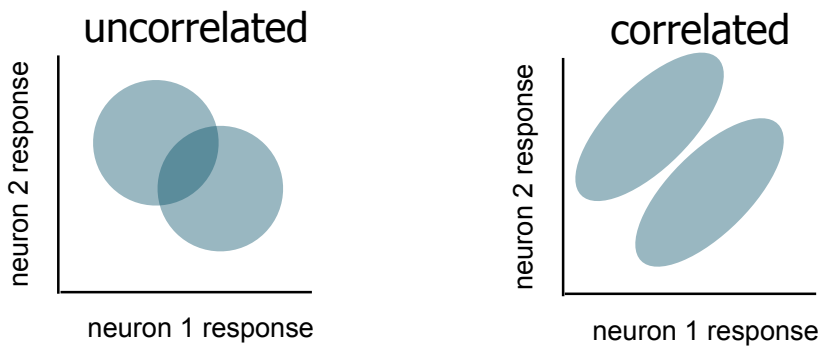
25

uncorrelated



when also have **negative signal correlation**.

26



Positive **noise correlation**
ENHANCES signal encoding.

...

when also have **negative**
signal correlation.

27

Neuron, Vol. 38, 649–657, May 22, 2003, Copyright ©2003 by Cell Press

Correlated Neuronal Discharges that Increase Coding Efficiency during Perceptual Discrimination

Ranulfo Romo,^{1,*} Adrián Hernández,¹
Antonio Zainos,¹ and Emilio Salinas²

¹Instituto de Fisiología Celular

rate increases similar to those observed in S1, but for other units the firing rate decreases monotonically as a function of stimulus frequency (Salinas et al., 2000).

OUTLINE

CONSEQUENCES OF CORRELATED SPIKING

Impact on coding

(a) Homogeneous populations: limits population averaging / degrades info

(b) Heterogeneous cell pairs ...

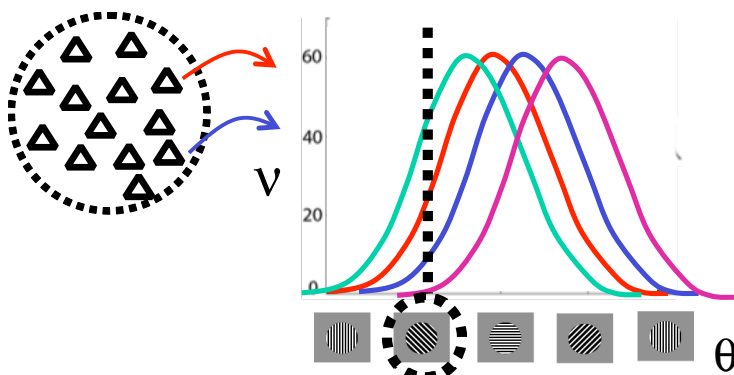


similar stimulus tuning: **DEGRADE CODING**

different stimulus tuning: **ENHANCE CODING**

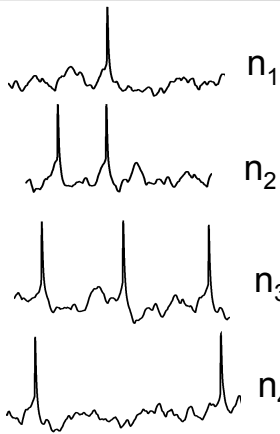
29

- Now, generalize to populations of more than two cells ...



30

Tuning curves [e.g Hubel+Wiesel, '60s]



$$\mathbf{n} = (n_1, n_2, n_3, \dots).$$

Stimulus x

Task: given $\mathbf{n}$, estimate x

Cramér-Rao Bound:

$$(\text{Estimation error})^2 \geq \frac{1}{I_{FISHER}}$$

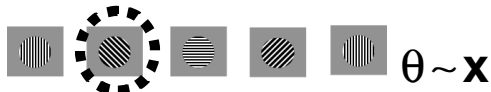
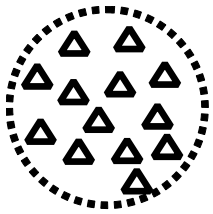
Usually, but not always saturated! (Berens et al COSYNE '11)



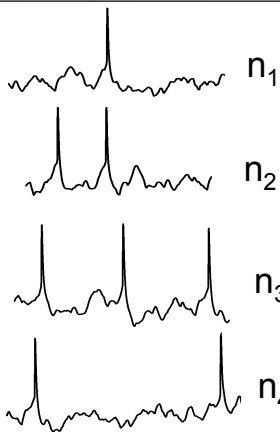
Each neuron i fires spike count $n_i = f_i(x) + \eta_i(x)$

Fisher Information

$$I_F(x) = \left\langle \frac{d^2}{dx^2} \log P[\mathbf{n}|x] \right\rangle$$



Implications: Tuning curves [e.g Hubel+Wiesel, '60s]



$$\mathbf{n} = (n_1, n_2, n_3, \dots).$$

Stimulus x

Task: given $\mathbf{n}$, estimate x

Cramér-Rao Bound:

$$(\text{Estimation error})^2 \geq \frac{1}{I_{FISHER}}$$

Each neuron i fires spike count $n_i = f_i(x) + \eta_i(x)$

Fisher Information

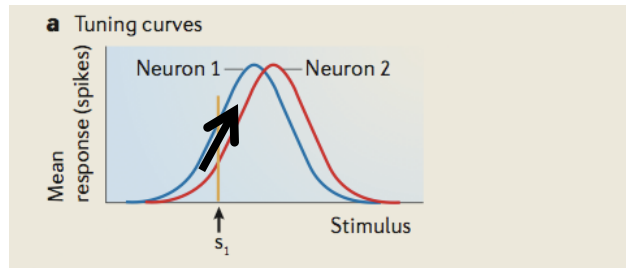
$$I_F = \left\langle \left(\frac{d}{dx} \log \hat{P}(n|x) \right)^2 \right\rangle$$

[Somplinsky et al]

Take η_i gaussian with: $Q_{i,j} = \delta_{i,j}v + (1 - \delta_{i,j})c \exp(-\alpha|i - j|)v$

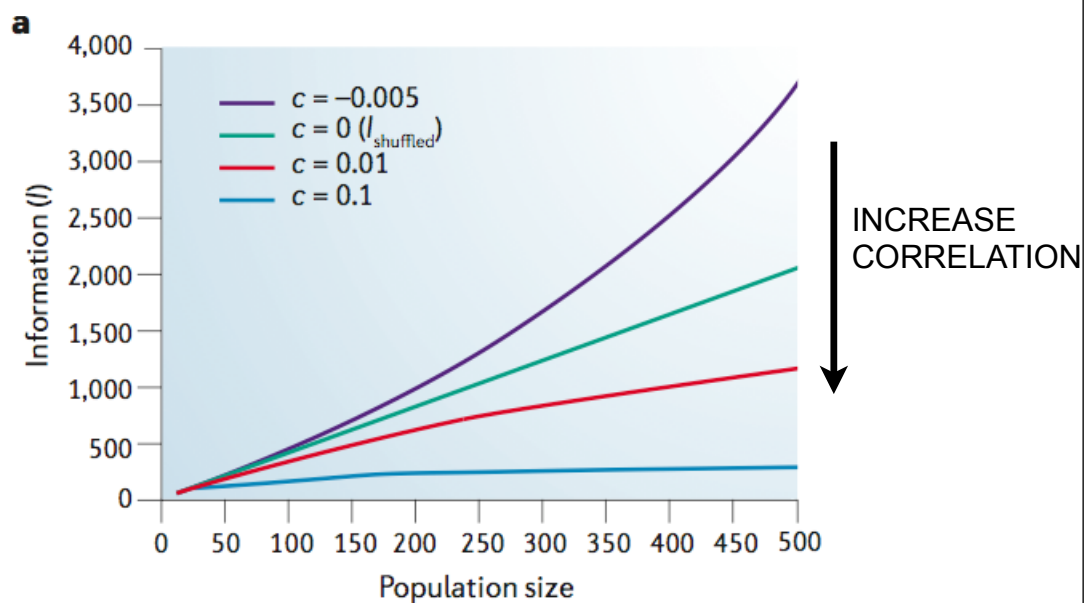
Interpret: positive correlations for “nearby” cells

... and nearby cells have positive signal correlations



... so, expect presence of correlations to DECREASE information

Averbeck et al 2006:



CONCLUDE: Here, correlations degrade coding.

In general, degrade OR enhance effect could dominate -- must examine case by case.

OUTLINE

CONSEQUENCES OF CORRELATED SPIKING

Impact on coding

(a) Homogeneous populations: limits population averaging / degrades info

(b) Heterogeneous cell pairs ...

similar stimulus tuning: **DEGRADE CODING**

different stimulus tuning: **ENHANCE CODING**

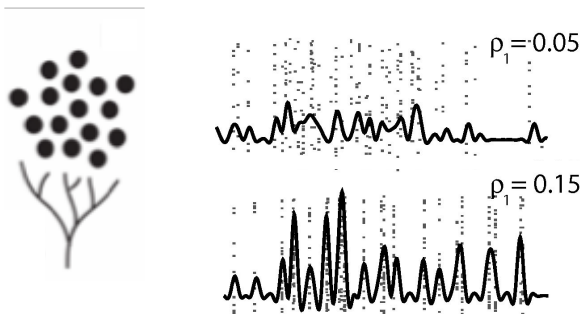
→ (c) Heterogeneous population ... mixed effects

Impact on signal propagation

35

Correlated variability modulates downstream rates

Salinas and Sejnowski, 2000

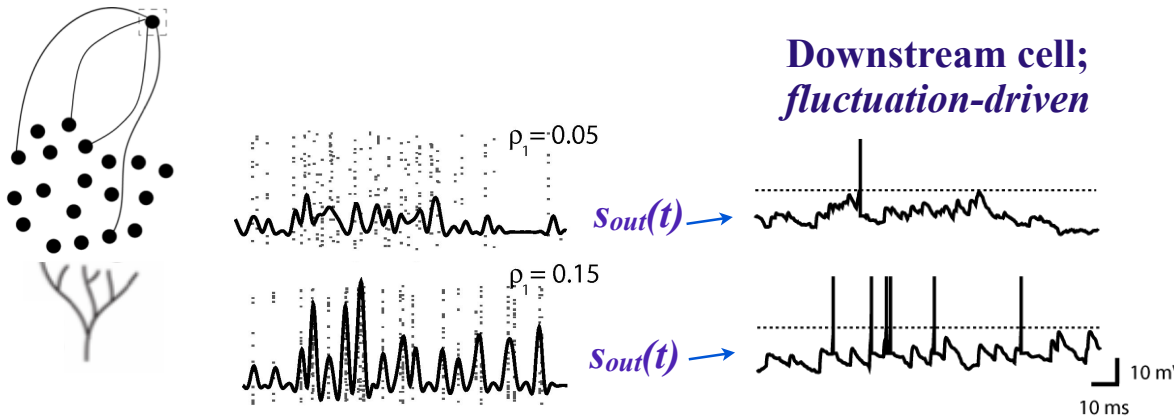


$$\text{std. dev.} \sim (\text{rate} \times \text{corr})^{1/2}$$

36

Correlated variability modulates downstream rates

Salinas and Sejnowski, 2000

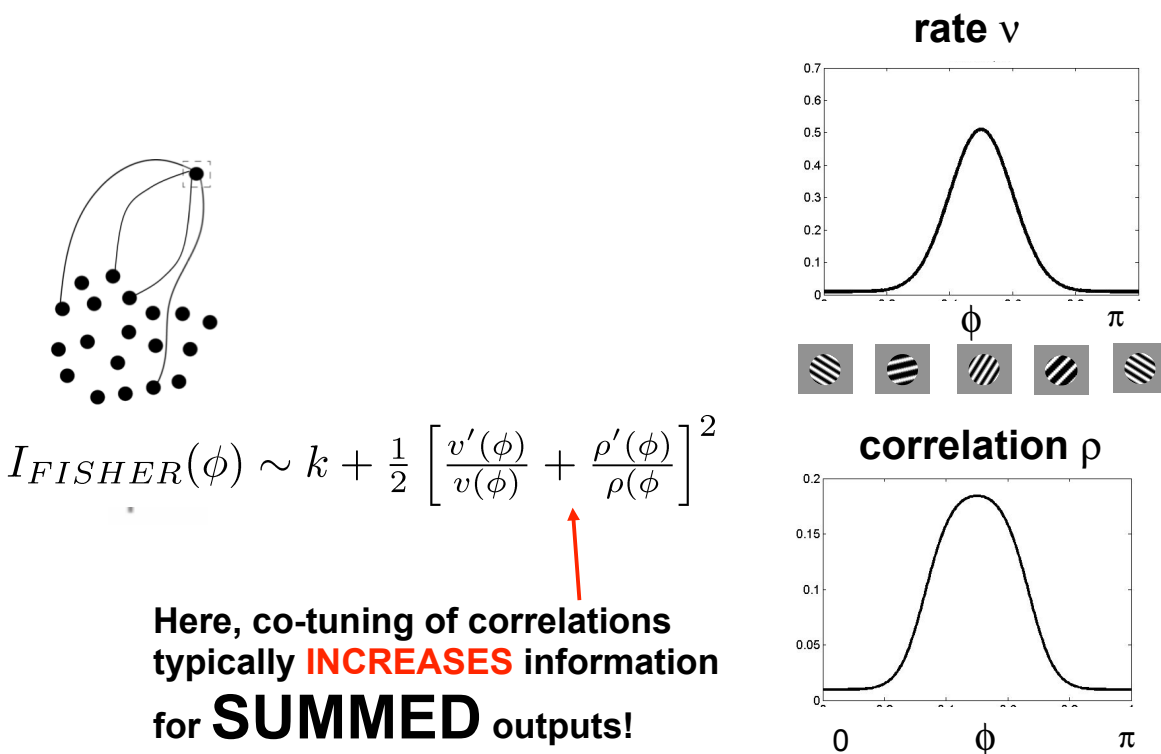


$$\text{std. dev.} \sim (\text{rate} \times \text{corr})^{1/2}$$

$$\text{rate} = f(\text{std dev})$$

37

What if correlations are *stimulus-dependent*?



Stimulus-dependent correlations - an example

LETTER

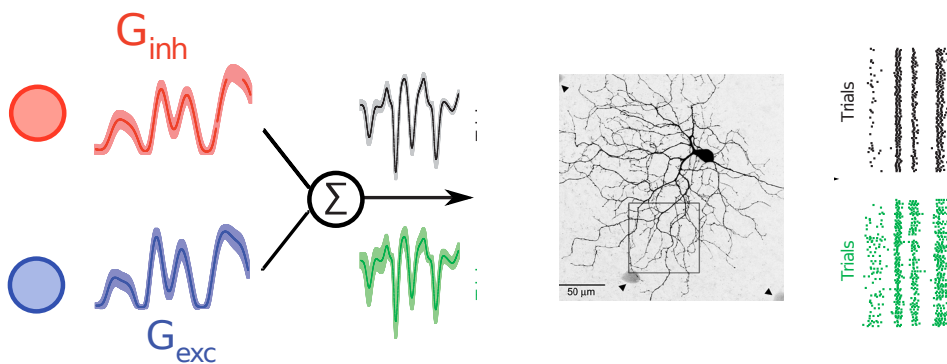
doi:10.1038/nature09570

Noise correlations improve response fidelity and stimulus encoding

Jon Cafaro² & Fred Rieke^{1,2}

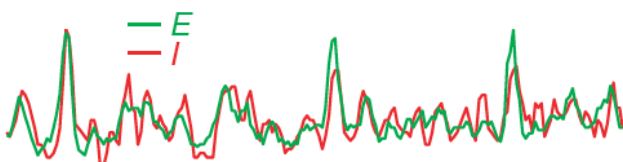
Stimulus-dependent correlations - an example

Cafaro and Rieke, Nature, 2010



Chen et al., J Phys, 2009

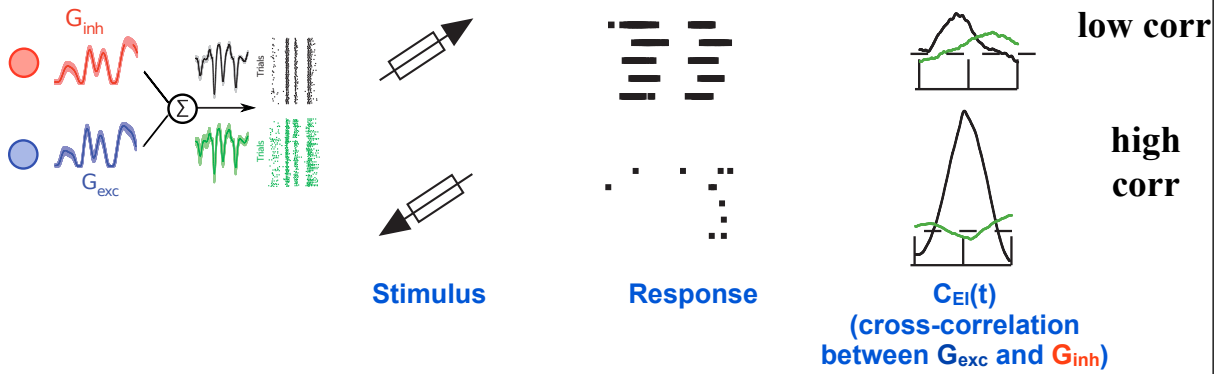
POSITIVE correlations b/w incoming conductances



→ **NEGATIVE** correlations
b/w incoming currents.

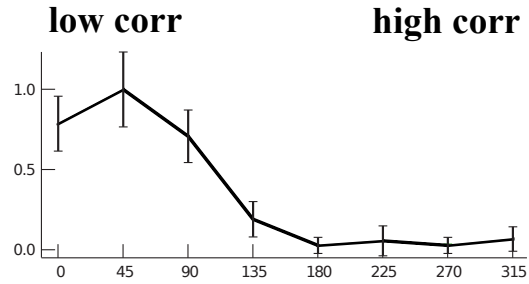
Fluctuations **cancel**.

Stimulus-dependent correlations - an example

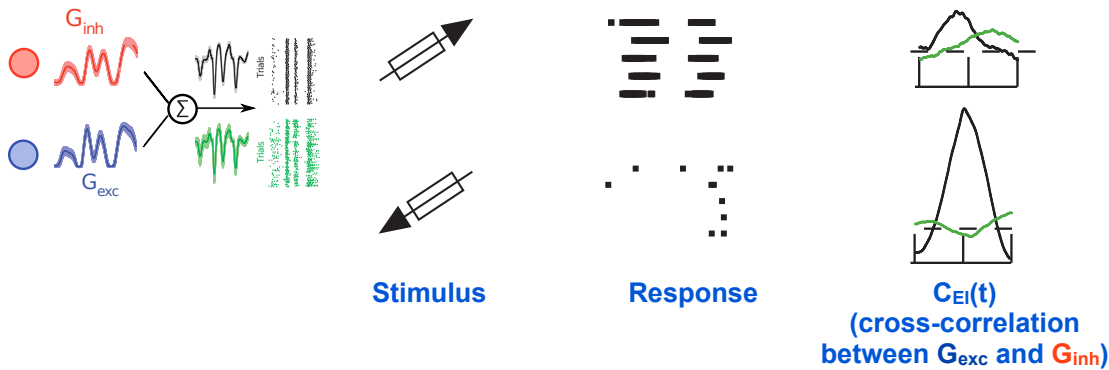


Correlated

adapted from
Cafaro and Rieke, Nature, 2010



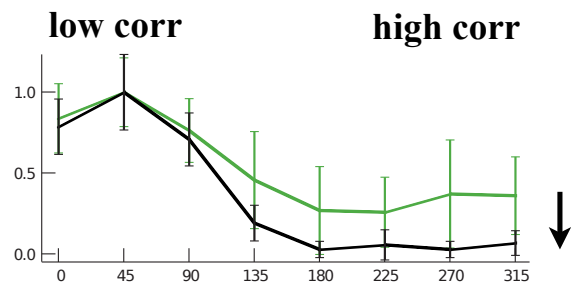
Stimulus-dependent correlations - an example



**Uncorrelated (trial
shuffled)**

Correlated

adapted from
Cafaro and Rieke, Nature, 2010



More neg. corr
→ **lower rates / variance**

OUTLINE

CONSEQUENCES OF CORRELATED SPIKING

Impact on coding

Impact on signal propagation

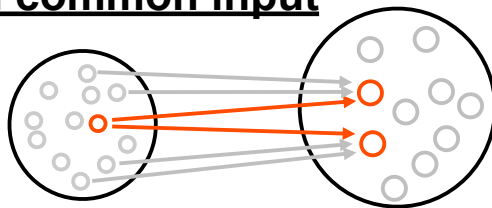
Positive correlation sets *gain*:

Downstream rate $\sim$ upstream rate $\times$ upstream correlation

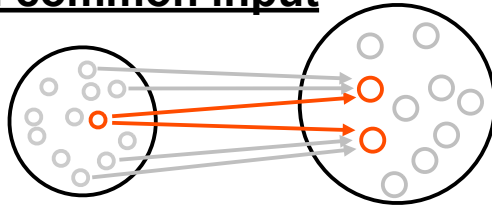
BASIC MECHANISMS FOR CORRELATED SPIKING

43

Correlations from common input



Correlations from common input



As in:

Shadlen and Newsome, *J. Nsci.* '98

Binder and Powers, *J. Neurophys.* '01

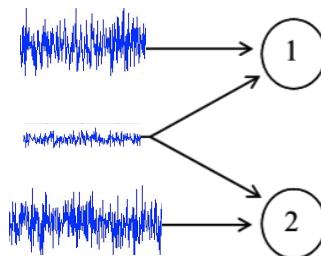
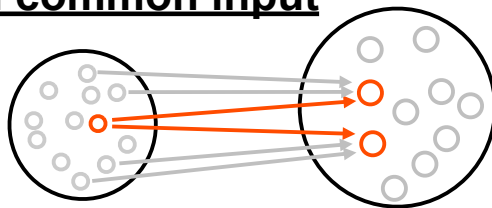
Tetzlaff, Geisel, and Diesmann, *Neurocomp.* '02

Moreno-Bote et al, *Phys. Rev. Lett.* '06

Galan et al, *J. Nsci.* '06

... and others

Correlations from common input



As in:

Shadlen and Newsome, *J. Nsci.* '98

Binder and Powers, *J. Neurophys.* '01

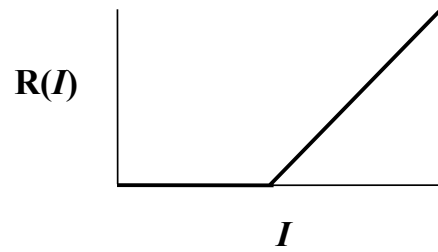
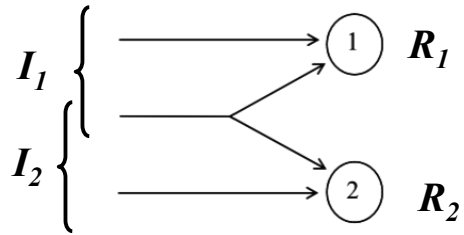
Tetzlaff, Geisel, and Diesmann, *Neurocomp.* '02

Moreno-Bote et al, *Phys. Rev. Lett.* '06

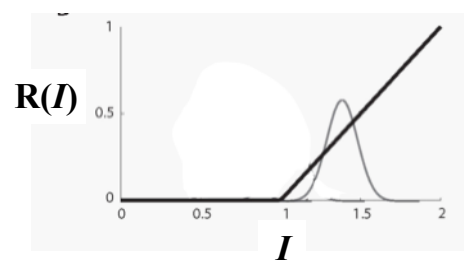
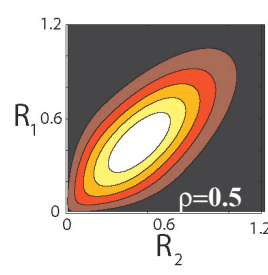
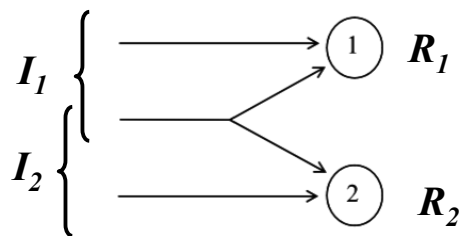
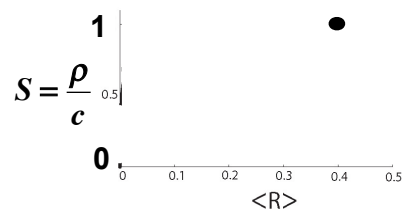
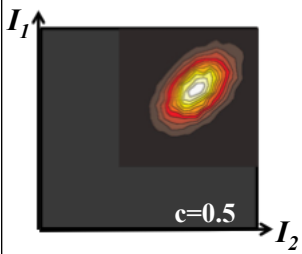
Galan et al, *J. Nsci.* '06

... and others

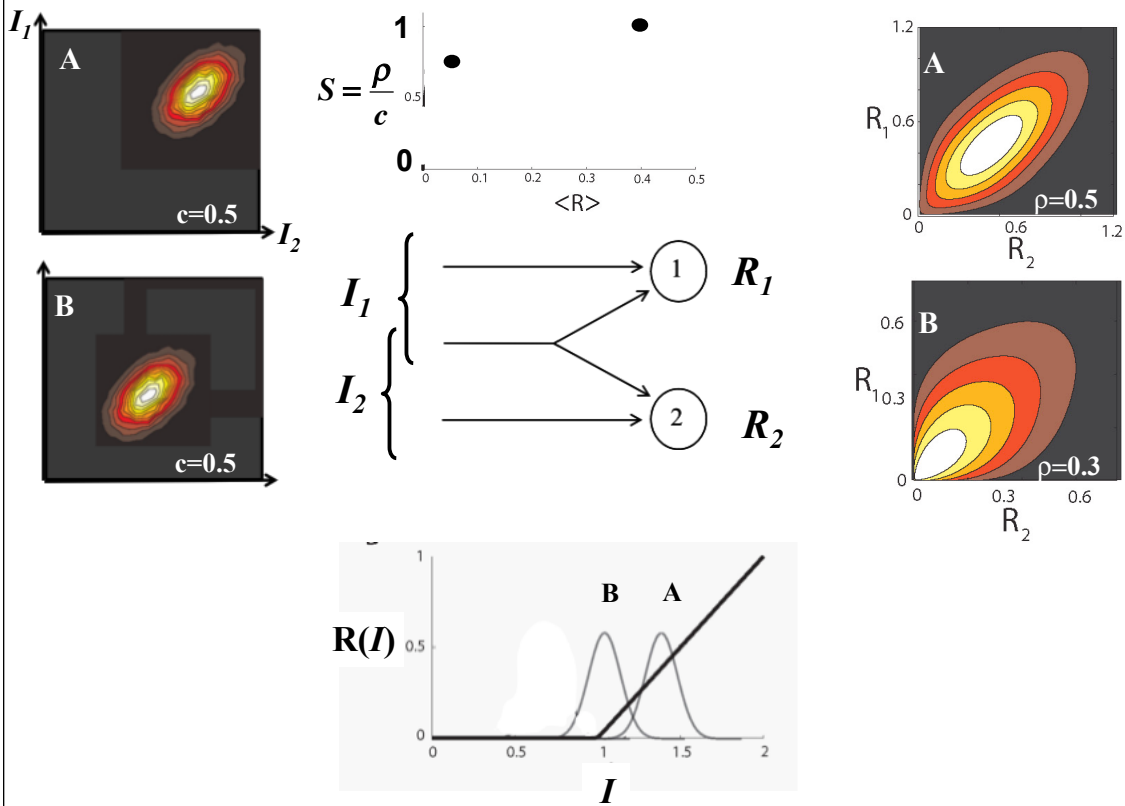
Simplest model



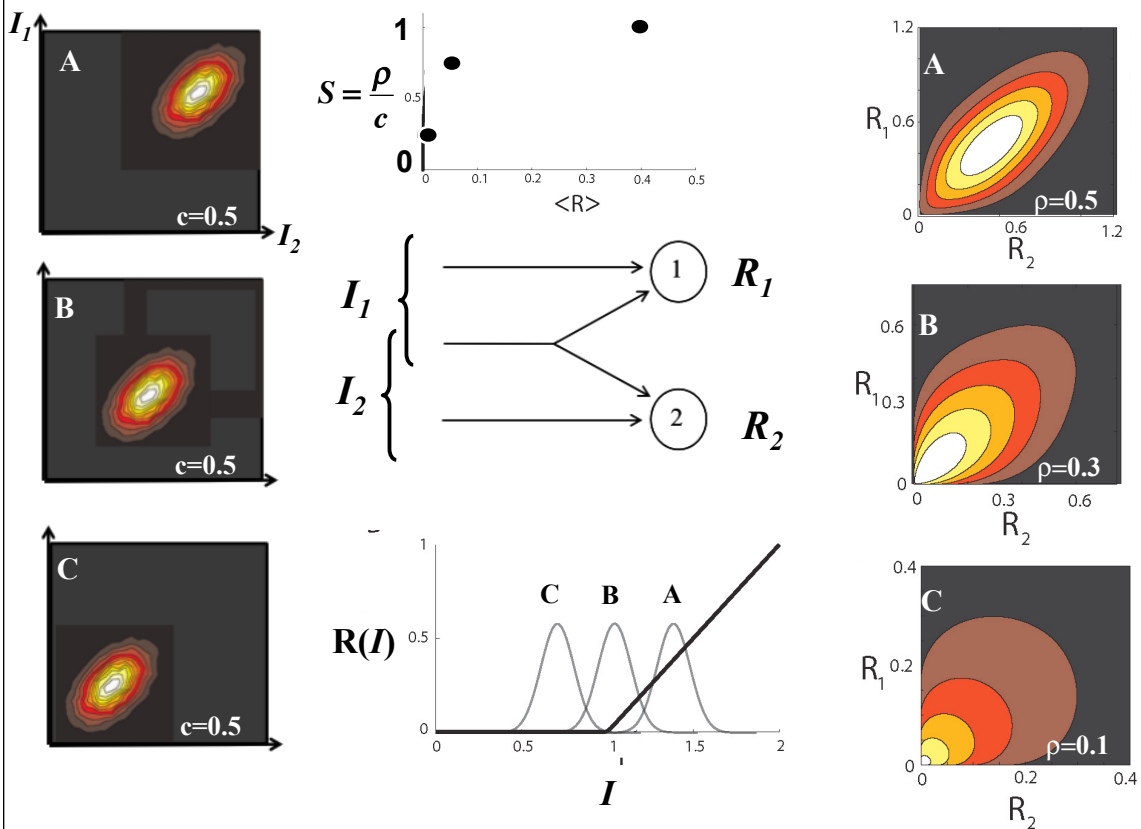
Simplest model



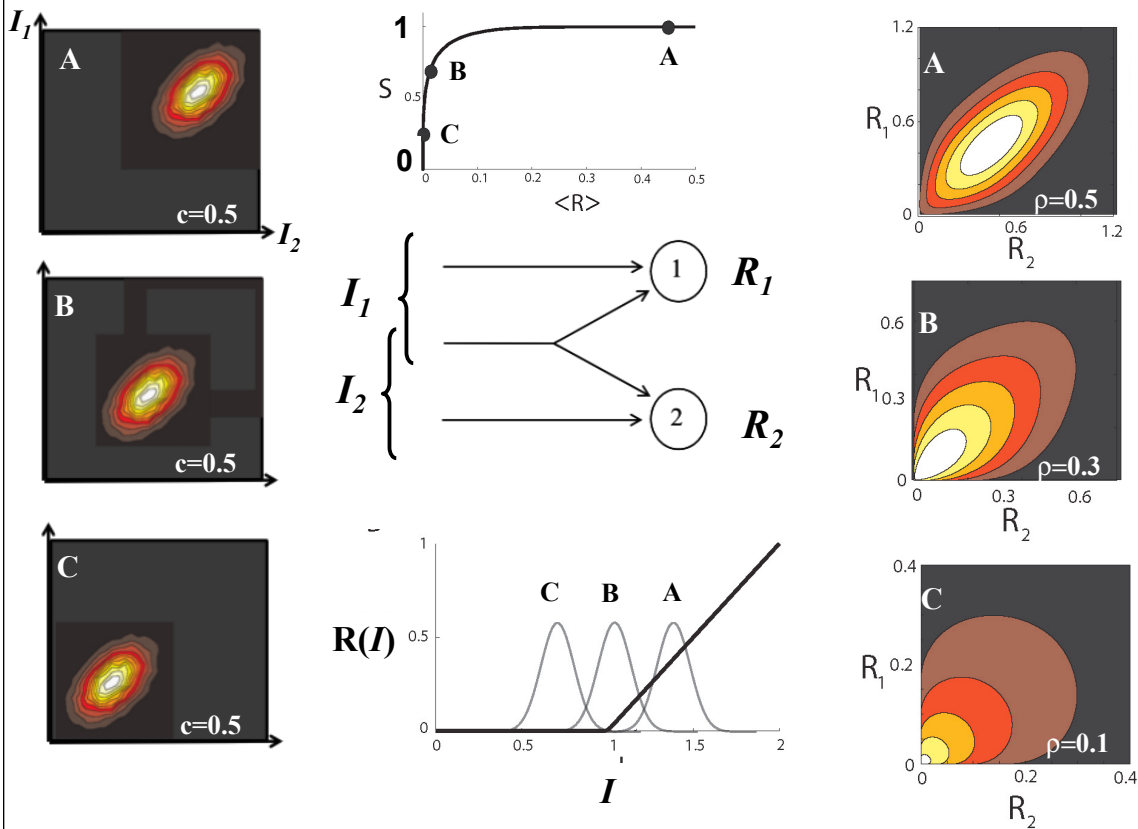
Simplest model



Simplest model

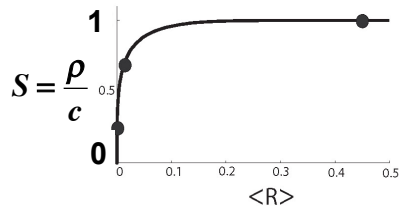


Simplest model



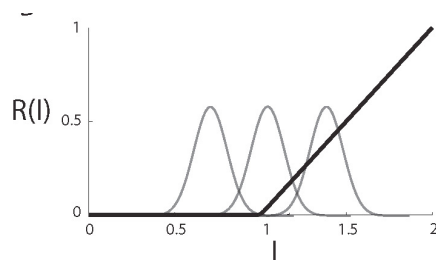
Simplest model

Integrate-and-fire model

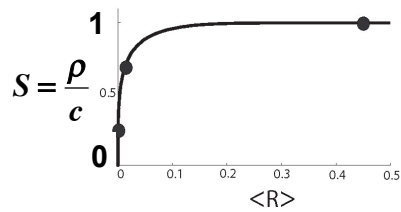


Correlations increase with rate

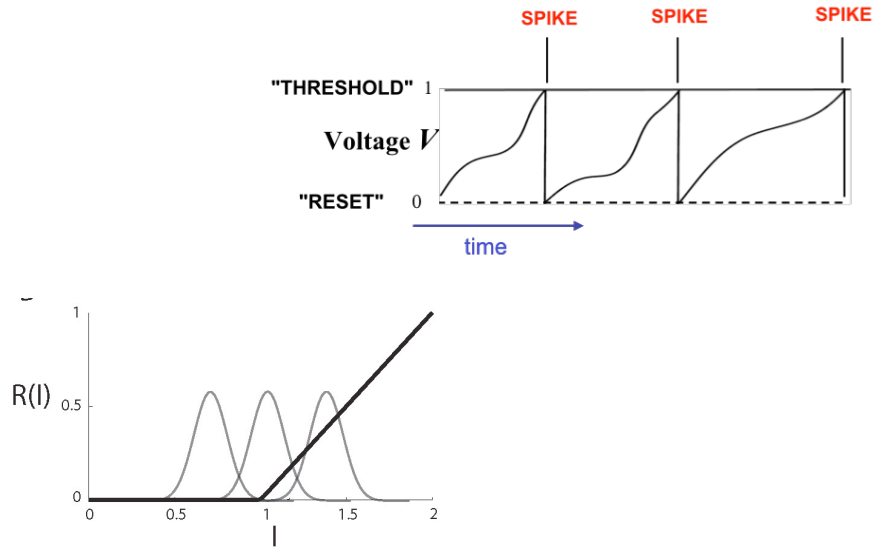
[de la Rocha, Doiron et al '07; Shea-Brown et al '08; Rosenbaum+Josic, '10]



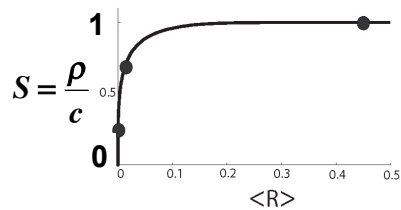
Simplest model



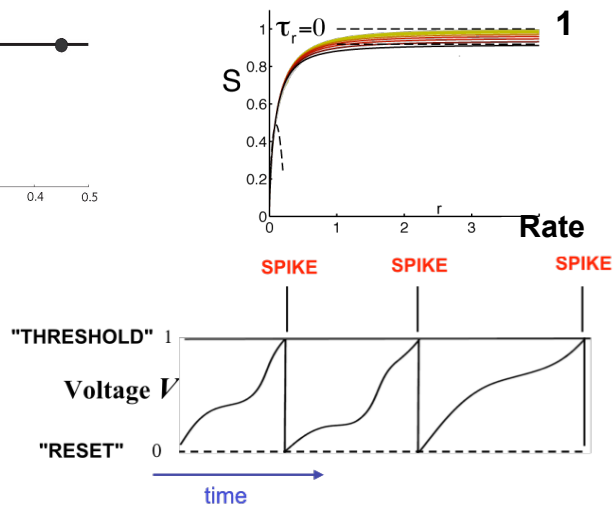
Integrate-and-fire model



Simplest model



Integrate-and-fire model

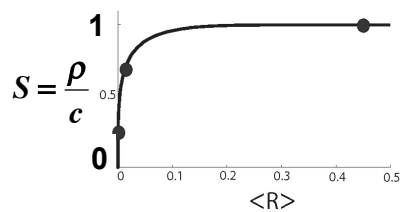


de la Rocha, Doiron et al, Nature '07
Shea-Brown et al, PRL '08

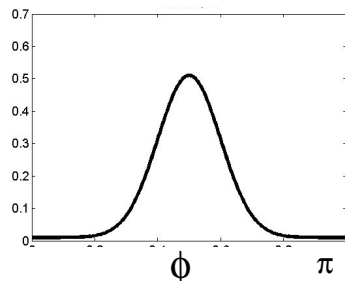
Spike generation and myriad other nonlinearities shape correlated spiking ...

... and can introduce *stimulus-dependent correlations*.

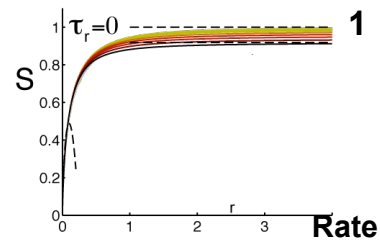
Simplest model



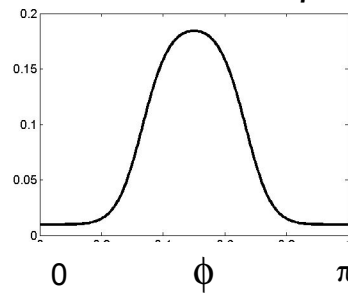
rate ν



Integrate-and-fire model



correlation ρ

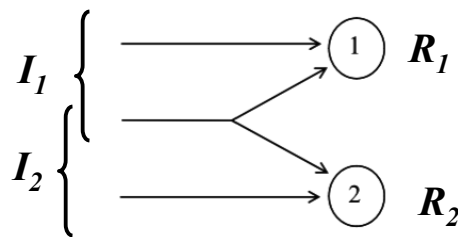
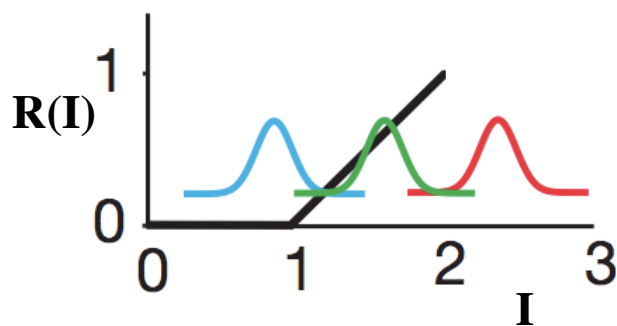


Spike generation and myriad other nonlinearities shape correlated spiking ...

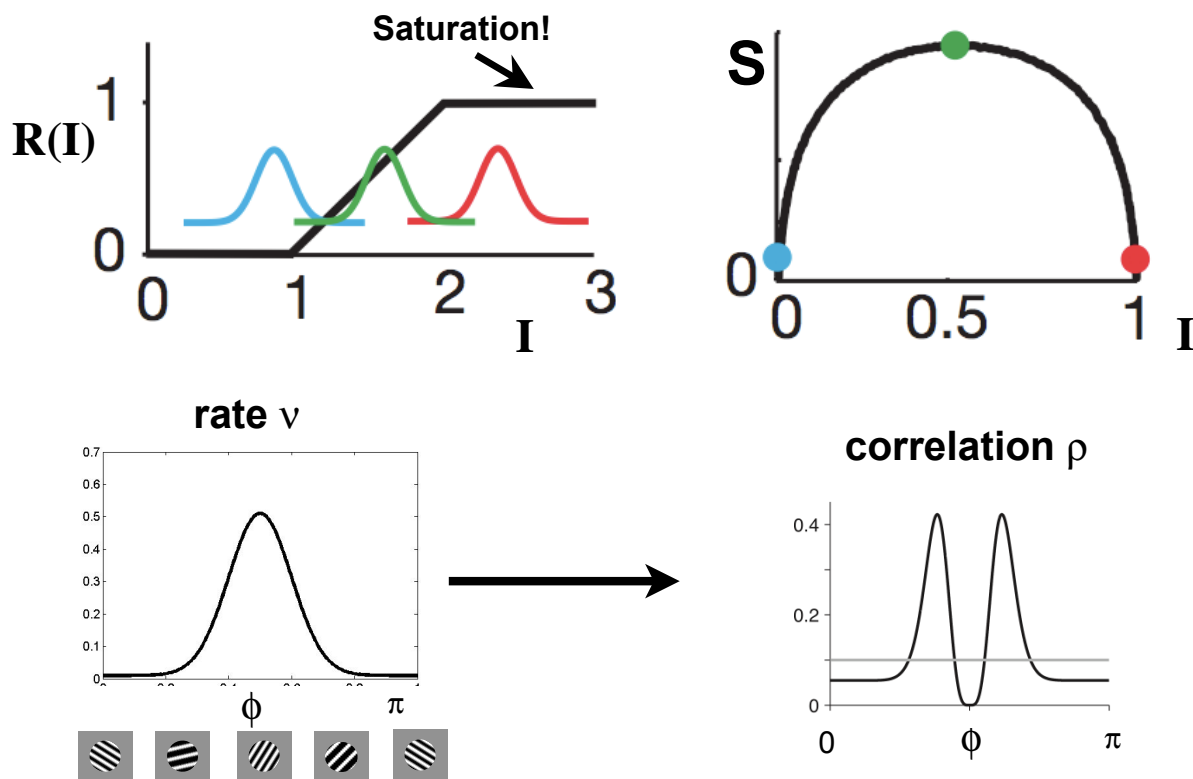
de la Rocha, Doiron et al, Nature '07
Shea-Brown et al, PRL '08

... and can introduce *stimulus-dependent correlations*.

Another simple (nonlinear) mechanism



Another simple (nonlinear) mechanism



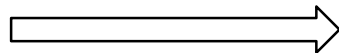
Pooling amplifies input correlations!

Renart, de la Rocha, Science '10;
Rosenbaum et al, Frontiers '10

Fraction p of
EPSPs is
common,
+ correlated w/
 r_{in}



Input
Currents,
correlation
 c



SPIKES,
correlation ρ

$$c = \frac{\text{cov}(\Sigma)}{\text{var}(\Sigma)}$$

$$c \gtrsim \frac{r_{in}N(N-1)}{N + r_{in}N(N-1)} \rightarrow 1$$

Denom in formula comes from
of correlated pairs ...
UNDEREST correlation if
ignore p

Numer is variance: variances
sum plus cross terms, p
makes not diff here ...

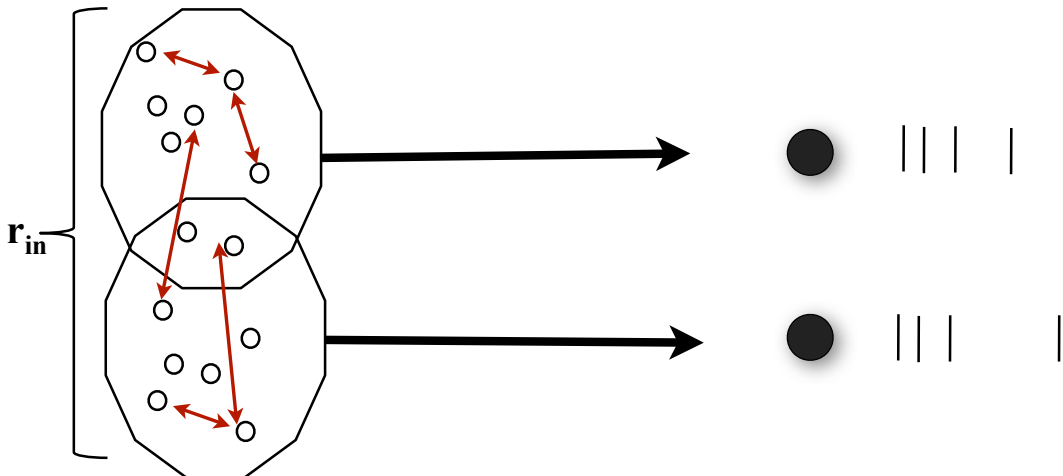


Recurrent connections: story gets surprising fast ...

e.g.

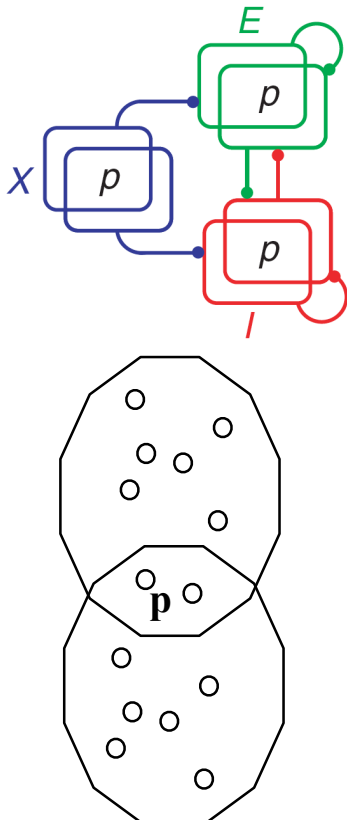
The Asynchronous State in Cortical Circuits

Alfonso Renart,^{1,†} Jaime de la Rocha,^{1,2,*} Peter Bartho,^{1,3} Liad Hollender,¹ Néstor Parga,⁴
Alex Reyes,² Kenneth D. Harris^{1,5,†}



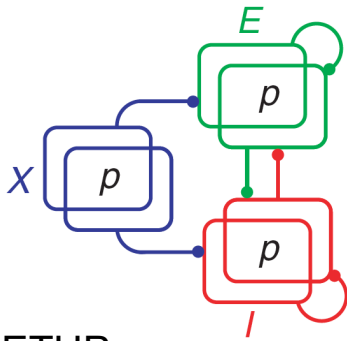
Suggests big network → big correlations?

Renart, de la Rocha,
Science '10



Suggests big network → big correlations?

Renart, de la Rocha,
Science '10



Main result: no.
Big network → small correlations

SETUP:

N cells / pop.

Connection proba $p=0.2$:

DENSE

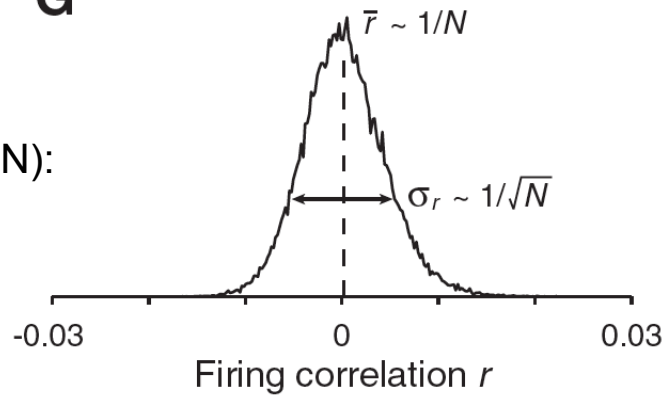
Connection strength $\sim 1/\sqrt{N}$:

STRONG

Mechanism: cancellation

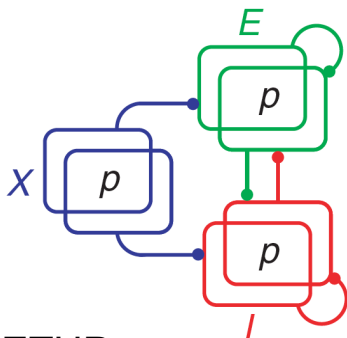
$$C = C_{EE} + C_{II} + 2C_{EI}$$

G



Suggests big network → big correlations?

Renart, de la Rocha,
Science '10



Main result: no.
Big network → small correlations

SETUP:

N cells / pop.

Connection proba $p=0.2$:

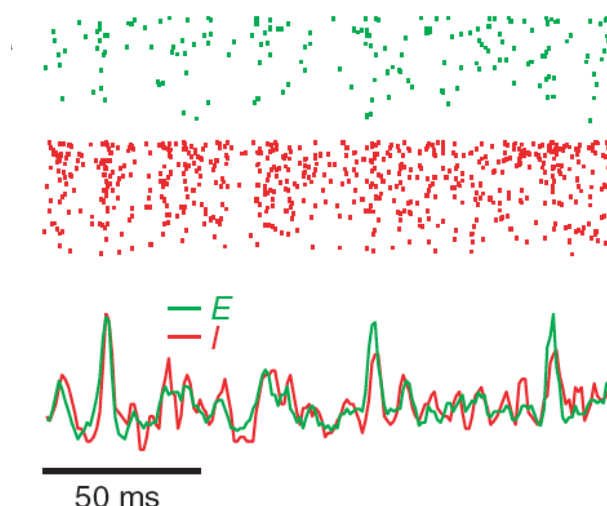
DENSE

Connection strength $\sim 1/\sqrt{N}$:

STRONG

Mechanism: cancellation

$$C = C_{EE} + C_{II} + 2C_{EI}$$



Time-integration:

Asynchronous firing

$$\bar{r} \sim 1/N$$

Pooling:

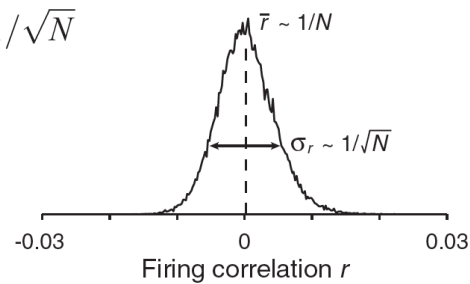
$$c \approx \frac{p + r_{in}(N - p)}{1 + r_{in}(N - 1)} \approx 1$$

Weakly asynchronous
total currents
 c , width $\sim 1/\sqrt{N}$

Synchronous
current components
 $c_{EE}, c_{II}, c_{EI}, c_{XX}$

Cancellation:

$$c = c_{EE} + c_{II} + 2c_{EI} \sim 1/\sqrt{N}$$



Mechanism: cancellation

$$c = c_{EE} + c_{II} + 2c_{EI}$$

OUTLINE

CONSEQUENCES OF CORRELATED SPIKING

Impact on coding

Impact on signal propagation

BASIC MECHANISMS FOR CORRELATED SPIKING

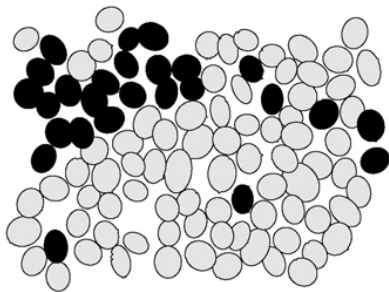
Common input → rate-dependent correlations

Pooling over correlated population → amplification of correlations

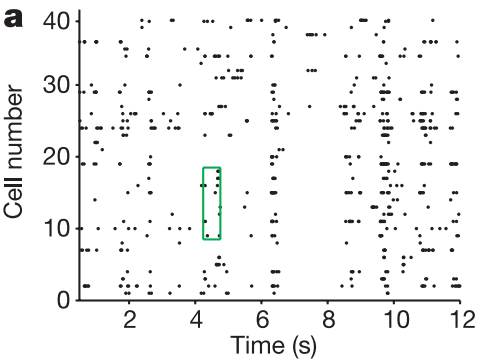
Recurrent balanced networks → cancellation of correlations

BEYOND CELL-PAIRS: HIGHER-ORDER CORRELATIONS

Population-wide
spiking dynamics

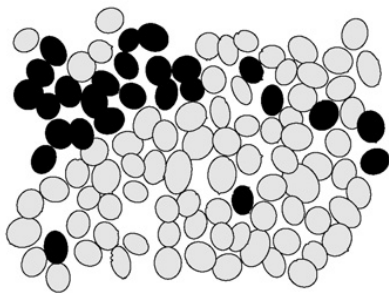


Graphic:
Shlens, Rieke and Chichilnisky, 2008

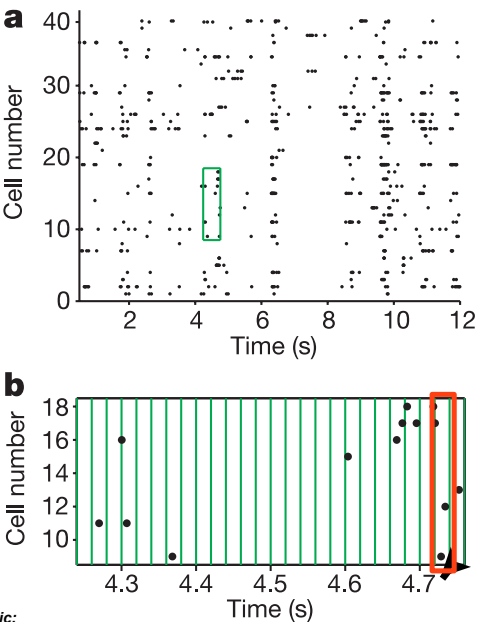


Graphic:
Schneidman et al. 2006

Population-wide
spiking dynamics



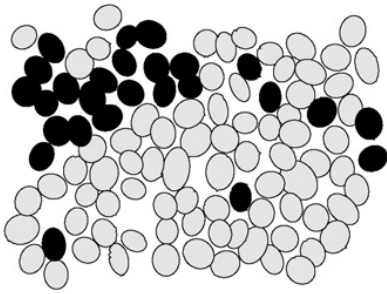
Graphic:
Shlens, Rieke and Chichilnisky, 2008



Graphic:
Schneidman et al. 2006

1001000010

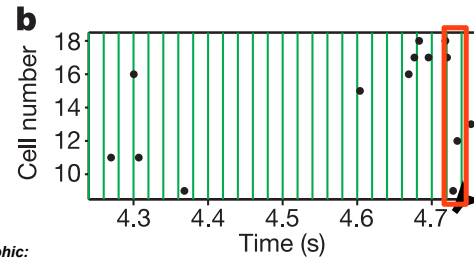
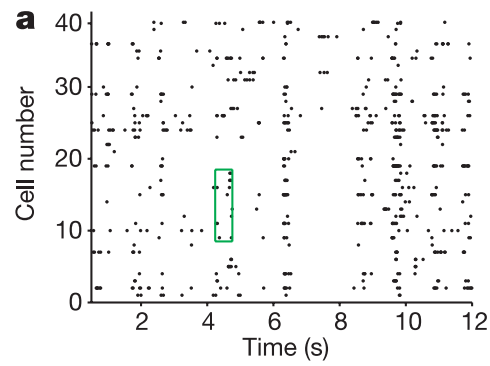
Population-wide spiking dynamics



Graphic:
Shlens, Rieke and Chichilnisky, 2008

$$x_j = \{0,1\}$$

$$P(x_1, x_2, \dots, x_N)$$

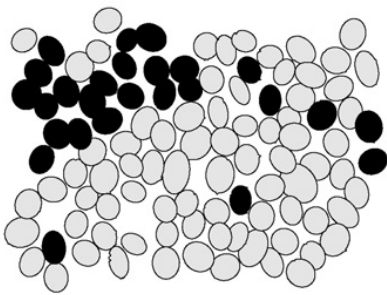


Graphic:
Schneidman et al. 2006

1001000010

67

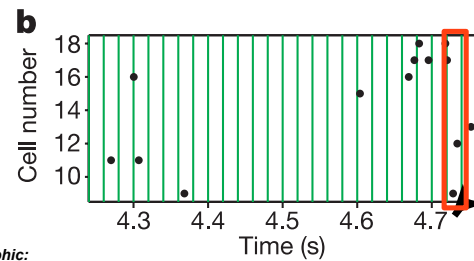
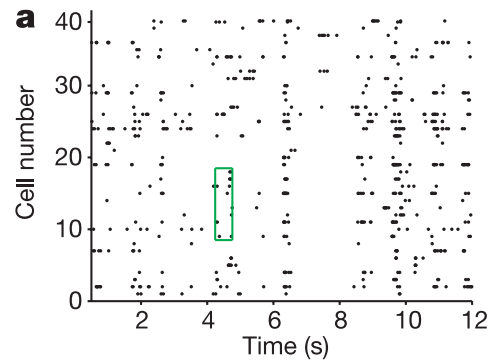
Population-wide spiking dynamics



Graphic:
Shlens, Rieke and Chichilnisky, 2008

$$x_j = \{0,1\}$$

$$P(x_1, x_2, \dots, x_N)$$



Graphic:
Schneidman et al. 2006

1001000010

68

Log-linear probability distribution

$$x_j = \{0,1\}$$

[Martignon et al, '95; Amari et al, '01; Schneidman et al, '06, Shlens et al, '06, '09, ...]

$$P(x_1, x_2, \dots, x_N) = \frac{1}{Z} \exp \left(\sum_i \lambda_i x_i + \sum_{i,j} \lambda_{ij} x_i x_j + \sum_{i,j,k} \lambda_{ijk} x_i x_j x_k + \dots \right)$$

2^N parameters (one for each state) $\rightarrow$ complete description

$N=100 \rightarrow 10^{30}$ parameters / impossibly complex

Maximum entropy approach:

Choose observables. $f_n(x_1, x_2, \dots, x_N)$.

Measure their averages: $\langle f_n \rangle$.

max:

$$H(P) = - \sum_{\{\vec{x} \in S\}} P(\vec{x}) \log P(\vec{x})$$

Fit λ parameters so $\langle f_n \rangle$ hold but minimal further assumptions.

$$\text{Get } P(x_1, x_2, \dots, x_N) = \frac{1}{Z} \exp \left(\sum_n \lambda_n f_n(x_1, x_2, \dots, x_N) \right)$$

Maximum entropy approach:

Choose observables. $f_n(x_1, x_2, \dots, x_N)$.

Measure their averages: $\langle f_n \rangle$.

Fit λ parameters so $\langle f_n \rangle$ hold but minimal further assumptions.

$$\text{Get } P(x_1, x_2, \dots, x_N) = \frac{1}{Z} \exp \left(\sum_n \lambda_n f_n(x_1, x_2, \dots, x_N) \right)$$

Choose $\{f_n(x_1, x_2, \dots, x_N)\} = \{x_1, x_2, \dots, x_1 x_2, \dots\}$.

Measure means + second-order moments $\langle x_1 \rangle, \dots, \langle x_1 x_2 \rangle, \dots$.

$$\text{Get } P(x_1, x_2, \dots, x_N) = \frac{1}{Z} \exp \left(\sum_i \lambda_i x_i + \sum_{i,j} \lambda_{ij} x_i x_j \right)$$

PAIRWISE MAXIMUM-ENTROPY MODEL P_2

Minimal-assumptions model that fits means + pairwise correlations

If accurate, declare: no “extra” beyond-pairwise correlations

“Accurate” means small Kullback-Leibler distance from true distribution P

$$D_{KL}(P, P_2) \equiv \sum_{\{\vec{x} \in S\}} P(\vec{x}) \log \left(\frac{P(\vec{x})}{P_2(\vec{x})} \right) \\ = H(P_2) - H(P)$$

Choose $\{f_n(x_1, x_2, \dots, x_N)\} = \{x_1, x_2, \dots, x_1 x_2, \dots\}$.

Measure means + second-order moments $\langle x_1 \rangle, \dots, \langle x_1 x_2 \rangle, \dots$.

$$\text{Get } P_2(x_1, x_2, \dots, x_N) = \frac{1}{Z} \exp \left(\sum_i \lambda_i x_i + \sum_{i,j} \lambda_{ij} x_i x_j \right)$$

PAIRWISE MAXIMUM-ENTROPY MODEL P₂

Minimal-assumptions model that fits means + pairwise correlations

If accurate, declare: no “extra” beyond-pairwise correlations

SUMMARY

CONSEQUENCES OF CORRELATED SPIKING

Impact on coding

- (a) Homogeneous populations: limits population averaging / degrades info
- (b) Heterogeneous cell pairs ...
 - similar stimulus tuning: **DEGRADE CODING**
 - different stimulus tuning: **ENHANCE CODING**
- (c) Heterogeneous populations: competing effects

Impact on signal propagation

Correlation sets *gain*: Downstream rate ~ upstream rate X upstream correlation

Review: Averbek et al, Nature Rev. Nsci. '06

BASIC MECHANISMS FOR CORRELATED SPIKING

- Common input → rate-dependent correlations
- Pooling over correlated population → amplification of correlations
- Recurrent balanced networks → cancellation of correlations

BEYOND CELL-PAIRS: HIGHER-ORDER CORRELATIONS

- Maximum-entropy methods measure via log-linear model
- Mixed results for presence and impact on coding