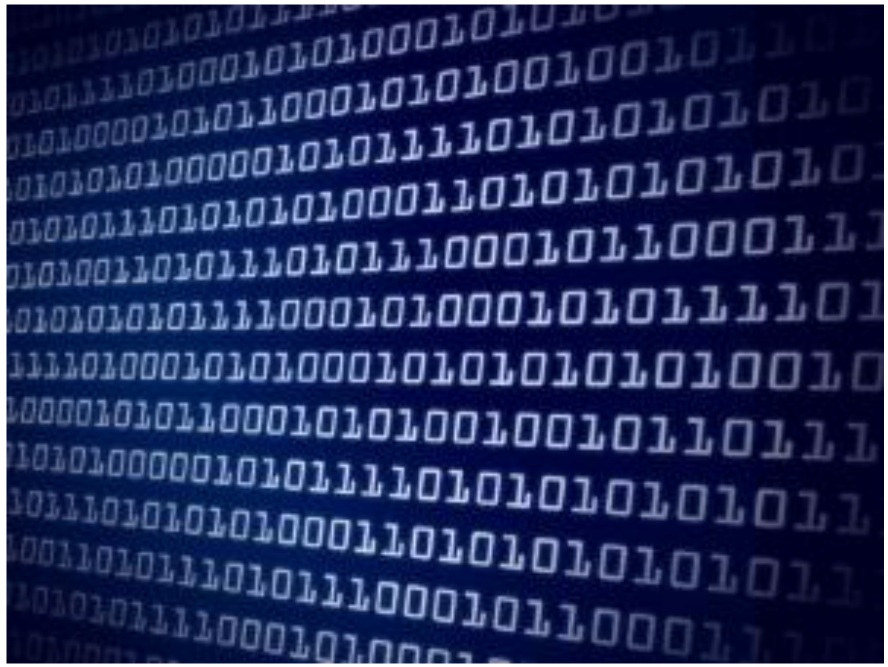
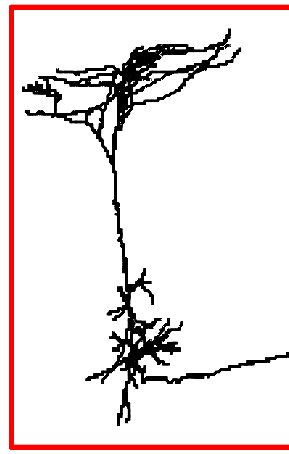
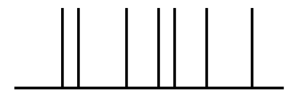
The Neural Code



Part 2: Population coding
Odelia Schwartz

Single neuron Encoding



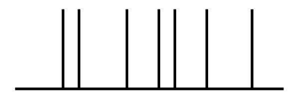


Response

Probability(Response | Stimulus)

Last time: encoding model

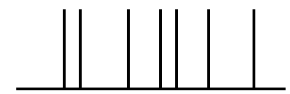
→ Linear, Nonlinear →



Poisson spiking

Last time: encoding model

→ Linear, Nonlinear →

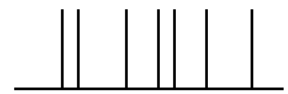


Poisson spiking

We talked about estimating the linear filter
(What filters??)

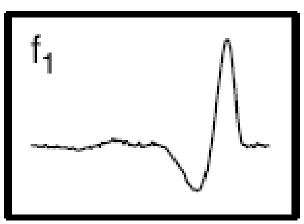
Last time: encoding model

→ Linear, Nonlinear →



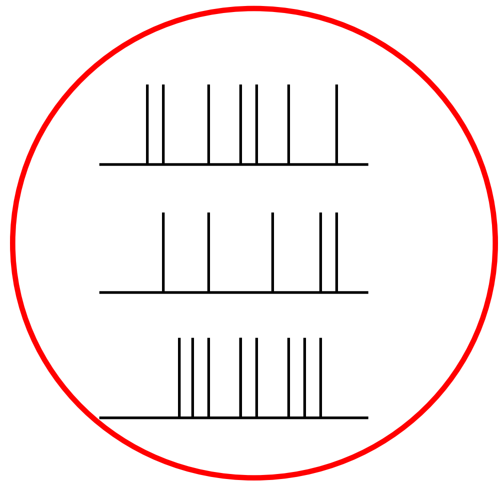
Poisson
spiking

We talked about estimating the linear filter
(orientation filter, time filter)



This class: Populations

→



Population responses

Probability(Responses | Stimulus)

This class: focus on Decoding

Decoding: the reverse problem…

Probability(Stimulus | Response)

Population Coding

Do brains use many or few neurons to represent the world and guide actions?

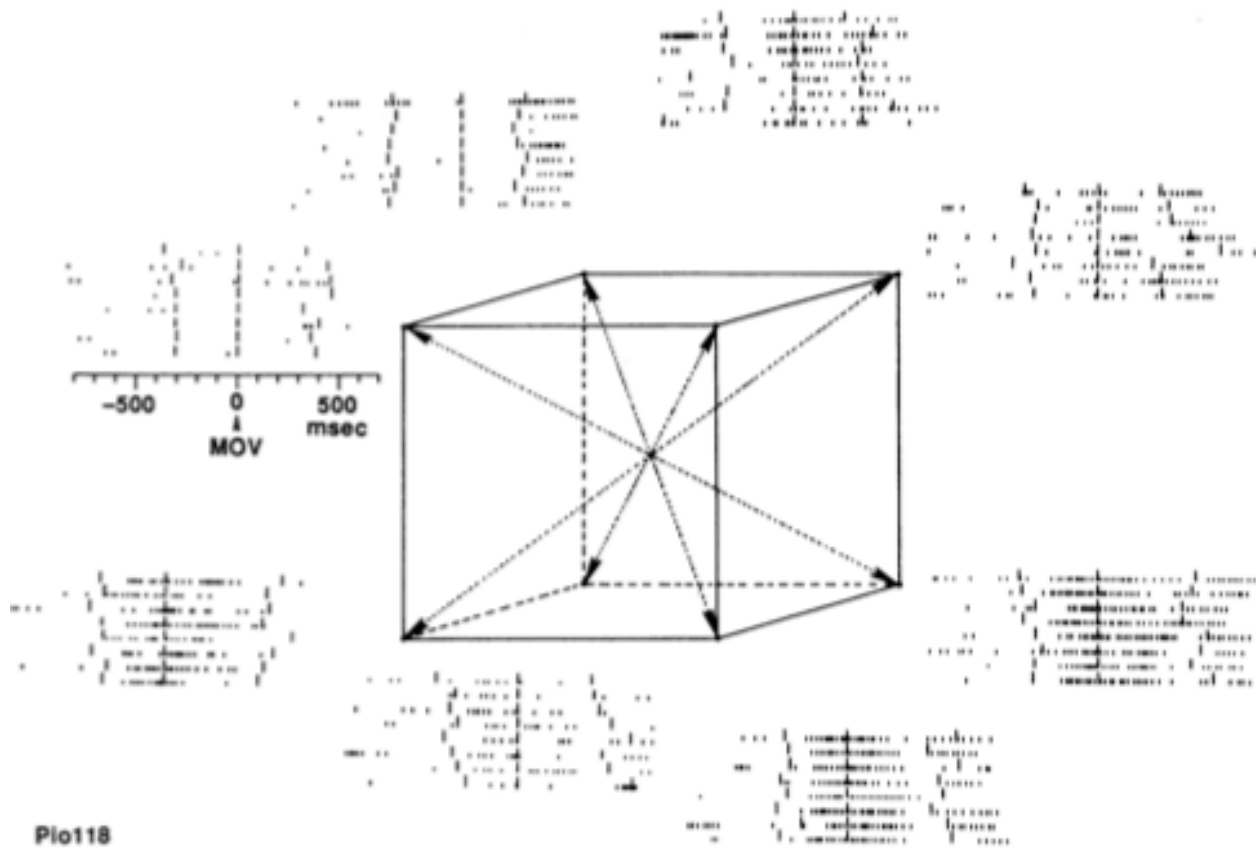
Population coding

Sparse representation

(Grandmother cells)

Selectivity leads to sparseness.

Population coding example



Example neuron in primary motor cortex
(from Schwartz & Georgopoulos 1986)

Sparse vs. Distributed Representations

Disadvantages (Advantages)

Distributed

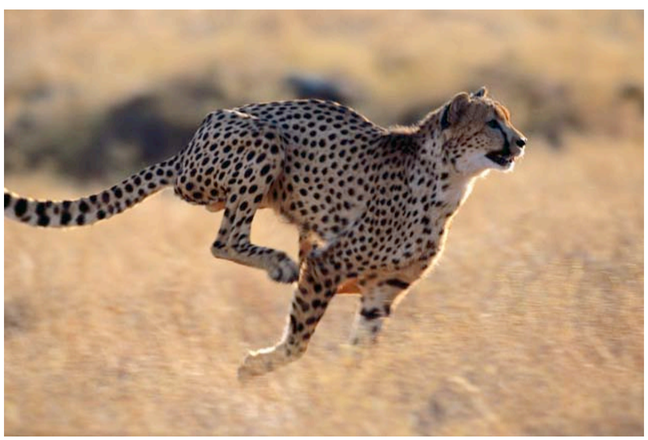
Sparse

Sparse vs. Distributed Representations

	Disadvantages (Advantages)
Distributed	*Metabolic cost* *Decoding/Links with other systems*
Sparse	*Cell Death* *Combinatorial explosion*

Why population codes

- Decoding properties of input (motion direction, color, warm day) likely involves neural population
- Linking between neural responses and perception/behavior



Types of questions with population codes

- How well can we (or populations of neurons) decode a given stimulus (what do we want to decode?)

- What encoding (and decoding) schemes are optimal in allowing us to best estimate properties of the stimulus?

Population codes

- Primary visual cortex (eg, orientation)
- Primary motor cortex (eg, arm movement)
- Higher areas…
- Hippocampus (self location)
- Cercal interneurons in cricket

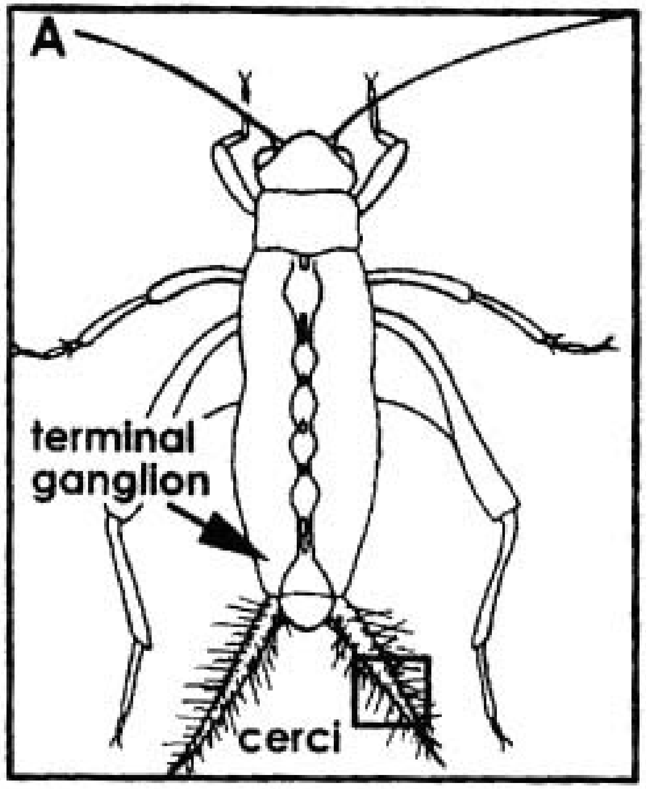
Population codes

- Primary visual cortex (eg, orientation)
- Primary motor cortex (eg, arm movement)
- Higher areas…
- Hippocampus (self location)
- Cercal interneurons in cricket

Population coding example

Decoding wind direction in the cricket cercal system

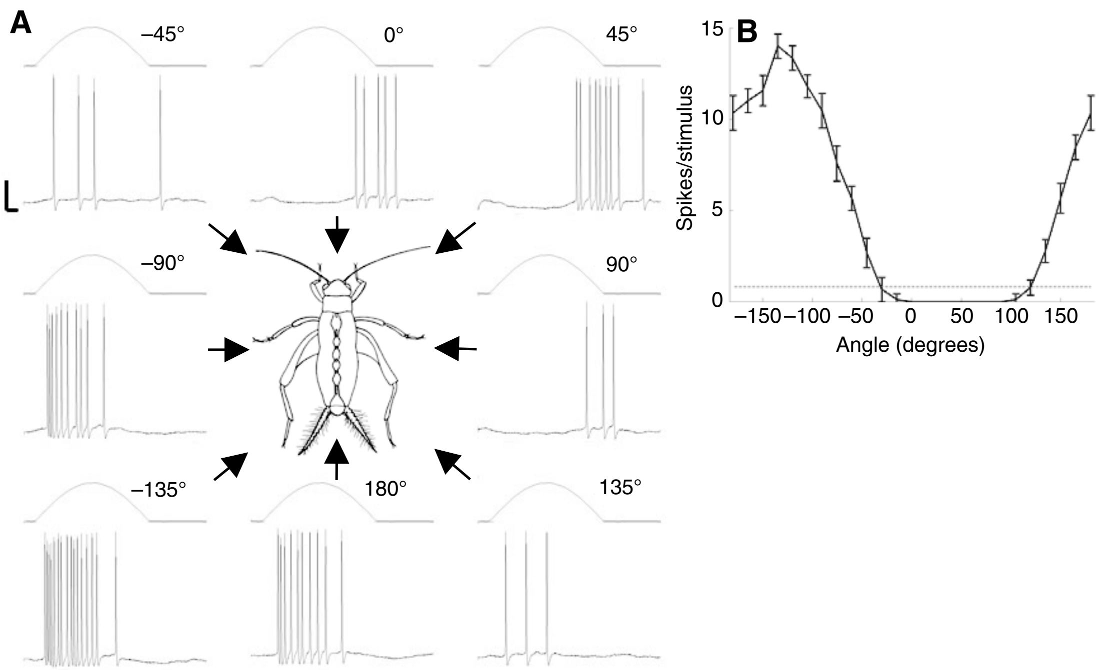
Population coding example



Neurons sensitive to wind angle

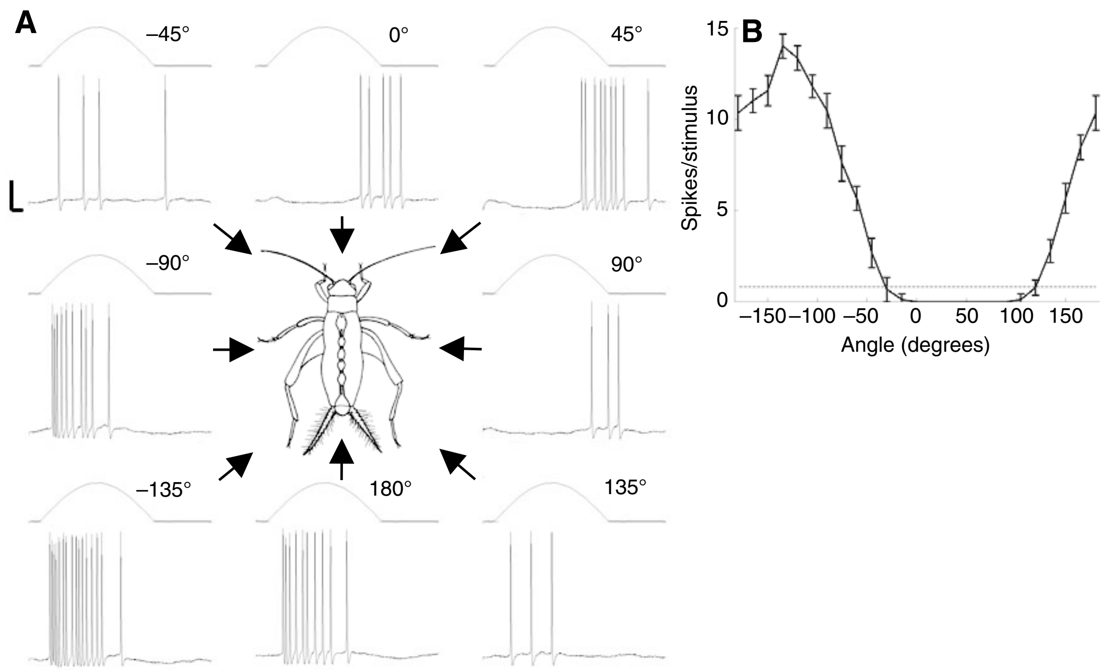
Decoding wind direction in the cricket cercal system (sensing direction of wind movement as warning against predators)

Population coding example
First look at one neuron



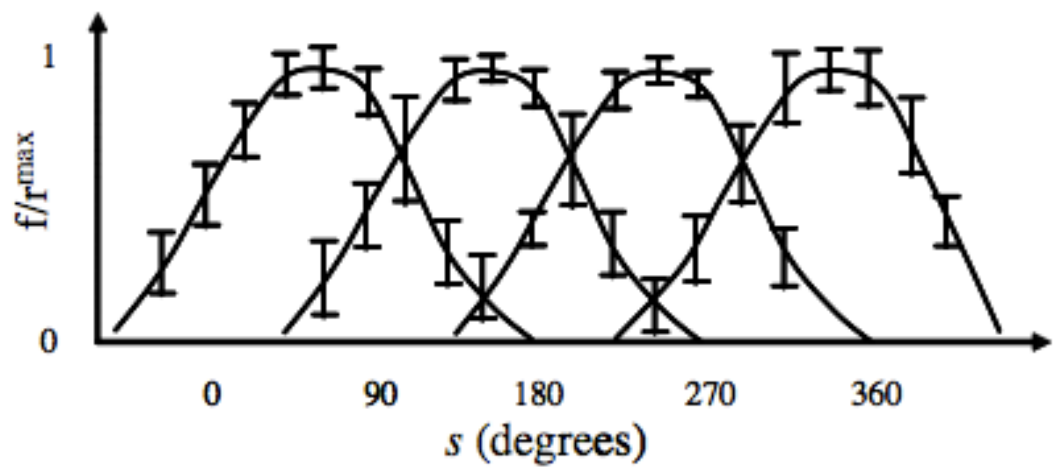
Is single neuron sufficient to "decode" wind direction?

Population coding example
First look at one neuron



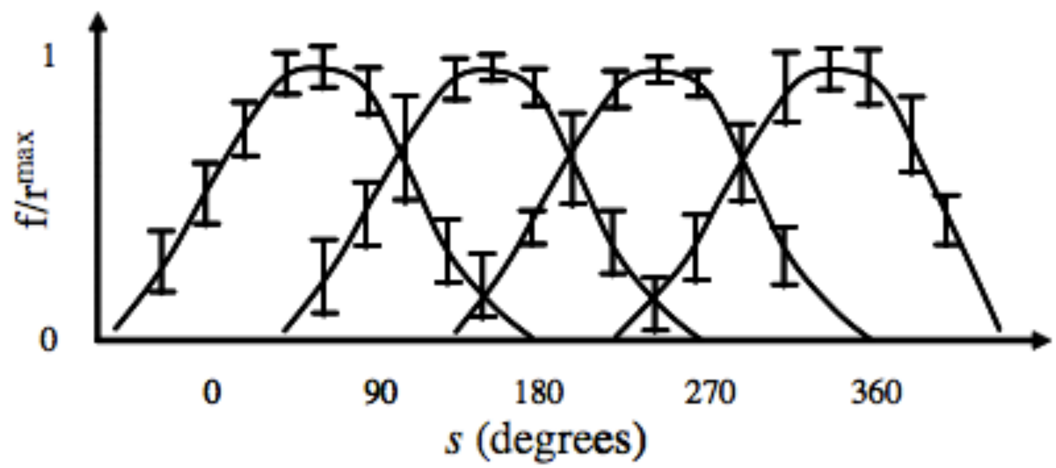
Is single neuron sufficient to "decode" wind direction? How many neurons needed?

Population coding example
4 neurons!



Wind orientation tuning curves include only four cardinal axes! (from Dayan and Abbott book)

Population coding example
4 neurons!

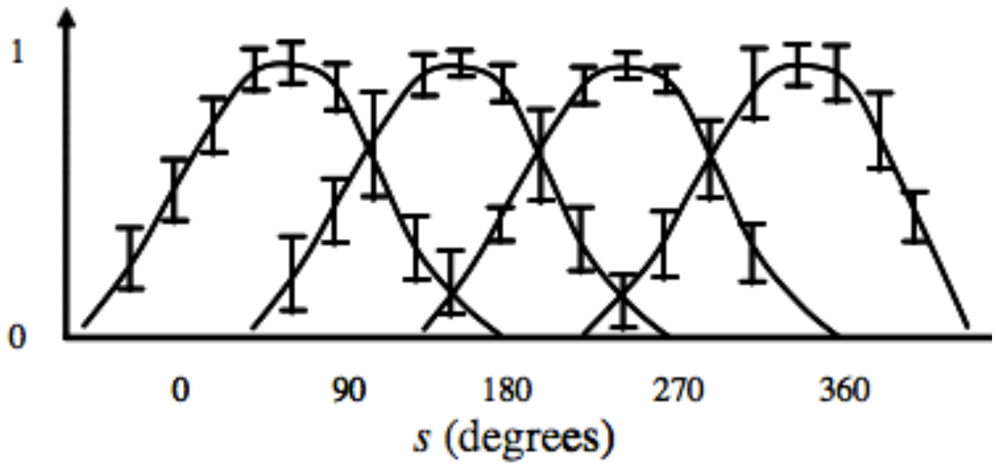


Each of the 4 neurons has a cosine tuning curve:

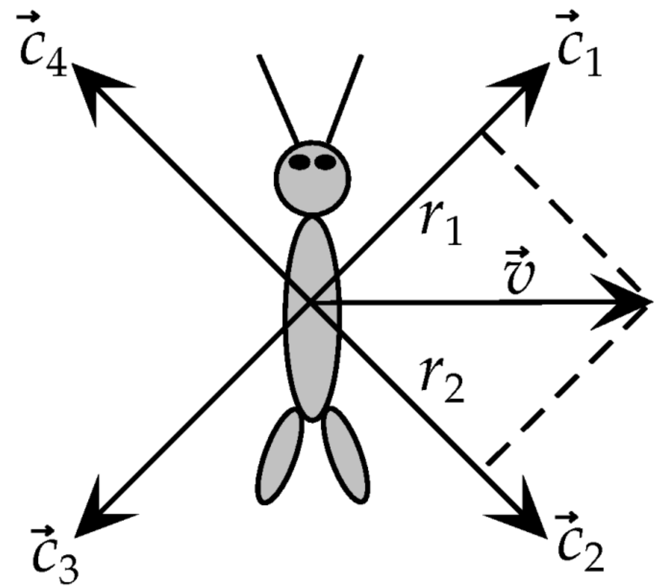
$$r_1 = \cos(\theta - \theta_1)$$

(wind direction minus preferred; and firing rate made positive)

Population coding example
Geometric depiction:



Neurons sensitive to angle of wind around vertical axis; and not elevation above horizontal plane



Decoding wind direction in the cricket cercal system with 4 interneurons (from Dayan and Abbott book);
On the board…

Population coding example

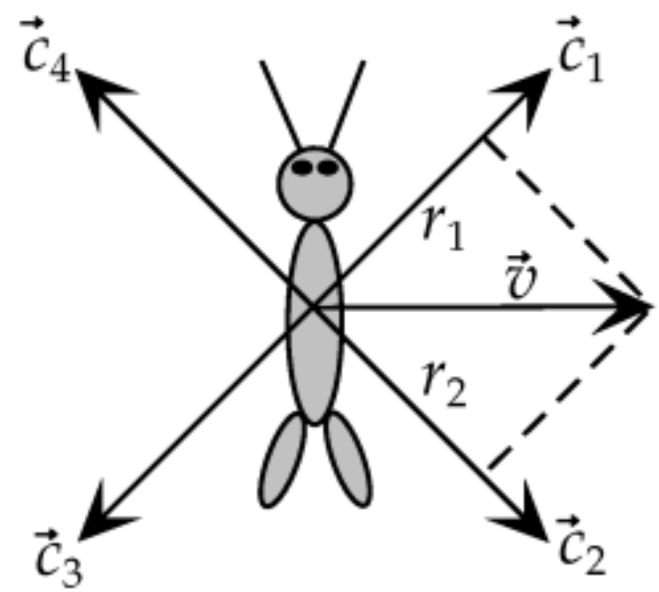
$\vec{V} = (V_1, V_2)$ Wind direction (we want to estimate)
Assume unit length vector.

$\overrightarrow{C_1}, \overrightarrow{C_2}, \overrightarrow{C_3}, \overrightarrow{C_4}$ Preferred wind direction each neuron (unit length)

r_1, r_2, r_3, r_4

Firing rate each neuron to given wind direction stimulus

$r_1 = \cos(\theta - \theta_1) = \overrightarrow{C_1}\vec{V}$



Population coding example

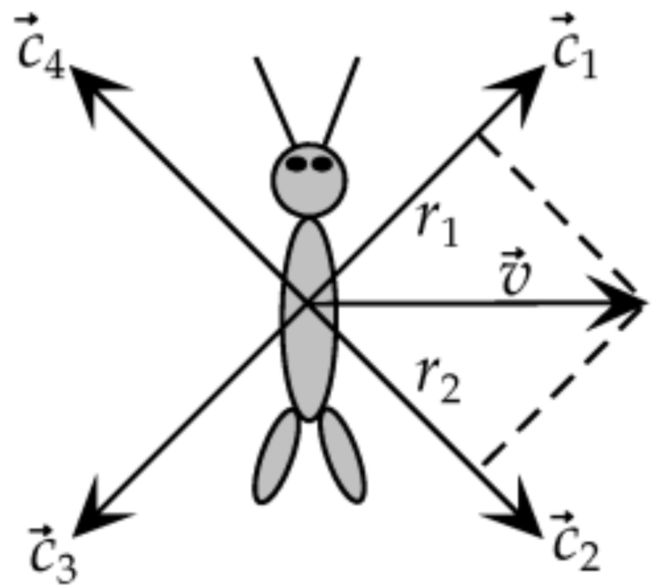
$\vec{V} = (V_1, V_2)$ Wind direction (want to estimate)
Assume unit length vector.

$\overrightarrow{C_1}, \overrightarrow{C_2}, \overrightarrow{C_3}, \overrightarrow{C_4}$ Preferred wind direction each neuron

$$\overrightarrow{C_1^T}\,\overrightarrow{C_2} = 0$$

$$\overrightarrow{C_3} = -\overrightarrow{C_1}$$

$$\overrightarrow{C_4} = -\overrightarrow{C_2}$$

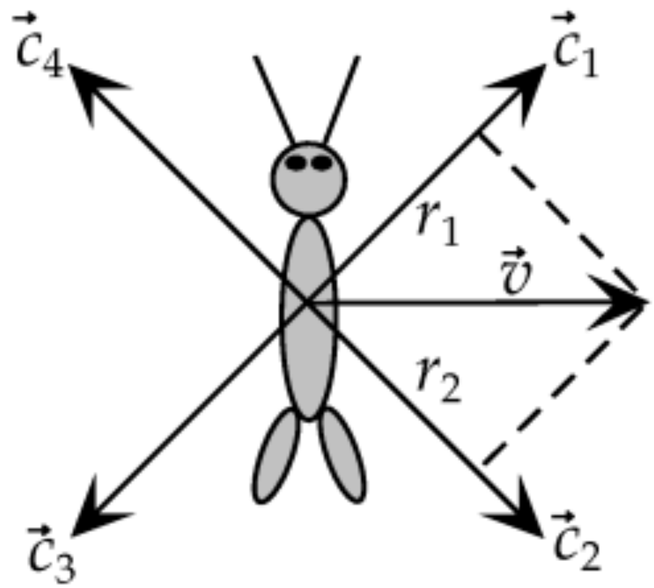


Population coding example

$\vec{V} = (V_1, V_2)$ Wind direction (want to estimate)
Assume unit length vector.

$\overrightarrow{C_1}, \overrightarrow{C_2}, \overrightarrow{C_3}, \overrightarrow{C_4}$ Preferred wind direction each neuron

$$\vec{V} = r_1\overrightarrow{C_1} + r_2\overrightarrow{C_2}$$



Population coding example

$\vec{V} = (V_1, V_2)$

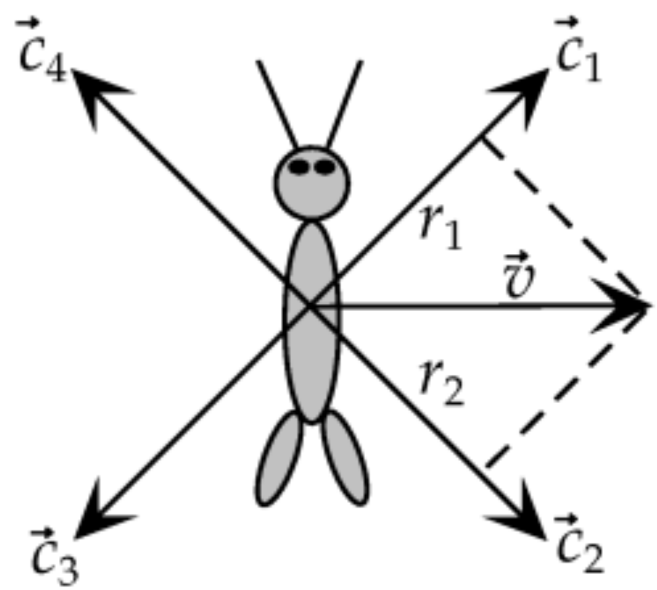
Wind direction (want to estimate)
Assume unit length vector.

$\overrightarrow{C_1}, \overrightarrow{C_2}, \overrightarrow{C_3}, \overrightarrow{C_4}$

Preferred wind direction each neuron

$$\vec{V} = r_1\overrightarrow{C_1} + r_2\overrightarrow{C_2}$$

In principle, two neurons could be enough for all directions. Why not?



Population coding example

$\vec{V} = (V_1, V_2)$

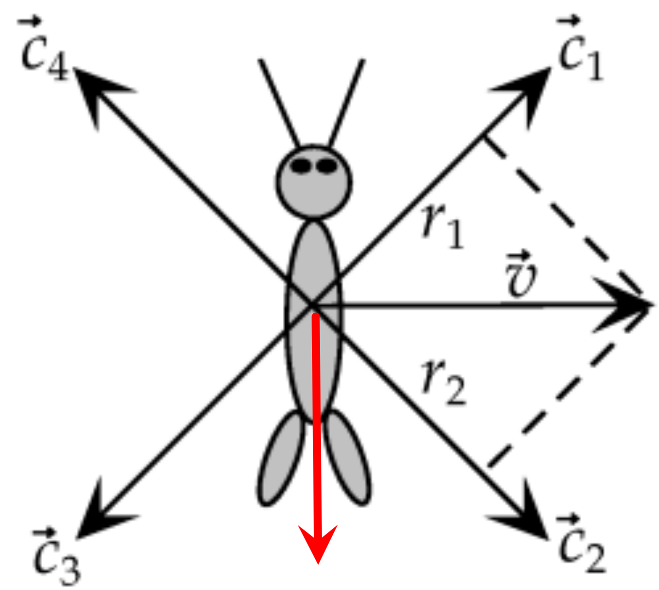
Wind direction (want to estimate)
Assume unit length vector.

$\overrightarrow{C_1}, \overrightarrow{C_2}, \overrightarrow{C_3}, \overrightarrow{C_4}$

Preferred wind direction each neuron

$$\vec{V} = r_1\overrightarrow{C_1} + r_2\overrightarrow{C_2}$$

In principle, two neurons could be enough for all directions. Why not?
Firing rates not negative,
Can't use C_1 if wind direction were the other way



Population coding example

$\vec{V} = (V_1, V_2)$

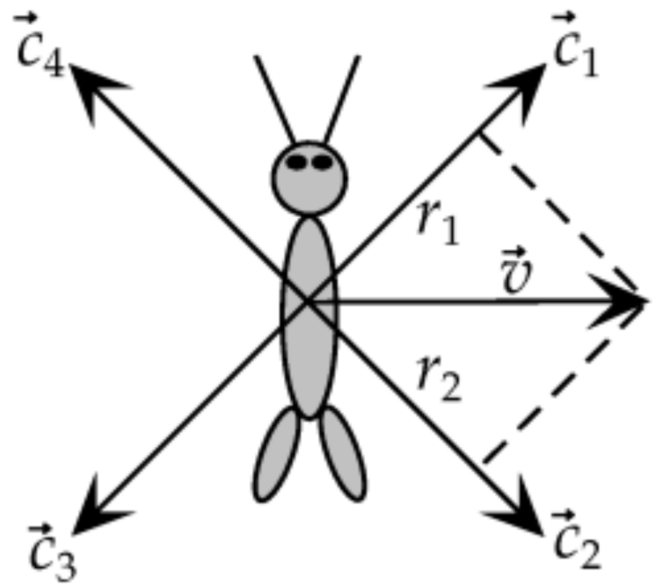
Wind direction (want to estimate)
Assume unit length vector.

$\overrightarrow{C_1}, \overrightarrow{C_2}, \overrightarrow{C_3}, \overrightarrow{C_4}$

Preferred wind direction each neuron

$$\vec{V} = r_1\overrightarrow{C_1} + r_2\overrightarrow{C_2} - r_3\overrightarrow{C_3} - r_4\overrightarrow{C_4}$$

4 neurons = just right!



Population coding example

$\vec{V} = (V_1, V_2)$

Wind direction (want to estimate)
Assume unit length vector.

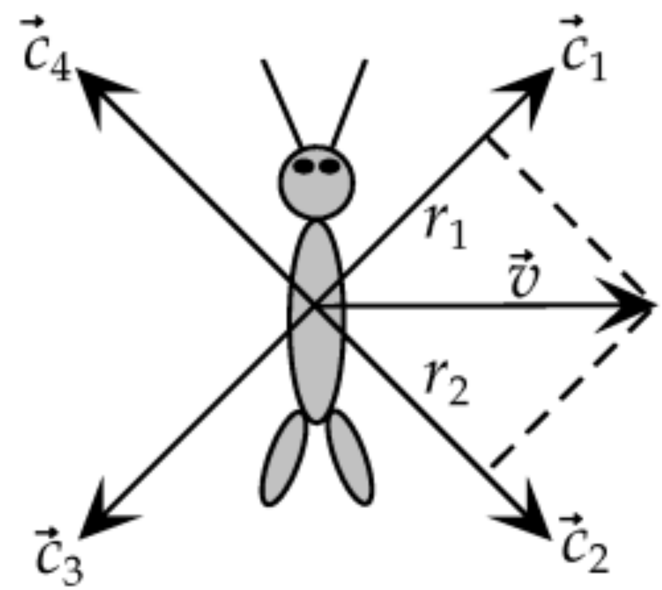
$\overrightarrow{C_1}, \overrightarrow{C_2}, \overrightarrow{C_3}, \overrightarrow{C_4}$

Preferred wind direction each neuron

$$\vec{V} = r_1\overrightarrow{C_1} + r_2\overrightarrow{C_2} - r_3\overrightarrow{C_3} - r_4\overrightarrow{C_4}$$

4 neurons = just right!

This is known as population vector decoding



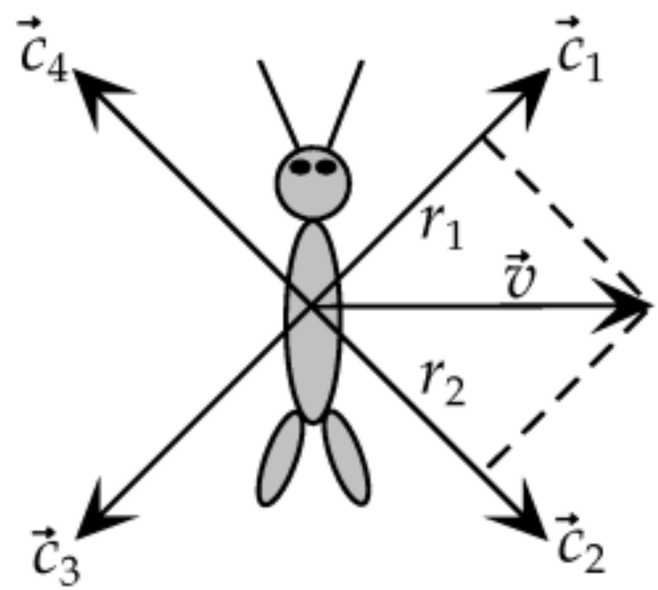
Population coding example

$\vec{V} = (V_1, V_2)$ — Wind direction (want to estimate) Assume unit length vector.

$\overrightarrow{C_1}, \overrightarrow{C_2}, \overrightarrow{C_3}, \overrightarrow{C_4}$ — Preferred wind direction each neuron

$$\vec{V} = \sum_{i=1}^{4} r_i' \overrightarrow{C_i}$$

This is known as population vector decoding (first used by Georgopoulos for motor system)



Population coding example

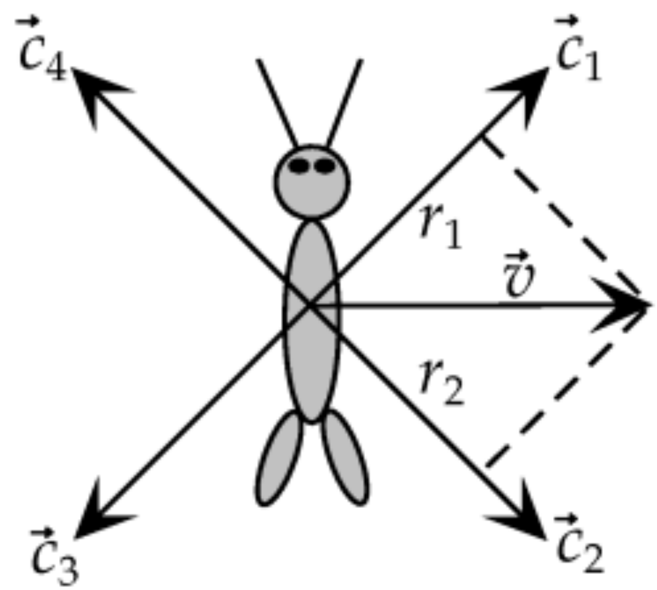
$\vec{V} = (V_1, V_2)$

Wind direction (want to estimate)
Assume unit length vector.

$\overrightarrow{C_1}, \overrightarrow{C_2}, \overrightarrow{C_3}, \overrightarrow{C_4}$

Preferred wind direction each neuron

$$\vec{V} = r_1\overrightarrow{C_1} + r_2\overrightarrow{C_2} - r_3\overrightarrow{C_3} - r_4\overrightarrow{C_4}$$



This is known as population vector decoding
Simple estimation!

Population coding example

$\vec{V} = (V_1, V_2)$

Wind direction (want to estimate)
Assume unit length vector.

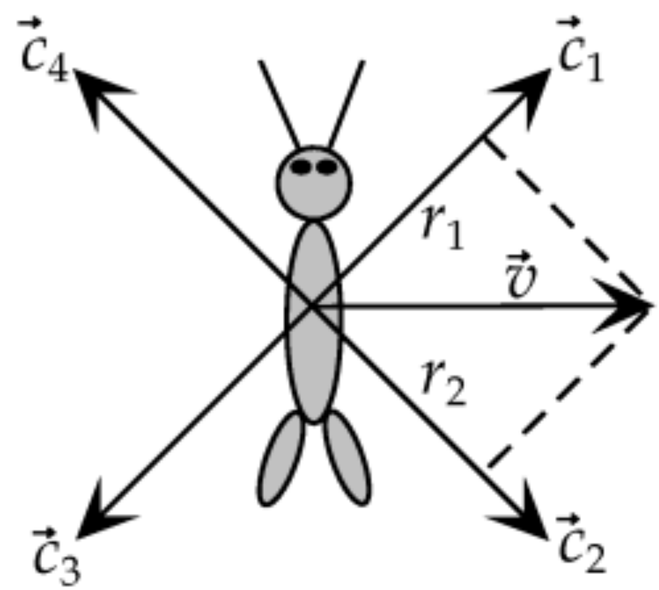
$\overrightarrow{C_1}, \overrightarrow{C_2}, \overrightarrow{C_3}, \overrightarrow{C_4}$

Preferred wind direction each neuron

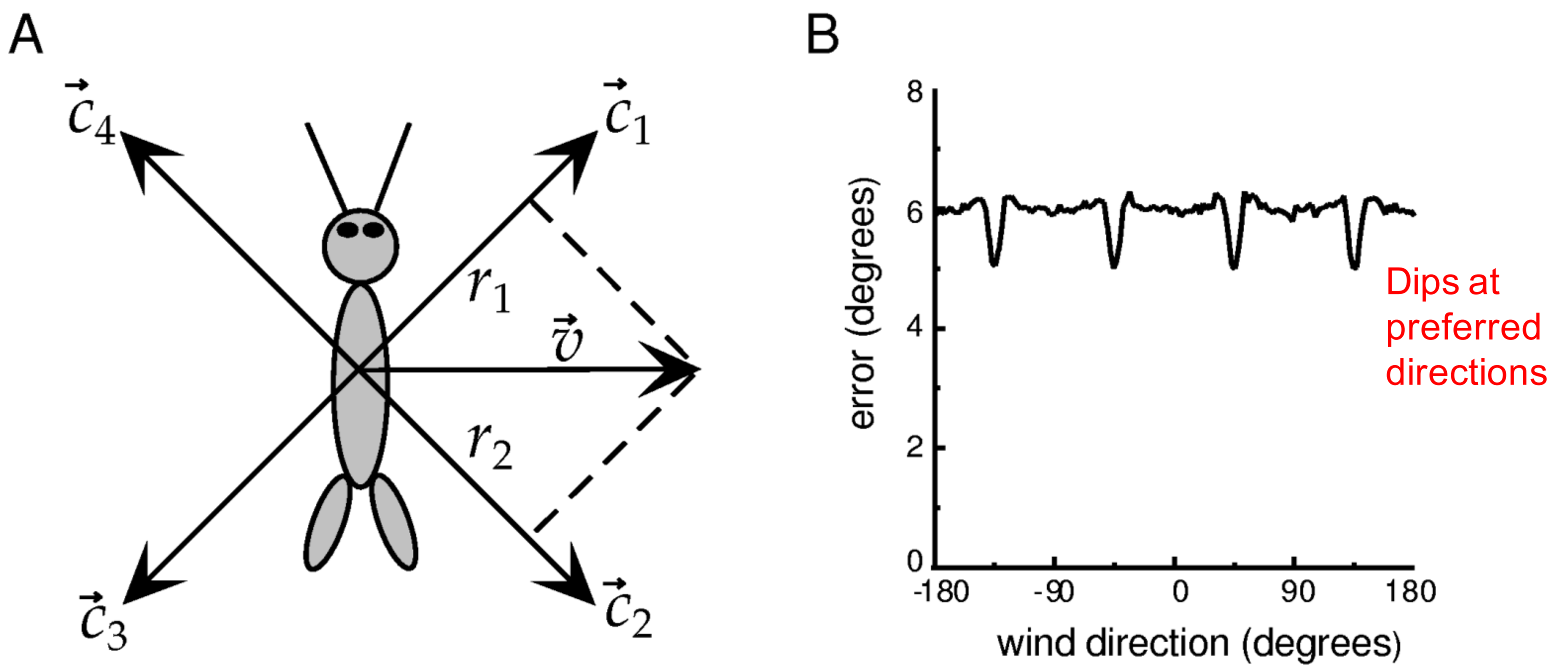
$$\vec{V} = r_1\overrightarrow{C_1} + r_2\overrightarrow{C_2} - r_3\overrightarrow{C_3} - r_4\overrightarrow{C_4}$$

This is known as population vector decoding

Cartesian coordinate system

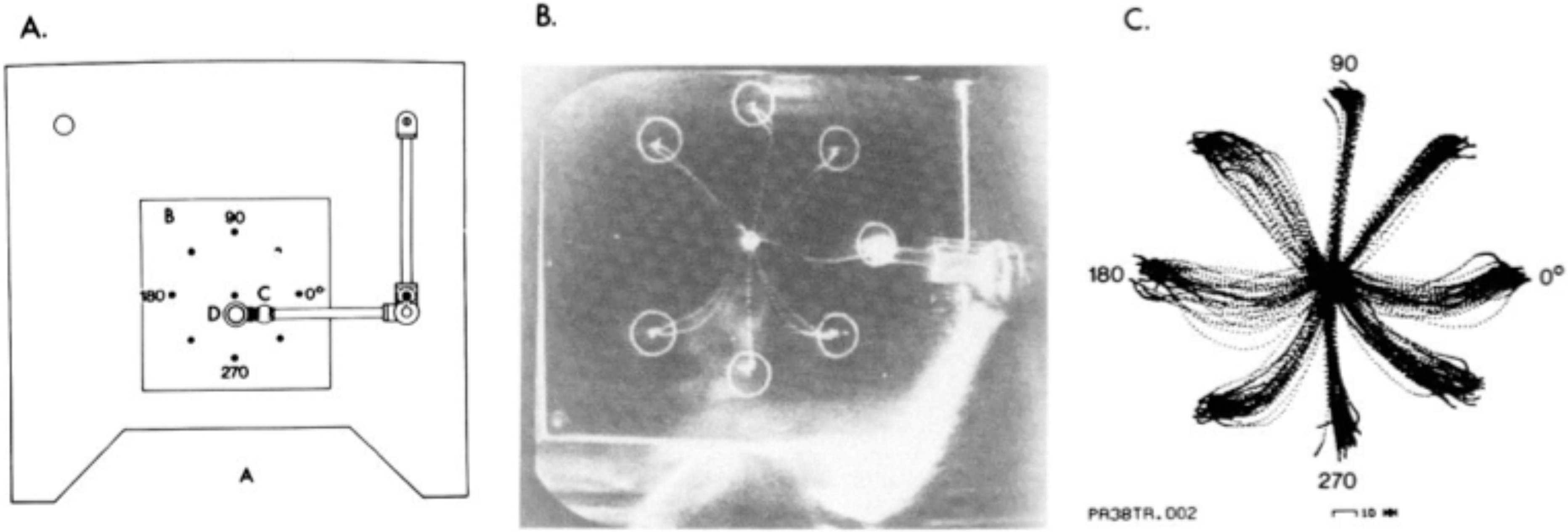


Population coding example



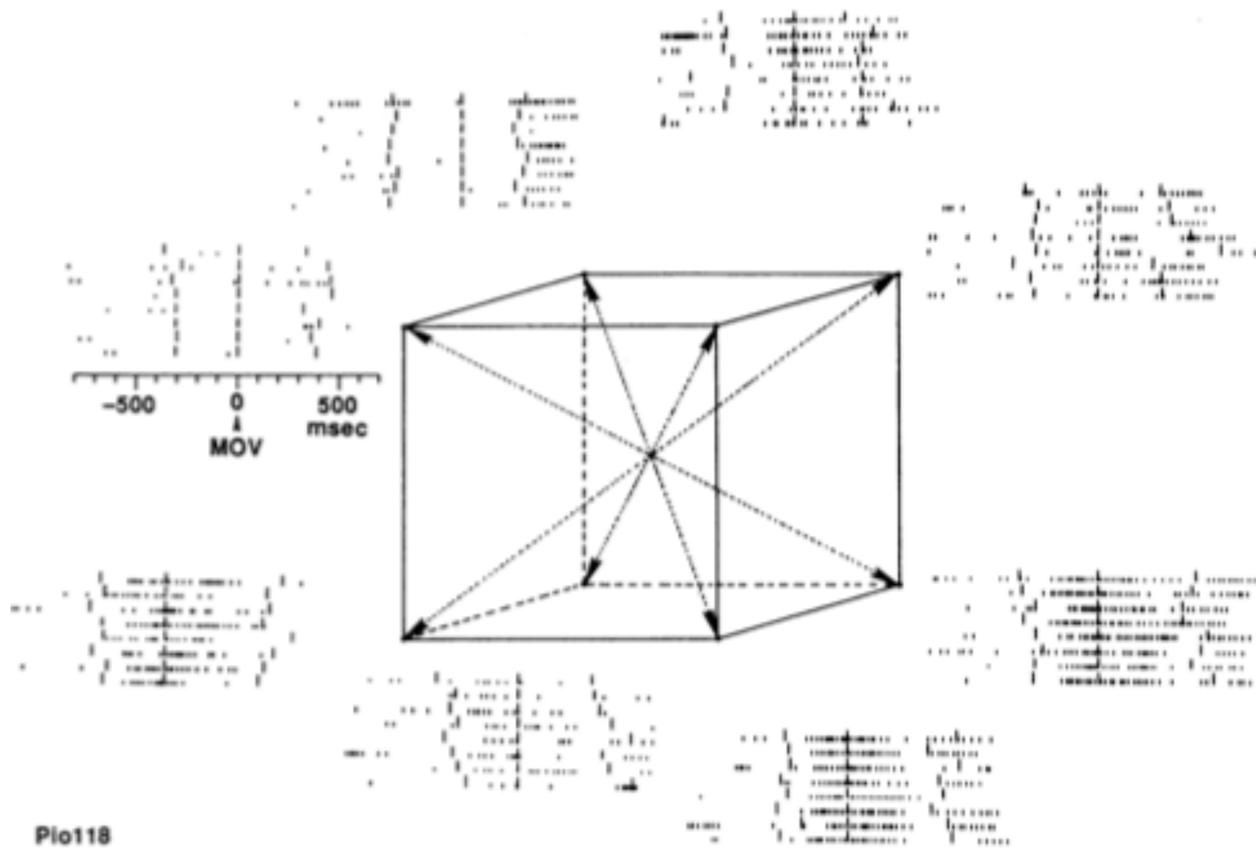
Decoding wind direction in the cricket cercal system with 4 interneurons (from Dayan and Abbott book)

Population coding example



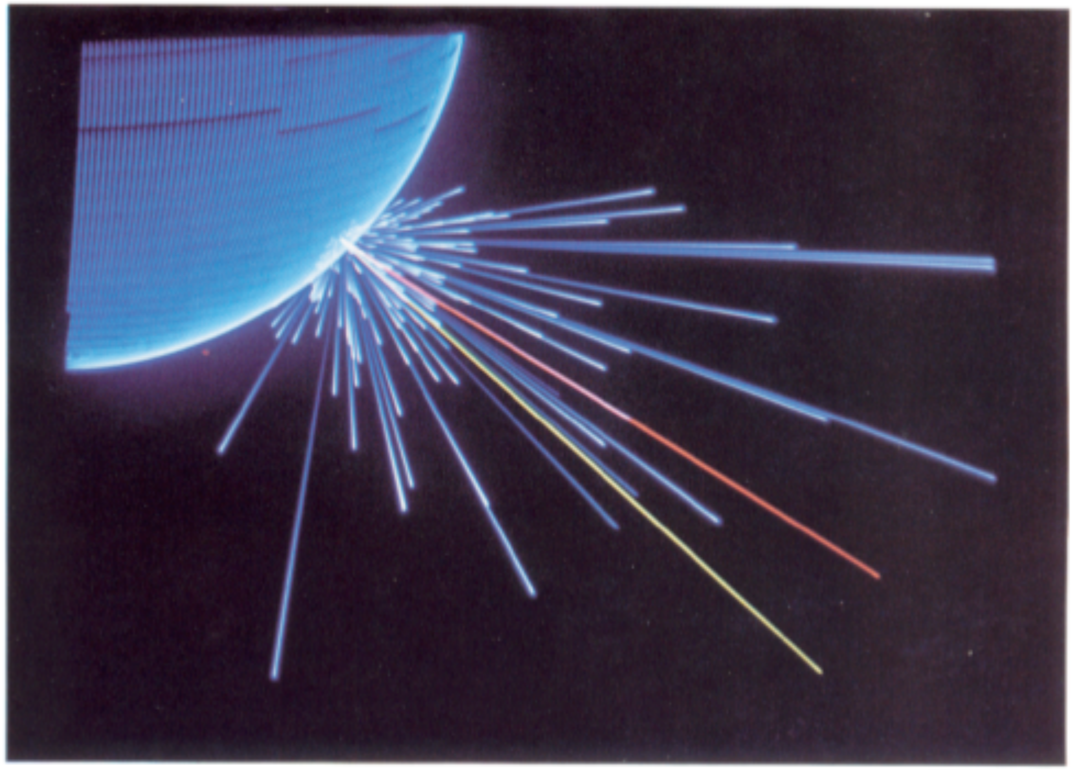
Decoding hand movement direction from primary motor cortex population (Georgopoulos et al. 1982)

Population coding example



Example neuron in primary motor cortex
(from Schwartz & Georgopoulos 1986)

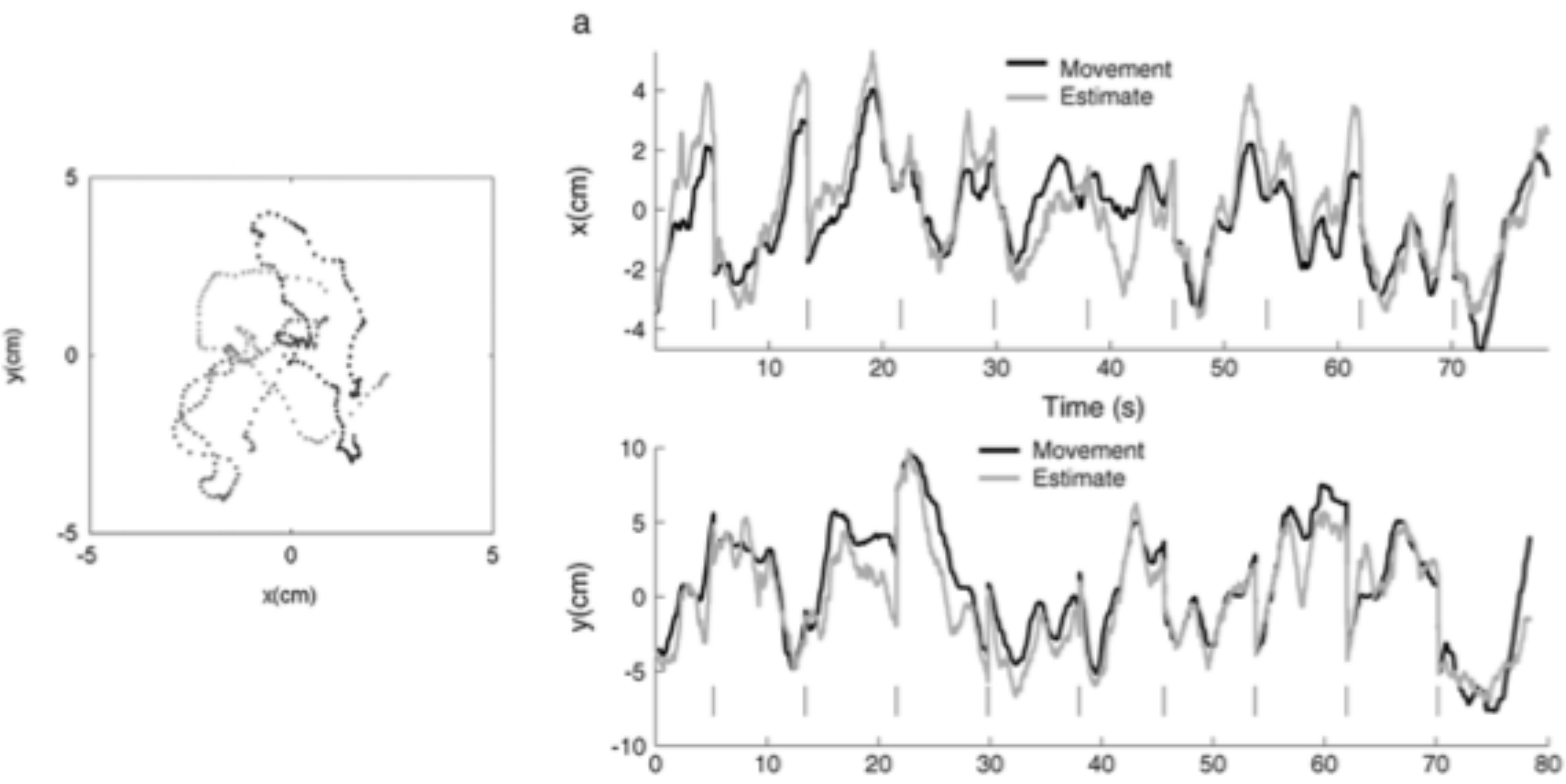
Population coding example



Decoding hand movement direction from primary motor cortex population (from Georgopoulos et al., 1988)

Population vector decoding

Population coding example



Decoding hand movement direction from primary motor cortex population of 17 neurons.
Shoham et al., 2005

Population coding

Let's look more generally at population decoding…

Population coding

Let's look more generally at population decoding…

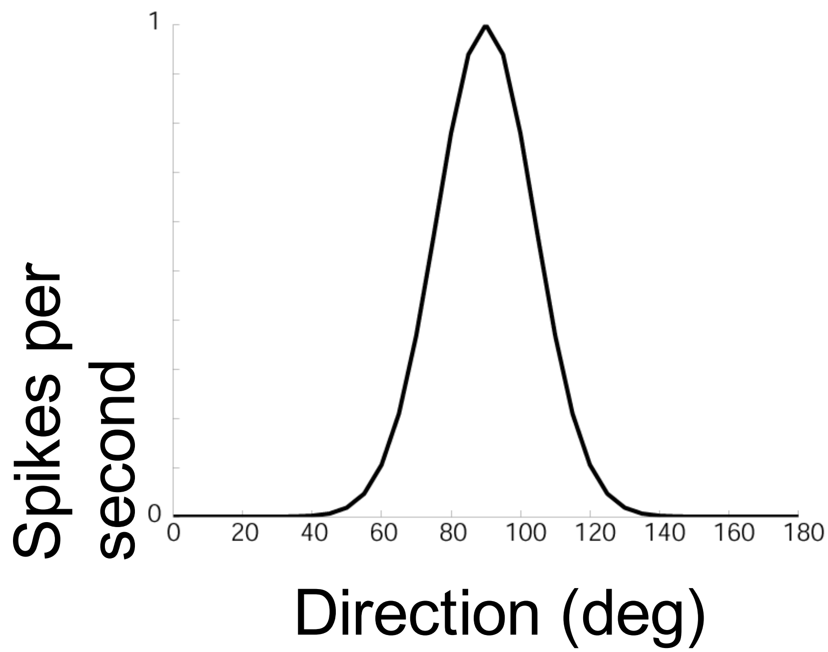
We have tuning curves, example:

Gaussian-like

Cosine-like

Population coding

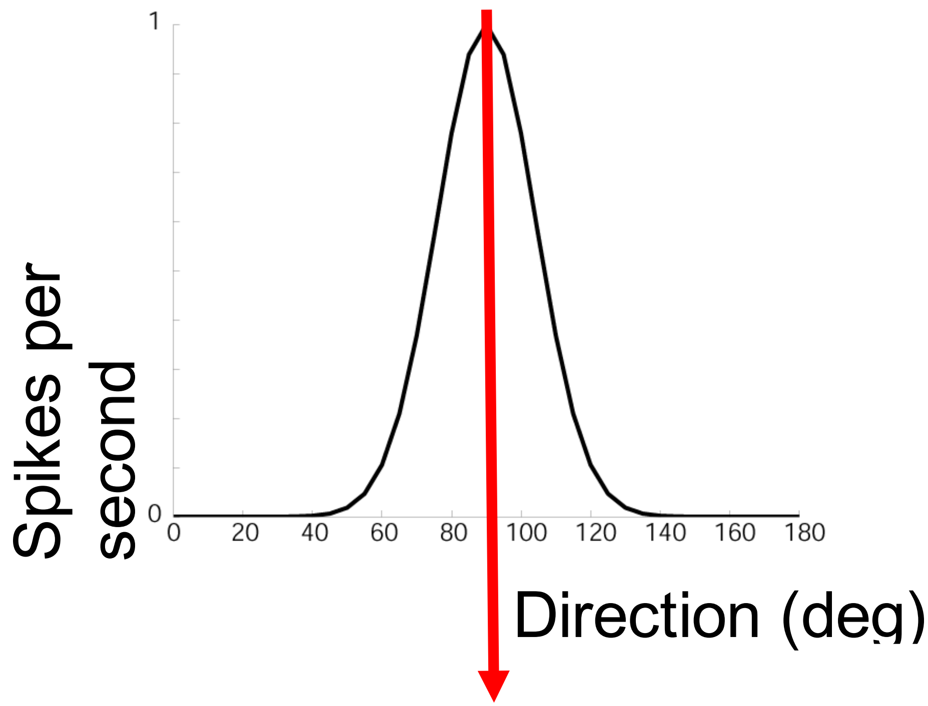
Example tuning curve for one neuron:



(this is an idealized depiction of a tuning curve)

Population coding

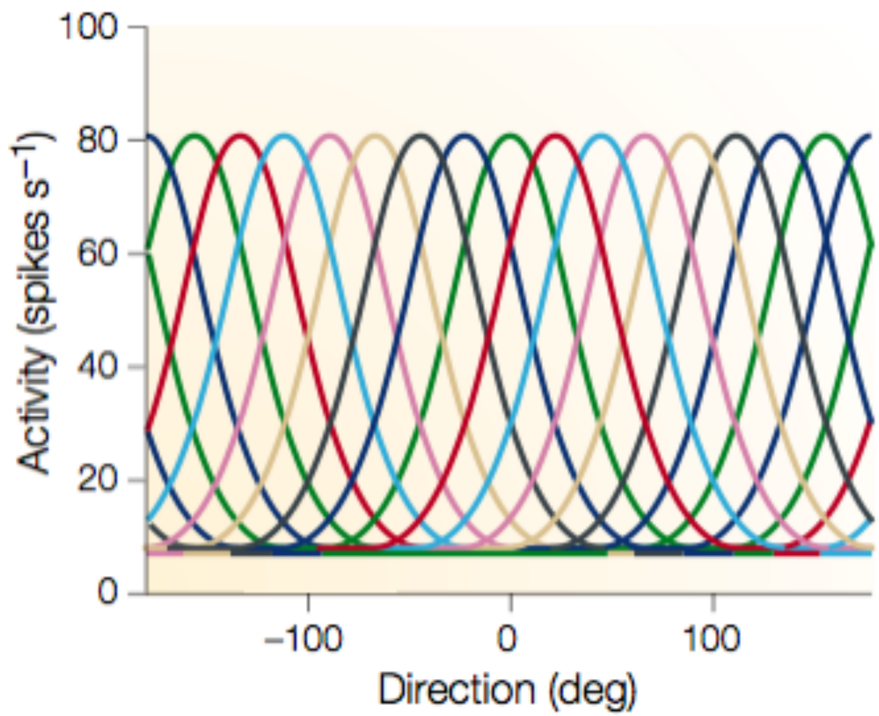
Example tuning curve for one neuron:



(this is an idealized depiction of a tuning curve)

Population coding

Tuning curve for population of neurons…

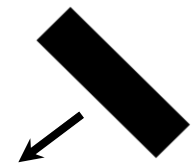


From Pouget, Dayan, Zemel, 2000

(again, idealized tuning curves)

Population coding

stimulus

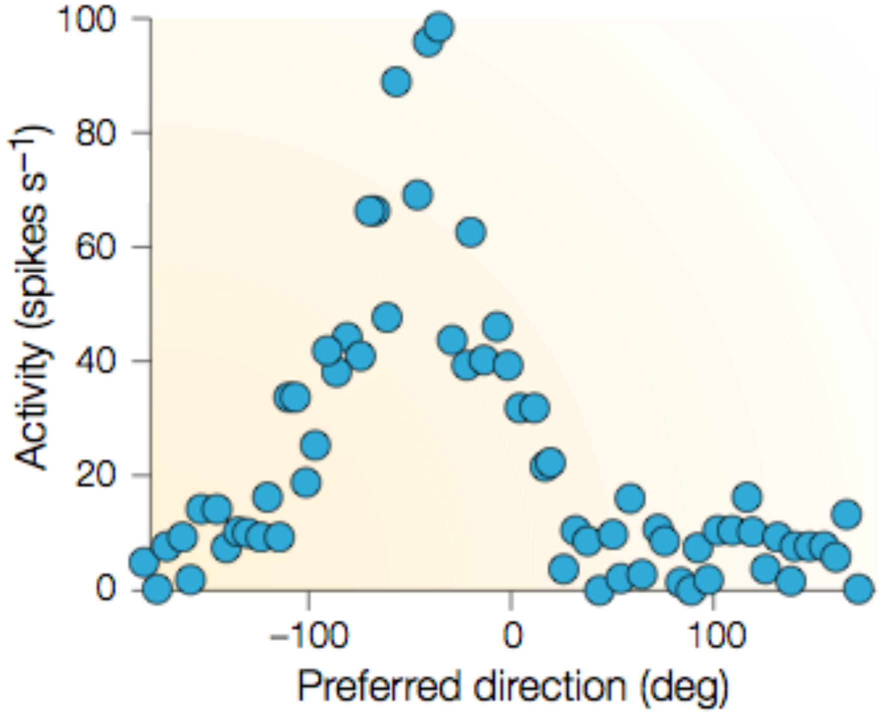


On the board…

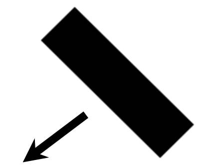
What is the population response to the stimulus?

From Pouget, Dayan, Zemel, 2000

Population coding



Stimulus S

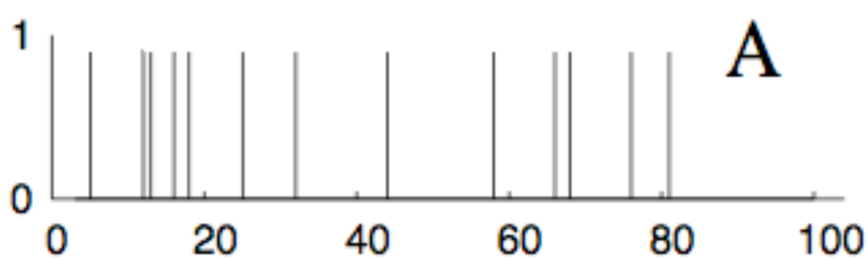


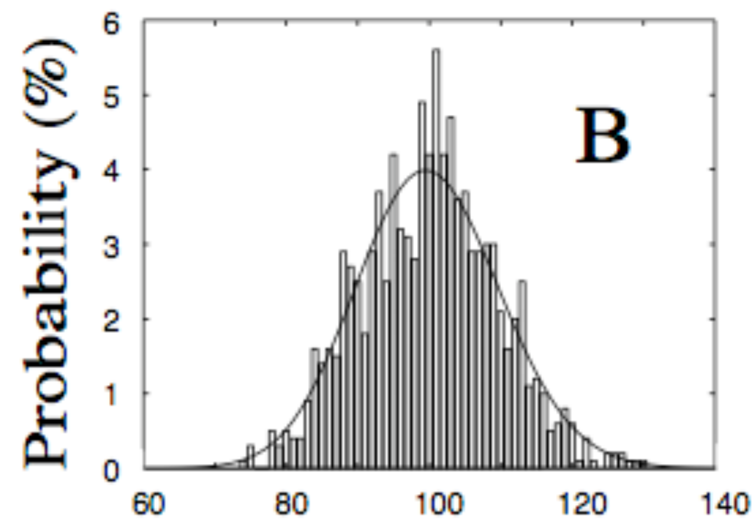
The activity is “noisy”… Why?

From Pouget, Dayan, Zemel, 2000

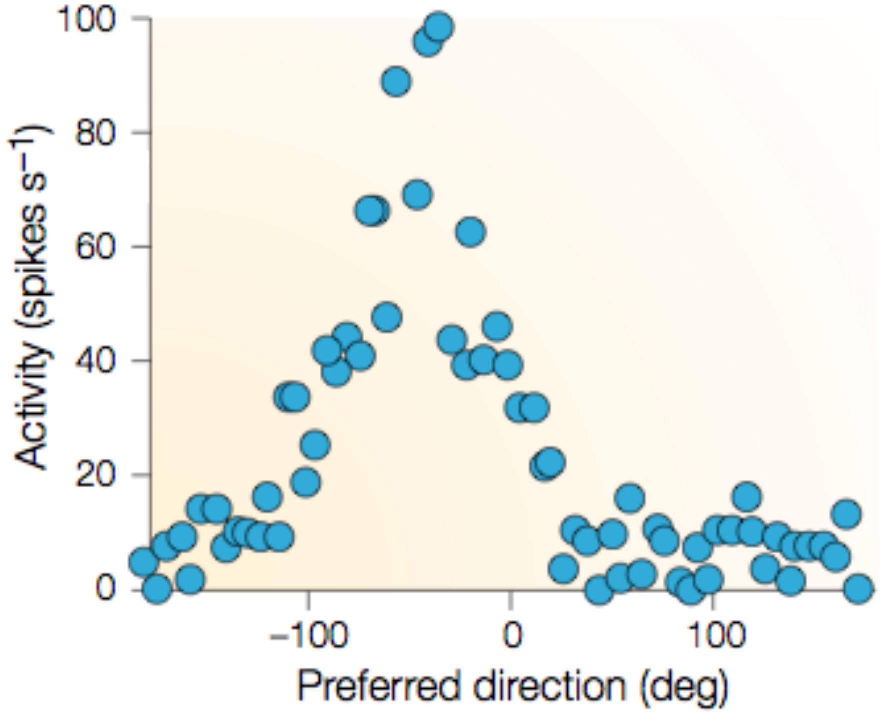
Poisson spike trains

Variability of neuronal spikes similar to a stochastic/random process,



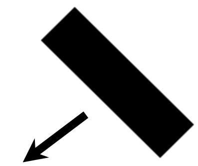


Population coding



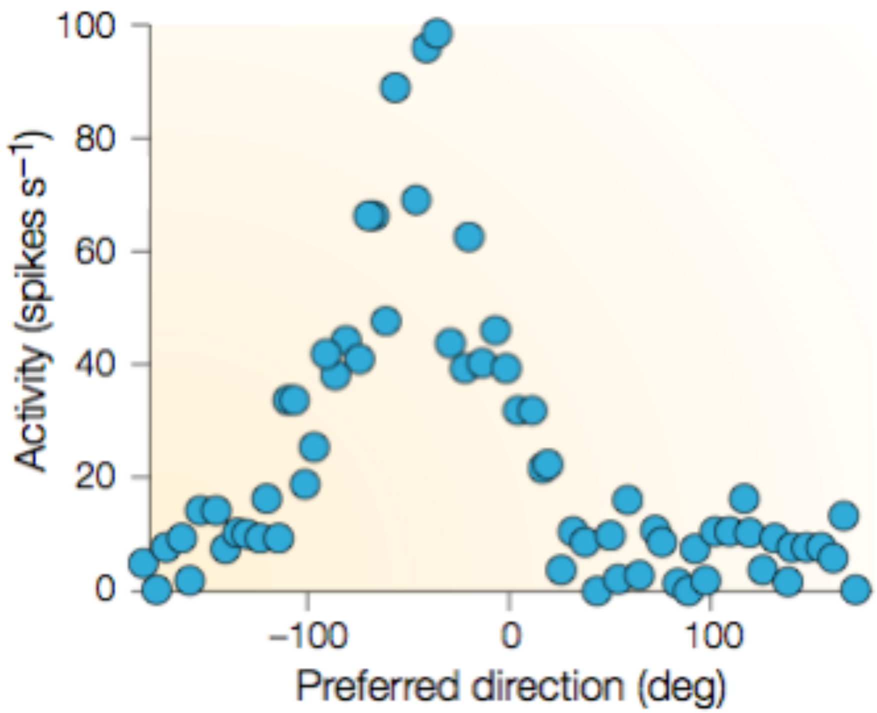
Noisy neural activity

Stimulus S



From Pouget, Dayan, Zemel, 2000

Population coding



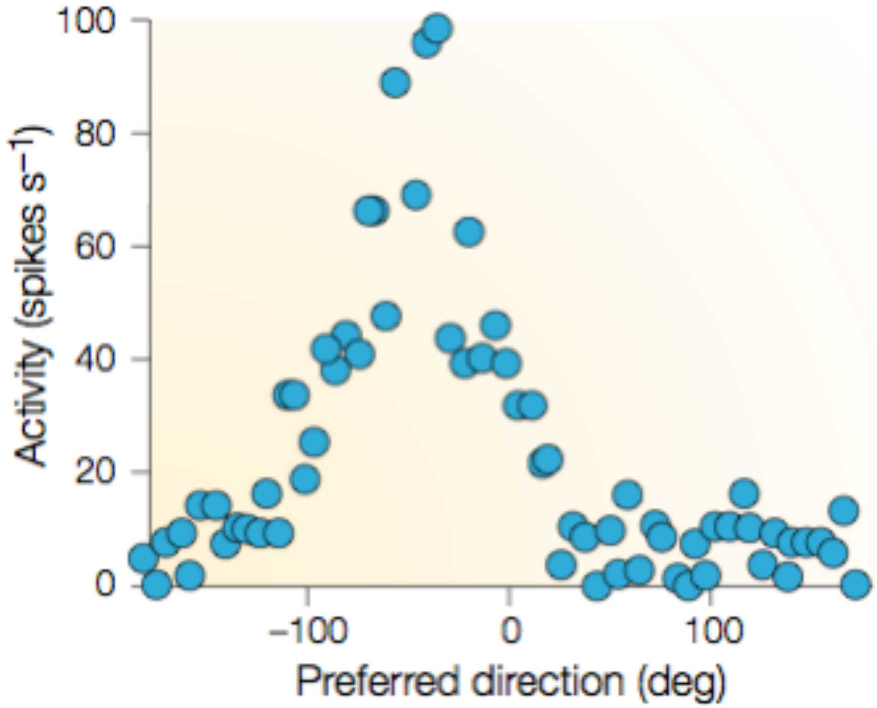
$\hat{S}$

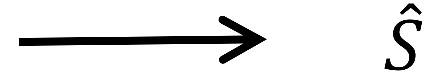
Decode
Readout
Estimate

Decoding: estimate signal (here direction of motion) given population activity

From Pouget, Dayan, Zemel, 2000

Population coding

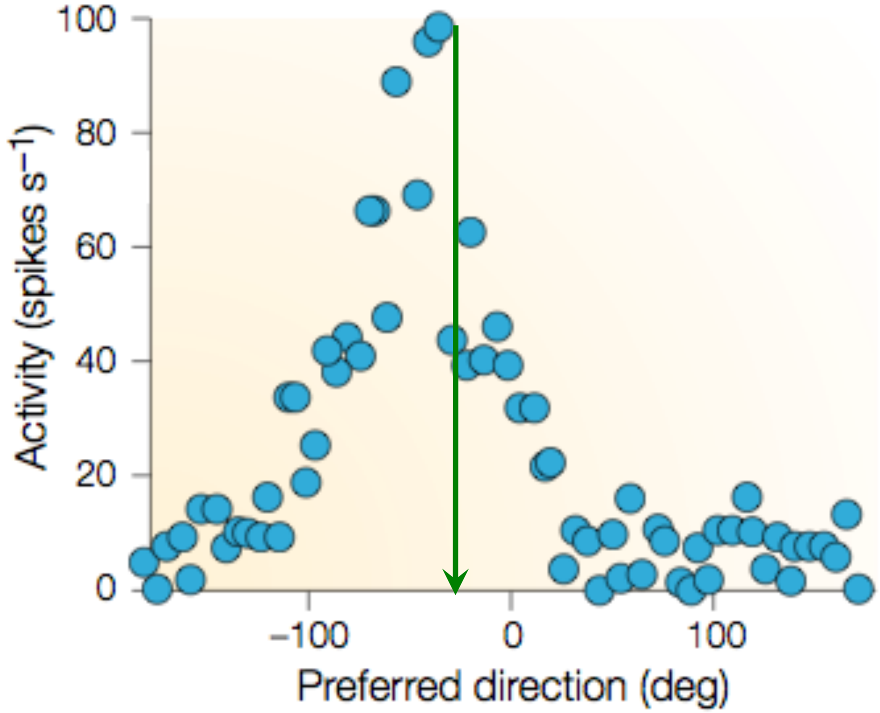




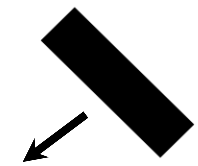
How should we decode??

From Pouget, Dayan, Zemel, 2000

Population coding



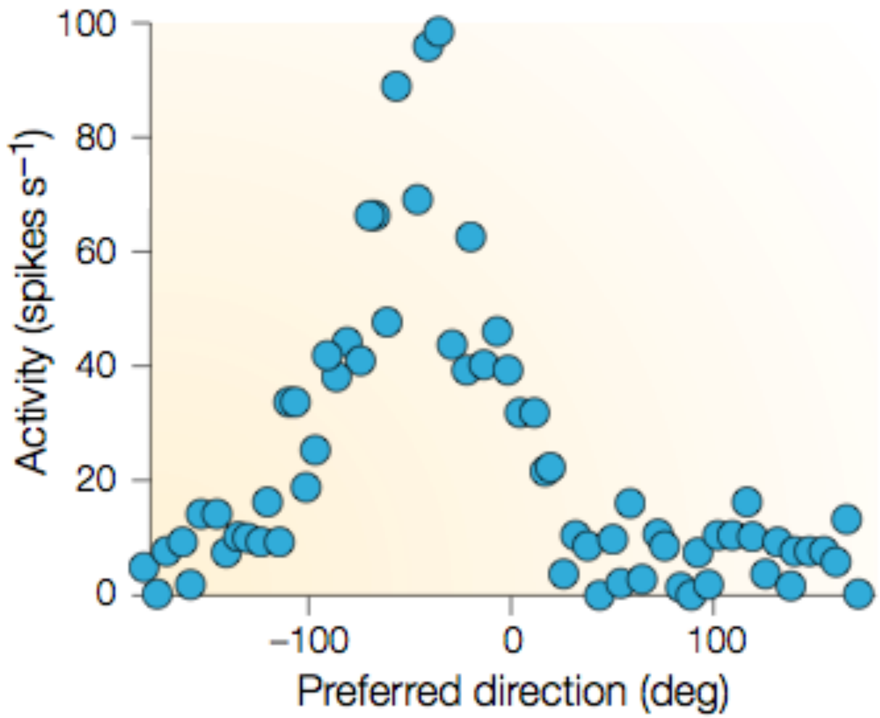
stimulus

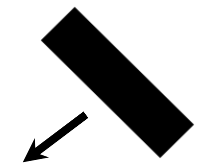


Decoding population activity: Center-of-mass (population vector)

From Pouget, Dayan, Zemel, 2000

Population coding

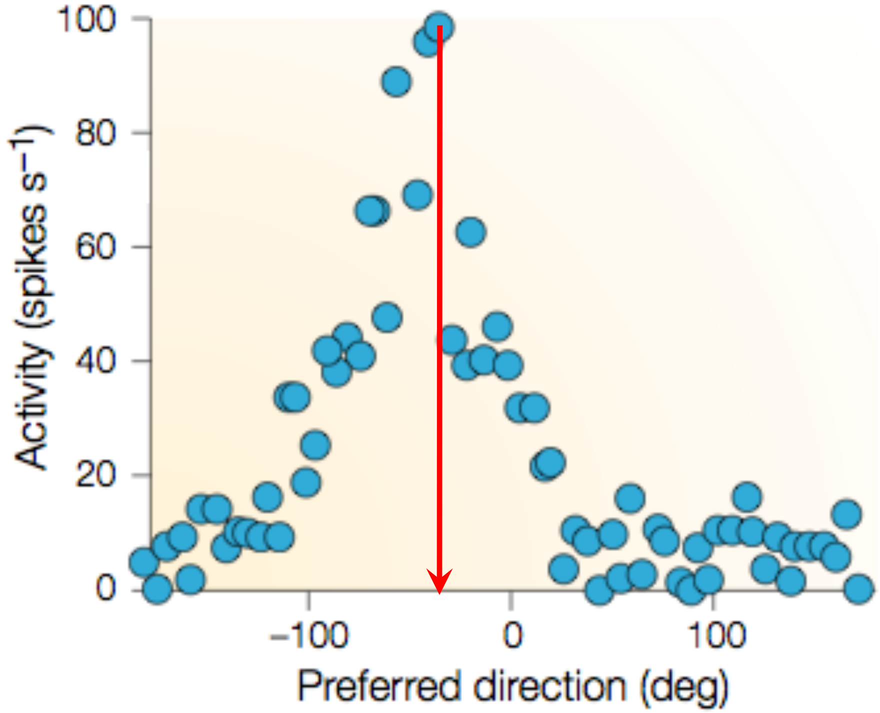


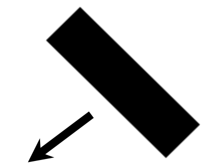


Other decoding schemes?

From Pouget, Dayan, Zemel, 2000

Population coding

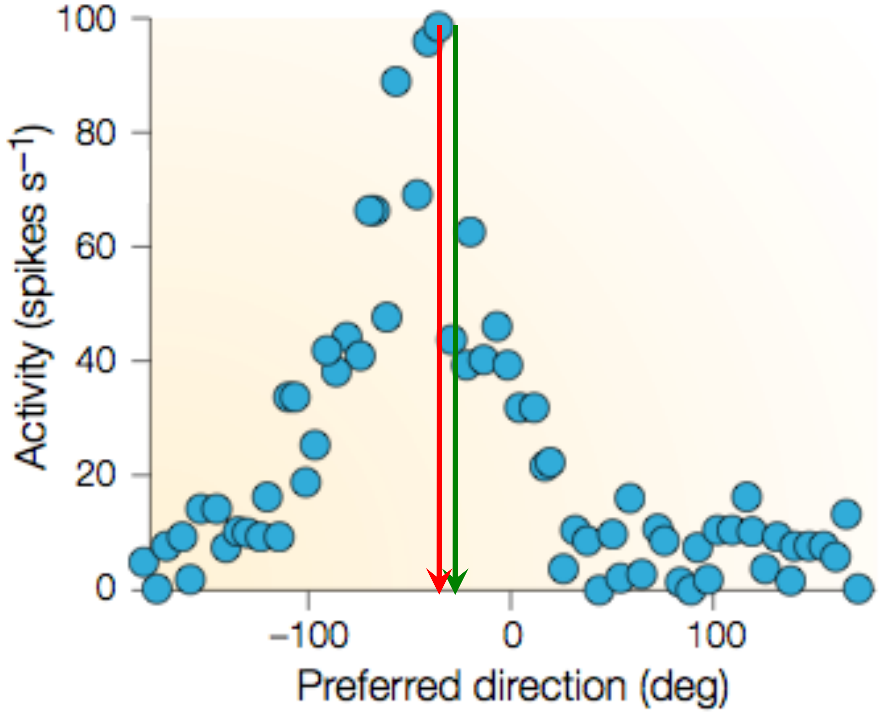


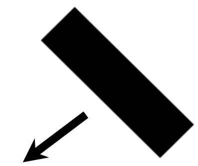


Decoding population activity: Maximum (winner-take-all)

From Pouget, Dayan, Zemel, 2000

Population coding





Decoding population activity: Different decoders can give different answers…

From Pouget, Dayan, Zemel, 2000

Population coding

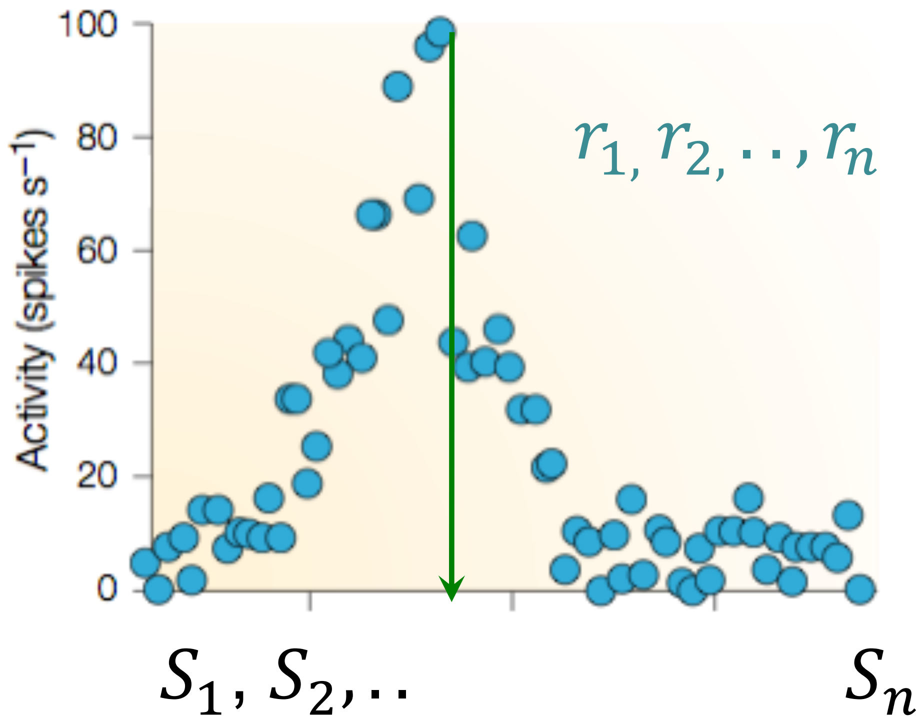
S	Stimulus we want to estimate
$r_1, r_2, .., r_n$	Firing rate activity of each neuron
$S_1, S_2, ..S_n$	Preferred stimulus each neuron

Population vector: each neuron "votes" for its preferred stimulus

$$\hat{S} = \sum_{i=1}^{n} r_i S_i$$

Has been useful for:
Cercal system
Motor cortex

Population coding

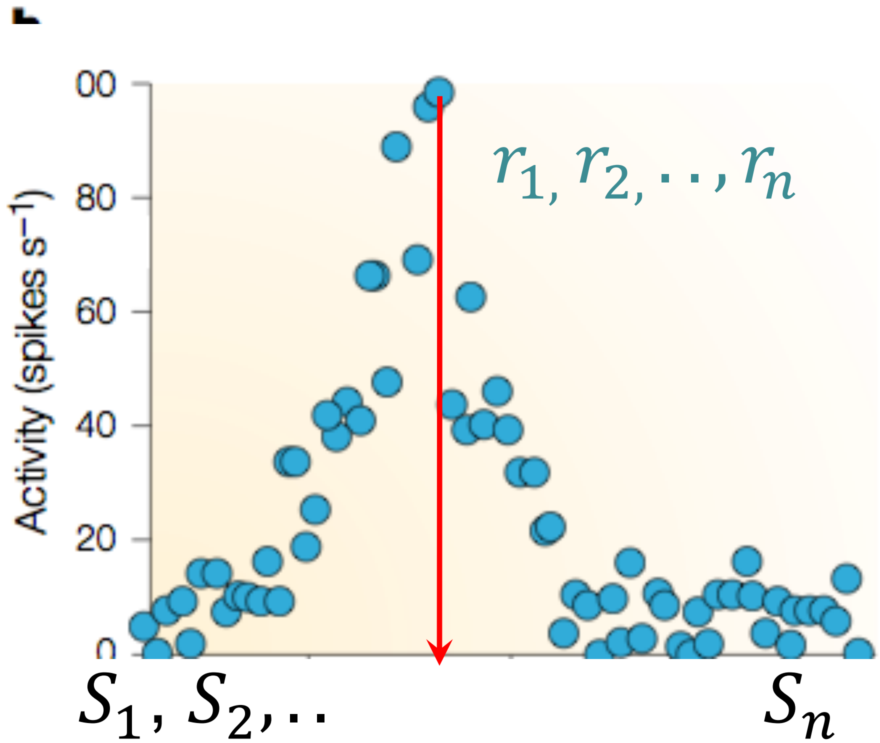


Population vector: each neuron “votes” for its preferred stimulus

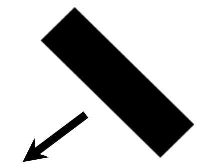
$$\hat{S} = \sum_{i=1}^{n} r_i S_i$$

Has been useful for:
Cercal system
Motor cortex

Population coding



stimulus



Winner take all: neuron with highest response “wins”

$$\hat{S} = S_j$$

j=argmax r_i

Based on Pouget, Dayan, Zemel, 2000

Population coding

Population and Winner take all properties:

- Simple!
- Does not take noise into account!
- Not necessarily optimal

Other methods?

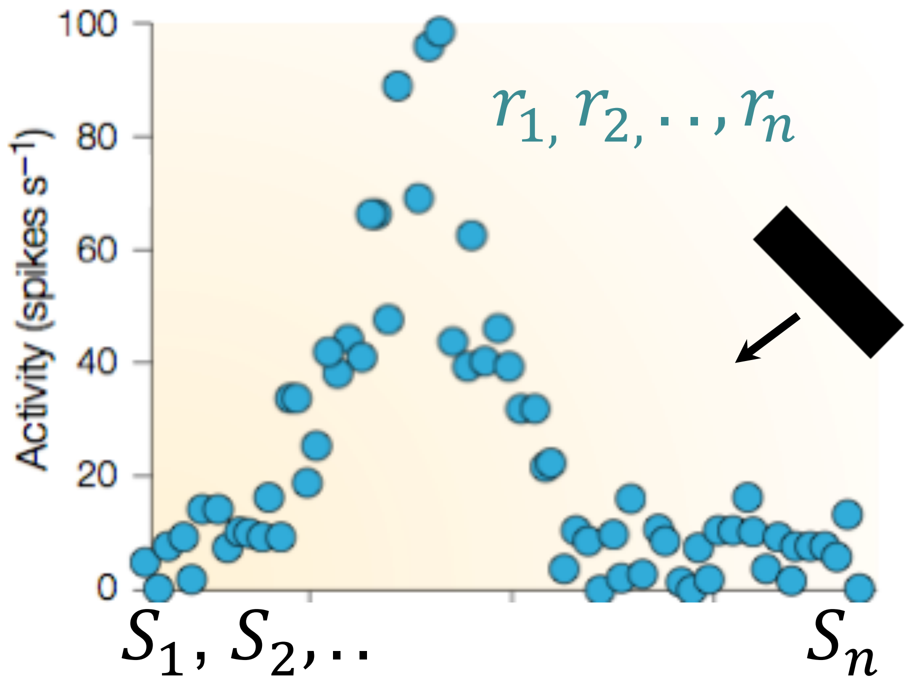
Population coding

S Stimulus we want to estimate

$r_{1,} r_{2,} .., r_n$ Firing rate activity of each neuron

Consider the distribution:

$$prob(r|S)$$



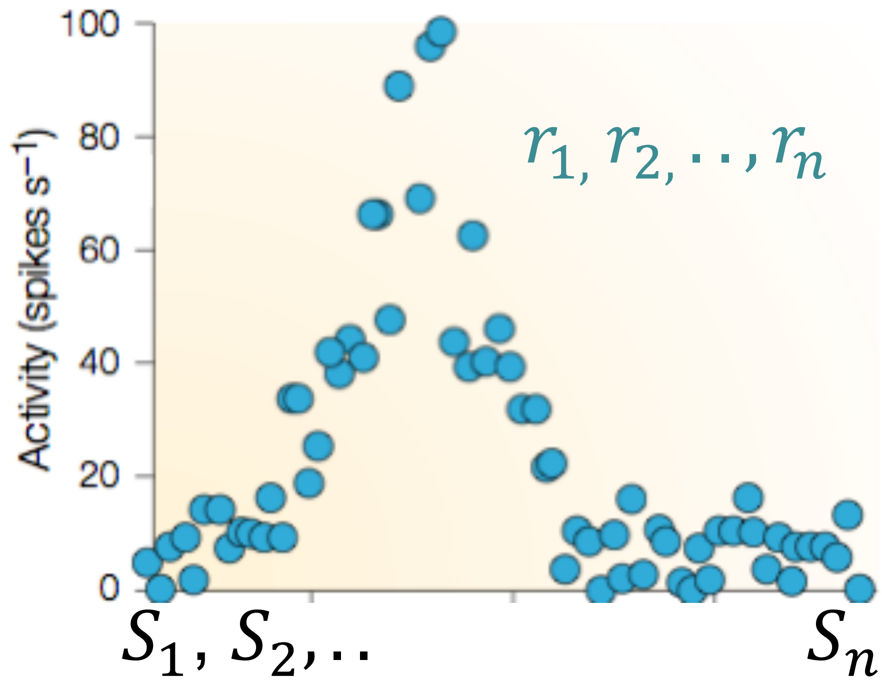
Population coding

S Stimulus we want to estimate

$r_{1,} r_{2,} .., r_n$ Firing rate activity of each neuron

Maximum likelihood:
Find S that maximizes

$$prob(r|S)$$



Population coding

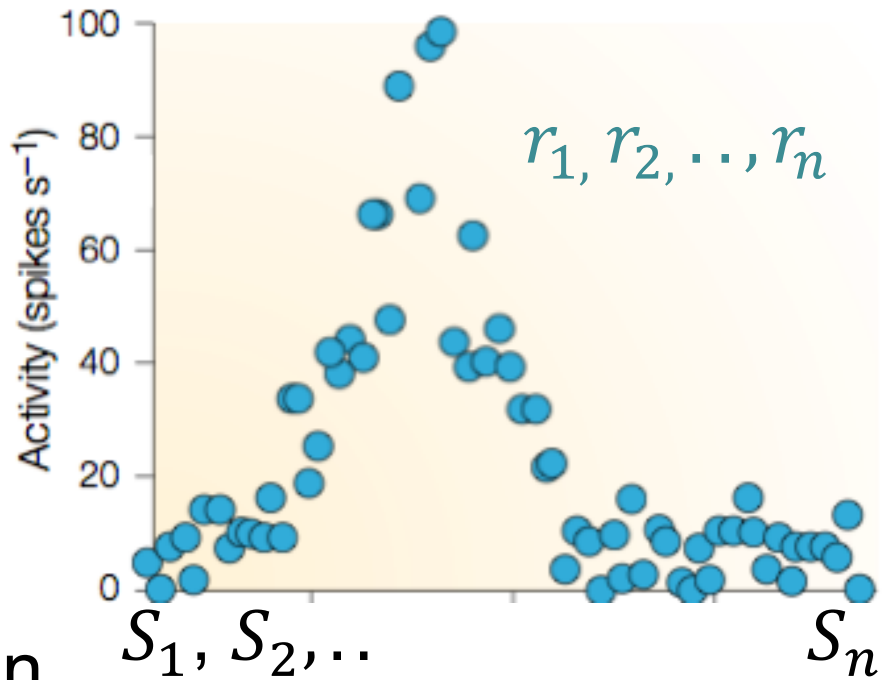
S Stimulus we want to estimate

$r_{1,} r_{2,} .., r_n$ Firing rate activity of each neuron

Maximum likelihood:
Find S that maximizes

$prob(r|S)$

We need to know or assume this distribution



Population coding

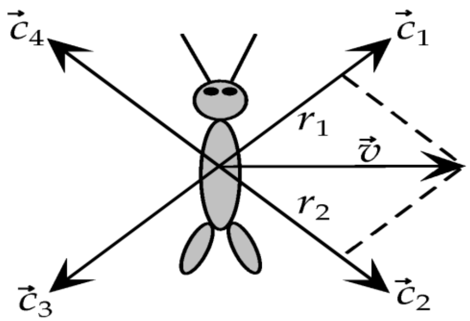
S Stimulus we want to estimate

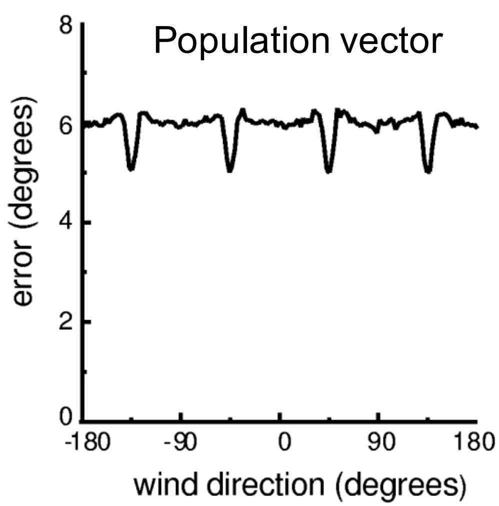
$r_{1,} r_{2,} .., r_n$ Firing rate activity of each neuron

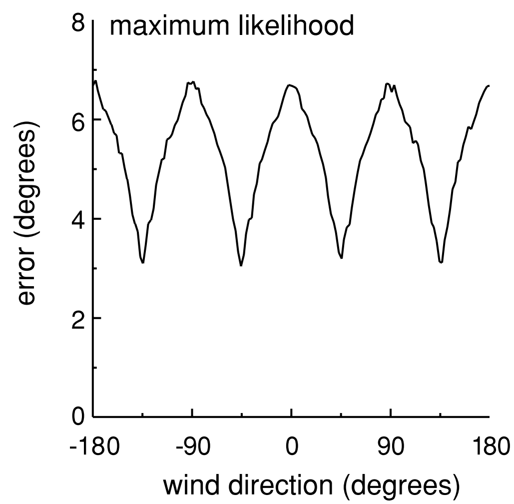
Maximum likelihood:
Find S that maximizes $prob(r|S)$

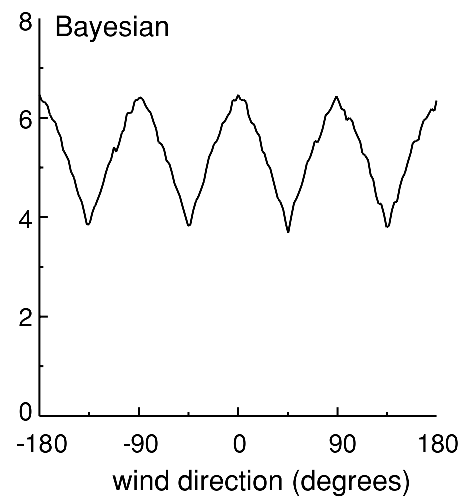
We can solve if we know noise distribution (eg, Poisson) and assume neurons independent (probabilities multiply); Set derivative to 0… Turns out similar to Population vector (see Dayan and Abbott book)

Population coding example









Dayan and Abbott book

Population coding (distributed)

How do we judge quality of a decoder??

Population coding (distributed)

How do we judge quality of a decoder?

If we estimate signal (e.g., direction of motion) many times given population response, desirable properties:

• unbiased estimate of stimulus (on average gets the right, e.g., direction)

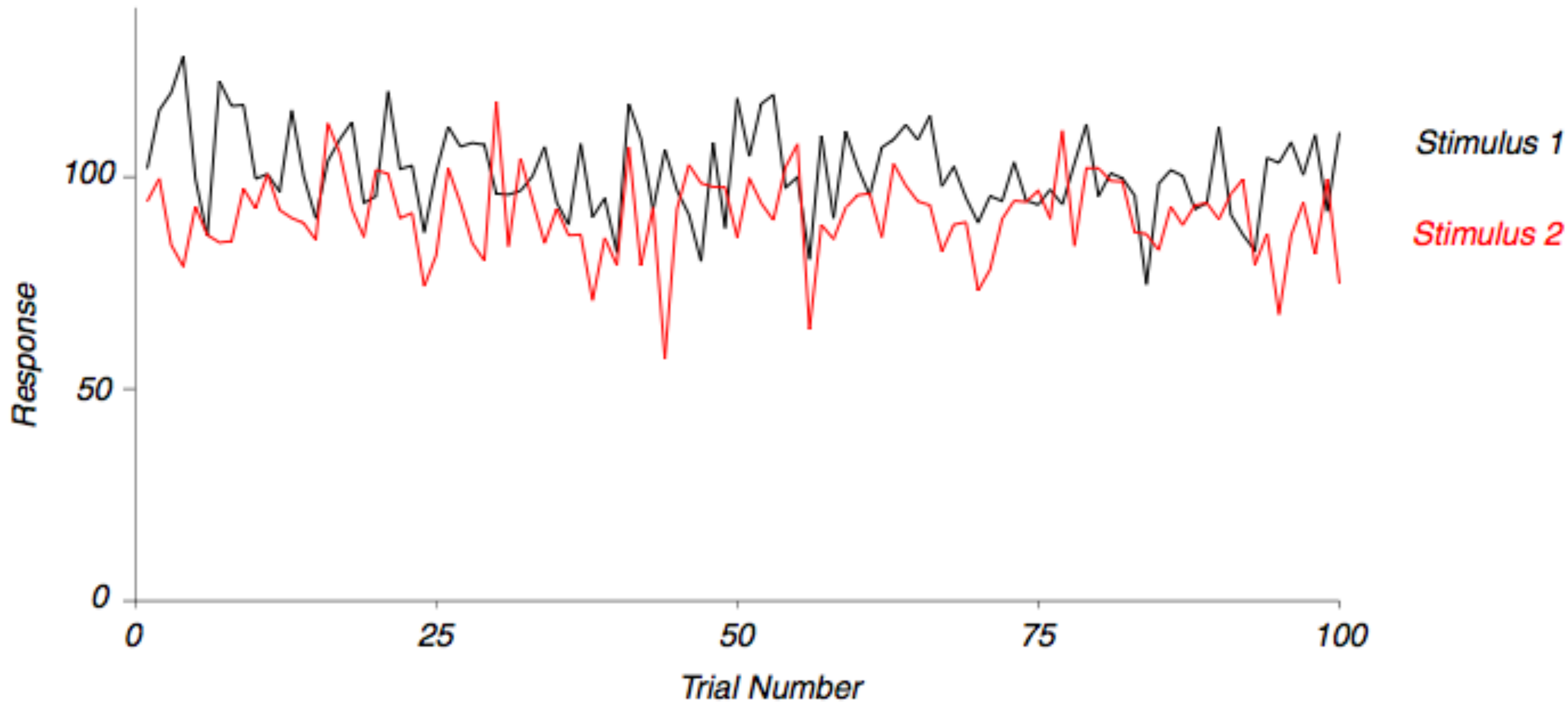
• low variance

• (theoretical work on bounds)

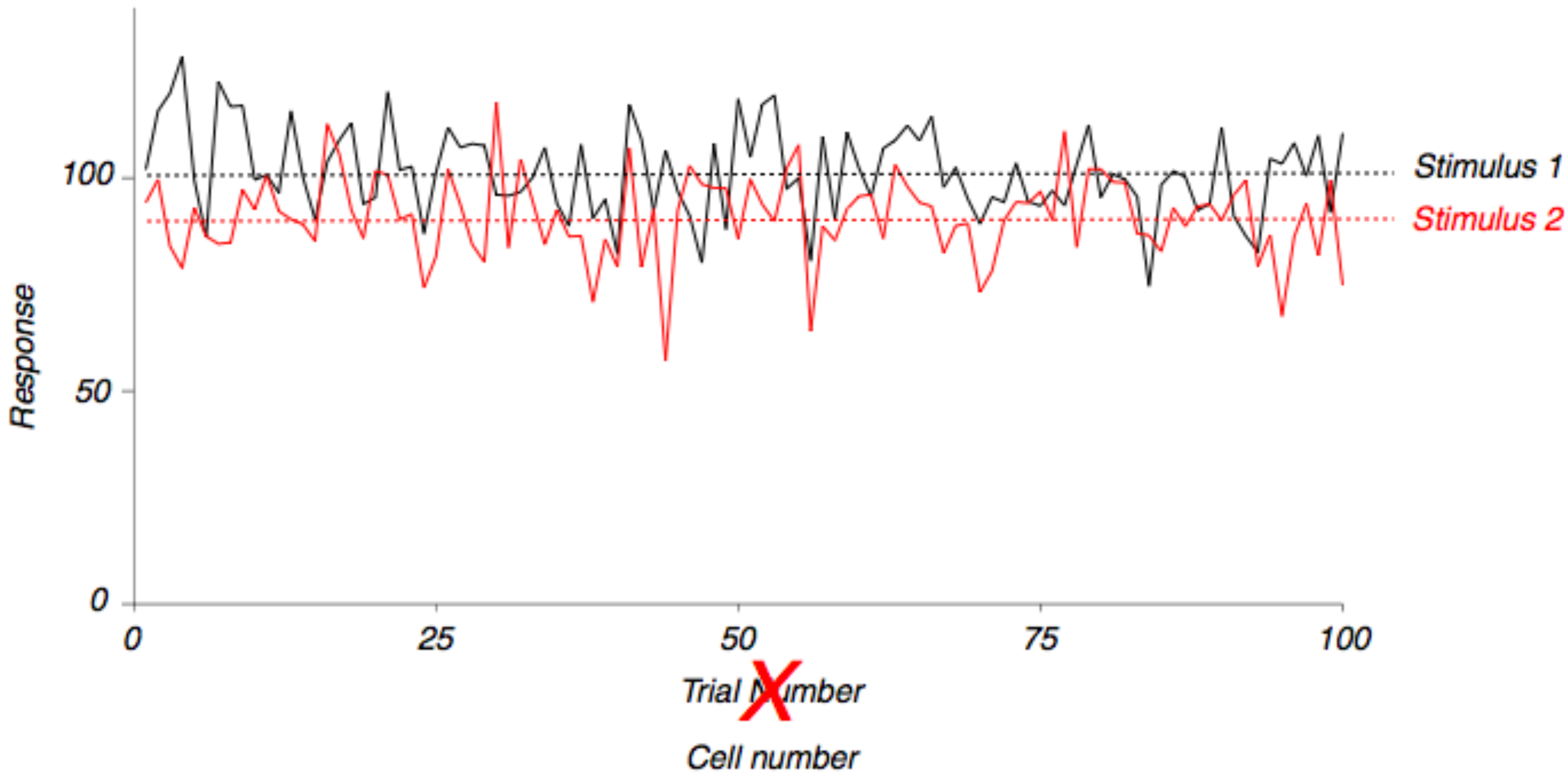
Correlated variability

- Revisiting noise in population codes!

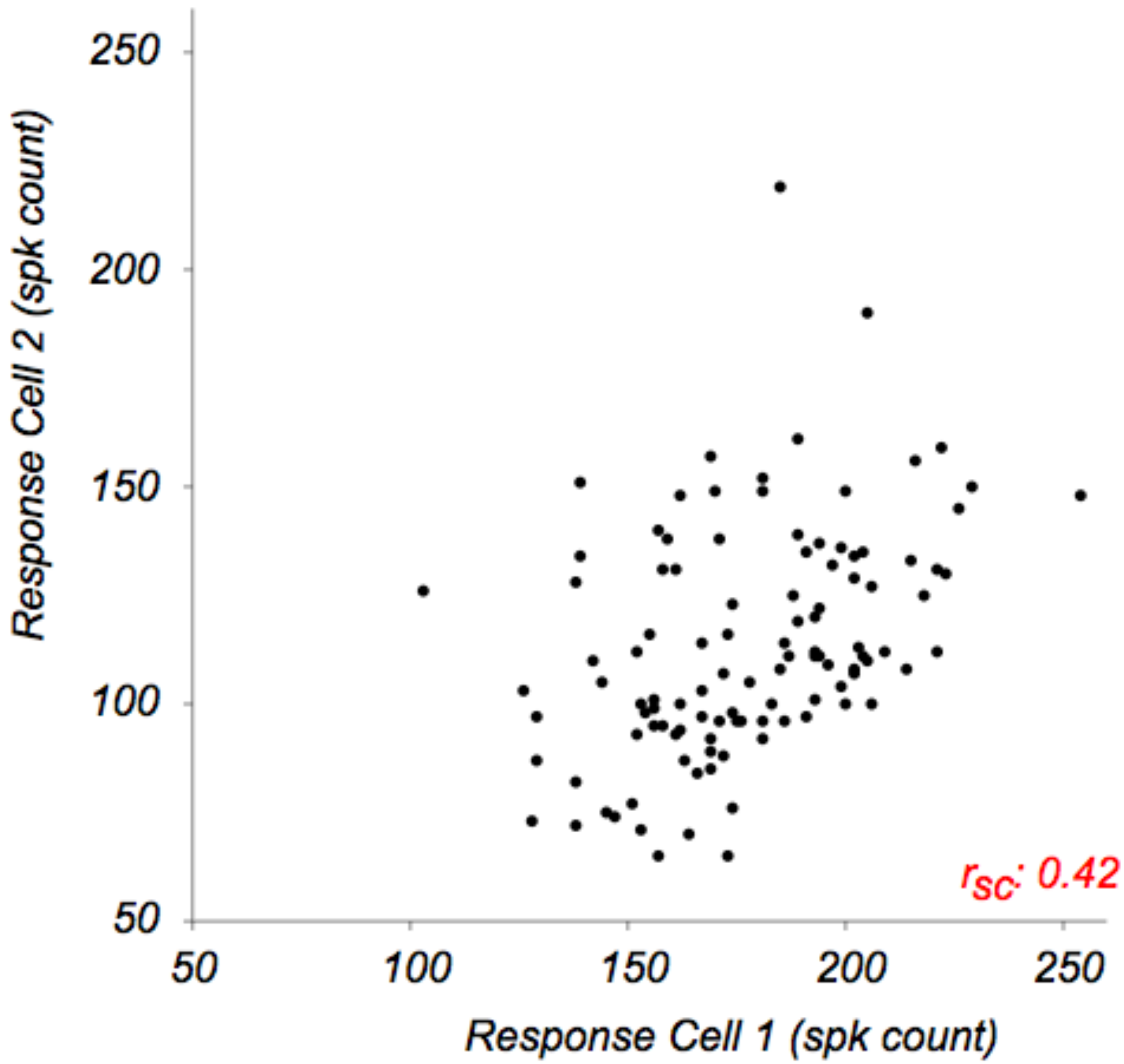
How to overcome noise?



How to overcome noise?



Response variability is correlated between cells



Correlated variability

- Revisiting noise in population codes!
- Noise and correlations affect how we read out neural populations (eg, independence assumption)
- Active area of research

Does pooling neuron outputs average noise and improve performance?

Thought experiment:

- Noisy machine that generates a number
- We want to estimate the number
- Unfortunately, it is corrupted by noise
- Generate number 10000 times and take average
- Lazy/biased machine: first sample independent of second; other 9998 same as second
- Assume independence and take average, when actually all 9998 numbers are not independent
- Good estimate?

from Averbeck et al. 2006

Thought experiment:

- Noisy machine that generates a number
- We want to estimate the number
- Unfortunately, it is corrupted by noise
- Generate number 10000 times and take average
- Lazy/biased machine: first sample independent of second; other 9998 same as second
- Assume independence and take average, when actually all 9998 numbers are not independent
- Good estimate?

from Averbeck et al. 2006

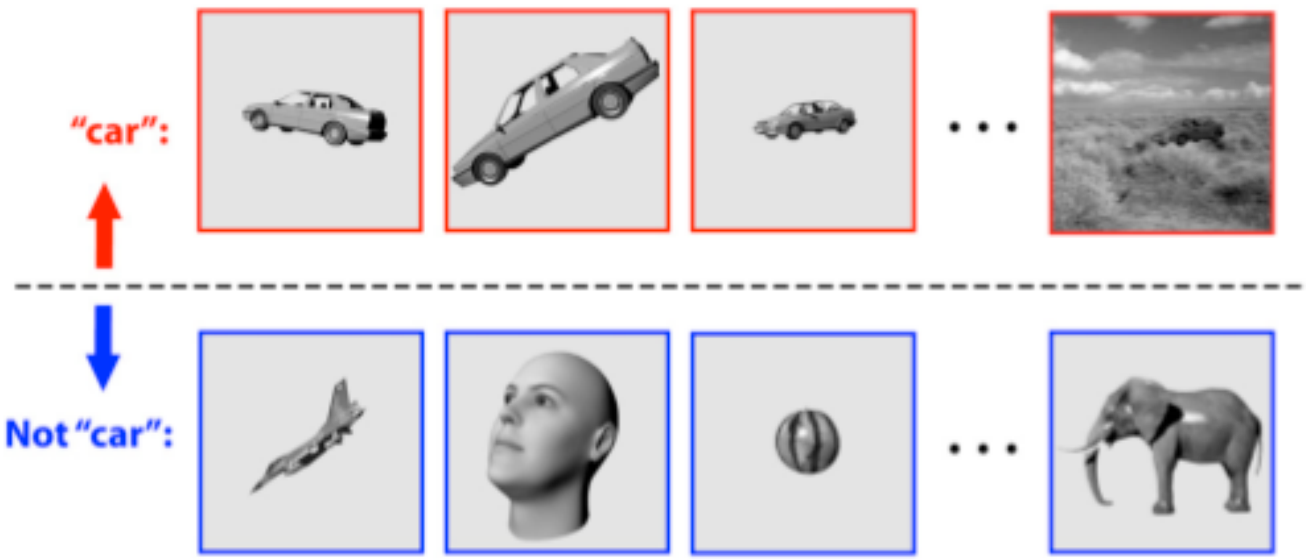
Optimal strategy:
Weight first sample by half
and all the rest by half

“Neurons face the same situation: they compute some function of the variables encoded in their inputs, and to perform this computation optimally they must know the correlations in the ~10,000 inputs that they receive. If they ignore the correlations, they may — or may not — pay a price in the form of suboptimal computations.”

from Averbeck et al. 2006

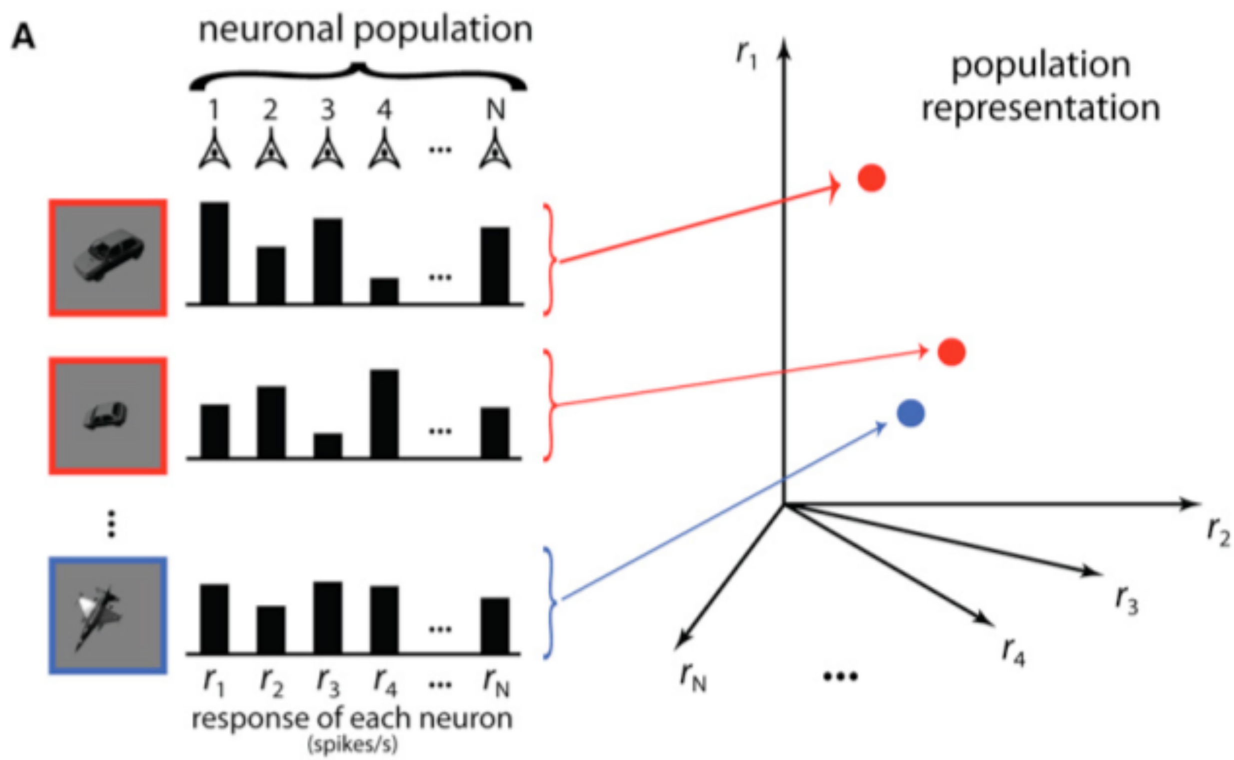
Optimal strategy:
Weight first sample by half
and all the rest by half

Other computations: discrimination (population)



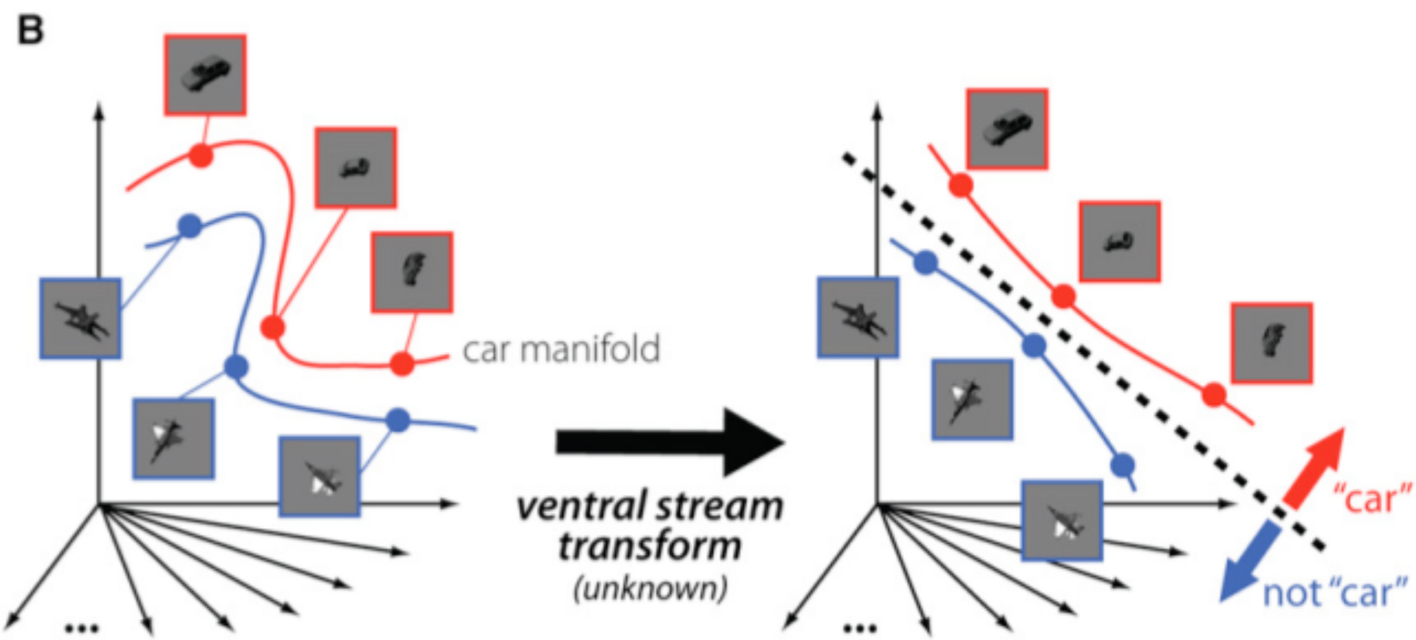
DiCarlo, Zocollan, Rust, 2012

Other computations: discrimination (population)



DiCarlo, Zocollan, Rust, 2012

Other computations: discrimination (population)



DiCarlo, Zocollan, Rust, 2012