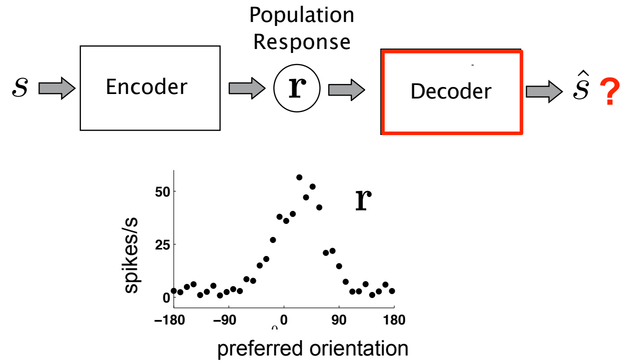
2. Decoding (continued)

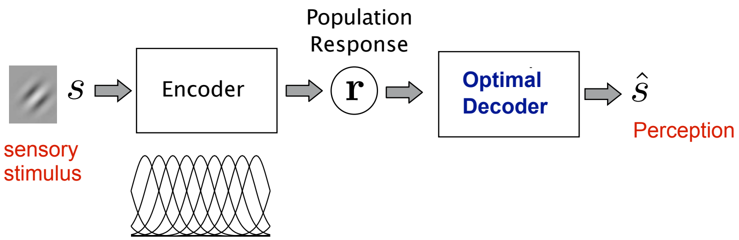
Readings: Decoding D&A ch.3

Decoding populations of neurons

In response to a stimulus with unknown orientation s, we observe a pattern of activity **r** (e.g. in V1). What can we say about s given **r**?



1. Optimal Decoding



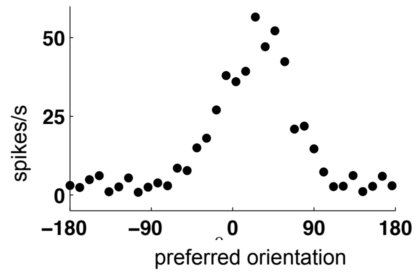
- optimality criterion?

$$MSE(s) = <(\hat{s} - s)^2>$$

1. Optimal Decoding (supposing we know P[r|s])

- Maximum Likelihood:
  if we know P[**r**|s] (the encoding model),
  choose the stimulus s that has maximal probability of having generated the observed response, **r**.

$$\hat{s} = \mathrm{argmax}_s P(\mathbf{r}|s)$$

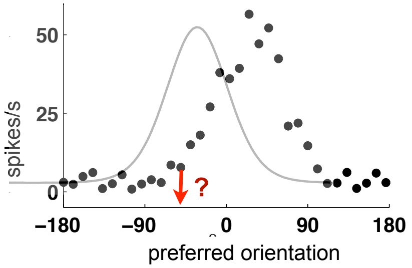


1. Optimal Decoding (supposing we know P[r|s])

- Maximum Likelihood:

  if we know P[**r**|s] (the encoding model),
  choose the stimulus s that has maximal probability of having generated the observed response, **r**.

$$\hat{s} = \mathrm{argmax}_s P(\mathbf{r}|s)$$

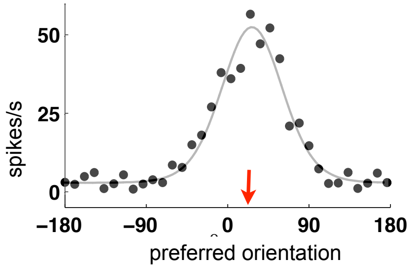


1. Optimal Decoding (supposing we know P[r|s])

- Maximum Likelihood:

  if we know P[**r**|s] (the encoding model),
  choose the stimulus s that has maximal probability of having generated the observed response, **r**.

$$\hat{s} = \mathrm{argmax}_s P(\mathbf{r}|s)$$

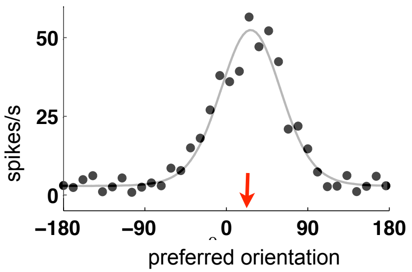


1. Optimal Decoding (supposing we know P[r|s])

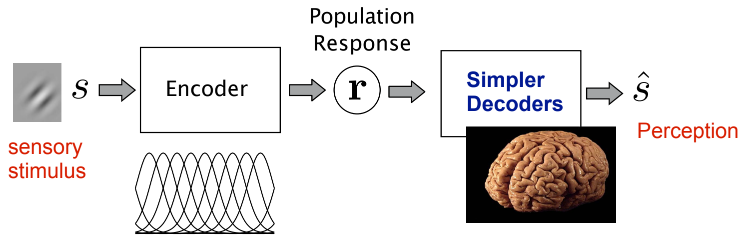
- Maximum a Posteriori:

  if we know P[**r**|s] and have a prior on s, P[s],
  choose the stimulus s that is most likely, given **r**.

$$\hat{s} = \mathrm{argmax}_s P(s|\mathbf{r}) = \mathrm{argmax}_s P[\mathbf{r}|s]P[s]$$



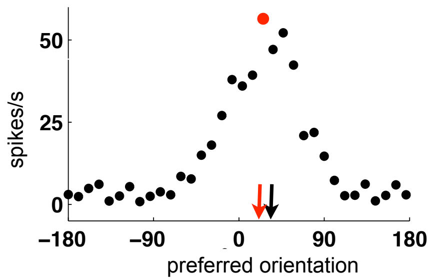
- It might be that, as experimentalist, we don't know have access to full knowledge of P[r|s].
- The brain itself might not be able to perform such complex computations when mapping information from one stage to another.
=> we need to consider simpler decoding strategies.



2. Simpler Decoding Strategies

Winner Take All :

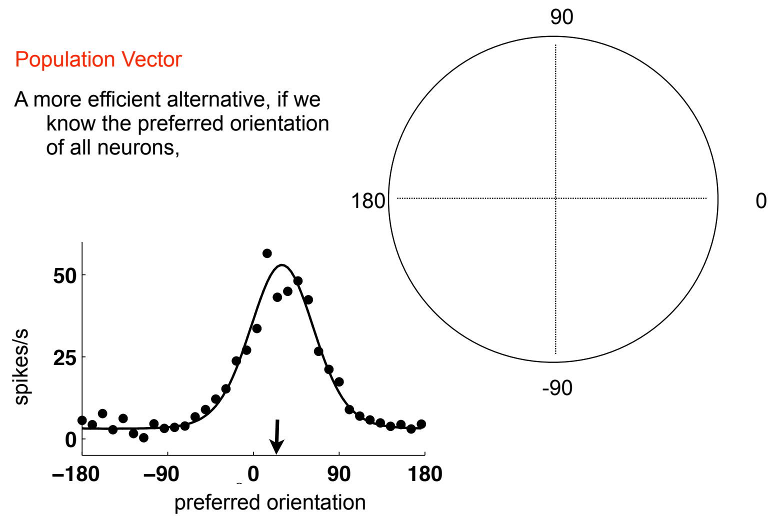
If we know the preferred orientation of all neurons,
choose the preferred orientation of the neuron that responds most.



2. Simpler Decoding Strategies

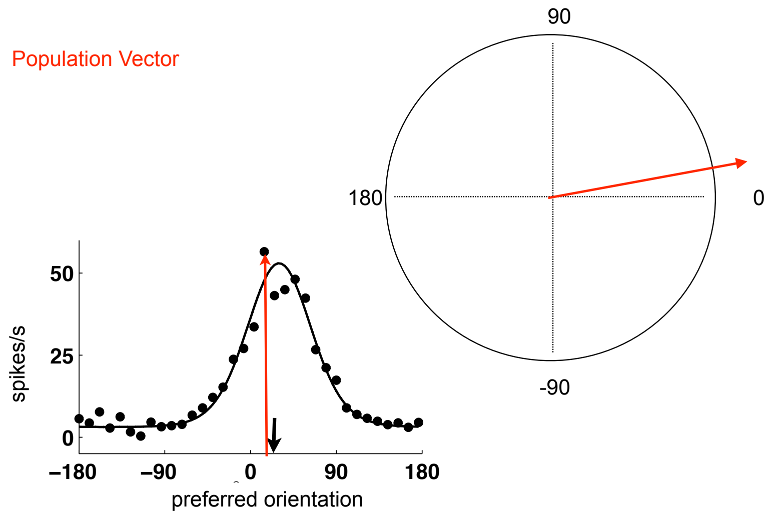
Population Vector

A more efficient alternative, if we
know the preferred orientation
of all neurons,



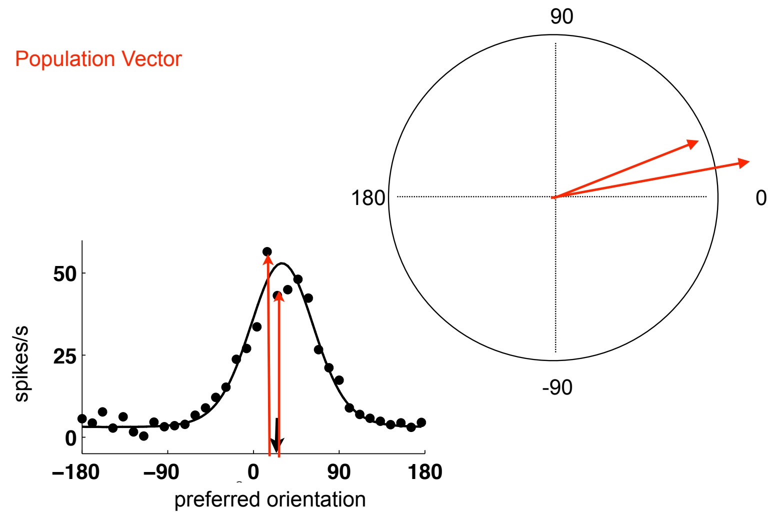
2. Simpler Decoding Strategies

Population Vector



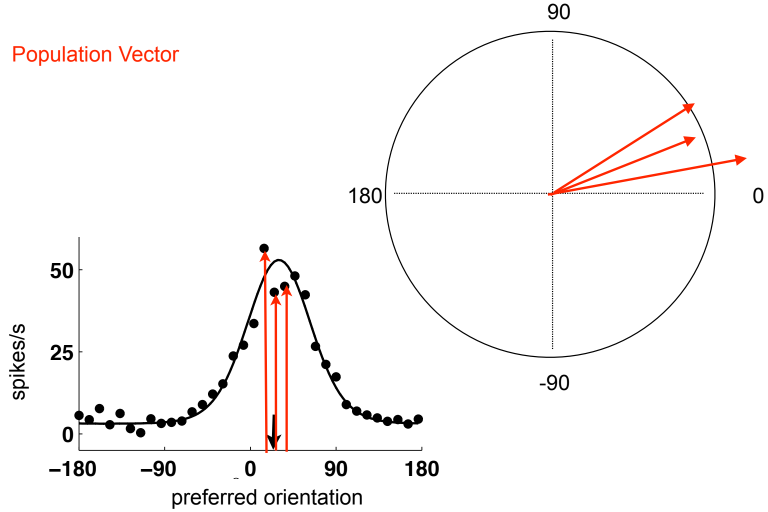
2. Simpler Decoding Strategies

Population Vector



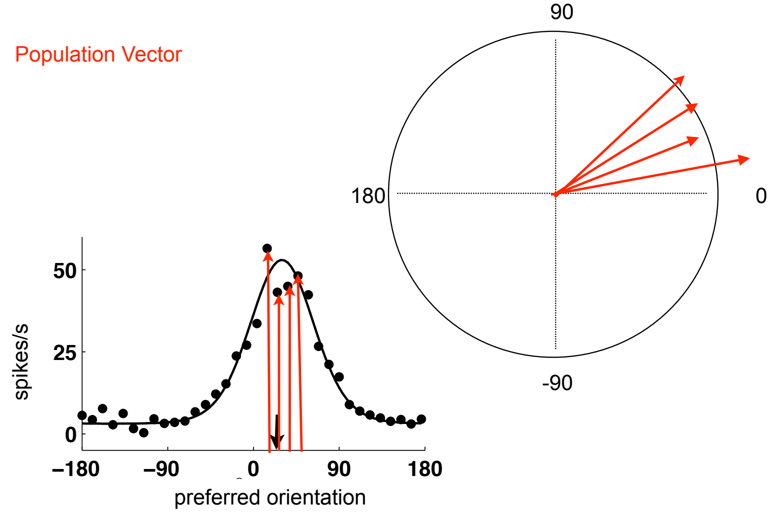
2. Simpler Decoding Strategies

Population Vector



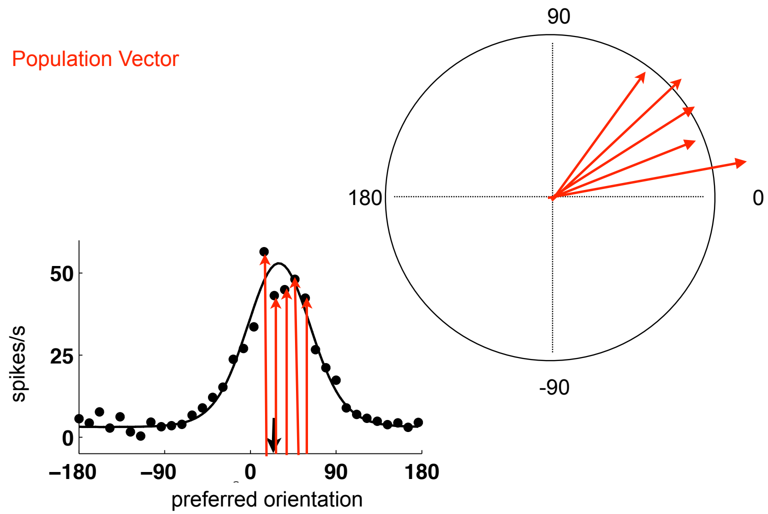
2. Simpler Decoding Strategies

Population Vector



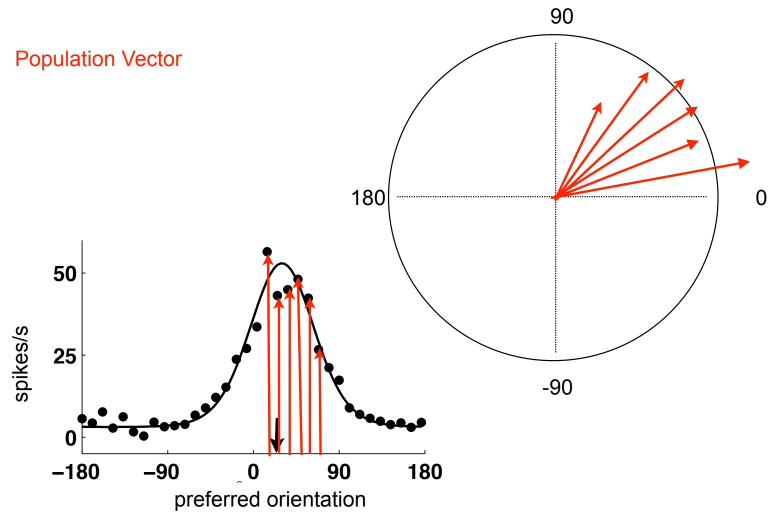
2. Simpler Decoding Strategies

Population Vector



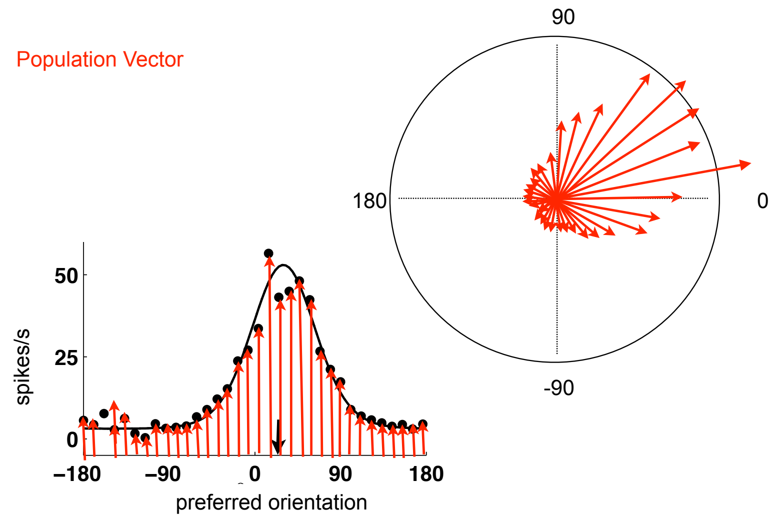
2. Simpler Decoding Strategies

Population Vector



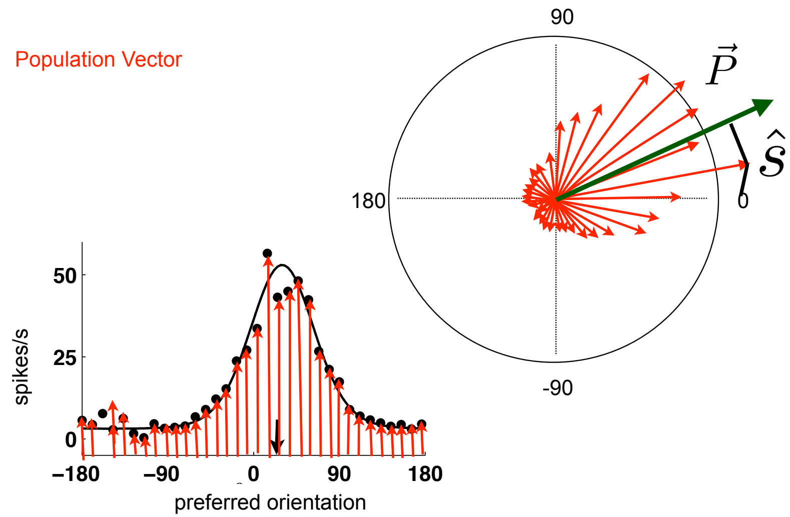
2. Simpler Decoding Strategies

Population Vector



2. Simpler Decoding Strategies

Population Vector



2. Simpler decoding strategies: Optimal Decoders within a class

Optimal decoders often requires much too much data (full model P[**r**|s]), seem too complex:

The question then is the cost of using non-optimal decoders, such as WTA or population vector, also sometimes considering optimal decoders within a class:

- Linear Decoders, eg. OLE, [Salinas and Abbott 1994] $\hat{s} = \sum_i w_i r_i$

- Decoders that ignore the correlations (decode with the "wrong model" which assumes independence) [Nirenberg & Latham 2000, Wu et al 2001, Seriès et al 2004]

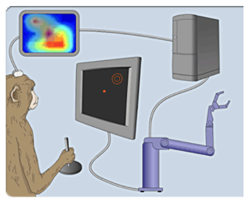
In practice also: supervised learning strategies, e.g. SVM.

Use of simple decoding methods for prosthetics

- Brain-machine interface usually use very simple decoding techniques, such as linear filters
- and they show promising results
- as well as surprising learning effects: dissociating BMI control from motor control, i.e., animals started controlling the robotic arm without any overt movement.

See eg. lab of M. Nicolelis @ Duke, and A. Schwartz @ Pittsburg

https://www.youtube.com/watch?v=_arFi-IfYc4



Use of simple decoding methods for prosthetics, in monkey and now also humans:

BrainGate technology:
- clinical trials 2004-2006 in individuals with tetraplegia, show they can control a cursor by thought.
- 2012: two people paralysed by brainstem stroke several years earlier were able to control robotic arms for reaching and grasping

nature Vol 442|13 July 2006|doi:10.1038/nature04970

ARTICLES

Neuronal ensemble control of prosthetic devices by a human with tetraplegia

Leigh R. Hochberg[1,2,4], Mijail D. Serruya[2,3], Gerhard M. Friehs[5,6], Jon A. Mukand[7,8], Maryam Saleh[9]†, Abraham H. Caplan[9], Almut Branner[10], David Chen[11], Richard D. Penn[12] & John P. Donoghue[2,9]

Neuromotor prostheses (NMPs) aim to replace or restore lost motor functions in paralysed humans by routeing movement-related signals from the brain, around damaged parts of the nervous system, to external effectors. To translate preclinical results from intact animals to a clinically useful NMP, movement signals must persist in cortex af spinal cord injury and be engaged by movement intent when sensory inputs and limb movement are long absent. Furthermore, NMPs would require that intention-driven neuronal activity be converted into a control signal that enabl useful tasks. Here we show initial results for a tetraplegic human (MN) using a pilot NMP. Neuronal ensemble activity recorded through a 96-microelectrode array implanted in primary motor cortex demonstrated that intended hand motion modulates cortical spiking patterns three years after spinal cord injury. Decoders were created, providing a 'neural cursor' with which MN opened simulated e-mail and operated devices such as a television, even while conversing. Furthermore, MN used neural control to open and close a prosthetic hand, and perform rudimentary actions with a multi-jointed robotic arm. These early results suggest that NMPs based upon intracortical neuronal ensemble spiking activity could provide a valuable new neurotechnology to restore independence for humans with paralysis.

http://www.bbc.co.uk/news/health-18092653

LETTER

doi:10.1038/nature1107

Reach and grasp by people with tetraplegia using a neurally controlled robotic arm

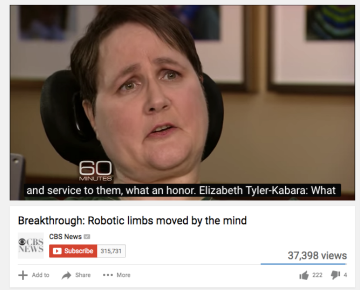
Leigh R. Hochberg[1,2,3,4], Daniel Bacher[2]*, Beata Jarosiewicz[1,5]*, Nicolas Y. Masse[5]*, John D. Simeral[1,2,3]*, Joern Vogel[6]*, Sami Haddadin[6], Jie Liu[1,2], Sydney S. Cash[3,4], Patrick van der Smagt[6] & John P. Donoghue[1,2,5]

Paralysis following spinal cord injury, brainstem stroke, amyotrophic lateral sclerosis and other disorders can disconnect the brain from the body, eliminating the ability to perform volitional movements. A neural interface system[1–5] could restore mobility and independence for people with paralysis by translating neuronal activity directly into control signals for assistive devices. We have previously shown that people with long-standing tetraplegia can use a neural interface

for algorithm development and interface testing, but she had n exposure to the DEKA arm before the sessions reported here. T participated in three DEKA arm sessions for similar developmen and testing before the session reported here but had no other experienc using the robotic arms.

To decode movement intentions from neural activity, electrical poten tials from each of the 96 channels were filtered to reveal extracellula

- **The story of Jan Sheuermann.**
- Diagnosed with spinocerebellar degeneration, she was only able to move her head and neck.
- in 2012, she seized an opportunity to turn her personal liability into an extraordinary asset for neuroscience: She elected to undergo brain surgery to implant two arrays of electrodes on her motor cortex
- used for increasingly complex tasks, in 2015 flying a F-35 flight Simulator with Her Mind



https://www.youtube.com/watch?v=Z3a5u6djGnE

TED Ideas worth spreading

Miguel Nicolelis: A monkey that controls a robot with its thoughts.

http://www.youtube.com/watch?v=CR_LBcZg_84

Neural Prosthetics: Krishna Shenoy at TEDxStanford

http://www.youtube.com/watch?v=ZuATvhlcUU4

Decoding in humans from fMRI signals

- Decoding techniques are increasingly used with fMRI signals, e.g. to decode the movie you're viewing (Jack Gallant, http://www.youtube.com/watch?v=6FsH7RK1S2E), or what class of objects you're thinking about, http://www.youtube.com/watch?v=6FsH7RK1S2E
- Classification task - pure machine learning ! (videolectures.net/fmri06_mitchell_odmsp/)
  Often used, support vector machines.

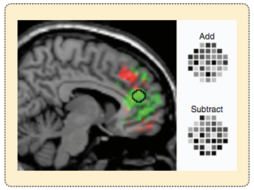


FIGURE 5 Haynes' research shows that it is possible to determine a subject's intentions—in this case, whether the person was preparing to perform an addition or a subtraction—by reading brain-activity patterns. Activity patterns in the green regions predicted covert intentions before the subject began to perform the calculation. The regions marked in red revealed intentions that were already being acted upon. (Photo courtesy of Bernstein Center for Computational Neuroscience.)

some ethical questions too …
lie detection: fMRI now better than polygraphs?

J Clin Psychiatry. 2016 Oct;77(10):1372-1380. doi: 10.4088/JCP.15m09785.

Polygraphy and Functional Magnetic Resonance Imaging in Lie Detection: A Controlled Blind Comparison Using the Concealed Information Test.

Langleben DD[1,2,3], Hakun JG[4], Seelig D[2], Wang AL[2], Ruparel K[2], Bilker WB[2], Gur RC[2,3].

⊕ Author information

Abstract

OBJECTIVE: Intentional deception is a common act that often has detrimental social, legal, and clinical implications. In the last decade, brain activation patterns associated with deception have been mapped with functional magnetic resonance imaging (fMRI), significantly expanding our theoretical understanding of the phenomenon. However, despite substantial criticism, polygraphy remains the only biological method of lie detection in practical use today. We conducted a blind, prospective, and controlled within-subjects study to compare the accuracy of fMRI and polygraphy in the detection of concealed information. Data were collected between July 2008 and August 2009.

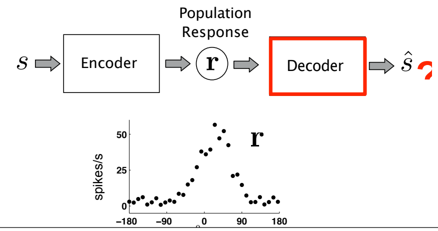
METHOD: Participants (N = 28) secretly wrote down a number between 3 and 8 on a slip of paper and were questioned about what number they wrote during consecutive and counterbalanced fMRI and polygraphy sessions. The Concealed Information Test (CIT) paradigm was used to evoke deceptive responses about the concealed number. Each participant's preprocessed fMRI images and 5-channel polygraph data were independently evaluated by 3 fMRI and 3 polygraph experts, who made an independent determination of the number the participant wrote down and concealed.

RESULTS: Using a logistic regression, we found that fMRI experts were 24% more likely (relative risk = 1.24, P < .001) to detect the concealed number than the polygraphy experts. Incidentally, when 2 out of 3 raters in each modality agreed on a number (N = 17), the combined accuracy was 100%.

CONCLUSIONS: These data justify further evaluation of fMRI as a potential alternative to polygraphy. The sequential or concurrent use of psychophysiology and neuroimaging in lie detection also deserves new consideration.

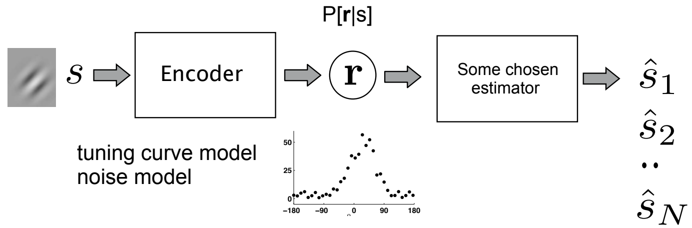
Decoding: Summary of previous slides

- Decoding: for neuro-prostheses and/or for understanding the relationship between the brain's activity and perception or action
- Different strategies are possible: optimal decoders (e.g. ML, MAP) vs simple decoders (e.g. winner take all, population vector), depending on what we know about the encoding model, and constraints.
- Exciting applications : Brain Machine interfaces (BMI), fMRI analysis.



Now: **Decoding as a tool for understanding the link between the properties of neurons, the precision of the code and behavioural performances**

From Population Codes to Psychophysical Performances



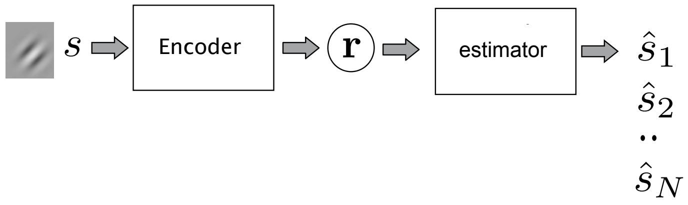
- **How can we relate this model of perception with measured psychophysical performance?**
- **Can we reverse-engineer information about the encoder and decoder based on experimental data ?**

A little detour

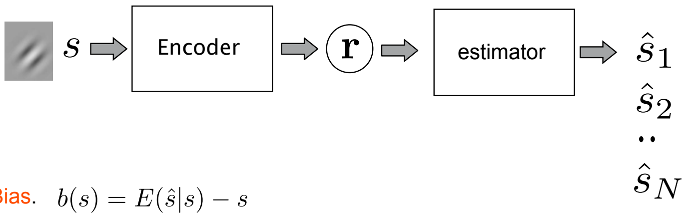
Estimation theory 101:

Estimation theory: the branch of statistics that deals with estimating the values of parameters based on measured empirical data that has a random component.

Performances of our model estimator (Estimation theory 101)



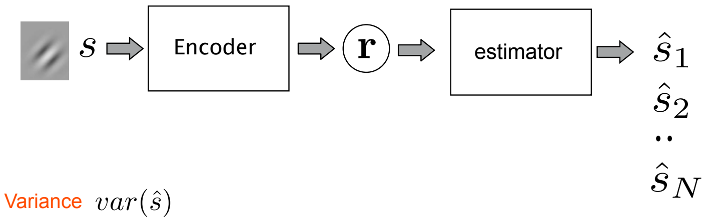
Measuring the performances of our model estimator (Estimation theory 101)



- 1) Bias. $b(s) = E(\hat{s}|s) - s$

  If $E(\hat{s}|s) = s$ the estimator is said to be <u>unbiased</u>.

Measuring the Performances of our model estimator (Estimation theory 101)



- 2) Variance $var(\hat{s})$

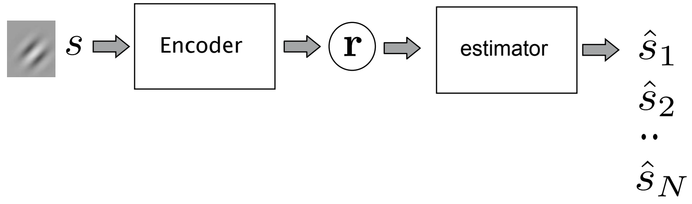
  If var = as small as possible, the (unbiased) estimator is said to be <u>efficient</u>

  The smallest possible variance is given by the Cramér-Rao Bound.
  The denominator is known as Fisher Information, a function of P[r|s].

$$var(\hat{s}) \geq \frac{(1 + b'(s))^2}{I_F(s)}$$

where $I_F(s) = - < \frac{\partial^2 \ln P[r|s]}{\partial s^2} >$

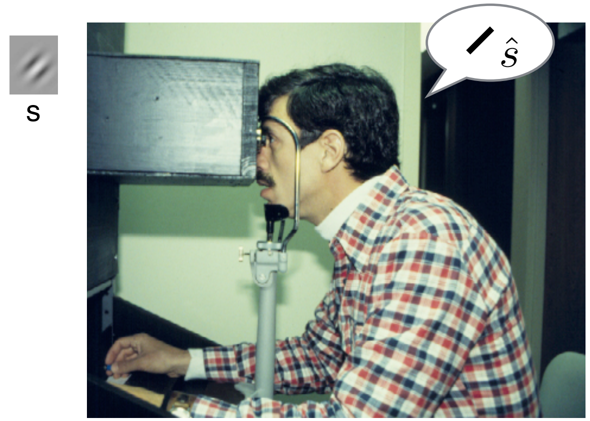
Measuring the Performances of our model estimator (Estimation theory 101)



- As a consequence, the best possible estimator is
  - unbiased,
  - have a variance defined by (the inverse of) Fisher Information.

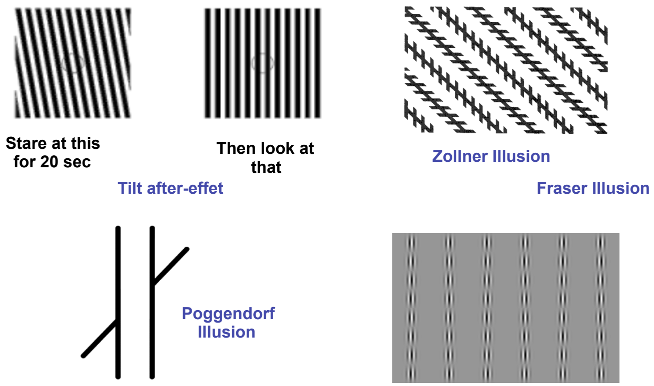
What is Being Measured in Psychophysics ?

Estimation tasks

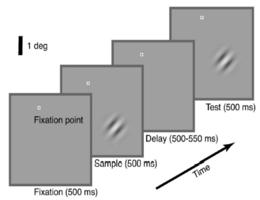


a) Estimation tasks

✳ The measured quantity is the (average) difference between the perceived orientation and the real orientation $< \hat{s} > - s$



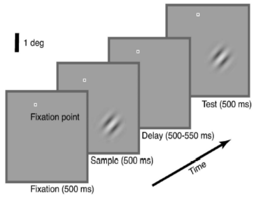
b) Discrimination Tasks



✳ **Is the second grating of the same orientation as the first grating, or a different orientation?**

✳ The measured quantity is the Discrimination Threshold a.k.a Just Noticeable difference (JND)
- on average detected on 76% of the trials.

b) Discrimination Tasks

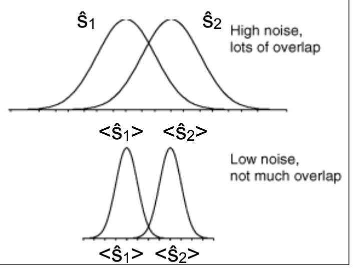


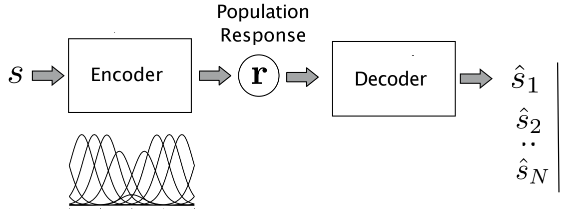
✳ **Is the second grating of the same orientation as the first grating, or a different orientation?**

✳ The measured quantity is the Discrimination Threshold a.k.a Just Noticeable difference (JND) - on average detected on 76% of the trials.

Discrimination threshold depends on the overlap between the internal 'representation' of the 2 stimuli: p[$\hat{s}_1$|**r**] and p[$\hat{s}_2$|**r**]:

- The bias of the internal representation (expansion/contraction of the 'distance' between the stimuli)
- How noisy the internal representation is (the variance of the estimates)





Linking the statistics of the model and psychophysics

$$\mathrm{b}(\hat{s}) = <\hat{s}> - s$$ ⟷ perceptual bias

$$\mathrm{var}(\hat{s})$$

$$\mathrm{threshold}(\hat{s}) = \frac{\mathrm{std}(\hat{s})}{1 + b'(\hat{s})}$$ ⟷ discrimination threshold (76% correct) just noticeable difference

Fisher information: the best possible discrimination performance for a given encoder model

✴ Fisher information: gives the discrimination threshold that would be obtained (asymptotically) by an optimal decoder, for eg. ML (units of var ^-1)

$$\mathrm{threshold}(\hat{s}) \geq \frac{1}{\sqrt{I_F(s)}}$$

✴ is expressed in terms of the encoding model P[**r**|s], i.e. in terms of the tuning curves and the noise

$$I_F(s) = - < \frac{\partial^2 \ln P[r|s]}{\partial s^2} >$$

e.g $P(\mathrm{r} = k|s) = \dfrac{e^{-f(s)} f(s)^k}{k!}$

$$I(s) = \sum_i \frac{f_i'(s)^2}{f_i(s)}$$

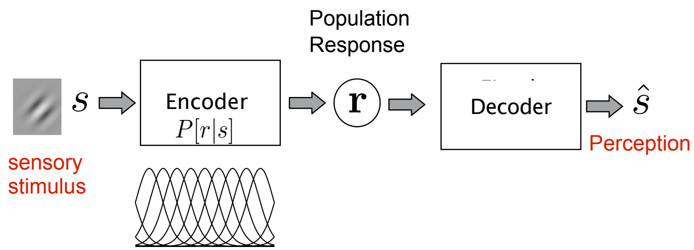
✴ Interpreted as a measure of 'information' in the responses;
✴ is related with Mutual information and Stimulus Specific Information (Brunel and Nadal 1998, Yarrow, Challis and Seriès 2012).

From Population Responses to Psychophysics

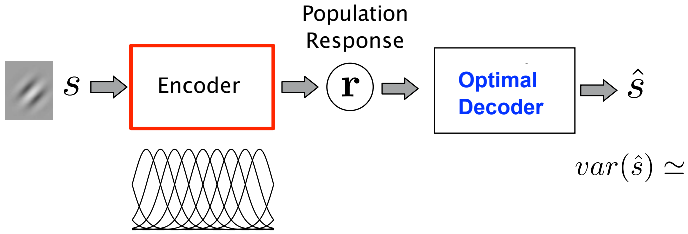
Two strategies:

✴ Assume the decoder is optimal: Compute Fisher information from P[r|s]. This gives us the minimal possible variance of any unbiased decoder, and the minimal threshold of any decoder (biased or unbiased).

✴ Construct explicitly the decoder (e.g. population vector). Compute explicitly bias, variance, and threshold of estimates.



From Population Responses to Psychophysics



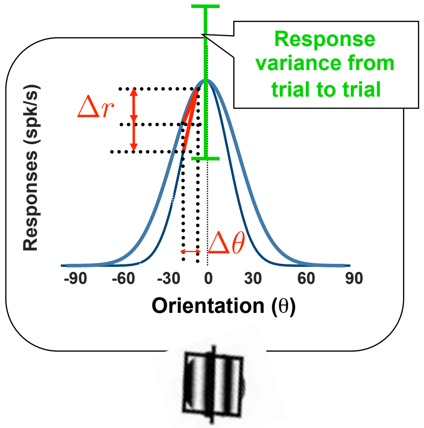
$$var(\hat{s}) \simeq \frac{1}{I_F(s)}$$

$$\mathrm{thres}(\hat{s}) \simeq \frac{1}{\sqrt{I_F(s)}}$$

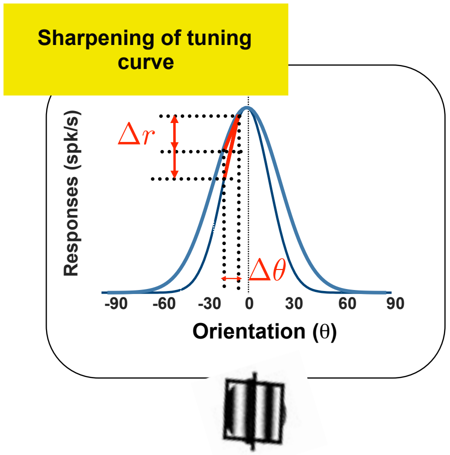
Questions that we can explore:
What changes in encoder would increase discrimination performances?

- Number of neurons?
- Tuning curves shape ?
- Noise correlations ?

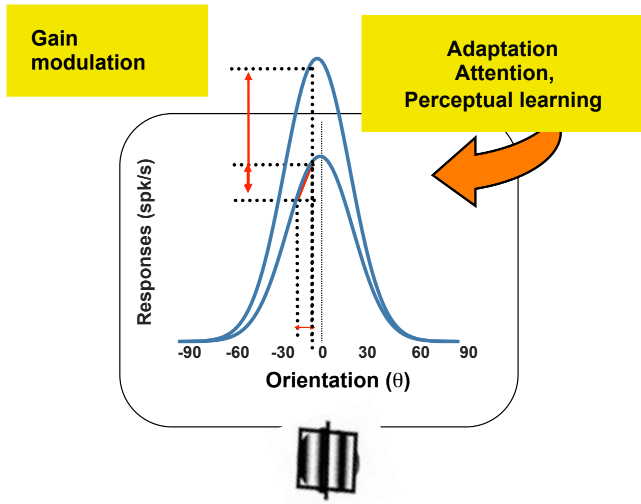
What are the factors that control performance?



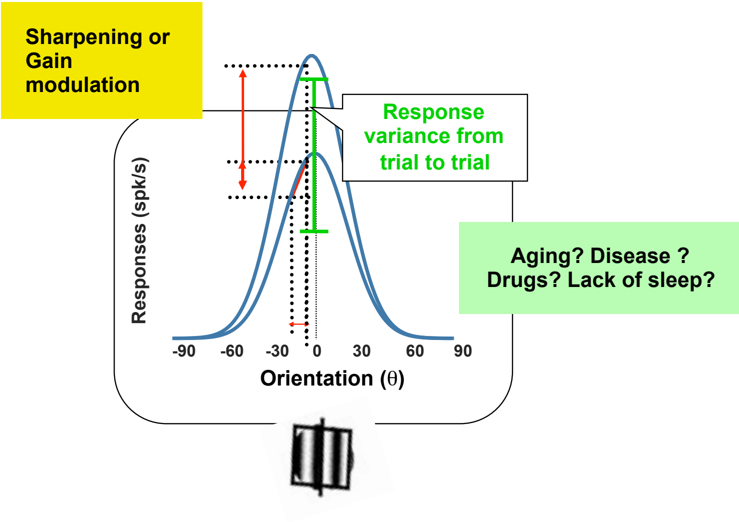
What are the factors that control performance?



What are the factors that control performance?



What are the factors that control performance?



What are the factors that control performance?

- Fisher information formalises intuition and provides a tool to explore these questions precisely.
- For Poisson noise

$$I_i(s) = \frac{f_i'(s)^2}{f_i(s)}$$

Slope [2] / **variance**

$$I(s) = \sum_i \frac{f_i'(s)^2}{f_i(s)}$$

For independent neurons, FI of the population is the sum of each neurons' FI

What are the factors that control performance?

For Gaussian uncorrelated noise:

$$I(s) = \sum_i \frac{f_i'(s)^2}{\sigma_i^2(s)}$$

For Gaussian correlated noise:

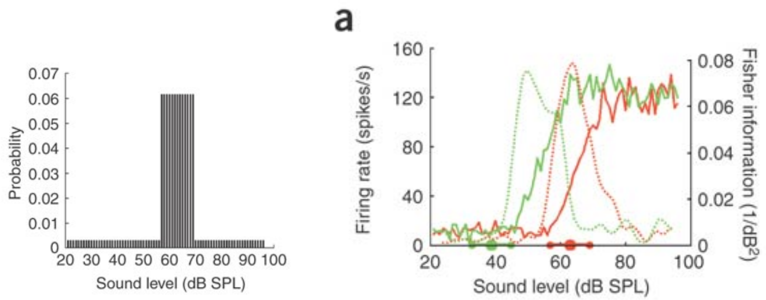
$$P[\mathbf{r}|s] = \frac{1}{\sqrt{(2\pi)^N|\mathbf{Q}(\mathbf{s})|}} e^{-\frac{1}{2}(\mathbf{r}-\mathbf{f}(\mathbf{s}))^{\mathbf{T}}\mathbf{Q}^{-1}(\mathbf{s})(\mathbf{r}-\mathbf{f}(\mathbf{s})))}$$

$$I_F(s) = \mathbf{f}'(\mathbf{s})\mathbf{Q}^{-1}(\mathbf{s})\mathbf{f}'(\mathbf{s}) + \frac{1}{2}\mathrm{Trace}[\mathbf{Q}^{-1}(\mathbf{s})\mathbf{Q}'(\mathbf{s})\mathbf{Q}^{-1}(\mathbf{s})\mathbf{Q}'(\mathbf{s})]$$

For correlated neurons, FI is modulated by correlations.

Research questions (1)

- What would be the 'optimal' shape for tuning curves?
- Are adaptation, attention and learning a step towards more 'optimal' tuning curves for the attended/trained stimulus ?

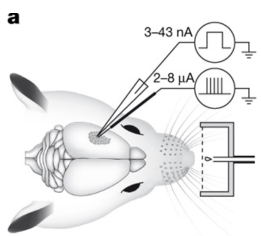


Neurons in auditory midbrain of the guinea pig adjust their response to improve the accuracy of the code close to the region of most commonly occurring sound levels.

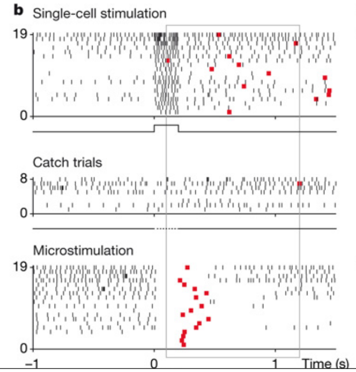
[Dean, Harper & McAlpine, Nature Neuro, 2005]

Research questions (2)

✴ How many neurons participate in a psychophysical task ? (see also, lab 1) 1, 10, 100, 10000? How can we find out ?
✴ comparing performance (e.g. MT: Britten et al 1992; Stuttgen & Schwartz 2008). stimulating (MT: Salzman, Britten, Newsome 1990).



Houweling & Brecht, Nature, 2008
Barrel cortex single cell stimulation.



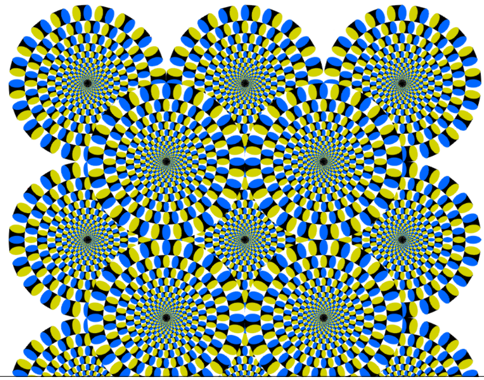
Research questions (3)

✴ Pooling from large populations of neurons thought to be a way to average out the noise.
✴ Pairs of neurons show correlations in their variability: does pooling more and more neurons increases (linearly) the accuracy of the representation?
or Is information saturating over a certain number of neurons ? [Zohary et al 1994]

✴ Could that be that adaptation, attention and perceptual learning act by changing correlations? [Cohen & Maunsell 2009; Gutnisky & Dragoi 2008, Gu et al 2011, Bejjanki et al 2011]

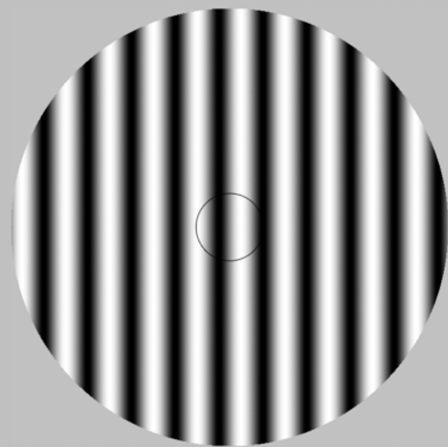
Research questions (4)

✴ Can the study of illusions inform us on the type of 'decoder' that is used in the brain? [Seriès, Stocker and Simoncelli 2009]



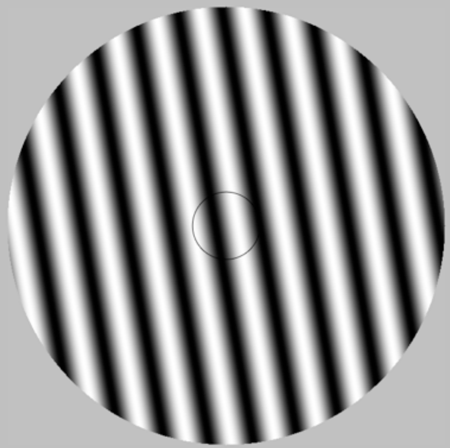
Sensory Adaptation

Verify that this grating is vertical



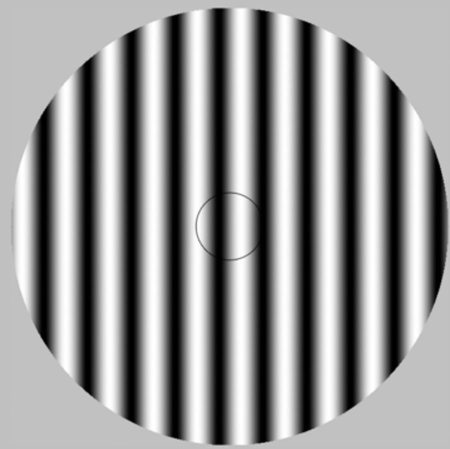
Sensory Adaptation

Fixate on the central circle for 30 sec



Sensory Adaptation

Now observe the grating again



The Tilt After-Effect

Visual Adaptation: Psychophysics

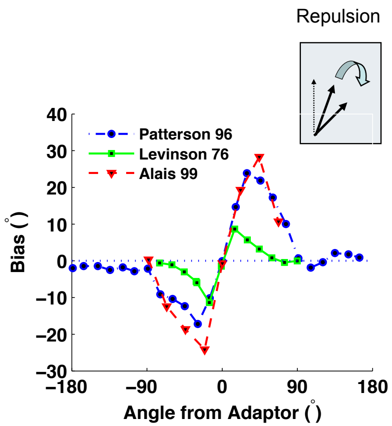
Visual adaptation leads to:

- estimation tasks: strong biases (mainly repulsion)

Visual Adaptation: Psychophysics

Visual adaptation leads to:

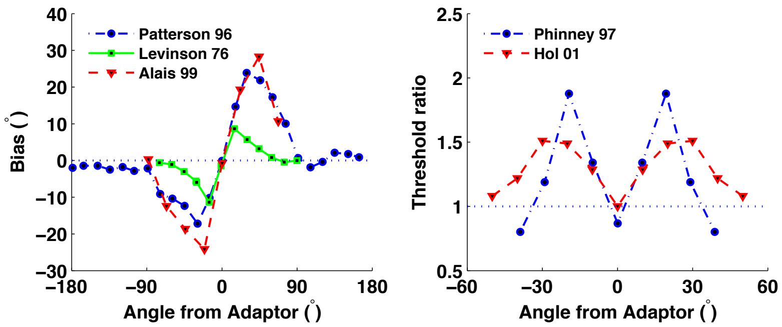
- estimation tasks: strong biases (mainly repulsion)



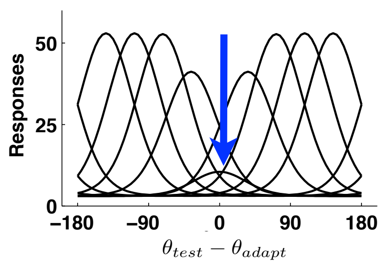
Visual Adaptation: Psychophysics

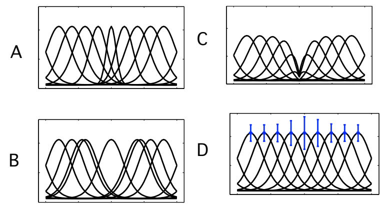
Visual adaptation leads to:

- ❖ estimation tasks: strong biases (mainly repulsion)
- ❖ discrimination tasks: changes in performance



Visual Adaptation: Physiology





Mainly a Gain change

[Van Wezel & Britten 2002, Krekelberg et al. 2006]

Other effects are controversial, dependent on time scale and area: shifts in preferred orientation, changes in width, changes in variability.

[Kohn & Movshon 2004, Dragoi et al, 2000]

ARTICLE — Communicated by Peter Dayan

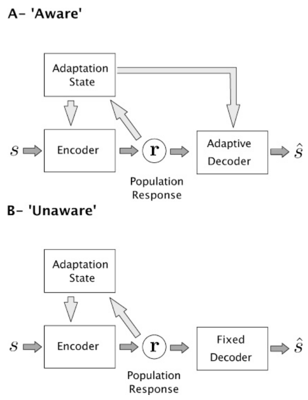
Is the Homunculus "Aware" of Sensory Adaptation?

Peggy Seriès
pseries@inf.ed.ac.uk
IANC, University of Edinburgh, Edinburgh EH8 9AB, U.K.

Alan A. Stocker
astocker@sas.upenn.edu
Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, U.S.A.

Eero P. Simoncelli
eero.simoncelli@nyu.edu
Howard Hughes Medical Institute, Center for Neural Science, and Courant Institute for Mathematical Sciences, New York University, New York, New York 10003, U.S.A.

Neural activity and perception are both affected by sensory history. The work presented here explores the relationship between the physiological effects of adaptation and their perceptual consequences. Perception is modeled as arising from an encoder-decoder cascade, in which the encoder is defined by the probabilistic response of a population of neurons, and the decoder transforms this population activity into a perceptual estimate. Adaptation is assumed to produce changes in the encoder, and we examine the conditions under which the decoder behavior is consistent with observed perceptual effects in terms of both bias and discriminability. We show that for all decoders, discriminability is bounded from below by the inverse Fisher information. Estimation bias, on the other hand, can arise for a variety of different reasons and can range from zero to substantial. We specifically examine biases that arise when the decoder is fixed, "unaware" of the changes in the encoding population (as opposed to "aware" of the adaptation and changing accordingly). We simulate the effects of adaptation on two well-studied sensory attributes, motion direction and contrast, assuming a gain change description of encoder adaptation. Although we cannot uniquely constrain the source of decoder bias, we find for both motion and contrast that an "unaware" decoder that maximizes the likelihood of the percept given by the preadaptation encoder leads to predictions that are consistent with behavioral data. This model implies that adaptation-induced biases arise as a result of temporary suboptimality of the decoder.



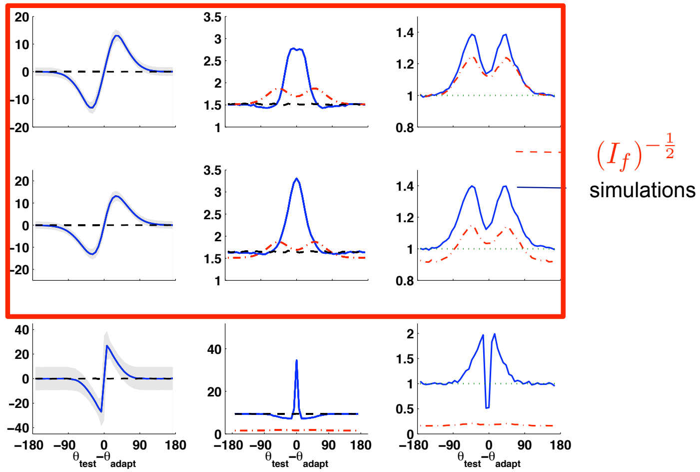
Results (2) -- 'unaware' read-out

Bias | [Standard Deviation] | Threshold Ratio

ML unaware

PV

WTA



[Series, Stocker and Simoncelli, 2009]

Summary

- The efficiency of Estimators / Decoders can be characterized by the bias and the variance.
- The bias and variance of estimators used to read-out neural responses can be easily compared with psychophysical performance (estimation biases, and discrimination threshold).

- Fisher Information is related to the minimal variance of a unbiased estimator.
- In a model of a population of neurons, Fisher Information can be expressed in terms of the tuning curves and the noise.
- Fisher information can be used to relate population responses and discrimination performances. It gives a bound on the discrimination threshold
- Fisher Information can be used to explore the factors that impact on the precision of the code / behavioral performances.