NOT NOISY, JUST WRONG

Alexandre Pouget

Department of Basic Neuroscience

UNIVERSITÉ DE GENÈVE

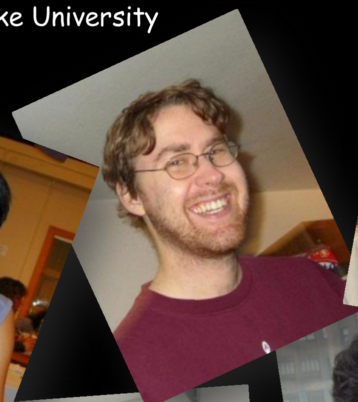
Department of Brain and Cognitive Sciences

UNIVERSITY of ROCHESTER

Collaborators

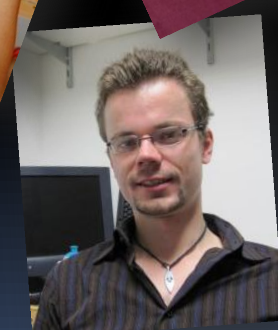
Peter Latham
Gatsby Computational Neuroscience Unit. London, UK

Jeffrey Beck
Duke University



Ruben Coen-Cagli
Universite de Geneve

Wei Ji Ma
NYU



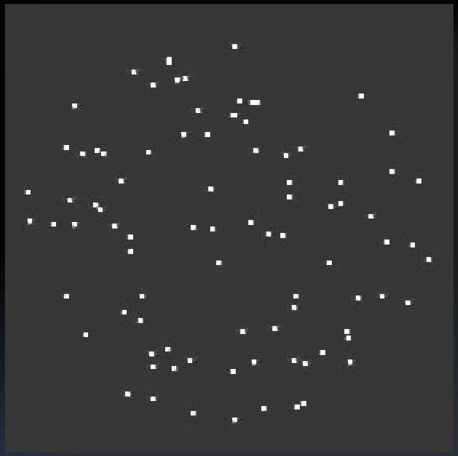
Ruben Moreno Bote
Foundation Sant Joan de deu
Universitat de Barcelona

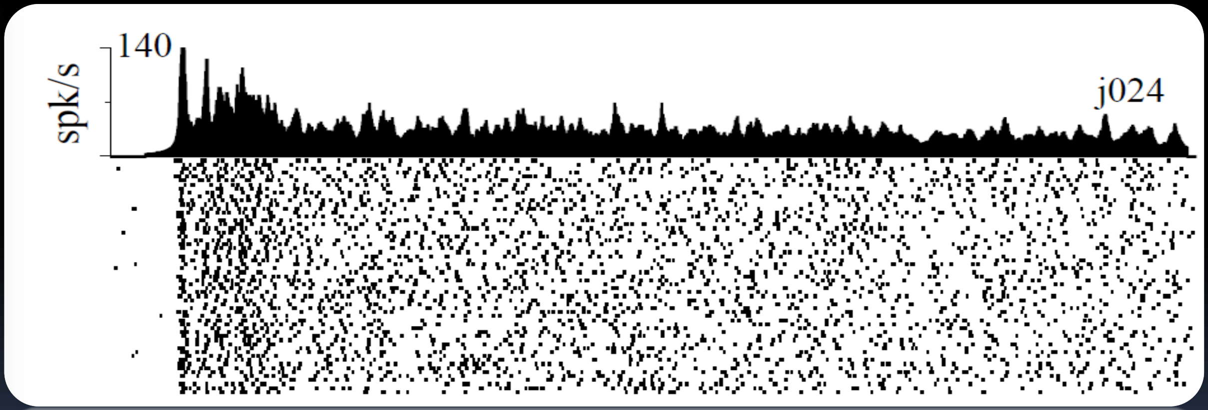
Xaq Pitkow
Baylor College of Medecine

Ingmar Kanitscheider
Universite de Geneve

Who is to blame?

Near-Poisson variability in spike trains



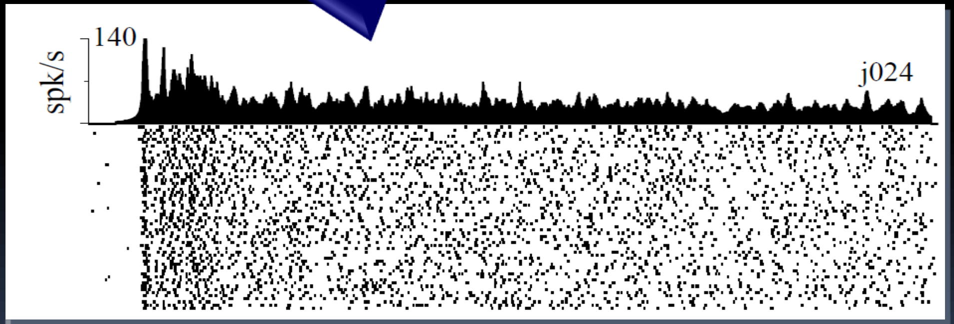


(Data from Newsome et al)

Who is to blame?

Near Poisson variability due to chaotic dynamics of balanced networks

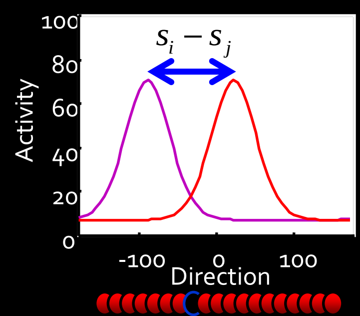
(Van Vreeswijk and Sompolinsky, 1996; Shadlen and Newsome, 1998; Banerjee et al, 2008, London et al, 2010)

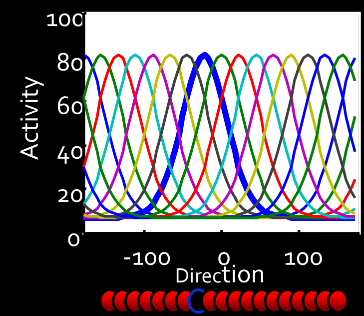


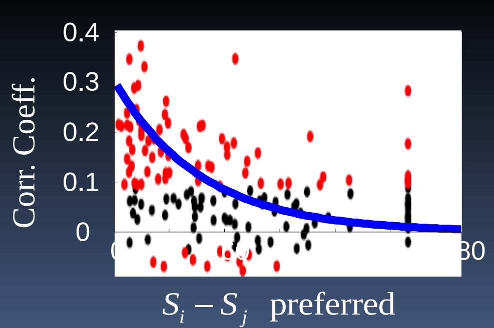
Near-Poisson variability → Behavioral variability

Who is to blame?

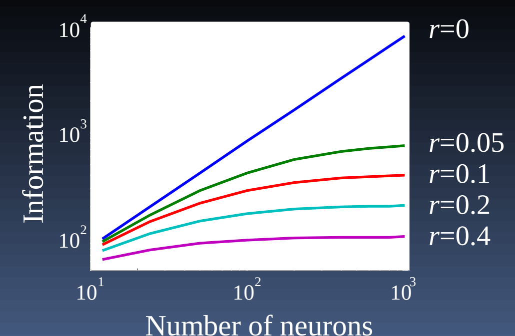
- It's not just Poisson variability, it's correlated Poisson variability







Huang and Lisberger, 2010



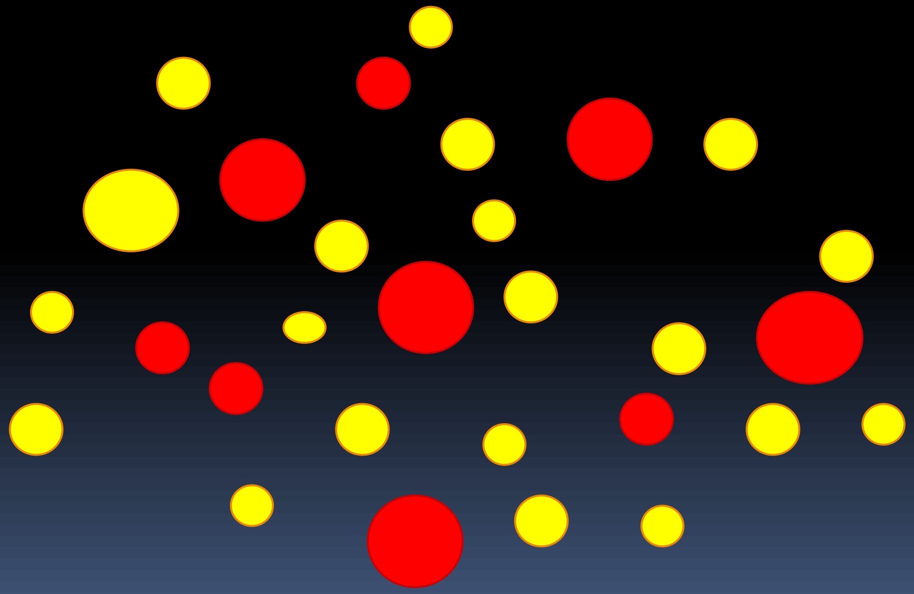
Zohary et al. 1994

Who is to blame?

- It’s not just Poisson variability, it’s correlated Poisson variability
- This variability explains, among other things, Weber’s law

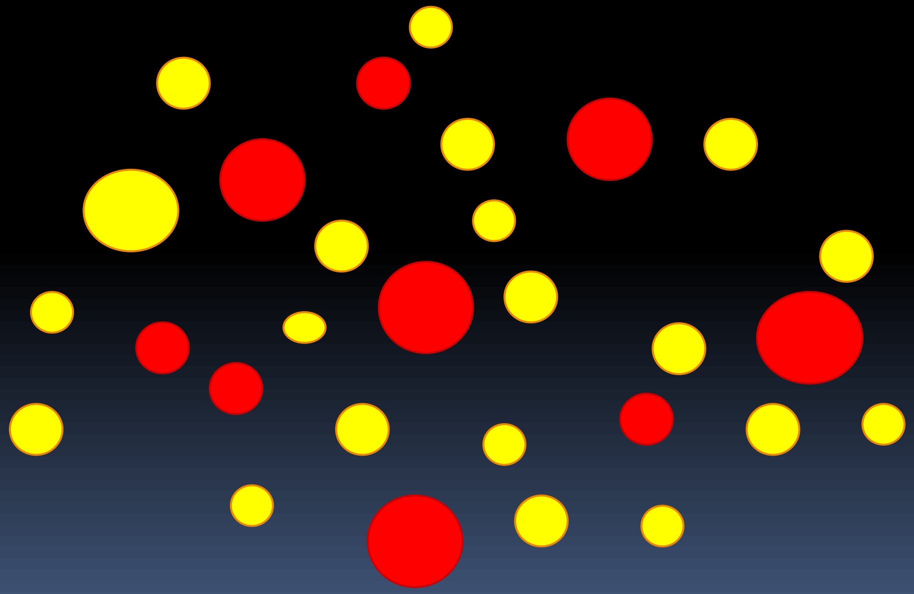
Weber's law

$$\frac{\Delta N}{N} = \text{constant}$$



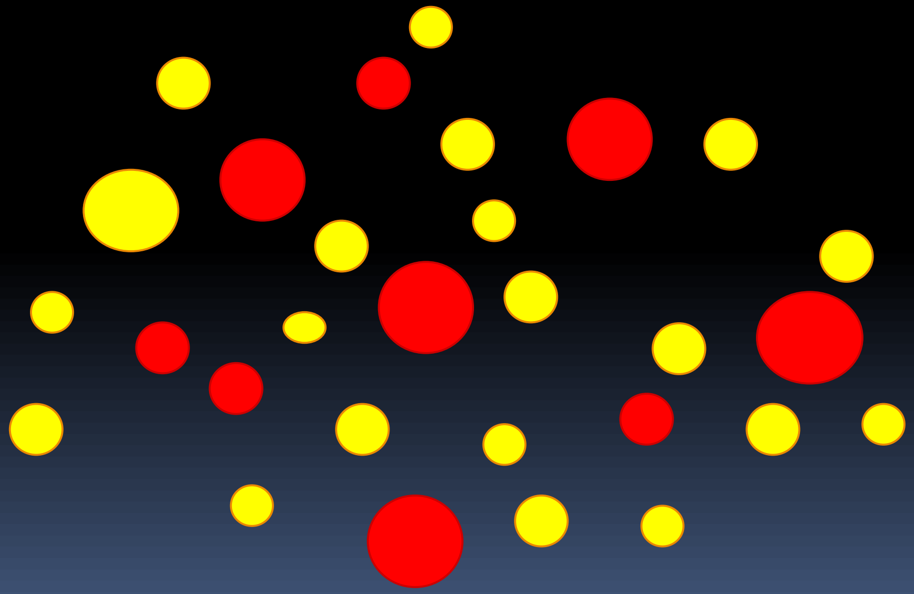
Weber's law

$$\frac{\Delta N}{N} = \frac{2}{10} = \text{constant}$$



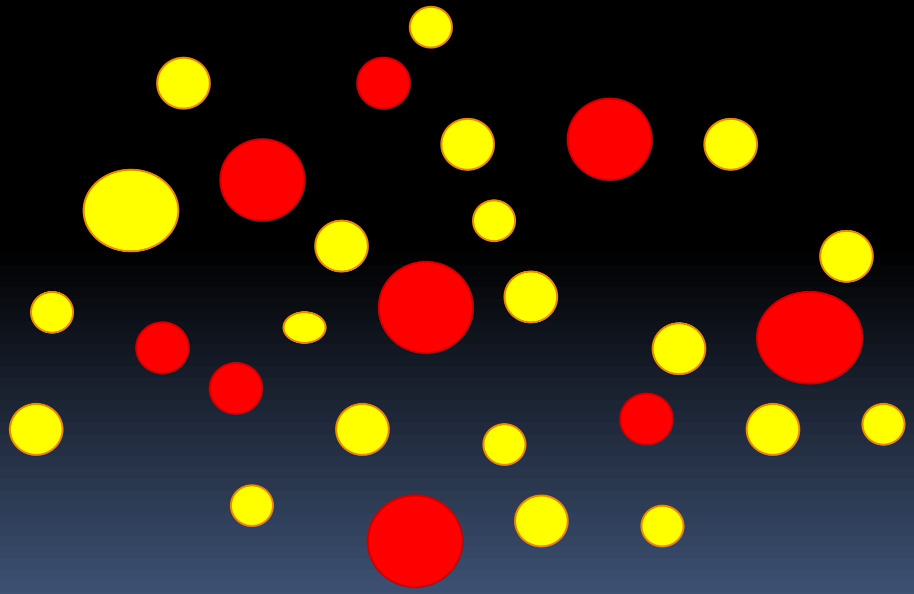
Weber's law

$$\frac{\Delta N}{N} = \frac{2}{10} = \frac{4}{20} = \text{constant}$$



Weber's law

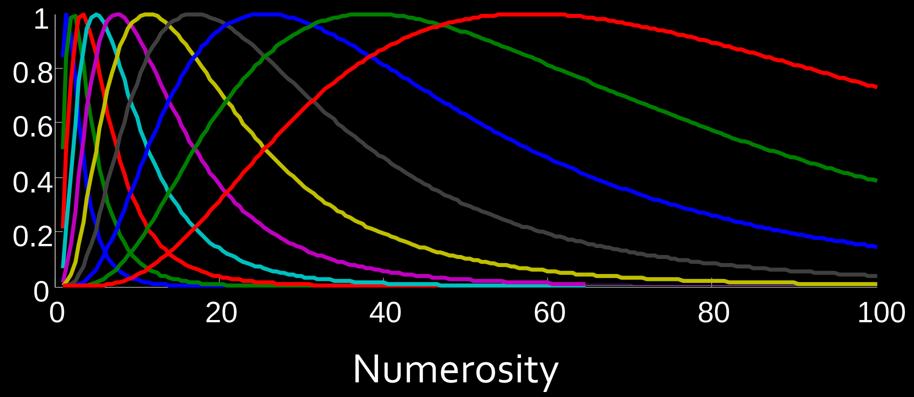
$$\frac{\Delta N}{N} = \frac{2}{10} = \frac{4}{20} = \frac{8}{40} = \text{constant}$$

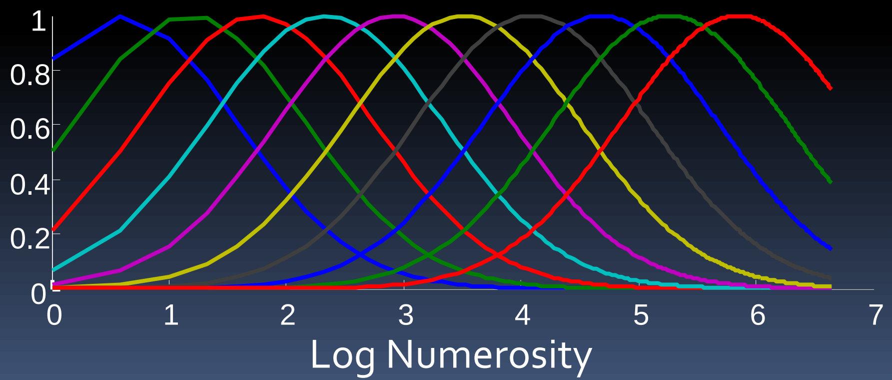


Weber's law

- Numerosity
- Line length
- Weights
- Luminance
- Speed of motion
- Distance traveled
- Time
- …

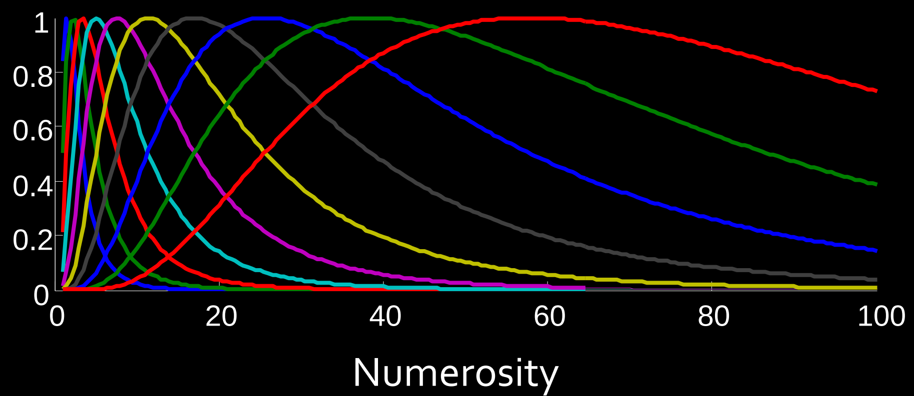
Weber’s law

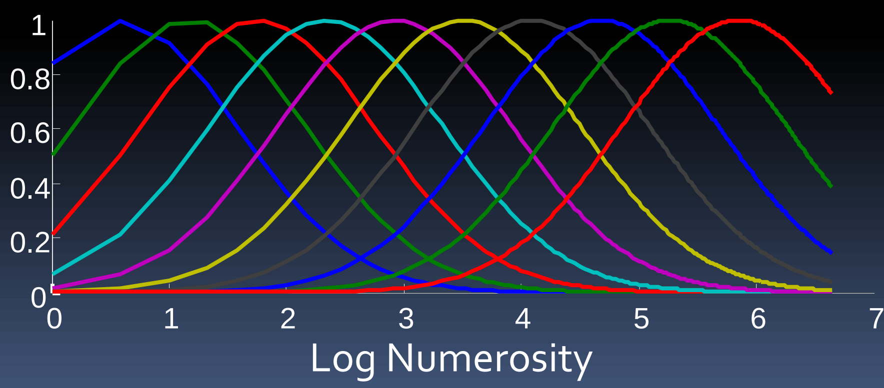




+ Poisson noise → Weber’s law

Weber’s law



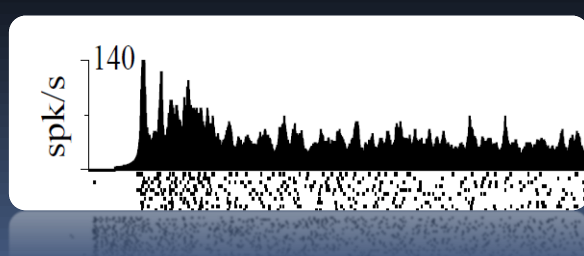


NO Poisson single cell variability

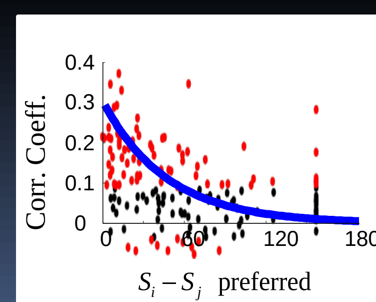
NO
Weber’s law

The origin of behavioral variability

Near-Poisson variability due to the chaotic dynamics of balanced networks of excitatory and inhibitory neurons, along with correlations inversely proportional to the difference in preferred stimuli, are the main causes of behavioral variability.



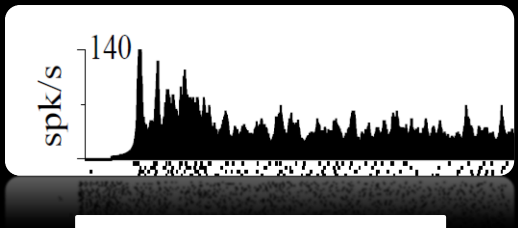
+



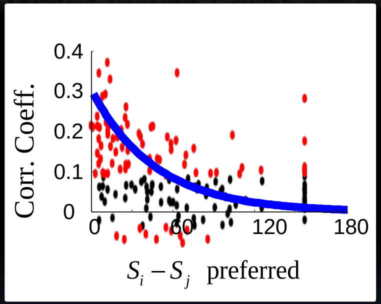
Alternative view

The brain is noisy but…

- The near-Poisson noise induced by chaotic dynamics of cortical circuits has little impact on behavior



- Correlations inversely proportional the difference in preferred stimuli do not necessarily limit information



- Most of behavioral variability comes from

1. Variable data from the world (which naturally leads to Weber's law)
2. Suboptimal inference

Roadmap

- Suboptimal inference can generate behavioral variability
- This cause dominates in most situations
- What this theory explains
- Implications for neuronal variability
- A normative view of Weber's law

Roadmap

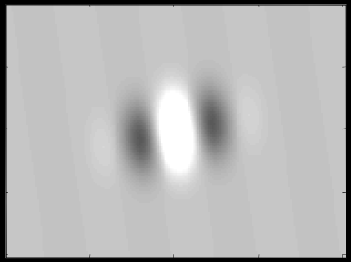
- Suboptimal inference can generate behavioral variability
- This cause dominates in most situations
- What this theory explains
- Implications for neuronal variability
- A normative view of Weber's law

Orientation discrimination

- Let’s consider a simple task:

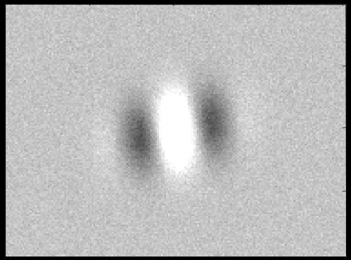
  Orientation discrimination

Orientation discrimination

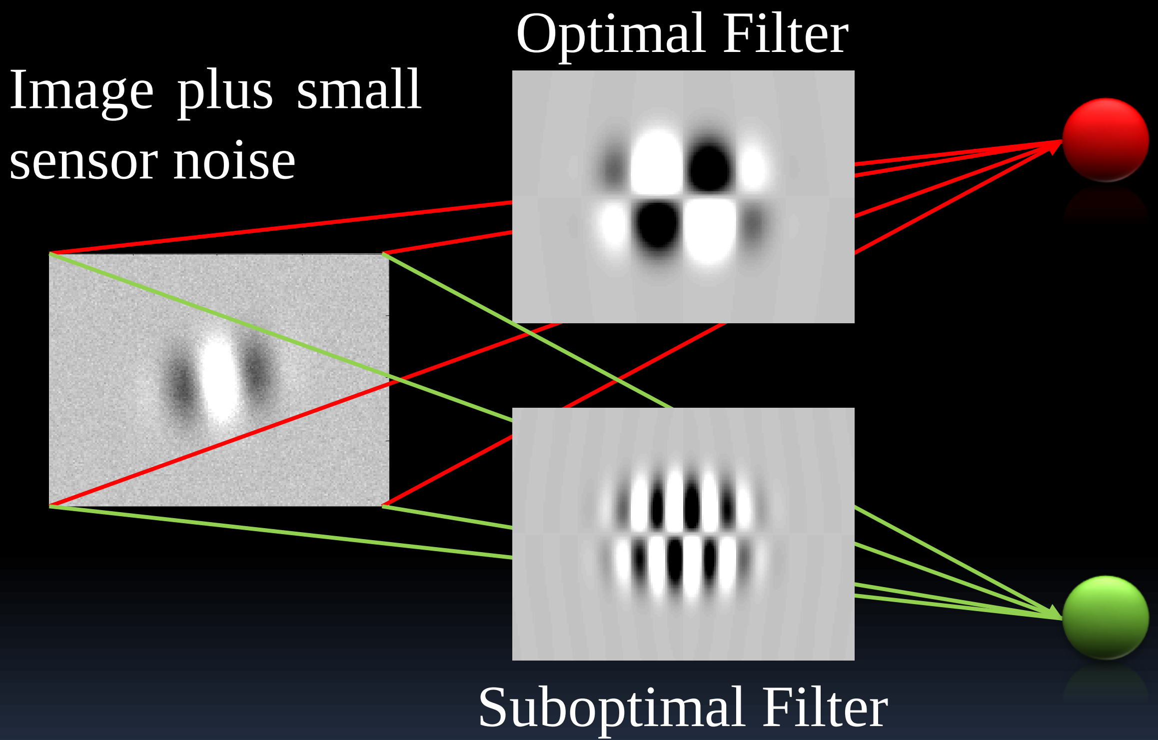


Orientation discrimination

Image plus small
sensor noise

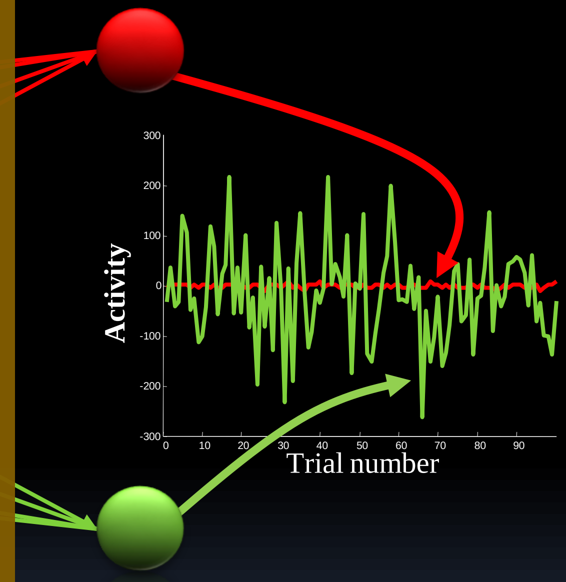


Orientation discrimination

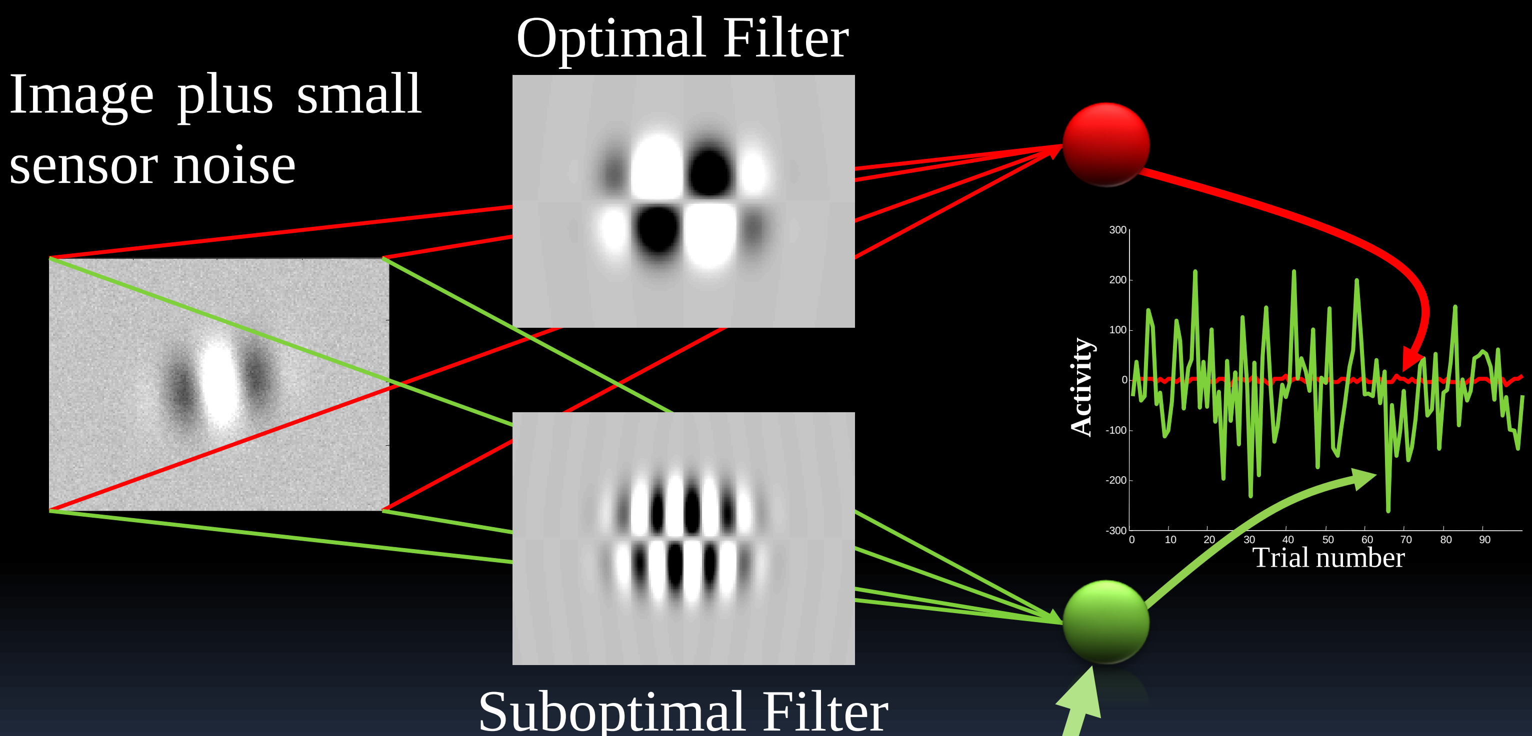


Orientation discrimination

NOTE:
The filters have been adjusted to ensure that the red and green units have the same mean activity

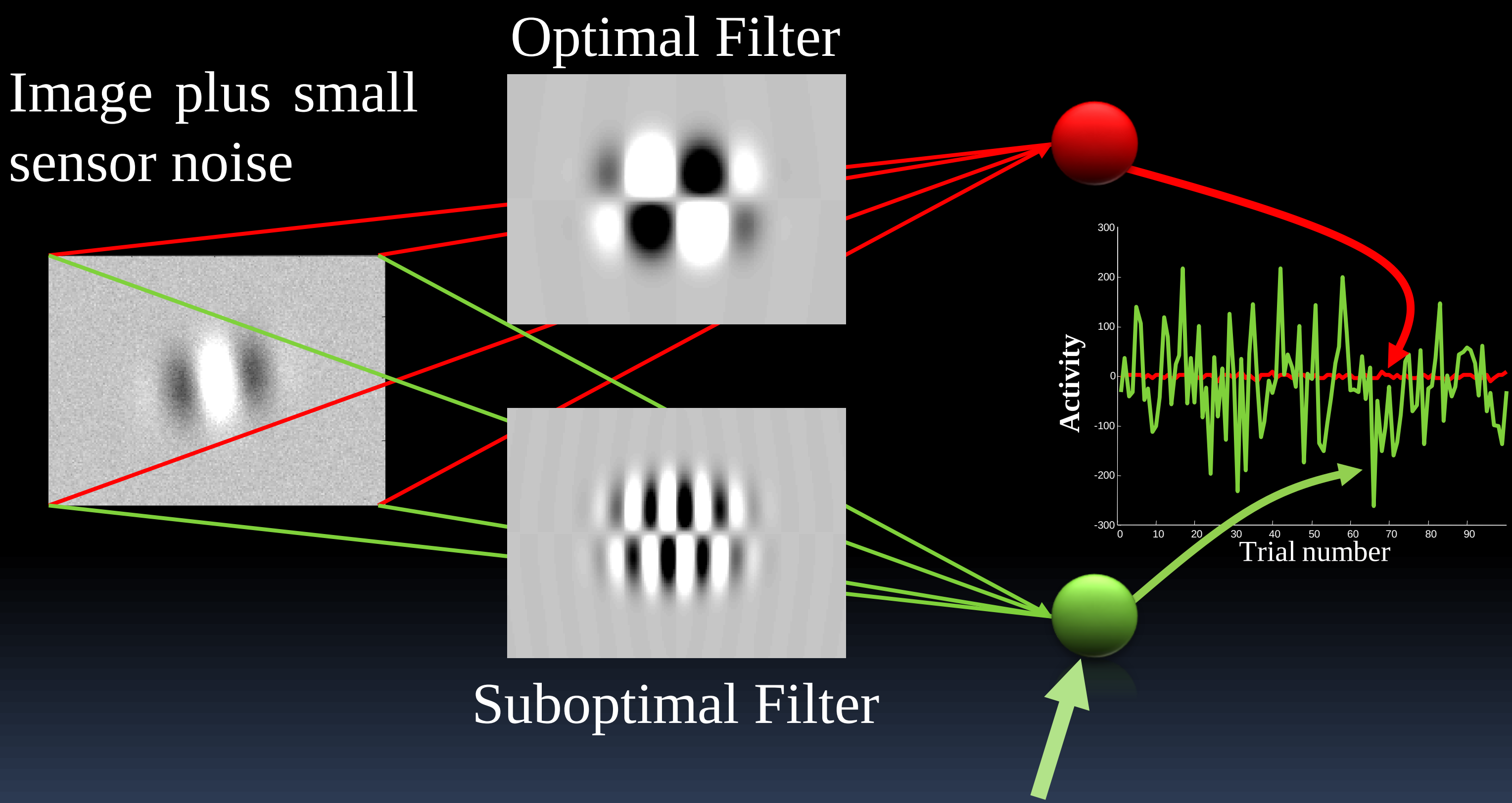


Orientation discrimination



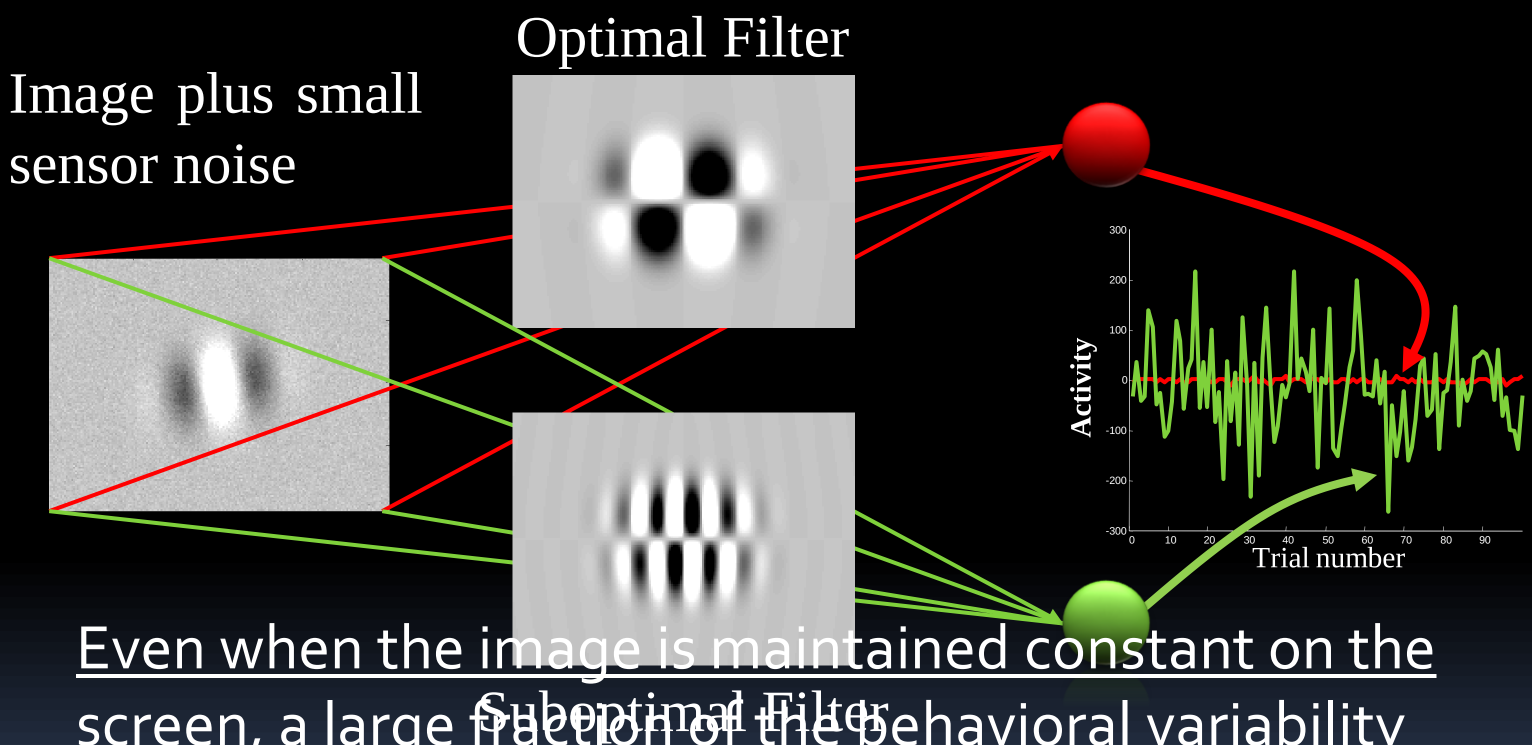
The extra variability comes from a *<u>deterministic</u>* approximation

Orientation discrimination



99% of the variability comes from the deterministic approximation!

Orientation discrimination



Even when the image is maintained constant on the screen, a large fraction of the behavioral variability might be due to suboptimal inference.

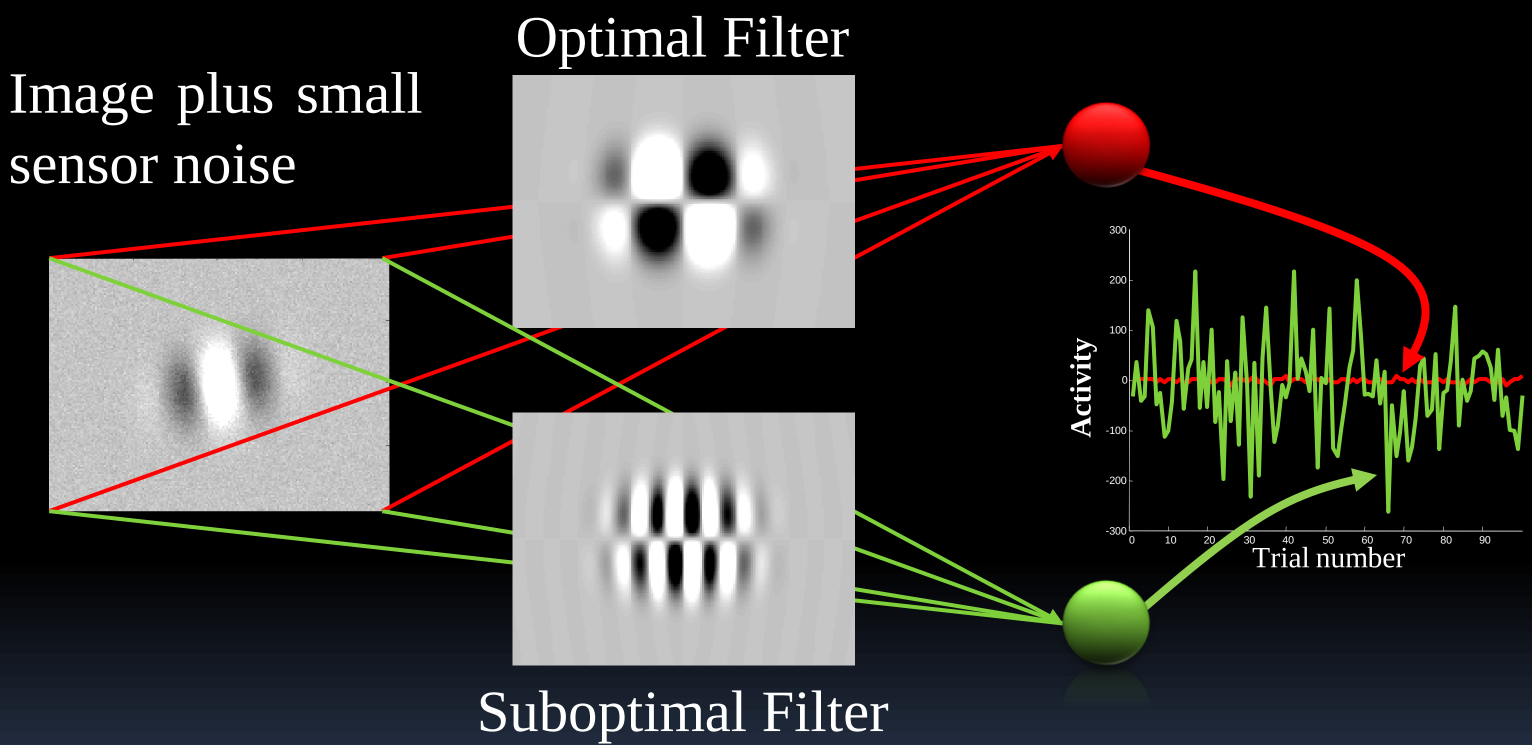
Causes of variability

- Suboptimal inference leads to extra variability.

- For complex problems, suboptimal approximations are unavoidable and dominate

(Beck et al, 2012)

Orientation discrimination

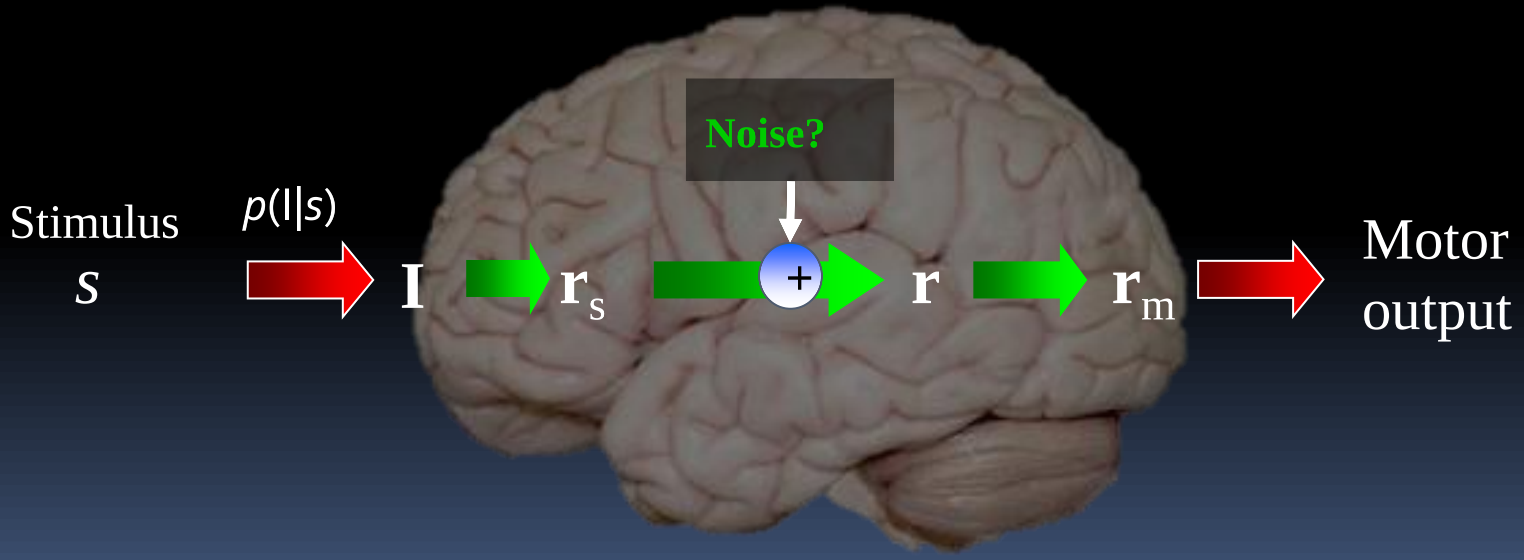


Motion estimation

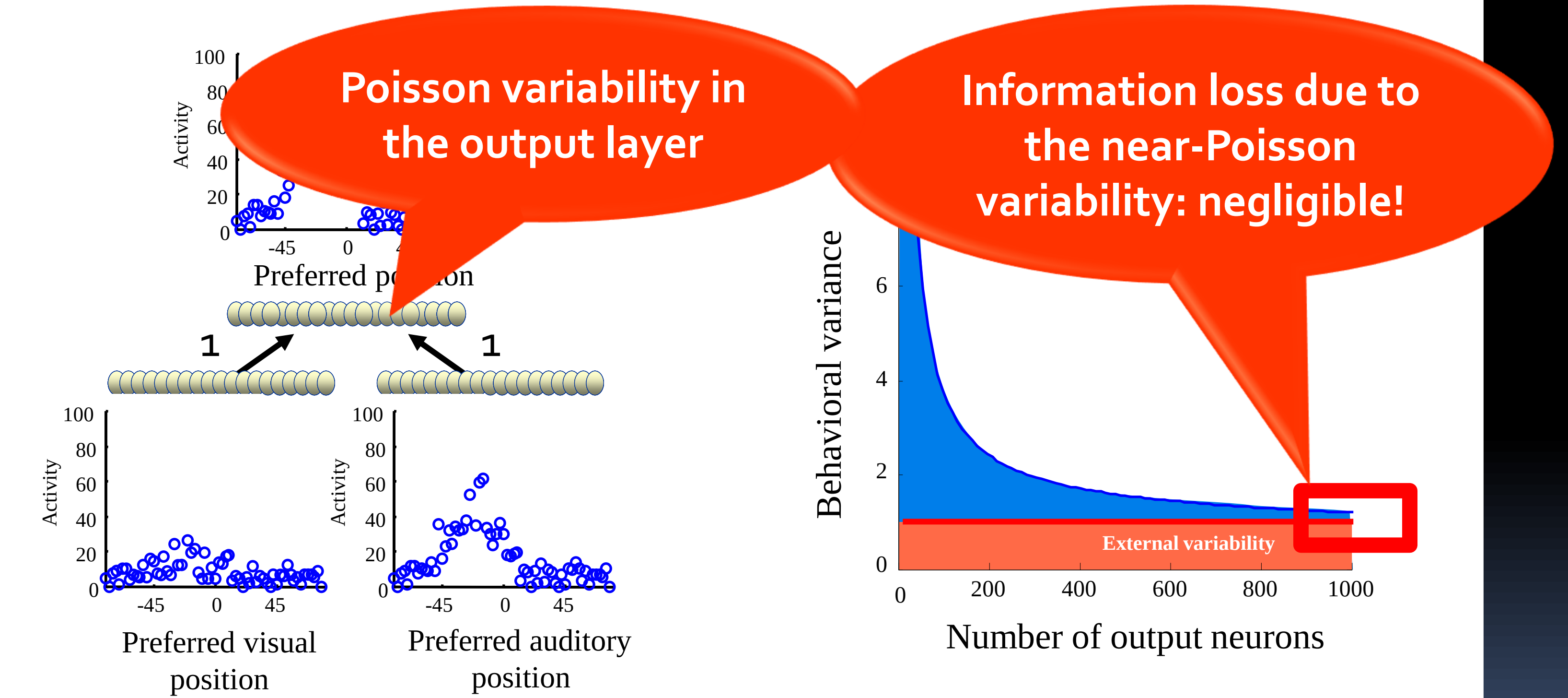
- For most problems of interest, we cannot know the generative model because it's too complex: suboptimality dominates.
- For very simple tasks, we might be able to learn the generative model (e.g. photon detection)

The General Case

- How about internal noise?



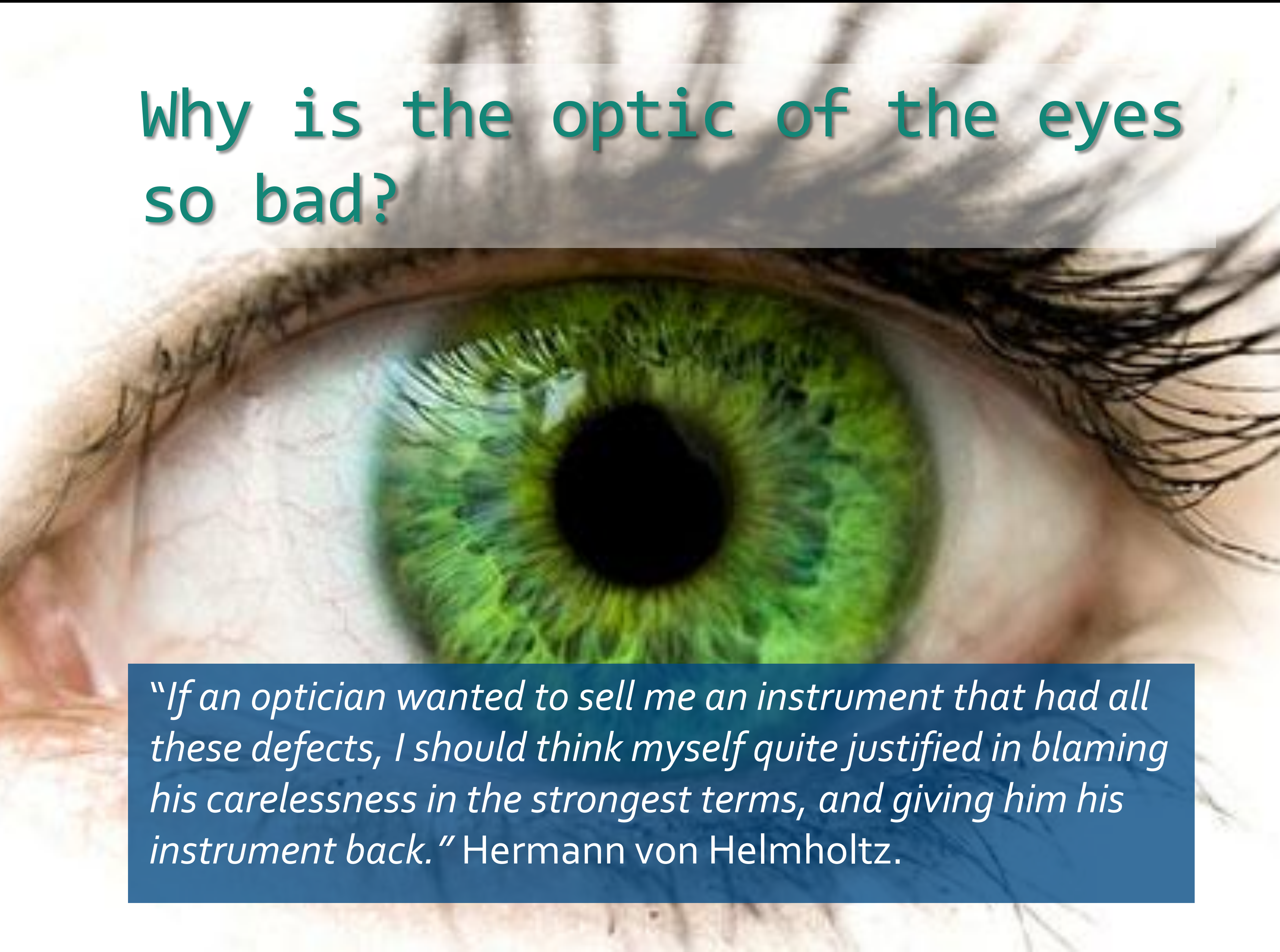
The impact of Poisson variability



Ma, Beck et al, 2006. Beck et al, 2012.

Roadmap

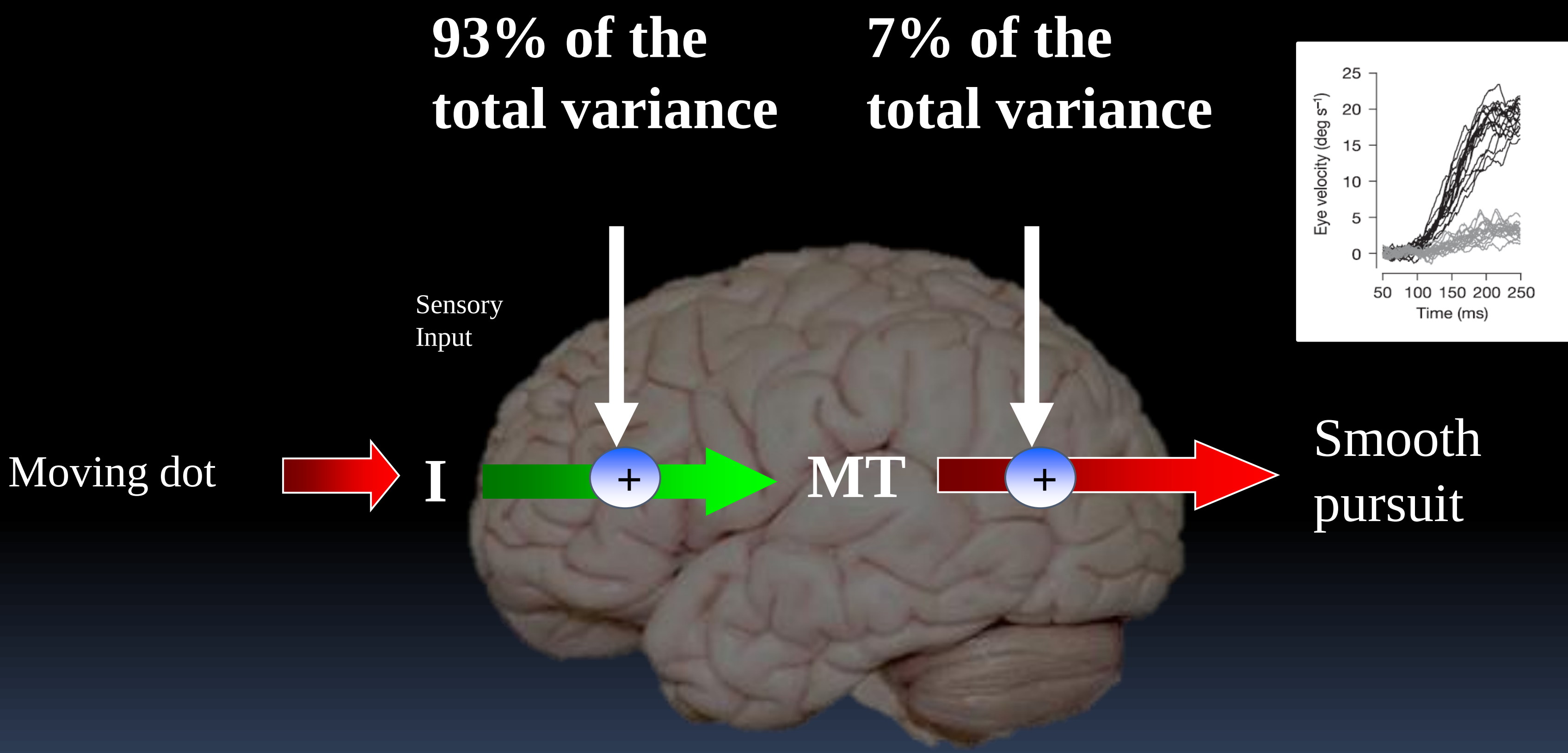
- Suboptimal inference can generate behavioral variability
- This cause dominates in most situations
- What this theory explains
- Implications for neuronal variability
- A normative view of Weber's law



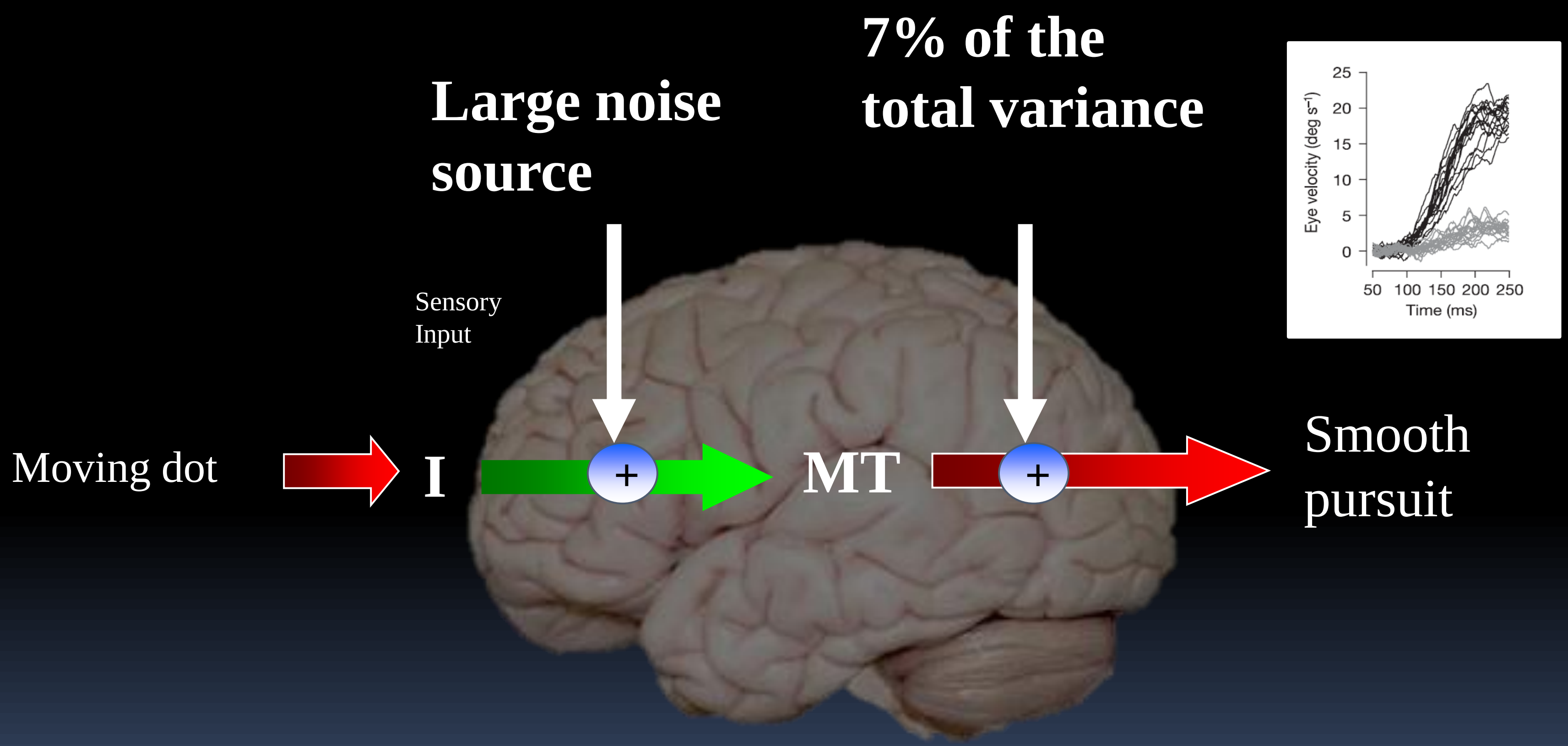
Why is the optic of the eyes so bad?

"If an optician wanted to sell me an instrument that had all these defects, I should think myself quite justified in blaming his carelessness in the strongest terms, and giving him his instrument back." Hermann von Helmholtz.

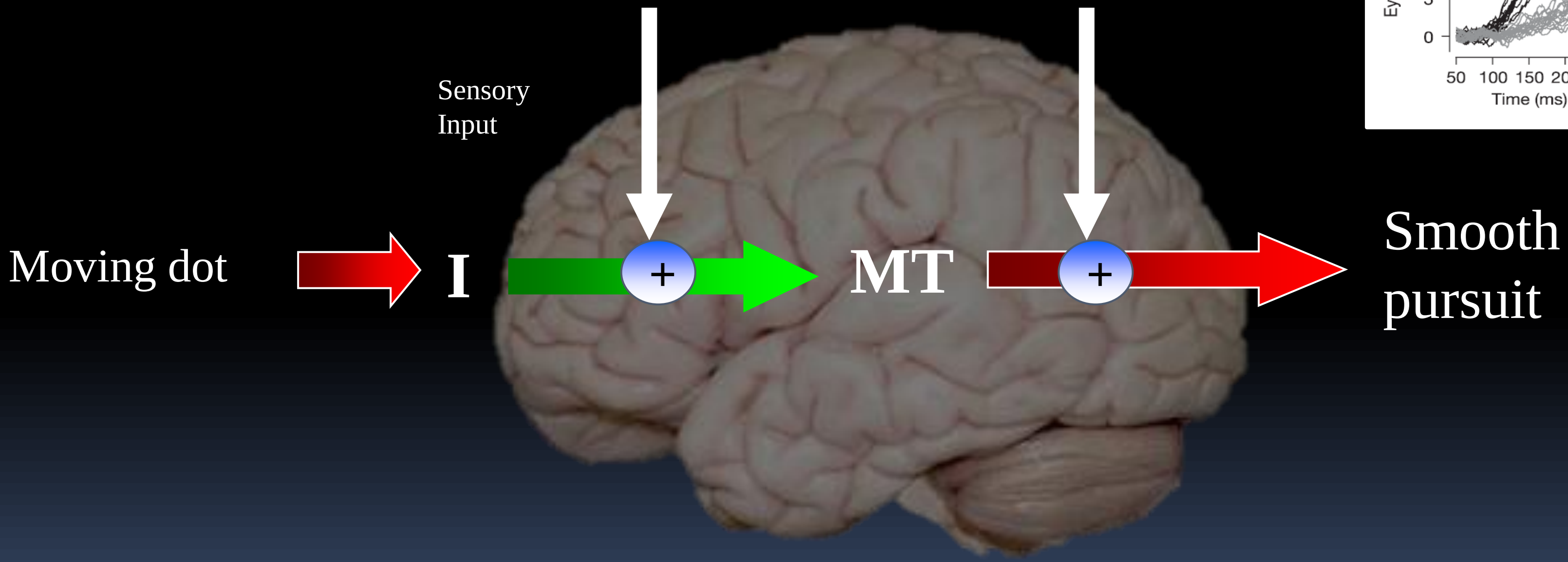
Osbourne et al, 2005

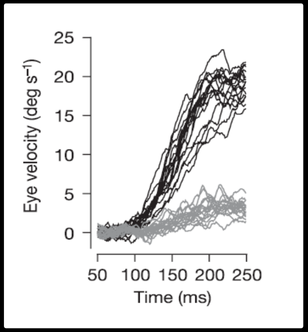


Osbourne et al, 2005

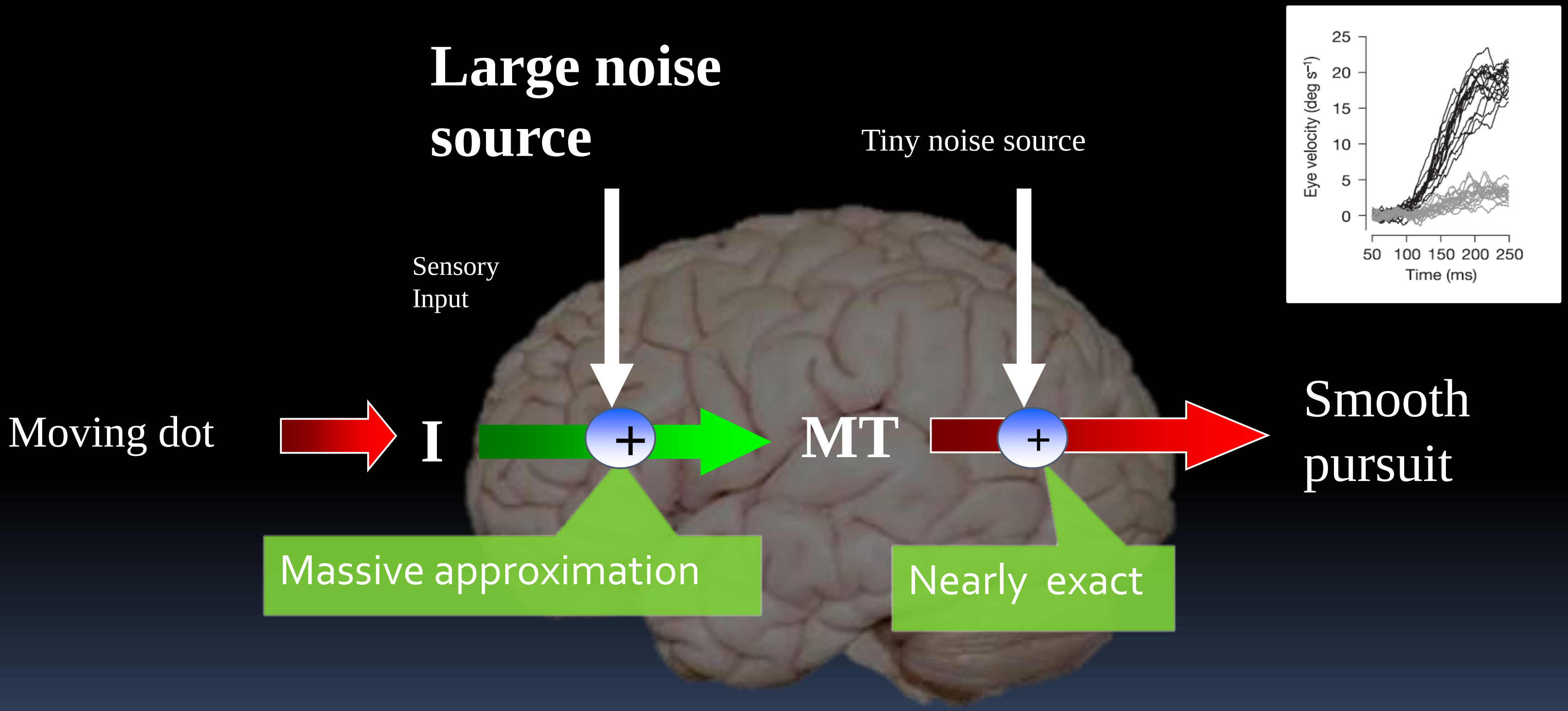


Osbourne et al, 2005





Osbourne et al, 2005



Roadmap

- Suboptimal inference can generate behavioral variability
- This cause dominates in most situations
- What this theory explains
- Implications for neuronal variability
- A normative view of Weber's law

Correlations and information

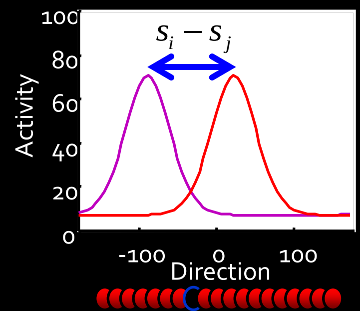
- Suboptimality increases behavioral variability.
- In other words, it decreases information.
- How is that reflected in neural responses?
- What limits information in neural codes?

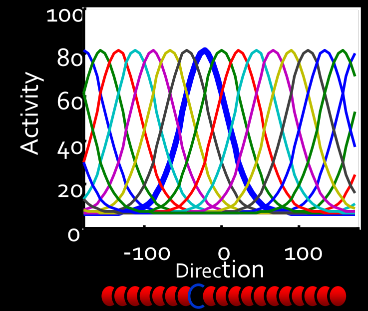
Correlations and information

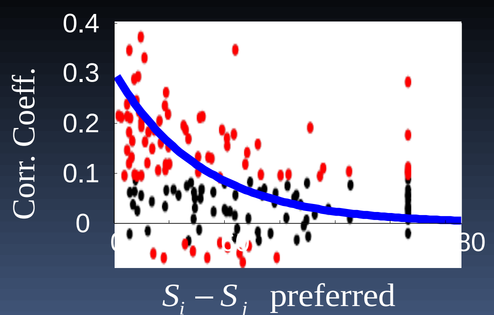
- What limits information in neural codes?
- Typical answer: positive correlations among neurons with similar preferred stimuli

Who is to blame?

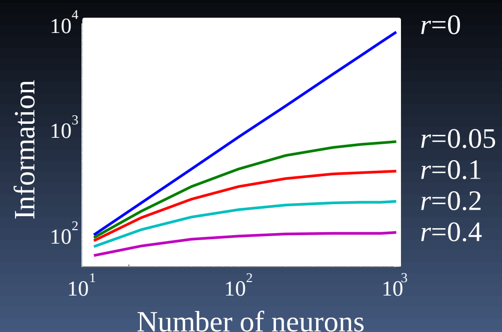
- It's not just Poisson variability, it's correlated Poisson variability







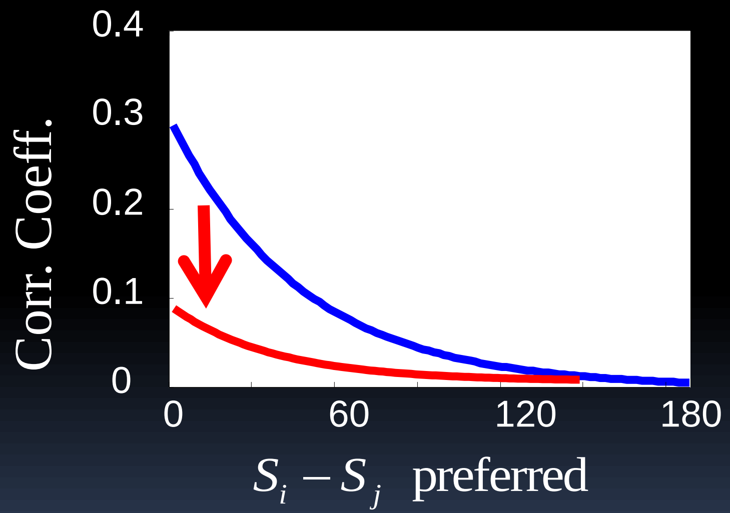
Huang and Lisberger, 2010

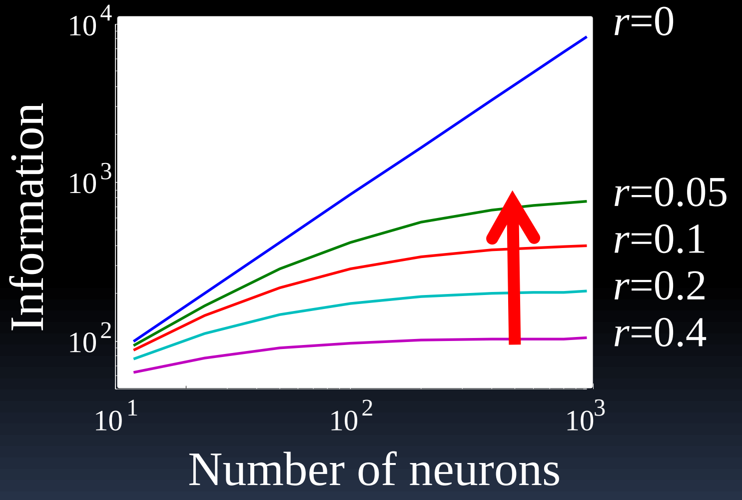


Zohary et al. 1994

Correlations and information

- Decorrelation = more information

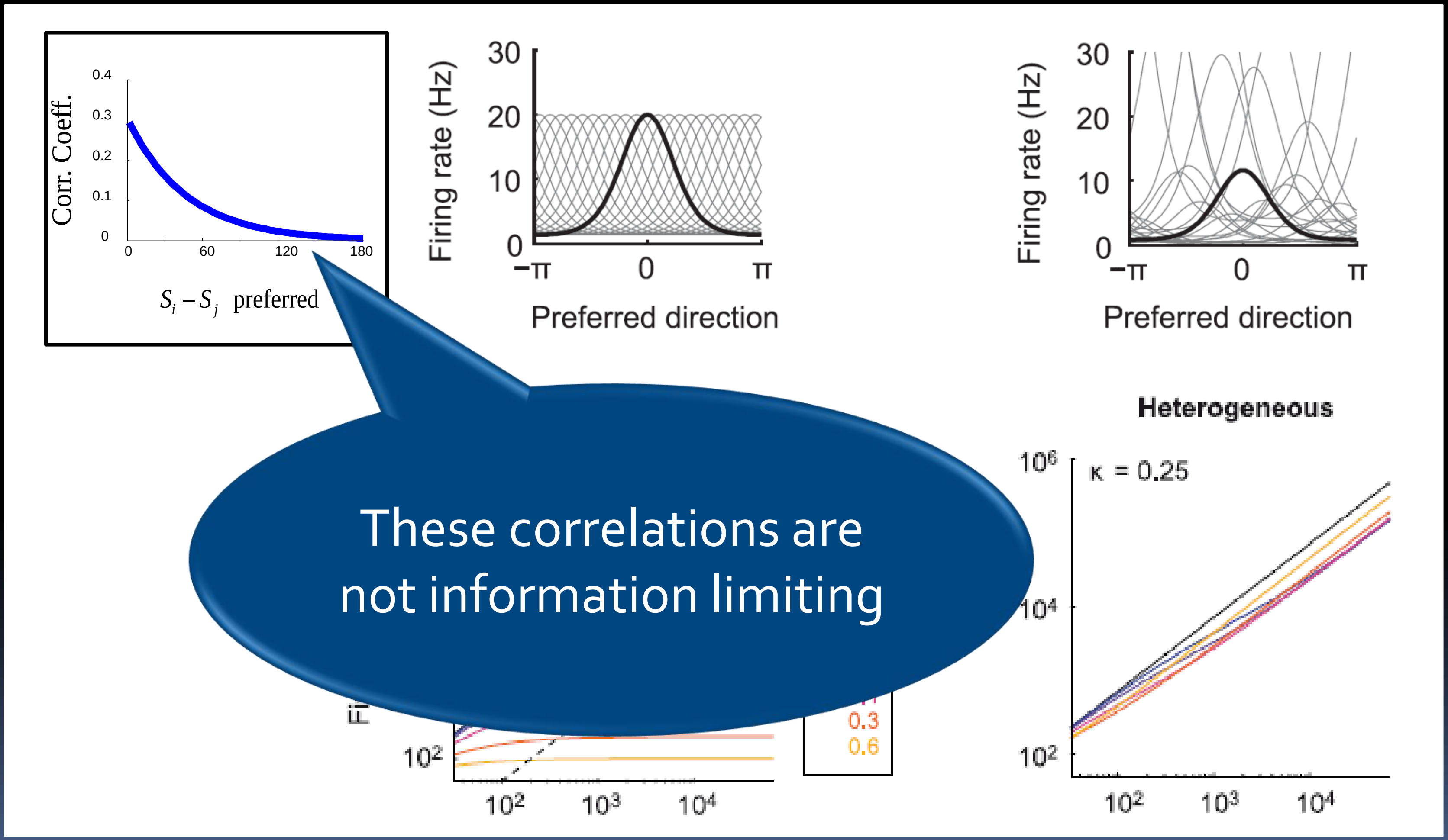




Correlation and information

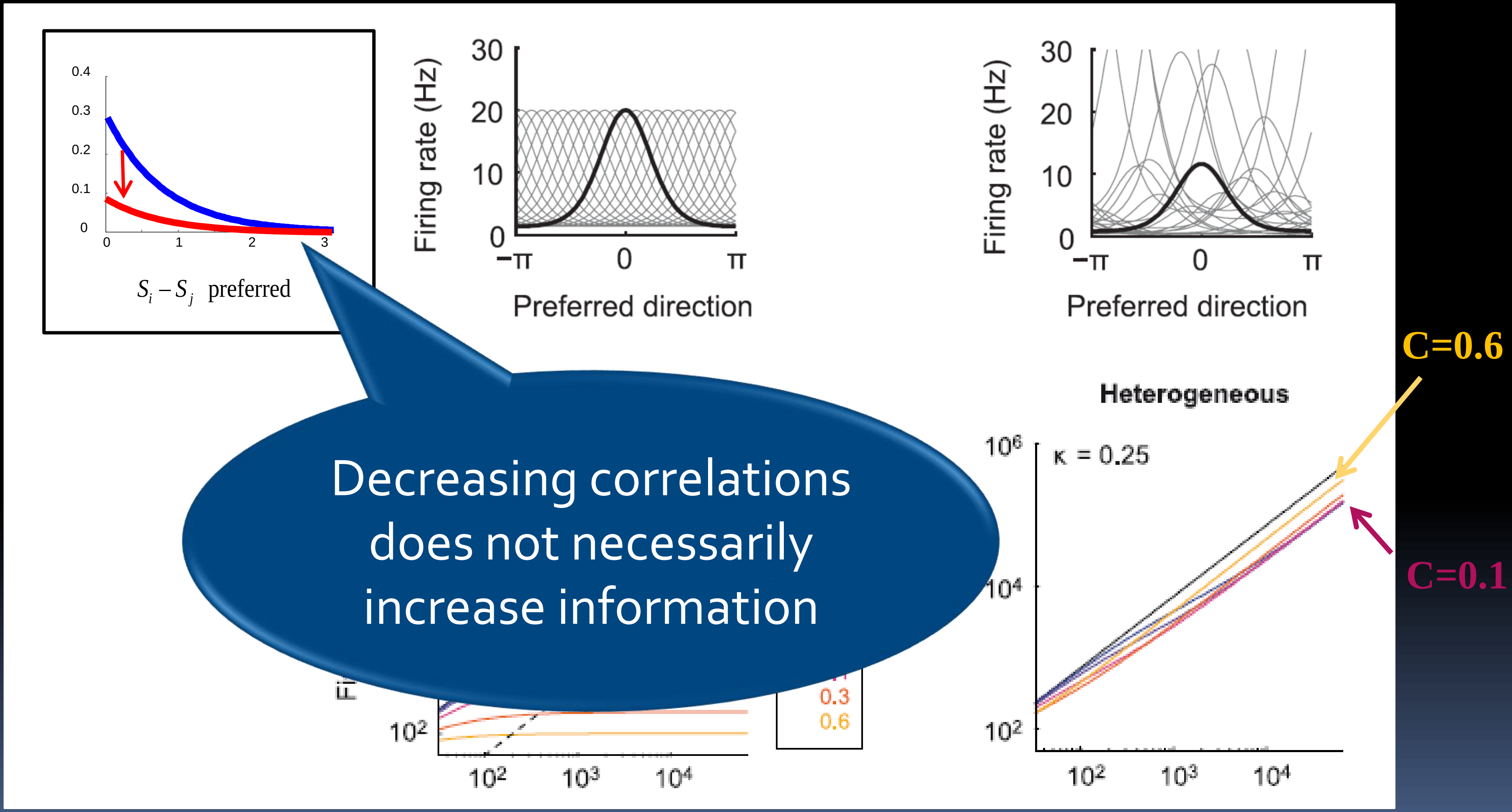
- Correlations inversely proportional to difference in preferred stimuli limit information
- Decorrelation = more information

- No, not necessarily.



Shamir and Sompolinsky, 2004. Ecker et al, 2012.

(See aso Series, Latham and Pouget. 2004.)



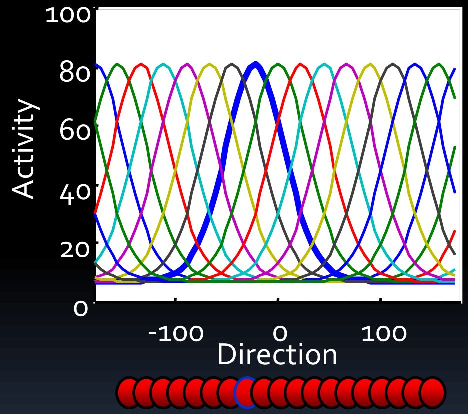
Shamir and Sompolinsky, 2004. Ecker et al, 2012.

Implications for Neural Coding

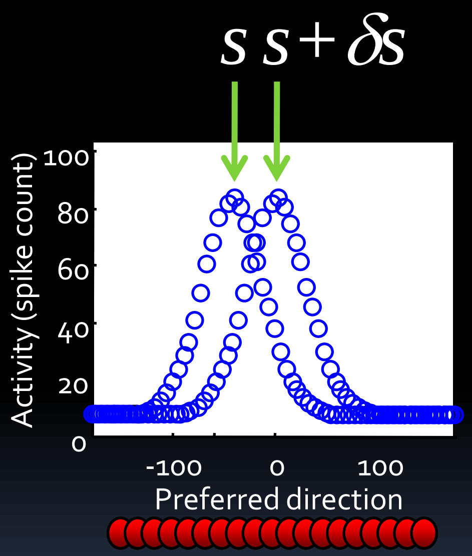
- What limits information in neural codes?
- Differential correlations

Differential correlations

Discrimination task

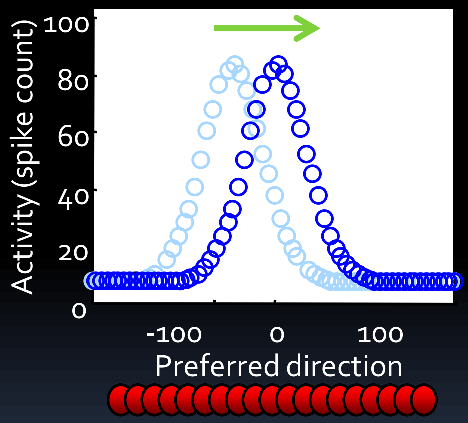


Differential correlations

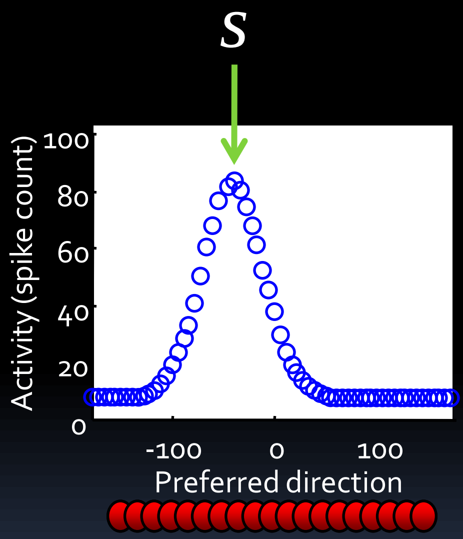


Differential correlations

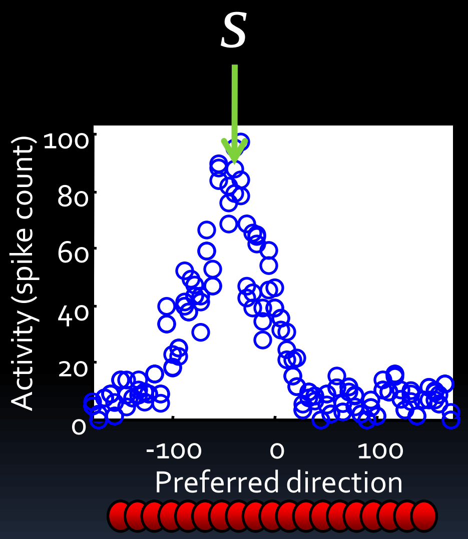
$$s \longrightarrow s + \delta s$$



Differential correlations

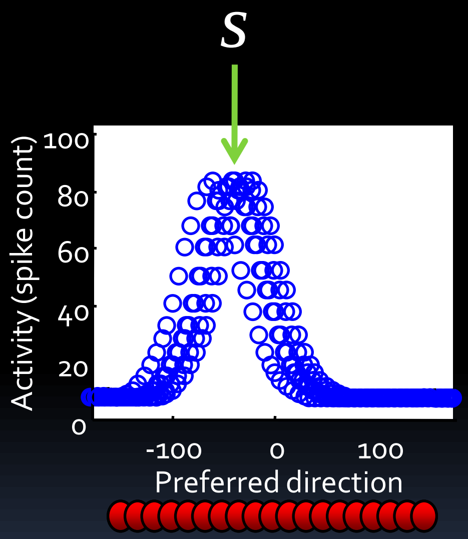


Differential correlations

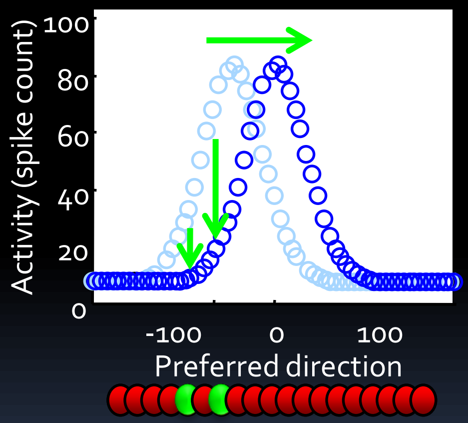


Differential correlations

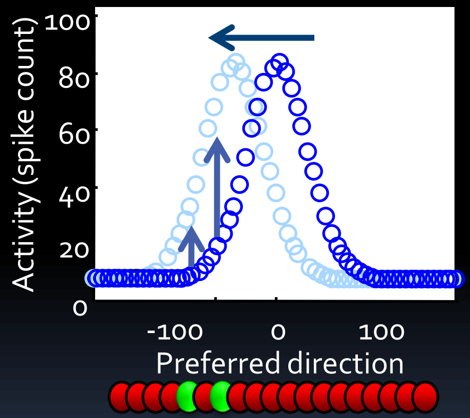
Differential correlations



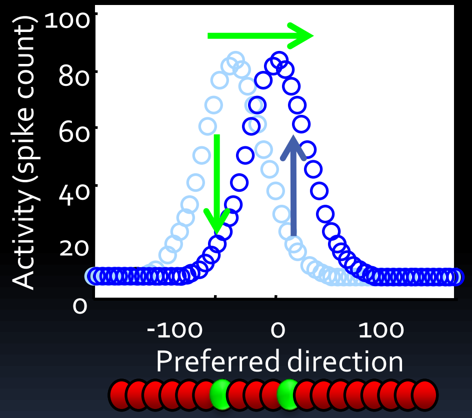
Differential correlations



Differential correlations

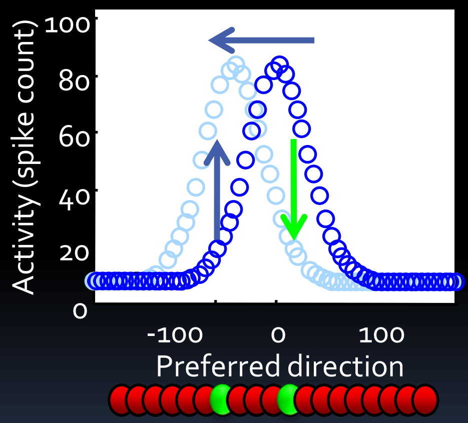


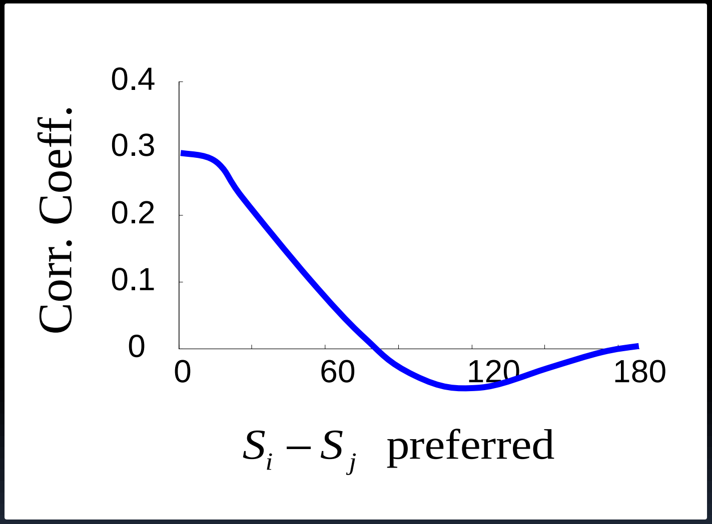
Differential correlations



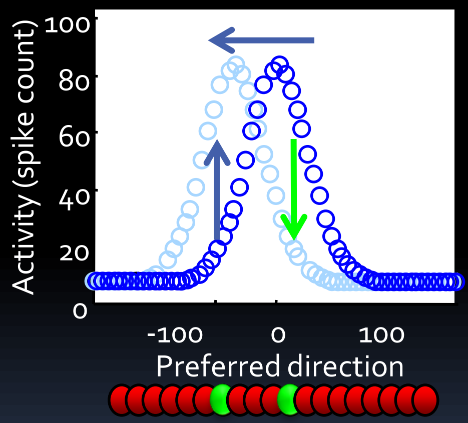
Differential correlations

$$\mathbf{f'f'}^T$$

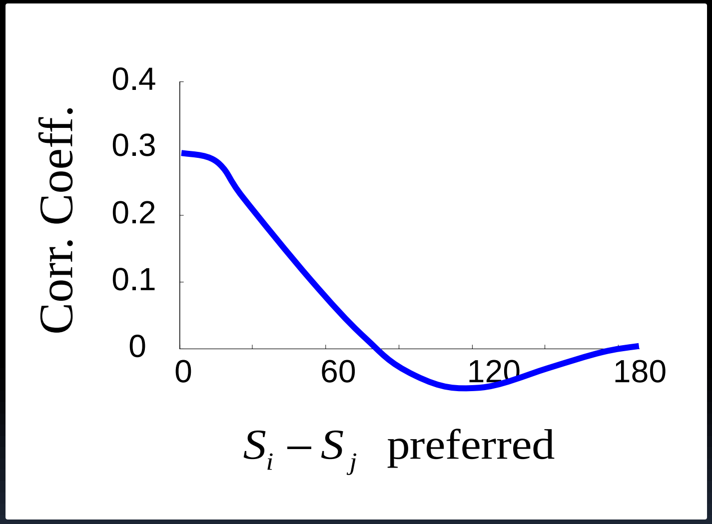




Differential correlations



Differential

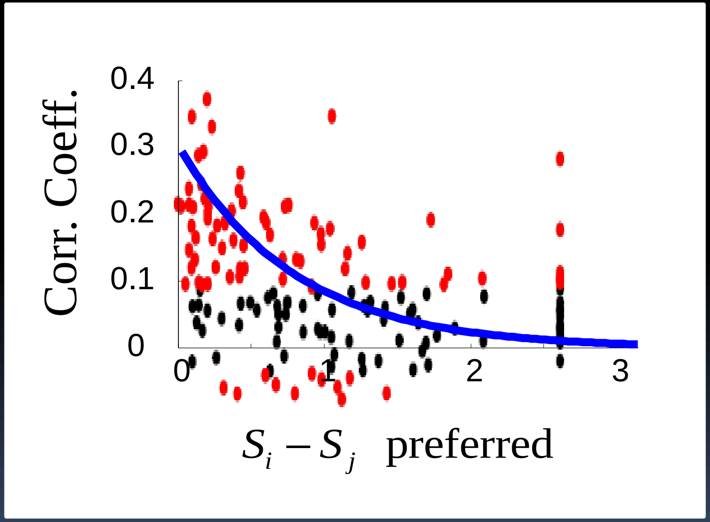


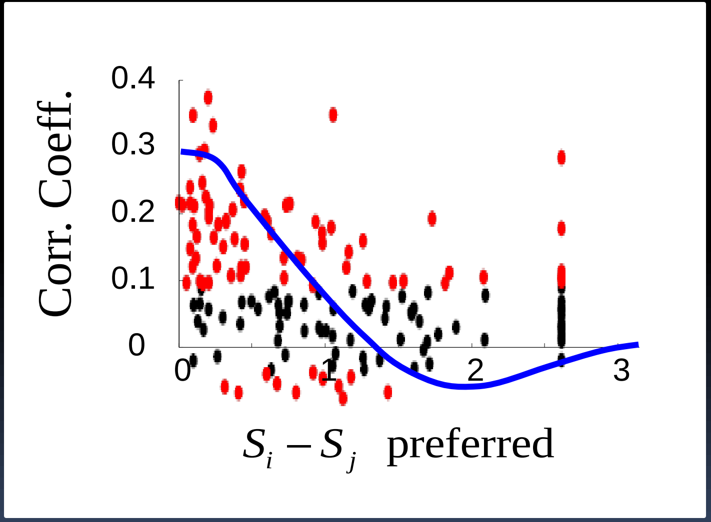
Differential correlations

- Information saturates because of differential correlations.

- But wait, nobody has ever found differential correlations.

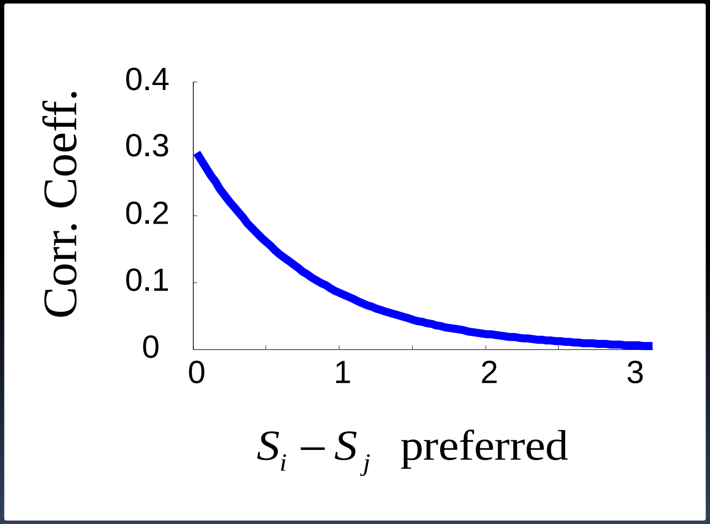
Differential

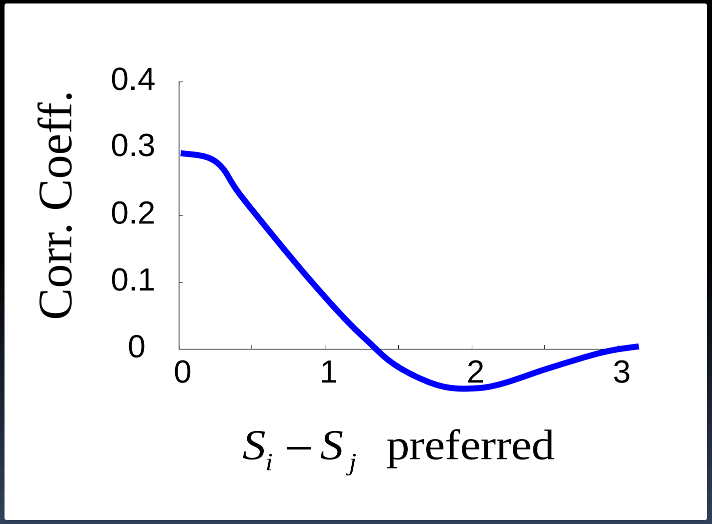




Huang and Lisberger, 2010

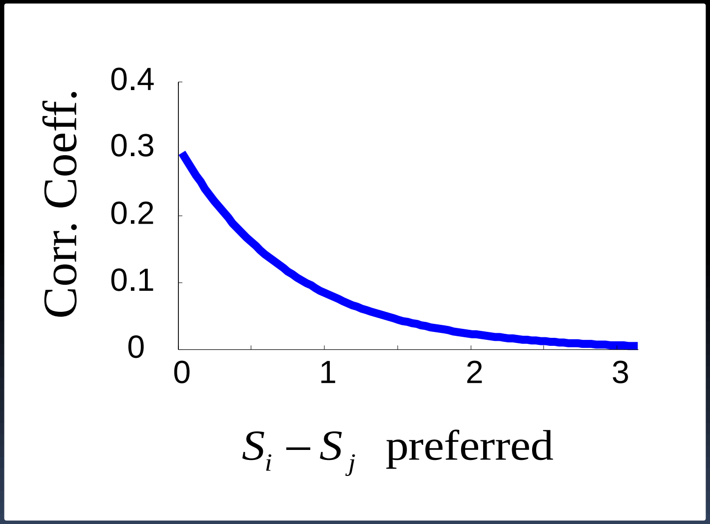
Differential

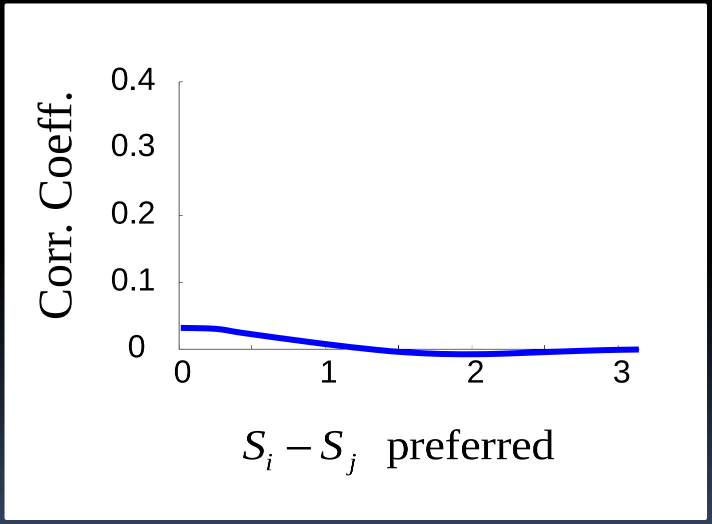




Huang and Lisberger, 2010

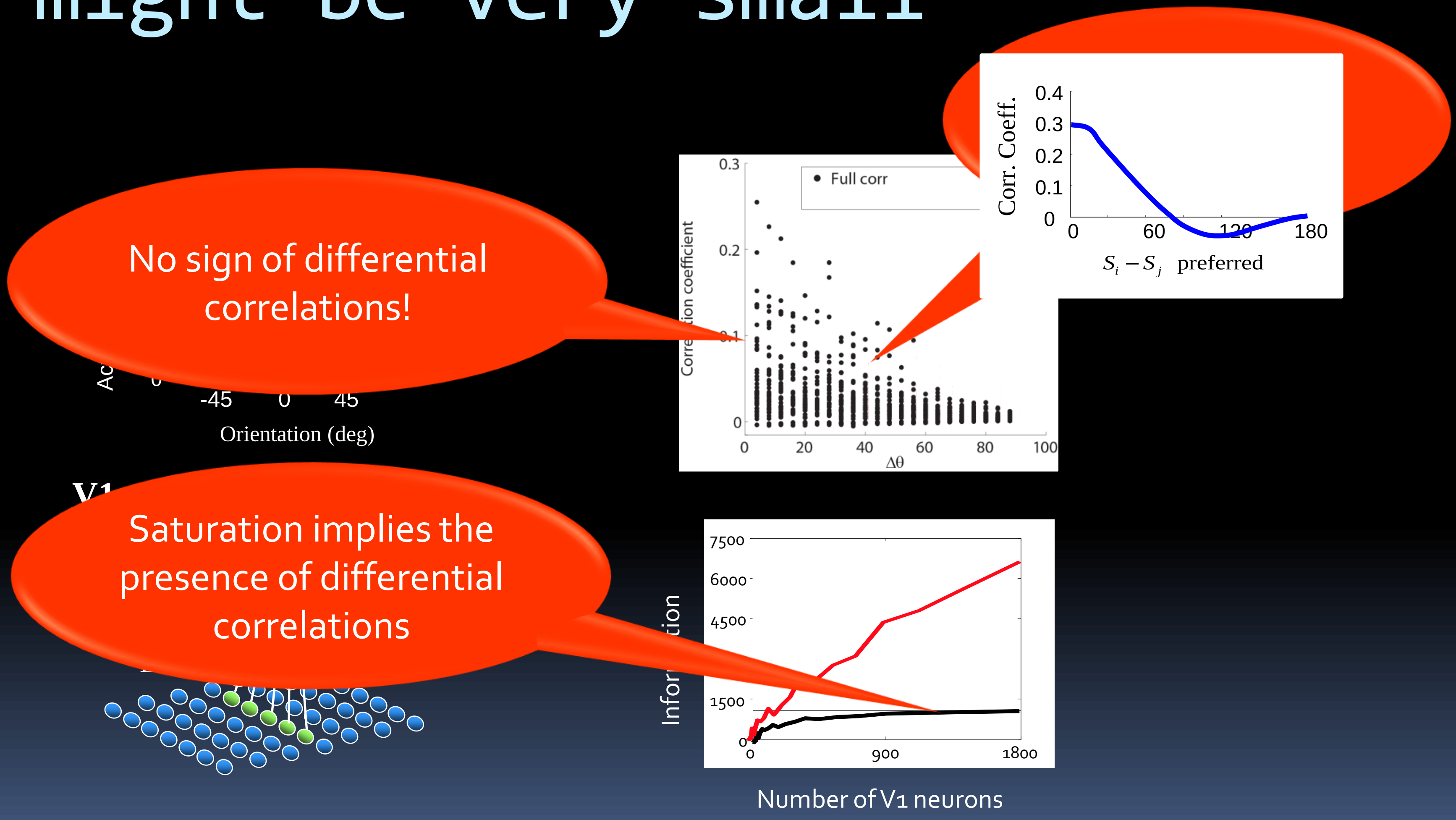
Differential





Huang and Lisberger, 2010

Differential correlations might be very small



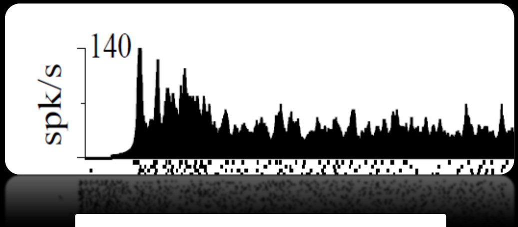
Information and correlations

- Information is limited by differential correlations
- Suboptimality increases differential correlations

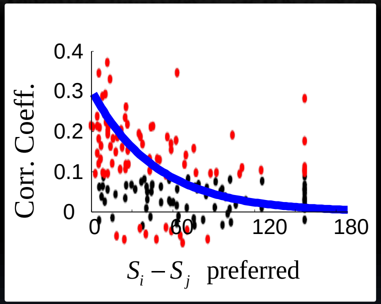
Alternative view

The brain is noisy but…

- The near-Poisson noise induced by chaotic dynamics of cortical circuits has little impact on behavior



- Correlations inversely proportional the difference in preferred stimuli do not necessarily limit information



- Most of behavioral variability comes from

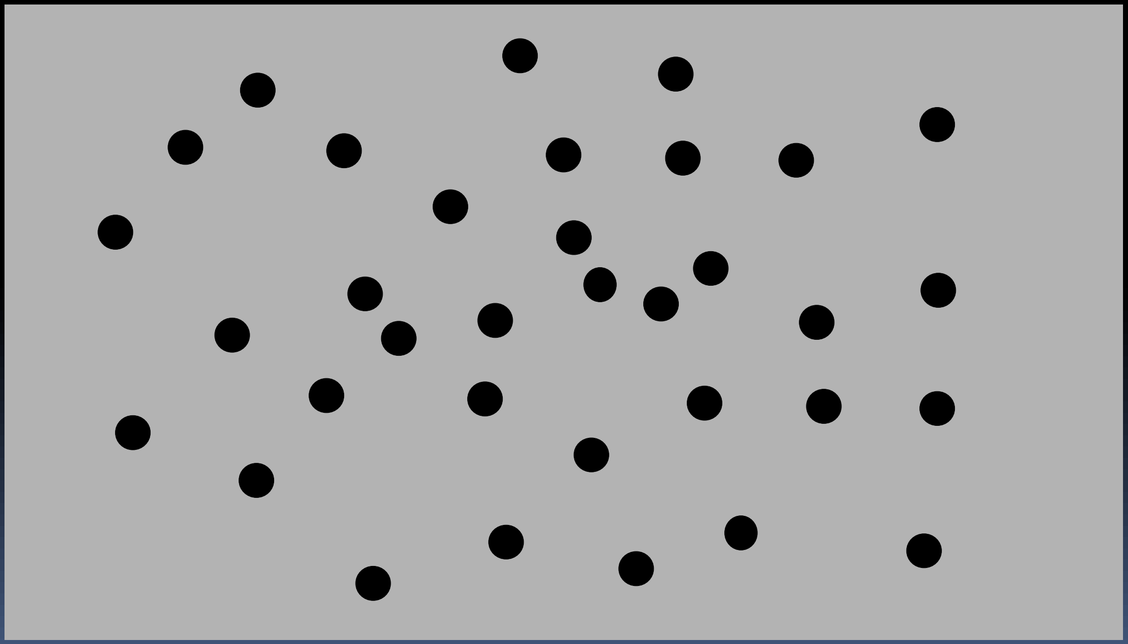
1. Variable data from the world (which naturally leads to Weber's law)
2. Suboptimal inference

Roadmap

- Suboptimal inference can generate behavioral variability
- This cause dominates in most situations
- What this theory explains
- Implications for neuronal variability
- A normative view of Weber's law

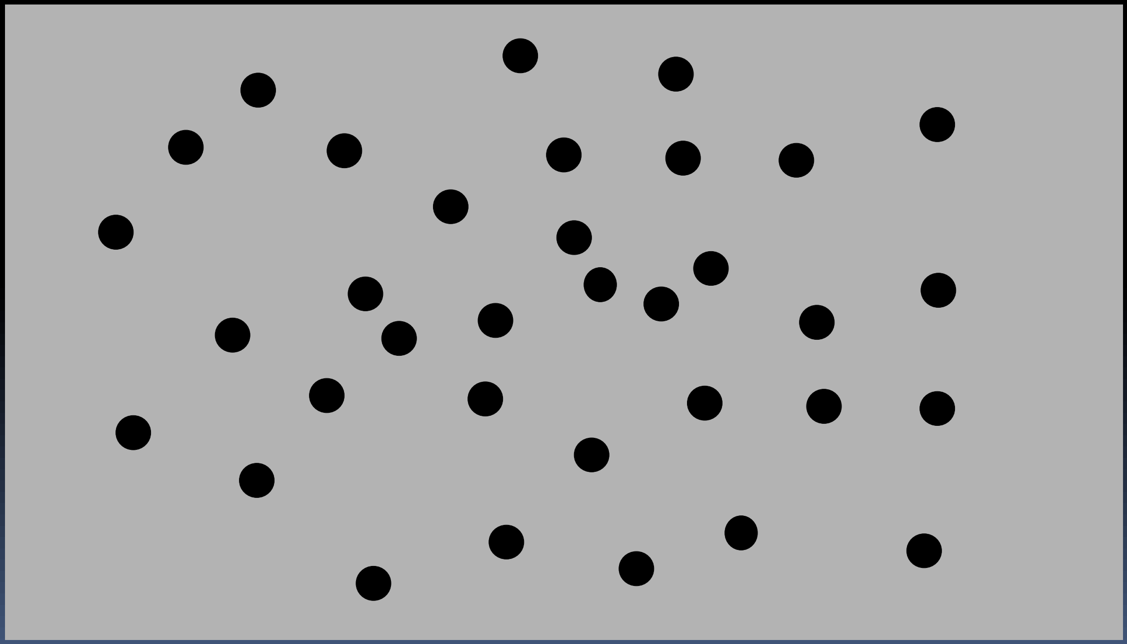
Weber's law

$$\frac{\Delta N}{N} = \text{constant}$$

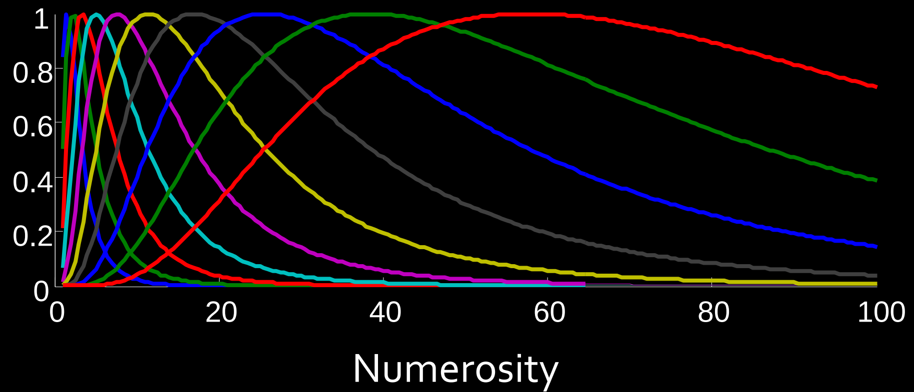


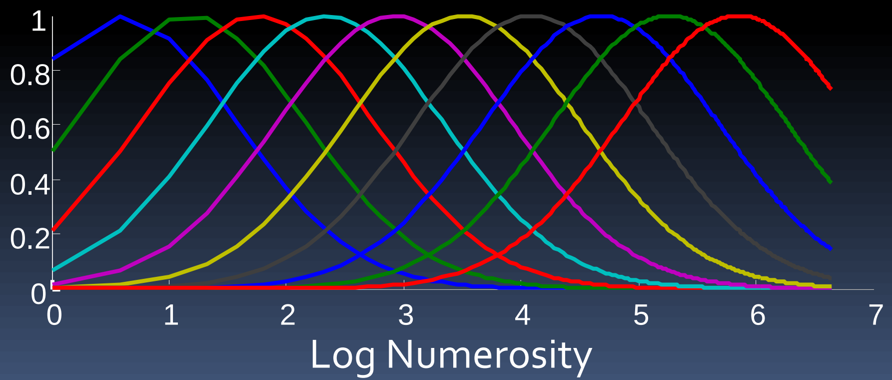
Weber's law

$$\sigma_N^2 \propto N^2$$



Weber’s law

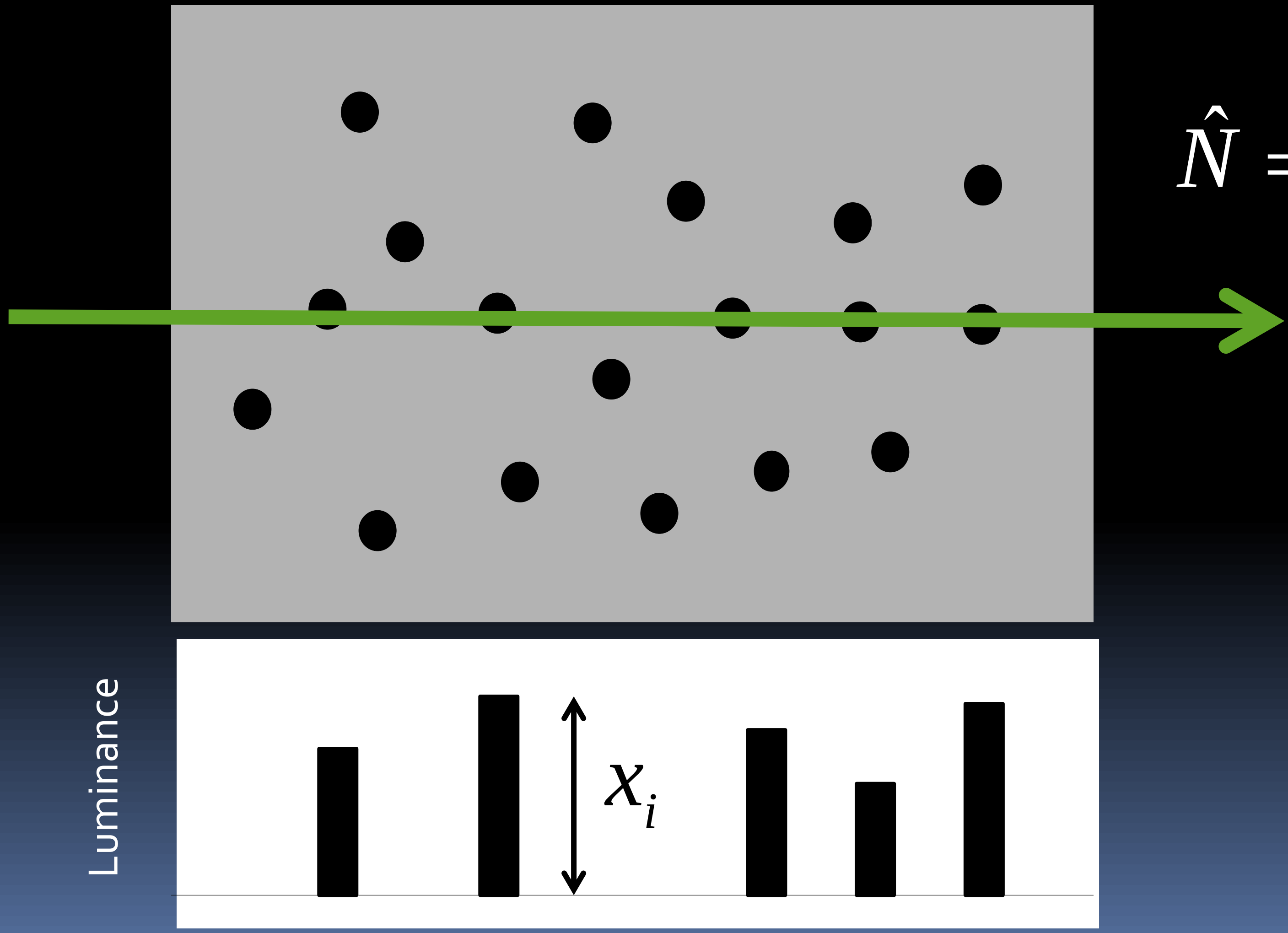




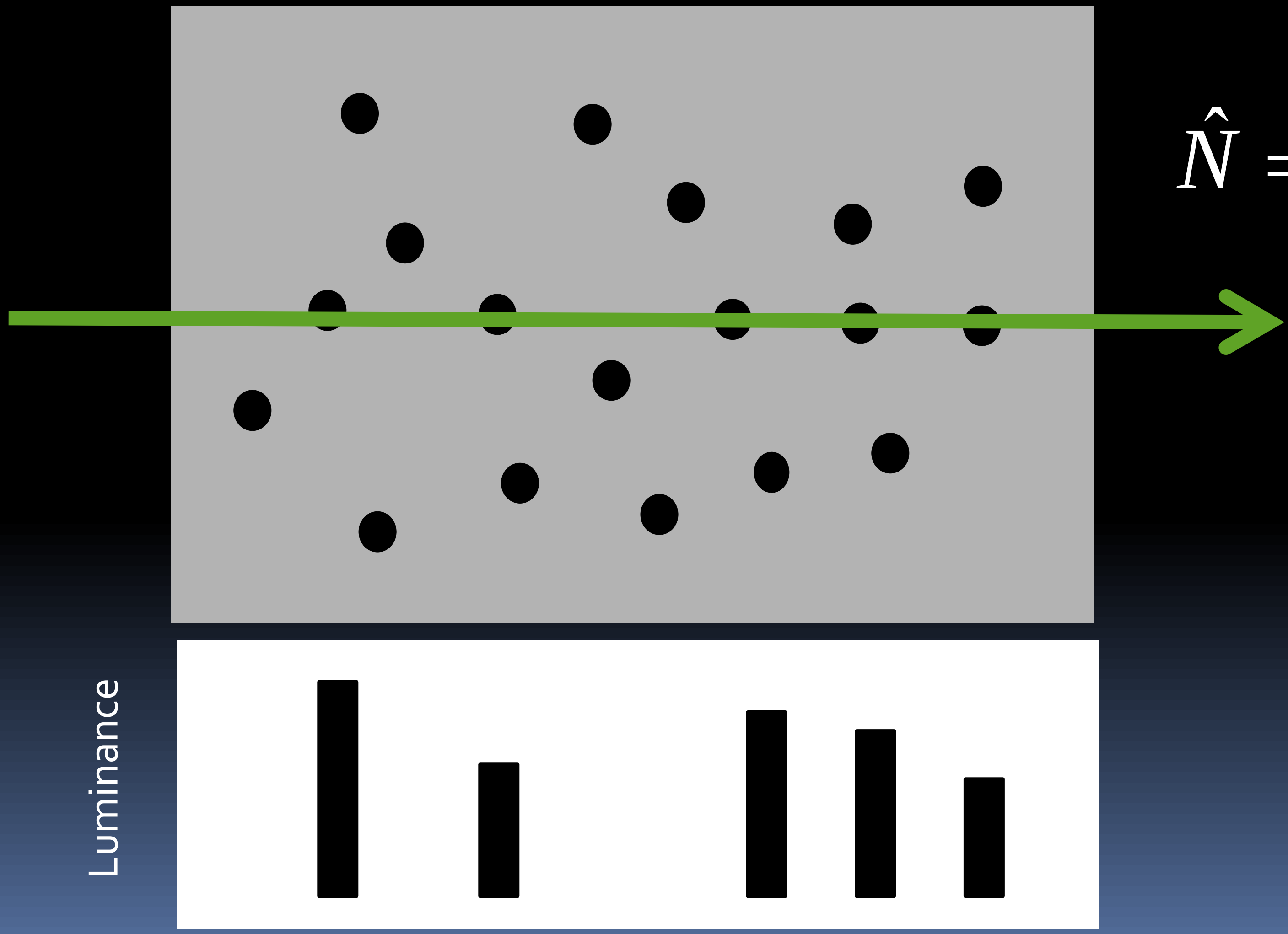
+ noise ➡ Weber’s law

Weber's law

$$\hat{N} = \sum_{i=1}^{N} x_i$$

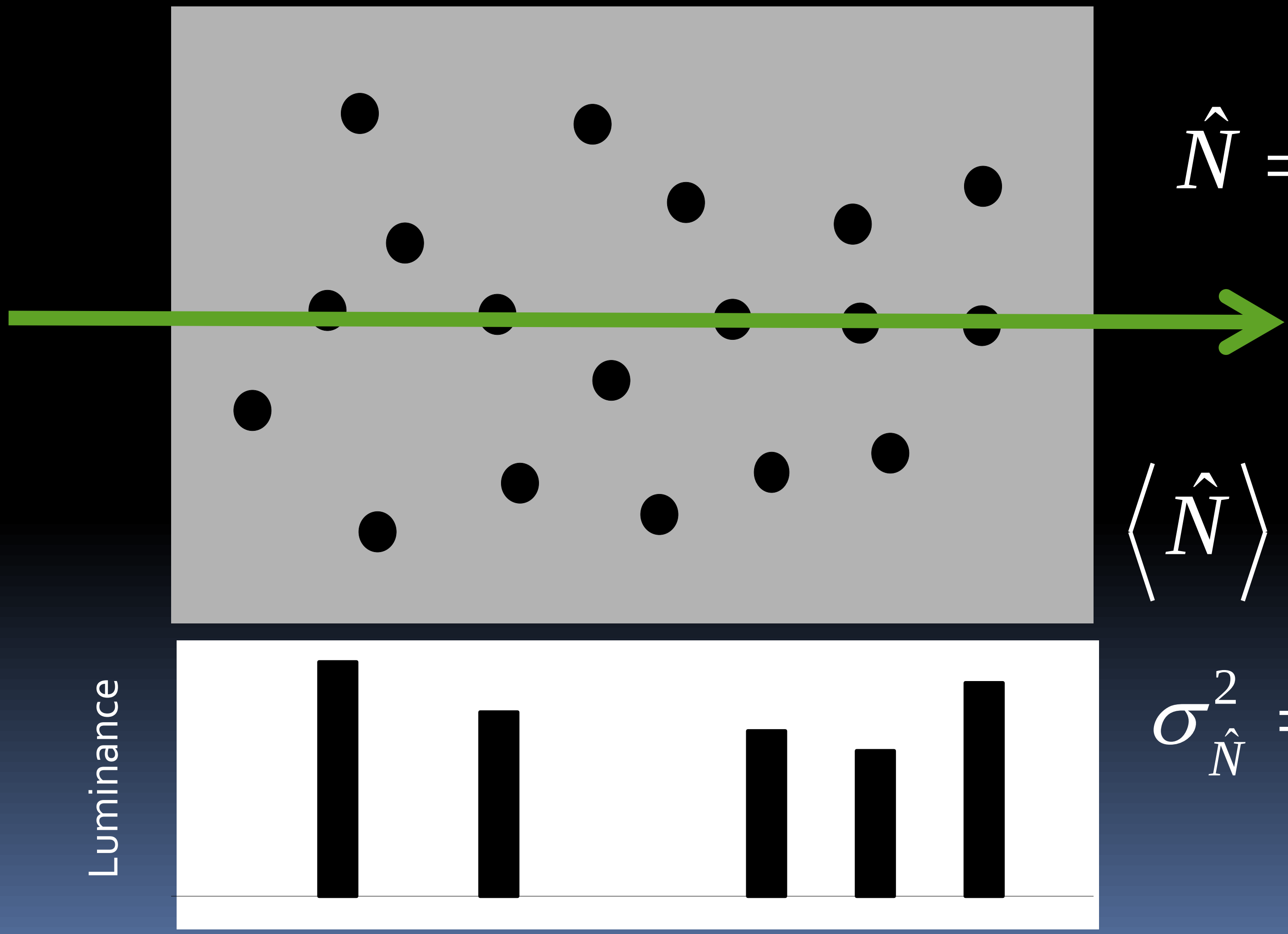


Weber's law



$$\hat{N} = \sum_{i=1}^{N} x_i$$

Weber's law

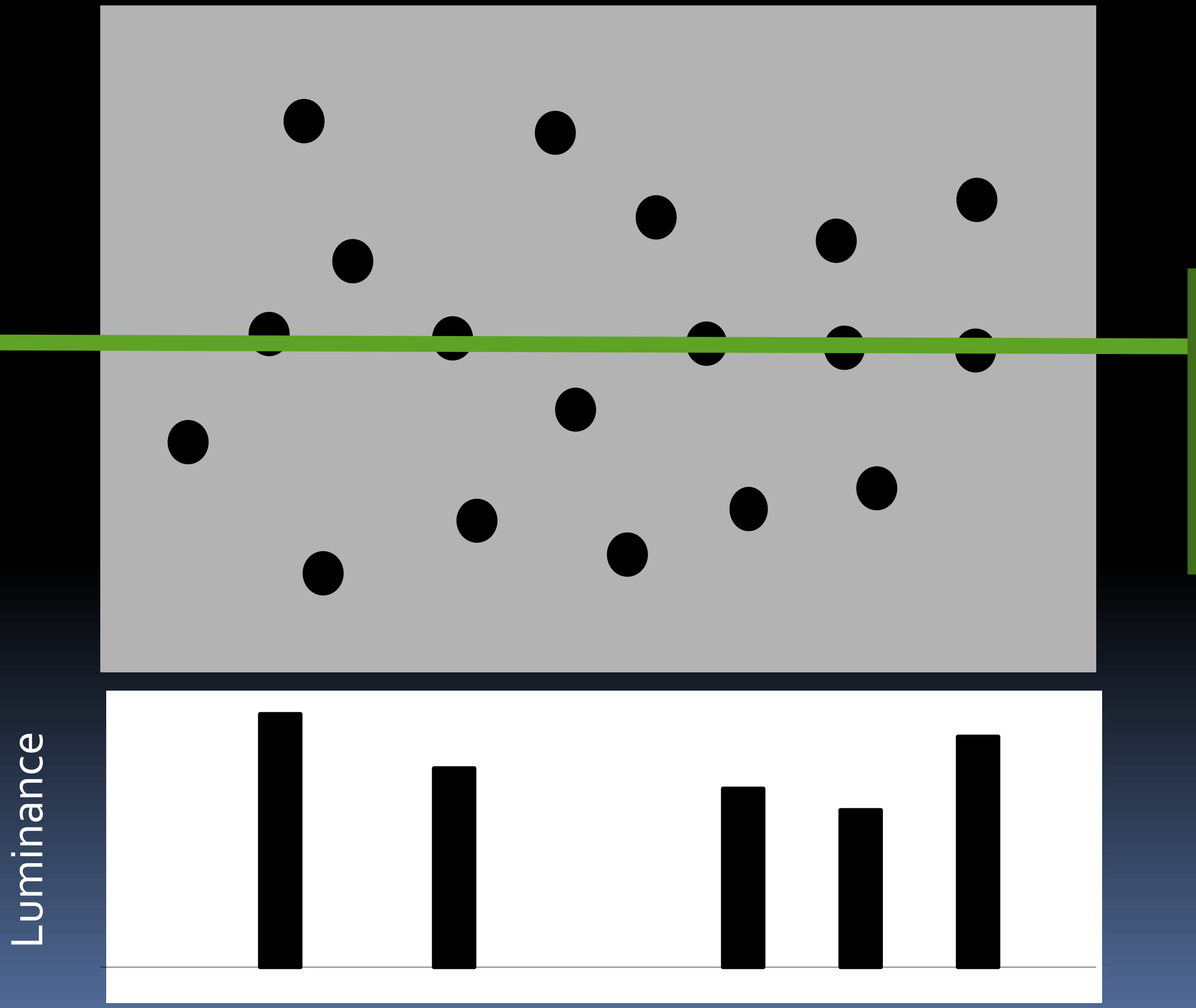


$$\hat{N} = \sum_{i=1}^{N} x_i$$

$$\left\langle \hat{N} \right\rangle \propto N$$

$$\sigma_{\hat{N}}^2 = \sum_{i=1}^{N} \sigma_{x_i}^2 \propto N$$

Weber's law

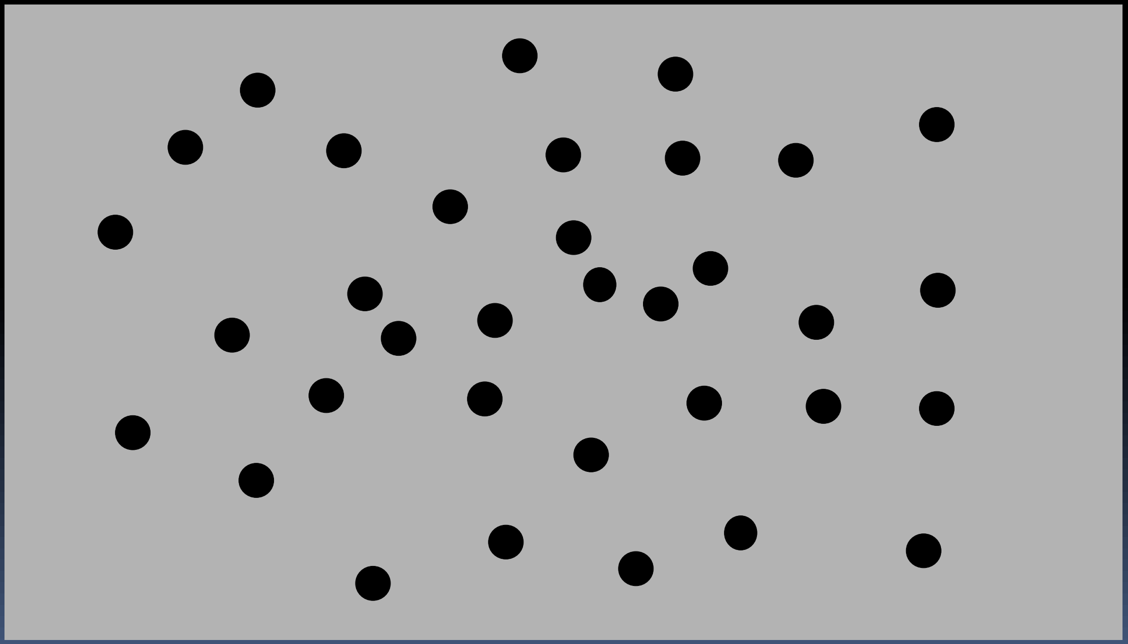


NOT
Weber's law!

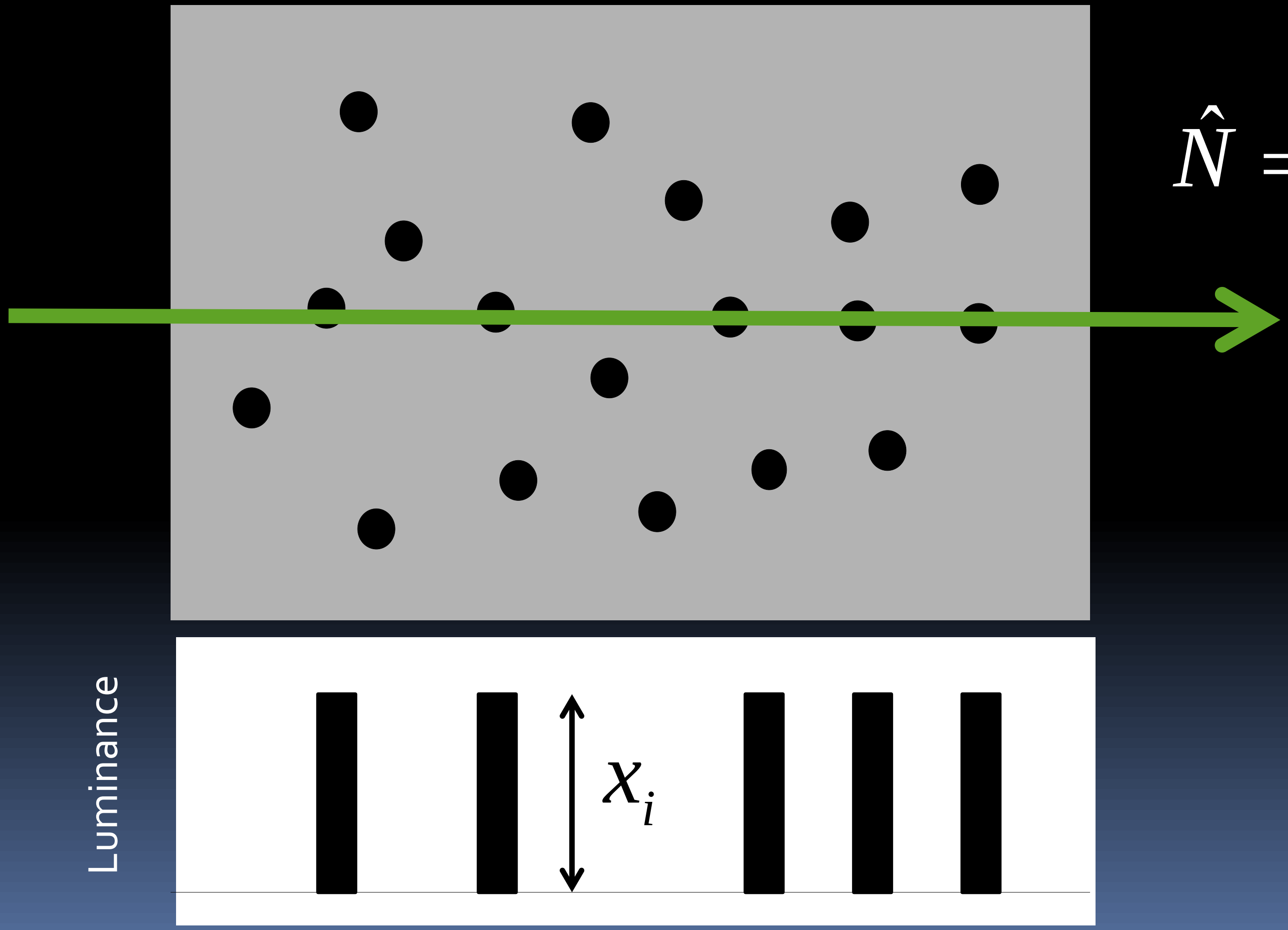
$$\sigma^2_{\hat{N}} \propto N$$

Weber's law

$$\sigma_N^2 \propto N^2$$

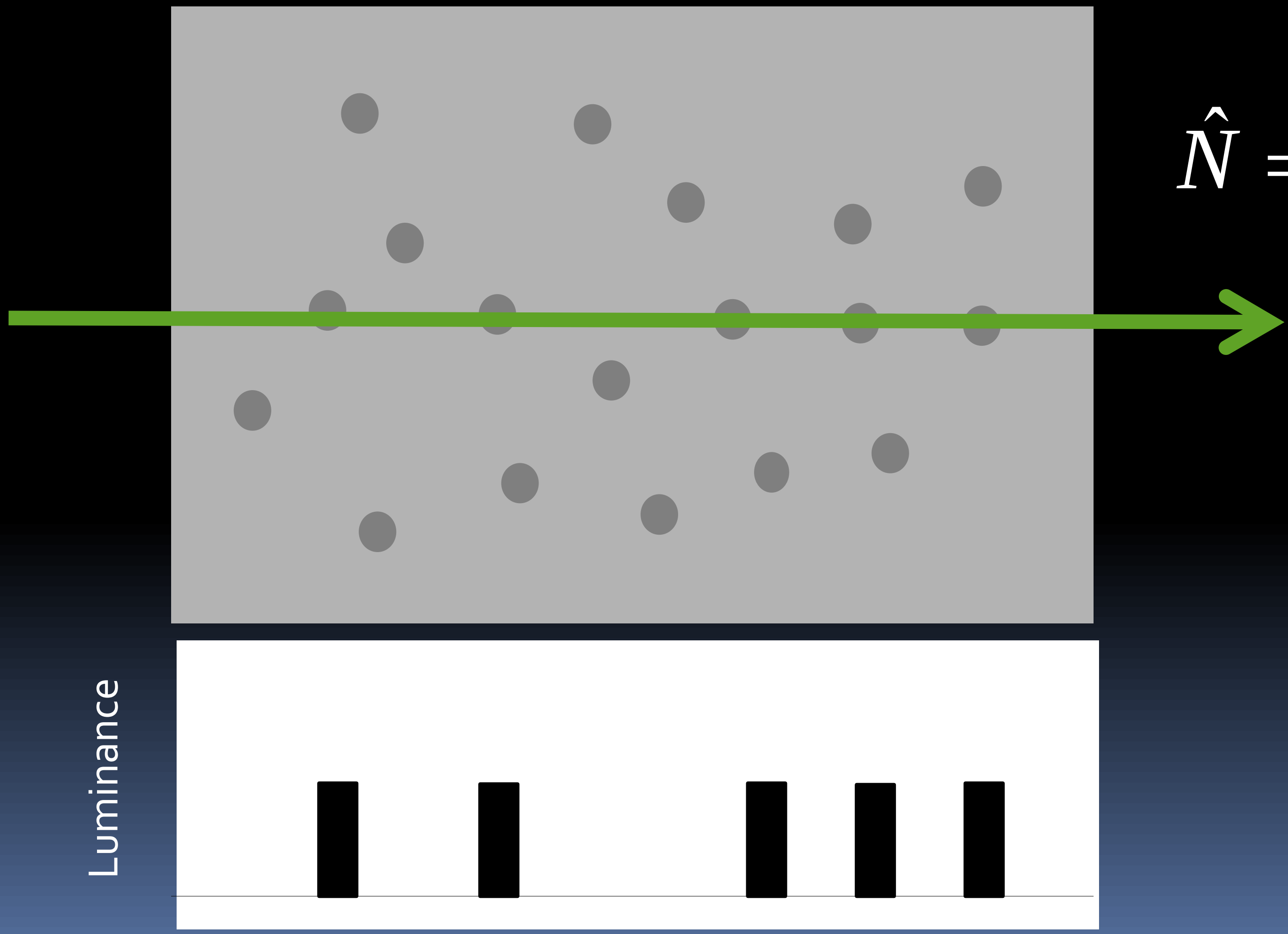


Weber’s law



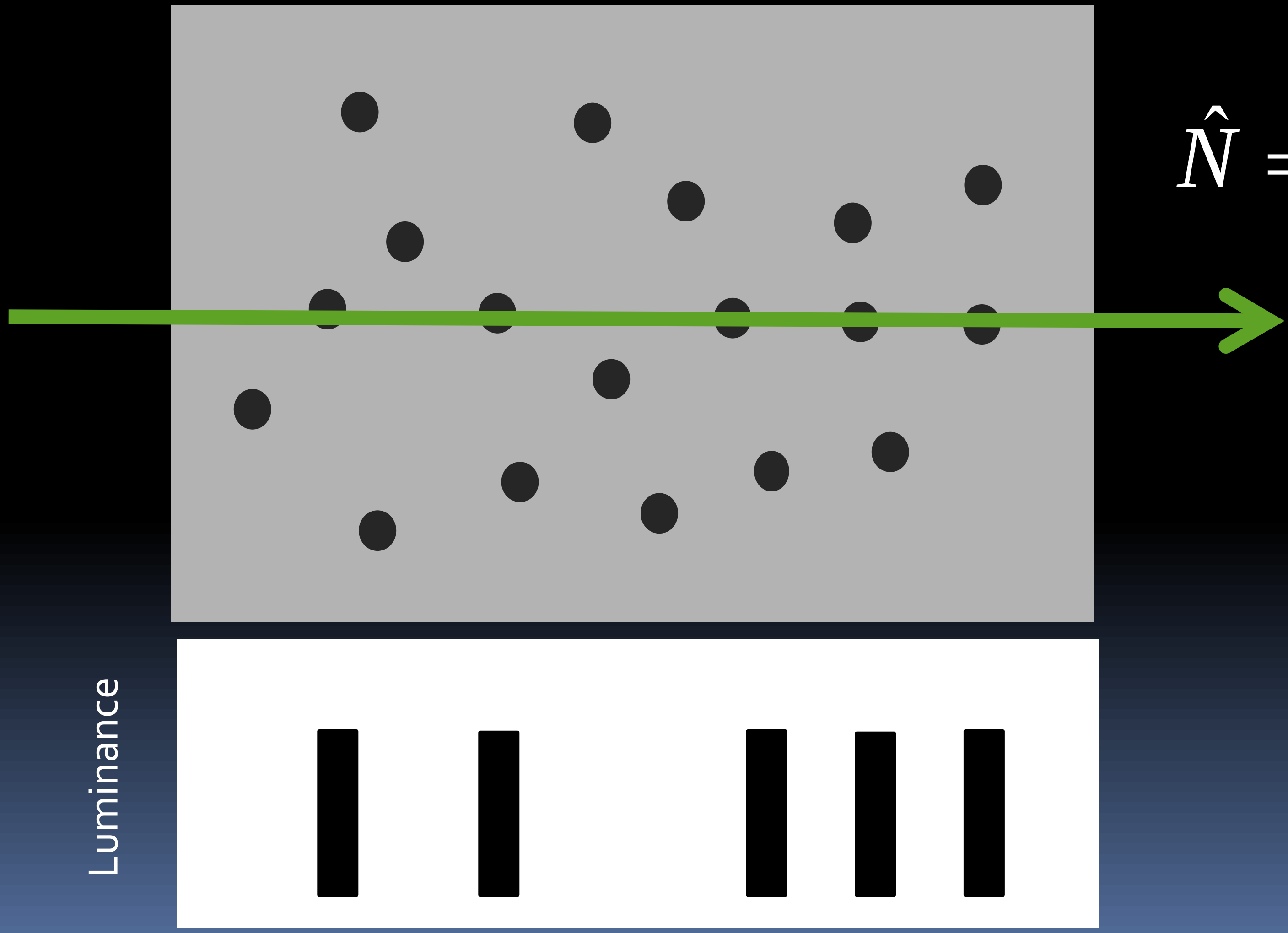
$$\hat{N} = \sum_{i=1}^{N} x_i$$

Weber's law



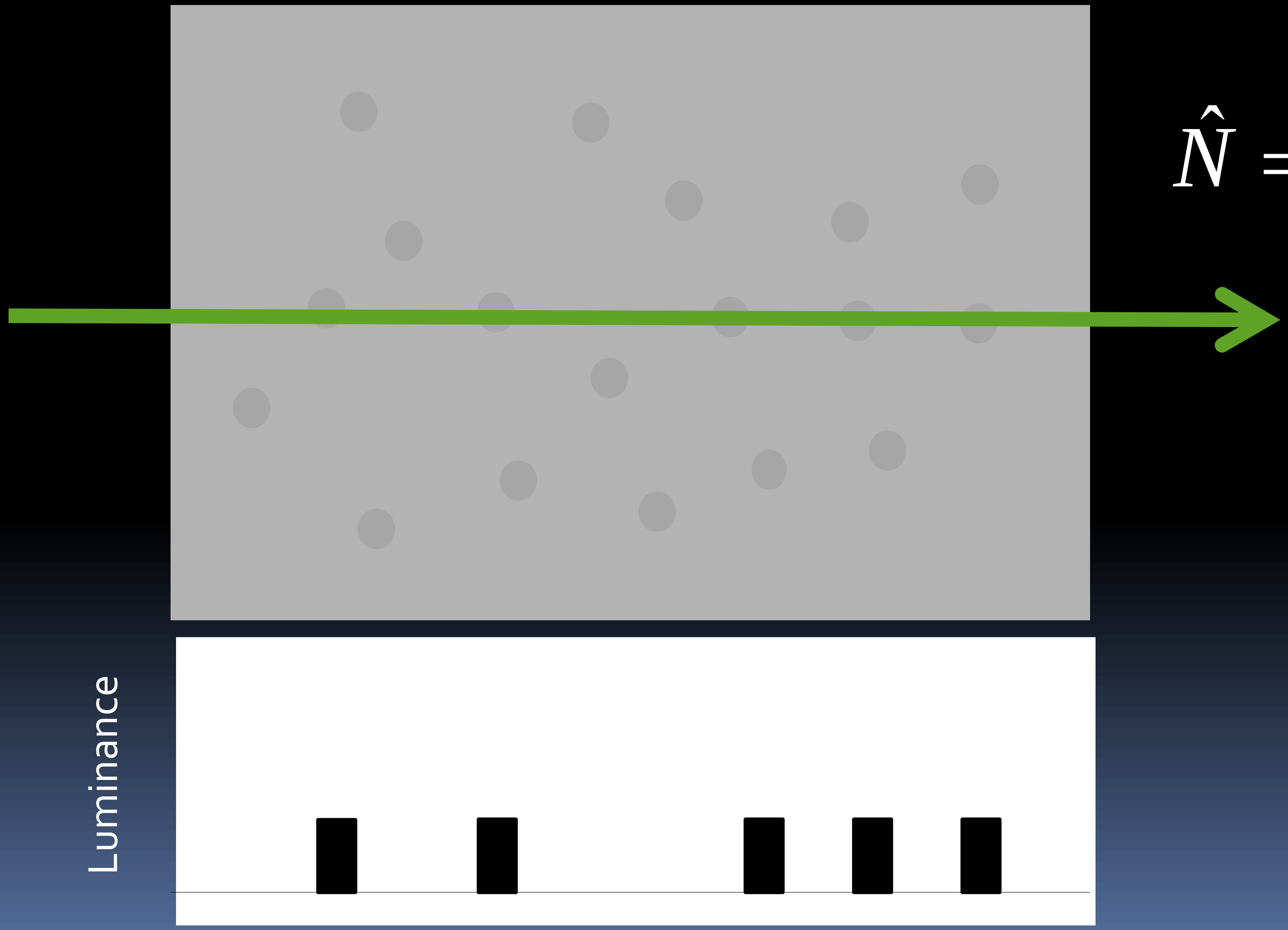
$$\hat{N} = \sum_{i=1}^{N} x_i$$

Weber’s law



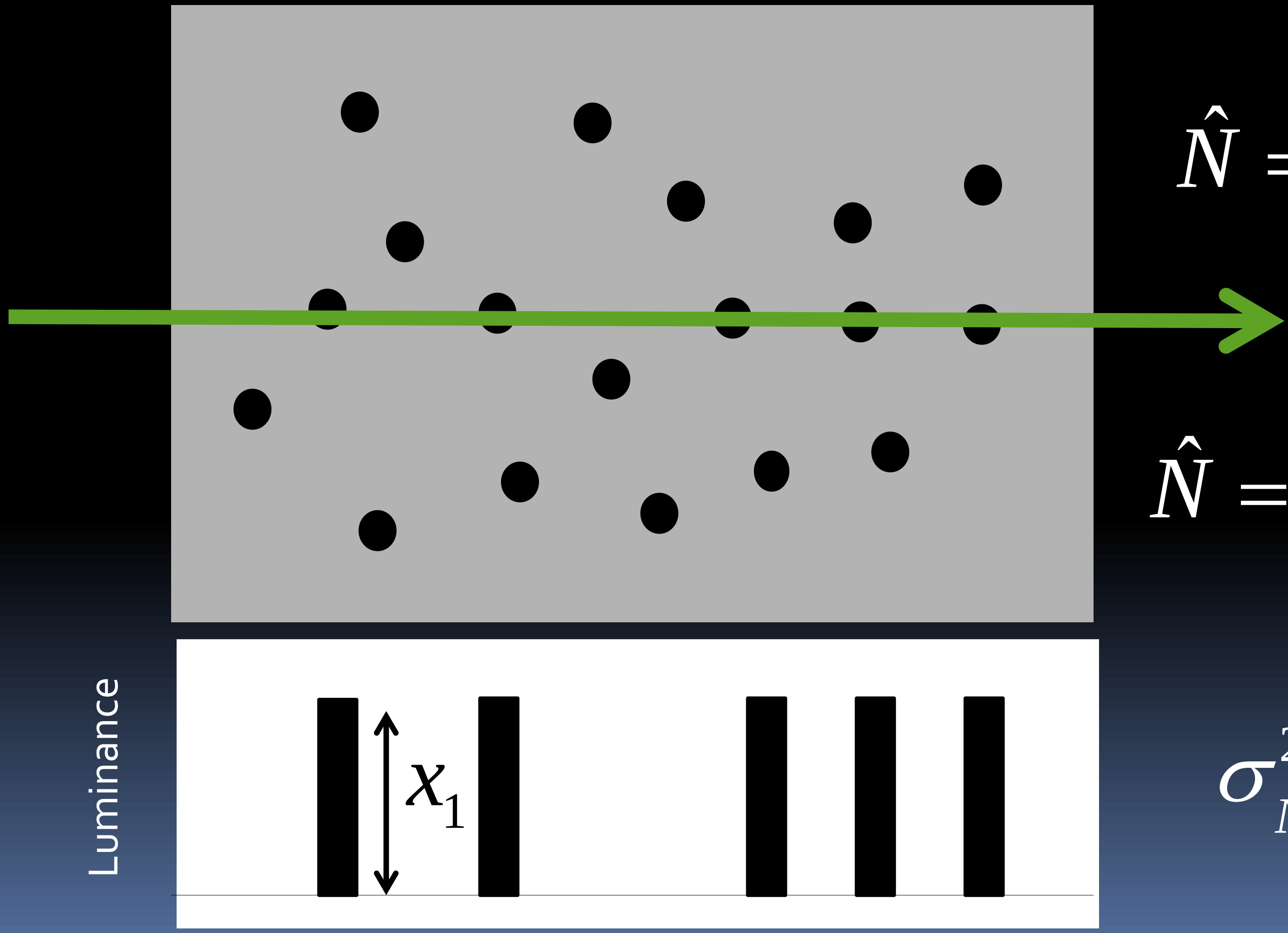
$$\hat{N} = \sum_{i=1}^{N} x_i$$

Weber's law



$$\hat{N} = \sum_{i=1}^{N} x_i$$

Weber's law

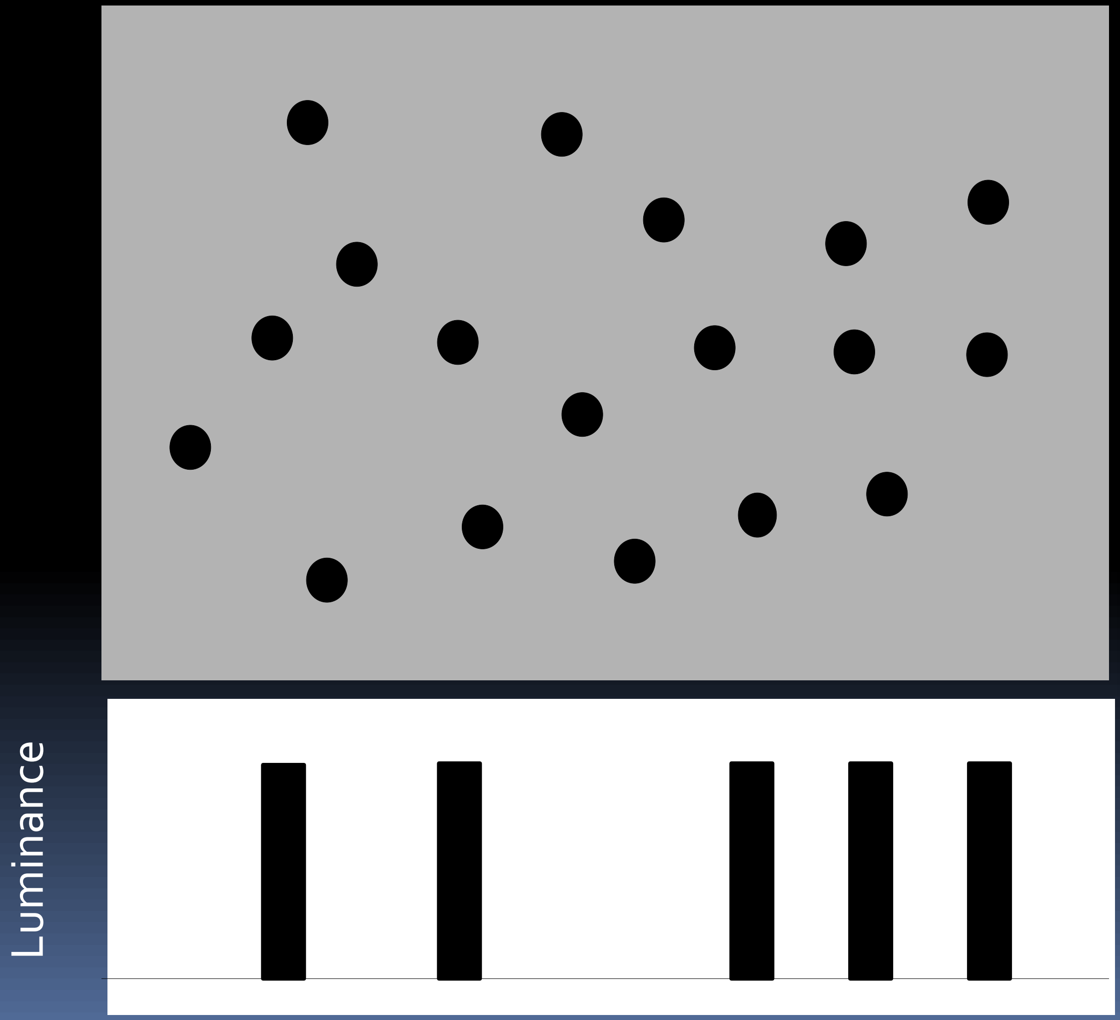


$$\hat{N} = \sum_{i=1}^{N} x_i$$

$$\hat{N} = \sum_{i=1}^{N} x_i \propto N x_1$$

$$\sigma_{\hat{N}}^2 \propto N^2 \sigma_{x_1}^2$$

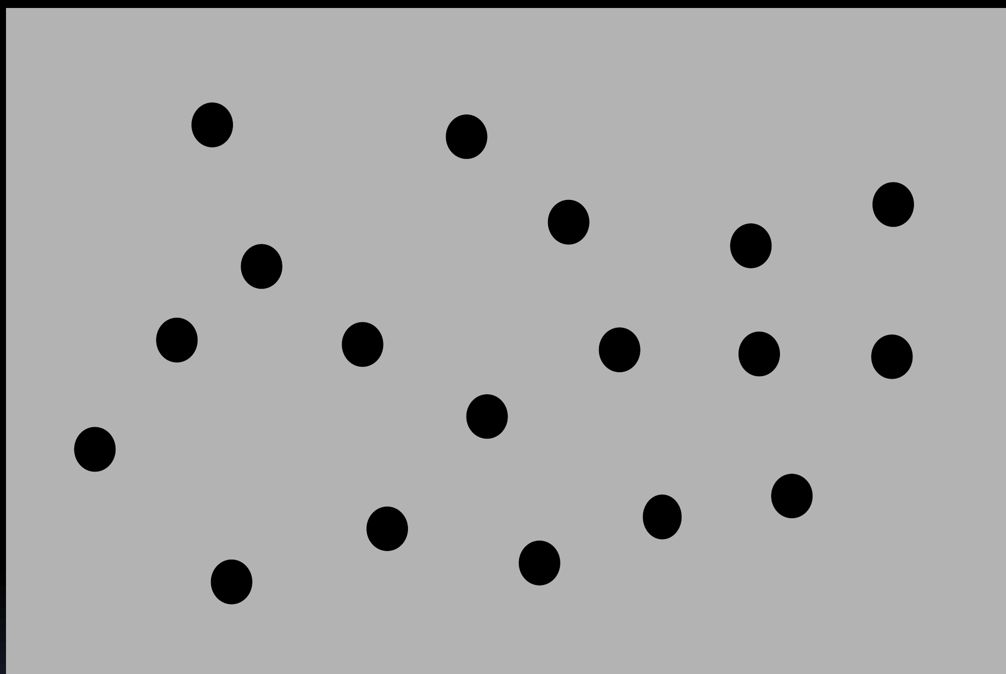
Weber's law

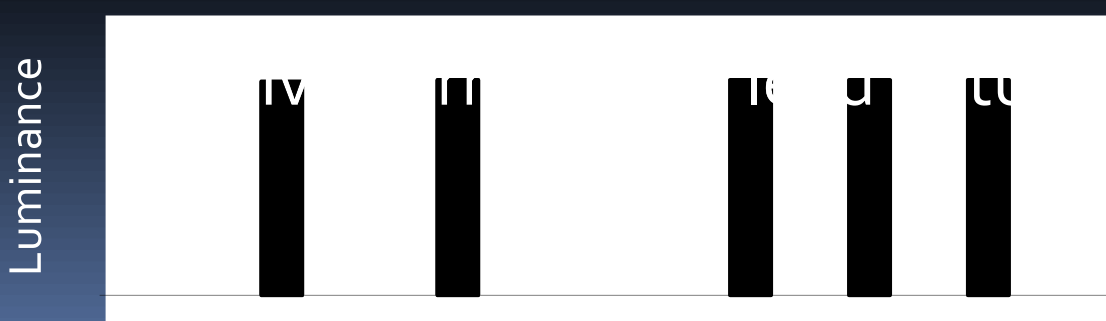


Weber's law!

$$\sigma^2_{\hat{N}} \propto N^2$$

Weber's law





Weber's law!

ly to Weber's law

$$\sigma_{\hat{N}}^2 \propto N^2$$

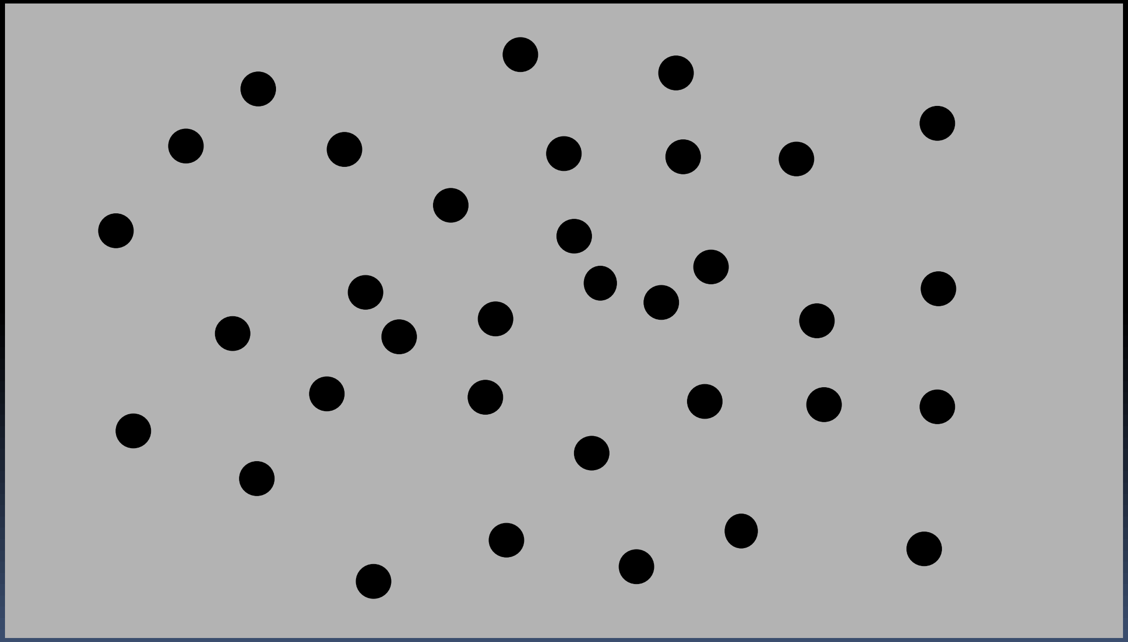
Weber's law

Sensory information is often scaled by global nuisance parameters such as

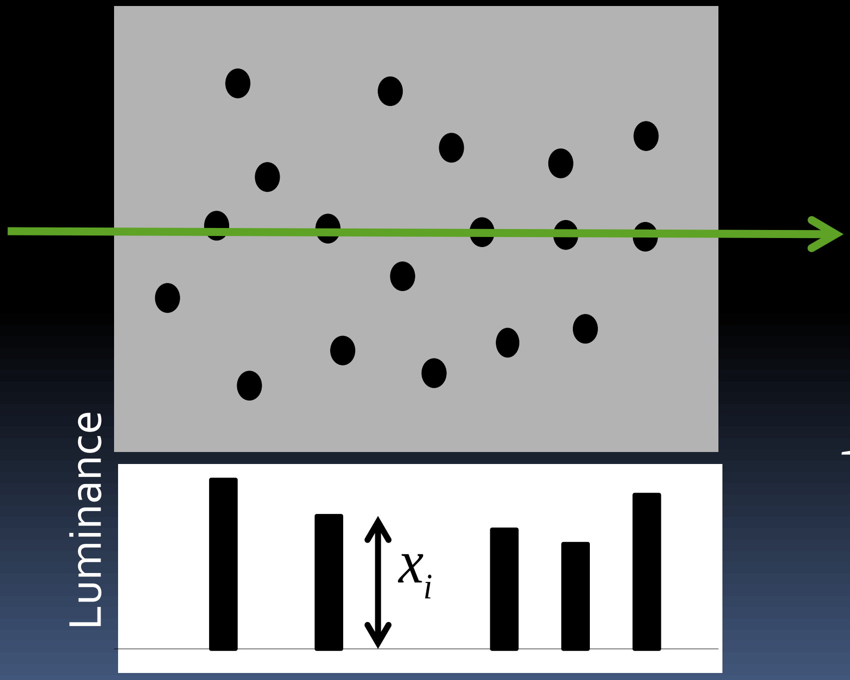
- Contrast
- Loudness
- Co-contraction of muscles
- Attention
- Learning

…etc

Don’t we get Weber’s law even when the contrast is maintained constant across images?



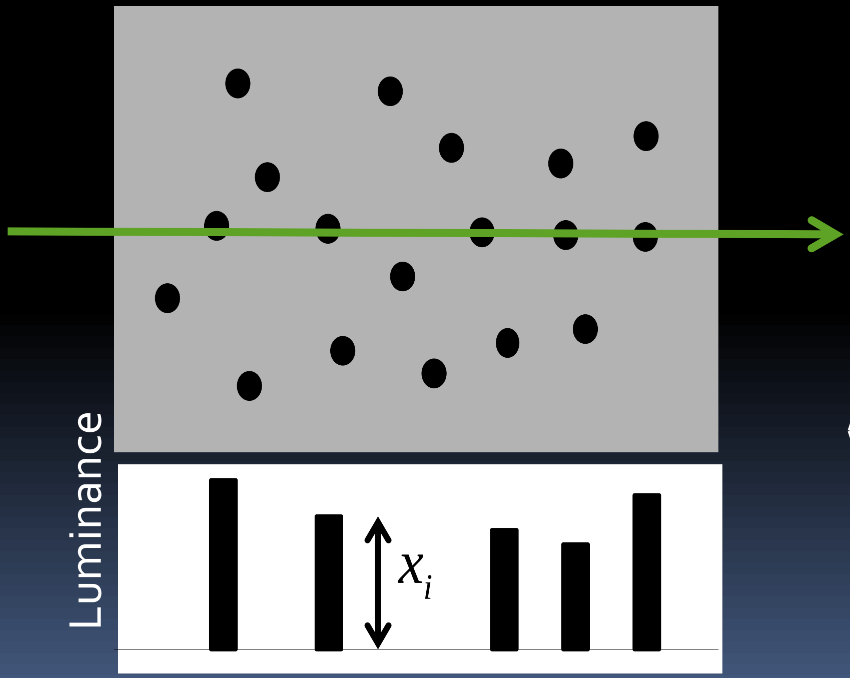
When contrast varies, the estimate of numerosity should use normalized luminance



$$\hat{N} = \sum_{i=1}^{N} g x_i$$

Contrast

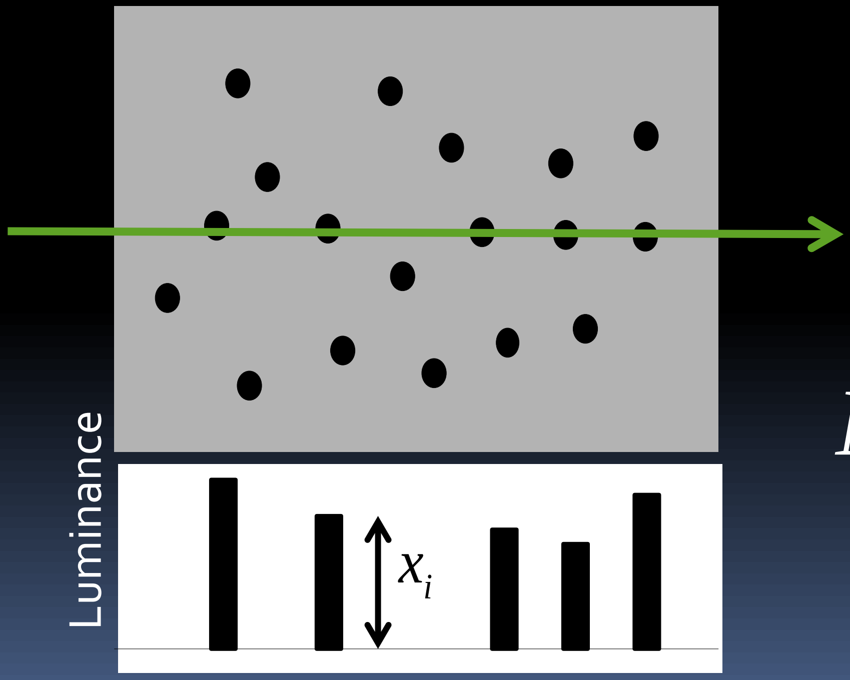
When contrast varies, the estimate of numerosity should use normalized luminance



$$\langle \hat{N} \rangle \propto gN$$

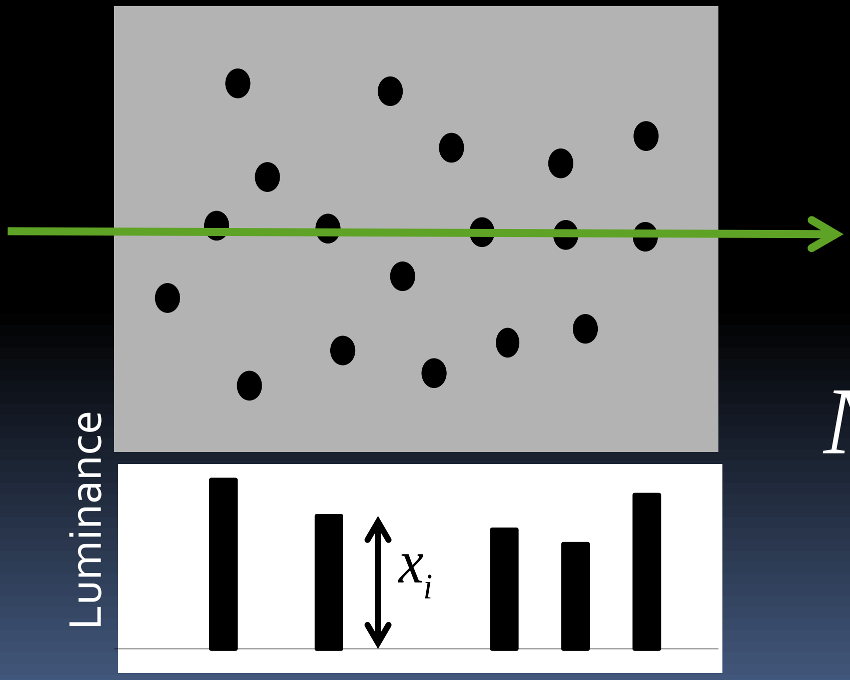
Contrast

When contrast varies, the estimate of numerosity should use normalized luminance



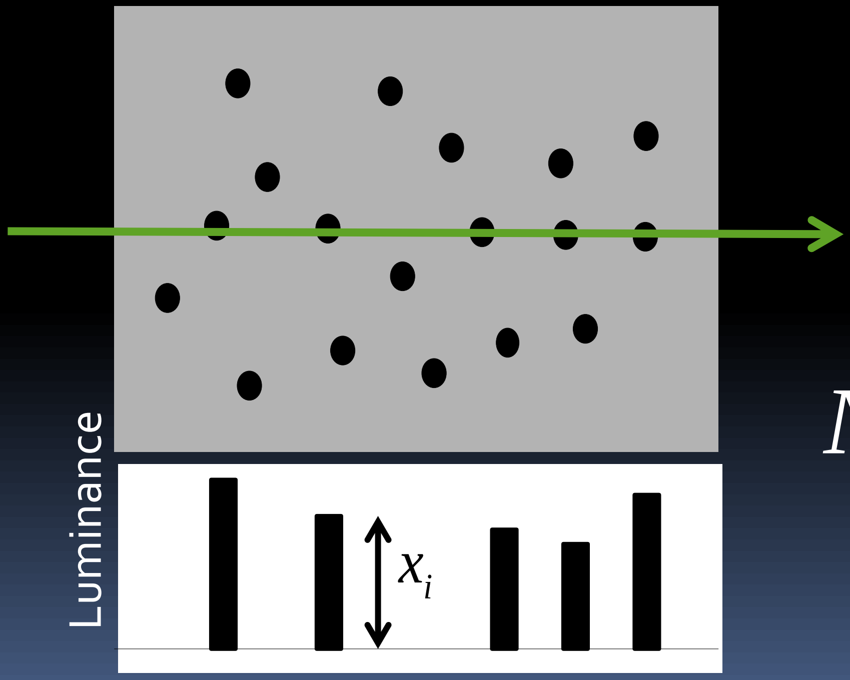
$$\hat{N} = g \sum_{i=1}^{N} x_i$$

When contrast varies, the estimate of numerosity should use normalized luminance



$$\hat{N} = \frac{g}{g} \sum_{i=1}^{N} x_i$$

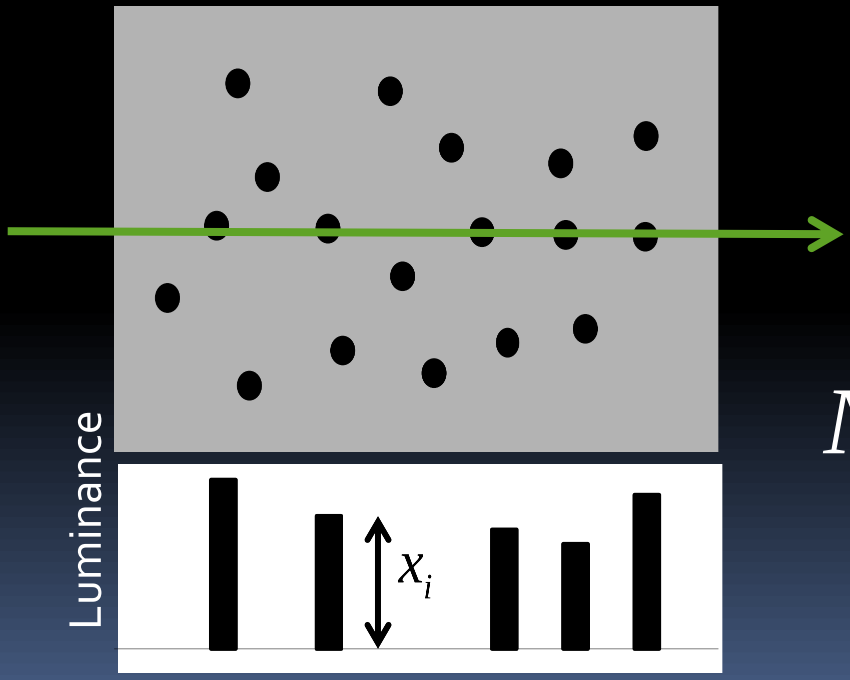
But subjects do not know contrast and must therefore use an estimate of contrast



$$\hat{N} = \frac{g}{\hat{g}} \sum_{i=1}^{N} x_i$$

Estimated Contrast

But subjects do not know contrast, and must therefore use an estimate of contrast



Induces positive correlations

$$\hat{N} = \frac{g}{\hat{g}} \sum_{i=1}^{N} x_i$$

- Even when contrast is constant, the estimate of contrast is likely to vary, and could greatly vary due to approximations.

Weber's law

Global scaling nuisance parameters induce correlations that naturally leads to Weber's law.

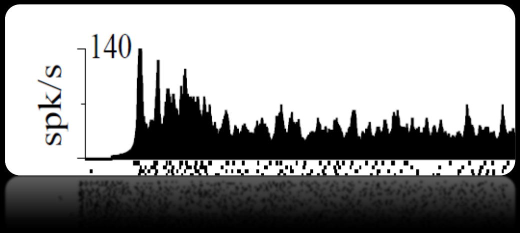
No need to invoke single cell Poisson variability plus log normal tuning curves

In fact, log normal tuning curves might be the <u>consequence</u> of Weber's law, not the cause (Dehaene et al, cosyne 2014).

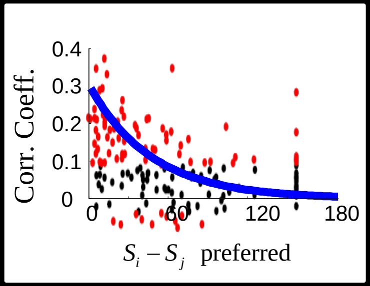
Conclusions

Where does behavioral variability come from?

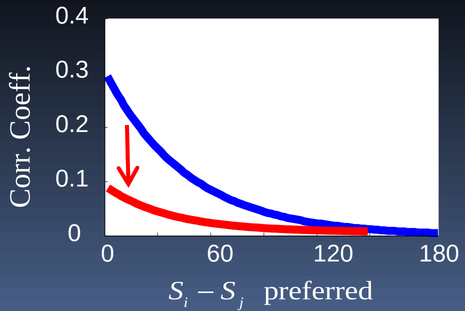
Poisson variability



Correlations



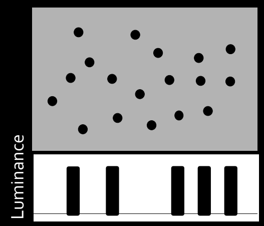
Decorrelation



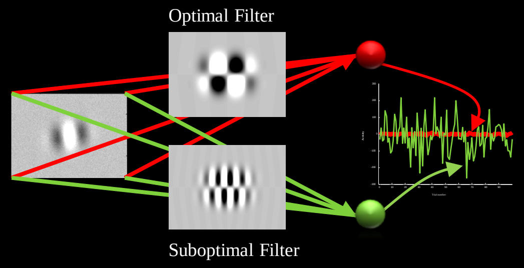
Conclusions

Where does behavioral variability come from?

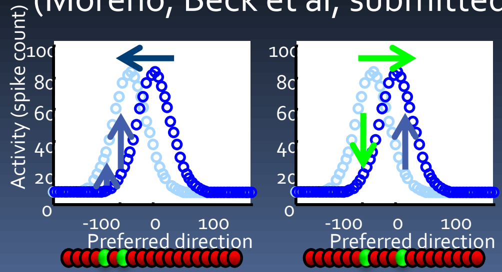
Variable sensory data
(Beck et al, in prep)



Suboptimal inference
(Beck et al, 2012)



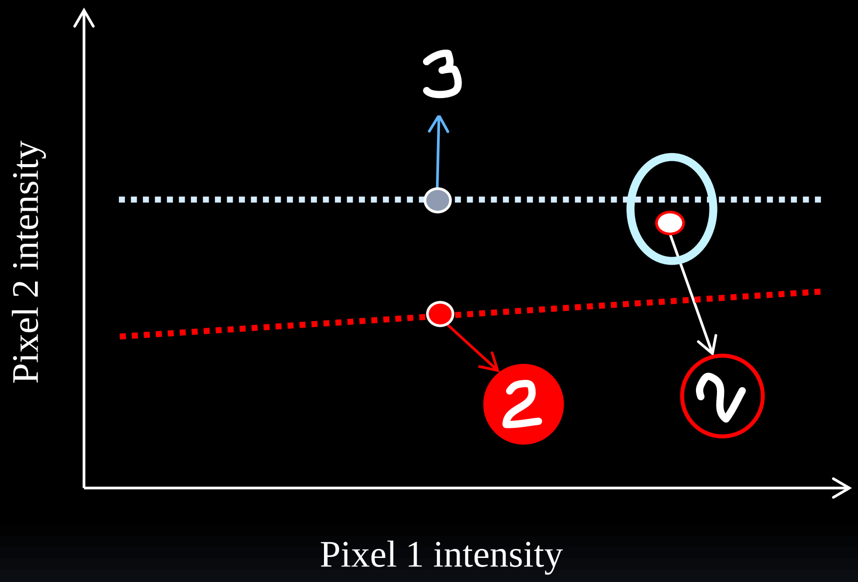
Differential correlations
(Moreno, Beck et al, submitted)



Positive roles of noise

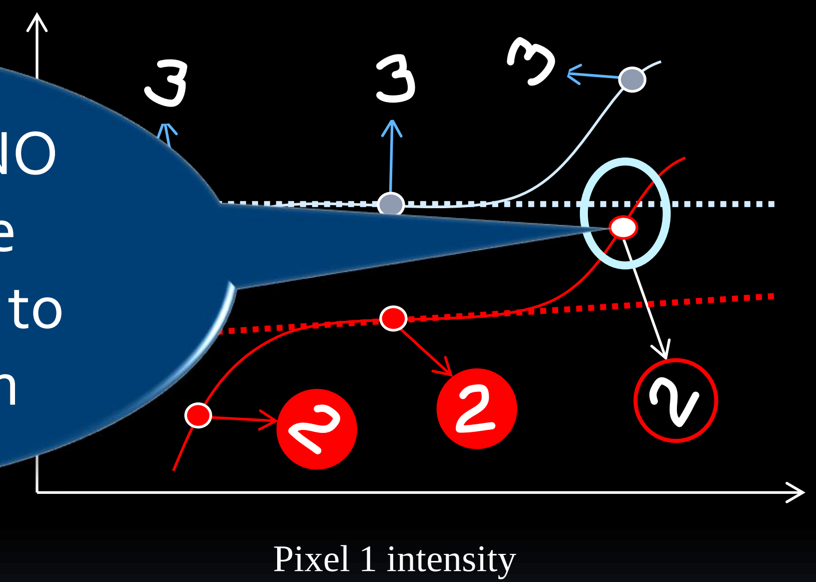
- Sampling
- Exploration
- Game theory

Object recognition

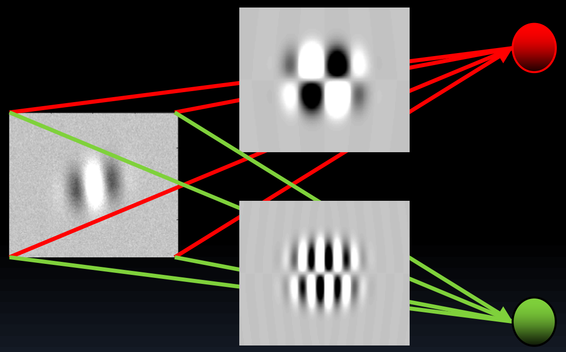


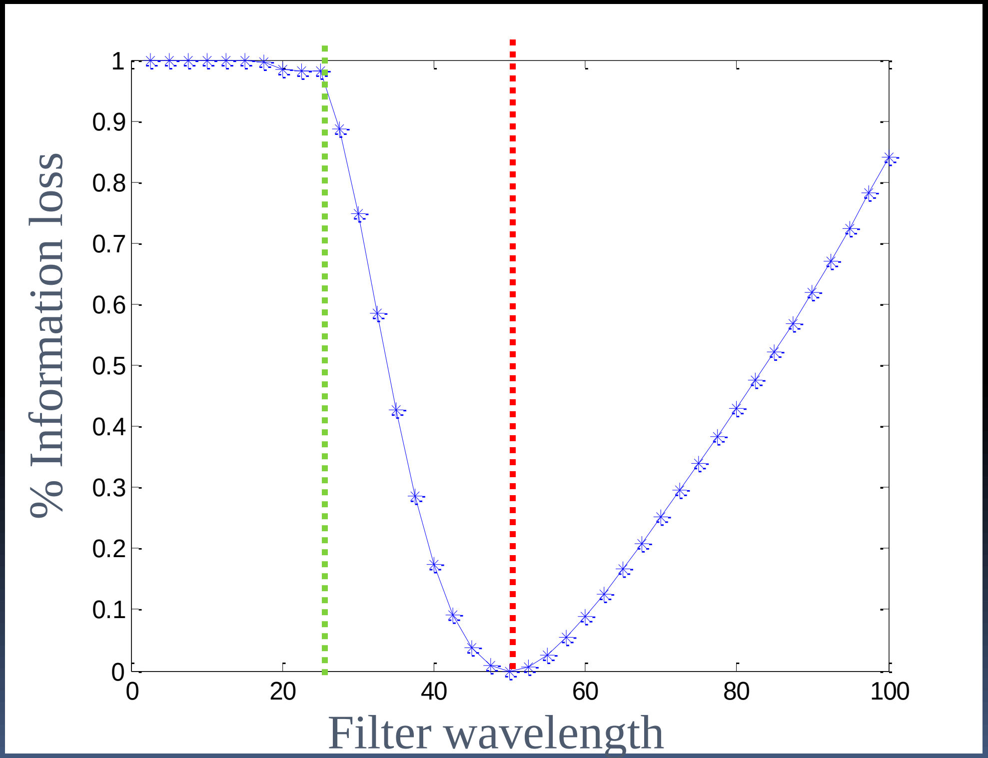
Object recognition

Note that there is NO sensory noise. The error is due entirely to the approximation

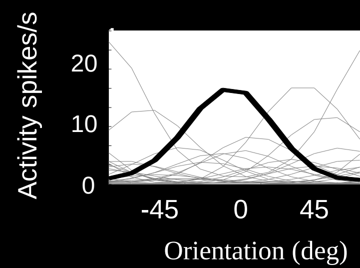


Orientation discrimination

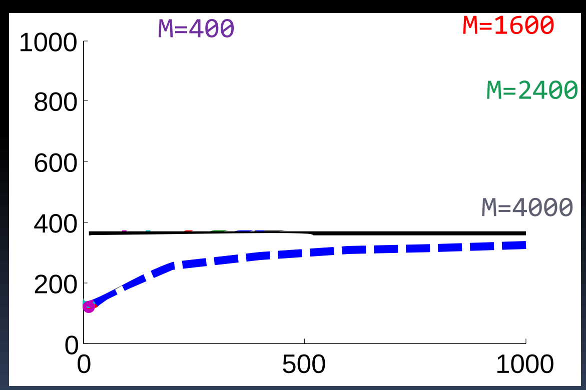




Experimental consequences



Explicit Computation

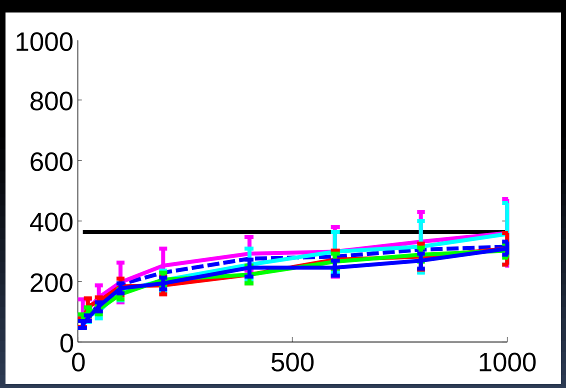


Information

Number of Neurons

$$\hat{\mathbf{f}}'\cdot\left(\hat{\Sigma}_{\varepsilon}\right)^{-1}\hat{\mathbf{f}}'$$

Information from decoding

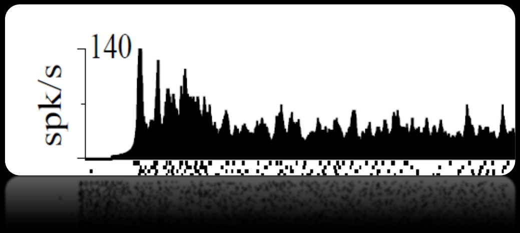


Number of Neurons

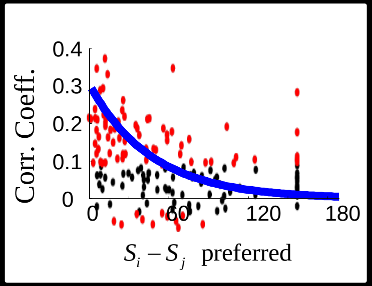
Conclusions

Where does behavioral variability come from?

Poisson variability



Correlations



Decorrelation