Modeling of behavior

Wei Ji Ma

New York University
Center for Neural Science and Department of Psychology

You can download the slides from www.cns.nyu.edu/malab (News), so no need to frantically copy/photograph them. Probably better to follow along.

On an entirely unrelated note, if you want to set up #growingupinscience in your own department, here are some pointers: www.growingupinscience.com (Get involved)

Seriously, a tutorial?

Focus on **motivation** and **methods**

not on results

And I'll make you work

Multisensory perception
Visual working memory
Categorization
Visual search
Perceptual organization
Horizontal-vertical illusion
Aperture problem
Proactive interference
Word recognition memory
Confidence ratings
Smooth pursuit eye movements
Information sampling in a trust game
Choosing between cookies and chips
Exploration/exploitation
Playing strategy games
Monkeys determining which female to mate with based on the color of her face

15 years of modeling of behavior

One can build a respectable career on a relatively limited skill set.

Why do we fit models?

From Ma lab survey by Bas van Opheusden, 201703

Why do we fit models?

To drown a conceptually uninteresting question in math

Maslow's hammer

Because we can and we are good at it

To get into a higher impact journal

Because Weiji says so

From Ma lab survey by Bas van Opheusden, 201703

Why do we fit models?

To make inferences about latent causes of behavior that we cannot observe directly

To get closer to a simplified form of people's cognitive processes

Because we want to infer latent variables/mechanisms

To say something about the potential computations involved when completing a task

Infer what's really happening inside the black box

To create order in the universe

Models let us ask questions that are hard to answer with experiments

To quantify evidence for our theories and hypotheses

To produce good models according to well-considered criteria

From Ma lab survey by Bas van Opheusden, 201703

“The purpose of models is not to fit the data but to sharpen the questions.”

— Samuel Karlin, R.A. Fisher Memorial Lecture, 1983

“If a principled model with a few parameters fits a rich behavioral data set, I feel that I have really understood something about the world” — Wei Ji Ma, CCN Tutorial, 2017

Agenda

1. Model building
2. Model fitting
3. Model comparison

Part 1: Model building

1a. What kind of model - descriptive or process?

1b. A special kind of process model - Bayesian

1c. Prior examples: visual illusions

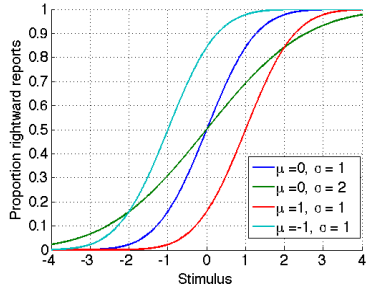
1d. Likelihood example: Gestalt perception

1e. How to actually do Bayesian modeling?

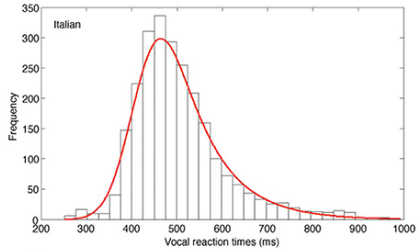
1a. What kind of model?

Descriptive model: summary of the data in a function with (usually a small number of) parameters

- Psychometric curve: cumulative Gaussian

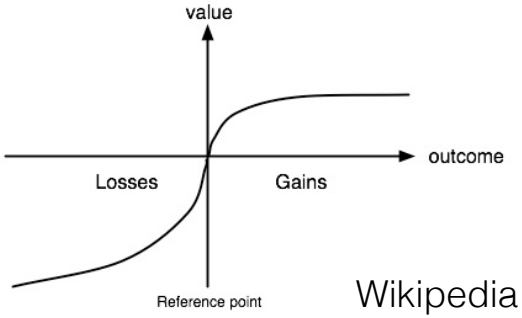


- Ex-Gaussian fit to reaction time distribution



Marinelli et al. 2014

- Prospect theory



Wikipedia

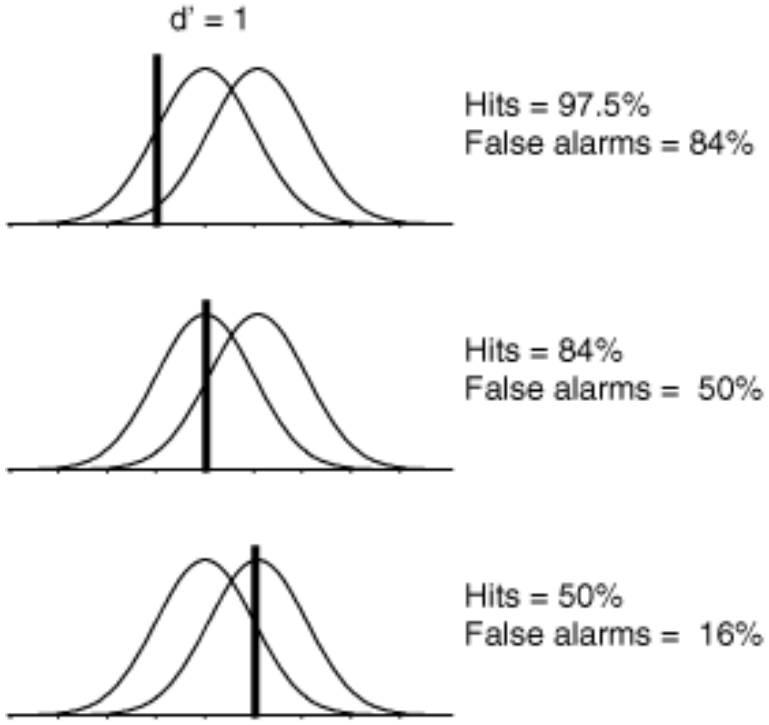
Fitting descriptive models is like doing laundry

1a. What kind of model?

- **Descriptive model:** summary of the data in a function with (usually a small number of) parameters
  - Danger: arbitrarily throwing parameters at it
- **Process model:** model based on a psychological hypothesis of how an observer/agent makes a decision
  - Interpretable! (Nicole Rust)

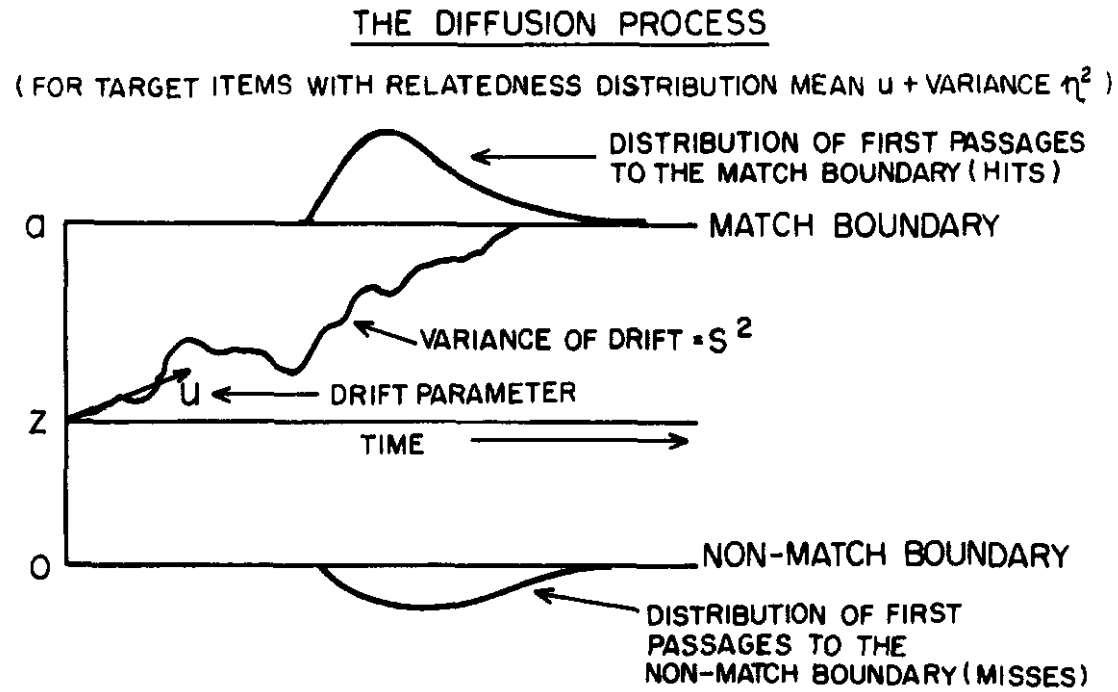
Process models

- Signal detection theory

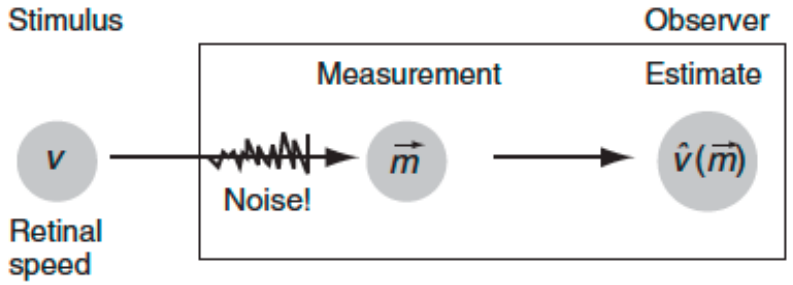


David Heeger lecture notes

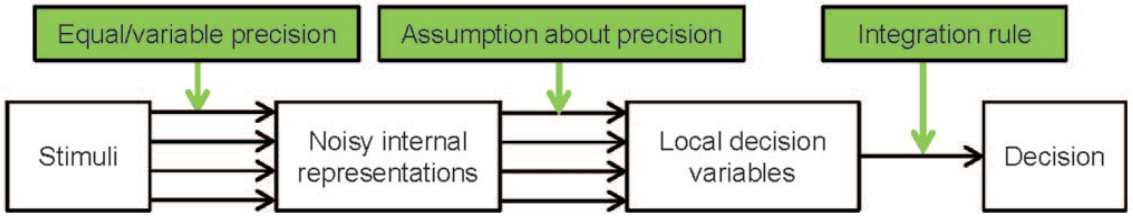
- Drift-diffusion model



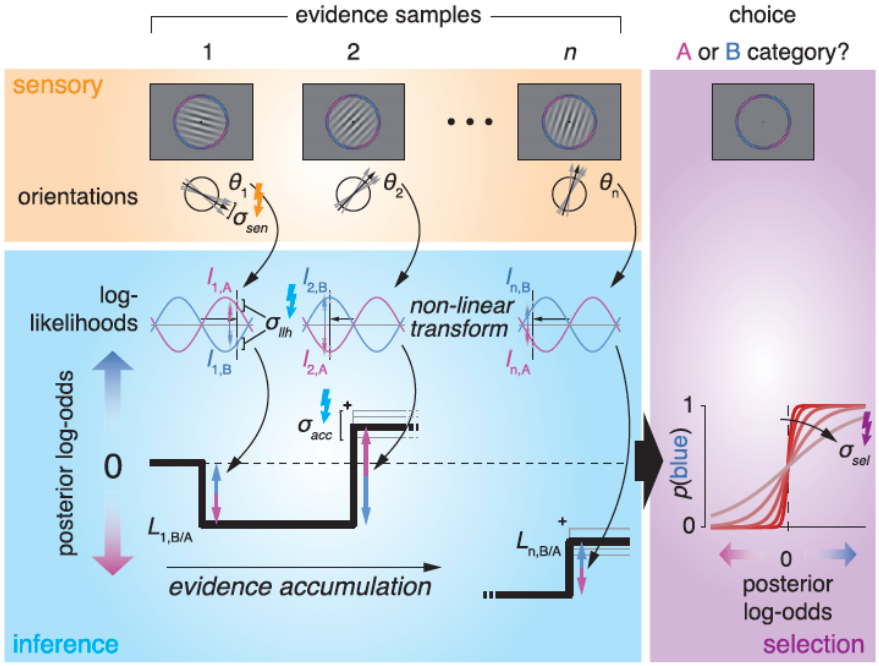
Ratcliff 1978



Stocker and Simoncelli, 2006



Keshvari et al., 2012



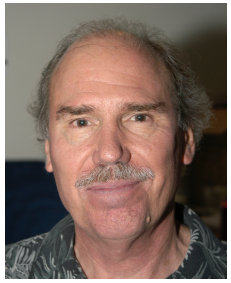
Drugowitsch et al., 2016

1b. A special kind of process model: Bayesian

- **State of the world unknown to decision-maker**
  - Uncertainty!
- **Decision-maker maximizes an objective function**
  - In categorical perception: accuracy
  - But could be hitting error, point rewards
- Stronger claim: brain represents probability distributions

1b. Why Bayesian models?

- **Evolutionary/philosophical**: Bayesian inference optimizes performance or minimizes cost. The brain might have optimized perceptual processes. This is handwavy but very cool if true.
- **Empirical**: in many tasks, people are close to Bayesian. This is hard to argue with.
- **Bill Geisler's couch argument:**



It is harder to come up with a good model sitting on your couch than to work out the Bayesian model.

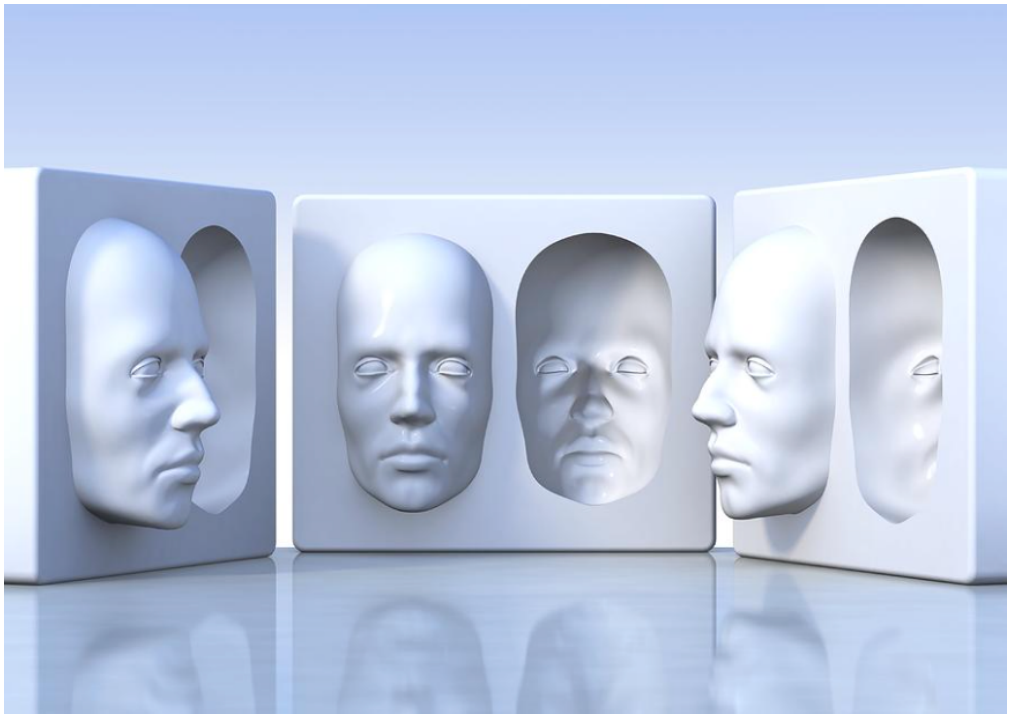
- **Basis for suboptimal models**: Other models can often be constructed by modifying the assumptions in the Bayesian model. Thus, the Bayesian model is a good starting point for model generation.

Where does uncertainty come from?

- Noise
- Ambiguity


3
GATHERING FOR
GARDNER

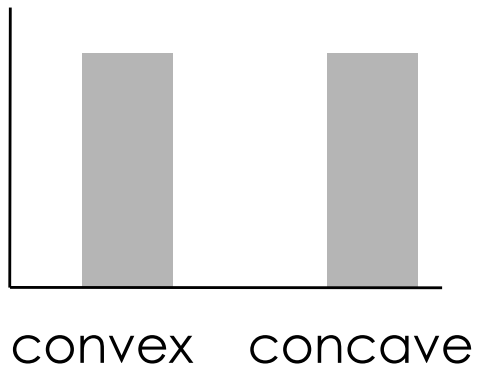
Hollow-face illusion



David Mack

Likelihood

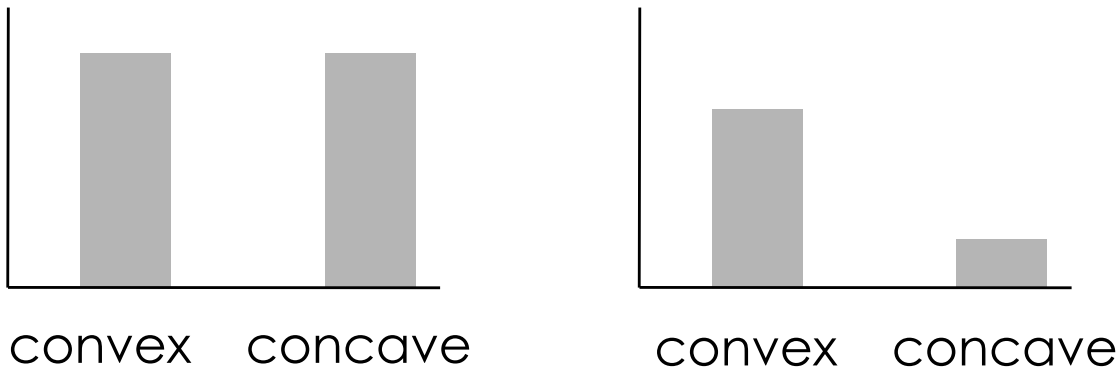
how probable are the retinal image is if the hypothesis were true



Likelihood x Prior

how probable are the retinal image is if the hypothesis were true

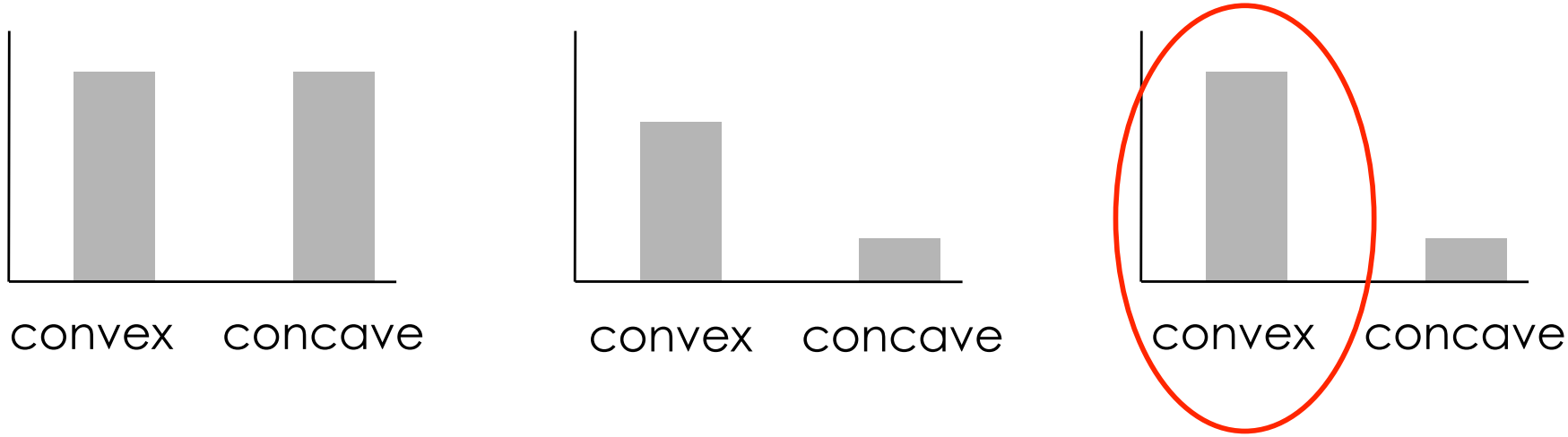
how much do you expect the hypothesis based on your experiences



Likelihood x Prior ∝ Posterior probability

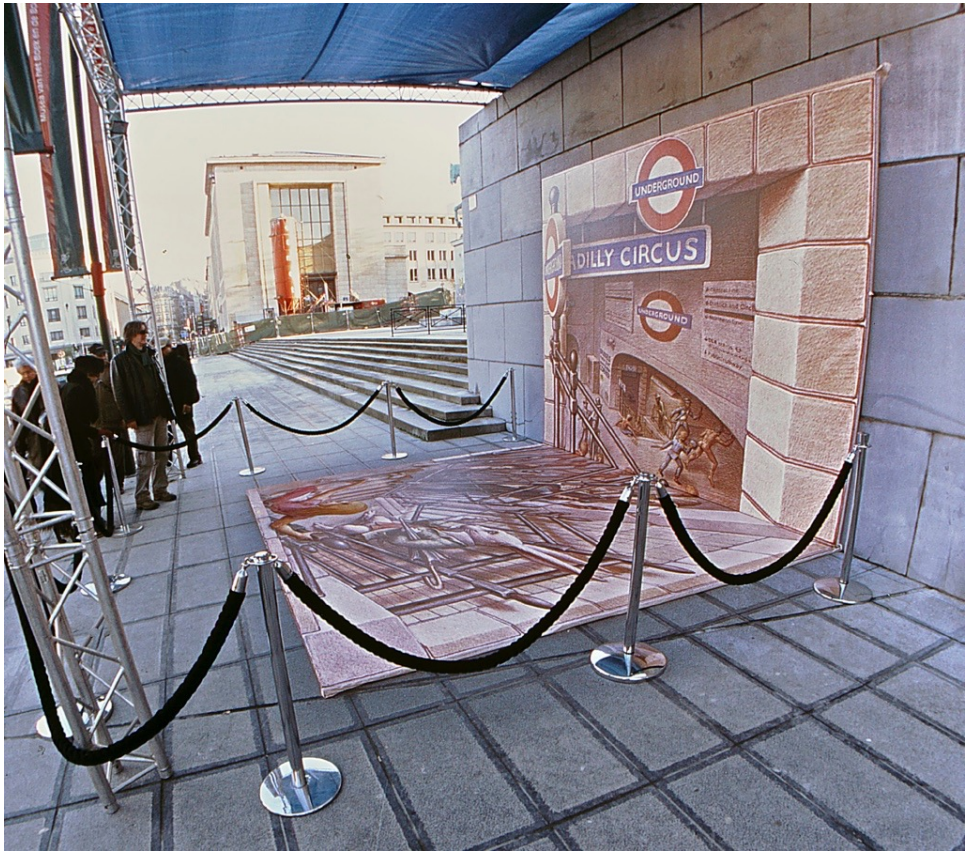
how probable are the retinal image is if the hypothesis were true

how much do you expect the hypothesis based on your experiences

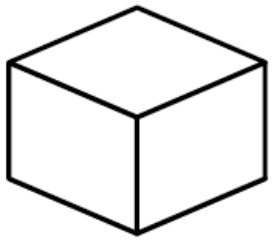


This hypothesis becomes your percept!

Anamorphic illusion by Kurt Wenner

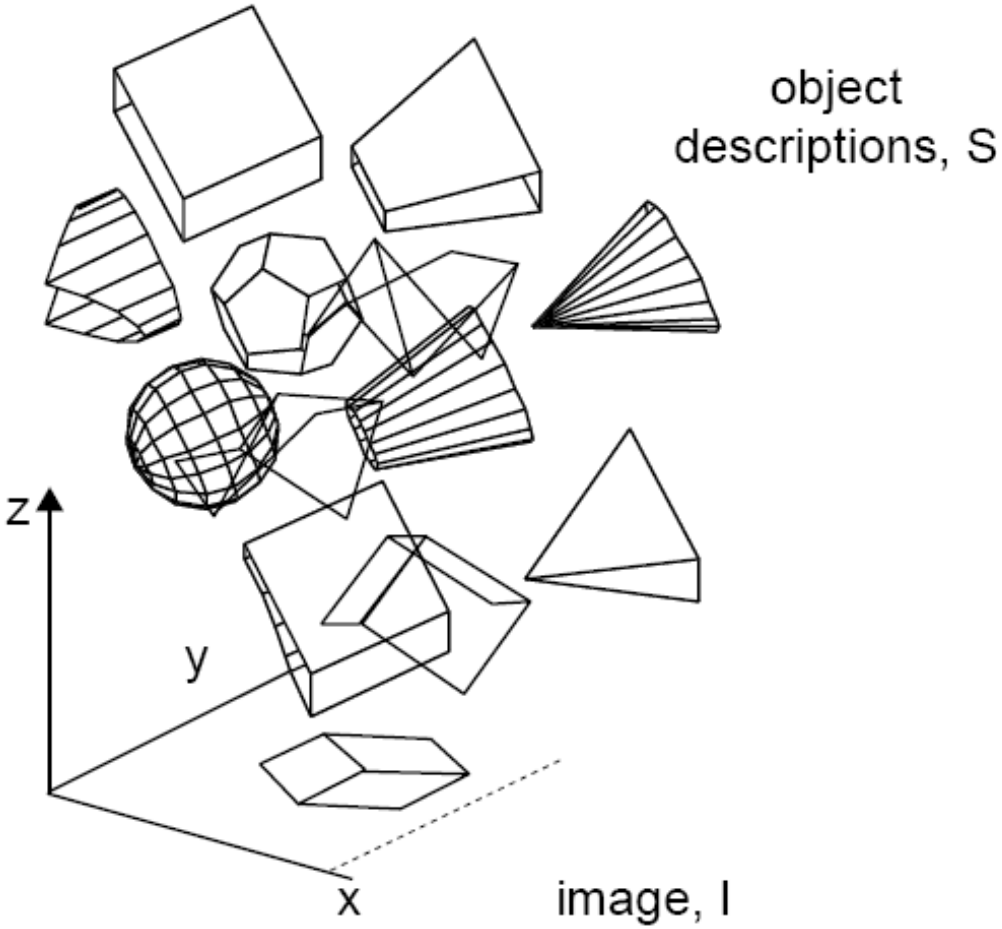


- Where is the ambiguity?
- What role do priors play?
- What happens if you view with two eyes, and why?



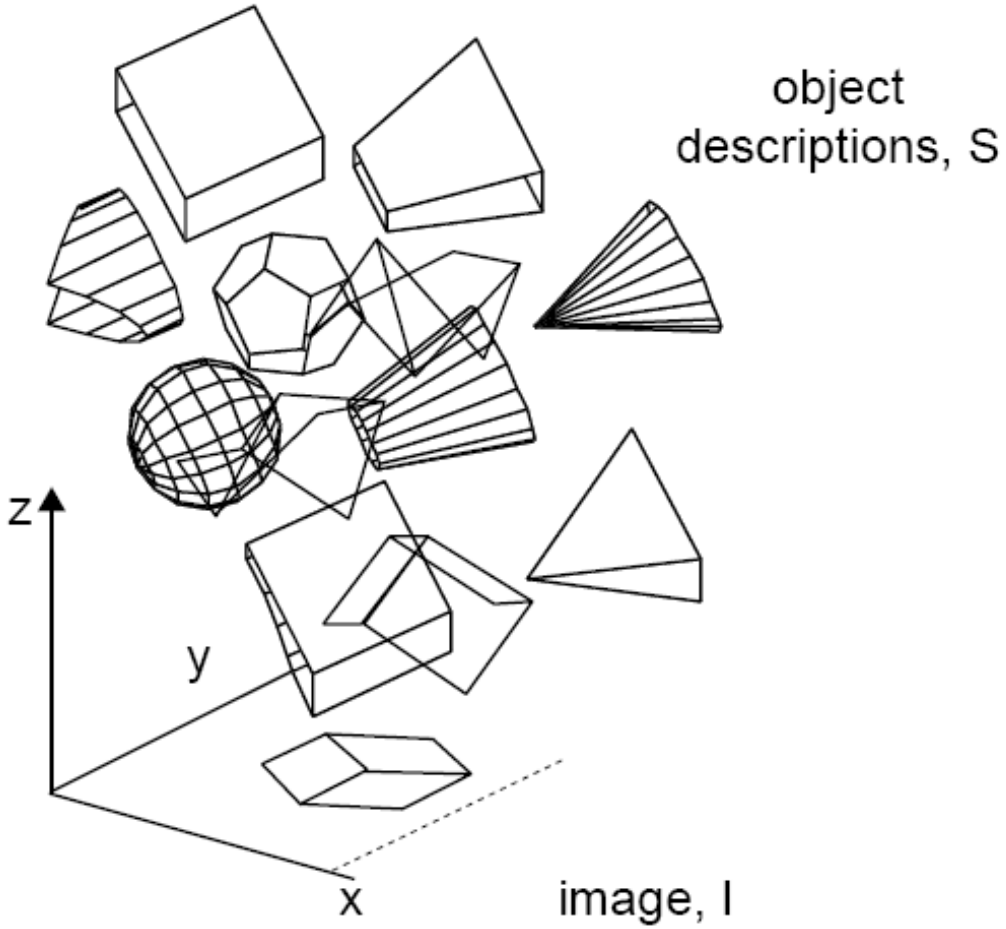
Prior over objects

$p(s)$



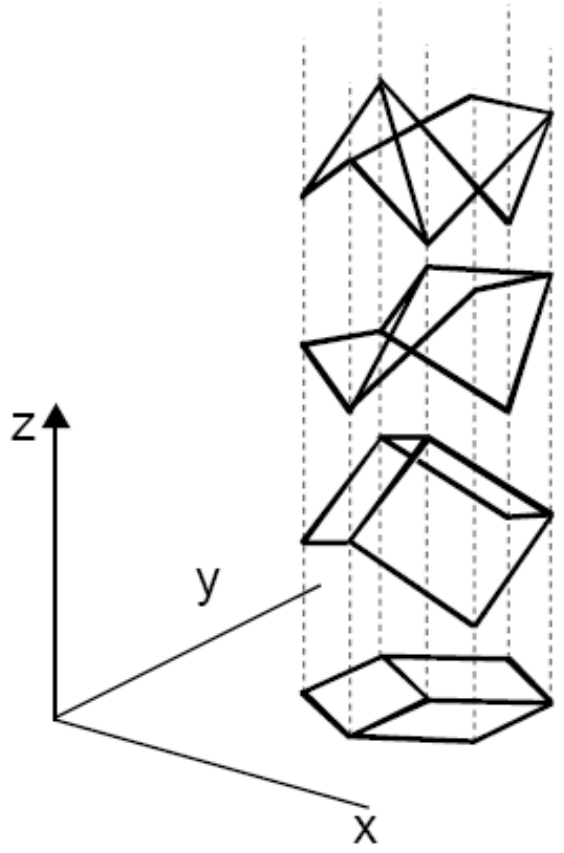
Prior over objects

$p(s)$



Likelihood over objects given 2D image

$L(s) = p(I|s)$



Kersten and Yuille, 2003

Examples of priors:

- Convex faces are more common than concave ones
- Priors at the object level (Kersten and Yuille)
- Light usually comes from above (Adams and Ernst)
- Slower speeds are more common (Simoncelli and Stocker)
- Cardinal orientations are more common (Landy and Simoncelli)

~~Bayesian models are about priors~~

Fake news

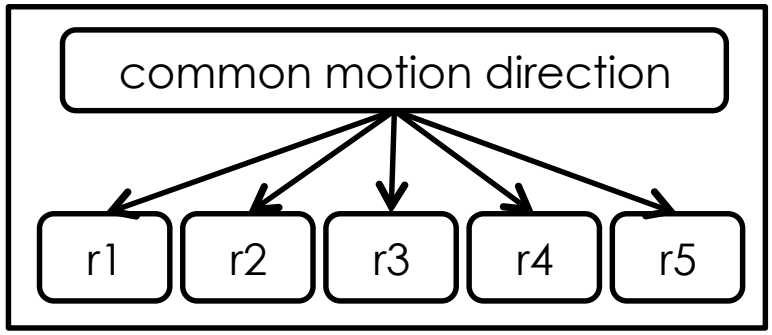
Bayesian models are about:

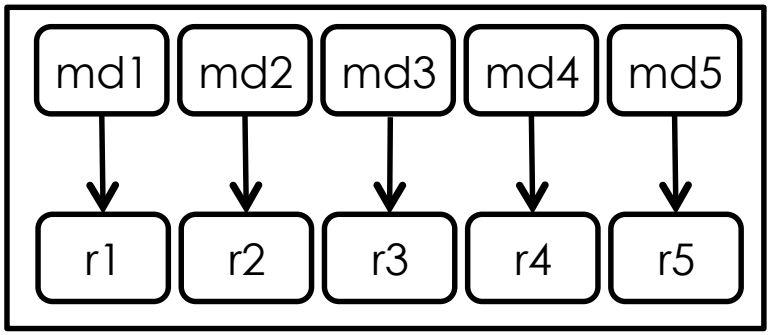
- the decision-maker making the best possible decision (given an objective function)
- the brain representing probability distributions

Law of common fate

Bayesian explanation?

Generative model

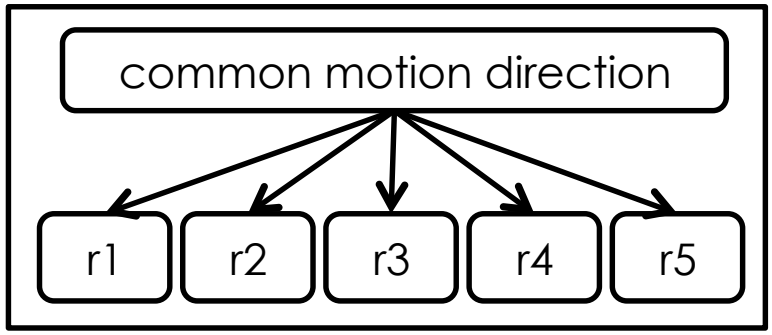




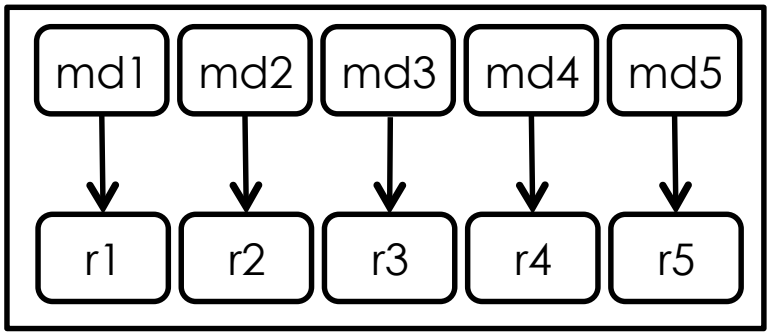
Scenario 1: All dots are part of the same object and they therefore always move together. They move together either up or down, each with probability 0.5.

Scenario 2: Each dot is an object by itself. Each dot independently moves either up or down, each with probability 0.5.

Scenario 1



Scenario 2

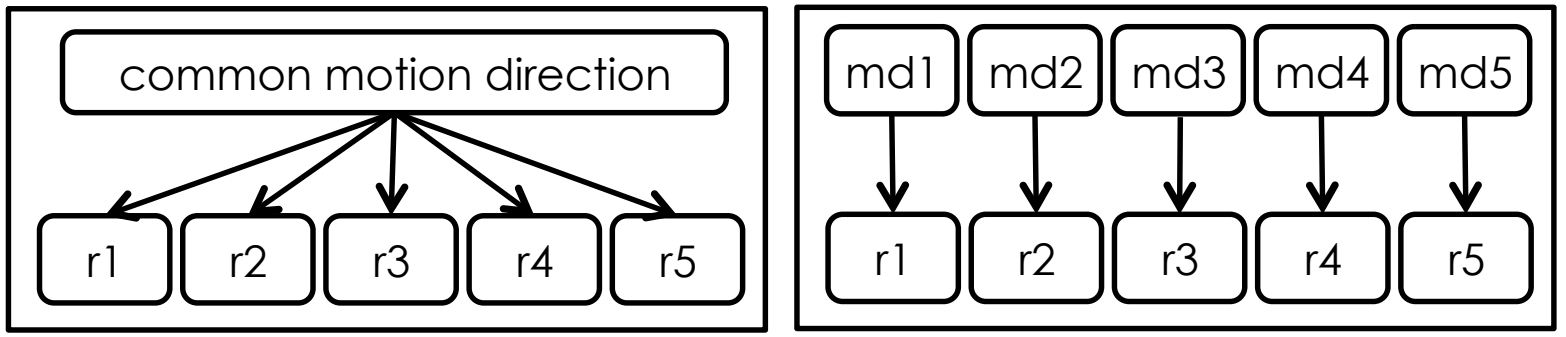


Scenario 1: All dots are part of the same object and they therefore always move together. They move together either up or down, each with probability 0.5.

Scenario 2: Each dot is an object by itself. Each dot independently moves either up or down, each with probability 0.5.

Sensory observation: all dots moving down.

The **likelihood** of a scenario is the probability of these sensory observations under the scenario. What is the likelihood of **Scenario 1**?

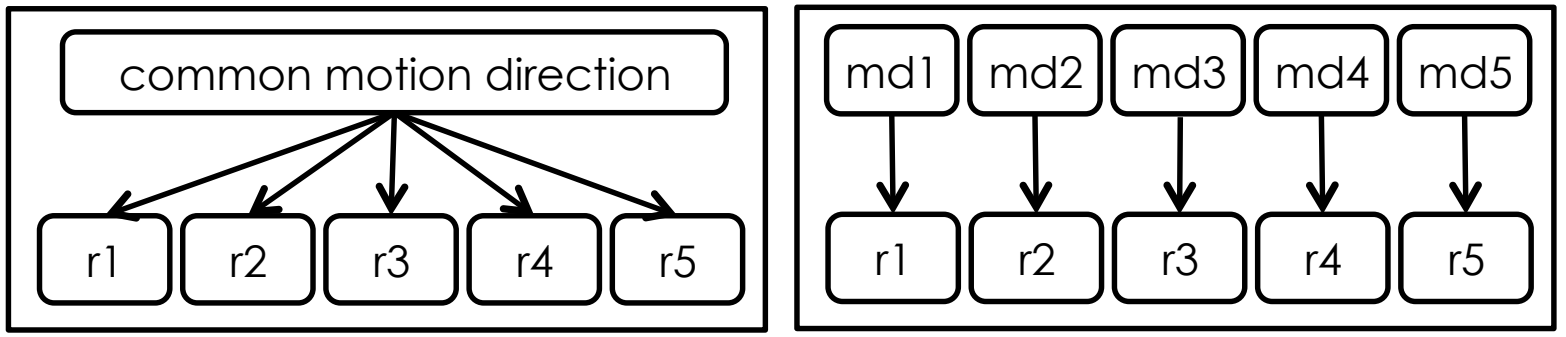


Scenario 1: All dots are part of the same object and they therefore always move together. They move together either up or down, each with probability 0.5.

Scenario 2: Each dot is an object by itself. Each dot independently moves either up or down, each with probability 0.5.

Sensory observation: all dots moving down.

The **likelihood** of a scenario is the probability of these sensory observations under the scenario. What is the likelihood of **Scenario 2**?



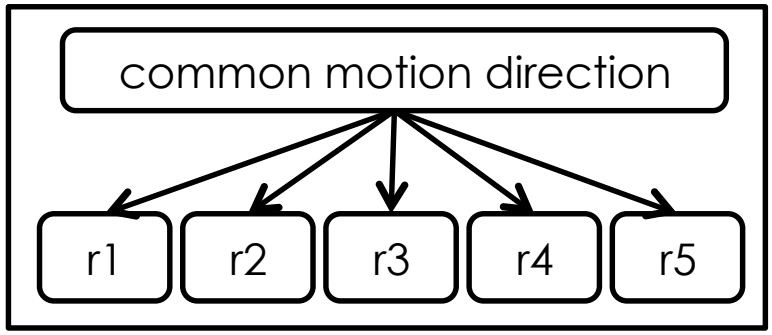
Scenario 1: All dots are part of the same object and they therefore always move together. They move together either up or down, each with probability 0.5.

Scenario 2: Each dot is an object by itself. Each dot independently moves either up or down, each with probability 0.5.

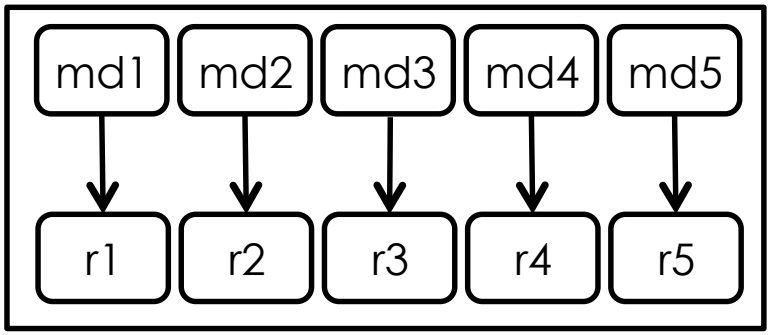
Sensory observation: all dots moving down.

How many times larger is the likelihood of Scenario 1 than of Scenario 2?

Scenario 1



Scenario 2



Scenario 1: All dots are part of the same object and they therefore always move together. They move together either up or down, each with probability 0.5.

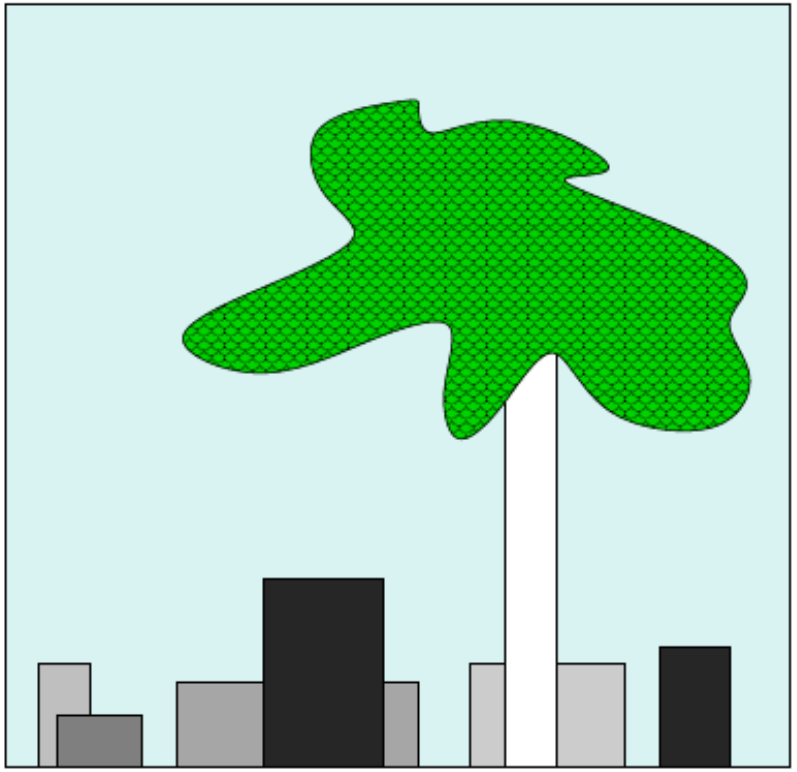
Scenario 2: Each dot is an object by itself. Each dot independently moves either up or down, each with probability 0.5.

Sensory observation: all dots moving down.

Say the priors are equal. How many times larger is the posterior probability of Scenario 1?

With likelihoods like these, who needs priors?

Bayesian models are about the *best possible* decision.



MacKay (2003), *Information theory, inference, and learning algorithms*, Sections 28.1-2

1e. How to actually do Bayesian modeling?

Good news: there is a general recipe
that you just need to follow.

The four steps of Bayesian modeling

Example: categorization task

STEP 1: GENERATIVE MODEL

a) Draw a diagram with each node a variable and each arrow a statistical dependency. Observation is at the bottom.
b) For each variable, write down an equation for its probability distribution. For the observation, assume a noise model. For others, get the distribution from your experimental design. If there are incoming arrows, the distribution is a conditional one.

World state of interest
$p(C)=0.5$

Stimulus
$p(s\,|\,C)=\mathrm{N}(s;\mu_C,\sigma_C^2)$

Observation
$p(x\,|\,s)=\mathrm{N}(x;s,\sigma^2)$

STEP 2: BAYESIAN INFERENCE (DECISION RULE)

a) Compute the posterior over the world state of interest given an observation. The optimal observer does this using the distributions in the generative model. Alternatively, the observer might assume different distributions (natural statistics, wrong beliefs). Marginalize (integrate) over variables other than the observation and the world state of interest.

$$p(C\,|\,s)\propto p(C)\,p(x\,|\,C)=p(C)\int p(x\,|\,s)\,p(s\,|\,C)\,ds=\ldots=\mathrm{N}(x;\mu_C,\sigma^2+\sigma_C^2)$$

b) Specify the read-out of the posterior. Assume a utility function, then maximize expected utility under posterior. (Alternative: sample from the posterior.) Result: decision rule (mapping from observation to decision). When utility is accuracy, the read-out is to maximize the posterior (MAP decision rule).

$$\hat{C}=1 \text{ when } \mathrm{N}(x;\mu_1,\sigma^2+\sigma_1^2)>\mathrm{N}(x;\mu_2,\sigma^2+\sigma_2^2)$$

STEP 3: RESPONSE PROBABILITIES

For every unique trial in the experiment, compute the probability that the observer will choose each decision option given the stimuli on that trial using the distribution of the observation given those stimuli (from Step 1) and the decision rule (from Step 2).

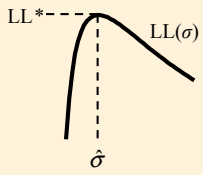
$$p\left(\hat{C}=1\,|\,x\right)=\mathrm{Pr}_{x|s;\sigma}\left(\mathrm{N}(x;\mu_1,\sigma^2+\sigma_1^2)>\mathrm{N}(x;\mu_2,\sigma^2+\sigma_2^2)\right)$$

- Good method: sample observation according to Step 1; for each, apply decision rule; tabulate responses. Better: integrate numerically over observation. Best (when possible): integrate analytically.
- Optional: add response noise or lapses.

STEP 4: MODEL FITTING AND MODEL COMPARISON

a) Compute the parameter log likelihood, the log probability of the subject's actual responses across all trials for a hypothesized parameter combination.
b) Maximize the parameter log likelihood. Result: parameter estimates and maximum log likelihood. Test for parameter recovery and summary statistics recovery using synthetic data.
c) Obtain fits to summary statistics by rerunning the fitted model.
d) Formulate alternative models (e.g. vary Step 2). Compare maximum log likelihood across models. Correct for number of parameters (e.g. AIC). (Advanced: Bayesian model comparison, uses log marginal likelihood of model.) Test for model recovery using synthetic data.
e) Check model comparison results using summary statistics.

$$\mathrm{LL}(\sigma)=\sum_{i=1}^{\#\text{trials}}\log p\left(\hat{C}_i\,|\,s_i;\sigma\right)$$



Ma, Kording, Goldreich, *Bayesian modeling of behavior*

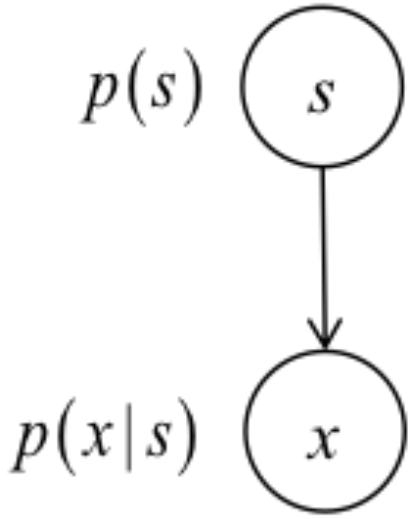
This will be a book published by Oxford University Press. It will appear in 2018.

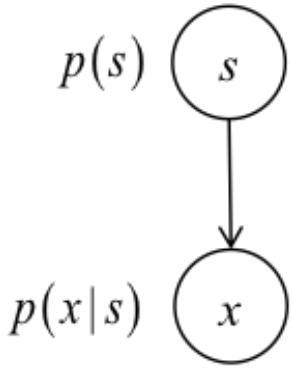
Really.

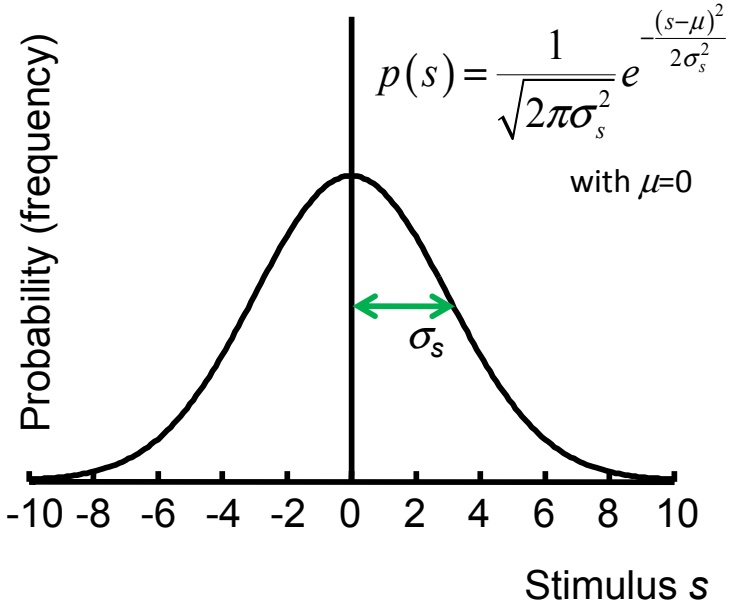
Sorry Konrad for my procrastination!!

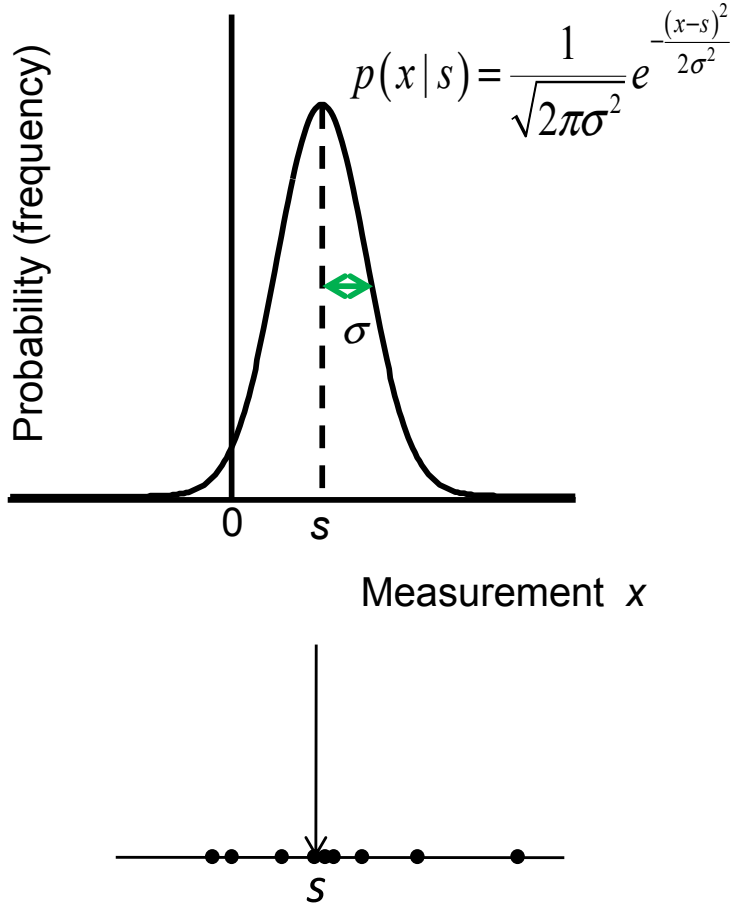
Example: auditory localization task

Step 1: Generative model



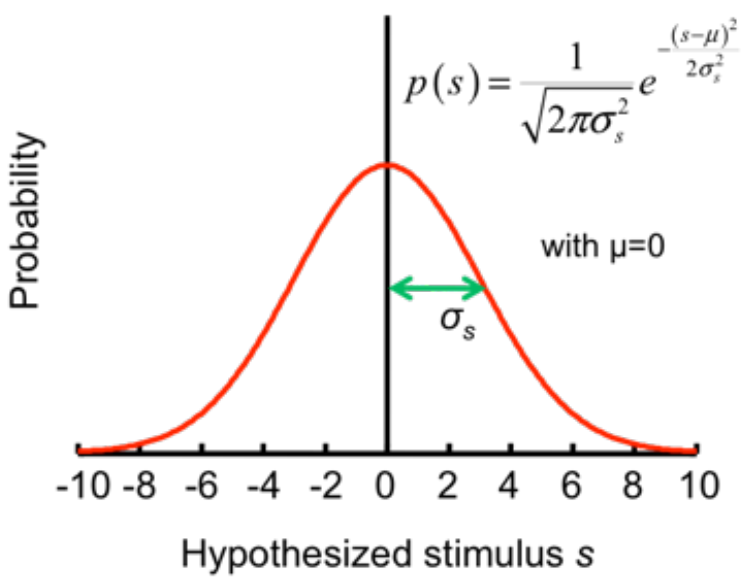

$p(s)$
s
$p(x|s)$
x


$p(s) = \frac{1}{\sqrt{2\pi\sigma_s^2}} e^{-\frac{(s-\mu)^2}{2\sigma_s^2}}$
with μ=0
Probability (frequency)
σ_s
-10 -8 -6 -4 -2 0 2 4 6 8 10
Stimulus s

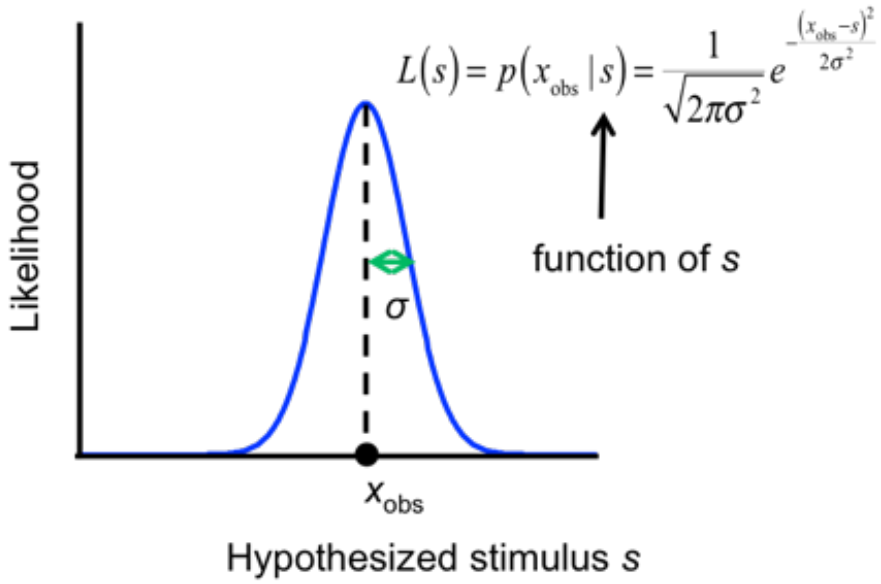

$p(x|s) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-s)^2}{2\sigma^2}}$
Probability (frequency)
σ
0
s
Measurement x
s

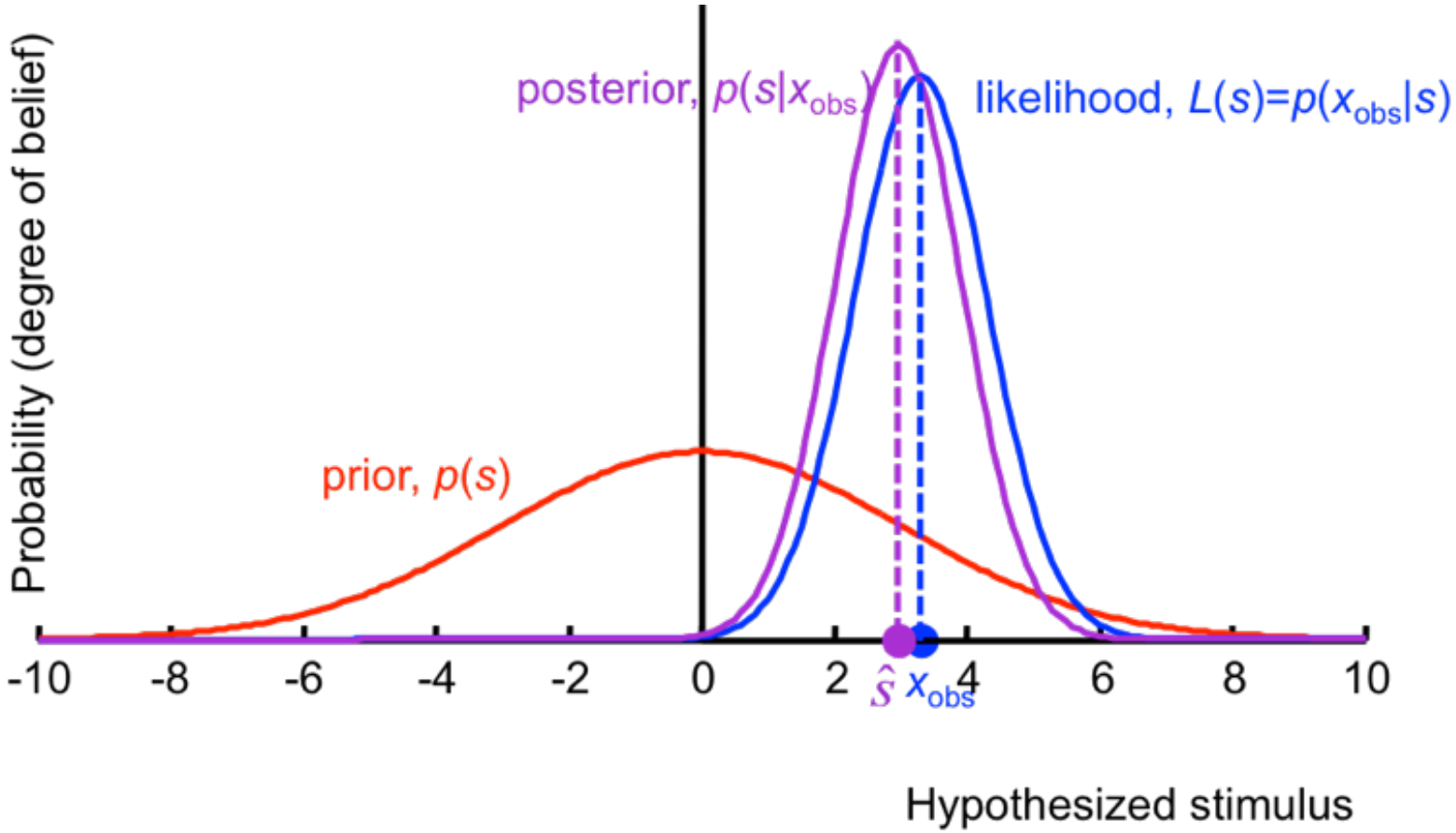
Step 2: Inference, deriving the decision rule

Prior

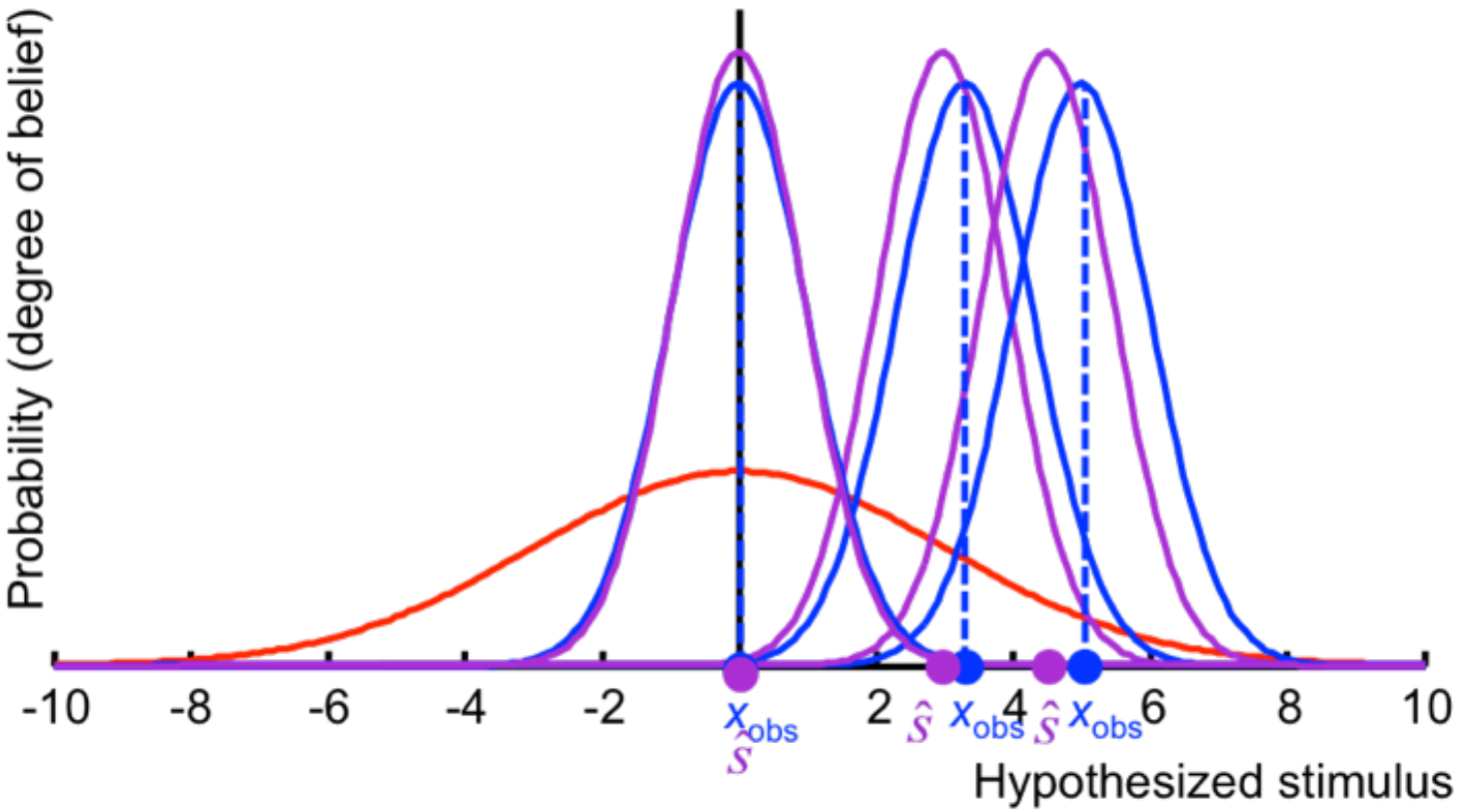


Likelihood





Does the model deterministically predict the posterior for a given stimulus and given parameters?


Probability (degree of belief)
-10
-8
-6
-4
-2
2
4
6
8
10
x_{obs}
$\hat{s}$
$\hat{s}$
x_{obs}
$\hat{s}$
x_{obs}
Hypothesized stimulus

Step 3: Response probabilities (predictions for your behavioral experiment)

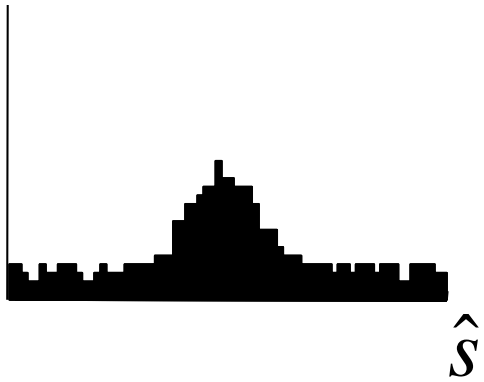
Decision rule: mapping $x \rightarrow \hat{s}$

But *x* is itself a random variable for given *s*

Therefore $\hat{s}$ is a random variable for given *s*

$$p(\hat{s}|s)$$

Can compare this to data!!



The four steps of Bayesian modeling

Example: categorization task

STEP 1: GENERATIVE MODEL

a) Draw a diagram with each node a variable and each arrow a statistical dependency. Observation is at the bottom.

b) For each variable, write down an equation for its probability distribution. For the observation, assume a noise model. For others, get the distribution from your experimental design. If there are incoming arrows, the distribution is a conditional one.

World state of interest
$p(C) = 0.5$

Stimulus
$p(s \mid C) = \mathrm{N}(s; \mu_C, \sigma_C^2)$

Observation
$p(x \mid s) = \mathrm{N}(x; s, \sigma^2)$

STEP 2: BAYESIAN INFERENCE (DECISION RULE)

a) Compute the posterior over the world state of interest given an observation. The optimal observer does this using the distributions in the generative model. Alternatively, the observer might assume different distributions (natural statistics, wrong beliefs). Marginalize (integrate) over variables other than the observation and the world state of interest.

$$p(C \mid s) \propto p(C)\, p(x \mid C) = p(C) \int p(x \mid s)\, p(s \mid C)\, ds = \ldots = \mathrm{N}(x; \mu_C, \sigma^2 + \sigma_C^2)$$

b) Specify the read-out of the posterior. Assume a utility function, then maximize expected utility under posterior. (Alternative: sample from the posterior.) Result: decision rule (mapping from observation to decision). When utility is accuracy, the read-out is to maximize the posterior (MAP decision rule).

$$\hat{C} = 1 \text{ when } \mathrm{N}(x; \mu_1, \sigma^2 + \sigma_1^2) > \mathrm{N}(x; \mu_2, \sigma^2 + \sigma_2^2)$$

STEP 3: RESPONSE PROBABILITIES

For every unique trial in the experiment, compute the probability that the observer will choose each decision option given the stimuli on that trial using the distribution of the observation given those stimuli (from Step 1) and the decision rule (from Step 2).

$$p(\hat{C} = 1 \mid x) = \Pr_{x \mid s;\sigma}\left(\mathrm{N}(x; \mu_1, \sigma^2 + \sigma_1^2) > \mathrm{N}(x; \mu_2, \sigma^2 + \sigma_2^2)\right)$$

- Good method: sample observation according to Step 1; for each, apply decision rule; tabulate responses. Better: integrate numerically over observation. Best (when possible): integrate analytically.
- Optional: add response noise or lapses.

STEP 4: MODEL FITTING AND MODEL COMPARISON

a) Compute the parameter log likelihood, the log probability of the subject's actual responses across all trials for a hypothesized parameter combination.

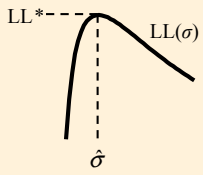
$$\mathrm{LL}(\sigma) = \sum_{i=1}^{\#\text{trials}} \log p(\hat{C}_i \mid s_i; \sigma)$$

b) Maximize the parameter log likelihood. Result: parameter estimates and maximum log likelihood. Test for parameter recovery and summary statistics recovery using synthetic data.

c) Obtain fits to summary statistics by rerunning the fitted model.

d) Formulate alternative models (e.g. vary Step 2). Compare maximum log likelihood across models. Correct for number of parameters (e.g. AIC). (Advanced: Bayesian model comparison, uses log marginal likelihood of model.) Test for model recovery using synthetic data.

e) Check model comparison results using summary statistics.

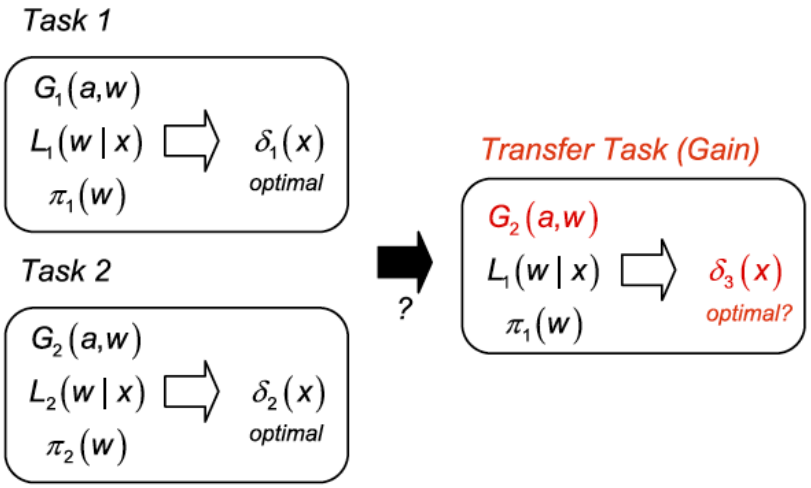


Bayesian models are about:

- the decision-maker making the best possible decision (given an objective function)
- **the brain representing probability distributions**

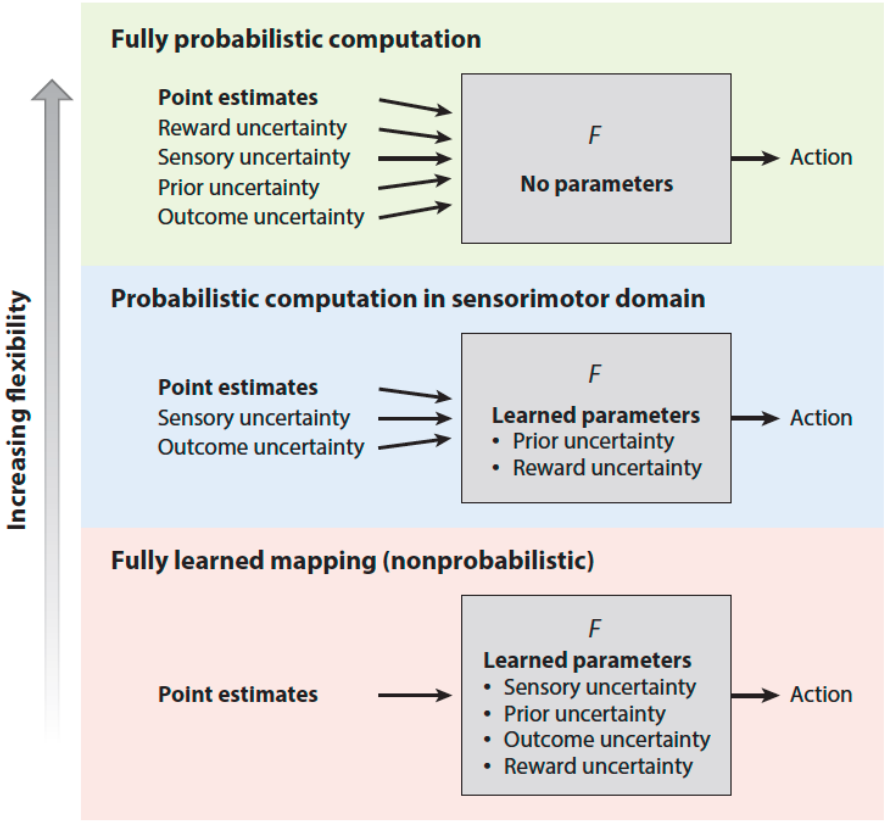
Does the brain represent probability distributions?

Bayesian transfer



Maloney and Mamassian, 2009

Different degrees of probabilistic computation



Ma and Jazayeri, 2014

Part 2: Model fitting

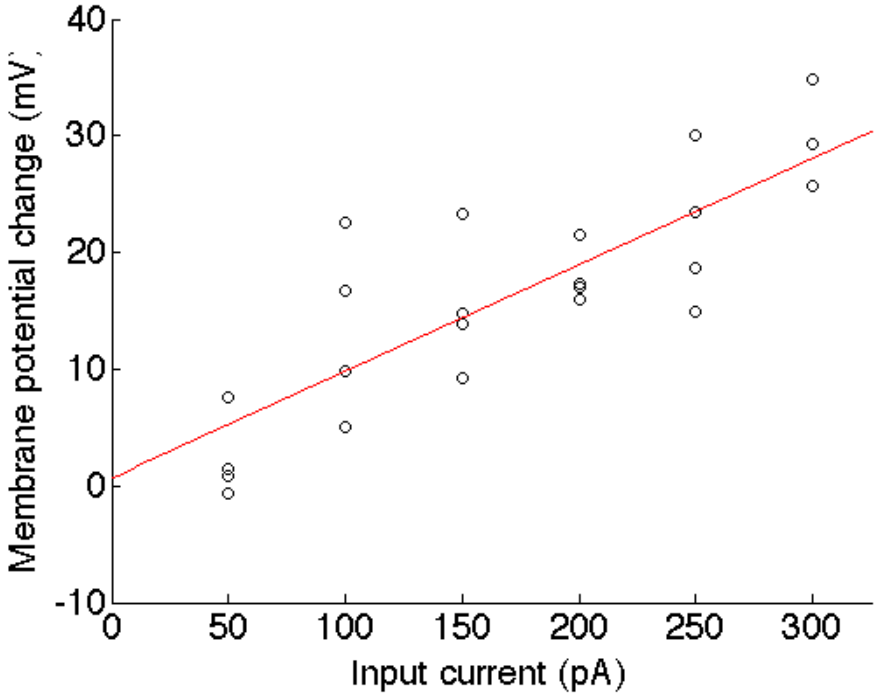
2a. What to minimize/maximize when fitting parameters?

2b. What fitting algorithm to use?

2c. Validating your model fitting method

2a. What to minimize/maximize when fitting a model?

Try #1: Minimize sum squared error



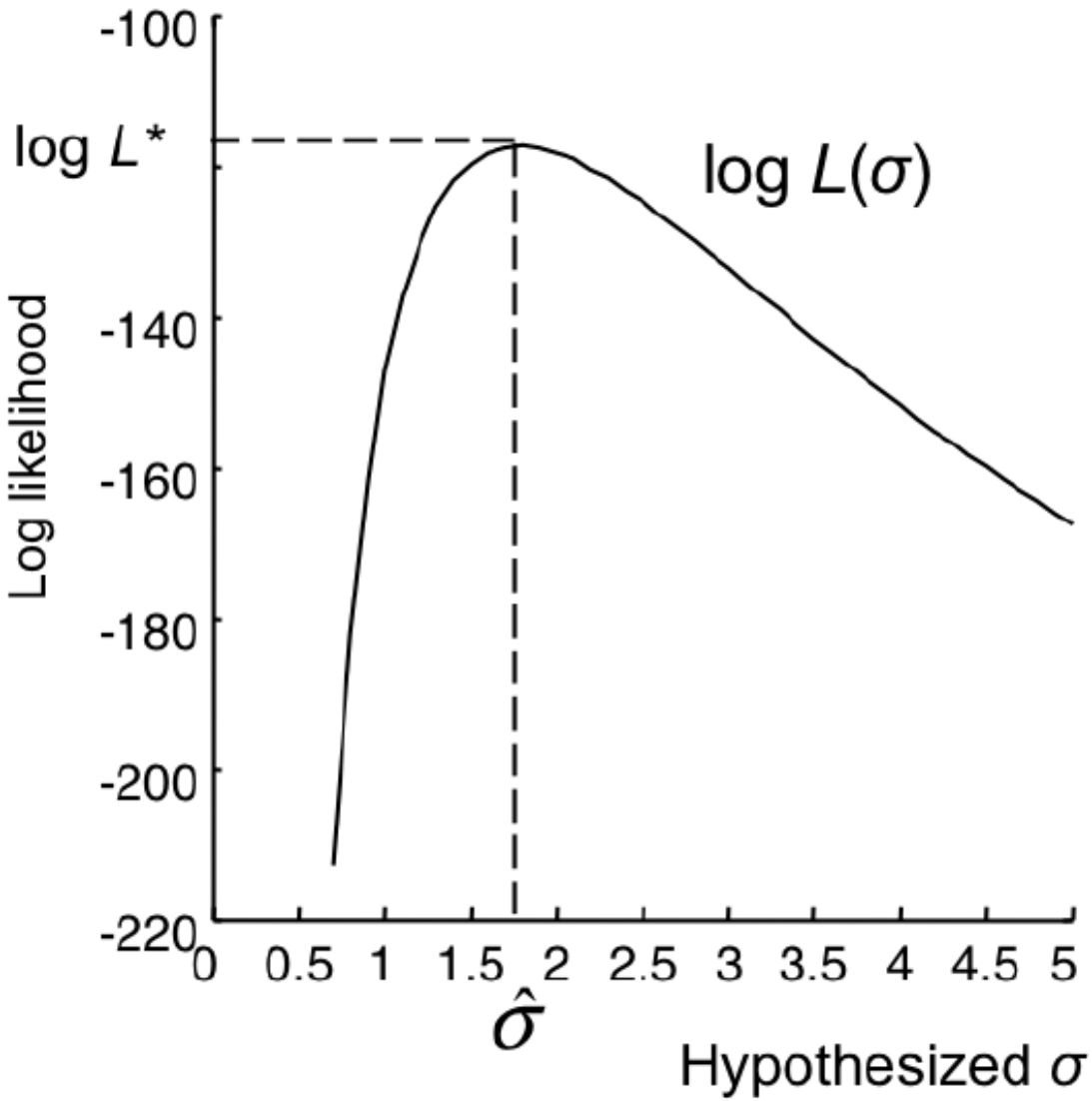
Only principled if your model is linear
Otherwise arbitrary and suboptimal

Try #2: Maximize likelihood

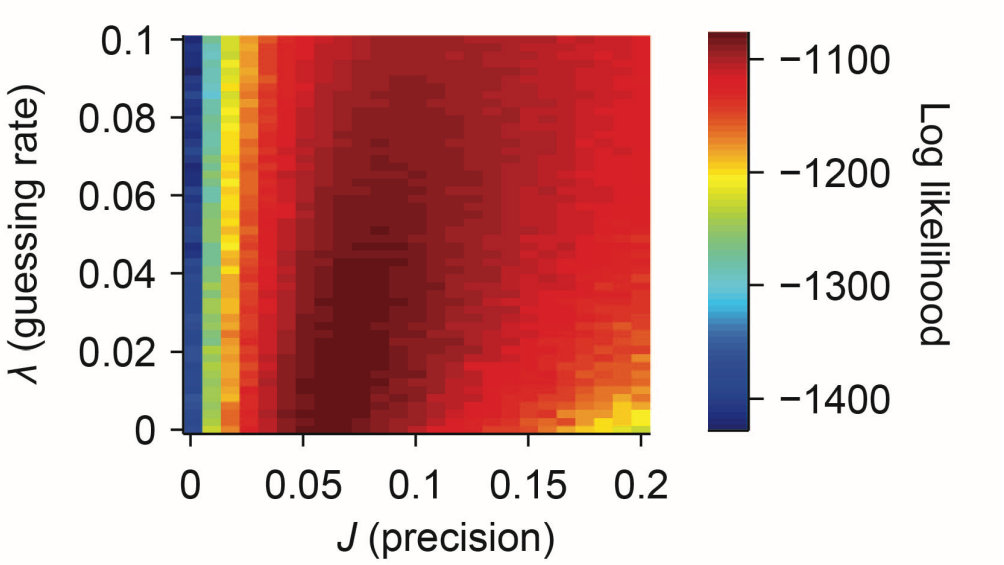
Output of Step 3:
***p*(response I stimulus, parameter combination)**

Likelihood of parameter combination
= *p*(data | parameter combination)

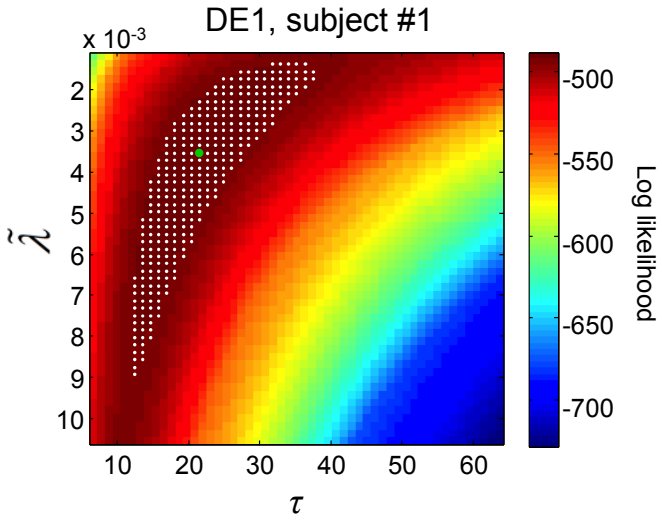
$$= \prod_{\text{trials } i} p\left(\text{response}_i \middle| \text{stimulus}_i, \text{ parameter combination}\right)$$


-100
log L*
log L(σ)
Log likelihood
-140
-160
-180
-200
-220
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
σ̂
Hypothesized σ

Parameter trade-offs



Shen and Ma, http://www.biorxiv.org/content/early/2017/06/22/153650



Van den Berg and Ma, data from http://www.biorxiv.org/content/early/2017/06/18/151365

2b. What fitting algorithm to use?

#usebads !

Bayesian Adaptive Direct Search

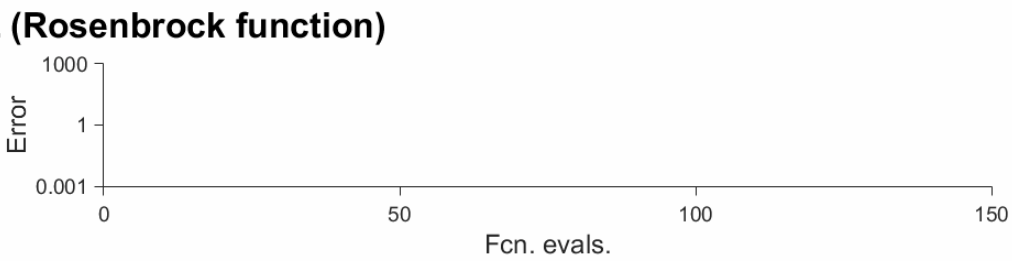
W61 Model Fitting Under Uncertainty: A Practical Analysis of Derivative-Free Optimization for Cognitive and Computational Neuroscience

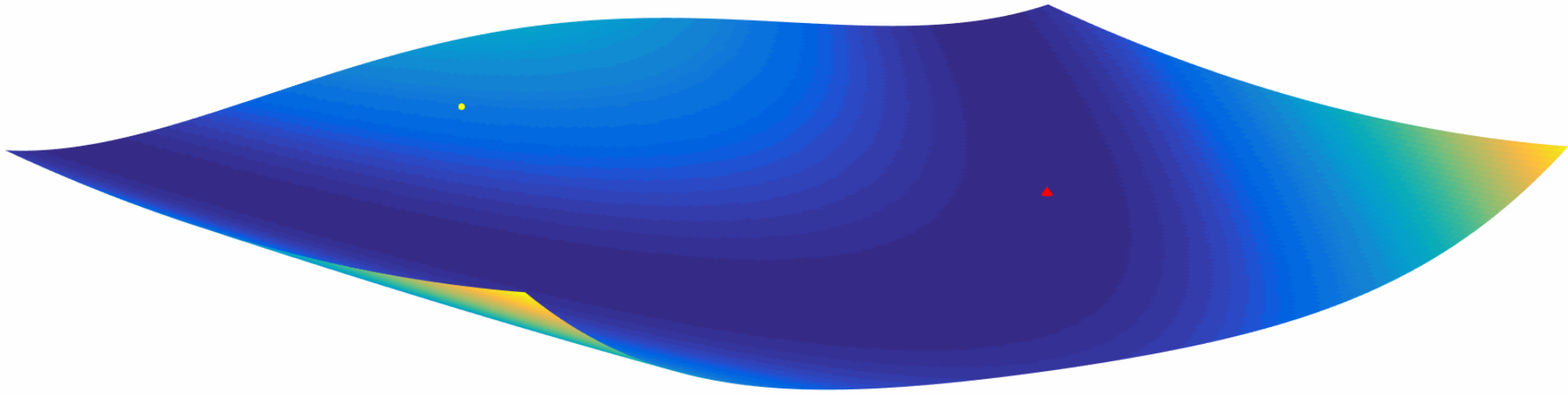
Co-Authors

Luigi Acerbi, Wei Ji Ma

https://arxiv.org/abs/1705.04405

fminsearch



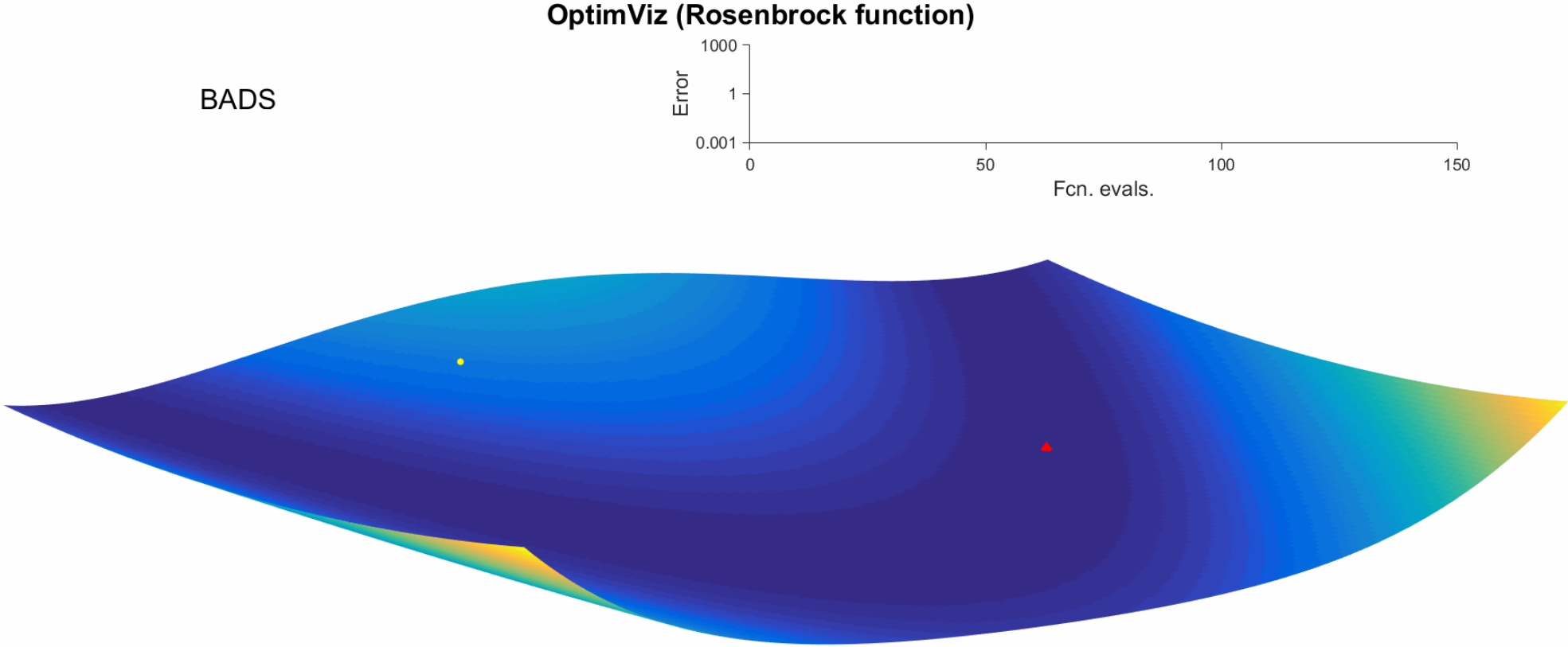


#useBADS

https://github.com/lacerbi/bads

https://github.com/lacerbi/optimviz

BADS



#useBADS

https://github.com/lacerbi/bads

https://github.com/lacerbi/optimviz

Bayesian Adaptive Direct Search (BADS) optimization algorithm for model fitting in MATLAB

optimization-algorithms bayesian-optimization log-likelihood noiseless-functions noisy-functions matlab

146 commits **2** branches **5** releases **1** contributor GPL-3.0

Branch: **master** New pull request Find file Clone or download

lacerbi committed on **GitHub** Update README.md Latest commit `fa8b054` on Jul 19

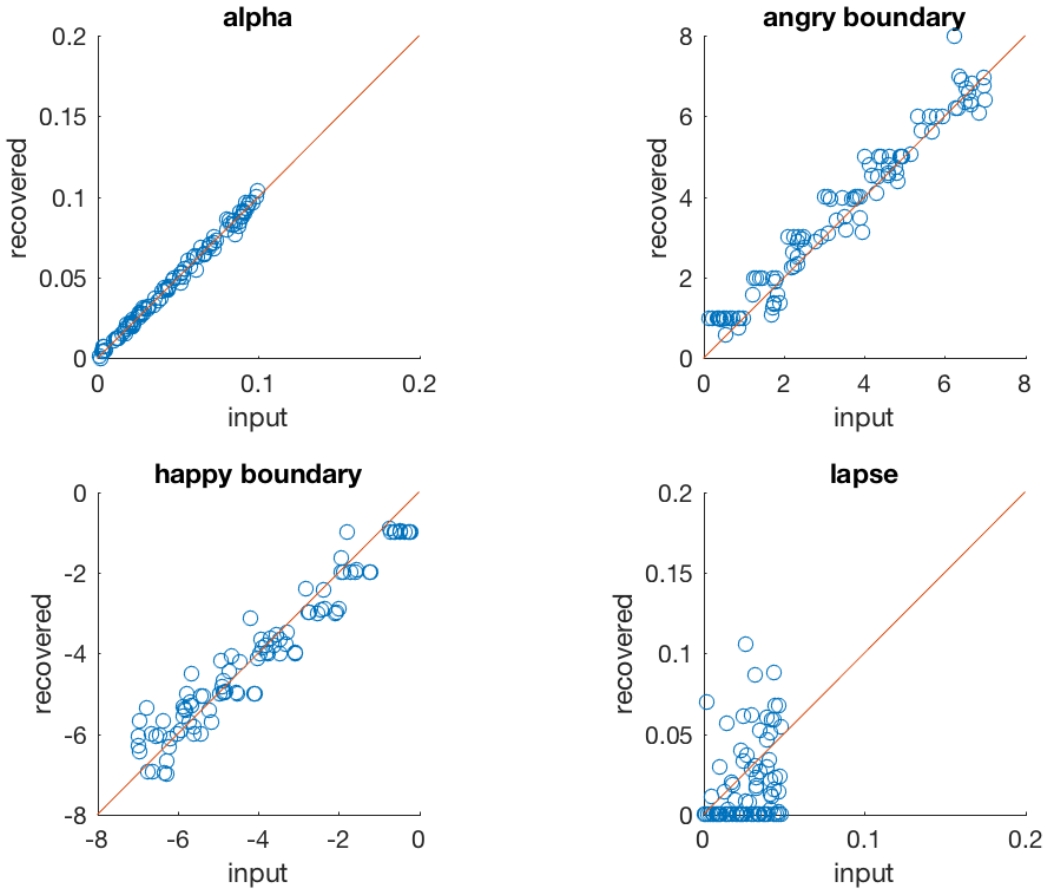
acq	Update README.md	4 months ago
docs	added arXiv reference	4 months ago
gpdef	fixed prctile1	4 months ago
gpml-matlab-v3.6-2015-07-07	moved files	6 months ago
gpml_fast	Update README.md	4 months ago
init	renamed file	4 months ago
poll	Update README.md	4 months ago
private	Fixed LB/UB starting point issue	2 months ago
search	Update README.md	4 months ago
utils	fixed prctile1	4 months ago
warp	Update README.md	4 months ago
.gitignore	first commit	6 months ago
LICENSE.txt	Create LICENSE.txt	5 months ago
README.md	Update README.md	2 months ago
bads.m	Fixed LB/UB starting point issue	2 months ago
bads_examples.m	fixed bug with fixed vars and output fcn	3 months ago
install.m	print readme after installation	4 months ago
rosenbrocks.m	improved documentation; added examples	4 months ago

Model fitting best practices

- If you can, maximize the likelihood (probability of single-trial responses) if you can.
  - *Do not* minimize squared error!
  - *Do not* fit summary statistics (but the raw data)!
- Use more than one algorithm
  - Grid search
  - Fmincon
  - BADS
- Multistart

2c. Validating your method: Parameter recovery

parameter recovery test



Jenn Laura Lee

Part 3: Model comparison

3a. Choosing a model comparison metric

3b. Validating your model comparison method

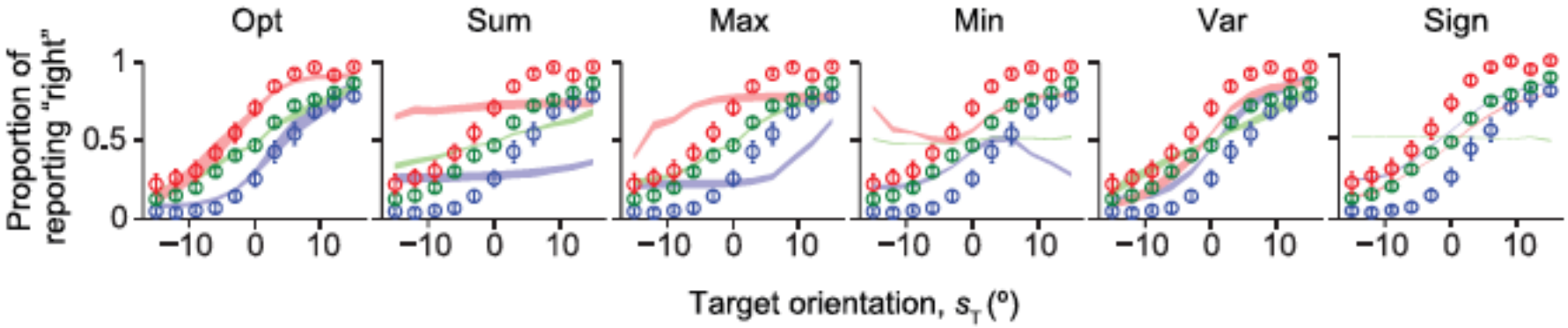
3c. Factorial model comparison

3d. Absolute goodness of fit

3e. Heterogeneous populations

3a. Choosing a model comparison metric

Try #1: Visual similarity to the data



Shen and Ma, 2016

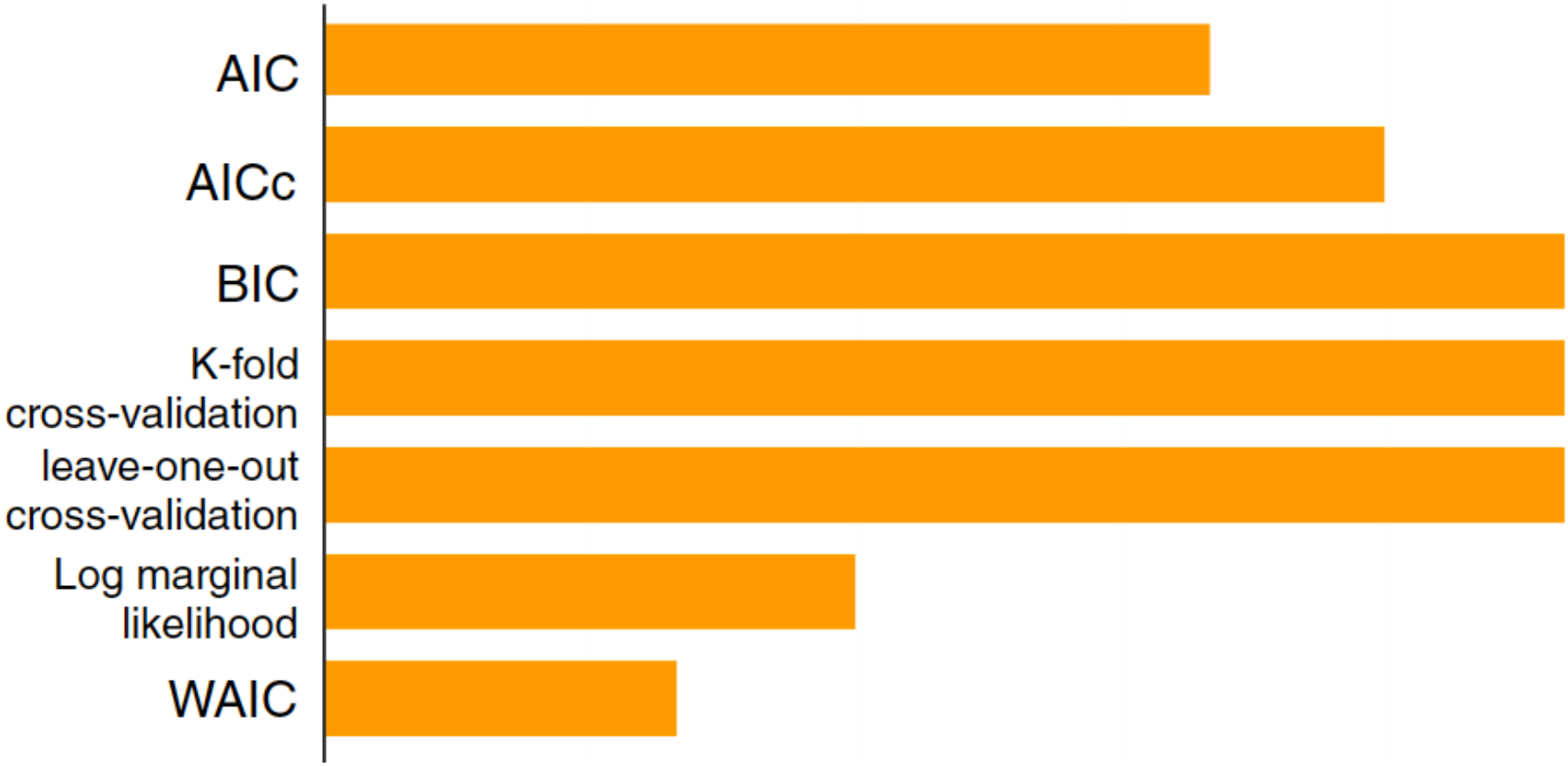
Fine, but not very quantitative

Try #2: R^2

- Just don’t do it
  - Unless you have only linear models
    - Which almost never happens

Try #3: Likelihood-based metrics

Good!
Problem: there are many!



From Ma lab survey by Bas van Opheusden, 201703

Metrics based on *maximum likelihood:*

- Akaike Information Criterion (AIC or AICc)
- Bayesian Information Criterion (BIC)

Metrics based on *the full likelihood function* (often sampled using Markov Chain Monte Carlo)*:*

- Marginal likelihood (model evidence, Bayes' factor)
- Watanabe-Akaike Information criterion

Cross-validation can be either

Metrics based on *explanation:*

- Bayesian Information Criterion (BIC)
- Marginal likelihoods (model evidence, Bayes' factors)

Metrics based on *prediction:*

- Akaike Information Criterion (AIC or AICc)
- Watanabe-Akaike Information criterion
- Most forms of cross-validation

Practical considerations:

- No metric is always unbiased for finite data.
- AIC tends to underpenalize free parameters, BIC tends to overpenalize.
- **Do not trust conclusions that are metric-dependent**. Report multiple metrics if you can.

	AICc*(model) − AICc*(VP)	BIC*(model) − BIC*(VP)		LML(model) − LML(VP)	
Model	Mean	Mean	Standard error of the mean	Mean	Standard error of the mean
IL					
M1	−125	−122	15	−121	15
M2	−183	−180	18	−180	18
M3	−167	−164	18	−163	18
Humans	−47.2	−45.7	6.8	−47.1	6.6
EP					
M1	−47.5	−44.8	9.2	−48.9	9.1
M2	−12.8	−10.1	4.6	−12.7	4.8
M3	−30.3	−27.6	7.8	−31.3	8.1
Humans	−12.9	−11.4	1.5	−14.4	1.7
EPF					
M1	−40.2	−40.2	7.9	−39.0	7.8
M2	−9.3	−9.3	4.4	−6.7	4.6
M3	−24.0	−24.0	6.7	−22.6	6.9
Humans	−7.6	−7.6	1.5	−6.2	1.6
VPF					
M1	−1.3	−4.18	0.83	1.5	1.5
M2	−2.2	−4.00	0.91	1.20	0.81
M3	−0.56	−3.2	1.5	2.0	1.1
Humans	−1.46	−3.00	0.32	−0.57	0.31

Devkar, Wright, Ma 2015

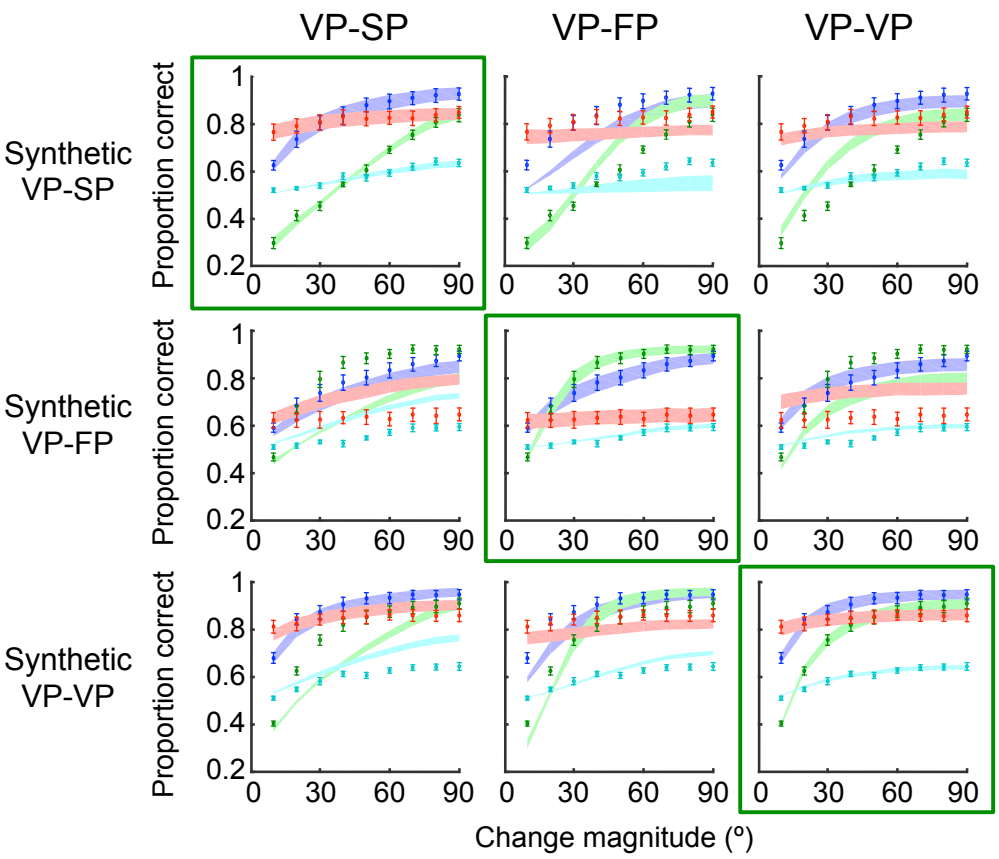
Challenge: your model comparison metric and how you compute it might have issues. How to validate it?

3b. Model recovery

Model recovery example

Data generation model

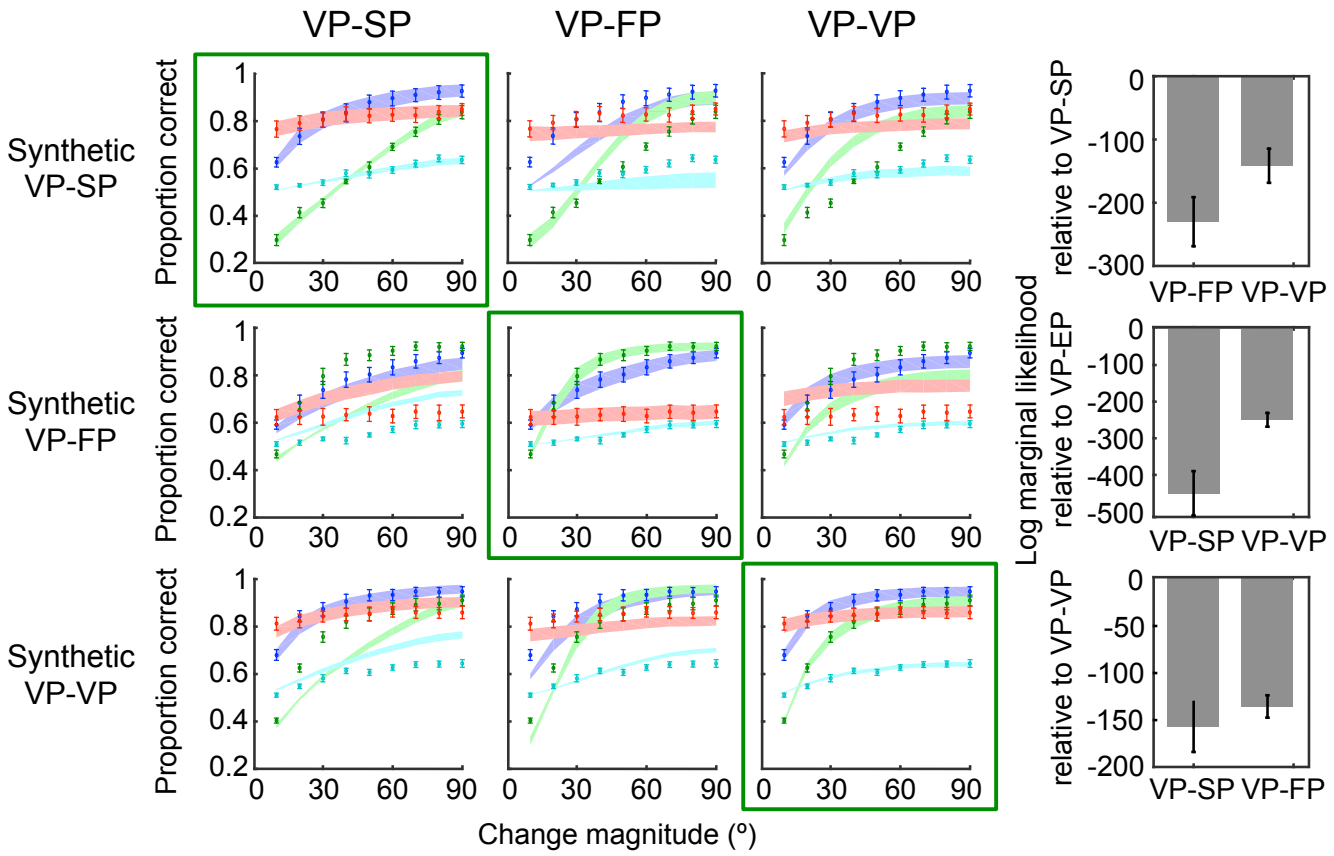
Fitted model



Devkar, Wright, Ma, Journal of Vision, in press

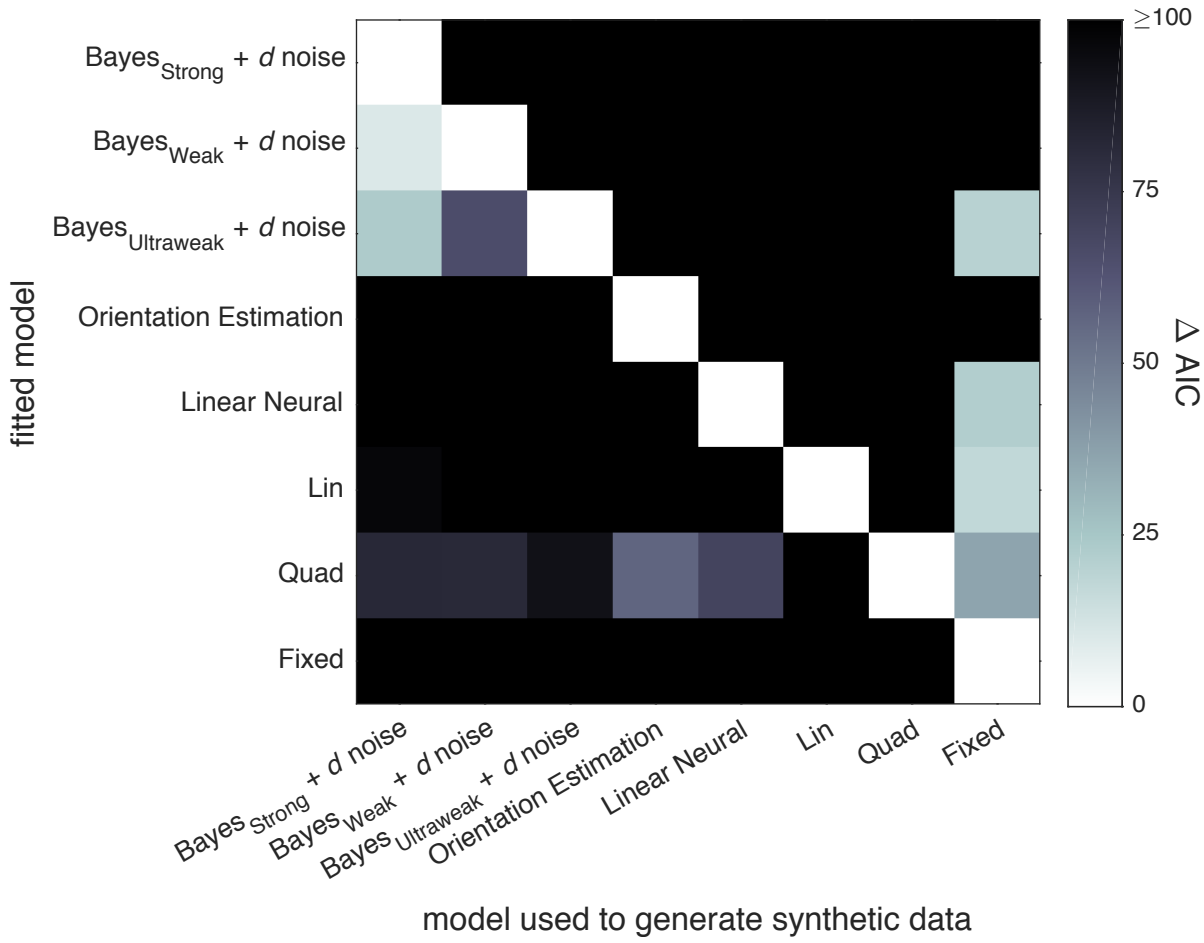
Fitted model

Data generation model



Devkar, Wright, Ma, Journal of Vision, in press

Model recovery



Will Adler, http://www.biorxiv.org/content/early/2016/12/11/093203

Challenge: how to avoid “handpicking” models?

3c. Factorial model comparison

3c. Factorial model comparison

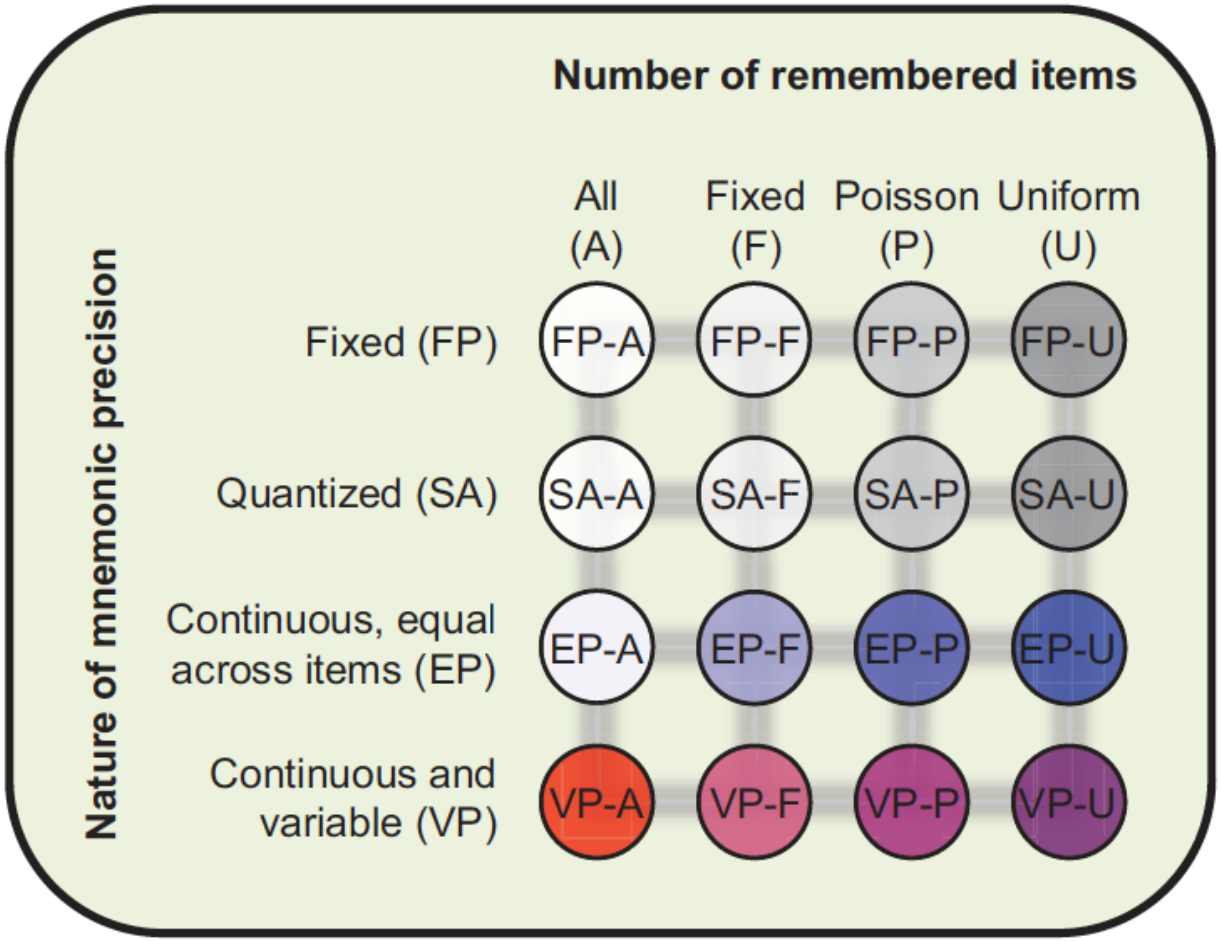
- Models often have many “moving parts”, components that can be in or out
- Similar to factorial design of experiments, one can mix and match these moving parts.
- References:
  - Acerbi, Vijayakumar, Wolpert 2014
  - Van den Berg, Awh, Ma 2014
  - See also Contributed Talk #14 (Mingyu Song)

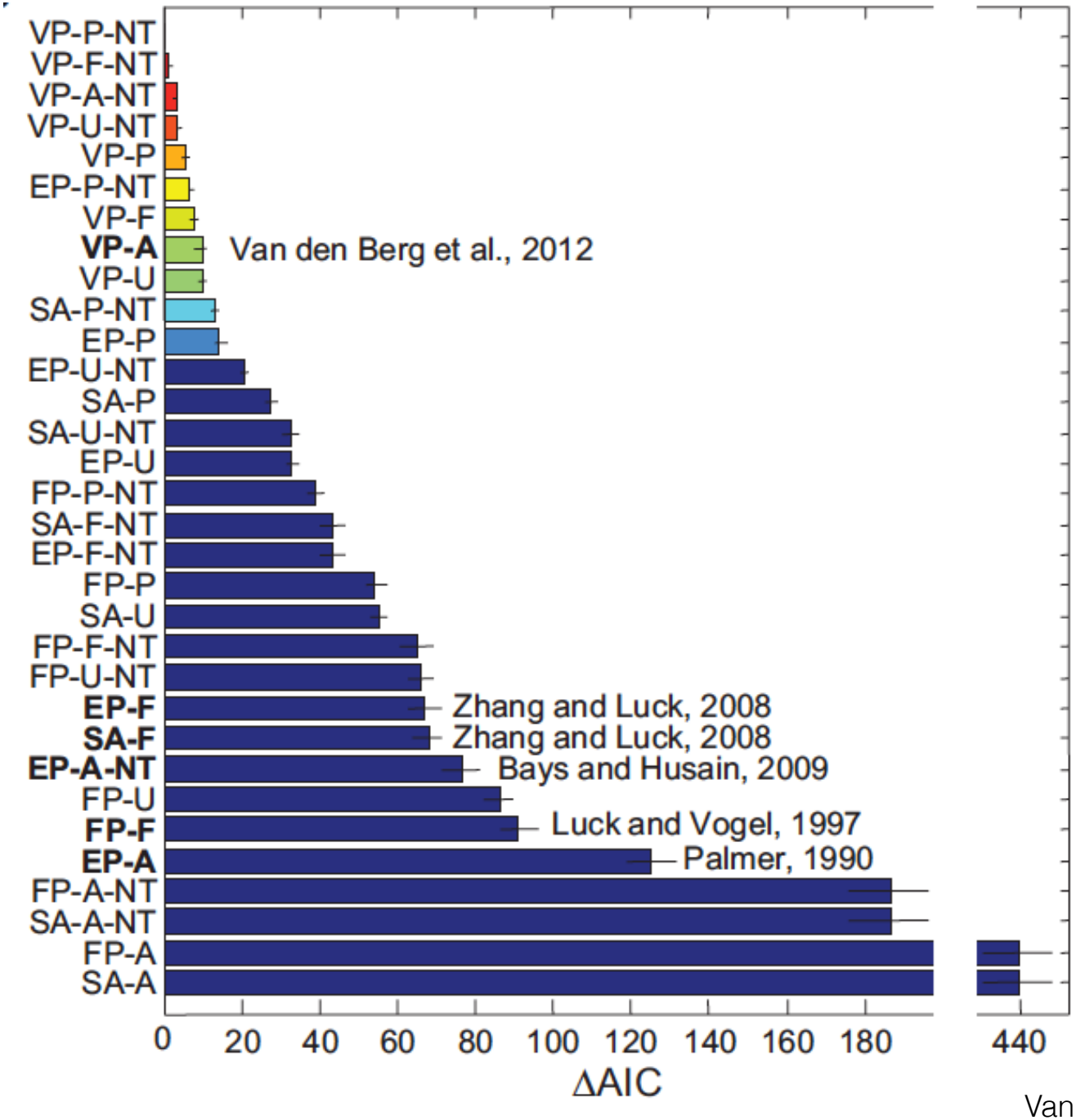
Factorial Comparison of Working Memory Models

Ronald van den Berg
University of Cambridge and Baylor College of Medicine

Edward Awh
University of Oregon

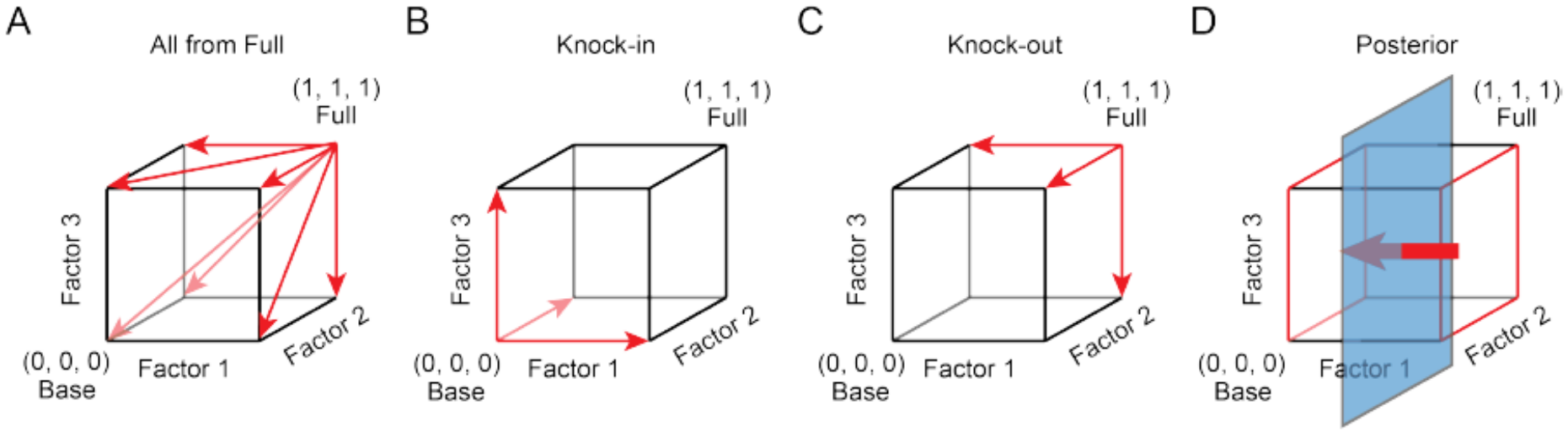
Wei Ji Ma
New York University and Baylor College of Medicine





Van den Berg, Awh, Ma 2014

Challenge: how to summarize the results?



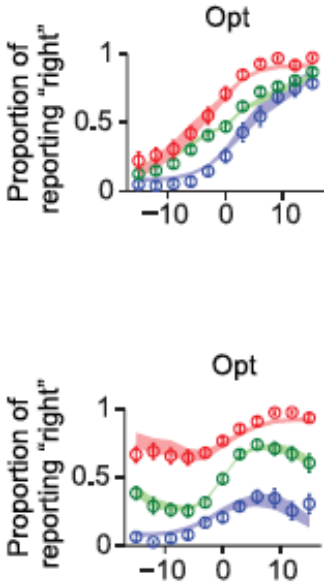
Shen and Ma, http://www.biorxiv.org/content/early/2017/06/22/153650

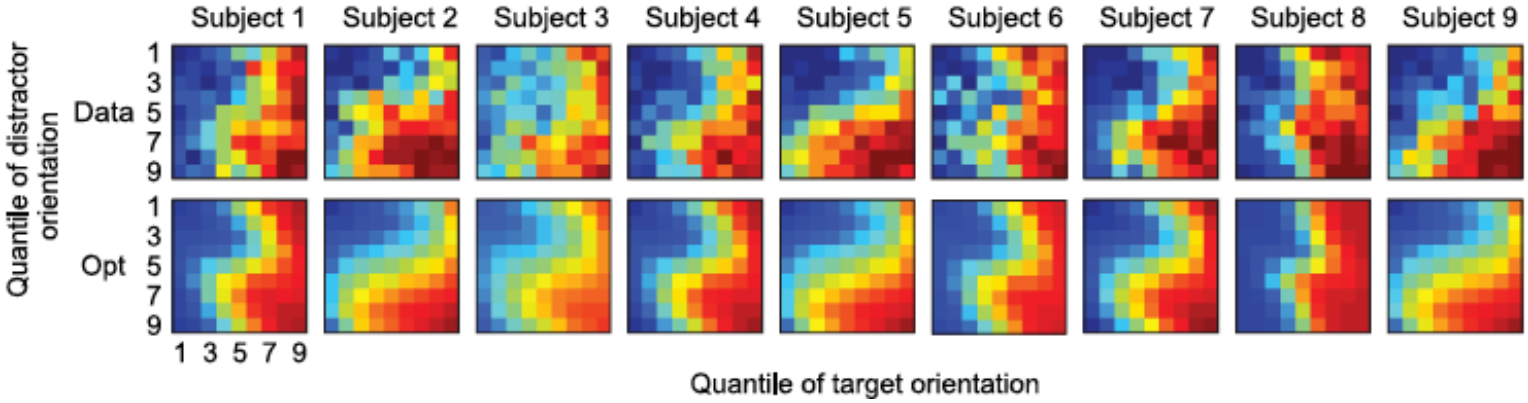
Challenge: the *best* model is not necessarily a *good* model.

3d. Absolute goodness of fit

Absolute goodness of fit

- How close is the *best* model to the data?
- Method 1: Visual inspection (model checking)

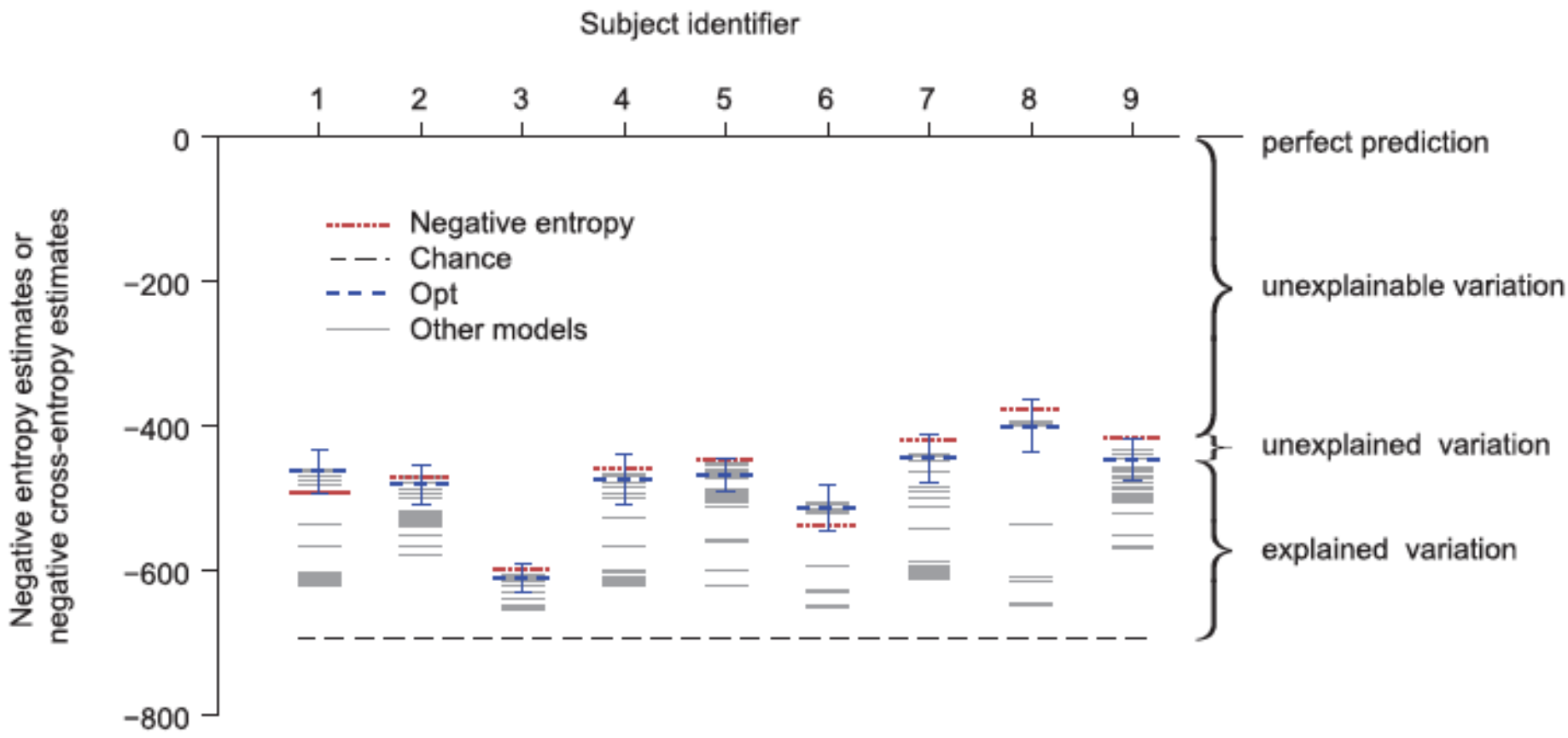




Shen and Ma, 2016

3d. Absolute goodness of fit

- Method 2: Deviance / negative entropy
  - There is irreducible, unexplainable variation in the data
  - This sets an upper limit on the goodness of fit of *any* model: negative entropy
  - How far away is a model from this upper bound?
  - Wichmann and Hill (2001)
  - Shen and Ma (2016)



Shen and Ma (2016)

Challenge: what if different subjects follow different models? (heterogeneity in the population)

3e. Hierarchical model selection

Consider all possible partitions of your population

Bayesian model selection for group studies

Klaas Enno Stephan [a,b,*], Will D. Penny [a], Jean Daunizeau [a], Rosalyn J. Moran [a], Karl J. Friston [a]

Neuroimage, 2009

Bayesian model selection for group studies – Revisited

L. Rigoux [a], K.E. Stephan [b,c], K.J. Friston [b], J. Daunizeau [a,b,*]

Neuroimage, 2014

- Returns probability that each model is the most common one in a population
- Returns posterior probability for each model
- Matlab code available online!
- Example application: Poster T25 (Maija Honig)

Model building

1a. What kind of model - descriptive or process?

1b. A special kind of process model - Bayesian

1c. Prior examples: visual illusions

1d. Likelihood example: Gestalt perception

1e. How to actually do Bayesian modeling?

Model fitting

2a. What to minimize/ maximize when fitting parameters?

2b. What fitting algorithm to use?

2c. Validating your model fitting method

Model comparison

3a. Choosing a model comparison metric

3b. Validating your model comparison method

3c. Factorial model comparison

3d. Absolute goodness of fit

3e. Heterogeneous populations

Good job everyone!!