Experimental Design
for Nonlinear Problems

Andrew Curtis

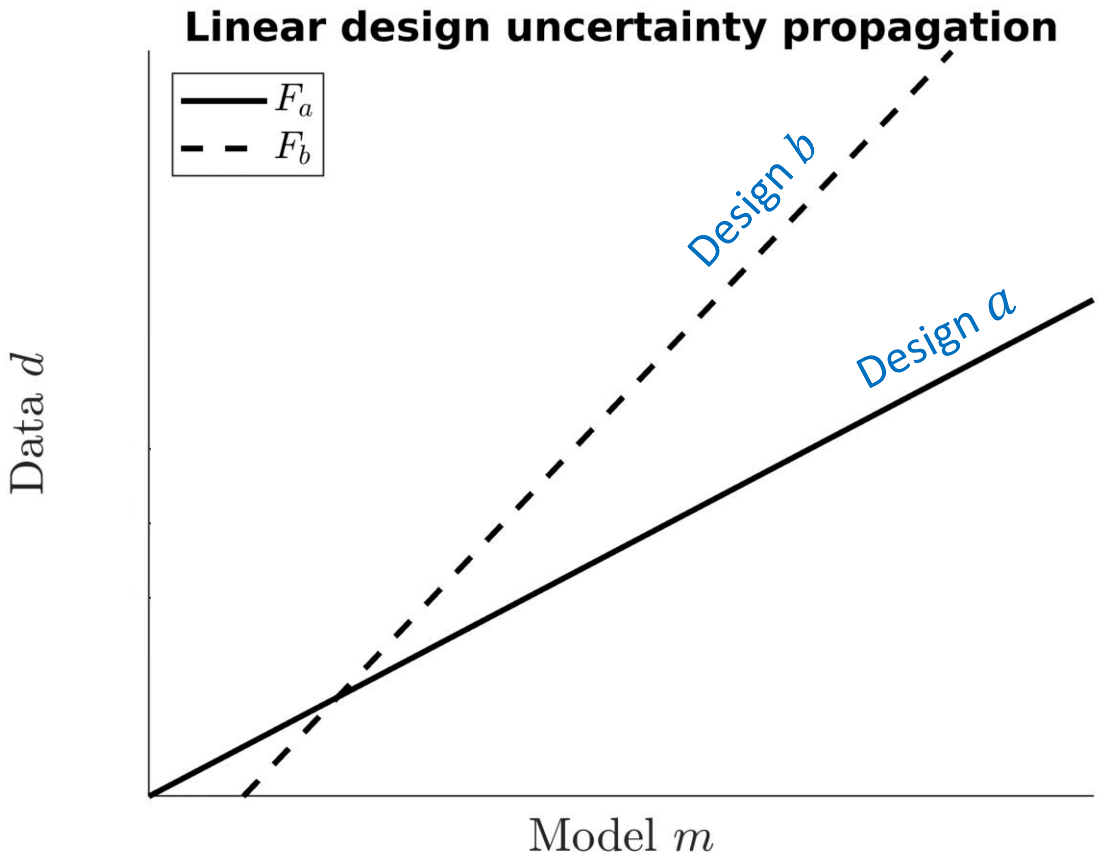
What is the best method to design receiver geometries?

- Experimental design concepts
  - Measures of 'design quality'
  - Design Algorithm → maximise quality
- Tests & Results

→ Choosing a Design Method

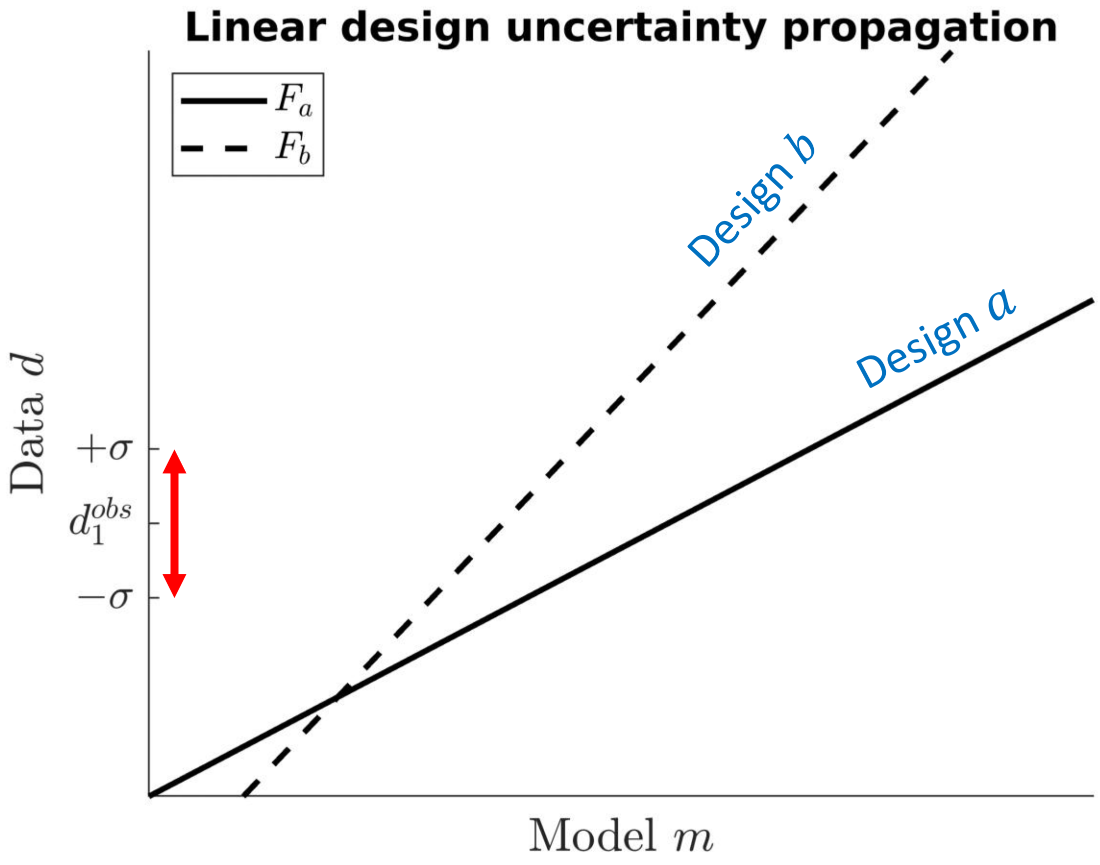
Experimental design theory

Experimental design theory



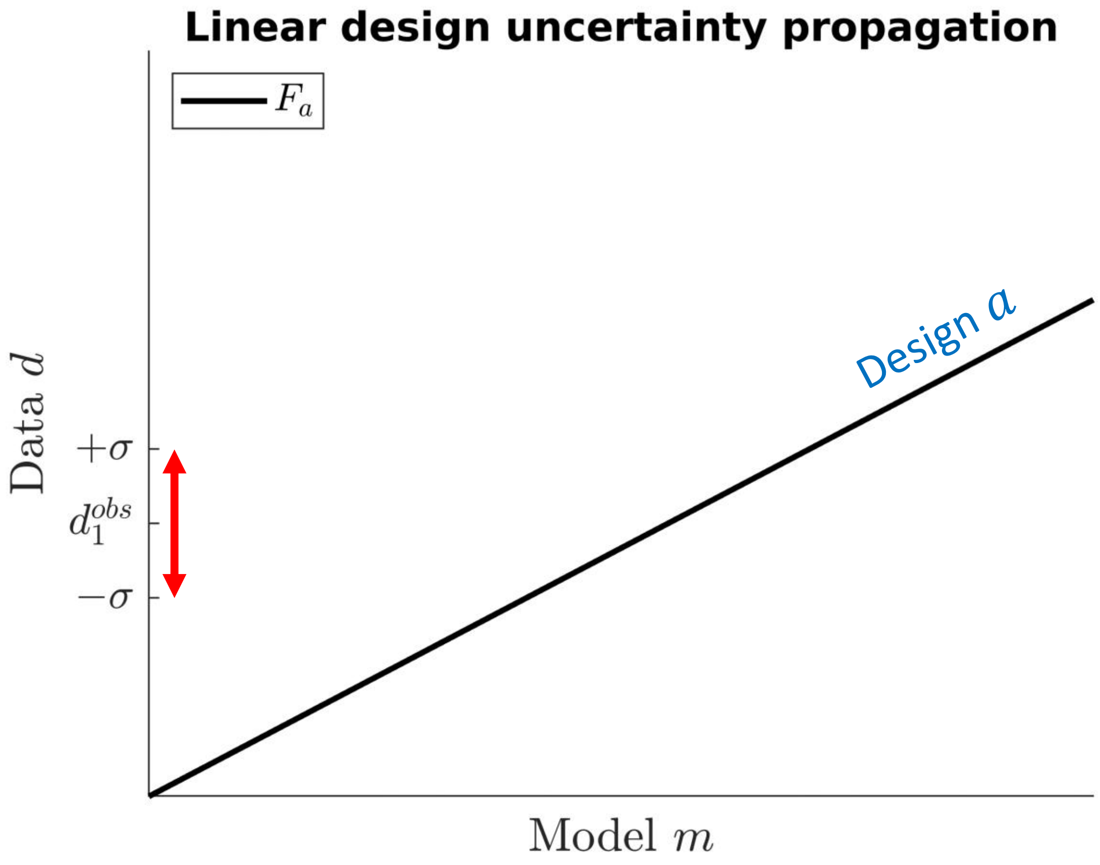
Which design is better?

Experimental design theory



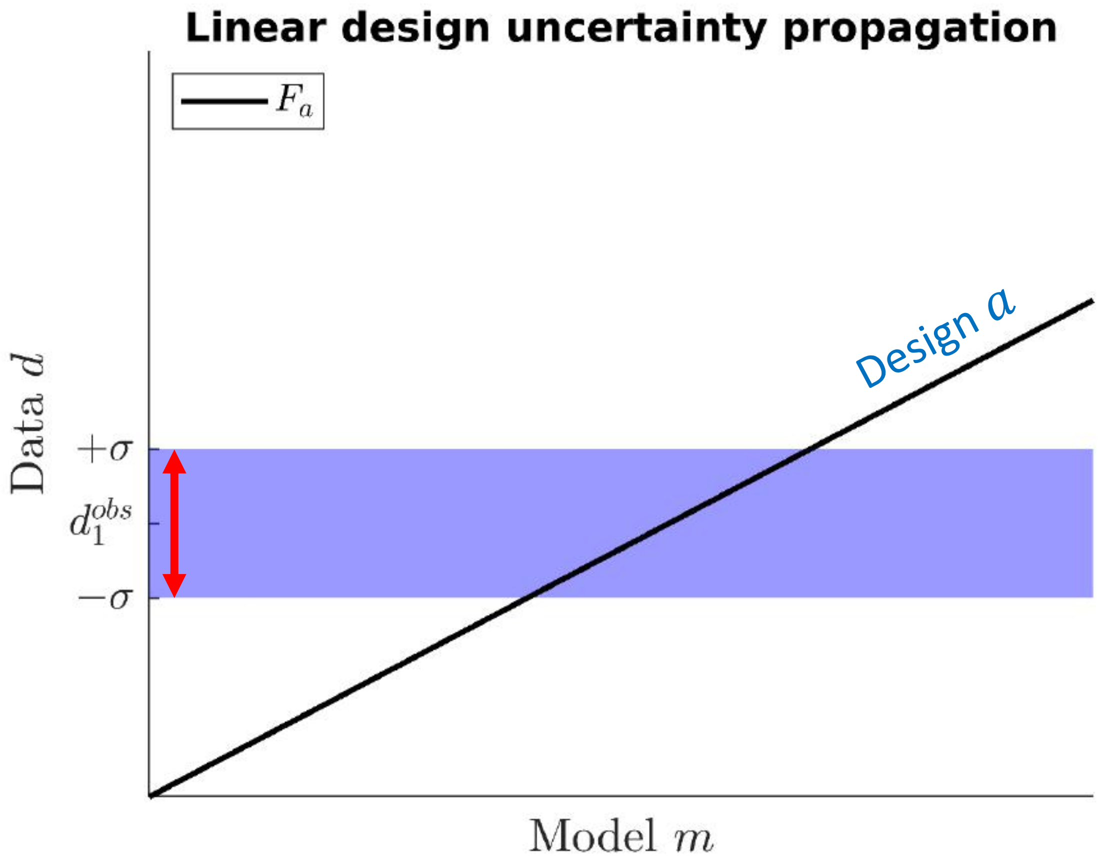
Which design is better?

Experimental design theory



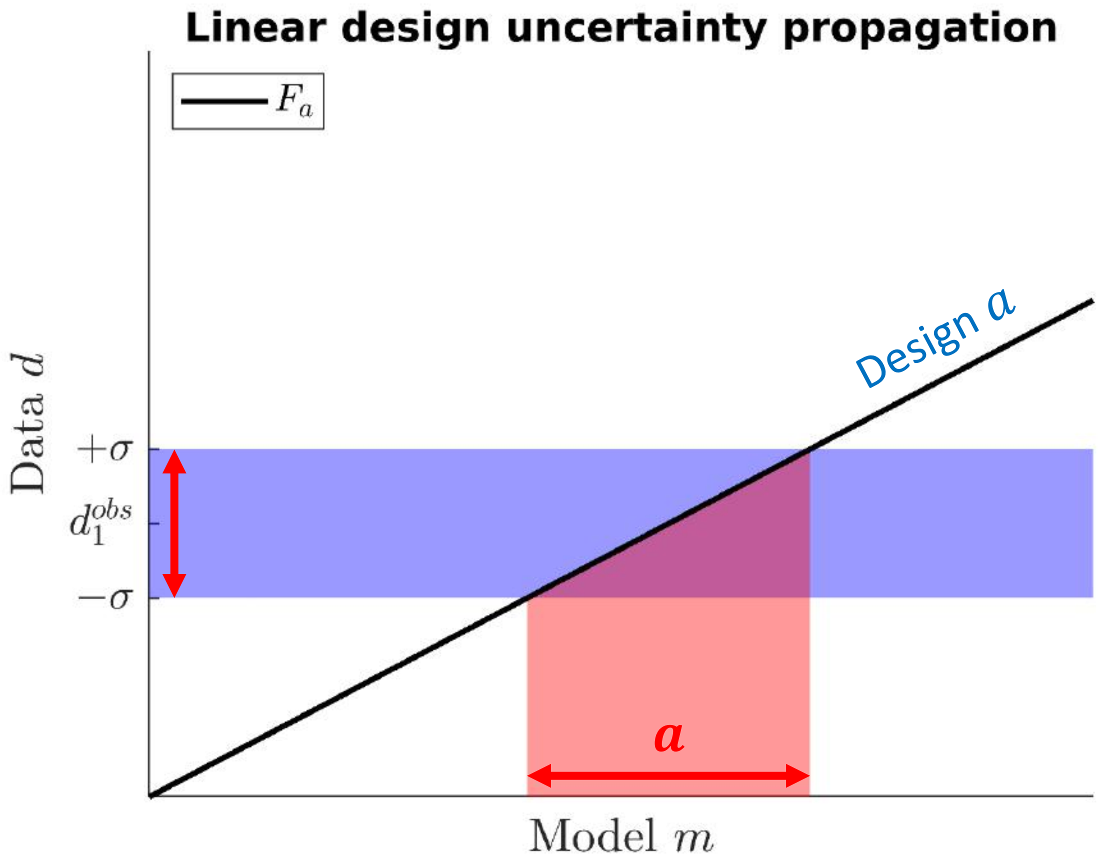
Which design is better?

Experimental design theory



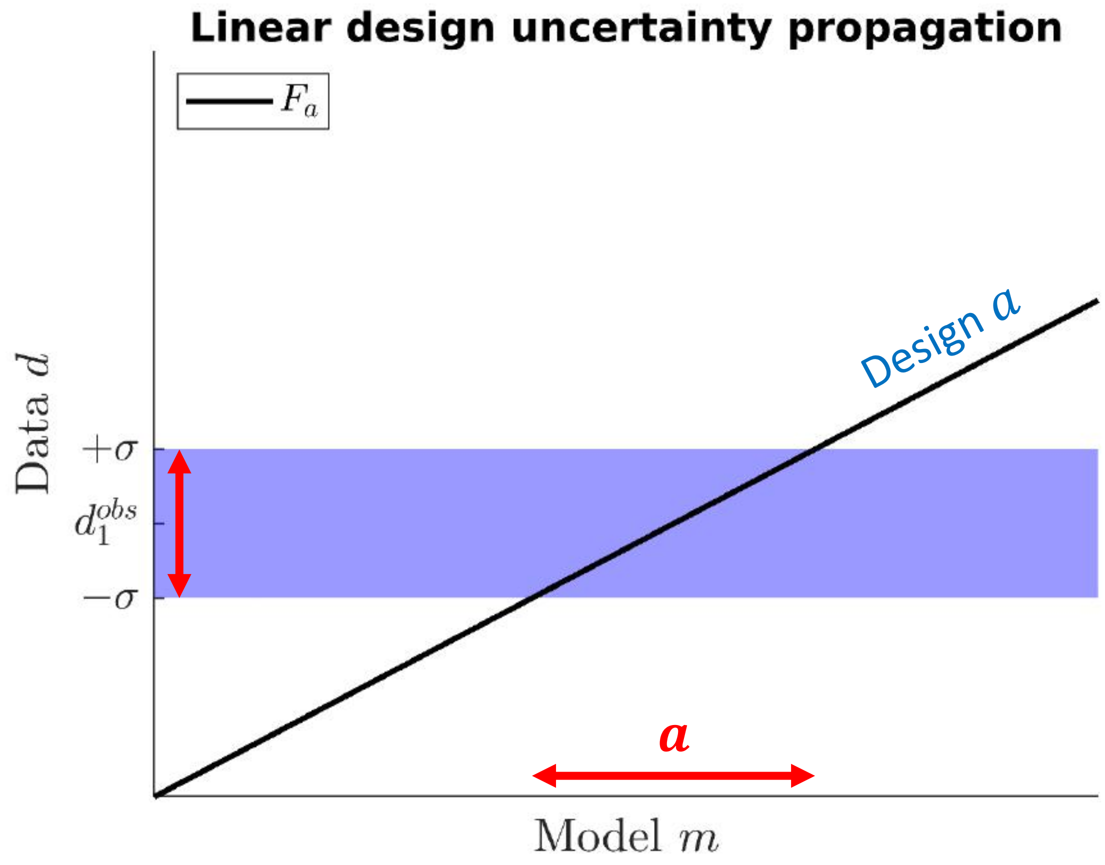
Which design is better?

Experimental design theory



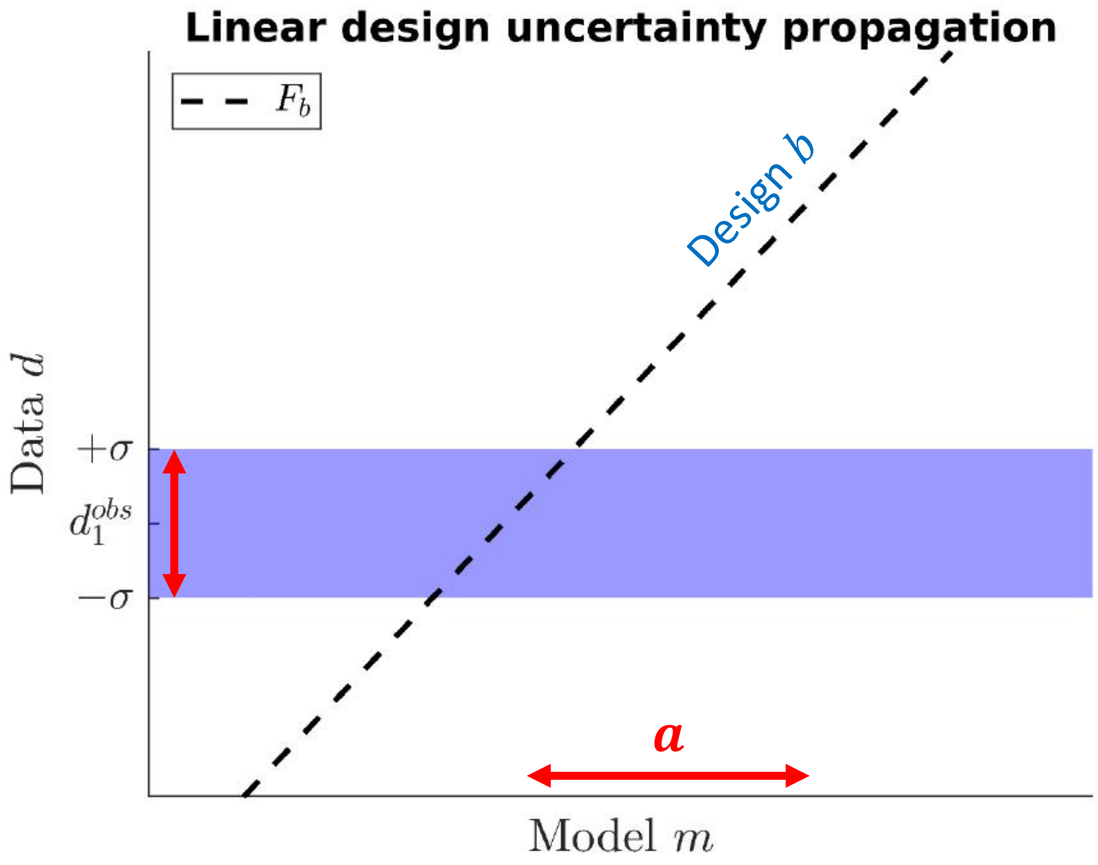
Which design is better?

Experimental design theory



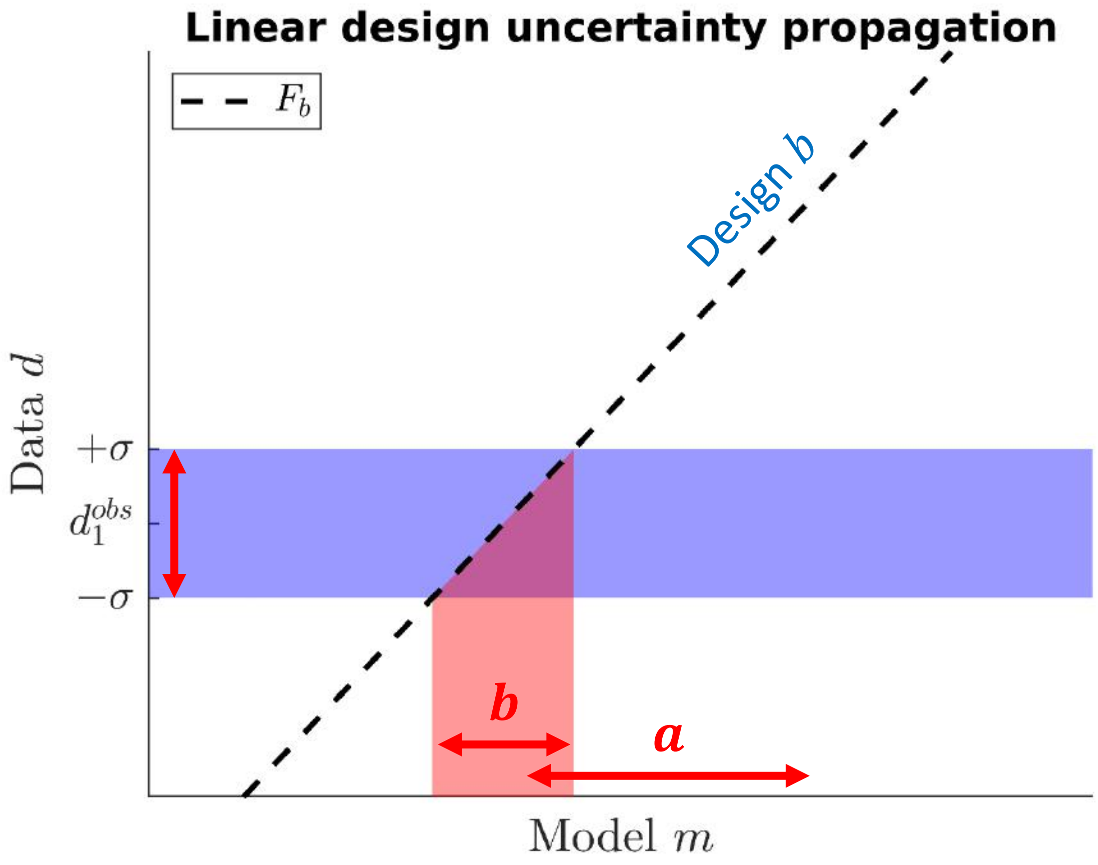
Which design is better?

Experimental design theory



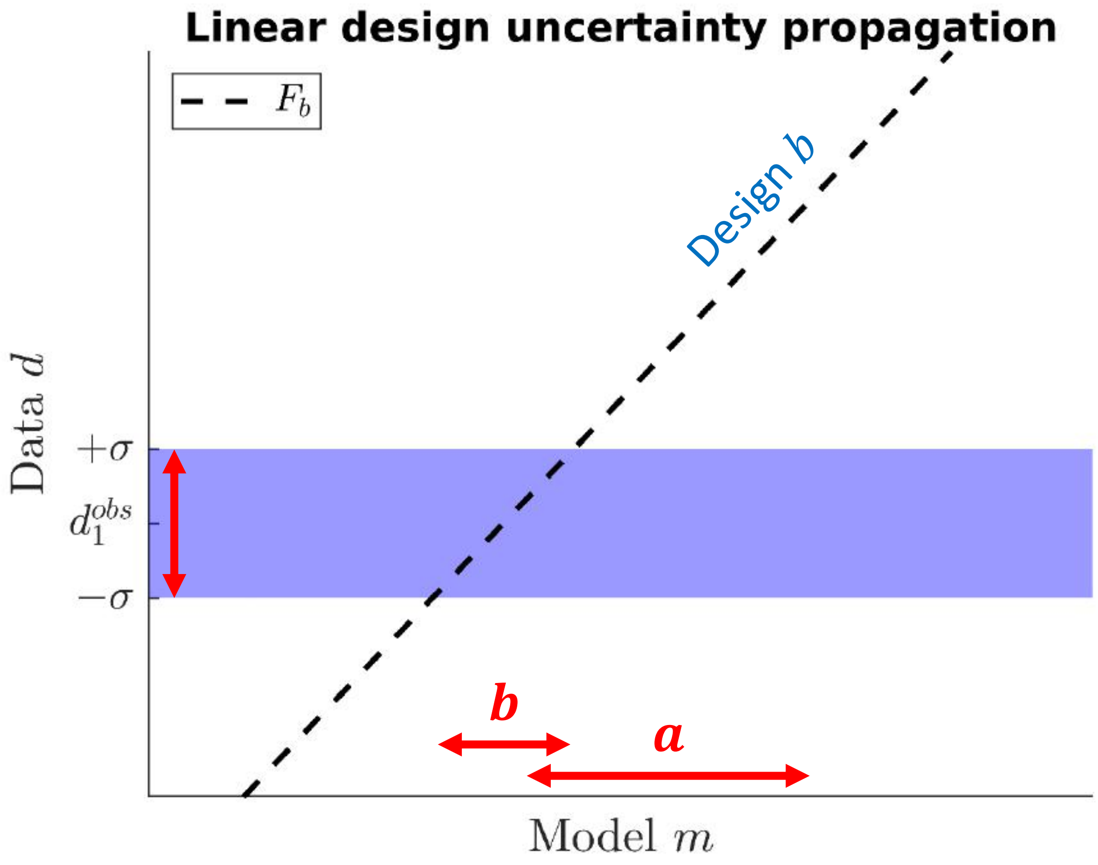
Which design is better?

Experimental design theory



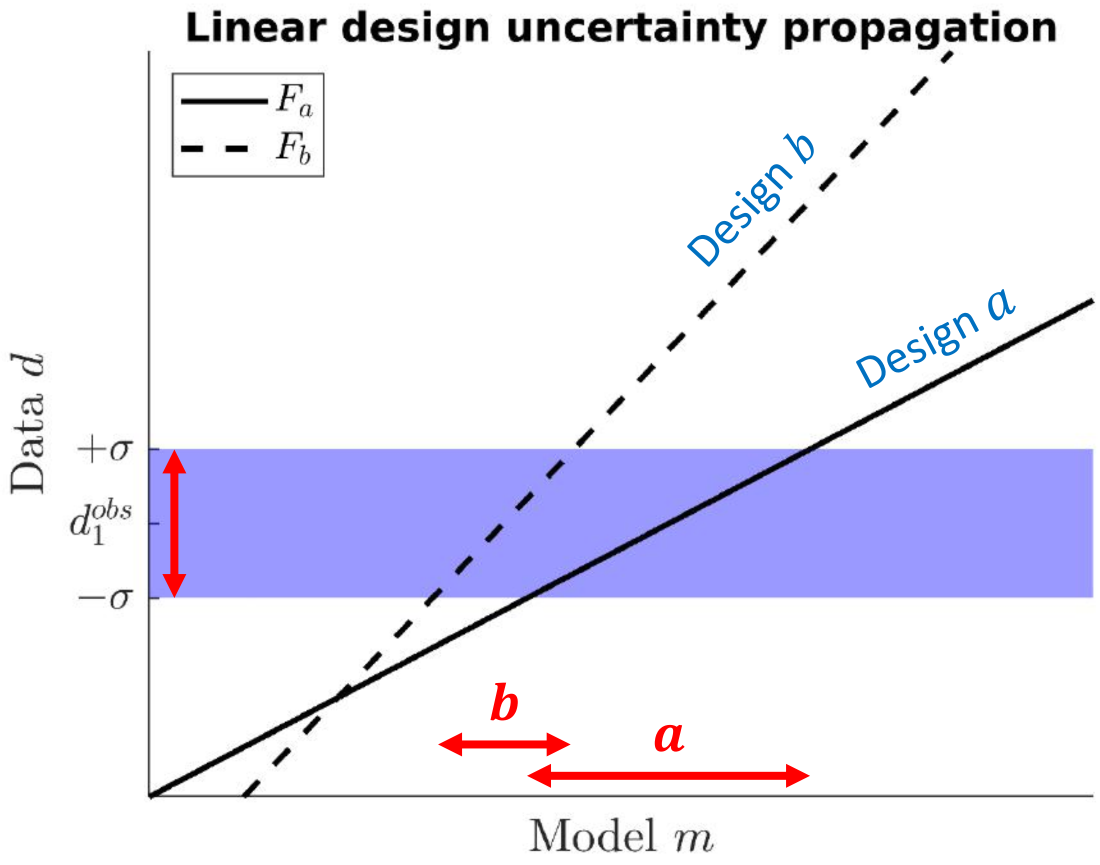
Which design is better?

Experimental design theory



Which design is better?

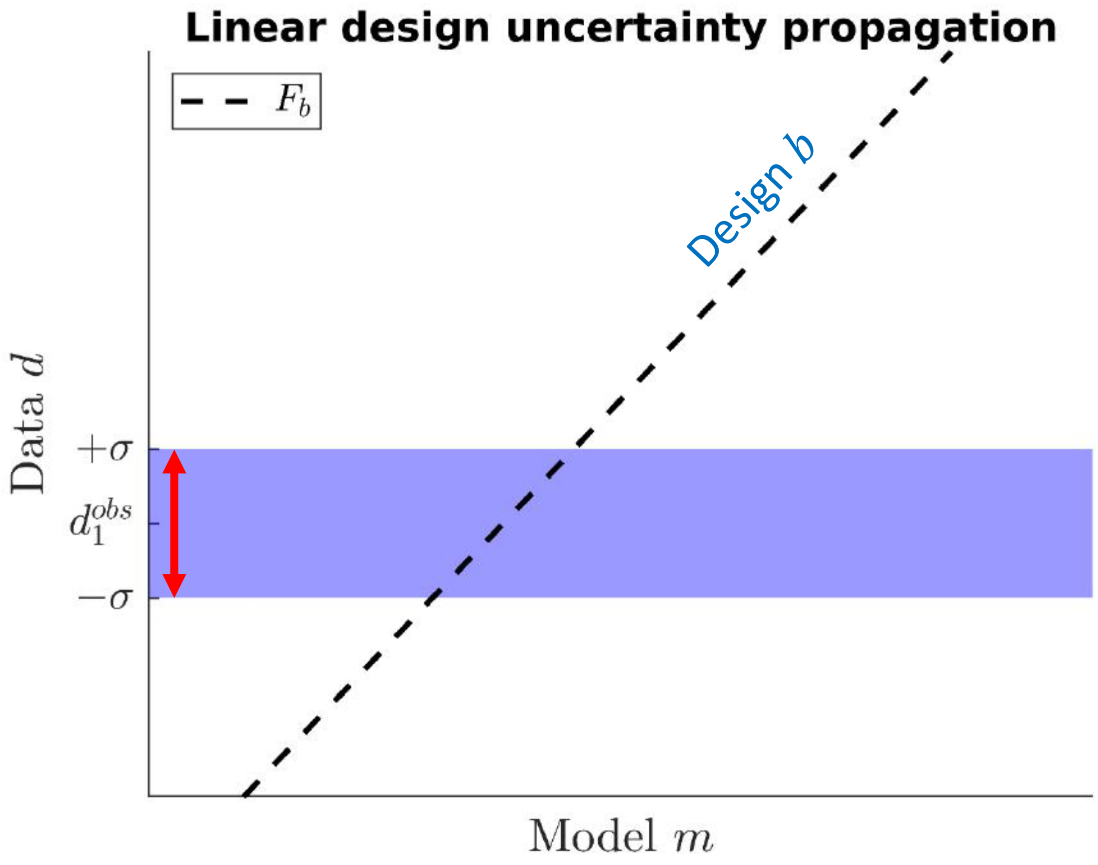
Experimental design theory



Which design is better?

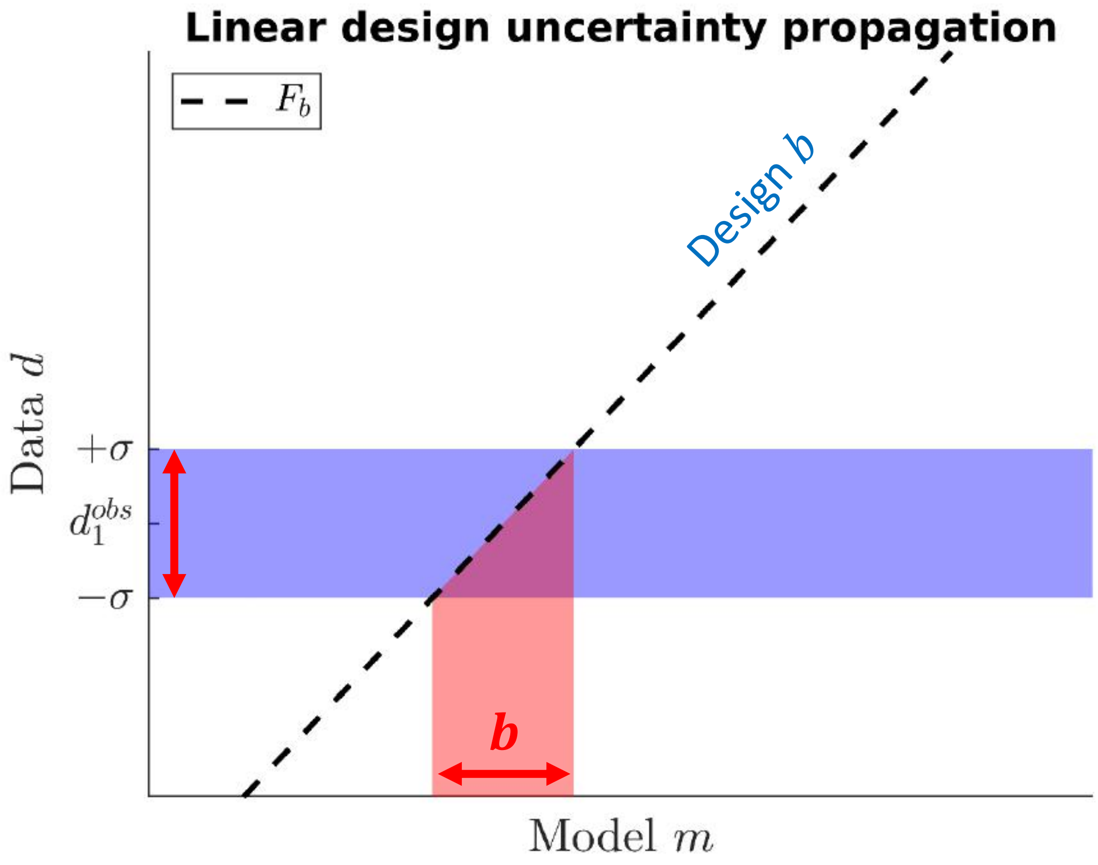
Design Heuristic (rule of thumb): Design with highest gradient is best

Experimental design theory



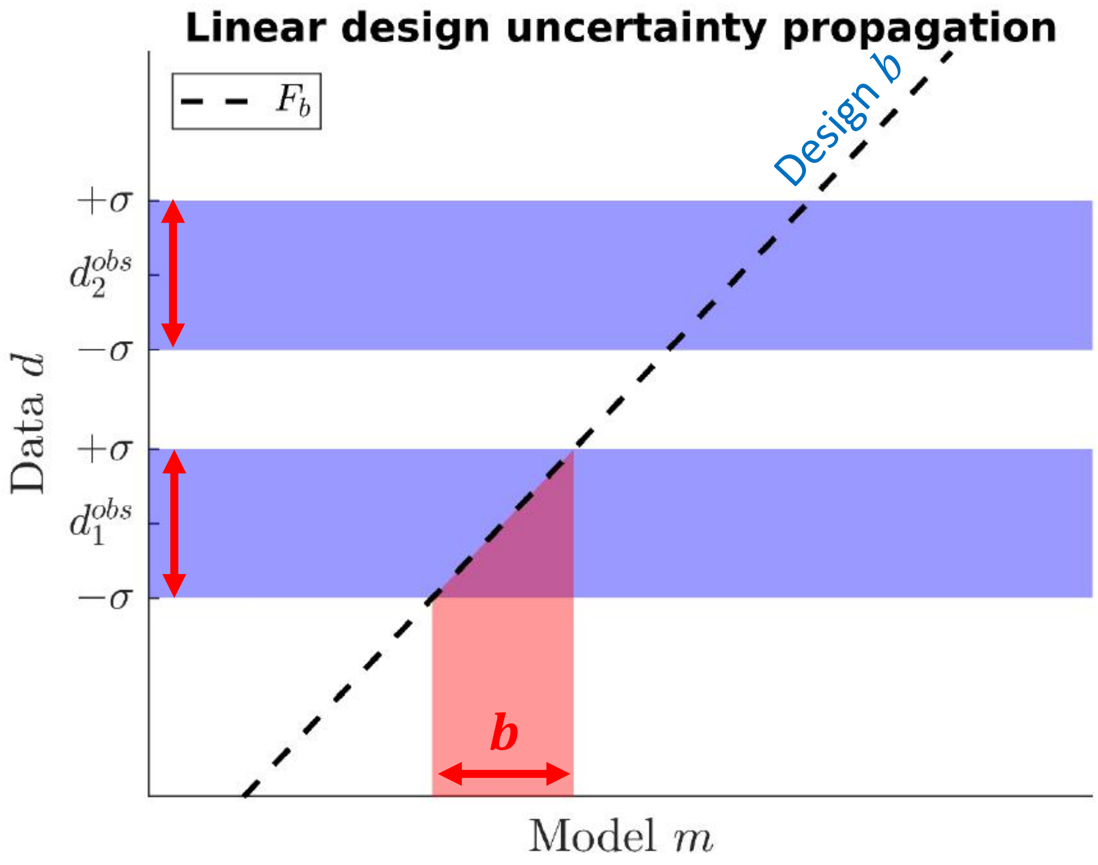
What if we recorded a different datum?

Experimental design theory



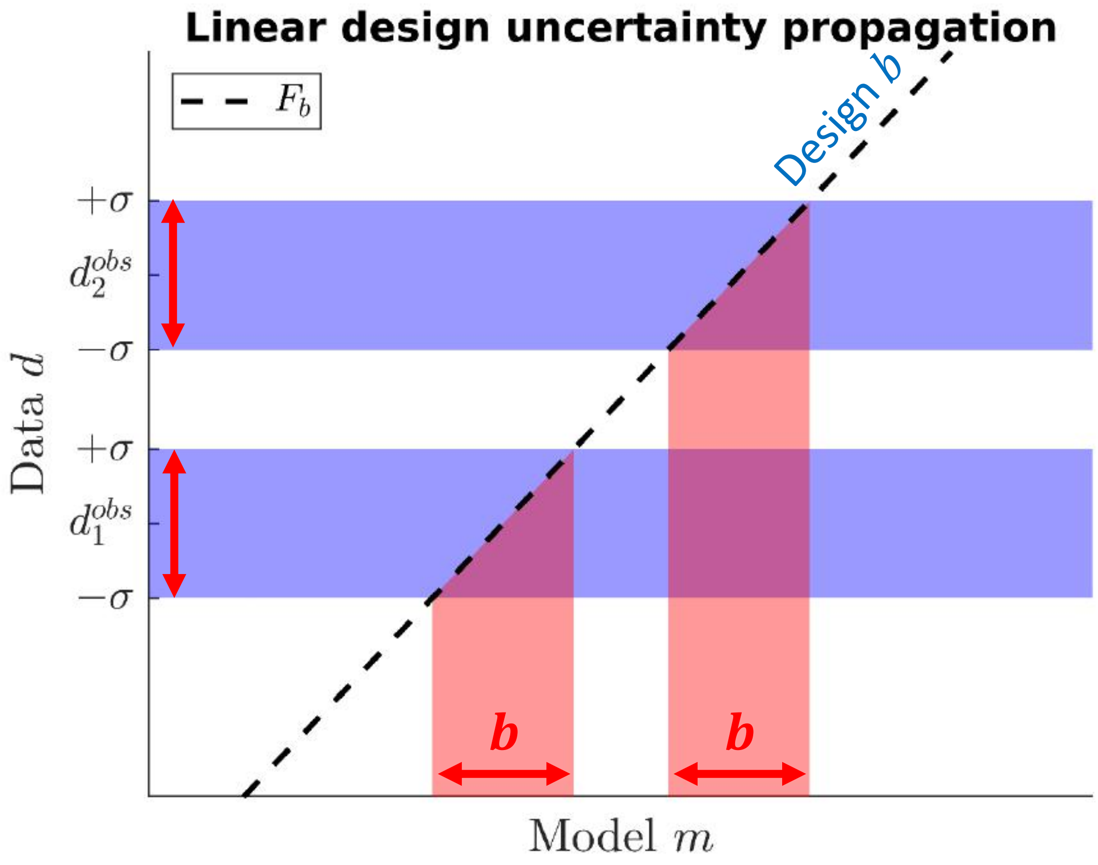
What if we recorded a different datum?

Experimental design theory



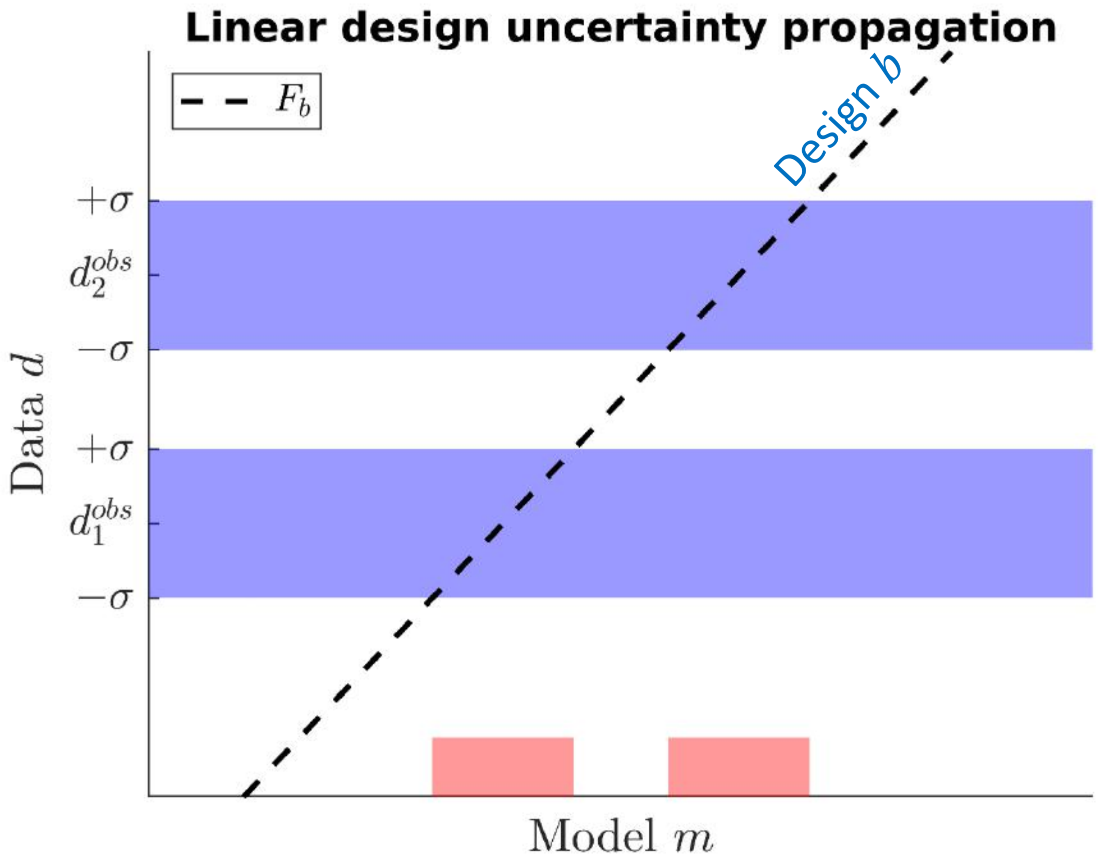
What if we recorded a different datum?

Experimental design theory



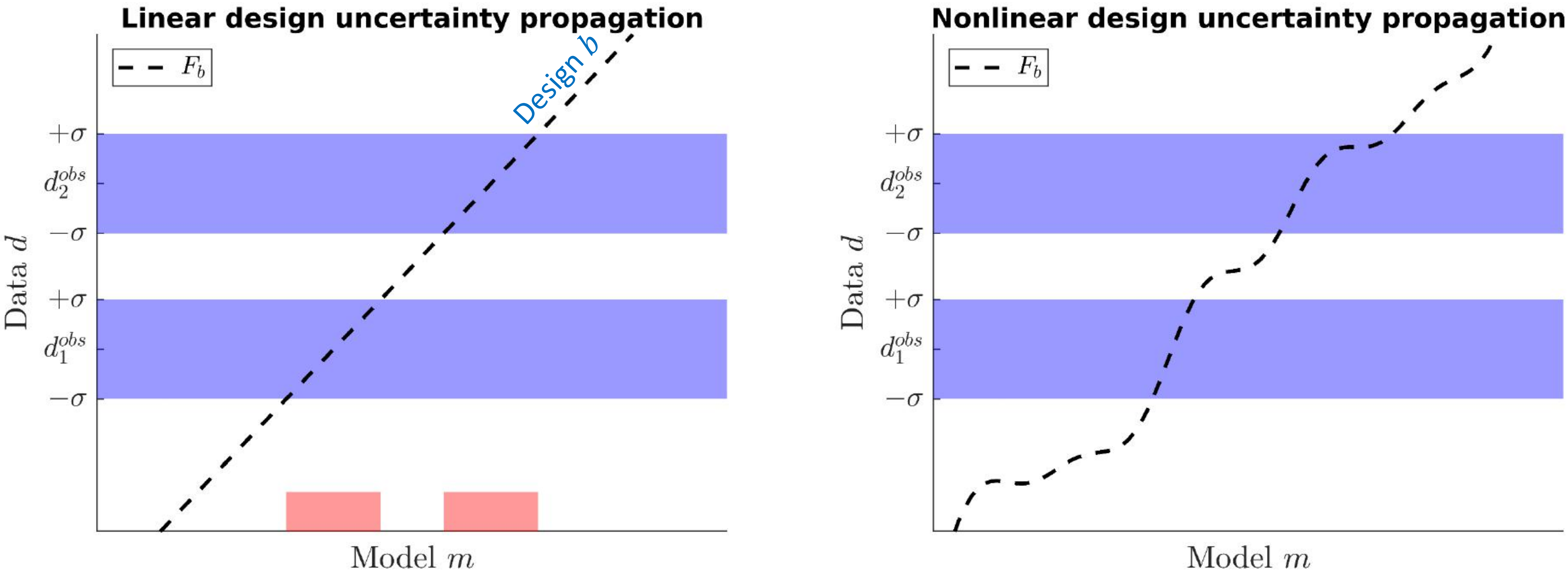
What if we recorded a different datum?

Experimental design theory



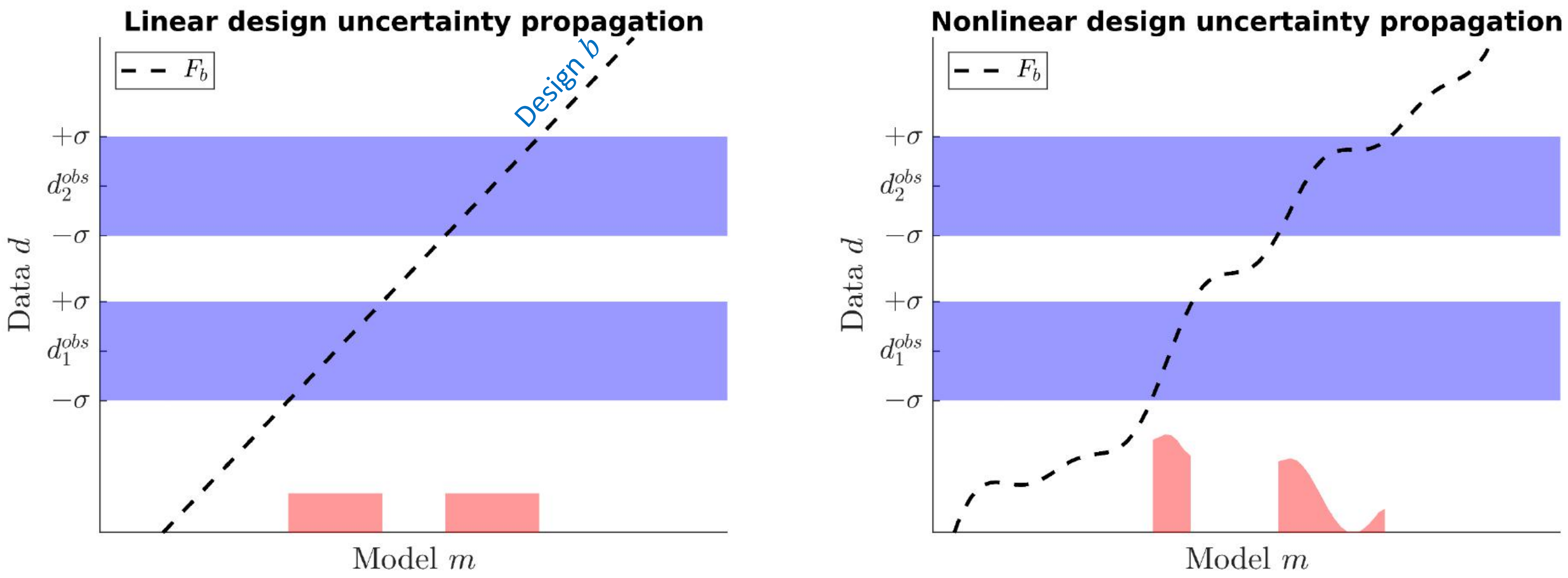
What if we recorded a different datum?

Experimental design theory



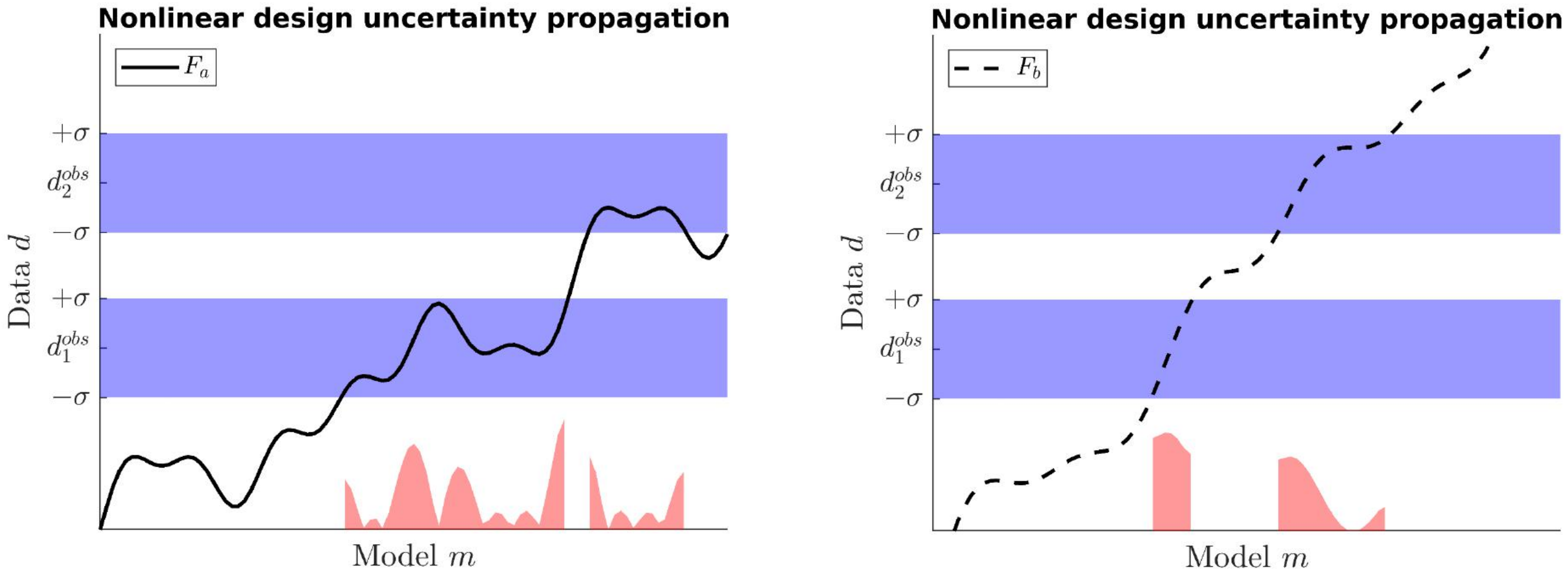
What if we recorded a different datum?

Experimental design theory



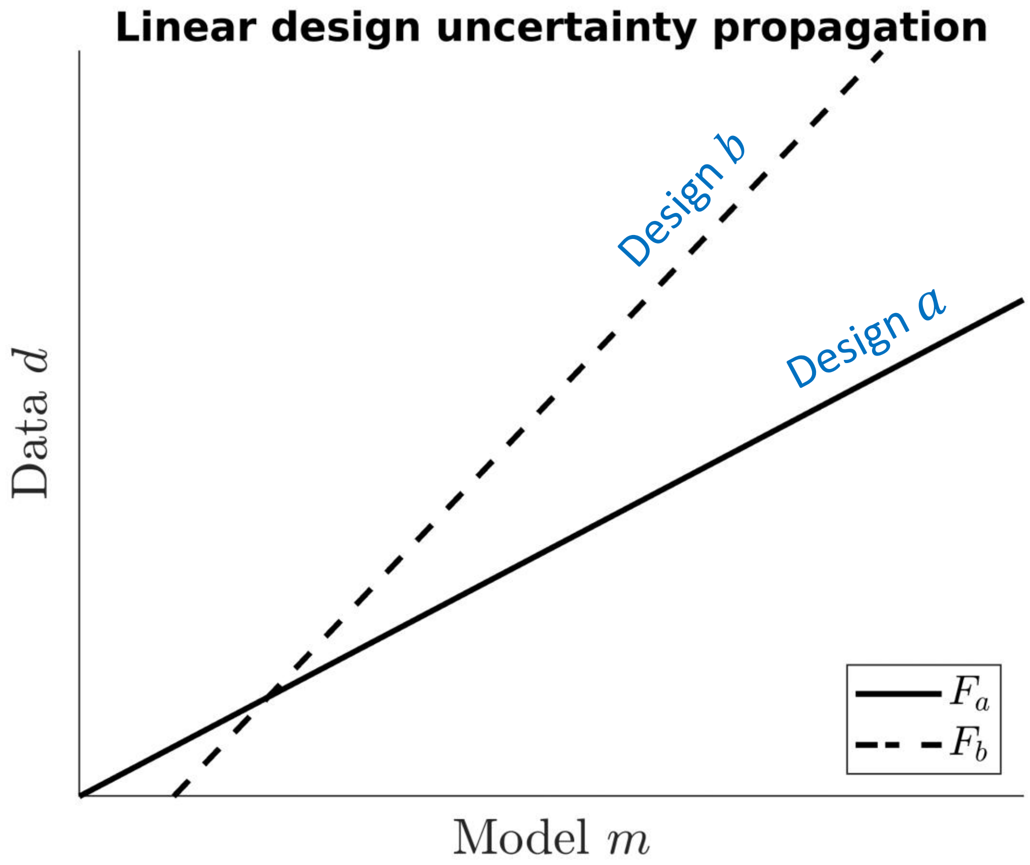
What if we recorded a different datum?

Experimental design theory $\arg_S \max\{ E[\, Inf(m)]\}$



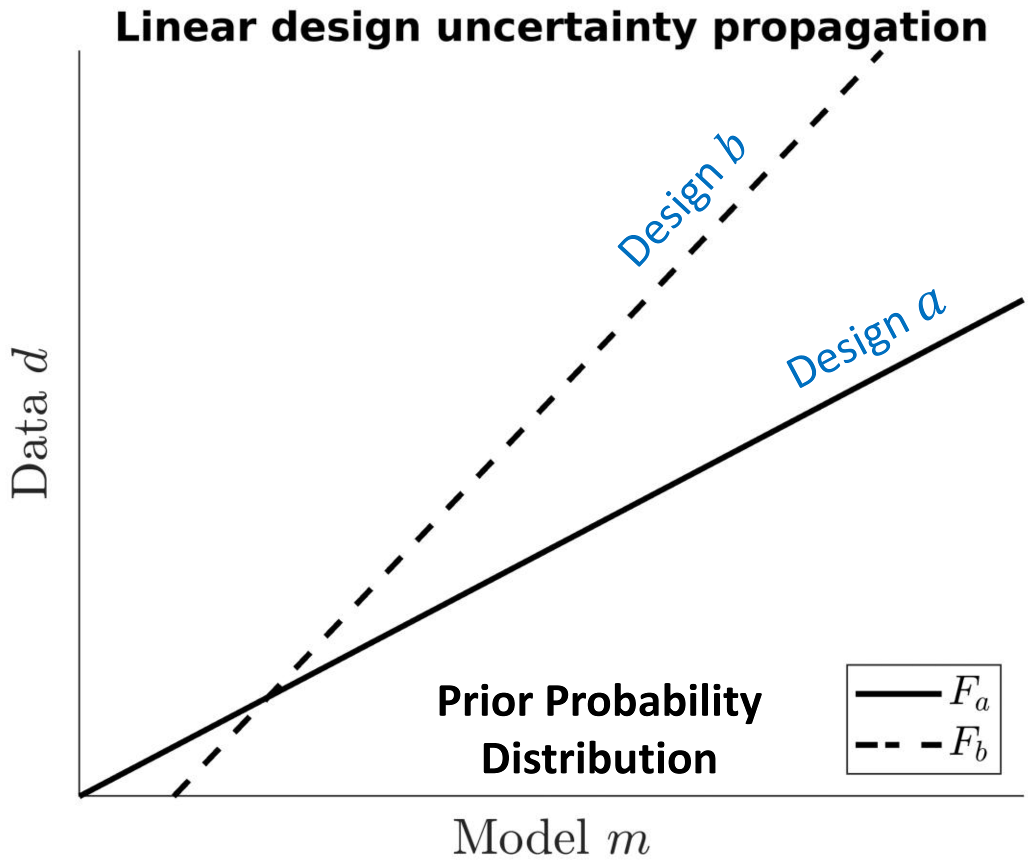
What if we recorded a different datum?

Experimental design theory



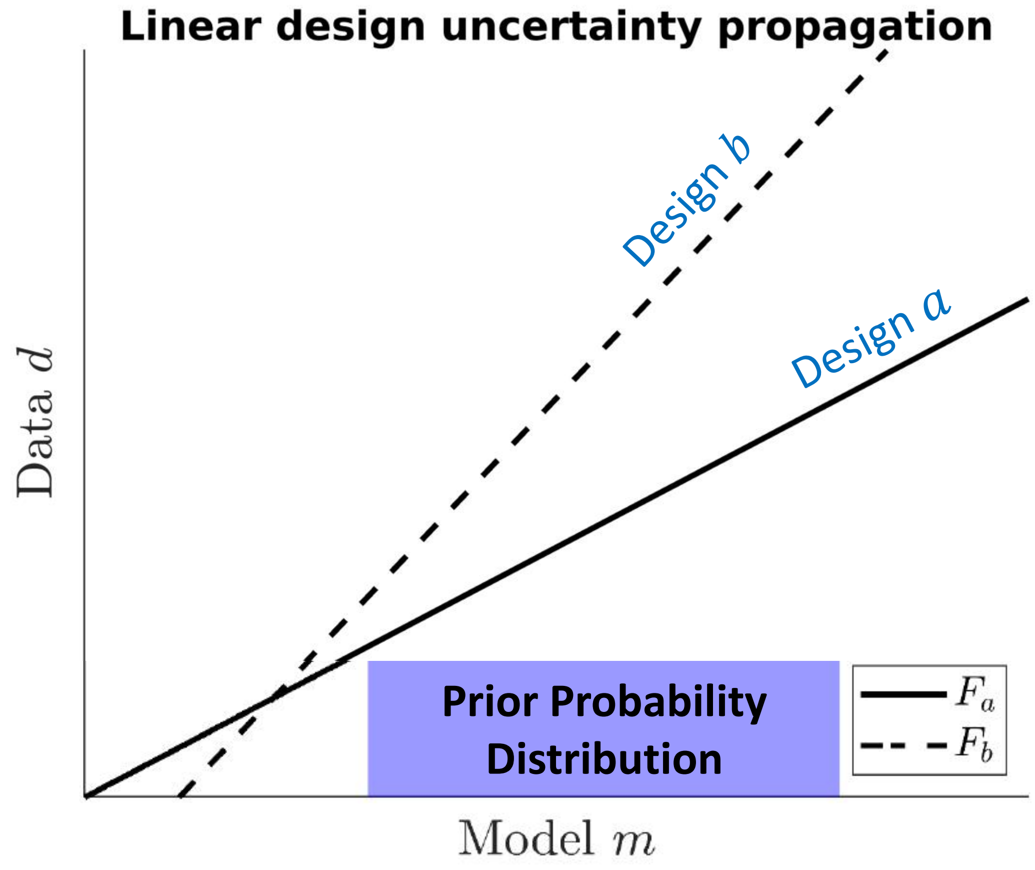
What to do before a survey?

Experimental design theory



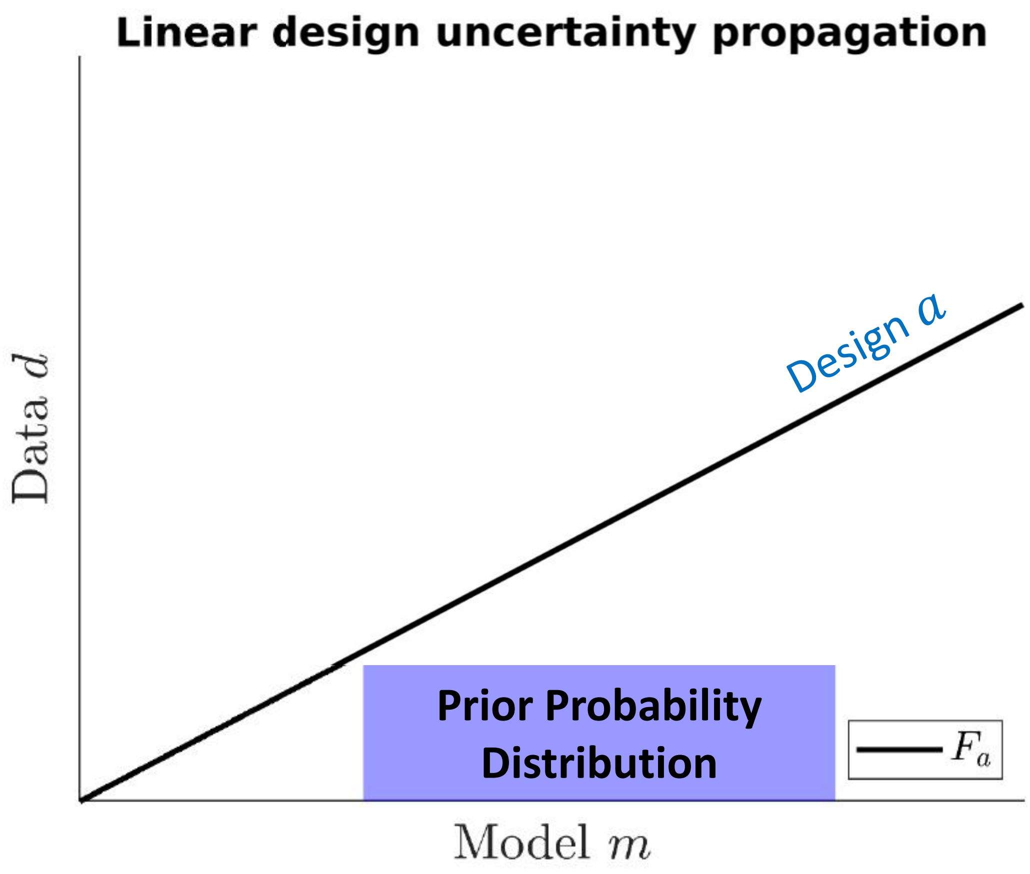
What to do before a survey?

Experimental design theory



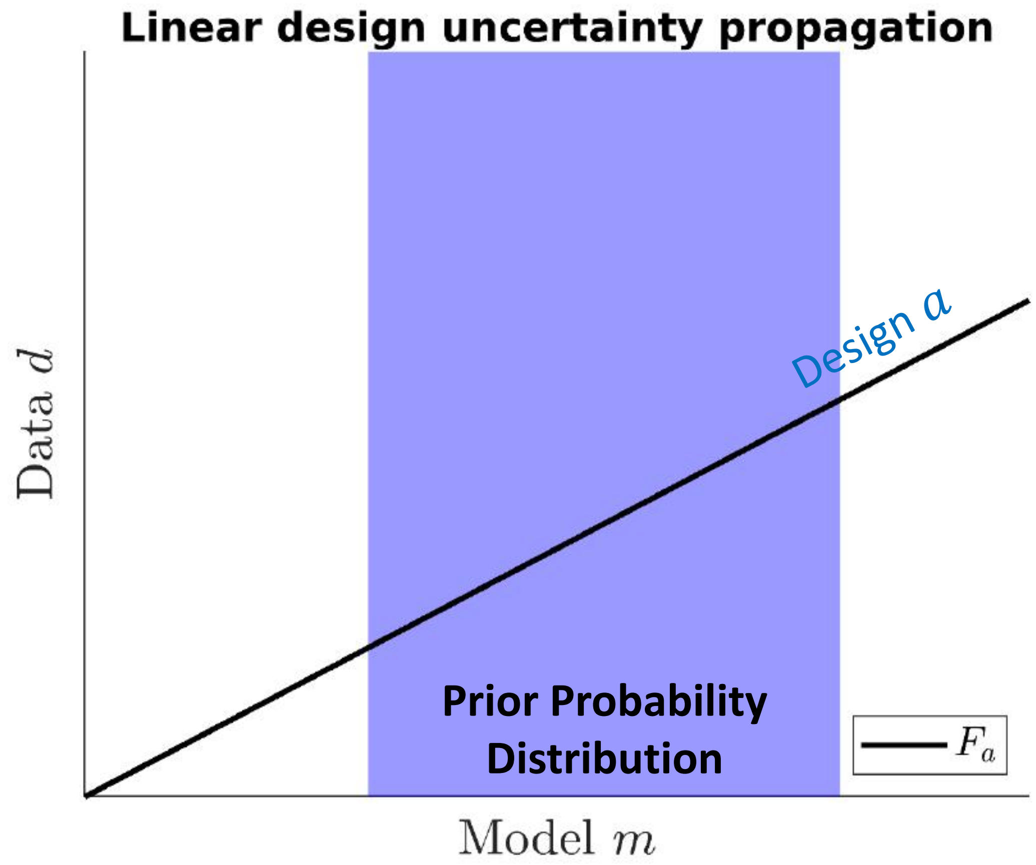
What to do before a survey?

Experimental design theory



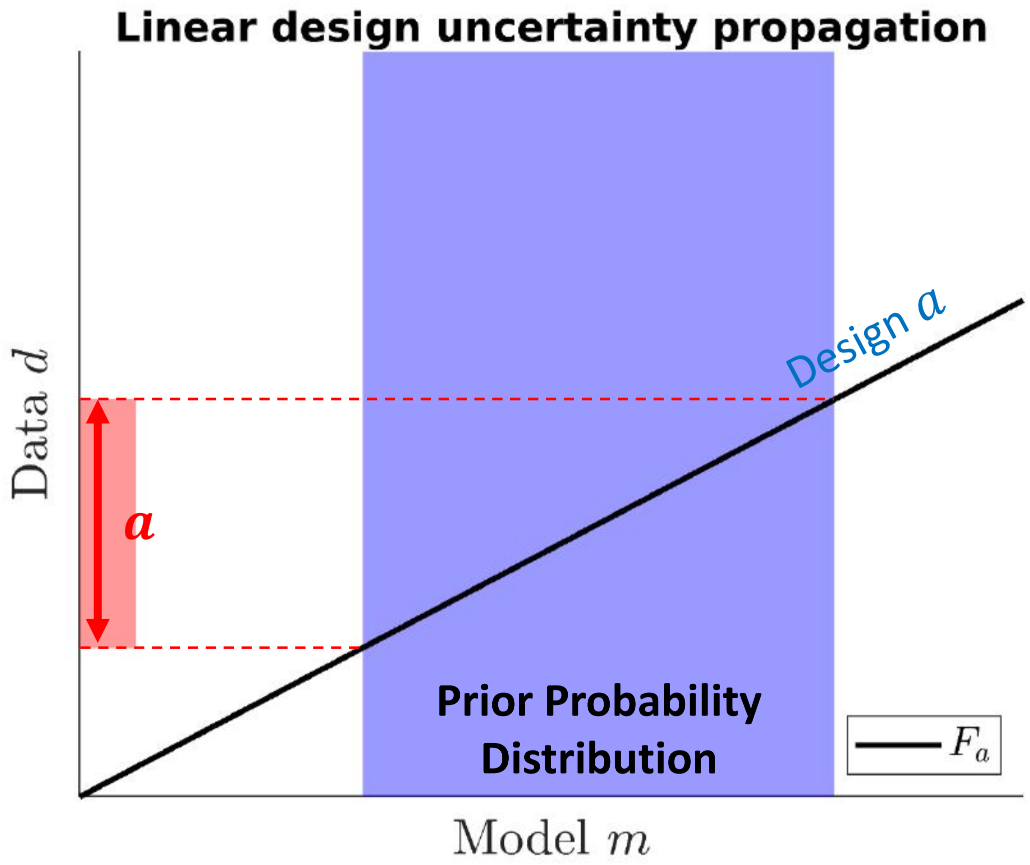
What to do before a survey?

Experimental design theory



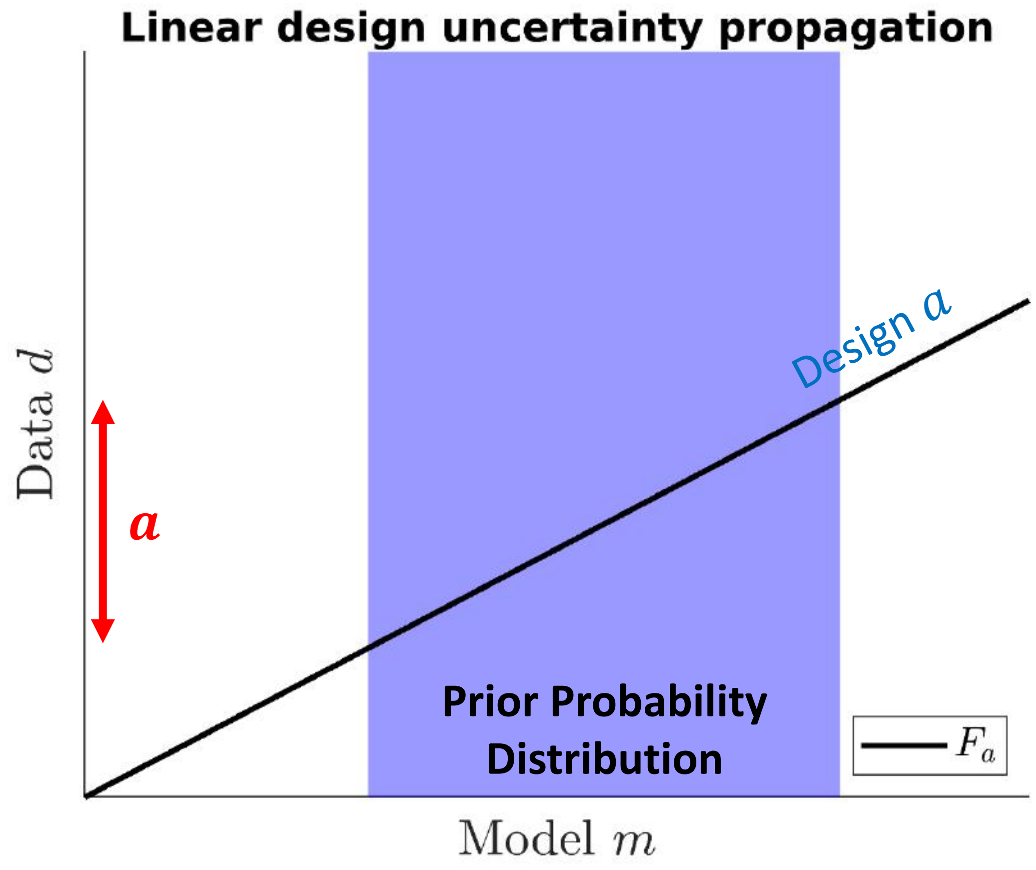
What to do before a survey?

Experimental design theory



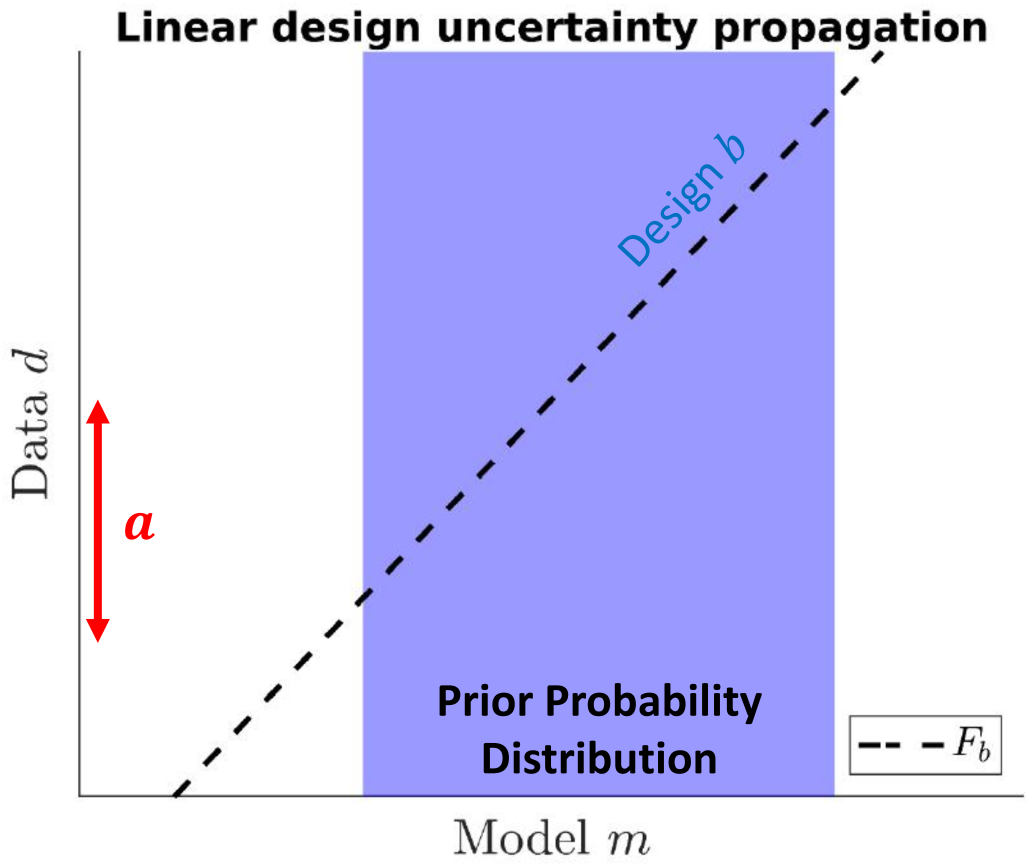
What to do before a survey?

Experimental design theory



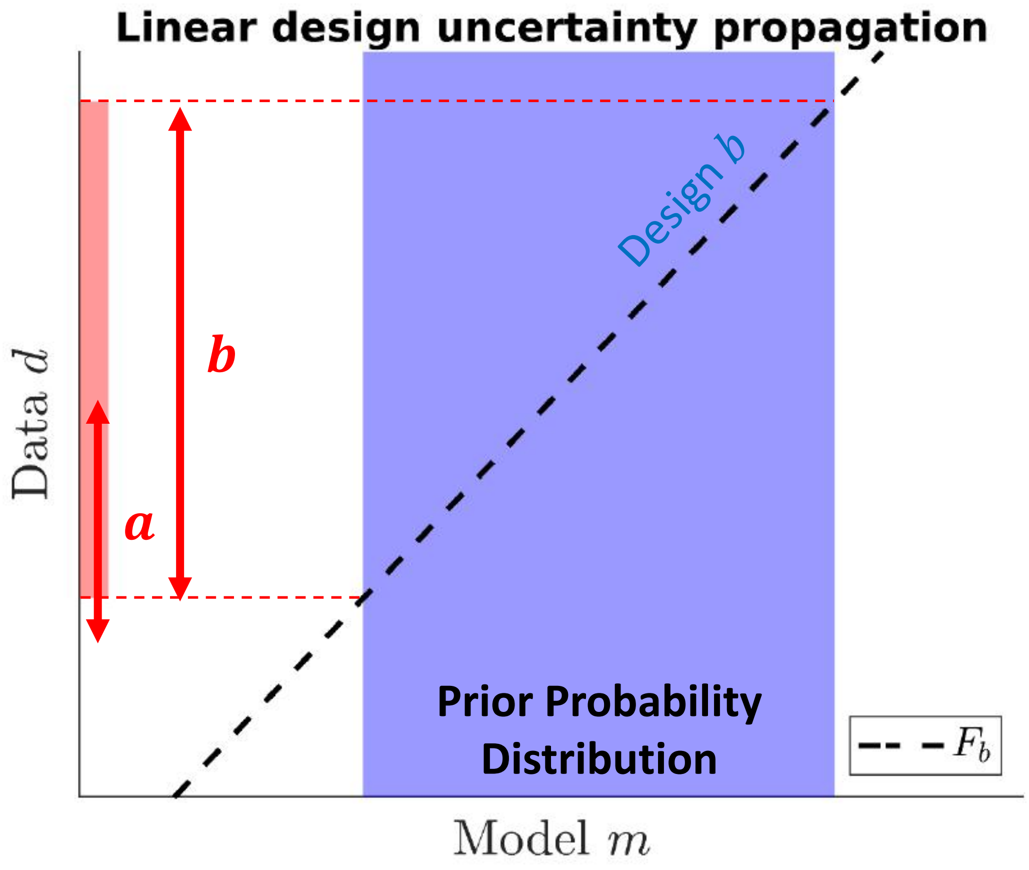
What to do before a survey?

Experimental design theory



What to do before a survey?

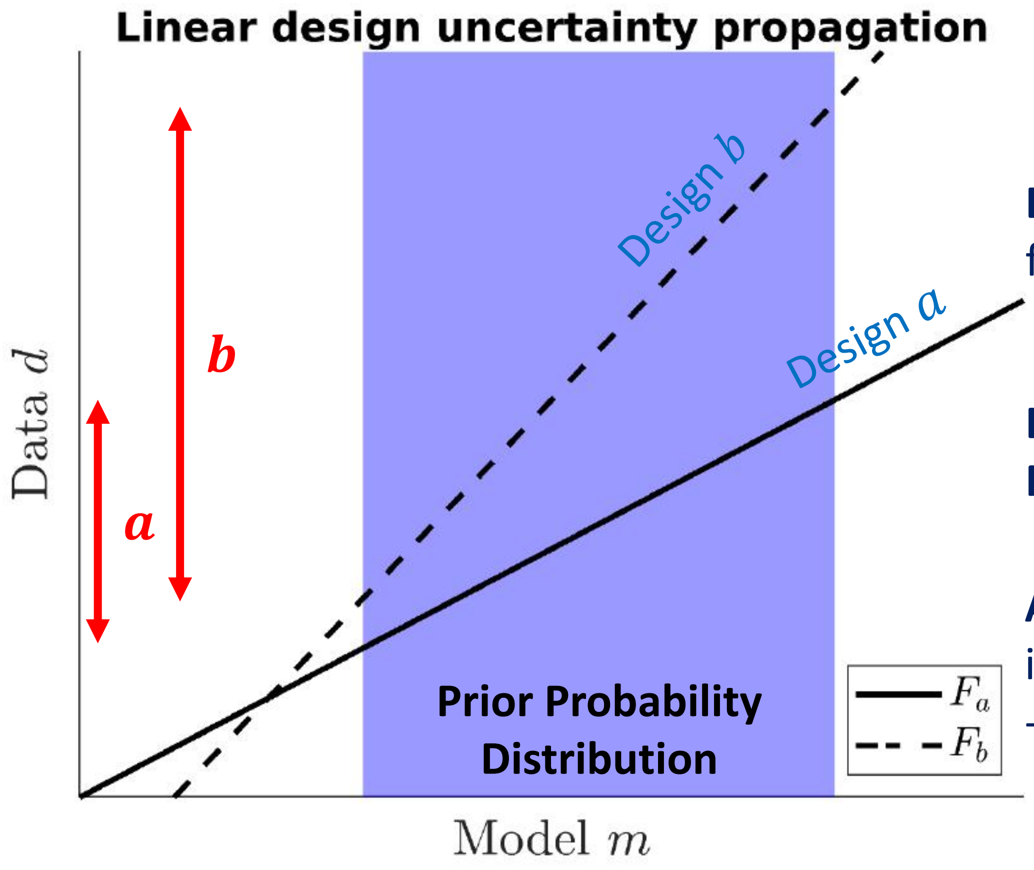
Experimental design theory



What to do before a survey?

Experimental design theory

Shewry and Wynn (1987) proved Heuristic 2 also works in fully nonlinear cases



Design Heuristic 1: Maximise forward model **gradient**

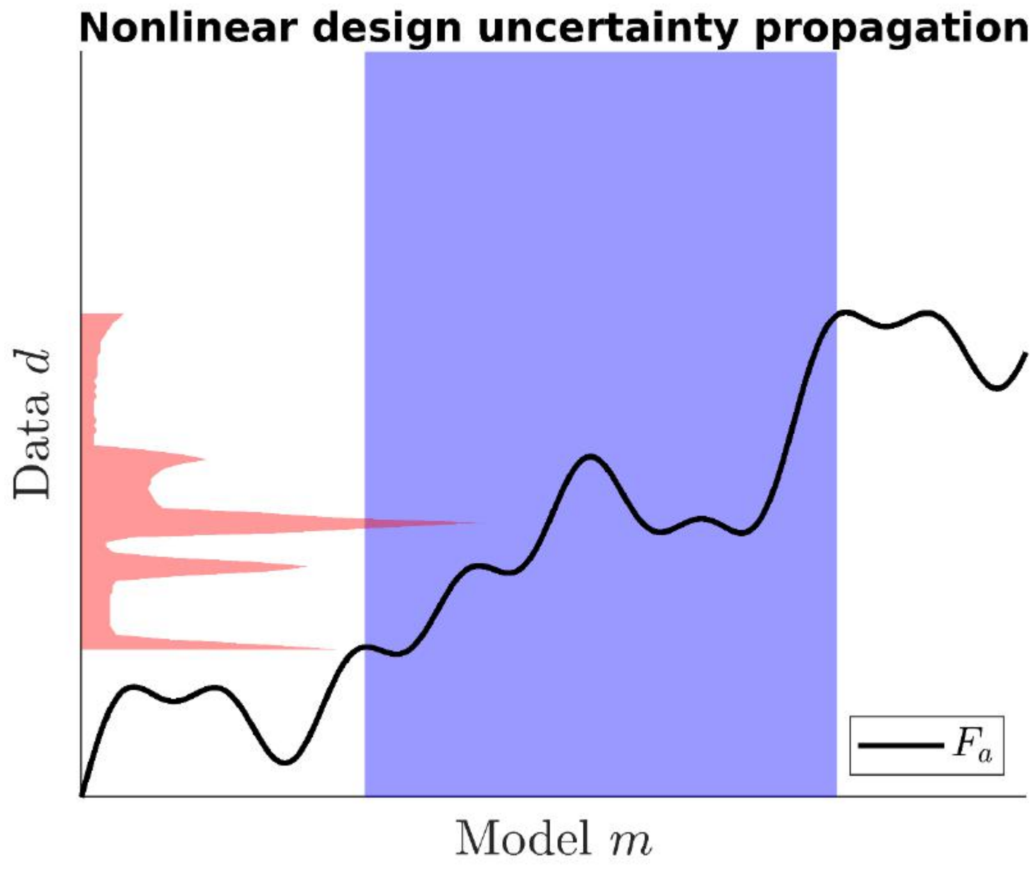
Design Heuristic 2: Maximise **Prior Data-Space** Uncertainty

Assumes: Joint [**Model+Data] space** information is constant

- **True** if data measurement uncertainties are independent of measured data values

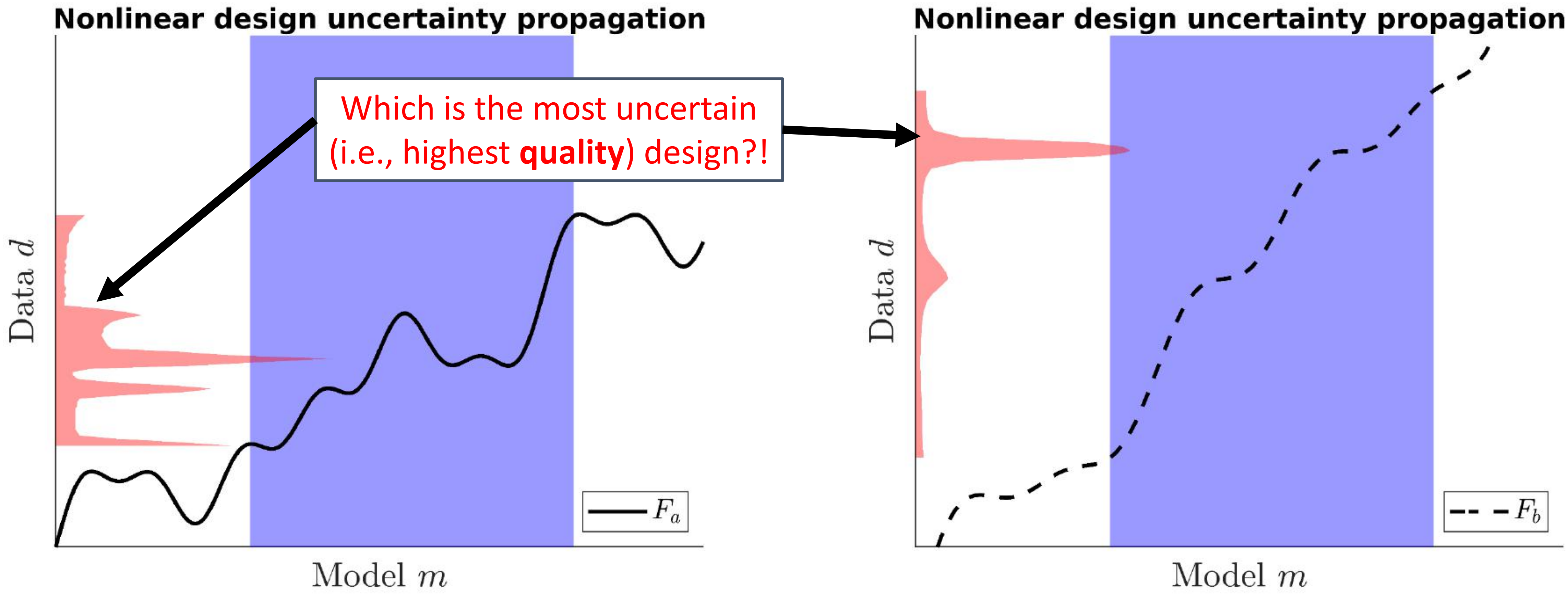
What to do before a survey?

Experimental design theory



Experimental design theory

$\arg_S \max\{\, E[\, Inf(m)]\}$



Maximise the Prior Data-Space <u>Entropy</u>: $\arg_S \max[Ent(d)]$

Quality measures

Quality measures

Optimal receiver locations design

$$\hat{S} = \arg_S \max[\dots]$$

Quality measure

Quality measures

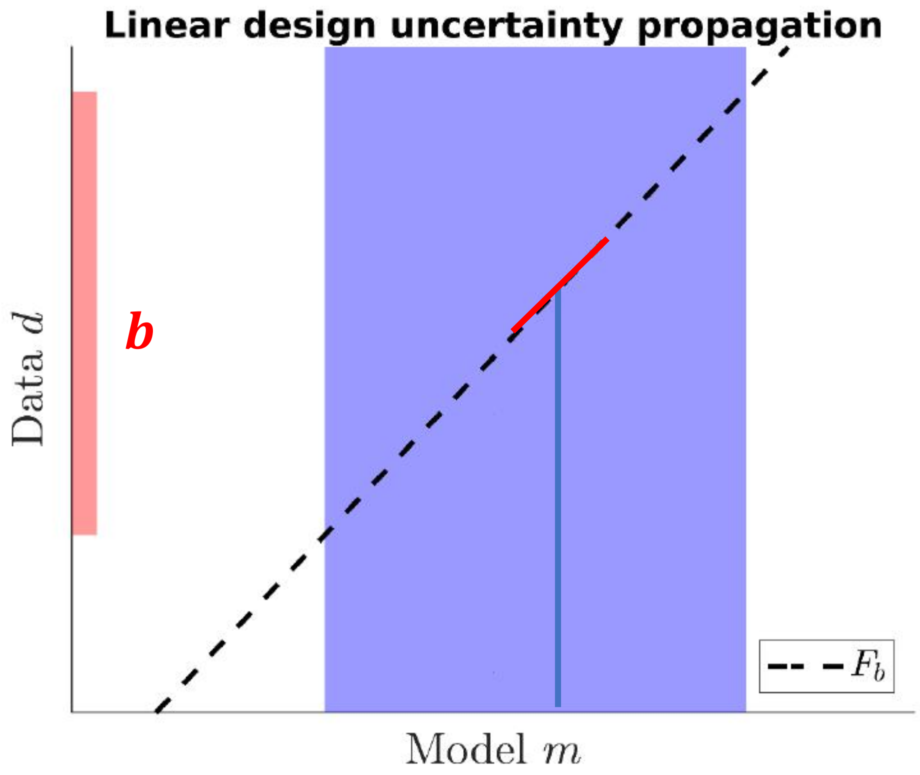
$\hat{S} = \arg_S \max[\ldots]$

- Linear design
  - **Bayesian D-optimisation:** maximise the gradient of F
- Nonlinear design
  - **Maximum Entropy Design:** $\max[Ent(D)]$
  - **$\boldsymbol{D_N}$-optimisation:** approximate $\max[Ent(D)]$

Linear design: D-optimisation

- Assumes *F* is linear
- Compute gradients of $\boldsymbol{d}$ w.r.t. $\boldsymbol{m}$

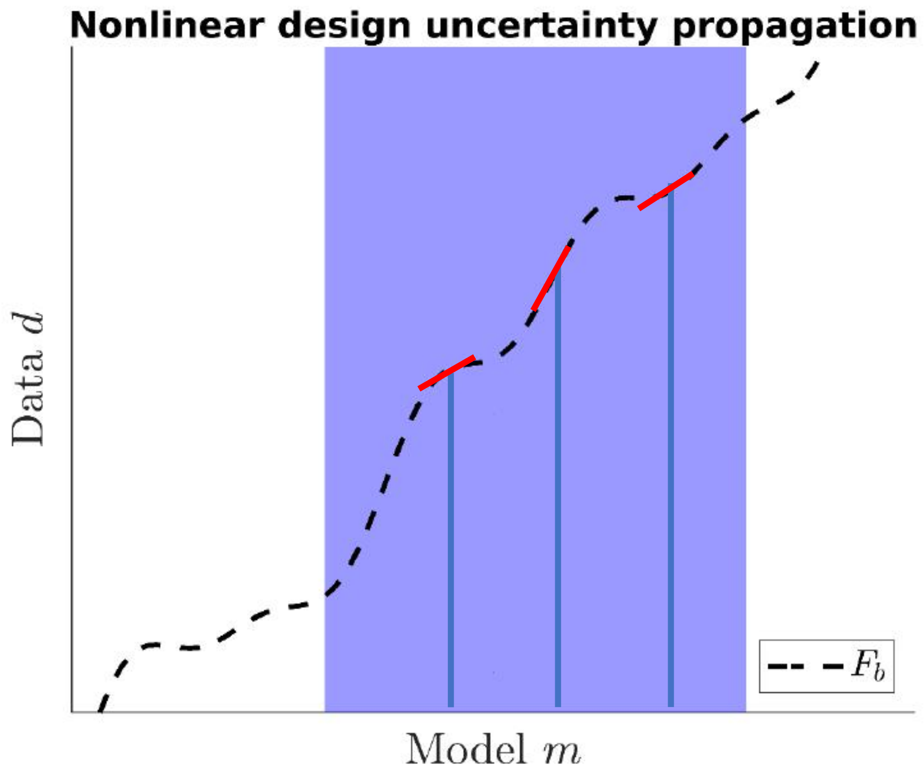
$$\hat{S} = \arg_S \max\left[a_j \ln\left(\left|\boldsymbol{A}_{S,j}^T \boldsymbol{A}_{S,j}\right|\right)\right]$$



Linear design: Bayesian D-optimisation

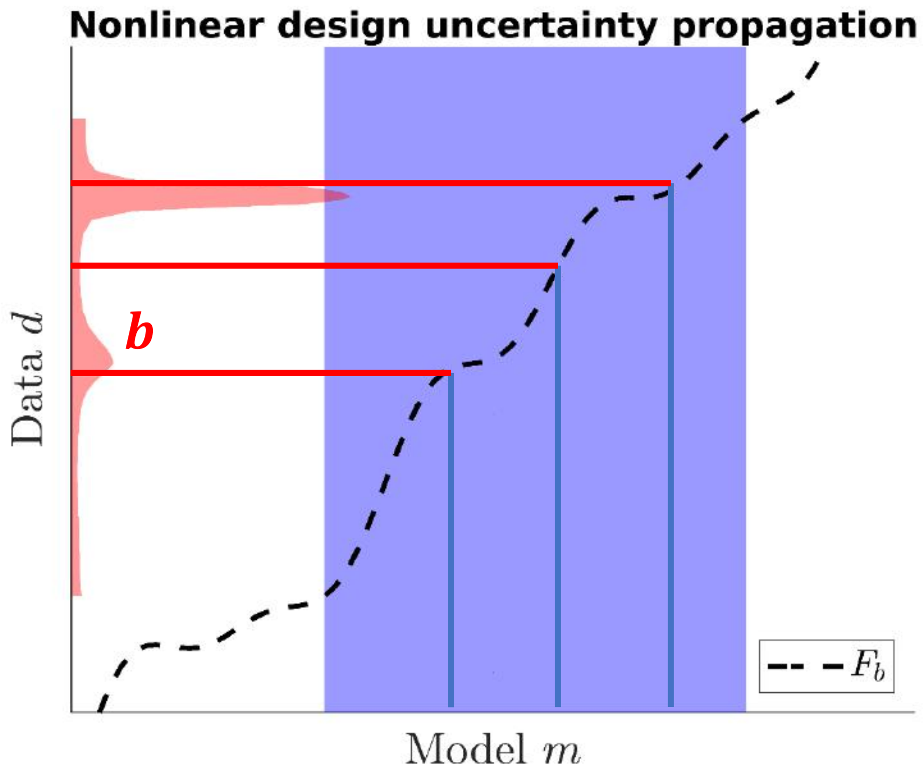
- Assumes F is linear
- Compute gradients of $\boldsymbol{d}$ w.r.t. $\boldsymbol{m}$

$$\hat{S} = \arg_S \max \left[\frac{1}{N} \sum_{j=1}^{N} a_j \ln\left(\left| \boldsymbol{A}_{S,j}^T \boldsymbol{A}_{S,j} \right| \right) \right]$$



Nonlinear design: Maximum Entropy Design

- The optimal design has the largest entropy in data space



$$\hat{S} = \arg_S \max\left[Ent\left(\rho(\boldsymbol{d}|S)\right)\right]$$

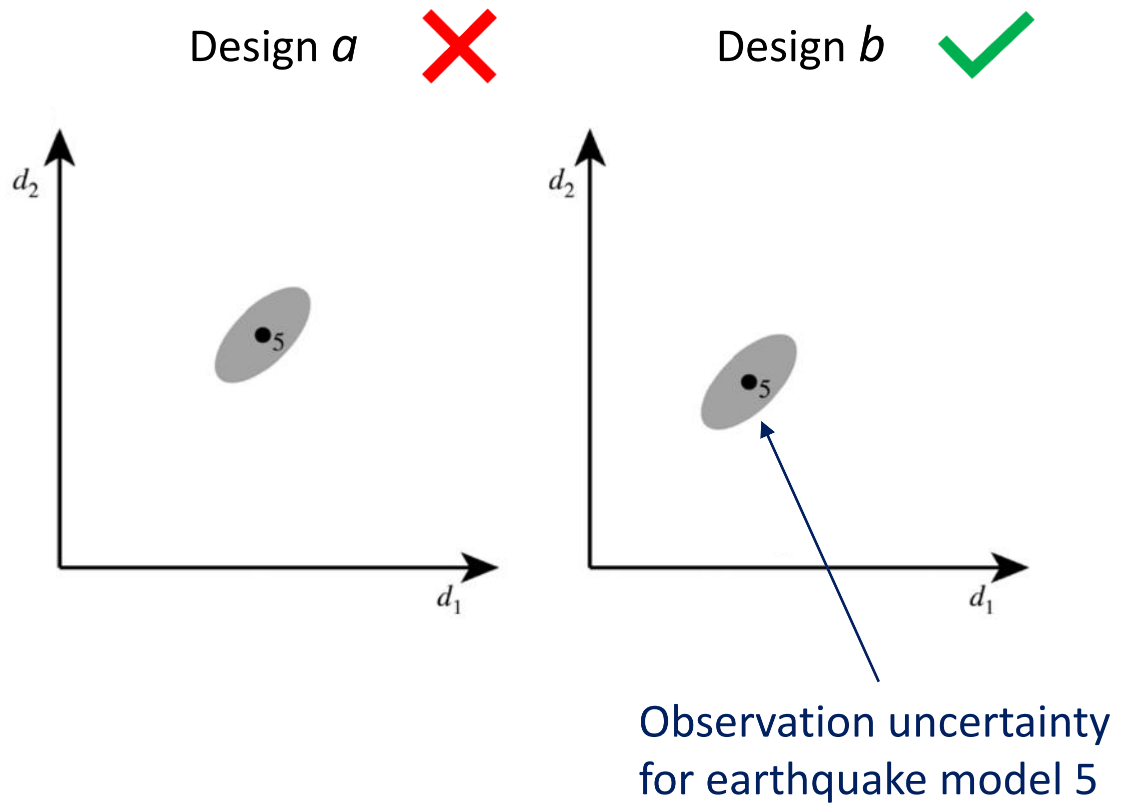
$$Ent = -\int_X f(x) \log f(x)\, dx$$

Nonlinear design: D_N-optimisation

- Entropy approximation
- Spread of data is defined by the 'ambiguity' in data space

$$\hat{S} = \arg_S \max[\ln|\Sigma(S)|]$$

With Σ the covariance matrix of the synthetic data samples

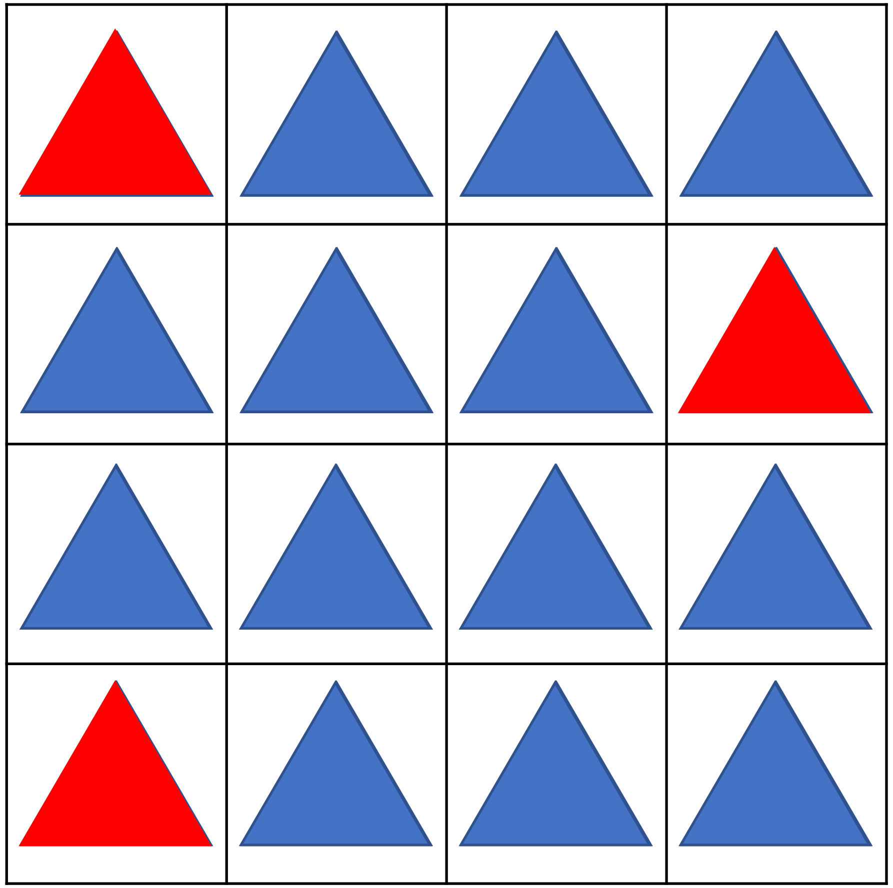


Design Optimisation

Sequential Design **Algorithm**

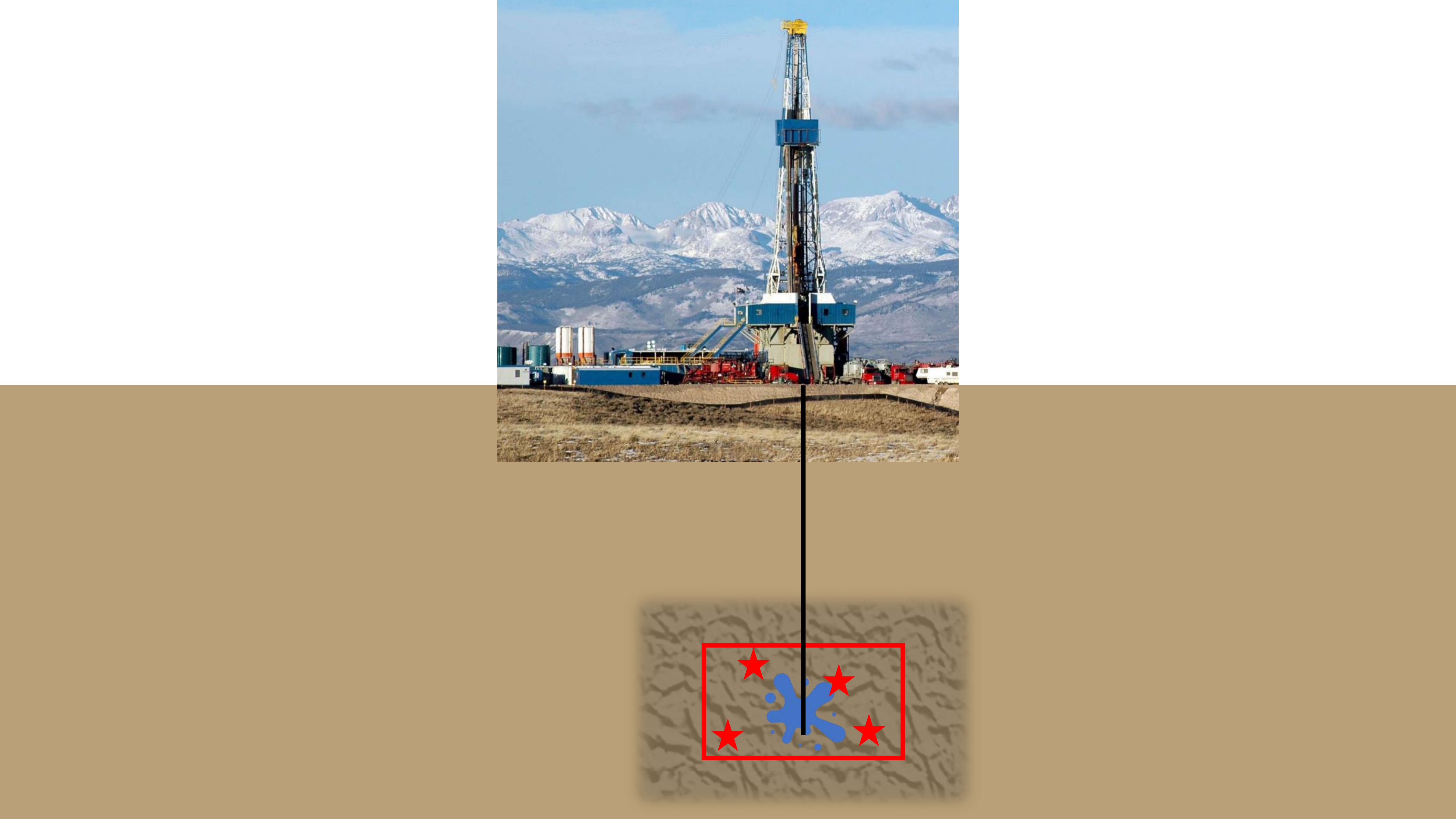
- Evaluating all possible designs is a combinatorial problem - too costly

1. *Calculate quality measure for all potential additional single receiver locations*
2. *Find maximum quality location*
3. *Add a receiver at that location to network*
4. *Repeat from 1*



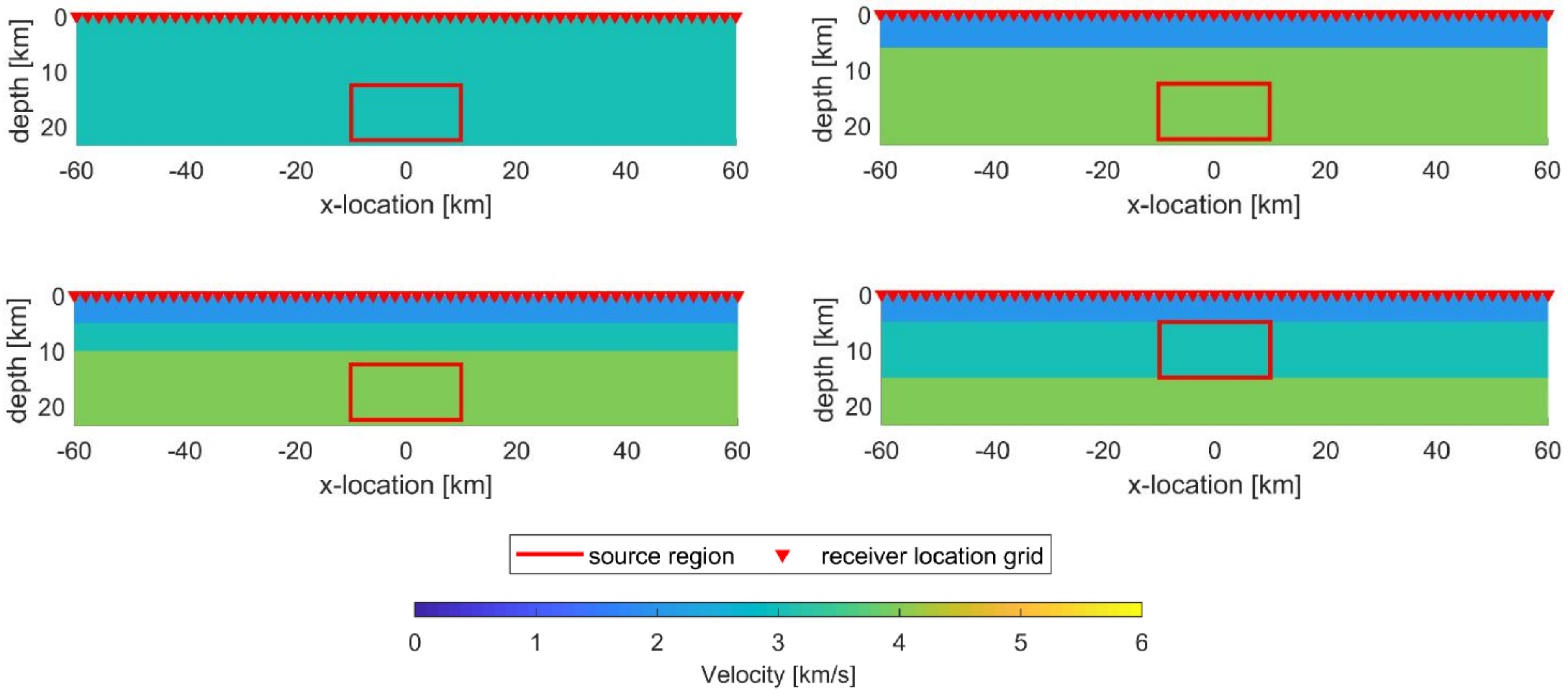
Method comparison

Method comparison



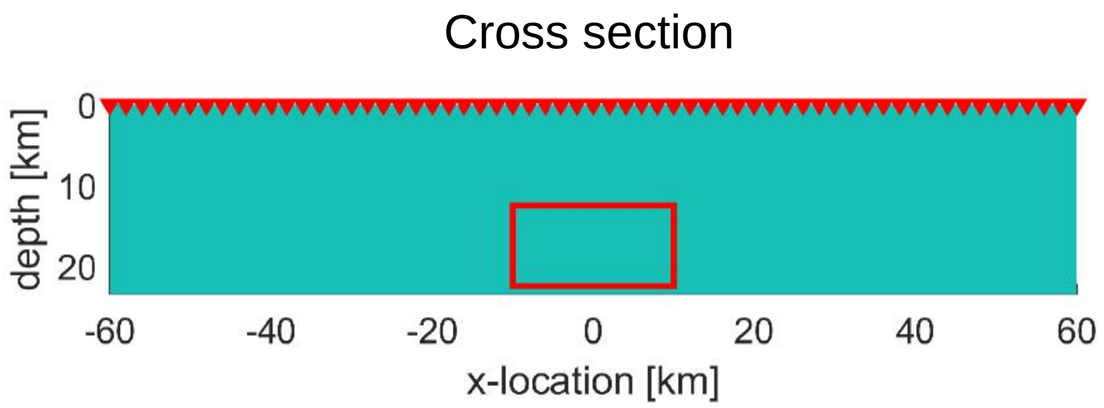
Method comparison

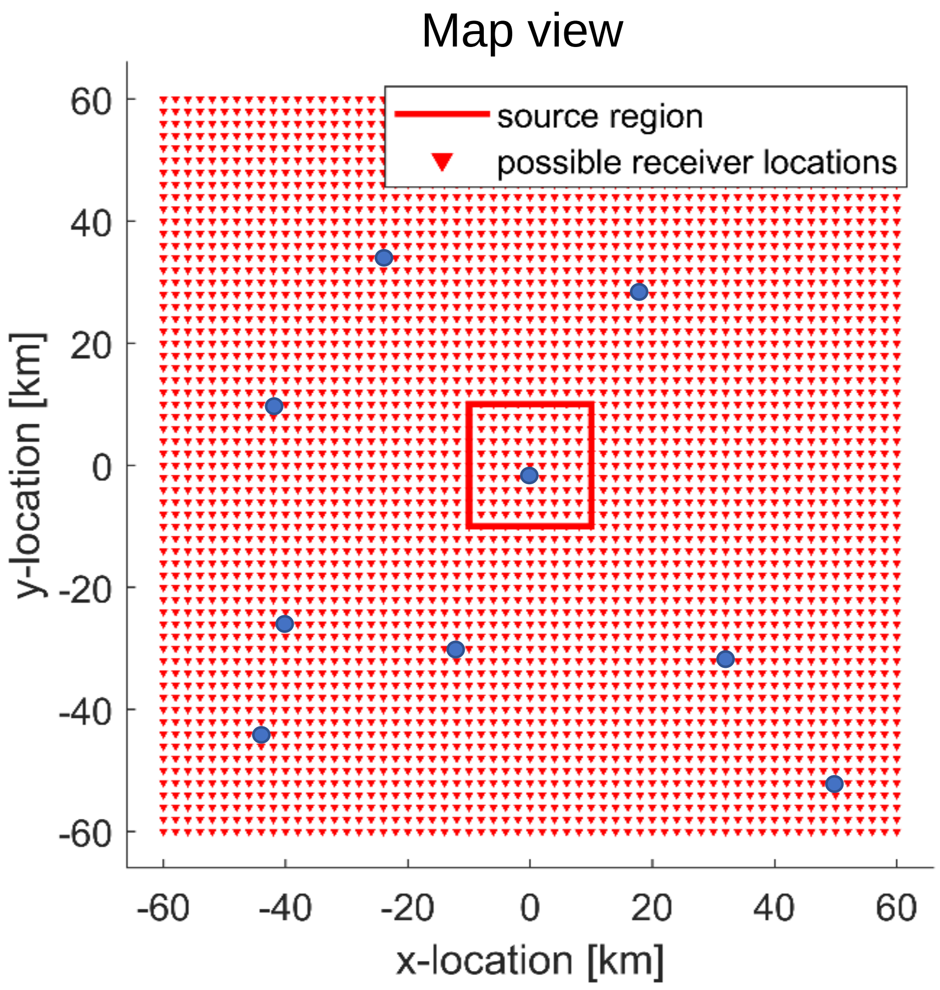
Vertical cross sections through four synthetic models



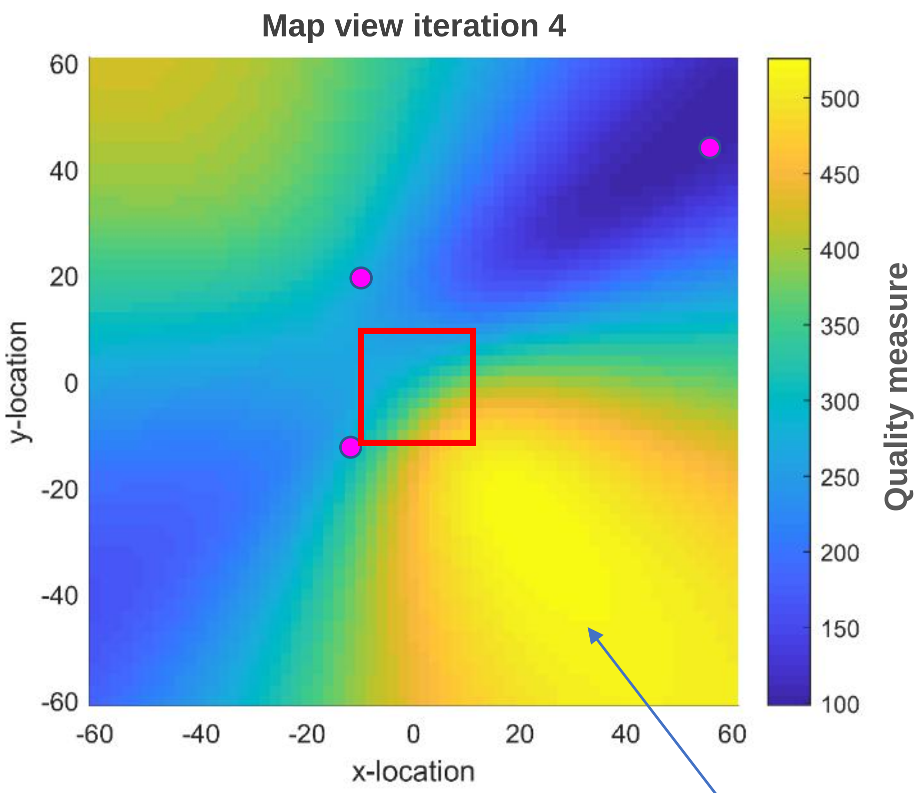
Method comparison

Potential receiver locations for network design



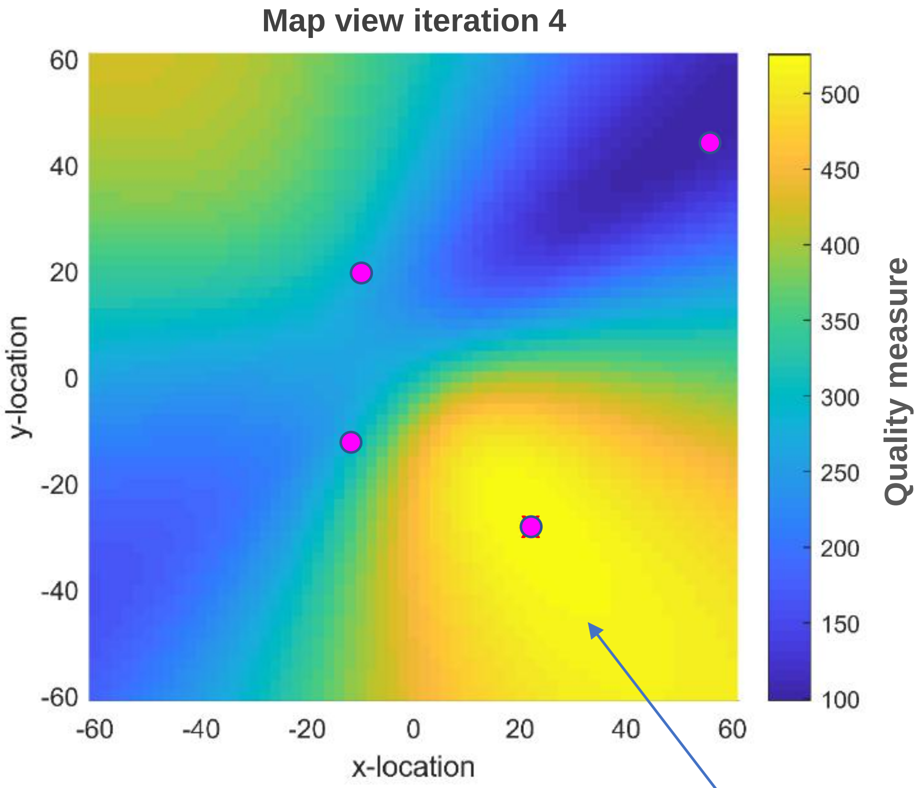


Quality measure – presentation (D-optimality)



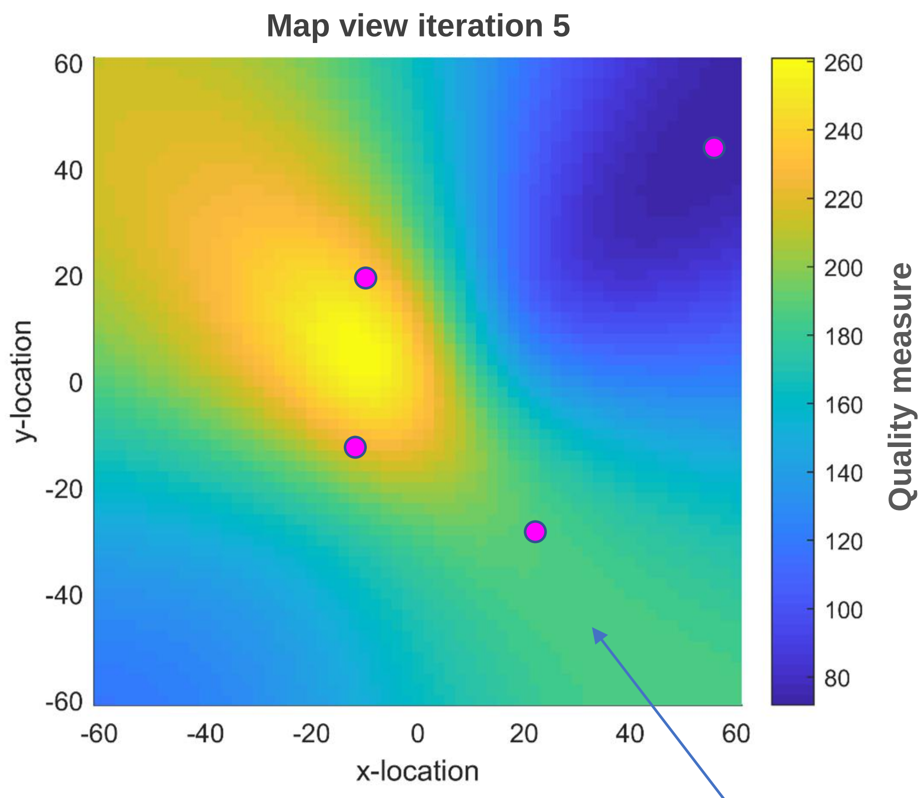
Colour is quality measure if we add a receiver at each location

Quality measure – presentation (D-optimality)



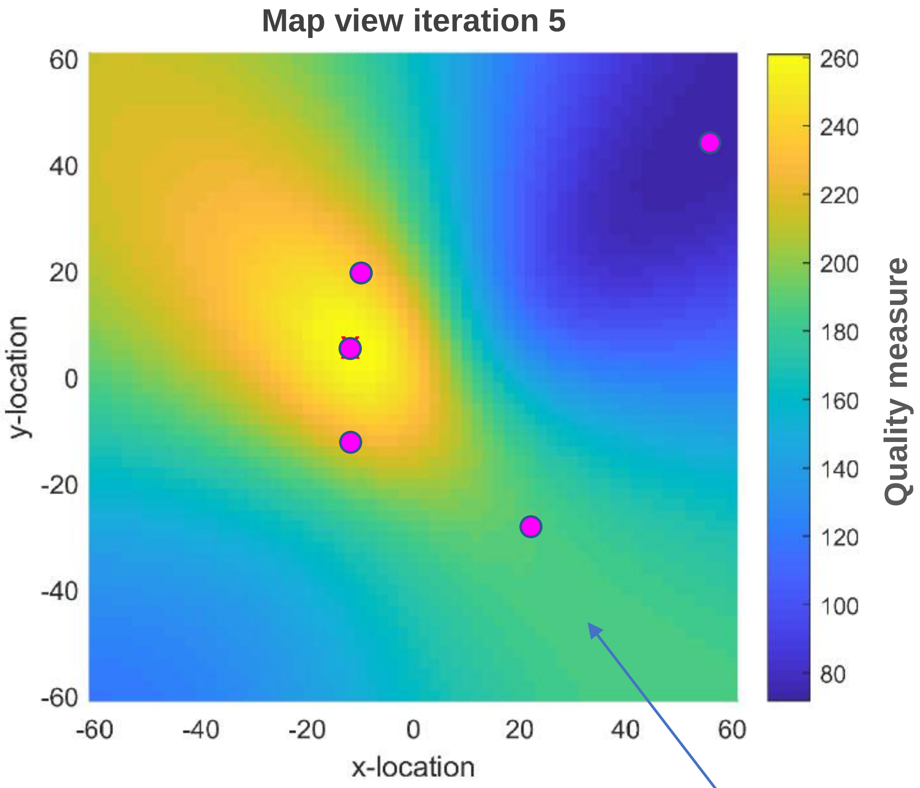
Colour is quality measure if we add a receiver at each location

Quality measure – presentation (D-optimality)



Colour is quality measure if we add a receiver at each location

Quality measure – presentation (D-optimality)



Colour is quality measure if we add a receiver at each location

Performance assessment

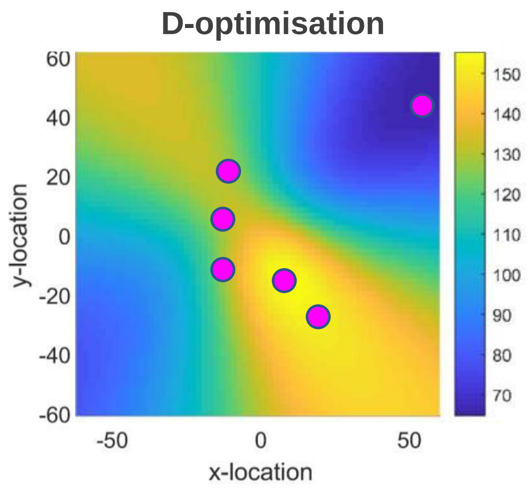
- Percentage of distinguishable sources
  - Take a source position
  - Calculate the arrival times for all other sources
  - Count the sources where $t_{arrival} \leq 0.1\ s$
  - Repeat for all pairs of sources

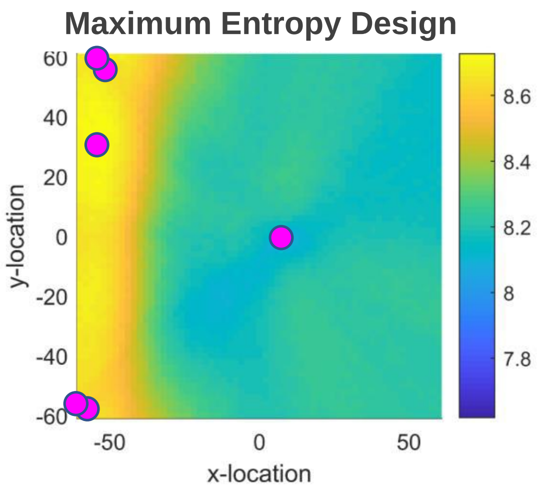
- Percentage of distinguishable sources

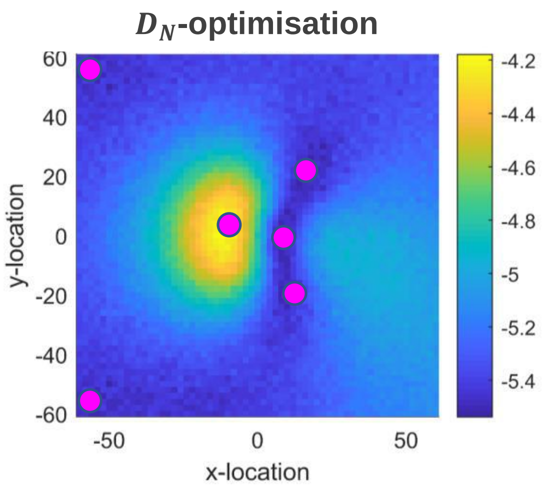
  ≈ measure('certainty') for an ideal inversion algorithm

Results

Three Designs from Three Measures…

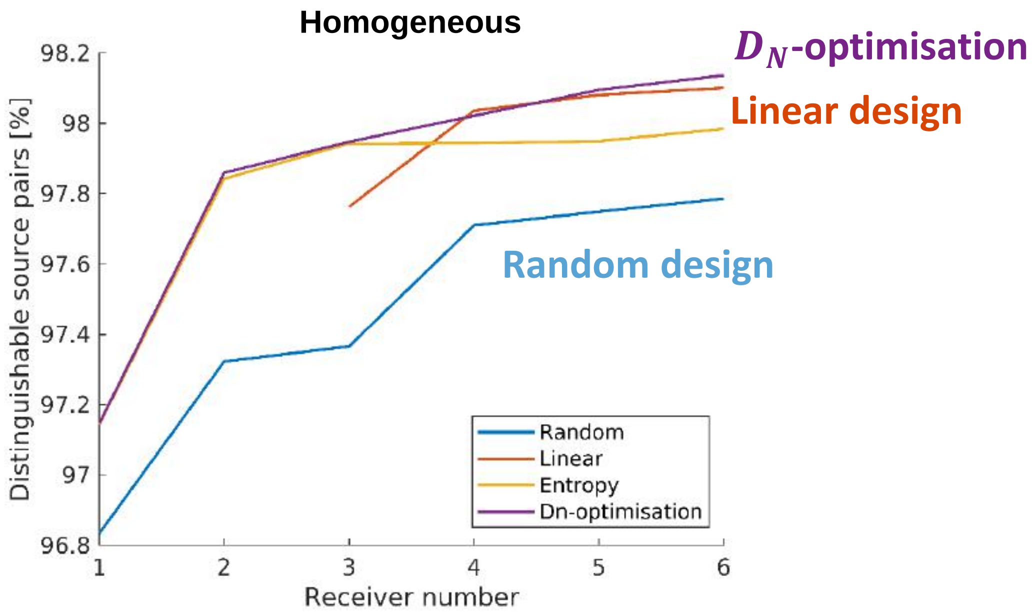


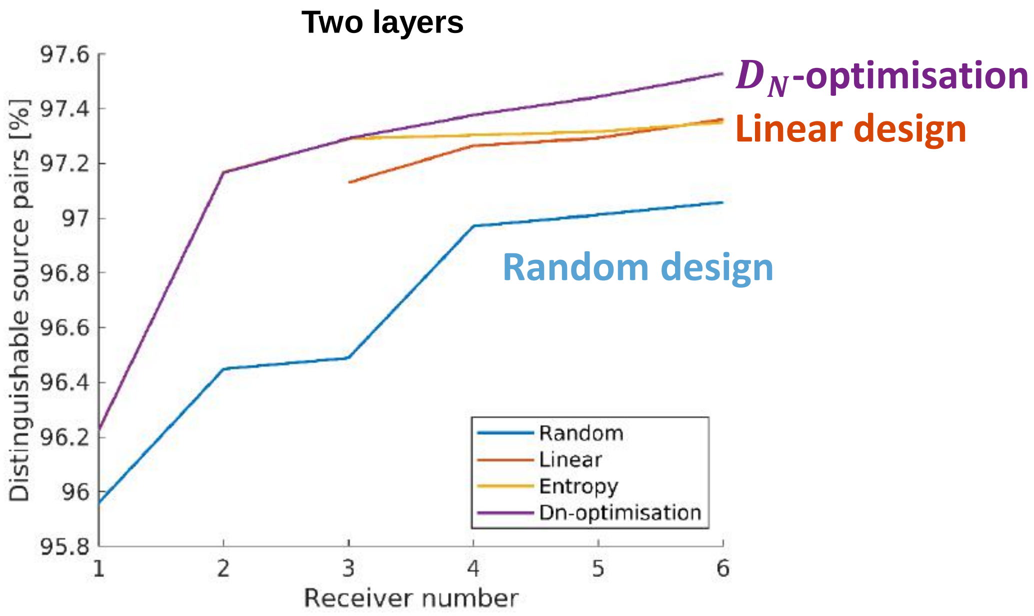


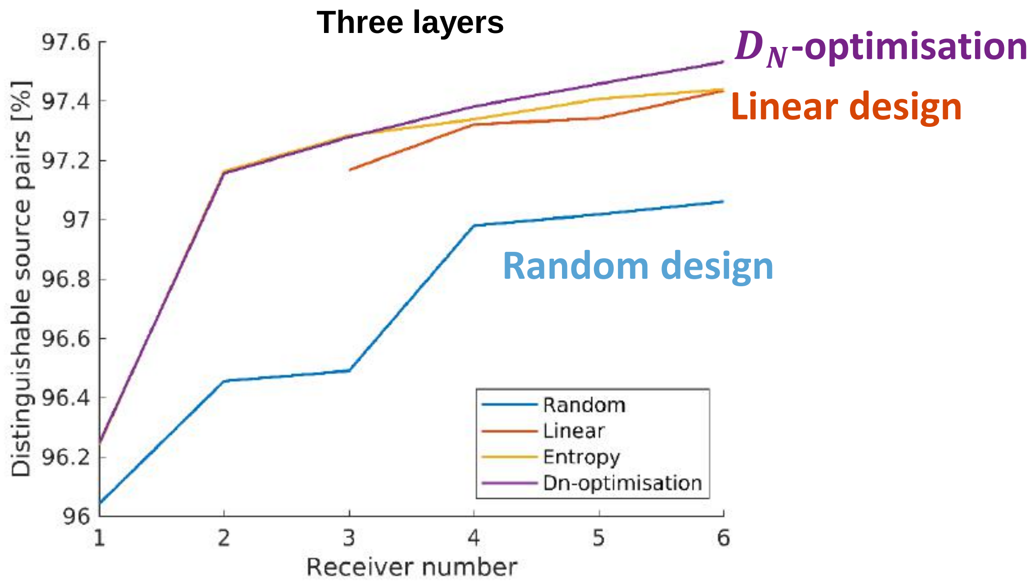


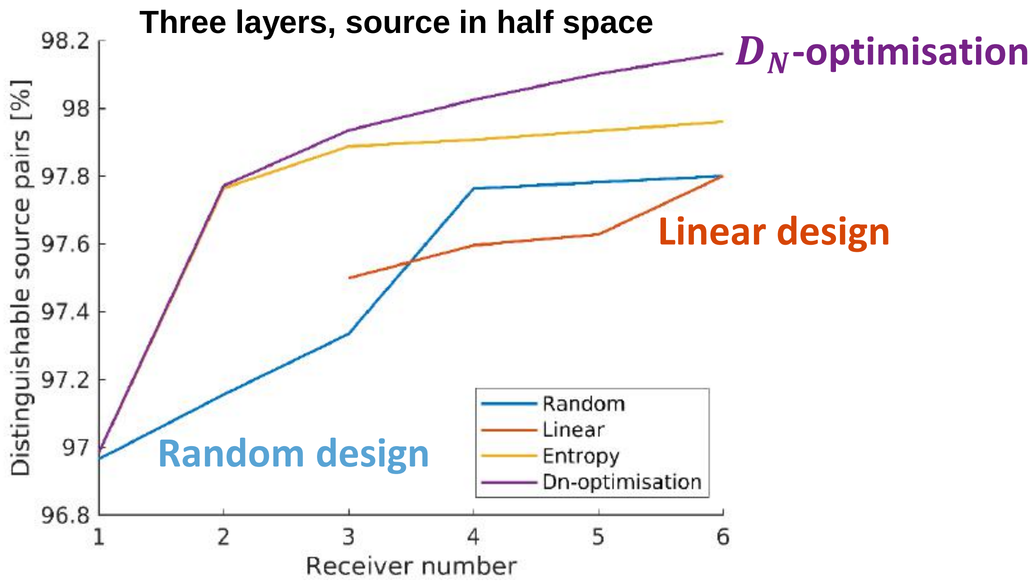
Which one is best?

Results: Design Performance

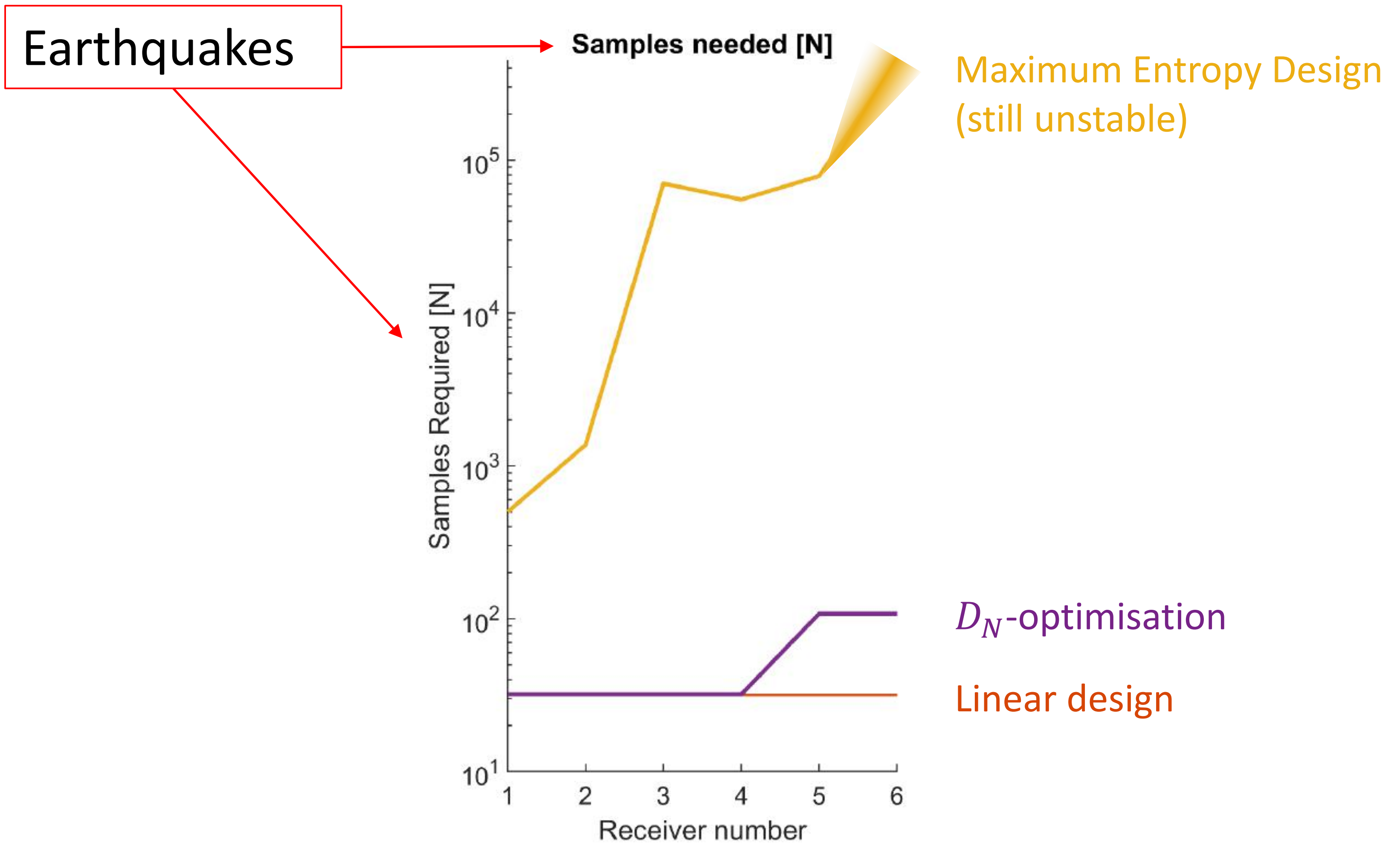




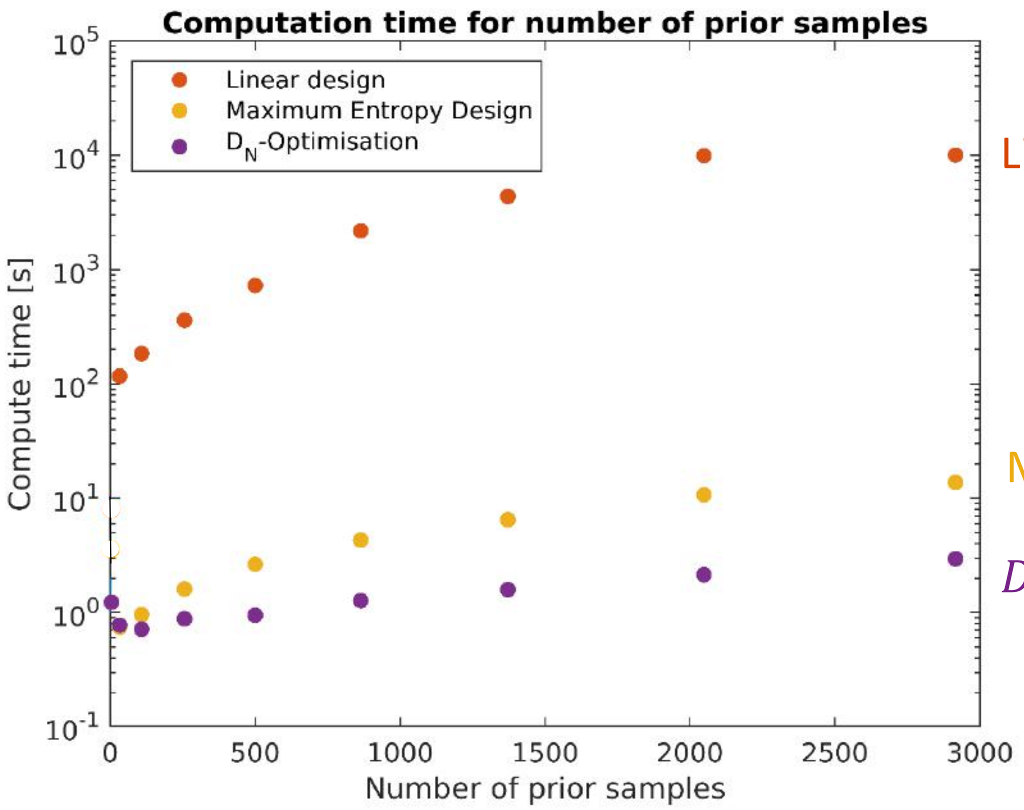




Results: Number of earthquakes for stability



Results: computation time



Linear design

Maximum Entropy Design

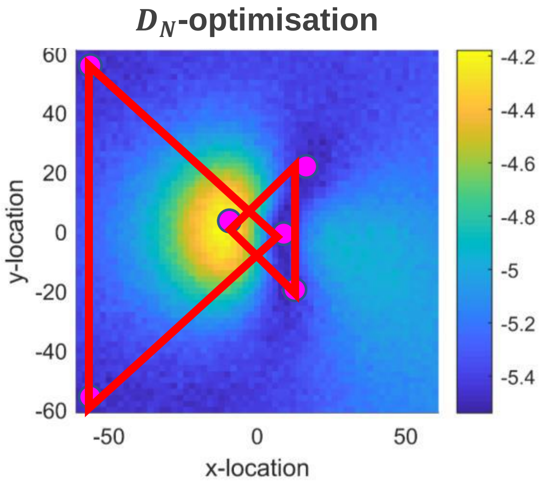
D_N-optimisation

Conclusion

- D-optimisation
  - Compute intensive
  - May *not* perform better than a random design in more complex situations
- Maximum Entropy Design
  - *Very* compute intensive
  - Requires *very* many prior samples
- D_N-optimisation
  - Quick to compute
  - Best performing networks

Darrel Coles & Curtis
- Geophysics 2011

Hugo Bloem, Curtis, Maurer
- Geophys. J. Int. 2020

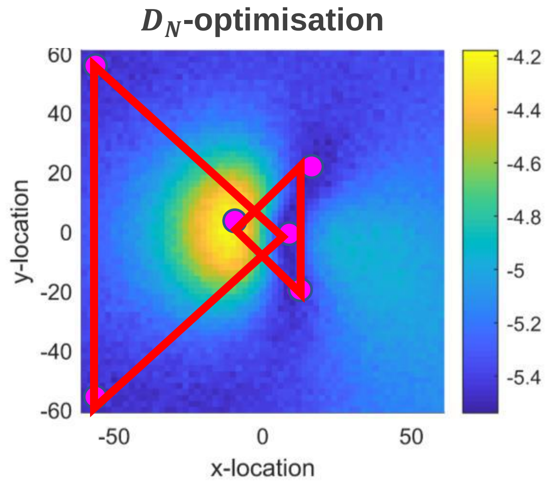


Conclusion

- D-optimisation
  - Compute intensive
  - May *not* perform better than a random design in more complex situation
- Maximum Entropy Design
  - *Very* compute intensive
  - Requires *very* many prior samples

Dominik Strutz
- Variational Design
- Maximise Information
- Interrogation Problems

- D_N-optimisation
  - Quick to compute
  - Best performing networks



The Leading Edge, 2004 – **Parts A (linear) and B (nonlinear)**
https://blogs.ed.ac.uk/curtis/

TUTORIAL

Theory of model-based geophysical survey and experimental design

Part A—linear problems

ANDREW CURTIS, Schlumberger Cambridge Research, Cambridge, U.K.

Enormous sums of money are invested by industry and scientific funding agencies every year in seismic, well logging, electromagnetic, earthquake monitoring and microseismic surveys, and in laboratory-based experiments. For each survey or experiment a design process must first take place. An efficient design is usually a compromise—a suitable trade-off between information that is expected to be retrieved about a model of interest and the cost of data acquisition and processing. In some fields of geophysics, advanced methods from design theory are used, not only to optimize the survey design, but also to shift this entire trade-off relationship between information and cost. In others, either crude rules of thumb are used or, indeed, expected model information is not optimized at all.

This is the first part of a two-part tutorial that provides a theoretical framework from the field of statistical experimental design (SED), within which model-based survey and experimental design problems and methods can be understood. Specifically, these two articles describe methods that are pertinent to the detection and inference of physical properties of rocks in the laboratory, or in the earth.

The choice of method to use when designing experiments depends greatly on how easily one can measure information. This in turn depends principally on whether the relationship between data that will be measured and model parameters of interest is approximately linear, or significantly nonlinear. Consequently, the first article focuses on the case where this relationship is approximately linear and the next (in next month's issue of *TLE*) deals with theory for nonlinear design.

surface to constrain optimally the shallow subsurface conductivity structure (Maurer and Boerner, GJI, 1998; Maurer et al., 2000); designing the interrogation of human experts to obtain optimal information to condition geophysical surveys (Curtis and Wood, 2004); designing nonlinear AVO surveys (van den Berg et al., 2003); planning crosswell seismic tomography surveys that illuminate the inter-well structure optimally (Curtis, 1999; Curtis et al., 2004); updating shallow resistivity survey designs in real-time as new data, and hence new information are acquired (Stummer et al., 2004); creating seismic acquisition geometries that maximize resolution of the earth model (Gibson and Tzimeas, 2002).

This tutorial considers the case where we would like to perform an experiment to collect data **d** (seismic, electromagnetic, logs, core, etc.) to constrain some model of earth properties or architecture described by a vector **m**. Say we define a set of basis functions $\{\mathbf{B}_j(\mathbf{x}): j=1,\ldots,P\}$ that describe elementary components of earth properties or architecture. Examples of such basis functions used in geophysics are rock properties in each of a set of mutually-exclusive spatial cells, discrete Fourier components over a finite band-width, scatterers of energy at a set of fixed locations, or statistical properties observed over a finite range of length scales. Possible models of the earth can then be expressed as:

$$M = \sum_{j=1}^{P} m_j \, \mathbf{B}_j(\mathbf{x}) \qquad (1)$$

The problem of estimating earth composition consists of estimating coefficients m_j.

THE UNIVERSITY of EDINBURGH

edinburgh imaging project

Thank you