Inverse Theory - a modern method of data analysis

Lecture 7

W. Dębski

1.06.2021

Inverse methods - some issues

- a marginal missing issue - errors

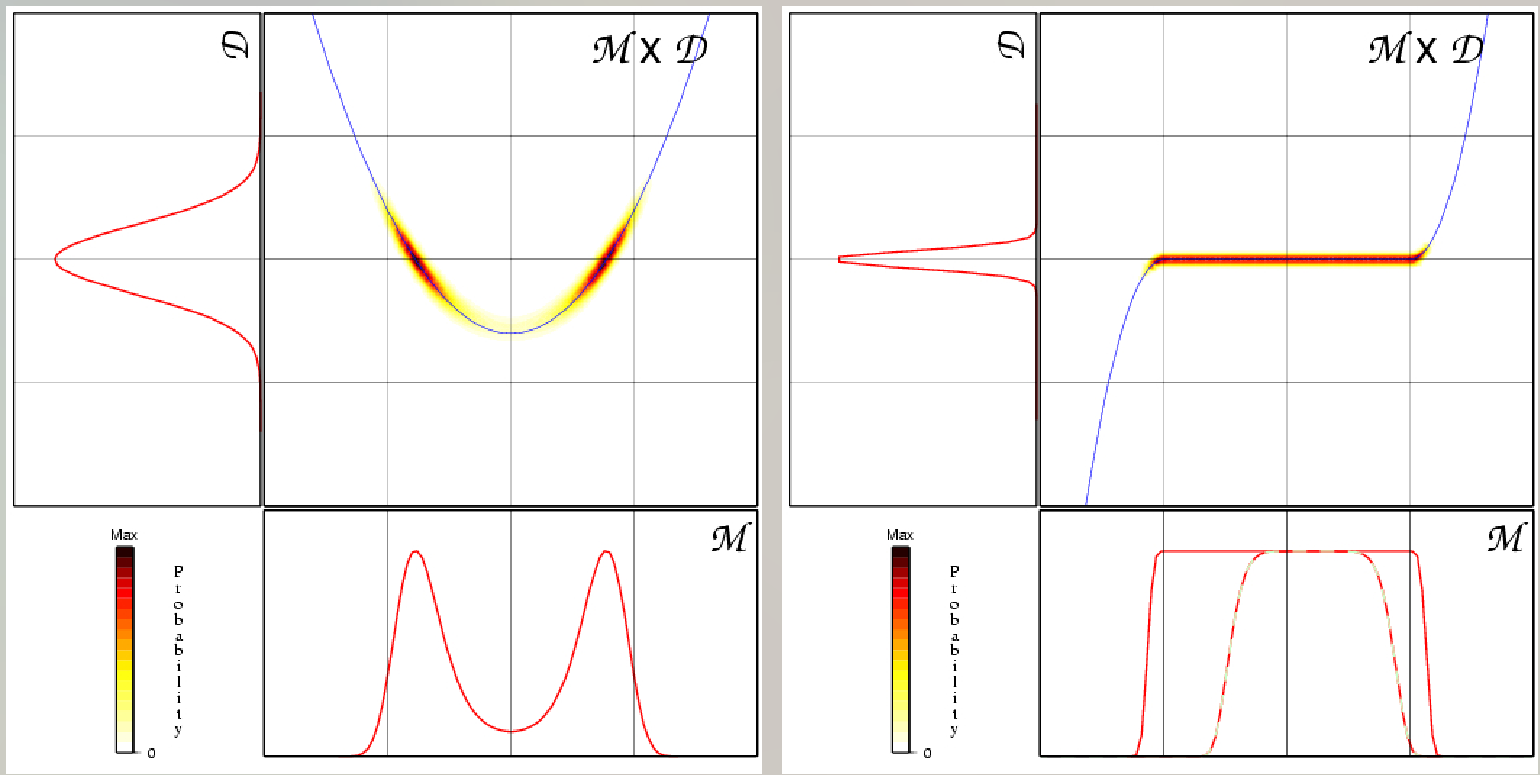
Optimization approach

- fully nonlinear method
- variety of existing optimization method
- Choice of $S(\mathbf{m})$ - different norms + additional constraints
- problem with error estimation
- is solution unique ?

Inverse methods - some issues

- a marginal missing issue - errors
- actually - the serious problem

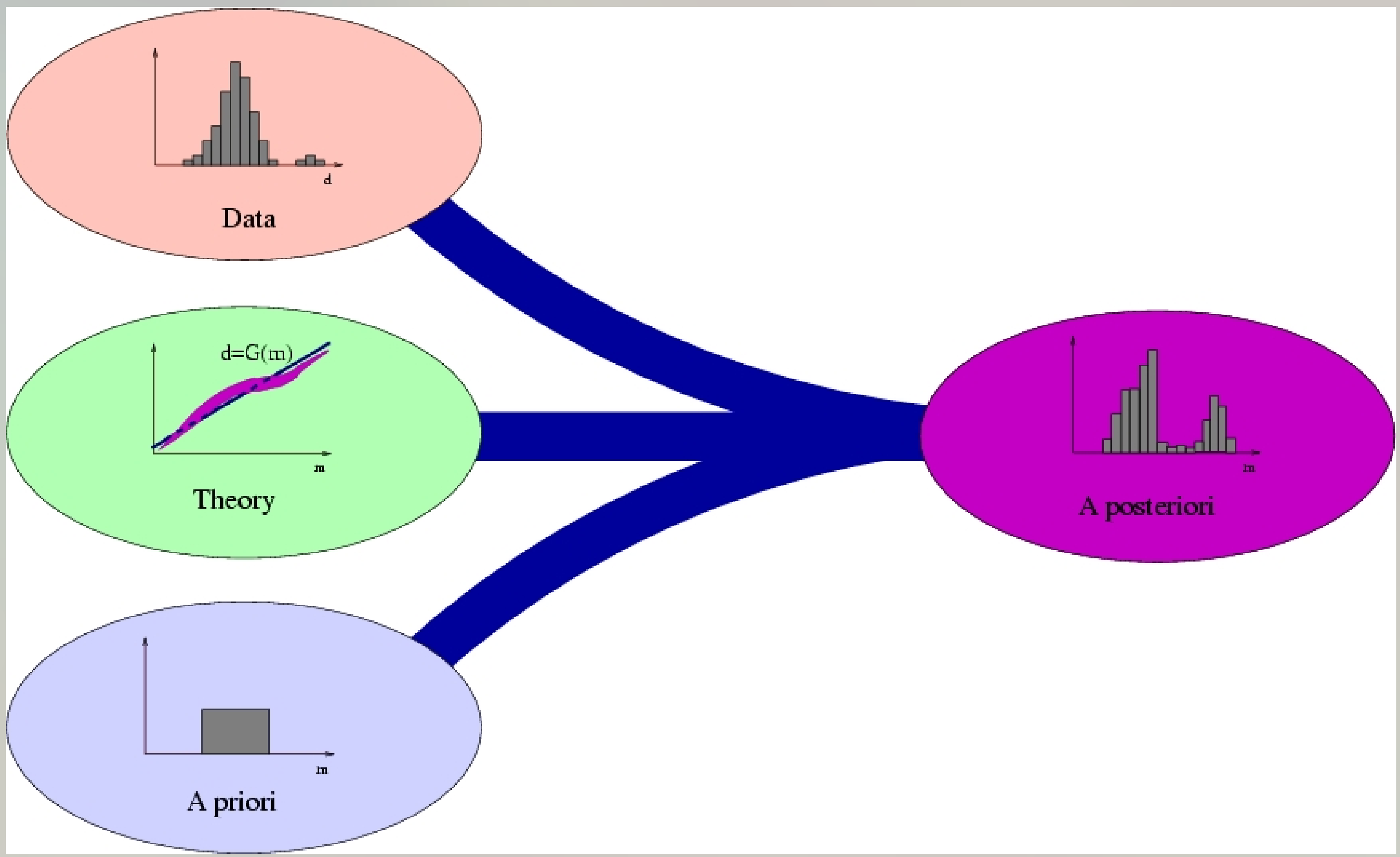
Uncertainties - inverse problems



Inverse methods - some issues

- a marginal missing issue - errors
- actually - the serious problem
- sources of final errors

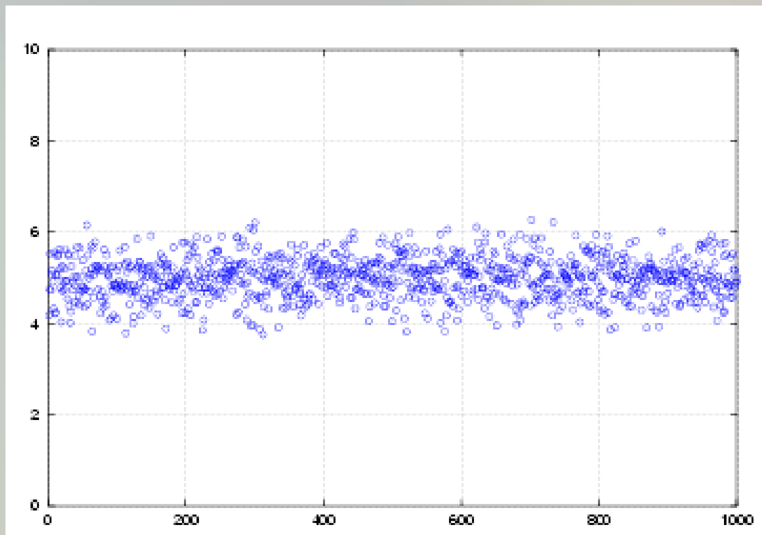
Source of *a posteriori* errors

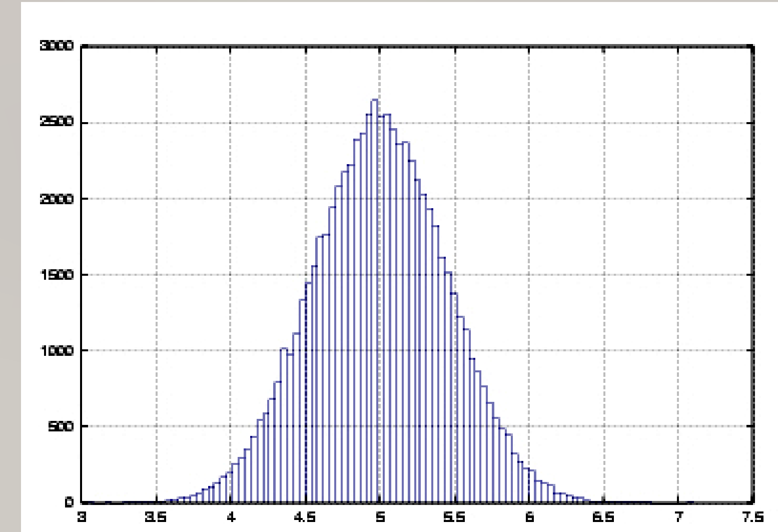


Inverse methods - some issues

- a marginal missing issue - errors
- actually - the serious problem
- final errors - sources
- error description - probability distribution

Uncertainties - probability distribution

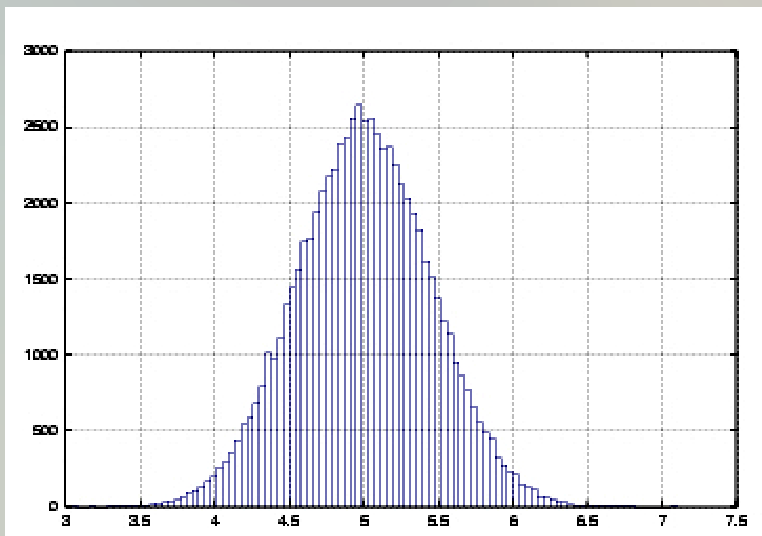




Output of experiment $\Longrightarrow$ probability distribution $p(\mathbf{d})$

Interpretation: $p(\mathbf{d})$ – probability that $\mathbf{d} = \mathbf{d}^{true}$

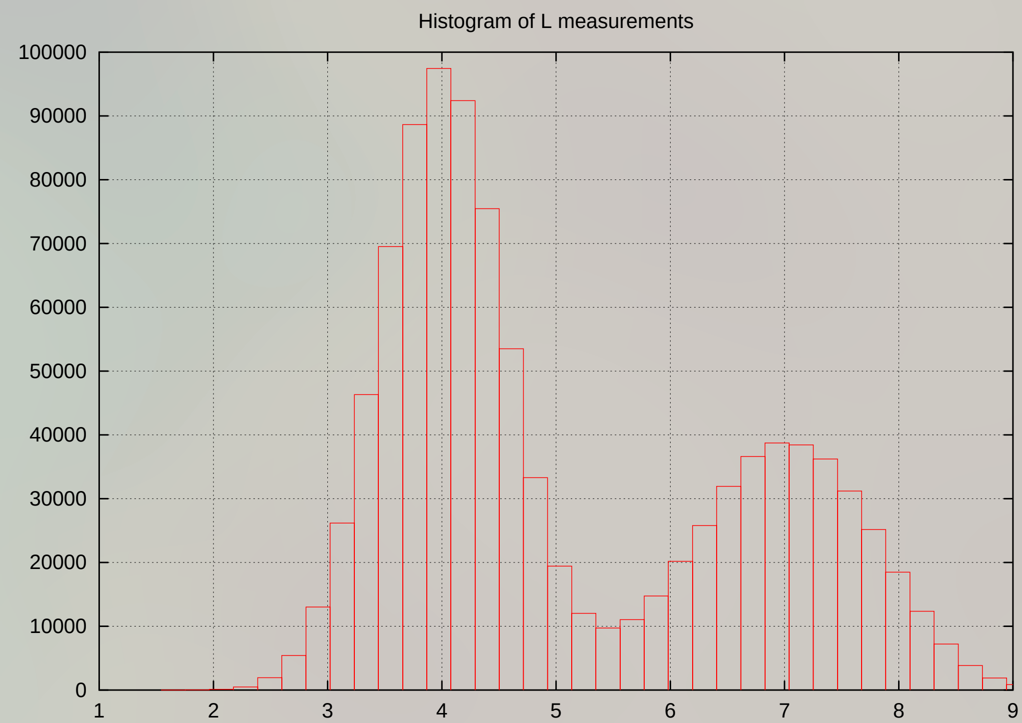
Gaussian Uncertainties



$$p(\mathbf{d}) \;=\; exp\left(-\frac{(\mathbf{d}-\bar{\mathbf{d}})^2}{2\delta^2}\right)$$

Output of experiment $\Longrightarrow \qquad \bar{d} \pm \delta d$

Uncertainties - probability distribution



Conclusion for inversion

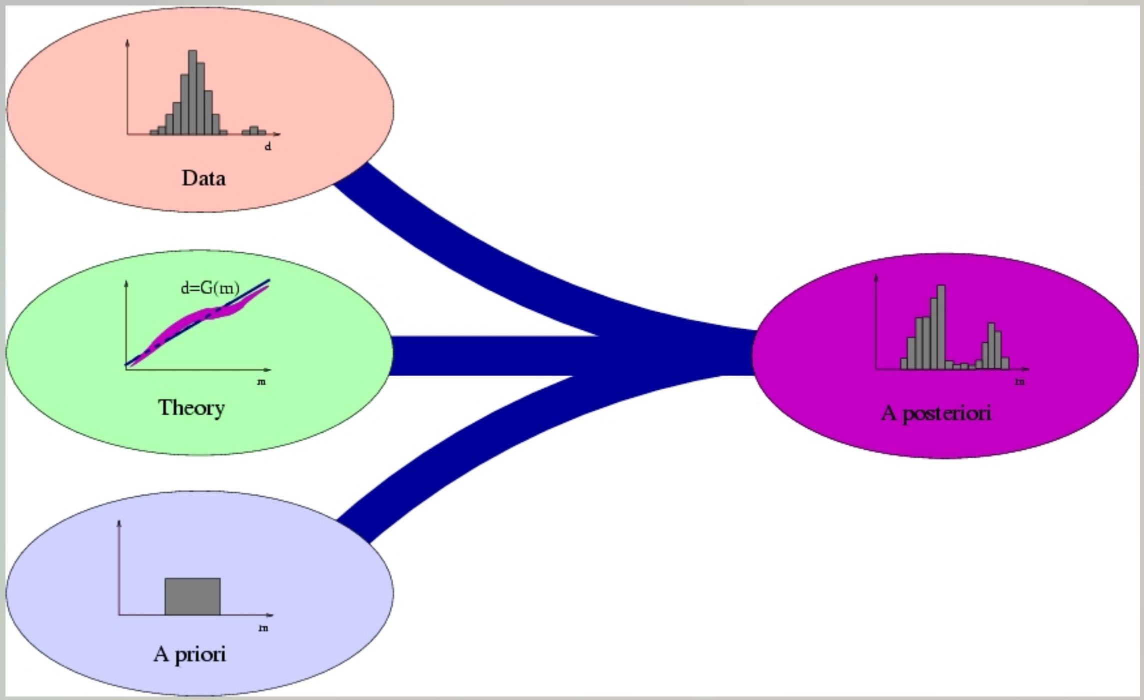
- Inverse problem $\Longleftrightarrow$ indirect measurement
- Output of experiment $\Longrightarrow$ probability distribution

$$\downarrow$$

Solution of inverse problem $\equiv$ probability distribution

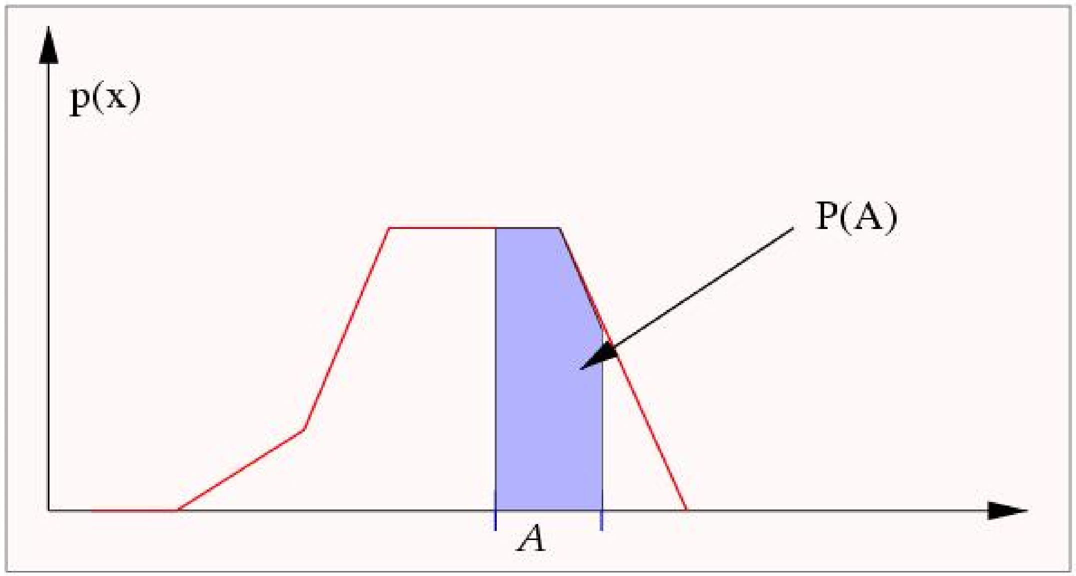
Interpretation: $\sigma(\mathbf{m})$ – probability that $\mathbf{m} = \mathbf{m}^{true}$

Probabilistic approach



HOW TO DO IT ?

Probability density



$$P(A) = \int_A p(x)dx$$

Probability density - features

- Normalization

$$\int_M p(m)dm = 1$$

- Average

$$\bar{m} =< m >= \int_M mp(m)dm$$

- Dispersion

$$\sigma^2 =< (m-\bar{m})^2 >= \int_M (m-\bar{m})^2 p(m)dm$$

Marginal and conditional probability distributions

$$p(\mathbf{m},\mathbf{d})$$

- $p_{\mathbf{m}}(\mathbf{m}) = \int_D p(\mathbf{m},\mathbf{d}) d\mathbf{d}$
- $p_{\mathbf{d}}(\mathbf{d}) = \int_M p(\mathbf{m},\mathbf{d}) d\mathbf{m}$
- $p_{\mathbf{m}|\mathbf{d}}(\mathbf{m}|\mathbf{d}) = p(\mathbf{m},\mathbf{d})/p_{\mathbf{d}}(\mathbf{d})$
- $p_{\mathbf{d}|\mathbf{m}}(\mathbf{d}|\mathbf{m}) = p(\mathbf{m},\mathbf{d})/p_{\mathbf{m}}(\mathbf{m})$

Marginal and conditional probability distributions (2)

$$p(\mathbf{m},\mathbf{d}) = p_{\mathbf{m}|\mathbf{d}}(\mathbf{m}|\mathbf{d})p_{\mathbf{d}}(\mathbf{d})$$

$$p(\mathbf{m},\mathbf{d}) = p_{\mathbf{d}|\mathbf{m}}(\mathbf{d}|\mathbf{m})p_{\mathbf{m}}(\mathbf{m})$$

$$p_{\mathbf{m}|\mathbf{d}}(\mathbf{m}|\mathbf{d})p_{\mathbf{d}}(\mathbf{d}) = p_{\mathbf{d}|\mathbf{m}}(\mathbf{d}|\mathbf{m})p_{\mathbf{m}}(\mathbf{m})$$

Bayes Theorem

$$p_{\mathbf{m}|\mathbf{d}}(\mathbf{m}|\mathbf{d}) = \frac{p_{\mathbf{d}|\mathbf{m}}(\mathbf{d}|\mathbf{m})\, p_{\mathbf{m}}(\mathbf{m})}{p_{\mathbf{d}}(\mathbf{d})}$$

Bayes Theorem - interpretation

$$\mathbf{d} \Longrightarrow \mathbf{d}^{obs}$$

$$p_{\mathbf{m}|\mathbf{d}}(\mathbf{m}|\mathbf{d}^{obs}) = \frac{p_{\mathbf{d}|\mathbf{m}}(\mathbf{d}^{obs}|\mathbf{m})p_{\mathbf{m}}(\mathbf{m})}{p_{\mathbf{d}}(\mathbf{d}^{obs})}$$

a posteriori probability distribution

$$p_{post}(\mathbf{m}|\mathbf{d}^{obs}) \sim p(\mathbf{d}^{obs}|\mathbf{m})\, p_{apr}(\mathbf{m})$$

$$\sigma(\mathbf{m}) = f(\mathbf{m})\, L(\mathbf{m}, \mathbf{d}^{obs})$$

Likelihood function

$$L(\mathbf{m}, \mathbf{d}^{obs}) = p(\mathbf{d}^{obs}|\mathbf{m})$$

theory			observation		
model		prediction	data		measured value
$\mathbf{m}$	$\rightarrow$	$G(\mathbf{m})$	$\mathbf{d}$	$\rightarrow$	$\mathbf{d}^{obs}$
	$\rho_{th}(\epsilon_{th})$			$\rho_o(\epsilon_o)$	
		$\rho_{th}(\mathbf{d} - G(\mathbf{m}))$			$\rho_o(\mathbf{d} - \mathbf{d}_o)$

$$L(\mathbf{m}, \mathbf{d}^{obs}) = \int_d \rho_{th}(\mathbf{d} - G(\mathbf{m}))\, \rho_o(\mathbf{d} - \mathbf{d}_o)\, d\mathbf{d}$$

Probabilistic solution

$$\sigma(\mathbf{m}) = f(\mathbf{m})\, L(\mathbf{m}, \mathbf{d}^{obs})$$

$$L(\mathbf{m}, \mathbf{d}^{obs}) \; = \; \int_d \rho_{th}(\mathbf{d} - G(\mathbf{m}))\, \rho_o(\mathbf{d} - \mathbf{d}_o)\, d\mathbf{d}$$

Example - exact theory

$$\rho_{th} = \delta(\mathbf{d} - \mathbf{G}(\mathbf{m}))$$

then

$$\sigma(\mathbf{m}) = f(\mathbf{m})\, \rho_o(\mathbf{d}^{obs} - \mathbf{G}(\mathbf{m}))$$

likelihood function represents uncertainties due to measurement errors

Example - exact measurements

$$\rho_o = \delta(\mathbf{d} - \mathbf{d}^{obs})$$

then

$$\sigma(\mathbf{m}) = f(\mathbf{m})\, \rho_{th}(\mathbf{d}^{obs} - \mathbf{G}(\mathbf{m}))$$

likelihood function represents uncertainties due to modelling errors

Example - missing theory

$$\rho_{th}(\mathbf{m},\mathbf{d}) \;=\; const.$$

then

$$\sigma(\mathbf{m}) = f(\mathbf{m}) \int_{\mathbf{d}} \rho_o(\mathbf{d}-\mathbf{d}_o)\, d\mathbf{d} \;\sim\; f(\mathbf{m})$$

solution is determined by *a priori* model+uncertainties

Example - missing measurement data

$$\rho_o(\mathbf{d}) \;=\; const.$$

then

$$L(\mathbf{m}) = \int_{\mathbf{d}} \rho_{th}(\mathbf{d} - G(\mathbf{m}))\, d\mathbf{d}$$

$$\sigma(\mathbf{m}) \;\sim\; f(\mathbf{m})$$

solution is determined by *a priori* model+uncertainties

Example - linear modelling + Gaussian errors

$$\mathbf{d} = \mathbf{G} \cdot \mathbf{m}$$

$$f(\mathbf{m}) = \exp\left(-(\mathbf{m} - \mathbf{m}^{apr})^T \, \mathbf{C}_m^{-1} \, (\mathbf{m} - \mathbf{m}^{apr})\right)$$

$$\rho_{th}(\mathbf{m}) = \exp\left(-(\mathbf{d} - \mathbf{G} \cdot \mathbf{m})^T \, \mathbf{C}_t^{-1} \, (\mathbf{d} - \mathbf{G} \cdot \mathbf{m})\right)$$

$$\rho_o(\mathbf{d}) = \exp\left(-(\mathbf{d} - \mathbf{d}^{obs})^T \, \mathbf{C}_d^{-1} \, (\mathbf{d} - \mathbf{d}^{obs})\right)$$

$$\sigma(\mathbf{m}) = \exp\left(-(\mathbf{m} - \mathbf{m}^{ml})^T \, \mathbf{C}_p^{-1} \, (\mathbf{m} - \mathbf{m}^{ml})\right)$$

$$\mathbf{m}^{ml} = \mathbf{m}^{apr} + \mathbf{C}_p^{-1}\mathbf{G}^T\mathbf{C}_d^{-1} \cdot \left(\mathbf{d}^{obs} - \mathbf{G} \cdot \mathbf{m}^{apr}\right)$$

$$\mathbf{C}_p = \left(\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G} + \mathbf{C}_m^{-1}\right)$$

Probabilistic solution - features

The solution: *a posteriori* probability distribution

- always exists
- fully nonlinear
- include all uncertainties
- possible full error analysis
- physical well define meaning (and role) of *a priori* term
- requires methods of exploring $\sigma(\mathbf{m})$
- non-parametric inverse problems?

Exploring *a posteriori* probability

- searching for maximum of $\sigma(\mathbf{m})$
- calculate point estimators
- marginal distributions
- sampling $\sigma(\mathbf{m})$

Exploring *a posteriori* probability: $\mathbf{m}^{ml}$ solution

Basic characteristic of $\sigma(\mathbf{m})$:

location of the (global) maximum
- the most likelihood $\mathbf{m}^{ml}$ value

$$\mathbf{m}^{ml} : \quad \sigma(\mathbf{m}) = \max$$

Problem reduced to optimization approach

Point estimators

Characterization of $\sigma(\mathbf{m})$ by its moments:

- average value: $\mathbf{m}^{avr} = \int_M \mathbf{m}\,\sigma(\mathbf{m})\mathrm{d}\mathbf{m}$
- covariance $C_{ij} = \int_M (m_i - m_i^{avr})(m_j - m_j^{avr})\,\sigma(\mathbf{m})\mathrm{d}\mathbf{m}$
- higher order moments

Require efficient methods of calculation multi-dimensional integrals

Marginal *a posteriori* distribution

- 1D marginals

$$\sigma_i(m_i) = \int_{\mathbf{m} \neq m_i} \sigma(\mathbf{m})\, \mathrm{d}\mathbf{m}$$

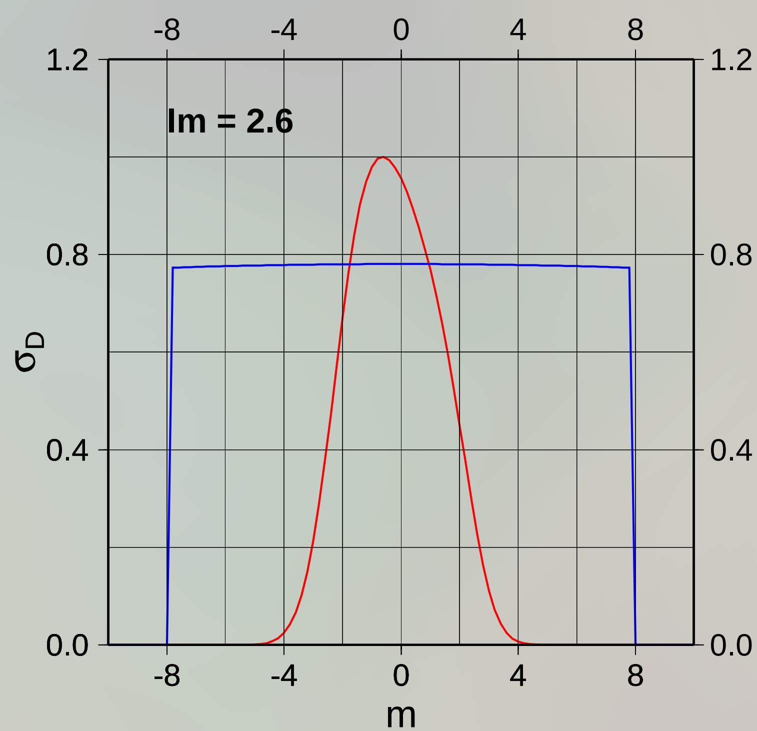
- 2D marginals

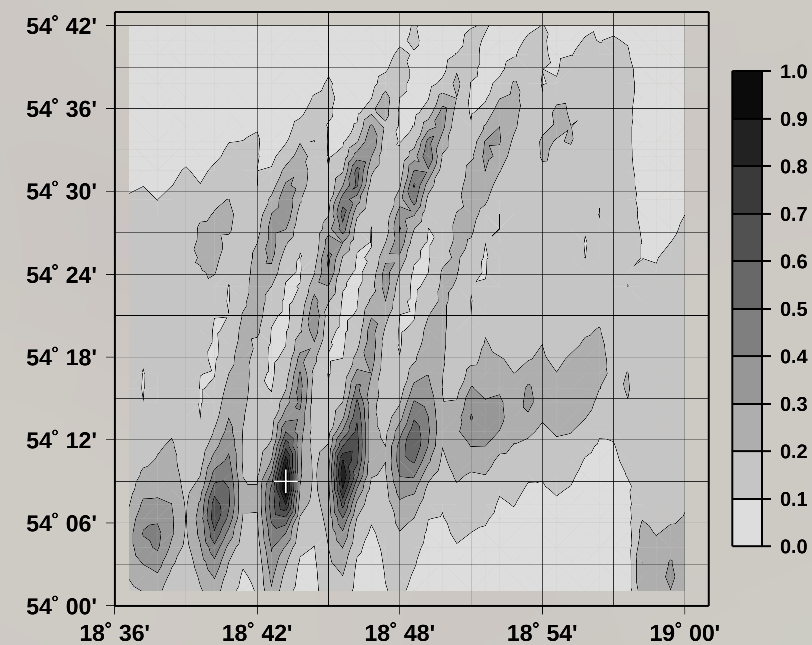
$$\sigma_{ij}(m_i, m_j) = \int_{\mathbf{m} \neq m_i, m_j} \sigma(\mathbf{m})\, \mathrm{d}\mathbf{m}$$

- higher dimension marginals

Require efficient methods of calculation multi-dimensional integrals

Marginal *a posteriori* distribution - examples





Sampling *a posteriori* distribution

- geometric sampling - grid search
- adaptive grid search (near neighborhood algorithm)
- stochastic (Monte Carlo sampling)

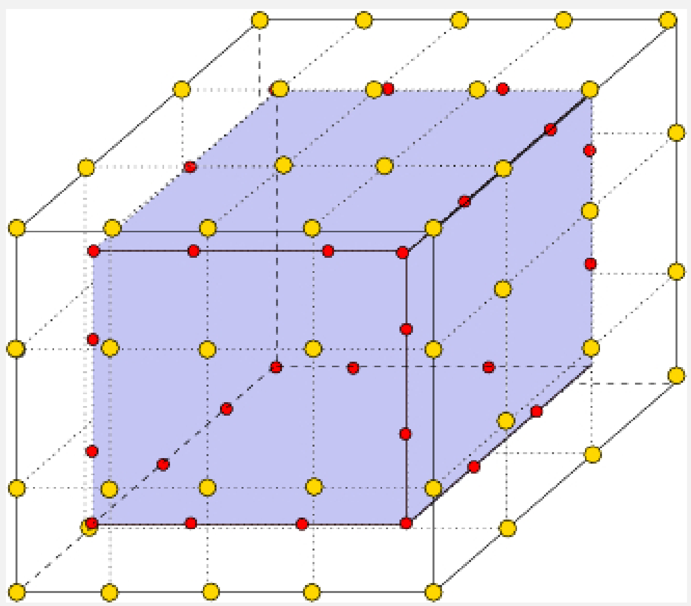
$$\sigma(\mathbf{m}) \Rightarrow \sigma_{i,j,k,\ldots} = \sigma(m_i, m_j, m_k, \ldots)$$

- very general
- only for small dimensional problem
- non-uniform sampling ...

Geometric sampling in multi-dimensional spaces

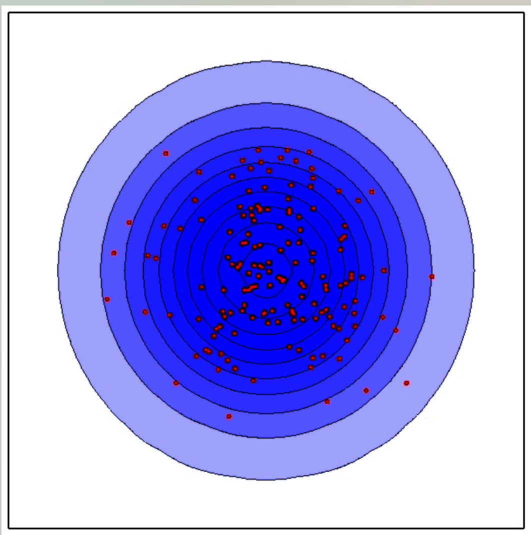
Non-uniform sampling:

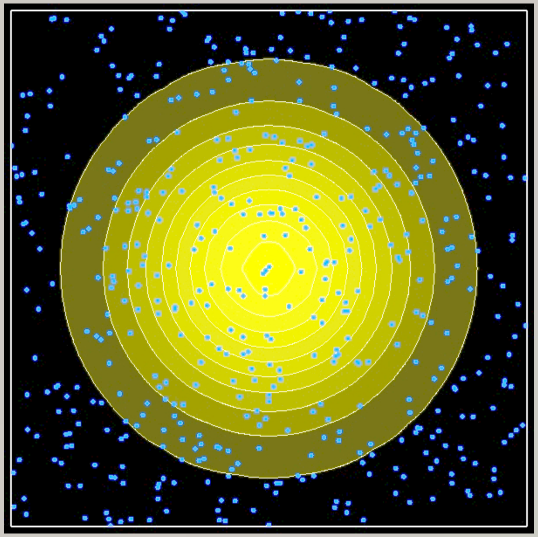
$$\frac{N_V}{N} = \left(\frac{p-2}{p}\right)^N \overset{p \gg 2}{\approx} e^{-2N/p} \xrightarrow{N} 0$$



Sampling *a posteriori* distribution

- geometric sampling - grid search
- adaptive grid search (near neighborhood algorithm)
- stochastic (Monte Carlo sampling)





End