Probabilistic Inverse Theory

Lecture 6

W. Dębski

19.04.2023

Optimization approach - direct search for best m

Optimization approach

$$S(\mathbf{m}) = ||\mathbf{d}^{obs} - \mathbf{d}^{th}(\mathbf{m})|| + ||\mathbf{m} - \mathbf{m}^{apr}||$$

solution: search for $\mathbf{m}^{ml}$ minimizing $S(\mathbf{m})$

$$S(\mathbf{m}^{ml}) = \min$$

Optimization approach - features

- 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 problem $\equiv$ indirect measurement

Errors are important!

Estimating measurement errors - what does it mean

When measuring parameter $\mathbf{p}$ we provide (or should do) result in the form

$$\mathbf{p} = \mathbf{p}^o \pm \delta\mathbf{p}$$

where $\delta\mathbf{p}$ is an error measurement.

What this notion does it mean?

Uncertainties - stochastic point of view

stochastic errors:

$$p_i \ = \ p^{true} + \epsilon$$

with given probability distribution $\rho(\epsilon)$ (usually unknown)

if sufficient number of measurements is available

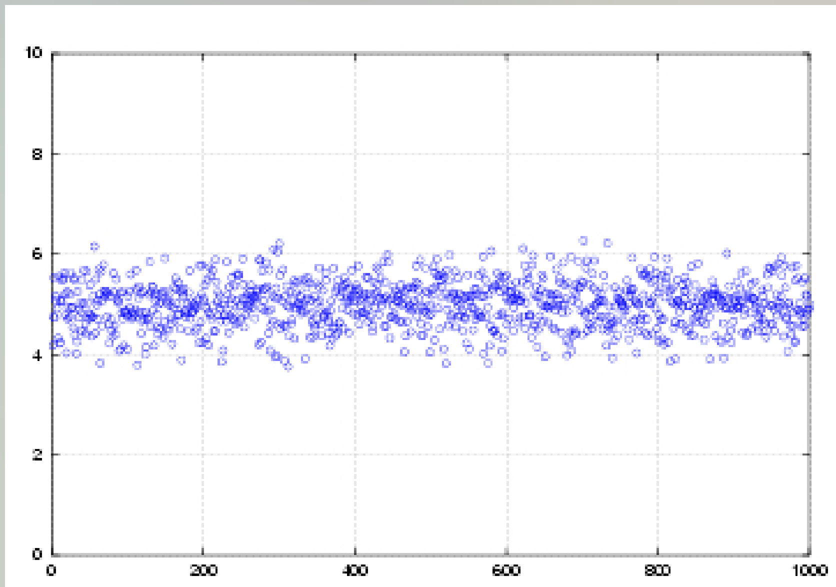
$$\rho(\epsilon) \approx H(p - \bar{p}) \qquad \bar{p} = \frac{1}{N}\sum_i^N p_i$$

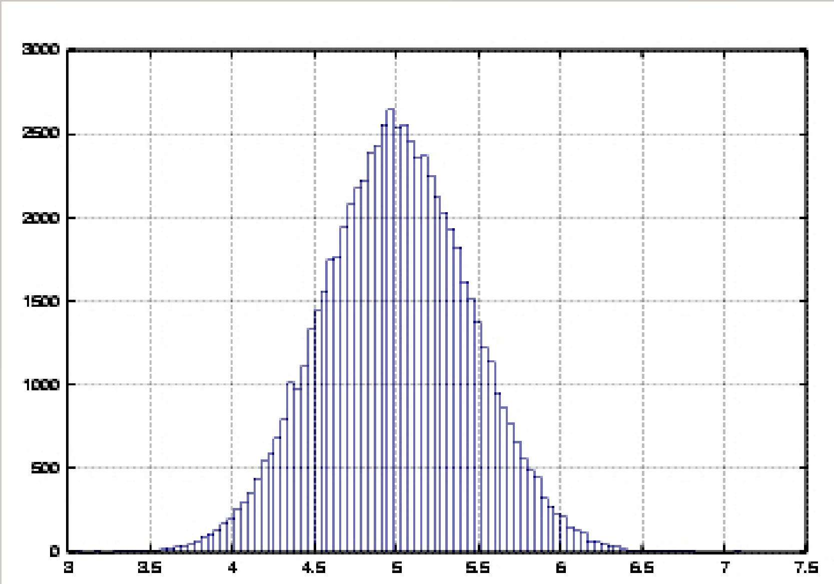
can be a good proxy of $\rho(\epsilon)$ and

$$\mathbf{p}^{true} \ = \bar{\mathbf{p}} \pm \delta\mathbf{p}$$

provided ...

Uncertainties - Gaussian errors

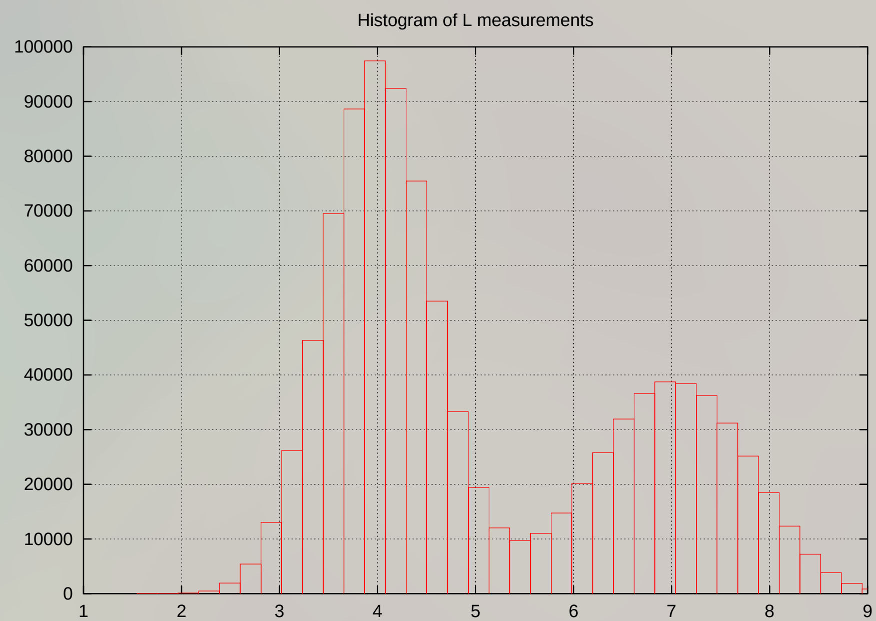




$$P(\bar{p} - \delta \; < \; p^{true} \; < \; \bar{p} + \delta) = 0.68$$

Uncertainties - non-Gausian errors

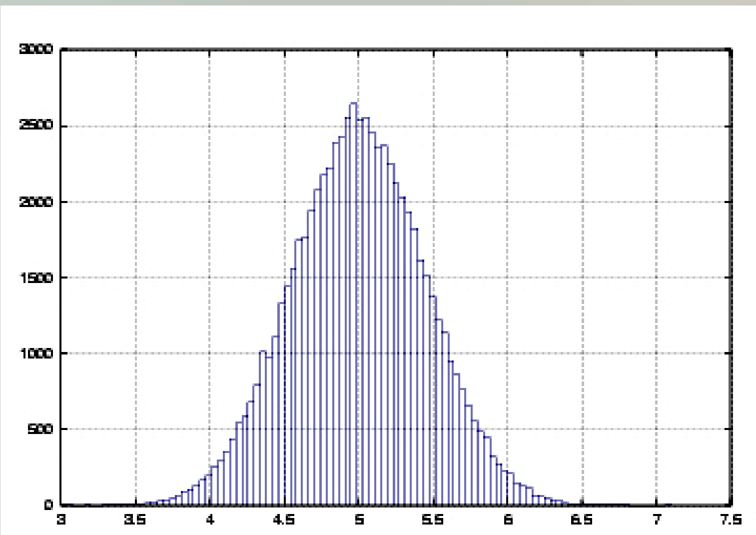
However, the error statistics can be quite complex, multi-modal, etc.



$$\mathbf{p}^{est} \neq \bar{\mathbf{p}} \pm \delta\mathbf{p}$$

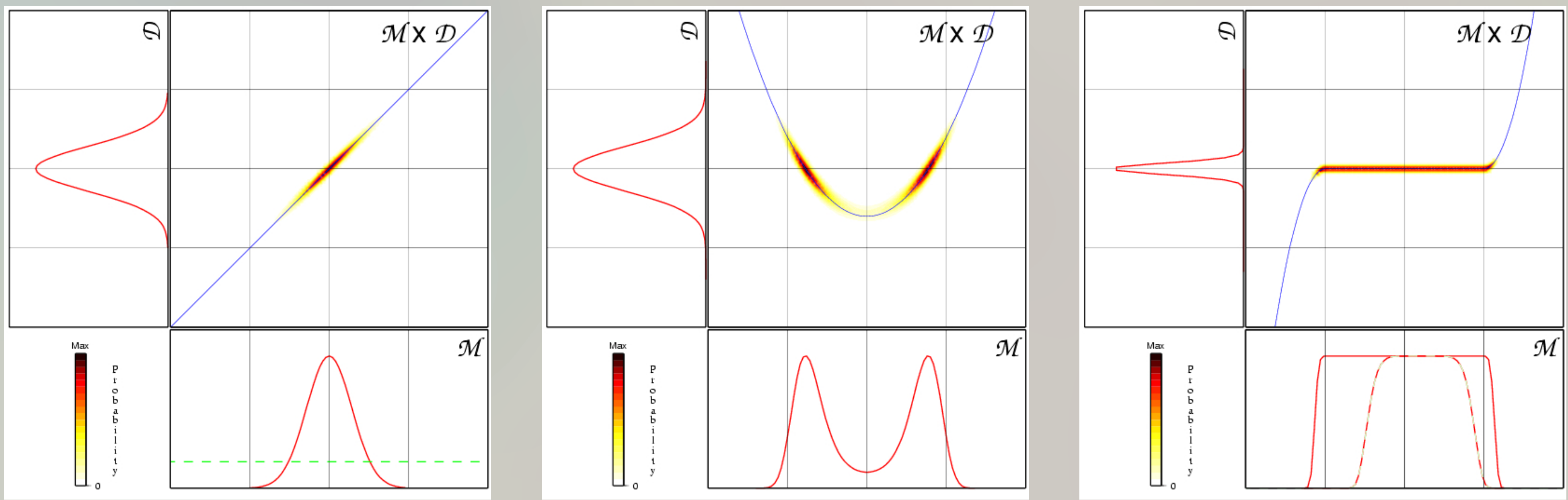
Uncertainties - generalization

The correct error estimation requires providing the **complete** description of error statistics $\rho(\epsilon)$



This is what we need in case of inverse problems!

Complication due to non-linearity



Uncertainties - conclusion

Nonlinearity of forward problem can significantly modify "observational" errors. If, in addition we consider that forward modelling can also be imprecise (e.g. due to theoretical simplification) we conclude that **inversion errors** can have complex distribution $\rho(\epsilon)$. Without its full knowledge the solution of inverse problem is

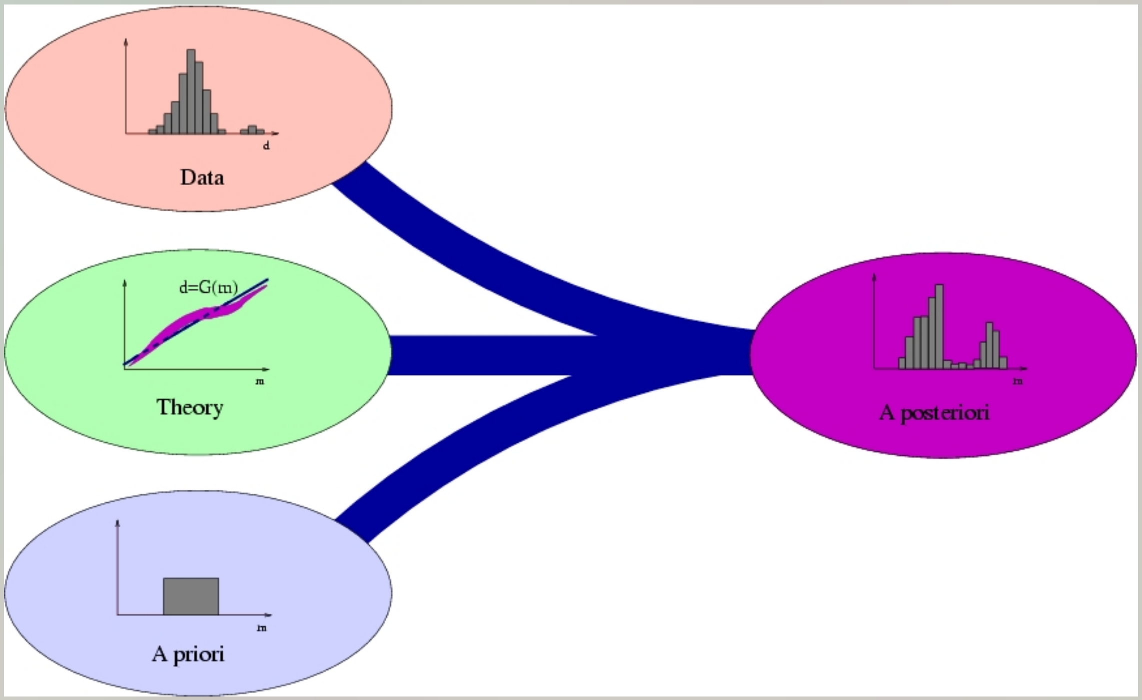
INCOMPLETE

Generalization - probabilistic approach

Solution ≡ searching for

probability distribution

describing *a posteriori* uncertainties



Generalization - probabilistic approach

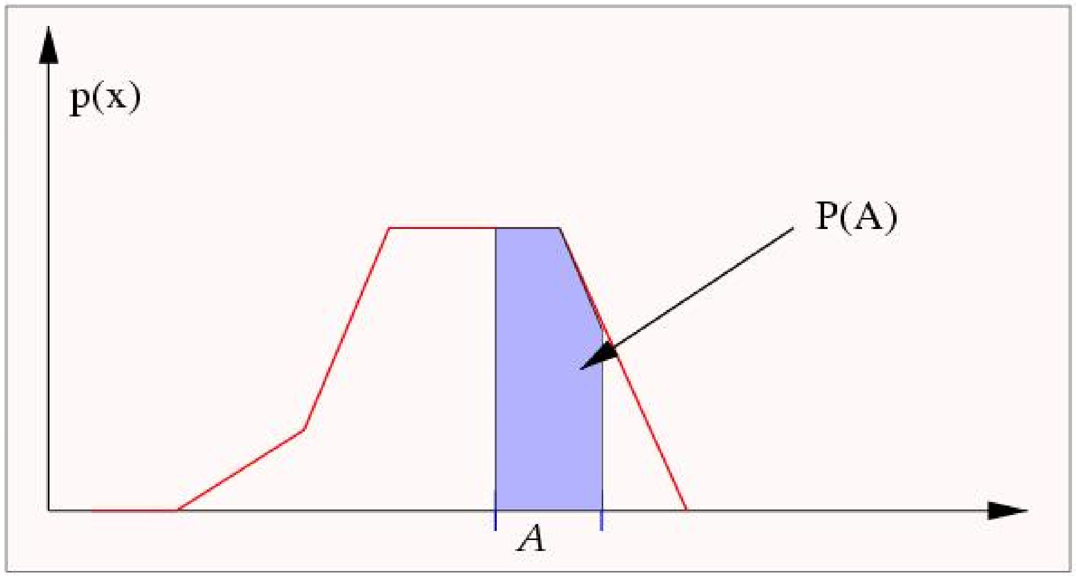
Method	*Solution*
Algebraic	model given by math formula
Optimization	single“optimum” model
$\Downarrow$	
Probabilistic	probability distribution

This *a posteriori* distribution is supposed to describe (at least) inversion errors

Probabilistic approach

How to construct this
a posteriori
probability distribution

Probability theory - elements



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

Probability theory - elements

- Normalization $[p(m) \geq 0]$

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

- Average (Expected value)

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

- Dispersion

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

Marginal and conditional probability distributions

$p(\mathbf{m},\mathbf{d})$: $\int_{MD} p(\mathbf{m},\mathbf{d}) d\mathbf{m} d\mathbf{d} = 1$

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

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
$\mathbf{m}$	$\rightarrow$	$G(\mathbf{m})$	$\mathbf{d}$	$\rightarrow$	$\mathbf{d}^{obs}$
	$\rho_{th}(\epsilon_{th})$			$\rho_o(\epsilon_o)$	
		$\rho_{th}(\mathbf{d} - G(\mathbf{m}))$			$\rho_o(\mathbf{d} - \mathbf{d}^{obs})$

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

Probabilistic solution

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

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

Example - exact theory

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

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

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

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

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

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

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

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

then

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

solution is determined by *a priori* model+uncertainties

Example - missing measurement data

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

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

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

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

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

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

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

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

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

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

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: 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:

- expected 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

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

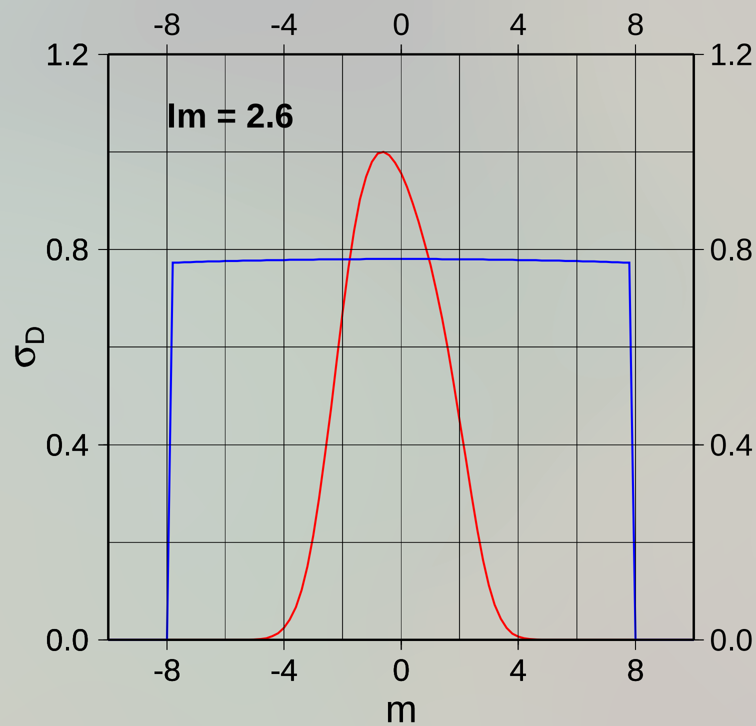
- 2D marginals

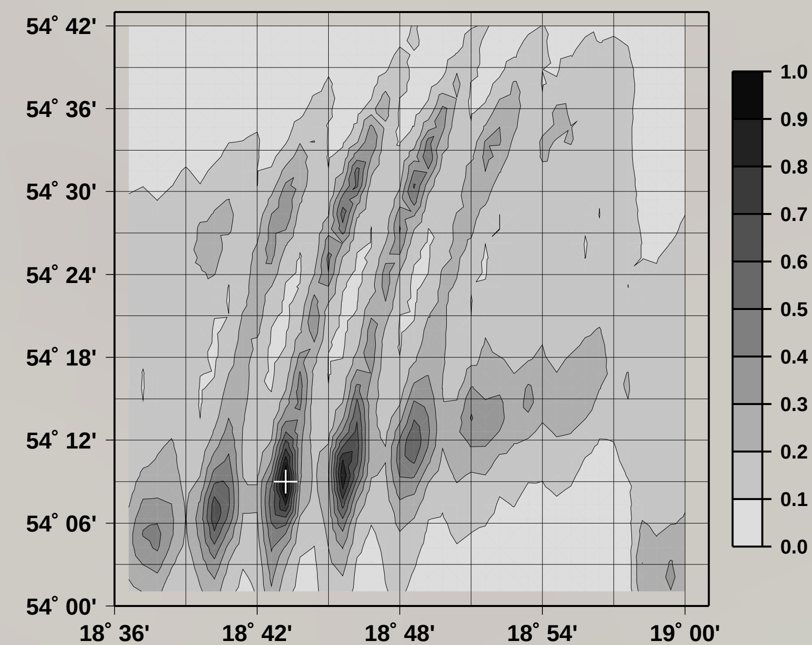
$$\sigma_{ij}(m_i, m_j) \; = \int\limits_{\mathbf{m} \neq m_i, m_j} \sigma(\mathbf{m}) \, \mathrm{d}\mathbf{m}$$

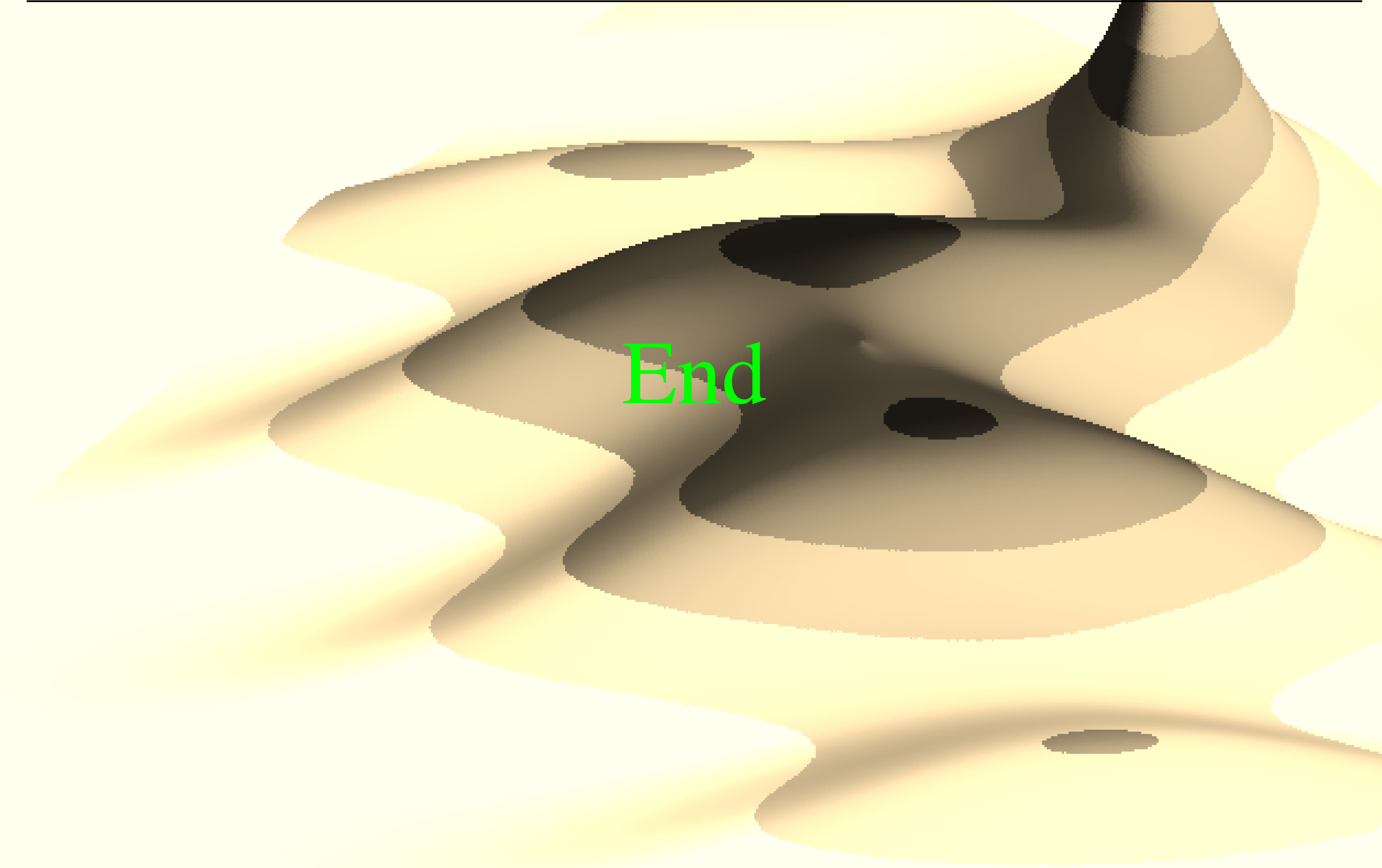
- higher dimension marginals

Require efficient methods of calculation multi-dimensional integrals

Marginal *a posteriori* distribution - examples







End