Probabilistic Inverse Theory

Lecture 8

W. Dębski

24.05.2023

Advanced issue (A): measuring continuous "data" misfit

$$||\mathbf{d}^{obs}(t) - \mathbf{d}^{th}(t)||^2 \; = \; \left\langle \mathbf{d}^{obs} - \mathbf{d}^{th}, \, \mathbf{d}^{obs} - \mathbf{d}^{th} \right\rangle$$

$$< h, g > = \; \mathrm{Re} \int_0^{+\infty} \frac{\bar{h}(f) * g(f)}{S_n(f)} \, \mathrm{d}f$$

$S_n(f)$ - observational noise power spectrum

Advanced issue (B): Time Reversal Method

$$\frac{1}{c^2}\frac{d^2p}{dt^2} = \Delta p + S$$

Inverse task:

$$p^{obs}(t) \rightarrow S(t)$$

TRM as the inverse problem

$$\mathbf{m} = S(t) \qquad \mathbf{d} = p(t)$$

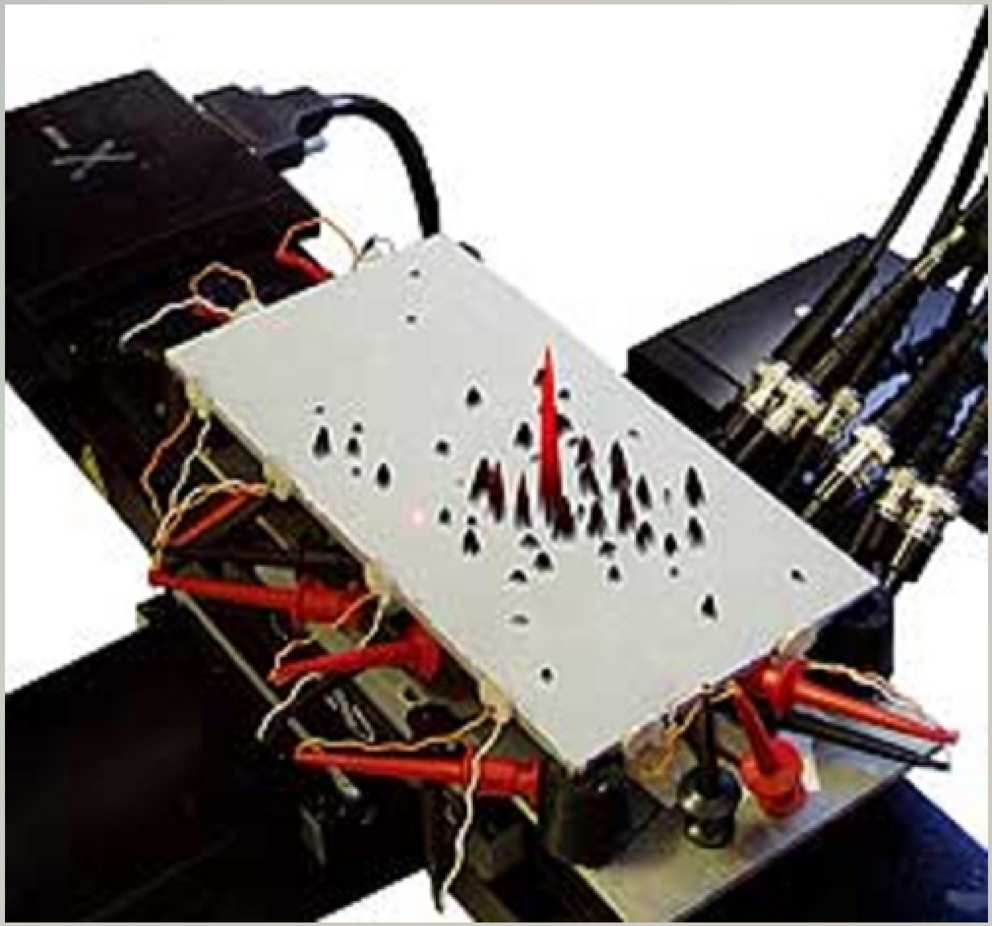
$$p(r,t) = \int G^{+}(r, t-t')\, S(t')\mathrm{d}t' \quad \Rightarrow \quad \mathbf{d} = \mathcal{G}\mathbf{m}$$

$$S(t) = \int G^{+}(r, t-t')\, p^{obs}(-t')\mathrm{d}t' \Rightarrow \quad \mathbf{m} = (\mathcal{G}^T\mathcal{G} + \lambda)^{-1}\mathcal{G}^T\mathbf{d}$$

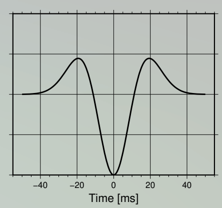
The method is computationally very efficient, but ...

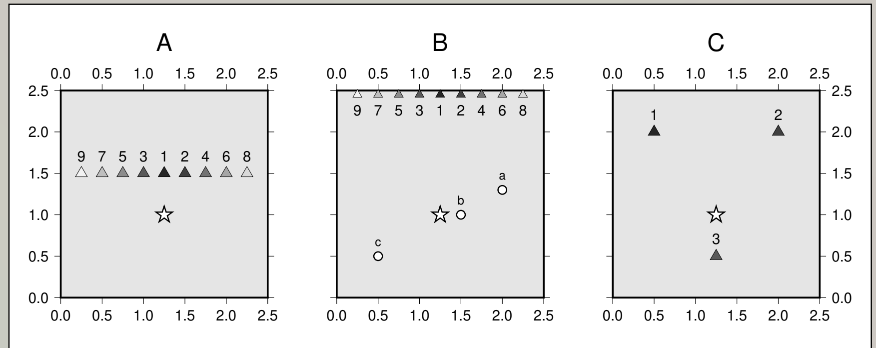
we need to know the medium: $G^{+}(\cdot,\cdot)$

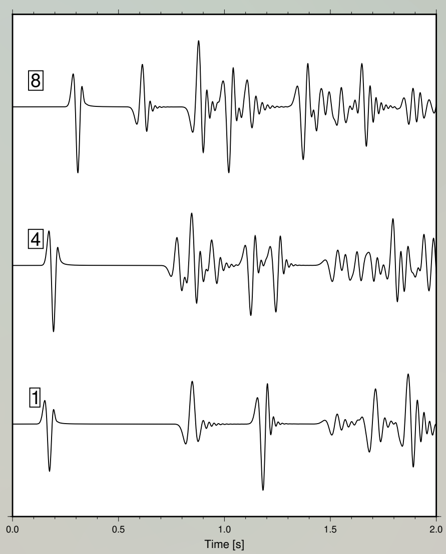
Experimental evidence (M. Fink)

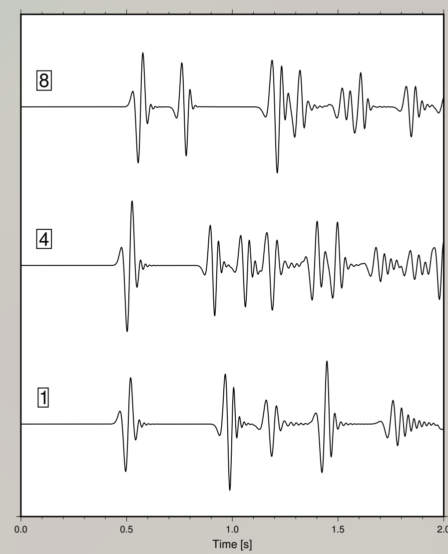


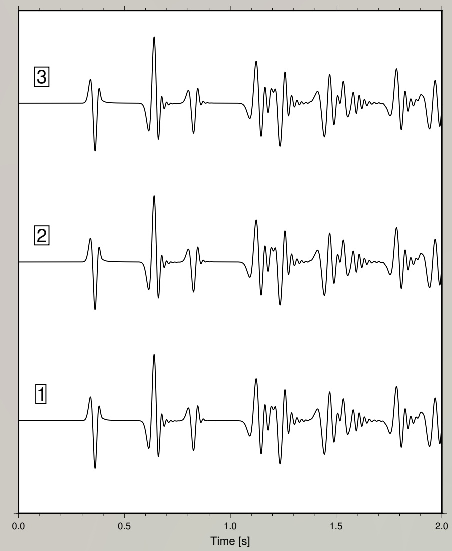
TRM and numerical simulations (K. Waskiewicz)



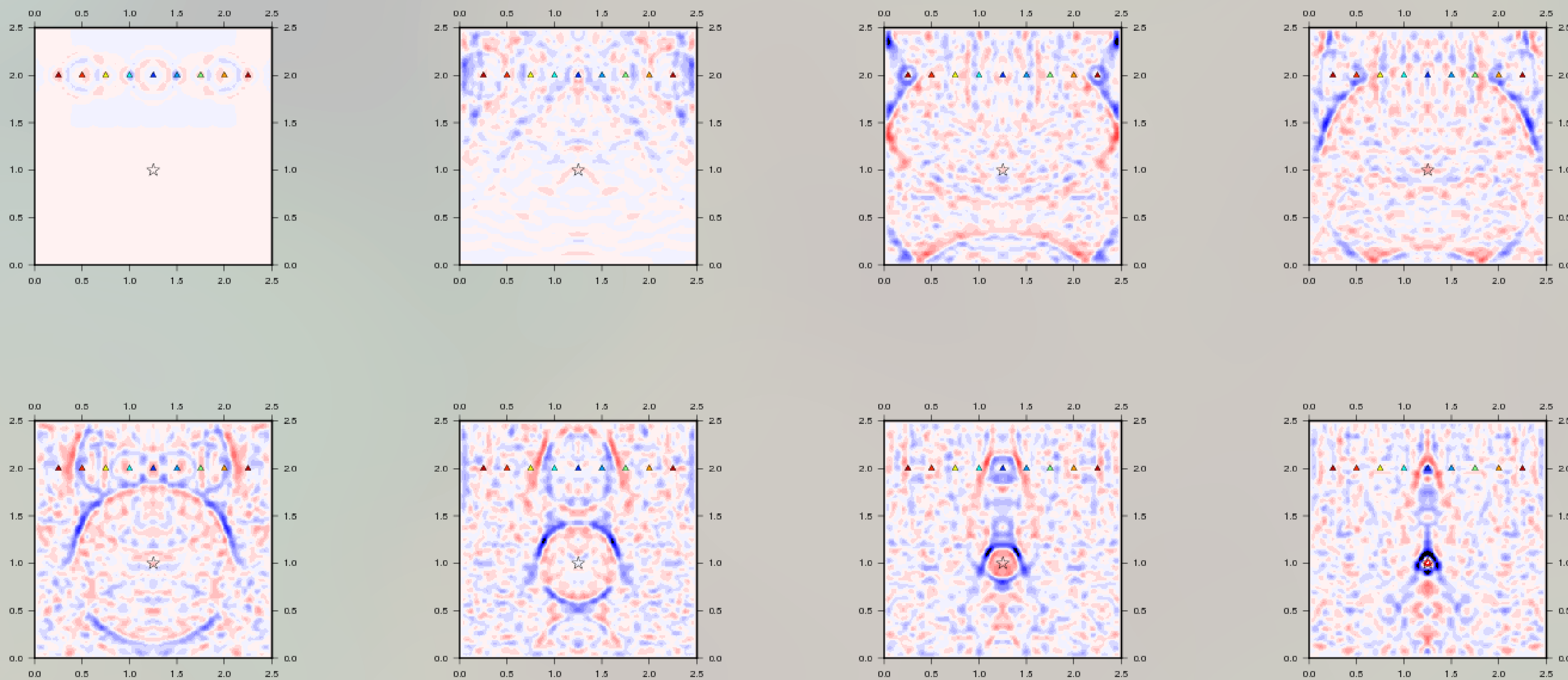




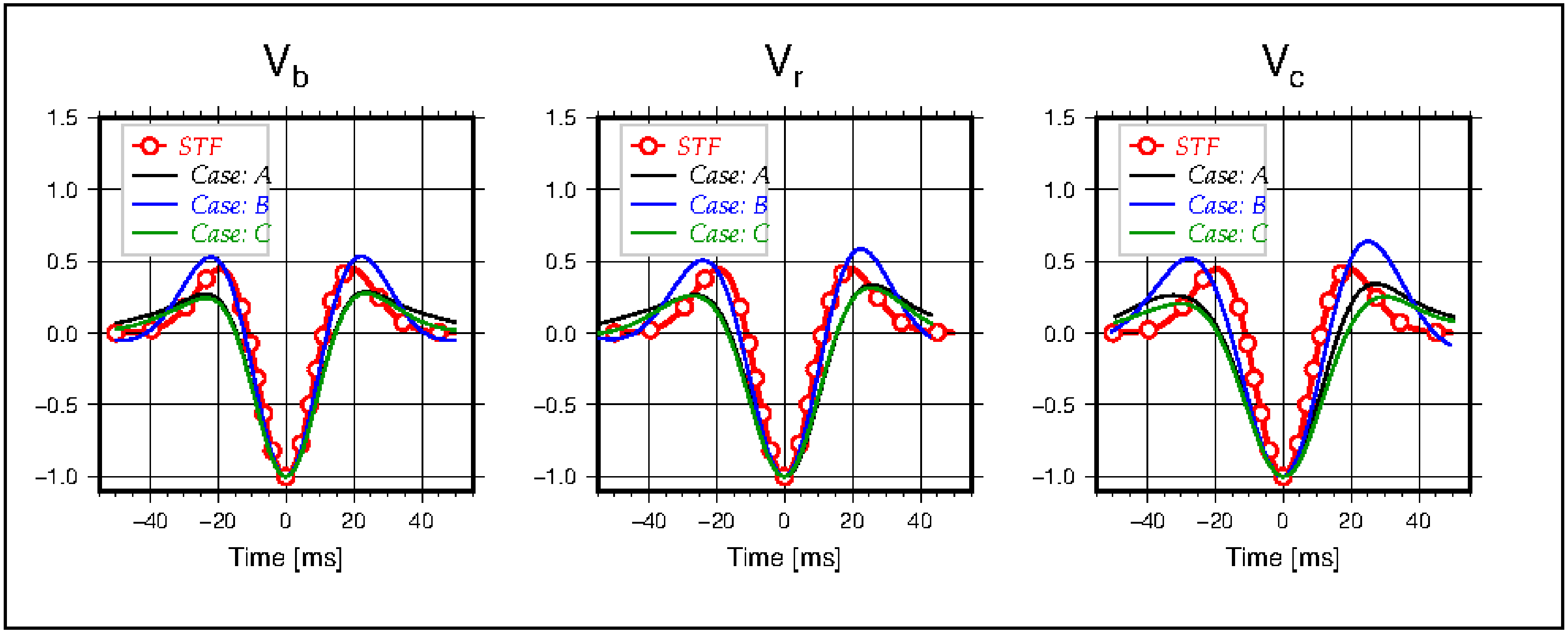




TRM and numerical simulations (K. Waskiewicz)



TRM and numerical simulations (K. Waskiewicz)

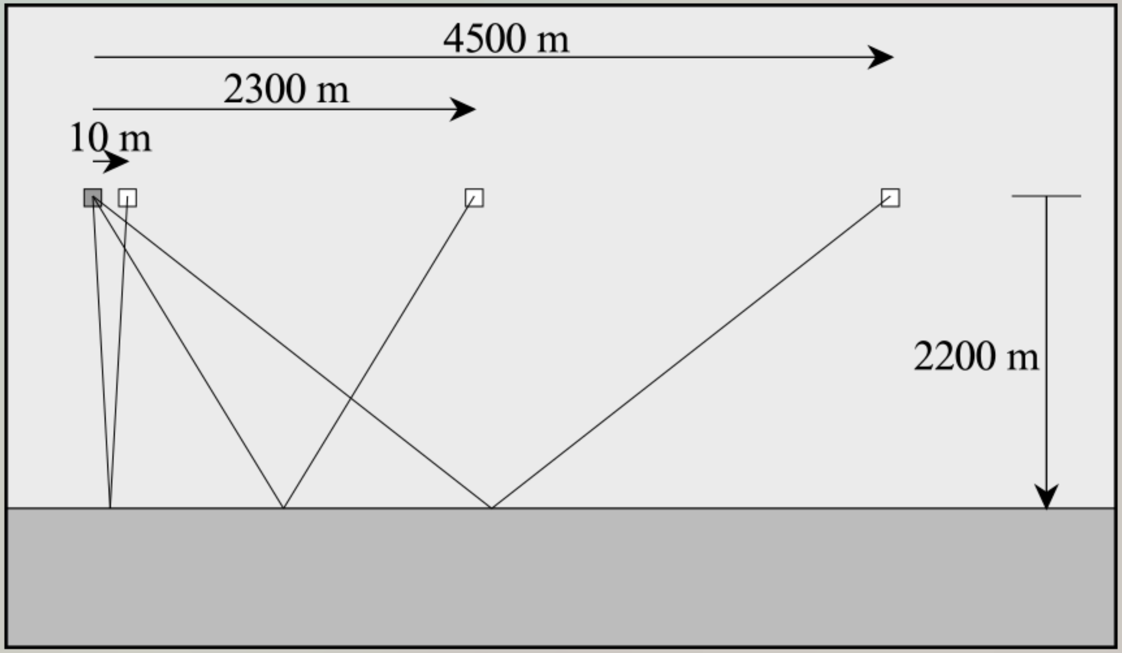


Advanced issue (C): Non-parameter-estimation inversion

Experiment: seismic waves reflection from a flat interface

Data: Amplitudes of reflected waves at 3 distances

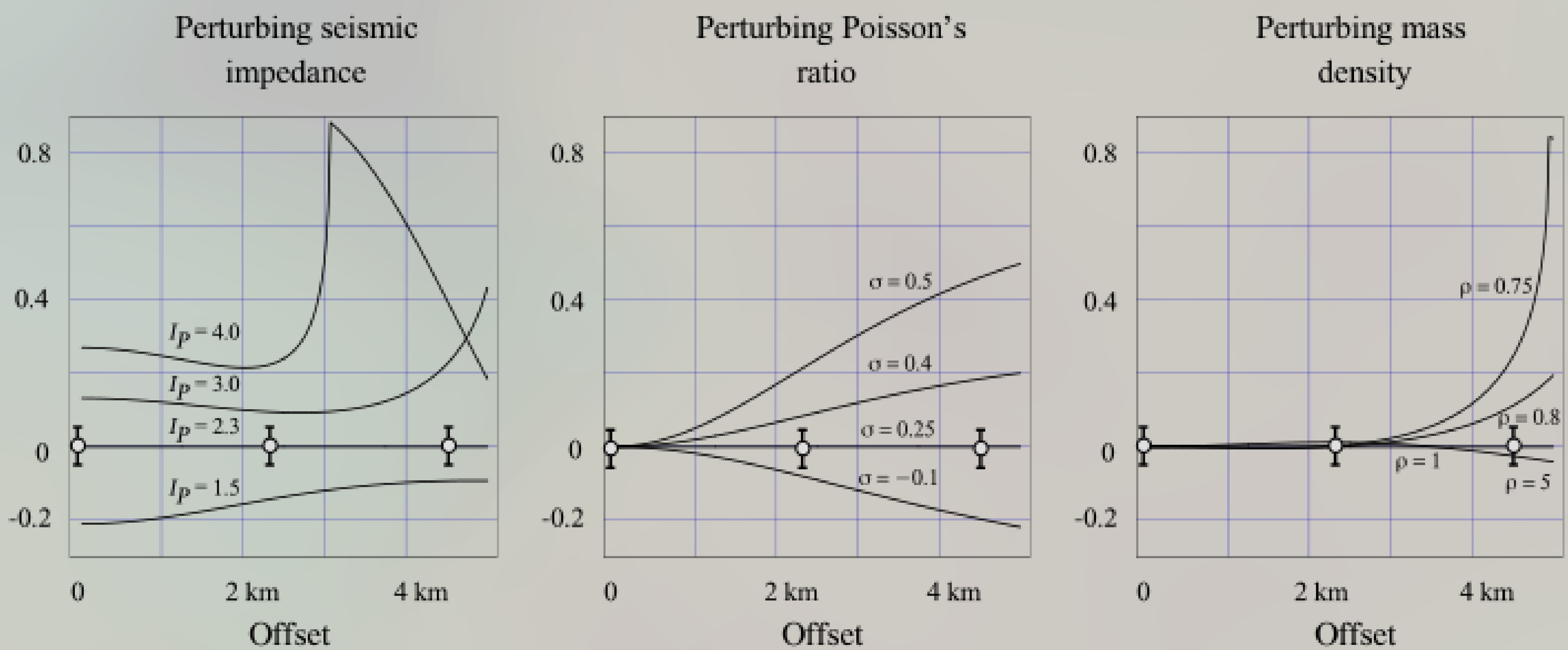
Goal: Find best parametrization of lawer layer



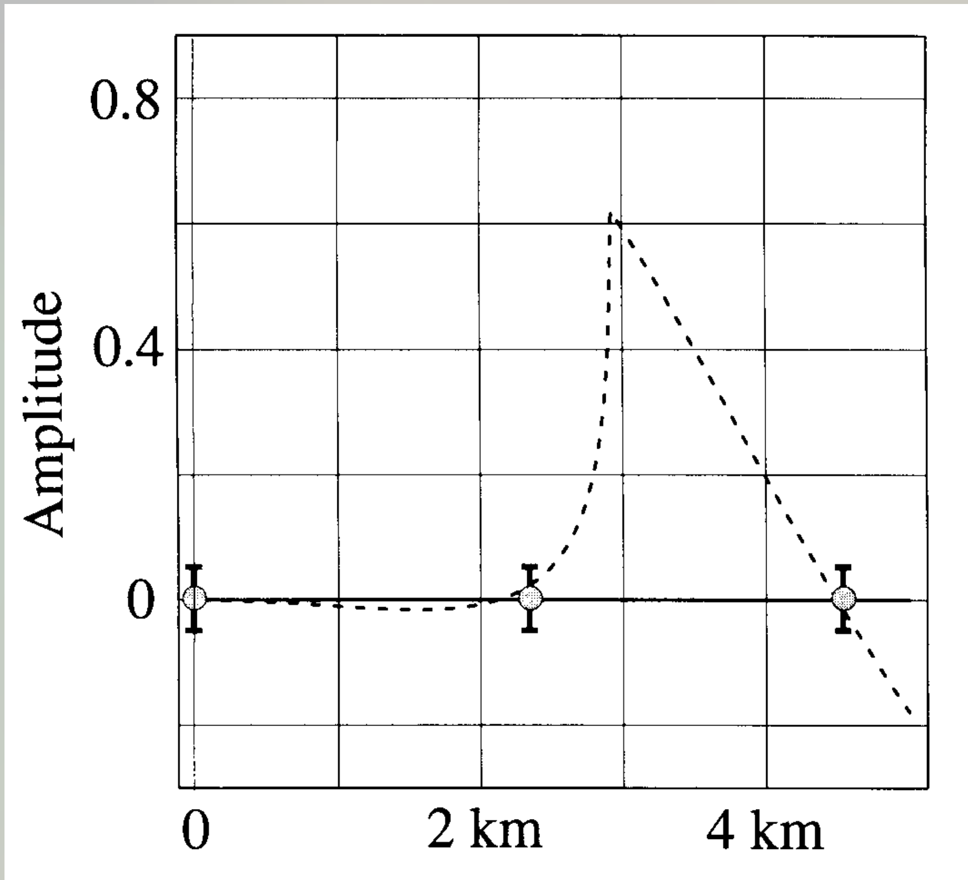
- ✦ (Vp, Vs, ρ)
- ✦ (Ip, Is, ρ)
- ✦ (Ip, S, ρ)

$$Ip = \rho Vp$$
$$Is = \rho Vs$$

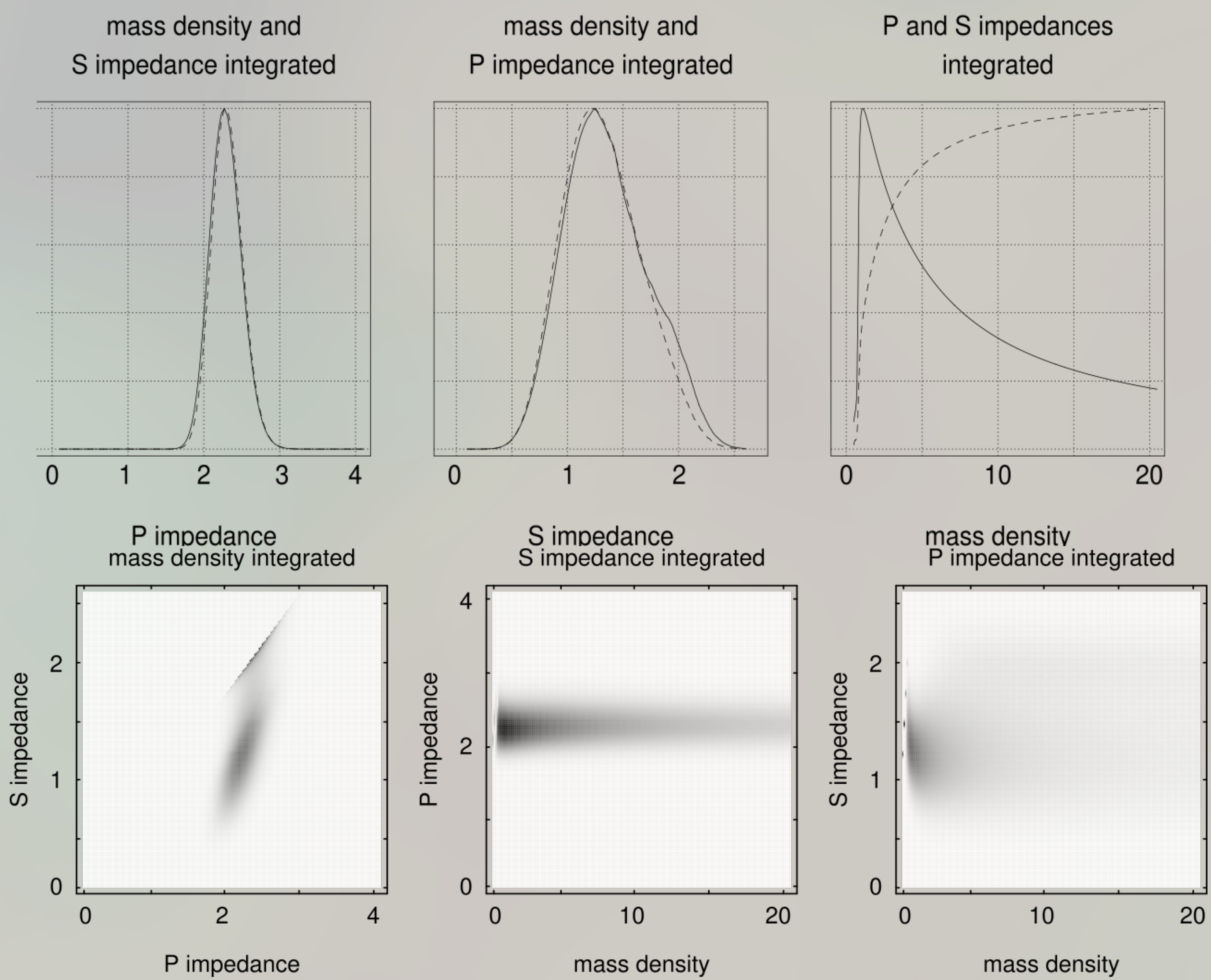
Theoretical predictions



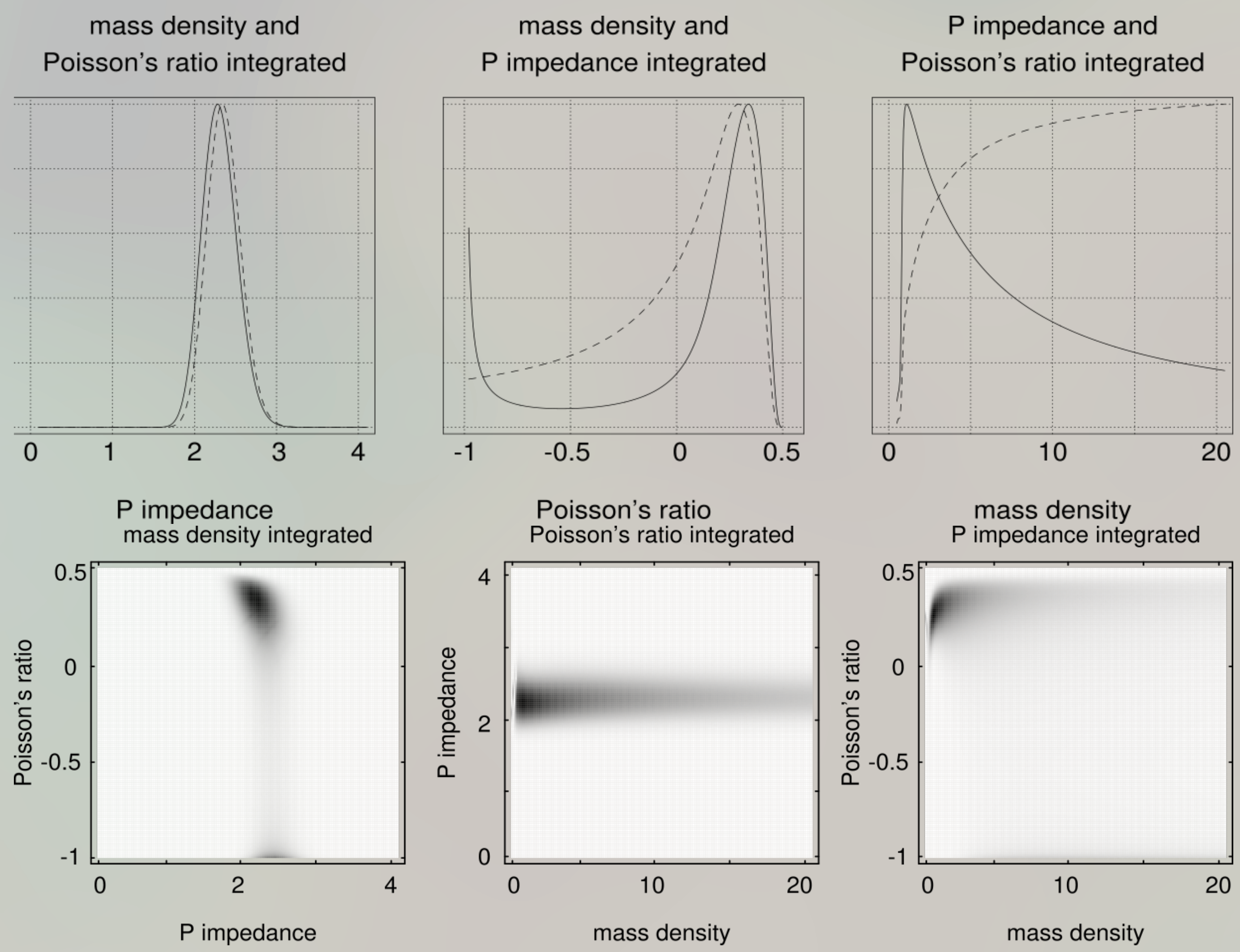
Non-uniqueness



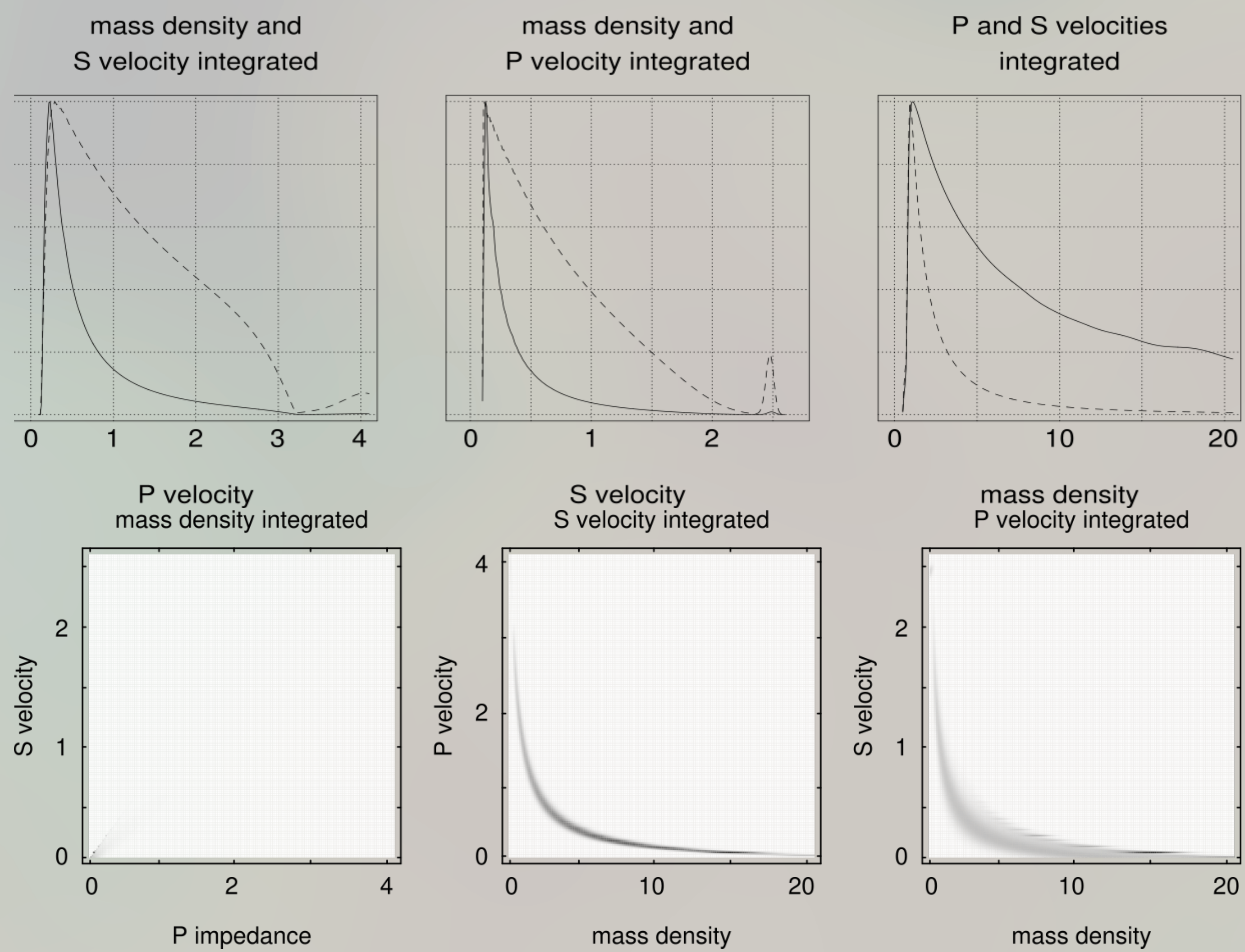
Parametrization: Ip, Is, M



Parametrization: Ip, S, M



Parametrization: Vp, Vs, M



Best parametrization

- mass is not resolved at all
- p(mass) depends significantly off *a priori* term
- Vp and Vs are strongly correlated wwith mass
- Ip and Is and correlated
- Ip and S are uncorrelated and independent of M

Parametrization (Ip, S,M) is optimum for tis experiment

Non-informative distribution

$$\sigma_m(m) = f(m) \cdot L(m, d^{obs})$$

$$L(\mathbf{m}, \mathbf{d}^{obs}) = \int_D p(\mathbf{d}, \mathbf{d}^{obs}) \frac{q(\mathbf{m}, \mathbf{d})}{\mu(\mathbf{m}, \mathbf{d})} \mathrm{d}\mathbf{d}$$

Joining information according to Tarantola

Two distributions describing different pieces of information about the same object

1. $p(x)$
2. $q(x)$

$$\mathbf{p} \wedge \mathbf{q}(\mathbf{x}) = \frac{\mathbf{p}(\mathbf{x})\,\mathbf{q}(\mathbf{x})}{\mu(\mathbf{x})}$$

$\mu(x)$ - non-informative probability

Non-informative distribution

$$q(x) \;=\; \mu(x)$$

$$\mathbf{p} \wedge \mathbf{q}(\mathbf{x}) = \frac{\mu(\mathbf{x})\,\mathbf{q}(\mathbf{x})}{\mu(\mathbf{x})} \;=\; \mathbf{p}(\mathbf{x})$$

$\mu(\cdot)$ - states of no information at all

Insufficient reasoning postulat (Laplace)

If x is a real unconstrained parameter then the state of lack of information on it is represented by a probability proportional to the volume (length) in $\mathcal{R}$.

$$p(A) = \int_A \mu(x) dx = vol(A),$$

where A stands for the volume (length) of the set A, ($A \subset \mathcal{R}$.

In a Cartesian coordinate system

$$\mu(x) = const.$$

Bounded parameters

Let x be a parameter whose values belong to a segment $\Omega = [a, b]$.

Next, let us introduce a new parameter x'

$$x' = g(x)$$

$g()$ - differentiable function

If $x' \in \mathcal{R}$ than $\mu(x') = const$

Using transformation properties of probability density function

$$\mu(x) \ = \ \left|\frac{dg}{dx}\right|$$

Example - positive defined parameters

Let f be a positive parameter (e.g. frequency) $f \in \mathcal{R}^+$ Then, one can take transformation function ($\mathcal{R}^+ \Rightarrow \mathcal{R}$)

$$g(f) = \ln(f)$$

What immediatelly leads to the noninformative pdf for f

$$\mu(f) = \frac{1}{f}$$

Example - positive defined parameters

However, we can prefer other parameter: period

$$T = 1/f$$

which represents exactly the same information like f.
Then, using transformation rule for parameter change

$$f \rightarrow T$$

We immediatelly get

$$\mu(T) \ = \ \mu(f(T)) \left| \frac{df}{dT} \right| \ = \ \frac{1}{T}$$

Exploring *a posteriori* probability

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

Efficient methods of calculation multi-dimensional integrals needed !

Marginal *a posteriori* distribution/Sampling

- 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

Sampling *a posteriori* distribution

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

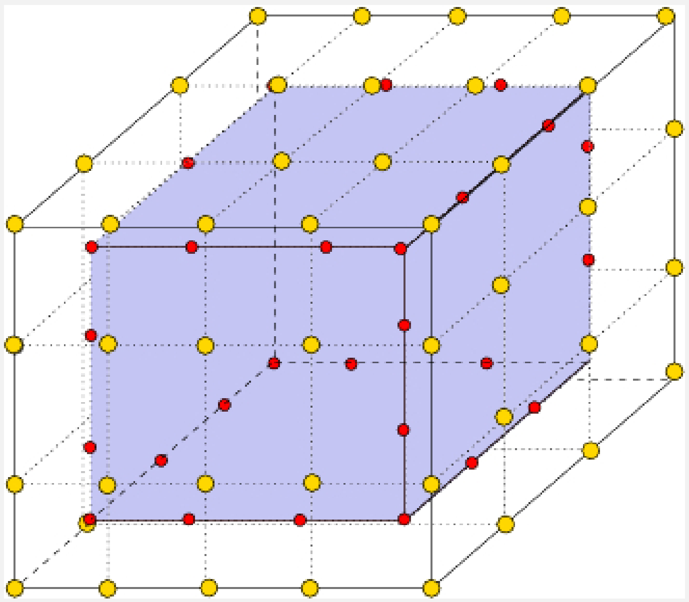
$$\sigma(\mathbf{m}) \Rightarrow \sigma_{i,j,k,\ldots} = \sigma(m_i, m_j, m_k, \ldots)$$

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

Geometric sampling in multi-dimensional spaces

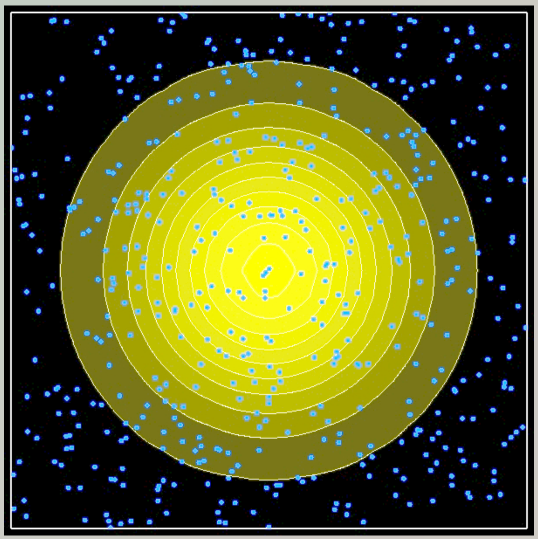
Non-uniform sampling:

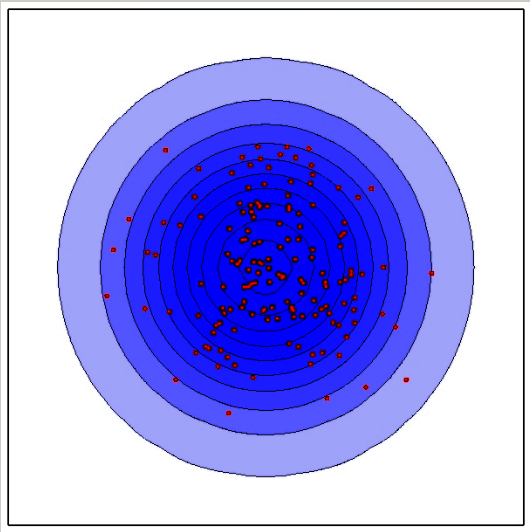
$$\frac{N_V}{N} = \left(\frac{p-2}{p}\right)^N \overset{p \gg 2}{\approx} e^{-2N/p} \xrightarrow{N} 0$$



Sampling *a posteriori* distribution

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





End