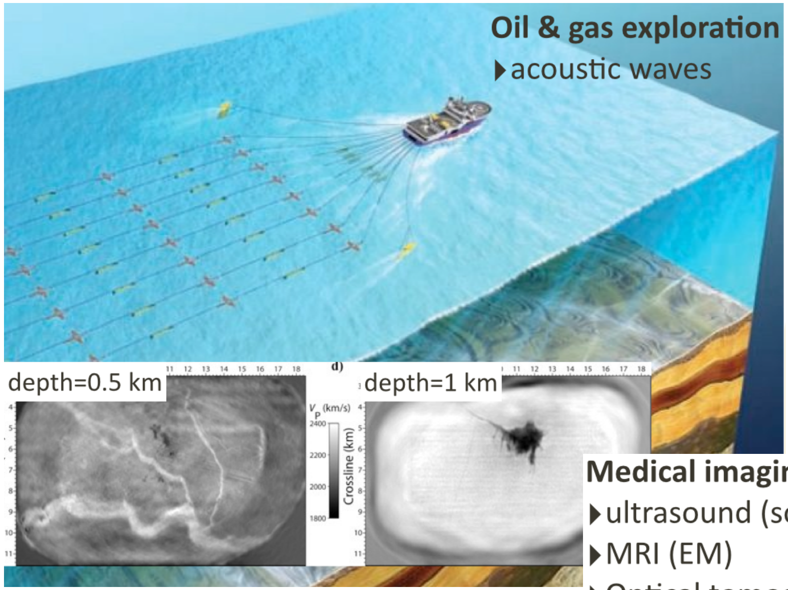
Wave-based inverse problems

reducing non-linearity and uncertainty quantification

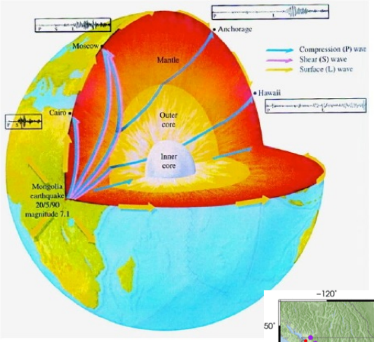
Tristan van Leeuwen

Centrum Wiskunde & Informatica
Utrecht University

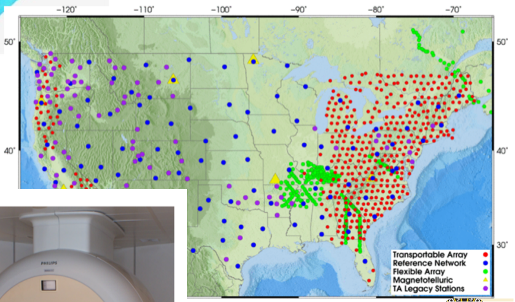
Imaging with waves



Oil & gas exploration
- acoustic waves



Seismology
- elastic waves

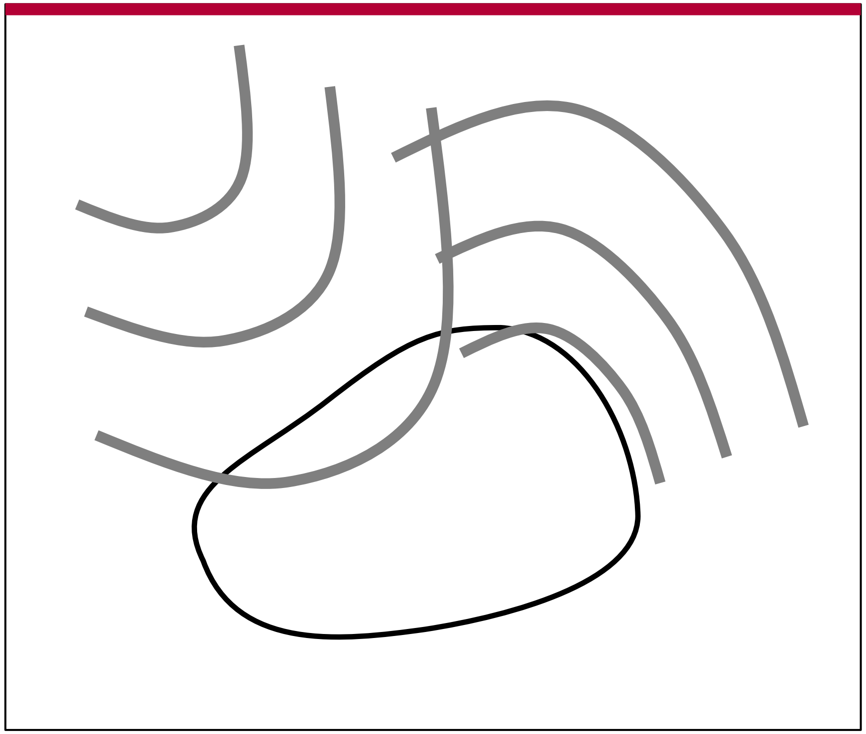


Medical imaging
- ultrasound (sound)
- MRI (EM)
- Optical tomography (light)

Overview

- The inverse problem
- Reducing non-linearity
- Uncertainty quantification
- Wrap-up

The inverse problem



PDE constrained optimisation

Classically formulated as

$$\min_{m,u} \frac{1}{2}\|Pu - d\|^2 \quad \text{s.t.} \quad A(m)u = q$$

which is often reduced to

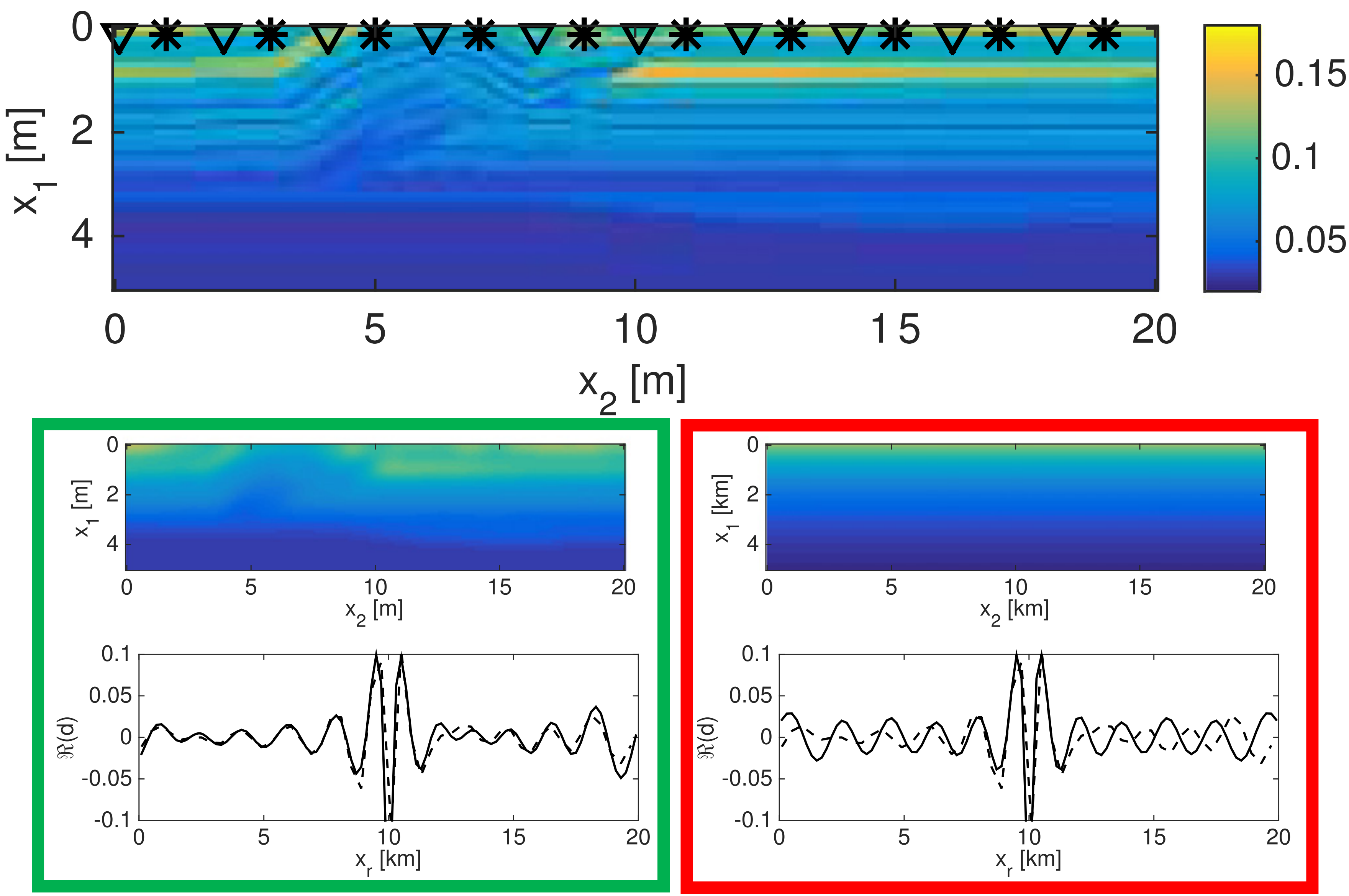
$$\min_{m} \frac{1}{2}\|PA(m)^{-1}q - d\|^2$$

Computational challenges

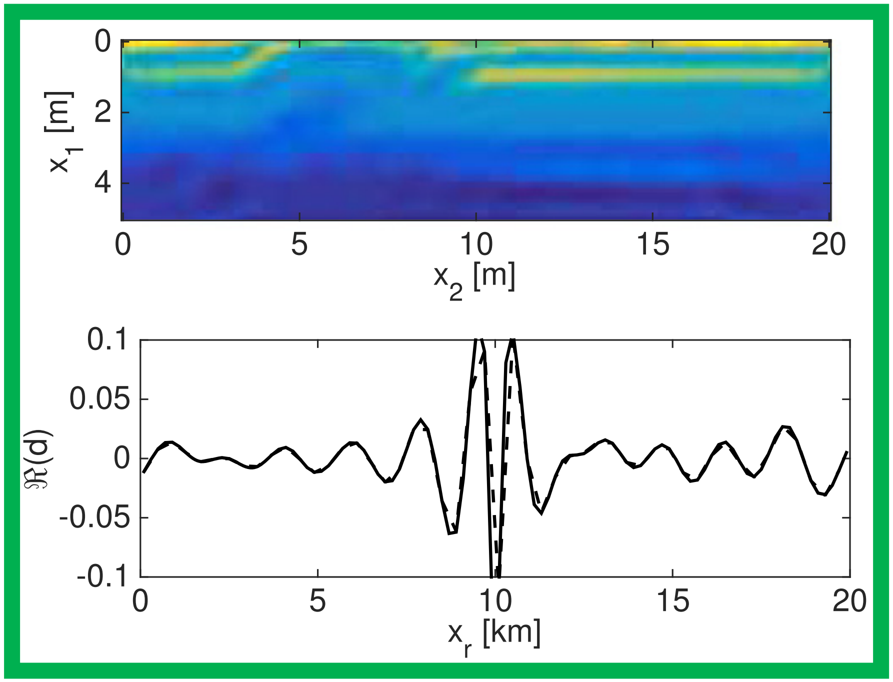
- **Non-linear, high-dimensional optimisation**
- **PDE-solves (many r.h.s., many wavelengths)**
- **Exact adjoint may not be readily available**

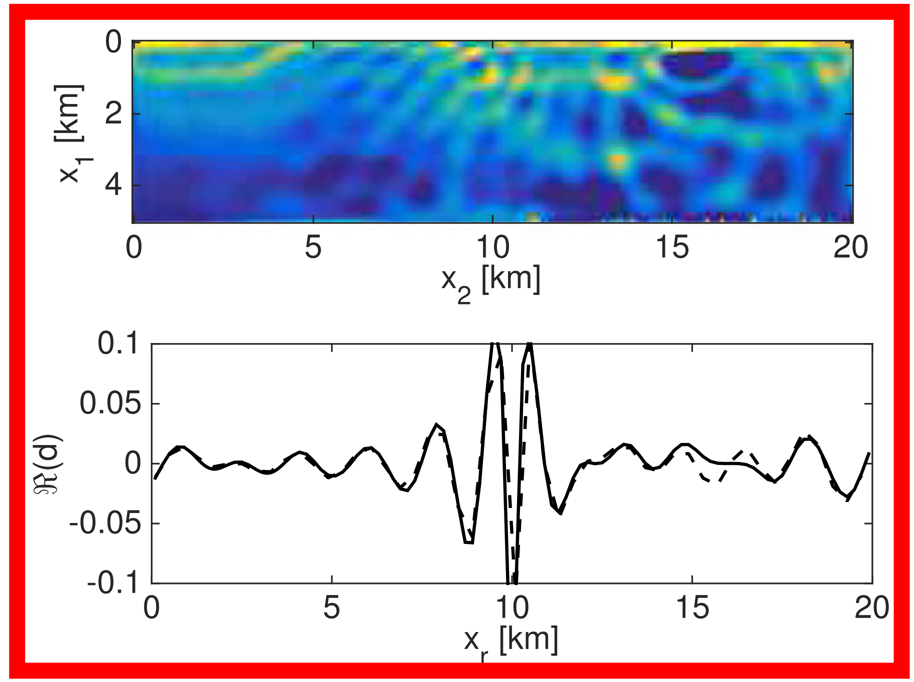
Sources of non-linearity

- Oscilatory nature of data
- Absence of low frequency data
- Limited aperture

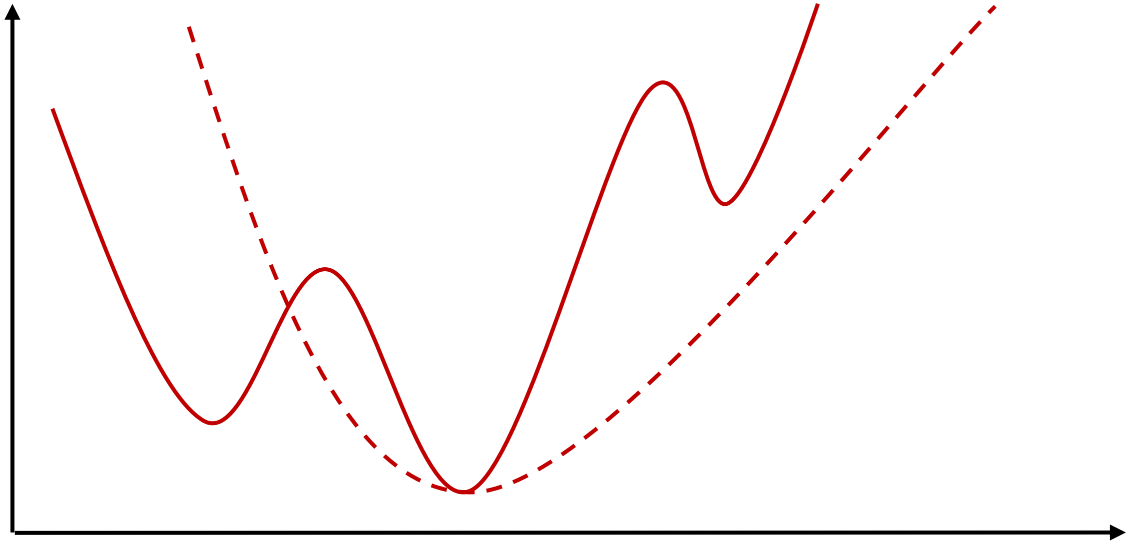


0
2
4
x1 [m]
0
5
10
15
20
x2 [m]
0.15
0.1
0.05
x1 [m]
x2 [m]
x1 [km]
x2 [km]
ℜ(d)
xr [km]
0.1
0.05
0
-0.05
-0.1


0
2
4
x1 [m]
0
5
10
15
20
x2 [m]
0.1
0.05
0
-0.05
-0.1
ℜ(d)
xr [km]


0
2
4
x1 [km]
0
5
10
15
20
x2 [km]
0.1
0.05
0
-0.05
-0.1
ℜ(d)
xr [km]

Reducing non-linearity



Joint parameter and state estimation

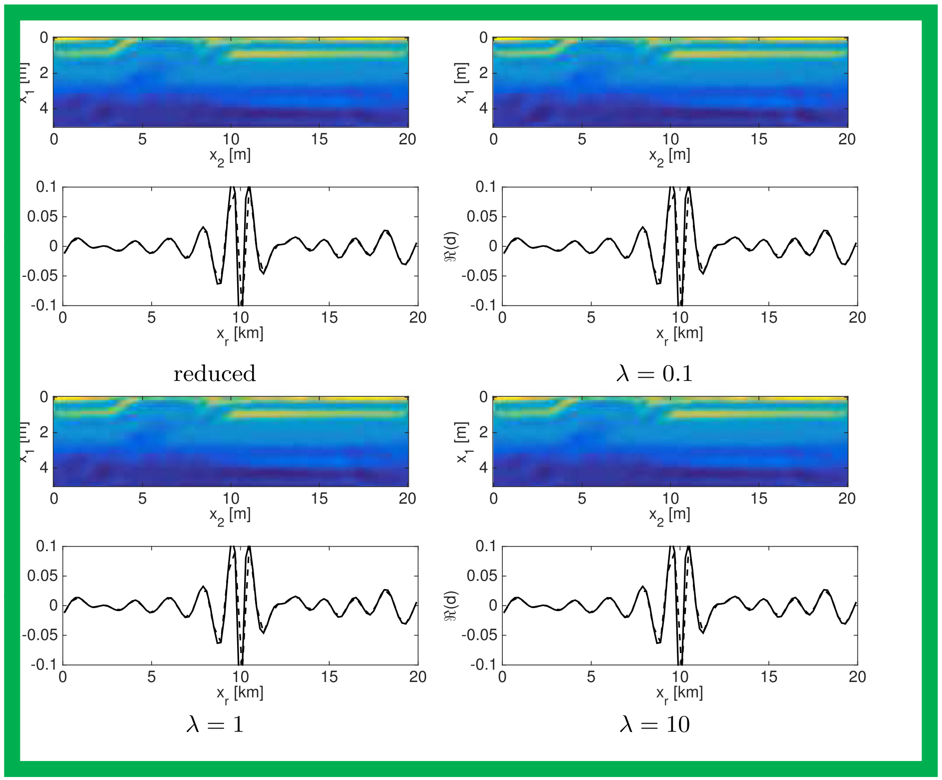
Use quadratic penalty instead

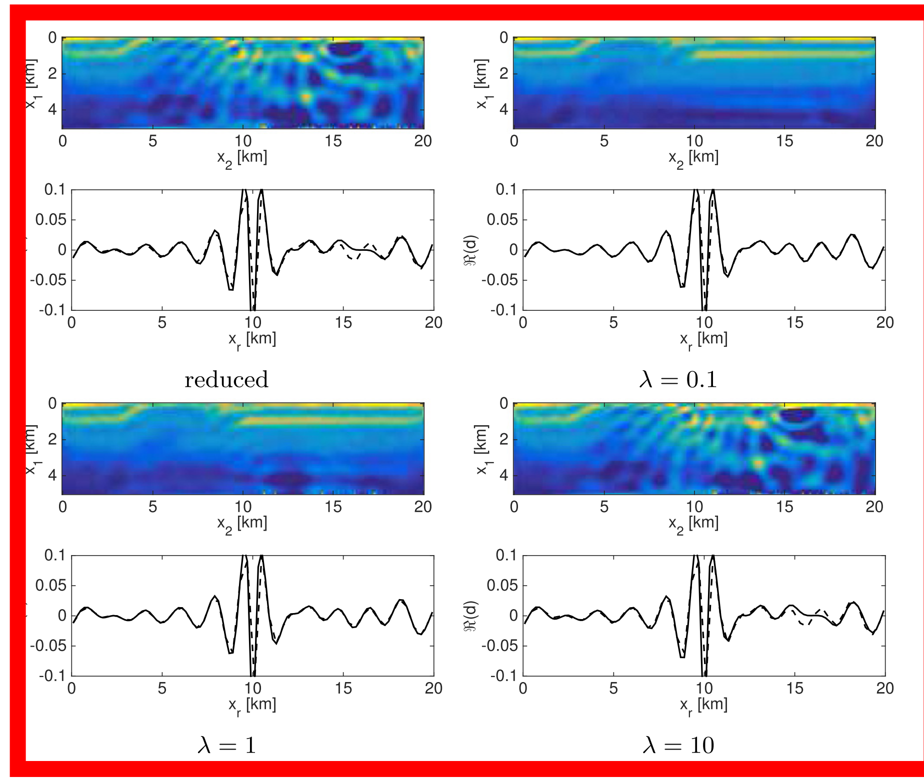
$$\min_{m,u} \frac{1}{2}\|Pu - d\|^2 + \frac{\rho}{2}\|A(m)u - q\|^2$$

and reduce via

$$u(m) = \left(\rho A(m)^T A(m) + P^T P\right)^{-1}\left(\rho A(m)^T q + P^T d\right)$$

to get a reduced penalty formulation.


reduced
λ = 0.1
λ = 1
λ = 10
x_1 [m]
x_2 [m]
ℜ(d)
x_r [km]

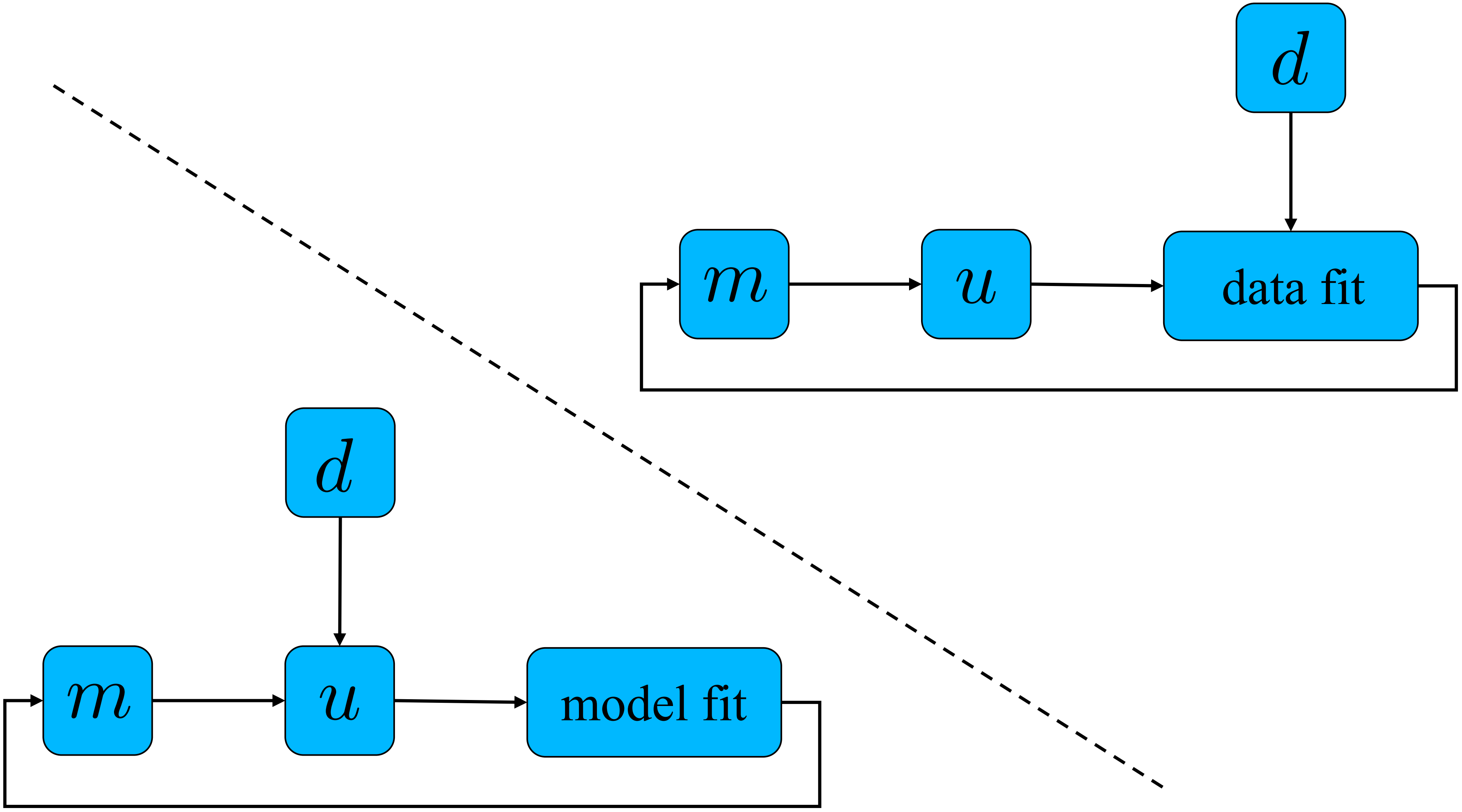

reduced
λ = 0.1
λ = 1
λ = 10
x_1 [km]
x_2 [km]
ℜ(d)
x_r [km]

- **How to choose trade-off parameter**
- **Implement data-assimilation step in a computationally efficient manner**
- **Understand limitations**

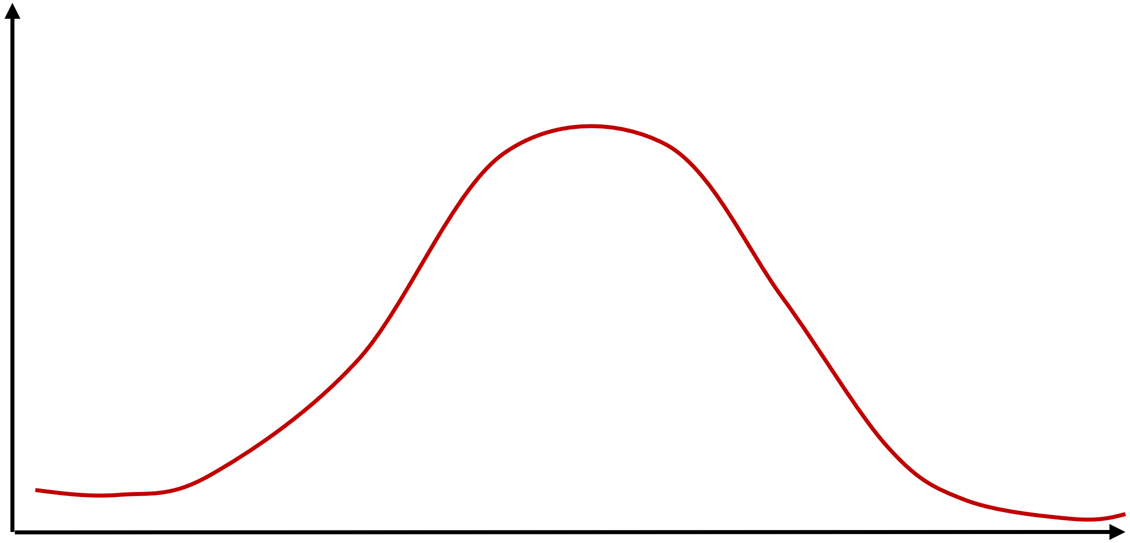
Other approaches

- **Other extensions / data fidelities**
- **Classical inverse scattering**
- **Wavefield redatuming**
- **Reduced-order-models**

Data-driven vs. model-driven



Uncertainty quantification



Bayesian approach

Formulate a prior and likelihood model, e.g.

$$\pi_{\text{prior}}(m) \propto \exp\left(-\frac{1}{2}\|m - \mu_m\|^2_{\Sigma_m}\right)$$
$$\pi_{\text{like}}(d|m) \propto \exp\left(-\frac{1}{2}\|PA^{-1}(m)q - d\|^2_{\Sigma_d}\right)$$

and generate samples from the posterior

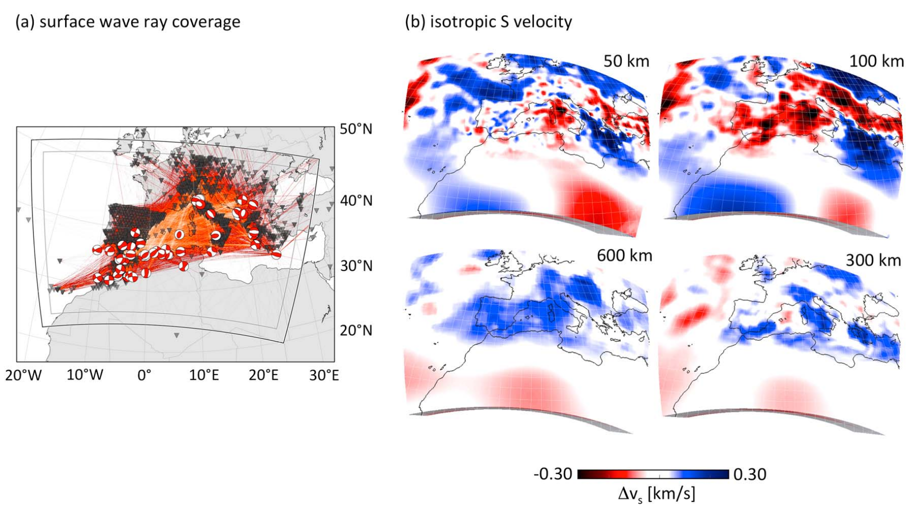
Uncertainty quantification

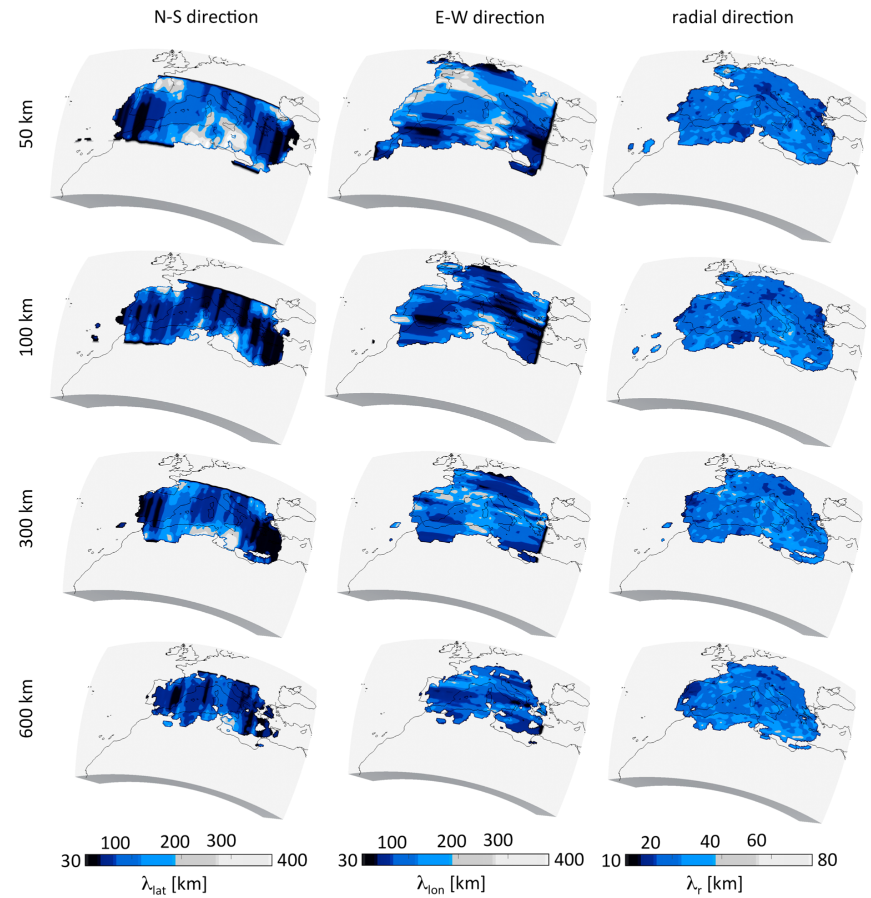
Want to go beyond computing moments:

- **Reliability of interpreted features**
- **Influence of prior of certain features**
- **Sensitivity to initial guess**

Hessian-based

Assume posterior is locally Gaussian and covariance is a blurring kernel





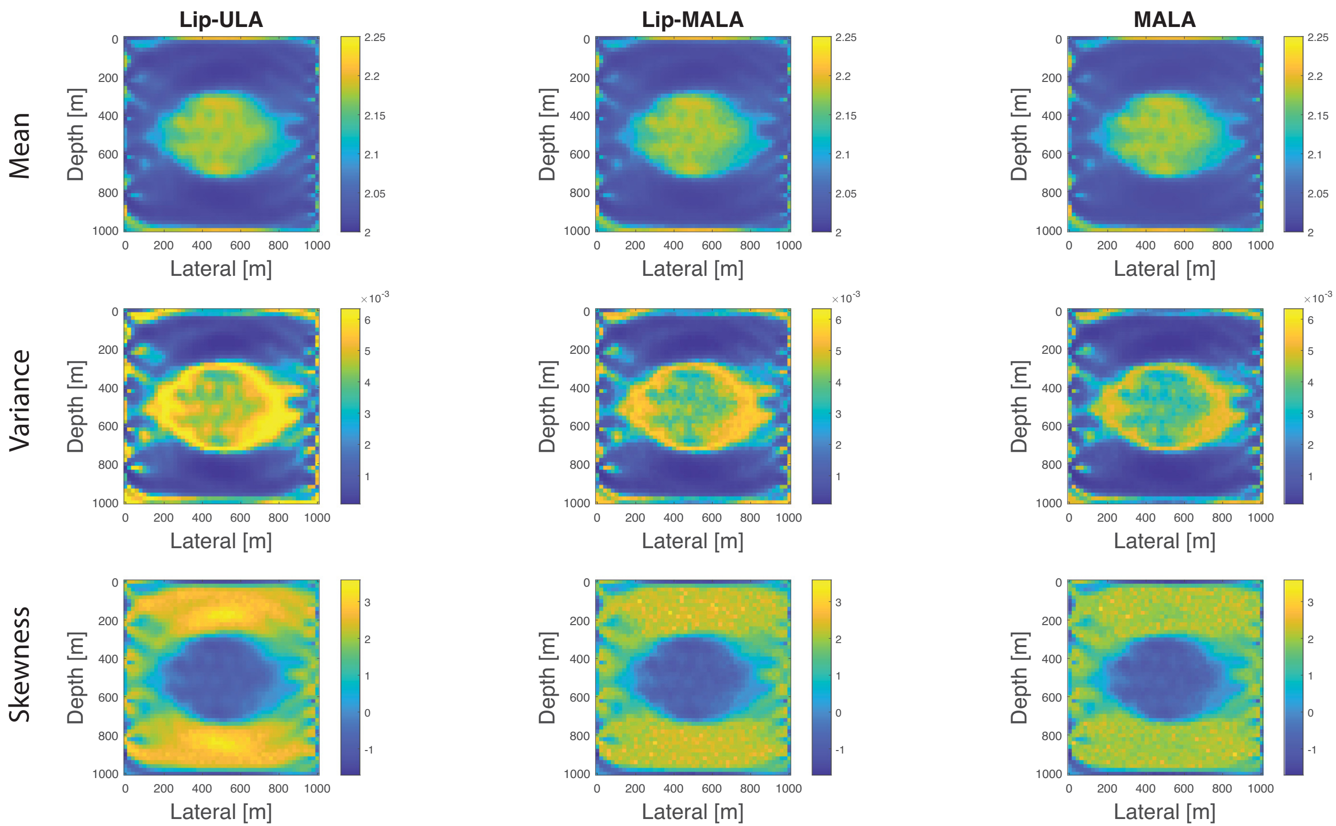
- **May be sufficient if problem is locally linear**

- **Gives only information on uncertainties of certain features**

- **Computationally feasible**

MCMC sampling

Sample from posterior using Langevin dynamics and use adaptive stepsize



- **Samples actual posterior (in theory)**
- **Sampling multi-modal could still be challenging**
- **Computationally more challenging**
- **What to do with all the samples?**

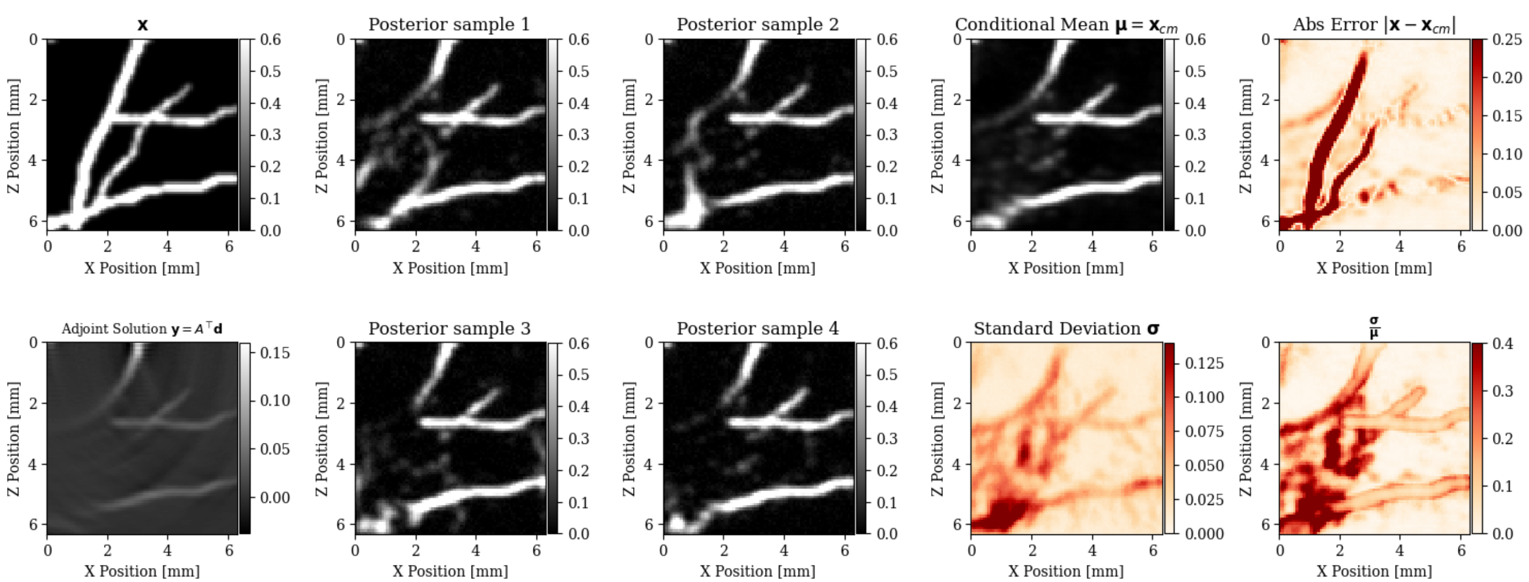
Normalising flows

Learn joint density from given samples

$$\min_{\boldsymbol{\theta}} \mathrm{KL}(p_{X,Y} \| p_{\boldsymbol{\theta}}) = \mathbb{E}_{\mathbf{x},\mathbf{y} \sim p_{X,Y}(\mathbf{x},\mathbf{y})} \frac{1}{2} \left\| f_{\boldsymbol{\theta}}(\mathbf{x},\mathbf{y}) \right\|^2 - \log \left| \det \mathrm{J}_{f_{\boldsymbol{\theta}}}(\mathbf{x},\mathbf{y}) \right|.$$

with

$$f_{\boldsymbol{\theta}}(\mathbf{x},\mathbf{y}) = (f_{\boldsymbol{\theta}}^{z_x}(\mathbf{x},\mathbf{y}), f_{\boldsymbol{\theta}}^{z_y}(\mathbf{y})).$$



- **Choosing appriopriate archtecture is not easy**
- **Requires many training samples**
- **Tied to particular acquisition and sampling**
- **What to do with all the samples?**

Wrap-up

- **Many computational challenges in solving non-linear problems and sampling in high-dimensions**
- **Many practical challenges in choosing parameters, architectures, etc.**
- **More fundamental problems with UQ; which distribution should we sample from and what to do with the samples?**
- **How do we merge data-driven approaches and UQ?**

Thanks!

References

Izzatullah, M., van Leeuwen, T., & Peter, D. (2021). Bayesian seismic inversion: a fast sampling Langevin dynamics Markov chain Monte Carlo method. *Geophysical Journal International*, *227*(3), 1523–1553.

van Leeuwen, T., & Herrmann, F. J. (2016). A penalty method for PDE-constrained optimization in inverse problems. *Inverse Problems*, *32*(1), 015007.

van Leeuwen, T., & van Leeuwen, T. (2019). A note on extended full waveform inversion. *ArXiv Preprint ArXiv:1904.00363*, 1–10.

van Leeuwen, T., Leeuwen, P. J. van, & Zhuk, S. (2021). Data-driven modeling for wave-propagation. In *Numerical Mathematics and Advanced Applications ENUMATH 2019* (pp. 683–691). Springer, Cham.

Orozco, R., Siahkoohi, A., Rizzuti, G., van Leeuwen, T., & Herrmann, F. J. (2021). Photoacoustic imaging with conditional priors from normalizing flows. *NeurIPS 2021 Workshop on Deep Learning and Inverse Problems*.

Fichtner, A., & van Leeuwen, T. (2015). Resolution analysis by random probing. *Journal of Geophysical Research: Solid Earth*, *120*(8), 5549–5573.

Tataris, A., & van Leeuwen, T. (2022). A Regularised Total Least Squares Approach for 1D Inverse Scattering. *Mathematics*, *10*(2), 216.