Inverse Problems in Geophysics
Part 7: Regularization

2. MGPY+MGIN

Thomas Günther

thomas.guenther@geophysik.tu-freiberg.de

TU BERGAKADEMIE FREIBERG

TUBAF

Die Ressourcenuniversität. Seit 1765.

Recap

- SVD provides a general tool, BUT:
  - can amplify noise for ill-conditioned problems (SV spectrum)
  - truncated SVD (limiting p) can suppress this
- explicit regularization to make solution unique
  - different strategies: smoothness, minimum norm
- choice of regularization strength (λ, p) is vital
  - watch model and misfit plots, discrepancy principle

Regularization scheme

Splitting into original matrix & data and constraints

$$\tilde{\mathbf{G}} = \begin{bmatrix} \mathbf{G} \\ \mathbf{C} \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{d}} = \begin{bmatrix} \mathbf{d} \\ \mathbf{c} \end{bmatrix}$$

Damping (minimum norm)

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Smoothness constraints

$$\mathbf{C} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

Weighting data vs. constraints

$$\Phi = \|\tilde{\mathbf{G}}\mathbf{m} - \tilde{\mathbf{d}}\|_2^2 = \|\mathbf{G}\mathbf{m} - \mathbf{d}\|^2 + \|\mathbf{C}\mathbf{m} - \mathbf{c}\|^2 = \rightarrow \min$$

$\mathbf{d}$ and $\mathbf{c}$ have different magnitudes & units, data maybe too weak or too strong $\Rightarrow$ weighting by regularization parameter λ:

$$\Phi = \|\mathbf{G}\mathbf{m} - \mathbf{d}\|^2 + \lambda\|\mathbf{C}\mathbf{m} - \mathbf{c}\|^2 = \Phi_d + \lambda\Phi_m \rightarrow \min$$

λ..regularization strength, Φ_d/Φ_m..data/model objective function

$$\Rightarrow \mathbf{m} = (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{C}^T\mathbf{C})^{-1}(\mathbf{G}^T\mathbf{d} + \lambda\mathbf{C}^T\mathbf{c})$$

Choice of regularization strength

Always have a look at your data fit and model plausibility.

- use different values and look at models (and misfit)
- try to determine the corner of the L-curve (maximum curvature)
- start large λ, decrease & stop when data misfit show no systematics

Discrepancy principle

Choose the highest λ value that is able to fit the data (χ^2=1)!

Regularization

- truncated singular value decomposition (TSVD)
- minimum norm: damped least squares
- minimum roughness: smoothness-constrained minimization

Damped least squares and SVD

$$\mathbf{m} = (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{G}^T\mathbf{d}$$

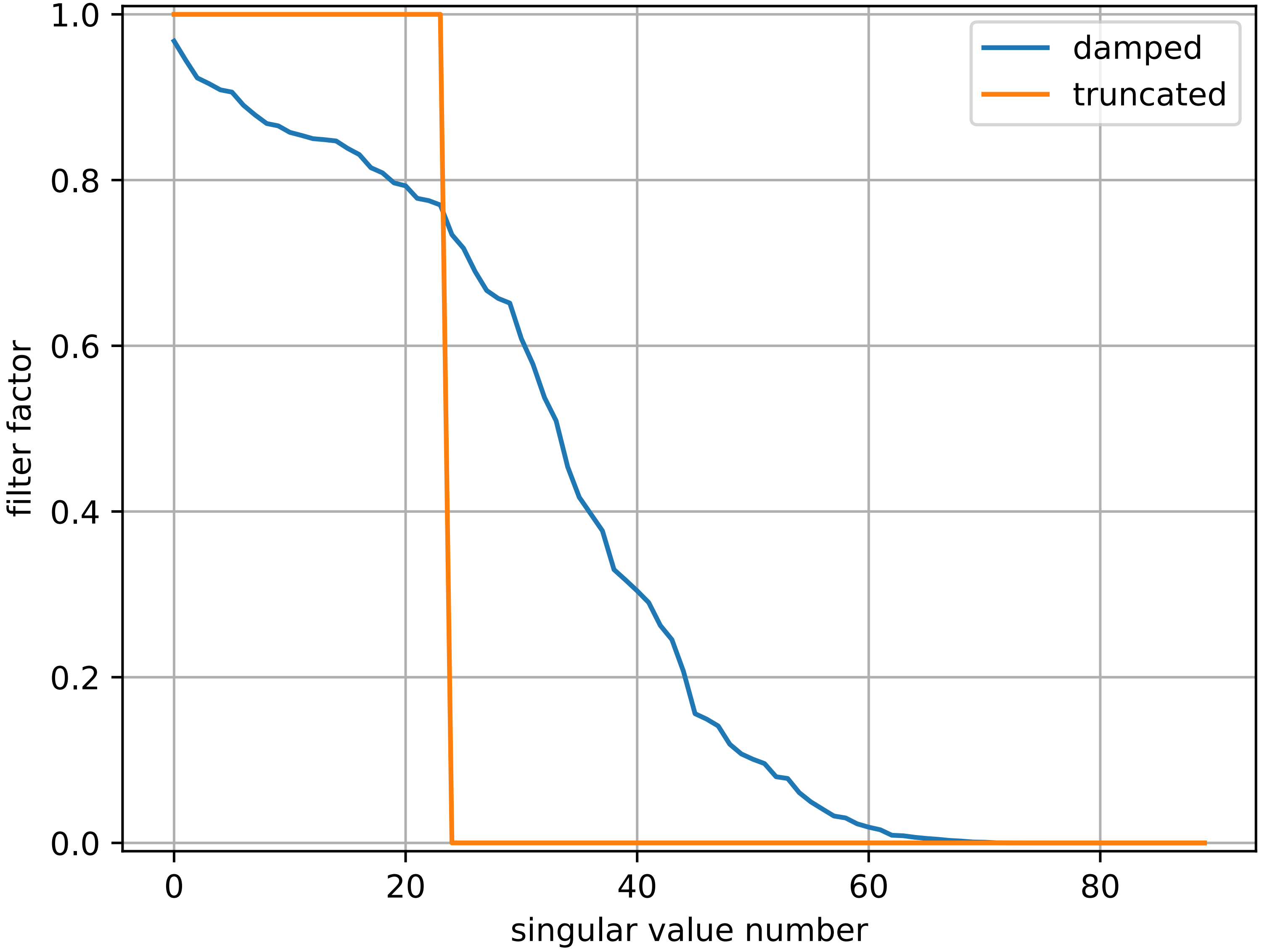
$$\mathbf{m} = (\mathbf{V\Sigma U}^T\mathbf{U\Sigma V}^T + \lambda\mathbf{I})^{-1}\mathbf{V\Sigma U}^T\mathbf{d}$$

$$\mathbf{m} = (\mathbf{V}\mathrm{diag}(s_i^2 + \lambda)\mathbf{V}^T)^{-1}\mathbf{V\Sigma U}^T\mathbf{d}$$

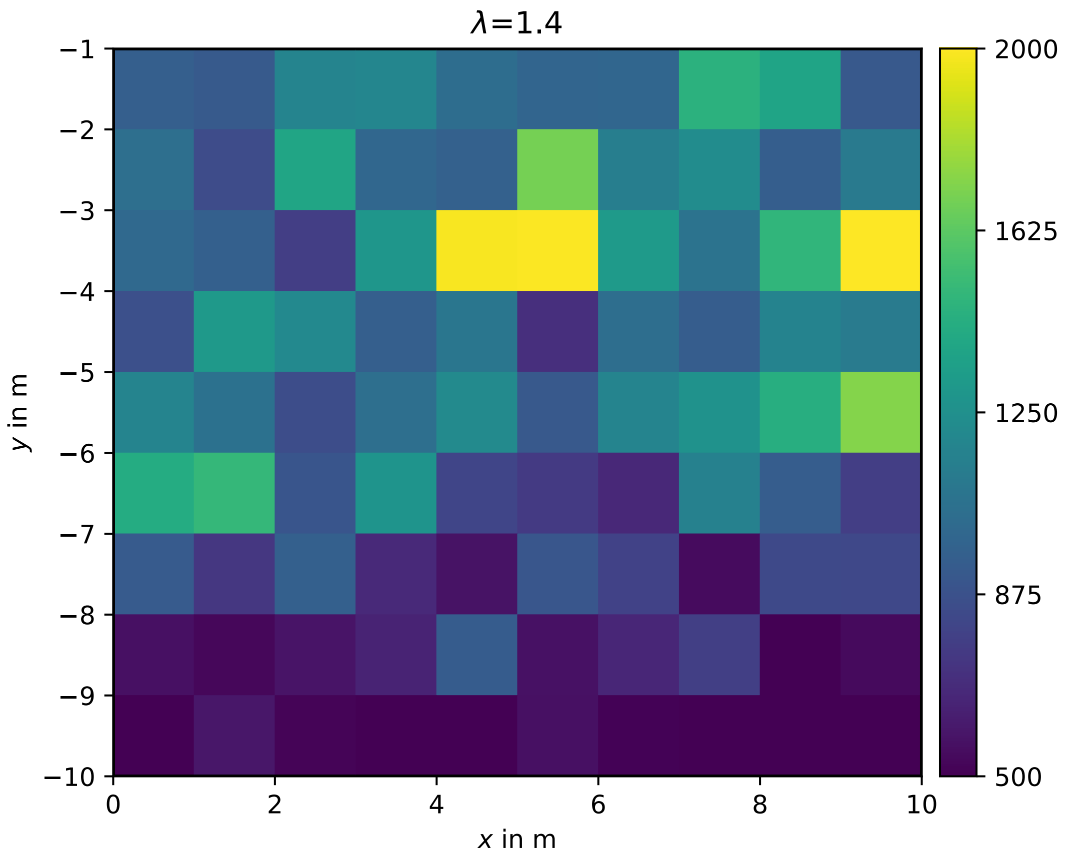
$$\mathbf{m} = \sum_i^r \frac{s_i}{s_i^2 + \lambda}\mathbf{u}_i^T\mathbf{d} \cdot \mathbf{v}_i = \sum_i^r \frac{s_i^2}{s_i^2 + \lambda}\frac{\mathbf{u}_i^T\mathbf{d}}{s_i}\mathbf{v}_i = \sum_i^r f_i \frac{\mathbf{u}_i^T\mathbf{d}}{s_i}\mathbf{v}_i$$

Small singular values are damped by filter factors $f_i = s_i^2/(s_i^2 + \lambda)$

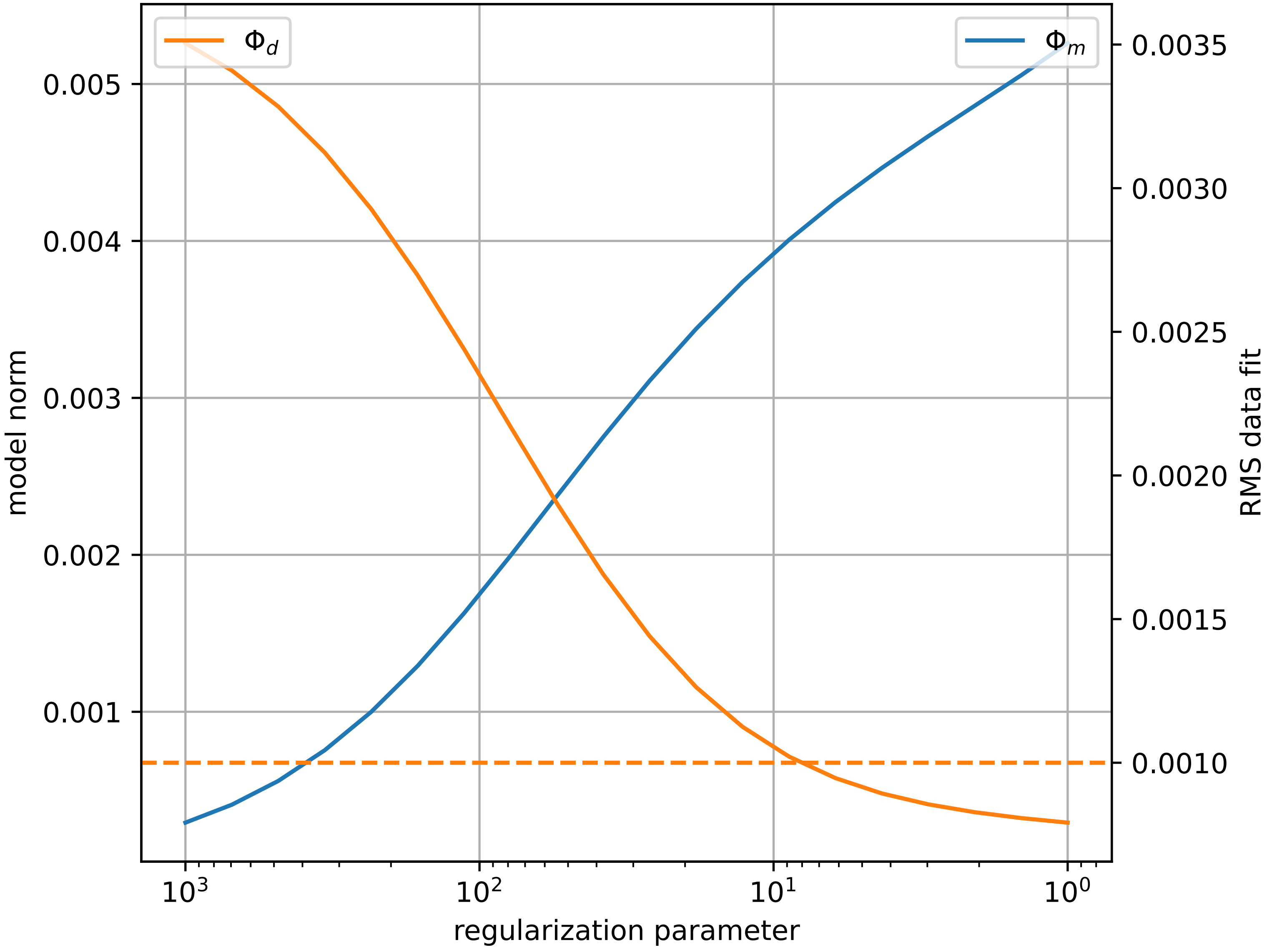
Filter factors



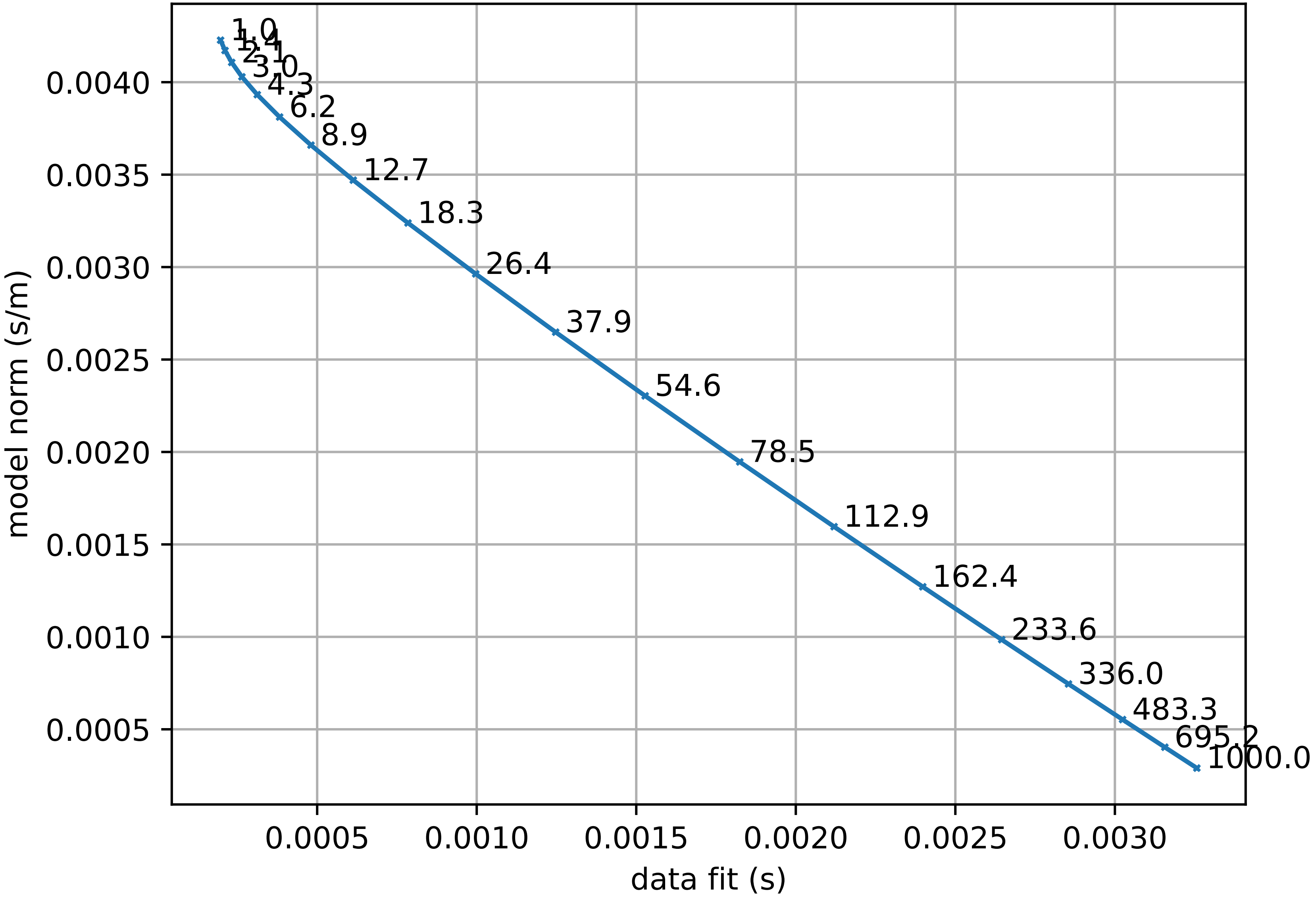
Inversion with damping



Choosing λ: Data and model norm



The L-curve



model norm (s/m)
0.0040
0.0035
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0005
0.0010
0.0015
0.0020
0.0025
0.0030
data fit (s)
1.0
1.4
2.1
3.0
4.3
6.2
8.9
12.7
18.3
26.4
37.9
54.6
78.5
112.9
162.4
233.6
336.0
483.3
695.2
1000.0

Resolution of regularized inverse problems

For $c = 0$ we have $\mathbf{G}^{\dagger} = (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{C}^T\mathbf{C})^{-1}\mathbf{G}^T$

$$\Rightarrow \mathbf{R}^M = \mathbf{G}^{\dagger}\mathbf{G} = (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{C}^T\mathbf{C})^{-1}\mathbf{G}^T\mathbf{G}$$

approaches $\mathbf{I}$ for $\lambda \to 0$ and deviates if λ grows

Resolution for damped normal equations

$$\mathbf{R}^M = \mathbf{V} \cdot \mathrm{diag}(f_i) \cdot \mathbf{V}^T$$

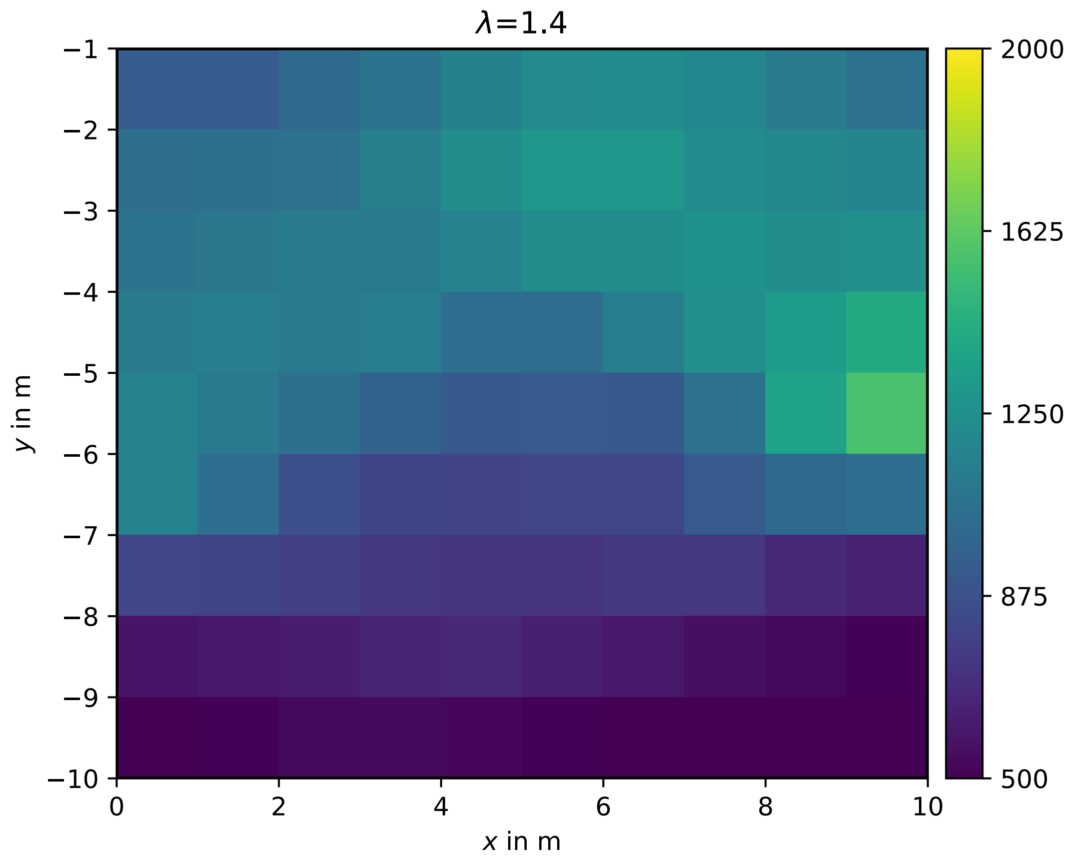
$$\mathbf{R}^D = \mathbf{U} \cdot \mathrm{diag}(f_i) \cdot \mathbf{U}^T$$

$\Rightarrow$ like for TSVD with $f_i = [1, \ldots, 1, 0, \ldots, 0]^T$

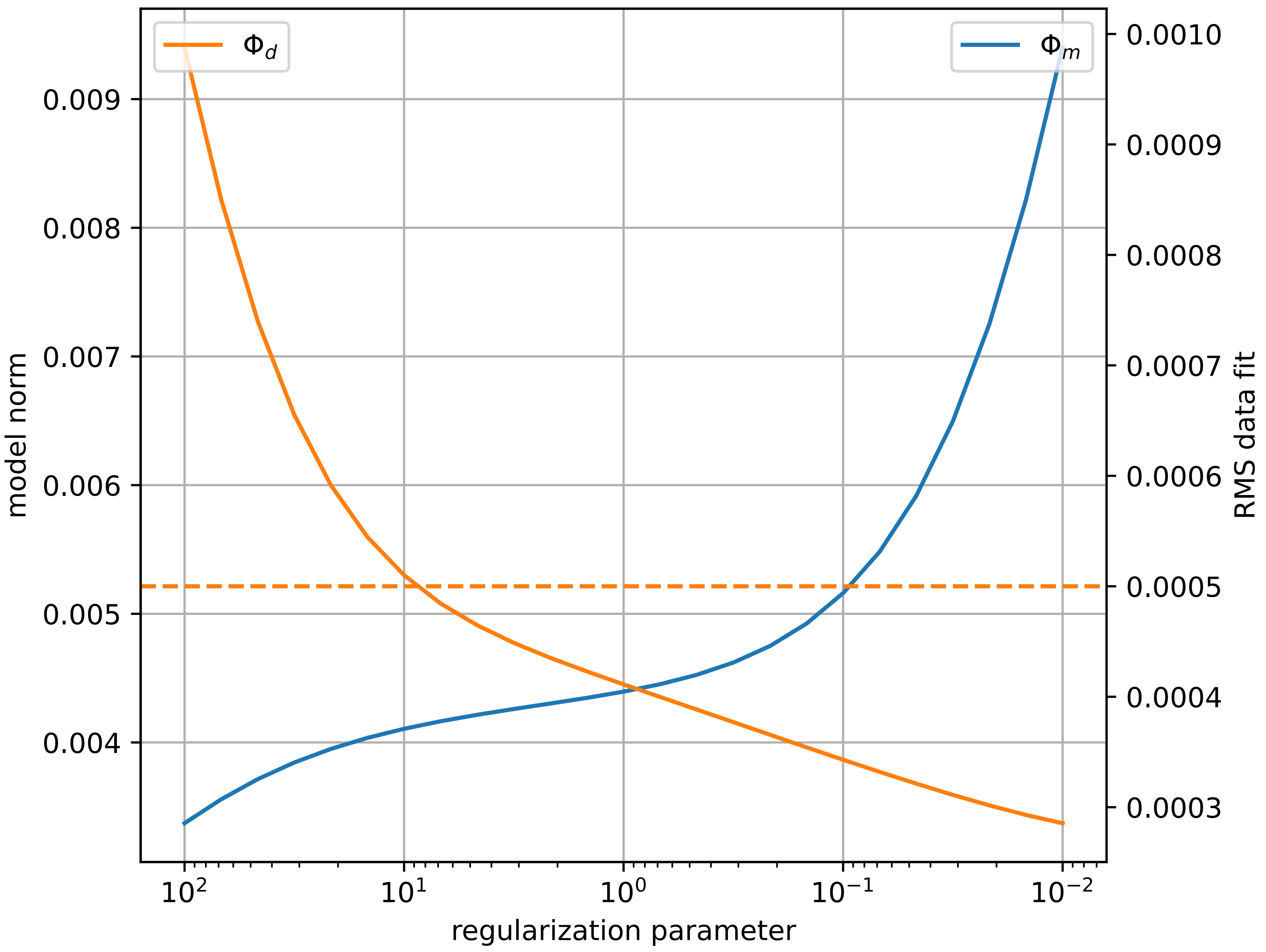
$$\mathbf{R}^M = \mathbf{V}_p \mathbf{V}_p^T \quad \text{and} \quad \mathbf{R}^D = \mathbf{U}_p \mathbf{U}_p^T$$

$$\mathbf{R}^M - \mathbf{I} = \mathbf{V}_p \mathbf{V}_p^T - \mathbf{V}\mathbf{V}^T = -\mathbf{V}_0 \mathbf{V}_0^T$$

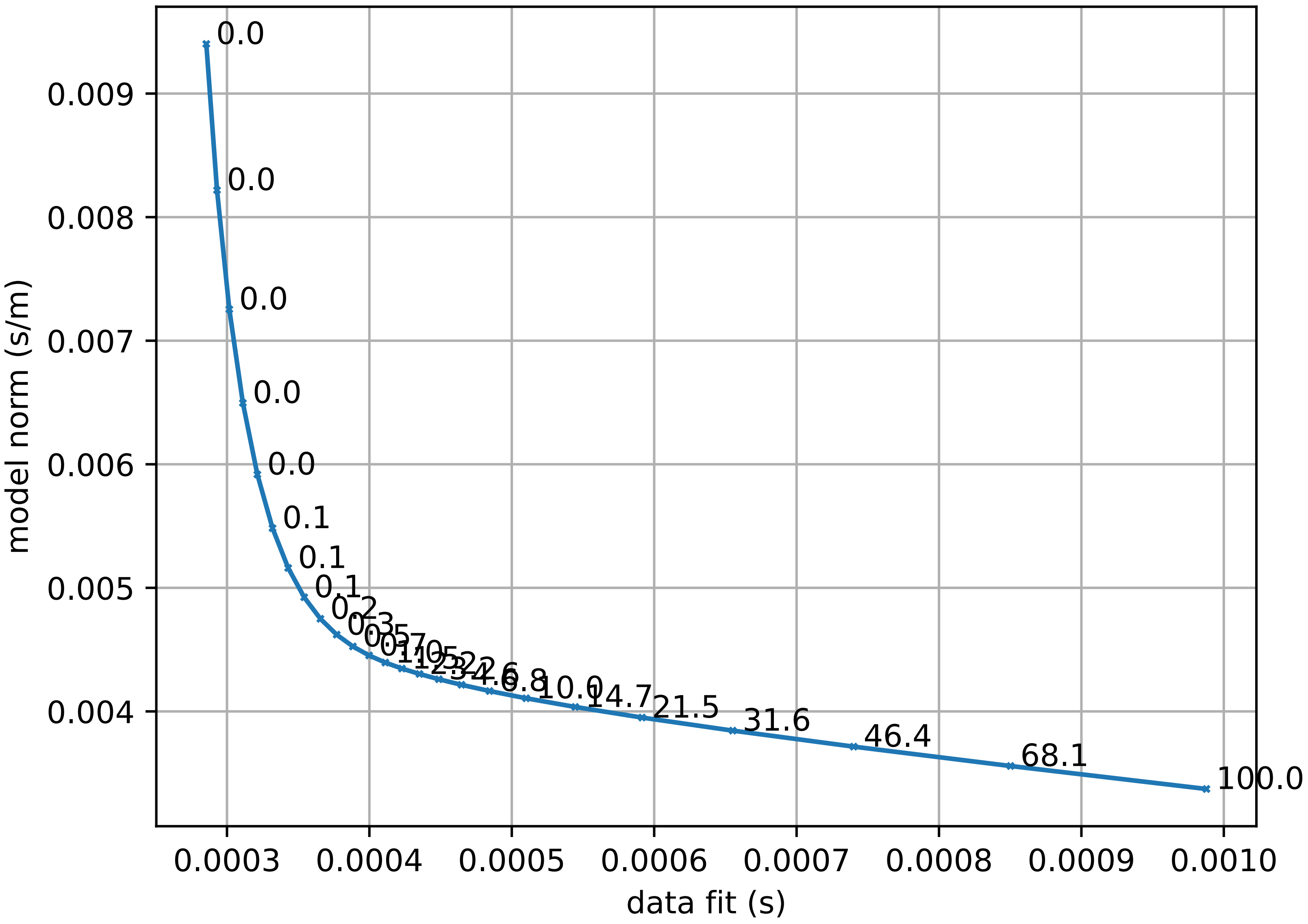
Inversion with smoothness constraints



Smoothness constraints: Data and model norm



Smoothness constraints: The L-curve

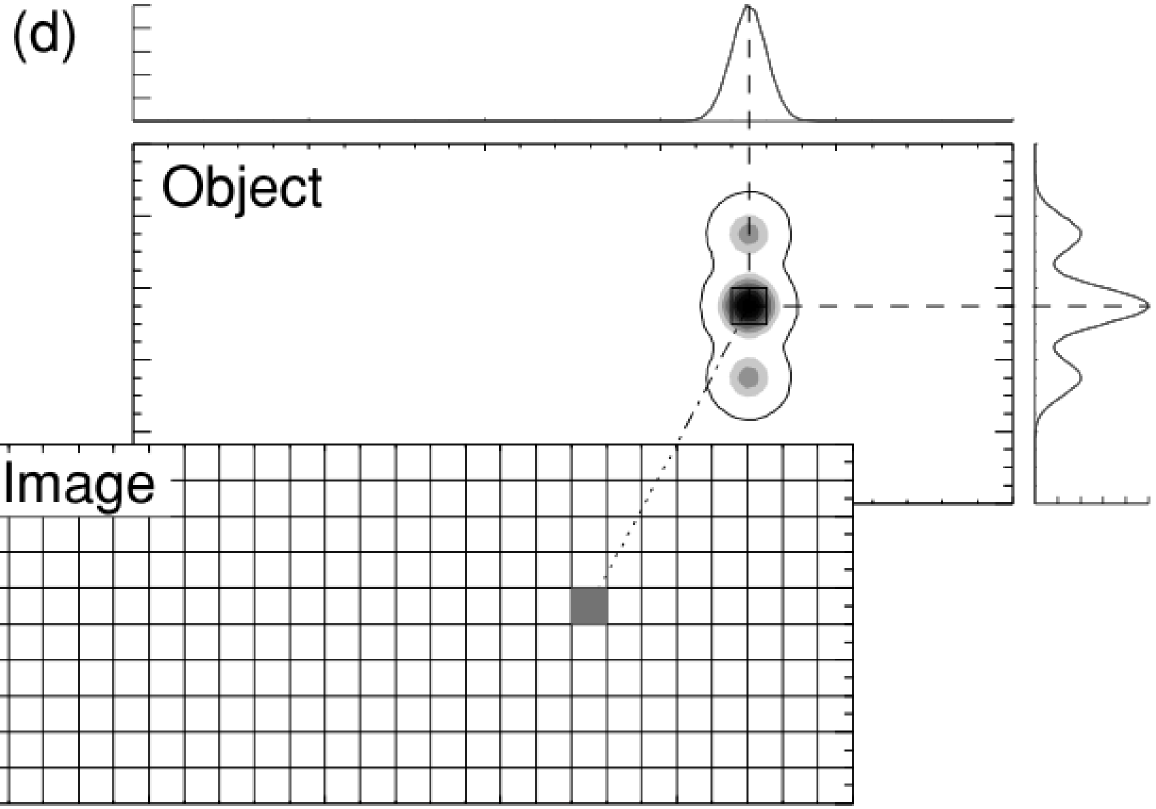


Resolution kernels (point spread functions)

$$\Rightarrow \mathbf{R}^M = \mathbf{G}^\dagger\mathbf{G} = (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{C}^T\mathbf{C})^{-1}\mathbf{G}^T\mathbf{G}$$

A column represents how a cell anomaly is “spread”.

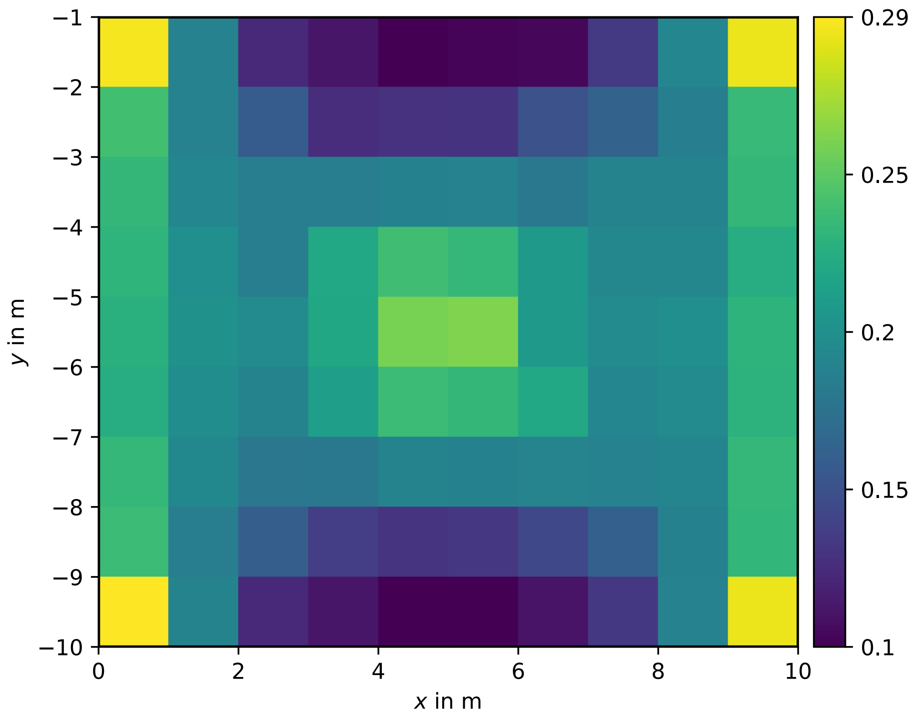
- Ideal imaging
- Contrast deficiency
- Geometrical distortion
- Side lobes



Model resolution radius

The diagonal of the resolution matrix

states how a single model cell can be retrieved



How big needs a sphere to be to integrate a value of 1 (perfect resolution)?

Resolution radius

$$r = \frac{r_i}{\sqrt{\mathbf{R}^M_{ii}}} = \frac{\sqrt{A_i/\pi}}{\sqrt{\mathbf{R}^M_{ii}}}$$

Model covariance matrix

Theorem

Sei $\mathbf{x}$ ein multivariabler, normalverteilter Zufallsvektor mit dem Erwartungswert μ und der Kovarianz $\mathbf{C}$ und sei $\mathbf{y} = \mathbf{A}\mathbf{x}$. Dann ist y ebenfalls ein multivariabler, normalverteilter Zufallsvektor mit dem Erwartungswert $E(y) = \mathbf{A}\mu$
und der Kovarianz $\mathrm{cov}(\mathbf{y}) = \mathbf{A}\mathbf{C}\mathbf{A}^T$

Inverse Probleme

$$E(\mathbf{m}) = E(\mathbf{G}^{\dagger}\mathbf{d}) = \mathbf{G}^{\dagger}E(\mathbf{d}) = \mathbf{R}^M\mathbf{m}^{true}$$

$$\mathrm{cov}(\mathbf{m}) = \mathbf{G}^{\dagger} \cdot \mathrm{cov}(\mathbf{d})(\mathbf{G}^{\dagger})^T$$

Beispiel Least-Squares mit einheitlicher Datenvarianz σ

$$\mathrm{cov}(\mathbf{m}) = \sigma^2(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1} = \sigma^2(\mathbf{G}^T\mathbf{G})^{-1}$$

Modern regularization

> ⓘ **Regularization is key to subsurface imaging**

$\Rightarrow$ Add any information about the subsurface:

- correlation lengths and angles
- structural boundaries (from boreholes, seismics, GPR)
- point information (samples, borehole logs,)
- limits of the possible parameters (e.g. positivity, natural bounds)