

## The program

### Part I: Examples

- Image deblurring
- Tomography
- Radiative transfer inversion
- Helioseismology

### Part II: Mainly direct methods

- Fourier transforms
- Singular value decomposition
- Backus-Gilbert or Mollifier method

### Part III: Mainly Iterative methods

- Noise and a priori knowledge
- Iteration algorithms
- Regularization by Tikhonov
- Nonlinear problems
- Genecode and Neural networks

## FT-inversion: Convolution kernels

Convolution problems of the type (including generalizations to higher dimensions,  $L$  is the size of the domain)

$$f(x) = \int_L K(x - x') g(x') dx'$$

can be solved in principle by Fourier transform. We define

$$\hat{f}(k) = \frac{1}{L} \int_L e^{-ikx} f(x) dx, \quad f(x) = \sum_k e^{ikx} \hat{f}(k) dx, \quad \frac{kL}{2\pi} \in \mathbb{Z}$$

An FT of the convolution type integral equation yields a simple algebraic relation between the respective Fourier coefficients

$$\hat{f}(k) = L \hat{K}(k) \hat{g}(k)$$

We therefore have a problem for which an analytic inversion formula exists

$$g(x) = \sum_k e^{ikx} \frac{\hat{f}(k)}{\hat{K}(k)}$$

Other inversion problems can be brought into the form of a convolution problem by means of variable transforms. The solar limb equation is an example for a kernel of the division type

$$f(x) = \int_a^b K\left(\frac{x}{x'}\right) g(x') dx'$$

Use  $x = e^y$ ,  $x' = e^{y'}$ ,  $dx' = e^{y'} dy'$  to obtain

$$f(x) = \int_{\ln a}^{\ln b} \underbrace{K(e^{y-y'})}_{K'(y-y')} \underbrace{g(e^{y'}) e^{y'}}_{g'(y')} dy'$$

Many other inverse problems have analytic inversion formulas.

## FT-inversion: X-ray transform

We have noted that the X-ray transform is also close to a convolution type problem  $\rightarrow$  we can also approach it by FT

$$\text{X-ray transform: } f(\mathbf{x}, \mathbf{e}_\theta) = \int_{-R}^R g(\mathbf{x} + s\mathbf{e}_\theta) ds$$

Insert the inverse 3D Fourier transform of  $g$

$$g(\mathbf{x}) = \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} \hat{g}(\mathbf{k}), \quad \mathbf{k} = \frac{\pi}{R} \mathbf{l}, \quad \mathbf{l} \in \mathbb{Z}^3$$

into the X-ray transform for fixed  $\mathbf{e}_\theta$

$$\int_{-R}^R g(\mathbf{x} + s\mathbf{e}_\theta) ds = \sum_{\mathbf{k}} \hat{g}(\mathbf{k}) \underbrace{\int_{-R}^R e^{i\mathbf{k}\cdot(\mathbf{x}+s\mathbf{e}_\theta)} ds}_{2R e^{i\mathbf{k}\cdot\mathbf{x}} \text{sinc}(R\mathbf{k}\cdot\mathbf{e}_\theta)} \sum_{\mathbf{k}_\perp \in \mathbf{e}_\theta} e^{i\mathbf{k}_\perp\cdot\mathbf{x}} \hat{g}(\mathbf{k}_\perp)$$

This is exactly the form of a 2D FT in the image plane.

$$\text{Hence } \hat{f}(\mathbf{k}, \mathbf{e}_\theta) = 2R \hat{g}(\mathbf{k}) \quad \text{for } \mathbf{k} \text{ in the plane } \perp \mathbf{e}_\theta.$$

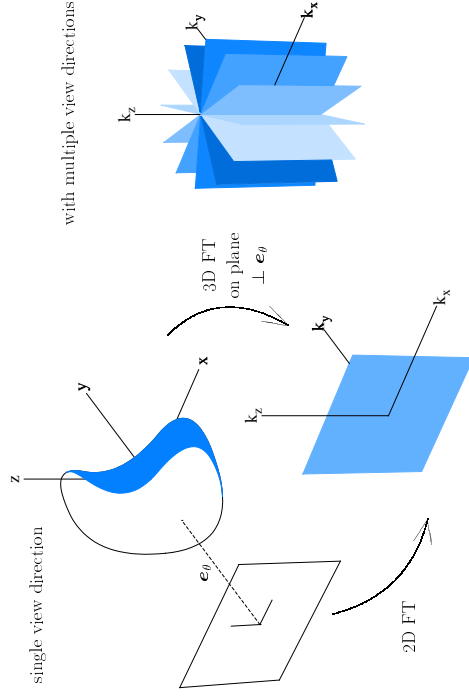


Illustration of the Fourier reconstruction of the X-ray transform

## FT-inversion: The noise problem

The practical usefulness of analytic inversion formulas is limited if the data is contaminated with noise

$$f(x) = \int_{-L}^L K(x-x') g(x') dx' + \epsilon(x)$$

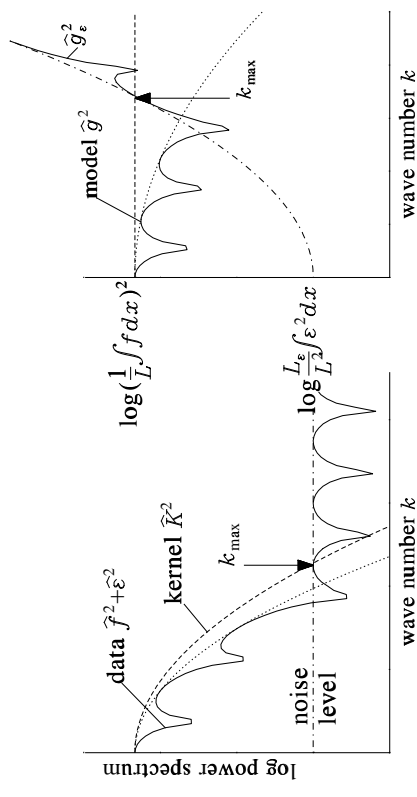
then

$$\hat{f}(k) = L \hat{K}(k) \hat{g}(k) + \hat{\epsilon}(k)$$

The noise is assumed of zero mean and correlation length  $L_\epsilon$ . Then its Fourier coefficients  $\hat{\epsilon}(k)$  are random complex numbers of zero mean and variance  $(L_\epsilon/L^2) \int_{-L}^L \epsilon^2 dx$  independent of  $k$  as long as  $k < 2\pi/L_\epsilon$ .

For the application of the analytic inversion formula the noise is a disaster:

$$g(x) + g_\epsilon(x) = \sum_k e^{ikx} \frac{\hat{f}(k)}{\hat{K}(k)} + \sum_k e^{ikx} \frac{\hat{\epsilon}(k)}{\hat{K}(k)}$$



Power spectra of data  $(f + \epsilon)$  (solid) and kernel  $K$  (dashed) before and after the inversion, i.e., division by  $\hat{K}$

## FT-inversion: Truncated spectrum

We have to limit the spectrum to wave numbers  $k_{\text{trunc}} < k_{\text{max}}$  where  $k_{\text{max}}$  is given by the intersection of the noise level with the kernel spectral power when both, data and kernel spectra are normalized to the same value at  $k = 0$

$$\frac{\hat{K}(0)}{\hat{K}(k_{\text{max}})} = \frac{\hat{f}^2(0) + \hat{\epsilon}^2(0)}{\hat{\epsilon}^2(0)} = 1 + \text{SNR}$$

The signal-to-noise ratio is here defined as

$$\text{SNR} = \frac{\hat{f}^2(0)}{\hat{\epsilon}^2(0)} = \frac{(1/L^2)(\int_L f dx)^2}{(L_\epsilon/L^2) \int_L \epsilon^2 dx}$$

As an example: image deblurring with Gaussian kernel in 1D

$$K(x) \simeq \exp -\frac{x}{2L_K}, \quad \hat{K}(k) \simeq \exp -\frac{1}{2}(kL_K)^2$$

then

$$\ln \left( \frac{\hat{K}(0)}{\hat{K}(k_{\text{max}})} \right) \simeq \frac{1}{2}(k_{\text{max}}L_K)^2 = 2\pi^2 \left( \frac{L_K}{L} \right)^2 l_{\text{max}}^2$$

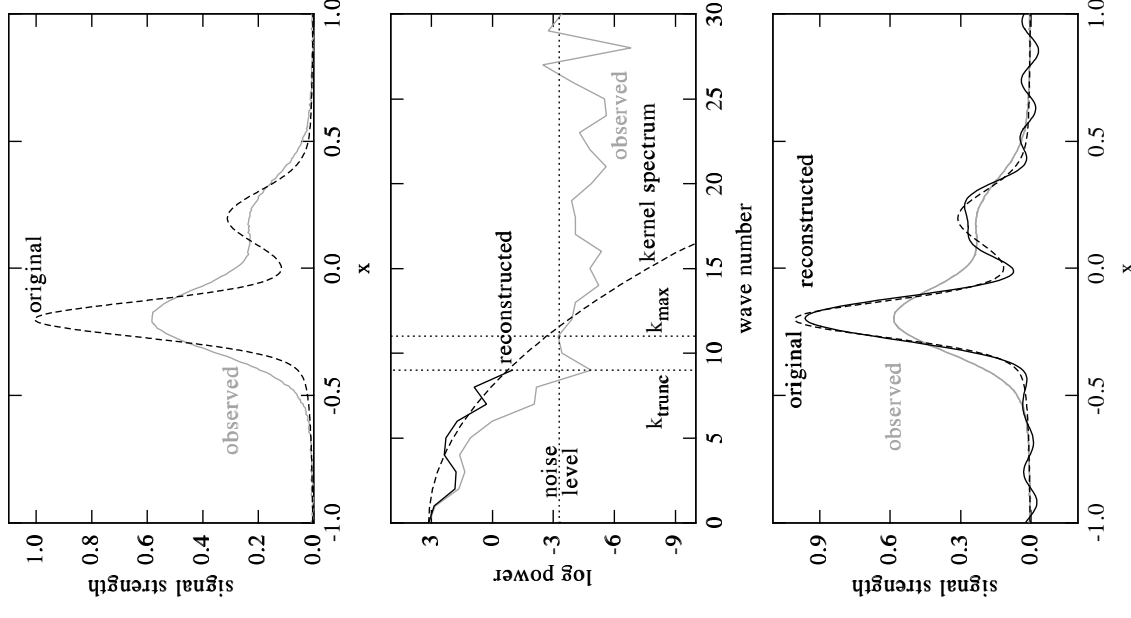
$l_{\text{max}}$  is the maximum number of complex Fourier coefficients of the reconstruction.

→ the number of independent image parameters of the reconstruction (factor 2 because the Fourier coefficients are complex)

$$2l_{\text{max}} \simeq \frac{L}{L_K} \sqrt{\ln(1 + \text{SNR})}$$

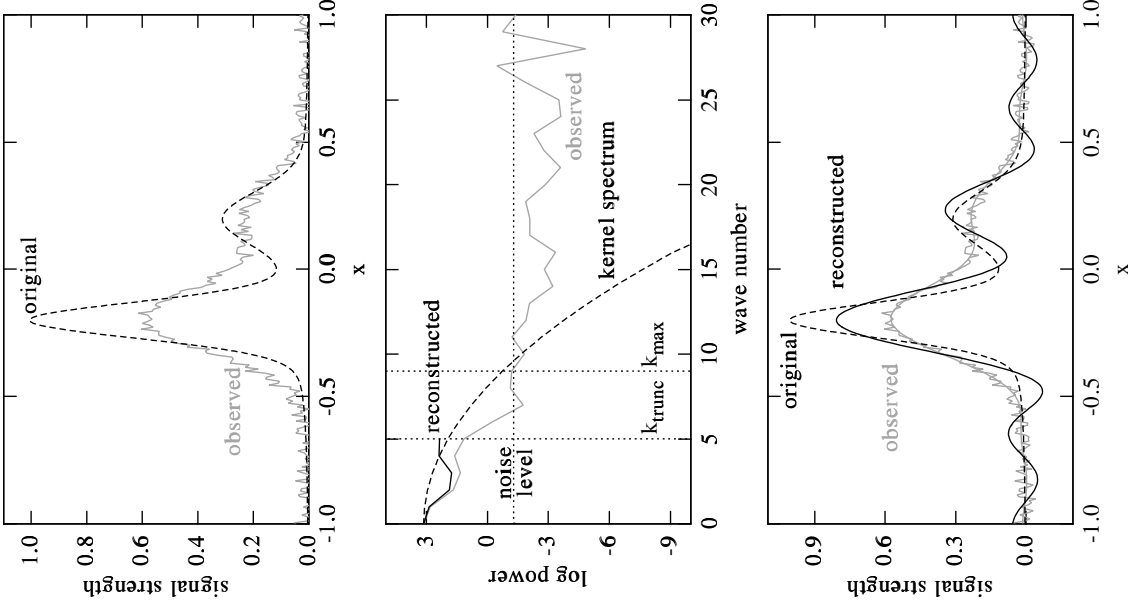
where  $L/L_K$  are the number of independent image parameters which have been measured. With the knowledge of the kernel we can enhance the number of independent image parameters depending on SNR.

## FT-inversion: Example with little noise



Example of a reconstruction. The Gaussian kernel has just a width corresponding to the distance of the two peaks in the original signal. Noise variance is  $0.005^2$ , the spectrum is truncated at  $3/4 k_{\text{max}}$ .

## FT-inversion: Example with more noise



Example of a reconstruction. The Gaussian kernel has just a width corresponding to the distance of the two peaks in the original signal. Noise variance is  $0.05^2$ , the spectrum is truncated at  $2/3 k_{\max}$ .

## SVD-inversion: The basics

We need a generalization of FT for more general inverse problems

$$f(x) = \int K(x, x') g(x') dx' \iff$$

$$\mathbf{f} = \mathbf{K} \mathbf{g}, \quad \mathbf{g} \in \mathbb{R}^m, \quad \mathbf{f} \in \mathbb{R}^n$$

The basic idea is to construct a symmetric matrix from  $\mathbf{K}$  which has a complete orthogonal set of eigenvectors and real eigenvalues

$$\begin{pmatrix} 0 & \mathbf{K} \\ \mathbf{K}^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u}_i \\ \mathbf{v}_i \end{pmatrix} = \lambda_i \begin{pmatrix} \mathbf{u}_i \\ \mathbf{v}_i \end{pmatrix} \quad \text{or} \quad \begin{matrix} \mathbf{K} \mathbf{v}_i = \lambda_i \mathbf{u}_i \\ \mathbf{K}^T \mathbf{u}_i = \lambda_i \mathbf{v}_i \end{matrix}$$

There are  $m$  vectors of the  $\mathbf{v}_i$  which orthogonally span the  $m$ -dimensional space  $\mathbb{R}^m$  and  $n$  vectors of the  $\mathbf{u}_i$  which orthogonally span the  $n$ -dimensional data space  $\mathbb{R}^n$  because

$$\mathbf{K}^T \mathbf{K} \mathbf{v}_i = \lambda_i^2 \mathbf{v}_i, \quad \mathbf{K} \mathbf{K}^T \mathbf{u}_i = \lambda_i^2 \mathbf{u}_i$$

→ For every nonzero  $\lambda_i$  there is also a negative one. There are at most  $i_{\text{nonzero}} = \min(n, m)$  pairs of nonzero eigenvalues  $\pm \lambda_i$

→ The action of  $\mathbf{K}$  is completely described by its singular value decomposition (to be read as a dyad)

$$\mathbf{K} = \sum_{i=1}^{i_{\text{nonzero}}} \mathbf{u}_i \lambda_i \mathbf{v}_i^T, \quad \mathbf{v}_i \in \mathbb{R}^m, \quad \mathbf{u}_i \in \mathbb{R}^n$$

where  $\mathbf{v}_i$  and  $\mathbf{u}_i$  are normalized to unity and all  $\lambda_i$  chosen positive and ordered so that  $\lambda_1 > \lambda_2 > \dots > \lambda_{i_{\text{nonzero}}} > 0$ .

- The decomposition  $\mathbf{K} \rightarrow \{\mathbf{v}_i, \mathbf{u}_i, \lambda_i\}$   $i = 1, \dots, i_{\text{nonzero}}$  can be found numerically for  $i_{\text{nonzero}} < \text{about } 1000$ .

## SVD-inversion: The nullspace

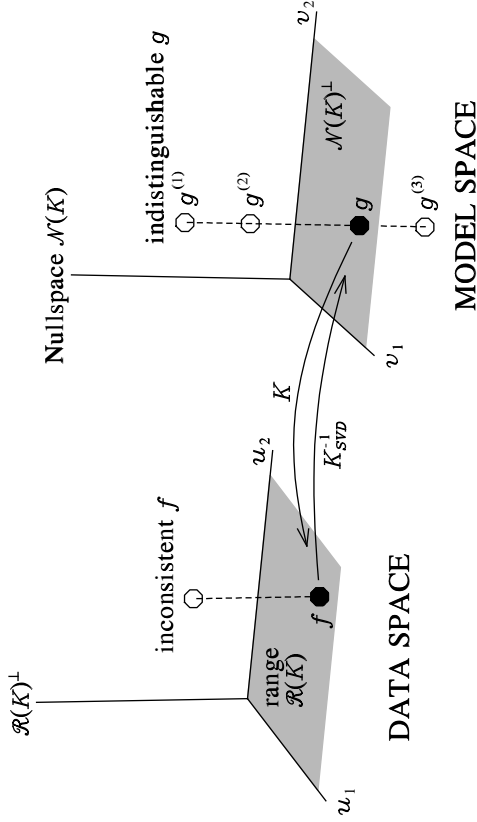
In most cases  $i_{\text{nzzero}} < \text{both } m \text{ and } n$ .

If  $i_{\text{nzzero}} < m$  not all model features are mapped into data space. There are nonequal models  $\mathbf{g}$  which cannot be distinguished by the observation operation of  $\mathbf{K}$ . The space spanned by  $\mathbf{v}_i$  with  $\lambda_i = 0$  is the nullspace  $\mathcal{N}(\mathbf{K})$  of the observation operator.

If  $i_{\text{nzzero}} < n$  the observations do not cover the total data space. There are inconsistent vectors  $\mathbf{f}$  which can impossibly be the result of an observation through  $\mathbf{K}$ . The space spanned by  $\mathbf{u}_i$  with  $\lambda_i \neq 0$  is the range  $\mathcal{R}(\mathbf{K})$  of observation operator.

As a generalized inverse to  $\mathbf{K}$  we define

$$\mathbf{K}_{\text{SVD}}^{-1} = \sum_{i=1}^{i_{\text{nzzero}}} \mathbf{v}_i \frac{1}{\lambda_i} \mathbf{u}_i$$



Mapping of  $\mathbf{K}$  and its generalized inverse  $\mathbf{K}_{\text{SVD}}^{-1}$  between data and model space. The “visible” part of model space is  $\mathcal{N}(\mathbf{K})^\perp$ , the orthogonal complement of the null space  $\mathcal{N}(\mathbf{K})$

## SVD-inversion: The noise problem

If noise is added, the data  $\mathbf{f} + \boldsymbol{\epsilon}$  is almost certainly inconsistent. We assume the noise  $\boldsymbol{\epsilon}$  to have zero mean and variance  $\sigma_\epsilon^2$ .

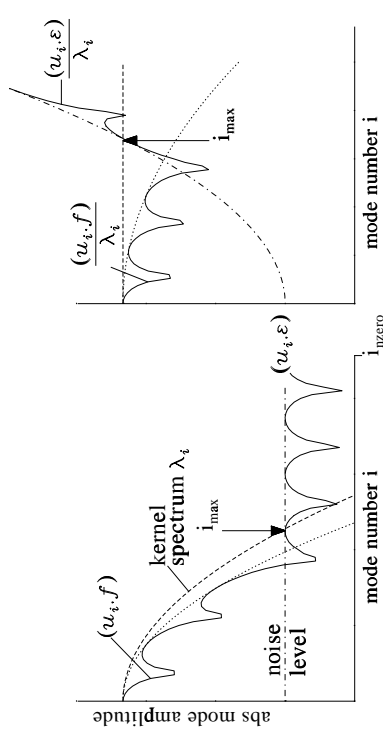
$\mathbf{K}_{\text{SVD}}^{-1}$  ignores the part of the noise which falls out of the range  $\mathcal{R}(\mathbf{K})$ . Yet the eigenvalues close to zero are problematic:

$$\mathbf{g} + \mathbf{g}_\epsilon = \mathbf{K}_{\text{SVD}}^{-1}(\mathbf{f} + \boldsymbol{\epsilon}) = \sum_{i=1}^{i_{\text{nzzero}}} \mathbf{v}_i \left[ \frac{(\mathbf{u}_i \cdot \mathbf{f})}{\lambda_i} + \frac{(\mathbf{u}_i \cdot \boldsymbol{\epsilon})}{\lambda_i} \right]$$

because  $(\mathbf{u}_i \cdot \boldsymbol{\epsilon})$  are random real numbers with zero mean and variance  $\sigma_\epsilon^2$  (the  $\mathbf{u}_i$  are normalized).

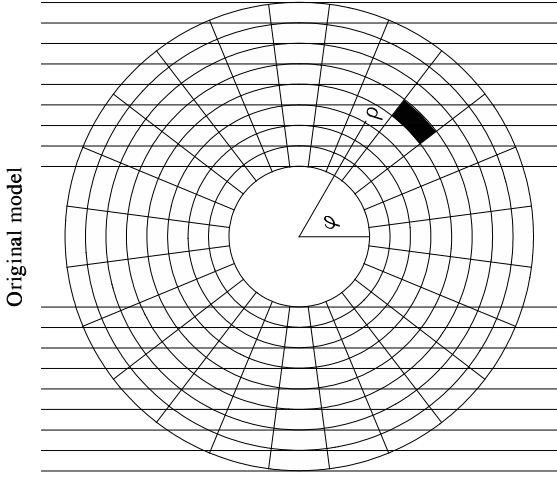
→ We have the same problem as in Fourier inversion. Depending on the noise level, we have to truncate the spectrum of  $\mathbf{K}_{\text{SVD}}^{-1}$  below mode  $i_{\text{max}}$  to be determined from

$$\frac{\lambda_1}{\lambda_{i_{\text{max}}}} \simeq \frac{\sqrt{(\mathbf{u}_1 \cdot \mathbf{f})^2 + \sigma_\epsilon^2}}{\sigma_\epsilon}$$

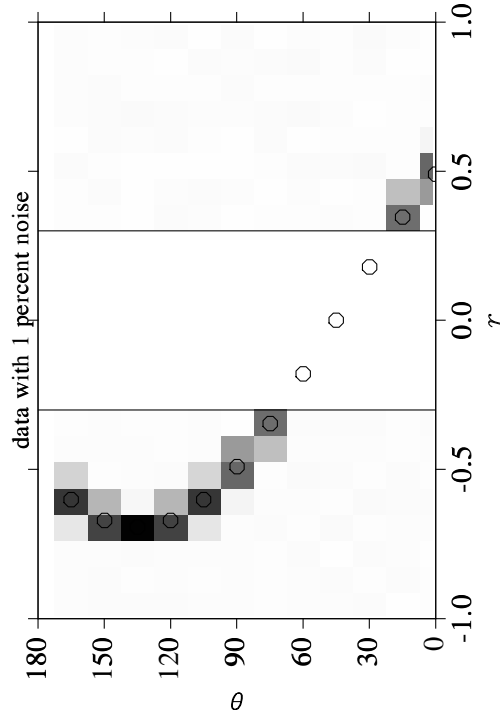


Normalized SVD spectra of data  $(\mathbf{f} + \boldsymbol{\epsilon})$  (solid) and kernel  $\mathbf{K}$  (dashed) before and after the inversion, i.e., division by  $\lambda_i$

## SVD-inversion: 2D tomography, model and data

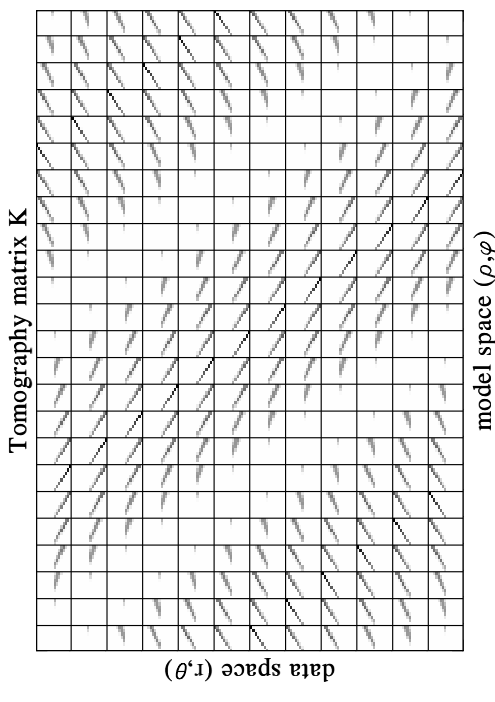


Tomography grid and model density. The grid is cylindrical with  $\varphi$  as azimuth angle and  $\rho$  as distance.

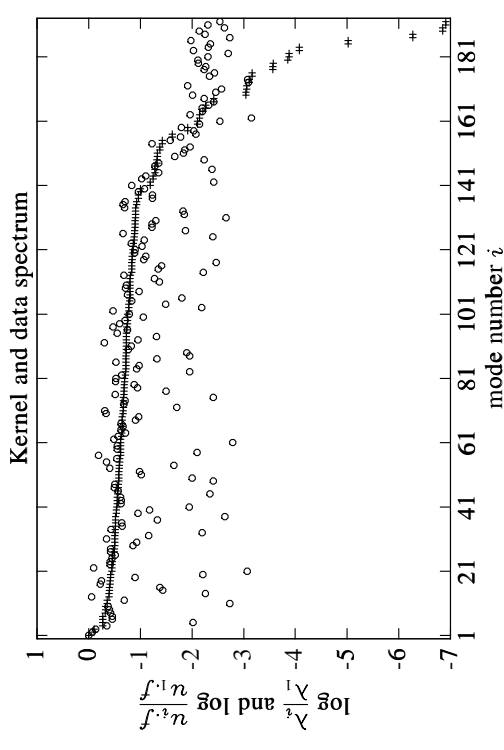


Data grid and image of original model with noise.  $\tau$  denotes the pixel number;  $\theta$  the view direction.

## SVD-inversion: 2D tomography, the kernel

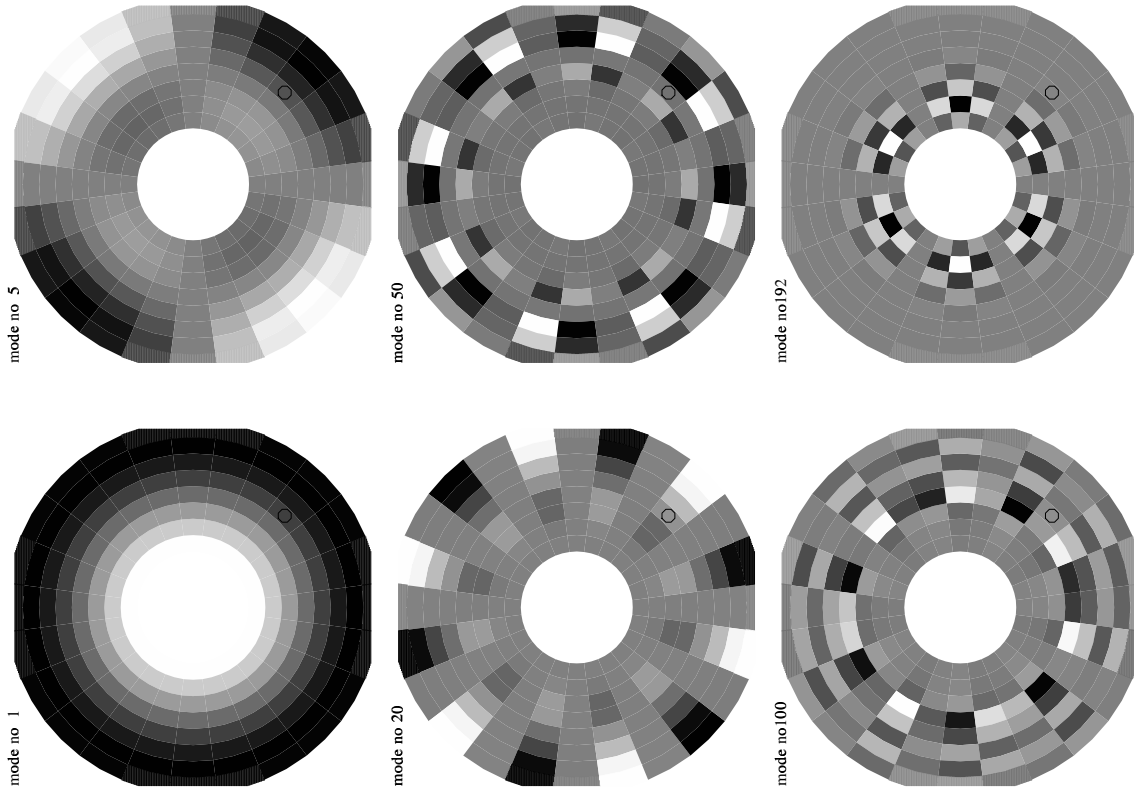


Tomography kernel matrix. Each subblock shows  $K$  for fixed  $\varphi$  and  $\theta$ . Zero elements are blank.



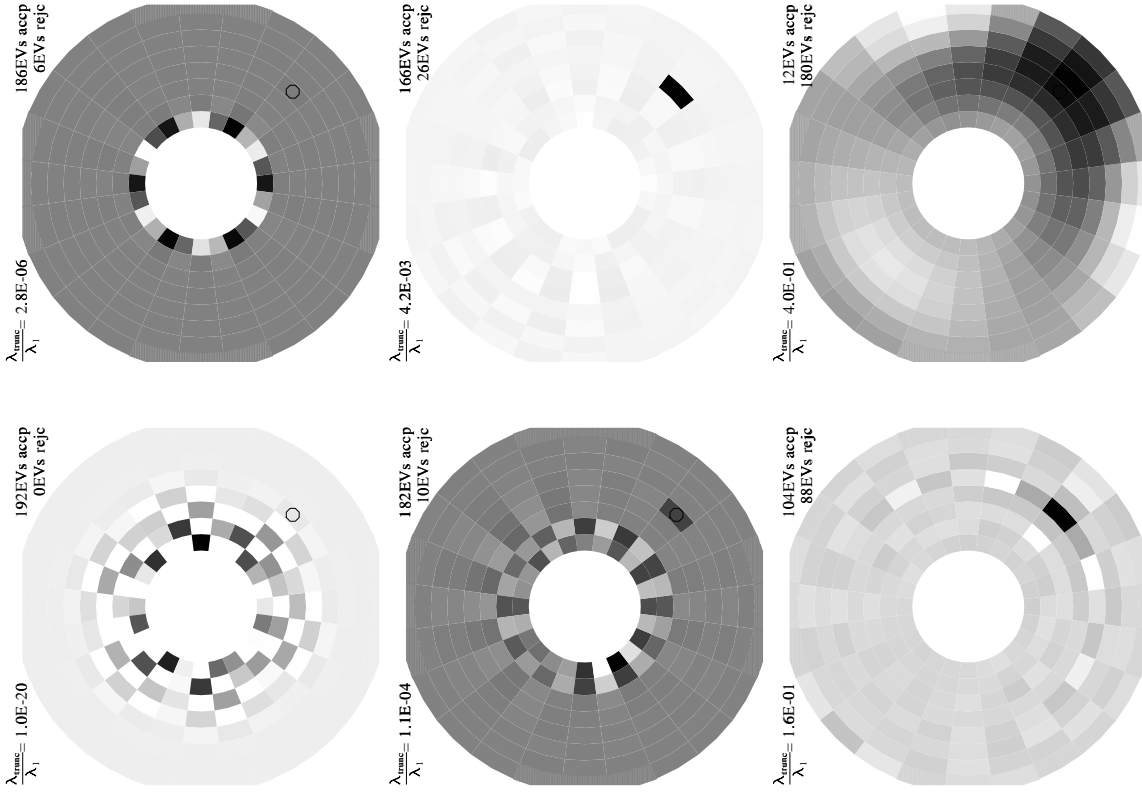
Spectrum of kernel (crosses) and data (circles)

## SVD-inversion: 2D tomography, the eigenfunctions



*Eigenfunctions  $v_i$  of some modes of the 2D tomography kernel*

## SVD-inversion: 2D tomography, reconstructions



*Reconstructions for various truncation levels  $\lambda_{\text{trunc}}/\lambda_1$*

## SVD-inversion: Generalized inverses

For most problems exact inverses  $\mathbf{K}^{-1}$  does not exist.

→ The concept of matrix inverses needs to be generalized. Generalized inverses  $\mathbf{K}_{\text{gen}}^{-1}$  are defined through the four Moore-Penrose criteria for generalized inverses :

Instead of being a unit matrix,  $\mathbf{K}_{\text{gen}}^{-1}\mathbf{K}$  and  $\mathbf{K}\mathbf{K}_{\text{gen}}^{-1}$  are only required to be symmetric

$$\begin{aligned} (\mathbf{K}_{\text{gen}}^{-1}\mathbf{K})^T &= \mathbf{K}_{\text{gen}}^{-1}\mathbf{K} && \text{(model resolution kernel)} \\ (\mathbf{K}\mathbf{K}_{\text{gen}}^{-1})^T &= \mathbf{K}\mathbf{K}_{\text{gen}}^{-1} && \text{(data resolution matrix)} \end{aligned}$$

and that they act as unit matrix at least in the “visible” model subspace  $\mathcal{N}(\mathbf{K})^\perp \subset \mathbb{R}^m$  and the range  $\mathcal{R}(\mathbf{K}) \subset \mathbb{R}^n$ , respectively,

$$\begin{aligned} \mathbf{K}\mathbf{K}_{\text{gen}}^{-1}\mathbf{K} &= \mathbf{K} \\ \mathbf{K}_{\text{gen}}^{-1}\mathbf{K}\mathbf{K}_{\text{gen}}^{-1} &= \mathbf{K}_{\text{gen}}^{-1} \end{aligned}$$

We find that  $\mathbf{K}\mathbf{K}_{\text{SVD}}^{-1}$  satisfies these criteria however its truncated version

$$\mathbf{K}_{\text{TSVD}}^{-1} = \sum_{i=1}^{i_{\text{trunc}}} \mathbf{v}_i \frac{1}{\lambda_i} \mathbf{u}_i$$

with  $i_{\text{trunc}} < i_{\text{zero}}$  satisfies only the first two Moore-Penrose criteria, because

$$\mathbf{K}_{\text{TSVD}}^{-1}\mathbf{K} = \sum_{i=1}^{i_{\text{trunc}}} \mathbf{v}_i \mathbf{v}_i^T, \quad \mathbf{K}\mathbf{K}_{\text{TSVD}}^{-1} = \sum_{i=1}^{i_{\text{trunc}}} \mathbf{u}_i \mathbf{u}_i^T$$

are projection operators onto only part of  $\mathcal{N}(\mathbf{K})^\perp$  and  $\mathcal{R}(\mathbf{K})$ , respectively.

## BG or mollifier inversion: Motivation

Assume we have a continuous model and a discrete number of observations, i.e.,  $g \in \text{Hilbert space}$  and  $\mathbf{f} \in \mathbb{R}^n$ . For each individual measurement  $i = 1, \dots, n$  we have

$$f_i = \int K_i(x') g(x') dx'$$

The problem is hopelessly underdetermined and a conventional inverse of  $\mathbf{K}(x)$  can never be achieved.

→ We only want to obtain an estimate of  $g(x)$  which should be a more or less localized average. Since the problem is linear this estimate must be a linear combination of the data. For each  $x$  find coefficients  $\mathbf{q}(x)$  with

$$\bar{g}(x) = \sum_{i=1}^n q_i(x) f_i = \underbrace{\int \sum_{i=1}^n q_i(x) K_i(x') g(x') dx'}_{\text{model resolution kernel } R(x, x')}$$

The resolution kernel (compare to Moore-Penrose definition) here is a suitable linear superposition of the individual forward kernels  $K_i$ .

The model resolution kernel  $R(x, x')$  should satisfy

- Localization within width  $w$

$$\begin{aligned} R_w(x, x') &\longrightarrow 0 && \text{for } |x - x'| > w \\ R_w(x, x') &\xrightarrow{w \rightarrow 0} \delta(x - x') \end{aligned}$$

- Normalization

$$\int R_w(x, x') dx' = 1$$

## BG or mollifier inversion: SVD and noise

Assume we were able to construct a SVD of the kernel functions

$$\mathbf{K}(x) = \sum_{i=1}^{i_{\text{nzzero}}} \mathbf{u}_i \lambda_i v_i(x), \quad p \leq n$$

then the  $v_i(x)$  span  $\mathcal{N}(\mathbf{K})^\perp$  completely and the resolution kernel  $R(x, x')$  has a representation in this basis. For simplicity we construct  $R(x, x')$  so that it is diagonal:

$$R(x, x') = \sum_{i=1}^{i_{\text{nzzero}}} v_i(x) r_i v_i(x')$$

then the equivalent inverse is

$$\mathbf{K}_{\text{moll}}^{-1}(x) = \sum_{i=1}^{i_{\text{nzzero}}} v_i(x) \frac{r_i}{\lambda_i} \mathbf{u}_i \quad \text{with} \quad \mathbf{K}_{\text{moll}}^{-1}(x) \mathbf{K} = R(x, x')$$

→ no truncation as in TSVD but gentle roll-off due to filter coefficients  $r_i$ .

If the observations  $f_i$  are contaminated with noise  $\epsilon_i$  then the estimate  $\bar{g}$  becomes affected as well:

$$\bar{g}(x) + \bar{g}_\epsilon(x) = \sum_{i=1}^{i_{\text{nzzero}}} q_i(x) f_i + \sum_{i=1}^{i_{\text{nzzero}}} q_i(x) \epsilon_i$$

If the noise has zero mean and variance  $\sigma_\epsilon^2$  the error  $\bar{g}_\epsilon$  of the estimate has zero mean and variance  $\sigma_\epsilon^2 |\mathbf{q}|^2$ . → To confine the error due to data noise we need as additional requirement:

- Shortest possible  $\mathbf{q}$

$$\sum_{i=1}^{i_{\text{nzzero}}} q_i^2(x) \rightarrow \text{minimum}$$

## BG or mollifier inversion: Mollification

For each  $x$  try to find coefficients  $q_i(x)$  so that  $R_w(x, x')$  comes close to a desired mollifier function  $\bar{\delta}_w(x, x')$  with width  $w$ , i.e., solve (usually by SVD)

$$\sum_{i=1}^n q_i(x) K_i(x') = \bar{\delta}_w(x, x') + \text{res}(x')$$

where  $\text{res}(x') \in \mathcal{N}(\mathbf{K})$  is the part of the mollifier which falls into the nullspace of  $\mathbf{K}$ . Tune width  $w$  so that the error  $\propto |\mathbf{q}|$  does not exceed given bounds.

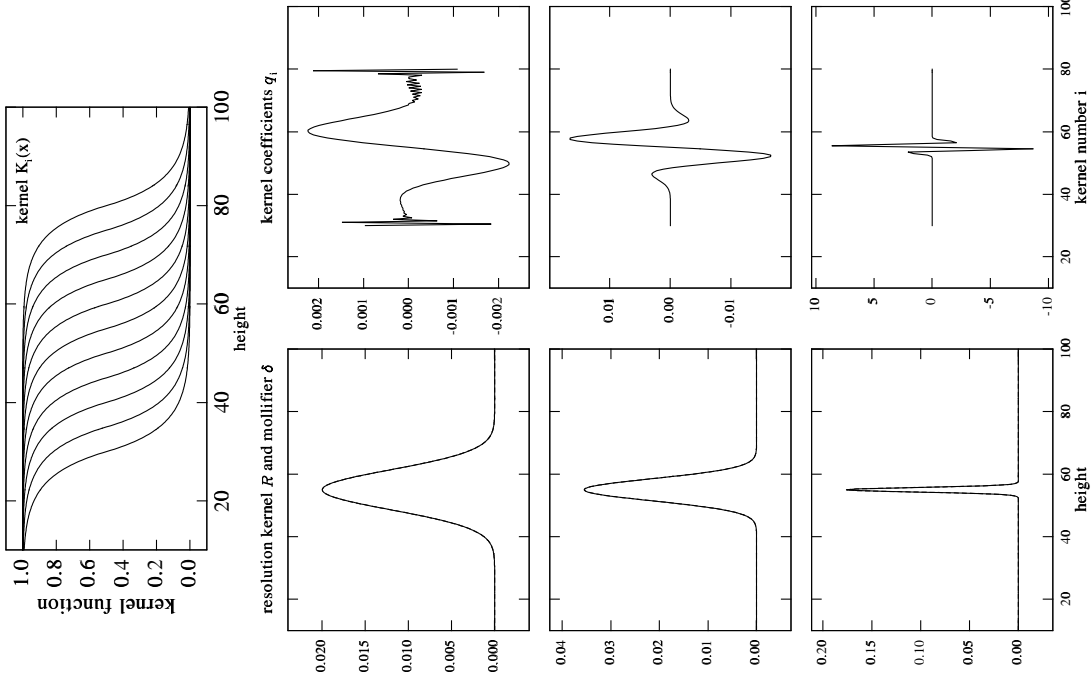
Disadvantages:

- The above equation has to be solved for every  $x$  at which an estimate  $\bar{g}$  is required. Note, however, that the above equation is much easier to solve than the original problem because there is no noise involved.
- The computational overhead is large unless symmetries of the system reduce the number of resolution kernels  $R_w(x, x')$  to be calculated.

Advantages:

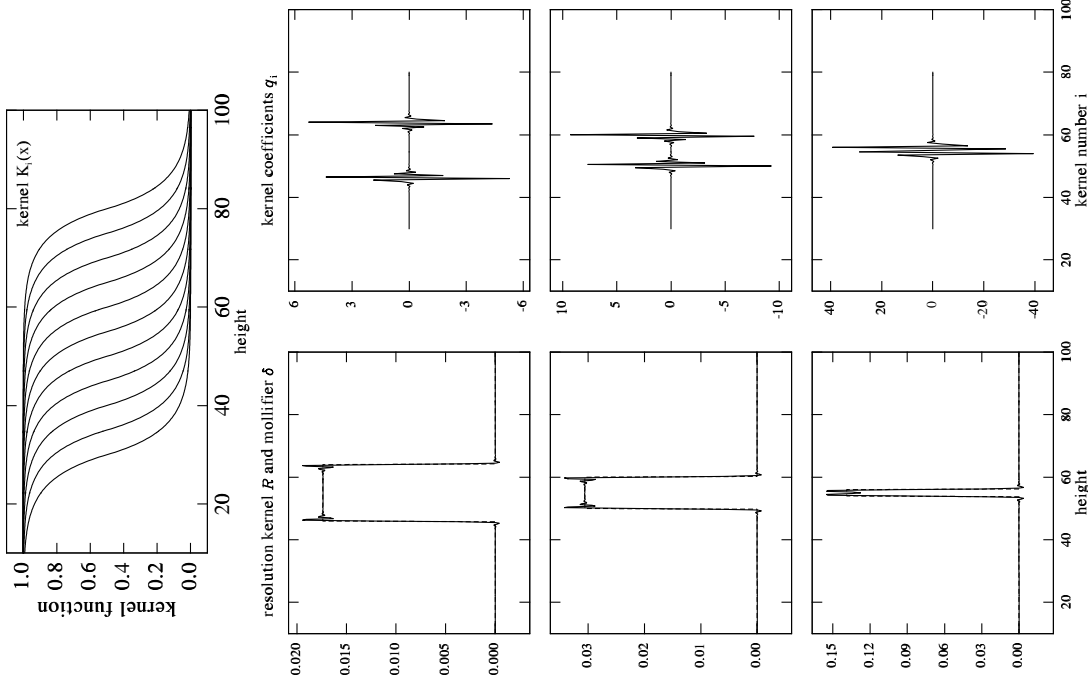
- For every  $x$  we not only obtain an estimate  $\bar{g}$  of the model but also a resolution kernel  $R_w(x, x')$  telling us which region  $\bar{g}(x)$  is representative of. We also obtain an individual error estimate  $\sigma_\epsilon |\mathbf{q}|$  for each  $\bar{g}$ .
- There is no need to discretize the model space
- The resolution kernels  $R_w(x, x')$  can be used again with different data if the kernels  $K_i(x')$  have not changed

## BG or mollifier inversion: Gaussian mollifier



Mollifiers  $\bar{\delta}$  and resolution kernels  $R_w$  (left) and kernel coefficients  $q_i$  (right) for the kernels  $K_i$  in the top diagram. Resolution kernels are derived for  $x$  (height) = 55 and different width  $w$ . Resolution kernels and mollifiers are practically identical.

## BG or mollifier inversion: Box-shape mollifier

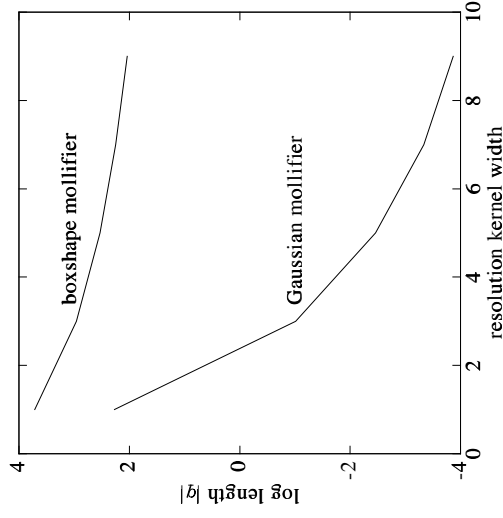


Mollifiers  $\bar{\delta}$  and resolution kernels  $R_w$  (left) and kernel coefficients  $q_i$  (right) for the kernels  $K_i$  in the top diagram. Resolution kernels are derived for  $x$  (height) = 55 and different width  $w$ . Mollifiers are exactly box-shape.

## BG or mollifier inversion

### Noise comparison

The amount of noise in the estimate  $\bar{g}(x)$  is  $\sigma_\epsilon |\mathbf{q}|$ . Here,  $\sigma_\epsilon$  is the standard deviation of the noise in the data, and  $|\mathbf{q}|$  is the length of the kernel coefficient vector given below.



*Length of  $\mathbf{q}$  vs width  $w$  of the resolution kernel for a Gaussian and box-shape mollifier*

## BG or mollifier inversion: Backus-Gilbert approach

We do not specify the shape of the resolution kernel but only try to concentrate its width around a given  $x$  by minimizing

$$\int (x-x')^2 R^2(x, x') dx' = \int (x-x')^2 \left( \sum_{i=1}^n q_i(x) K_i(x') \right)^2 dx$$

$$= \sum_{i,j=1}^n q_i(x) q_j(x) \int (x-x')^2 K_i(x') K_j(x') dx \equiv (\mathbf{q}(x) \cdot \mathbf{W}(x) \cdot \mathbf{q}(x))$$

This expression has to be minimized along with  $\sigma_\epsilon^2 |\mathbf{q}(x)|^2$  (noise reduction) under the normalization constraint

$$1 = \int R(x, x') dx' = \sum_{i=1}^n q_i(x) \int K_i(x') dx \equiv (\mathbf{q}(x) \cdot \mathbf{p})$$

Using Lagrangian multipliers  $\alpha$  and  $\beta$ , the coefficient vector  $\mathbf{q}(x)$  is determined by

$$(\mathbf{q} \cdot \mathbf{W} \mathbf{q}) + \alpha \sigma_\epsilon^2 (\mathbf{q} \cdot \mathbf{q}) + \beta [(\mathbf{q} \cdot \mathbf{p}) - 1] \rightarrow \text{minimum}$$

for known matrix  $\mathbf{W}$  and vector  $\mathbf{p}$ .

The result is 
$$\mathbf{q} = \frac{1}{(\mathbf{p} \cdot [\mathbf{W} + \alpha \sigma_\epsilon^2 \mathbf{1}]^{-1} \mathbf{p})} [\mathbf{W} + \alpha \sigma_\epsilon^2 \mathbf{1}]^{-1} \mathbf{p}$$

which has to be solved for every  $x$ . The parameter  $\alpha$  serves to balance resolution vs noise and stabilize the inversion of the  $n \times n$  matrix  $\mathbf{W} + \alpha \sigma_\epsilon^2 \mathbf{1}$ .

- The resulting  $R(x, x')$  is well concentrated around  $x$  but yet may not be well centered on  $x$ . Therefore, an additional constraint is sometimes used to obtain well centered resolution kernels

## BG or mollifier inversion: Tomography

In tomography the index  $i$  stands for pixel number  $r$  and view direction  $\theta$ . In 2D:

$$f_{(r,\theta)} = \int K_{(r,\theta)}(x') g(x') dx', \quad \text{where } x \in \mathbb{R}^2, \quad r \in \mathbb{R}$$

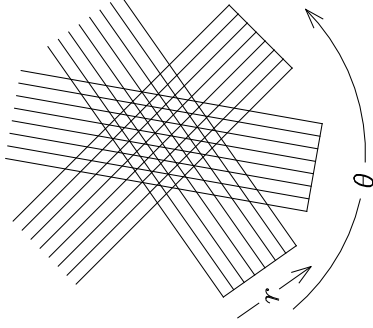
and  $K_{(r,\theta)}(x')$  is the beam from pixel  $r$  into direction  $e_\theta$

$$K_{(r,\theta)}(x') = \begin{cases} 1 & \text{if } x' \text{ inside the beam } (r, \theta) \\ 0 & \text{else} \end{cases}$$

The mollifier method seeks  $\bar{g}(x) = \sum_{r,\theta} q_{(r,\theta)}(x) f_{(r,\theta)}$  with

$$q_{(r,\theta)} \text{ so that } R(x, x') = \sum_{r,\theta} q_{(r,\theta)}(x) K_{(r,\theta)}(x') \longrightarrow \delta(x - x')$$

hence, for each  $x$  find coefficients  $q_{(r,\theta)}(x)$  so that the resulting superposition of beams approaches a  $\delta$  function at  $x$ .



In filtered backprojection tomography the special choice is

$$q_{(r,\theta)}(x) = w_{r-r_x} K_{(r_x,\theta)}(x) \quad \text{where } r_x \text{ so that } K_{(r_x,\theta)}(x) \neq 0$$

This gives a symmetric resolution kernel  $R(x, x')$  and

$$\bar{g}(x) = \underbrace{\sum_{\theta} K_{(r_x,\theta)}(x)}_{\text{backproj}} \underbrace{\sum_r w_{r-r_x} f_{(r,\theta)}}_{\text{filter}}$$

## Conclusions: What is the problem with inverse problems ?

Kernel functions are “smooth” in the sense (Riemann-Lebesgue)

$$\int K(x, x') \begin{cases} \cos kx' \\ \sin kx' \end{cases} dx' \longrightarrow 0 \quad \text{as } k \longrightarrow \infty$$

- $f$  insensitive to the short wavelength structure in  $g$
- solving for  $g$  is an ill-posed problem (Hamadard):
  - $g$  is either not unique (nullspaces)
  - $g$  changes discontinuously with  $f$  (small eigenvalues of  $K$ )

What is the solution to the problem with inverse problems?  
Replace the original problem by a series of solvable problems:

$$f(x) = \int K_{\gamma}(x, x') g_{\gamma}(x') dx' \quad \text{with} \quad \lim_{\gamma \rightarrow 0} K_{\gamma}(x, x') = K(x, x')$$

and set  $g = \lim_{\gamma \rightarrow 0} g_{\gamma}$ . Examples for the regularization parameter:

$$\begin{aligned} \gamma &= 1/k_{\text{trunc}} && \text{in FT inversion} \\ &= 1/i_{\text{trunc}} && \text{in SVD inversion} \\ &= \text{width } w && \text{in mollification} \\ &= \alpha && \text{in Backus-Gilbert inversion} \end{aligned}$$

In practical cases, however, we have to stop at a finite  $\gamma$  due to noise. The key problems are:

- to find the optimum value  $\gamma^*$  of  $\gamma$ ,
- to understand which features of  $g_{\gamma^*}$  will survive if we could let  $\gamma \longrightarrow 0$
- which features  $\lim_{\gamma \rightarrow 0} g_{\gamma}$  might have which  $g_{\gamma^*}$  does not have.
- which contribution from  $\mathcal{N}(K)$  has to be added to  $g_{\gamma^*}$ .