Lecture 23

Exemplary Inverse Problems including Earthquake Location

Syllabus

Lecture 01	Describing Inverse Problems
Lecture 02	Probability and Measurement Error, Part 1
Lecture 03	Probability and Measurement Error, Part 2
Lecture 04	The L_2 Norm and Simple Least Squares
Lecture 05	A Priori Information and Weighted Least Squared
Lecture 06	Resolution and Generalized Inverses
Lecture 07	Backus-Gilbert Inverse and the Trade Off of Resolution and Variance
Lecture 08	The Principle of Maximum Likelihood
Lecture 09	Inexact Theories
Lecture 10	Nonuniqueness and Localized Averages
Lecture 11	Vector Spaces and Singular Value Decomposition
Lecture 12	Equality and Inequality Constraints
Lecture 13	L_1 , L_∞ Norm Problems and Linear Programming
Lecture 14	Nonlinear Problems: Grid and Monte Carlo Searches
Lecture 15	Nonlinear Problems: Newton's Method
Lecture 16	Nonlinear Problems: Simulated Annealing and Bootstrap Confidence Intervals
Lecture 17	Factor Analysis
Lecture 18	Varimax Factors, Empircal Orthogonal Functions
Lecture 19	Backus-Gilbert Theory for Continuous Problems; Radon's Problem
Lecture 20	Linear Operators and Their Adjoints
Lecture 21	Fréchet Derivatives
Lecture 22	Exemplary Inverse Problems, incl. Filter Design
Lecture 23	**Exemplary Inverse Problems, incl. Earthquake Location**
Lecture 24	Exemplary Inverse Problems, incl. Vibrational Problems

Purpose of the Lecture

solve a few exemplary inverse problems

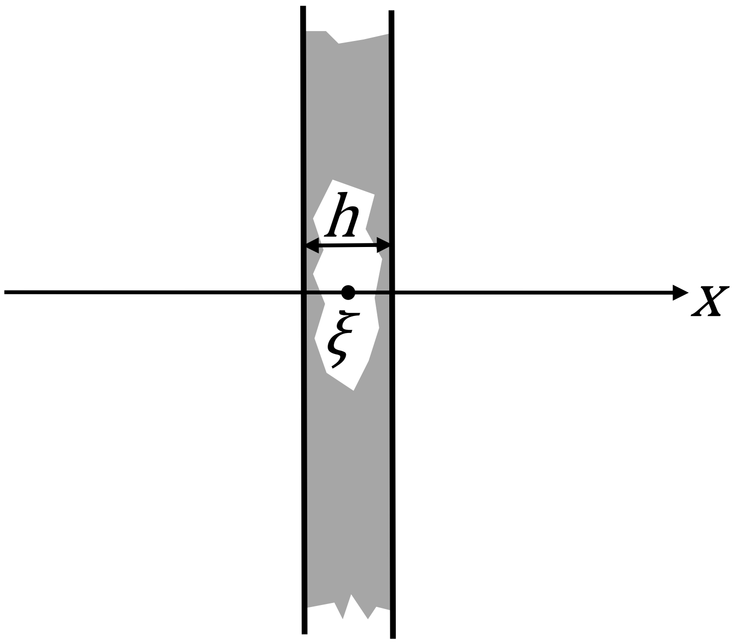
thermal diffusion
earthquake location
fitting of spectral peaks

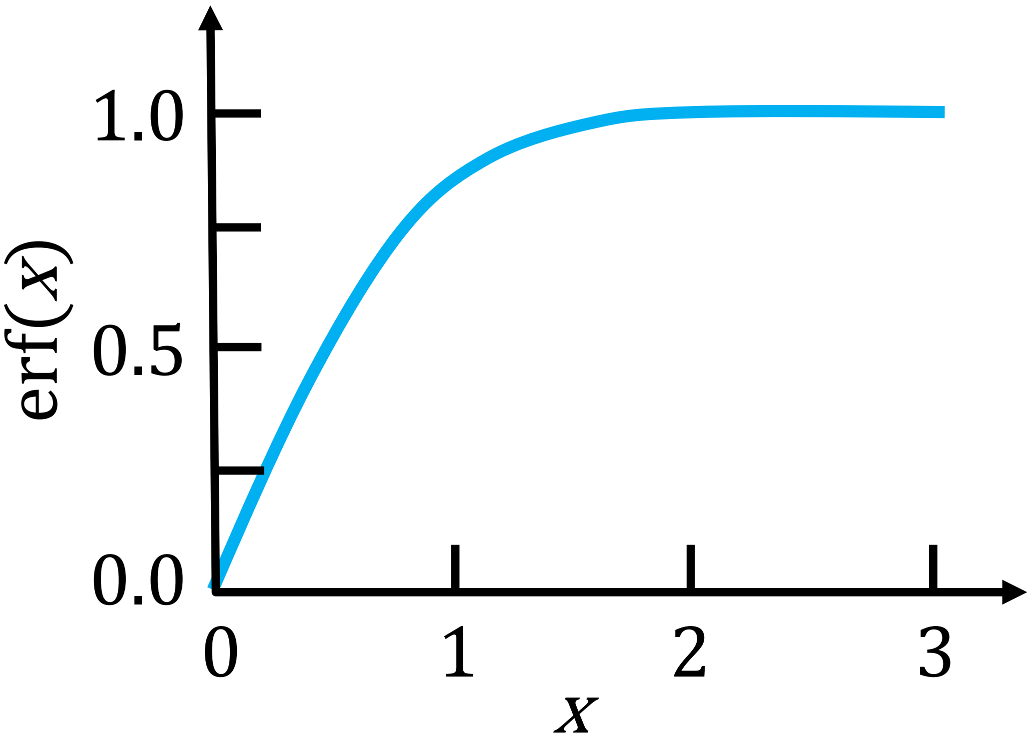
Part 1

thermal diffusion

temperature in a cooling slab

$$T(x,t) = \tfrac{1}{2}T_0\left\{\operatorname{erf}\left[\frac{x-(\xi-\tfrac{1}{2}h)}{\sqrt{t}}\right]-\operatorname{erf}\left[\frac{x-(\xi+\tfrac{1}{2}h)}{\sqrt{t}}\right]\right\} = T_0\, g(x,t,\xi)$$





temperature due to M cooling slabs (use linear superposition)

$$T(x_i, t) = \sum_{j=1}^{M} g(x_i, t, \xi_j)\ m_j$$

temperature due to M slabs each with initial temperature m_j

$$T(x_i, t) = \sum_{j=1}^{M} g(x_i, t, \xi_j)\ m_j$$

temperature measured at time $t>0$

initial temperature

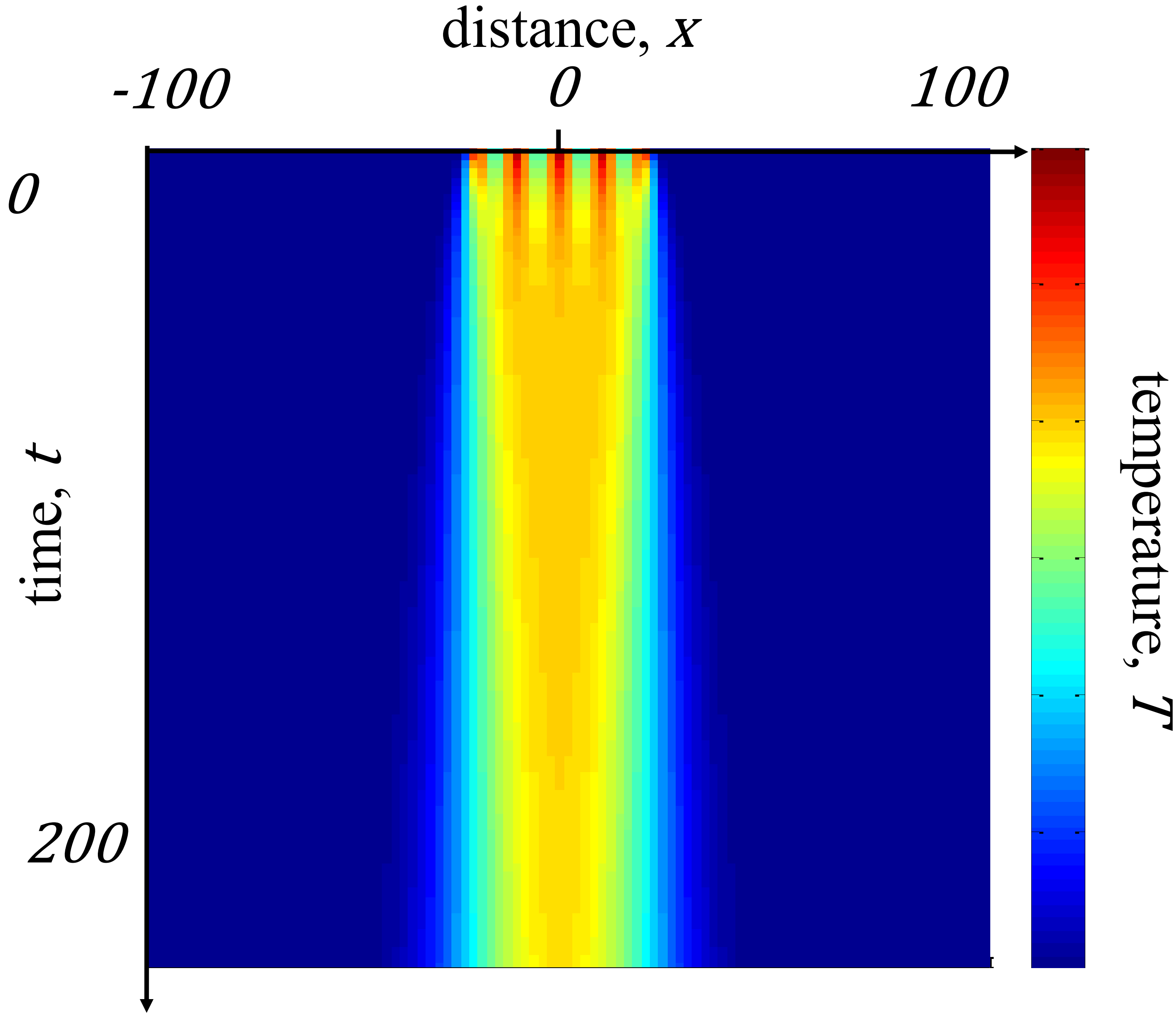
inverse problem
infer initial temperature **m**
using temperatures measures at a suite of *x*s
at some fixed later time t

$$T(x_i, t) = \sum_{j=1}^{M} g(x_i, t, \xi_j)\ m_j$$

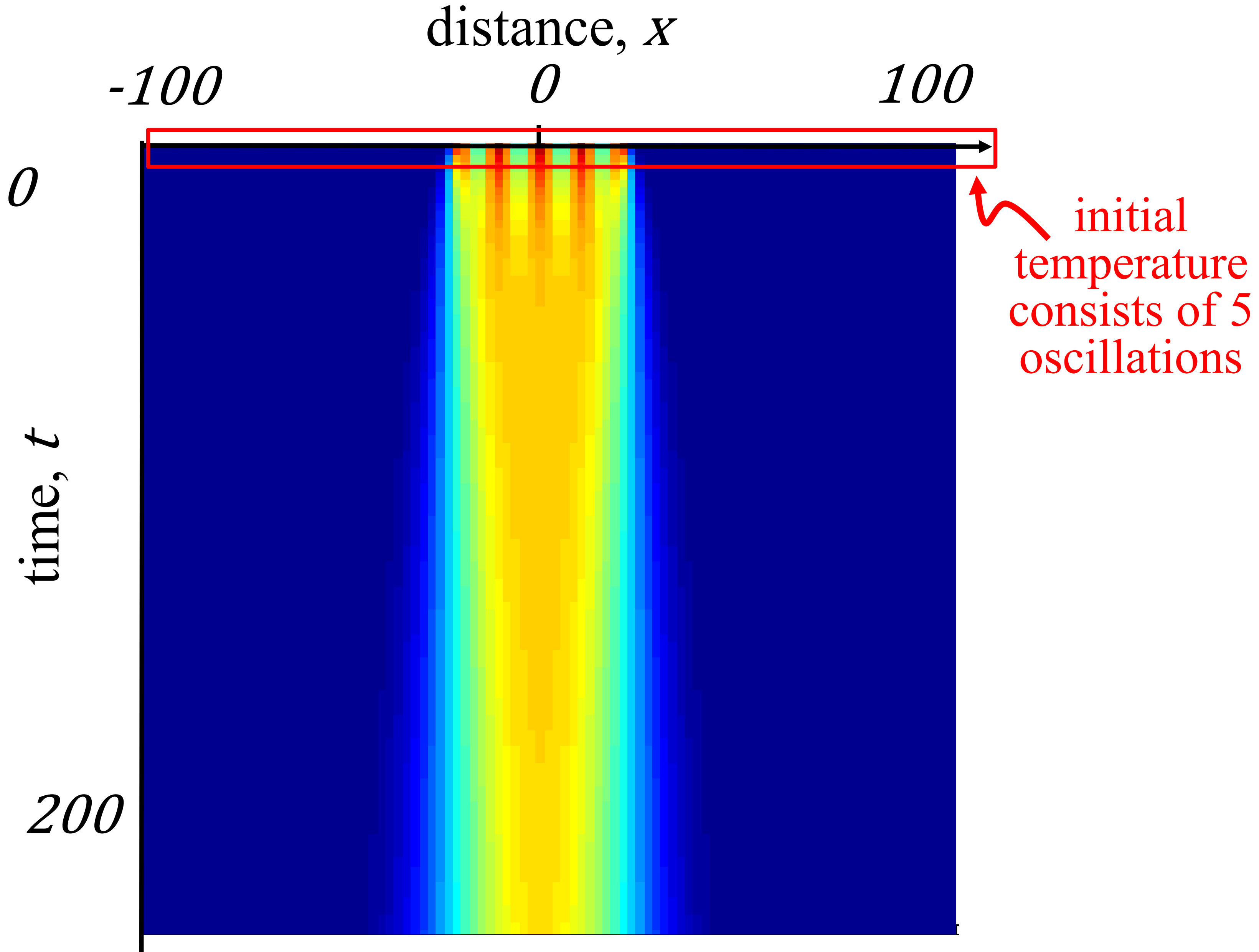
data

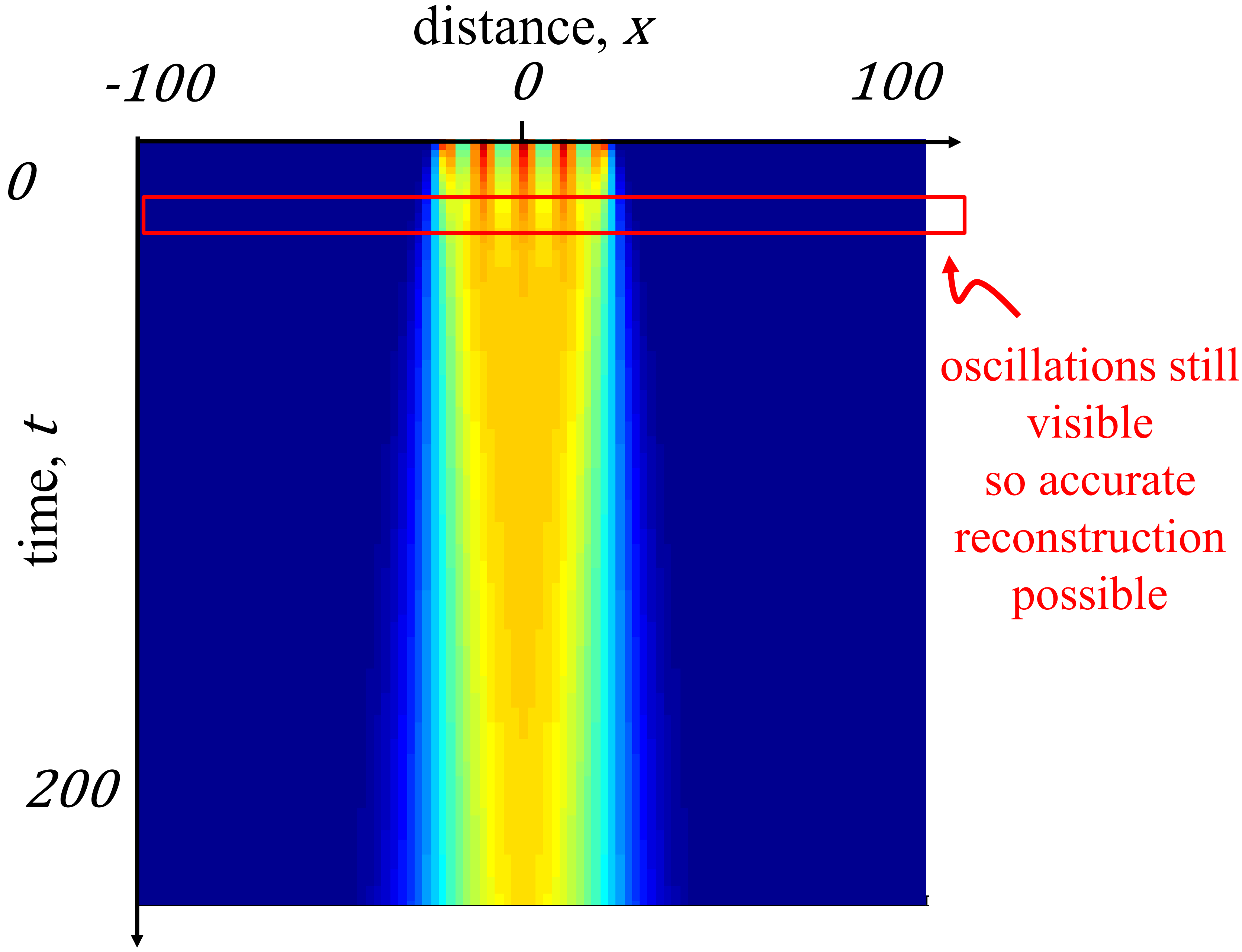
model parameters

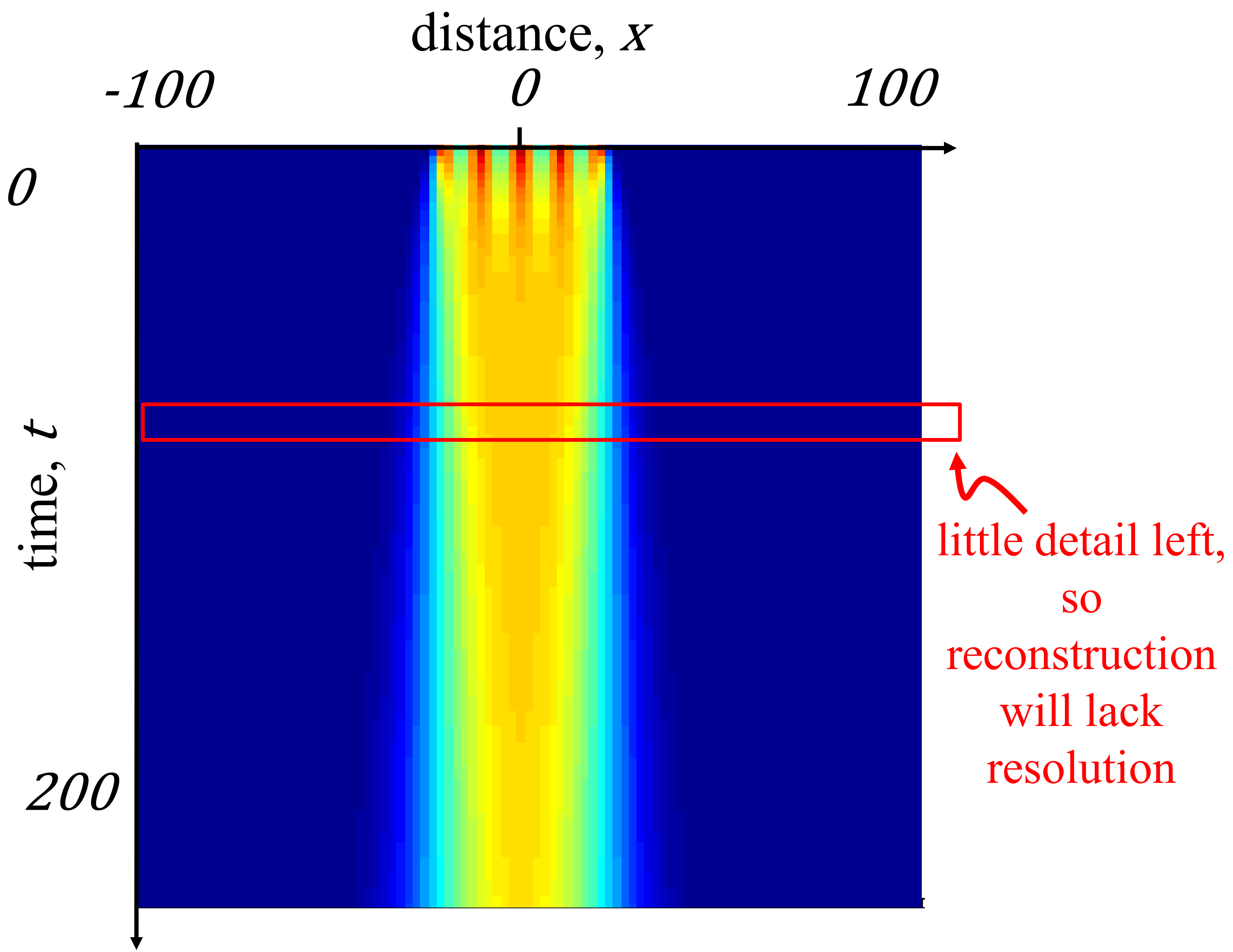
$$\mathbf{d} = \mathbf{G}\,\mathbf{m}$$



distance, x
-100
0
100
0
200
time, t
temperature, T


distance, x
-100
0
100
0
time, t
200
initial temperature consists of 5 oscillations


distance, x
-100
0
100
0
time, t
200
oscillations still visible so accurate reconstruction possible


distance, x
-100
0
100
0
time, t
200
little detail left,
so
reconstruction
will lack
resolution

What Method ?

The resolution is likely to be rather poor, especially when data are collected at later times

damped least squares

$$\mathbf{G}^{-g} = [\mathbf{G}^T\mathbf{G}+\varepsilon^2\mathbf{I}]^{-1}\mathbf{G}^T$$

damped minimum length

$$\mathbf{G}^{-g} = \mathbf{G}^T\,[\mathbf{G}\mathbf{G}^T+\varepsilon^2\mathbf{I}]^{-1}$$

Backus-Gilbert

What Method ?

The resolution is likely to be rather poor, especially when data are collected at later times

damped least squares

$$\mathbf{G}^{-g} = [\mathbf{G}^T\mathbf{G}+\varepsilon^2\mathbf{I}]^{-1}\mathbf{G}^T$$

damped minimum length

$$\mathbf{G}^{-g} = \mathbf{G}^T [\mathbf{G}\mathbf{G}^T+\varepsilon^2\mathbf{I}]^{-1}$$

actually, these generalized inverses are equal

Backus-Gilbert

What Method ?

The resolution is likely to be rather poor, especially when data are collected at later times

damped least squares

$$\mathbf{G}^{-g} = [\mathbf{G}^T\mathbf{G}+\varepsilon^2\mathbf{I}]^{-1}\mathbf{G}^T$$

damped minimum length

$$\mathbf{G}^{-g} = \mathbf{G}^T [\mathbf{G}\mathbf{G}^T+\varepsilon^2\mathbf{I}]^{-1}$$

Backus-Gilbert ← might produce solutions with fewer artifacts

Try both

damped least squares

Backus-Gilbert

Solution Possibilities

1. Damped Least Squares:
   Matrix **G** is not sparse
   no analytic version of $\mathbf{G}^T\mathbf{G}$ is available
   M=100 is rather small
   experiment with values of ε^2
   ```
   mest=(G'*G+e2*eye(M,M))\(G'*d)
   ```

2. Backus-Gilbert
   use standard formulation, with damping α
   experiment with values of α
   ```
   GMG = zeros(M,N);
   u = G*ones(M,1);
   for k = [1:M]
       S = G * diag(([1:M]-k).^2) * G';
       Sp = alpha*S + (1-alpha)*eye(N,N);
       uSpinv = u'/Sp;
       GMG(k,:) = uSpinv / (uSpinv*u);
   end
   ```

Solution Possibilities

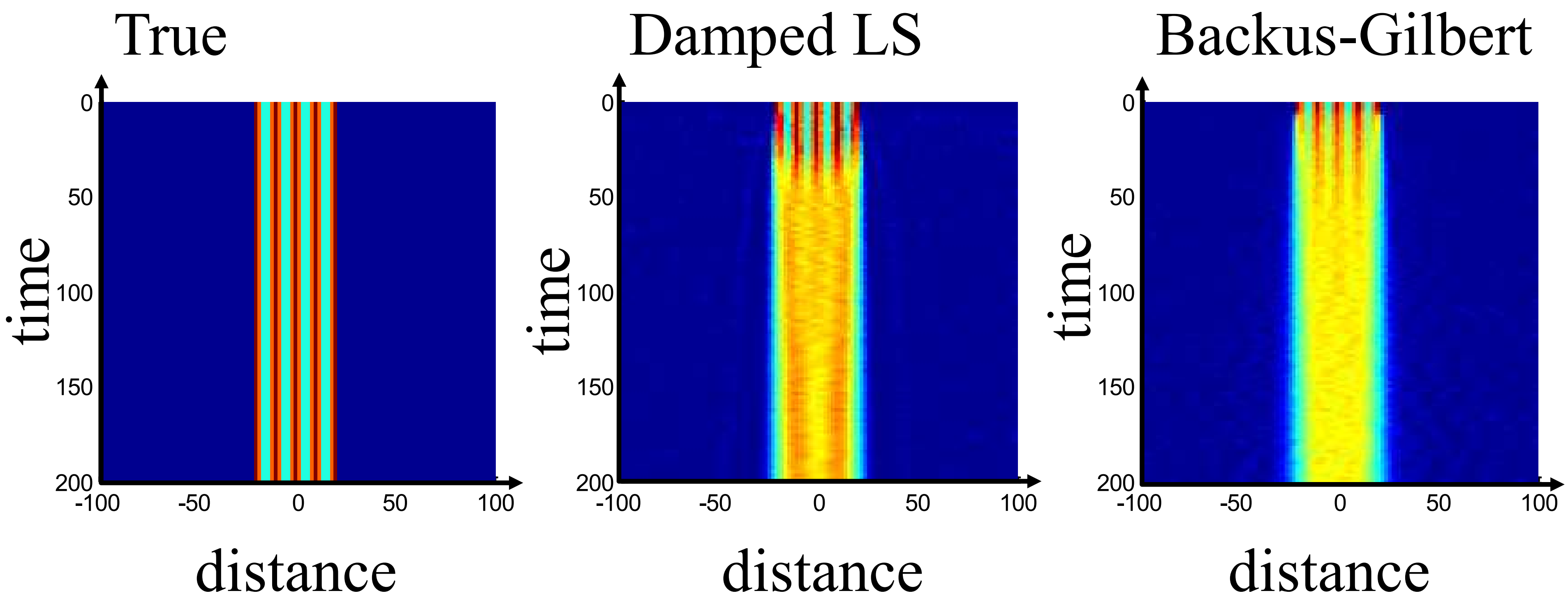
1. Damped Least Squares:
   Matrix **G** is not sparse
   no analytic version of $\mathbf{G}^T\mathbf{G}$ is available
   M=100 is rather small
   experiment with values of ε^2
   mest=(G'*G+e2*eye(M,M))\(G'*d)

2. Backus-Gilbert
   use standard formulation, with damping α
   experiment with values of α

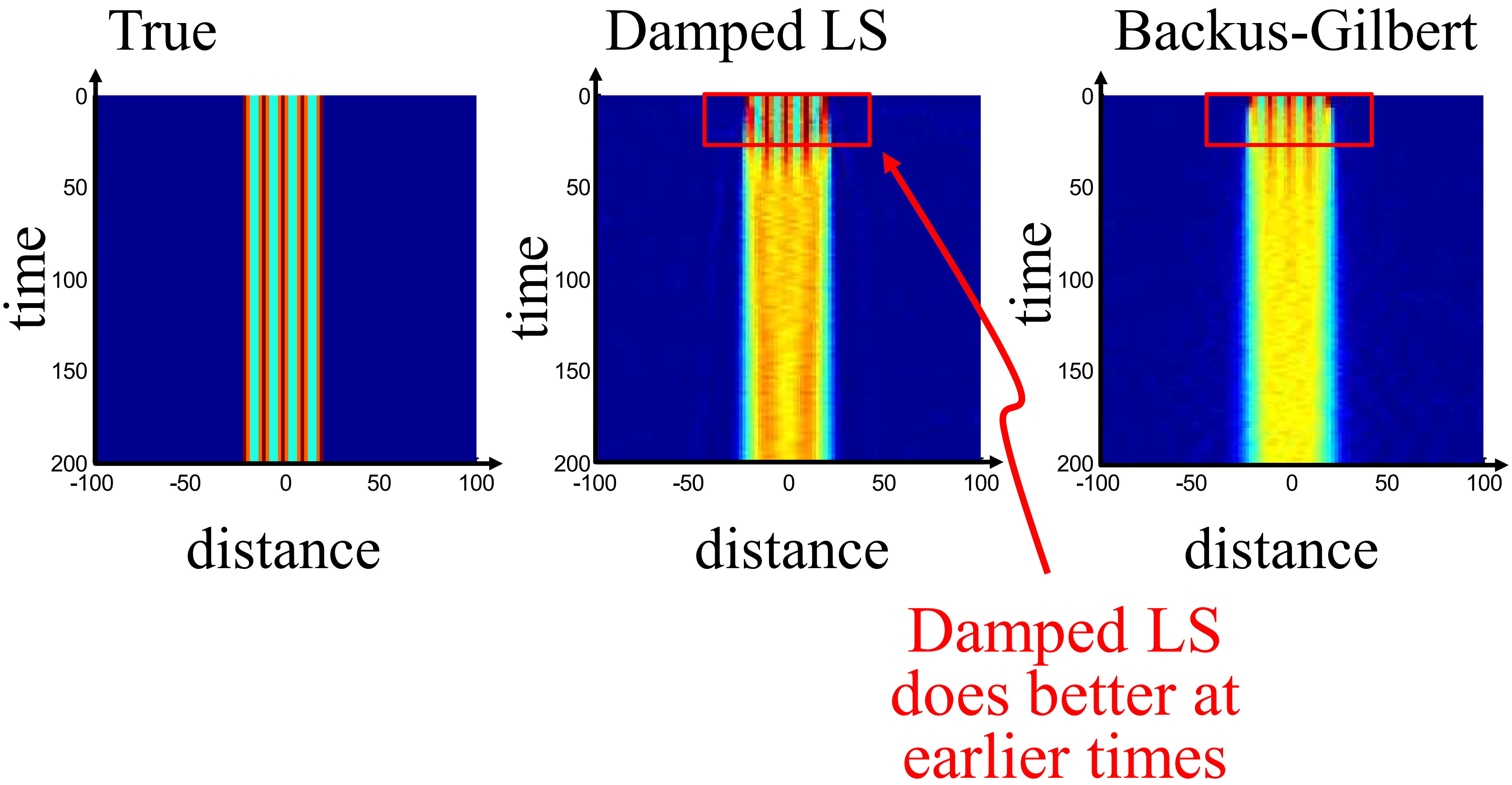
```
GMG = zeros(M,N);
u = G*ones(M,1);
for k = [1:M]
    S = G * diag(([1:M]-k).^2) * G';
    Sp = alpha*S + (1-alpha)*eye(N,N);
    uSpinv = u'/Sp;
    GMG(k,:) = uSpinv / (uSpinv*u);
end
```

try both

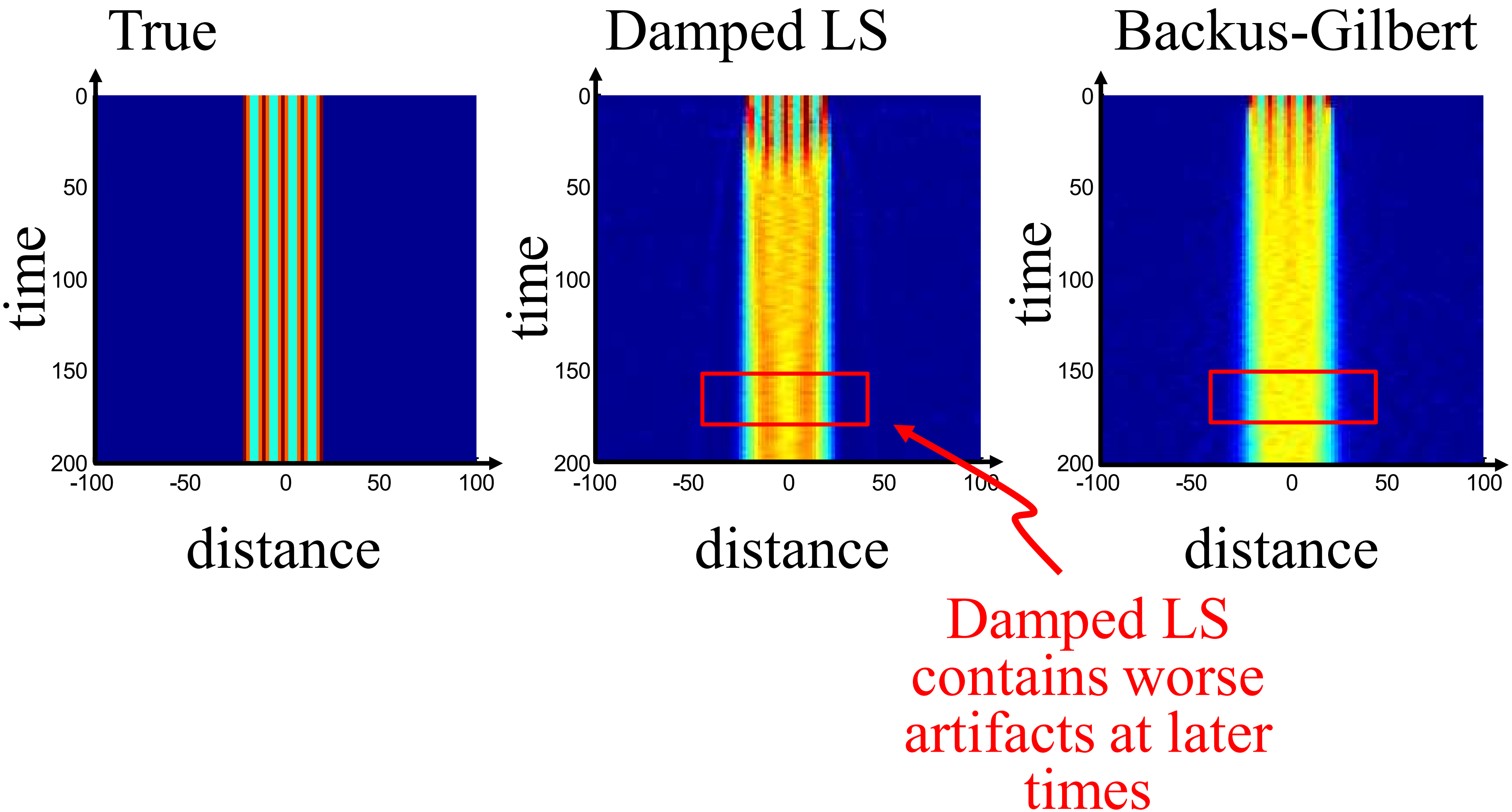
estimated initial temperature distribution as a function of the time of observation



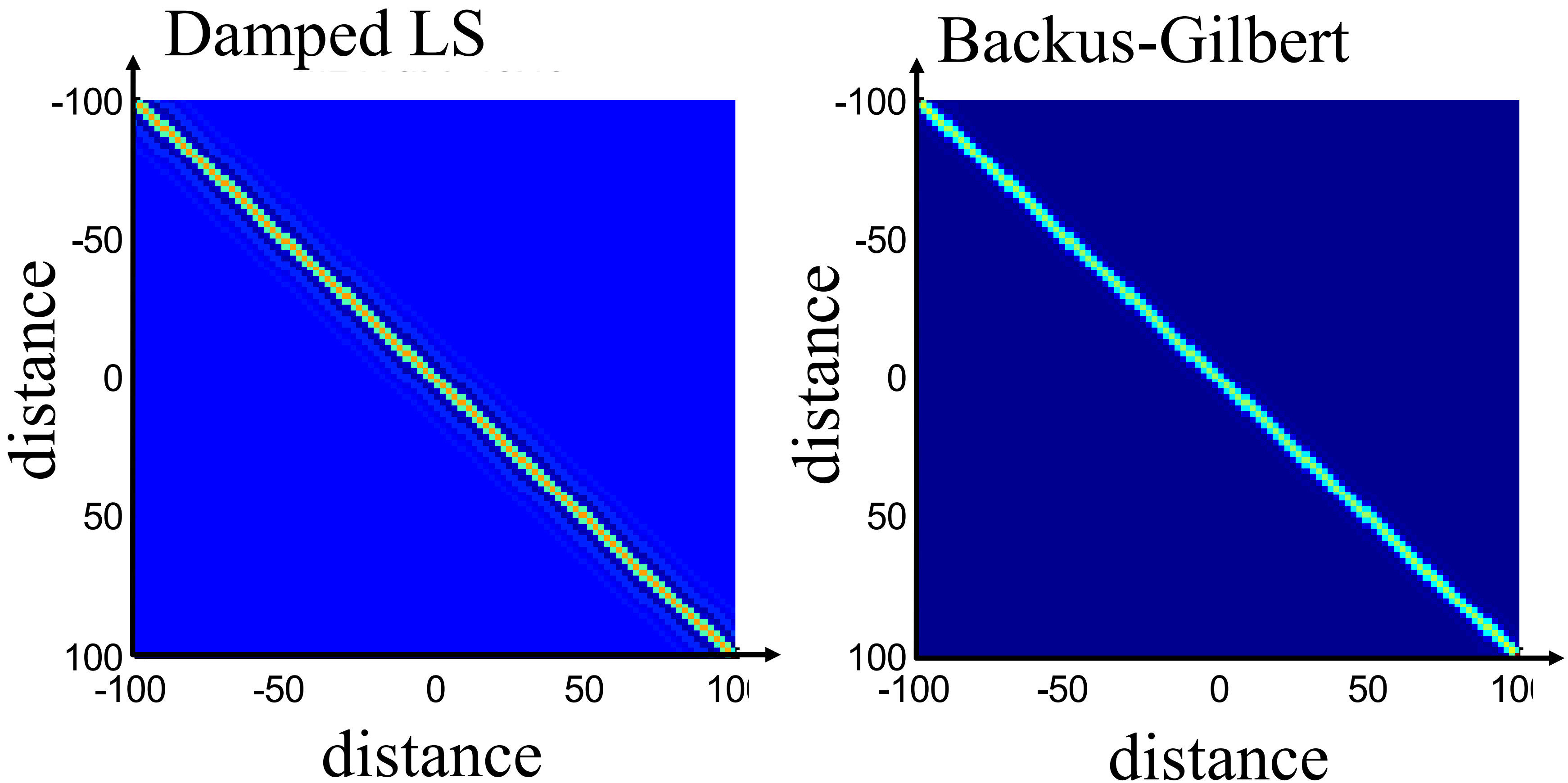
estimated initial temperature distribution as a function of the time of observation



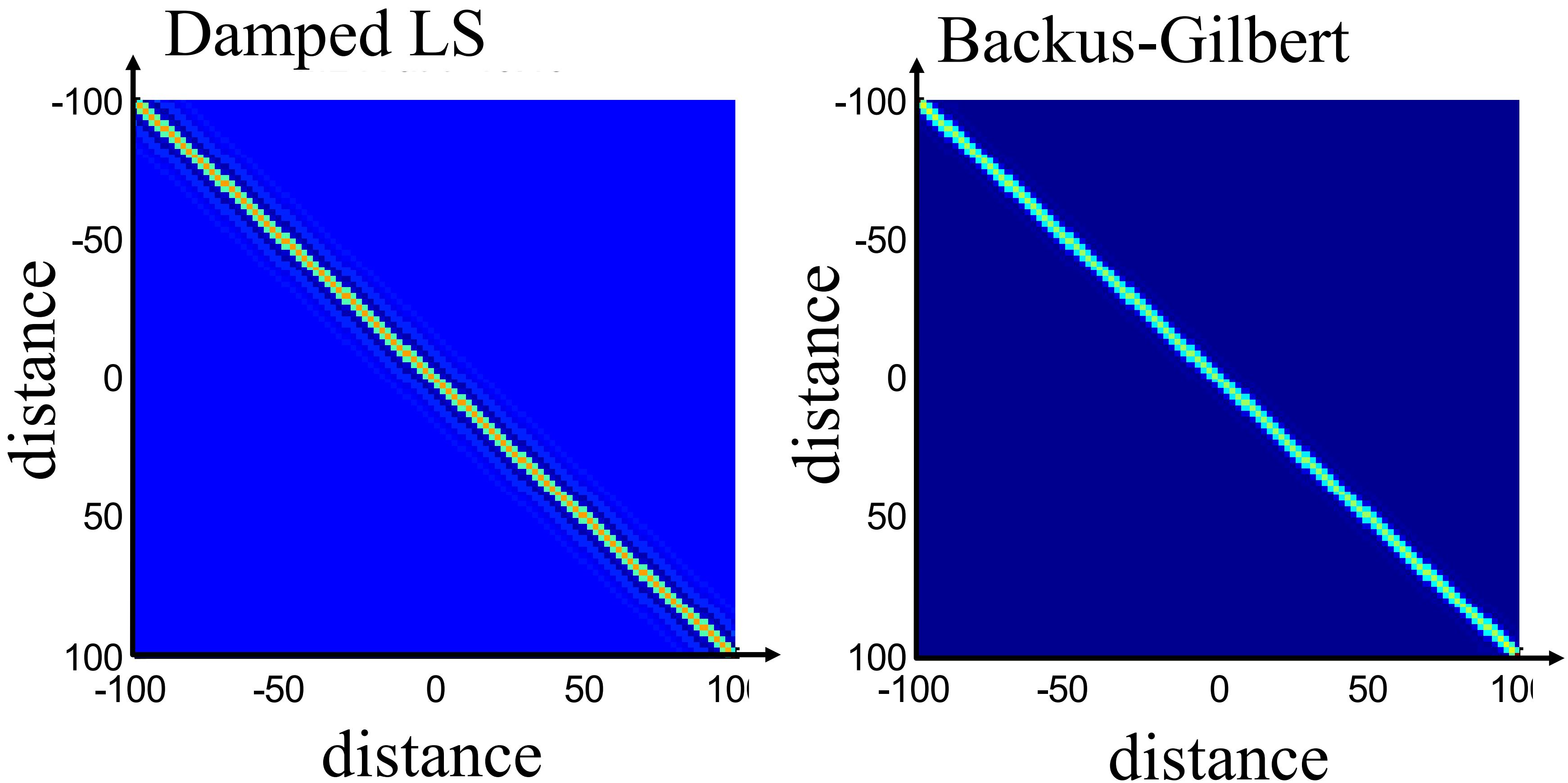
estimated initial temperature distribution as a function of the time of observation



model resolution matrix when for data collected at t=10

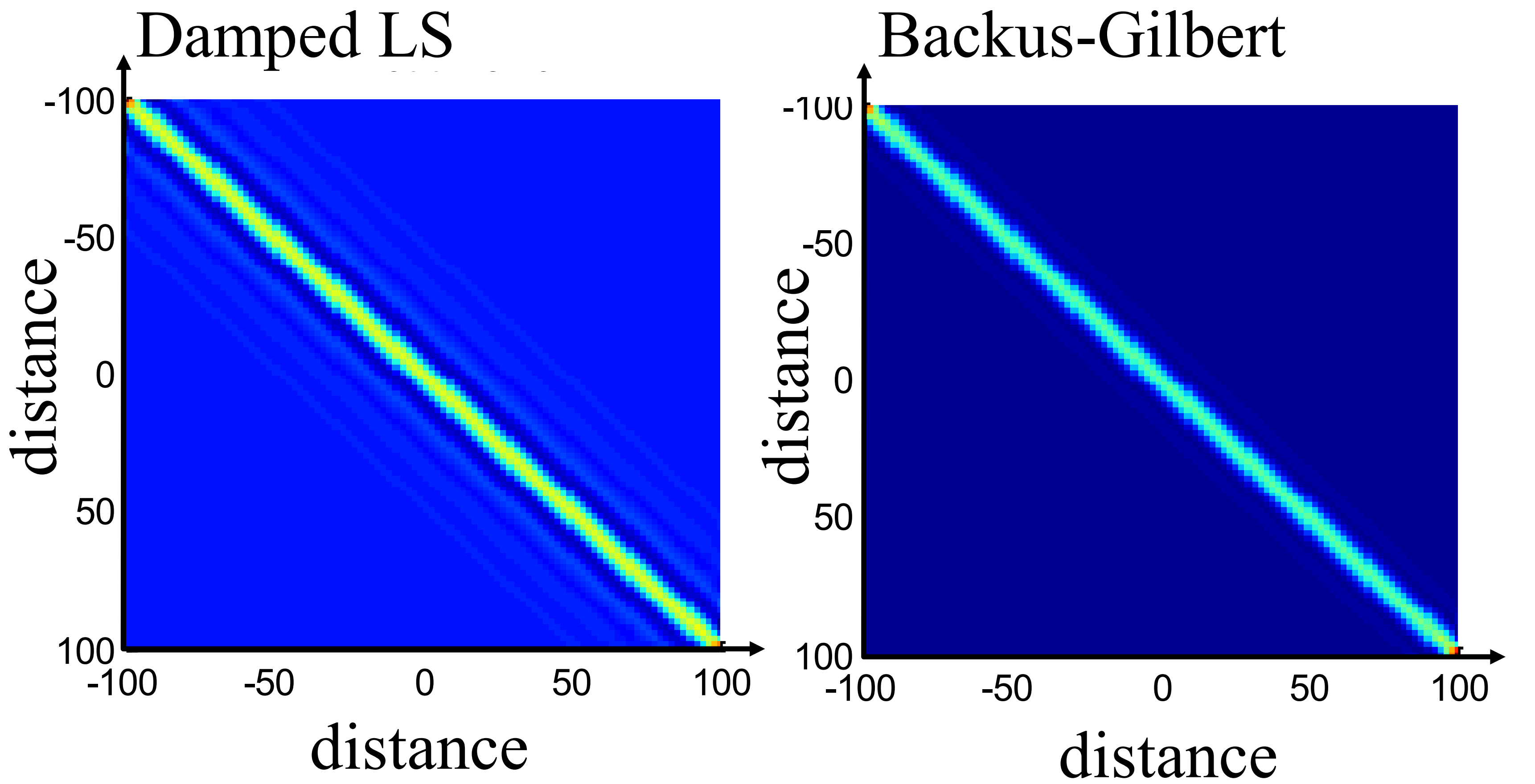


model resolution matrix when for data collected at t=10

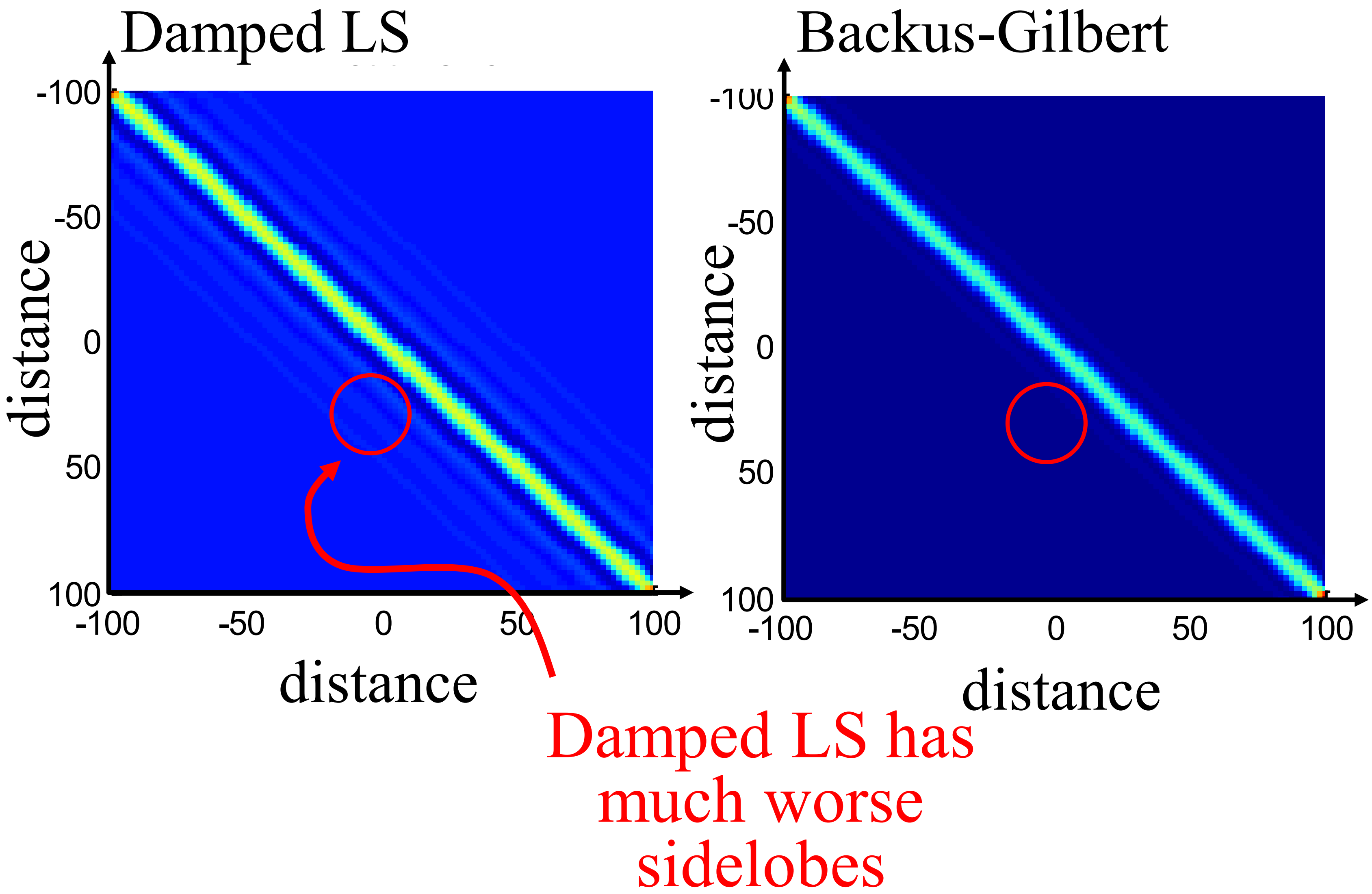


resolution
is similar

model resolution matrix when for data collected at $t=40$



model resolution matrix when for data collected at t=40

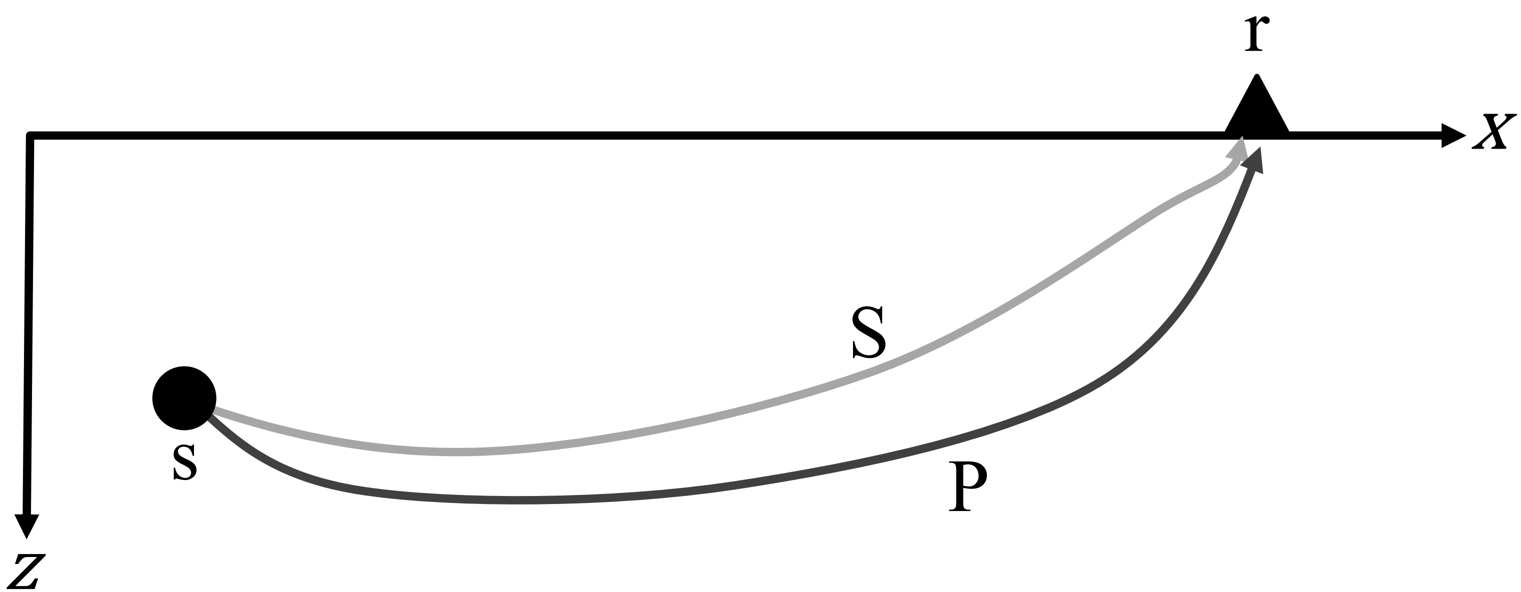


Part 2

earthquake location

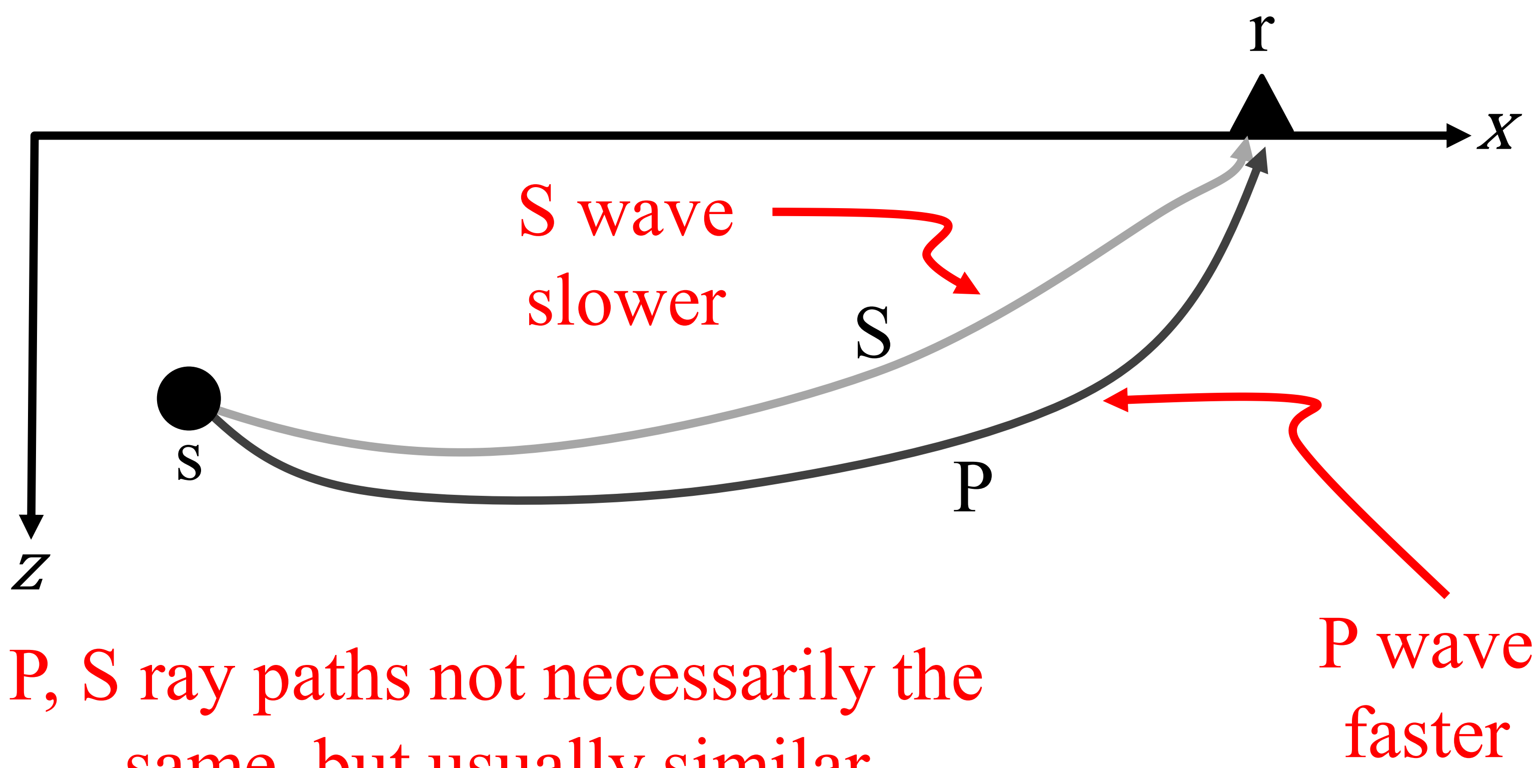
ray approximation

vibrations travel from source to receiver along curved rays



ray approximation

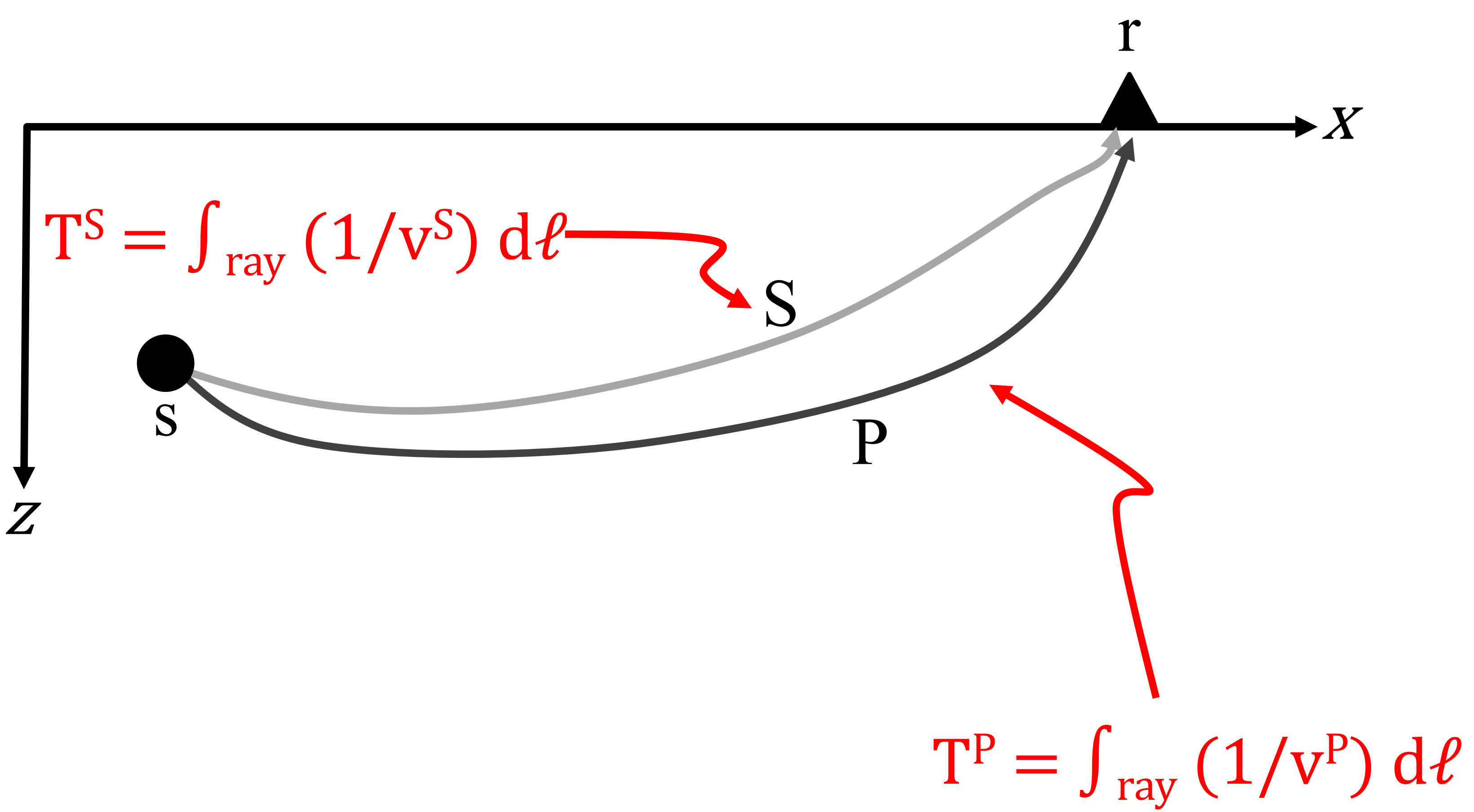
vibrations travel from source to receiver along curved rays



P, S ray paths not necessarily the same, but usually similar

travel time T

integral of slowness along ray path



arrival time = travel time along ray + origin time

$$t_i^P = T_i^P\left(\mathbf{x}^{(0)}, \mathbf{x}^{(\mathrm{i})}\right) + t_0 \quad \text{and} \quad t_i^S = T_i^S\left(\mathbf{x}^{(0)}, \mathbf{x}^{(\mathrm{i})}\right) + t_0$$

arrival time = travel time along ray + origin time

$$t_i^P = T_i^P\left(\mathbf{x}^{(0)}, \mathbf{x}^{(\mathrm{i})}\right) + t_0 \quad \text{and} \quad t_i^S = T_i^S\left(\mathbf{x}^{(0)}, \mathbf{x}^{(\mathrm{i})}\right) + t_0$$

data

data

earthquake location 3 model parameters

earthquake origin time 1 model parameter

arrival time = travel time along ray + origin time

$$t_i^P = T_i^P\left(\mathbf{x}^{(0)}, \mathbf{x}^{(\mathrm{i})}\right) + t_0 \quad \text{and} \quad t_i^S = T_i^S\left(\mathbf{x}^{(0)}, \mathbf{x}^{(\mathrm{i})}\right) + t_0$$

explicit nonlinear equation

4 model parameters
up to 2 data per station

arrival time = travel time along ray + origin time

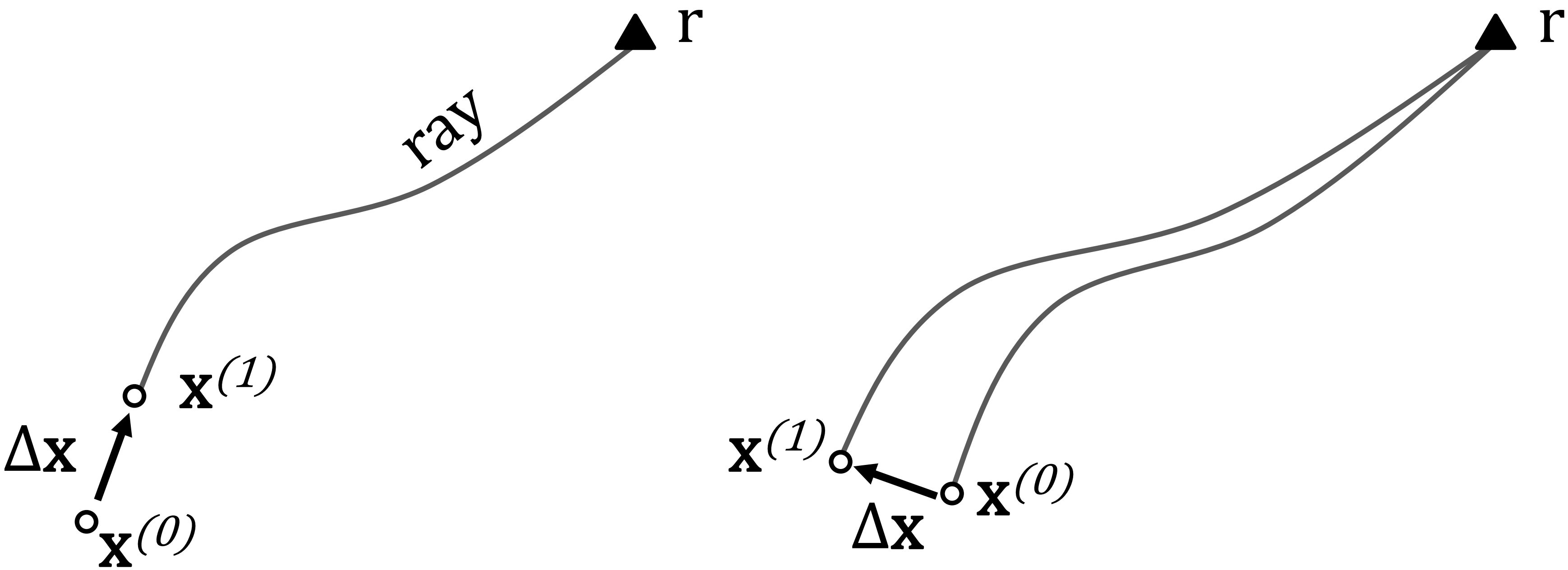
$$t_i^P = T_i^P\left(\mathbf{x}^{(0)}, \mathbf{x}^{(\mathrm{i})}\right) + t_0$$

linearize around trial source location $\mathbf{x}^{(\mathrm{p})}$

$$\mathrm{t_i}^{\mathrm{P}} = \mathrm{T_i}^{\mathrm{P}}(\mathbf{x}^{(\mathrm{p})},\mathbf{x}^{(\mathrm{i})}) + [\nabla \mathrm{T_i}^{\mathrm{P}}] \bullet \Delta \mathrm{x} + \mathrm{t_0}$$

trick is computing this gradient

Geiger's principle



$$[\nabla T_i^P] = -\mathbf{s}/v$$

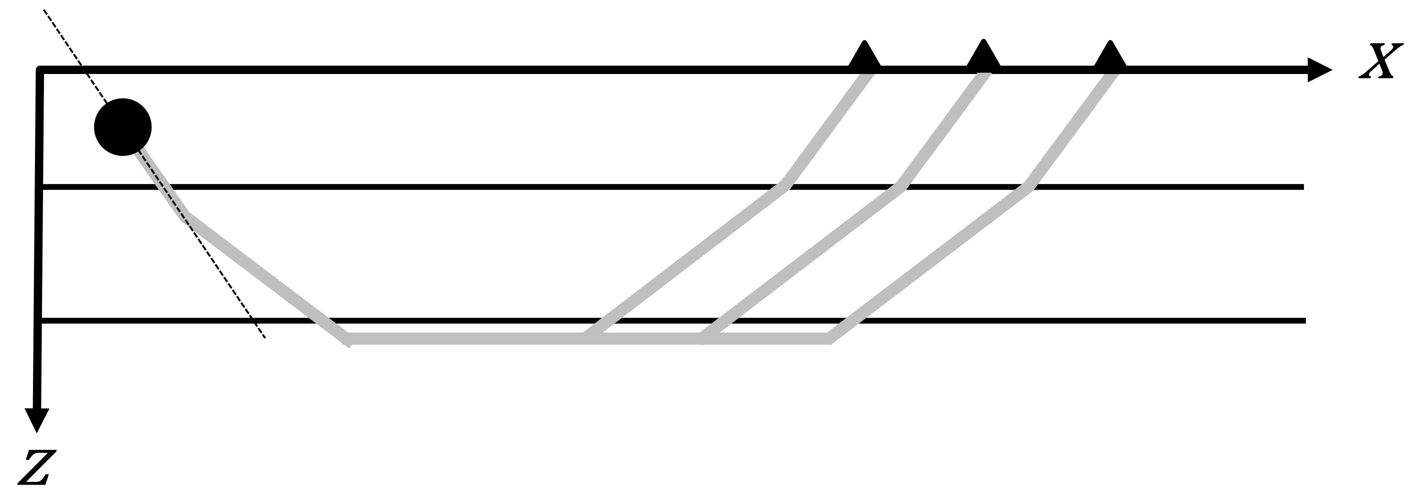
unit vector parallel to ray pointing away from receiver

linearized equation

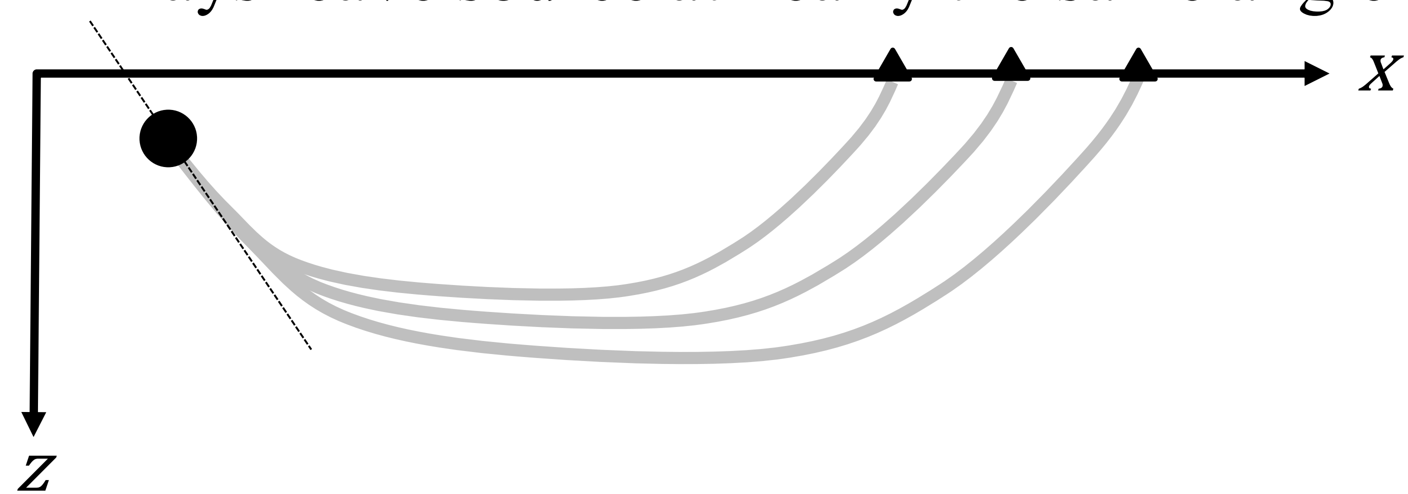
$$\begin{bmatrix} -s_1^{(1)P}/v_0^P & -s_2^{(1)P}/v_0^P & -s_3^{(1)P}/v_0^P & 1 \\ \cdots & \cdots & \cdots & \cdots \\ -s_1^{(N)P}/v_0^P & -s_2^{(N)P}/v_0^P & -s_3^{(N)P}/v_0^P & 1 \\ -s_1^{(1)S}/v_0^S & -s_2^{(1)S}/v_0^S & -s_3^{(1)S}/v_0^S & 1 \\ \cdots & \cdots & \cdots & \cdots \\ -s_1^{(N)S}/v_0^S & -s_2^{(N)S}/v_0^S & -s_2^{(N)S}/v_0^S & 1 \end{bmatrix} \begin{bmatrix} \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \\ t_0 \end{bmatrix} = \begin{bmatrix} t_1^P - T_1^P \\ \cdots \\ t_N^P - T_N^P \\ t_1^S - T_1^S \\ \cdots \\ t_N^S - T_N^S \end{bmatrix}$$

Common circumstances when earthquake far from stations

All rays leave source at the same angle



All rays leave source at nearly the same angle



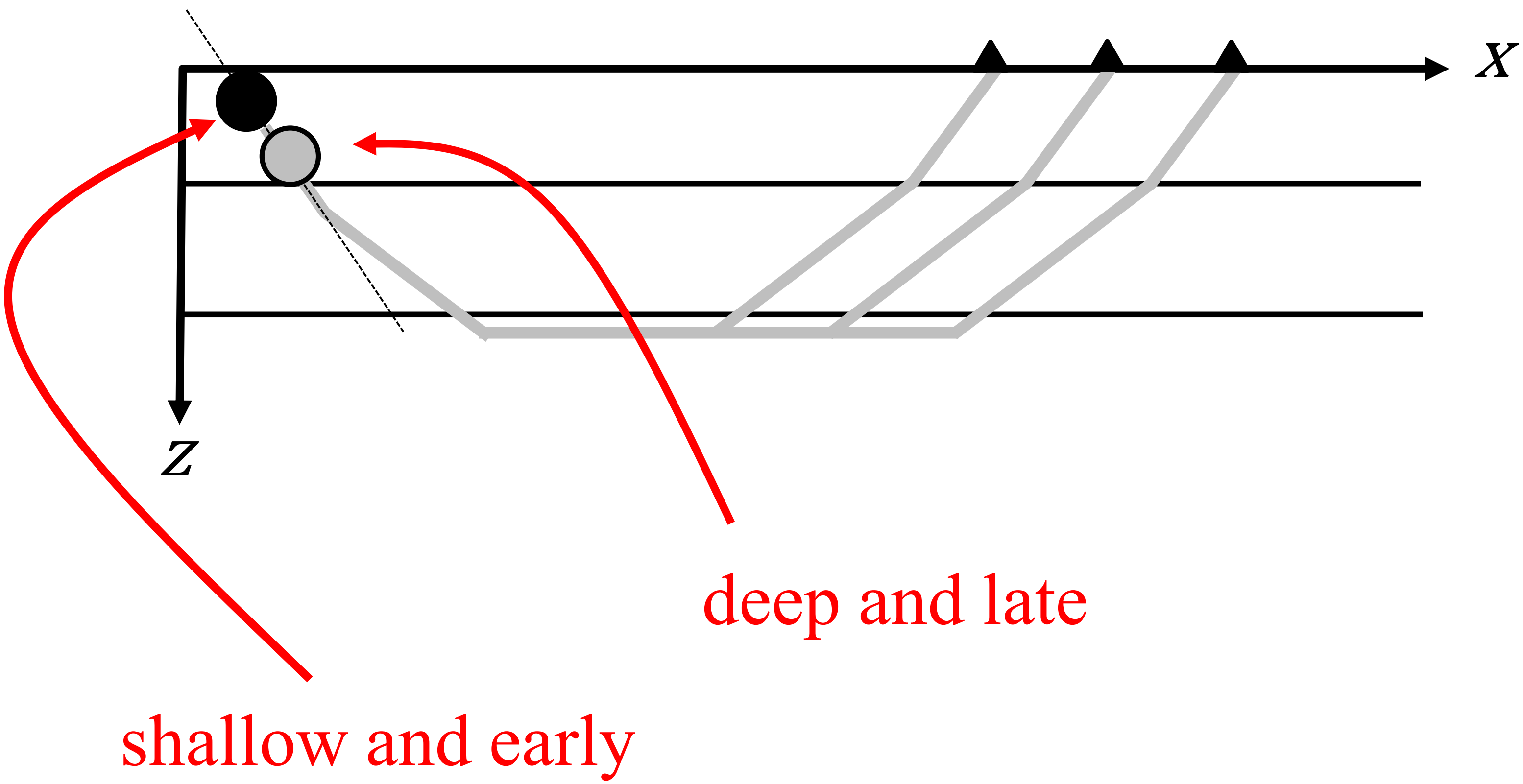
then, if only P wave data is available

$$\begin{bmatrix} -s_1^{(1)P}/v_0^P & -s_2^{(1)P}/v_0^P & -s_3^{(1)P}/v_0^P & 1 \\ \cdots & \cdots & \cdots & \cdots \\ -s_1^{(N)P}/v_0^P & -s_2^{(N)P}/v_0^P & -s_3^{(N)P}/v_0^P & 1 \\ . & & & \\ & & & . \\ . & & & \end{bmatrix} \begin{bmatrix} \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \\ t_0 \end{bmatrix} = \begin{bmatrix} t_1^P - T_1^P \\ \cdots \\ t_N^P - T_N^P \\ t_1^S - T_1^S \\ \cdots \\ t_N^S - T_N^S \end{bmatrix}$$

(no S waves)

these two columns are proportional to one-another

depth and origin time trade off

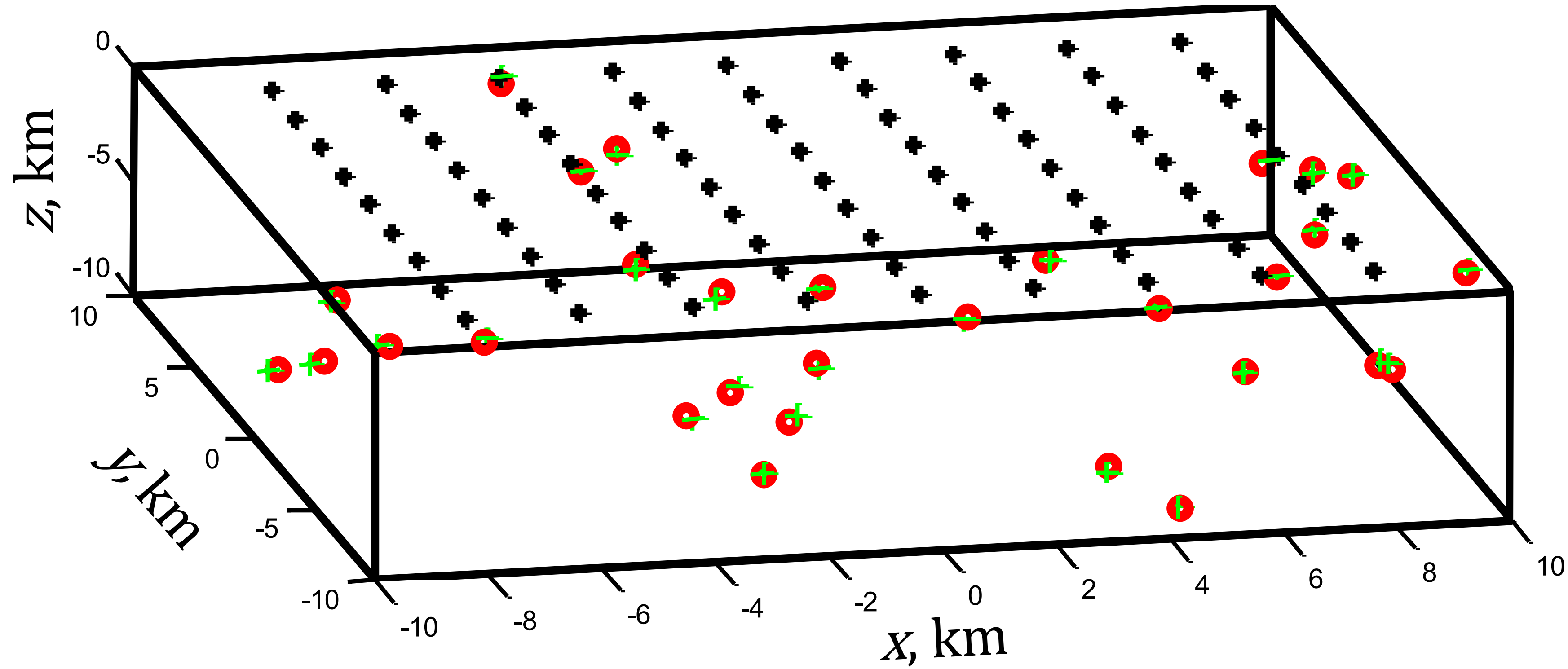


Solution Possibilities

1. Damped Least Squares:
   Matrix **G** is not sparse
   no analytic version of $\mathbf{G}^T\mathbf{G}$ is available
   M=4 is tiny
   experiment with values of ε^2
   `mest=(G'*G+e2*eye(M,M))\(G'*d)`

test case has earthquakes "inside of array"

2. Singular Value Decomposition
   to detect case of depth and origin time trading off

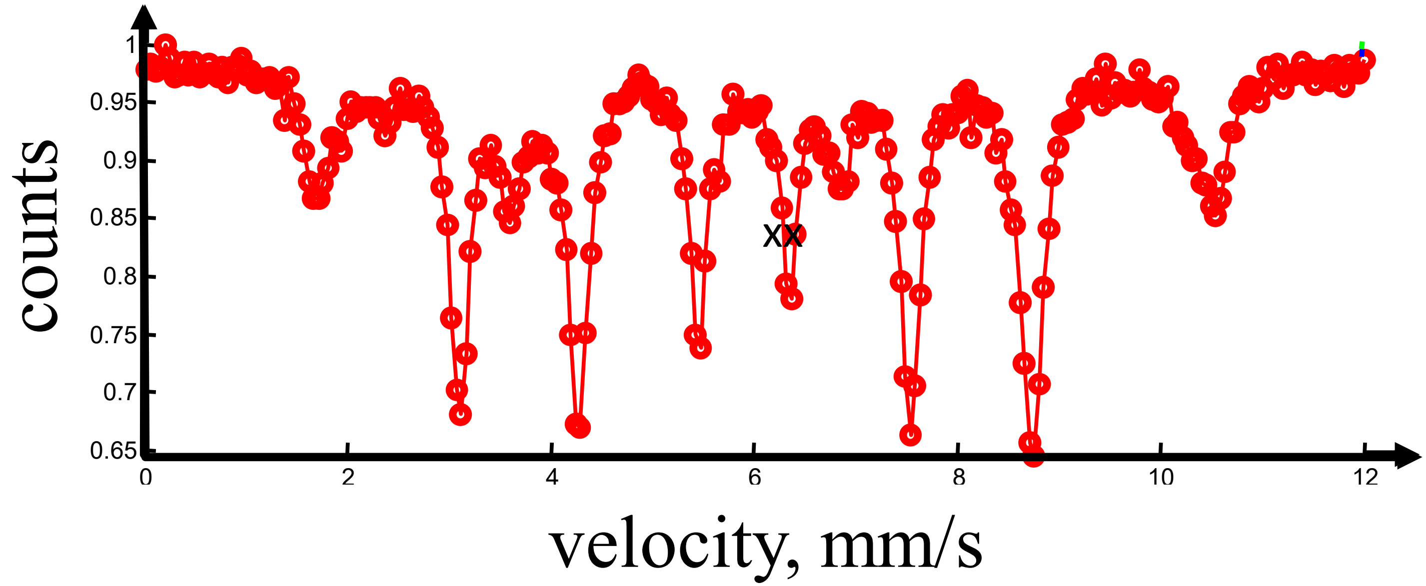


0
-5
-10
z, km
10
5
0
-5
-10
y, km
-10
-8
-6
-4
-2
0
2
4
6
8
10
x, km

Part 3

fitting of spectral peaks

typical spectrum consisting of overlapping peaks



what shape are the peaks?

Gaussian: $$d_i = \sum_{j=1}^{q} \frac{A_j}{(2\pi)^{½} c_j} \exp\left[-\frac{\left(z_i - f_j\right)^2}{2c_j^2}\right]$$

Lorentzian: $$d_i = \sum_{j=1}^{q} \frac{A_j c_j^2}{\left(z_i - f_j\right)^2 + c_j^2}$$

what shape are the peaks?

Gaussian: $$d_i = \sum_{j=1}^{q} \frac{A_j}{(2\pi)^{1/2} c_j} \exp\left[-\frac{\left(z_i - f_j\right)^2}{2c_j^2}\right]$$

Lorentzian: $$d_i = \sum_{j=1}^{q} \frac{A_j c_j^2}{\left(z_i - f_j\right)^2 + c_j^2}$$

try both

use *F* test to test whether one is better than the other

what shape are the peaks?

3 unknowns per peak

data

Gaussian: $$d_i = \sum_{j=1}^{q} \frac{A_j}{(2\pi)^{1/2} c_j} \exp\left[-\frac{(z_i - f_j)^2}{2c_j^2}\right]$$

data

Lorentzian: $$d_i = \sum_{j=1}^{q} \frac{A_j c_j^2}{(z_i - f_j)^2 + c_j^2}$$

3 unknowns per peak

both cases:
explicit nonlinear problem

linearize using analytic gradient

Gaussian:

$$\partial g_i/\partial A_j = \left[1/(2\pi)^{1/2} c_j\right] \exp\left[-(z_i - f_j)^2/2c_j^2\right]$$

$$\partial g_i/\partial f_j = \left[\frac{A_j}{(2\pi)^{1/2} c_j}\right]\left[(z_i - f_j)/c_j^2\right] \exp\left[-(z_i - f_j)^2/2c_j^2\right]$$

$$\partial g_i/\partial c_j = \left[\frac{A_j}{(2\pi)^{1/2} c_j^2}\right]\left[\left((z_i - f_j)^2/c_j^2\right) - 1\right] \exp\left[-(z_i - f_j)^2/2c_j^2\right]$$

linearize using analytic gradient

Lorentzian:

$$\partial g_i/\partial A_j = c_j^2/\left[(z_i - f_j)^2 + c_j^2\right]$$

$$\partial g_i/\partial f_j = 2A_j c_j^2 (z_i - f_j)/\left[(z_i - f_j)^2 + c_j^2\right]^2$$

$$\partial g_i/\partial c_j = 2A_j c_j/\left[(z_i - f_j)^2 + c_j^2\right] - 2A_j c_j^3/\left[(z_i - f_j)^2 + c_j^2\right]^2$$

issues

how to determine

number q of peaks

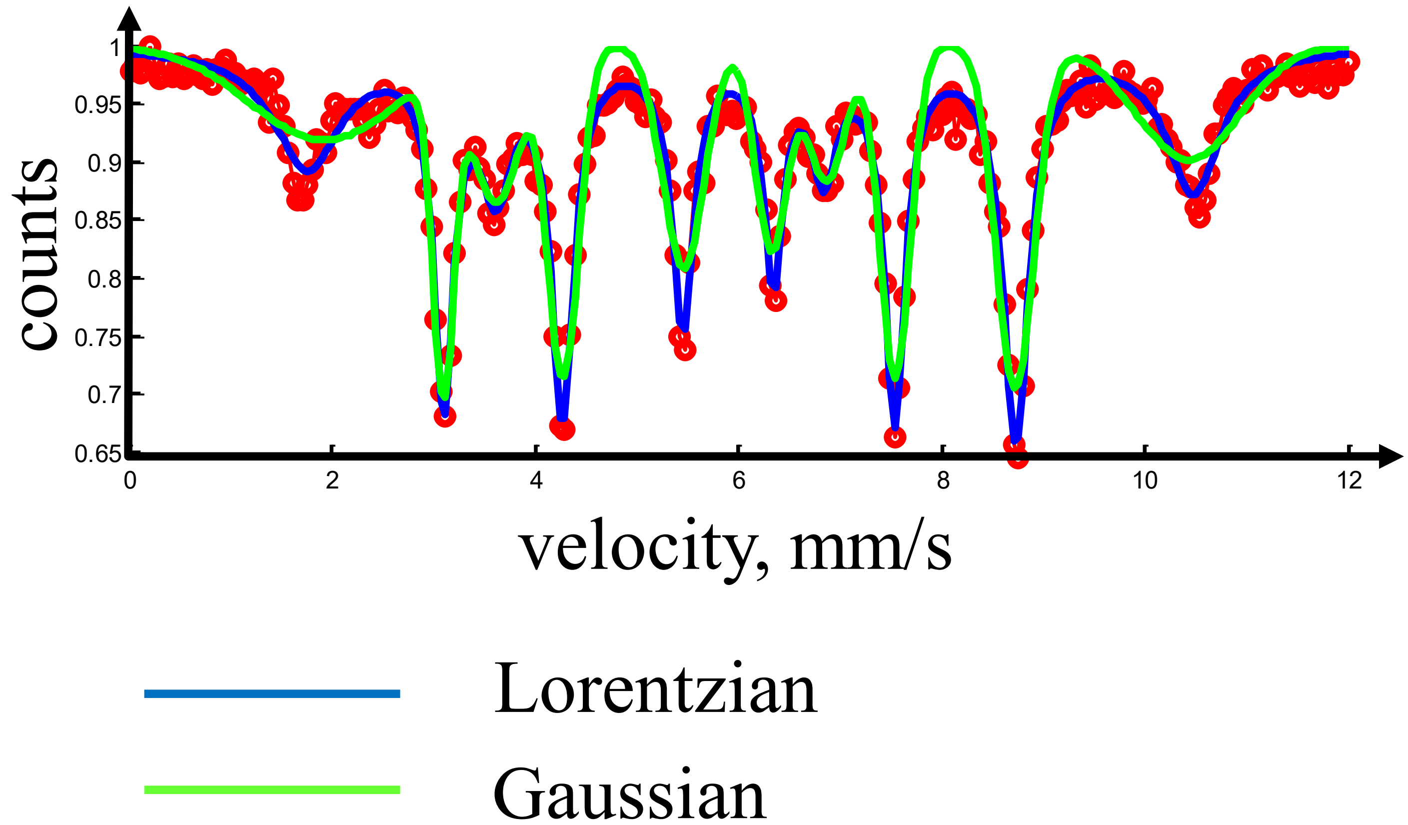
trial A_i c_i f_i of each peak

our solution

have operator click mouse
computer screen
to indicate position of each peak

MatLab code for graphical input

```
K=0;
for k = [1:20]
    p = ginput(1);
    if( p(1) < 0 )
        break;
    end
    K=K+1;
    a(K) = p(2)-A;
    v0(K)=p(1);
    c(K)=0.1;
end
```


counts
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0
2
4
6
8
10
12
velocity, mm/s
Lorentzian
Gaussian

Results of F test

```
Fest = E_normal/E_lorentzian: 4.230859
P(F<=1/Fest||F>=Fest) = 0.000000
```

Lorentzian better fit
to 99.9999% certainty