CAPON's f-k estimator (1967,1969)
a.k.a.
high-resolution f-k
maximum likelihood f-k estimate
minimum variance distortionless look estimate

Recalling signal model and conventional beamforming

signal model $$X_i(\omega) = S(\omega)\exp(j\vec{k}_0\vec{r}_i) + n_i(\omega)$$

shift and sum $$B(\omega,\vec{k}) = \sum_{i=1}^{N} X_i(\omega)\exp(-j\vec{k}\vec{r}_i)$$

$$B(\omega,\vec{k}) = \sum_{i=1}^{N} \left[S(\omega)\exp(j\vec{k}_0\vec{r}_i)\exp(-j\vec{k}\vec{r}_i) + n_i(\omega)\exp(-j\vec{k}\vec{r}_i)\right]$$

$$B(\omega,\vec{k}) = \sum_{i=1}^{N} \left[S(\omega)\exp(j(\vec{k}_0 - \vec{k})\vec{r}_i) + \tilde{n}_i(\omega)\right]$$

with $$\tilde{n}_i(\omega) = n_i(\omega)\exp(-j\vec{k}\vec{r}_i)$$

discrete spatial sampling of a **continuous process**

consequences: **aliasing (sampling theorem)**

at least 3 samples per period, wavelength

time domain $\Delta T < T_{min}/2$

spatial domain $\Delta x < \lambda^{\star}_{min}/2$ *** apparent**

spectral resolution limit

time domain $\Delta\omega = 2\pi/((N-1)\Delta T)$

spatial domain $\Delta k = 2\pi/((N-1)d_{min}) = 2\pi/D_{max}$

discrete spatial sampling of a continuous process

So far we dicussed as consequences:

aliasing (sampling theorem)

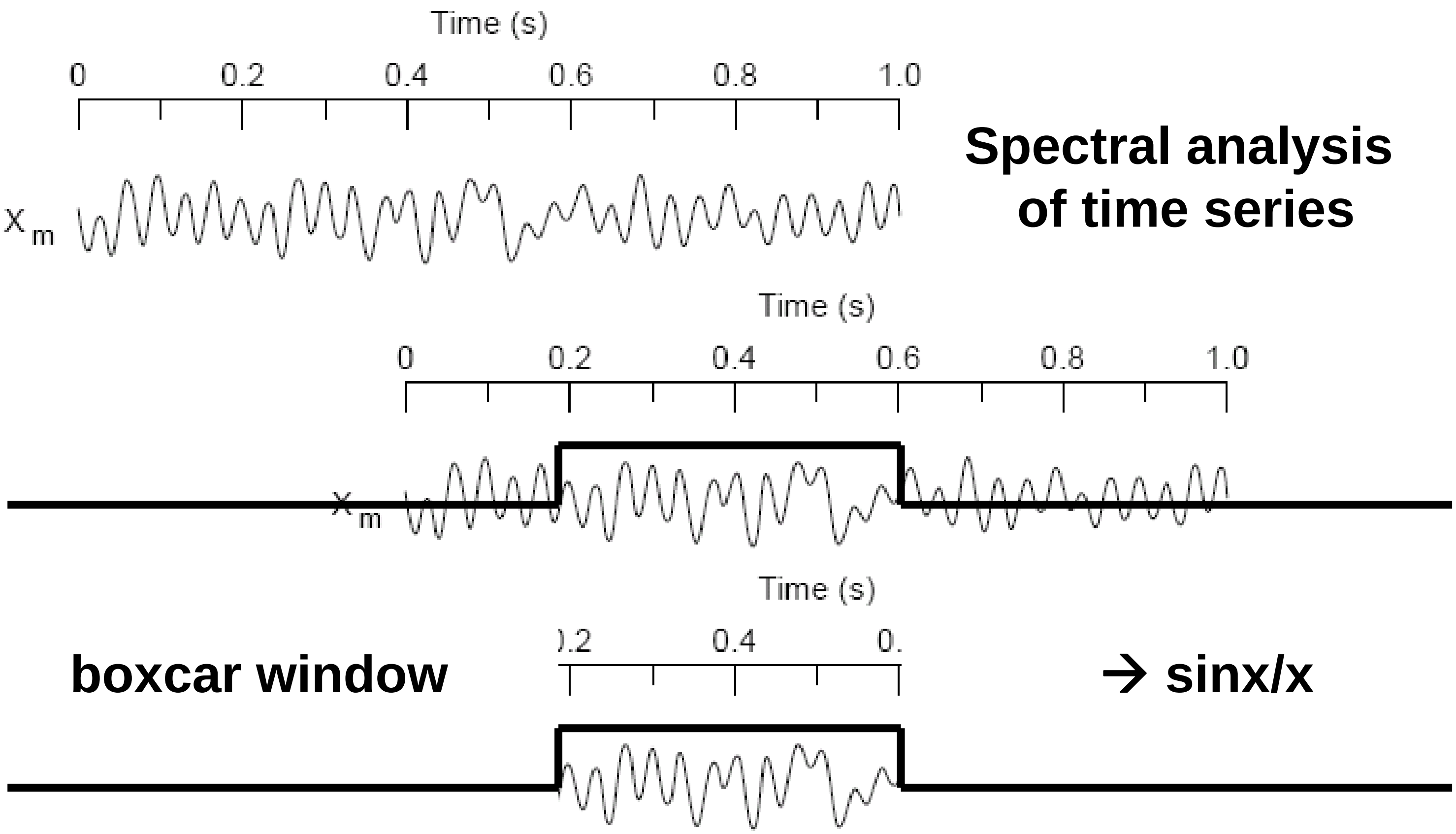
spectral resolution limit

No comment so far to the typically observed shape:

e.g. for 1D-equidistant array typical sinx/x
Where does this come from?

from continuous infinite to finite discrete...

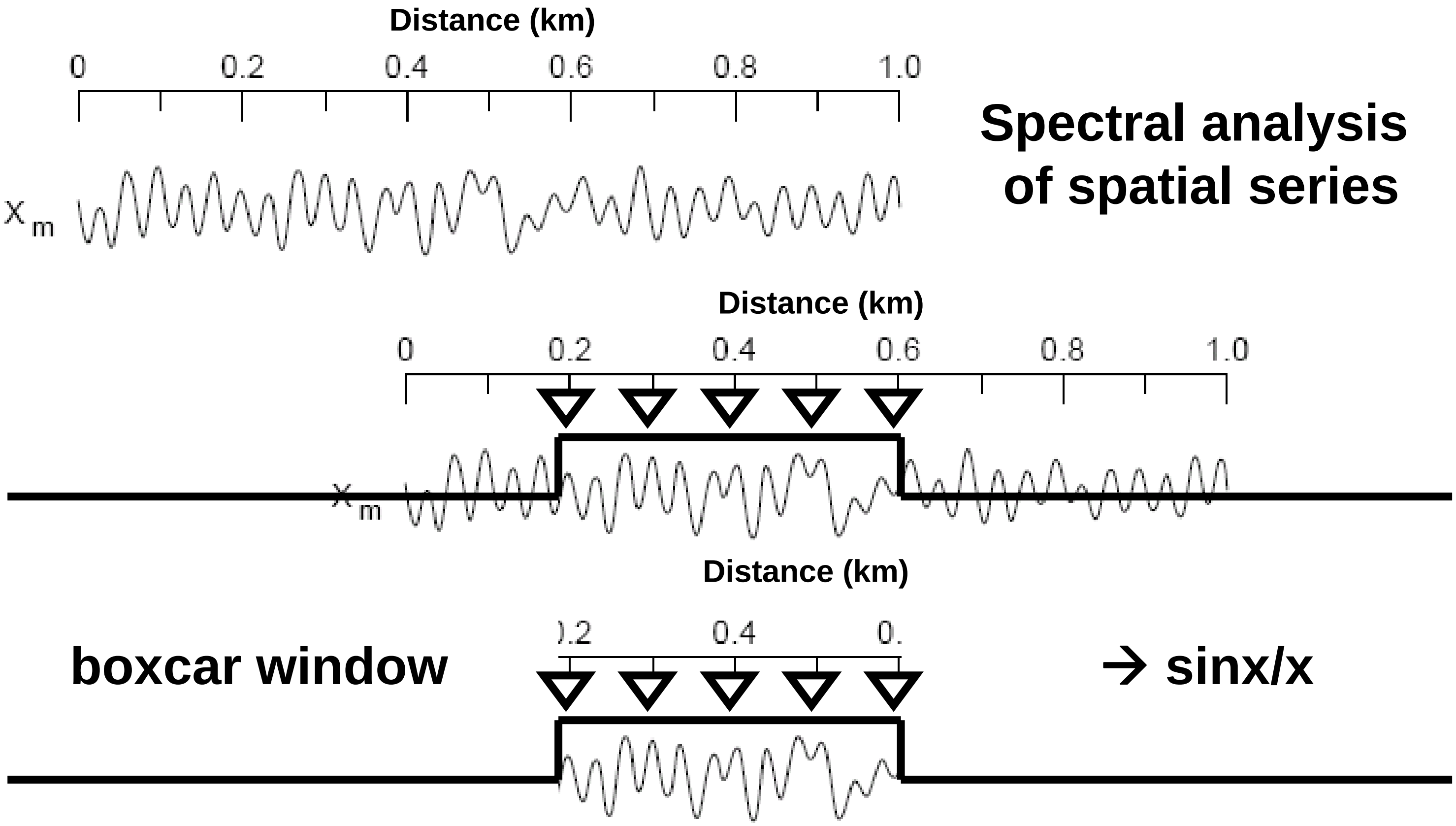
Spectral analysis of time series



boxcar window

→ sinx/x

from continuous infinite to finite discrete...

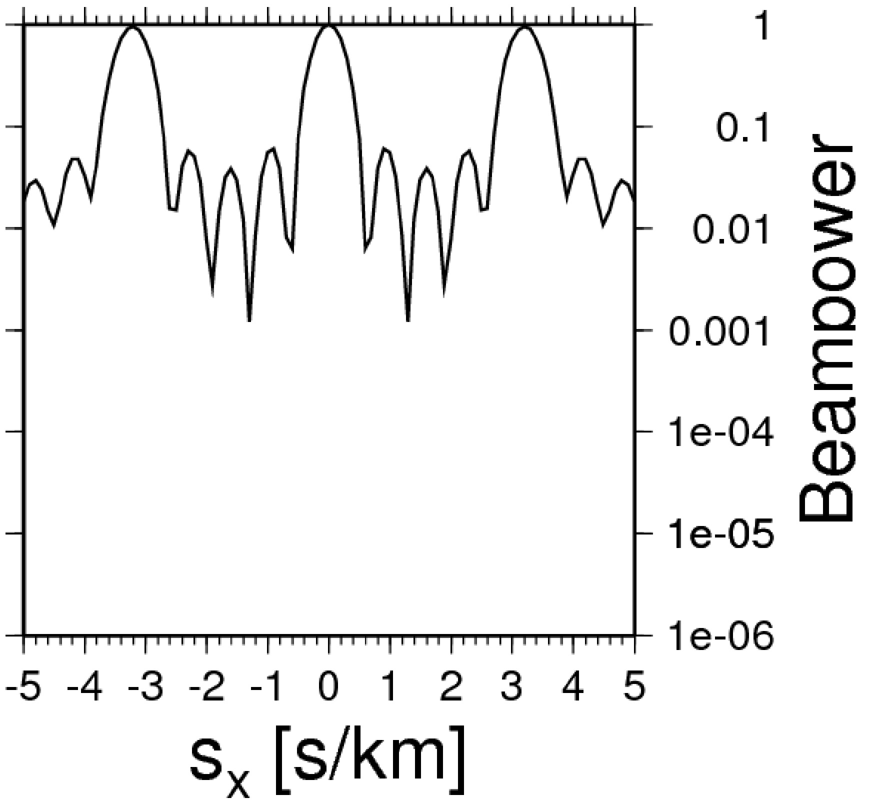


Spectral analysis of spatial series

boxcar window

→ sinx/x

from continuous infinite to finite discrete...



sinx/x shape in spectral domain (wavenumber)

Introduce sensor weights as spatial taper equivalent to windowing function in time series spectral analysis!

$$B(\omega,\vec{k}) = \sum_{i=1}^{N} X_i(\omega)\exp(-j\vec{k}\vec{r}_i)$$

Generalization of array output: weighted shift and sum

$$B(\omega,\vec{k}) = \sum_{i=1}^{N} W_i(\omega) X_i(\omega)\exp(-j\vec{k}\vec{r}_i)$$

Introducing matrix notation for compact writing

$$\vec{X}(\omega) = \begin{bmatrix} X_1(\omega) \\ X_2(\omega) \\ X_3(\omega) \\ \vdots \\ X_N(\omega) \end{bmatrix} \qquad \vec{A}(\omega) = \begin{bmatrix} W_1(\omega)\exp(j\vec{k}\vec{r}_1) \\ W_2(\omega)\exp(j\vec{k}\vec{r}_2) \\ W_3(\omega)\exp(j\vec{k}\vec{r}_3) \\ \vdots \\ W_N(\omega)\exp(j\vec{k}\vec{r}_N) \end{bmatrix}$$

$$B(\omega,\vec{k}) = \sum_{i=1}^{N} W_i(\omega)X_i(\omega)\exp(-j\vec{k}\vec{r}_i)$$

$$\longrightarrow \quad B(\omega,\vec{k}) = \vec{A}(\omega)^H \vec{X}(\omega)$$

From beam (array output) to beampower ...

$$B(\omega,\vec{k}) = \vec{A}(\omega)^H \vec{X}(\omega)$$

$$\left|B(\omega,\vec{k})\right|^2 = \vec{A}(\omega)^H \vec{X}(\omega)(\vec{A}(\omega)^H \vec{X}(\omega))^\star = \vec{A}(\omega)^H \underbrace{\vec{X}(\omega)\vec{X}(\omega)^H}\vec{A}(\omega)$$

using abbreviation $\longrightarrow$ $\underline{R}(\omega) = \vec{X}(\omega)\vec{X}(\omega)^H$

$$\left|B(\omega,\vec{k})\right|^2 = \vec{A}(\omega)^H \underline{R}(\omega)\vec{A}(\omega)$$

= generalized beamformer

abbreviated quantity ...

$$\underline{R}(\omega) = \vec{X}(\omega)\vec{X}(\omega)^H$$

is the primary source of information for all f-k estimates

$\underline{R}(\omega)$ **has many names:**

cross spectral matrix
spatio-spectral matrix
spatial correlation matrix
sensor covariance matrix....

Capon's method (1969)

Based on the formulation of the generalized beamformer

$$\left|B(\omega,\vec{k})\right|^2 = \vec{A}(\omega)^H \underline{R}(\omega)\vec{A}(\omega)$$

Capon's significant contribution consists now in the following idea (and of course in the consistent mathematical formulation of a solution to it):

Find optimum weights for the generalized beamformer which provide an f-k estimate which has unity gain at the true wavenumber and is minimized elsewhere!

Find optimum weights for the generalized beamformer which provide an f-k estimate which has unity gain at the true wavenumber and is minimized elsewhere!
Put into other words: f-k estimate should approximate at best a 3D delta function (in kx, ky and f) or overall cross spectral power is to be minimized:

$$\underline{W R W^H} \longrightarrow \text{to be minimized}$$

Further requirement:

Array output equals observation for true wavenumber

Array output equals observation for true wavenumber

... formulated in maths ...

$$W_i(\omega) X_i(\omega) \exp(-j\vec{k}\vec{r}_i) = X_i(\omega)$$

... leads to the condition ...

$$\vec{W}\vec{A}'^T = 1$$

$$\vec{W}(\omega) = \begin{bmatrix} W_1(\omega) \\ W_2(\omega) \\ W_3(\omega) \\ \vdots \\ W_N(\omega) \end{bmatrix} \qquad \vec{A}'(\omega) = \begin{bmatrix} \exp(j\vec{k}\vec{r}_1) \\ \exp(j\vec{k}\vec{r}_2) \\ \exp(j\vec{k}\vec{r}_3) \\ \vdots \\ \exp(j\vec{k}\vec{r}_N) \end{bmatrix}$$

Problem to be solved:

Minimize $\underline{W\,R\,W}^{H}$ **with constraint** $\vec{W}\vec{A}'^{T} = 1$

A minimization task with subject to constraints can be solved using the formalism of Lagrangian multipliers

For the given problem, the following solution is obtained:

$$\underline{W} = \{\underline{R}^{-1}\underline{A}'^{T}\}/\{(\underline{A}'^{H})^{T}\underline{R}^{-1}\underline{A}'^{T}\}$$

Inserting this result in order to obtain a formulation for the f-k spectral estimate, leads to:

$$P_{Capon}(\omega, \vec{k}) = \frac{1}{\vec{A}'^{\star}\underline{R}^{-1}\vec{A}'}$$

NOTE:

The weights are adaptive and tune the shape of the spatial taper function in dependence on the wavenumber

$$\underline{W} = \{\underline{R}^{-1}\underline{A}'^{T}\}/\{(\underline{A}'^{H})^{T}\underline{R}^{-1}\underline{A}'^{T}\}$$

But: the weights don‘t have to be explicitely estimated!

$$P_{Capon}(\omega, \vec{k}) = \frac{1}{\vec{A}'^{\star}\underline{R}^{-1}\vec{A}'}$$

The information is contained in the structure of the cross spectral matrix

Capon's method put into practice:

- Estimate cross spectral matrix
- Invert cross spectral matrix
- Sweep over wavenumber space using trial steering vectors by applying:

$$P_{Capon}(\omega, \vec{k}) = \frac{1}{\vec{A}'^{\star} \underline{R}^{-1} \vec{A}'}$$

High-resolution or not? Yes, but not always...

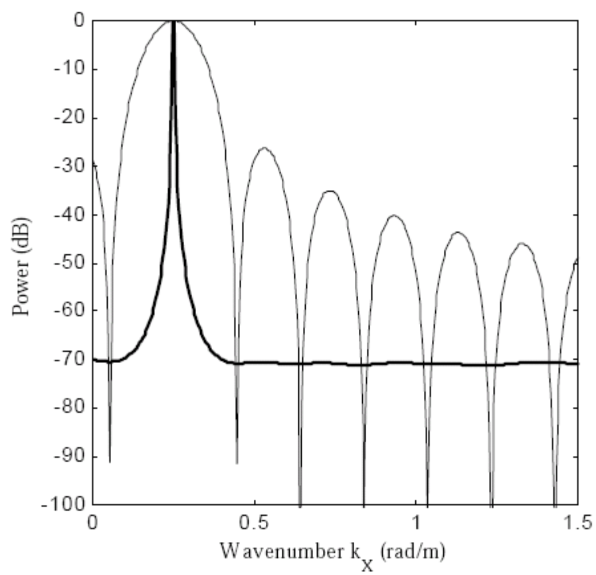


Figure 4.12 Power Output for the Minimum Variance Distortionless Look (MVDL) Method for a Single Wave at k_x = 0.25 rad/m Propagating Along the Main Axis of the 16 Sensor Uniform Linear Array. The MVDL output is shown with the dark line, and the FDBF output is shown with the light line for reference.

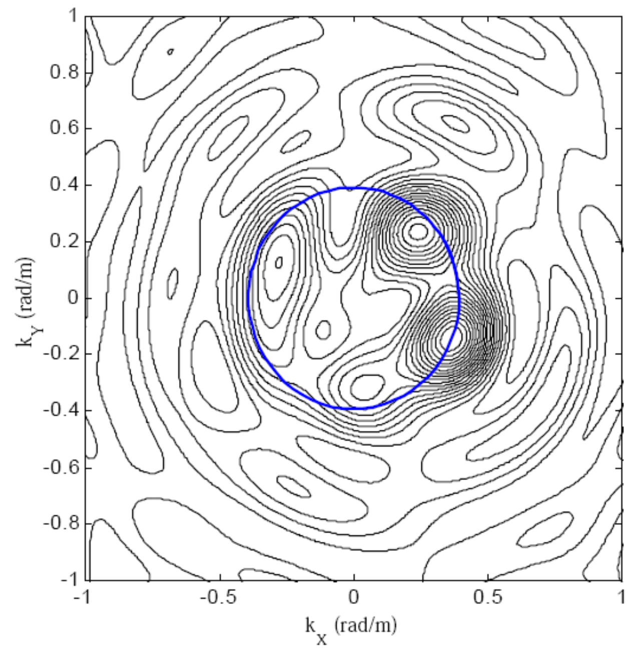


Figure 8.17 Example of Multiple Signals Arriving at a Frequency = 9.875 Hz.

from Zywicki, 1999, Ph.D. Thesis, http://www.zywicki.com

due to necessary stabilization of CSM inversion and ‚goodness' of CSM estimate from data!

Estimation of cross spectral matrix

howto?

$$\hat{R}(f) = \begin{pmatrix} \hat{X}_0(f)\hat{X}_0^*(f) & \hat{X}_0(f)\hat{X}_1^*(f) & \hat{X}_0(f)\hat{X}_2^*(f) & \cdots & \hat{X}_0(f)\hat{X}_{M-1}^*(f) \\ \hat{X}_1(f)\hat{X}_0^*(f) & \hat{X}_1(f)\hat{X}_1^*(f) & \hat{X}_1(f)\hat{X}_2^*(f) & \cdots & \hat{X}_1(f)\hat{X}_{M-1}^*(f) \\ \hat{X}_2(f)\hat{X}_0^*(f) & \hat{X}_2(f)\hat{X}_1^*(f) & \hat{X}_2(f)\hat{X}_2^*(f) & \cdots & \hat{X}_2(f)\hat{X}_{M-1}^*(f) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \hat{X}_{M-1}(f)\hat{X}_0^*(f) & \hat{X}_{M-1}(f)\hat{X}_1^*(f) & \hat{X}_{M-1}(f)\hat{X}_2^*(f) & \cdots & \hat{X}_{M-1}(f)\hat{X}_{M-1}^*(f) \end{pmatrix}$$

Matrix elements contain phase difference between sensors!

In order to obtain good estimates of the phase differences an averaging procedure is required! Further, the assumption of stationarity of the wavefield has to be introduced.

Variance reduction of cross spectral matrix estimates are usually obtained by the block averaging procedure:

$$(\underline{R})_{ij} = \frac{1}{M} \sum_{m=1}^{M} X_i(\omega) X_j(\omega)^{\star}$$

$$(\hat{\underline{R}})_{ij} = \frac{1}{M} \sum_{m=1}^{M} \frac{X_i(\omega) X_j(\omega)^{\star}}{\sqrt{X_i(\omega) X_i(\omega)^{\star}} \sqrt{X_j(\omega) X_j(\omega)^{\star}}}$$

block-averaging

number of blocks at least = number of stations!

**If not:
CSM may be singular!**

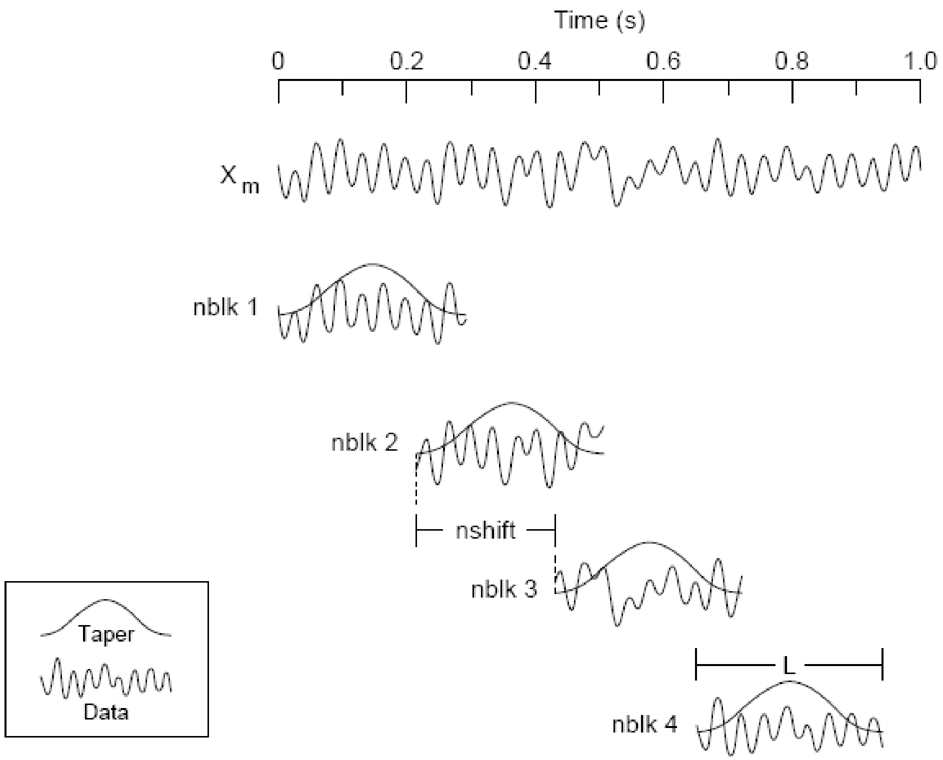


Figure 3. Segmentation and overlapping of a time-domain component of the signal vector. Each tapered block is transformed into the frequency domain via an FFT.

block-averaging in time domain equivalent to smoothing over some bandwidth in frequency domain

**as power spectrum computation
in spectral analysis of time series
(periodogram of frequency smoothing used
for variance reduction)**

In both cases large overall time windows are needed to obtain sufficient samples for averaging/(smoothing)

Implementation in geopsy: frequency smoothing

Stabilization of cross spectral matrix before inversion (avoiding singular CSM)

- **By averaging/smoothing procedure**

- **By diagonal loading**

$$\underline{\hat{R}} = (1-\lambda)\underline{\ddot{R}} + \lambda\underline{I}$$

- **By adding gaussian noise to all sensor observations**