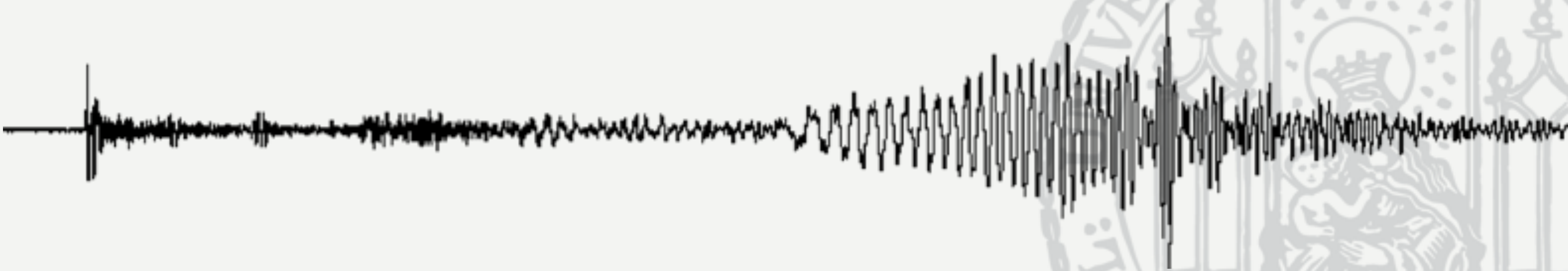
LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Ceri Nunn

Geophysical Data Analysis

L09 – Correlation



Scope

Scope: Understand how the correlation gives us information about how similar signals are, and where. ...

What is a correlation?

What is a correlation?

Scope

- *Correlation of time series*
  - *Similarity*
  - *Time shifts*

Example Applications

Brainstorming …

Applications?

… on the board.

Correlation Applications

- *Applications*
  - *Correlation of rotations/strains and translations*
  - *Ambient noise correlations*
  - *Accurate phase picking*
  - *Searching for signals (using a template)*
  - *Autocorrelations*

Correlation

Convolution:

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$

Cross-correlation:

$$(f \star g)(\tau) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f^*(t)\ g(t+\tau)\, dt,$$

where f_* denotes the complex conjugate of f, t is the time and τ is the lag.

Correlation

Convolution:

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$

Cross-correlation:

$$(f \star g)(\tau) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f^*(t)\; g(t+\tau)\, dt,$$

The cross-correlation of functions f(t) and g(t) is equivalent to the convolution of f(–t) and g(t):*

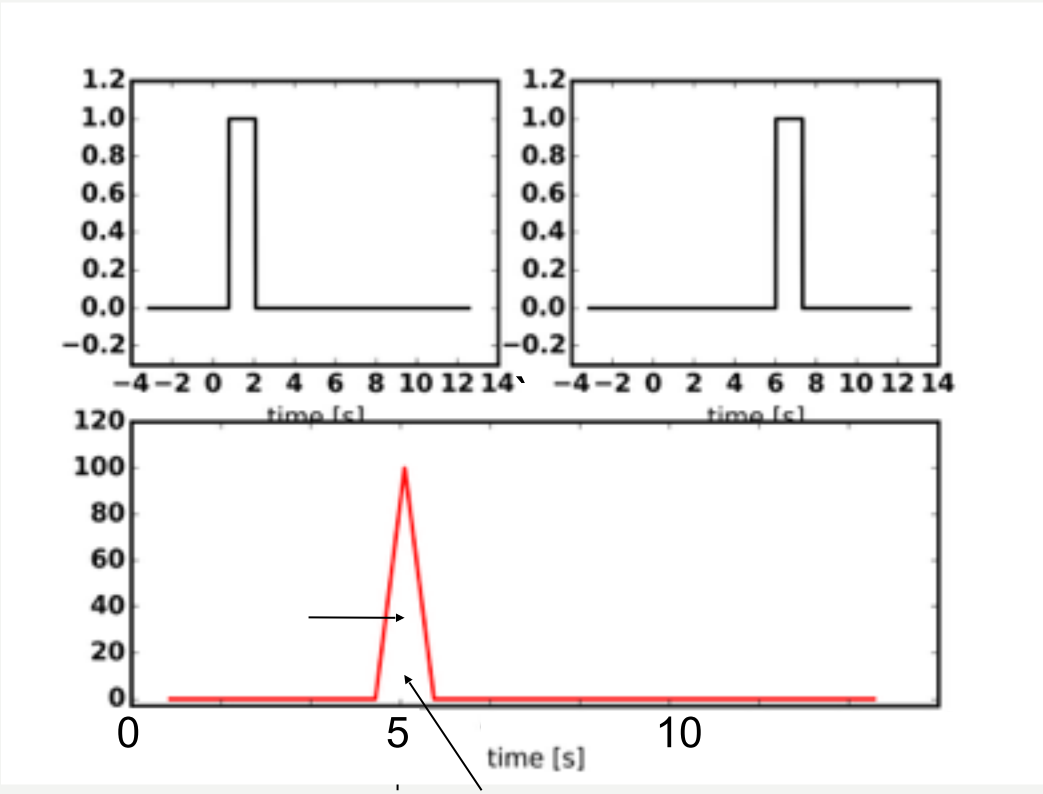
$$f \star g = f^*(-t) * g.$$

For discrete signals:

$$(f * g)[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} f[m]\, g[n-m]$$

$$(f \star g)[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\; g[m+n].$$

Correlation



Similarity between two functions

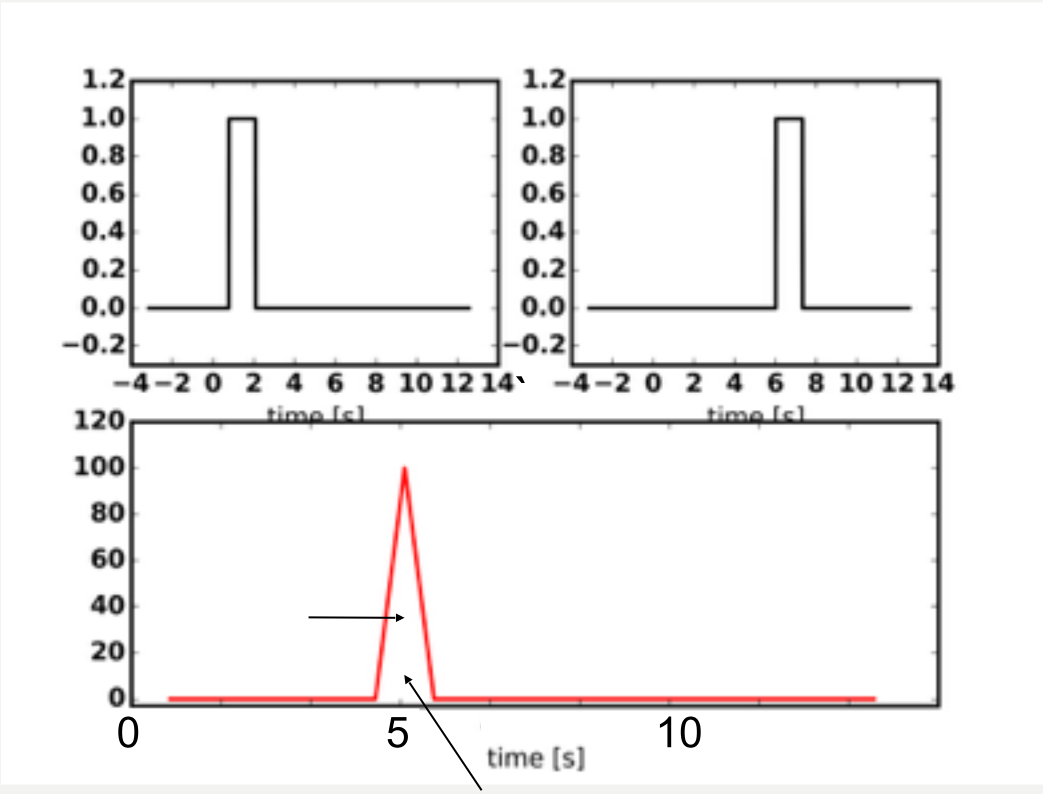
Cross-correlation:

$$(f \star g)(\tau) \stackrel{\mathrm{def}}{=} \int_{-\infty}^{\infty} f^*(t)\ g(t+\tau)\,dt,$$

cross correlation * (star!)

$$(f \star g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

Correlation



Lag between two functions

Cross-correlation:

$$(f \star g)(\tau) \stackrel{\mathrm{def}}{=} \int_{-\infty}^{\infty} f^*(t)\ g(t+\tau)\,dt,$$

cross correlation * (star!)

$$(f \star g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

Correlation

Convolution:

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$

Cross-correlation:

$$(f \star g)(\tau) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f^*(t)\ g(t+\tau)\, dt,$$

For discrete signals: :

$$(f * g)[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} f[m]\, g[n-m]$$

$$(f \star g)[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

So how would you calculate the cross correlation?

Correlation

Convolution:

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$

Cross-correlation:

$$(f \star g)(\tau) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f^*(t)\ g(t+\tau)\, dt,$$

For discrete signals: :

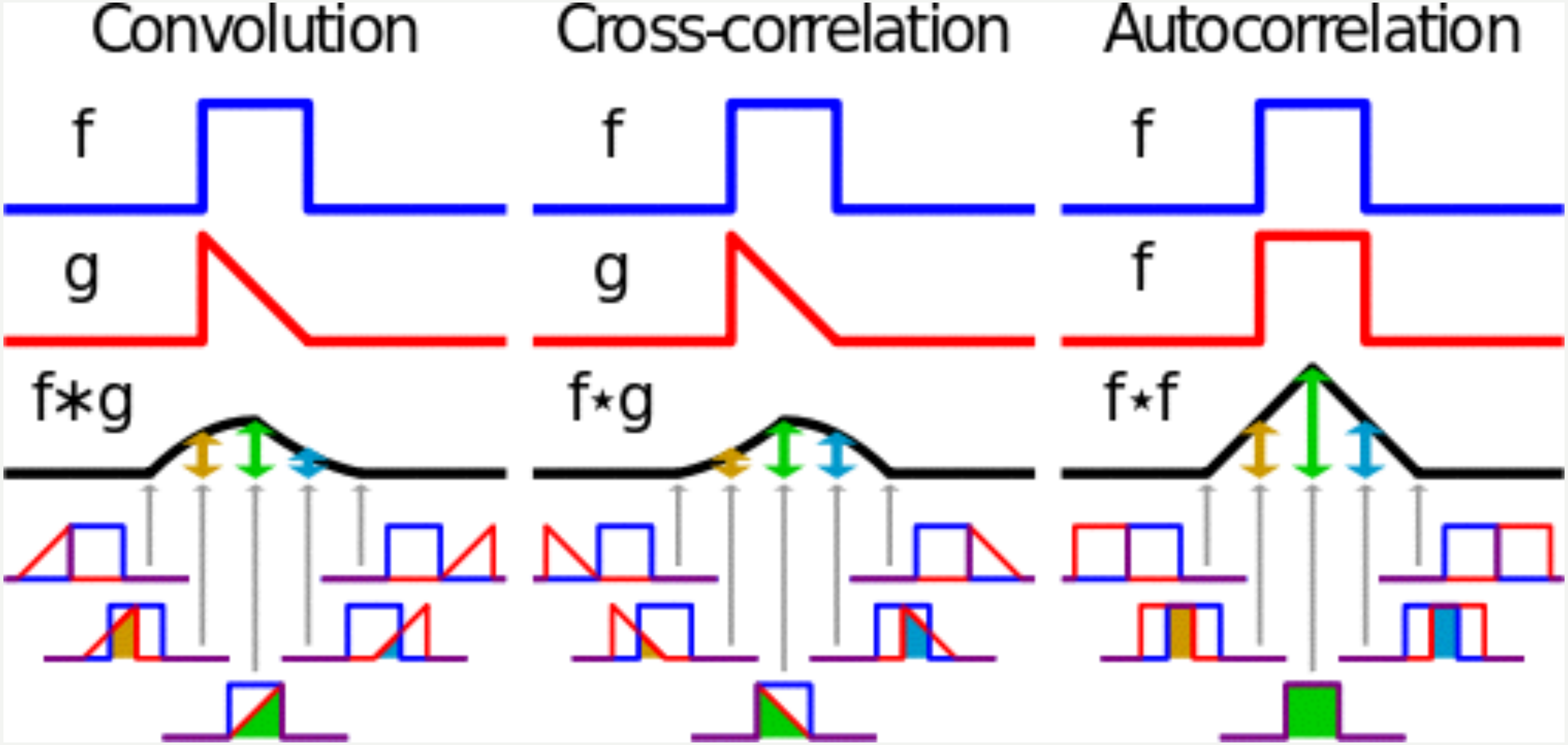
$$(f * g)[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} f[m]\, g[n-m]$$

$$(f \star g)[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

- : Flip, shift, sum.

+! : don't flip, shift, sum.

Correlation



Pictoral representation

$$y_k = \sum_{i=0}^{m} g_i f_{k-i}$$

$$k = 0, 1, 2, \ldots, m+n$$

	x					y		x*y
		0	1	0	0			
					1	2	1	**0**
		0	1	0	0			
				1	2	1		**0**
		0	1	0	0			
			1	2	1			**1**
		0	1	0	0			
		1	2	1				**2**
		0	1	0	0			
	1	2	1					**1**
		0	1	0	0			
1	2	1						**0**

Correlation

$$(f \star g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

+! : don't flip, shift, sum.

a = [0, 1, 0, 3, -2, 0]   b = [0, 3, -2, 0, 2, 0]

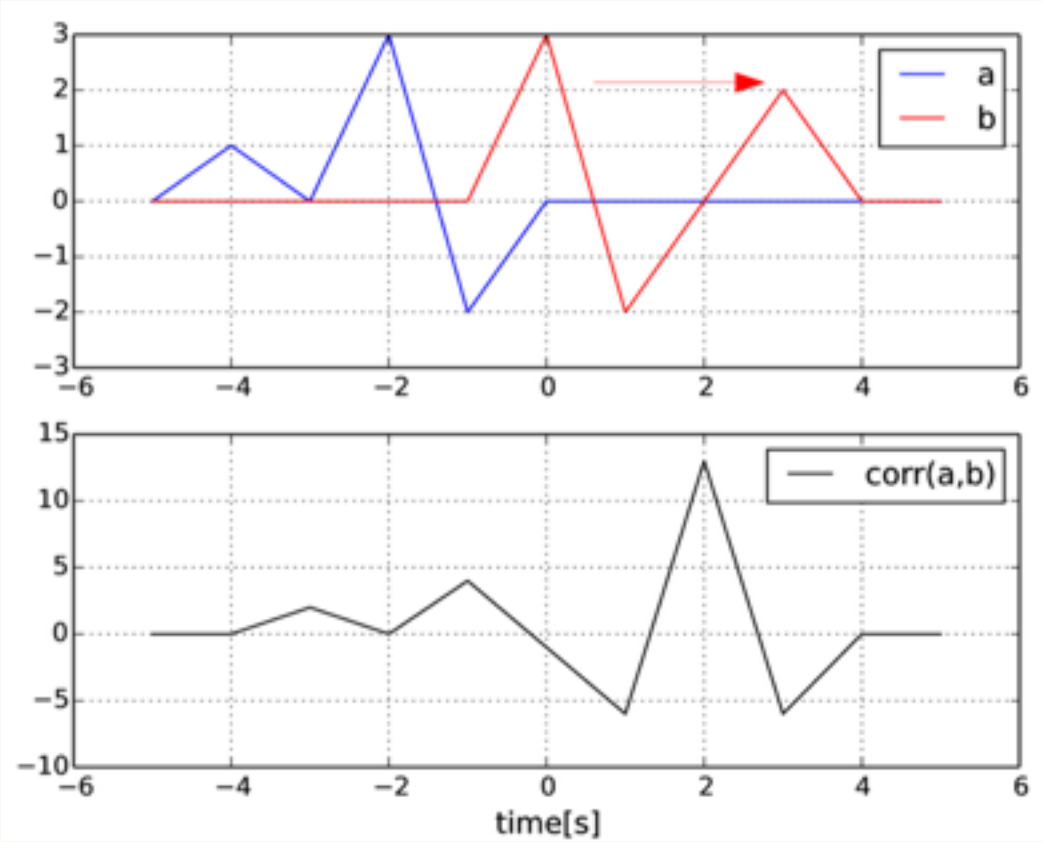
corr(a,b) =??

Correlation

$$(f \star g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

+! : don't flip, shift, sum.

a = [0, 1, 0, 3, -2, 0] b = [0, 3, -2, 0, 2, 0]



Total: 6

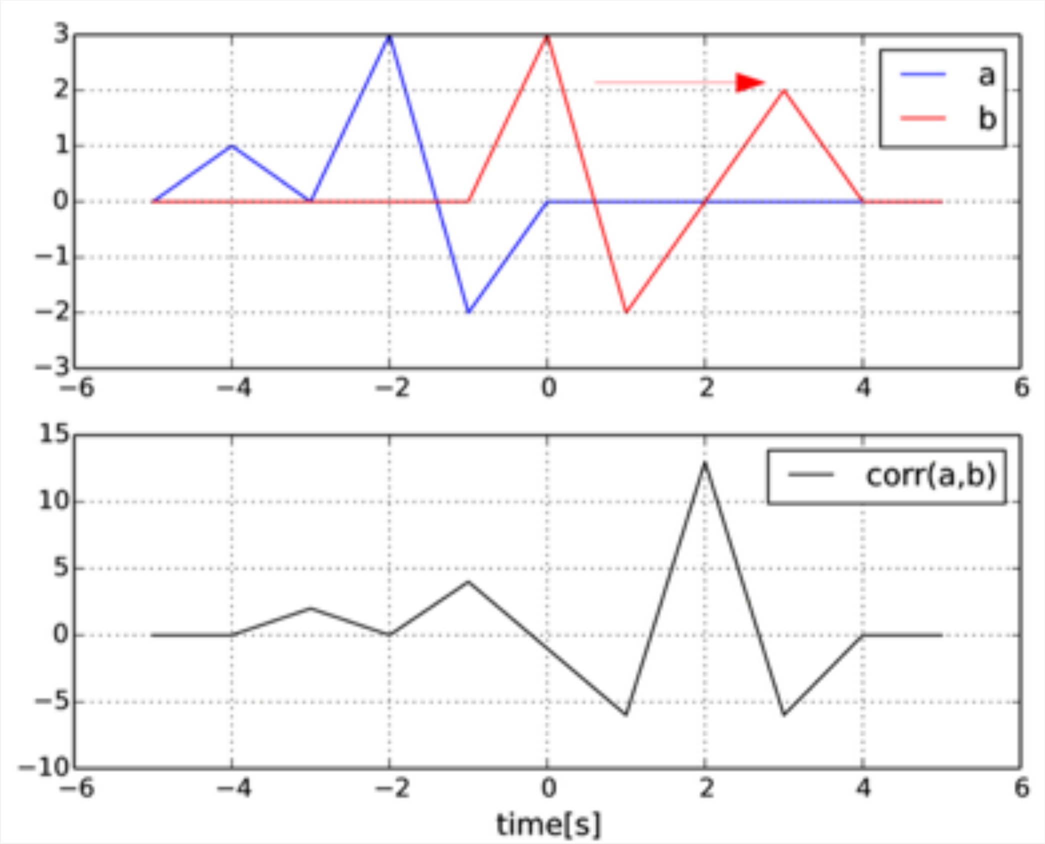
corr(a,b) = [0, 0, 2, 0, 4, -1, -6, 13, -6, 0, 0]

N=0 N=2

Cross-correlation: commutative?

a = [0, 1, 0, 3, -2, 0]   b = [0, 3, -2, 0, 2, 0]

corr(a,b):



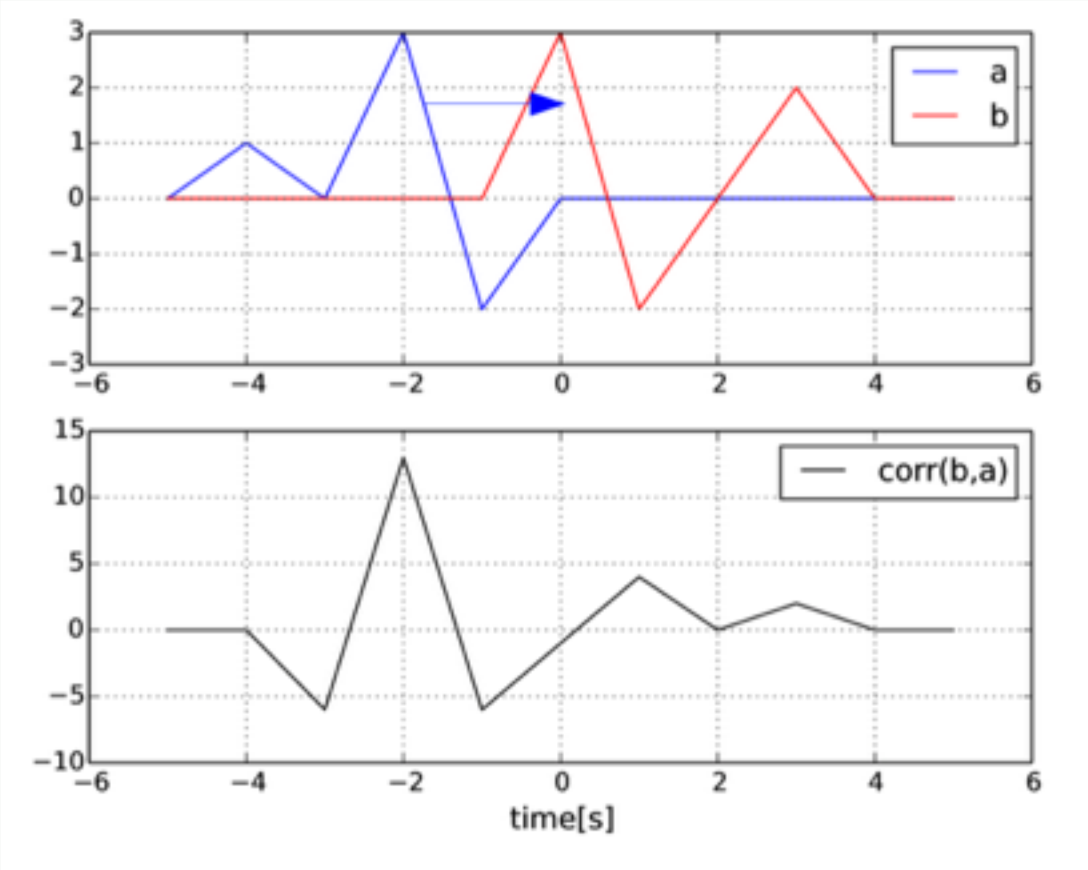
Total: 6

corr(a,b) = [0, 0, 2, 0, 4, -1, -6, 13, -6, 0, 0]

Cross-correlation: commutative?

a = [0, 1, 0, 3, -2, 0] b = [0, 3, -2, 0, 2, 0]

corr(a,b)



Total: 6

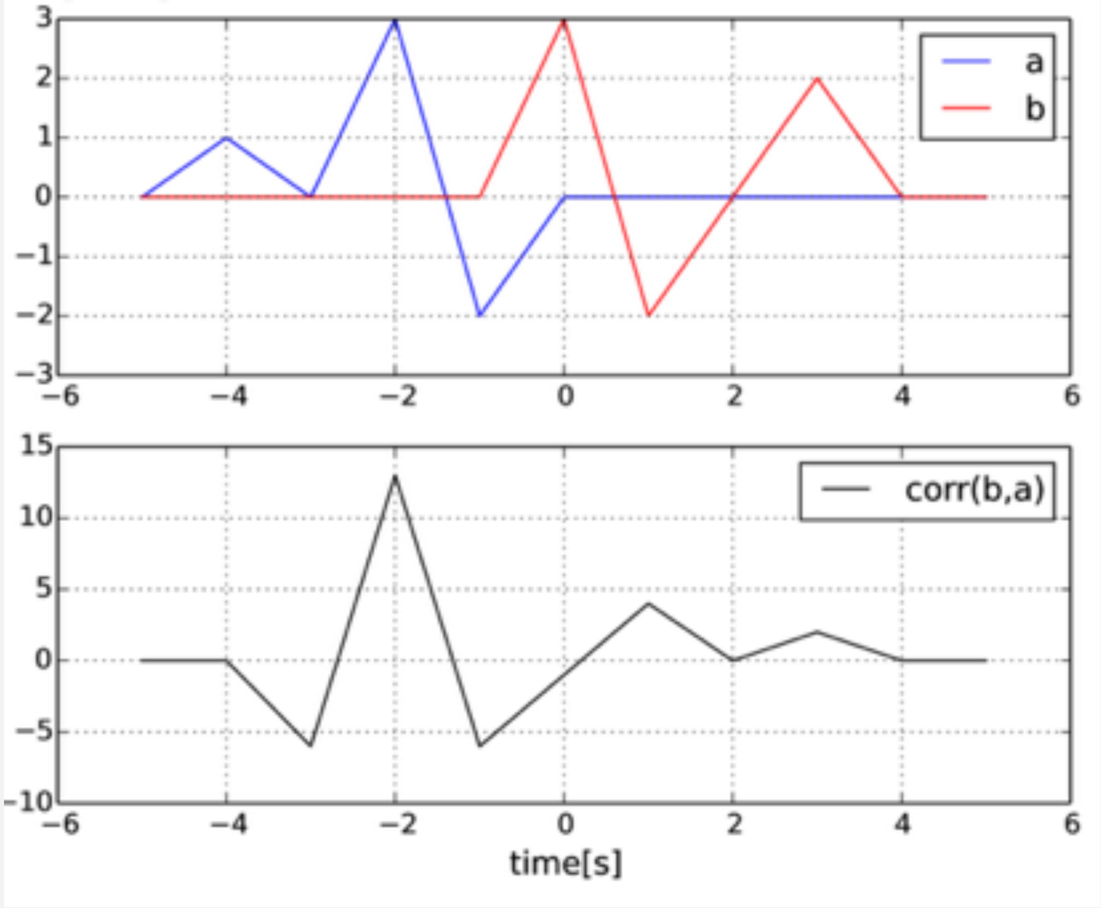
corr(a,b) = [0, 0, 2, 0, 4, -1, -6, 13, -6, 0, 0]

Correlation

Cross-correlation: commutative?

a = [0, 1, 0, 3, -2, 0]   b = [0, 3, -2, 0, 2, 0]

corr(b,a) =/= corr(a,b)



Total: 6

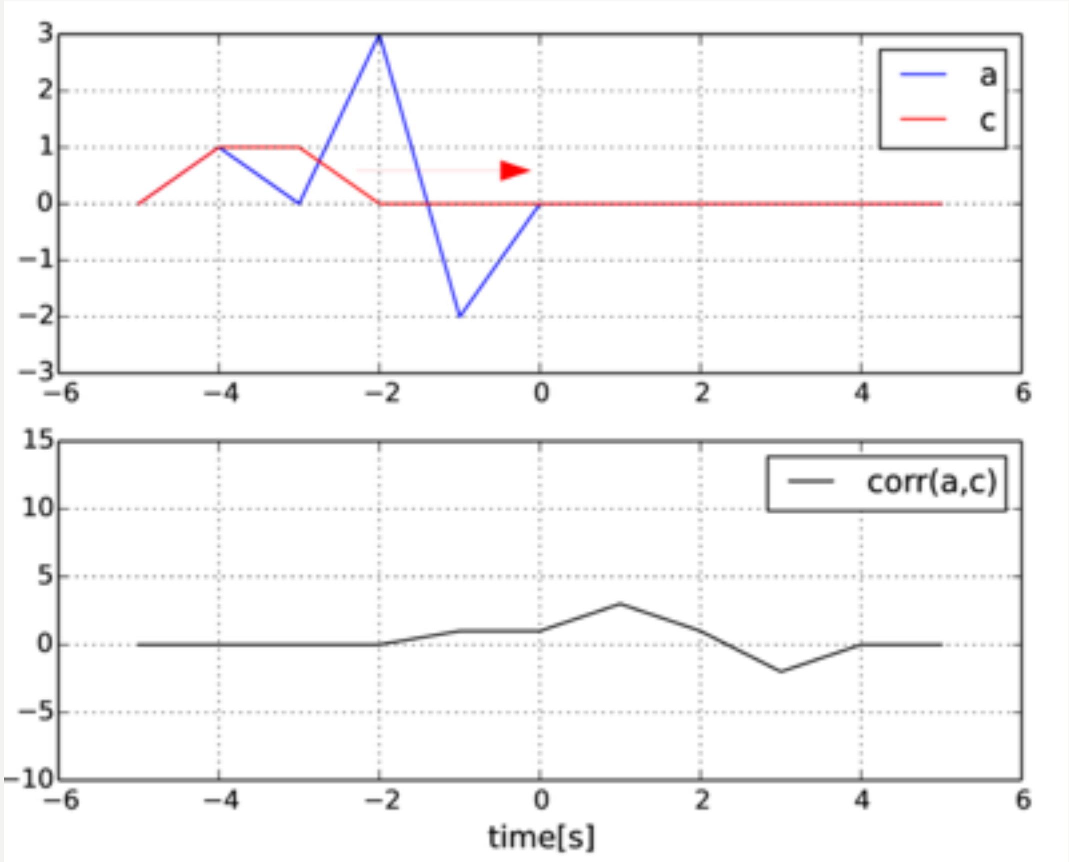
corr(a,b) = [0, 0, 2, 0, 4, -1, -6, 13, -6, 0, 0]

Correlation

$$(f \star g)[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

+! : don't flip, shift, sum.

a = [0, 1, 0, 3, -2, 0] b = [0, 3, -2, 0, 2, 0] c = [0, 1, 1, 0, 0, 0]



Total: 4

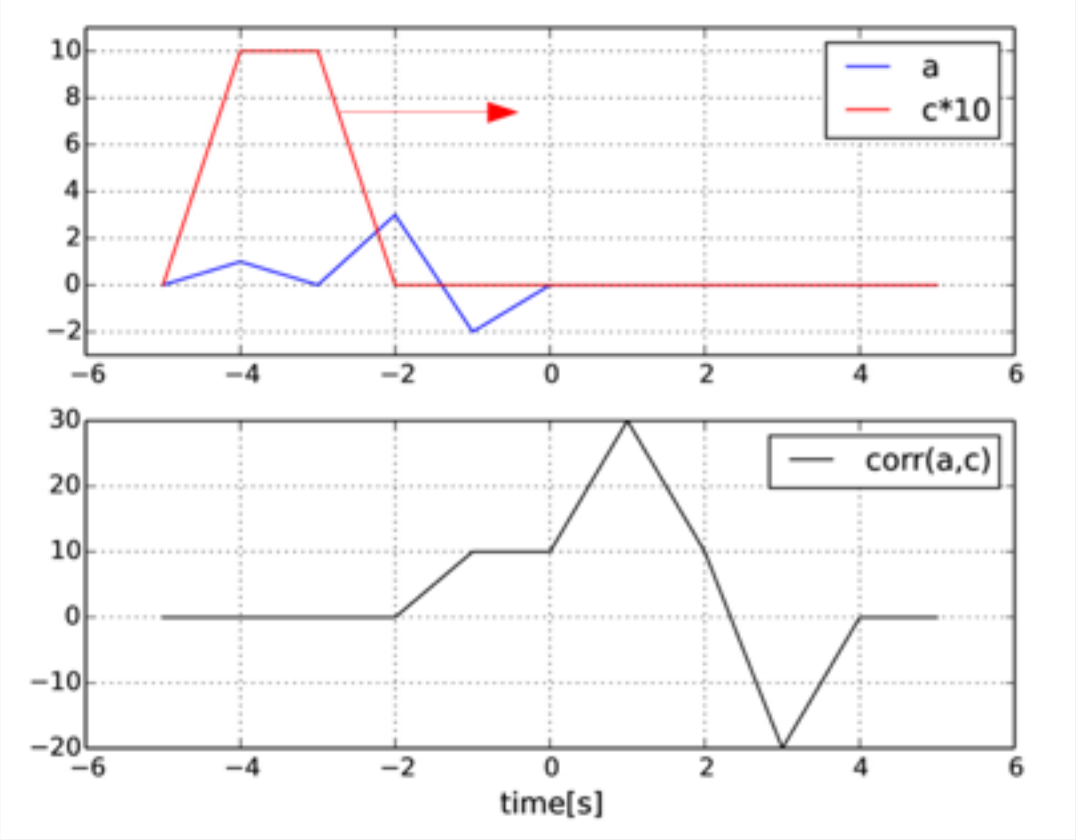
corr(a,c) = [0, 0, 0, 0, 1, 1, 3, 1, -2, 0, 0]

Correlation

$$(f \star g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

+! : don't flip, shift, sum.

a = [0, 1, 0, 3, -2, 0]   b = [0, 3, -2, 0, 2, 0]   c * 10 = [0, 10, 10, 0, 0, 0]



Total: 40

corr(a, c*10) = [0, 0, 0, 0, 10, 10, 30, 10, -20, 0, 0]

Correlation

$$(f \star g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

+! : don't flip, shift, sum.

a = [0, 1, 0, 3, -2, 0]   b = [0, 3, -2, 0, 2, 0]   c * 10 = [0, 10, 10, 0, 0, 0]

$$\mathrm{CC} = \frac{\sum_{m=-\infty}^{\infty} f_m^* g_{m+n}}{\sqrt{\sum f_m^2 \sum g_m^2}}$$

Normalized according to energy contained in both signals

CC can have values from [-1:1]

sum corr(a,b): 6
sum corr(b,a): 6
sum corr(a,-b): -6
sum corr(a,c): 4
sum corr(a,c*10): 40

Correlation Coefficient

$$(f \star g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

+! : don't flip, shift, sum.

a = [0, 1, 0, 3, -2, 0]   b = [0, 3, -2, 0, 2, 0]   c * 10 = [0, 10, 10, 0, 0, 0]

$$CC = \frac{\sum_{m=-\infty}^{\infty} f_m^* g_{m+n}}{\sqrt{\sum f_m^2 \sum g_m^2}}$$

Normalized according to energy contained in both signals

CC can have values from [-1:1]

sum corr(a,b): 6
sum corr(b,a): 6
sum corr(a,-b): -6
sum corr(a,c): 4
sum corr(a,c*10): 40

corrcoef: -0.14
corrcoef: -0.14
corrcoef: 0.14
corrcoef: 0.08
corrcoef: 0.08

Correlation Coefficient

$$(f \star g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\; g[m+n].$$

+! : don't flip, shift, sum.

a = [0, 1, 0, 3, -2, 0]   b = [0, 3, -2, 0, 2, 0]   c * 10 = [0, 10, 10, 0, 0, 0]

$$\mathrm{CC} = \frac{\sum\limits_{m=-\infty}^{\infty} f_m^* g_{m+n}}{\sqrt{\sum f_m^2 \sum g_m^2}}$$

Normalized according to energy contained in both signals

CC can have values from [-1:1]

Correlation

Correlation … faster??

$$(f \star g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

Correlation

Correlation theorem:

$$(f \star g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m]\ g[m+n].$$

$$\mathcal{F}\{f \star g\} = (\mathcal{F}\{f\})^* \cdot \mathcal{F}\{g\}$$

Convolution theorem:

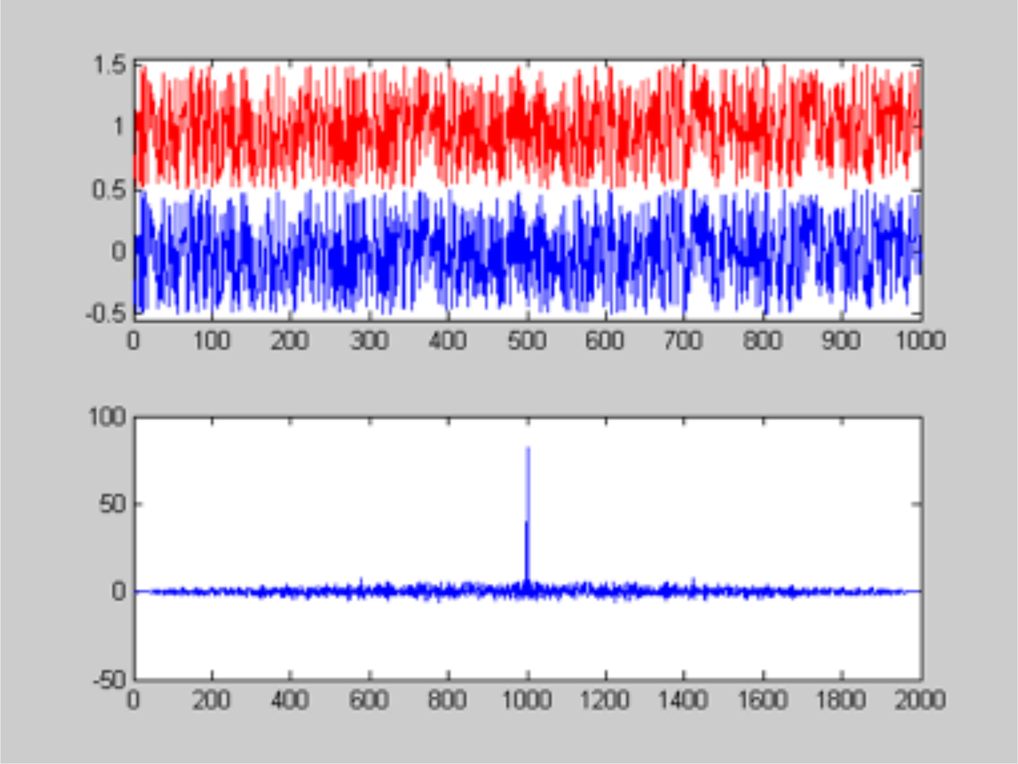
$$(f * g)[n] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} f[m]\, g[n-m]$$

$$\mathcal{F}\{f * g\} = \mathcal{F}\{f\} \cdot \mathcal{F}\{g\}$$

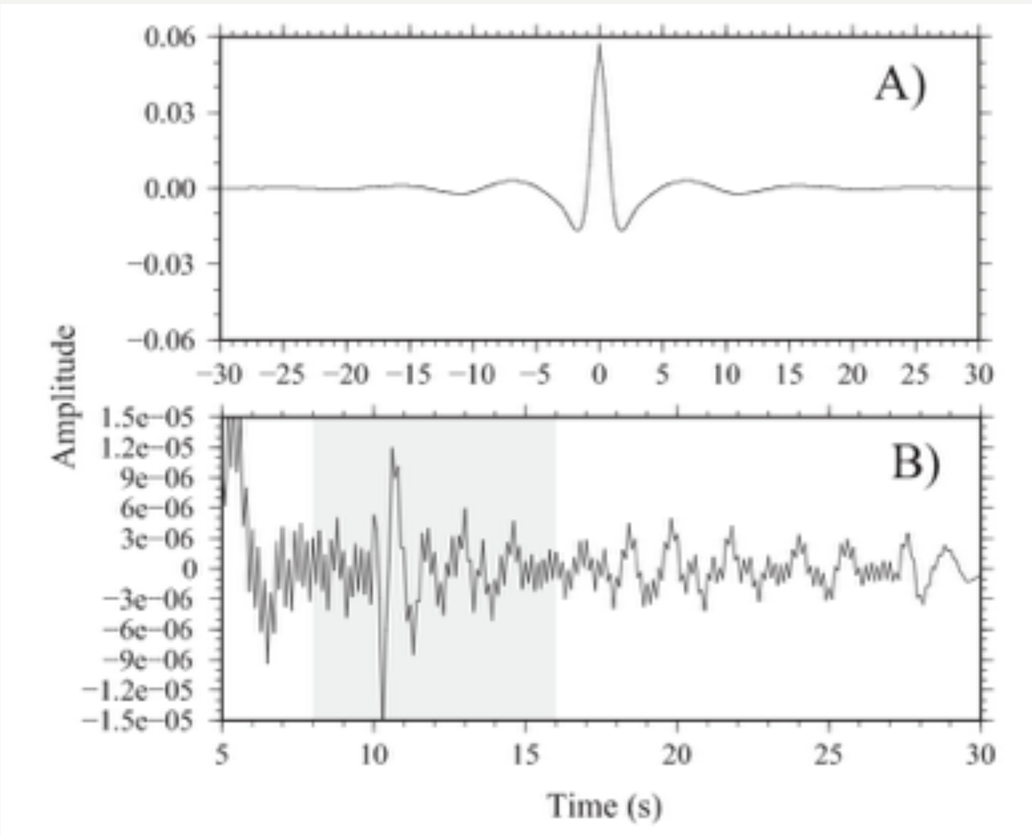
Correlation theorem:
Multiplying the FT of one function by the complex conjugate of the FT of the other gives the FT of their correlation

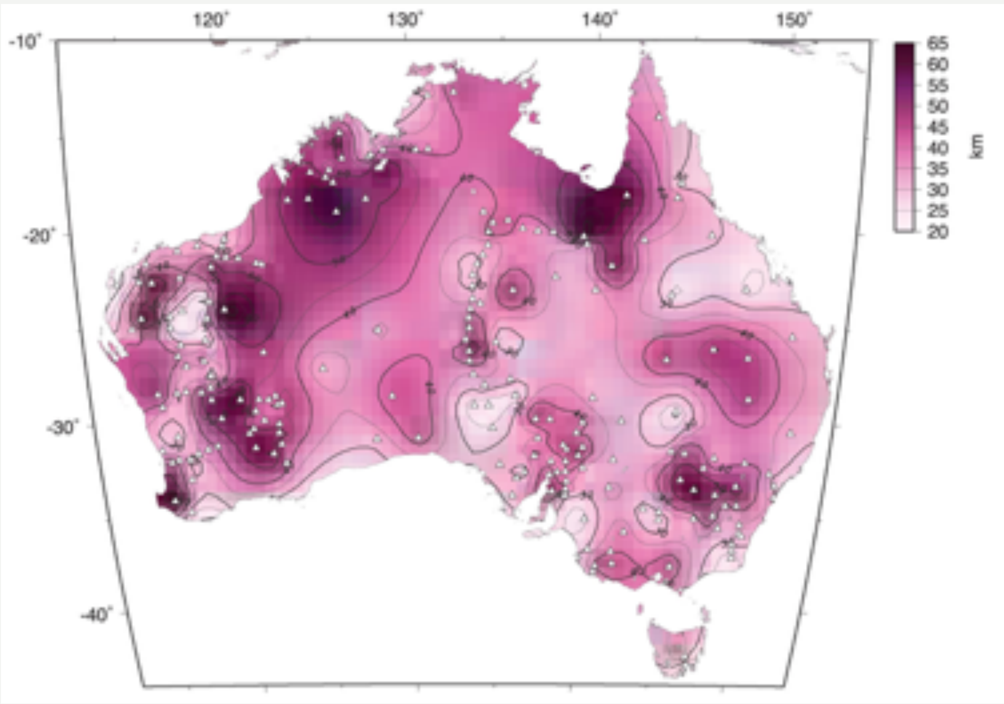
Special case: Autocorrelation

Correlation of a signal with itself: $f \star f$



Special case: Autocorrelation

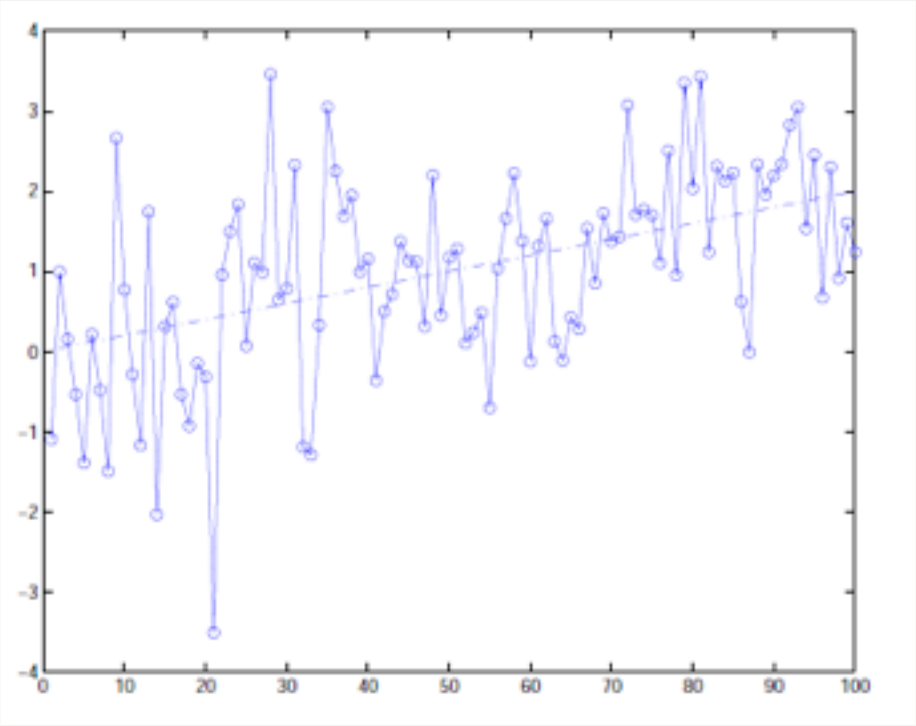




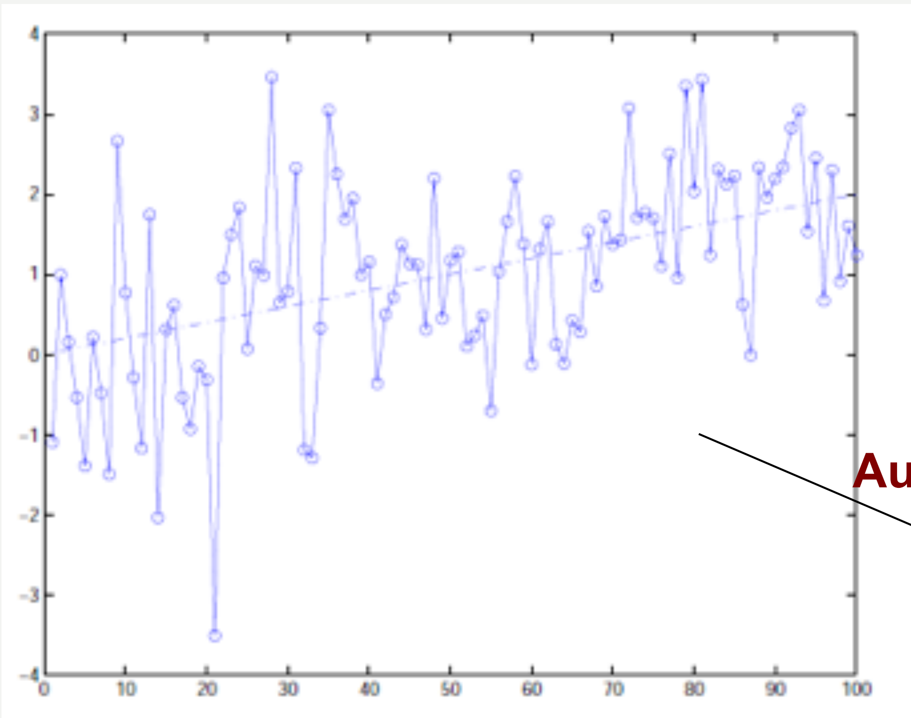
From Gorbatov et al., GJI, 2013

Autocorrelation

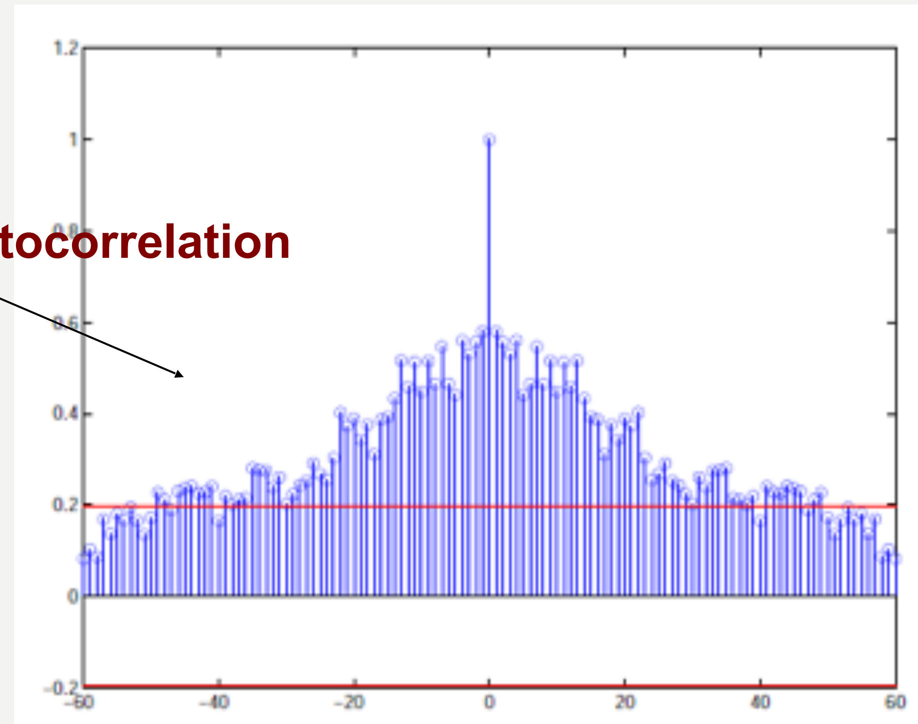
Autocorrelation of signal with trend:



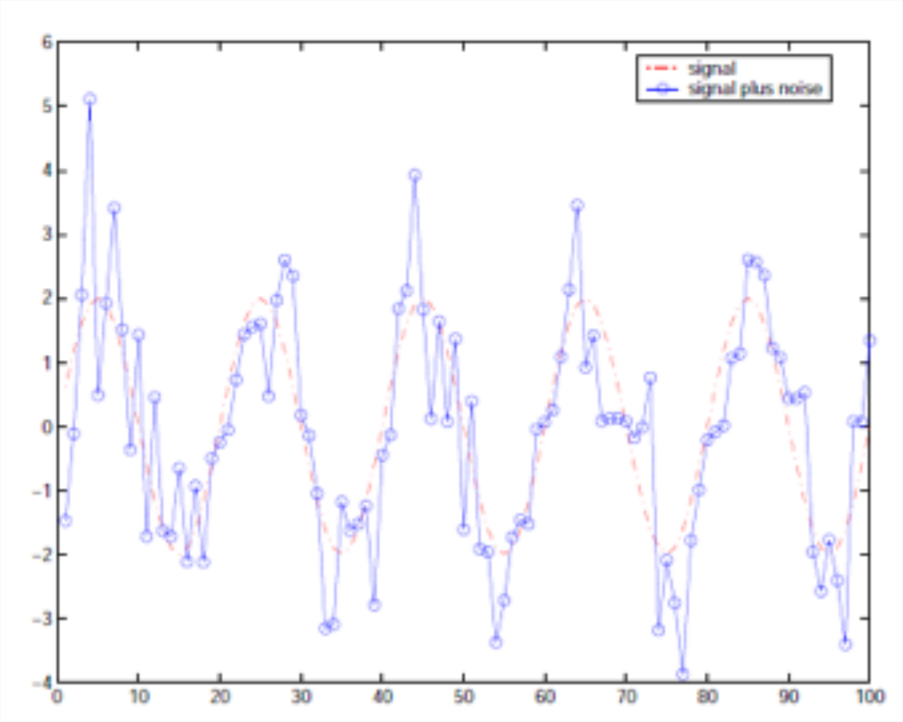
Autocorrelation of signal with trend:



Autocorrelation

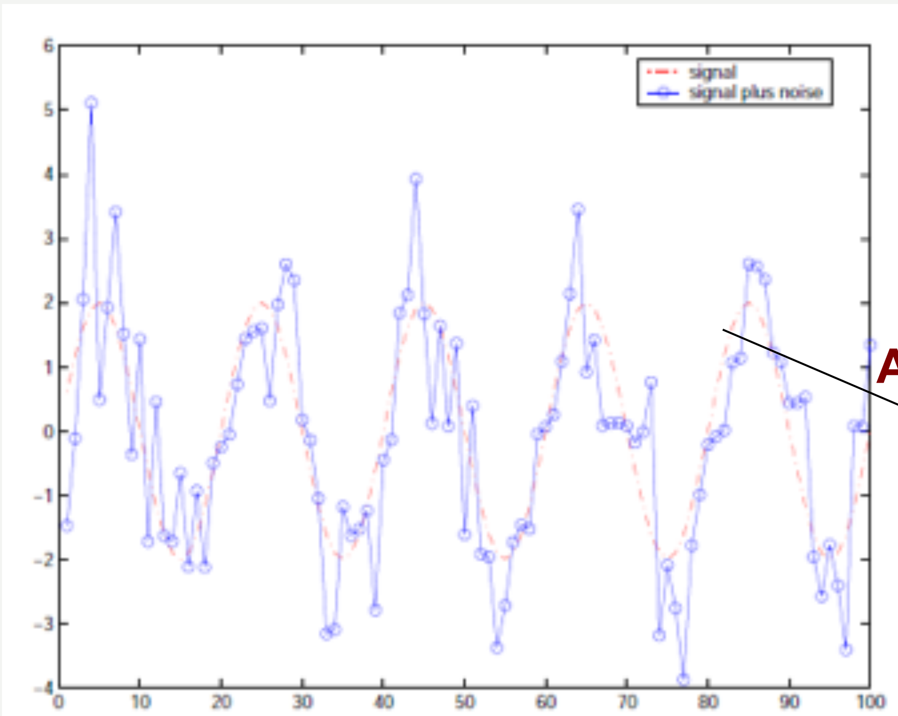


Autocorrelation of signal with periodic component:

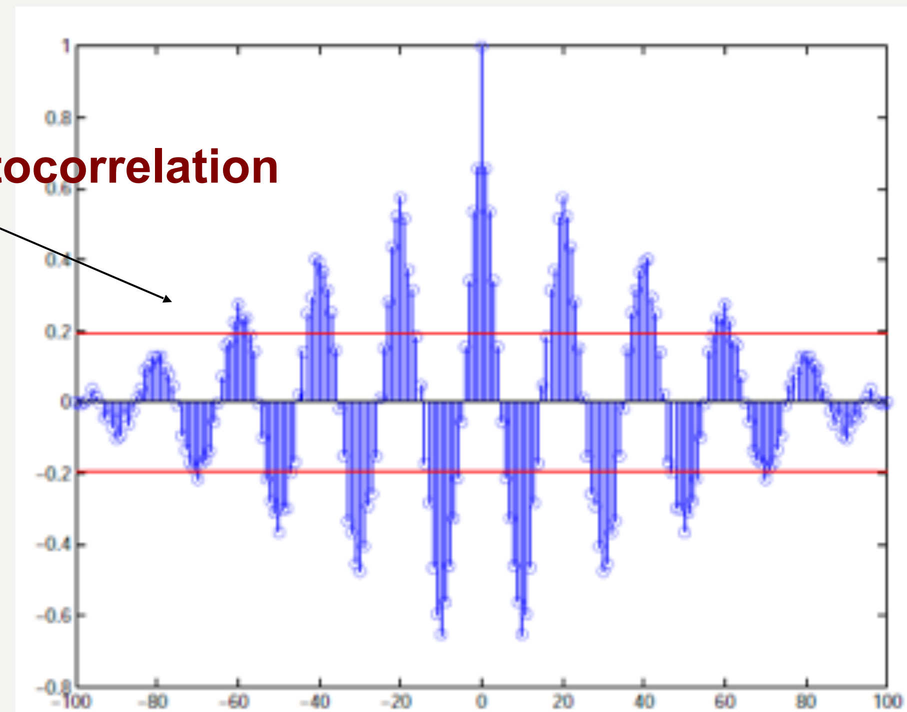


Autocorrelation

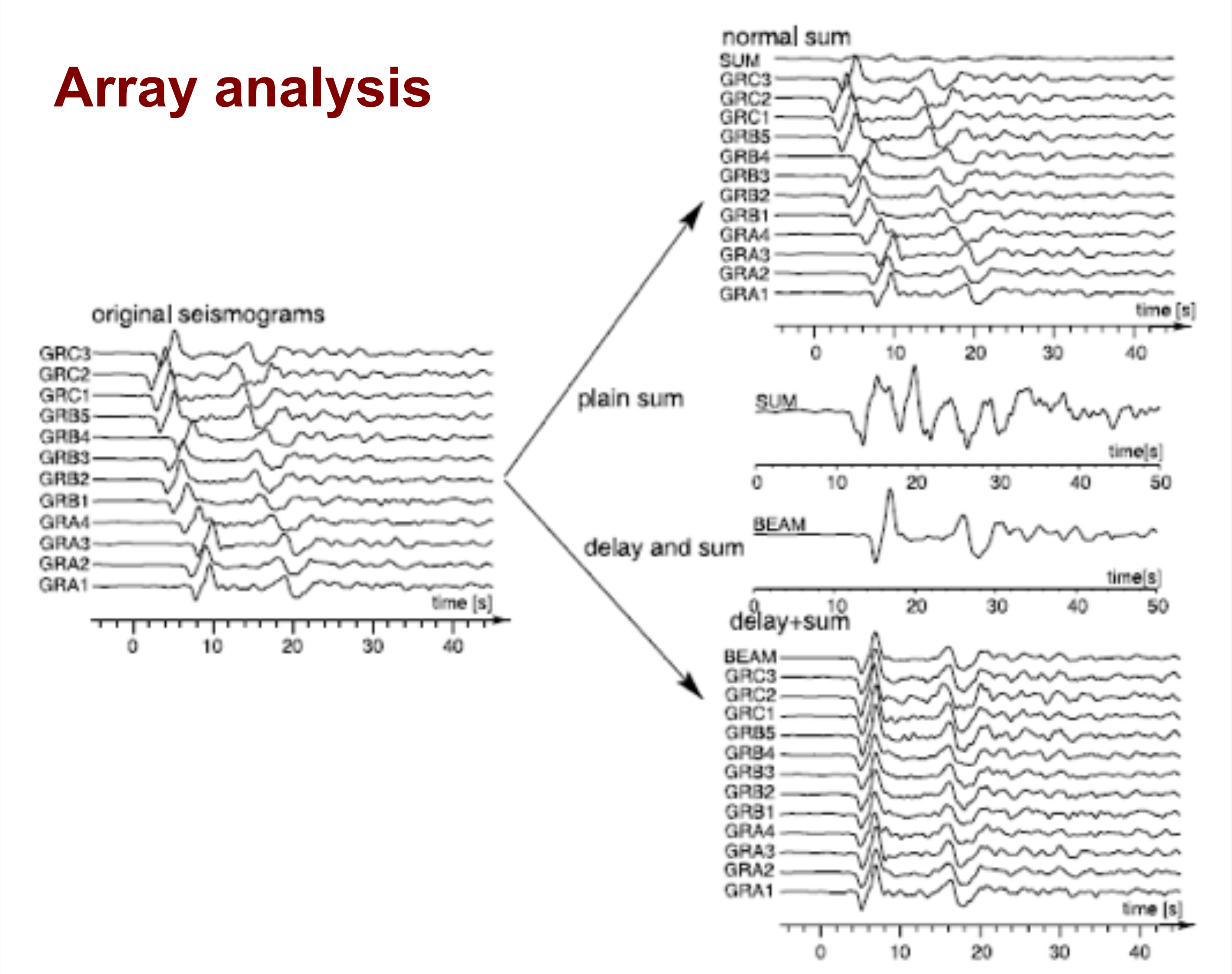
Autocorrelation of signal with periodic component:



Autocorrelation

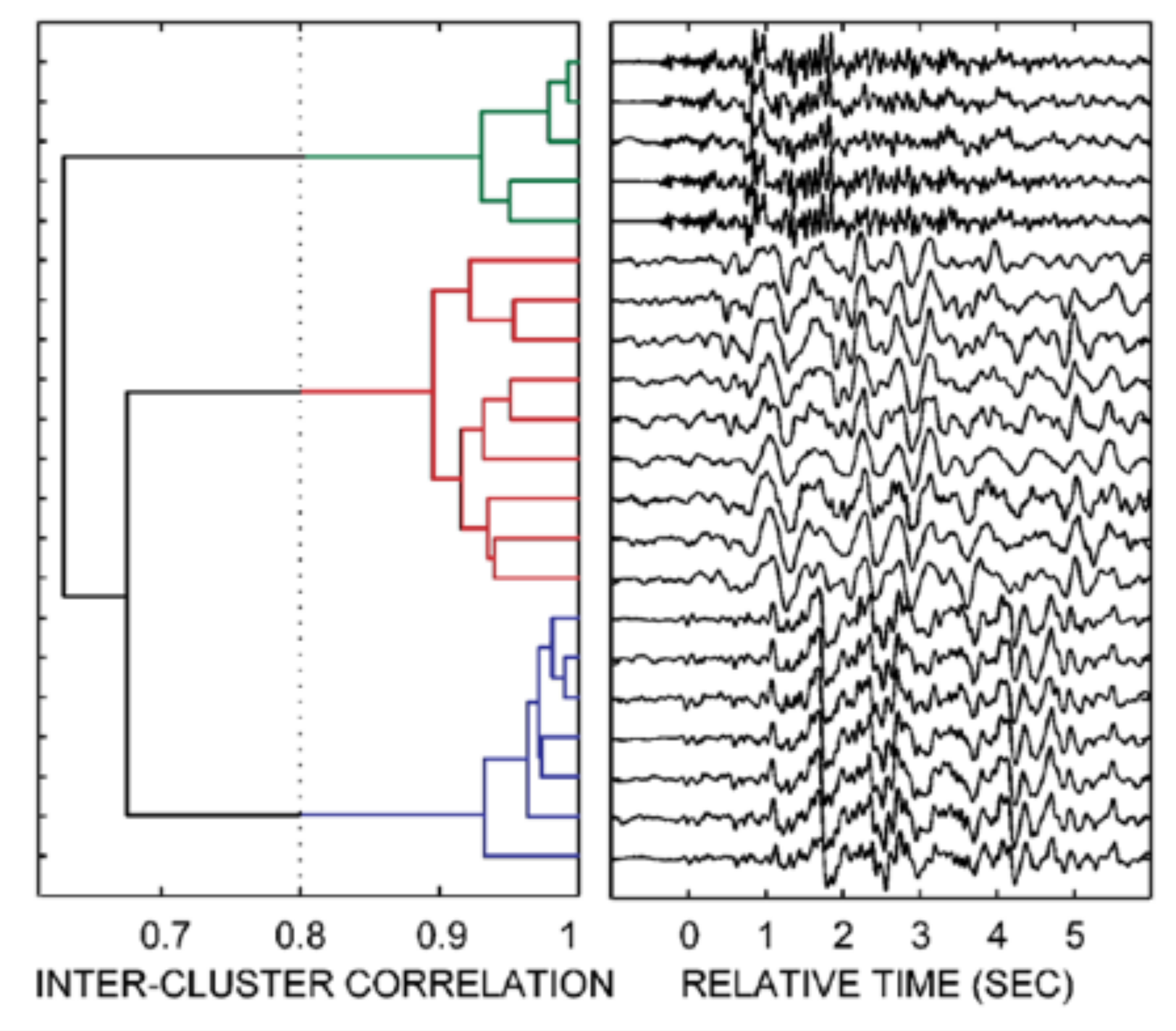


Correlation … Applications



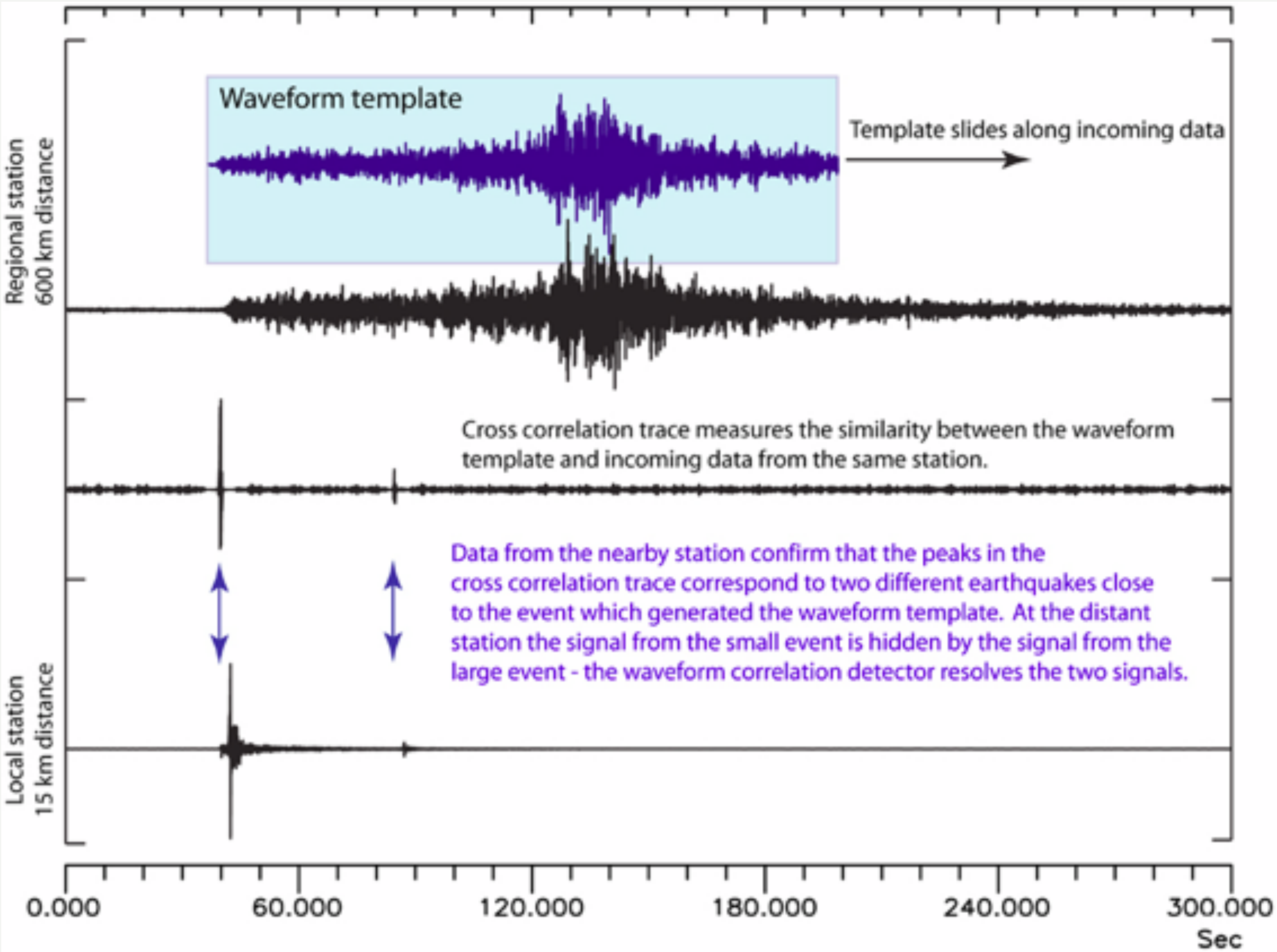
Correlation … Applications

Event clustering...



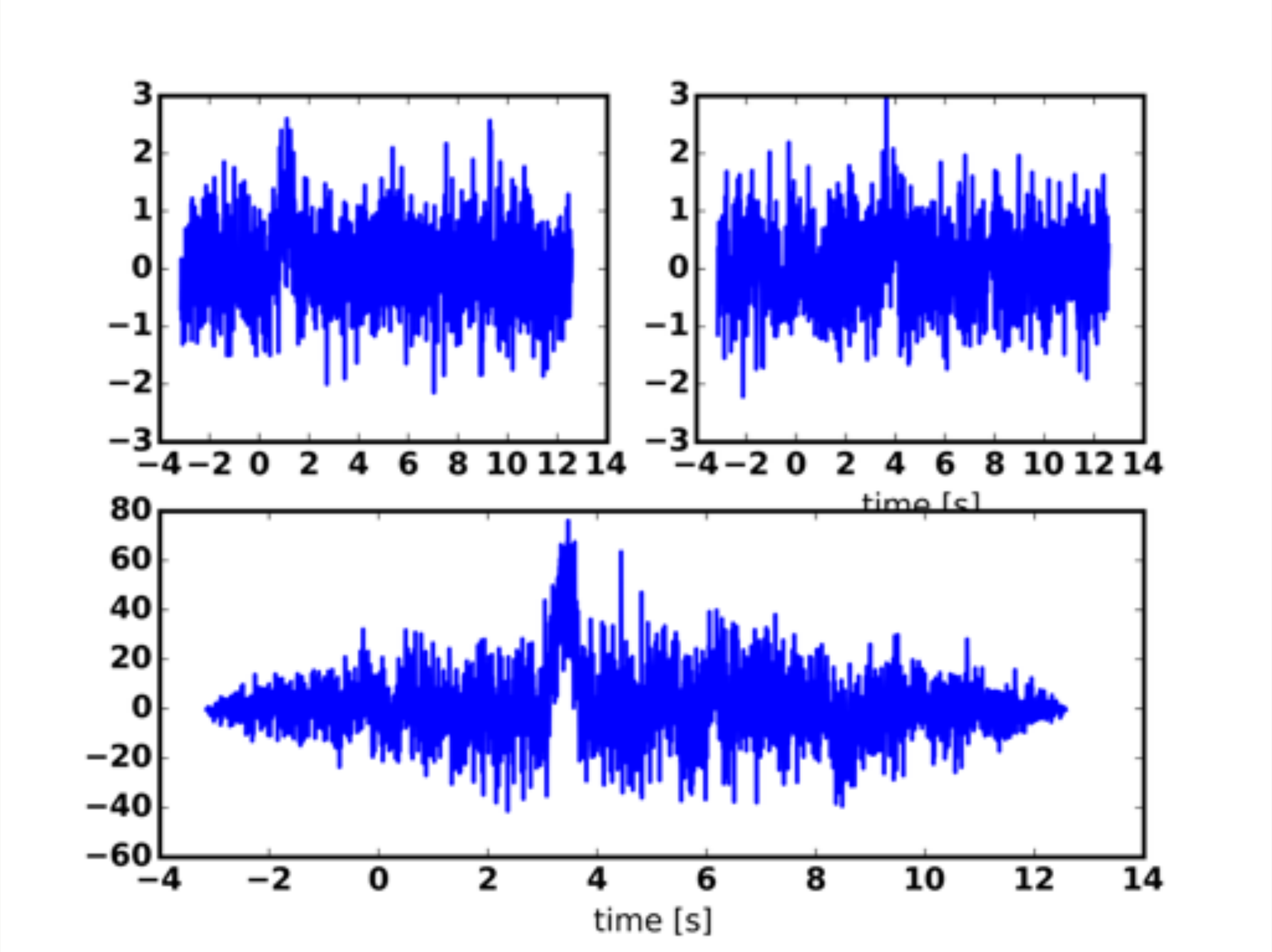
Correlation … Applications

Sophisticated event detection...



Correlation … Applications

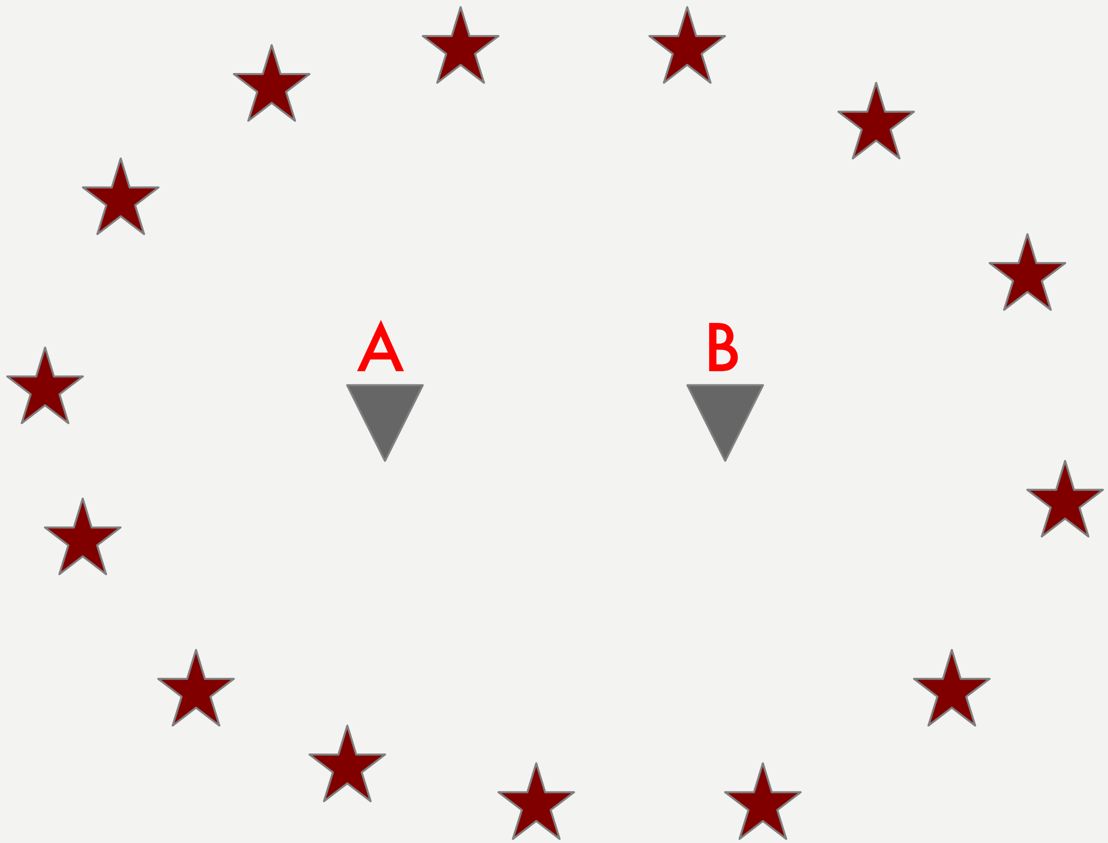
Correlation .. can help identify coherent signals:



Ambient noise correlations

… creating signal from noise

ideal case: noise is a random field + average over long time



Noise sources surround the receivers

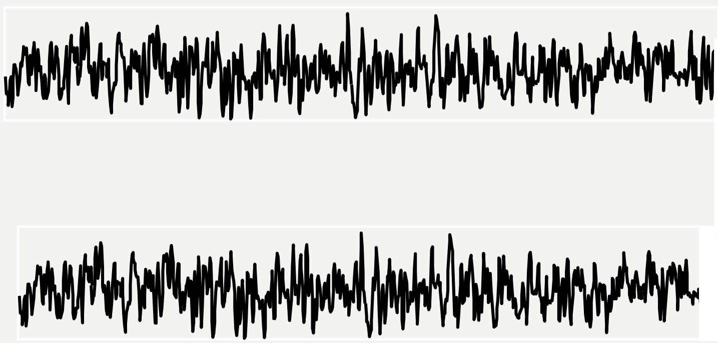
Correlation … Applications

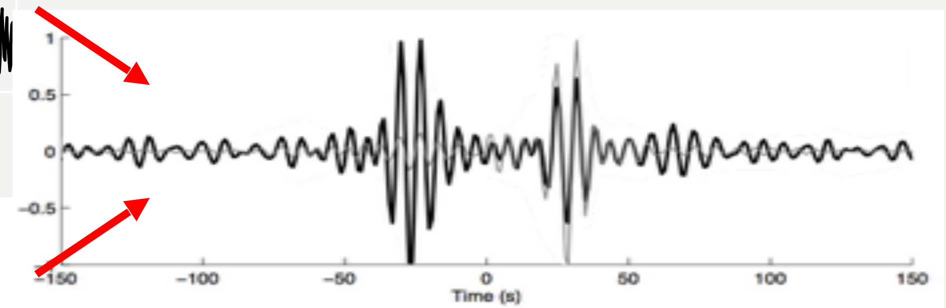
Ambient Noise Cross-correlation

ideal case: noise is a random field + average over long time

A

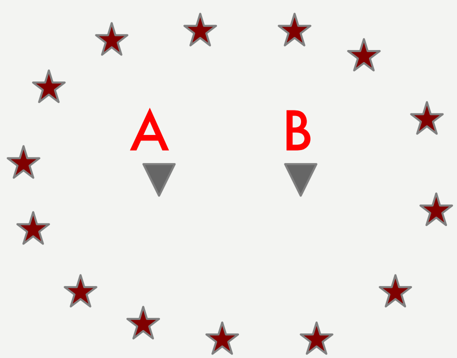
B





$C(h_A, h_B)$

$GF_{BA} + GF_{AB}$



Correlation of field in A and B

=

Green function between A and B

Correlation … Applications

Ambient Noise Cross-correlation

GF_{AB}

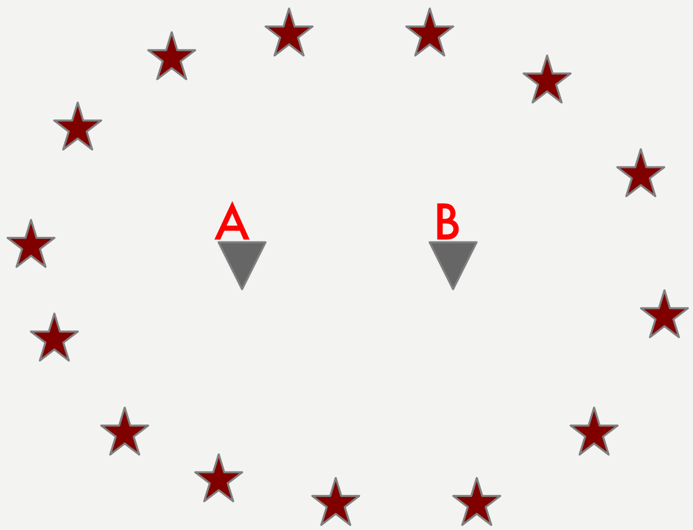
Ambient Noise Cross-correlation

Ambient Noise Cross-correlation

Advantages:

Green's function any*time*

Green's function any*where*



Correlation … Applications

Let's try that in California:

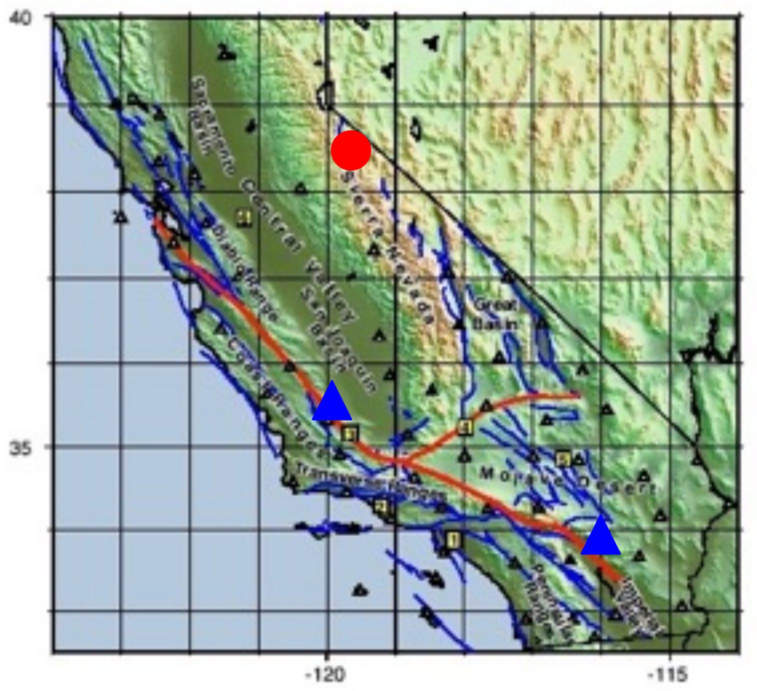
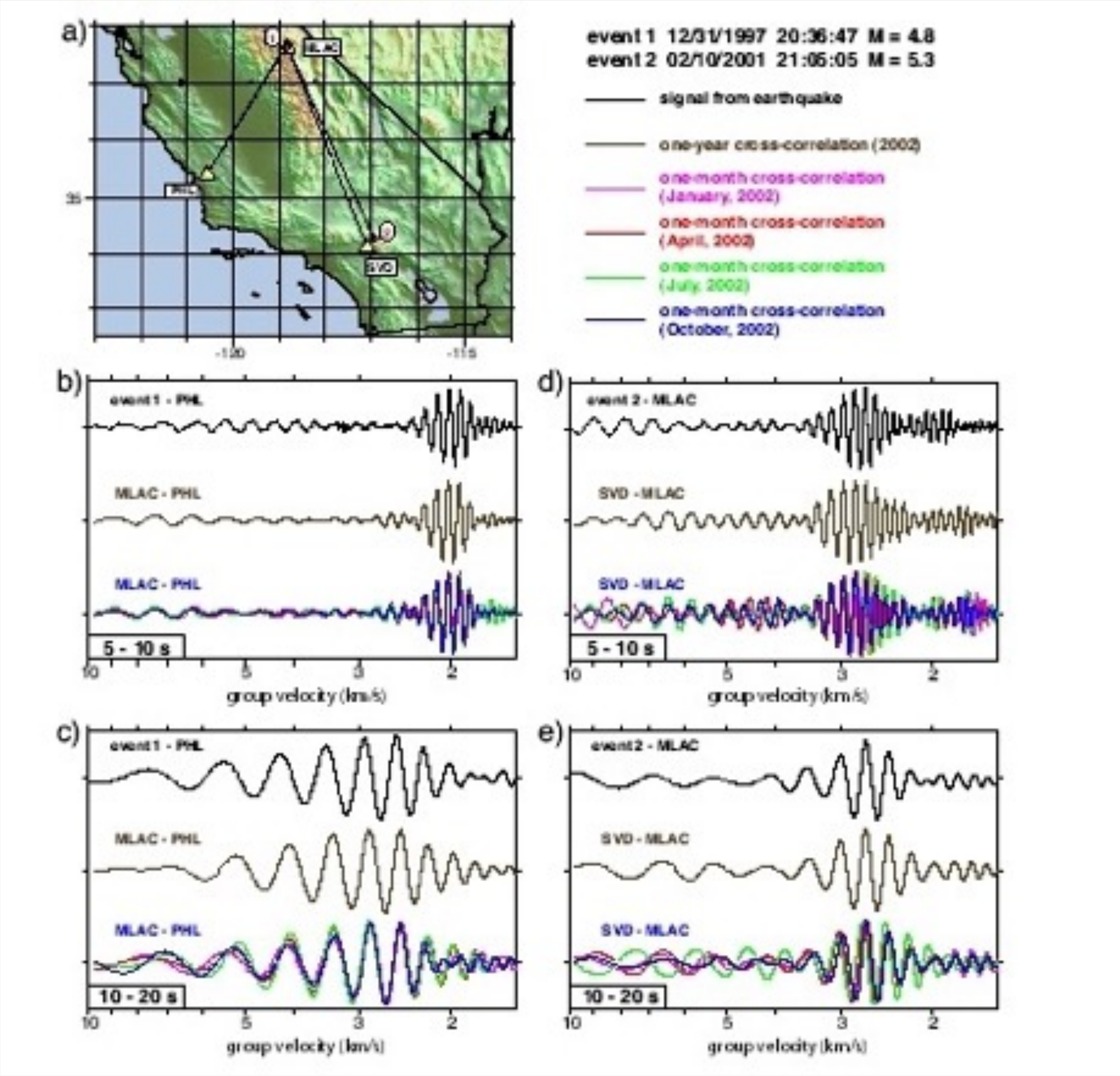


Figure 1: Reference map showing the locations of principal geographical and geological features discussed in the test. White triangles show the locations of the USArray stations used in this study (5 of the of 62 stations are located north of 40N and are not shown in this map). Blue and red solid lines show locations of known active faults. Yellow rectangles with digits indicate the following features: (1) Los Angeles Basin; (2) Ventura Basin; (3) San Andreas fault; (4) Garlock fault; (5) Mojave shear zone; (6) Stockton Arch.

Correlation … Applications



Signal from event

Signal from noise Cross- correlation

Signal from event

Signal from noise Cross- correlation

Correlation … Applications

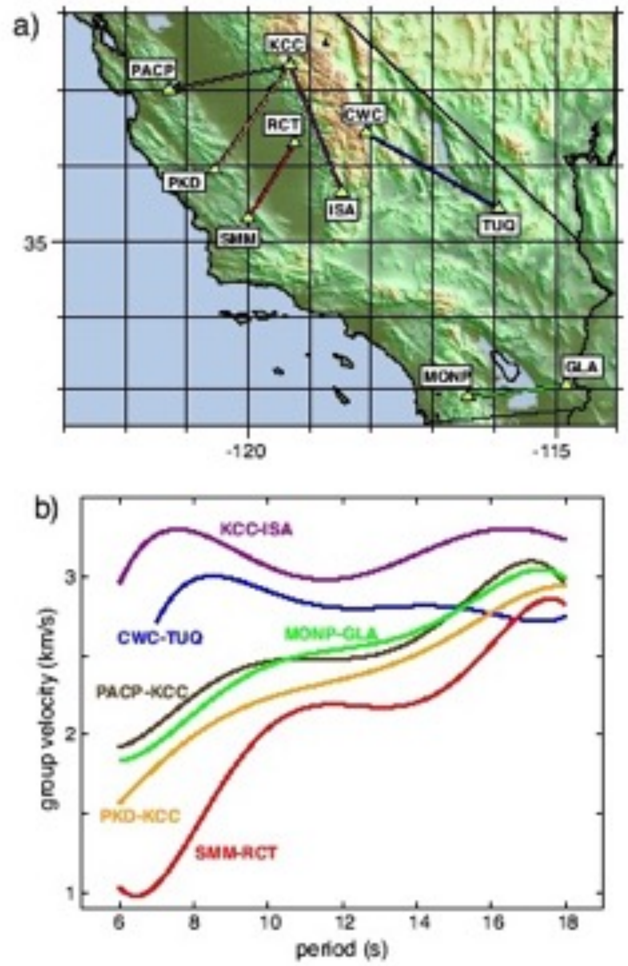


Figure 3: Group speed curves measured in different parts of California by cross-correlating 30 days of ambient noise between USArray stations. (a) Map showing the stations locations and inter-station paths. (b) Group speed dispersion curves between periods of 6 s and 18 s.

Correlation … Applications

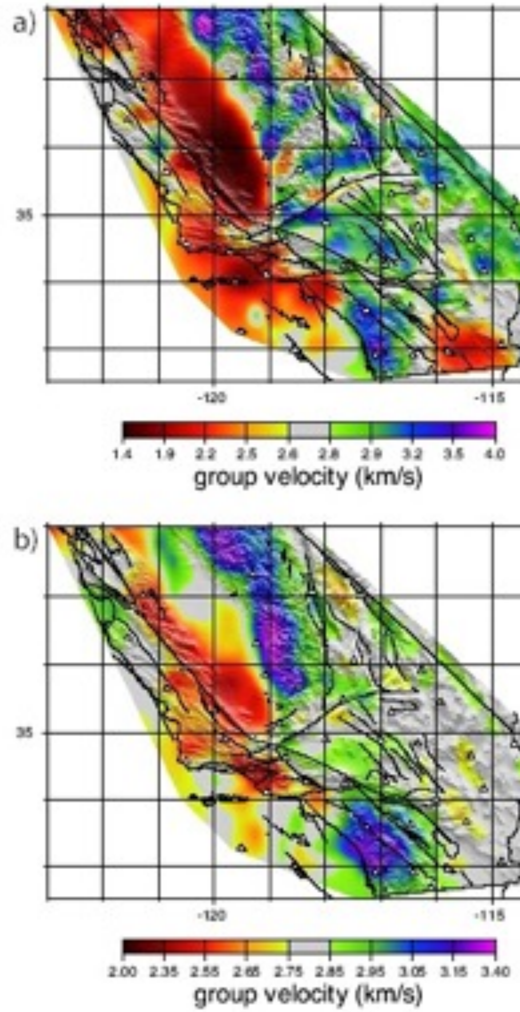
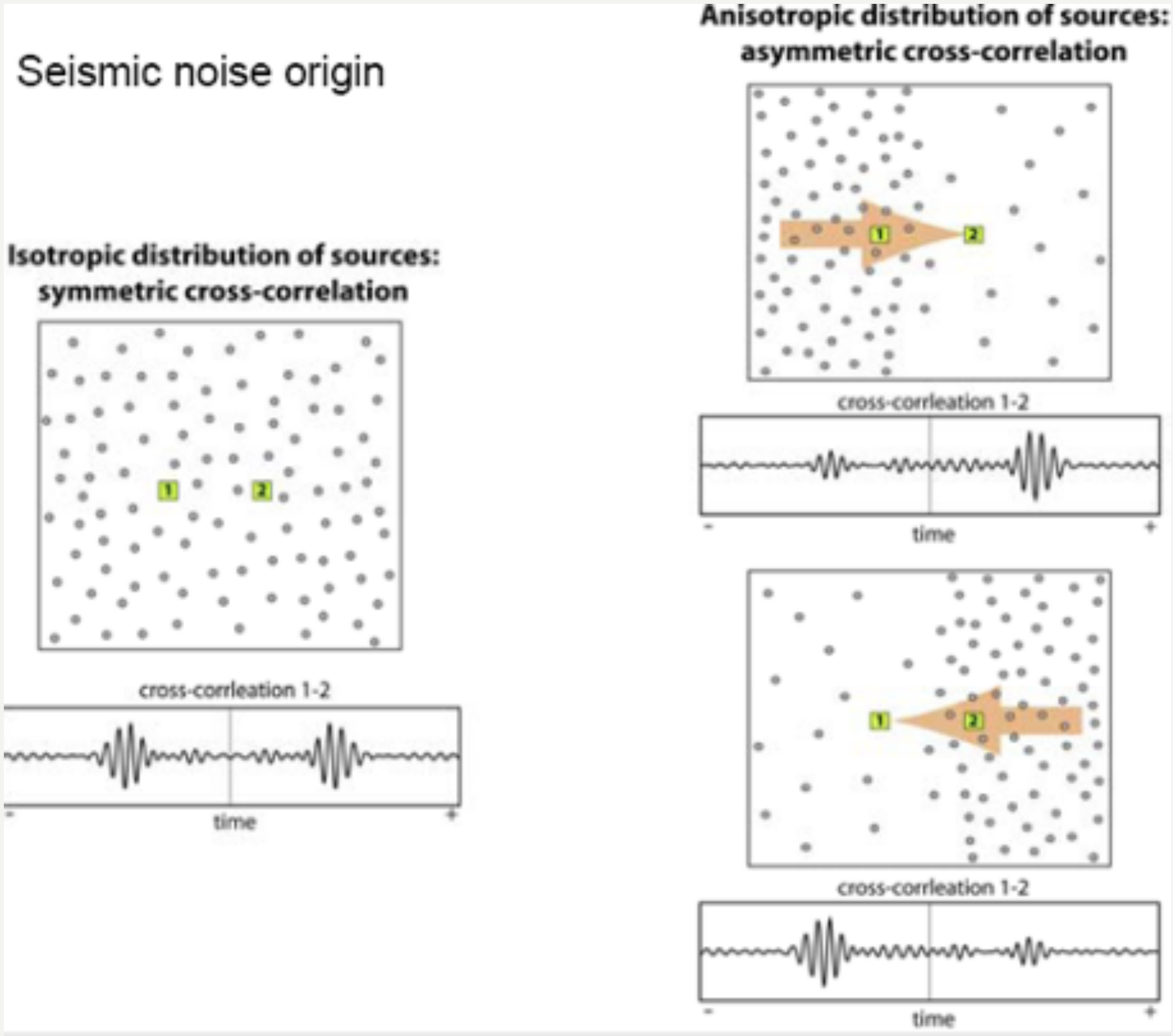


Figure 4: **Group speed maps constructed by cross-correlating 30 days of ambient noise between USArray stations.** (a) 7.5 s period Rayleigh waves. (b) 15 s period Rayleigh waves. Black solid lines show known active faults. White triangles show locations of USArray stations used in this study.

Correlation … Applications



Cross-correlation: commutative? No!

Relating rotation rate and transverse acceleration:

Plane transversely polarized wave propagating in x-direction with phase velocity c

$$u_y(x,t) = f(kx - \omega t) \qquad c = \omega / k$$

Acceleration
$$a_y(x,t) = \ddot{u}_y(x,t) = \omega^2 f''(kx - \omega t)$$

Rotation rate
$$\dot{\Omega}(x,t) = \frac{1}{2} \nabla \times [0, \dot{u}_y, 0] = \left[0, 0, -\frac{1}{2} k \omega f''(kx - \omega t)\right]$$

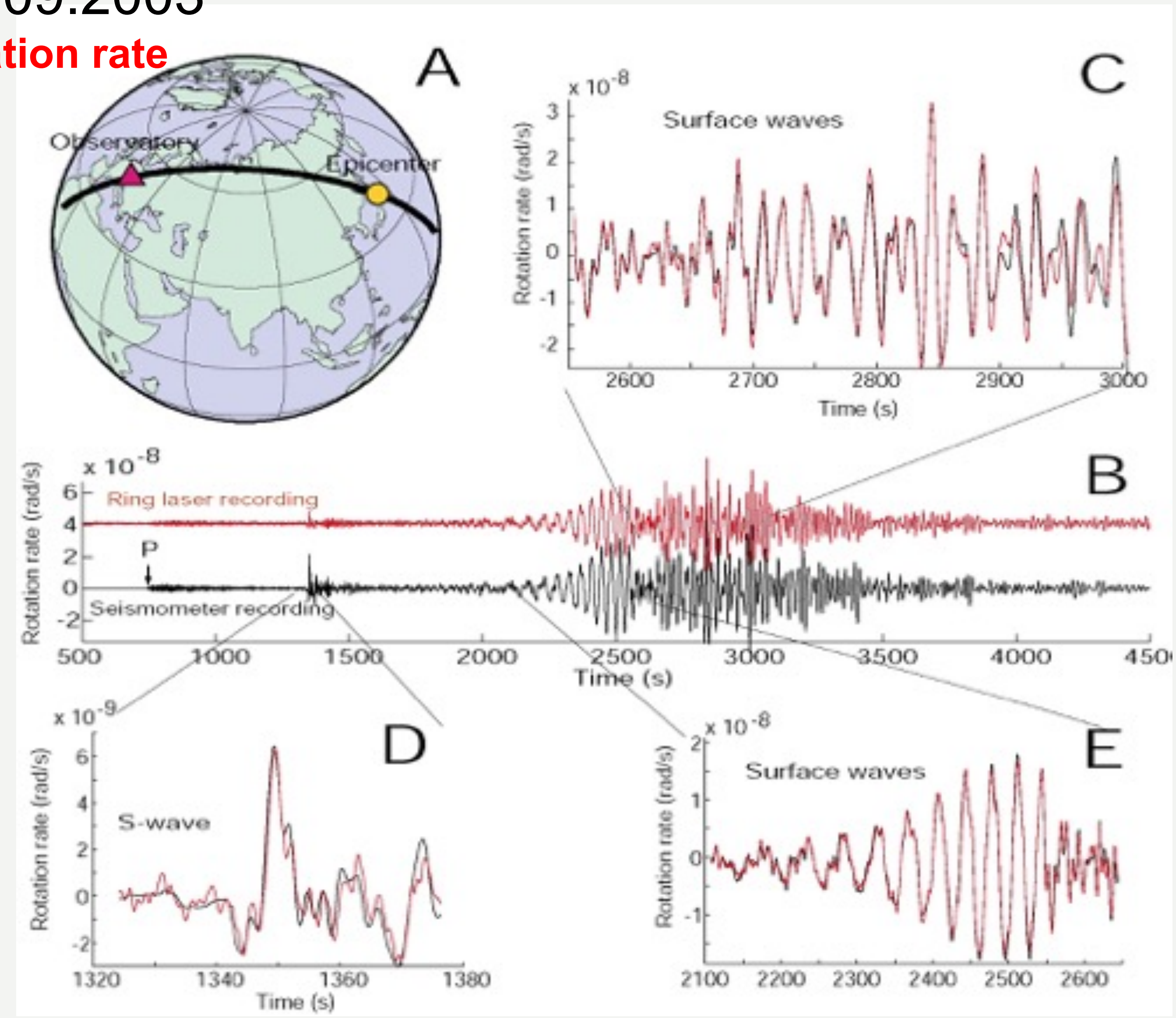
$$\Rightarrow \quad a(x,t) / \dot{\Omega}(x,t) = -2c$$

Rotation rate and acceleration should be in phase and the amplitudes scaled by two times the horizontal phase velocity

Correlation … Applications

Mw = 8.3 Tokachi-oki 25.09.2003

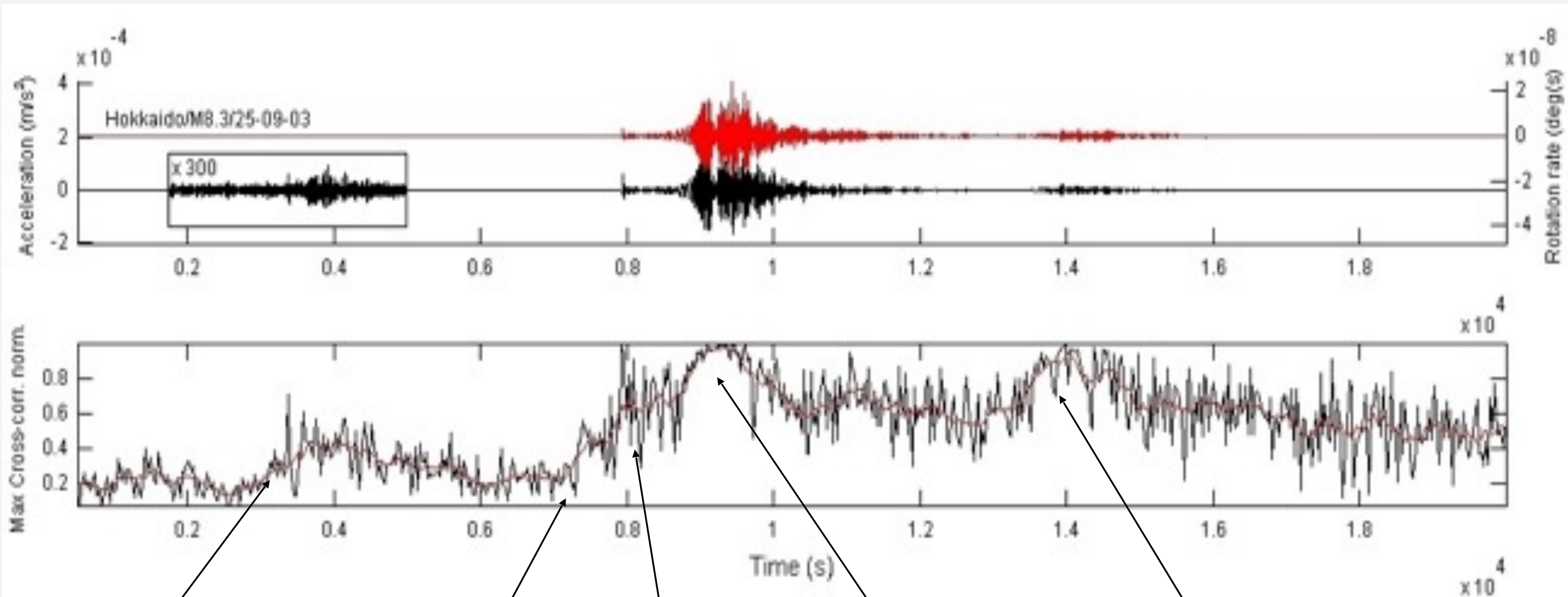
transverse acceleration – rotation rate



From Igel et al., GRL, 2005

Correlation … Applications

Max. cross-corr. coefficient in sliding time window **transverse acceleration** – **rotation rate**



Small teleseismic event

P-onset

S-wave

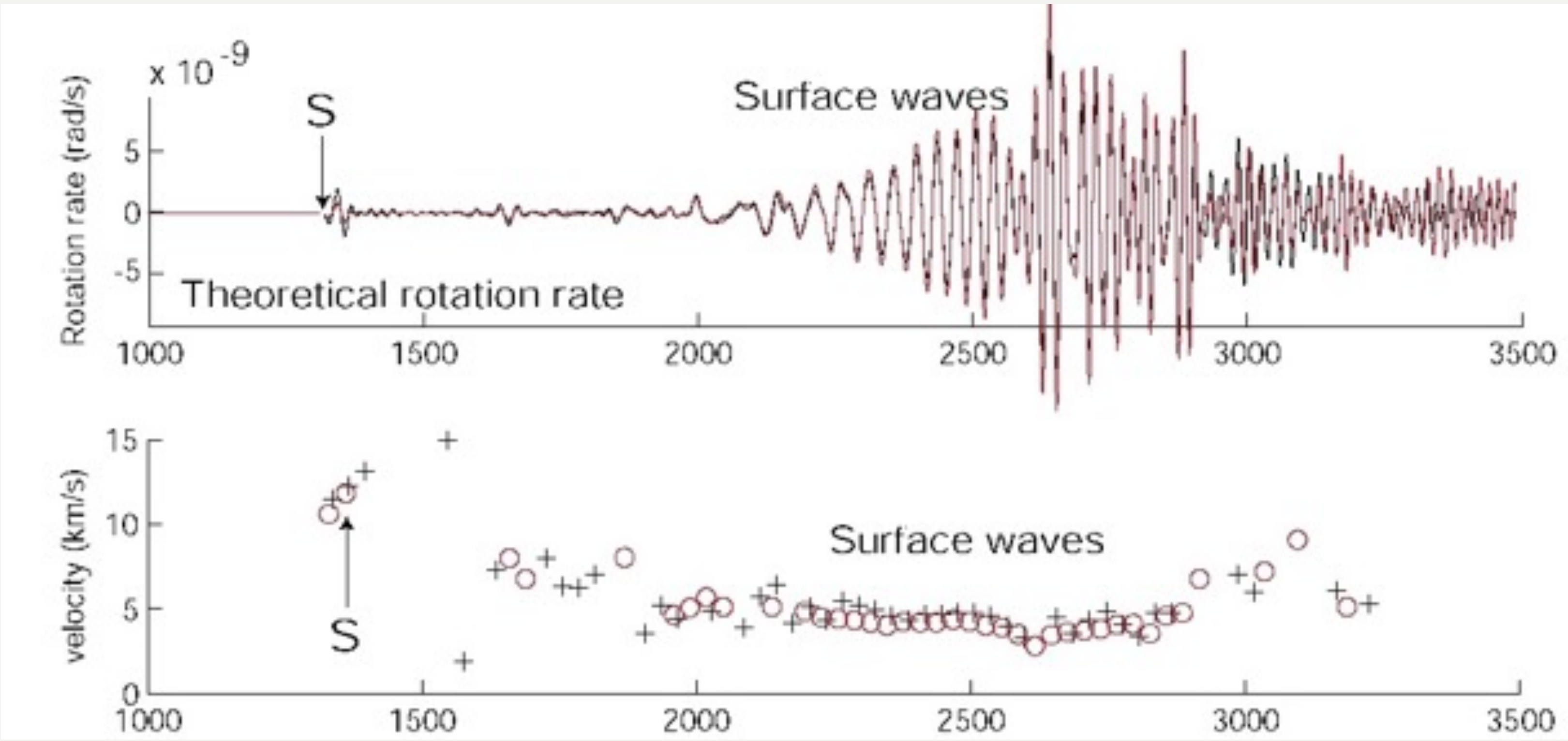
Love waves

Aftershock

Correlation … Applications

M8.3 Tokachi-oki, 25 September 2003

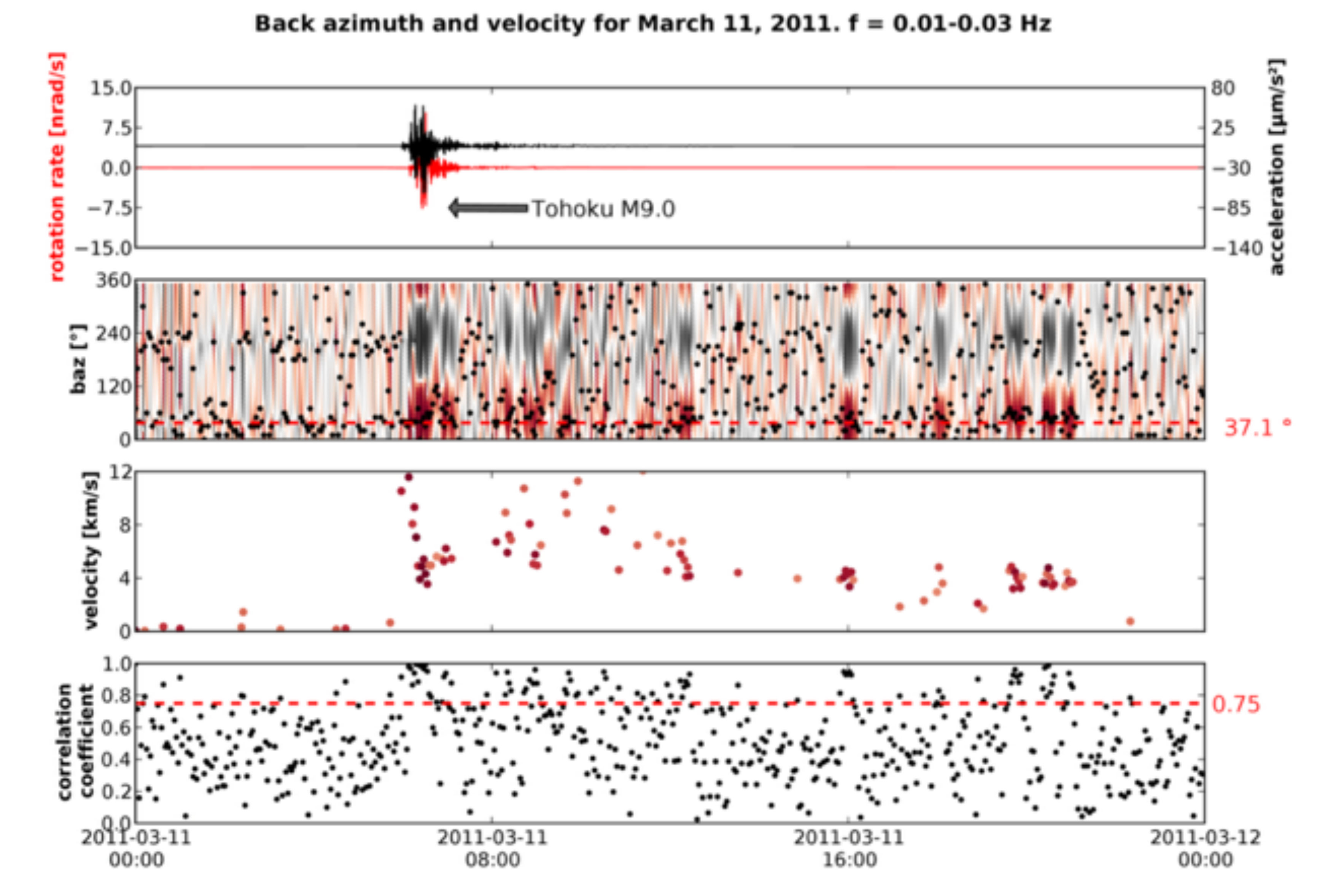
phase velocities (+ observations, o theory)



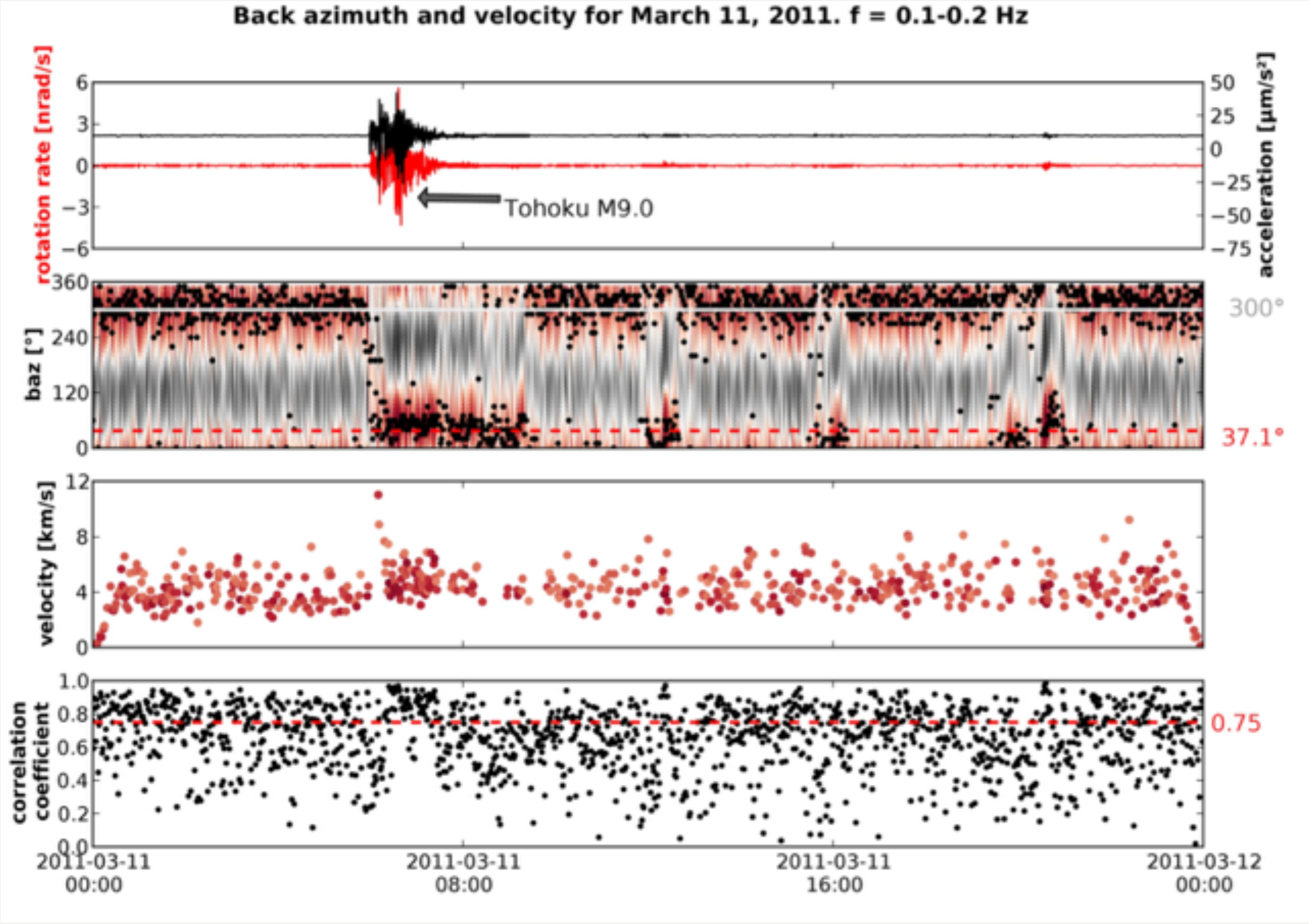
Horizontal phase velocity in sliding time window

From Igel et al. (GRL, 2005)

Correlation … Applications



Correlation … Applications



Correlation Summary

- *Correlation plays a central role in the study of time series. In general, correlation gives a quantitative estimate of the degree of similarity between two functions, as well as their lag time.*

- *Correlation in the time domain is multiplication with the complex conjugate in the frequency domain*

- *Correlation of noisy seismograms from two stations allows in principle the reconstruction of the Green's function between the two stations*

- *The ideal tool to quantify similarity (e.g., frequency dependent) between various signals (e.g., rotations, strains with translations)*

- *Correlation has many applications in geophysics*