Christian-Albrechts-Universität zu Kiel

Signal Processing for Medical Applications –

Frequency Domain Analyses

Muthuraman Muthuraman

Christian-Albrechts-Universität zu Kiel
Department of Neurology / Faculty of Engineering
Digital Signal Processing and System Theory

Coherence

- The coherence analysis is an extensively used method to study the correlations in frequency domain, between two simultaneously measured signals.
- Let $x(t)$ and $y(t)$ be two simultaneously recorded data sets of length N
- We estimate the short-time power spectra of s_{xx}, s_{yy}, and cross-spectrum, s_{xy} which is the Fourier transform of the cross-correlation function of the signals $x(t)$ and $y(t)$ in each segment.
- Finally , we average the power spectra and the cross-spectrum across all the segments and calculate the coherence as follows:

$$C(\omega) = \frac{\left|s_{xy}(\omega)\right|^2}{s_{xx}(\omega)\, s_{yy}(\omega)} \qquad (4.1)$$

Significance of Coherence

- The coherence spectrum represents the strength of correlation between two signals, $x(t)$ and $y(t)$.
- The confidence limit for coherence at the $100\%\alpha$ is given by

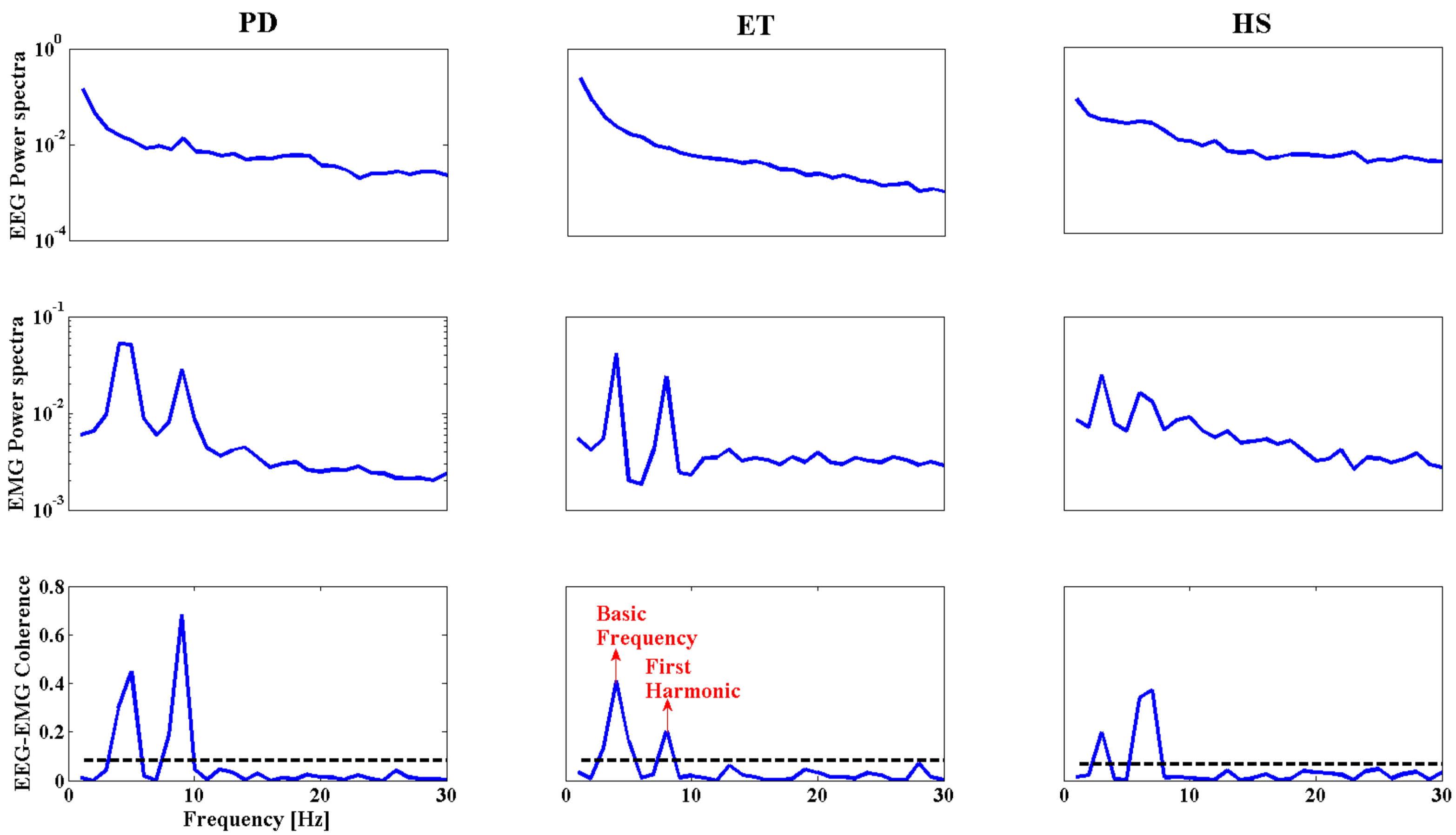
$$C_L = 1-(1-\alpha)^{1/(M-1)} \tag{4.2}$$

α= 0.99; and M is the number of disjoint segements;

hence the confidence limit is $1-0.01^{1/(M-1)}$.

- If the signal length N =10000; D=1000; C_L = ?;
- Frequency resolution ➡ If f_s (i.e. the number of data points sampled per second) is the sampling frequency, then the frequency resolution is f_s/D .
- Thus, one should optimally choose the value of D depending on the purpose of analysis, to compromise between sensitivity and reliability.

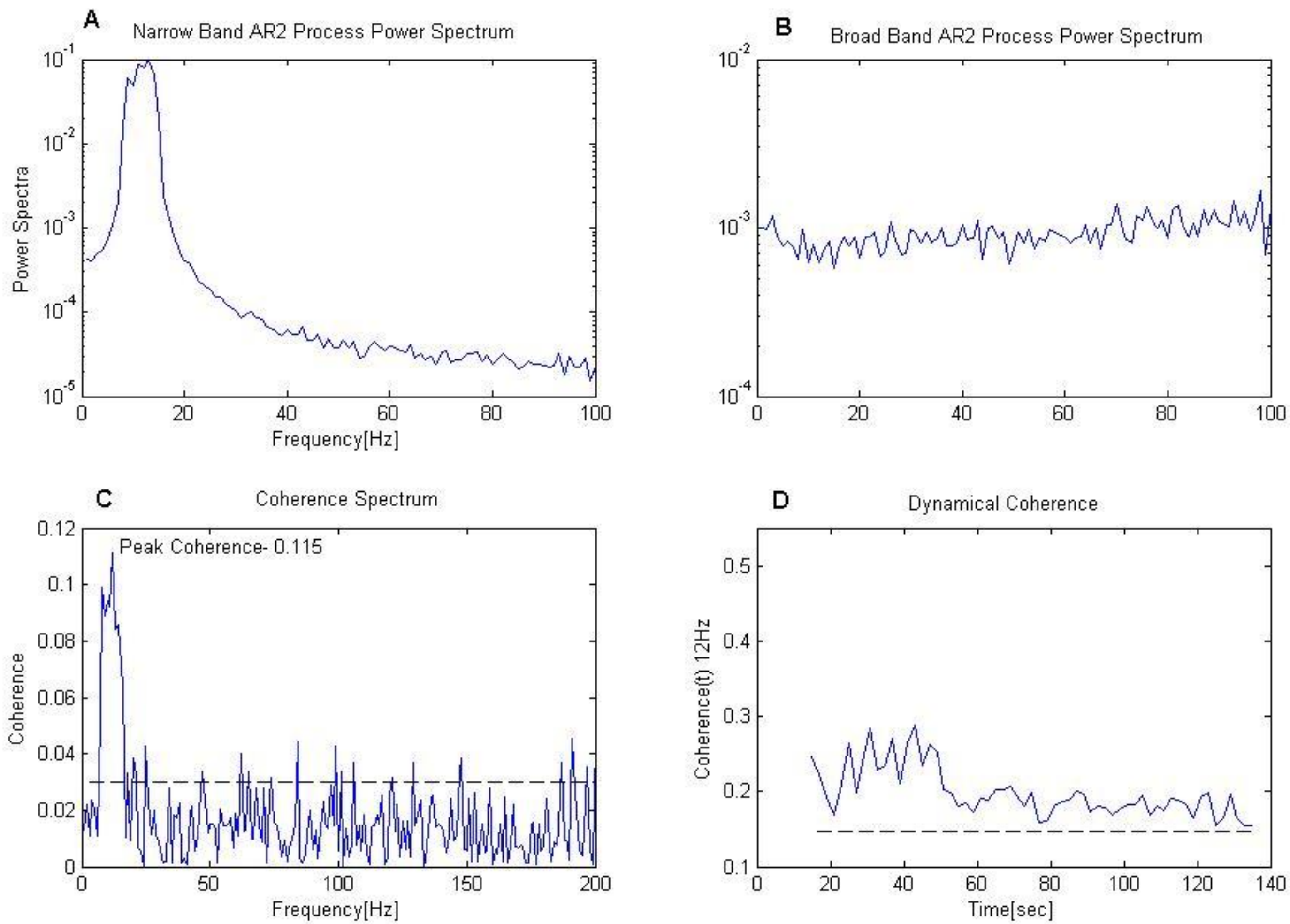
Significance of Coherence



Dynamical Coherence

- The dynamical coherence analysis is done by estimating the coherence spectra for a moving 30-second windows with an overlap of 28-seconds, resulting in an apparent time resolution of 2s.
- A model is created by coupling two AR2 processes $y[n]=a_1y[n-1]+a_2y[n-2]+\eta[n]$. One AR2 (V1) had narrow band characteristics $(a_1=1.9691, a_2=-0.9753)$ and the other (V2) had broadband spectral characteristics $(a_1=0.37486, a_2=-0.36788)$.
- These two processes were simulated for a duration of 150 seconds at a sampling rate of 1,000 Hz. The narrow band AR2 was then band-pass filtered around its spectral peak between 8 and 15 Hz and then combined by point-by-point summation with the broadband AR2 (V2) as follows: V=V2+0.2 V1.
- Independent white noises were added to V and V1 whose amounts were tuned so that overall coherence between V and V1 was around 0.1.

Dynamical Coherence

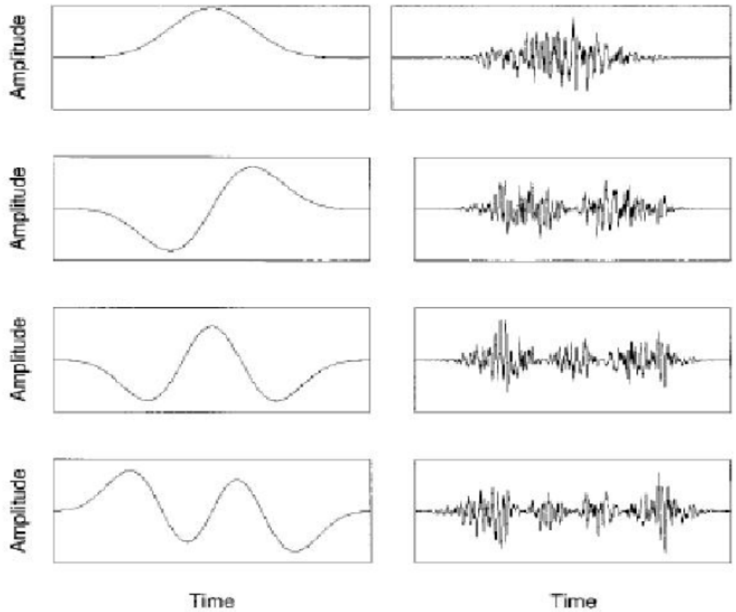


Welch periodogram and Multitaper Method

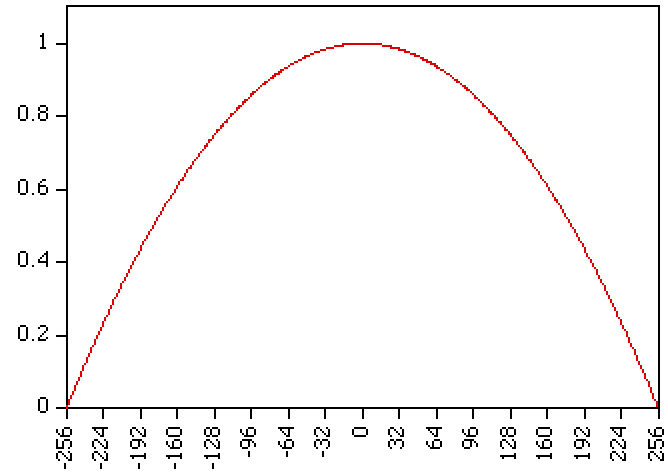
- Estimation of coherence with these two methods which work on the same principle

➡ Multitaper Method with single hanning taper ➡ Hanning window in Welch periodogram Method

Tapers



Hanning Window



Increasing Time Resolution

- If we consider a finite length sample of a discrete time process $x(t)$, $t = 1, 2, \cdots N$. Let us assume a spectral representation for the process,

$$x(t) = \int_{-1/2}^{1/2} X(f) \exp(2\pi i f t)\, df \tag{4.3}$$

- The Fourier transform $\tilde{x}(f)$ of the data sequence is therefore given by

$$\tilde{x}(f) = \sum_{1}^{N} x(t) \exp(-2\pi i f t) = \int_{-1/2}^{1/2} K(f - f', N) X(f')\, df' \tag{4.4}$$

- The Welch periodogram method is capable of also analysing the signals using larger time windows in the time domain which inturn gives a good estimation of the signal components.

- However, when shorter time windows need to be used, some disadvantages of this method become evident when applied to non-linear signals.

- In this case for a stationary process, the spectrum is given by

$$S(f)\, df = E\left[\left| X(f) \right|^2 \right] \tag{4.5}$$

Increasing Time Resolution

- A simple estimation of the spectrum (apart from the normalization constant) is obtained by squaring the Fourier transform of the data sequence, i.e; $|\tilde{x}(f)|^2$.

- This suffers from two difficulties:

- Firstly, $\tilde{x}(f)$ is not equal to $X(f)$, except when the data length is infinite, in which case the kernel in equation (4.5) becomes a delta function. Rather it is related to $X(f)$ by a convolution as given in equation (4.4).

- This problem is usually referred to as „bias“ corresponding to a mixing of information from different frequencies of the underlying process due to a finite window length.

- Secondly, if the data are stochastic, then the squared Fourier transform of a time series is an inconsitent estimator of the spectrum, because it does not converge to the „true“ spectrum when the data series tends to infinite length.

- Inorder to overcome all these disadvantages, the signals can be analysed with the multitaper and the extended continous wavelet-transform method.

Multitaper Method

- If $x(t)$ is the signal then the spectrum in this method is calculated by multiplying the data with several orthogonal tapers (windows)

$$S_{MT}(\omega) = \frac{1}{K}\sum_{k=1}^{K}\left|\tilde{X}_k(\omega)\right|^2 \qquad (4.6)$$

where $\tilde{X}_k(\omega)$ is the Fourier transform $\left[X(\omega) = \int_{-\infty}^{\infty} x(t)\exp(-2\pi i\omega t)dt\right]$ of the data $x(t)$

$$\tilde{X}_k(\omega) = \sum_{t=1}^{N} w_t(k)\, x_t \exp(-2\pi i\omega t) \qquad (4.7)$$

where $w_t(k)(k=1,2,\ldots K)$ are the K orthogonal tapers.

- A particular choice of these taper functions, with optimal spectral concentration properties, is given by the discrete prolate spheroidal sequences (DPSS).

Multitaper Method

- Let $w_t(k,W,N)$ be the k^{th} DPSS of length N and frequency bandwidth parameter W

- Consider a sequence w_t of length N whose Fourier transform is given by $U(\omega)=\sum_1^N w_t \exp(-2\pi i\omega t)$, we find the sequences w_t so that the spectral amplitude $U(\omega)$ is maximally concentrated in the interval $[-W,W]$, i.e.

$$\int_{-W}^{W} |U(f)|^2 df \qquad (4.8)$$

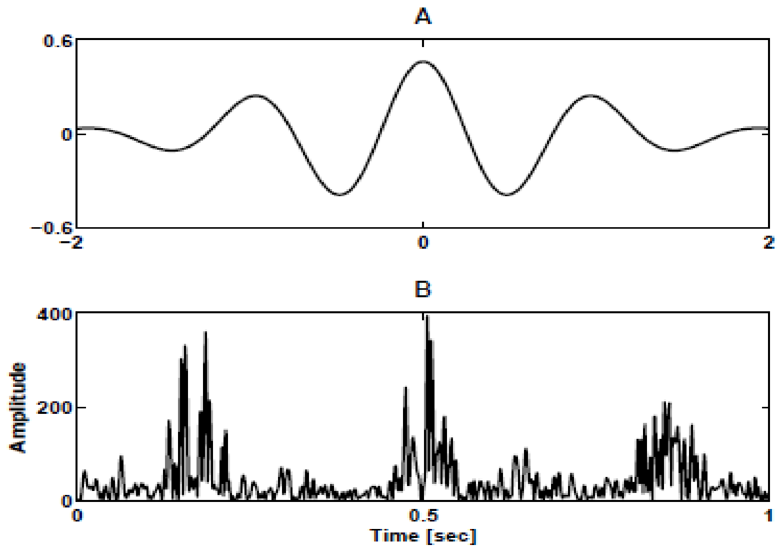
is maximised.

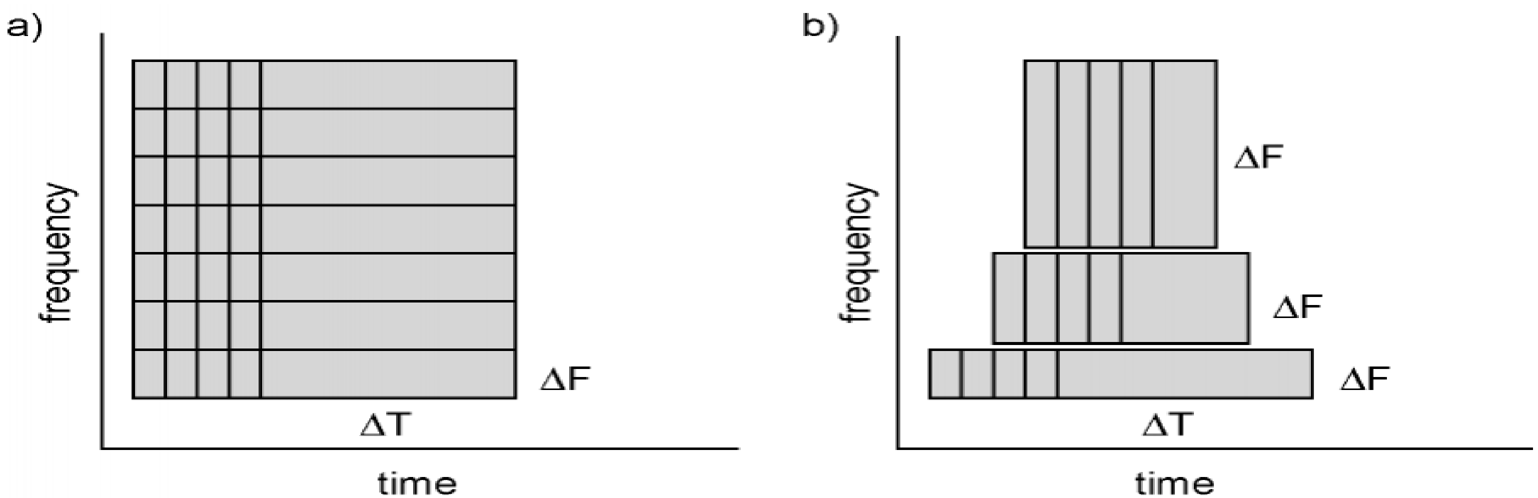
- The maximisation problem leads to the matrix eigenvalue equation

$$\sum_{t'}^{N} \frac{\sin[2\pi W(t-t')]}{\pi(t-t')} w_{t'} = \lambda w_t \qquad (4.9)$$

- Eigen vector – Let A be a square matrix, a non-zero vector C is called a eigen vector of A if and only if there exists a number (real/complex) λ such that $AC=\lambda C.$

Extended Continous Wavelet Transform Method





Extended Continous Wavelet Transform Method

- If $x(t)$ is the signal then the continuous wavelet transform can be written as

$$CWT_x(\tau, a) = \frac{1}{a}\int x(t)\, h^*\left(\frac{t-\tau}{a}\right) dt \qquad (4.10)$$

where h^* is the wavelet function and a is the scaling factor.

- The wavelet used in this method is the Morlet wavelet which is defined as

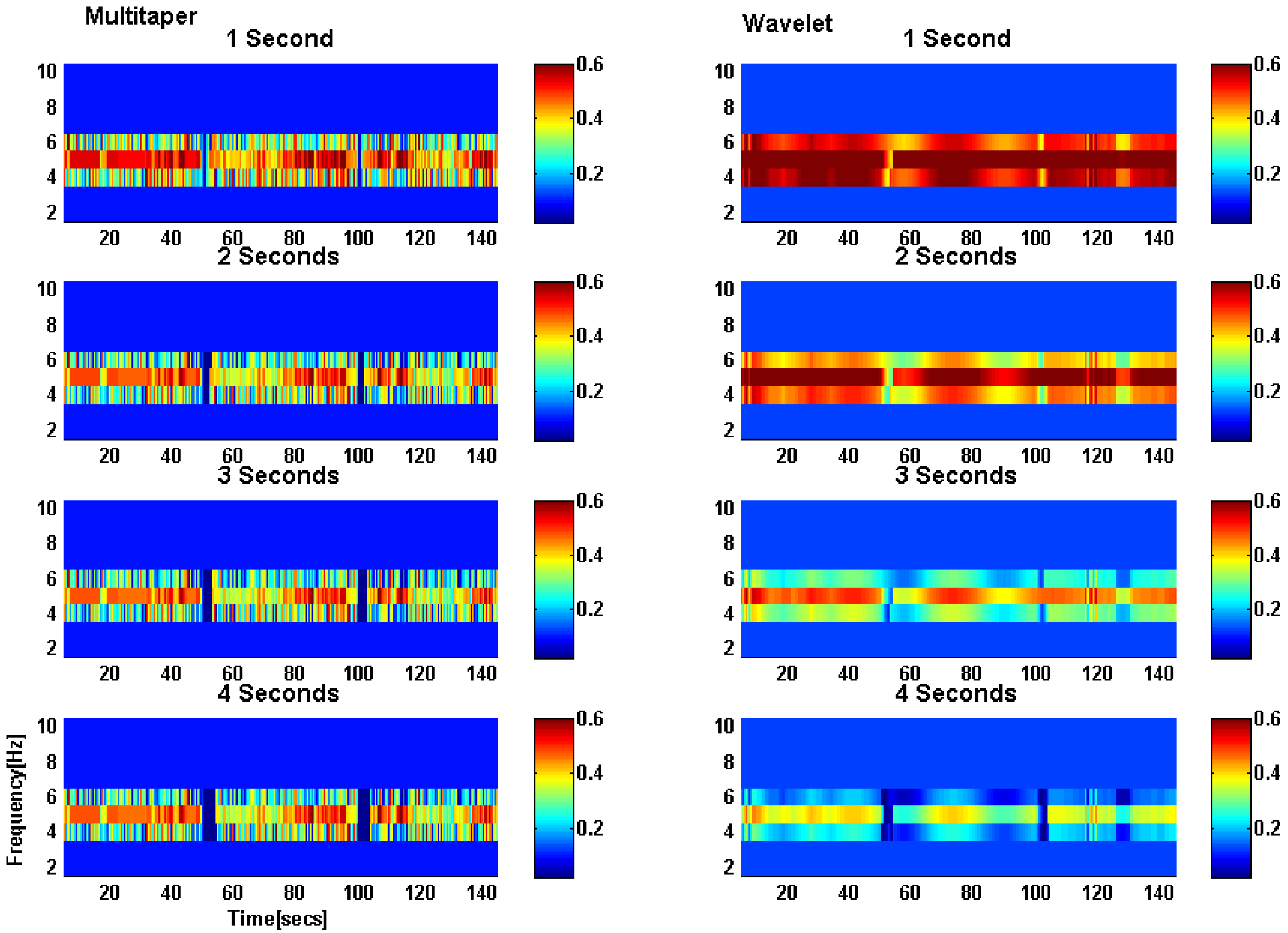
$$h(t) = \exp(jct)\exp(-\beta t^2/2) \qquad (4.11)$$

- The relative bandwidth of this wavelet can be defined as

$$BW_{rel} = \frac{2\sqrt{2\beta}}{c} \qquad (4.12)$$

- By adjusting the ratio of $\sqrt{\beta}/c$ which gives the flexibility in having a particular frequency resolution at a particular frequency.

Model data


Multitaper
Wavelet
1 Second
2 Seconds
3 Seconds
4 Seconds
Frequency[Hz]
Time[secs]
10
8
6
4
2
20
40
60
80
100
120
140
0.6
0.4
0.2

Partial Coherence

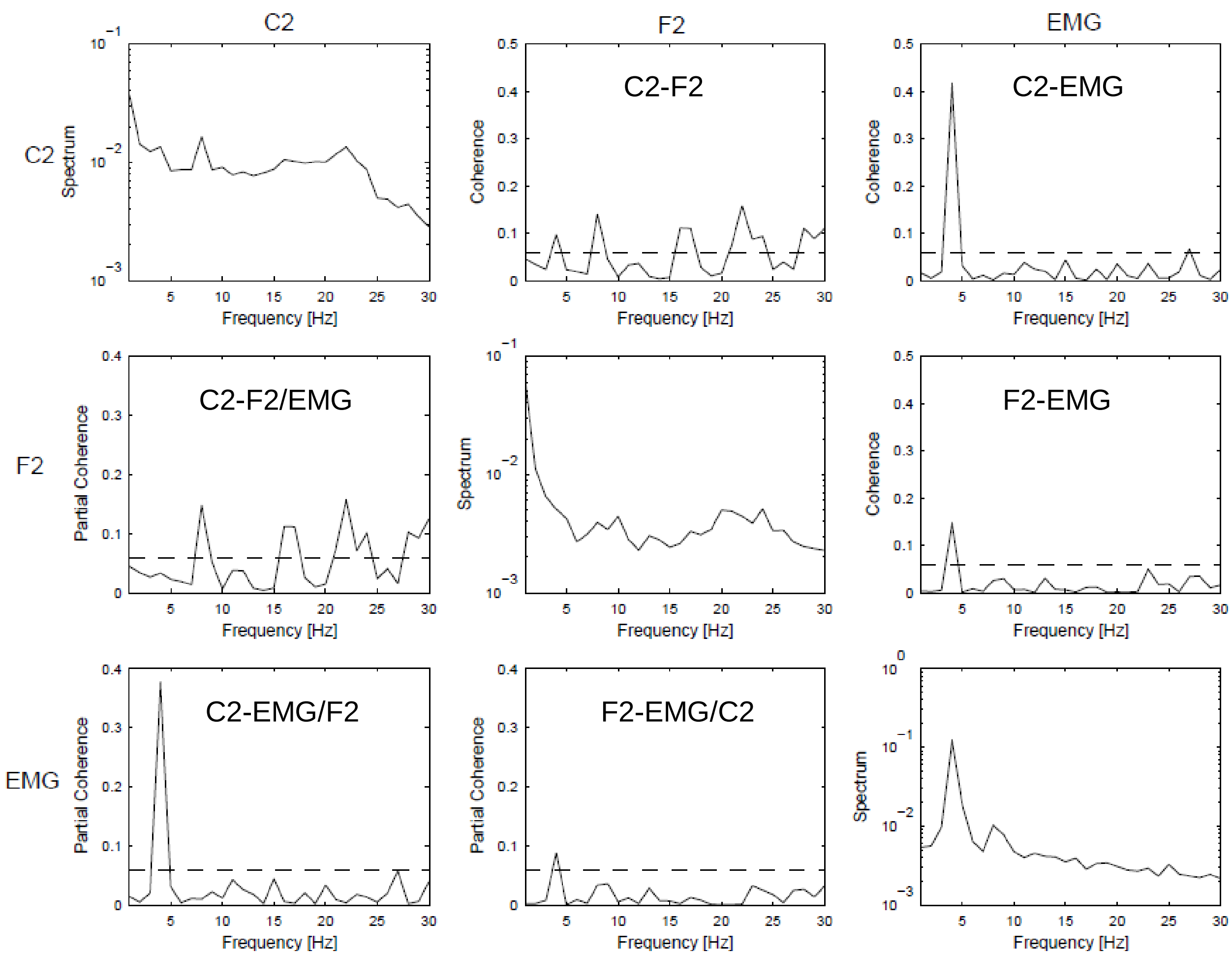
- Let $x(t), y(t)$ and $z(t)$ be three simultaneously measured signals of length N.
- The partial coherence is estimating by first calculating the power spectra S_{xx}, S_{yy}, S_{zz} and cross-spectra S_{xy}, S_{xz} and S_{yz} in each of the disjoint windows. Finally we average these quantities across all the segments to get the estimate of the same.
- We estimate the partial coherence between the signals $x(t), y(t)$ and $z(t)$ as follows:

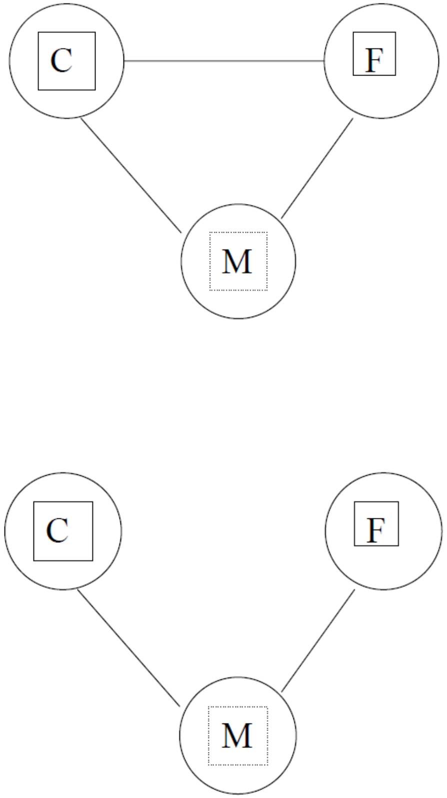
$$C_{xy/z}(\omega) = \frac{\left| CY_{xy}(\omega) - CY_{xz}(\omega)\, CY_{zy}(\omega) \right|^2}{(1 - C_{xz}(\omega))(1 - C_{zy}(\omega))} \qquad (4.13)$$

where $CY_{ij}(\omega)$ is a complex-valued function whose magnitude is called coherency between the two signals i and j.

- The confidence limit for the partial coherence at $100\%\ \alpha$ is $1-(1-\alpha)^{1/(M-2)}$.

Partial Coherence





Methods for finding direction of information flow

Non-Parametric Methods	**Parametric Methods** **(Based on modeling of system by linear VAR processes)**
Partial Cross spectrum	Granger Causality Index
Partial Coherence	Directed Transfer Function
	Partial Directed Coherence

Partial Directed Coherence

- Partial directed coherence (PDC) can be formulated as follows:

$$x(t) = \sum_{r=1}^{p} a(r) \cdot x(t-r) + \eta(t) \qquad (4.14)$$

$$A(\omega) = I - \sum_{r=1}^{p} a(r) \cdot e^{-i\omega r} \qquad (4.15)$$

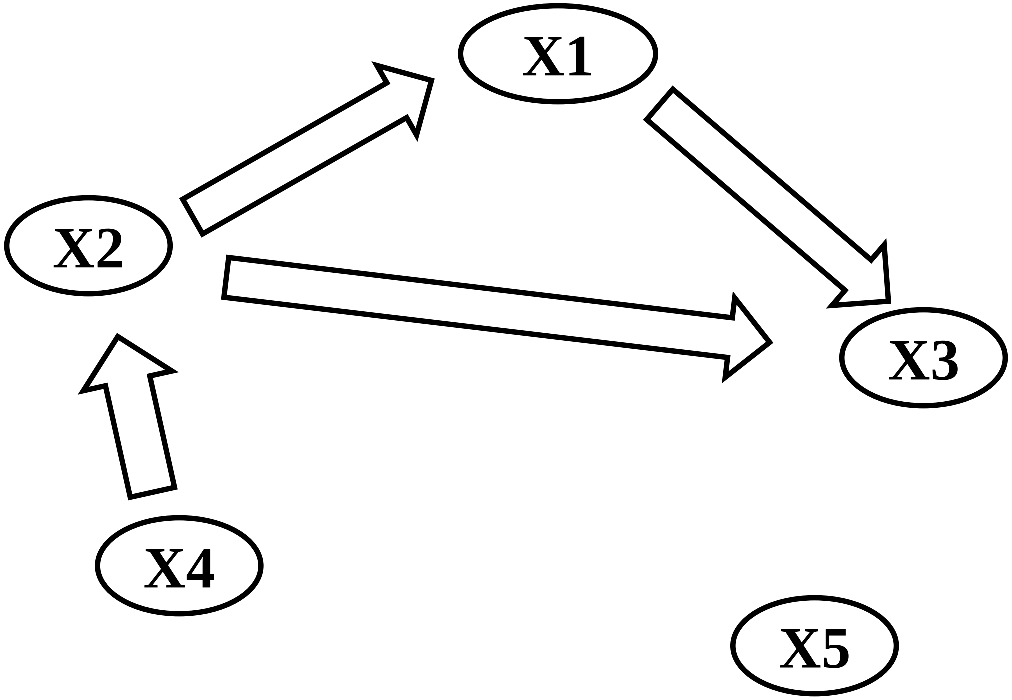
PDC:

$$\left|\pi_{i \leftarrow j}(\omega)\right| = \frac{\left|A_{ij}(\omega)\right|}{\sqrt{\sum_k \left|A_{kj}(\omega)\right|^2}} \qquad (4.16)$$

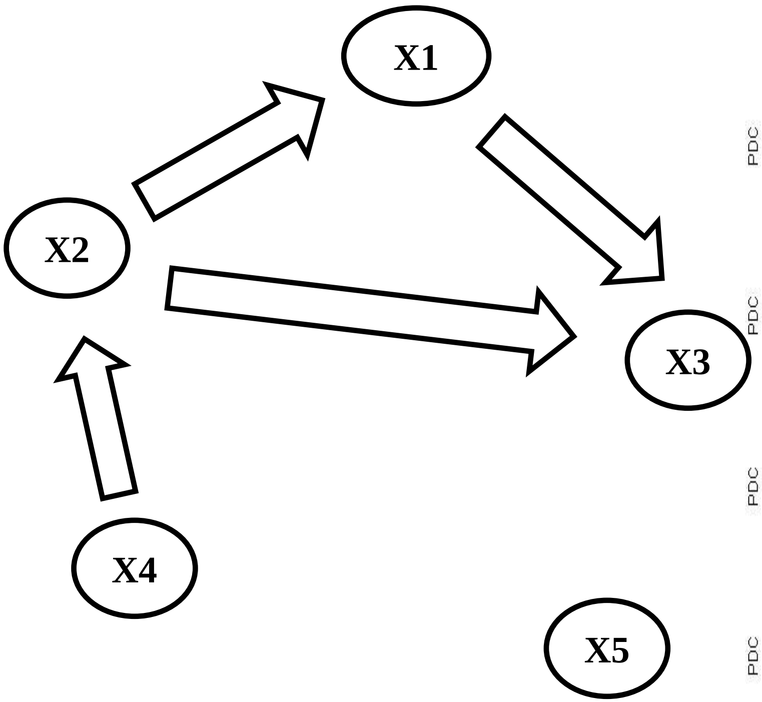
- Inorder to calculate PDC, we need to find the appropriate order ‚p‘ for the underlying process.
- Optimum order can be found by Akaike Information Criterion (AIC).
- After finding the approprite ‚p‘, we need to find the coefficients for autoregressive model which closely depicts the process.

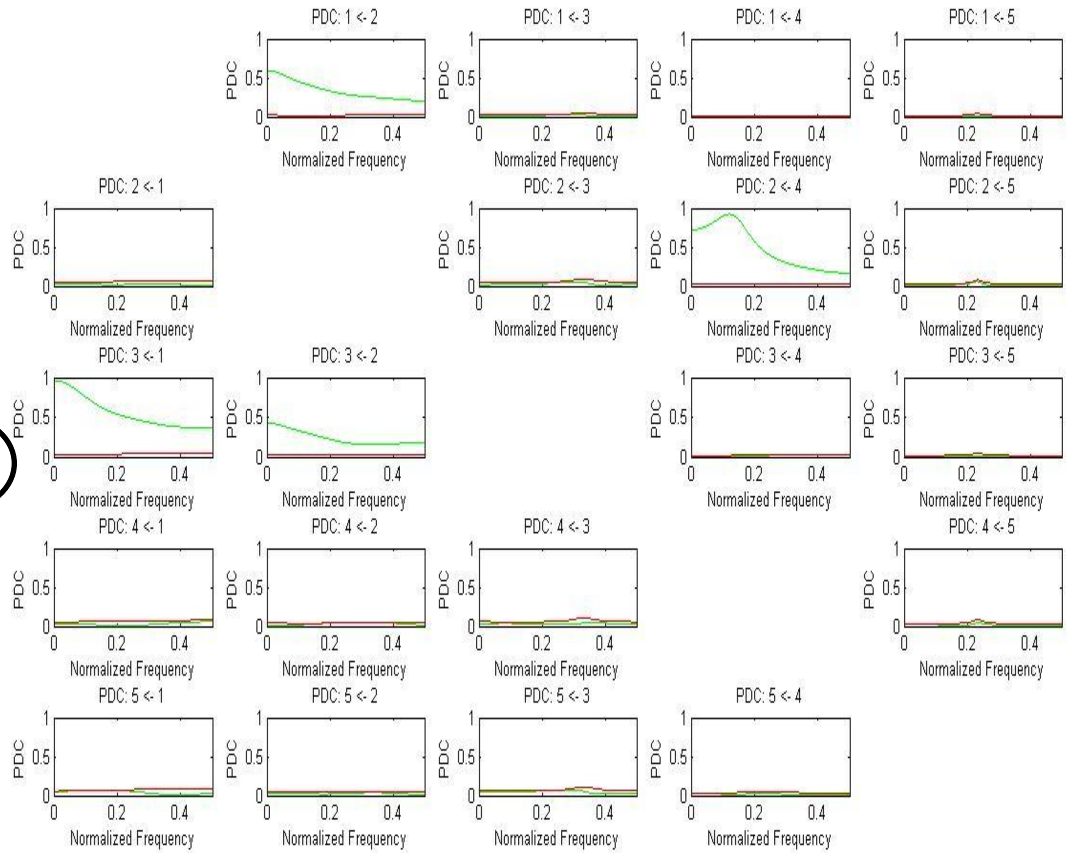
Partial Directed Coherence

- Application of PDC on model data with AR process of order p=5
- The five time series as the information flow between them as given in the figure below:



Partial Directed Coherence





Granger Causality Index

- The principle of Granger causality states that if some series $y(t)$ contains information in past terms that helps in the prediction of series $x(t)$, then $y(t)$ is said to cause $x(t)$.

- For predicting a value of $x(t)$ using p previous values of the series X only, we get a prediction error e:

$$X(t) = \sum_{i=1}^{p} A_{11}^{'}(j)\, X(t-j) + e(t) \qquad (4.17)$$

- If we try to predict a value of $X(t)$ using p previous values of the series X and p previous values of Y we get another predicton error e_1:

$$X(t) = \sum_{j=1}^{p} A_{11}(j)\, X(t-j) + \sum_{j=1}^{p} A_{12}(j) Y(t-j) + e_1(t) \qquad (4.18)$$

- If the variance of e_1 (after including series Y to the prediction) is lower than the variance of e we say that Y causes X in the sense of Granger causality.

Granger Causality Index

- In the same way we can say that X causes Y in the sense of Granger causality when the variance of e_2 is reduced after including series X in the prediction of series Y:

$$Y(t) = \sum_{j=1}^{p} A_{22}(j)\, X(t-j) + \sum_{j=1}^{p} A_{21}(j) Y(t-j) + e_2(t) \tag{4.19}$$

- Granger causlity index is based directly on the defintion of causality, namely it shows, if the information contributed by second channel improves the prediction of the first channel.

- The logarithm ratio of the residual variances for one and two-channel models is computed:

$$GCI_{1\to 2} = \ln(e/e_1) \tag{4.20}$$

- This defintion can be extended to multichannel case by considering how the inclusion of the given channels changes the residual variance ratios.

- GCI is an estimator in the time domain.

Multivariate autoregressive Model

- Granger causality was defined for two channels, however, he later stated that the causality principle holds only, if there are no other channels influencing the process.

- To account for the whole multivariate structure of a process of k channels the multichannel autoregressive model (MVAR) has to be considered. For the MVAR k- channel process $X(t)$:

$$X(t) = (X_1(t), X_2(t), \ldots, X_k(t)) \qquad (4.21)$$

  The model takes the form

$$X(t) = \sum_{i=1}^{p} A(j)\, X(t-j) + E(t) \qquad (4.22)$$

  where $E(t)$ are vectors of size k and the coefficients A are $k \times k$- sized matrices

- Equation (4.22) can be easily transformed to describe relations in the frequency domain. After changing the sign of the A and application of Z transform we get:

$$E(f) = A(f)\, X(f) \qquad (4.23)$$

Multivariate autoregressive Model

- It can be derived as follows:

$$X(f) = A^{-1}(f)E(f) = H(f)E(f) \tag{4.24}$$

$$H(f) = \left(\sum_{m=0}^{p} A(m)\exp(-2\pi i m f \Delta t) \right)^{-1} \tag{4.25}$$

- From the form of the above equations we can consider the model as a linear filter with white noises $E(f)$ on its input and the signals $X(f)$ on its output. The matrix of filter coefficients $H(f)$ is called the transfer matrix of the system.

- It contains information about all realtions between channels in the given set including the phase relations between the signals.

Directed Transfer Function

- The formulation of the DTF is based on the properties of the transfer function of multivariate autoregressive process as:

$$DTF^2_{j\to i}(f) = \frac{|H_{ij}(f)|^2}{\sum_{m=1}^{k} |H_{im}(f)|^2} \qquad (4.24)$$

- The DTF describes causal influence of channel j on channel i at frequency f.
- The above equation defines a normalized version of DTF, which takes values from 0 to 1 producing a ration between the inflow of from channel j to channel i to all the inflows to channel i.