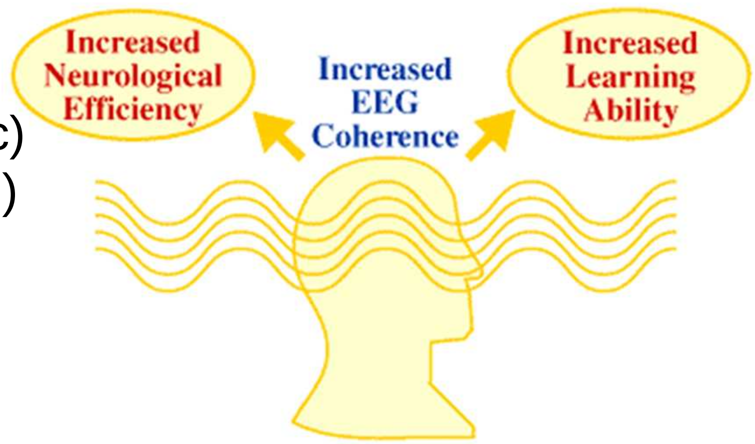
Analisi spettrale multivariata

Spesso interessa quantificare la relazione mutua tra due segnali, che possono essere sia di natura diversa, es: variabilità del battito cardiaco e pressione sanguigna, che della stessa natura, es. EEG prelevato da diverse posizioni dello scalpo.

Con riferimento al secondo esempio, l'analisi di coerenza permette di capire se due aree della corteccia cerebrale sono funzionalmente correlate, ad es. se hanno generatori sottocorticali comuni. E' stata impiegata fin dagli anni 60 in molteplici studi, tra cui

- fasi del sonno
- disordini psichiatrici
- effetto di farmaci
- patologie (emicrania, ischemia cerebrale ecc)
- in assenza di gravità (missione Gemini GT-7)
- meditazione trascendentale
- processi cognitivi



Se interpretiamo i due segnali come realizzazioni di due processi aleatori X e Y a tempo discreto, stazionari e a media nulla, la relazione mutua tra i due processi può essere quantificate sia nel dominio del tempo che della frequenza usando varie funzioni

- Cross-correlazione $R_{xy}(k) = E[x(n)\, y(n+k)]$

- Cross spettro $P_{xy}(\omega) = FT[R_{xy}(k)]$

- Coerenza $$C^2_{xy}(\omega) = \frac{\left|P_{xy}(\omega)\right|^2}{P_x(\omega) \cdot P_y(\omega)}$$

Cross correlazione: proprietà

La cross correlazione non è in generale una funzione pari, ma dall'ipotesi di stazionarietà segue

$$\boxed{R_{xy}(k)=E[x(n)y(n+k)]=E[x(n-k)y(n)]=E[y(n)x(n-k)]=R_{yx}(-k)}$$

1. Processi indipendenti (e a media nulla) $\boxed{R_{xy}(k)=0}$

2. Processi coincidenti X=Y $\boxed{R_{xy}(k)=R_x(k)=R_y(k)}$

3. Processi legati da una relazione lineare $\boxed{R_{xy}(k)=h(k)\otimes R_x(k)}$

$$y(n)=\sum_{j=-\infty}^{n} h(j)x(n-j)$$

Dim:

$$R_{xy}(k)=E\left[x(n)\sum_j h(j)x(n+k-j)\right]=$$

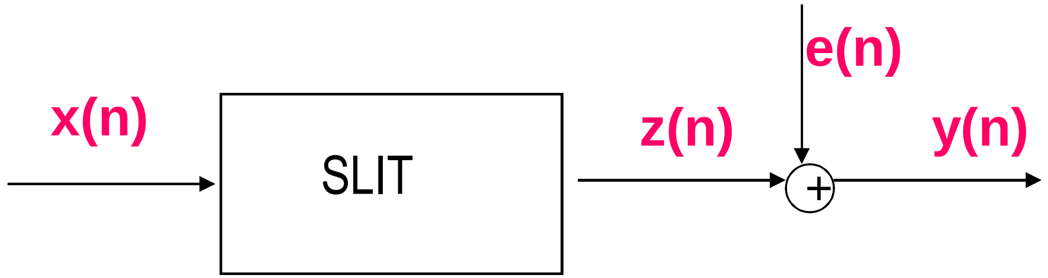
$$=\sum_j h(j)E[x(n)x(n+k-j)]=\sum_j h(j)R_x(k-j)=$$

$$=h(k)\otimes R_x(k)$$

Cross correlazione: proprietà

4. Processo Y è la somma di due componenti, la prima (Z) legata da una relazione lineare al processo X, l'altra (E) indipendente da X

$$\boxed{R_{xy}(k) = h(k) \otimes R_x(k)}$$



$$y(n) = z(n) + e(n) = \sum_{k=-\infty}^{n} h(k)x(n-k) + e(n) = h(n) \otimes x(n) + e(n)$$

Dim:

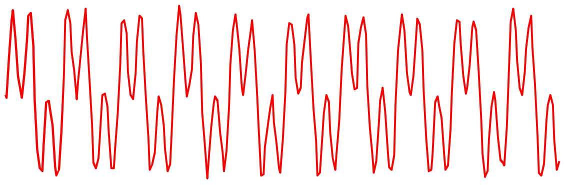
$$R_{xy}(k) = E[x(n)(z(n+k) + e(n+k))] = E[x(n)z(n+k)] + E[x(n)e(n+k)] =$$
$$= R_{xz}(k) = h(k) \otimes R_x(k)$$

La cross correlazione isola il legame tra X e Z, cioè tra X e la componente di Y linearmente legata a X

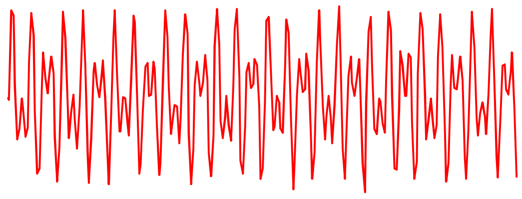
Cross spettro

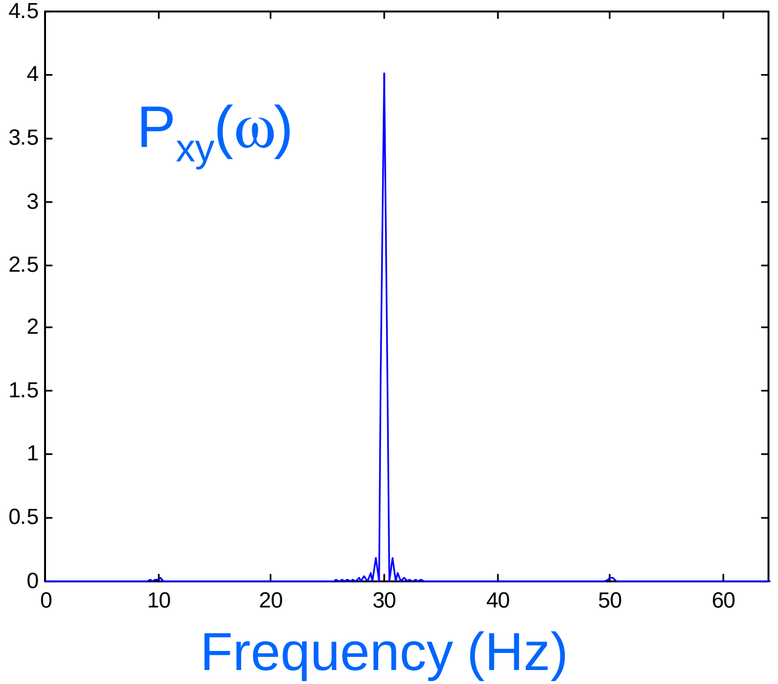
E' la FT della funzione di cross correlazione. Evidenzia le componenti spettrali comuni ai due processi. Per esemplificare, se consideriamo due segnali deterministici che hanno in comune una componente a 30 Hz

$x(t)=A_1\sin(2\pi 10t)+A_2 \sin(2\pi 30t)$

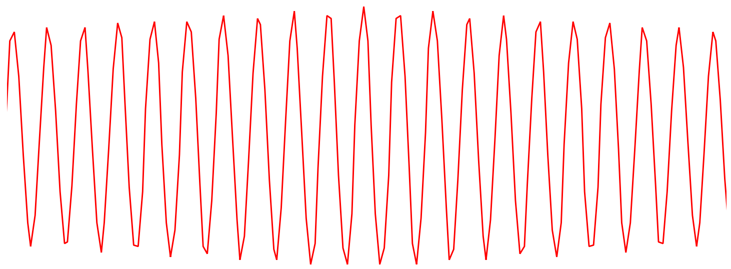


$y(t)=A_1\sin(2\pi 30t)+ A_2\sin(2\pi 50t)$





Il cross spettro evidenzia la componente a 30 Hz



Cross spettro: proprietà

E' in generale una funzione complessa, dato che la cross correlazione non è in generale pari

1. Processi indipendenti (e a media nulla) $\boxed{P_{xy}(\omega)=0}$

2. Processi coincidenti X=Y $\boxed{P_{xy}(\omega)=P_x(\omega)=P_y(\omega)}$

3. Processi legati da una relazione lineare $\boxed{P_{xy}(\omega) = H(\omega)P_x(\omega)}$

4. Processo Y somma di due componenti, la prima (Z) legata da una relazione lineare al processo X, l'altra (E) indipendente da X $\boxed{P_{xy}(\omega) = P_{xz}(\omega) = H(\omega)P_x(\omega)}$

Il cross spettro coincide con lo spettro della componente di Y linearmente legata a X.

Funzione coerenza

$$C_{xy}^2(\omega) = \frac{|P_{xy}(\omega)|^2}{P_x(\omega)\cdot P_y(\omega)} = C_{yx}^2(\omega)$$

E' una funzione reale, che normalizza il modulo al quadrato dello spettro incrociato tra due processi rispetto agli spettri dei due processi

1. Processi indipendenti (e a media nulla) $\boxed{C_{xy}^2(\omega)=0}$

2. Processi coincidenti X=Y $\boxed{C_{xy}^2(\omega)=1}$

3. Processi legati da una relazione lineare $\boxed{C_{xy}^2(\omega)=1}$

Dim: $C_{xy}^2(\omega) = \frac{|P_{xy}(\omega)|^2}{P_x(\omega)\cdot P_y(\omega)} = \frac{|H(\omega)P_x(\omega)|^2}{P_x(\omega)\cdot|H(\omega)|^2 P_x(\omega)} = 1$

La funzione coerenza è una funzione reale di ω, compresa tra zero (processi indipendenti) e 1 (processi linearmente legati)

Funzione coerenza

4. Processo Y somma di due componenti, la prima (Z) legata da una relazione lineare al processo X, l'altra (E) indipendente da X

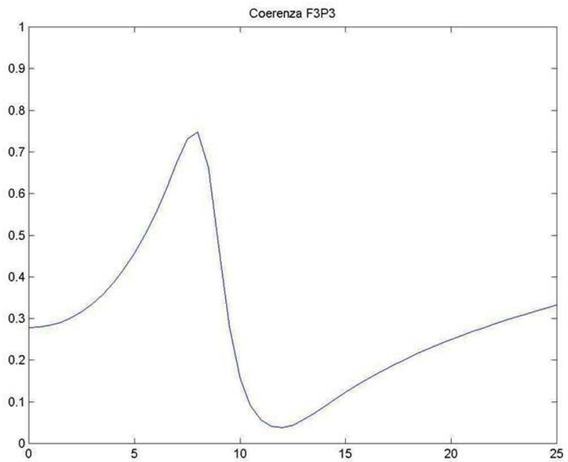
$$\boxed{C_{xy}^2(\omega) = \frac{\left|H(\omega)P_x(\omega)\right|^2}{P_x(\omega)P_y(\omega)} = \frac{P_z(\omega)}{P_y(\omega)}}$$

Pertanto, essendo: $P_y(\omega) = P_z(\omega) + P_E(\omega)$

$$C_{xy}^2(\omega)P_y(\omega) = P_z(\omega)$$

$$\left[1 - C_{xy}^2(\omega)\right]P_y(\omega) = P_E(\omega)$$

La funzione coerenza consente di separare il processo Y nelle due componenti, per ogni valore di pulsazione!!



Esempio: funzione coerenza tra EEG misurati da elettrodi diversi, nel range 0-20 hz

I due segnali hanno un valore di coerenza prossimo ad uno ad una frequenza attorno ai 9 hz (ritmo alfa) quindi se adottiamo la relazione lineare+rumore, possiamo concludere che i due processi sono legati i modo pressochè lineare in banda alfa (prevale la componente z) ma non nelle altre bande, dove prevale la componente scorrelata

Funzione coerenza: come si stima?

Per calcolare gli spettri P_x, P_y, P_{xy}:

- Metodi tradizionali
  (basati sulla Trasformata di Fourier)

- Metodi parametrici
  (basati sull'impiego di modelli multivariati)

Metodi FT

La stima della funzione coerenza con i metodi FT richiede alcune cautele. Infatti, se ad es. sotto ipotesi di ergodicità, si stimano gli spettri P_x, P_y, con il metodo del periodogramma e si estendono principi analoghi (definizione di uno stimatore della cross correlazione, e poi sua FT) per il calcolo di P_{xy} si ottiene un risultato non informativo

$$\hat{R}_{xy}(k) = \frac{1}{N}\sum_{n=0}^{N-1-k} x(n)y(n+k) = \frac{1}{N} x(k) \otimes y(-k)$$

$$\hat{P}_{xy}(\omega) = \frac{1}{N} X(\omega) Y(-\omega)$$

$$\hat{P}_{x}(\omega) = \frac{1}{N} |X(\omega)|^2$$

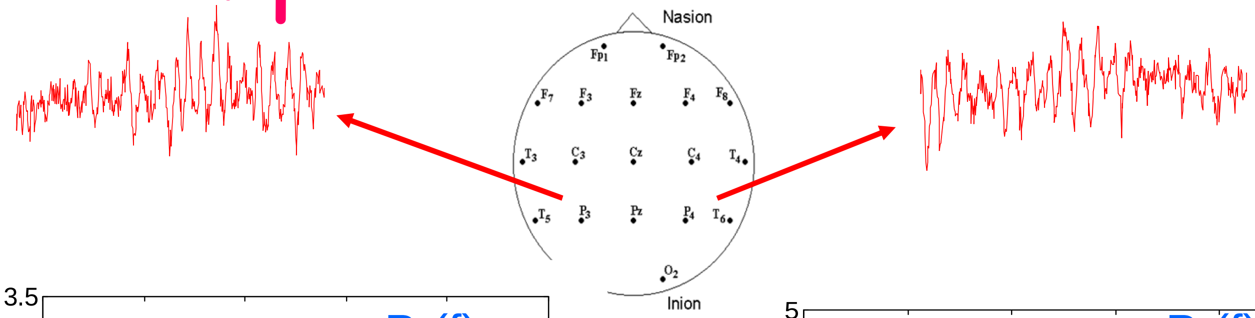
$$\hat{P}_{y}(\omega) = \frac{1}{N} |Y(\omega)|^2$$

$$\Downarrow$$

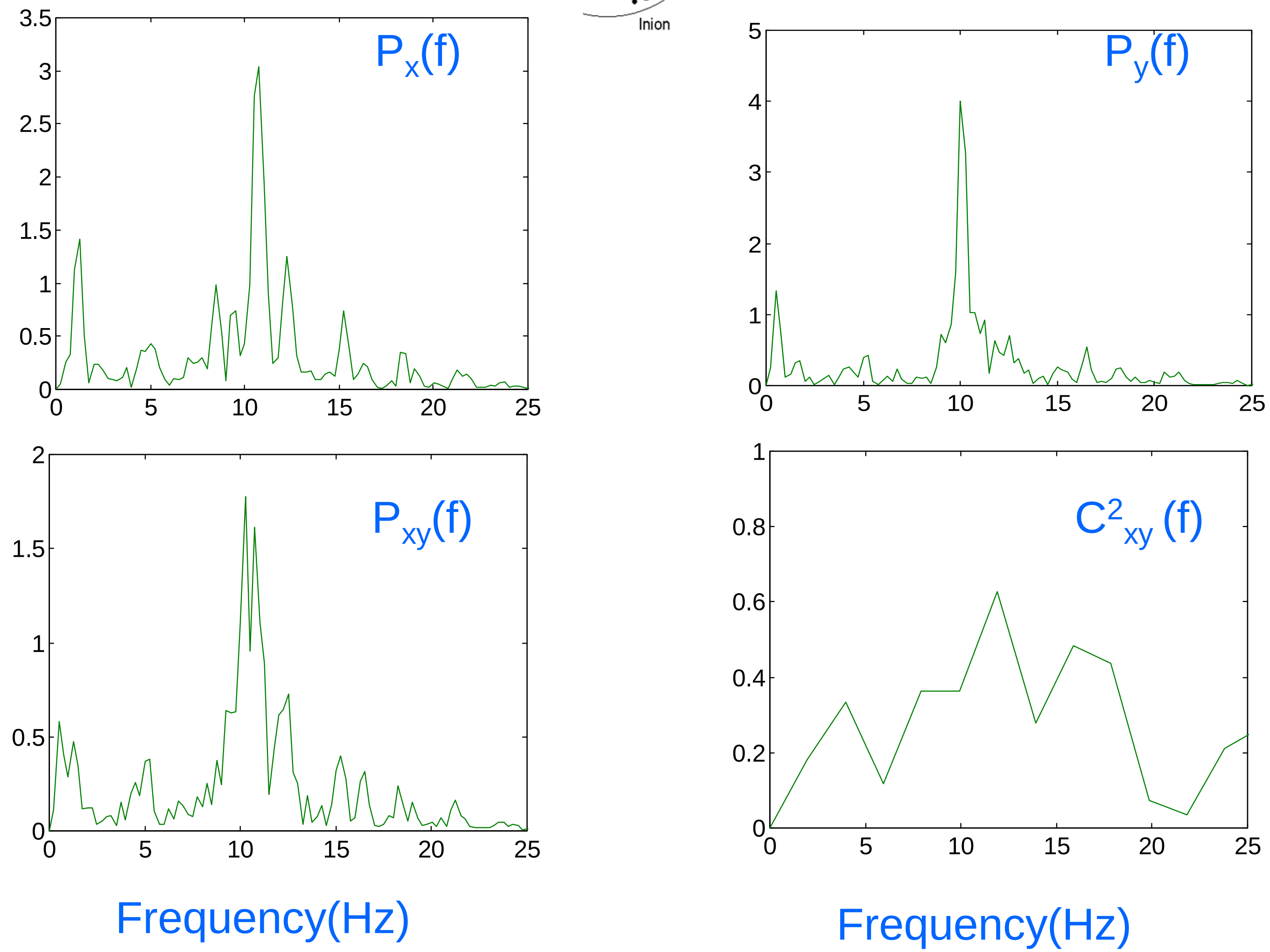
$$\boxed{C_{xy}^2(\omega) = 1}$$

Quindi per avere dei risultati informativi è indispensabile applicare ai metodi di stima degli spettri la tecnica di averaging per migliorare le proprietà degli stimatori!!!

Esempio



60 sec EEG divisi in 30 epoche di 2 sec
Fc=256 Hz, N=512



Gli spettri sono stati calcolando per ogni epoca e poi sono stati mediati (stazionarietà??)

La DFT fornisce 512 campioni equispaziati in un intervallo pari alla frequanza di campionamento, quindi la risoluzione spettrale è di 0.5 Hz

Rimangono comunque i problemi di leakage…

Metodi parametrici

I due processi X e Y vengono descritti come le uscite di un modello AR multivariato, pilotato da due processi bianchi, a media nulla. Se ora indichiamo i due processi con y_1 e y_2, il modello stabilisce un legame tra y_1 all'istante n e non solo y_1 in p istanti precedenti e il rumore u_1 all'istante attuale, ma anche y_2 nei p istanti precedenti. Analogamente per y_2

$$y_1(n) = -\sum_{k=1}^{P} a_k^{(1,1)} y_1(n-k) - \sum_{k=1}^{P} a_k^{(1,2)} y_2(n-k) + u_1(n)$$

$$y_2(n) = -\sum_{k=1}^{P} a_k^{(2,1)} y_1(n-k) - \sum_{k=1}^{P} a_k^{(2,2)} y_2(n-k) + u_2(n)$$

Modello AR multivariato

$$y_1(n) = -\sum_{k=1}^{P} a_k^{(1,1)} y_1(n-k) - \sum_{k=1}^{P} a_k^{(1,2)} y_2(n-k) + u_1(n)$$

$$y_2(n) = -\sum_{k=1}^{P} a_k^{(2,1)} y_1(n-k) - \sum_{k=1}^{P} a_k^{(2,2)} y_2(n-k) + u_2(n)$$

In forma matriciale :

$$\underline{y}(n) = \sum_{k=1}^{p} A_k \cdot \underline{y}(n-k) + \underline{u}(n)$$

$\underline{y}$(n) è il vettore [2 x 1] di processi aleatori

$\underline{A}_k$ matrice [2 x 2] dei coefficienti, es

$$A_1 = \begin{vmatrix} a_1^{(1,1)} & a_1^{(1,2)} \\ a_1^{(2,1)} & a_1^{(2,2)} \end{vmatrix}$$

$\underline{u}$(n) è il vettore [2 x 1] di rumore bianco a media nulla e matrice di varianza:

$$\Sigma_u = \begin{vmatrix} \boldsymbol{\sigma}_1^{\,2} & 0 \\ 0 & \boldsymbol{\sigma}_2^{\,2} \end{vmatrix}$$

Per identificare i parametri si possono scrivere equazioni analoghe a quelle di Yule Walker, in funzione di $\underline{R}$(k)=E[$\underline{y}$(n)$\underline{y}$(n+k)] matrice di autocorrelazione [2 x 2], che ha sulla diagonale principale le correlazioni dei due processi e fuori della diagonale le cross-correlazioni

Matrice di densità spettrale

Una volta identificato il modello, è possibile stimare la matrice [2x2] delle densità spettrali

$$\underline{P}_y(\omega) = \begin{vmatrix} P_1(\omega) & P_{12}(\omega) \\ P_{21}(\omega) & P_2(\omega) \end{vmatrix}$$

usando la seguente relazione che lega la matrice della densità spettrale del vettore di processi in uscita Py alla matrice di trasferimento H (valutata per $z=e^{jw}$) e alla matrice della densità spettrale del vettore di rumore bianco in ingresso Σ_u

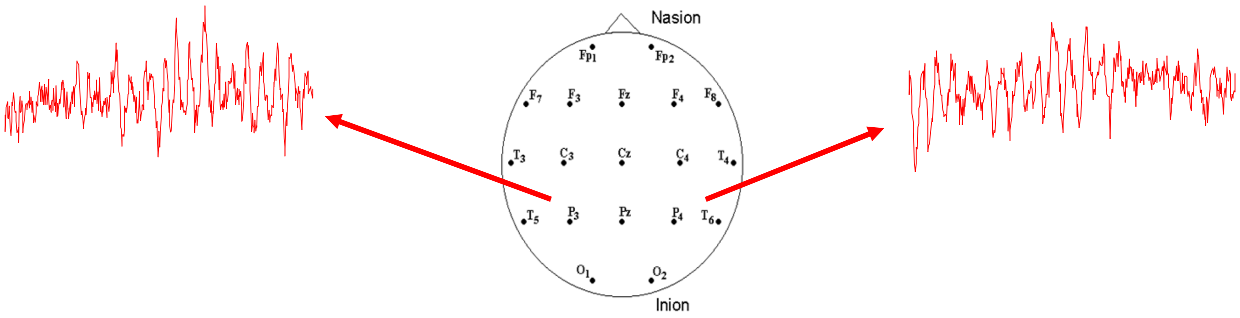
$$P_y(\omega) = \underline{H}(\omega) \cdot \Sigma_u \cdot \underline{H}^T(\omega) \quad \text{con} \quad \underline{H}(\omega) = \left(I - \sum_{k=1}^{p} A_k z^{-k} \right)^{-1}_{z=e^{iw}}$$

La matrice P_y è Hermitiana, gli elementi sulla diagonale sono reali è rappresentano gli auto spettri, mentre quelli fuori diagonale sono complessi e rappresentano i cross-spettri.

Coerenza

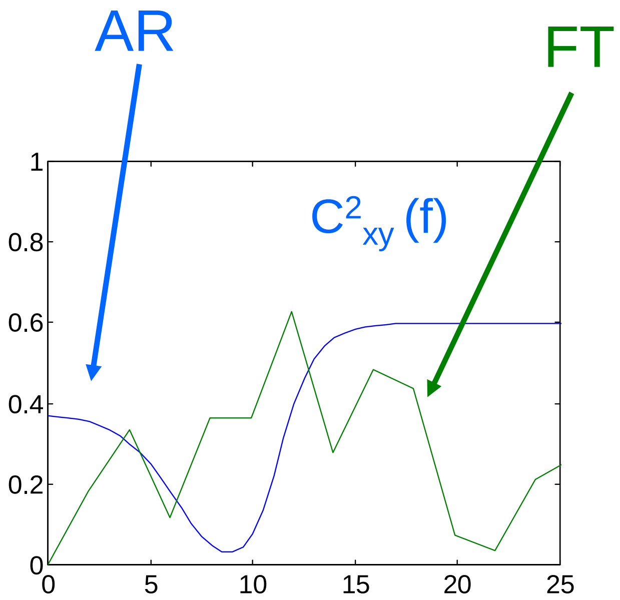
$$C_{12}^2(\omega) = \frac{\left|P_{12}(\omega)\right|^2}{P_1(\omega) \cdot P_2(\omega)} = C_{21}^2(\omega)$$

Esempio

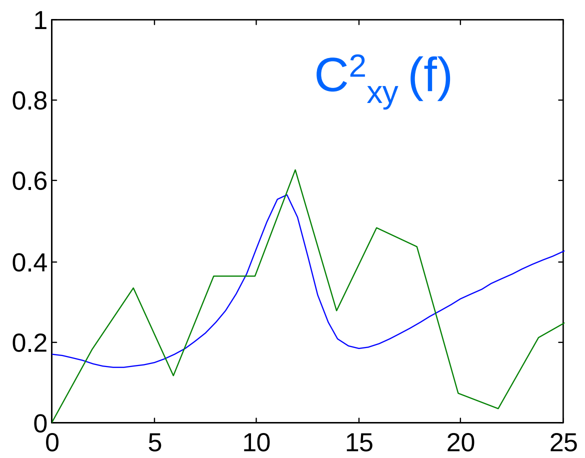


2 sec EEG
Fc=256 Hz, N=512

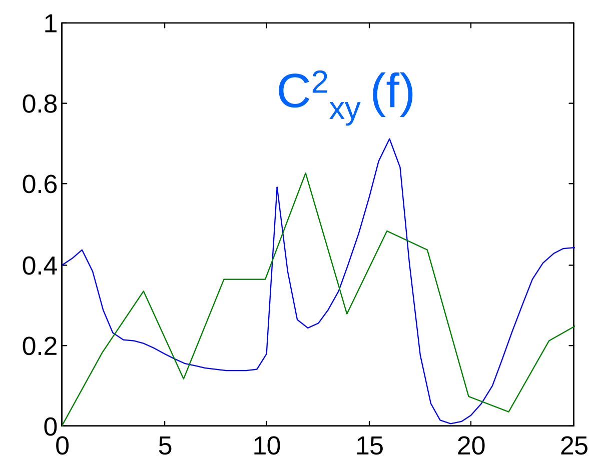
Scelta dell'ordine



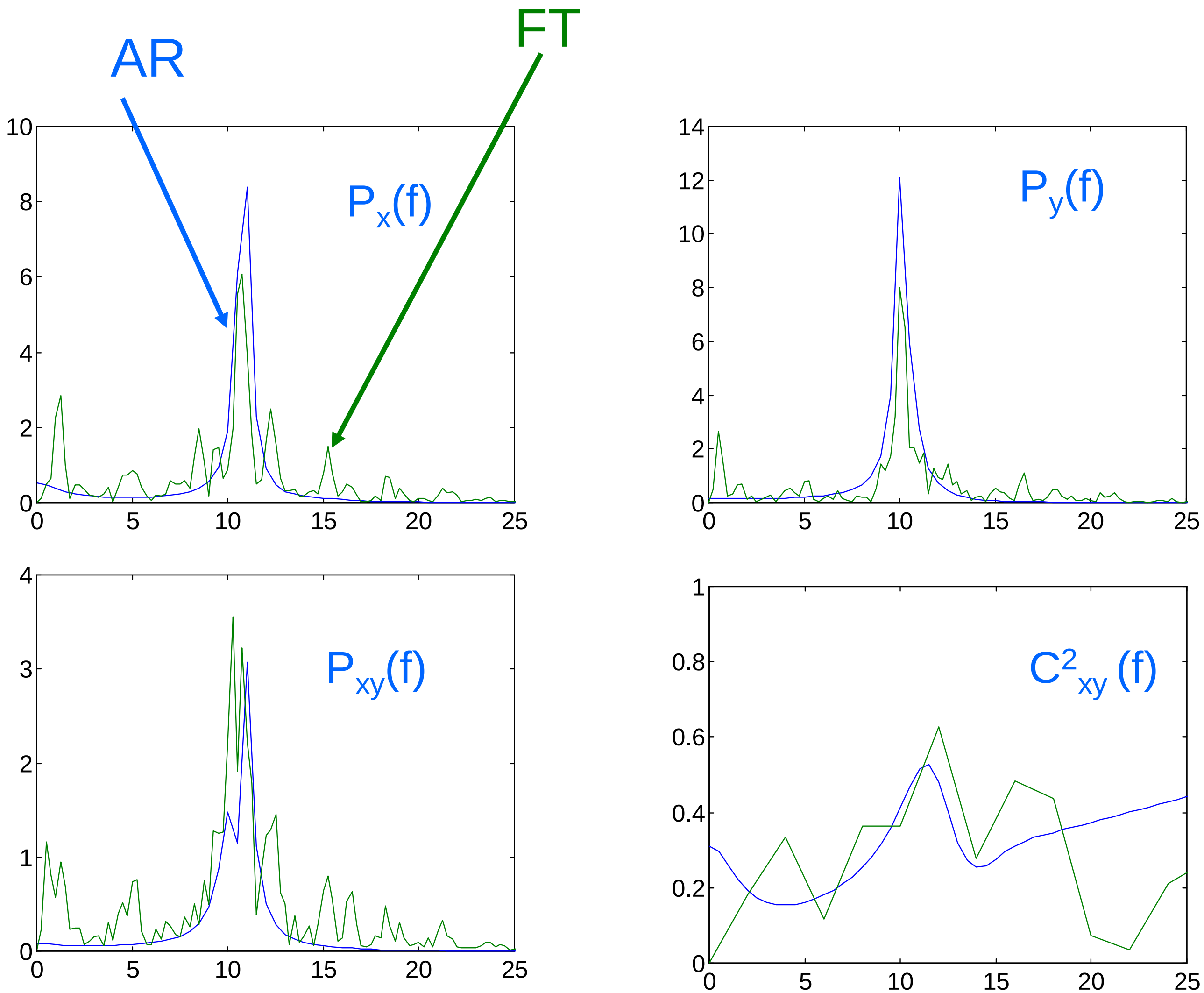
Low Order
p=3



Good Order
p=6



High Order
p=20



Parametri

$$A_1 = \begin{bmatrix} 1.60 & -0.40 \\ 0.09 & 1.30 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} -0.55 & 0.71 \\ -0.14 & 0.39 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} -0.01 & -0.20 \\ 0.36 & -0.67 \end{bmatrix}$$

$$A_4 = \begin{bmatrix} -0.29 & -0.02 \\ -0.25 & -0.04 \end{bmatrix}$$

$$A_5 = \begin{bmatrix} 0.28 & -0.29 \\ 0.30 & -0.37 \end{bmatrix}$$

$$A_6 = \begin{bmatrix} -0.15 & 0.18 \\ -0.22 & 0.34 \end{bmatrix}$$

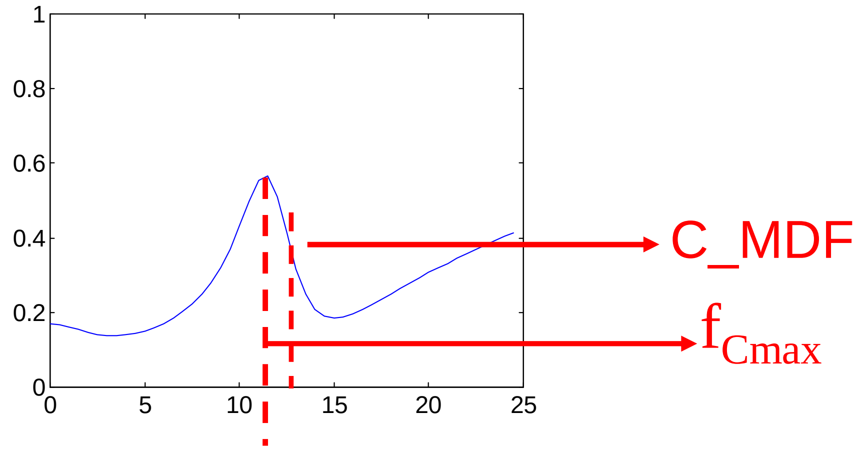
Σ=diag[1.30,1.20]

Indici di coerenza

In modo del tutto analogo a quanto visto per lo spettro del segnale EEG, anche a partire dalla funzione coerenza si possono definire degli indici quantitativi, ad es.

- Frequenza media (C_MDF)

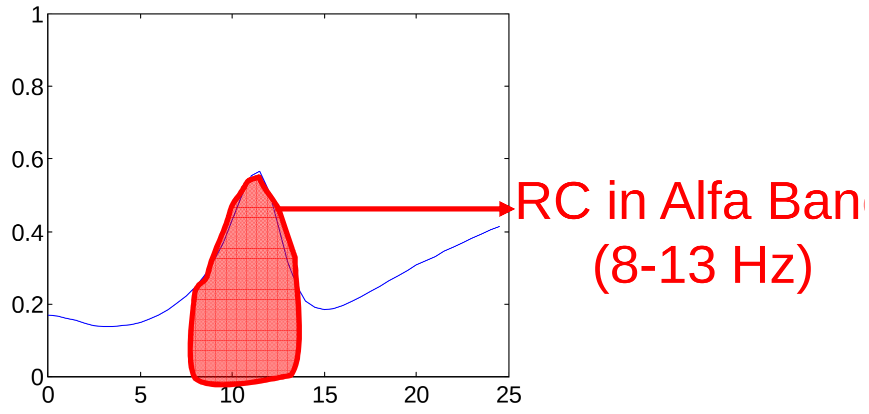
$$C\_MDF = \frac{\int_{0.5}^{f_{max}} f \cdot C_{xy}^2(f)df}{\int_{0.5}^{f_{max}} C_{xy}^2(f)df}$$



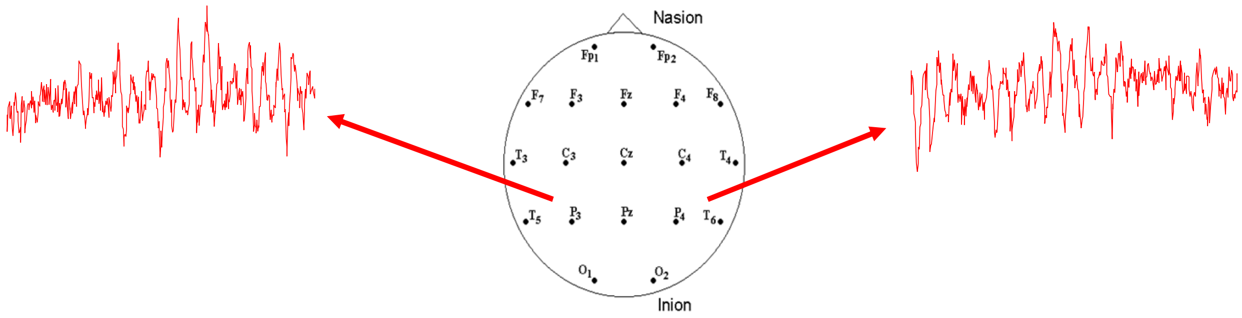
- Frequenza del massimo F_{max}

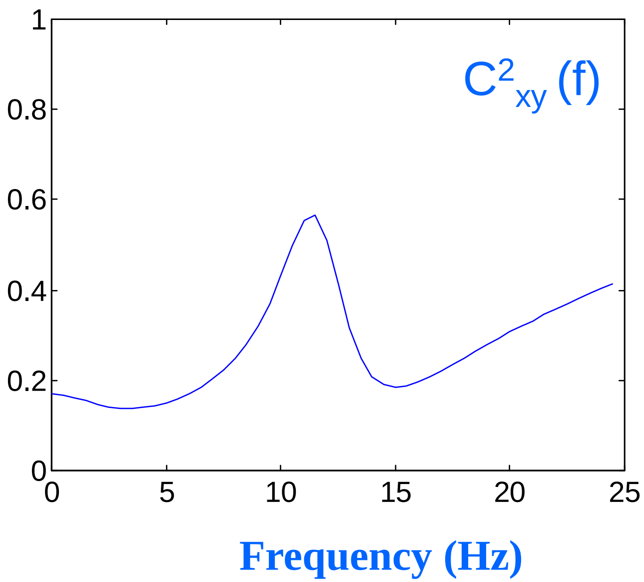
- Coerenza relativa nelle varie bande

$$RC = 100 \cdot \frac{\int_{low}^{high} C_{xy}^2(f)df}{\int_{0.5}^{f_{max}} C_{xy}^2(f)df}$$



Esempio





Coherence Indices

C_MDF	fCmax	Deta RC	Theta RC	Alfa RC	Beta RC
14.46	11.50	8.33	9.86	19.36	57.16

high=4 Hz, low =0.5 Hz for δ band
high=8 Hz, low =4 Hz for θ band
high=13 Hz, low =8 Hz for α band
high=25 Hz, low =13 Hz for β band

Riassumendo:

La funzione coerenza, compresa tra zero e uno, misura il grado di legame (lineare) tra due segnali, in funzione della frequenza.

E' una misura di sincronismo tra due segnali, cioè permette di individuare le frequenze in cui i due processi risultano avere un certo grado di associazione, ma non permette di introdurre alcuna misura di causalità, cioè di capire quale dei due segnali pilota l'altro. Questo rappresenta un limite della coerenza.

Per ovviare a ciò, esistono :

- definizioni alternative della funzione coerenza, ad esempio la coerenza diretta (DTF)

$$\mathrm{DTF}_{1\to 2}(\omega) = \frac{\sigma_1 H_{21}(\omega)}{\sqrt{\sigma_2^2 |H_{22}(\omega)|^2 + \sigma_1^2 |H_{21}(\omega)|^2}}$$

$$\mathrm{DTF}_{2\to 1}(\omega) = \frac{\sigma_2 H_{12}(\omega)}{\sqrt{\sigma_1^2 |H_{11}(\omega)|^2 + \sigma_2^2 |H_{12}(\omega)|^2}}$$

A differenza dalla coerenza ordinaria, la DTF fornisce indicazioni sulla direzione dell'interazione tra due segnali

- Indici di causalità di Granger

Causalità secondo Granger

La definizione è stata formulata nel 1969 da Granger (premio Nobel per l'economia nel 2003),

Dati due segnali x e y, si può dire che x influenza y se:

- l'informazione di x è utile a predire y (1° criterio)
- includendo i valori passati di x nella stima regressiva di y, si migliora la qualità della stima, ovvero si riduce l'errore di predizione (2° criterio)

Causalità di Granger -1° criterio

Il modello AR multivariato viene utilizzato non solo come passo intermedio per il calcolo dello spettro incrociato e quindi della funzione coerenza, ma anche per capire se, in presenza di una coerenza non trascurabile, è il primo segnale che pilota il secondo, oppure il secondo che pilota il primo, oppure ci sono controlli incrociati tra I due segnali. A tal fine vengono osservati I valori degli elementi fuori della diagonale principale delle matrici $\underline{A_k}$

$$A_k = \begin{vmatrix} a_k^{(1,1)} & a_k^{(1,2)} \\ a_k^{(2,1)} & a_k^{(2,2)} \end{vmatrix}$$

Si possono avere 4 situazioni:

y_1 causa y_2 ma non viceversa se almeno un elemento in posizione (2,1) è significativamente diverso da zero mentre tutti gli elementi (1,2) non lo sono

y_2 causa y_1 ma non viceversa se almeno un elemento in posizione (1,2) è significativamente diverso da zero mentre tutti gli elementi (2,1) non lo sono

y_1 causa y_2 e y_2 causa y_1 se almeno un elemento in posizione (1,2) e almeno un elemento in posizione (2,1) sono significativamente diversi da zero

y_1 e y_2 sono scorrelati tra loro se tutti i cofficienti in posizione (1,2) e (2,1) non sono significativamente diversi da zero

Causalità di Granger - 2° criterio

L'indice di causalità viene definito come

$$GC_{x \to y} = \ln \frac{\operatorname{var}(y)}{\operatorname{var}(y|x)}$$

dove

- ln è il logaritmo naturale
- $var(y)$ è la varianza del rumore di ingresso per il modello AR monovariato, che misura la possibilità di predire y sulla base della sua storia passata
- $var(y|x)$ è la varianza del rumore di ingresso per il modello multivariato, che misura la possibilità di predire a partire dalla storia passata non solo di y ma anche di x.

a) GC è sempre definito non negativo

b) GC = 0 quando nessun legame causale esiste tra i segnali : $var(y) = var(y|x)$

c) GC significativamente diverso da zero indica la presenza di una relazione di casualità da x a y

d) Il valore di GC è una misura della forza del legame

e) Una differenza significativa tra gli indici di causalità $x \to y$ e $y \to x$ indica il verso della relazione

Come valutare la significatività?

Test statistico (F-test)

Ipotesi nulla H_0 : *var(y) = var(y|x)*

Statistica F

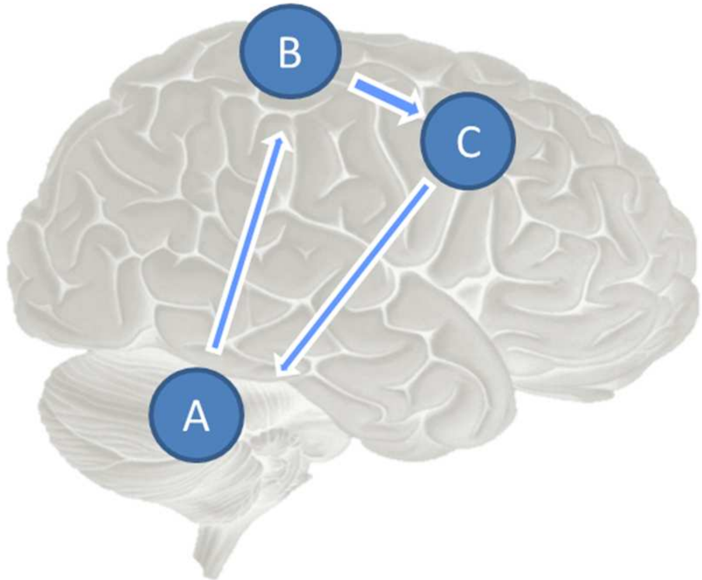
$$F = \frac{\dfrac{\mathrm{var}(y) - \mathrm{var}(y \mid x)}{p}}{\dfrac{\mathrm{var}(y \mid x)}{(L - 2p - 1)}}$$

P = dimensione del modello MVAR
L= numero di campioni

Sotto HO F è distribuita secondo la statistica F di Fisher-Snedecor con df pari a p e L-2p-1

Fissato un intervallo di significatività (es. Alfa=0,05) si valuta se accettare o rifiutare l'ipotesi nulla

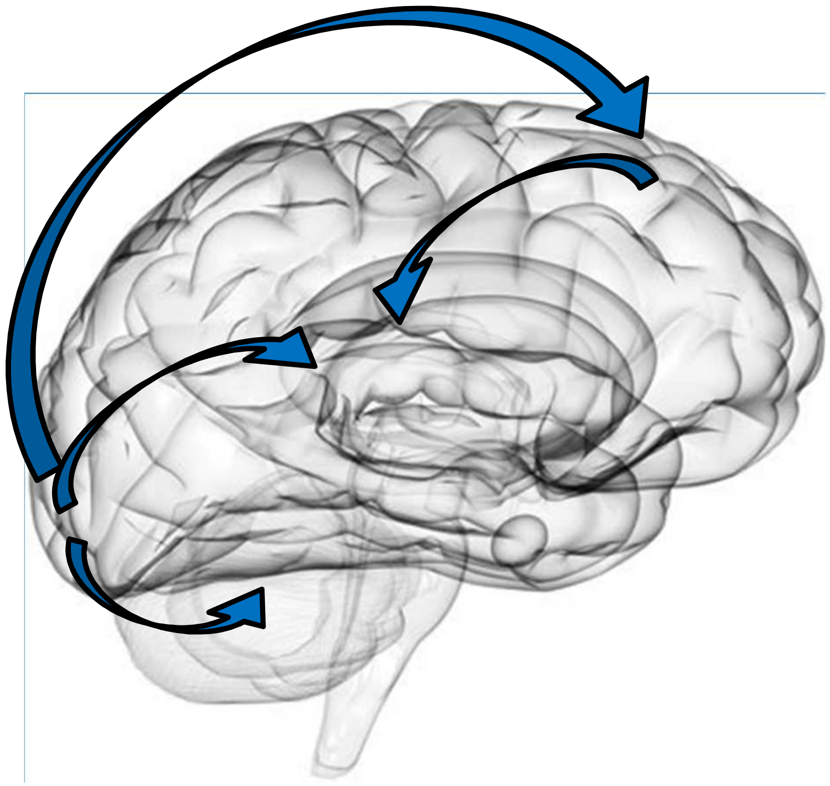
Applicazione: Analisi di coerenza dell'EEG per lo studio della connettività cerebrale



Why connectivity?

«One of the most important goals of neuroscience is to establish precise structure-function relationships in the brain.»

Stephan KE (2004). On the role of general system theory for functional neuroimaging. J.Anat. 205, 443-470.

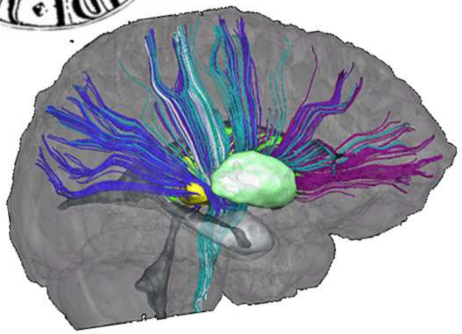


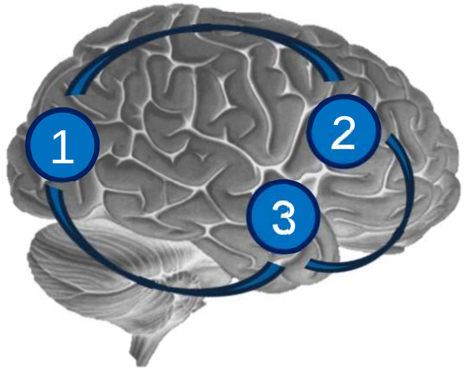
Connectivity variability ⟹ **PLASTICITY**

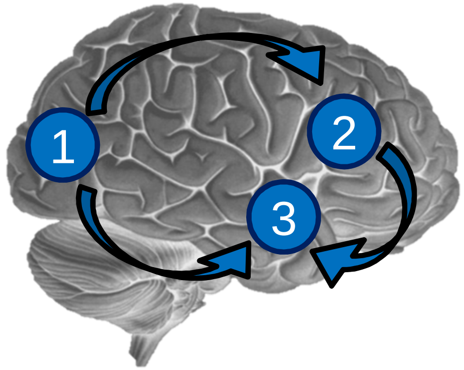
- Development
- Learning
- Perception
- Adaptive response to injury

Connectivity

STRUCTURAL, FUNCTIONAL & EFFECTIVE







structural connectivity
presence of axonal connections

by using MRI and DTI
(fiber tracking)

functional connectivity
statistical dependencies between
regional time series

by using
fMRI and
EEG/MEG

effective connectivity
causal (directed) influences between
neurons or neuronal populations

How to evaluate connectivity?

TOPOLOGY

STRENGTH

CAUSALITY
RECTED/INDIRECTED

Two approaches

Model based

- A priori knowledge on
  - topology
  - causality
- Inference of
  - strength

SEM (Structural Equation Modeling)
implemented in commercial (Lisrel) and free (R) packages

Data driven

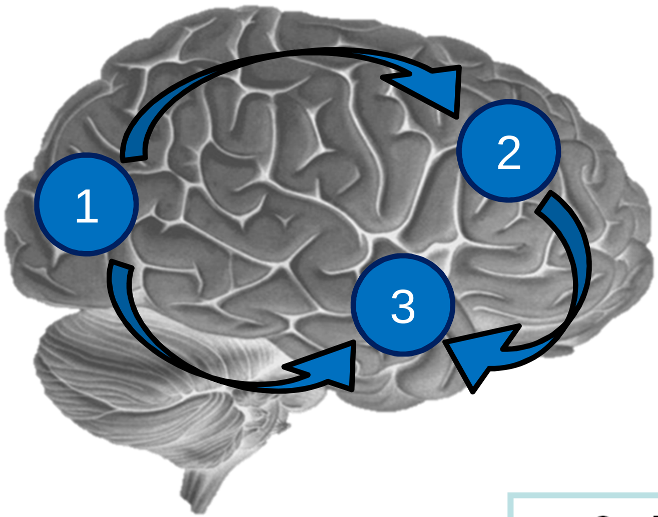
- Inference of
  - topology
  - causality
  - strength

MVAR (Multi-Variate Autoregressive Models)
implemented in MATLAB tools

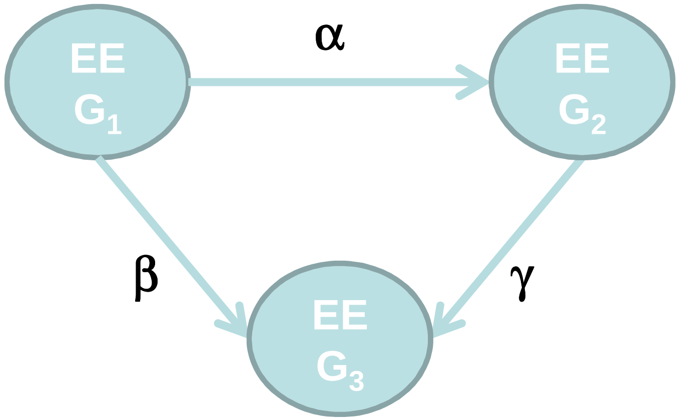
SEM

Assumptions

1. Known network topology



Example



2. Linear structural equation

Example:

$$EEG_2(n) = \alpha EEG_1(n) + e_2(n)$$
$$EEG_3(n) = \beta EEG_1(n) + \gamma EEG_2(n) + e_3(n)$$

Estimate

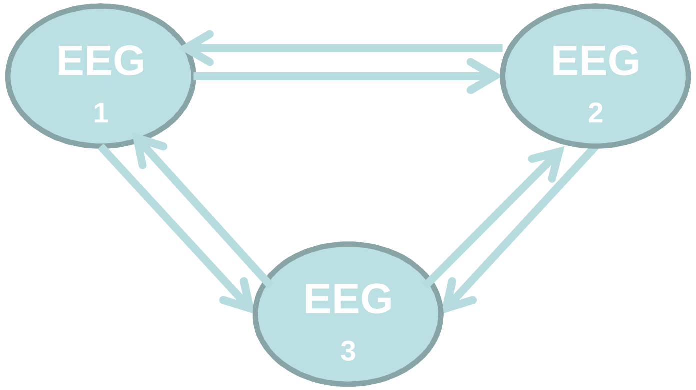
α β γ are estimated by maximum likelihood, based on the comparison between:

- sample covariance (estimated from the data)
- model covariance (expressed in terms of **α β γ**)

MVAR

Assumptions

MVAR Model



Example:

$$EEG_1(n) = \sum_{k=1}^{P} a_k^{(1,1)} EEG_1(n-k) + \sum_{k=1}^{P} a_k^{(1,2)} EEG_2(n-k) + \sum_{k=1}^{P} a_k^{(1,3)} EEG_3(n-k) + e_1(n)$$

Estimate

- a_{ij} coefficients are estimated by using weighted least squares methods
- model order **P** is selected based on validation criteria

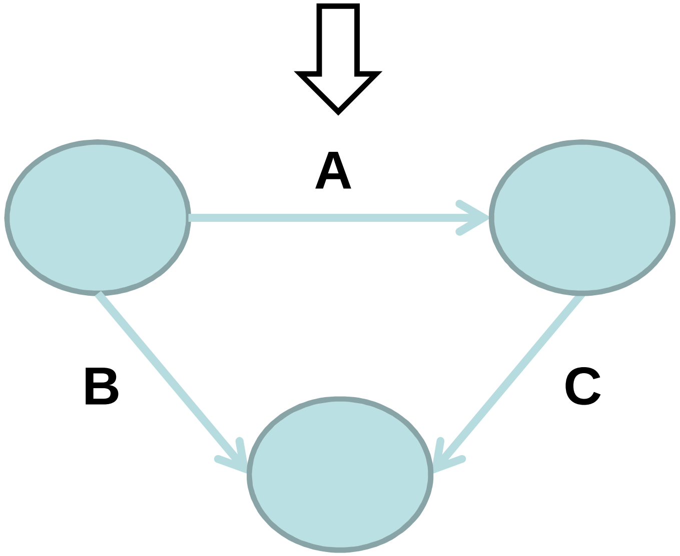
From MVAR

- Granger indices
- Coherence and other frequency indices (DTF)

How well do SEM and MVAR reveal brain functional connections?

Simulation study

1. EEG signals are generated based on known connectivities.

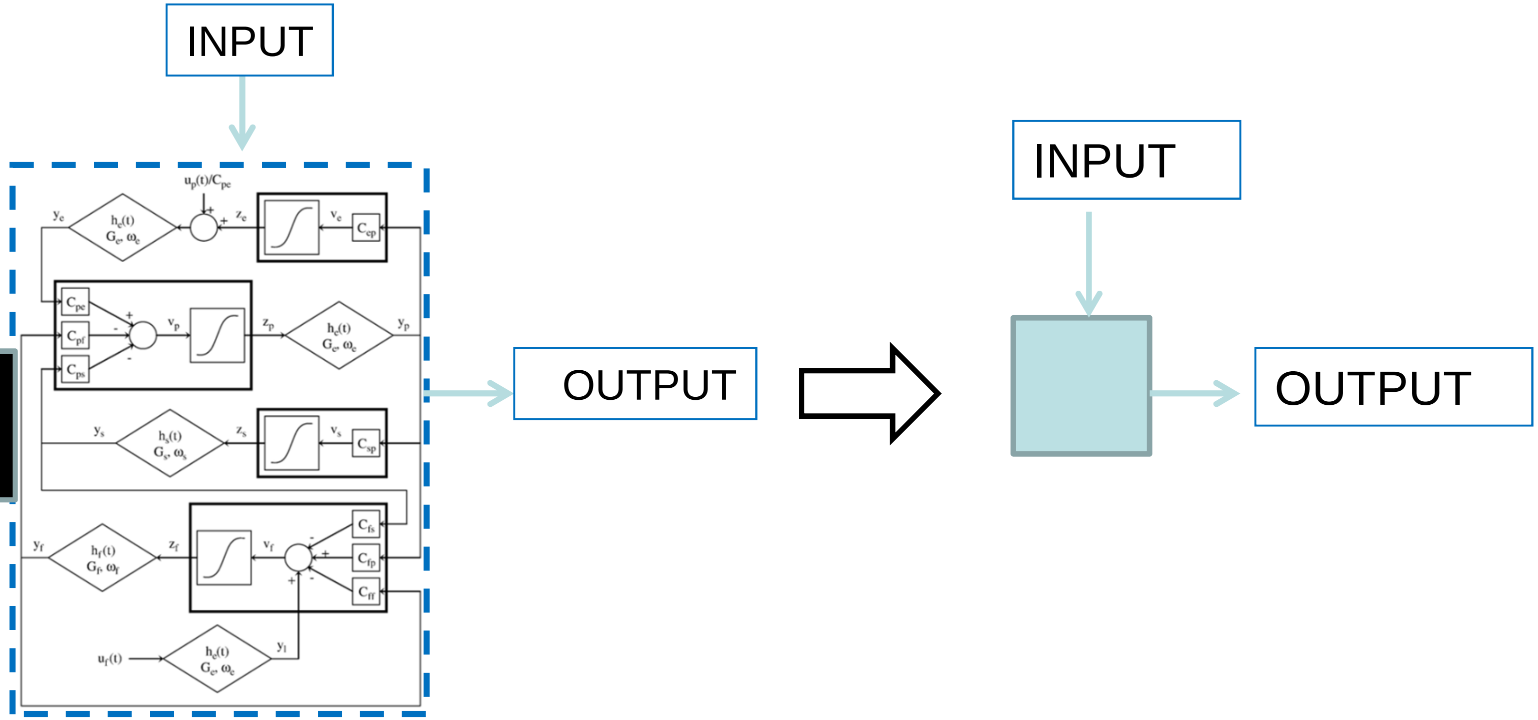


2. On these data, SEM parameters and MVAR indexes were estimated.

3. Comparison between the true and the estimated networks.

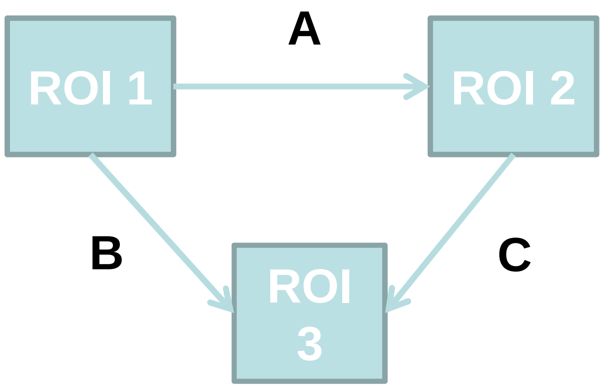
1. Neural mass model (Ursino et al, 2006)

ONE REGION



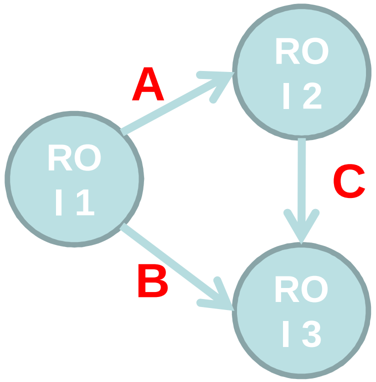
NETWORK

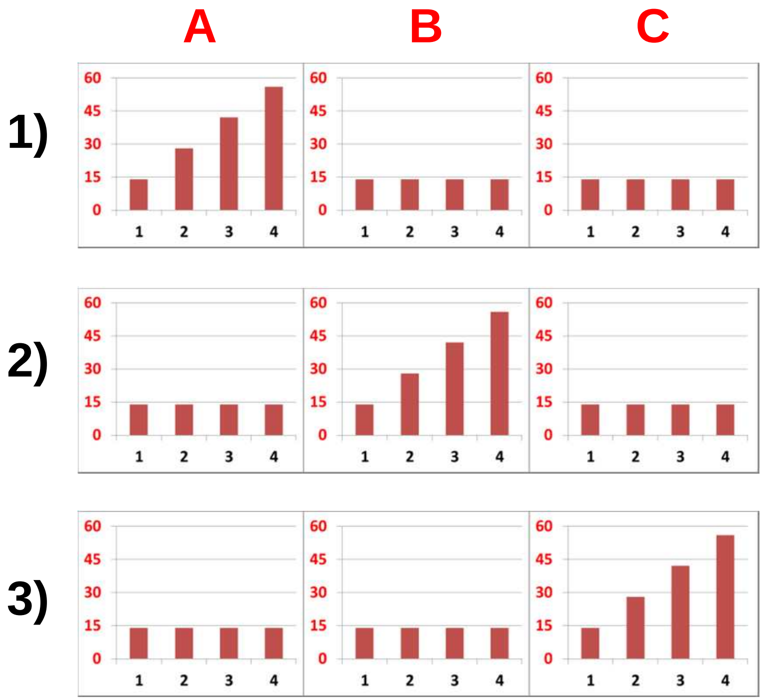
Known connecting weigths A, B, C



e.g. for ROI 2:
White noise is the endogenous input
EEG in ROI1 is the exogenous input
EEG in ROI2 is the output

2. Implementation

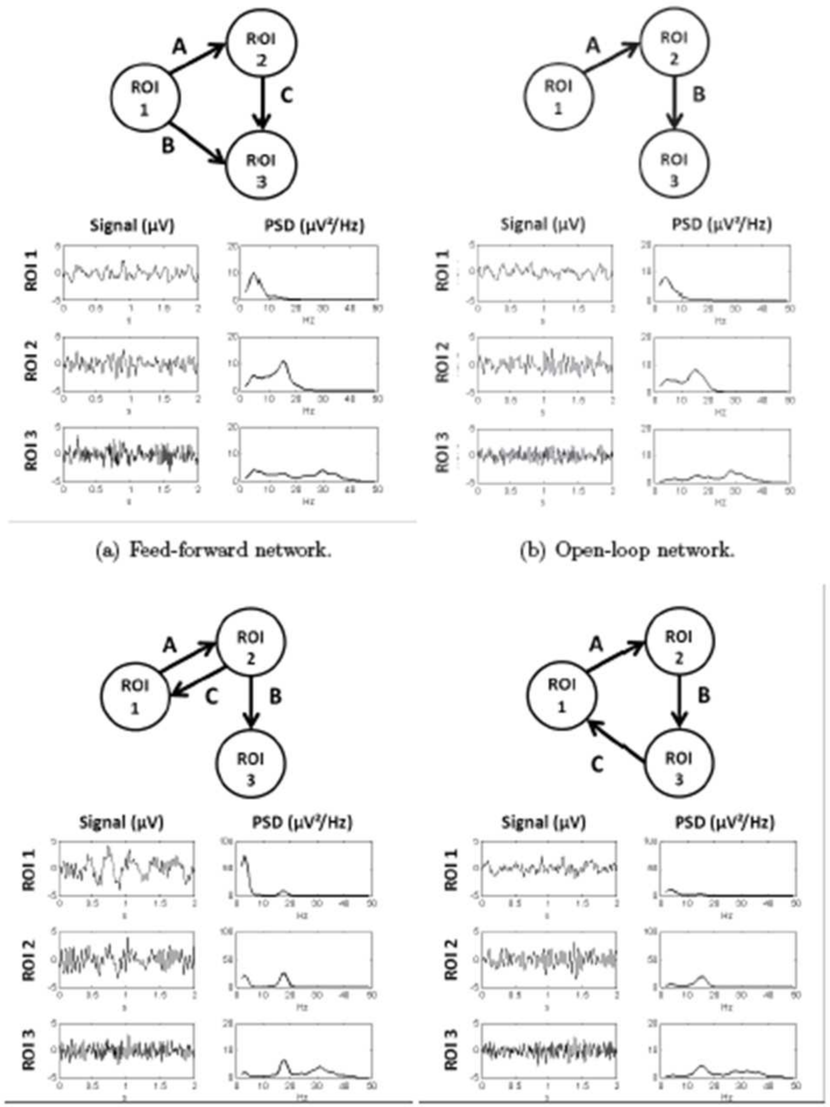




Different networks are simulated, characterized by: different topologies different values of A, B, C

For each model: 100 realizations are generated

3. Examples of EEG data



(a) Feed-forward network.

(b) Open-loop network.

(c) Network with feed-back link.

(d) Cycle network.

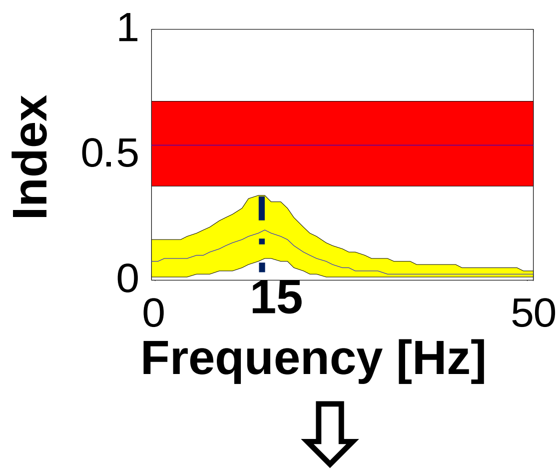
5. MVAR – GC and Frequency indexes

On each data set (triplet of EEGs from the three ROIs) a MVAR model is estimated

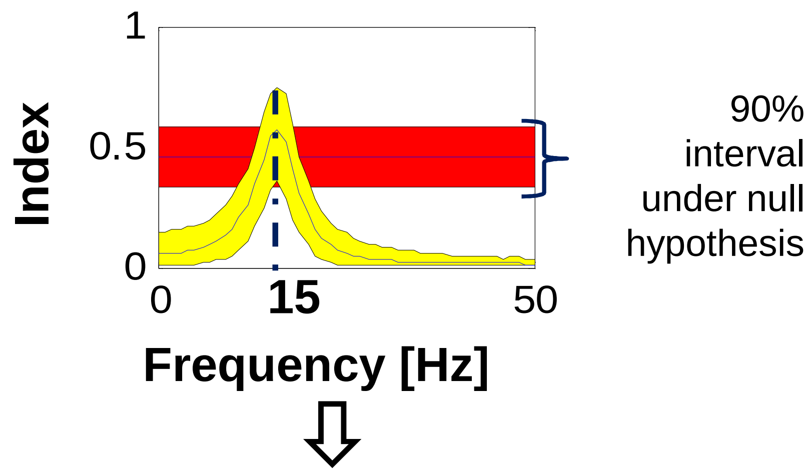
From MVAR:

GC is evaluated. It provides information on: 1. topology (significant GC means that a connection is present) 2. strength (CG value) 3. direction

Coherence is evaluated, and by elaborating on it, DTF and PDC are defined to measure the **LINEAR** coupling A→ B at various frequencies.



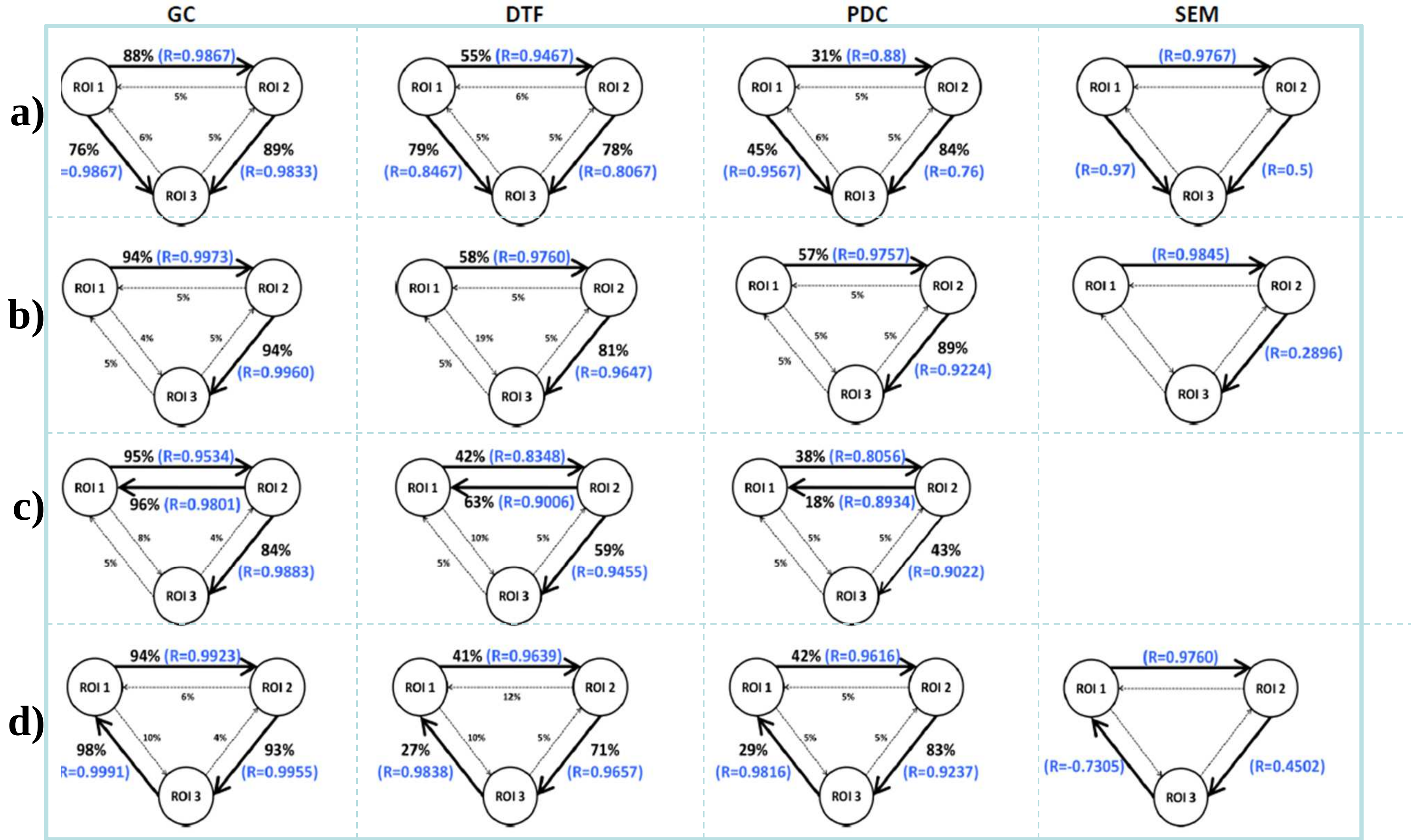
No significative coupling



Significative coupling at ≈15 Hz

4. Results

Inferred topologies by Granger causality index (GC), frequency indices (DTF and PDC) and Structural Equation Modeling (SEM). Continuous lines indicate true connections and the associated numbers quantify the percentage of true positives, averaged across datasets. Dashed lines indicate absent connections, and the associated numbers quantify the percentage of false positives averaged across dataset. The blue numbers associated with continuous lines indicate the correlation coefficient between true (i.e. assumed in the model) and estimated values.



Conclusion

	TOPOLOGY	STRENGTH	IRECTED/INDIRECTED CAUSALITY
SEM	**known**	✓	**known**
Granger Causality	✓	✓	✓
Frequency indices	✓	✓	✗

Granger causality and frequency indexes are not competitive but complementary techniques → **Use together!**