



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede amministrativa: Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

SCUOLA DI DOTTORATO DI RICERCA IN: Ingegneria dell'Informazione

INDIRIZZO: Bioingegneria

CICLO XXV

**PARAMETRIC CONNECTIVITY ANALYSIS IN
TIME AND FREQUENCY DOMAIN FROM IN
SILICO AND EEG DATA**

Direttore della Scuola: Ch.mo Prof. Matteo Bertocco

Coordinatore: Prof. Giovanni Sparacino

Supervisore: Ch.ma Prof.ssa Gianna Maria Toffolo

Contents

Summary	iii
Sommario	ix
1. Multivariate methods for connectivity analysis	1
1.1. MVAR	2
1.1.1. MVAR model identification	2
1.1.2. Coupling with Granger causality	3
1.1.3. Coupling with frequency indexes	6
1.2. SEM	10
2. Assessment on in silico data: the neural mass model for data simulation	13
2.1. The neural mass model	14
2.2. NMM parameters analysis	19
2.3. Simulated dataset	22
3. Data Analysis	27
3.1. Simulation	27
3.2. Connectivity estimation	27
3.3. Statistical significance	28
3.3.1. F-test for Granger causality	28
3.3.2. Null hypothesis test for DTF and PDC	28
3.4. Assessment of estimated indexes	29

Contents

4. Assessment on in silico data: results	31
4.1. Simulation	31
4.1.1. NMM parameters	31
4.1.2. Model predicted EEG signals	32
4.2. Connectivity estimation	37
4.2.1. Feed-forward network	37
4.2.2. Other network models	53
4.3. Summary	57
5. Application to EEG data: assessment of hepatic encephalopathy	61
5.1. Study design	63
5.1.1. Dataset	63
5.1.2. Spectral analysis	63
5.1.3. EEG connectivity analysis	64
5.1.4. Statistical analysis	64
5.2. Results	65
5.2.1. Spectral analysis	65
5.2.2. EEG connectivity estimation	65
6. Discussion	73
6.1. Simulation study	73
6.2. Network connectivity estimation	75
6.3. Conclusions	77
A. SEM analysis in frequency domain	81
A.1. Path coefficients interpretation	81
A.1.1. Parameters interpretation using the Neural Mass Model	83
Bibliography	87

Summary

In this decade, establishing structure-function relationships in human brain has become one of the most influential concepts in modern cognitive neuroscience since interactions among cerebral components are fundamental to explain cortical activities ([1]; [2]; [3]).

In literature such relationships have been defined in terms of structural, functional and effective connectivity. This distinction, mainly focused on the theoretic concept, is also related to the different measurement instruments and analytical tools used for acquiring and processing the data. The structural connectivity refers to a pattern of anatomical links among brain regions. Its analysis aims to characterize the architecture of complex networks underlying the cerebral functional organization. Magnetic Resonance Imaging and especially Diffusion Tensor Imaging can be used to convey information concerning the physical connection between neuronal populations. Functional/effective connectivity aims at identifying the presence and the strength of connections in terms of statistically significant dependency. The former is defined as the temporal correlation between neurophysiological events occurring in distributed neuronal groups and areas. The latter describes the causal influence that one neural system exerts over another either directly or indirectly in terms of temporal precedence and physical control ([4];[5]). Functional and effective connectivity can be estimated exploiting both Functional Magnetic Resonance Imaging (fMRI) and electrophysiological signals, such as Electroencephalography (EEG) and Magnetoencephalography (MEG), with different advantages and drawbacks, respectively. fMRI provides high spa-

Summary

tial resolution (mm) but poor temporal precision (s) while EEG/MEG has more limited spatial resolution (cm) and higher temporal precision (ms). Because functional and effective connectivity are largely estimated over time, EEG and MEG are more suitable for calculating such connectivity.

In literature several methods have been developed to characterize brain connectivity in terms of network topology, connections strength and causality, following two main approaches: the data-driven, where topology, causality and strength are all inferred from data, and the neural model-based, where the model topology is postulated from a priori knowledge and only the connections strength is estimated from the data.

Data driven approach. The data driven approach includes linear, non-linear and information-based techniques.

- The linear ones provide a battery of indices derived by multivariate autoregressive models (MVAR) based on Granger causality principles ([6]) or MVAR frequency response ([7]). Such are Ordinary Coherence, Partial Coherence, Directed Transfer Function (DTF) and Partial Directed Coherence (PDC). These indexes measure the strength of the linear coupling between two signals; in addition DTF and PDC provide information about causal influence ([8]).
- Among the non-linear techniques, phase synchronization has been shown to be very efficient in detecting interactions between oscillators. The phase locking values approach assumes that two dynamic systems may have their phases synchronized even if their amplitude are zero correlates ([9]).
- The most representative information-based technique is the cross mutual information that measures the mutual dependence between two signals by quantifying the amount of information gained about one signal from measuring the other, as a function of delay between these two signals ([10]).

Neural model based approach. Representative methods are the Structural Equations Modelling (SEM) and the Dynamic Causal Modelling (DCM) ([11]; [12]). They are multivariate technique used to test hypothesis regarding the influences among interacting variables, but different concepts underlies these two methods. SEM approach assumes that neuronal dynamics are very fast in relation to signals fluctuations and, hence, is based on a static neuronal model. This case, the neuronal activity has reached steady-state and changes in connectivity are led directly by changes in the covariance structure of the observed time series ([13]). On the other hand, in DCM the observed time series are modelled as a deterministic dynamical system in which external inputs causes changes in neural activity and therefore in connectivity values ([14]).

Most approaches, like those based on Granger causality principles, have been examined in literature to quantify their ability in revealing cerebral connections ([15];[16]; [11]) but their simulation studies do not provide a comprehensive analysis because they use *in silico* data generated by self-referential linear methods which do not reproduce the complexity of brain.

To overcome this issue, an innovative simulation approach has been developed in this work, based on a nonlinear neural mass model ([17]) totally independent of SEM and MVAR linear equation and able to address the complexity of neural networks. This no-self referential approach was exploited to generate *in silico* network data to be used as a benchmark, to quantitatively compare obtained results with true connections.

The main objective of this work was to understand limits and advantages of MVAR indexes and SEM by exploiting the simulation study. Thus, it mainly serves as a proof-of-concept for connectivity measures under ideal conditions. Our purpose was to derive from simulation results some practical procedures in order

Summary

to classify different brain states to support both cognitive research and clinical activity. First, research activity was focused to address connectivity on simulated data obtained on three regions networks characterized by different strength connections and based on different levels of non linearity. Second, a dataset, made available by Department of Medicine, University of Padova was used to explore application of these methods to real data by applying the simulation study suggestions.

This thesis consists of three main section.

The first one includes Chapter 1-2-3 describing in detailed the considered connectivity measures, such are those based on Multivariate Autoregressive models and the Structural Equation Modelling, and the simulation study. The second part depicts *in silico* results and the application to EEG data. Finally, comments are reported in Discussion and Conclusions.

Chapter 1 explains how the connecting parameters of MVAR and SEM models are identified on EEG data and describes procedures commonly exploited to analyse connectivity. **Chapter 2** reports an overview about the principal models used to generate *in silico* data, namely the neural mass models, and described the neural mass model exploited in this work. Finally, it characterizes network models adopted to simulate data and lists the procedure followed to generate *in silico* datasets. **Chapter 3** summarizes the computations implemented to have more insights on our data by analysing the output of each methods. It describes the procedure used to evaluate the statistical significance of each index results, such are the F-test for Granger causality index and the null distribution threshold using surrogate data for MVAR frequency indexes. **Chapter 4** illustrates the results obtained with the simulation study. First, we reported the complete analysis for a representative subset of experiments, then for all datasets we showed topology and strength estimates. **Chapter 6** delineates the procedure followed to study the connectivity in case of hepatic encephalopathy. **Chapter 7** covers

the Discussion and Conclusions. The Appendix is a parallel work aimed to understand the meaning of connectivity indexes computed via Structural Equation Modelling. By exploiting the neural mass model used to simulate cortical data, the objective is to quantify which measure its estimates represent.

We demonstrated that Granger causality is a good estimator with high values both of sensitivity and specificity, while frequency indexes, DTF and PDC, are too much affected by the threshold choice and their interpretation in terms of absolute strength connection is not clear.

As regard SEM, we proved the difficulty of its approach to describe just simple situations. Even if SEM is based on linear regression as well as MVAR models, it differently assumes there is no connection with past information, as if brain connectivity could describe time series relationships by the instant we observe it. Hence, it is not sufficiently robust to characterize neuronal dynamic activity.

Sommario

Negli ultimi decenni le varie tecniche e metodiche sviluppate per lo studio dell'attività cerebrale hanno dimostrato che le diverse regioni neuronali del cervello non operano in isolamento ma interagiscono tra loro formando una complessa rete di connessioni. Lo studio di queste relazioni/connessioni esistenti tra le diverse regioni corticali, tramite l'elaborazione sia di segnali elettrofisiologici, come l'EEG, sia di immagini, come l'fMRI, è generalmente denominato come studio della connettività. La definizione di connettività può essere classificata in tre principali categorie: anatomica, funzionale ed effettiva. La connettività anatomica è strettamente associata alla presenza di connessioni assoniche tra i vari neuroni; la connettività funzionale è definita come la correlazione temporale tra eventi neurofisiologici appartenenti a diverse regioni neuronali; la connettività effettiva è definita come l'influenza che una regione neuronale esercita attraverso una relazione causa-effetto su un'altra regione.

In letteratura sono presenti due principali approcci per lo studio della connettività: l'uno di tipo esplorativo, basato esclusivamente sui dati da cui estrarre informazioni sia sulla topologia sia sulla forza; l'altro che prevede la conoscenza a priori di un modello di rete per ottenere informazioni circa l'intensità degli accoppiamenti.

L'obiettivo di questa tesi si è focalizzato sulla validazione e implementazione di alcuni dei metodi più utilizzati: quelli basati sui modelli autoregressivi multivariati (MVAR), come la Directed Transfer Function (DTF), la Partial Directed Coherence (PDC), e sui principi della causalità di Granger e il metodo detto

Sommario

Structural Equation Modeling (SEM). Questi metodi sono ampiamente esaminati in letteratura per quantificare la loro capacità di rilevare le connessioni cerebrali, ma gli studi di simulazione proposti sono basati su modelli di generazione dei dati *in silico* che semplificano molto la reale complessità del cervello [15] e che si basano sui modelli autoregressivi stessi. Per superare questo problema è stata sviluppata una simulazione con un approccio innovativo basato sull'utilizzo di un Neural Mass Model[17]. L'obiettivo consiste nel generare dati simulati completamente indipendenti dalle equazioni lineari dei metodi che poi si vanno a testare e, al contempo, in grado di simulare la complessità delle reti neurali. Brevemente, la simulazione consiste delle seguenti fasi:

- diversi set di dati *in silico* sono simulati utilizzando il modello neurale di massa con diversi modelli di topologia, livelli di non linearità e intensità di connessioni;
- per ogni set dei suddetti parametri, 100 realizzazioni di segnali di 2 secondi vengono generati;
- le reti stimate a partire dai parametri di connettività calcolati con i metodi considerati vengono confrontate con le reti vere.

Per analizzare le prestazioni dell'indice di causalità di Granger e degli indici in frequenza Directed Transfer Function (DTF) e Partial Directed Coherence (PDC) sono state effettuate simulazioni Monte Carlo in modo da ottenere una statistica delle performance. Si è osservato che l'indice di Granger è il più affidabile con elevata percentuali di sensibilità e bassa frequenza di falsi positivi e negativi. Per analizzare la stima delle forze, sono stati confrontati i valori dei pesi imposti con i risultati degli indici dei metodi MVAR e le stime ottenute dal SEM mediante regressione lineare. Si è osservato che il SEM è il metodo meno affidabile, mentre i risultati ottenuti con gli indici MVAR presentano una buona correlazione lineare con i pesi veri. Anche in questo caso l'indice di Granger dà i migliori risultati

correlando sempre con $R \sim 0.99$.

I risultati hanno rivelato che l'indice di causalità di Granger è un accurato stimatore della topologia di rete in quanto si è dimostrato in accordo con le reti vere nella maggior parte degli esperimenti simulati, mentre DTF e PDC, oltre a presentare alcune imprecisioni, risultano più difficili da interpretare in termini di forze assolute. Questi risultati suggeriscono di utilizzare l'indice di causalità di Granger come strumento esplorativo per definire sia la topologia della rete sia l'intensità delle forze. Poi, le informazioni in frequenza provenienti dai diversi metodi (DTF, PDC) devono essere integrate per migliorare l'affidabilità dei risultati sulle intensità delle connessioni. L'obiettivo principale di questo studio di simulazione è quello di fornire una procedura robusta da usare per l'analisi della connettività del cervello umano, in grado di classificare i diversi stati del cervello in supporto sia della ricerca in ambito cognitivo e sia dell'attività clinica. L'analisi effettuata sui segnali EEG riportata è un esempio di applicazione a dati reali, in cui si esamina l'effetto dell'iperammonemia indotta da un carico amminoacidico su pazienti cirrotici e soggetti sani sulla riorganizzazione funzionale del segnale EEG (Dott. Amodio, Dipartimento di Medicina, Università degli Studi di Padova).

Questa tesi si sviluppa in sei capitoli di seguito brevemente riassunti. Nel **Capitolo 1** si definiscono sia i modelli multivariati autoregressivi e gli indici derivati per stimare la connettività in termini di causalità di Granger e nel dominio della frequenza, sia il metodo SEM. Nel **Capitolo 2** si presenta il modello utilizzato per la generazione dei dati simulati analizzati in questa tesi e si descrivono le caratteristiche principali delle reti di simulazione considerate. Nel **Capitolo 3** vengono descritti sia i metodi impiegati per la valutazione della significatività statistica dei vari stimatori sia la procedura per valutare l'accuratezza delle stime con il confronto sulle reti vere. Nel **Capitolo 4** si presentano i principali risultati dello studio della connettività corticale ottenuti mostrando dapprima l'intera analisi

Sommario

su un sottoinsieme di simulazioni, poi sintetizzando i risultati su tutti i dataset. Nel **Capitolo 5** viene presentata una possibile applicazione dei metodi prima esposti su un problema di tipo clinico, riguardante l'analisi di EEG su pazienti affetti da Encefalopatia epatica. Infine, nel **Capitolo 6** si discutono i risultati presentati nel capitolo 5 evidenziando limiti e vantaggi dei vari metodi e il loro range di applicabilità in modo da visualizzare globalmente le loro prestazioni. L'**Appendice** riporta un lavoro parallelo eseguito per studiare il significato dei coefficienti di connettività stimati con il metodo SEM utilizzando le equazione del neural mass model descritto in precedenza.

1. Multivariate methods for connectivity analysis

Joyful, joyful, Lord, we adore Thee, God of Glory, Lord of love! Hearts unfold like flowers before Thee, hail Thee as the sun above!

In brain connectivity analysis two main approaches can be distinguished: the neural model based, where the model topology is postulated from a priori knowledge and connection strength only is estimated from the data, and the data driven, where topology, causality and strength are all inferred from the data. The most prevalent data driven methods are those based on Granger causality principles, and those based on neural model are the Dynamic Causal Modeling (DCM) and the Structural Equation Modeling (SEM). Thanks to their simplicity, Granger causality and SEM have been largely applied but, on the other hand, some criticisms have been arisen in literature concerning their assumptions ([13]).

Both methods are based on multivariate linear regression models while they differ for the discount of temporal information. Granger causality is computed by using Multivariate Autoregressive Models, where correlations among measurements at different time lags are used to quantify coupling. SEM models instantaneous interactions among variables and ignores the influence previous states have on current responses.

This chapter explains how the connecting parameters of MVAR and SEM models are identified on EEG data and describes procedures commonly exploited to

analyse connectivity. Representative examples will be used to quantify the implications of various assumptions.

1.1. MVAR

1.1.1. MVAR model identification

The MVAR model with N variables is expressed as:

$$\mathbf{y}(n) = - \sum_{i=1}^p \mathbf{A}(i)\mathbf{y}(n-i) + \mathbf{e}(n) \quad (1.1)$$

where $\mathbf{y}(n)$ is the data vector of dimension N related to n -sample, $\mathbf{e}(n)$, called *prediction error*, is the multivariate uncorrelated white noise process with diagonal covariance matrix Σ_e , $\mathbf{A}(i)$ are the $N \times N$ matrices of model coefficients and p is the model order.

The model coefficients and the covariance matrix are identified on time series data by applying the correlation approach known as the multichannel Yule-Walker method which minimizes the mean square *prediction error* to find the optimum MVAR parameters set.

Under the orthogonality requirement, stated as

$$E\{\mathbf{e}(n)\mathbf{y}(n-k)^T\} = 0, \quad (1.2)$$

where $()^T$ denotes matrix or vector transpose, considering the autocorrelation definition

$$\mathbf{R}_y(k) = E\{\mathbf{y}(n)\mathbf{y}(n-k)^T\}, \quad (1.3)$$

we can obtain the following set of equations

$$\mathbf{R}_y(k) = - \sum_{i=1}^p \mathbf{A}(i)\mathbf{R}_y(k-i) + \Sigma_e \delta(k), \quad (1.4)$$

where $\delta(k)$ is Kronecker delta function ($\delta(k) = 1$ when $k = 0$, otherwise $\delta(k) = 0$). Therefore, let \mathbf{s} be a set of N time series:

$$\mathbf{s}(n) = [\mathbf{s}_1(n), \mathbf{s}_2(n), \dots, \mathbf{s}_N(n)], \quad n = 1, \dots, L. \quad (1.5)$$

using the sample correlations $\hat{R}_s(k)$ and the set of equations 1.4 with $(p+1)$ matrix equations and $(p+1)$ unknown matrix parameters, the MVAR coefficient matrices $(\mathbf{A}(i), \sum_e)$ can be solved by using the Levinson-Durbin recursion extended to the multivariate Yule-Walker. Finally, model order is chosen so as the residuals are uncorrelated white noise. In practice, several MVAR models are calculated while varying p , and the best order is the one minimizing a parsimony criterion. We considered the Akaike Information Criterion (AIC) given by

$$AIC(p) = L \log(\det \sum_e) + 2pN^2 \quad (1.6)$$

where L is the number of data points applied in the estimation and N is the number of time series. For reliable parameter identification, the number of parameters must be significantly smaller than the number of data points available, i.e. $pN^2 \ll NL$.

1.1.2. Coupling with Granger causality

The Wiener-Granger causality principle can be implemented using linear MVAR models. The origin of this concept in time series analysis arose in statistical field, when Wiener (1956) ([18]) recognized the role of temporal ordering in the inference of cause-effect relationship between two simultaneously measured time series. Coupling is defined in terms of ability of one time series to better predict a second time series by incorporating knowledge of the first one. Later, Granger

(1969) ([6]) formalized this notion for linear regression models of stochastic processes specifying that a decreasing in prediction error variance of the second time series implies a driver-response relationship between them. Different implementations of this concept were applied to study the feedback relation between input and output variables and to multivariate autoregressive processes introducing the concept of conditional causality ([19],[20]). In the 1990s neurobiological applications rapidly spread, growing interest in studying the effect that one part of the nervous system has on another, either in the absence of identifiable behavioural events or in the context of task performances ([21]).

Theoretical considerations

The original bivariate Granger definition is generalized to interactions among sets of interdependent variables taking into account the data variance. Hence, \mathbf{y}_i causes \mathbf{y}_j , with $i \neq j$, if the prediction error of \mathbf{y}_j estimated with a MVAR model including all the N time series of \mathbf{y} is lower than the one estimated with a MVAR model including all the N time series of \mathbf{y} but \mathbf{y}_i . Defining as $\tilde{\mathbf{y}}$ the $N-1$ time series set where \mathbf{y}_i has been excluded, the Granger causality from \mathbf{y}_i to \mathbf{y}_j respect to all the other inputs is measured as:

$$GC_{y_i \rightarrow y_j | y} = \ln \frac{\text{var}(y_j | \tilde{y})}{\text{var}(y_j | y)} \tag{1.7}$$

Statistical significance can be determined via F-statistic:

$$F_{y_i \rightarrow y_j | y} = \frac{\frac{RSS_r - RSS_c}{p}}{\frac{RSS_c}{(L - 2p - 1)}} \tag{1.8}$$

where RSS_r and RSS_c are the Residuals Sum of Squared of restricted and complete models, respectively, and p and $(L - 2p - 1)$ are the degrees of freedom of numerator and denominator. A significant F-statistic is interpreted as evidence

that the complete model provides a better prediction compared to what does the restricted one ([21]).

As an example, let us consider three jointly distributed, stationary multivariate stochastic processes \mathbf{y}_1 , \mathbf{y}_2 and \mathbf{y}_3 . To measure the causality from \mathbf{y}_2 to \mathbf{y}_1 given \mathbf{y}_3 , the complete MVAR model is the following:

$$\begin{aligned}
 y_1(n) &= - \sum_{k=1}^p a_{1,1}(k) y_1(n-k) - \sum_{k=1}^p a_{1,2}(k) y_2(n-k) - \sum_{k=1}^p a_{1,3}(k) y_3(n-k) + e_1(n) \\
 y_2(n) &= - \sum_{k=1}^p a_{2,1}(k) y_1(n-k) - \sum_{k=1}^p a_{2,2}(k) y_2(n-k) - \sum_{k=1}^p a_{2,3}(k) y_3(n-k) + e_2(n) \\
 y_3(n) &= - \sum_{k=1}^p a_{3,1}(k) y_1(n-k) - \sum_{k=1}^p a_{3,2}(k) y_2(n-k) - \sum_{k=1}^p a_{3,3}(k) y_3(n-k) + e_3(n)
 \end{aligned} \tag{1.9}$$

with covariance matrix $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12}^2 & \sigma_{13}^2 \\ \sigma_{21}^2 & \sigma_2^2 & \sigma_{23}^2 \\ \sigma_{31}^2 & \sigma_{32}^2 & \sigma_3^2 \end{bmatrix}$

while the restricted MVAR model is described by the following equations

$$\begin{aligned}
 y_1(n) &= - \sum_{k=1}^p \tilde{a}_{1,1}(k) y_1(n-k) - \sum_{k=1}^p \tilde{a}_{1,3}(k) y_3(n-k) + \tilde{e}_1(n) \\
 y_3(n) &= - \sum_{k=1}^p \tilde{a}_{3,1}(k) y_1(n-k) - \sum_{k=1}^p \tilde{a}_{3,3}(k) y_3(n-k) + \tilde{e}_2(n)
 \end{aligned} \tag{1.10}$$

with covariance matrix $\tilde{\Sigma} = \begin{bmatrix} \rho_1^2 & \rho_{13}^2 \\ \rho_{31}^2 & \rho_3^2 \end{bmatrix}$.

The Granger causality $\mathbf{y}_2 \rightarrow \mathbf{y}_1$, defined in eq. 1.7, is expressed from the elements of Σ and $\tilde{\Sigma}$:

$$GC_{y_2 \rightarrow y_1 | y_3} = \ln \frac{\rho_1^2}{\sigma_1^2}$$

This index is positive when the prediction error of \mathbf{y}_1 estimated in the complete model is lower than the one estimated in the restricted model, whilst is close to zero when \mathbf{y}_2 does not improve the regression.

1.1.3. Coupling with frequency indexes

The spectral representation of a MVAR model gives useful tools for the analysis of stochastic processes, based on MVAR model (eq. 1.1) transformation into \mathbf{Z} domain:

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{E}(z) \quad (1.11)$$

where $\mathbf{H}(z)$ is the system transfer matrix:

$$\mathbf{H}(z) = (\mathbf{I} + \sum_{i=1}^p \mathbf{A}(i) z^{-i})^{-1} \quad (1.12)$$

and $\mathbf{E}(z)$ is the *prediction error* \mathbf{Z} -transform.

The $N \times N$ frequency response matrix can be expressed as:

$$\mathbf{H}(f) = \mathbf{H}(z)|_{z=e^{i2\pi fT}} \quad (1.13)$$

where T is the sampling period, and the cross-spectral matrix can be derived as follows:

$$\mathbf{S}(f) = \mathbf{H}(f) \Sigma_e \mathbf{H}(f)^H, \quad (1.14)$$

where $(*)^H$ stands for the Hermitian transpose and $\Sigma_e = \text{diag}(\sigma_i^2)$, *prediction error* covariance matrix.

The most traditional function proposed to detect cooperative neuronal activity in a couple of electro-physiological signals, y_i and y_j , is coherence:

$$\text{Coh}_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f) S_{jj}(f)}} \quad (1.15)$$

where $S_{ij}(f)$ and $S_{ii}(f)$, $S_{jj}(f)$ are the cross and the auto - spectra, respectively, varying in the range 0 – 1. High values of coherence between two EEG signals are interpreted as evidence for ongoing cooperation and long-range synchronization.

This is a consolidate index to describe the linear coupling, but it furnishes a symmetrical information and hence cannot distinguish dependencies' direction. To overcome this problem, several measures have been suggested [22], and the most applied are Directed Transfer Function ([23]) and Partial Directed Coherence ([24]).

Directed Transfer Function (DTF)

Let us consider the coherence between two signals y_i and y_j , eq.1.15, rewritten as follows:

$$Coh_{ij}(f) = \sum_{n=1}^N \frac{\sigma_n H_{in}(f)}{\sqrt{S_{ii}(f)}} \frac{\sigma_n H_{jn}^*(f)}{\sqrt{S_{jj}(f)}} \quad (1.16)$$

The first factor contains the generalized version of DTF index, which is defined as:

$$\begin{aligned} \gamma_{j \rightarrow i}(f) &= \frac{\sigma_j H_{ij}(f)}{\sum_{n=1}^N \sqrt{\sigma_n^2 |H_{in}|^2(f)}} \\ &= \frac{\sigma_j H_{ij}(f)}{\sqrt{\sigma_n^2 H_{i:}(f) H_{i:}^H(f)}} \end{aligned} \quad (1.17)$$

with $H_{i:}(f)$ being the i -th row of $H(f)$.

Unlike coherence, DTF is able to identify dependency direction, since it exclusively depends on the frequency response, which is a non symmetrical matrix, $H_{ij} \neq H_{ji}$. It expresses the influence of y_j on y_i as the ratio between the inflow from j to i to all the inflows to i . Since DTF is normalized, it varies in the interval $[0, 1]$, where 0 means no significant connections and positive values describe the presence of connection.

Partial Directed Coherence (PDC)

Unlike DTF, PDC relies on the inverse of the frequency response matrix, written as:

$$H(f)^{-1} = A(f) \quad (1.18)$$

where

$$A_{ij}(f) = \sum_{k=1}^p a_{ij}(k) e^{-j2\pi fk}$$

directly based on the MVAR model coefficients, and is defined as follows:

$$\begin{aligned} \pi_{j \rightarrow i}(f) &= \frac{A_{ij}(f)}{\sum_{n=1}^N \sqrt{\sigma_n^2 |A_{nj}|^2(f)}} \\ &= \frac{A_{ij}(f)}{\sqrt{A_{:j}^H(f) A_{:j}(f)}} \end{aligned} \quad (1.19)$$

with $A_{:j}(f)$ being the j -th column of $A(f)$.

It describes the influence of y_j on y_i as the ratio between the outflow from y_j to y_i to all the outflows from the source y_j . As for DTF, PDC is a normalized index, ranging from 0 to 1.

Comparison between DTF and PDC

Even if both DTF and PDC operate in frequency domain, they assume different meanings. DTF can be interpreted in terms of spectral density as the power spectrum of y_i coming from y_j normalized to all the contributions to y_i at frequency f . Unlike DTF, PDC has not a direct correspondence with the power spectrum. It depends on the inverse of the frequency response matrix (eq. 1.13) which does not reflect spectral information. Moreover, they differ in the ability of distinguishing direct and indirect dependencies: DTF shows not only direct, but also cascade flows, whereas PDC shows only direct flows. As before, let us consider three stochastic processes $\mathbf{y}_1, \mathbf{y}_2$ and \mathbf{y}_3 described by the MVAR model of eq.1.9 with frequency response rewritten as follows:

$$\mathbf{H}(f) = \frac{1}{\det(\mathbf{A}(f))} \begin{bmatrix} A_{22}A_{33} - A_{23}A_{32} & -(A_{12}A_{33} - A_{13}A_{32}) & A_{12}A_{23} - A_{13}A_{22} \\ -(A_{12}A_{33} - A_{13}A_{32}) & A_{11}A_{33} - A_{13}A_{31} & -(A_{11}A_{23} - A_{13}A_{21}) \\ A_{21}A_{32} - A_{22}A_{31} & -(A_{11}A_{32} - A_{12}A_{31}) & A_{11}A_{22} - A_{12}A_{21} \end{bmatrix} \quad (1.20)$$

Coupling between variables i and j described with DTF, eq. 1.17, results in a linear combination of the elements of the $\mathbf{A}(f)$ matrix 1.18, whilst PDC, eq. 1.19, considers the single $A_{ij}(f)$ element. In case of absence of direct connection between i and j , then $a_{ij}(k) = 0$ for each k and, hence, PDC equals zero. Differently DTF reveals some connections in any case due to alternative indirect paths linking those two variables.

Therefore, DTF can be treated as a global index which describes interaction between i and j throughout both direct and indirect connections, while PDC only reveals direct ones.

Statistical test for significance

To examine the statistical significance of DTF and PDC a null hypothesis test is performed for each pair of signals. Specifically, the null distribution of these measures is determined using phase randomization: each data series, transformed in frequency domain via FFT, is randomly shuffled in order to change phase information and then reported in time domain via iFFT. This procedure is iterated $M = 100$ times, and for each dataset DTF and PDC are computed. Maximum values of the one-hundred frequency functions are considered to estimate null distribution and threshold is fixed at 95 – *th* percentile. Indexes values beyond it indicate the existence of strong dependency between each pair of signals.

DTF and PDC indexes

DTF and PDC frequency function magnitude is usually evaluated at the peak frequency and the more these functions rise the more connection strength increases. In order to quantitatively sum up their frequency information, we derived some indexes by AUC integrals calculation:

$$AUC_{DTF} = \int_0^{f_{max}} \gamma_{j \rightarrow i}(f) df \quad , \quad AUC_{PDC} = \int_0^{f_{max}} \pi_{j \rightarrow i}(f) df \quad (1.21)$$

for each pairs of signals. In addition, to evaluate connection strength in classical EEG bands, AUC integrals are computed considering δ , θ , α , β and γ frequency intervals.

1.2. SEM

SEM grew out of geneticist, social-science and economics fields from 1920s onwards and has been used in functional imaging since the early 1990s. It was firstly applied to animal autoradiographic data and then extended to human PET data to identify task-dependent differential activation of the dorsal and ventral visual pathways ([12];[25]). Since then, other researchers have used SEM to analyse fMRI and EEG data ([26];[15]).

SEM is a static multivariate regression model used to estimate connections within a defined network. It is based on the hypothesis that the topology of the network in terms of interconnections among interacting variables is a priori known and that inter variables coupling is linear time invariant.

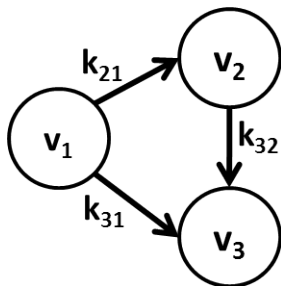


Figure 1.1.: An example of structural model with three regions and three connections. Each region is associated to a variable, v_i , $i = 1,2$ and 3 . Arrows indicate causal relationships that are assumed a-priori and strength connections are defined by the scalar k_{ij} .

Considering as an example the network shown in Figure 1.1, these assumptions

are translated in the following equations:

$$\begin{aligned} v_2(t) &= k_{21}v_1(t) + e_2(t) \\ v_3(t) &= k_{31}v_1(t) + k_{32}v_2(t) + e_3(t) \end{aligned} \tag{1.22}$$

where $v_n(t)$ is the model prediction for cortical activity associated with variable n , with $n = 1, 2, 3$, k_{ij} is the path coefficient from variable j to variable i and $e_n(t)$ is a residual term of covariance Σ_e , interpreted as driving each variable stochastically and assumed to be uncorrelated with $v_n(t)$.

The path coefficients k_{ij} and the covariance matrix Σ_e are identified on time series data by minimizing the difference between the covariance matrix estimated from the data, and the covariance matrix implied by the structural model in Figure 1.1. Let \mathbf{s} be a set of three time series:

$$\mathbf{s}(n) = [\mathbf{s}_1(n), \mathbf{s}_2(n), \mathbf{s}_3(n)], \quad n = 1, \dots, L. \tag{1.23}$$

The 3x3 covariance matrix estimated from the data is:

$$S = \frac{\mathbf{s}^T \mathbf{s}}{L - 1} \tag{1.24}$$

where L is the number of observations. Covariance matrix implied by the model, respect to Eq. 1.22, is calculated as:

$$\Sigma_v = (1 - \mathbf{k})^{-T} \Sigma_e (1 - \mathbf{k})^{-1} \tag{1.25}$$

where $\mathbf{k} = \begin{bmatrix} 0 & 0 & 0 \\ k_{21} & 0 & 0 \\ k_{31} & k_{32} & 0 \end{bmatrix}$.

Multivariate methods for connectivity analysis

The maximum likelihood (ML) objective function to be optimized is:

$$F = \ln|\Sigma_v| - \text{tr}(S\Sigma_v^{-1}) - \ln|S| \quad (1.26)$$

The ML objective function of Eq. 1.26 is optimized by means of a fitting criterion which employs a Newton-type algorithm based on an analytic gradient. The starting values can be estimated using ordinary least square. Statistical inference takes into account two aspects: the goodness of the overall fit of the model and the difference between alternative models, called stacked-model approach. Briefly, the χ^2 statistic difference test is used to compare two models. A so-called *null-model* is constructed where the path coefficients are constrained to zero. The alternative model allows these parameters to assume different values. The significance of the difference between the two models is expressed by the difference in the χ^2 goodness of fit indicator ([26]).

2. Assessment on in silico data: the neural mass model for data simulation

Melt the clouds of sin and sadness, drive the dark of doubt away, Giver of immortal gladness, fill us with the light of day!

EEG results mainly from extracellular current flow, associated with massively summed postsynaptic potentials in synchronously activated and vertically oriented neurons. Modeling these neurophysiological mechanisms can rely upon simplifying assumptions and empirical priors. The last decades literature sees a rising interest in this issue and several ways have been developed to model neural signals. The most feasible approach is based on mathematical tools, called *neural mass models* (NMM). This approach describes the processes generating EEG signals arranging in series and in parallel simplified blocks which simulate the key mechanisms only. A neural mass model of EEG is a surrogate of a cortical area. It usually comprises a small number of neural populations interacting each other and uses only one or two state of variables to represent the mean activity of each single neural population. Their synapses dynamics are described under the assumptions that neurons in the same population share similar inputs and synchronize their activity. Therefore by tuning the kinetics parameters of each population, this procedure is able to design specific signal rhythms and reproduce responses seen empirically. One of the first proposed model is the Wilson-Cowan

Assessment on *in silico* data: the neural mass model for data simulation

oscillator to study synchronization among neural oscillations ([27]). After that, a two population model has been developed to simulate the α rhythm in the thalamus by using a feedback loop incorporating excitatory and inhibitory neuron groups ([28]). These models have been subsequently improved with the Jansen's model which introduces three neural populations with different synaptic kinetics ([29]). These equations are still frequently used especially to investigate on brain connectivity ([30], [31],[32], [33]). An important improvement in the use of neural mass models has been provided by Wendling ([34]). Adding another population to Jansen's model to account the presence of fast interneurons allows to simulate the dynamics of real EEG signals measured with intracerebral electrodes in the hippocampus during epileptic seizures. The majority of neural mass models of EEG responses have been designed to model alpha rhythms; recent studies have emphasised the necessity to produce rhythms in different bands. Some of the literary works cited above show that the kinetic of inhibitory populations have a focal influence on signals generation in particular to generate a γ rhythm ([17]). Hence, in the last decade more attention has been drawn to simulate several rhythms coexisting in the same cortical area ([30], [35], [17]).

The model exploited to simulate our *in silico* data aims to render as much realistic as possible the complexity of a cerebral network. It has been developed by Ursino's equipe and consistent literature works prove its effectiveness in reproducing EEG signal behaviour ([35], [36],[37],[38],[17],[39]).

2.1. The neural mass model

This model simulates a signal of one cortical region and produces an intrinsic rhythm that can vary its frequency band by changing the synaptic kinetics parameters. One region model consists of four neural groups, representing the pyramidal neurons, the excitatory interneurons and inhibitory interneurons with

slow and fast synaptic kinetics. The dynamic of each group is reproduced with a general model consisting of three key blocks in cascade, as shown with a synthetic formalism in Fig.2.1.

Each block is characterized by an input-output relationship. The first one receives from the other groups the so-called post-synaptic potentials y_i and combines them linearly by multiplying for a constant C_{ij} . It results in a average post-synaptic membrane potential v_i which subsequently is converted into an average density of spikes fired by the neurons, z_i . The presence of non-linear mechanisms, such as inhibition and saturation, is simulated with a sigmoidal relationship. Then, the last block reproduces the synaptic kinetics with a second order system, with different parameter values each group. These concepts are summarized by the following equations:

$$\begin{aligned}
 v_i &= \sum_j C_{ij} y_j \\
 z_i &= \frac{2e_0}{1 + e^{-rv_i}} - e_0 \quad i = p, e, s, f \\
 \ddot{y}_i &= G_i \omega_i z_i - 2\omega_i \dot{y}_i - \omega_i^2 y_i
 \end{aligned} \tag{2.1}$$

where the subscript j refers to a presynaptic neural group, y_j is the post-synaptic potential change induced by a unitary synapse coming from other groups, C_{ij} represents the connectivity constant from the j th group to the i th one; parameters e_0 and r , assumed equal for all groups, set the maximal saturation and the slope of the sigmoidal relationship; G_i and ω_i represent the strength and the reciprocal of the time constant of the individual synapses. Different values of the connectivity constants and synapses kinetic parameters can mimic the impulse response of each neuron groups, denoted with the subscript p , e , s and f : pyramidal cells, excitatory interneurons, slow and fast inhibitory interneurons, respectively. As shown in the sigmoidal relationship, the model is maintained in the linear region. This choice implies that all quantities have zero mean and avoids that the equilibrium point in same group shifts to the sub-threshold or to the saturation

region due to an excessive connection.

A particular attention should be drawn to the scheme adopted for the fast inhibitory interneurons, Fig.2.2. This group synapses with itself and is powered by an external input. The rationale of this choice stands in previous work ([40]) which implements self loop to generate γ rhythms. The addition to this neural mass model of a feedback loop with fast inhibitory interneurons allows producing γ rhythm per se without the contribution of the other groups.

Connecting the four groups gives the complete scheme of the neural mass model, as shown in Fig.2.3. It consists of three general blocks cascade of Fig. 2.1 for the pyramidal cells, the excitatory interneurons and the slow inhibitory interneurons plus the scheme of Fig. 2.2 for the fast inhibitory interneurons. An important aspect of the model is the external inputs, targeting the excitatory and the fast inhibitory interneurons. In a physiological context, these comprehend all external signals coming from the other cortical areas. In order to study connectivity between two cortical areas, a linear relationship is assumed between the averaged spike density of pyramidal neurons of the pre-synaptic area, z_p^k , and the input of the post-synaptic area, u_j^h , as follows:

$$u_j^h(t) = n_j^h(t) + N * W_j^{jk} z_p^k(t - \omega) \quad j = p, f \quad (2.2)$$

where $n_j(t)$ represents Gaussian white noise, W_j is the weight factor and ω is the time delay. $N = 20$ tunes the input signal amplitude.

For brevity, in this work the model will be described with a condensed mathematical formalism to highlight the two different impulse responses, $h_{ex}(t)$ and $h_{in}(t)$, for excitatory and inhibitory inputs, respectively.

2.1 The neural mass model

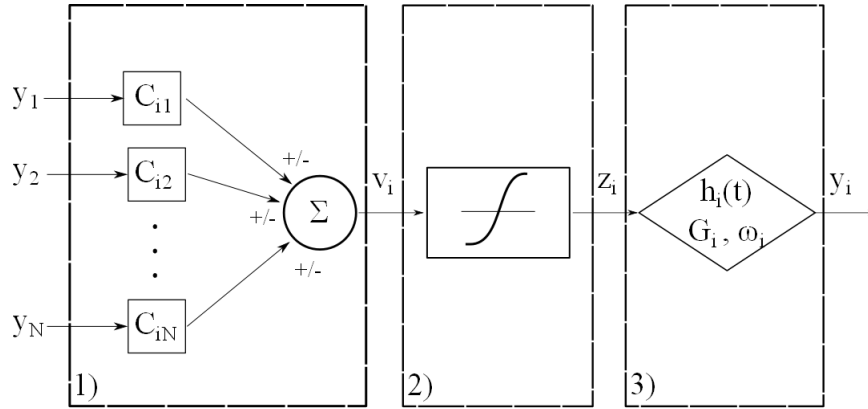


Figure 2.1.: General three blocks cascade to simulate the synapses junction and the information transmission. 1) Pre-synaptic average of potentials coming from the other groups. 2) Conversion to average density of spikes fired by neurons. The non-linear behaviour such as inhibition and saturation is simulated by means of a sigmoidal relationship. 3) Information transmission through synapses: each neuronal group is characterized by specific synaptic kinetics modelled by the impulse response $h_i(t)$, where G_i is the gain and ω_i is the natural frequency. Modified Fig. 1 of [17].

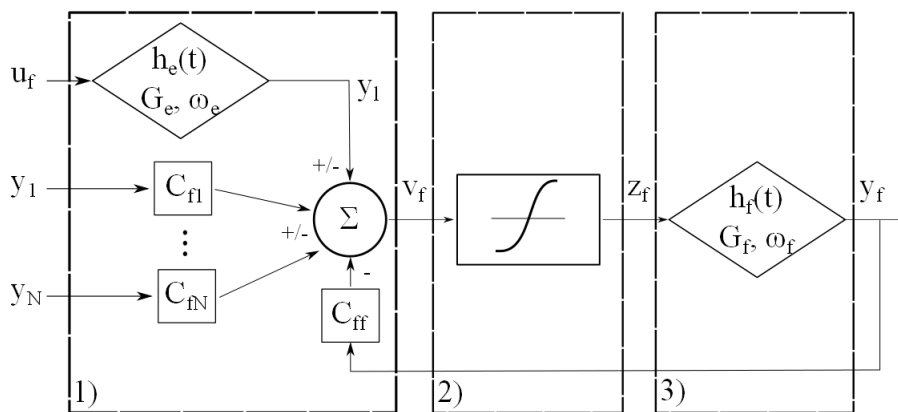


Figure 2.2.: Scheme of fast inhibitory interneurons. It consists of three blocks in cascade as for the general model in Fig. 2.1 but, in addition, the first block is powered by an external input $u_f(t)$ and by the feedback loop. Modified Fig. 2 of [17].

Assessment on in silico data: the neural mass model for data simulation

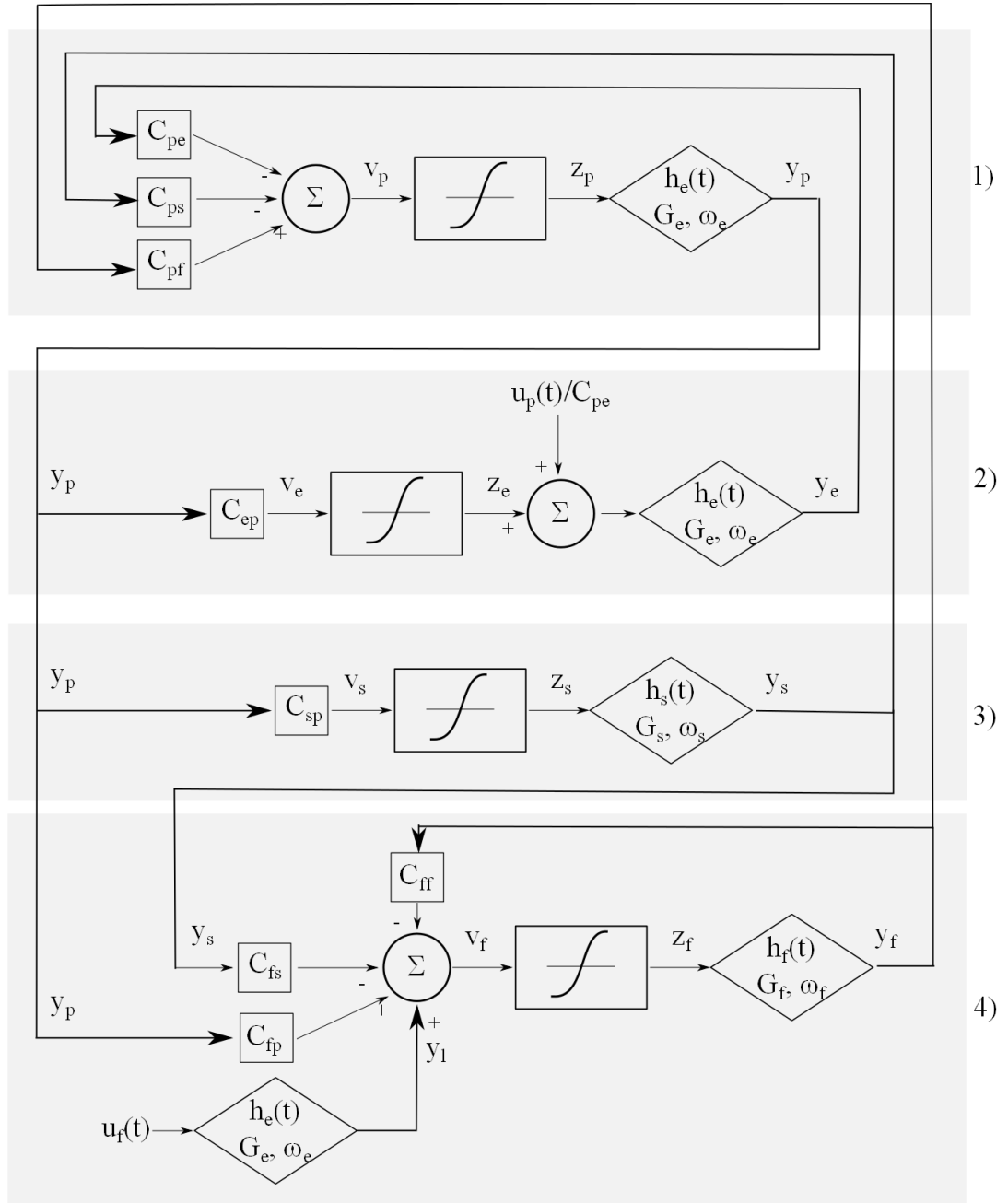


Figure 2.3.: Layout of the complete model for one region. Four neural groups, designed as shown in Fig. 2.1, communicating via excitatory and inhibitory synapses: 1) Pyramidal cells. 2) Excitatory interneurons. 3) Slow inhibitory interneurons. 4) Fast inhibitory interneurons. Modified Fig. 3 of [17].

2.2. NMM parameters analysis

In order to generate electro-physiological signals, ranging in EEG band 0 – 30 Hz, the NMM equations [17] have been studied by varying the kinetic parameters within the set (0, 5, 30, 55, 56, 80, 126, 130) and fixing the others (ω_e , ω_s , ω_f , G_e , G_s , G_f) to the values in Table 1 of [17].

To identify the optimal parameter set, several simulations have been performed to analyse both the system stability and its frequency response. To this purpose, the NMM equations, after been linearised, have been described as multi input-output (MIMO) system of linear differential equations with the state-space representation:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{v} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}\end{aligned}\tag{2.3}$$

where \mathbf{x} is a n by 1 vector representing the state, \mathbf{u} and \mathbf{v} are the input and the output, respectively. The matrices \mathbf{A} (n by n), \mathbf{B} (n by m), and \mathbf{C} (r by n) determine the relationships between the state and input and output variables.

In our case, there are ten first-order differential equations, two inputs, $u_p(t)$ and $u_f(t)$, and four outputs $v_p(t)$, $v_e(t)$, $v_s(t)$, $v_f(t)$. System matrices \mathbf{A} , \mathbf{B} and \mathbf{C} are the following:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -\omega_e^2 & k_1 C_{pe} & -k_1 C_{ps} & -k_1 C_{pf} & -2\omega_e & 0 & 0 & 0 & 0 & 0 \\ k_1 C_{ep} & -\omega_e^2 & 0 & 0 & 0 & -2\omega_e & 0 & 0 & 0 & 0 \\ k_2 C_{sp} & 0 & -\omega_s^2 & 0 & 0 & 0 & -2\omega_s & 0 & 0 & 0 \\ k_3 C_{fp} & 0 & -k_3 C_{fs} & -k_3 C_{ff} - \omega_f^2 & 0 & 0 & 0 & -2\omega_f & k_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\omega_e^2 & -2\omega_e \end{bmatrix}$$

where $k_1 = G_e \omega_e \frac{e_0 r}{2}$, $k_2 = G_s \omega_s \frac{e_0 r}{2}$ and $k_3 = G_f \omega_f \frac{e_0 r}{2}$,

$$\mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{G_e \omega_e}{C_{pe}} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & G_e \omega_e \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 0 & C_{pe} & -C_{ps} & -C_{pf} & 0 & 0 & 0 & 0 & 0 & 0 \\ C_{ep} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ C_{sp} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ C_{fp} & 0 & -C_{fs} & -C_{ff} & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

To evaluate the system stability, we have verified that all eigenvalues of \mathbf{A} lie in the left-hand side, while to analyse the frequency response, we have re-written the state-space representation into Laplace domain, as follows:

$$\begin{aligned} s\mathbf{X}(s) &= \mathbf{A}\mathbf{X}(s) + \mathbf{B}\mathbf{U}(s) \\ (s\mathbf{I} - \mathbf{A})\mathbf{X}(s) &= \mathbf{B}\mathbf{U}(s) \\ \mathbf{X}(s) &= (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{U}(s) \end{aligned}$$

and

$$\begin{aligned} \mathbf{V}(s) &= \mathbf{C}[(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}]\mathbf{U}(s) \\ &= \mathbf{H}(s)\mathbf{U}(s) \end{aligned}$$

where $\mathbf{H}(s)$ is the *matrix transfer function* relating the output vector $\mathbf{V}(s)$ to the input vector $\mathbf{U}(s)$:

$$\mathbf{H}(s) = \frac{\mathbf{C} \text{adj}(s\mathbf{I} - \mathbf{A})\mathbf{B} + \det(s\mathbf{I} - \mathbf{A})\mathbf{D}}{\det(s\mathbf{I} - \mathbf{A})}.$$

In our case, $\mathbf{H}(s)$ is 4 by 2 matrix whose elements are the individual transfer functions relating a given component of the output $\mathbf{V}(s)$ to a component of the inputs $\mathbf{U}(s)$, as described in the following set of equations:

$$\begin{bmatrix} V_p(s) \\ V_e(s) \\ V_s(s) \\ V_f(s) \end{bmatrix} = \begin{bmatrix} H_{pp}(s) & H_{pf}(s) \\ H_{ep}(s) & H_{ef}(s) \\ H_{sp}(s) & H_{sf}(s) \\ H_{fp}(s) & H_{ff}(s) \end{bmatrix} \begin{bmatrix} U_p(s) \\ U_f(s) \end{bmatrix}.$$

Among the subsets of parameter values assuring the system stability, spectral analysis of the frequency response has been performed in order to find those generating signals with well-defined frequency peaks. We focused the attention on the transfer functions relating the inputs with $V_p(s)$, since $v_p(t)$ represents the cortical pyramidal cells signal. By evaluating their response magnitude:

$$|H_{pp}(\omega)|^2 \quad \text{and} \quad |H_{pf}(\omega)|^2 \tag{2.4}$$

we choose those sets giving peak frequency around 5 Hz, 15 Hz and 30 Hz.

Since these two transfer functions refer to excitatory and inhibitory cells, respectively, to highlight their meaning they are called $H_{ex}(\omega)$ and $H_{in}(\omega)$ and, hence $h_{ex}(t)$ and $h_{in}(t)$ in time domain.

2.3. Simulated dataset

The NMM described in 2.1 is used to simulate a ROI cortical EEG, where the input is white noise $n(t)$ - with zero mean and variance $\sigma^2 = 5$ for both impulse response $h_{ex}(t)$ and $h_{in}(t)$ and the output is $v_n(t)$, corresponding to the NMM $v_p(t)$, represented in condensed formalism in Fig.2.4.

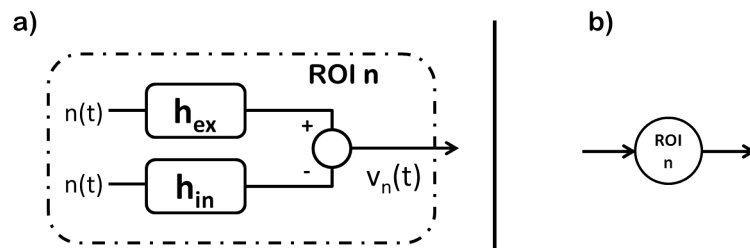


Figure 2.4.: One ROI model. Each ROI is characterized by two impulse responses, $h_{ex}(t)$ and $h_{in}(t)$, for excitatory and inhibitory inputs, $n(t)$, which are assumed gaussian with zero mean and variance $\sigma^2 = 5$. $v_n(t)$ corresponds to the pyramidal output $v_p(t)$ in the NMM (a). Picture in (b) shows the symbolism used for synthetic representations.

By combining three populations - called ROI 1, ROI 2 and ROI 3, connected by weight parameters A, B, C and characterized by different synaptic kinetics, four different network models have been analysed. The first one is a feed-forward network, shown Fig.2.5.

Three datasets are generated with different non linearity level defined by the slope adopted in each sigmoid relationship of the model, assuming the following values:

- dataset 1 with $r = 0.36$;
- dataset 2 with $r = 0.56$;
- dataset 3 with $r = 0.66$.

For each dataset, four conditions are performed comprehending a basal one with unit value in A, B and C and three obtained by fixing different values of weight

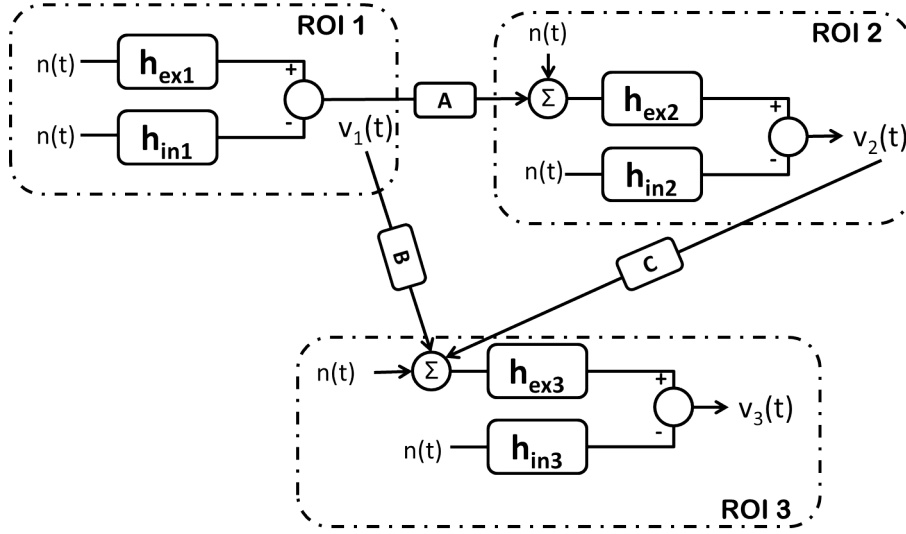


Figure 2.5.: Layout of the neural network model used to simulate data. The connection intensity between coupled ROIs is described by weight parameters A, B and C.

parameters as follows and reported in Tab.2.1:

- basal condition: $A=B=C=1$;
- condition i): A assumes the values 2, 3 and 4, B and C are fixed at 1;
- condition ii): B assumes the values 2, 3 and 4, A and C are fixed at 1;
- condition iii): C assumes the values 2, 3 and 4, B and C are fixed at 1.

Summing up, a total of thirty experiments are simulated, for each of them one hundred realizations of 2 s are generated for the three ROIs.

The other networks, obtained by combining the same three ROIs, have been tested performing a reduced number of experiments. An open-loop network with two directed links, connecting ROI 1 to ROI 2 and ROI 2 to ROI 3, have been firstly analysed; then, we added a feedback link from ROI 2 to ROI 1 in order to examining the effect of a reciprocal connection; finally, we considered a cycle network. For each of them, two experiments have been performed, comprehending a basal condition with all weights fixed at 1 and the other one with B assuming value 2, as shown in Tab.2.2.

Assessment on in silico data: the neural mass model for data simulation

Networks	Weight values
	<ul style="list-style-type: none"> • $A=B=C=1$
	<ul style="list-style-type: none"> • $A=2 \ B=C=1$ • $A=3 \ B=C=1$ • $A=4 \ B=C=1$
	<ul style="list-style-type: none"> • $B=2 \ A=C=1$ • $B=3 \ A=C=1$ • $B=4 \ A=C=1$
	<ul style="list-style-type: none"> • $C=2 \ A=B=1$ • $C=3 \ A=B=1$ • $C=4 \ A=B=1$

Table 2.1.: Feed-forward network schemes and weight values assumed for each experiment.

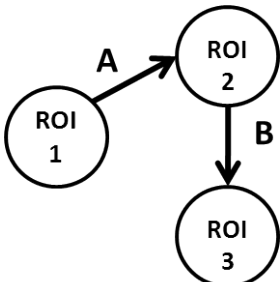
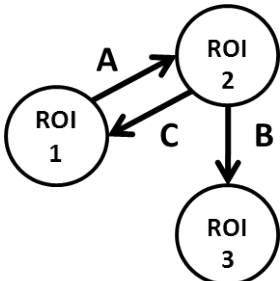
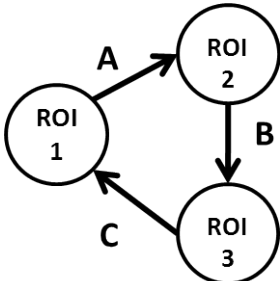
Networks	Weight values
	<ul style="list-style-type: none"> • $A=B=C=1$ • $A=1 B=2$
	<ul style="list-style-type: none"> • $A=B=C=1$ • $A=1 B=2 C=1$
	<ul style="list-style-type: none"> • $A=B=C=1$ • $A=1 B=2 C=1$

Table 2.2.: Additional network models, open-loop (*upper panel*), network with feed-back link (*middle panel*) and cycle network (*lower panel*), and weight values assumed for each experiment.

3. Data Analysis

*All Thy works with joy surround Thee, earth and heaven reflect Thy rays,
Stars and angels sing around Thee, center of unbroken praise.*

3.1. Simulation

Analysis of NMM parameters to evaluate system stability and frequency response of transfer functions have been performed using the Matlab package reported in Appendix. After identifying those parameter sets generating well-defined frequency peak in low (5 Hz), medium (16 Hz) and high frequency (32 Hz), *in silico* EEG have been simulated using an *ad-hoc* Matlab code provided by Ursino's equipe.

For each experiment described in 2.3 one hundred realizations 2 s long of three joined time series are generated, with sampling frequency $F_s = 200Hz$.

3.2. Connectivity estimation

A multivariate model is fitted to each simulation data by means of the Matlab package ARFIT, based on stepwise least square algorithm, [41], selecting the best order Akaike's information criterion. Then, Granger causality estimation is achieved using the Matlab toolbox GCCA, [42], with the ordinary-least-squares option, and frequency indexes computation is performed by applying the Matlab toolbox implemented in [43]. Estimation of SEM path coefficients is accomplished

by analysing data in R using its package "sem", [44]. Obtained results are averaged over the one hundred realizations for each experiment.

3.3. Statistical significance

3.3.1. F-test for Granger causality

Statistic described in eq. 1.8 is an application to regression problems of classical F-test:

$$F = \frac{\textit{explained variance}}{\textit{unexplained variance}}$$

where explained variance is associated to the difference of the Residual Sum of Squares of restricted and unrestricted models and the unexplained variance is the Residual Sum of Squares of unrestricted model, corrected for numerator and denominator degrees of freedom, eq.1.8. Model with more parameters will always be able to fit the data at least as well as the model with fewer parameters. Thus typically unrestricted model will give a lower prediction error than the restricted one. In order to determine whether the unrestricted model gives a significantly better fit to the data the F calculated from the data should be greater than the critical value of the F-distribution for some desired false-rejection probability. These tests are corrected for multiple comparisons via Bonferroni correction. Hence, the applied threshold is $\frac{0.05}{n(n-1)}$, where $n=3$ in our case.

3.3.2. Null hypothesis test for DTF and PDC

Statistical test applied for DTF and PDC is a modified version of the surrogate data strategy proposed in [45]. Instead of shuffling the time series, their phases are mixed up randomly in order to break time relationships. Then, a three-variate autoregressive model is fit to surrogate data to compute DTF and PDC.

3.4 Assessment of estimated indexes

Under the null hypothesis, for each experiment, performed as described in 2.3, the null distribution is estimated by pooling together the peak value reached in each realization, as shown in fig. 3.1. This choice has two main reasons. First, it is cautious since considers high indexes values avoiding as much as possible spurious connections. Second, it is a global value independent of the frequency structure of the data. Threshold at $P < 0.05$ fixes γ_0 critical value at its 95-th percentile. Hence, there exists a connection for a specific frequency f_o between two signals if DTF/PDC function overcomes threshold line in correspondence to f_o .

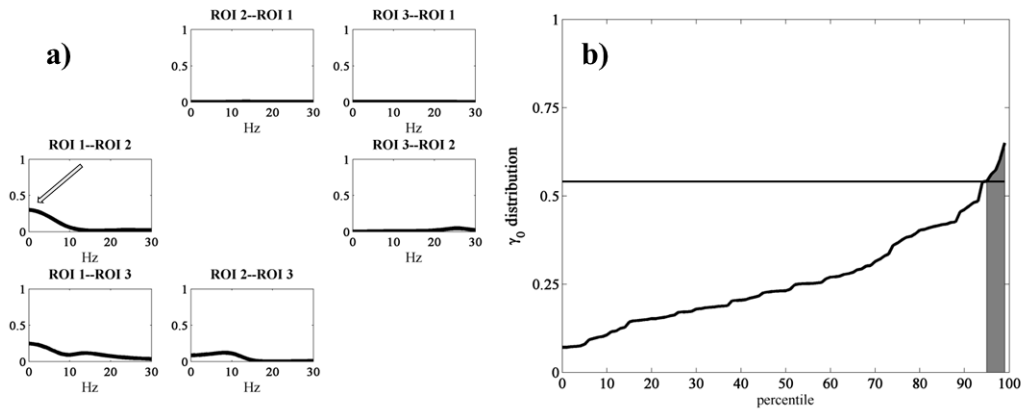


Figure 3.1.: Representative example of null distribution resulting from 100 surrogate realizations for DTF. a) DTF functions among ROI 1, ROI 2 and ROI 3 when causal influence is absent: grey arrow indicates the peak value. b) DTF distribution by combining results from all one hundred realizations.

3.4. Assessment of estimated indexes

Topology analysis. Network topology is estimated by means of MVAR indexes. GC, DTF and PDC are calculated for each realization and statistical tests are performed to evidence significant values. Comparison with true network gives the amount of false negatives and false positives results and statistical power of GC, DTF and PDC is described in terms of sensitivity and specificity:

Data Analysis

		Condition		
		True	False	
Outcome	Positive	TP	FP	Type I error (α)
	Negative	FN	TN	Type II error (β)
		Sensitivity	Specificity	

Table 3.1.: Relationships between actual condition (true or false) and predicted outcomes (positive or negative). False positives (FP) and negatives (FN) provide type I (α) and type II (β) error rate, respectively. Ratio of true positives (TP) to combined TP and FN gives the sensitivity amount, while ratio of true negative (TN) to combined FP and TN furnishes the specificity.

$$Sensitivity = \frac{TP}{TP+FN}$$

and

$$Specificity = \frac{TN}{FP+TN}$$

where sensitivity relates to the ability of identifying true connections; specificity refers to the test ability of identifying absence of connections, Tab.4.5.

Two-way analysis of variance (ANOVA) for repeated measures is used to examine the influence of the network model parameters, such are sigmoid slope and connection weights, on the topology estimation, quantified by the percentage of false negatives and positives.

Strength analysis. Network strength connections are evaluated by considering the output scores of GC, eq. 1.7, DTF and PDC AUC integrals, eq. 1.21, and SEM path coefficients. Since estimates and true weights are measured with different scales, they are compared using linear regression to verify the existence of a linear relationship between them.

4. Assessment on in silico data: results

*Field and forest, vale and mountain, flowery meadow, flashing sea,
Singing bird and flowing fountain call us to the joyous rejoice in Thee.*

This chapter consists of two parts. The first one shows *in silico* data obtained by tuning network model parameters, such are slopes of sigmoid relationships, connection weights and linking direction. The second one reports connectivity outputs of each method. Results obtained with *in silico* data aim to quantify the accuracy of MVAR and SEM methods. Since they are based on linear equations, it is a crucial aspect measuring to what extent they approximate connectivity of non linear signals.

4.1. Simulation

4.1.1. NMM parameters

System stability analysis has been performed considering 7^7 different combinations of NMM parameter values. Among the sets assuring the stability, a subset of 288 have been considered to study the transfer functions frequency response, $H_{ex}(f)$ and $H_{in}(f)$, in order to find three sets producing well-defined frequency peak in low (5 Hz), medium (16 Hz) and high frequency (32 Hz), Fig.4.1. Parameter values adopted in NMM equations are reported in Tab.4.1.

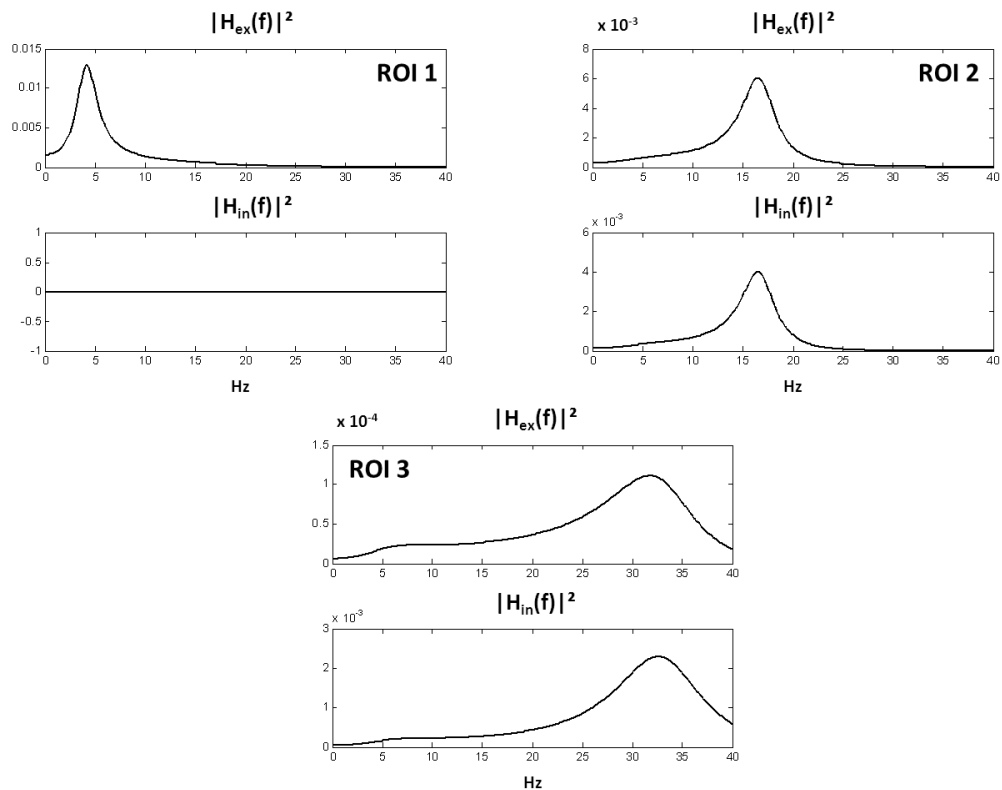


Figure 4.1.: Frequency response magnitude of excitatory and inhibitory impulse response, $H_{ex}(f)$ and $H_{in}(f)$, of each ROI computed on the linearized model described in 2.2.

4.1.2. Model predicted EEG signals

Examples of signals of the different network model are shown in Fig. 4.2, evidencing how the linking direction influences the frequency content. In feed-forward and open-loop networks Fig. 4.2 (a) and (b), ROI 2 clearly exhibits spectral contribute coming from ROI 1, while in ROI 3 there are not significant differences between the two network models, since its intrinsic gain is very low in the pass-band of ROI 1 and ROI 2 and, hence, frequency contents of inflowing ROIs are less evident. Differently, networks in Fig.4.2 (c) and (d) produce an increasing in ROI 3 PSD, thanks to the synergic connection of the feed-back link.

Focusing the attention on the feed-forward network, Figs. 4.3 and 4.4 show the

(a) Common parameters.

Parameters	Symbols	
Average gain (mV)	G_e	5.17
	G_s	4.45
	G_f	57.1
Time Constant reciprocal (s^{-1})	w_e	75
	w_s	30
	w_f	75
Sigmoid saturation (s^{-1})	e_0	2.5
Time delay (ms)	ω	10
Input noise variance	σ^2	5

(b) Regions' parameters.

Parameters	Symbols	ROI 1	ROI 2	ROI 3
Number of synaptic contacts	C_{ep}	55	5	130
	C_{pe}	5	5	5
	C_{sp}	5	5	105
	C_{ps}	55	55	130
	C_{fp}	56	56	80
	C_{fs}	5	5	126
	C_{pf}	0	5	30
	C_{ff}	5	5	30

Table 4.1.: Network model kinetics parameters. a) Common parameters are average gains, time constant reciprocals, sigmoid saturations and time delays. Values are the same of [38]. b) Regions' parameters are the synaptic contact numbers.

model prediction in time and frequency domain in a representative realization characterized by different values of network sigmoid slope, r . This parameter influences frequency contents, since it directly modifies the intrinsic gain of each ROI.

Assessment on in silico data: results

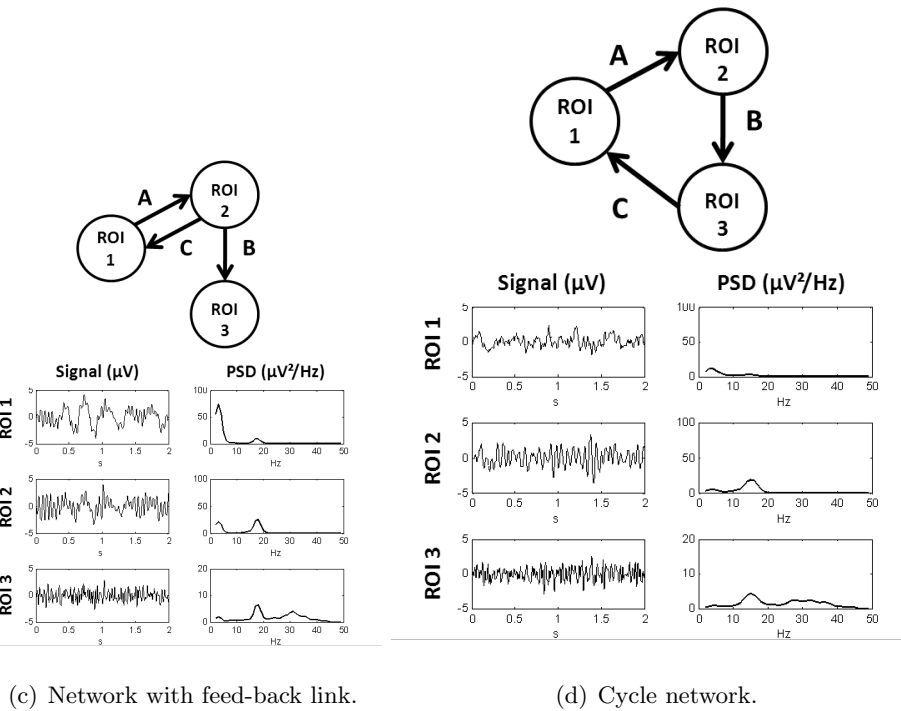
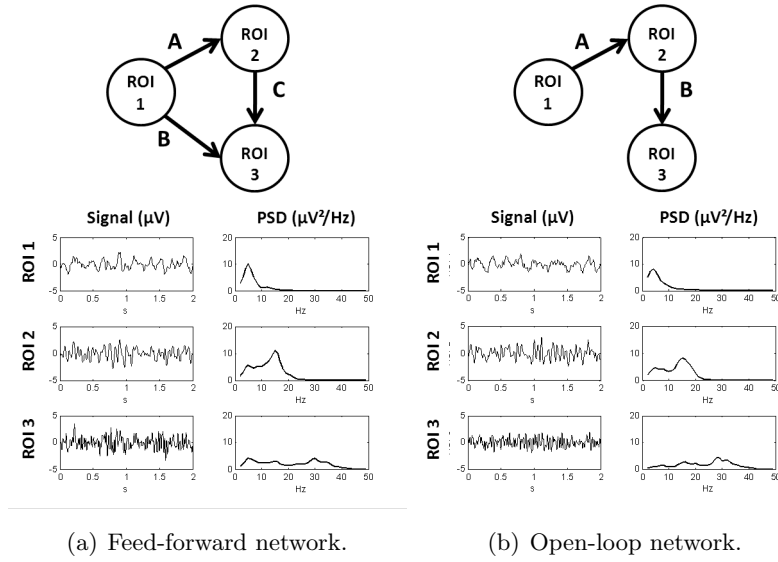


Figure 4.2.: Model predicted EEG signals in time and frequency domain for each network model in basal condition, where all weights are equal to 1 and $r = 0.56$.

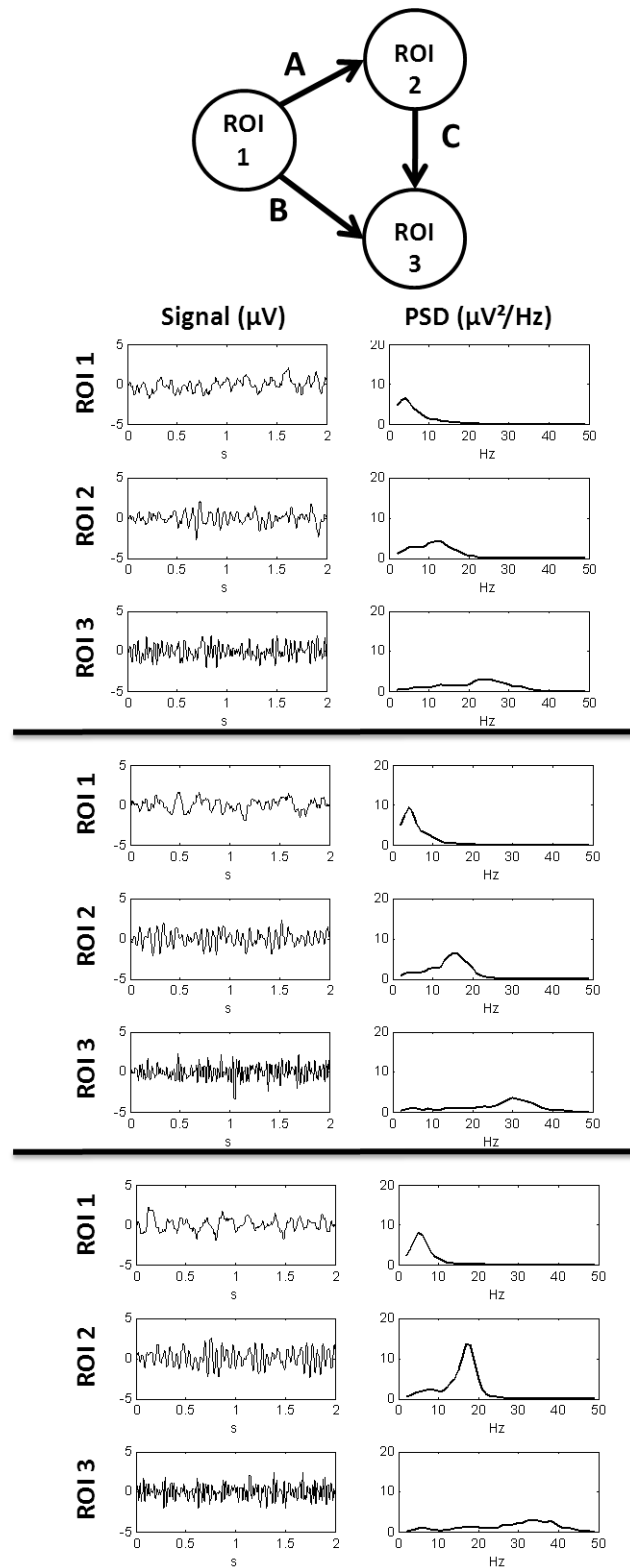


Figure 4.3.: Basal condition ($A=B=C=1$). Model predicted EEG signals in time and frequency domain in a representative realization for each dataset. (Upper panel $r = 0.36$; middle panel $r = 0.56$; lower panel $r = 0.66$.)

Assessment on in silico data: results

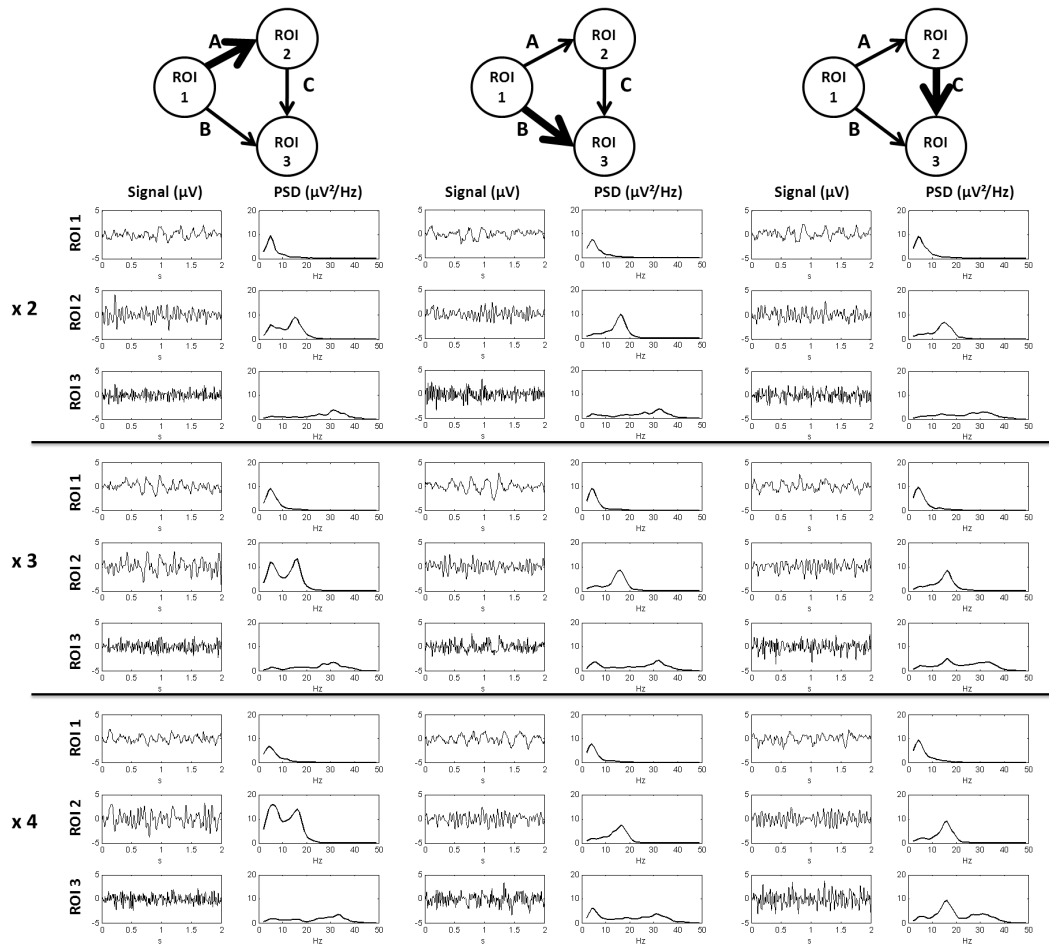


Figure 4.4.: Model predicted EEG signals in time and frequency domain in a representative realization of experiments characterized by different values of A, B and C parameters of the network bold link ($r = 0.56$ dataset).

4.2. Connectivity estimation

In order to estimate connectivity via Granger causality and frequency indexes, a MVAR model has been fitted to the time series obtained by simulation. Optimum model order has been selected by Akaike’s criterion with values varying within the range 5-7.

4.2.1. Feed-forward network

Analysis of a single condition

Results reported below show the complete list of outcomes produced by connectivity estimation for the representative condition with increasing strength in the link between ROI 1 and ROI 2. GC estimates of strength and percent of true and false positives are displayed in Tab. 4.2. DTF and PDC frequency functions are illustrated in Figs. 4.2.1 and 4.2.1, respectively. Their AUC values and percent of true and false positives are summarized in Tabs. 4.3 and 4.4. Statistical performances of each index in terms of type I error ($\alpha\%$), type II error ($\beta\%$), sensitivity and specificity are reported in Tab.4.5. Finally, SEM outputs are show in Tab. 4.6.

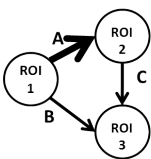
	Links	A=2 B=C=1	A=3 B=C=1	A=4 B=C=1
	$ROI1 \rightarrow ROI2$	0.16 (100%)	0.27 (100%)	0.36 (100%)
	$ROI1 \rightarrow ROI3$	0.03 (72%)	0.03 (57%)	0.02 (51%)
	$ROI2 \rightarrow ROI3$	0.03 (98%)	0.04 (100%)	0.04 (100%)
	$ROI2 \rightarrow ROI1$	< 0.001 (7%)	< 0.001 (6%)	< 0.001 (9%)
	$ROI3 \rightarrow ROI1$	< 0.001 (3%)	0.002 (7%)	< 0.001 (2%)
	$ROI3 \rightarrow ROI2$	< 0.001 (3%)	< 0.001 (4%)	0.003 (9%)

Table 4.2.: Average GC index estimates with the percentage of true and false positives).

Assessment on in silico data: results

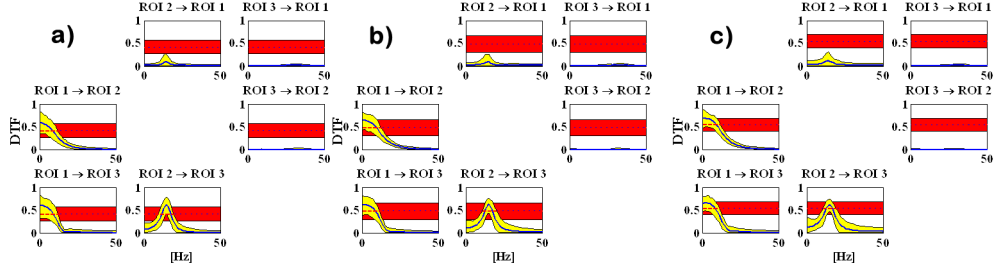


Figure 4.5.: DTF index as a function of frequency. Panels (a), (b) and (c) are relative to different values of A , the weight parameter linking ROI 1 to ROI 2. Yellow and red bands represent 100 realizations of DTF and threshold, blue and dotted red lines are their mean functions and average thresholds, respectively.

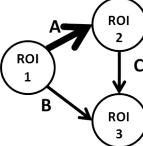
Links	$A=2$ $B=C=1$	$A=3$ $B=C=1$	$A=4$ $B=C=1$
			
$ROI1 \rightarrow ROI2$	7.60 (83%)	9.60 (81%)	10.55 (92%)
$ROI1 \rightarrow ROI3$	7.10 (89%)	7.88 (76%)	8.66 (80%)
$ROI2 \rightarrow ROI3$	7.25 (94%)	8.29 (85%)	9.14 (73%)
$ROI2 \rightarrow ROI1$	1.22 (9%)	1.50 (6%)	1.97 (3%)
$ROI3 \rightarrow ROI1$	0.46 (5%)	0.52 (5%)	0.49 (3%)
$ROI3 \rightarrow ROI2$	0.30 (5%)	0.23 (5%)	0.21 (3%)

Table 4.3.: DTF index estimates as AUC integrals computed from the mean DTF function with percentages of true and false positives).

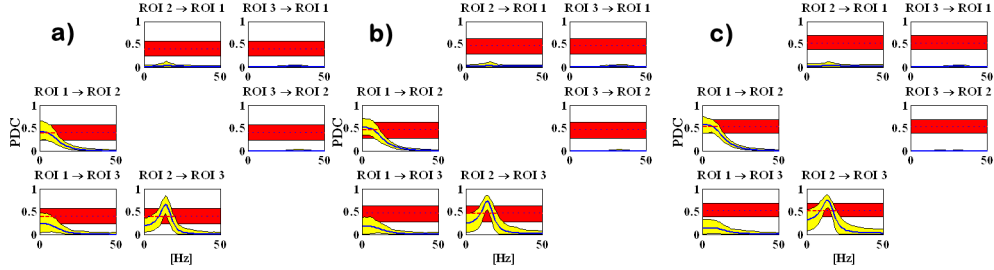


Figure 4.6.: PDC index as a function of frequency. Panels (a), (b) and (c) are relative to different values of A , the weight parameter linking ROI 1 to ROI 2. Yellow and red bands represent 100 realizations of PDC and threshold, blue and dotted red lines are their mean functions and average thresholds, respectively.

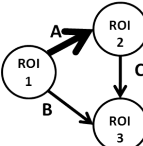
Links	$A=2$ $B=C=1$	$A=3$ $B=C=1$	$A=4$ $B=C=1$
			
$ROI1 \rightarrow ROI2$	6.22 (59%)	8.25 (55%)	9.24 (62%)
$ROI1 \rightarrow ROI3$	3.86 (18%)	3.39 (8%)	3.26 (3%)
$ROI2 \rightarrow ROI3$	8.92 (96%)	10.75 (100%)	12.40 (95%)
$ROI2 \rightarrow ROI1$	0.87 (5%)	1.03 (5%)	1.28 (3%)
$ROI3 \rightarrow ROI1$	0.47 (5%)	0.53 (5%)	0.52 (3%)
$ROI3 \rightarrow ROI2$	0.28 (5%)	0.23 (5%)	0.17 (3%)

Table 4.4.: PDC index estimates as AUC integrals computed from the mean PDC function with percentages of true and false positives).

4.2 Connectivity estimation

GC				
		Condition		
		True	False	
Outcome	Positive	2929	189	$\alpha = 5\%$
	Negative	671	3411	$\beta = 18\%$
		Sensitivity	Specificity	
		0.81	0.94	
DTF				
		Condition		
		True	False	
Outcome	Positive	2863	149	$\alpha = 4\%$
	Negative	737	3451	$\beta = 20\%$
		Sensitivity	Specificity	
		0.79	0.95	
PDC				
		Condition		
		True	False	
Outcome	Positive	2005	124	$\alpha = 4\%$
	Negative	1595	3476	$\beta = 44\%$
		Sensitivity	Specificity	
		0.55	0.96	

Table 4.5.: Statistical measures of GC, DTF and PDC performances.

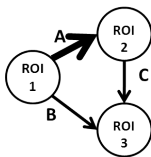
	Links	A=2 B=C=1	A=3 B=C=1	A=4 B=C=1
	$ROI1 \rightarrow ROI2$	0.36	0.45	0.51
	$ROI1 \rightarrow ROI3$	0.14	0.12	0.10
	$ROI2 \rightarrow ROI3$	0.13	0.20	0.25
	$ROI2 \rightarrow ROI1$	0.03	0.03	0.03
	$ROI3 \rightarrow ROI1$	0.04	0.03	0.04
	$ROI3 \rightarrow ROI2$	0.03	0.03	0.04

Table 4.6.: SEM average estimates.

Analysis of all experiments

Following paragraphs report overall results throughout all experiments performed on the feed-forward network, describing both topology and strength estimate.

Topology. Network topology is inferred by means of statistical test responses, such are F-test for GC and comparison with the null hypothesis threshold for DTF and PDC.

Percentage of positive connections identified between each pair of ROIs, performed by considering a total number of 3000 tests, are listed in Tabs. 4.7-4.8-4.9, for the three indexes, respectively. As expected, percentage of identified connections is larger in links where the connection is true and in each condition the strongest link clearly emerged, since it always has the highest score. Results of two way ANOVA computed to analyse the influence of network model parameters reveal that there are no statistical differences among datasets and conditions i), ii) and iii), except for false negatives amount of GC outputs. A significant difference ($P < 0.05$) appears within experiments only, where false negatives percentage relative to predominant links is smaller respect to those characterized by equal strength connections. Results averaged by pooling datasets and conditions are condensed in Fig.4.7, reporting the mean percentage of statistical significant connections for each index. Green arrows are true connections and the red ones represent links where connection is absent. For each index amount of false connections is about 5%, while true recognition varies depending on the kind of connection, which can be direct or linear mixed with interacting ROIs, and the statistical testing strategy used. Statistical performance measures are summarized in Tab. 4.10 reporting probability of error I less than 5% but high false negative rates for each index. Great value of specificity, about 95% for each index, respect to sensitivity suggests the methods are highly conservative.

4.2 Connectivity estimation

	Links	DATASET 1	DATASET 2	DATASET 3						
		$A = B = C = 1$								
	$ROI1 \rightarrow ROI2$	63	97	98						
	$ROI1 \rightarrow ROI3$	46	88	98						
	$ROI2 \rightarrow ROI1$	3	4	5						
	$ROI2 \rightarrow ROI3$	51	92	100						
	$ROI3 \rightarrow ROI1$	5	4	9						
	$ROI3 \rightarrow ROI2$	3	3	1						
	Links	DATASET 1	DATASET 2	DATASET 3						
		x2	x3	x4	x2	x3	x4	x2	x3	x4
	$ROI1 \rightarrow ROI2$	99	100	100	100	100	100	100	100	100
	$ROI1 \rightarrow ROI3$	33	40	14	72	57	51	90	72	67
	$ROI2 \rightarrow ROI1$	4	2	4	7	6	9	4	8	12
	$ROI2 \rightarrow ROI3$	56	75	73	98	100	100	100	99	100
	$ROI3 \rightarrow ROI1$	3	7	4	3	7	2	5	7	6
	$ROI3 \rightarrow ROI2$	7	6	1	3	4	9	2	10	10
	Links	DATASET 1	DATASET 2	DATASET 3						
		x2	x3	x4	x2	x3	x4	x2	x3	x4
	$ROI1 \rightarrow ROI2$	73	48	43	86	80	74	96	90	84
	$ROI1 \rightarrow ROI3$	98	99	100	100	100	100	100	100	100
	$ROI2 \rightarrow ROI1$	2	6	3	5	6	4	6	4	2
	$ROI2 \rightarrow ROI3$	54	41	50	90	96	95	100	100	99
	$ROI3 \rightarrow ROI1$	10	3	2	3	6	7	6	13	11
	$ROI3 \rightarrow ROI2$	6	5	3	3	5	10	6	2	1
	Links	DATASET 1	DATASET 2	DATASET 3						
		x2	x3	x4	x2	x3	x4	x2	x3	x4
	$ROI1 \rightarrow ROI2$	65	65	79	98	97	96	99	97	98
	$ROI1 \rightarrow ROI3$	41	39	46	84	78	85	93	93	96
	$ROI2 \rightarrow ROI1$	1	6	6	2	4	4	3	5	6
	$ROI2 \rightarrow ROI3$	100	100	100	100	100	100	100	100	100
	$ROI3 \rightarrow ROI1$	2	4	4	7	8	7	7	1	4
	$ROI3 \rightarrow ROI2$	1	3	7	4	3	6	5	12	5

Table 4.7.: GC index: percent of true and false positives in each experiment. Bold arrow in network model indicates the link with true strength multiplied by factors 2, 3 and 4.

Assessment on in silico data: results

	Links	DATASET 1	DATASET 2	DATASET 3						
		$A = B = C = 1$								
	$ROI1 \rightarrow ROI2$	76	83	82						
	$ROI1 \rightarrow ROI3$	85	98	97						
	$ROI2 \rightarrow ROI1$	14	7	6						
	$ROI2 \rightarrow ROI3$	82	100	100						
	$ROI3 \rightarrow ROI1$	5	0	0						
	$ROI3 \rightarrow ROI2$	3	0	0						
	Links	DATASET 1	DATASET 2	DATASET 3						
		x2	x3	x4	x2	x3	x4	x2	x3	x4
	$ROI1 \rightarrow ROI2$	80	94	92	83	81	92	71	83	72
	$ROI1 \rightarrow ROI3$	53	82	58	89	76	80	80	86	76
	$ROI2 \rightarrow ROI1$	6	3	3	9	6	3	4	3	5
	$ROI2 \rightarrow ROI3$	43	46	35	94	85	73	90	92	74
	$ROI3 \rightarrow ROI1$	5	3	3	5	5	3	4	3	5
	$ROI3 \rightarrow ROI2$	5	3	3	5	5	3	4	3	5
		Links	DATASET 1	DATASET 2	DATASET 3					
			x2	x3	x4	x2	x3	x4	x2	x3
$ROI1 \rightarrow ROI2$		46	19	29	38	41	43	40	33	43
$ROI1 \rightarrow ROI3$		93	97	99	96	100	99	100	99	100
$ROI2 \rightarrow ROI1$		9	5	6	9	9	9	5	5	3
$ROI2 \rightarrow ROI3$		33	19	18	84	69	60	95	91	66
$ROI3 \rightarrow ROI1$		10	2	6	8	8	8	4	4	3
$ROI3 \rightarrow ROI2$		9	2	5	8	8	8	4	4	3
		Links	DATASET 1	DATASET 2	DATASET 3					
			x2	x3	x4	x2	x3	x4	x2	x3
	$ROI1 \rightarrow ROI2$	38	35	43	41	30	44	32	31	36
	$ROI1 \rightarrow ROI3$	55	50	58	72	46	60	68	54	57
	$ROI2 \rightarrow ROI1$	8	7	3	7	5	6	10	5	5
	$ROI2 \rightarrow ROI3$	94	98	100	100	100	100	100	99	100
	$ROI3 \rightarrow ROI1$	8	7	3	7	4	4	10	5	5
	$ROI3 \rightarrow ROI2$	8	8	3	7	4	4	10	5	5

Table 4.8.: DTF index (AUC): percent of true positives and false positives in each experiment. Bold arrow in network model indicates the link with true strength multiplied by factors 2, 3 and 4.

4.2 Connectivity estimation

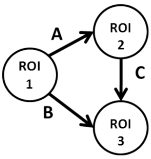
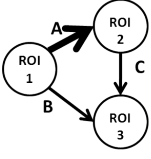
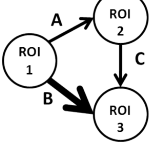
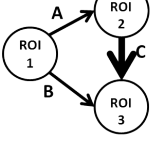
	Links	DATASET 1	DATASET 2	DATASET 3	
		$A = B = C = 1$	$A = B = C = 1$	$A = B = C = 1$	
	$ROI1 \rightarrow ROI2$	68	60	25	
	$ROI1 \rightarrow ROI3$	72	86	84	
	$ROI2 \rightarrow ROI1$	7	0	0	
	$ROI2 \rightarrow ROI3$	84	100	100	
	$ROI3 \rightarrow ROI1$	6	0	0	
	$ROI3 \rightarrow ROI2$	3	0	0	
	Links	DATASET 1	DATASET 2	DATASET 3	
		x2 x3 x4	x2 x3 x4	x2 x3 x4	
	$ROI1 \rightarrow ROI2$	69 92 90	59 55 62	25 33 35	
	$ROI1 \rightarrow ROI3$	11 8 4	18 8 3	16 8 6	
	$ROI2 \rightarrow ROI1$	5 3 3	5 5 3	4 3 5	
	$ROI2 \rightarrow ROI3$	51 45 45	96 100 95	97 99 96	
	$ROI3 \rightarrow ROI1$	5 3 3	5 5 3	4 3 5	
	$ROI3 \rightarrow ROI2$	5 3 3	5 5 3	4 3 5	
		Links	DATASET 1	DATASET 2	DATASET 3
			x2 x3 x4	x2 x3 x4	x2 x3 x4
$ROI1 \rightarrow ROI2$		22 4 8	8 8 8	4 4 3	
$ROI1 \rightarrow ROI3$		82 96 98	93 99 99	94 98 100	
$ROI2 \rightarrow ROI1$		9 5 6	8 8 8	4 4 3	
$ROI2 \rightarrow ROI3$		34 27 30	88 85 69	96 98 85	
$ROI3 \rightarrow ROI1$		10 2 6	8 8 8	4 4 3	
$ROI3 \rightarrow ROI2$		9 2 5	8 8 8	4 4 3	
		Links	DATASET 1	DATASET 2	DATASET 3
			x2 x3 x4	x2 x3 x4	x2 x3 x4
	$ROI1 \rightarrow ROI2$	34 27 42	18 15 27	11 7 16	
	$ROI1 \rightarrow ROI3$	24 22 14	34 15 13	28 11 19	
	$ROI2 \rightarrow ROI1$	8 7 3	7 4 4	10 5 5	
	$ROI2 \rightarrow ROI3$	97 100 100	100 100 100	100 100 100	
	$ROI3 \rightarrow ROI1$	8 7 3	7 4 4	10 5 5	
$ROI3 \rightarrow ROI2$	8 8 3	7 4 4	10 5 5		

Table 4.9.: PDC index (AUC): percent of true positives and false positives in each experiment. Bold arrow in network model indicates the link with true strength multiplied by factors 2, 3 and 4.

(a) GC outcomes.					
		Positive Condition	Negative Condition		
Positive Outcomes		7574	456	False Positive Rate (α)	5%
Negative Outcomes		1426	8544	False Negative Rate (β)	16%
	Sensitivity	84%	Specificity		
(b) DTF outcomes.					
		Positive Condition	Negative Condition		
Positive Outcomes		6354	476	False Positive Rate (α)	5%
Negative Outcomes		2646	8524	False Negative Rate (β)	29%
	Sensitivity	71%	Specificity		
(c) PDC outcomes.					
		Positive Condition	Negative condition		
Positive Outcomes		4819	443	False Positive Rate (α)	5%
Negative Outcomes		4181	8557	False Negative Rate (β)	46%
	Sensitivity	54%	Specificity		

Table 4.10.: Statistical measures of the performances of GC (a), DTF (b) and PDC (c) throughout the experimental conditions.

4.2 Connectivity estimation

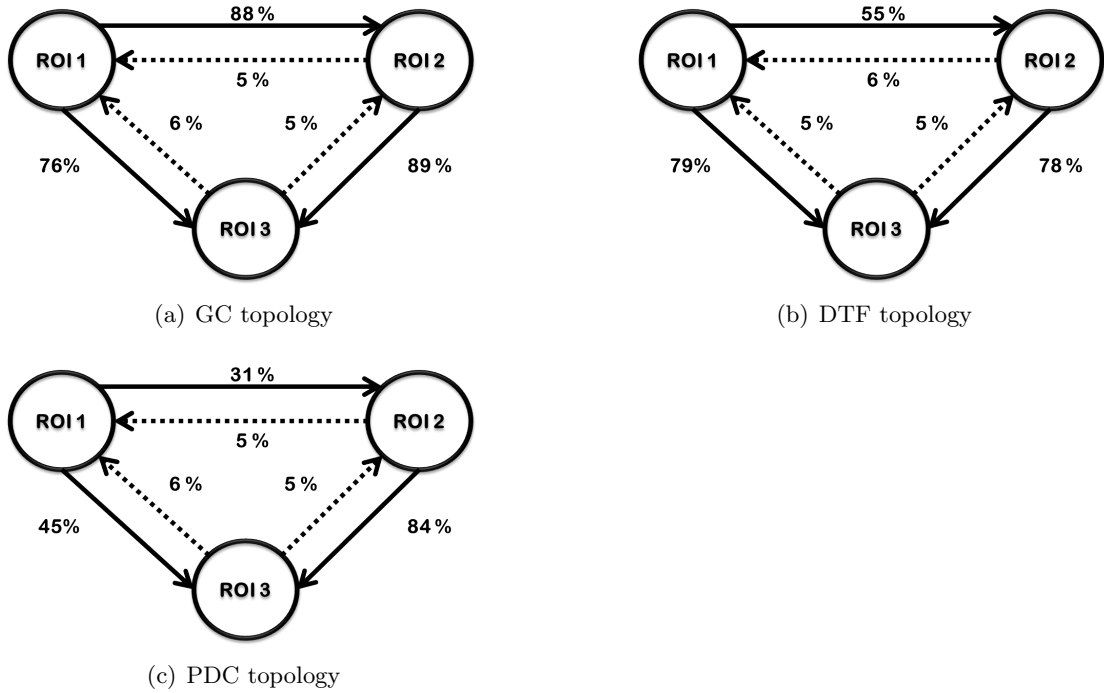


Figure 4.7.: Estimated topology for each index: overall true (bold line) and false (dotted line) positives in percent throughout the experiments.

Strength estimation. Estimates of strength parameters for all experiments by each index are reported in Tabs. 4.11,4.12,4.13,4.14.

Since estimates and true weights are measured with different scales, as shown in a representative realization in Fig.4.8, estimates are analysed in terms of their correlation with true value, so as to assess the ability of each index to reproduce strength proportionality within networks.

		DATASET 1 DATASET 2 DATASET 3 $A = B = C = 1$ $A = B = C = 1$ $A = B = C = 1$							
Links									
$ROI1 \rightarrow ROI2$	0.03		0.07		0.08				
$ROI1 \rightarrow ROI3$	0.02		0.05		0.06				
$ROI2 \rightarrow ROI3$	0.03		0.07		0.09				
Links									
DATASET 1		DATASET 2		DATASET 3					
$x2$	$x3$	$x4$	$x2$	$x3$	$x4$				
$ROI1 \rightarrow ROI2$	0.10	0.15	0.24	0.16	0.27	0.36	0.20	0.31	0.40
$ROI1 \rightarrow ROI3$	0.01	0.01	0.01	0.03	0.03	0.02	0.05	0.04	0.03
$ROI2 \rightarrow ROI3$	0.03	0.03	0.03	0.03	0.04	0.04	0.10	0.11	0.12
Links									
DATASET 1		DATASET 2		DATASET 3					
$x2$	$x3$	$x4$	$x2$	$x3$	$x4$				
$ROI1 \rightarrow ROI2$	0.03	0.02	0.02	0.05	0.04	0.04	0.07	0.05	0.04
$ROI1 \rightarrow ROI3$	0.05	0.11	0.16	0.12	0.20	0.27	0.17	0.26	0.33
$ROI2 \rightarrow ROI3$	0.02	0.02	0.02	0.06	0.06	0.06	0.08	0.09	0.07
Links									
DATASET 1		DATASET 2		DATASET 3					
$x2$	$x3$	$x4$	$x2$	$x3$	$x4$				
$ROI1 \rightarrow ROI2$	0.03	0.028	0.031	0.063	0.058	0.062	0.076	0.074	0.071
$ROI1 \rightarrow ROI3$	0.015	0.015	0.015	0.045	0.038	0.044	0.062	0.06	0.059
$ROI2 \rightarrow ROI3$	0.081	0.144	0.217	0.174	0.281	0.372	0.22	0.344	0.416

Table 4.11.: GC index: average strength estimates in each experiment. Bold arrow in network model indicates the link with true strength multiplied by factors 2, 3 and 4.


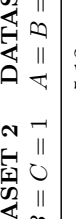

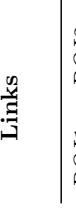
Links	DATASET 1		DATASET 2		DATASET 3					
	$A = B = C = 1$		$A = B = C = 1$		$A = B = C = 1$					
	$ROI1 \rightarrow ROI2$	2.45	4.45	5.12						
	$ROI1 \rightarrow ROI3$	3.18	6.28	8.16						
	$ROI2 \rightarrow ROI3$	2.93	6.62	8.17						
Links	DATASET 1		DATASET 2		DATASET 3					
	$x2$	$x3$	$x4$	$x2$	$x3$	$x4$				
	$ROI1 \rightarrow ROI2$	5.32	6.61	8.26	7.60	9.60	10.55	8.99	10.61	11.12
	$ROI1 \rightarrow ROI3$	3.44	3.68	4.47	7.08	7.88	8.66	9.66	10.00	9.98
	$ROI2 \rightarrow ROI3$	3.68	3.93	4.88	7.25	8.29	9.14	8.57	9.26	10.15
Links	DATASET 1		DATASET 2		DATASET 3					
	$x2$	$x3$	$x4$	$x2$	$x3$	$x4$				
	$ROI1 \rightarrow ROI2$	2.26	2.24	2.44	4.07	3.92	4.23	5.38	5.05	4.98
	$ROI1 \rightarrow ROI3$	4.82	7.70	9.53	9.52	11.94	13.87	11.75	14.19	15.98
	$ROI2 \rightarrow ROI3$	2.41	1.96	1.79	5.57	4.27	3.63	6.61	5.64	4.21
Links	DATASET 1		DATASET 2		DATASET 3					
	$x2$	$x3$	$x4$	$x2$	$x3$	$x4$				
	$ROI1 \rightarrow ROI2$	2.62	2.47	2.46	4.42	3.84	4.40	4.84	4.78	4.95
	$ROI1 \rightarrow ROI3$	2.74	2.53	2.43	5.76	4.70	4.97	7.30	6.47	6.01
	$ROI2 \rightarrow ROI3$	6.41	9.22	11.09	11.22	14.07	14.90	12.91	15.07	16.01

Table 4.12.: DTF index: strength estimates computed via AUC of the mean function in the frequency domain for each experiment. Bold arrow in network model indicates the link with true strength multiplied by factors 2, 3 and 4.

	Links	DATASET 1	DATASET 2	DATASET 3
		$A = B = C = 1$	$A = B = C = 1$	$A = B = C = 1$
	$ROI1 \rightarrow ROI2$	2.18	3.24	3.19
	$ROI1 \rightarrow ROI3$	2.64	5.05	6.94
	$ROI2 \rightarrow ROI3$	3.14	7.78	9.97
Links	DATASET 1	DATASET 2	DATASET 3	
	$x2$	$x4$	$x2$	$x3$
$ROI1 \rightarrow ROI2$	4.86	6.18	7.85	6.22
$ROI1 \rightarrow ROI3$	2.08	1.62	1.50	3.86
$ROI2 \rightarrow ROI3$	3.94	4.30	5.63	8.92
			10.75	12.40
			11.44	13.24
			14.40	14.80
Links	DATASET 1	DATASET 2	DATASET 3	
	$x2$	$x3$	$x4$	$x2$
$ROI1 \rightarrow ROI2$	1.79	1.44	1.43	2.00
$ROI1 \rightarrow ROI3$	4.40	7.56	9.56	9.83
$ROI2 \rightarrow ROI3$	2.77	2.42	2.29	7.07
			6.08	5.46
			9.08	8.42
			6.77	
Links	DATASET 1	DATASET 2	DATASET 3	
	$x2$	$x3$	$x4$	$x2$
$ROI1 \rightarrow ROI2$	2.43	2.25	2.39	3.32
$ROI1 \rightarrow ROI3$	2.00	1.75	1.35	4.35
$ROI2 \rightarrow ROI3$	6.99	9.99	12.17	13.26
			16.16	17.48
			15.96	18.46
			19.30	
			3.17	3.44
			3.74	3.76
			4.74	3.76
			18.46	19.30

Table 4.13.: PDC index: strength estimates computed via AUC of the mean function in the frequency domain for each experiment. Bold arrow in network model indicates the link with true strength multiplied by factors 2, 3 and 4.

Links	DATASET 1	DATASET 2	DATASET 3
	$A = B = C = 1$		
$ROI1 \rightarrow ROI2$	0.16	0.21	0.23
$ROI1 \rightarrow ROI3$	0.10	0.15	0.18
$ROI2 \rightarrow ROI3$	0.10	0.08	0.07

Links	DATASET 1	DATASET 2	DATASET 3
	$x2 \quad x3 \quad x4$		
$ROI1 \rightarrow ROI2$	0.29	0.36	0.42
$ROI1 \rightarrow ROI3$	0.09	0.07	0.05
$ROI2 \rightarrow ROI3$	0.14	0.18	0.23

Links	DATASET 1	DATASET 2	DATASET 3
	$x2 \quad x3 \quad x4$		
$ROI1 \rightarrow ROI2$	0.16	0.17	0.17
$ROI1 \rightarrow ROI3$	0.18	0.26	0.32
$ROI2 \rightarrow ROI3$	0.11	0.13	0.15

Links	DATASET 1	DATASET 2	DATASET 3
	$x2 \quad x3 \quad x4$		
$ROI1 \rightarrow ROI2$	0.16	0.16	0.16
$ROI1 \rightarrow ROI3$	0.09	0.09	0.08
$ROI2 \rightarrow ROI3$	0.15	0.19	0.24

Table 4.14.: SEM index: average strength estimates in each experiment. Bold arrow in network model indicates the link with true strength multiplied by factors 2, 3 and 4.

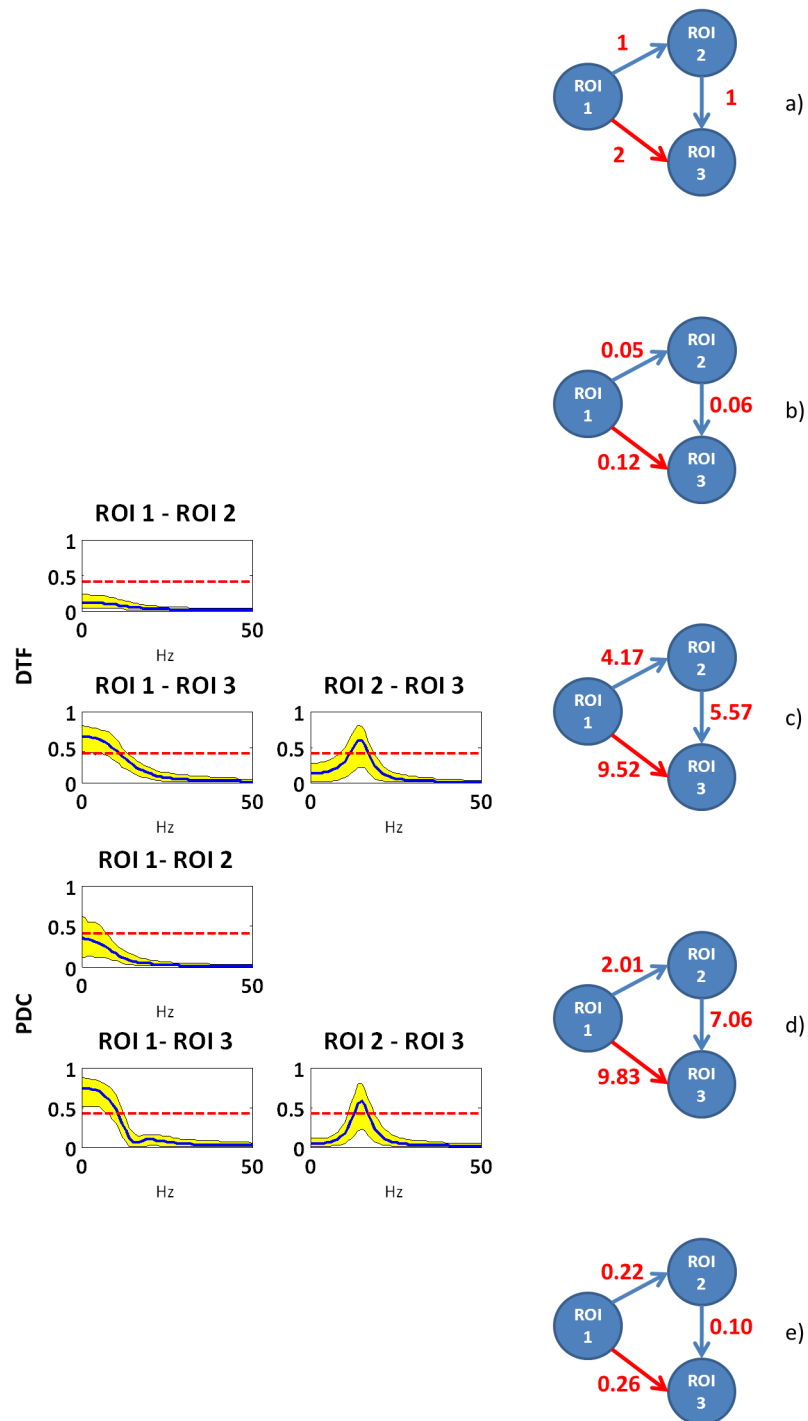


Figure 4.8.: Strength outputs in a representative realization for dataset with $r = 0.56$. True network model is depicted in a), while b),c) d) and e) report the averaged results obtained with GC, DTF, PDC and SEM methods. As regards c) and d), the panel in the right side shows DTF and PDC estimates as function of frequency; yellow bands represent 100 realizations, blue and dotted red lines are their mean functions and average thresholds, respectively. Network graph on the left side reports the corresponding AUC value computed from the mean function.

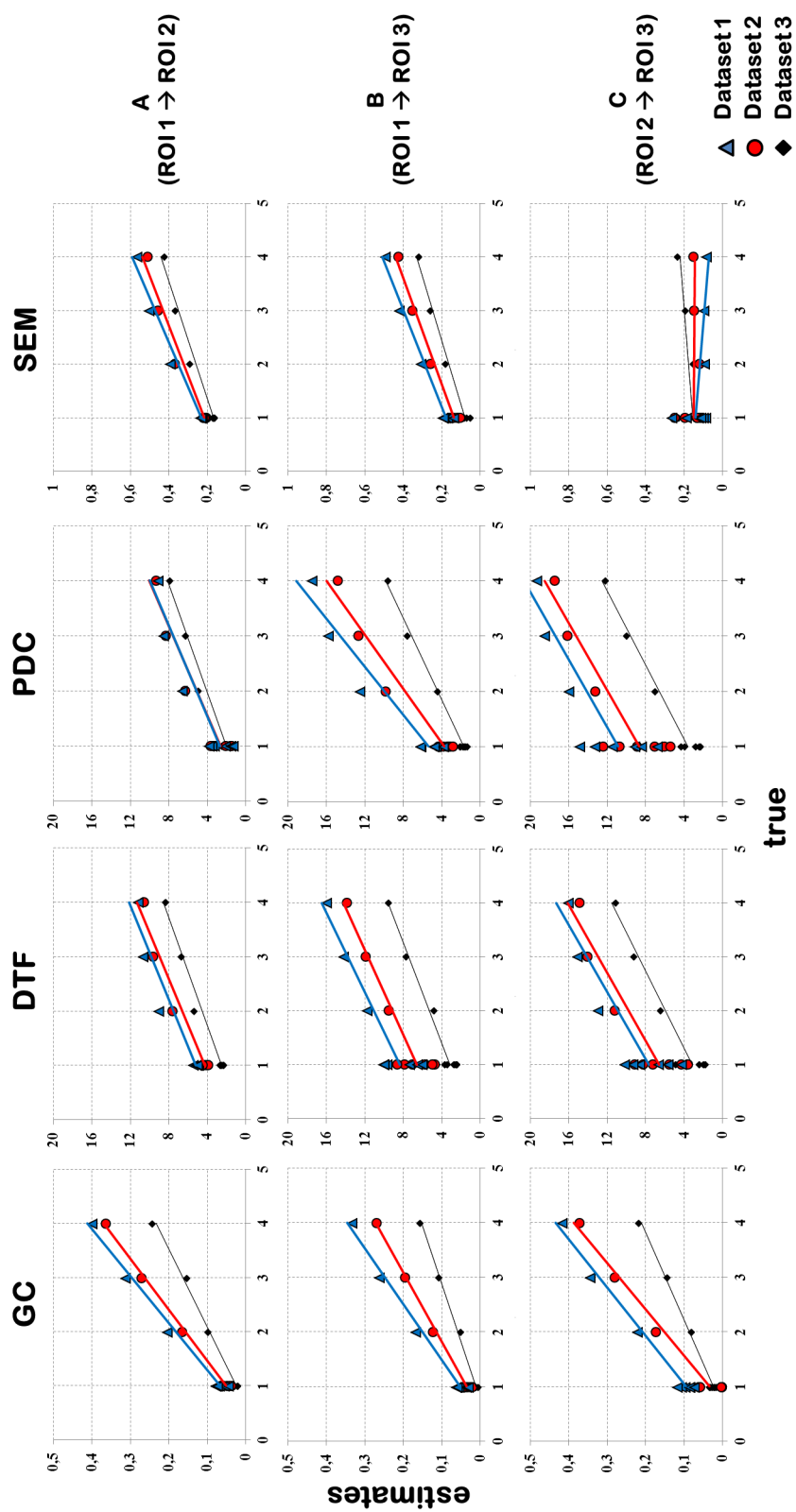


Figure 4.9.: Linear regression between GC, DTF, PDC and SEM estimates and true connections weights.

Assessment on in silico data: results

Fig. 4.2.1 displays scatter-plots showing regression between estimates and true weights, pooling together results over experimental conditions for all methods and datasets. Tab.4.15 reports the correlation coefficient R for each regression line. Correlation evidences high significant linear relationship with $R > 0.9$ ($P < 0.05$) for all GC estimates, while it slightly decreases for DTF and PDC, $R \sim 0.80$ and $R \sim 0.7$, in links connecting ROI 3. As regards SEM, estimates are well correlated with true weights in direct links, while it fails in estimating the link from ROI 2 to ROI 3.

Performances of regression lines are compared to study the interaction between strength estimates and network model sigmoid slope. This comparison shows that DTF and PDC estimates get worse with higher values of sigmoid slope (dataset 2 and 3), in particular in estimating strength of link connecting ROI 2 and ROI 3, while GC is highly correlated in all experiments.

Link from ROI 1 to ROI 2			
	DATASET 1	DATASET 2	DATASET 3
GC	0.99	0.99	0.98
DTF	0.97	0.95	0.92
PDC	0.95	0.85	0.84
SEM	0.98	0.98	0.97
Link from ROI 1 to ROI 3			
	DATASET 1	DATASET 2	DATASET 3
GC	0.99	0.99	0.98
DTF	0.92	0.82	0.80
PDC	0.99	0.95	0.93
SEM	0.97	0.97	0.97
Link from ROI 2 to ROI 3			
	DATASET 1	DATASET 2	DATASET 3
GC	0.99	0.98	0.98
DTF	0.90	0.77	0.75
PDC	0.90	0.72	0.66
SEM	< 0.5	< 0.5	< 0.5

Table 4.15.: Regression coefficients relative to scatter-plots in Fig.4.2.1 for each index and dataset.

4.2.2. Other network models

Results reported below show the outcomes produced by connectivity estimation for the other three network models described in 2.3. GC estimates of strength and percent of true and false positives are displayed in Tab. 4.16. DTF and PDC frequency functions are illustrated in Figs. 4.10, 4.11 and 4.12. Their AUC values and percent of true and false positives are summarized in Tabs. 4.17 and 4.18. Finally, SEM outputs are show in Tab. 4.19.

Topology and strength estimates confirm results obtained in the previous

Network	Links	Estimates	
	$ROI1 \rightarrow ROI2$	$A = B = 1$ 0.06 (92%)	$A = 1 \quad B = 2$ 0.06 (91%)
	$ROI2 \rightarrow ROI3$	0.05 (89%)	0.16 (100%)
	$ROI1 \rightarrow ROI3$	< 0.001 (1%)	< 0.001 (0%)
	$ROI2 \rightarrow ROI1$	< 0.001 (2%)	< 0.001 (0%)
	$ROI3 \rightarrow ROI1$	< 0.001 (0%)	< 0.001 (0%)
	$ROI3 \rightarrow ROI2$	< 0.001 (0%)	< 0.001 (0%)
	$ROI1 \rightarrow ROI2$	$A = B = C = 1$ 0.08 (98%)	$A = C = 1 \quad B = 2$ 0.11 (100%)
	$ROI2 \rightarrow ROI1$	0.11 (100%)	0.28 (100%)
	$ROI2 \rightarrow ROI3$	0.05 (85%)	0.03 (71%)
	$ROI1 \rightarrow ROI3$	< 0.001 (1%)	< 0.001 (3%)
	$ROI3 \rightarrow ROI1$	< 0.001 (2%)	< 0.001 (1%)
	$ROI3 \rightarrow ROI2$	< 0.001 (2%)	< 0.001 (1%)
	$ROI1 \rightarrow ROI2$	$A = B = C = 1$ 0.06 (95%)	$A = C = 1 \quad B = 2$ 0.06 (95%)
	$ROI2 \rightarrow ROI3$	0.09 (86%)	0.17 (100%)
	$ROI3 \rightarrow ROI1$	0.05 (100%)	0.08 (97%)
	$ROI1 \rightarrow ROI3$	< 0.001 (0%)	< 0.001 (0%)
	$ROI3 \rightarrow ROI2$	< 0.001 (1%)	< 0.001 (0%)
	$ROI2 \rightarrow ROI1$	< 0.001 (2%)	< 0.001 (0%)

Table 4.16.: GC index: average strength estimates and percent of true and false positives in each experiment. Bold arrow in network graph indicates the link with doubled strength in the second experiment.

section. GC recognizes very well network topology and demonstrates its ability in distinguishing direct and indirect connections, reciprocal links and cycled relationships. Considering the open loop net in Tab. 4.16, upper panel, characterized by two direct coupling ($ROI1 \rightarrow ROI2$; $ROI2 \rightarrow ROI3$) and an indirect link ($ROI1 \rightarrow ROI3$), we can observe a causality increase only between the signals

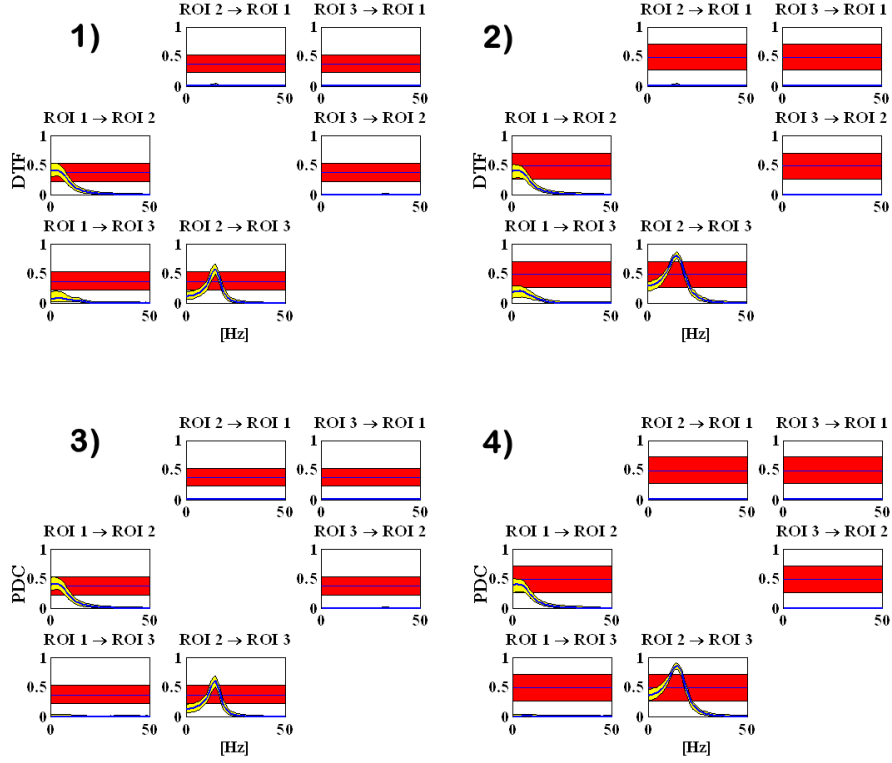


Figure 4.10.: DTF and PDC index as function of frequency (open-loop network). Panels in the left show to the basal condition, ($A=B=1$), while the other two are relative to the experiment with $A=1$ and $B=2$. Yellow and red bands represent 100 realizations and their threshold respectively, while blue and dotted red lines are their mean functions.

involved in the connection with weight B . In presence of reciprocal links, as in the second network graph, Tab. 4.16 middle panel, characterized by connection between ROI 1 and ROI 2 in both ways, direct link between ROI 2 and ROI 3 and the indirect connection from ROI 1 to ROI 3, GC index, not only correctly describes the net topology in more than 70% of times, but also gives a quantitative measurement of the internal connection intensity. Indeed it is possible to derive numerically the causality increase recorded in the link $\text{ROI } 2 \rightarrow \text{ROI } 1$ when the corresponding true weight B of the connection itself is doubled. Observing results obtained with the cycle network, Tab. 4.16 lower panel, the B weight increase influences the strength only in the connection from ROI 2 to ROI 3. Other significant links remain almost unvaried, despite the cyclic topology of

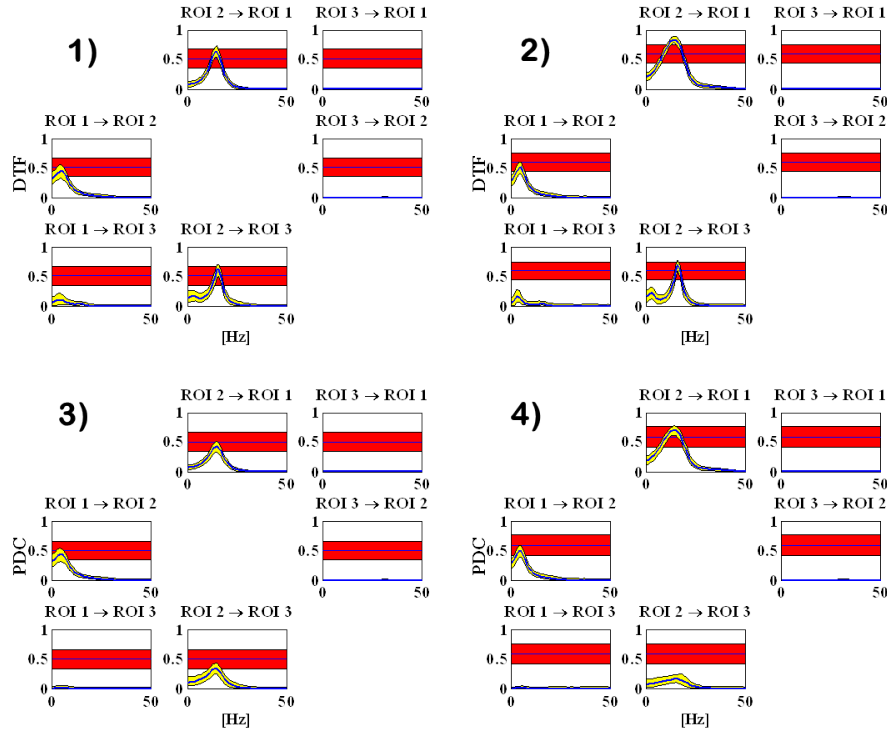


Figure 4.11.: DTF and PDC index as function of frequency (network with feed-back link). Panels in the left show to the basal condition, ($A=B=C=1$), while the other two are relative to the experiment with $A=C=1$ and $B=2$. Yellow and red bands represent 100 realizations and their threshold respectively, while blue and dotted red lines are their mean functions.

the network.

Information obtained from DTF and PDC allow to observe the causality flow in frequency domain, Fig. 4.10 - 4.11 - 4.12. Analyzing the results computed via AUC, Tab. 4.17, we can notice that when connection weight increase, DTF value reflects an increase in signal connectivity related to that connection both in direct and indirect way. This limit, already described in previous section, is noticeably identified in Fig. 4.10, panels (1)-(2). Even if not recognized as significant, DTF index shows the influence of signal from ROI 1 to ROI 3 due to the presence of an indirect connection. Connectivity analysis using PDC instead allows to distinguish direct connections from indirect ones. Indeed, with reference to the same network, Fig. 4.10, panels (3)-(4), the contribution due to indirect connection from ROI 1 to ROI 3 results heavily weakened, Tab. 4.10, upper panel.

Assessment on in silico data: results

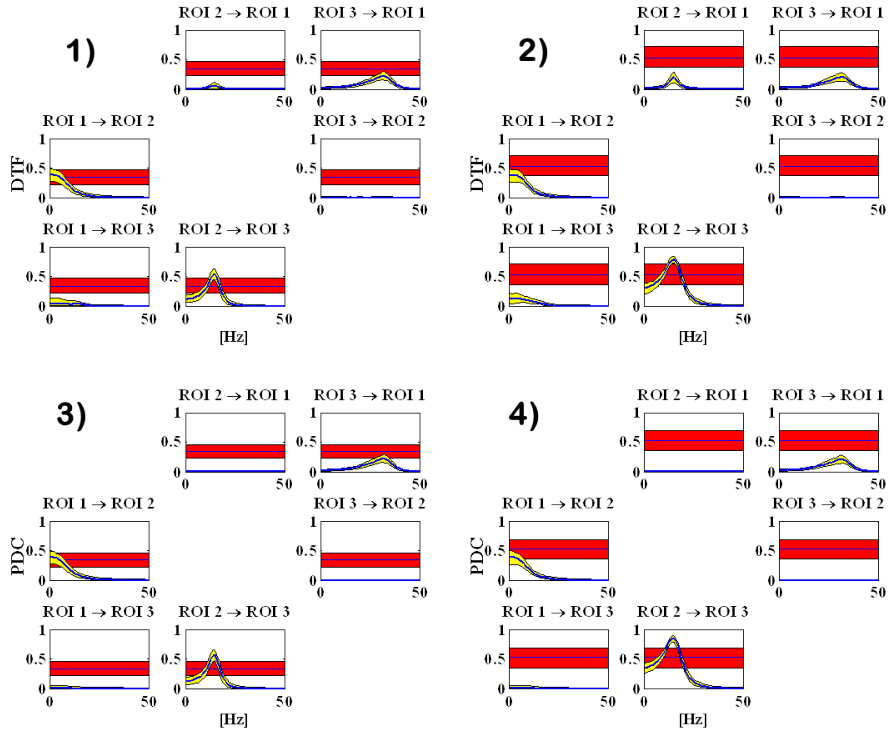


Figure 4.12.: DTF and PDC index as function of frequency (cycle network). Panels in the left show to the basal condition, ($A=B=C=1$), while the other two are relative to the experiment with $A=C=1$ and $B=2$. Yellow and red bands represent 100 realizations and their threshold respectively, while blue and dotted red lines are their mean functions.

As regard SEM, its results evidence the difficulty in estimating connection series. Considering the scores of the first network in Tab.4.19, upper panel, SEM is able to recognize strength increasing from ROI 2 to ROI 3 between the two experiment, but in the basal condition, where weights are equal, it is not in agreement with true network. Performances get worse in the cycle network estimation, where SEM confuses the direction of one link, and, finally, fall down in estimating networks with reciprocal connection.

Network	Links	Estimates	
		$A = B = 1$	$A = 1 \quad B = 2$
	$ROI1 \rightarrow ROI2$	5.01 (71%)	5.02 (41%)
	$ROI2 \rightarrow ROI3$	6.48 (93%)	12.48 (96%)
	$ROI1 \rightarrow ROI3$	1.60 (2%)	2.88 (6%)
	$ROI2 \rightarrow ROI1$	0.29 (0%)	0.28 (0%)
	$ROI3 \rightarrow ROI1$	0.13 (0%)	0.14 (0%)
	$ROI3 \rightarrow ROI2$	0.16 (0%)	0.15 (0%)
		$A = B = C = 1$	$A = C = 1 \quad B = 2$
	$ROI1 \rightarrow ROI2$	1.04 (81%)	1.02 (80%)
	$ROI2 \rightarrow ROI1$	6.84 (42%)	13.44 (16%)
	$ROI2 \rightarrow ROI3$	6.44 (81%)	6.74 (47%)
	$ROI1 \rightarrow ROI3$	1.72 (0%)	1.97 (0%)
	$ROI3 \rightarrow ROI1$	0.21 (0%)	0.23 (0%)
	$ROI3 \rightarrow ROI2$	0.22 (0%)	0.27 (0%)
		$A = B = C = 1$	$A = C = 1 \quad B = 2$
	$ROI1 \rightarrow ROI2$	4.74 (73%)	4.80 (23%)
	$ROI2 \rightarrow ROI3$	6.23 (94%)	12.56 (92%)
	$ROI3 \rightarrow ROI1$	4.07 (17%)	3.94 (0%)
	$ROI1 \rightarrow ROI3$	0.61 (0%)	1.80 (0%)
	$ROI3 \rightarrow ROI2$	0.33 (0%)	0.31 (0%)
	$ROI2 \rightarrow ROI1$	1.21 (0%)	2.26 (0%)

Table 4.17.: DTF index: strength estimates computed via AUC of the mean function in the frequency domain and percent of true and false positives in each experiment. Bold arrow in network graph indicates the link with doubled strength in the second experiment.

4.3. Summary

So far, a gold standard for connectivity studies is not already emerged in literature, therefore, analysis of *in silico* results is needed to understand limits and advantages of each index. Our simulation clearly shows that GC can be used as a stand-alone index to explore significant connections underlying data and to give insights on the strength up to scaling factor. As regard DTF and PDC, their results are less accurate in terms of topology and strength estimates via AUC integral, but are necessary to obtain information in frequency domain, even if their combined use is preferred to have a better characterization of the coupling. These results have been transferred into a practical procedure to analyse EEG data, following three steps summarized in flowchart in Fig. 4.3. GC index is used to both to explore network topology, assuming the existence of connection when its value is statistically significant, and to infer connection intensities. Finally,

Assessment on in silico data: results

Network	Links	Estimates	
	$ROI1 \rightarrow ROI2$	$A = B = 1$ 4.97 (69%)	$A = 1 \quad B = 2$ 4.96 (41%)
	$ROI2 \rightarrow ROI3$	6.79 (97%)	13.69 (97%)
	$ROI1 \rightarrow ROI3$	0.53 (0%)	0.66 (0%)
	$ROI2 \rightarrow ROI1$	0.25 (0%)	0.20 (0%)
	$ROI3 \rightarrow ROI1$	0.13 (0%)	0.14 (0%)
	$ROI3 \rightarrow ROI2$	0.17 (0%)	0.15 (0%)
			$A = B = C = 1$
	$ROI1 \rightarrow ROI2$	5.15 (44%)	5.05 (16%)
	$ROI2 \rightarrow ROI1$	5.28 (36%)	11.84 (60%)
	$ROI2 \rightarrow ROI3$	5.11 (12%)	3.60 (0%)
	$ROI1 \rightarrow ROI3$	0.81 (0%)	0.93 (0%)
	$ROI3 \rightarrow ROI1$	0.21 (0%)	0.23 (0%)
	$ROI3 \rightarrow ROI2$	0.22 (0%)	0.27 (0%)
			$A = B = C = 1$
	$ROI1 \rightarrow ROI2$	4.70 (71%)	4.74 (25%)
	$ROI2 \rightarrow ROI3$	6.45 (98%)	13.51 (97%)
	$ROI3 \rightarrow ROI1$	4.10 (19%)	4.04 (1%)
	$ROI1 \rightarrow ROI3$	0.18 (0%)	0.18 (0%)
	$ROI3 \rightarrow ROI2$	0.14 (0%)	0.13 (0%)
	$ROI2 \rightarrow ROI1$	0.60 (0%)	0.65 (0%)
			$A = B = C = 1$

Table 4.18.: PDC index: strength estimates computed via AUC of the mean function in the frequency domain and percent of true and false positives in each experiment. Bold arrow in network graph indicates the link with doubled strength in the second experiment.

Network	Links	Estimates	
	$ROI1 \rightarrow ROI2$	$A = B = 1$ 0.21	$A = 1 \quad B = 2$ 0.22
	$ROI2 \rightarrow ROI3$	0.05	0.10
			$A = B = C = 1$
	$ROI1 \rightarrow ROI2$	—	—
	$ROI2 \rightarrow ROI1$	—	—
	$ROI2 \rightarrow ROI3$	—	—
			$A = B = C = 1$
	$ROI1 \rightarrow ROI2$	$A = B = C = 1$ 0.18	$A = C = 1 \quad B = 2$ 0.15
	$ROI2 \rightarrow ROI3$	0.06	0.11
	$ROI3 \rightarrow ROI1$	-0.05	-0.075
			$A = B = C = 1$

Table 4.19.: SEM index: average strength in each experiment. Bold arrow in network graph indicates the link with doubled strength in the second experiment.

in order to study the spectral coupling, frequency patterns of DTF and PDC together are evaluated.

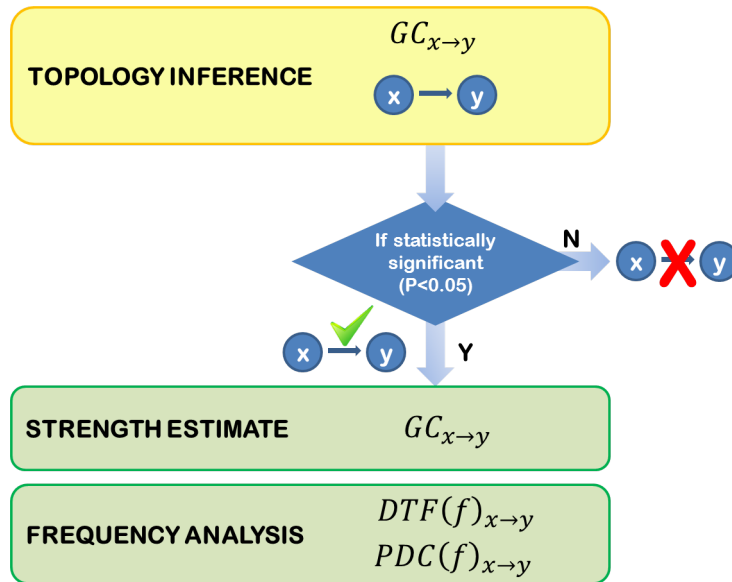


Figure 4.13.: Flowchart procedure to analyse EEG data connectivity.

5. Application to EEG data: assessment of hepatic encephalopathy

Thou art giving and forgiving, ever blessing, ever blessed, Wellspring of the joy of living, ocean depth of happy rest! Thou our Father, Christ our Brother, all who live in love are Thine, Teach us how to love each other, lift us to the joy divine.

Hepatic encephalopathy (HE) is a worsening of brain function caused by liver failure. When liver is affected by some disorders, such as cirrhosis or hepatitis, it is no longer able to change toxic substances that are either made by the body or taken into the body. Hence, it releases intestinal neurotoxic substances in the bloodstream reaching the systemic circulation and the brain. The toxic substances are not entirely known, but certainly a central role is determined by ammonia, which is a product of the digestion of proteins, [46].

HE clinical manifestations range from a slightly altered mental state to coma; it may be either acute or chronic. The condition has been graded into five stages based on deteriorations in mental and psychometric function. The identification of patients with mild or subclinical encephalopathy is a problem in hepatology but quantitative analysis of EEG provides objective measures of HE, such as frequency analysis of the wake EEG, [47], [48]. Patients with liver disease do not show any obvious EEG abnormalities until impairment of liver function is sufficient to give rise to episodes of confusion. In the early phase of HE, the EEG shows progressive slowing and disorganization of the alpha rhythm which becomes

Application to EEG data: assessment of hepatic encephalopathy

mixed with random diffuse theta activity, often most marked in the temporal regions. As clouding of consciousness becomes more evident the theta components become dominant and the dominant frequency falls. Further progression of liver failure is associated with diminished amplitude of this delta activity which becomes arrhythmic and finally disappears so that prior to death, the EEG shows a generalized absence of activity of any kind, [46].

A reproducible human model to assess the risk of HE development in patients with cirrhosis is made with the induction of hyperammonaemia by the oral administration of a mixture of amino acids, called amino acid challenge (AAC). Studies in literature already investigated the clinical, psychometric and wake/sleep EEG correlation with induced hyperammonaemia by analyzing EEG frequency contents and its topographic distribution, [49]. They demonstrated that the wake EEG in posterior brain areas is sensitive to hyperammonaemia and, hence, it can be used to help the HE neurophysiological definition [50].

To obtain a more detailed picture about the effect of HE on brain circuits, this study aims to explore EEG of cirrhotic patients in terms of connectivity. So far, the altered brain state activity in patient with cirrhosis and different level of HE has been analyzed during resting-state functional MRI studies [51], hence, this study is a completely novel application. We focused our attention on a subset of four electrodes each hemisphere, namely $F3$, $P3$, $C3$, $O1$ and $F4$, $P4$, $C4$, $O2$ to preliminary explore a plausible connectivity network from the occipital to the frontal brain areas. In particular, we analyzed the effect of hyperammonaemia on the wake EEG both of patients with cirrhosis and healthy volunteers, in order to understand to what extent this condition affects the EEG rhythms distribution and how it modifies the electrophysiological information flux.

5.1. Study design

5.1.1. Dataset

Two groups were studied: 10 patients with cirrhosis (9 men; 1 woman; age 54 ± 14 years; mean \pm SD) and ten healthy volunteers (age 49 ± 13 years), served as reference.

Ten min of relaxed EEG was recorded in basal condition and after the administration of AAC, according to the International 10 – 20 system, sampled at 256 Hz and band-pass filtered in the range 0.33 – 70 Hz. Then, derivations were referenced to the average of all derivations. EEG was visually inspected to exclude focal activity and any muscular artefacts, hence 40–seconds sections were selected for analysis.

5.1.2. Spectral analysis

Spectral analysis was performed by using the parametric approach based on the univariate autoregressive model identification by least squares, i.e.:

$$x(n) = - \sum_{i=1}^p a_i x(n-i) + e(n), \quad n = 1, \dots, M-1 \quad (5.1)$$

where p is the model order, a_i are model parameters and $e(n)$ is a white Gaussian noise having zero mean and variance equal to σ^2 . The best model order is chosen by applying the Akaike's criterion. From AR model of eq.5.1 the power spectral density is described as follows:

$$S(f) = \sigma^2 |H(f)| \quad (5.2)$$

where

$$H(f) = \frac{1}{1 + \sum_{i=1}^p a_i e^{j2\pi i \frac{f}{Fs}}} \quad (5.3)$$

Spectral estimates are obtained for electrodes F4, P4, C4, O4 and F3, P3, C3, O3 analysed in the frequency range 1.5 – 26.5 Hz.

5.1.3. EEG connectivity analysis

As suggested by the simulation study, to estimate connectivity relationships among the selected electrodes we followed flowchart in Fig. 4.3, but since signals have been analysed in the narrow range 6 – 13 Hz, we considered GC estimation only. To do that a MVAR model was identified for each subject, choosing the optimal model order via Akaike's criterion in the range 1 – 10, and then, for each group an average network was obtained, including the connection significant in more than 65% of the subjects, revealed by the index. Hence, scalar values of GC in significant links have been considered to have insights on coupling intensity. As for simulation study, estimates give an indirect measure of the connection intensity, hence we can infer on the strength only up to a scalar factor.

5.1.4. Statistical analysis

Comparisons within and between population were performed using non parametric tests. Wilcoxon test for paired data was used to compare results between conditions before and after AAC within group. Mann-Whitney test was used to determine the significance of differences between volunteers and patients. $P < 0.05$ was considered to be statistically significant.

5.2. Results

5.2.1. Spectral analysis

Optimal MVAR model orders were chosen via Akaike's criterion in the range 1 – 15. Fig. 5.1 shows signal of each ROI in left hemisphere in time and frequency domain for a representative healthy volunteer before and after AAC, respectively, while Fig. 5.2 reports the same conditions for a representative patient.

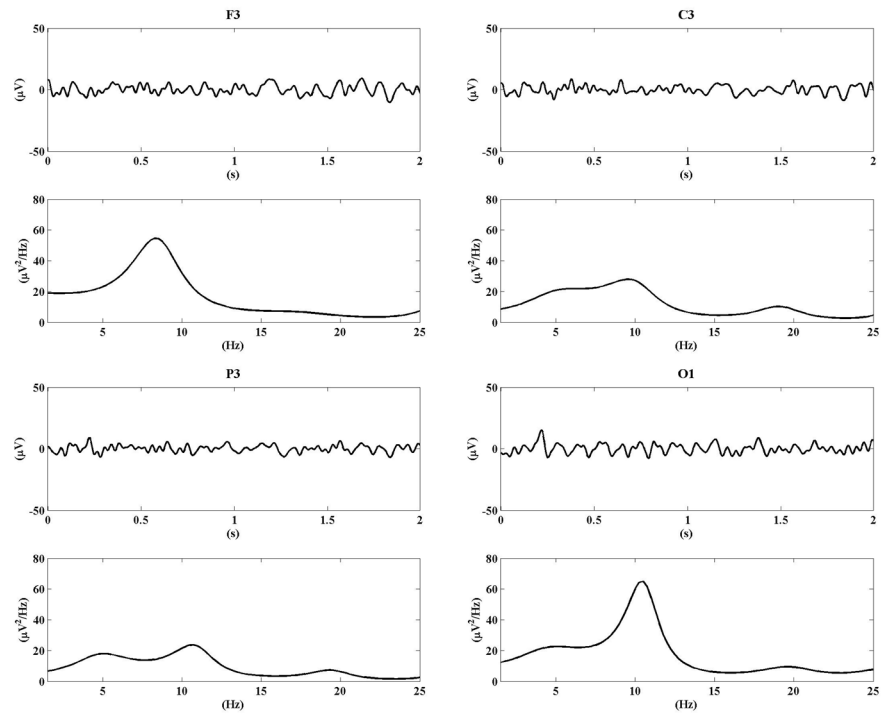
Inspection of the power spectrum revealed the prevalence of the dominant EEG activity in the occipital area in both population, but magnitude of the spectra was higher in patients in compared to healthy volunteers. Administration of AAC induced an increase in power over the most of the scalp in healthy volunteers, while less evident changes were observed in patients.

5.2.2. EEG connectivity estimation

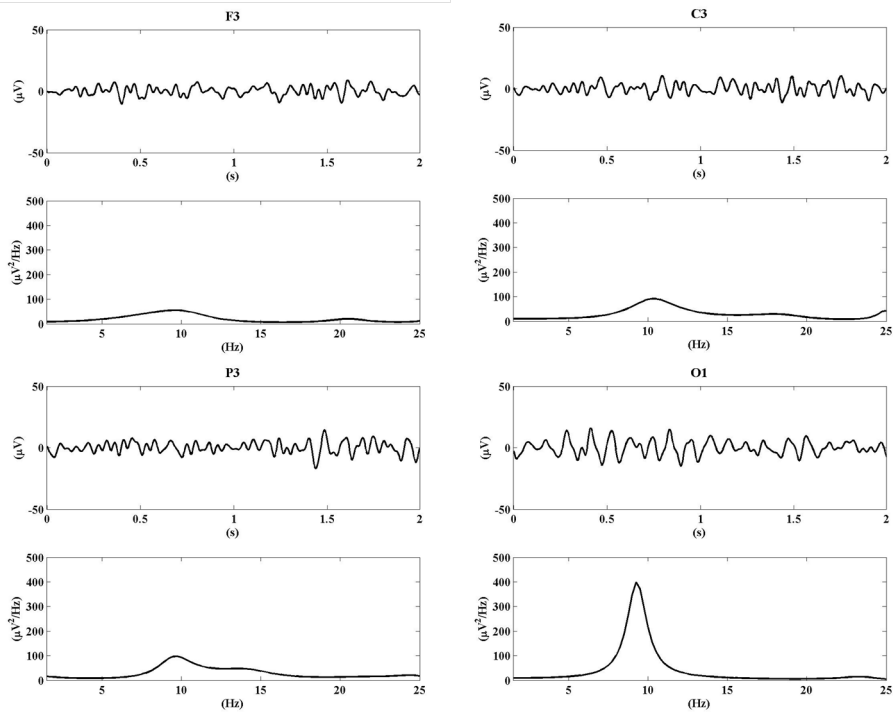
Figs. 5.3, 5.4 and 5.5, 5.6 depict the average network for both populations and hemisphere, before and after AAC. Both in human volunteers and patients, topology presents different patterns between hemispheres, showing a denser network especially in the right hemisphere after AAC. Intensity of each connection is estimated by exploited GC numerical values and Fig. ?? reports the differences between the two conditions for both groups and hemispheres.

Differences between hemisphere both in topology and strength within groups evidence the lateral asymmetry of the brain, showing a deeper influence of AAC in right connectivity network. Comparing human volunteers with patients displays that the latter group does not change significantly after AAC, while the first one exhibits a re-organization of the network connectivity, giving more support to the idea that patients stand constantly in hyperammonaemia condition. This is well evident examining Fig. 5.2.2 which show the global effect of AAC in increasing connection intensities.

Application to EEG data: assessment of hepatic encephalopathy

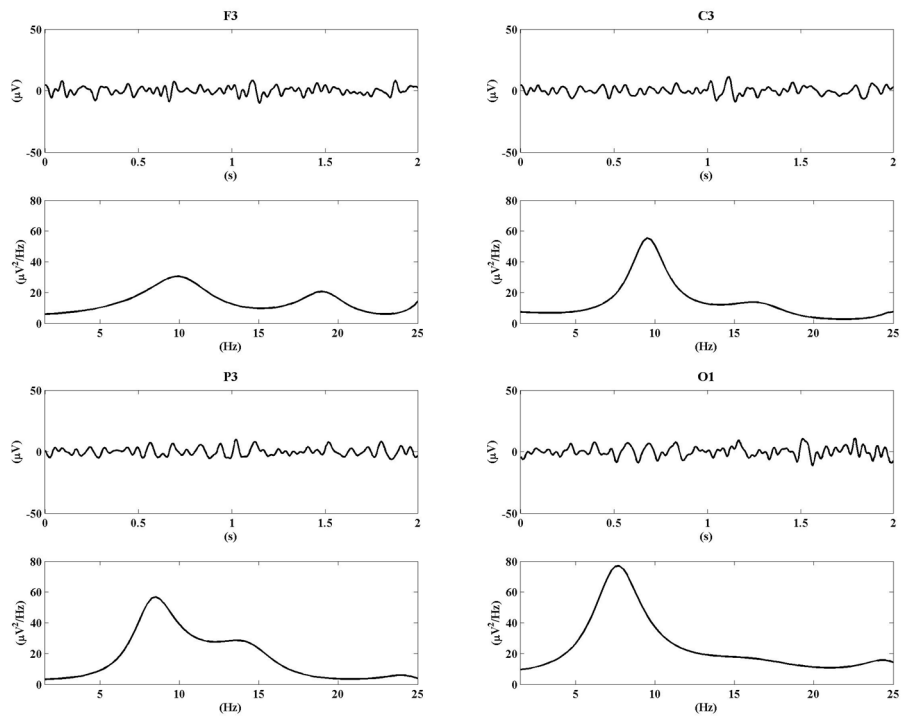


(a) Basal condition, before AAC.

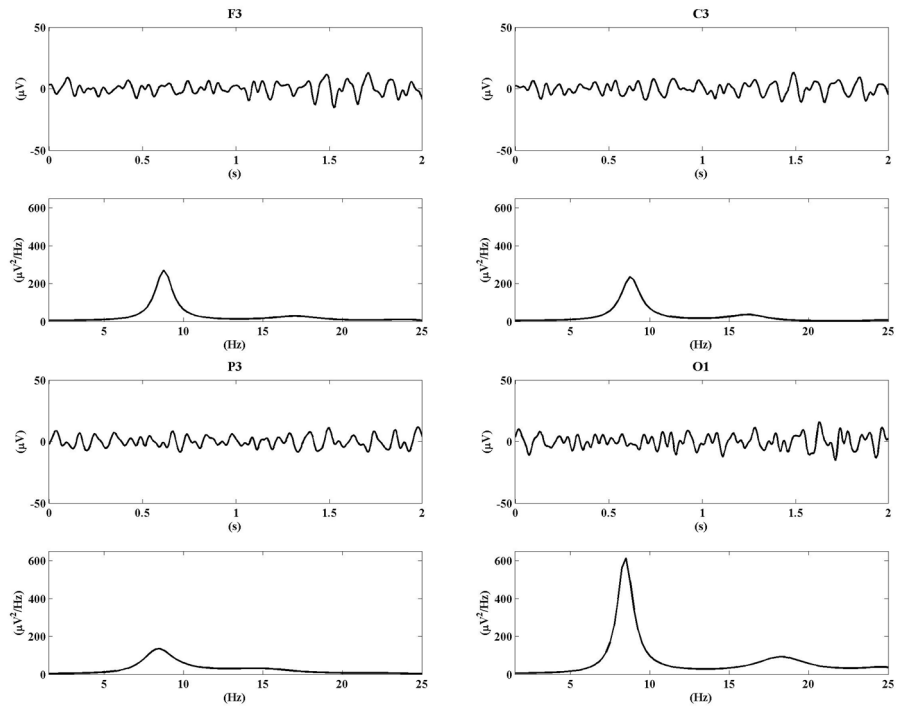


(b) Condition after AAC.

Figure 5.1.: ROIs signals of the left hemisphere in time and frequency domain for a representative healthy volunteer.



(a) Basal condition, before AAC.



(b) Condition after AAC.

Figure 5.2.: ROIs signals of the left hemisphere in time and frequency domain for a representative patient.

Application to EEG data: assessment of hepatic encephalopathy

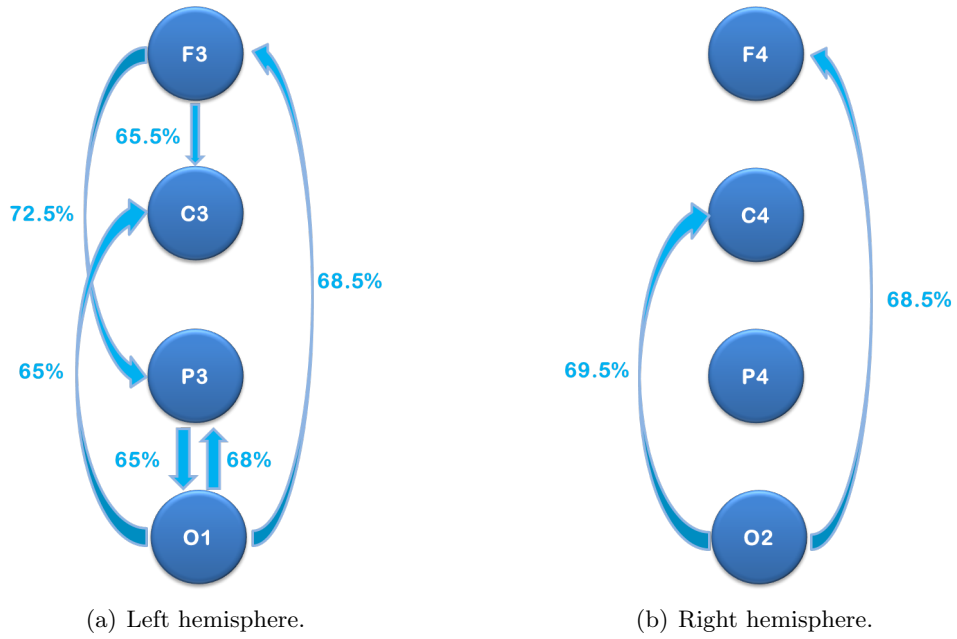


Figure 5.3.: Estimated topology overall human volunteers in both hemisphere in basal condition (before AAC).

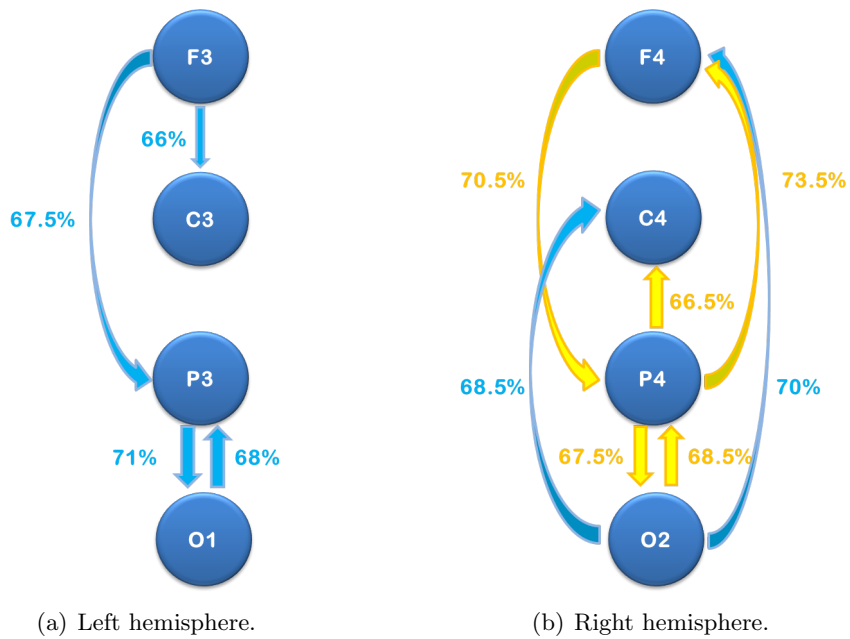


Figure 5.4.: Estimated topology overall human volunteers in both hemisphere after AAC.

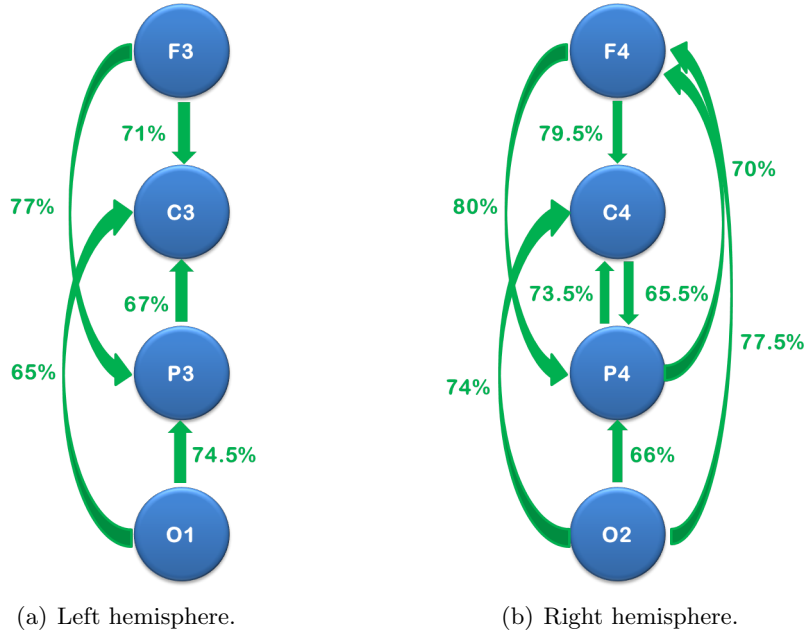


Figure 5.5.: Estimated topology overall patients in both hemisphere in basal condition (before AAC).

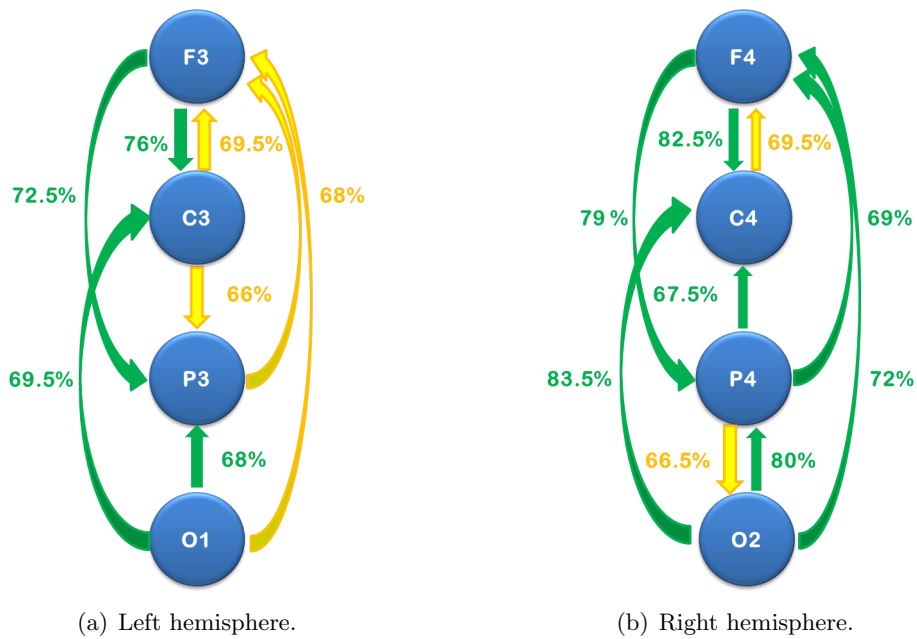


Figure 5.6.: Estimated topology overall human volunteers in both hemisphere after AAC.

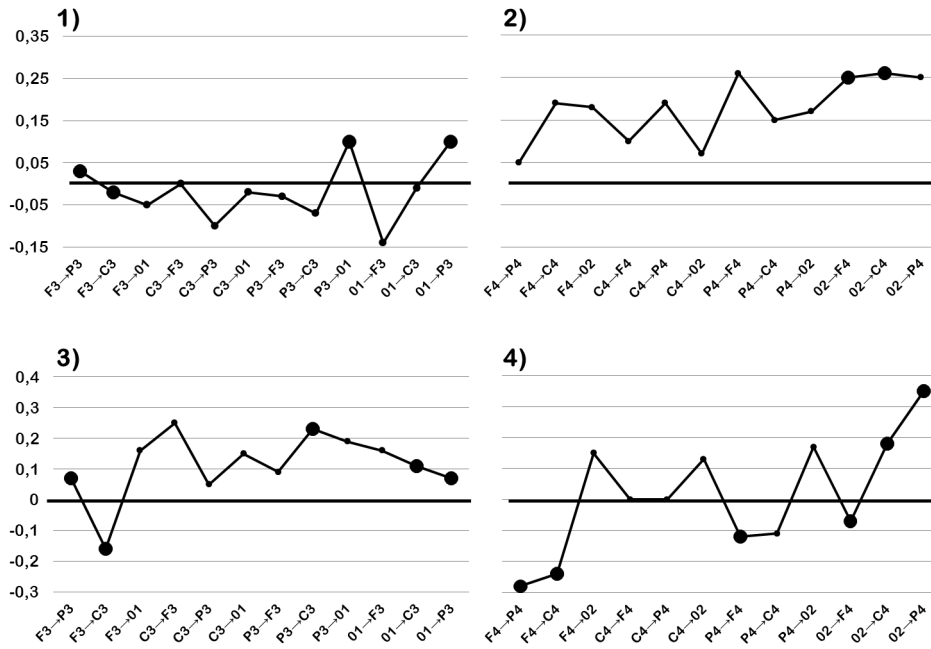


Figure 5.7.: GC index: difference of the average GC estimates between the two conditions (after AAC - before AAC) for both hemispheres and groups (*upper panels*: human volunteers; *lower panels*: patients). Black circles represent the significant links in both conditions; bold line stands for zero.

This findings improve those obtained in previous work in which effects of AAC were observed in changes in the power of the dominant EEG rhythms, only. Application of connectivity measures provides more detailed in explaining the behaviour of the scalp EEG consequent to hyperammonaemia, since they are able to describe the coupling between signals and hence they furnish a network map to decode the power topographic distribution of EEG.

6. Discussion

Mortals, join the happy chorus, which the morning stars began, Father love is reigning o'er us, brother love binds man to man. Ever singing, march we onward, victors in the midst of strife, Joyful music leads us Sunward in the triumph song of life!

6.1. Simulation study

The purpose of this study was to assess the ability of commonly used measures of brain connectivity using *in silico* data. Our strategy was to reproduce plausible neurophysiological processes in which we could manipulate coupling among ROIs, each of them simulated by a neural mass model generating real power spectra very similar to the empirical ones.

We tested two connectivity methods, MVAR indexes and SEM, both based on linear regression equations, but different in describing data dynamics. MVAR models consider past data information, whilst SEM describes variables interaction using present instants, only. In connectivity estimate, the first approach was used to infer both topology and strength connections, while the second one provides estimates of strength, since it assumes topology to be known. The aim was to compare their performances in order to define a practical tool, useful in clinical application with real data. Hence, firstly, we wanted to establish their validity with a simulation study. Our approach uses *ad-hoc* network models, built taken into account the number of ROIs and links, the different non linearity degree due

Discussion

to the neural mass model equations and the connection weights imposed between each pair of ROIs.

Two interconnected ROIs is the simplest situation and at the most can be exploited to assess on the connectivity index ability to identify the influences direction. Differently, working with at least three ROIs allows to analyse cascade flows, reciprocal connections and cycle graphs. This issue is crucial in comparing the performances of DTF and PDC, which mainly differ in distinguishing direct from indirect connections.

Changes in sigmoid slope of the neural mass model in eq. 2.1 and in coupling strength (2.3) have had an important impact on data characterizations.

Changing sigmoid slope values is twofold. First, it allows to vary the non-linearity degree since the greater the slope is, the smaller the range of sigmoid linear approximation gets, approaching the climax over time faster. Second, since it affects the intrinsic gain of each neural group, increasing in power content modifies the amount of frequency coupling without changing weight connections. Applying *in silico* data obtained with a non-linear procedure to MVAR indexes and SEM, based on linear equations, furnishes a quantitative description to what extent they are able to describe a phenomenon typically non-linear.

Varying connections weights simulates different strength of coupling. This approach enable both to evaluate methods sensitivity to different level of strength and to understand which measures is more influenced by the mixing of interacting signals.

By modifying these parameters, *in silico* data emulate key features observed in empirical data, then estimate networks were computed. In addition, using surrogate data we determined the null distribution of each measure in order to evaluate their statistical significance.

6.2. Network connectivity estimation

As regards topology, as expected our analysis suggests the absence of causal influence where effectively no connection exists, i.e. $ROI2 \rightarrow ROI1$, $ROI3 \rightarrow ROI1$ and $ROI3 \rightarrow ROI2$, since the percentage of revealed connections by all methods is much lower than those of the three real direct connections, i.e. $ROI1 \rightarrow ROI2$, $ROI1 \rightarrow ROI2$ and $ROI2 \rightarrow ROI3$, fig. 4.7. No significant differences emerge among datasets, but within each experiment comparing the predominant link with those characterized by the same strength, Tabs. 4.7,4.8,4.9. All couple of ROIs are related to each other according to a feed forward scheme characterized by asymmetric information flow. The time domain signal mixing consequently sums up overlapping frequency contents, obscuring underlying connectivity patterns. Inevitably the MVAR model loses weaker connections and describes the prevalent ones, only. Analysing statistical performances, Tab. 4.10, all methods provide high values of specificity, meaning they clearly recognize where connection does not exist, as reported above. As regards sensitivity, GC reaches 84%, while DTF and PDC provide less powerful performances, demonstrating they are less robust to frequency overlapping due to signal mixing. Hence, GC can be considered a good network estimator. Moreover, results obtained with the additional network models, 4.2.2, show its ability to recognize not only coupling direction, but also to locate direct connection contributions in case of reciprocal and cycle links. Furthermore, results according to DTF and PDC, greatly vary depending on the statistical testing strategy used. Unlike GC, which statistical test is based on evaluation of the F-statistic and results depend on p-value only, there exist different way to assess DTF and PDC connectivity significance. Alongside well-known methods based on phase-randomization, adopted in this work, surrogate data can be generated shuffling time series samples or using multivariate ARMA. Other innovative approaches are proposed in literature for PDC analysis, such are based on asymptotic statistic [52] and anti-symmetrisation testing

Discussion

[53]. An *ad-hoc* analysis to understand the optimal strategy was performed comparing permutation methods, both time samples and phase, and those based on simulation with multivariate ARMA processes,[43]. Our experience evidenced that, unlike phase-randomisation, thresholds obtained with sample-shuffling and ARMA simulation provided smaller false negatives rate, but much more false positives. For this reason, in this work we preferred to be more cautious, adopt the first one surrogate strategy with the disadvantage of having smaller sensitivity values. Moreover, the average value of the threshold distribution that comes out is always equal to 0.5, a well-established threshold value in literature to determine the coherence significance, [54].

As regards strength, since true network and estimates are measured with different scales, linear regression analysis is used to understand if they are sensitive to strength modulations. Examining fig. 4.2.1, correlation evidences the ability of MVAR index to quite well reproduce connection intensity, while SEM reveals considerable difficulty in estimating connection between *ROI2* and *ROI3*. SEM weakness to describe just simple connectivity patterns is due to the over-simplified model underlying SEM equations. Unlike MVAR approach, including past information of each time series, SEM computes connection taking into account present information, only. This is an implausible assumption, causing troubles in removing confounding effects, in particular in case of reciprocal and cycle nets.

Both DTF and PDC have shown to be able to give information about network topology and connection intensity, even if results are less accurate than GC performances. DTF allows to obtain good results, with a lower chance of error for direct connections only. In case of indirect connections, DTF performances and results reliability heavily depend on the method used to choose the significance threshold. PDC, instead, allows to recognize direct from indirect connections but in some instances, in presence of more patterns connecting the same ROIs, the weaker connection is obscured, i.e. in Figs. ?? and Tab. ?? link from ROI 1 to ROI 3 is not considered significant.

Even if this study explored simple networks, some useful suggestions emerged. As regard topology, we can confirmed that GC can be used as a good estimator, supported by its high values of sensitivity and specificity. As regard strength, SEM is definitely too simple to describe complex interactions, such as neurophysiological processes, and hence MVAR indexes are more appropriate. On the other hand, GC, DTF and PDC do not allow clear conclusions on the absolute strength of the coupling, because they are not a direct measure of strength connectivity, but the regression analysis confirms the existence a linear relationship among true and estimated intensity. Hence some processing are needed to evaluate their connectivity scores, which results greatly depend on the strategy adopted. This work evidences that GC is a stand-alone estimator for both topology and strength. In order to study connectivity in frequency domain, DTF and PDC should be used together, to obtain a robust network characterization.

6.3. Conclusions

The aim of this work was to critical assess a variety of brain connectivity methods to provide a validated protocol to support both cognitive research and clinical activity. To address the principal issues in characterizing connectivity among cortical areas, we performed a simulation study, where three ROIs interacted each other. It is remarkable that other authors performed such analysis using *in silico* datasets, but we exploited an innovative strategy based on a non linear neural mass model able to generate different EEG rhythms instead of simulating a system using a MVAR model. This was a crucial aspect for our validation, to understand to what extent the linear methods we considered, such are those based on MVAR and SEM, are able to describe connectivity of non linear signals, using an approach totally independent of MVAR and SEM linear equation and, hence, avoiding any risk of self-reference.

Discussion

In this work we modelled four simple network graphs: feed-forward, open-loop with two links, open-loop including a reciprocal connection and cycle closed-loop, by combining three ROIs which coupling strength was tuned by varying the linking parameters. The resulting signals were characterized by intrinsic rhythms (in alpha, beta and gamma range) and some received from the connected ROIs. To understand methods robustness and precision, we performed an extensive simulation considering the feed forward network: three datasets were generated with different non linearity degree, by changing the network mass model parameters. Then, using these datasets network topology and connectivity strength were estimated and compared with true networks.

We demonstrated that Granger causality is a good estimator with high values both of sensitivity and specificity, while frequency indexes, DTF and PDC, are too much affected by the threshold choice and their interpretation in terms of absolute strength connection is not clear.

As regard SEM, we proved the difficulty of its approach to describe just simple situations. Even if SEM is based on linear regression as well as MVAR models, it differently assumes there is no connection with past information, as if brain connectivity could describe time series relationships by the instant we observe it. Hence, it is not sufficiently robust to characterize neuronal dynamic activity.

Finally, our results suggest that Granger causality and DTF can be combined in a practical procedure to be followed in clinical research, since they are more robust to signal non linearity than PDC. Application described here showed how connectivity analysis can add further information on neuronal signals behaviour previously study by spectral analysis, only. Their results gave plausible explanations on the processes relationships, quantifying the different brain states classification: healthy volunteer Vs patients, before Vs after AAC.

We believe that the simulation study was very useful to highlight advantages and disadvantages of such methods. Further efforts should be done to cover a greater realism and complexity in the network model, introducing feedback loop and non

linear relationships in equations modelling the linking between ROIs. In addition, other multivariate methods should be evaluated to describe connectivity not only in terms of frequency coupling.

As we know, there are many debates on the recent upsurge of Granger causality and MVAR indexes, since they are not able to describe connectivity in terms of physical control. In this work, our aim is to explore data and then estimate a plausible network able to describe their relationships in terms of temporal precedence. There are many ways to intend coupling between time series, but each method must be used properly. Brain connectivity is still a highly promising framework in modern theories in neuroscience and cognition, hence, more and more advance techniques could be applied to neurobiological data and bring advantages to clinical applications.

A. SEM analysis in frequency domain

A.1. Path coefficients interpretation

Path coefficients represent the dependent variable response to a unit change in an explanatory variable, whilst the other variables in the model are held constant ([55]). The coefficient sign reveals what kind of covariance relationship exists between network components. A positive coefficient means a synergic connection, conversely, a negative one implies that the increasing activity in one variable leads to a decrease in the activity of the variable it projects to.

Another plausible explanation of their meaning can be obtained in frequency domain, considering each path coefficient of the model with Eq. 1.22 as a general LTI system as follows:

$$\begin{aligned}v'_2(t) &= h_A(t) * v'_1(t) + e'_2(t) \\v'_3(t) &= h_B(t) * v'_1(t) + h_C(t) * v'_2(t) + e'_3(t) \\ &= h_B(t) * v'_1(t) + h_C(t) * h_A(t) * v'_1(t) + h_C(t) * e'_2(t) + e'_3(t) \\ &= h_D(t) * v'_1(t) + h_C(t) * e'_2(t) + e'_3(t)\end{aligned}\tag{A.1}$$

where $h_A(t)$, $h_B(t)$ and $h_C(t)$ are the connections impulse responses of the network shown in Figure 1.22 and $h_D(t) = h_B(t) + h_C(t) * h_A(t)$.

In frequency domain, the linear system in Eq. A.1 is translated into the following spectral density functions:

$$S'_2(f) = |H_A(f)|^2 S'_1(f) + (\sigma'_2)^2\tag{A.2}$$

SEM analysis in frequency domain

$$S'_3(f) = |H_D(f)|^2 S'_1(f) + |H_C(f)|^2 (\sigma'_2)^2 + (\sigma'_3)^2, \quad (\text{A.3})$$

where

$$\begin{aligned} |H_D(f)|^2 &= |H_D(f)||H_D(f)| \\ &= |H_B(f) + H_C(f)H_A(f)||H_B(f) + H_C(f)H_A(f)|, \end{aligned} \quad (\text{A.4})$$

The above expressions evidence the action each variable exerts over the others. Applying proper algebraic tricks, Eq. A.3 can be developed in order to highlight the power contribution coming from each variable separately. By means of the properties $|zw| = |z||w|$ and $|z + w| = |z| + |w|$,

$$|H_C(f)H_A(f)| = |H_C(f)||H_A(f)|, \quad (\text{A.5})$$

and

$$|H_D(f)|^2 \leq (|H_B(f)| + |H_C(f)H_A(f)|)(|H_B(f)| + |H_C(f)H_A(f)|) \quad (\text{A.6})$$

can be substituted in Eq. A.3, obtaining an approximation of the spectral density of $v'_3(t)$:

$$\begin{aligned} S'_3(f) &\leq (|H_B(f)|^2 + |H_C(f)|^2 |H_A(f)|^2 + 2|H_B(f)||H_C(f)||H_A(f)|)S'_1(f) \\ &\quad + |H_C(f)|^2 (\sigma'_2)^2 + (\sigma'_3)^2. \end{aligned} \quad (\text{A.7})$$

Eq.A.7 shows that power inflow of variable $n = 3$ is a cross-combination of all the connection impulse responses. The linear equations which this method is based introduce the cross-correlation effect when at least the same two variables are linked with different paths, i.e. in the model 1.1 such are variables $n = 1$ and $n = 3$.

Matching the power spectral density expressions of eq. A.2 and eq. A.7 with the following, obtained by means of the Fourier Transform of the autocorrelation

A.1 Path coefficients interpretation

function of eq. 1.22:

$$S_1(f) = FT[R_{v_1}(\tau)] \quad (\text{A.8})$$

$$S_2(f) = FT[R_{v_2}(\tau)] = (k_{21}^2)S_1(f) + \sigma_2^2 \quad (\text{A.9})$$

$$\begin{aligned} S_3(f) &= FT[R_{v_3}(\tau)] \\ &= (k_{31}^2)S_1(f) + (k_{21} k_{32})^2 S_1(f) + 2(k_{31} k_{21} k_{32})S_1(f) + \\ &\quad (k_{32})^2 \sigma_2^2 + \sigma_3^2 \end{aligned} \quad (\text{A.10})$$

In terms of power, the scalar path coefficients k_{ij} of SEM can be approximated with the following integrals:

$$k_{21} \approx \left(\frac{1}{f_{max} - f_{min}} \int_{f_{min}}^{f_{max}} |H_A(f)|^2 df \right)^{1/2} \quad (\text{A.11})$$

$$k_{31} \approx \left(\frac{1}{f_{max} - f_{min}} \int_{f_{min}}^{f_{max}} |H_B(f)|^2 + |H_B(f)||H_C(f)||H_A(f)| df \right)^{1/2} \quad (\text{A.12})$$

$$k_{32} \approx \left(\frac{1}{f_{max} - f_{min}} \int_{f_{min}}^{f_{max}} |H_C(f)|^2 |H_A(f)|^2 + |H_B(f)||H_C(f)||H_A(f)| df \right)^{1/2} \quad (\text{A.13})$$

k_{21} exclusively depends on the transfer function, $H_A(f)$, associated to the respective direct connection. k_{31} and k_{32} result in the combination of the transfer functions included in all the paths targeting *ROI3*. This interpretation suggests that path coefficients are composite measures, representing not only the direct effect between two variables, but also the cross-effect mediated by the other variables in the linking pattern.

A.1.1. Parameters interpretation using the Neural Mass Model

In order to find a relationship among path coefficients and weights A, B and C, as explained in §. A.1, the impulse responses of the network model can be rewritten

SEM analysis in frequency domain

after linearization of the neural mass model, as formalized in Fig. 2.5:

$$\begin{aligned}
 v_1(t) &= (h_{ex_1}(t) + h_{in_1}(t)) * n(t) \\
 v_2(t) &= h_A(t) * v_1(t) + (h_{ex_2}(t) + h_{in_2}(t)) * n(t) \\
 v_3(t) &= h_B(t) * v_1(t) + h_C(t) * v_2(t) + (h_{ex_3}(t) + h_{in_3}(t)) * n(t)
 \end{aligned} \tag{A.14}$$

where

$$\begin{aligned}
 h_A(t) &= h_{ex_2}(t) * A \\
 h_B(t) &= h_{ex_3}(t) * B \\
 h_C(t) &= h_{ex_3}(t) * C
 \end{aligned}$$

Expressions of eq. A.11, eq. A.12 and eq. A.13 allow to calculate the approximation of SEM parameters in terms of A, B, C and network model transfer functions as follows:

$$k_{21}^* = \left(\frac{1}{f_{max} - f_{min}} \int_{f_{min}}^{f_{max}} A^2 |H_{ex_2}(f)|^2 df \right)^{1/2} \tag{A.15}$$

$$\begin{aligned}
 k_{31}^* &= \left(\frac{1}{f_{max} - f_{min}} \right)^{1/2} \\
 &\left(\int_{f_{min}}^{f_{max}} B^2 |H_{ex_3}(f)|^2 + B |H_{ex_3}(f)| C |H_{ex_3}(f)| A |H_{ex_2}(f)| df \right)^{1/2}
 \end{aligned} \tag{A.16}$$

$$\begin{aligned}
 k_{32}^* &= \left(\frac{1}{f_{max} - f_{min}} \right)^{1/2} \\
 &\left(\int_{f_{min}}^{f_{max}} C^2 |H_{ex_3}(f)|^2 A^2 |H_{ex_2}(f)|^2 + B |H_{ex_3}(f)| C |H_{ex_3}(f)| A |H_{ex_2}(f)| df \right)^{1/2}.
 \end{aligned} \tag{A.17}$$

where H_{ex_n} , with $n = 1, 2, 3$, are the Fourier transform of the ROIs impulse responses.

The same condition of the simulation study are considered to compute SEM path coefficients and the ones derived by the model parameter, eqs. A.15, A.16, A.17.

Fig. A.1.1 reports the linear correlation between estimated k_{ij} values via SEM (mean \pm SD; $N = 100$) and k_{ij}^* values derived via Eqs. A.15, A.16 and A.17. High scores of R^2 evidence the existence of a strong linear relationship between

A.1 Path coefficients interpretation

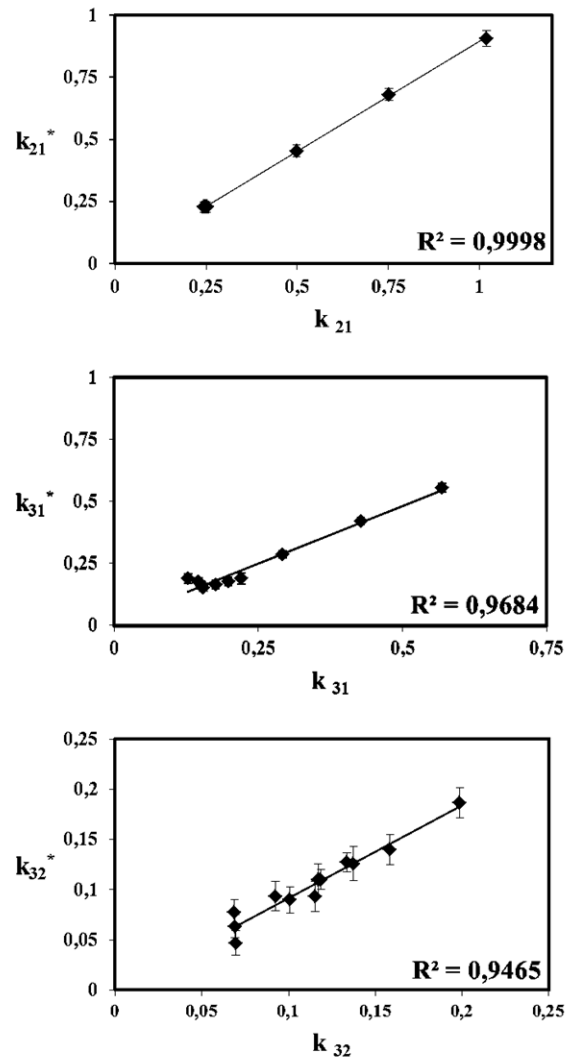


Figure A.1.: Comparison between estimated k_{ij} values via SEM (mean \pm SD; $N = 100$) and k_{ij}^* values derived via Eqs. A.15,A.16 and A.17.

SEM analysis in frequency domain

them and confirm that path coefficients can be explain as composite measures taking into account of all network contribute.

Bibliography

- [1] K. E. Stephan. On the role of general system theory for functional neuroimaging. *J. Anat.*, 205(6):443–470, Dec 2004.
- [2] B. Horwitz and S. G. Horovitz. Introduction to research topic - brain connectivity analysis: investigating brain disorders. Part 1: the review articles. *Front Syst Neurosci*, 6:3, 2012.
- [3] S. G. Horovitz and B. Horwitz. Introduction to research topic - brain connectivity analysis: investigating brain disorders. Part 2: original research articles. *Front Syst Neurosci*, 6:4, 2012.
- [4] L. Lemieux, J. Daunizeau, and M. C. Walker. Concepts of connectivity and human epileptic activity. *Front Syst Neurosci*, 5:12, 2011.
- [5] P. A. Valdes-Sosa, A. Roebroeck, J. Daunizeau, and K. Friston. Effective connectivity: influence, causality and biophysical modeling. *Neuroimage*, 58(2):339–361, Sep 2011.
- [6] C.W.J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):768–778, 1969.
- [7] K. J. Blinowska. Review of the methods of determination of directed connectivity from multichannel data. *Med Biol Eng Comput*, 49(5):521–529, May 2011.
- [8] A. Schlogl and G. Supp. Analyzing event-related EEG data with multivariate autoregressive parameters. *Prog. Brain Res.*, 159:135–147, 2006.
- [9] F. Mormann, K. Lehnertz, P. David, and E. Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. *Physica D Nonlinear Phenomena*, 144(3-4):358–369, October 2000.
- [10] J. Rissanen and M. Wax. Measures of mutual and causal dependence between two time series. *IEEE Trans. Inf. Theor.*, 33(4):598–601, July 1987.
- [11] K. J. Friston, L. Harrison, and W. Penny. Dynamic causal modelling. *Neuroimage*, 19(4):1273–1302, Aug 2003.
- [12] A. R. McIntosh and F. Gonzalez-Lima. Structural equation modeling and its application to network analysis in functional brain imaging. *Hum. Brain Mapp*, 2(22), 1994.
- [13] K. J. Friston. Functional and effective connectivity: a review. *Brain Connect*, 1(1):13–36, 2011.
- [14] W. D. Penny, K. E. Stephan, A. Mechelli, and K. J. Friston. Modelling functional integration: a comparison of structural equation and dynamic causal models. *Neuroimage*, 23 Suppl 1:S264–274, 2004.
- [15] L. Astolfi, F. Cincotti, D. Mattia, S. Salinari, C. Babiloni, A. Basilisco, P. M. Rossini, L. Ding, Y. Ni, B. He, M. G. Marciani, and F. Babiloni. Estimation of the effective and functional human cortical connectivity with structural equation modeling and directed transfer function applied to high-resolution EEG. *Magn Reson Imaging*, 22(10):1457–1470, Dec 2004.

Bibliography

- [16] A. Boucard, A. Marchand, and X. Nogues. Reliability and validity of structural equation modeling applied to neuroimaging data: a simulation study. *J. Neurosci. Methods*, 166(2):278–292, Nov 2007.
- [17] M. Ursino, F. Cona, and M. Zavaglia. The generation of rhythms within a cortical region: analysis of a neural mass model. *Neuroimage*, 52(3):1080–1094, Sep 2010.
- [18] N. Wiener. *The Theory of Prediction*. McGraw-Hill New York, 1956.
- [19] J. Geweke. Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77:304–313, 1982.
- [20] J. Geweke. *Inference and causality in economic time series models*. in: *Handbook of Econometrics*. Elsevier, 1984.
- [21] S. L. Bressler and A. K. Seth. Wiener-Granger causality: a well established methodology. *Neuroimage*, 58(2):323–329, Sep 2011.
- [22] A. Schlogl and G. Supp. Analyzing event-related EEG data with multivariate autoregressive parameters. *Prog. Brain Res.*, 159:135–147, 2006.
- [23] M. J. Kaminski and K. J. Blinowska. A new method of the description of the information flow in the brain structures. *Biol Cybern*, 65(3):203–210, 1991.
- [24] L. A. Baccala and K. Sameshima. Partial directed coherence: a new concept in neural structure determination. *Biol Cybern*, 84(6):463–474, Jun 2001.
- [25] A. R. McIntosh, C. L. Grady, L. G. Ungerleider, J. V. Haxby, S. I. Rapoport, and B. Horwitz. Network analysis of cortical visual pathways mapped with PET. *J. Neurosci.*, 14(2):655–666, Feb 1994.
- [26] C. Buchel and K. J. Friston. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cereb. Cortex*, 7(8):768–778, Dec 1997.
- [27] H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12(1):1–24, Jan 1972.
- [28] F. H. Lopes da Silva, A. Hoeks, H. Smits, and L. H. Zetterberg. Model of brain rhythmic activity. The alpha-rhythm of the thalamus. *Kybernetik*, 15(1):27–37, May 1974.
- [29] B. H. Jansen, G. Zouridakis, and M. E. Brandt. A neurophysiologically-based mathematical model of flash visual evoked potentials. *Biol Cybern*, 68(3):275–283, 1993.
- [30] O. David and K. J. Friston. A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage*, 20(3):1743–1755, Nov 2003.
- [31] O. David, L. Harrison, and K. J. Friston. Modelling event-related responses in the brain. *Neuroimage*, 25(3):756–770, Apr 2005.
- [32] A. Babajani and H. Soltanian-Zadeh. Integrated MEG/EEG and fMRI model based on neural masses. *IEEE Trans Biomed Eng*, 53(9):1794–1801, Sep 2006.
- [33] M. Zavaglia, L. Astolfi, F. Babiloni, and M. Ursino. The effect of connectivity on EEG rhythms, power spectral density and coherence among coupled neural populations: analysis with a neural mass model. *IEEE Trans Biomed Eng*, 55(1):69–77, Jan 2008.
- [34] F. Wendling, F. Bartolomei, J. J. Bellanger, and P. Chauvel. Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *Eur. J. Neurosci.*, 15(9):1499–1508, May 2002.

- [35] M. Zavaglia, L. Astolfi, F. Babiloni, and M. Ursino. A neural mass model for the simulation of cortical activity estimated from high resolution EEG during cognitive or motor tasks. *J. Neurosci. Methods*, 157(2):317–329, Oct 2006.
- [36] M. Ursino, M. Zavaglia, L. Astolfi, and F. Babiloni. Use of a neural mass model for the analysis of effective connectivity among cortical regions based on high resolution EEG recordings. *Biol Cybern*, 96(3):351–365, Mar 2007.
- [37] M. Ursino, M. Zavaglia, E. Magosso, A. Serino, and G. di Pellegrino. A neural network model of multisensory representation of peripersonal space: effect of tool use. *Conf Proc IEEE Eng Med Biol Soc*, 2007:2735–2739, 2007.
- [38] M. Zavaglia, F. Cona, and M. Ursino. A neural mass model to simulate different rhythms in a cortical region. *Comput Intell Neurosci*, page 456140, 2010.
- [39] F. Cona, M. Zavaglia, M. Massimini, M. Rosanova, and M. Ursino. A neural mass model of interconnected regions simulates rhythm propagation observed via TMS-EEG. *Neuroimage*, 57(3):1045–1058, Aug 2011.
- [40] R. J. Moran, S. J. Kiebel, K. E. Stephan, R. B. Reilly, J. Daunizeau, and K. J. Friston. A neural mass model of spectral responses in electrophysiology. *Neuroimage*, 37(3):706–720, Sep 2007.
- [41] Arnold Neumaier and Tapio Schneider. Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans. Math. Softw.*, 27(1):27–57, 2001.
- [42] A. K. Seth. A MATLAB toolbox for Granger causal connectivity analysis. *J. Neurosci. Methods*, 186(2):262–273, Feb 2010.
- [43] F. Piccolo. Analisi parametrica multivariata del segnale EEG per la stima della connettività corticale. Master’s thesis, Università degli Studi di Padova, Dipartimento di Ingegneria dell’Informazione, 2011.
- [44] John Fox. TEACHER’S CORNER: Structural Equation Modeling With the sem Package in R. *Structural Equation Modeling: A Multidisciplinary Journal*, 13(3):465–486, June 2006.
- [45] M. Kaminski, M. Ding, W. A. Truccolo, and S. L. Bressler. Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance. *Biol Cybern*, 85(2):145–157, Aug 2001.
- [46] S. J. Munoz. Hepatic encephalopathy. *Med. Clin. North Am.*, 92(4):795–812, Jul 2008.
- [47] P. Amodio, R. Orsato, P. Marchetti, S. Schiff, C. Poci, P. Angeli, A. Gatta, G. Sparacino, and G. M. Toffolo. Electroencephalographic analysis for the assessment of hepatic encephalopathy: comparison of non-parametric and parametric spectral estimation techniques. *Neurophysiol Clin*, 39(2):107–115, Apr 2009.
- [48] P. Marchetti, C. D’Avanzo, R. Orsato, S. Montagnese, S. Schiff, P. W. Kaplan, F. Piccione, C. Merkel, A. Gatta, G. Sparacino, G. M. Toffolo, and P. Amodio. Electroencephalography in patients with cirrhosis. *Gastroenterology*, 141(5):1680–1689, Nov 2011.
- [49] A. Bersagliere, I. D. Raduazzo, M. Nardi, S. Schiff, A. Gatta, P. Amodio, P. Achermann, and S. Montagnese. Induced hyperammonemia may compromise the ability to generate restful sleep in patients with cirrhosis. *Hepatology*, 55(3):869–878, Mar 2012.

Bibliography

- [50] A. Bersagliere, I. D. Raduazzo, S. Schiff, A. Gatta, C. Merkel, P. Amodio, P. Achermann, and S. Montagnese. Ammonia-related changes in cerebral electrogenesis in healthy subjects and patients with cirrhosis. *Clin Neurophysiol*, Sep 2012.
- [51] R. Qi, Q. Xu, L. J. Zhang, J. Zhong, G. Zheng, S. Wu, Z. Zhang, W. Liao, Y. Zhong, L. Ni, Q. Jiao, Z. Zhang, Y. Liu, and G. Lu. Structural and functional abnormalities of default mode network in minimal hepatic encephalopathy: a study combining DTI and fMRI. *PLoS ONE*, 7(7):e41376, 2012.
- [52] J. Toppi, F. Babiloni, G. Vecchiato, F. Cincotti, F. De Vico Fallani, D. Mattia, S. Salinari, and L. Astolfi. Testing the asymptotic statistic for the assessment of the significance of Partial Directed Coherence connectivity patterns. *Conf Proc IEEE Eng Med Biol Soc*, 2011:5016–5019, 2011.
- [53] S. Haufe, V. V. Nikulin, K. R. Muller, and G. Nolte. A critical assessment of connectivity measures for EEG data: a simulation study. *Neuroimage*, 64:120–133, Jan 2013.
- [54] Julius S. Bendat and Allan G. Piersol. *Random Data: Analysis & Measurement Procedures*. Wiley-Interscience, 2000.
- [55] KA Bollen. *Structural equations with latent variables*. New York: John Wiley, 1989.