

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
BIOINGEGNERIA

Ciclo XXV

Settore Concorsuale di afferenza: 09/G2

Settore Scientifico disciplinare: ING-INF/06

TITOLO TESI

Methods for the Estimation of the Cortical Activity and
Connectivity during Cognitive Tasks in Humans

Presentata da: JLENIA TOPPI

Coordinatore Dottorato

Relatore

Prof. Angelo Cappello

Prof.ssa Serenella Salinari

Esame finale anno 2013



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA



SAPIENZA
UNIVERSITÀ DI ROMA

Università di Bologna “Alma Mater Studiorum”

Università degli Studi di Roma “La Sapienza”

Corso di Dottorato di Ricerca in Bioingegneria – XXV ciclo

Coordinatore: Prof. Angelo Cappello

A.A. 2011/2012

Settore Scientifico/Disciplinare ING-INF/06

TESI DI DOTTORATO

**METHODS FOR THE ESTIMATION OF CORTICAL
ACTIVITY AND CONNECTIVITY DURING COGNITIVE
TASKS IN HUMANS**

Candidato: **Ing. Jlenia Toppi**

Supervisore: **Prof.ssa Serenella Salinari**

Co-Supervisore: **Ing. Laura Astolfi**

Controrelatori: **Prof. Mauro Ursino**

Prof. Luiz A. Baccalà

Table of Contents

GENERAL INTRODUCTION.....	5
<u>SECTION I: STATISTICAL ASSESSMENT OF STATIONARY FUNCTIONAL CONNECTIVITY PATTERNS ESTIMATED ON NON-INVASIVE EEG MEASURES</u>	
INTRODUCTION	9
STATISTICAL ASSESSMENT OF CONNECTIVITY PATTERNS: STATE OF THE ART	13
Multivariate Methods for the Estimation of Connectivity	13
Partial Directed Coherence	13
Statistical tests for the assessment of connectivity patterns	15
Shuffling Method: Empirical Distribution	17
Asymptotic Statistic Method: Theoretical Distribution	17
Reducing the occurrence of Type I errors in assessment of connectivity patterns	18
False Discovery Rate	19
Bonferroni adjustment	19
COMPARING METHODS FOR THE STATISTICAL ASSESSMENT OF CONNECTIVITY PATTERNS: A SIMULATION STUDY	21
The Simulation Study	21
Signal Generation	22
Evaluation of performances	23
Statistical Analysis	24
Results	26
Discussion	36
CONCLUSION.....	39
SECTION I REFERENCES	41
<u>SECTION II: TIME-VARYING APPROACH FOR THE ESTIMATION OF FUNCTIONAL CONNECTIVITY</u>	
INTRODUCTION	44
ESTIMATION OF TIME-VARYING FUNCTIONAL CONNECTIVITY: STATE OF THE ART	47
Adaptive Partial Directed Coherence	47
The Recursive Least Square	47
The General Linear Kalman Filter	49
COMPARING METHODS FOR THE ESTIMATION OF TIME-VARYING FUNCTIONAL CONNECTIVITY: A SIMULATION STUDY.....	51
The Simulation Study	51
Signal Generation	52
Evaluation of performances	54

Statistical Analysis	57
Results	58
Discussion	67
CONCLUSION.....	70
SECTION II REFERENCES	72
<u>SECTION III: GRAPH THEORY APPROACH: A CONSISTENT DESCRIPTION OF THE RELEVANT PROPERTIES OF THE BRAIN NETWORKS</u>	
INTRODUCTION.....	73
GRAPH THEORY APPROACH: STATE OF THE ART	76
Concept at the basis of Graph Theory	76
Adjacency Matrix Extraction	77
Graph Theory Indexes	77
BEYOND THE STATE OF THE ART: HOW THE STATISTICAL VALIDATION OF FUNCTIONAL CONNECTIVITY PATTERNS CAN PREVENT ERRONEOUS DEFINITION OF SMALL-WORLD PROPERTIES OF A BRAIN CONNECTIVITY NETWORK.....	81
Methods	82
Datasets used in the experiment	82
Signal Processing	83
Analysis of Variance	84
Results	84
Simulated data	84
Mannequin data	89
Discussion	93
CONCLUSION.....	96
SECTION III REFERENCES	98
<u>SECTION IV: FUNCTIONAL CONNECTIVITY FOR THE STUDY OF RESTING STATE AND COGNITIVE PROCESSES IN HUMAN</u>	
INTRODUCTION.....	103
DESCRIBING RELEVANT INDEXES FROM THE RESTING STATE ELECTROPHYSIOLOGICAL NETWORKS.....	105
Introduction	105
Material and Methods	106
Experimental Design	106
Pre-Processing and Functional Connectivity Analysis	106
Results	107
Discussion	111
FUNCTIONAL CONNECTIVITY FOR INVESTIGATING THE BASIS OF COGNITIVE PROCESSES ..	113

Introduction	113
Material and Methods	114
Experimental Design	114
Visual Oddball Task	114
Sternberg Task	115
Pre-Processing of EEG traces	117
Results	117
EEG Visual Oddball Task	117
EEG Sternberg Task	124
Discussion	131
CONCLUSION OF SECTION IV	133
SECTION IV REFERENCES	134
GENERAL CONCLUSION	136

General Introduction

The development of a methodology for the estimation of the information flow between different and differently specialized cerebral areas, starting from non-invasive measurements of the neuro-electrical brain activity, has gained more and more importance in the field of Neuroscience, as an instrument to investigate the neural basis of cerebral processes. In fact, the knowledge of the complex cerebral dynamics underlying human cognitive processes cannot be fully achieved by the reconstruction of temporal/spectral activations of different parts of the brain. The description of the brain circuits involved in the execution of a task is a crucial point for understanding the mechanisms at the basis of the specific function under examination, as well as for the development of applications in the clinical and rehabilitation fields.

In the last two decades, several approaches were developed and applied to neuroelectrical signals in order to estimate the connectivity patterns elicited among cerebral areas. Several studies investigated the properties of all the available methods providing different solutions for different application fields and highlighting the best approaches able to reproduce the brain circuits related to non-invasive EEG measurements.

The theory of Granger Causality (1969), based on the statistical studies on causality by Norbert Wiener (1956), is one of the most generally adopted in EEG based connectivity studies, due to its many advantages in terms of generality, easiness, and possibility to provide a fully multivariate analysis of brain circuits in terms of existence, strength, direction of the functional links. At the same time, this method avoids the necessity of *a priori* information on the brain circuits under investigation. Since it provides not a mere description of a synchronicity between distant brain structures, but an hypothesis on the brain circuits, including the influence exerted by a neural system to the others, it can be considered an estimator of the so-called *effective* connectivity.

Notwithstanding the advancements provided in this respect during the last twenty years, the main problem still unsolved regards the stability and reliability of the connectivity patterns obtained from Granger Causality-based approaches. Such issue still needs to be solved in order to provide an instrument really able to fulfill clinical and applicative purposes, where the reliability of the results and their consistence among the population are mandatory.

For this reason, my main aim during the three years of my PhD course was to study, develop and refine methodologies for effective connectivity estimation allowing to overcome the limitations of existent procedures with the aim to produce a valid tool able to reliably describe human connectivity networks and their global and local properties in different application fields.

The main aims of my PhD thesis can be summarized by two main goals:

1. METHODOLOGICAL AIMS

Development of a stable, consistent and reproducible procedure for effective connectivity estimation with a high impact on neuroscience field in order:

1. To avoid the **a priori selection** of the channels or brain regions to be included in the model
2. To assess the significance of functional **connections**, including the **corrections** for multiple comparisons
3. To consistently describe **relevant properties** of the brain networks
4. To accurately describe the **temporal evolution** of effective connectivity patterns

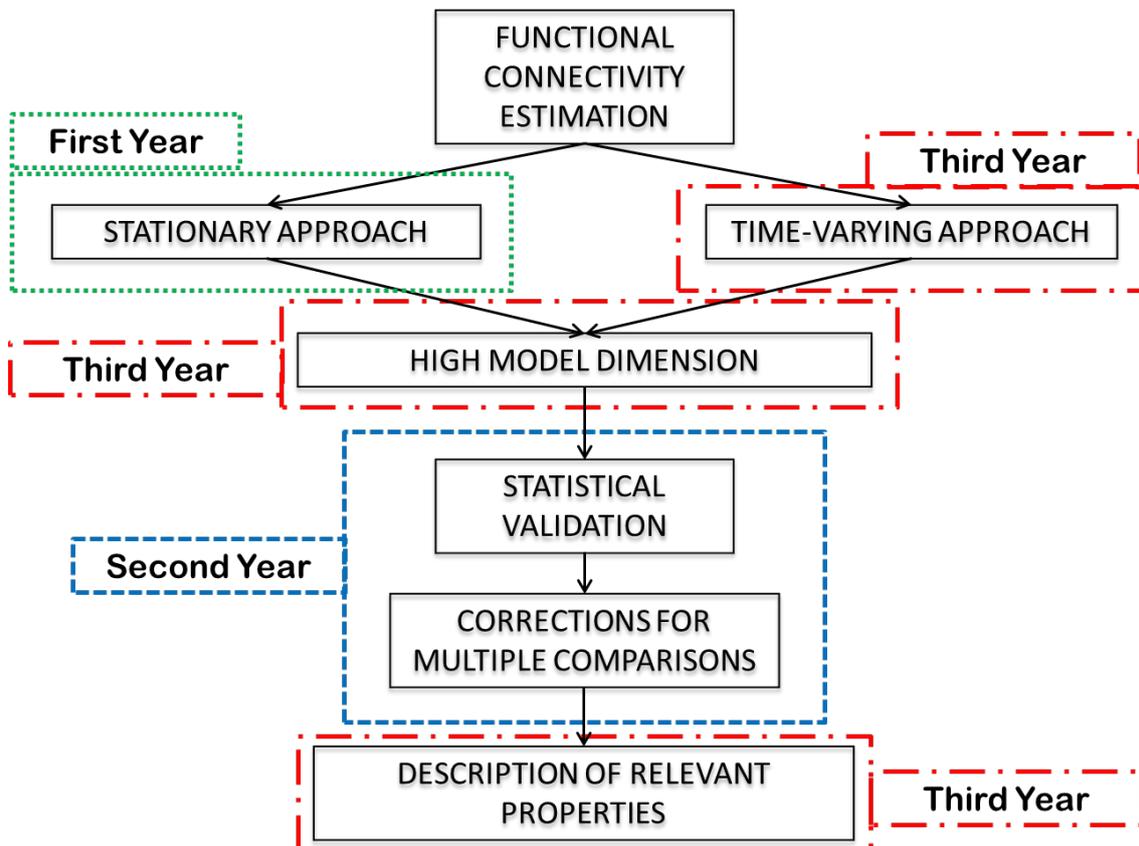
2. APPLICATIVE AIMS

1. Testing the new methodologies on real data in order to evaluate their application field.
2. to find **quantifiable descriptors**, based on brain connectivity indexes, of resting condition and cognitive processes (attention, memory)

In order to reach these two main goals I followed the roadmap reported in the figure below.

In particular, the first year was dedicated to the study of the state of the art of all methodologies for effective connectivity estimation and to the refining of existent approaches for the estimation of connectivity patterns in the stationary case, i.e. when the statistical properties of the data can be considered stable during the observation window of our experiments. During the second year I faced the problem of statistically validating the estimated connectivity patterns, by comparing the existent assessing methodologies in an extensive simulation study with the aim to understand which was the best approach to be used in relation to the quality of the available data. Moreover, for the first time in connectivity field, I introduced corrections for multiple comparisons in the statistical validation process in

order to avoid the occurrence of false positives which might affect the topology of validated connectivity networks. The third year was dedicated to three important issues regarding the necessities i) to include all the sources in the model used for connectivity estimation, ii) to follow the temporal dynamics of connectivity patterns with a time varying approach and iii) to reliably extract indexes characterizing in a stable way the main properties of investigated networks.



The structure of the thesis reflects the route followed during these three years with the aim to produce a valid tool able to consistently and reliably estimate connectivity patterns and investigate their properties.

In Section I, after a description of the state of the art of stationary connectivity, an extensive simulation study aiming at comparing the two existent assessing methods under different conditions of data quality was described.

In Section II, a simulation study comparing the two more advanced methodologies for the estimation of time-varying Granger Causality connectivity was proposed.

In Section III, after a brief introduction on the concepts at the basis of graph theory, a study demonstrating how the procedure for the extraction of graph indexes from the

connectivity matrix can affect the topological properties of the investigated networks was provided.

In Section IV, all the developed methodologies were tested on real data in two neuroscience applications regarding the investigation of resting state brain networks and cognitive processes such as attention and memory in humans.

Section I

Statistical assessment of stationary functional connectivity patterns estimated on non-invasive EEG measures

INTRODUCTION

STATISTICAL ASSESSMENT OF CONNECTIVITY PATTERNS: STATE OF THE ART

Multivariate Methods for the Estimation of Connectivity

Partial Directed Coherence

Statistical tests for the assessment of connectivity patterns

Shuffling Method: Empirical Distribution

Asymptotic Statistic Method: Theoretical Distribution

Reducing the occurrence of Type I errors in assessment of connectivity patterns

False Discovery Rate

Bonferroni adjustment

COMPARING METHODS FOR THE STATISTICAL ASSESSMENT OF CONNECTIVITY PATTERNS: A SIMULATION STUDY

The Simulation Study

Signal Generation

Evaluation of performances

Statistical Analysis

Results

Discussion

CONCLUSION

SECTION I REFERENCES

Introduction

In neuroscience field, the concept of brain connectivity (i.e. how the cortical areas communicate one to each other) is central for the understanding of the organized behavior of cortical regions beyond the simple mapping of their activity (Horwitz, 2003; Lee et al., 2003). In the last two decades several studies have been carried on in order to understand neuronal networks at the basis of mental processes, which are characterized by lots of interactions between different and differently specialized cortical sites. Connectivity estimation techniques aim at describing interactions between electrodes or cortical areas as connectivity patterns

holding the direction and the strength of the information flow between such areas. The functional connectivity between cortical areas is then defined as the temporal correlation between spatially remote neurophysiologic events and it could be estimated by using different methods both in time as well as in frequency domain based on bivariate or multivariate autoregressive models (Blinowska, 2011; Kaminski and Blinowska, 1991; Sameshima and Baccalá, 1999; Turbes et al., 1983). Several studies have proved the higher efficiency in estimating functional connectivity of methods, such as Directed Transfer Function (DTF) (Kaminski and Blinowska, 1991) or Partial Directed Coherence (PDC) (Baccalá and Sameshima, 2001), which are defined in frequency domain and based on the use of multivariate autoregressive (MVAR) models built on original time-series (Kus et al., 2004). In fact the bivariate approach is affected by a high number of false positives due to the impossibility of the method in discarding a common effect on a couple of signals of a third one acquired simultaneously (Blinowska et al., 2004). Moreover the bivariate methods give rise to very dense patterns of propagation, thus it is impossible to find the sources of propagation (Blinowska et al., 2010; David et al., 2004). The PDC technique has been demonstrated (Baccalá and Sameshima, 2001) to rely the concept of Granger causality between time series (Granger, 1969), according to which an observed time series $x(n)$ causes another series $y(n)$ if the knowledge of $x(n)$'s past significantly improves prediction of $y(n)$. Moreover, the PDC is also of particular interest because it can distinguish between direct and indirect connectivity flows in the estimated connectivity pattern better than Directed Transfer Function (DTF) and its direct modified version, the dDTF (Astolfi et al., 2007).

Random correlations between signals induced by environmental noise or by chance can lead to false detection of links or to the loss of true existing connections in the connectivity estimation process. In order to minimize this phenomenon, the process of functional connectivity estimation has to be followed by a procedure able to assess the statistical significance of estimated networks, distinguishing the estimated connectivity patterns from the null case. As PDC functions have a highly nonlinear relation to the time series data from which they are derived, the distribution of such estimator in the null case has not been well established for several years. Therefore, several methods have been introduced in order to build the PDC distribution in an empirical way (Kaminski et al., 2001; Sameshima and Baccalá, 1999). The first one, the Spectral Causality Criterion, consisted in the use of the threshold value 0.1 for all the connections, directions and frequency samples included in the analysis (Sameshima and Baccalá, 1999). Its non-dependence from the data subjected to effective connectivity analysis led to the development of a data-driven approach, the shuffling

procedure, able to empirically build the null-case distribution by shuffling the phase of the investigated signals. Such approach had the advantage to increase the accuracy of estimation approach but at the same time the disadvantage to reduce the speed of validation process, being based on a time consuming approach. Only in 2007 the real distribution of PDC in the null-case was asymptotically theorized (Takahashi et al., 2007), allowing to increase the accuracy of validation process and reducing the computational time required for empirical methods.

All the assessing methods required the definition of a significance level to be applied in order to evaluate the statistical threshold for each possible connection between signals in the multivariate dataset and for each frequency sample. Due to the high number of statistical assessments performed simultaneously, it is necessary to apply corrections to the significance level imposed in the validation process, in order to prevent the occurrence of type I errors (Nichols and Hayasaka, 2003). In fact, statistical theory offers a lot of solutions for adequately manage the occurrence of type I errors during the execution of multiple univariate tests, such as the traditional Bonferroni adjustment (Bonferroni, 1936) or the more recent False Discovery Rate (FDR) (Benjamini and Yekutieli, 2001; Benjamini and Hochberg, 1995).

Even if the Shuffling method for statistical validation of connectivity patterns has been proposed in literature since 2001, a comprehensive analysis of its performances under different conditions of signal-to-noise ratio (factor SNR) and signal lengths (factor LENGTH) was not available in literature before my PhD. In addition, the definition of a new assessing method allowed to make comparisons between the two procedures in terms of percentages of type I and type II errors occurred during the validation process in order to highlight the best assessing procedure to be applied also in relation with the quality (SNR level) and amount of available data. Moreover, for the first time in the functional connectivity field, I proposed the introduction of corrections for multiple comparisons in the assessment process of connectivity patterns, also studying their effects on the validated networks.

In the present section of my thesis, first of all I briefly describe the methods available at the moment for assessing the estimated functional connectivity patterns and for adjusting the significance level of statistical test by taking into account multiple comparisons. This overview was followed by a description of the simulation study performed in order to compare the performances achieved by the two validation methods in terms of both false positives and false negatives under different condition of signal to noise ratio (SNR) and

amount of data available for the analysis. In particular the study allowed to address the following specific questions:

- 1) Is the Shuffling procedure a reliable method for preventing the occurrence of type I and type II errors?
- 2) How the statistical corrections for multiple comparisons affect the performances of Shuffling procedure?
- 3) Known the high computational complexity of the procedure based on empirical reconstruction of null-case distribution (shuffling), can the method based on its theoretical definition (Asymptotic Statistic) substitute it in the assessment of connectivity patterns?
- 4) Does the definition of PDC estimator (normalization according to columns or rows) influence the performances of the two assessing methods?
- 5) How are the performances of the two methods influenced by different factors affecting the recordings, like the signal-to-noise ratio (factor SNR) and the amount of data at disposal for the analysis (factor LENGTH)?

In order to answer these questions, a simulation study was performed, on the basis of a predefined connectivity scheme which linked several modeled cerebral areas. Functional connections between these areas were retrieved by the estimation process under different experimental SNR and LENGTH conditions. Analysis of variance (ANOVA) and Duncan's pairwise comparisons were applied in order to: i) evaluate the effect of different factors such as SNR and data length on the percentages of false positives and false negatives occurred during the validation process; ii) compare the existent Shuffling procedure with the new Asymptotic Statistic approach in order to understand the best assessing method also in relation to the quality of data available for the analysis.

Statistical Assessment of Connectivity Patterns: State of the Art

Multivariate Methods for the Estimation of Connectivity

Let Y be a set of signals, obtained from non-invasive EEG recordings or from the reconstruction of neuro-electrical cortical activities based on scalp measures:

$$Y = [y_1(t), y_2(t), \dots, y_N(t)]^T \quad (1.1)$$

where t refers to time and N is the number of electrodes or cortical areas considered.

Supposing that the following MVAR process is an adequate description of the dataset Y :

$$\sum_{k=0}^p \Lambda(k)Y(t-k) = E(t) \quad \text{with} \quad \Lambda(0) = I \quad (1.2)$$

where $Y(t)$ is the data vector in time, $E(t) = [e_1(t), \dots, e_n(t)]^T$ is a vector of multivariate zero-mean uncorrelated white noise processes, $\Lambda(1), \Lambda(2), \dots, \Lambda(p)$ are the $N \times N$ matrices of model coefficients and p is the model order, usually chosen by means of the Akaike Information Criteria (AIC) for MVAR processes (Akaike, 1974).

Once an MVAR model is adequately estimated, it becomes the basis for subsequent spectral analysis. To investigate the spectral properties of the examined process, Eq. (1.2) is transformed to the frequency domain:

$$\Lambda(f)Y(f) = E(f) \quad (1.3)$$

where:

$$\Lambda(f) = \sum_{k=0}^p \Lambda(k)e^{-j2\pi f \Delta t k} \quad (1.4)$$

and Δt is the temporal interval between two samples.

Partial Directed Coherence

The Partial Directed Coherence (PDC) (Baccalá and Sameshima, 2001) is a full multivariate spectral measure, used to determine the directed influences between pairs of signals in a multivariate data set. PDC is a frequency domain representation of the existing multivariate

relationships between simultaneously analyzed time series that allows the inference of functional relationships between them. This estimator was demonstrated to be a frequency version of the concept of Granger causality (Granger, 1969), according to which a time series $x[n]$ can be said to have an influence on another time series $y[n]$ if the knowledge of past samples of x significantly reduces the prediction error for the present sample of y .

It is possible to define PDC as

$$\pi_{ij}(f) = \frac{\Lambda_{ij}(f)}{\sqrt{\sum_{k=1}^N \Lambda_{kj}(f) \Lambda_{kj}^*(f)}} \quad (1.5)$$

which bears its name by its relation to the well-known concept of partial coherence (Baccalá, 2001). The PDC from node j to node i , $\pi_{ij}(f)$, describes the directional flow of information from the signal $y_j(n)$ to $y_i(n)$, whereupon common effects produced by other signals $y_k(n)$ on the latter are subtracted leaving only a description that is exclusive from $y_j(n)$ to $y_i(n)$. PDC squared values are in the interval $[0;1]$ and the normalization condition

$$\sum_{n=1}^N |\pi_{nj}(f)|^2 = 1 \quad (1.6)$$

is verified. According to this normalization, $\pi_{ij}(f)$ represents the fraction of the information flow of node j directed to node i , as compared to all the j 's interactions with other nodes.

Even if this formulation derived directly from information theory, the original definition was modified in order to give a better physiological interpretation to the estimate results achieved on electrophysiological data. In particular, a new type of normalization, already used for another connectivity estimator such as Directed Transfer Function (Kaminski and Blinowska, 1991) was introduced. Such normalization consisted in dividing each estimated value of PDC for the root squared sums of all elements of the relative row, obtaining the following definition:

$$\pi_{ij}^{row}(f) = \frac{\Lambda_{ij}(f)}{\sqrt{\sum_{k=1}^N \Lambda_{ik}(f) \Lambda_{ik}^*(f)}} \quad (1.7)$$

Even in this formulation PDC squared values are in the range $[0;1]$ but the normalization condition is as follows:

$$\sum_{n=1}^N |\pi_{in}^{row}(f)|^2 = 1 \quad (1.8)$$

Moreover, a squared formulation of PDC has been introduced and can be defined as follows for the two types of normalization:

$$sPDC_{ij}^{col}(f) = \frac{|\Lambda_{ij}(f)|^2}{\sum_{k=1}^N |\Lambda_{kj}(f)|^2} \quad (1.9)$$

$$sPDC_{ij}^{row}(f) = \frac{|\Lambda_{ij}(f)|^2}{\sum_{k=1}^N |\Lambda_{ik}(f)|^2} \quad (1.10)$$

The main difference with respect to the original formulation is in the interpretation of these estimators. Squared PDC can be put in relationship with the power density of the investigated signals and can be interpreted as the fraction of i^{th} signal power density due to the j^{th} measure. The higher performances of squared methods in respect to simple PDC have been demonstrated in a simulation study. Such study revealed higher accuracy, for the methods based on squared formulation of PDC, in the estimation of connectivity patterns on data characterized by different lengths and signal to noise ratio (SNR) and in distinction between direct and indirect pathways (Astolfi et al., 2006).

Statistical tests for the assessment of connectivity patterns

Random correlation between signals induced by environmental noise or by chance can lead to the presence of spurious links in the connectivity estimation process. In order to assess the significance of estimated patterns, the value of functional connectivity for a given pair of signals and for each frequency, obtained by computing PDC, has to be statistically compared with a threshold level which is related to the lack of transmission between considered signals (Fig. 1.1). Due to the high nonlinear dependence of the PDC from the estimated MVAR parameters, the theoretical distribution for such estimator in the null case has been not known until few years ago. For this reason different procedures for empirically reconstructing the distribution of the PDC in the null case (absence of connection) were developed in order to extract statistical thresholds to be compared with the inferred patterns by means of statistical tests for a fixed significance level α .

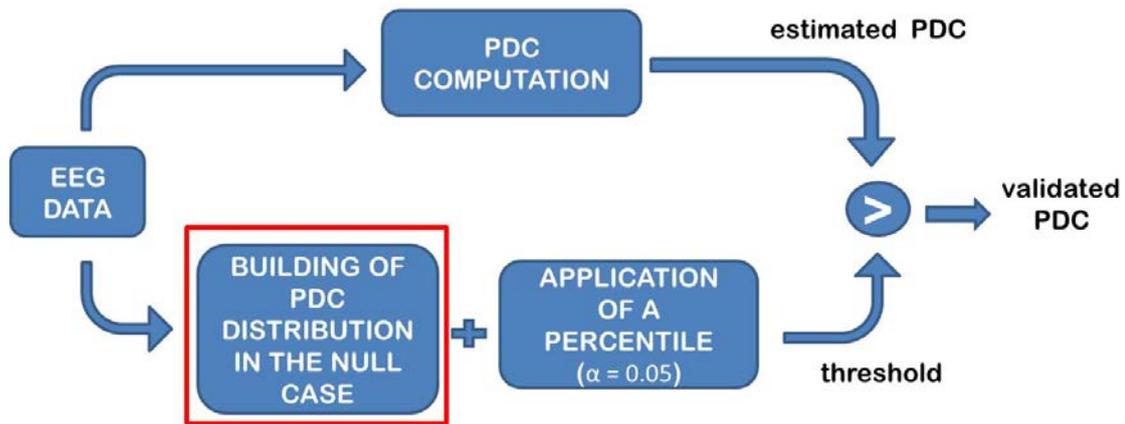


Figure 1.1 – Flow-chart describing the statistical validation process of functional connectivity

The first assessing method was introduced by Schnider for the directed coherence estimator (Schnider et al., 1989) and then was applied to PDC by Sameshima and Baccalà (Sameshima and Baccalà, 1999) with the name of Spectral Causality Criterion (SCC). It consists in the use of the same threshold, set to 0.1, for all the frequencies and for all the couples and directions, below which the connectivity value is considered due to chance. This criterion, which is simplex to be applied and does not require any computation, has been demonstrated corresponding to the application of a percentile of 99% on the null case distribution achieved empirically by estimating functional connectivity on simulated couples of independent white noise (Gourévitch et al., 2006).

In order to improve the accuracy of the validation process, new methods based on the computation of a significance threshold, depending on the data, for each link and for each frequency were developed. The shuffling is a time consuming procedure, introduced in 2001 (Kaminski et al., 2001), which allows to achieve a distribution for the null case by iterating the PDC estimation on different surrogate data sets obtained by shuffling the original traces in order to disrupt the temporal relations between them. The shuffling procedure may involve the phases of the signals in the frequency domain or the samples of data in time domain (Faes et al., 2009). The necessity to reduce the time required for the computation of statistical thresholds in Shuffling procedure, led to the development of a recent validation approach based on the theoretical distribution of the PDC, which tends asymptotically to a χ^2 -distribution in the null case (lack of transmission) (Takahashi et al., 2007; Schelter et al., 2006). A detailed description of the two validation procedures can be found below.

Shuffling Method: Empirical Distribution

Shuffling method is based on the generation of sets of surrogate data (Theiler et al., 1992), obtained shuffling the time series of each channel. Data shuffling can be computed in different ways, although the most used is the phase shuffling. In particular, original data are transformed from time domain to frequency domain by means of Fourier Transform; the data phases are mixed without modifying their amplitude and shuffled signals are taken back in the time domain. This procedure is able to save the amplitude of the power spectrum, but at the same time to disrupt causal links between signals. A MVAR model is fitted to surrogate data set and connectivity estimates are derived from the model. Iterating this process many times, each time on a new surrogate data set, allows to build an empirical distribution of the null hypothesis for the causal estimator (Kamiński et al., 2001). Once obtained the empirical distribution, the significance of the estimated connectivity patterns for a fixed significance level is assessed. In particular, the threshold value, below which the estimated connection is due to the case, is evaluated for each couple of signals and for each frequency by applying a percentile, corresponding to a predefined significance level, on the null-case empirical distribution computed for the correspondent link. The only connections whose values exceed the thresholds are considered as not due to the case.

Asymptotic Statistic Method: Theoretical Distribution

The only way to reduce the computational time required for statistical assessment process and to improve its accuracy is to use the theoretical distribution of FC for the null case instead building its empirical approximation by a time consuming procedure. Recently, Schelter et al. introduced the concept that PDC estimator for the not-null hypothesis is asymptotically normally distributed, while it tends to a χ^2 -distribution in the null case (Schelter et al., 2006). Relying on this statement, it is possible to derive the null-case distribution of PDC from the acquired signals by applying a Monte Carlo method able to reshape the data on a χ^2 -distribution to be used in the assessment process. Details about the method can be found in Takahashi et al. (Takahashi et al., 2007). Once obtained the null case distribution, the procedure of validation is applied as already explained in the previous section for the shuffling method.

Reducing the occurrence of Type I errors in assessment of connectivity patterns

The statistical validation process has to be applied on each couple of signals for each direction and for each frequency sample. This necessity leads to the execution of a high number of simultaneous univariate statistical tests with evident consequences in the occurrence of type I errors. The statistical theory provides several techniques that could be usefully applied in the context of the assessment of connectivity patterns in order to avoid the occurrence of false positives.

The family-wise error rate represents the probability of observing one or more false positives after carrying out simultaneous univariate tests. Supposing to have m null hypotheses H_1, H_2, \dots, H_m . Each hypothesis could be declared significant or non-significant by means of a statistical test. Tab.1.1 summarizes the situation after multiple significance tests are simultaneously applied:

	Null hypothesis is True	Alternative hypothesis is True	Total
Declared significant	V	S	R
Declared non-significant	U	T	$m - R$
Total	m_0	$m - m_0$	m

Table 1.1 – Table explaining the concept of family-wise error rate

where:

- m_0 is the number of true null hypotheses
- $m - m_0$ is the number of true alternative hypotheses
- V is the number of false positives (Type I error)
- S is the number of true positives
- T is the number of false negatives (Type II error)
- U is the number of true negatives
- R is the number of rejected null hypotheses

The FWER is the probability of making even one type I error in the family:

$$FWER = \Pr(V \geq 1) \tag{1.11}$$

Many methodologies are available for preventing type I errors (Nichols and Hayasaka, 2003), but in the following sections I limit the discussion to False Discovery Rate (FDR) and Bonferroni adjustments, which are the most used methodologies in neuroscience field.

False Discovery Rate

The false discovery rate (FDR), suggested by Benjamini and Hochberg (1995) is the expected proportion of erroneous rejections among all rejections. Considering V as the number of false positives and S as the number of true positives, the FDR is given by:

$$FDR = E \left[\frac{V}{V + S} \right] \quad (1.12)$$

where $E[\cdot]$ is the symbol for expected value.

In the following I reported the False Discovery Rate controlling procedure described by Benjamini and Hochberg in 1995. Let H_1, H_2, \dots, H_m be the null hypotheses, with m as the number of univariate tests to be performed, and p_1, p_2, \dots, p_m their corresponding p-values. Let order in ascending order the p-values as $p(1) \leq p(2) \leq \dots \leq p(m)$ and then select the largest i ($i=k$) for which the condition

$$P_{(i)} \leq \frac{i}{m} \alpha \quad (1.13)$$

is verified.

At the end, the hypotheses H_i with $i=1, \dots, k$ have to be rejected.

In the case of independent tests, an approximation for evaluating corrected significance level has been introduced (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001):

$$\beta^* = \frac{(m+1)}{2m} \alpha \quad (1.14)$$

In this case the new level of significance is β^* . Such value guarantees that each test is performed with the imposed significance α .

Bonferroni adjustment

The Bonferroni adjustment (Bonferroni, 1936) starts from the consideration that if I perform N univariate tests, each one of them with an unknown significant probability α , the probability p that at least one of the test is significant is given by (Zar, 2010):

$$p < N\alpha \quad (1.15)$$

In other words this means that if $N = 20$ tests are performed with the usual probability $\alpha = 0.05$, at least one of them became significant statistically by chance alone. However, the Bonferroni adjustment required that the probability p for which this event could occur (i.e., one result will be statistically significant by chance alone) could be equal to α . By using the Eq. (1.15), the single test will be performed at a probability

$$\beta^* = \alpha / N \tag{1.16}$$

This β^* is the actual probability at which the statistical tests are performed in order to conclude that all of the tests are performed at α level of statistical significance, Bonferroni adjusted for multiple comparisons. The Bonferroni adjustment is quite flexible since it does not require the hypothesis of independence of the data to be applied.

Comparing Methods for the Statistical Assessment of Connectivity Patterns: A Simulation Study

The Simulation Study

In order to compare the two methodologies available at the moment for assessing the significance of connectivity patterns, I performed a simulation study in which the performances achieved for the two methods were evaluated under different conditions of SNR and amount of data available for the analysis.

The simulation study was composed by the following steps:

- 1) Generation of several sets of test signals simulating activations at scalp or cortical levels. These datasets were generated in order to fit a predefined connectivity model and to respect imposed levels of some factors. These factors were the SNR (factor SNR) and the total length of the data (factor LENGTH).
- 2) Estimation of the cortical connectivity patterns obtained in different conditions of SNR and data LENGTH by means of sPDCcol (squared PDC normalized according to columns), sPDCrow (squared PDC normalized according to rows) and sPDCnn (squared PDC not normalized). The normalization was applied before and after the validation process (factor NORMTYPE).
- 3) Application of the two different methods, the shuffling and asymptotic statistic procedures (factor VALIDTYPE), for assessing significance of estimated connectivity patterns. The evaluation of significant thresholds in both methods was computed applying a significance level of 0.05 in three different cases: no correction, corrected for multiple comparisons by means of FDR and Bonferroni adjustments (factor CORRECTION).
- 4) Computation of the total percentage of false positives and false negatives occurred in the assessment of significance of connectivity patterns for all the considered factors.
- 5) Statistical analysis of percentage of both false positives and negatives by means of ANOVA for repeated measures in order to evaluate the effects of some factors (SNR, LENGTH, CORRECTION, NORMTYPE) on the performances achieved by means of the two validation methods.

Signal Generation

Different simulated datasets were generated fitting a predefined model, which is reported in Fig. 1.2, composed by 4 cortical areas and imposing different levels of Signal to Noise Ratio (factor SNR) and data length (factor LENGTH).

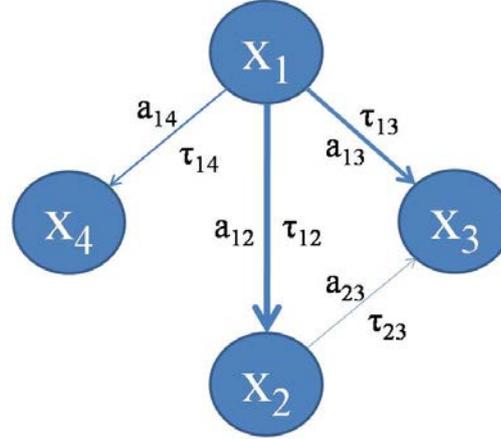


Figure 1.2 - Connectivity model imposed in the generation of testing dataset. x_1, \dots, x_4 represent the signals of four electrodes or cortical regions of interest. a_{ij} represents the strength of the imposed connection between nodes i and j , while τ_{ij} represents the delay in transmission applied between the two signals x_i and x_j in the generation of the dataset. The values chosen for connections strength are $a_{12}=0.5$, $a_{13}=0.4$, $a_{14}=0.2$, $a_{23}=0.08$, while the values set for delays in transmission are $\tau_{12}=10s$, $\tau_{13}=10s$, $\tau_{14}=5s$, $\tau_{23}=20s$ at sampling rate of 200Hz.

$x_1(t)$ is a real signal acquired at the scalp level during a high resolution EEG recording session (61 channels) involving a healthy subject during the rest condition. The other signals $x_2(t)$, ..., $x_4(t)$ were iteratively achieved according to the predefined scheme reported in Fig. 1.2. In particular, the signal $x_j(t)$ is obtained adding uncorrelated Gaussian white noise to all the contributions of other signals $x_i(t)$ (with $i \neq j$), each of which amplified of a_{ij} and delayed of τ_{ij} . The scheme is composed by 3 direct arcs ($1 \rightarrow 2$; $1 \rightarrow 4$; $2 \rightarrow 3$), a direct-indirect arc ($1 \rightarrow 2 \rightarrow 3$) between node 1 and node 3 through node 2 and 8 “null” arcs, i.e. 8 pairs of ROIs which are not linked, not directly nor indirectly. Coefficients for connection strengths used in the imposed model are $a_{12}=0.5$, $a_{13}=0.4$, $a_{14}=0.2$, $a_{23}=0.08$. These values are chosen in a realistic range which is typical of connectivity patterns estimated on data recorded during memory, motor and sensory tasks (Astolfi et al., 2009; Büchel and Friston, 1997). In particular the low value chosen for the connection between nodes 2 and 3 was introduced with the aim to test the accuracy of validation process even for weak connections. The values used for the delay in transmission are $\tau_{12}=2$, $\tau_{13}=2$, $\tau_{14}=1$, $\tau_{23}=4$ data samples which correspond to $\tau_{12}=10$ ms, $\tau_{13}=10$ ms, $\tau_{14}=5$ ms, $\tau_{23}=20$ ms at a sampling rate of 200 Hz.

The procedure of signal generation was repeated under the following conditions:

- SNR factor = [0.1; 1; 3; 5; 10]
- LENGTH factor = [3000; 10000; 20000; 30000] data samples corresponding to a signal length of [15; 50; 100; 150] s, at a sampling rate of 200 Hz.

The levels chosen for both SNR and LENGTH factors cover the typical range for the cortical activity estimated during high resolution EEG experiments.

The MVAR models was estimated by means of Nuttall-Strand method, or multivariate Burg algorithm, which is one of the most common estimators for MVAR models and has been demonstrated to provide the most accurate results (Marple, 1987; Kay, 1988; Schlögl, 2006).

Evaluation of performances

A statistical evaluation of accuracy in assessing significance of estimated connectivity patterns was required to compare the two considered validation methods. The accuracy was quantified by means of two indicators, which are the percentages of false positives and of false negatives committed during the statistical assessment of estimated networks. These percentages were obtained comparing the results of connectivity estimation with the imposed connection scheme.

The percentage of false positives was computed considering the number of nodes pairs between which the estimated connectivity value was significantly different from null case while the imposed value in the predefined model was zero (no connection). In particular the percentage of false positives ($FP_{\%}$) was defined as follows:

$$FP_{\%} = \frac{\sum_{f=f_{start}}^{f_{end}} \sum_{n=1}^{n_{TOT}} K^+(n, f)}{n_{model}^+ \cdot (f_{end} - f_{start})} \quad (1.17)$$

where $[f_{start}, f_{end}]$ is the frequency range in which the PDC was computed, n_{model}^+ is the number of not “null” connections in the imposed model and $K^+(n, f)$ is

$$K^+(n, f) = \begin{cases} 1 \rightarrow (\tilde{a}_{ij}(n, f) - a_{ij}(n, f)) > 0 \\ 0 \rightarrow (\tilde{a}_{ij}(n, f) - a_{ij}(n, f)) \leq 0 \end{cases} \quad (1.18)$$

where $\tilde{a}_{ij}(n, f)$ and $a_{ij}(n, f)$ are the boolean expressions for the estimated and the imposed PDC values respectively. In fact $\tilde{a}_{ij}(n, f)$ is set to 1 when the PDC value for the link n and at frequency f exceeds the significance threshold and is set to 0 when PDC value for the link n and at frequency f is below the statistical threshold, while $a_{ij}(n, f)$ is set to 1 when the

imposed value is different from zero and is set to 0 when the imposed value in the model is zero.

The percentage of false negatives was computed considering the number of nodes pairs between which the estimated connectivity value was not significantly different from null case while its imposed value was different from zero in the predefined model. In particular the percentage of false negatives (FN%) was defined as follows:

$$FN_{\%} = \frac{\sum_{f=f_{start}}^{f_{end}} \sum_{n=1}^{n_{TOT}} K^{-}(n, f)}{n_{model}^{-} * (f_{end} - f_{start})} \quad (1.19)$$

where $[f_{start} f_{end}]$ is the frequency range in which the PDC was computed, n_{model}^{-} is the number of “null” connections in the imposed model and $K^{-}(n, f)$ is

$$K^{-}(n, f) = \begin{cases} 1 \rightarrow (\tilde{a}_{ij}(n, f) - a_{ij}(n, f)) < 0 \\ 0 \rightarrow (\tilde{a}_{ij}(n, f) - a_{ij}(n, f)) \geq 0 \end{cases} \quad (1.20)$$

where $\tilde{a}_{ij}(n, f)$ and $a_{ij}(n, f)$ are the boolean expressions for the estimated and the imposed PDC values respectively, define as describe above.

Both indexes were computed for the two validation methods (factor VALIDTYPE), for different normalization types (factor NORMTYPE) and for a significance level of 0.05 not corrected and corrected with FDR and Bonferroni adjustments (factor CORRECTION).

Simulations were performed by repeating for 100 times each generation-estimation procedure, in order to increase the robustness of the following statistical analysis.

Statistical Analysis

The percentages of false positives and false negatives were subjected to separate ANOVAs. First, I performed a four-way ANOVA aiming at studying the effect of different normalizations applied to PDC spectral estimator (NORMTYPE) on the percentages of false positives and false negatives occurred during estimation process, taking into account several factors such as the SNR, the length of data used for the processing (LENGTH) and the use of different methods of corrections for multiple comparisons (CORRECTION). The within main factors of the ANOVA were NORMTYPE (with five levels: NoNorm \rightarrow no normalization, NormRowPre \rightarrow rows normalization before statistical validation, NormRowPost \rightarrow rows normalization after statistical validation, NormColPre \rightarrow columns normalization before statistical validation, NormColPost \rightarrow columns normalization after statistical validation),

SNR (with five levels: 0.1, 1, 3, 5, 10), LENGTH (with four levels: [3000, 10000, 20000 30000] data samples, corresponding to a signal length of [15, 50, 100, 150] s, at a sampling rate of 200 Hz) and CORRECTION (with three levels: no correction, FDR and Bonferroni). The dependent variables were both the percentages of false positives and false negatives occurred during the validation process. Duncan's pairwise comparisons were then performed in order to better understand the significance between different levels of the same factor or between different factors within the same level. This analysis was executed on the connectivity networks validated applying both Shuffling and Asymptotic Statistic procedures.

The second analysis was a four-way ANOVA aiming at comparing the two validation procedures (VALIDTYPE) in terms of both percentages of false positives and false negatives occurred during estimation process, taking into account different factors such as the SNR, the length of the data subject to the processing (LENGTH) and the use of different methods of corrections for multiple comparisons (CORRECTION). The within main factors of the ANOVA were VALIDTYPE (with two levels: Shuffling and Asymptotic Statistic procedures), SNR (with five levels: 0.1, 1, 3, 5, 10), LENGTH (with four levels: [3000, 10000, 20000 30000] data samples, corresponding to a signals length of [15, 50, 100, 150] s, at a sampling rate of 200 Hz) and CORRECTION (with three levels: no correction, FDR and Bonferroni). The dependent variables were both the percentages of false positives and false negatives occurred during the validation process. Duncan's pairwise comparisons were then performed. This analysis was executed on the connectivity patterns achieved by applying each type of normalization (rows, columns or none).

The third analysis was a four-way ANOVA aiming at comparing the two validation procedures (VALIDTYPE) in terms of percentages of false positives occurred during estimation process focalizing the attention on each link (LINK) and taking into account different factors such as the SNR, the length of data used for the processing (LENGTH) and the use of different methods of corrections for multiple comparisons (CORRECTION). The within main factors of the ANOVAs were LINK (with eight levels: 2→1, 3→1, 4→1, 3→2, 4→2, 4→3, 2→4, 3→4), SNR (with five levels: 0.1, 1, 3, 5, 10), LENGTH (with four levels: [3000, 10000, 20000 30000] data samples, corresponding to a signals length of [15, 50, 100, 150] s, at a sampling rate of 200 Hz) and CORRECTION (with three levels: no correction, FDR and Bonferroni). The dependent variable was the percentage of false positives occurred during the validation process. Duncan's pairwise comparisons were then performed. This

analysis was executed on the connectivity patterns validated applying both Shuffling and Asymptotic Statistic procedure.

		FACTORS	p	F
ASYMPTOTIC STATISTIC	FALSE POSITIVES	NORMTYPE	0.89	0.28
		NORMTYPE x SNR	0.056	1.69
		NORMTYPE x LENGTH	0.9	0.55
		NORMTYPE x CORRECTION	0.84	0.52
	FALSE NEGATIVES	NORMTYPE	0.59	0.69
		NORMTYPE x SNR	0.11	1.46
		NORMTYPE x LENGTH	0.94	0.46
		NORMTYPE x CORRECTION	0.11	1.64
SHUFFLING	FALSE POSITIVES	NORMTYPE	< 0.00001	1232
		NORMTYPE x SNR	< 0.00001	169.21
		NORMTYPE x LENGTH	< 0.00001	47.05
		NORMTYPE x CORRECTION	< 0.00001	427.46
	FALSE NEGATIVES	NORMTYPE	< 0.00001	5406.2
		NORMTYPE x SNR	< 0.00001	473.24
		NORMTYPE x LENGTH	< 0.00001	458.5
		NORMTYPE x CORRECTION	< 0.00001	1202.9

Table 1.2 – Results of the ANOVA performed on the percentages of false positives and false negatives occurred during the validation process executed by means of Asymptotic Statistic and Shuffling methods. The within main factors were the normalization type (NORMTYPE), the signal to noise ratio (SNR), the data length (LENGTH) and the statistical corrections used for imposing the significance level (CORRECTION)

Results

Several sets of signals were generated as described in the previous section, in order to fit the connectivity pattern shown in Fig. 1.2. The imposed connectivity model contained three indirect arcs, one direct-indirect arc between nodes 1 and 3 through node 2 characterized by a

weak connection between nodes 2 and 3, and eight “null” arcs. This topology was chosen in order to test the performances of estimation/validation process in:

- preventing estimates of nonexistent arcs (false positives)
- avoiding inversions of direct paths (false positives)
- discarding existent links (false negatives)
- distinguishing real weak connections from links due to the case (false negatives)

A MVAR model of order 16 was fitted to each set of simulated data, which were in the form of several trials of the same length (1s). The procedure of signal generation, estimation of connectivity patterns by means of sPDCnn, sPDCcol, sPDCrow and statistical validation for not corrected and adjusted statistical significance level 0.05 was carried out 100 times for each level of factors SNR and LENGTH, as indicated in the previous paragraphs, in order to increase the robustness of the subsequent statistical analysis.

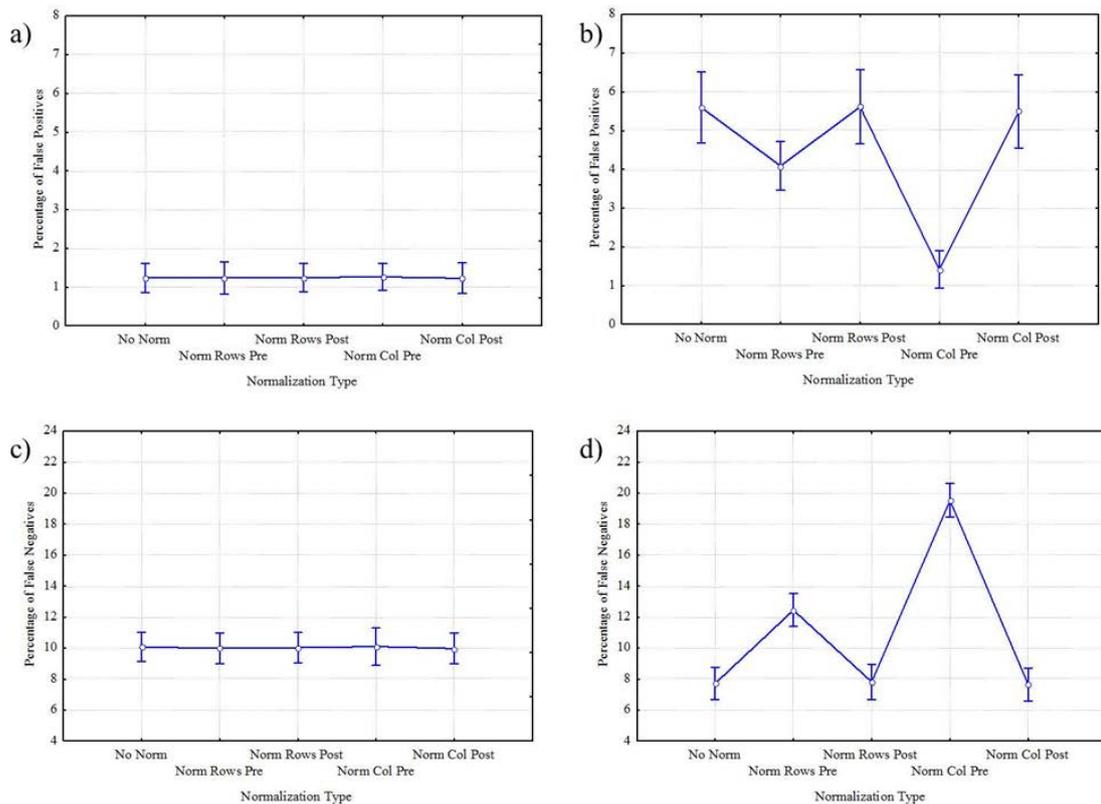


Figure 1.3 - Results of ANOVA performed on the percentages of false positives (first row) and false negatives (second row) occurred applying Asymptotic Statistic (first column) and Shuffling procedure (second column) respectively, using NORMTYPE as within main factor. The diagram shows the mean value for the percentages obtained not normalizing and normalizing according to rows and columns, before and after the validation process the inferred connectivity patterns. The bar represented their relative 95% confidence interval.

The indexes of performances used were the percentages of false positives and false negatives. They were computed for each generation-estimation procedure performed, and then subjected to different ANOVAs.

A four-way ANOVA was applied separately to the percentages of false positives and false negatives occurred during the validation process performed by means of Asymptotic Statistic and Shuffling procedures. The within main factors were the normalization type (NORMTYPE), the signal to noise ratio (SNR), the data length (LENGTH) and the statistical corrections used for imposing the significance level (CORRECTION). The corresponding results were reported in Tab.1.2. In particular they revealed no statistical influence of the main factor NORMTYPE on percentages of false positives and false negatives committed using Asymptotic Statistic validation procedure. Although high statistical influence of the main factors NORMTYPE and its interactions NORMTYPE x SNR, NORMTYPE x LENGTH, NORMTYPE x CORRECTION resulted on percentages of both false positives and false negatives occurred applying Shuffling validation method.

In Fig. 1.3 I reported results of ANOVA performed on the percentages of false positives (first row) and false negatives (second row) occurred with the application of Asymptotic Statistic (first column) and Shuffling procedure (second column) respectively, using NORMTYPE as within main factor. The diagram shows the mean value for the percentages obtained not normalizing and normalizing according to rows and columns, before and after the validation process, the inferred connectivity patterns. The bar represented their relative 95% confidence interval. Diagrams in the two columns of Fig. 1.3 revealed different behavior of the two validation methods in discarding both false positives and false negatives during the statistical assessment procedure in relation to the applied normalization type. In fact, panel a and c showed no significant differences, for Asymptotic Statistic method, between the normalization types in the percentages of type I and type II errors, which remained strictly around 1.2% and 10% respectively, as confirmed by Duncan's pairwise comparisons. Panel b and d revealed a strong influence of the normalization type on the occurrence of type I and type II errors during the application of Shuffling method. In particular, in panel b, the percentage of false positives remained within the range [4% – 6%] except for the normalization according to columns applied before the validation procedure, for which the percentage highly decreased down to 1.5%. Duncan's pairwise comparisons highlighted significant differences ($p < 0.00001$) between the normalization types, with exceptions of the cases no normalization, normalization according to row and according to column both after

the validation method. In panel d, the percentage of false negatives ranged between 7.5% and 12% except for the case Normalization according to columns, for which the value reached the 20%. Duncan's pairwise comparisons highlighted significant differences ($p < 0.00001$) between the normalization types, with exceptions of the cases no normalization, normalization according to row and according to column both executed after the validation method.

FACTORS	FALSE POSITIVES		FALSE NEGATIVES	
	p	F	p	F
VALIDTYPE	< 0.00001	3996	< 0.00001	625
SNR	< 0.00001	762	< 0.00001	117
LENGTH	< 0.00001	10	< 0.00001	11300
CORRECTION	< 0.00001	6598	< 0.00001	10200
VALIDTYPE x SNR	< 0.00001	549	< 0.00001	206
VALIDTYPE x LENGTH	< 0.00001	7	< 0.00001	35
VALIDTYPE x CORRECTION	< 0.00001	248	< 0.00001	216
SNR x LENGTH	< 0.00001	7	< 0.00001	232
SNR x CORRECTION	< 0.00001	85	< 0.00001	260
LENGTH x CORRECTION	< 0.00001	26	< 0.00001	818
VALIDTYPE x SNR x CORRECTION	< 0.00001	34	< 0.00001	26
VALIDTYPE x LENGTH x CORRECTION	< 0.00001	33	< 0.00001	158

Table 1.3 – Results of the ANOVA computed considering the percentages of false positives and false negatives occurred during the validation processes as dependent variables respectively and as within main factors the validation type (VALIDTYPE), the SNR, the data length (LENGTH) and the corrections used for multiple comparisons (CORRECTION).

In order to investigate the effects of different factors, such as SNR and data length, on the quality of validation process, I computed two separate ANOVAs using the percentages of false positives and false negatives occurred during the validation processes executed through Shuffling and Asymptotic Statistic procedure as dependent variables respectively. In both ANOVAs the within main factors were the validation type (VALIDTYPE), the SNR, the data length (LENGTH) and the corrections used for multiple comparisons (CORRECTION).

Such analysis has been applied on connectivity patterns estimated by means of sPDCnn, sPDCrow, sPDCcol. I reported in Tab.1.3 the results obtained by using normalization

according to rows before validation process, but the same trends could be noted for the other normalization types.

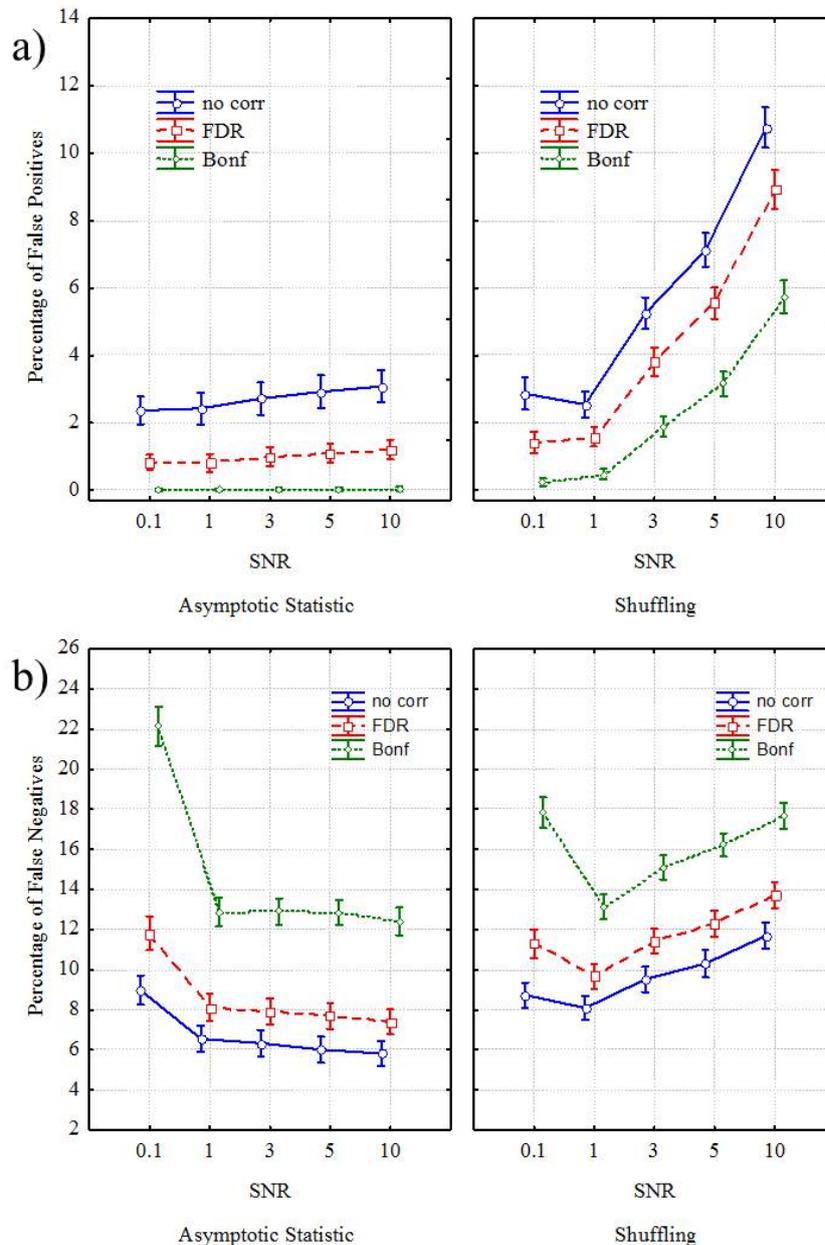


Figure 1.4 - Results of ANOVA performed on the percentages of false positives (a) and false negatives (b) occurred during the validation procedure, using VALIDTYPE, SNR and CORRECTION as within main factors. The diagram shows the mean value for the percentages, achieved by normalizing according to rows the inferred connectivity patterns before the validation process, for different values of SNR and for different correction types (no correction in blue, FDR in red and Bonferroni in green). The bar represents their relative 95% confidence interval.

In Fig.1.4 I reported the results of ANOVA performed on the percentages of false positives (panel a) and false negatives (panel b) occurred during the validation procedure, using VALIDTYPE, SNR and CORRECTION as within main factors. The diagram shows the mean value of the two performance indexes achieved on connectivity patterns normalized according

to rows before the validation process, for different values of SNR and for different correction types (no correction in blue, FDR in red and Bonferroni in green) used for preventing type I errors in the statistical assessment procedure. The bar represented their relative 95% confidence interval.

Panel a highlighted statistical differences between the two validation methods in the occurrence of false positives across different SNR levels. The percentage of type I errors, achieved by applying Asymptotic Statistic procedure, for the SNR ranging from 0.1 to 10, increased from 2% to 4% for the no correction case, but remained constant around 1% for FDR case and 0.1% for Bonferroni adjustment. Post hoc analysis revealed statistical differences ($p < 0.00001$) between all the SNR levels in the no correction case and with the exceptions of SNR 1 and 3 in FDR case, although this differences disappeared completely in Bonferroni case. The percentage of false positives, achieved by applying Shuffling procedure, was at 3% for SNR 0.1, decreased to 2% for SNR 1, then increased up to 11% for SNR 10 in the no correction case. The same trend but with lower percentages resulted for FDR and Bonferroni cases where the percentages increased up to 9% and 6% respectively along SNR levels. Duncan's pairwise comparisons revealed statistical differences ($p < 0.00001$) along all the levels of SNR for the three CORRECTION cases. The differences between the two validation types in the occurrence of false positives are statistical significant across all the SNR imposed and all the corrections applied, as stated by the Duncan's pairwise comparisons.

Panel b showed statistical differences between the two validation types in the occurrence of type II errors across different SNR levels. In particular in the first diagram of panel b, for SNR values ranging from 0.1 to 1, the percentage of false negatives achieved applying Asymptotic Statistic procedure, decreased from 22% to 13% for the Bonferroni correction, from 12% to 8% for FDR adjustment and from 9% to 6% for no correction case. For SNR values ranging from 1 to 10 the percentages of type II errors remained constant around 13% for Bonferroni correction, 8% for FDR and 6% for no correction case. Duncan's pairwise comparisons revealed statistical differences ($p < 0.00001$) between all the SNR levels in all the three CORRECTION cases. In the second diagram of panel b, for SNR values ranging from 0.1 to 1, the percentage of false negatives reached applying Shuffling procedure, decreased from 18% to 13% for the Bonferroni correction, from 11% to 10% for FDR adjustment and from 9% to 8% for no correction case. For SNR values ranging from 1 to 10 the percentage of type II errors moved from 13% to 18% for Bonferroni correction, from 10% to 14% for FDR

and from 8% to 12% for no correction case. Pairwise comparisons revealed statistical differences ($p < 0.00001$) between all the SNR levels in all the three CORRECTION cases. The differences between the two validation types in the occurrence of false negatives were statistical significant across all the SNR and all the corrections applied, as stated by the Duncan's pairwise analysis.

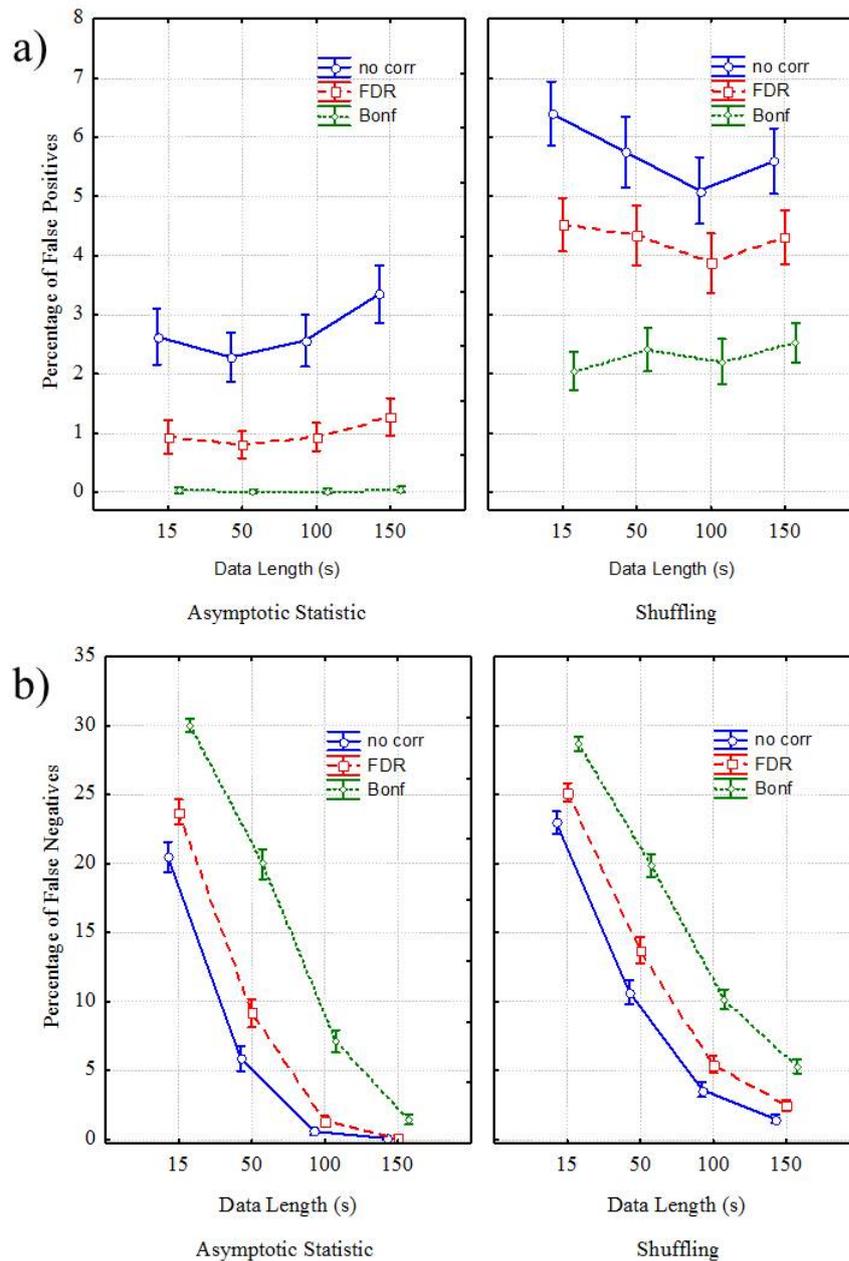


Figure 1.5 - Results of ANOVA performed on the percentages of false positives (a) and false negatives (b) occurred during the validation procedure, using VALIDTYPE, LENGTH and CORRECTION as within main factors. The diagram shows the mean value for the percentages, achieved by normalizing according to rows the inferred connectivity patterns before the validation process, for different values of data length (in ms) and for different correction types (no correction in blue, FDR in red and Bonferroni in green). The bar represents their relative 95% confidence interval.

In Fig.1.5 I showed results of ANOVA performed on the percentages of false positives (panel a) and false negatives (panel b) occurred during the validation procedure, using

VALIDTYPE, LENGTH and CORRECTION as within main factors. The diagram shows the mean value for both indexes, achieved on connectivity patterns normalized according to rows before the validation process, for different values of data length (in ms) and for different correction types (no correction in blue, FDR in red and Bonferroni in green). The bar represented their relative 95% confidence interval.

Panel a highlighted statistical differences between the two validation methods in the occurrence of false positives across different data lengths. In fact, in the first diagram of panel a, the percentage of type I errors, achieved by applying Asymptotic Statistic procedure, for the data length ranging from 15s to 150s, increased from 2.5% to 3.5% for the no correction case, from 1% to 1.5% for FDR case but remained constant around 0.1% for Bonferroni adjustment. Post hoc analysis revealed statistical differences ($p < 0.00001$) between all the different data lengths in the no correction case and with the exceptions of 50s and 100s in FDR case, although this differences disappeared completely in Bonferroni case. In the second diagram of panel a, the percentage of false positives, achieved by applying Shuffling procedure, for data length from 15s to 100s, moved from 6.5% to 5% for no correction case, from 4.5% to 4% for FDR case and remained around 2.5% for Bonferroni adjustment. For data length ranging from 100s to 150s, the percentages of false positives increased from 5% to 5.5% for no correction case, from 4% to 4.5% for FDR case and from 2 to 2.5% for Bonferroni adjustment. Duncan's pairwise comparisons revealed statistical differences ($p < 0.00001$) across all the different data lengths for all the three CORRECTION cases. The differences between the two validation types in the occurrence of false positives are statistical significant across all the data lengths and all the corrections applied, as stated by the Duncan's test.

Panel b highlighted statistical differences between the two validation methods in the occurrence of false negatives across different data lengths. In fact, in the first diagram of panel b, the percentage of type II errors, achieved by applying Asymptotic Statistic procedure, for the data length ranging from 15s to 150s, decreased from 30% to 2% for Bonferroni adjustment, from 25% to 1% for FDR, from 20% to 0% for the no CORRECTION case. Post hoc analysis revealed statistical differences ($p < 0.00001$) between all the different data lengths in all the three correction cases. In the second diagram of panel b, the percentage of false negatives, achieved by applying Shuffling procedure, for data length from 15s to 150s, decreased from 28% to 5% for Bonferroni adjustment, from 25% to 2.5% for FDR case and from 23% to 2% for no correction case. Duncan's analysis revealed statistical differences

($p < 0.00001$) across all the different data lengths for all the three CORRECTION cases. The differences between the two validation types in the occurrence of false negatives are statistically significant across all the data lengths and all the corrections applied, as stated by the Duncan's pairwise comparisons.

In order to understand which type of links are responsible for the occurrence of type I errors, a new ANOVA was applied using the percentage of false positives as dependent variable and LINK, SNR and LENGTH as within main factors. I considered only the links corresponding to the imposed 'null' arcs because type I errors could occur only in their correspondence. A specific ANOVA was applied for each validation procedure and for each correction type. For the brevity of the text I reported only results achieved applying both Asymptotic Statistic and Shuffling methods and imposing a significance level FDR corrected, but the same conclusions can be inferred even for the other two correction cases. Results showed high statistical influence of the main factors LINK ($P < 0.00001$, $F = 15.33$) and LINK x SNR ($P < 0.00001$, $F = 7.33$) on percentage of false positives occurred during the Asymptotic Statistic procedure and high statistical influence of the main factors LINK ($P < 0.00001$, $F = 2346$), LINK x LENGTH ($P < 0.00001$, $F = 16$), LINK x SNR ($P < 0.00001$, $F = 270$) on percentage of false positives occurred during the Shuffling procedure.

In Fig.1.6 I reported results of ANOVA performed on the percentage of false positives, occurred during the validation procedure computed by Asymptotic statistic (panel a) and Shuffling (panel b) methods, using LINK as within main factor. The diagram shows the mean value for the percentage achieved on each considered arc of the networks and its relative 95% confidence interval. Panel a showed the percentage of false positives, occurred during the Asymptotic Statistic procedure. In particular, such percentage was around 3.5% for links $2 \rightarrow 1$, $3 \rightarrow 1$ and $4 \rightarrow 1$ and around 2% for links $3 \rightarrow 2$, $4 \rightarrow 2$, $4 \rightarrow 3$, $2 \rightarrow 4$, $3 \rightarrow 4$. Duncan's pairwise comparisons confirmed statistical differences ($p < 0.00001$) between the links $2 \rightarrow 1$, $3 \rightarrow 1$ and $4 \rightarrow 1$ each other and with the other considered links.

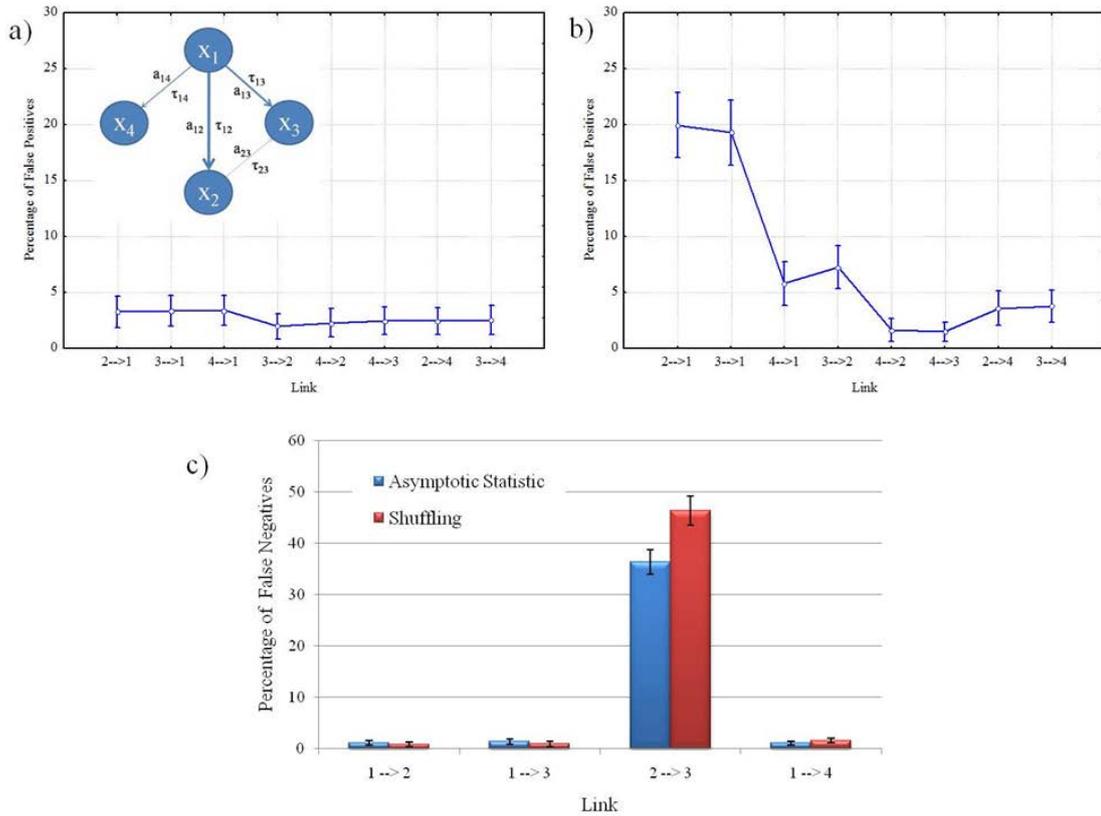


Figure 1.6 – (a) and (b) Results of ANOVA performed on the percentage of false positives, occurred during the validation procedure computed through Asymptotic statistic (a) and Shuffling (b) methods, using LINK as within main factor. I considered only the links corresponding to the imposed ‘null’ arcs because they can be origin of type I errors. The diagram shows the mean value for the percentage of false positives on each considered arc of the networks and their relative 95% confidence interval. (c) The bar diagram shows the mean value, evaluated on different data lengths (in ms) and different SNR, for the percentages of type II errors, achieved by normalizing according to rows the inferred connectivity patterns before the validation process and validating them through Asymptotic statistic (blue columns) and Shuffling (red columns) methods for a significance level of 5% FDR corrected.

Panel b showed a percentage of false positives, occurred during Shuffling validation process. In particular, such percentage was around 20% for links $2 \rightarrow 1$, $3 \rightarrow 1$, around 6% for link $4 \rightarrow 1$, around 7.5% for link $3 \rightarrow 2$ and from 1% to 4% for links $4 \rightarrow 2$, $4 \rightarrow 3$, $2 \rightarrow 4$, $3 \rightarrow 4$. Duncan’s analysis confirmed statistical differences ($p < 0.00001$) between the links $2 \rightarrow 1$, $3 \rightarrow 1$, $4 \rightarrow 1$ and $3 \rightarrow 2$ each other and with the other considered links.

The bar diagram in Fig.1.6c showed the mean value, computed across different data lengths and SNR values, for the percentages of type II errors, achieved on connectivity patterns normalized according to rows before the validation process and validated through Asymptotic Statistic (blue columns) and Shuffling (red columns) for a significance level of 5% FDR corrected. The percentage of false negative is almost due to the link $2 \rightarrow 3$ whose imposed connection strength was really weak.

Discussion

In order to evaluate the performances of the two methods in the assessment of functional connectivity patterns estimated on non-invasive EEG measurements, I performed several statistical analysis (ANOVAs and Duncan post-hoc tests) on different simulated datasets, built according to a predefined scheme, under different conditions of SNR and data length whose variability ranges are similar to those found during real EEG recordings. The imposed connectivity model was built in order to test the validation processes in some situations typical of cerebral networks such as the presence of direct, indirect and ‘null’ arcs with strengths in ranges consistent with those estimated in previous studies characterized by a large sample of subjects performing memory, motor, and sensory tasks. In particular I evaluated the capacity of such assessing methods in i) correctly estimating the existent connections; ii) preventing the inversion of existent links and iii) avoiding the deletion of real weak connections.

Thus, the occurrence of both type I and type II errors was evaluated for the two validation methods in relation to the type of estimator chosen for the connectivity estimation (type of normalization, before or after the validation process), to the SNR level, to the length of data and to the statistical corrections applied on the significance level imposed during the validation process computed with the two methods.

The simulations provided the following answers to the questions raised in the Introduction:

- 1) The Shuffling procedure is a quite reliable method in preventing the occurrence of both false positives and false negatives if applied together with methods for preventing the occurrence of type I errors. In fact, considering the FDR correction for multiple comparisons, the percentages of false positives remained around 5% for all the considered data lengths and for SNR values below 3. The percentages of false negatives exceeded 5% for all the SNR and data length values considered.
- 2) The application of corrections for multiple comparisons improved the performances of Shuffling procedure in preventing type I errors but increased the percentage of false negatives. In fact, the Bonferroni correction led the percentage of false positives around 5% for all the SNR and all data lengths but considerably increased the percentages of false negatives especially for low values of SNR (below 1) and short data lengths.
- 3) The high performances in terms of percentages of false positives and false negatives of Asymptotic Statistic procedure has been already highlighted in (Toppi et al., 2011).

The comparison between the recent Asymptotic Statistic method and the already existent Shuffling procedure performed in this study demonstrated that Asymptotic Statistic method can substitute the Shuffling approach in the validation of estimated connectivity patterns due to its higher performances in preventing type I and type II errors. In fact the percentages of false positives, without applying any corrections, were already below 5% for all the considered SNR and data values. The percentages of false negatives were below 5% for SNR values above 1 and data length above 50s. The application of corrections for multiple comparisons, in particular of Bonferroni adjustment, reduced the percentage of false positives but highly increased the number of false negatives.

- 4) The definition of PDC estimator affected only the percentages of type I and type II errors occurred by using Shuffling procedure. In particular for normalization according to rows and columns both before the validation process resulted low percentages of false positives and high percentages of false negatives. No effects of the PDC formulation on the performances achieved during the validation process executed by means of Asymptotic Statistic.
- 5) The data length has a reliable effect on percentages of type I and type II errors. In fact percentages below 5% can be achieved for a number of samples above 15s and 50s for Asymptotic Statistic and Shuffling methods respectively. The SNR has a reliable effect on the performances of the validation process executed by means of Asymptotic Statistic. In fact for SNR above 1 the performances remained around 5%. Conversely the SNR has not a reliable effect on the performances of the validation process executed by means of Shuffling procedure. High values of SNR increased the number of type I errors but decrease the number of type II errors occurred for Shuffling method. SNR between 1 and 3 seems to be the right condition for applying Shuffling procedure.

In conclusion, the ANOVA results (integrated with the Duncan's test performed at $p < 0.05$) indicated a clear influence of different levels of the main factors SNR and LENGTH on the efficacy of the validation process of connectivity patterns via Shuffling and Asymptotic Statistic. The trends are in agreement with previous simulation studies performed on the effect of factors such as SNR level and data length on the estimation error (Astolfi et al., 2006). The new method for the assessment of PDC significance is a valid tool for the validation of connectivity patterns. In fact, considering SNR and LENGTH values largely met, for instance, in EEG recordings, the occurrence of type I errors is below 6% for all the three

CORRECTION levels. The shuffling procedure showed high percentages of false positives for high values of SNR and small amount of data. Such false positives were due to the inversion of existent links imposed in the model, as highlighted in a specific analysis computed on the percentages of type I errors occurred for each link.

High percentages of type II errors resulted for both validation methods for SNR values below 1 and data length below 100s. However, it must be noted that the presence of a very weak connection in the model imposed to simulated data (2->3) could be responsible of the increase of the number of false negatives.

As expected, as the severity of correction method increased (from no corrections to FDR and Bonferroni) the percentage of false positives is reduced but the percentage of false negatives is increased. In particular, the FDR method seems to provide the best compromise in preventing both type I and type II errors.

In conclusion, the estimation of connectivity patterns on high quality data (good SNR or huge amount of samples) can assure low percentages of both type I and type II errors even without considering severe statistical corrections such as Bonferroni. If the data are characterized by low SNR or signals length, statistical corrections are requested for controlling the estimation performances in terms of type I errors. However, it should be taken into account that this could lead to a loss of weaker connections.

Conclusion

The necessity to obtain stable and reliable connectivity patterns has led in the last ten years to the development of different assessing methods able to discard all the spurious links estimated by chance in the network. The first proposed approaches allowed to empirically assess the inferred connectivity patterns by means of time-consuming procedures based on the empirical reconstruction of PDC distribution in the null case. Recent advancements in the effective connectivity field allowed to develop a new approach based on the theoretical description of the properties of PDC estimator which is characterized by a normal distribution in the not null case and tends to a χ^2 distribution for the null case.

Although all the advancements achieved in the field, a complete characterization of the two most used approaches, the “empirical” Shuffling procedure and the “theoretical” Asymptotic Statistic method have not been provided so far. For this reason, in this section of my thesis I reported the results of a simulation study aimed at characterizing the performances of the newly introduced asymptotic statistics compared to those of the previous shuffling procedure, in terms of false positives and false negatives, under different conditions of SNR and amount of data available for the estimate. Moreover, I included corrections for multiple comparisons in the validation process, demonstrating that the False Discovery Rate represents a good compromise in preventing both type I and type II errors.

The comparison between the two assessing methods allowed to characterize the application field of the two approaches. Moreover, it confirmed the validity of the new asymptotic statistic method in preventing both false positives and false negatives under different conditions of data quality.

Notwithstanding the good results achieved by the simulation study, it is necessary to take into account that in such study the linear approach was used both for generating the surrogate data and for estimating the connectivity pattern. This means that the study took into account only estimation errors without evaluating the occurrence of errors in the estimate due to the choice of a simple linear model for modeling a complex phenomenon. In the future, similar results should be achieved on more complex data models including the nonlinearity of the investigated phenomena.

The characterization of the two assessing methods and the introduction of multiple comparisons in the statistical assessment of connectivity patterns allow to increase the

stability in the estimation procedure, solving one of the open problems in the connectivity field. However, further investigations in that direction are required to allow comparisons between connectivity patterns associated to different experimental conditions (task vs baseline). In fact, the statistical assessment process allows only to discard the connections due to chance, without providing any tool for statistically comparing different investigated patterns.

The development of new methods for comparing patterns associated to two conditions and the refining of existent procedure based on the bootstrapping approach are necessary in the future for investigating more complex experimental situations. Such advancement should be provided starting from what it was learned until now about the statistical assessment.

Section I References

- Akaike, H., 1974. A new look at statistical model identification. *IEEE Trans Automat Control* 19, 716–723.
- Astolfi, L., Cincotti, F., Mattia, D., Marciani, M.G., Baccalà, L.A., De Vico Fallani, F., Salinari, S., Ursino, M., Zavaglia, M., Babiloni, F., 2006. Assessing cortical functional connectivity by partial directed coherence: simulations and application to real data. *IEEE Trans Biomed Eng* 53, 1802–1812.
- Astolfi, L., Cincotti, F., Mattia, D., Marciani, M.G., Baccala, L.A., De Vico Fallani, F., Salinari, S., Ursino, M., Zavaglia, M., Ding, L., Edgar, J.C., Miller, G.A., He, B., Babiloni, F., 2007. Comparison of different cortical connectivity estimators for high-resolution EEG recordings. *Hum Brain Mapp* 28, 143–157.
- Astolfi, L., De Vico Fallani, F., Cincotti, F., Mattia, D., Marciani, M.G., Salinari, S., Sweeney, J., Miller, G.A., He, B., Babiloni, F., 2009. Estimation of effective and functional cortical connectivity from neuroelectric and hemodynamic recordings. *IEEE Trans Neural Syst Rehabil Eng* 17, 224–233.
- Baccalà, L.A., 2001. On the efficient computation of partial coherence from multivariate autoregressive model. *Proceed 5th World Conf Cybernetics Systemics Informatics*.
- Baccalà, L.A., Sameshima, K., 2001. Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics* 84, 463–474.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 289–300.
- Benjamini, Y., Yekutieli, D., 2001. The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics* 29, 1165–1188.
- Blinowska, K.J., 2011. Review of the methods of determination of directed connectivity from multichannel data. *Med Biol Eng Comput* 49, 521–529.
- Blinowska, K.J., Kaminski, M., Kaminski, J., Brzezicka, A., 2010. Information processing in brain and dynamic patterns of transmission during working memory task by the SDTF function. *Conf Proc IEEE Eng Med Biol Soc 2010*, 1722–1725.

- Blinowska, K.J., Kuś, R., Kamiński, M., 2004. Granger causality and information flow in multivariate processes. *Phys Rev E Stat Nonlin Soft Matter Phys* 70, 050902.
- Bonferroni, C., 1936. *Teoria statistica delle classi e calcolo delle probabilità*. Pubblicazioni dell'Istituto Superiore di Scienze Economiche e Commerciali di Firenze.
- Büchel, C., Friston, K.J., 1997. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex* 7, 768–778.
- David, O., Cosmelli, D., Friston, K.J., 2004. Evaluation of different measures of functional connectivity using a neural mass model. *Neuroimage* 21, 659–673.
- Faes, L., Porta, A., Nollo, G., 2009. Surrogate data approaches to assess the significance of directed coherence: application to EEG activity propagation. *Conf Proc IEEE Eng Med Biol Soc 2009*, 6280–6283.
- Gourévitch, B., Bouquin-Jeannès, R.L., Faucon, G., 2006. Linear and nonlinear causality between signals: methods, examples and neurophysiological applications. *Biol Cybern* 95, 349–369.
- Granger, C.W.J., 1969. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica* 37, 424–438.
- Horwitz, B., 2003. The elusive concept of brain connectivity. *NeuroImage* 19, 466–470.
- Kaminski, M., Ding, M., Truccolo, W.A., Bressler, S.L., 2001. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological Cybernetics* 85, 145–157.
- Kamiński, M., Ding, M., Truccolo, W.A., Bressler, S.L., 2001. Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance. *Biol Cybern* 85, 145–157.
- Kaminski, M.J., Blinowska, K.J., 1991. A new method of the description of the information flow in the brain structures. *Biol. Cybern.* 65, 203–210.
- Kay, M.S., 1988. *Modern spectral estimation*. Prentice Hall, Englewood Cliffs, NJ.
- Kus, R., Kaminski, M., Blinowska, K.J., 2004. Determination of EEG activity propagation: pair-wise versus multichannel estimate. *IEEE Transactions on Biomedical Engineering* 51, 1501–1510.

- Lee, L., Harrison, L.M., Mechelli, A., 2003. A report of the functional connectivity workshop, Dusseldorf 2002. *NeuroImage* 19, 457–465.
- Marple, S.L., 1987. *Digital spectral analysis with applications*. Prentice Hall, Englewood Cliffs, NJ.
- Nichols, T., Hayasaka, S., 2003. Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat Methods Med Res* 12, 419–446.
- Sameshima, K., Baccalá, L.A., 1999. Using partial directed coherence to describe neuronal ensemble interactions. *Journal of Neuroscience Methods* 94, 93–103.
- Schelter, B., Winterhalder, M., Eichler, M., Peifer, M., Hellwig, B., Guschlbauer, B., Lücking, C.H., Dahlhaus, R., Timmer, J., 2006. Testing for directed influences among neural signals using partial directed coherence. *Journal of Neuroscience Methods* 152, 210–219.
- Schlögl, A., 2006. A comparison of multivariate autoregressive estimators. *Signal processing* 86, 2426–2429.
- Schnider, S.M., Kwong, R.H., Lenz, F.A., Kwan, H.C., 1989. Detection of feedback in the central nervous system using system identification techniques. *Biol. Cybern.* 60.
- Takahashi, D. Y., Baccalà, L. A., Sameshima, K., 2007. Connectivity Inference between Neural Structures via Partial Directed Coherence. *Journal of Applied Statistics*, vol. 34, no. 10, pp. 1259–1273.
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., Doynne Farmer, J., 1992. Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena* 58, 77–94.
- Toppi, J., Babiloni, F., Vecchiato, G., Cincotti, F., De Vico Fallani, F., Mattia, D., Salinari, S., Astolfi, L., 2011. 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.
- Turbes, C.C., Schneider, G.T., Morgan, R.J., 1983. Partial coherence estimates of brain rhythms. *Biomed Sci Instrum* 19, 97–102.
- Zar, J.H., 2010. *Biostatistical analysis*. Prentice Hall.

Section II

Time-Varying Approach for the Estimation of Functional Connectivity

INTRODUCTION

ESTIMATION OF TIME-VARYING FUNCTIONAL CONNECTIVITY: STATE OF THE ART

Adaptive Partial Directed Coherence

The Recursive Least Square

The General Linear Kalman Filter

ESTIMATION OF TIME-VARYING FUNCTIONAL CONNECTIVITY: A SIMULATION STUDY

The Simulation Study

Signal Generation

Evaluation of performances

Statistical Analysis

Results

Discussion

CONCLUSION

SECTION II REFERENCES

Introduction

All the MVAR based methodologies for the functional connectivity estimation require the hypothesis of stationarity of the signals included. Thus, the temporal dynamics of the influences between cerebral areas are completely lost. To overcome this limitation, different algorithms for the estimation of MVAR models with time dependent coefficients were recently developed (Blinowska, 2011). In particular, the first proposed methodology was based on a short-time window approach, assuming the stationarity of signals in short time intervals (Ding et al., 2000). In 2001 the multi-trial Recursive Least Square (RLS) method with Forgetting Factor was introduced. It consisted in the adaptive estimation of the MVAR

model by means of a recursive algorithm involving a weighted influence of the past of the investigated signals (Astolfi et al., 2008; Möller et al., 2001). However, even if the RLS overcomes the limits of the short time approach, providing a more accurate estimation of the dynamics of non-stationary data, due to its computational complexity it presents a limitation in the number of signals to be considered at the same time in the estimation (Milde et al., 2010; Weiss et al., 2008). Since few years ago, the problem of the model dimension has been solved by reducing the number of electrodes time series to be included in the model (Zhu et al., 2011; Milde et al., 2010; Weiss et al., 2008) or by using cortical waveforms derived for some regions of interest from high resolution EEG data (Astolfi et al., 2008). However, the need to reduce the model dimension introduces a significant error due to the “hidden source dilemma”. In fact, each time a relevant source of information of the problem is removed from the autoregressive model, spurious connectivity links are introduced in the estimation, degrading the reconstruction of the connectivity network. Moreover, the a priori selection of sources, both at scalp or cortical levels, could not be applied every time the dynamics at the basis of the investigated cerebral processes are not completely known and the selection of a subgroup of areas mostly involved in the investigated task is not so obvious. In 2010, a new method based on the Kalman filter, the General Linear Kalman Filter (GLKF), was provided as good alternative to RLS in the description of temporal evolution of information flows between the nodes of a network and as a solution to the limitation in the number of signals to be considered simultaneously in estimation process (Milde et al., 2010).

Even if the RLS method for time varying functional connectivity estimation has been proposed in literature from 2001, its performances under different conditions of signal-to-noise ratio (SNR) and amount of data available were studied only on datasets composed by a small number of signals (Astolfi et al., 2008). The GLKF method has been introduced and qualitatively compared with the RLS by means of a simulated time varying MVAR of dimension 20 (Milde et al., 2011). Rigorous demonstrations of RLS limitation in the number of signals to be considered simultaneously in the estimation and a description of the performances achieved by the GLKF under different conditions of SNR and number of trials have not been performed before my PhD. In fact, a comparison between the two methods in terms of estimation accuracy and speed of adaptation to transitions is necessary to understand their limitations and the conditions (SNR, number of trials, adaptation parameters) in which each method can be used.

In the present Section of my thesis, first of all I briefly described the two methodologies used, at the state of the art, for estimating the temporal evolution of connectivity patterns achieved on non-invasive EEG measurements, highlighting both their advantages and limits. Then I introduced the simulation study performed for comparing the two approaches in terms of accuracy in the estimation process and of speed in the adaptation to the temporal evolution of the estimated patterns, under different conditions of number of nodes included in the network (factor NODE), SNR (factor SNR), amount of trials to be used in the analysis (factor TRIAL) and adaptations constants set for applying the methods (factor COST). Such constants are in fact responsible for the estimate accuracy and for the speed of adaptation to temporal transitions of investigated patterns.

In particular the simulation study allowed to address the following specific questions:

- 1) Does the RLS procedure show a real limitation in estimating time-varying information flows in networks characterized by a high number of nodes ($N=60$)?
- 2) Once demonstrated the limits of RLS method related to the model dimension, can the GLKF approach substitute it in the estimation of temporal evolution of connectivity patterns characterized by a high number of nodes?
- 3) What is the influence of the two adaptation constants to be set in the application of GLKF on its performances, and which can be a criterion for the choice of their optimum values?
- 4) How are the performances of the two methods influenced by different factors affecting the recordings, like the signal-to-noise ratio (factor SNR) and the amount of trials at disposal for the analysis (factor TRIAL)?

In order to answer these questions, a simulation study was performed, on the basis of a predefined connectivity scheme which linked several modeled cerebral areas. Functional connections between these areas were retrieved by the estimation process under different experimental SNR and TRIAL conditions. Analysis of variance (ANOVA) and Duncan's pairwise comparisons were applied in order to compare the two methods and evaluate the effect of different factors on the performances achieved in terms of estimation accuracy and adaptation speed.

Estimation of Time-Varying Functional Connectivity: State of the Art

Adaptive Partial Directed Coherence

The PDC (Baccalá and Sameshima, 2001) is a spectral measure, used to determine the directed influences between any given pair of signals in a multivariate data set. A description of all its possible definitions has been already provided in Section I.

The original formulation of such estimator is based on the hypothesis of stationarity of signals included in the estimation process. Such hypothesis leads to a complete loss of the information about the temporal evolution of estimated information flows.

For overcoming this limitation, a time varying adaptation of squared PDC was introduced. The adaptation consisted in modifying the original formulation of PDC by including dependence from the time in the MVAR coefficients. Thus, the adaptive squared PDC estimator can be defined as follows:

$$sPDC_{ij}^{row}(f, t) = \frac{|\Lambda_{ij}(f, t)|^2}{\sum_{k=1}^N |\Lambda_{ik}(f, t)|^2} \quad (2.1)$$

$$sPDC_{ij}^{col}(f, t) = \frac{|\Lambda_{ij}(f, t)|^2}{\sum_{k=1}^N |\Lambda_{kj}(f, t)|^2} \quad (2.2)$$

where t refers to a dependence of the MVAR coefficients from time and $\Lambda_{ij}(f, t)$ represents the ij entry of the matrix of model coefficients Λ at frequency f and time t .

The estimation of time-varying MVAR parameters can be performed by means of two different approaches available at the moment, the Recursive Least Square and the General Linear Kalman Filter, which will be described below.

The Recursive Least Square

The Recursive Least Square with forgetting factor is a method, based on the minimization of the squared prediction error, introduced for the analysis of event related EEG data. An extended version to the multi-trials case was provided by Moller et al. in 2001 (Möller et al., 2001).

A set of EEG trials, recorded according to a certain stimulus, can be seen as several realizations of the same stochastic process. Consider a stochastic process Y composed by N repetitions (trials) of M trajectories (signals). At the time point t , the process observation can be defined as follows

$$Y_t = \begin{bmatrix} y_1^{(1)}(t) & \cdots & y_1^{(N)}(t) \\ y_2^{(1)}(t) & \cdots & y_2^{(N)}(t) \\ \vdots & \ddots & \vdots \\ y_M^{(1)}(t) & \cdots & y_M^{(N)}(t) \end{bmatrix} \quad (2.3)$$

where $y_m^{(n)}(t)$ is the n -th repetition ($n=1, \dots, N$) of the m -th trajectory ($m=1, \dots, M$) at time t ($t=1, \dots, T$). The trajectories Y_t will be fitted by the MVAR model, defined in Section I.

In order to infer the adaptive estimation $\tilde{\Lambda}_t$ of parameters Λ_t , the instantaneous prediction error should be minimized:

$$E_t = \sum_{i=1}^T (1-c)^{T-i} \|Z_i\|^2 \quad (2.4)$$

where

$$Z_t = Y_t - W_t \tilde{\Lambda}'_{t-1} \quad (2.5)$$

Z_t is the instantaneous prediction error and describes the difference between the desired response Y_t and the estimation $W_t \tilde{\Lambda}'_{t-1}$, the matrix $W_t = (Y_{t-1}, \dots, Y_{t-p})$ includes the last p observations of the time series. The introduction of the exponential $(1-c)^{t-i}$, with $0 \leq c < 1$, allows to forget the distant past of the signals in order to follow their non-stationary environment (Haykin, 1986). The constant c controls the compromise between adaptation speed and the estimation accuracy. In fact, values of c close to zero led to slower adaptation but higher stability in the estimation and vice versa. The application range of c and its best value to be set for the application of RLS algorithm to EEG signals were defined for different conditions of SNR and amount of data available for the analysis in a simulation study performed on a predefined model composed by three nodes (Astolfi et al., 2008).

Further details about the algorithm can be found in Möller et al., 2001.

The General Linear Kalman Filter

In the GLKF an adaptation of the Kalman filter to the case of multi-trial time series is provided. In particular:

$$\begin{aligned} Q_t &= G_{t-1}Q_{t-1} + V_t \\ O_t &= H_t Q_t + W_t \end{aligned} \quad (2.6)$$

where O_t represents the observation, Q_t is the state process, H_t and G_t are the transition matrices and V_t and W_t are the additive noises. To obtain the connection with the time-varying MVAR it is necessary to make the following associations:

$$Q_t = \begin{bmatrix} \Lambda_1(t)^N \\ \vdots \\ \Lambda_p(t)^N \end{bmatrix}, \quad O_t = \begin{pmatrix} y_1^{(1)}(t) & \cdots & y_M^{(1)}(t) \\ \vdots & \ddots & \vdots \\ y_1^{(N)}(t) & \cdots & y_M^{(N)}(t) \end{pmatrix} = Y_t \quad (2.7)$$

$$G_{t-1} = I_{dp}, \quad H_t = (O_{t-1}, \dots, O_{t-p}) \quad (2.8)$$

where N denotes the number of trials, whereas M is the dimension of the measured process (Milde et al., 2011). In particular, for $t=p+1:T$ the following steps were repeated:

$$Q_{t-1}^+ = Q_{t-1} + K_t(O_t - H_t Q_{t-1}) \quad (2.9)$$

$$Q_t = G_{t-1} Q_{t-1}^+ \quad (2.10)$$

$$K_t = P_{t-1} H_t^T / S_t \quad (2.11)$$

$$S_t = H_t P_{t-1} H_t^T + \text{tr}(\bar{W}_t) I_k \quad \text{where } \bar{W}_t = E[W_t W_t^T] \quad (2.12)$$

$$P_{t-1}^+ = (I_{dp} - K_t H_t) P_{t-1} \quad (2.13)$$

$$P_t = G_{t-1} P_{t-1}^+ G_{t-1}^T + \bar{V}_t \quad \text{where } \bar{V}_t = E[V_t V_t^T] \quad (2.14)$$

The rationale behind this set of equations is the demand for a linear and recursive estimator for the state $Q_t(Q_{t-1}, O_t)$ as a function of the previous state Q_{t-1} and of the actual observation matrix O_t . In particular starting from the coefficient matrix at time $t-1$, its estimation is updated to Q_{t-1}^+ by means of eq.2.9 and used for estimating the values corresponding to the following time sample Q_t by means of eq.2.10. K_t is called the Kalman gain matrix and it weights the prediction error. Its formulation is reported in eq.2.11 where P_{t-1} is the covariance matrix whose evolution is estimated by means of eq.2.13 and 2.14.

The quality of estimation is related to the definition of two parameters, $c1$ and $c2$, which regulate the compromise between the estimation accuracy and the speed of adaptation to transitions. A description of the application range of $c1$ and $c2$ and their suggested values for EEG signals under different condition of SNR and amount of trials available for the analysis was not available before my PhD thesis (Toppi et al., 2012).

Comparing Methods for the Estimation of Time-Varying Functional Connectivity: A Simulation Study

The Simulation Study

A simulation study was designed in order to compare the accuracy of the two time-varying estimation methods and their adaptation speed to information flows transitions, under different conditions of SNR and amount of trials at disposal for the estimation.

The simulation study involved the following steps:

- 1) Generation of different simulated datasets fitting a predefined model Λ composed by 60 nodes and achieved imposing different levels of trials number (factor TRIAL: 10, 20, 30, 50, 100, each trial was composed by 500 samples) and Signal to Noise Ratio (factor SNR: 0.1, 1, 3, 5). The levels chosen for both SNR and TRIAL factors cover the typical ranges for the cortical activity reconstructed by means of high resolution EEG.
- 2) Evaluation, for each dataset, of AMVAR coefficients by means of RLS and GLKF methods and estimation of related time varying PDC. The AMVAR coefficients estimation was repeated for different values of constants c for RLS and $c1$ and $c2$ for GLKF, required for the application of the two methods. In RLS case COST: $c=[0.02, 0.03, 0.04, 0.05]$, while in the GLKF case COST: $c1=c2=[0.01, 0.03, 0.05, 0.09]$. The levels for the factor COST were chosen on the basis of previous studies aiming at describing the properties of both time varying estimation methods. In particular, Astolfi et al. identified an interval of values 0.02-0.05 as a valid range for high quality time varying functional connectivity estimates by means of RLS algorithm for different values of SNR and number of trials used in the estimation process (Astolfi et al., 2008). Milde et al. suggested to set, for the GLKF method, the two adaptation constants with the same value and around 0.03 (Milde et al., 2011).
- 3) Evaluation of performance indexes, which will be defined below, allowing to compare the two methods in terms of estimation accuracy and speed of adaptation to transitions.

- 4) Analysis of Variance (ANOVA) for repeated measures of the performance indexes, in order to evaluate the effects of the factors TRIAL, SNR and COST on the performances of the analyzed methods.

Simulations were performed by repeating for 50 times each generation-estimation procedure, in order to increase the robustness of the following statistical analysis.

Signal Generation

The first step of datasets generation consisted in the creation of a predefined time varying model. Due to the necessity to test the performances of the two methods in the estimation of time varying connectivity on datasets composed by a high number of signals, the process of model generation was completely automated. The high number of signals to be generated required the use of different independent real EEG sources. In fact creating a dataset of 60 signals each one achieved as linear combination of its preceding signals led to high differences in amplitude between the firsts and the lasts which could induce spurious links in the network. The refresh of sources allowed to limit the amplitude differences between the source and the signals generated as linear combination of the preceding signals.

The time varying model Δ was generated as a juxtaposition of two stationary models \mathcal{A} and Σ , each one repeated for 250 samples (Fig.2.1a). In Fig.2.1b an example of the generated \mathcal{A} model was showed. A sub-model Λ_i was generated for each one of the S sources ($i=1, \dots, S$). Then the models were concatenated across the main diagonal of matrix \mathcal{A} , as showed in Fig.2.1b. Each model Λ_i consisted in a matrix whose not-null entries were filled with values uniformly distributed in the range $[0.2 - 0.5]$ in order to limit the differences in amplitude of signals included in the estimation process. The model Σ was built by applying small variations of the values imposed in the model \mathcal{A} to 30% of not-null arcs. In particular, I applied 15% of positive transitions and 15% of negative transitions as follows:

$$\Sigma_{ij} = (2 + 0.5r)\Lambda_{ij} \quad \text{if } 0.1 \leq \Lambda_{ij} \leq 0.25 \quad (2.15)$$

$$\Sigma_{ij} = (0.5 - 0.5r)\Lambda_{ij} \quad \text{if } 0.25 < \Lambda_{ij} \leq 0.5 \quad (2.16)$$

where r is a random number defined within the range $[0 \div 1]$.

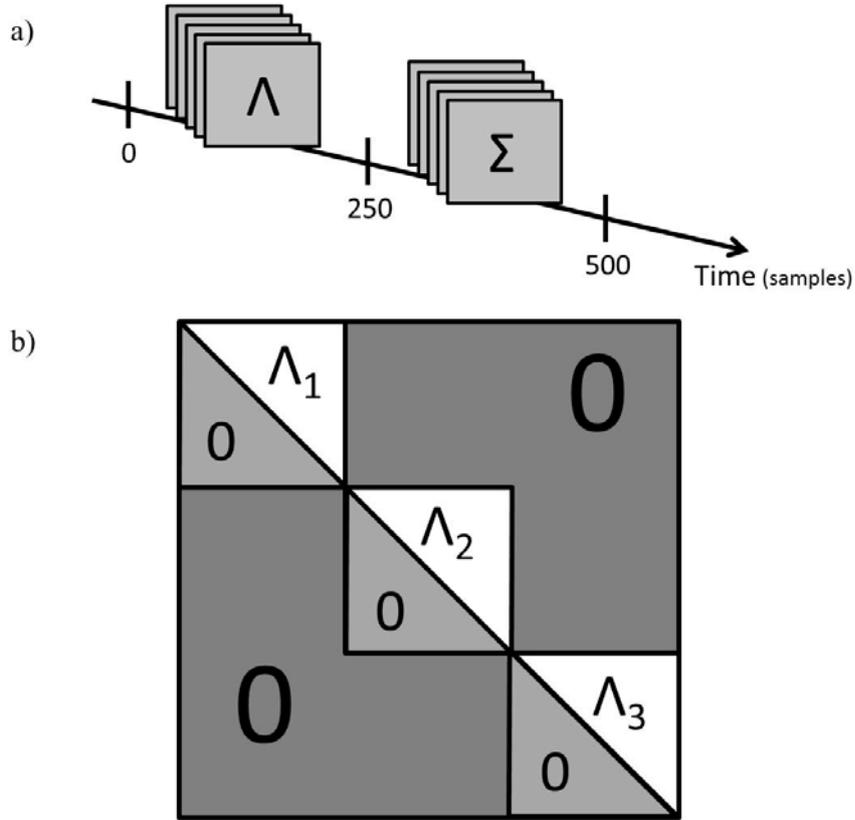


Figure 2.1 – a) The time-varying model is defined as juxtaposition of two different models Λ and Σ , each one replicated for 250 samples. Σ was achieved by applying positive or negative transitions to 30% of not null arcs of Λ ; b) Graphical representation of model Λ , achieved concatenating across the main diagonal Λ_i sub-models, consisting in an upper triangular matrix whose not-null entries were filled with values uniformly distributed in the range $[0.2 - 0.5]$

Once defined the time-varying model to be imposed, the corresponding dataset was achieved by applying the following formula:

$$y_j(t) = \sum_{i=1}^M \Delta_{ij}(t) \cdot y_i(t - \tau) + n_j(t) \quad (2.17)$$

where $\Delta_{ij}(t)$ is the ij entry of the imposed matrix of parameters Δ , achieved according the procedure described above, τ is the corresponding imposed delay for the propagation from j to i expressed in samples and $n_j(t)$ is the residual representing the part of the j^{th} signal not depending from other signals, simulated with a uncorrelated Gaussian white noise. Its amplitude was defined according to the imposed level of SNR. Delays applied were randomly chosen of one or two samples, corresponding to delays of 4 and 8 ms, respectively, at a sampling rate of 256 Hz. The number of external sources included in the model was different in relation to the model dimension. In particular I used [1, 2, 3, 6] sources for model dimensions of [10, 20, 30, 60] nodes respectively. The sources corresponded to EEG signals acquired from different electrodes of the same subject in different periods of the recording. I

decided to select signals from a single EEG recording of a subject involved in different tasks for avoiding differences in the amplitudes of the sources, which if included in the model could affect the estimation process with spurious links.

All the procedures of signal generation were repeated 50 times, each one under different conditions of:

- number of NODES = [10; 20; 30; 60]
- SNR factor = [0.1; 1; 3; 5]
- TRIAL factor = [10; 20; 30; 50; 100].

The levels chosen for both SNR and TRIALS factors cover the typical range for the cortical activity estimated during high resolution EEG experiments. The number of nodes included in the study was the one typical of standard EEG montages.

Evaluation of performances

In order to compare from different point of views the two methods used for the estimation of time-varying connectivity patterns, I defined several parameters able to consider the accuracy achieved in the estimation process and the speed of adaptation to the transitions imposed in the time evolution of the investigated networks.

In particular for the evaluation of the accuracy achieved during the estimation process, I defined three different types of errors as reported in Fig. 2.2a. Such errors derived directly from the way in which the matrix of parameters Λ was built. In particular, I adapted the definition of relative error for defining the following error types:

- 1) Non-zero arcs of the matrix Δ (Non Zero Error) for evaluating the errors committed in the estimation of the specific imposed values:

$$Error_{NonZero}(i, j) = \frac{1}{(f_2 - f_1) \cdot (t_1 - 1)} \sum_{t=0}^{t_1-1} \sum_{f=f_1}^{f_2} \frac{\|\Delta_{ij}(f, t) - \tilde{\Delta}_{ij}(f, t)\|}{\|\Delta_{ij}(f, t)\|} \quad (2.18)$$

with $i, j \in \text{triup}(\Delta_s(f, t))$, $s=1, \dots, S$

where $[f_1 f_2]$ is the frequency range in which the PDC estimator was computed, t_1 is the sample at which the model was subjected to the transition as reported in Fig. 2.2b, $\Delta_{ij}(f, t)$ and $\tilde{\Delta}_{ij}(f, t)$ correspond to the entries ij of the matrix of parameters Δ and its estimate $\tilde{\Delta}$ respectively at time t and frequency f . The symbol $\|\cdot\|$ indicates the frobenius norm, while the symbol $\text{triup}(\Delta_s(f, t))$ codes for the upper triangular matrix

of each sub-matrix A_s or Σ_s composing A or Σ with $s=1, \dots, S$ (number of independent sources included in the model).

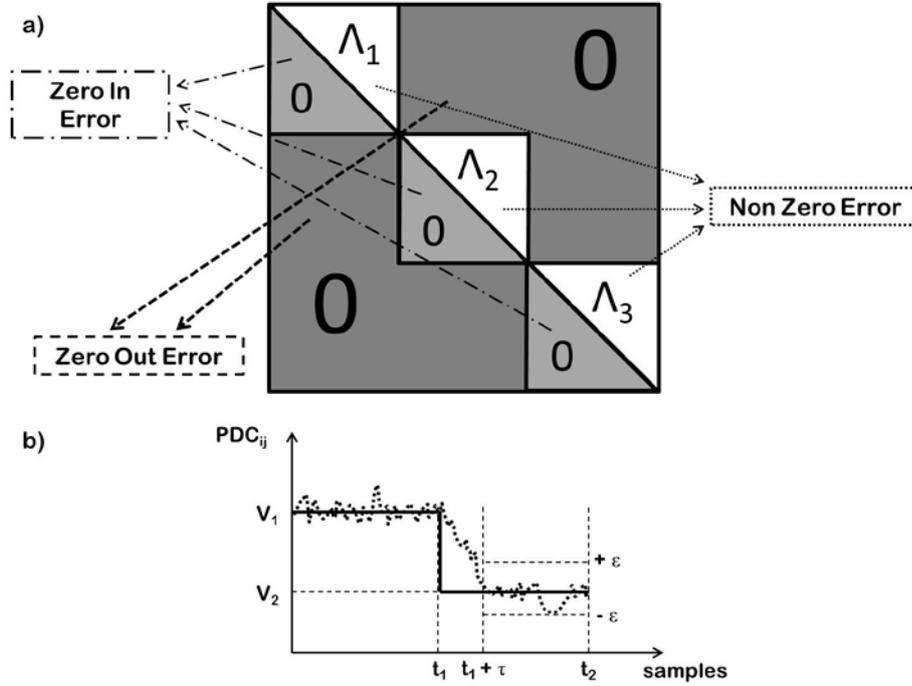


Figure 2.2 – a) Graphical representation of the matrix of parameters A , helpful for defining the three types of errors for: not-null arcs (--- line), null-arcs within the sub-matrices Λ_i (- · - line) and null-arcs outside the sub-matrices Λ_i (· · · line); b) Plot of the time evolution of a generic information flow directed from j to i imposed in the model (solid line) and estimated by means one of the two investigated method (dotted line) in the period $[0 t_2]$. At time t_1 ($n=250$ samples), a transition in the model was imposed. The estimate try to follow the transition with a specific delay τ .

- 2) Null arcs within the sub-matrices Δ_i (Zero In Error) for evaluating the amount of inversion occurred during the estimation process:

$$Error_{ZeroIn}(i, j) = \frac{1}{(f_2 - f_1) \cdot (t_1 - 1)} \sum_{t=0}^{t_1-1} \sum_{f=f_1}^{f_2} \|\Delta_{ij}(f, t) - \tilde{\Delta}_{ij}(f, t)\| \quad (2.19)$$

with $i, j \in \text{trilow}(\Delta_s(f, t))$, $s=1, \dots, S$

where $[f_1 f_2]$ is the frequency range in which the PDC estimator was computed, t_1 is the sample in which the model was subjected to the transition as reported in Fig. 2.2b, $\Delta_{ij}(f, t)$ and $\tilde{\Delta}_{ij}(f, t)$ correspond to the entry ij of the matrix of parameters Δ and its estimate $\tilde{\Delta}$ respectively at time t and frequency f . The symbol $\|\cdot\|$ indicates the frobenius norm, while the symbol $\text{trilow}(\Delta_s(f, t))$ codes for the lower triangular matrix of each sub-matrix A_s or Σ_s composing A or Σ with $s=1, \dots, S$ (number of independent sources included in the model).

- 3) Null arcs outside the sub-matrices Δ_i (Zero Out Error) for evaluating the errors occurred in links involving signals generated from different EEG sources (one for each sub-matrices Δ_i):

$$Error_{ZeroOut}(i, j) = \frac{1}{(f_2 - f_1) \cdot (t_1 - 1)} \sum_{t=0}^{t_1-1} \sum_{f=f_1}^{f_2} \|\Delta_{ij}(f, t) - \tilde{\Delta}_{ij}(f, t)\| \quad (2.20)$$

with $i, j \notin \Delta_s(f, t)$, $s=1, \dots, S$

where $[f_1, f_2]$ is the frequency range in which the PDC estimator was computed, t_1 is the sample in which the model was subjected to the transition as reported in Fig. 2.2b, $\Delta_{ij}(f, t)$ and $\tilde{\Delta}_{ij}(f, t)$ correspond to the entries ij of the matrix of parameters Δ and its estimate $\tilde{\Delta}$ respectively at time t and frequency f . The symbol $\|\cdot\|$ indicates the frobenius norm.

In order to evaluate the speed in adaptation to transitions I defined the samples at settling parameter as the first instant τ (following the transition) after which the error keeps definitely below the 20% of the transition amplitude. In particular the condition was:

$$|\Delta_{ij}(f, t) - \tilde{\Delta}_{ij}(f, t)| < \varepsilon \cdot (V_2 - V_1) / 100 \quad (2.21)$$

where ε was set to 20, $\Delta_{ij}(f, t)$ and $\tilde{\Delta}_{ij}(f, t)$ correspond to the entries ij of the matrix of parameters Δ and its estimate $\tilde{\Delta}$ respectively at time t and frequency f , V_1 and V_2 are the values of PDC related to the arc directed from j to i before and after the transition applied at time t_1 as reported in Fig. 2.2b. Once calculated τ for each node in which the transition was applied and for each frequency sample, it is possible to define the samples at settling as follows:

$$ST = \frac{1}{N_T \cdot F} \sum_{i,j=1}^{N_T} \sum_{f=1}^F \tau_{ij}(f) \quad (2.22)$$

where N_T is the number of nodes in which a transition was applied at time t_1 , F is the amplitude of the frequency range used for PDC estimation and $\tau_{ij}(f)$ is the settling time for the specific connection directed from j to i at frequency f .

Statistical Analysis

The parameters defined for comparing the two methodologies in terms of accuracy in estimation process and of speed of adaptation to transitions characterizing the temporal evolution of investigated connectivity patterns were subjected to separate ANOVAs.

First, I performed a three-way ANOVA aiming at studying the effect of different adaptation parameters (COST) set during the estimation process on the performance indexes defined above, taking into account several factors such as the SNR and the number of trials (TRIAL) available for the analysis. The within main factors of the ANOVA were COST (with four levels: $c_1=c_2=0.01$, $c_1=c_2=0.03$, $c_1=c_2=0.05$, $c_1=c_2=0.09$), SNR (with four levels: 0.1, 1, 3, 5) and TRIAL (with five levels: [10, 20, 30, 50, 100]). The dependent variables were the samples at settling and the three types of error, NonZero, ZeroIn and ZeroOut. Duncan's pairwise comparisons were then performed in order to better understand the significance between different levels of the same factor or between different factors within the same level. This analysis was executed on the time-varying connectivity networks estimated by means of GLKF. The model used for this specific analysis was composed by 60 nodes, in order to describe the behavior of GLKF method for models of high dimensions.

Secondly, I performed two separate four-way ANOVAs aiming at studying the effect of model dimension (NODE) during the estimation process performed by means of both approaches on the performance indexes defined above, taking into account several factors such as the different adaptation parameters (COST) set during the estimate, the SNR and the number of trials (TRIAL) available for the analysis. The within main factors of the ANOVA were COST (with four levels: $c_1=c_2=0.01$, $c_1=c_2=0.03$, $c_1=c_2=0.05$, $c_1=c_2=0.09$ for GLKF; $c=0.02$, $c=0.03$, $c=0.04$, $c=0.05$ for RLS), SNR (with four levels: 0.1, 1, 3, 5), NODES (with four levels: 10, 20, 30, 60) and TRIAL (with five levels: [10, 20, 30, 50, 100]). The dependent variables were the samples at settling and the three types of error, NonZero, ZeroIn and ZeroOut. For the ZeroOut error the level 10 nodes for the factor NODE was not considered, in fact that type of error could not be computed on estimates performed on 10 nodes model, being built by using only one independent EEG source. The correction of Greenhouse-Gasser for the violation of the spherical hypothesis was used. Duncan's pairwise comparisons were then performed in order to better understand the significance between different levels of the same factor or between the same level of different factors. This analysis was separately executed on the time-varying connectivity networks estimated by means of GLKF and RLS.

Results

Several sets of signals were generated as described in the previous paragraph, in order to fit the connectivity pattern built as described in Fig. 2.1a,b.

A time varying MVAR model of order 20 was fitted to each set of simulated data, which were in the form of several trials of the same length (500 samples), with the transition occurring after 250 samples from the beginning. The procedure of signal generation/estimation of time-varying connectivity patterns by means of RLS and GLKF methods was carried out 50 times for each level of factors SNR, TRIAL, COST and NODES as indicated in the previous paragraphs, in order to increase the robustness of the subsequent statistical analysis. The indexes of performances used were the samples at settling, the NonZero Error, the ZeroIn Error and the ZeroOut Error. They were computed for each generation-estimation procedure performed, and then subjected to different ANOVAs.

FACTORS	SAMPLES AT SETTLING		NON ZERO ERROR		ZERO IN ERROR		ZERO OUT ERROR	
	p	F	p	F	p	F	p	F
COST	<0.00001	198.1	<0.00001	8471	<0.00001	9687	<0.00001	58115
SNR	<0.00001	391.3	<0.00001	4096	<0.00001	10705	<0.00001	1143.2
TRIAL	<0.00001	1792	<0.00001	19274	<0.00001	2223	<0.00001	10693
COST x SNR	<0.00001	83.2	<0.00001	430	<0.00001	1616	<0.00001	1684.6
COST x TRIAL	<0.00001	61.3	<0.00001	751.1	<0.00001	899	<0.00001	6984.7
SNR x TRIAL	<0.00001	102.1	<0.00001	603.8	<0.00001	167.3	<0.00001	385.4
COST x SNR x TRIAL	<0.00001	17.7	<0.00001	145.7	<0.00001	141.5	<0.00001	522.2

Table 2.1 – Results of three-way ANOVA performed on the performance indexes computed on the connectivity patterns estimated by means of GLKF applied on a 60 nodes imposed model. The within main factors were the adaptation parameters set in the estimation process (COST), the signal to noise ratio (SNR) and the number of trials (TRIAL) available for the analysis.

A three-way ANOVA was applied separately to the performance indexes defined above and computed on the connectivity patterns estimated by means of GLKF applied on a 60 nodes imposed model. The within main factors were the adaptation parameters set in the estimation process (COST), the signal to noise ratio (SNR) and the number of trials (TRIAL) available for the analysis. The influence of the main factors and their interactions on the investigated performance parameters were reported in Tab.2.1.

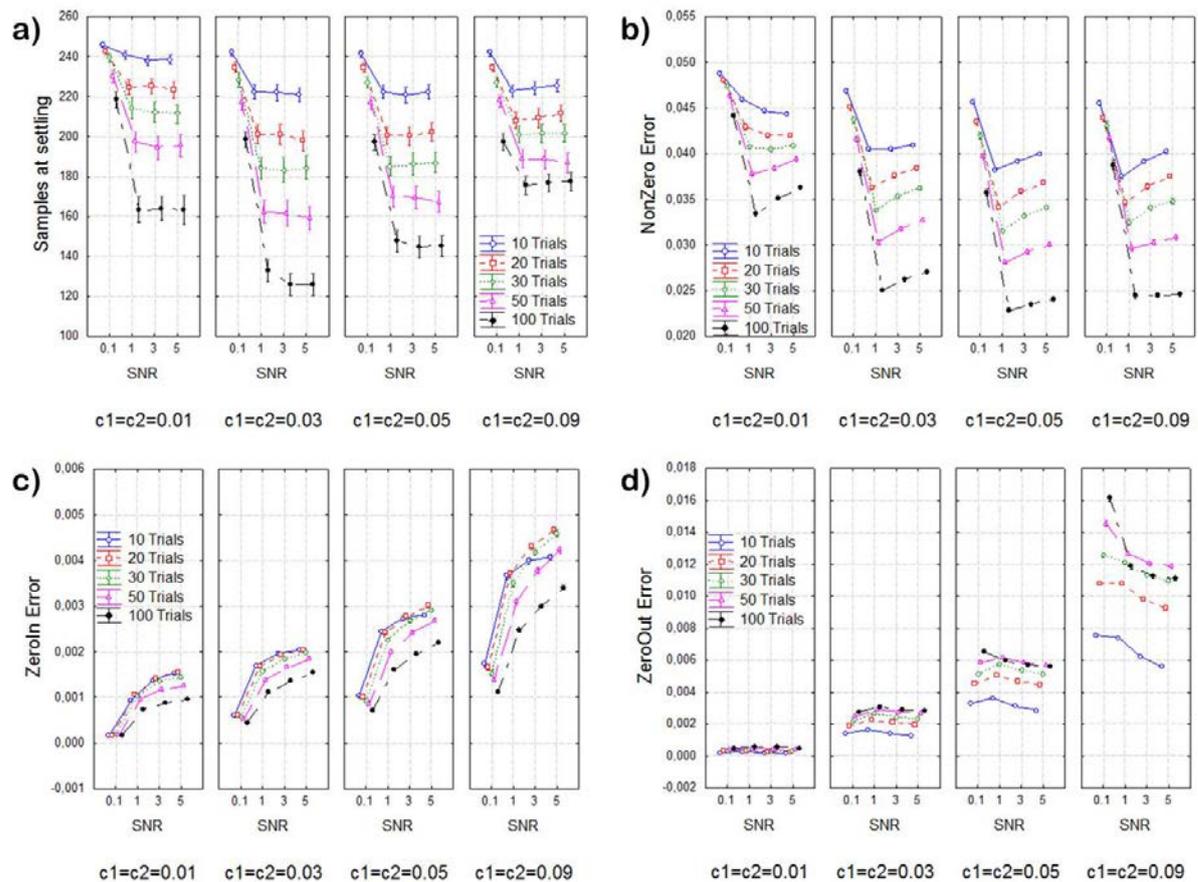


Figure 2.3 - Results of ANOVA performed on samples at settling (panel a), NonZero Error (panel b), ZeroIn Error (panel c) and ZeroOut Error (panel d) computed on 60 nodes time-varying connectivity patterns estimated by means of GLKF, using COST x SNR x TRIAL as within main factor. The diagram shows the mean value of the investigated performance index. The bar represented their relative 95% confidence interval.

In Fig. 2.3 I reported the results of the ANOVAs performed on samples at settling (panel a), NonZero Error (panel b), ZeroIn Error (panel c) and ZeroOut Error (panel d) computed on 60 nodes time-varying connectivity patterns estimated by means of GLKF, using COST x SNR x TRIAL as factor. The diagram shows the mean value of the investigated performance indexes. The bar represented their relative 95% confidence interval. Diagrams reported in Fig.2.3 highlighted the good performances achieved by GLKF in estimating time-varying connectivity patterns on high dimensional models. In fact, the time required for following the imposed transition and the errors committed in respect to the predefined model are acceptable even in bad conditions of low SNR or few trials available for the analysis. In particular in each panel I reported the results of ANOVA for the within main factor COST x TRIAL x SNR. The plot in each panel is composed by four sub-plots, one for each adaptation constant set in the estimation process: $c_1=c_2=0.01$ in the first subplot; $c_1=c_2=0.03$ in the second subplot; $c_1=c_2=0.05$ in the third subplot; $c_1=c_2=0.09$ in the fourth subplot. On the x-axis of the four subplots different values of SNR in ascending order can be found. The results

achieved for each number of trials used in the estimation process were reported with different colors: 10 trials in blue, 20 trials in red, 30 trials in green, 50 trials in pink and 100 trials in black. Panel a showed a high influence of the factors SNR and TRIAL on samples at settling parameter, which decreased with the increase of the SNR and of the number of trials used in the estimate. In particular a high decrease of the parameter resulted for an increase of SNR from 0.1 to 1, then its value remained constant for values of SNR in the range from 1 to 5. Moreover, the decrease resulted also in relation with the increase in the number of trials available for the analysis. The reduction of the time at settling is mainly evident (around 40% from 10 to 100 trials) for values of SNR above 1. No high differences between the different values of adaptation constants set during the estimation process for bad conditions of signals like low SNR and few trials available for the analysis. Although if the number of trials is above 30 or the SNR is above 1, some differences between the adaptation constants resulted in the samples at settling, with better performances achieved for the case $c_1=c_2=0.03$, as confirmed by Duncan's pairwise comparisons.

Panel b showed a high influence of the factors SNR and TRIAL on NonZero Error parameter. Its values decrease with the increase of the SNR from 0.1 to 1 and then increase for values of SNR above 1. Moreover values of NonZero Error decrease with the increase of the number of trials available for the analysis. However the error remains below 5% for all the values of factors SNR and TRIAL. Fixing the level of SNR and the number of trials included in the estimate, the increase of the two adaptation constants reduces the NonZero error.

Panel c showed a high influence of the factors SNR, TRIAL and COST on ZeroIn Error parameter. Duncan's pairwise comparisons highlighted statistically significant increase of ZeroIn error with the increase of SNR and the decrease of number of trials available for the estimate, for each adaptation constant set in the process. The lowest values of ZeroIn error resulted for adaptation constants $c_1=c_2=0.01$, the error increases with the increase of both c_1 and c_2 . However the values reported for the ZeroIn error are below 1% for all the values of SNR, for all the trials number and for all the adaptation parameters set in the process.

Panel d showed a high influence of the factors TRIAL and COST on ZeroOut error parameter. The Duncan's pairwise analysis revealed an increase of the ZeroOut error with the increase of number of trials and the increase of the adaptation constant set for applying the GLKF method. However, even for this parameter, its values remain below 1%.

FACTORS	SAMPLES AT SETTLING		NON ZERO ERROR		ZERO IN ERROR		ZERO OUT ERROR	
	p	F	p	F	p	F	p	F
NODES	p<0.0001	15	p<0.0001	2943.2	p<0.0001	1215	p<0.0001	6429
COST	p<0.0001	176	p<0.0001	2362	p<0.0001	5349	p<0.0001	53610
SNR	p<0.0001	402	p<0.0001	4796	p<0.0001	4198	p<0.0001	2134
TRIAL	p<0.0001	516	p<0.0001	4706	p<0.0001	2400	p<0.0001	808
NODES x COST	p<0.0001	73	p<0.0001	1463	p<0.0001	1007	p<0.0001	2188
NODES x SNR	0.0002	3.6	p<0.0001	662	p<0.0001	872.3	p<0.0001	459.3
COST x SNR	p<0.0001	81.1	p<0.0001	1228	p<0.0001	2234	p<0.0001	2174
COST x TRIAL	p<0.0001	108.4	p<0.0001	77.2	p<0.0001	2302	p<0.0001	787.8
NODES x TRIAL	p<0.0001	13.7	p<0.0001	395	p<0.0001	398.6	p<0.0001	1219
SNR x TRIAL	0.0002	3.2	p<0.0001	231.5	p<0.0001	257.6	p<0.0001	215.9
NODES x COST x SNR	p<0.0001	15.3	p<0.0001	533	p<0.0001	685.9	p<0.0001	242.4
NODES x COST x TRIAL	p<0.0001	20.3	p<0.0001	70.4	p<0.0001	511.9	p<0.0001	1130
NODES x SNR x TRIAL	p<0.0001	6.4	p<0.0001	75.8	p<0.0001	50.8	p<0.0001	78.5
COST x SNR x TRIAL	p<0.0001	12.2	p<0.0001	113.2	p<0.0001	273.3	p<0.0001	213.6

Table 2.2 – Results of four-way ANOVA performed on the performance indexes computed on the connectivity patterns estimated by means of GLKF. The within main factors were the adaptation parameters set in the estimation process (COST), the signal to noise ratio (SNR), the number of trials (TRIAL) available for the analysis and the number of nodes included in the model (NODES).

Several four-way ANOVAs were applied separately to the performance indexes defined above and computed on the connectivity patterns estimated by means of GLKF and RLS methods. The within main factors were the adaptation parameters set in the estimation process (COST), the signal to noise ratio (SNR), the number of trials (TRIAL) available for the analysis and the number of nodes included in the model (NODES). In particular, the results of the ANOVAs computed on investigated performance parameters extracted from the time-varying connectivity patterns estimated by means of GLKF and RLS were reported in Tab.2.2 and Tab.2.3 respectively.

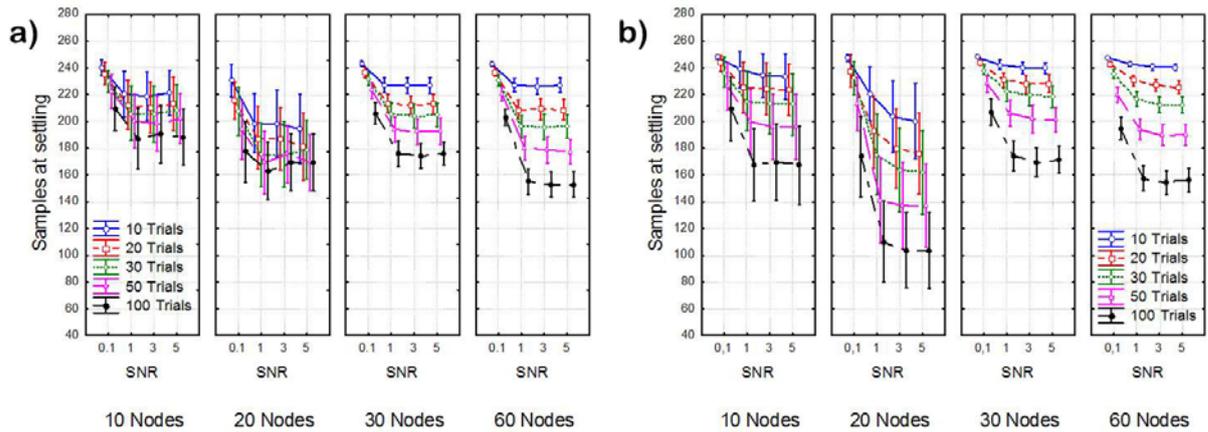


Figure 2.4 - Results of ANOVA performed on samples at settling parameter computed on time-varying connectivity patterns estimated by means of GLKF (panel a) and RLS (panel b), using NODES x SNR x TRIAL as within main factor. The diagram shows the mean value of the investigated performance index. The bar represented their relative 95% confidence interval.

In Fig. 2.4 I reported the results of the two ANOVAs performed on samples at settling parameter computed on time-varying connectivity patterns estimated by means of GLKF (panel a) and RLS (panel b), using NODES x SNR x TRIAL as within main factor. The diagram shows the mean value of the investigated performance indexes. The bar represented their relative 95% confidence interval. The plot in each panel is composed by four sub-plots, one for each model dimension: 10 nodes in the first subplot; 20 nodes in the second subplot; 30 nodes in the third subplot; 60 nodes in the fourth subplot. On the x-axis of the four subplots different values of SNR in ascending order can be found. The results achieved for each number of trials used in the estimation process were reported with different colors: 10 trials in blue, 20 trials in red, 30 trials in green, 50 trials in pink and 100 trials in black.

Both panels showed a high influence of the factors NODES, SNR and TRIAL on samples at settling parameter, which decreases with the increase of both SNR and number of trials used in the estimate. In particular a high decrease of the parameter resulted for an increase of SNR from 0.1 to 1 and then its value remained constant for values of SNR above 1. Moreover, the decrease of such parameter is also related to the increase in the number of trials available for the analysis. The reduction of the time at settling is mainly evident (around 30% from 10 to 100 trials) for values of SNR above 1. Fixing the SNR value and the number of trials included in the estimates, significant differences resulted in the time at settling for different number of nodes, as confirmed by Duncan's pairwise comparisons. In particular the lowest values of the parameter were found for 20 and 60 nodes for GLKF (panel a) and for 20 nodes for RLS (panel b), especially for a high number of trials included in the analysis. For high quality data

(high SNR and high number of trials) the number of samples needed for settling is around 180, guarantying a good speed in adaptation to transitions.

FACTORS	SAMPLES AT SETTLING		NON ZERO ERROR		ZERO IN ERROR		ZERO OUT ERROR	
	p	F	p	F	p	F	p	F
NODES	p<0.0001	29.7	p<0.0001	75	p<0.0001	277	p<0.0001	212.3
COST	p<0.0001	1149	p<0.0001	66237	p<0.0001	1849	p<0.0001	709.5
SNR	p<0.0001	416	p<0.0001	401.8	p<0.0001	934	p<0.0001	512
TRIAL	p<0.0001	988	p<0.0001	556.9	p<0.0001	3509	p<0.0001	441
NODES x COST	p<0.0001	12	p<0.0001	68	p<0.0001	156.6	p<0.0001	221.5
NODES x SNR	p<0.0001	31.6	p<0.0001	59.5	p<0.0001	131.2	p<0.0001	174
COST x SNR	p<0.0001	31.3	p<0.0001	397.4	p<0.0001	671.2	p<0.0001	474.4
COST x TRIAL	p<0.0001	46.2	p<0.0001	590.4	p<0.0001	1337.2	p<0.0001	440.2
NODES x TRIAL	p<0.0001	11.3	p<0.0001	62.9	p<0.0001	206.3	p<0.0001	122.7
SNR x TRIAL	p<0.0001	38	p<0.0001	376	p<0.0001	126.4	p<0.0001	318.4
NODES x COST x SNR	p<0.0001	3.2	p<0.0001	52	p<0.0001	105.4	p<0.0001	174
NODES x COST x TRIAL	p<0.0001	5.6	p<0.0001	54	p<0.0001	138.3	p<0.0001	138.3
NODES x SNR x TRIAL	0.0028	1.8	p<0.0001	49.9	p<0.0001	84.9	p<0.0001	98.5
COST x SNR x TRIAL	p<0.0001	2.9	p<0.0001	343	p<0.0001	434	p<0.0001	286.8

Table 2.3 – Results of four-way ANOVA performed on the performance indexes computed on the connectivity patterns estimated by means of GLKF. The within main factors were the adaptation parameters set in the estimation process (COST), the signal to noise ratio (SNR), the number of trials (TRIAL) available for the analysis and the number of nodes included in the model (NODES).

The comparison between the two approaches was performed by means of a visual inspection of the trends, in fact an ANOVA between the two methods could not be computed due to their difference in the adaptation constant to be set during the estimation process. Fixing the level of SNR, the number of nodes and the amount of trials included in the estimate, the time required for settling was similar except for the case of 20 nodes, in which the RLS showed better performances in terms of adaptation speed, especially for high number of trials.

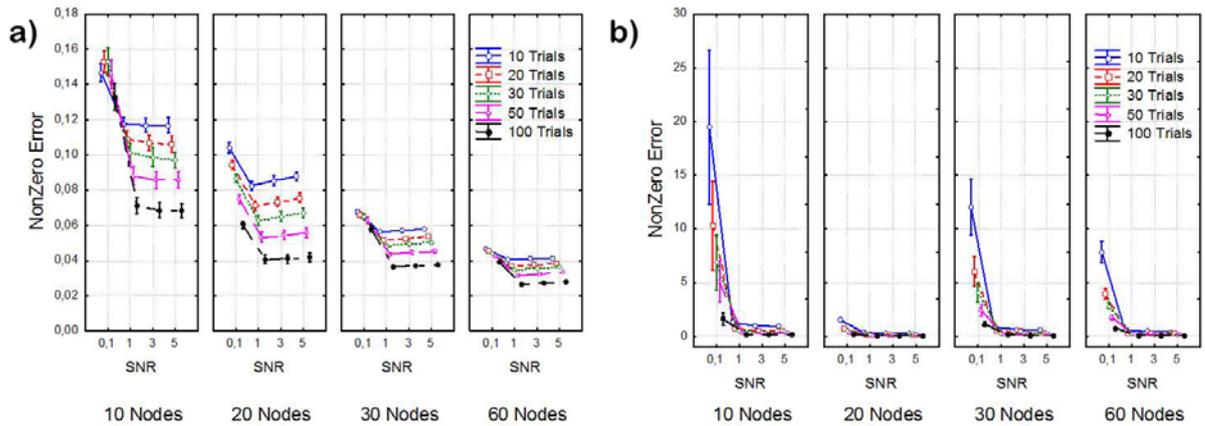


Figure 2.5 - Results of ANOVA performed on NonZero error parameter computed on time-varying connectivity patterns estimated by means of GLKF (panel a) and RLS (panel b), using NODES x SNR x TRIAL as within main factor. The diagram shows the mean value of the investigated performance index. The bar represented their relative 95% confidence interval.

In Fig. 2.5 I reported the results of the two ANOVAs performed on NonZero error parameter computed on time-varying connectivity patterns estimated by means of GLKF (panel a) and RLS (panel b), using NODES x SNR x TRIAL as within main factor. The diagram shows the mean value of the investigated performance indexes. The bar represented their relative 95% confidence interval. The plot in each panel is composed by four sub-plots, one for each model dimension: 10 nodes in the first subplot; 20 nodes in the second subplot; 30 nodes in the third subplot; 60 nodes in the fourth subplot. On the x-axis of the four subplots different values of SNR in ascending order can be found. The results achieved for each number of trials used in the estimation process were reported with different colors: 10 trials in blue, 20 trials in red, 30 trials in green, 50 trials in pink and 100 trials in black. Both panels revealed a strong influence of the factors NODE, SNR and TRIAL on the NonZero error, which decreasing with the increase of SNR level and with the amount of trial included in the estimate. The trends reported for the NonZero error in relation to SNR and TRIAL factors are similar to those achieved for the samples at settling. Fixing the SNR level and the amount of trials included in the estimate, a reduction of the NonZero error according to an increase in the number of nodes included in the model resulted for GLKF. Similar trends across different number of nodes were shown for RLS. Comparing by visual inspection the trends achieved for NonZero error by applying both methods, the performances achieved by GLKF are better than those obtained for RLS. In fact the NonZero Error remained below 15% in case of low quality data or below 5% in case of high quality data for the GLKF, while it reached really high values (above 100%) for the RLS, confirming what was already known in literature about RLS.

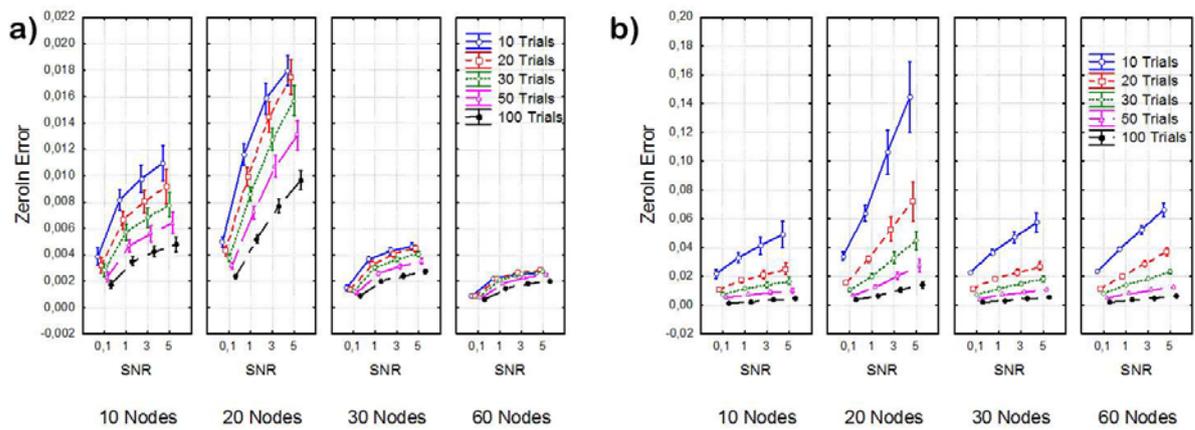


Figure 2.6 - Results of ANOVA performed on ZeroIn error parameter computed on time-varying connectivity patterns estimated by means of GLKF (panel a) and RLS (panel b), using NODES x SNR x TRIAL as within main factor. The diagram shows the mean value of the investigated performance index. The bar represented their relative 95% confidence interval.

In Fig. 2.6 I reported the results of the two ANOVAs performed on ZeroIn error parameter computed on time-varying connectivity patterns estimated by means of GLKF (panel a) and RLS (panel b), using NODES x SNR x TRIAL as within main factor. The diagram shows the mean value of the investigated performance indexes. The bar represented their relative 95% confidence interval. The plot in each panel is composed by four sub-plots, one for each model dimension: 10 nodes in the first subplot; 20 nodes in the second subplot; 30 nodes in the third subplot; 60 nodes in the fourth subplot. On the x-axis of the four subplots different values of SNR in ascending order can be found. The results achieved for each number of trials used in the estimation process were reported with different colors: 10 trials in blue, 20 trials in red, 30 trials in green, 50 trials in pink and 100 trials in black. Both panel revealed a strong influence of the factors NODE, SNR and TRIAL on the ZeroIn error, which decreases with the increase in the amount of trial included in the estimate and increases with the increase of SNR level. The ZeroIn error decreases with the increase in the number of nodes included in the estimate especially for GLKF. Similar trends among different number of nodes resulted for RLS, except for the case of 20 nodes, where such error is higher than in the other cases. As in the NonZero error the comparison between the two methods revealed better performances for the GLKF in respect to the RLS for all the conditions of SNR, amount of trials and number of nodes.

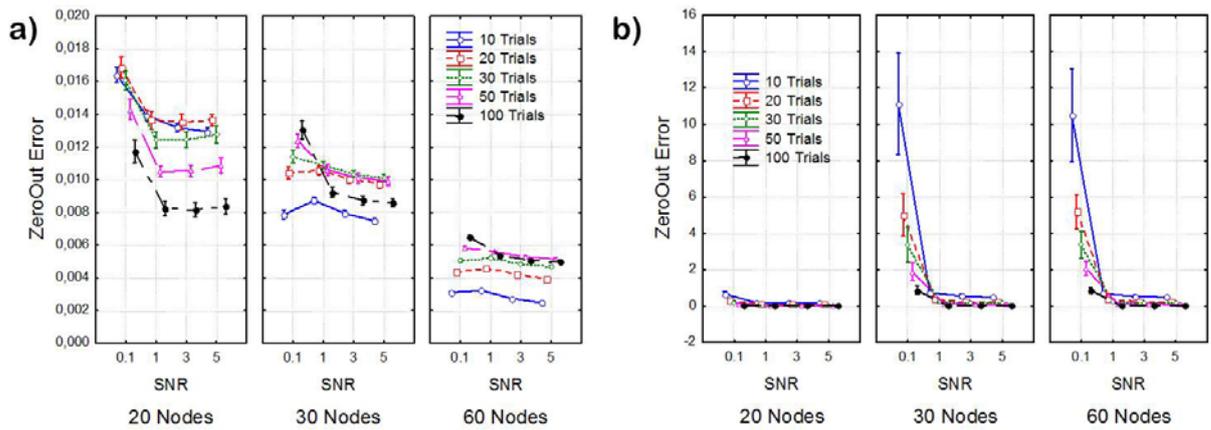


Figure 2.7 - Results of ANOVA performed on ZeroOut error parameter computed on time-varying connectivity patterns estimated by means of GLKF (panel a) and RLS (panel b), using NODES x SNR x TRIAL as within main factor. The diagram shows the mean value of the investigated performance index. The bar represented their relative 95% confidence interval.

In Fig. 2.7 I reported the results of the two ANOVAs performed on ZeroOut error parameter computed on time-varying connectivity patterns estimated by means of GLKF (panel a) and RLS (panel b), using NODES x SNR x TRIAL as within main factor. The diagram shows the mean value of the investigated performance indexes. The bar represented their relative 95% confidence interval. The plot in each panel is composed by three sub-plots, one for each model dimension: 20 nodes in the first subplot; 30 nodes in the second subplot; 60 nodes in the third subplot. On the x-axis of the three subplots different values of SNR in ascending order can be found. The results achieved for each number of trials used in the estimation process were reported with different colors: 10 trials in blue, 20 trials in red, 30 trials in green, 50 trials in pink and 100 trials in black. Both panel revealed a strong influence of the factors NODE, SNR and TRIAL on the ZeroOut error. The trends reported for the ZeroOut parameter computed on connectivity patterns estimated by means of RLS is similar to those already described for the samples at settling and for the NonZero error. In fact the error reduces with the increase of SNR level and of the amount of trials considered in the analysis. The GLKF showed different behaviors in relation to the values of the factors taken in account in the simulation study. A reduction of the error related to the increase of the SNR level resulted for all the number of nodes. In the 20 Nodes case, the error decreases also in relation to the increase in the number of trials considered in the estimate. For 30 and 60 nodes, the error increases with the increase of the amount of data included in the analysis. However, the ZeroOut error computed on time-varying patterns estimated by means of GLKF was lower than the error committed with the application of RLS method, highlighting, also in this case, the best performances of the GLKF.

Discussion

In order to evaluate the performances of the two methods in the estimation of time-varying functional connectivity patterns, I performed several statistical analysis (ANOVAs and Duncan's pairwise comparisons) on different simulated datasets, built according to predefined schemes composed by different number of nodes (from 10 to 60), under different conditions of SNR and number of trials included in the estimate whose variability ranges are similar to those found during real EEG recordings. The imposed connectivity model was built in order to test the validation processes in some situations typical of cerebral networks such as the presence of direct, indirect and 'null' arcs with strengths in ranges consistent with the ones estimated in previous studies characterized by a large sample of subjects performing memory, motor, and sensory tasks. In particular I evaluated the performances of these two methods in terms of accuracy in the estimation process and in terms of speed of adaptation to transitions applied on the temporal evolution of connectivity patterns, also putting them in relation with the model dimension. Thus, the performance parameters were evaluated for the two time-varying approaches in relation to the SNR level, to the amount of trials included in the estimate, to the adaptation constants set during the estimation process and to the number of nodes characterizing the investigated networks.

The study was conducted with the aims of: i) characterizing the new GLKF method and describe its performances in relation to the adaptation constants to be set during its application; ii) verifying the limit of RLS in estimating time-varying connectivity patterns characterized by a high number of nodes; iii) comparing the performances of the two approaches under different conditions of signal quality and model dimension.

The simulations provided the following answers to the questions raised in the Introduction:

- 1) The RLS method showed limits in the estimation of connectivity patterns characterized by a high number of nodes. The errors remained below 50% only if the signals included in the estimates were characterized by an elevate SNR level ($\text{SNR} \geq 3$) and a high amount of trials available for the analysis ($\#\text{trials} \geq 100$). The decline of the quality of data strongly reduced the RLS performances, leading the NonZero error over 100%.
- 2) The comparison between the recent GLKF method and the already existent RLS approach performed in this study demonstrated that GLKF method can substitute the RLS approach in the time-varying estimation of connectivity patterns due to its higher performances in following the temporal dynamics of investigated networks. In fact the

time required for settling was below 500 ms for all the SNR levels, number of trials and number of nodes considered in the study. Such time drastically reduced (around 300 ms) if the signals included in the estimate were of high quality ($\text{SNR} \geq 3$ and $\#\text{trials} \geq 30$). Moreover, the accuracy in estimation process showed by GLKF was really high, considering that the NonZero error remained below the 20%, the ZeroIn and the ZeroOut errors below 2% for all the SNR levels, number of trials and number of nodes considered in the study. As expected, the NonZero error decreased below 5% for high quality data. The high performances of GLKF persisted also for models of high dimension, confirming from a quantitative point of view what has been qualitatively demonstrated in literature (Milde et al., 2010).

- 3) The two adaptation constants allowed to regulate the speed of adaptation to transitions in temporal dynamics of connectivity patterns. High values of c_1 and c_2 increased the speed of time-varying estimates but at the same time decreased their accuracy. If we are interested in increase the adaptation speed the best choice is $c_1=c_2=0.03$, if we are interested instead in maximizing the accuracy of the estimate it's preferable the choice $c_1=c_2=0.01$. However the simulation study provided some guidelines allowing to select the best value for the two constants c_1 and c_2 in relation to the quality of the data and to the amount of trials available for the analysis (Toppi et al., 2012).
- 4) Both SNR level and the number of trials available for the estimate have a reliable effect on performance indexes computed on the estimates achieved by means of both methods. In fact a consistent reduction of the time at settling and an increase of the estimation accuracy resulted in relation to the increase of SNR level and of the amount of trials included in the process. However, the performances achieved with GLKF can be considered as acceptable also for low SNR and few trials included in the analysis. The RLS showed acceptable performances only on high quality data.

In conclusion, the ANOVA results (integrated with the Duncan's test performed at $p < 0.05$) confirmed what has been qualitatively introduced in previous studies about the issues of time-varying connectivity and limits in the dimension of connectivity patterns. The results are in agreement with previous simulation studies performed on the effect of factors such as SNR level and number of trials on the samples at settling and accuracy parameters (Astolfi et al., 2008).

The new method for the estimation of time-varying connectivity is a valid tool able to overcome the limits of existent procedures. In fact, the method is able to provide good time-

varying estimates characterized by a high accuracy and elevate speed of adaptation to transitions also for networks composed by a high number of nodes. Such results confirmed what has been introduced by Milde et al., 2010.

The simulation study highlighted high performances of GLKF in conditions which are generally obtained in many standard EEG recordings of event-related activity in humans, usually characterized by values of SNR ranging from 3 (movement related potentials) to 10 (sensory evoked potentials) and by a number of trials ranging from 30 to 50. In the same conditions the RLS showed bad performances especially in the accuracy of estimated patterns characterized by a high number of nodes.

Conclusion

The characterization of event-related tasks whose timing is regulated by an external trigger requires the use of time-varying approaches able to follow the temporal dynamics of investigated neuronal processes. The first method proposed ten years ago and based on the assumption of stationarity of signals in short time windows provided a not sufficient solution to the problem, offering time-varying solution with low temporal resolution. Few years later, a more advanced approach based on a recursive algorithm involving a weighted influence of the past of the investigated signal was proposed. Such approach showed a good speed of adaptation to temporal transitions but it presented a limitation in the number of signals to be included simultaneously in the estimate. The necessity to overcome this limit led to the development of a new approach based on Kalman filter able to provide accurate time-varying estimates also for high dimensional model. In fact, the limitation in the dimension of the model used for the adaptive connectivity estimation could affect the achieved connectivity patterns with spurious links due to “hidden sources”.

A quantitative description of the properties of the two time-varying methods and of their application field was not available so far. In fact in order to understand the advantages and the limits of both approaches under different conditions of SNR, amount of trials included in the analysis and model dimension, a simulation study was conducted and described in this section.

The results achieved in this section of the thesis highlighted the possibility to have a methodology able to estimate the temporal evolution of connectivity patterns with high performances also including all the sources in the estimate. In particular, the study provided a set of results to be used as guidelines for the application of time-varying methods without any a-priori selection of the sources of electrical activity. Such guidelines were reported in relation to the quality of data available and the number of sources included in the analysis. Even in this case, as for the results achieved in the connectivity assessment in Sec.I, it is necessary to take into account that a linear approach is used in both generation of datasets and estimation processes. For this reason, these results did not take into account all the errors due to a wrong choice of the model but only those related to the inaccurate estimates. In the future, the same investigations should be replicated on datasets generated including the nonlinearity of investigated signals.

Moreover, notwithstanding the advancement achieved by means of the new Kalman-based approach, further investigations are necessary for increasing the temporal resolution of time-varying estimates. In fact, the application of such promising procedure based on Kalman filter leads to a reduction of the temporal resolution typical of EEG recordings. Even if the estimation is performed for each sample in the observation window, the transitions were described by means of an oscillating trend followed by a slow adaptation towards the updated state. Such process which is the one typical of models perturbed by an external agent, is described by a group of consecutive samples (more than 20 samples) not allowing to get instantaneously the variations. For this reason, in the future new methodologies will be necessary for reducing the time required for the settling of the algorithm with the aim to reduce the delay in the estimation of temporal transitions of information flows.

Section II References

- Astolfi, L., Cincotti, F., Mattia, D., De Vico Fallani, F., Tocci, A., Colosimo, A., Salinari, S., Marciani, M.G., Hesse, W., Witte, H., Ursino, M., Zavaglia, M., Babiloni, F., 2008. Tracking the time-varying cortical connectivity patterns by adaptive multivariate estimators. *IEEE Trans Biomed Eng* 55, 902–913.
- Baccalá, L.A., Sameshima, K., 2001. Partial directed coherence: a new concept in neural structure determination. *Biol Cybern* 84, 463–474.
- Haykin, S., 1986. *Adaptive Filter Theory*. Prentice Hall.
- Milde, T., Leistriz, L., Astolfi, L., Miltner, W.H.R., Weiss, T., Babiloni, F., Witte, H., 2010. A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate AR models and its application in analysis of laser-evoked brain potentials. *Neuroimage* 50, 960–969.
- Milde, T., Putsche, P., Schwab, K., Wacker, M., Eiselt, M., Witte, H., 2011. Dynamics of directed interactions between brain regions during interburst–burst EEG patterns in quiet sleep of full-term neonates. *Neuroscience Letters* 488, 148–153.
- Möller, E., Schack, B., Arnold, M., Witte, H., 2001. Instantaneous multivariate EEG coherence analysis by means of adaptive high-dimensional autoregressive models. *J. Neurosci. Methods* 105, 143–158.
- Toppi, J., Babiloni, F., Vecchiato, G., Cincotti, F., De Vico Fallani, F., Mattia, D., Salinari, S., Astolfi, L., 2012. Towards the Time Varying Estimation of Complex Brain Connectivity Networks by means of General Linear Kalman Filter. *Conf Proc IEEE Eng Med Biol Soc*.

Section III

Graph Theory Approach: A consistent description of the relevant properties of the brain networks

INTRODUCTION

GRAPH THEORETICAL APPROACH: STATE OF THE ART

Concept at the basis of Graph Theory

Adjacency Matrix Extraction

Graph Theory Indexes

BEYOND THE STATE OF THE ART: HOW THE STATISTICAL VALIDATION OF FUNCTIONAL CONNECTIVITY PATTERNS CAN PREVENT ERRONEOUS DEFINITION OF SMALL-WORLD PROPERTIES OF A BRAIN CONNECTIVITY NETWORK

Methods

Datasets used in the experiment

Signal Processing

Analysis of Variance

Results

Simulated data

Mannequin data

Discussion

CONCLUSION

SECTION III REFERENCES

Introduction

The methodological advancement in the functional connectivity field has been leading to the description of neurological mechanisms at the basis of complex cerebral processes involving a high number of sources. Once qualitatively described the connectivity pattern achieved for the investigated condition, a quantitative characterization of its main properties is necessary in order to synthesize the huge amount of information derived from the application of such advanced methodologies. The extraction of indexes describing global and local properties of the investigated networks could open the way to several different applications in neuroscience field. For example, such indexes if achieved from EEG data recorded during human resting condition can be used for: i) describing changes in resting properties due to ageing (Zhu et al., 2011); ii) characterizing properties of the healthy population to be used as model for the

description of some pathological conditions and their recovery (Toppi et al., 2012); iii) predicting the performances achieved by the subject in executing a particular task (Grosse-Wentrup and Schölkopf, 2012).

The extraction of salient characteristics from brain connectivity patterns is thus a challenging topic which needed to be improved and refined in order to be performed in a consistent, stable and repetitive way. In the last ten years, a graph theoretical approach was proposed for the characterization of the topographical properties of real complex networks (Sporns et al., 2004; Strogatz, 2001). In fact, it was demonstrated that tools already implemented and used for the treatments of graphs as mathematical objects could be applied to functional connectivity networks estimated from electroencephalographic (EEG), magnetoencephalographic (MEG) or hemodynamic (fMRI) recordings (De Vico Fallani et al., 2012; Bullmore and Sporns, 2009; Stam and Reijneveld, 2007; Stam, 2004). The use of characteristic indexes, borrowed by graph theory, allows the evaluation of real networks in terms of density of connections incoming or outgoing from a node, tendency to cluster, centrality of some nodes or edges and distances between nodes (Fagiolo, 2007; Sporns et al., 2004; Latora and Marchiori, 2001).

The computation of graph indexes can be performed on adjacency matrices achieved by applying a threshold on the estimated connectivity values obtained by means of different estimators. The application of a thresholding procedure allows to convert the connectivity values into edges. An edge connecting two nodes exists if the connectivity value between those nodes is above a certain threshold; otherwise the edge is null. The choice of the threshold should not depend on the application, and, if done in an arbitrary way, could affect the results. In fact, the threshold influences the number of connections considered for the subsequent graph analysis and thus affects the indexes extracted from the network (Van Wijk et al., 2010). Different methodologies are available for defining such threshold. A possible approach is to select a fixed threshold. In this respect, three criteria are typically adopted: 5% significant level as a threshold fixed for discarding connectivity values from the random case (Ferrarini et al., 2009; Chen et al., 2008; Salvador et al., 2005); an arbitrary value in order to discard the weak connections (Van den Heuvel et al., 2009); the largest possible threshold allowing all nodes to be connected at least to another node in the network (Bassett et al., 2006). The second way to extract a threshold is to fix the average degree within the networks in order to maximize the small-world properties of the network (Rubinov and Sporns, 2010; Wang et al., 2009; de Haan et al., 2009; Ferri et al., 2008; He et al., 2008). A third way to define a threshold is to fix the edge density of the network, i.e. the number of existing edges

divided by the number of possible edges (Wang et al., 2009). This approach is useful if we are interested in comparing different conditions but can produce modifications in the topology of the studied network (Van Wijk et al., 2010).

All the approaches described above are empirical and do not take into account the intrinsic statistical significance of the estimator used in functional connectivity estimation process. In fact, when the adjacency matrix is achieved by imposing a threshold fixing the number of residual connections of the network, we cannot exclude a-priori that a percentage of such residual connections is estimated by chance. My idea was thus to take into account the statistical significance of the estimator used for functional connectivity estimation in the construction of adjacency matrix. In the case of PDC, the threshold is extracted by applying a percentile, for a defined significance level, on the distribution achieved for such estimator in the null case. In this way, an edge exists in the adjacency matrix describing the considered network if and only if it is statistically different from the null case.

In the present section of my thesis, first of all I briefly introduced the concepts at the basis of graph theory, I described the way in which the adjacency matrix can be extracted and all the possible indexes which can be defined. In the second part, I described a simulation study conducted for investigating how the methods used for extracting the adjacency matrix on which is based the graph theory could affect the global properties of the investigated networks.

Graph Theory Approach: State of the Art

Concept at the basis of Graph Theory

A graph is a mathematical object consisting in a set of vertices (or nodes) and edges (or connections) indicating the presence of some sort of interaction between the vertices. The adjacency matrix A contains the information about the connectivity structure of the graph and can be derived directly from the investigated network as reported in Fig.3.1. When a directed edge exists from the node j to the node i , the corresponding entry of the adjacency matrix is $A_{ij} = 1$ in binary graphs or $A_{ij} = v$ (where v is the value achieved by the estimator) in weighted graphs otherwise $A_{ij} = 0$. The adjacency matrix can be used for the extraction of salient information about the characteristic of the investigated network by defining several indexes based on its elements.

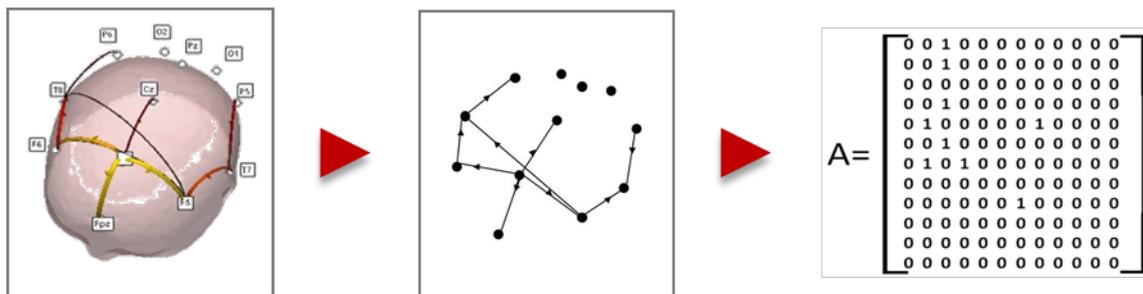


Figure 3.1 – Flowchart describing the process of adjacency matrix extraction starting from the estimated connectivity pattern.

In relation to the estimator used for building the network, the associated graph could be:

- undirected → if the estimator is able to extract only the value of the information flows and not the direction. In this case the adjacency matrix is symmetric ($A_{ij} = A_{ji}$).
- directed → if the estimator allows to reconstruct not only the magnitude but also the direction of the connection. In this case the adjacency matrix is asymmetric ($A_{ij} \neq A_{ji}$).

If the estimator used for the analysis is based on multivariate approach, as the case of Partial Directed Coherence (Baccalá and Sameshima, 2001), the corresponding graph is binary/weighted and directed.

Adjacency Matrix Extraction

Once the functional connectivity pattern is estimated, it is necessary to define an associated adjacency matrix for each network, on which extract salient indexes able to characterize the network properties. The generic ij^{th} entry of a directed binary adjacency matrix is equal to 1 if there is a functional link directed from the j^{th} to the i^{th} signal and to 0 if no link exists. As explained in the Introduction section, the construction of an adjacency matrix can be performed by comparing each estimated connectivity value to its correspondent threshold value. In particular:

$$G_{ij} = \begin{cases} 1 \rightarrow A_{ij} \geq \tau_{ij} \\ 0 \rightarrow A_{ij} < \tau_{ij} \end{cases} \quad (3.1)$$

where G_{ij} and A_{ij} represent the entry (i,j) of an adjacency matrix G and a connectivity matrix A , respectively, and τ_{ij} is the corresponding threshold. It is possible to derive the adjacency matrix simply by applying the same threshold for all the links of the network. In this case, eq. (3.1) becomes

$$G_{ij} = \begin{cases} 1 \rightarrow A_{ij} \geq \tau \\ 0 \rightarrow A_{ij} < \tau \end{cases} \quad (3.2)$$

where τ represents the threshold to be applied to all the links in the network.

Different approaches have been developed for evaluating the threshold values, as already described in the Introduction paragraph, most of them based on qualitative assumptions aiming at fixing the number of edges or the degree of some nodes or at maximizing some properties of the investigated networks. The selection of the threshold to be used for the extraction of adjacency matrices is a crucial step of the graph theory approach; in fact the type of threshold used in the process might affect the structure and the topological properties of the investigated networks. For this reason, further investigations are needed in this field with the aim to establish a consistent and stable procedure to be applied, able to keep the real global and local properties of the connectivity patterns under analysis.

Graph Theory Indexes

Different indexes can be defined on the basis of the adjacency matrix extracted from a given connectivity pattern. The most commonly used, will be described in the following:

Global Efficiency. The global efficiency is the average of the inverse of the geodesic length and represents the efficiency of the communication between all the nodes in the network (Latora and Marchiori, 2001). It can be define as follows

$$E_g = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \quad (3.3)$$

where N represents the number of nodes in the graph and d_{ij} the geodesic distance between i and j (defined as the length of the shortest path between i and j).

Local Efficiency. The local efficiency is the average of the global efficiencies computed on each sub-graph G_i belonging to the network and represents the efficiency of the communication between all the nodes around the node i in the network (Latora and Marchiori, 2001). It can be defined as follows

$$E_l = \frac{1}{N} \sum_{i=1}^N E_g(G_i) \quad (3.4)$$

where N represents the number of nodes in the graph and G_i the sub-graph achieved deleting the i^{th} row and the i^{th} column from the original graph.

Degree. The degree of a node consists in the number of links connected directly to it. In directed networks, the indegree is the number of inward links and the outdegree is the number of outward links. Connections weight is ignored in calculations (Sporns et al., 2004). Degree can be defined as follows

$$k_i = \sum_{j \in N} a_{ij} \quad (3.5)$$

where a_{ij} represents the entry ij of the Adiacency matrix A .

Betweenness Centrality. Node betweenness centrality is the fraction of all shortest paths in the network containing a given node. Nodes with high values of betweenness centrality participate in a large number of shortest paths (Kintali, 2008; Freeman, 1978). It can be defined as follows

$$b_i = \frac{1}{(n-1)(n-2)} \sum_{\substack{h, j \in N \\ h \neq j, h \neq i, j \neq i}} \frac{\rho_{hj}(i)}{\rho_{hj}} \quad (3.6)$$

where ρ_{hj} is the number of shortest paths between h and j and $\rho_{hj}(i)$ is the number of shortest paths between h and j that pass through i .

Characteristic Path Length. The characteristic path length is the average shortest path length in the network, where the shortest path length between two nodes is the minimum number of edges that must be traversed to get from one node to another. It can be defined as follows

$$L = \frac{1}{n} \sum_{i \in N} L_i = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}}{n-1} \quad (3.7)$$

where L_i is the average distance between node i and all other nodes and d_{ij} is the geodesic distance between node i and node j (Sporns et al., 2004).

Clustering Coefficient. The clustering coefficient describes the intensity of interconnections between the neighbors of a node (Watts and Strogatz, 1998). It is defined as the fraction of triangles around a node or the fraction of node's neighbors that are neighbors of each other. The binary directed version of Clustering Coefficient is defined as follows (Fagiolo, 2007)

$$C = \frac{1}{n} \sum_{i \in N} C_i = \frac{1}{n} \sum_{i \in N} \frac{t_i}{(k_i^{out} + k_i^{in})(k_i^{out} + k_i^{in} - 1) - 2 \sum_{j \in N} g_{ij} g_{ji}} \quad (3.8)$$

where t_i represents the number of triangles involving node i , k_i^{in} and k_i^{out} are the number of incoming and outgoing edges of nodes i respectively and g_{ij} is the entry ij of adjacency matrix.

Small-Worldness. A network G is defined as small-world network if $L_G \cong L_{rand}$ and $C_G \gg C_{rand}$ where L_G and C_G represent the characteristic path length and the clustering coefficient of a generic graph and L_{rand} and C_{rand} represent the correspondent quantities for a random graph (Watts and Strogatz, 1998). On the basis of this definition, a measure of small-worldness of a network can be introduced as follows

$$S = \frac{C_G / C_{rand}}{L_G / L_{rand}} \quad (3.9)$$

So a network is said to be a small world network if $S > 1$ (Humphries and Gurney, 2008).

The indexes described above do not allow to characterize the network in terms of symmetries or influences between two different areas of the scalp. For this purpose, during my PhD I defined two new indexes: symmetry and influence, able to describe the topological properties of the investigated networks (Toppi et al., 2012). For the evaluation of such indexes it is

important to modify the adjacency matrix A in A' by disposing in the first N_1 rows and N_1 columns the connectivity values related to the first cerebral area and in the second N_2 rows and N_2 columns the connectivity values related to the second cerebral area.

Symmetry. The symmetry index is the percentage difference in the number of internal connections between two different spatial regions. It could assume values in the range $[-1 ; 1]$, where 0 means same number of internal connections within the two regions, 1 means connections only within the first region and -1 means only connections within the second region. It can be defined as follows

$$S = \frac{\sum_{i=1}^{N_1} \sum_{j=1}^{N_1} A'_{ij}}{N_1^2} - \frac{\sum_{i=N_1+1}^{N_2} \sum_{j=N_1+1}^{N_2} A'_{ij}}{N_2^2} \quad (3.10)$$

where N_1 and N_2 indicates the number of electrodes belonging to the two scalp regions respectively and A'_{ij} represents the adjacency matrix modified according to the scalp's regions to be compared.

Influence. The influence index is the percentage difference in the number of inter-connections between two different spatial regions. It could assume values in the range $[-1 ; 1]$, where 0 means no connections or same number of inter-connections between the two regions, 1 means connections only from the second region to the first and -1 means only connections from the first region to the second.

$$I = \frac{\sum_{i=1}^{N_1} \sum_{j=N_1+1}^{N_2} A'_{ij} - \sum_{i=N_1+1}^{N_2} \sum_{j=1}^{N_1} A'_{ij}}{N_1 \cdot N_2} \quad (3.11)$$

where N_1 and N_2 indicates the number of electrodes belonging to the two scalp regions respectively and A'_{ij} represents the adjacency matrix modified according to the scalp's regions to be compared.

The last two indexes could be used in neuroscience field for investigating the symmetries and influences between for example the two hemispheres or between frontal and parietal areas.

Beyond the State of the Art: How the Statistical Validation of Functional Connectivity Patterns Can Prevent Erroneous Definition of Small-World Properties of a Brain Connectivity Network

The extension of graph theory approach to the clinical field required the investigation of a reliable and consistent procedure able to extract stable indexes whose variations could be attributed only to the experimental conditions and not to the instability of the methodologies used. For this reason, during my PhD I performed a study with the aim to reinforce the weak link in the graph processing chain by proposing an alternative method for adjacency matrix extraction, based on a statistical approach, able to overcome the limitations of existing procedure.

The general aim of the study is to understand how the methods for extracting the adjacency matrix could affect the graph theory indexes and their interpretation, in order to define a reliable approach for the derivation of salient indexes from connectivity networks estimated by means of multivariate methods. In particular I used two different datasets with the purpose of comparing one of the methods extensively used in graph theory applications for extracting adjacency matrices from the connectivity patterns (i.e. the method based on fixing the edge density) with the new proposed method based on the thresholds extracted by means of the statistical validation of connectivity patterns. The first dataset I used consisted of a set of random uncorrelated signals, which should represent a null model for functional connectivity estimates and a random case for graph theory indexes. In fact, since no correlation exists between signals, the connectivity estimation process should discard almost entirely the information flows between signals, leaving only a few percentage of connections, estimated by chance and organized according to a random network. This dataset can be seen as an ideal “null case” model, but it does not take into account some factors strictly related to an electroencephalographic recording, such as the existence of a correlation between the recorded signals, due to effects of volume conduction, to the spatial positions of electrodes disposed on the scalp and to the location of the reference (Nunez et al., 1997). For this reason, I introduced a second dataset, composed by signals recorded from a mannequin head during a pseudo experiment. This situation represents the null model for functional connectivity estimates inferred by applying Partial Directed Coherence on EEG signals recorded at scalp

level. In fact, the absence of physiological content in the recorded signals allows to model the absence of information flows between electrodes, but at the same time, the use of a real EEG cap, with electrodes positioned as 10-20 systems and references placed at the earlobes, models the effects of some factors typical of an EEG recording situation.

I estimated the functional connectivity patterns associated to both applications and I extracted the correspondent adjacency matrices by means of two approaches: fixed edge density k and shuffling procedure for a significance level of 5%. This second approach was explored by applying no corrections for multiple comparisons and by applying False Discovery Rate. Several graph indexes were computed on binary adjacency matrices achieved with both methodologies. The results, achieved on the two different datasets by means of the two methods, were normalized by means of 100 random graphs with the same number of connections of the graphs obtained on simulated and mannequin data. A statistical analysis of variance (ANOVA) was performed on the results obtained by the two approaches in each dataset to study the effect of the methodology applied on the properties extracted from the networks.

Methods

Datasets used in the experiment

Simulated data. The first dataset I used to compare the two approaches was generated to build the null case (complete lack of correlation between the signals). To this purpose, I generated random datasets of signals with the same average amplitude and the same standard deviation of the data acquired on the mannequin head (see following paragraph for details) to avoid differences between the two datasets due to different signals amplitudes. In particular, each dataset is composed by 20 signals segmented in 50 trials of 3s each. 20 electrodes is the typical number of sensors used for connectivity measures estimated by means of multivariate method on scalp EEG signals. In the following, I will refer to this dataset as “simulated data”.

Mannequin data. I simulated an EEG recording on a head of a synthetic mannequin by using a 61-channel system (Brain Amp, Brainproducts GmbH, Germany). The sampling frequency was set to 200 Hz. In order to keep the impedance below the 10 k Ω , the mannequin was equipped with a cap positioned over a humidified towel. It must be noted that there were not electromagnetic sources inserted within the mannequin’s head, which is instead composed only by polystyrene. Thus, the mannequin head cannot produce any possible electromagnetic signals on the electric sensors disposed on the recording cap. The only environmental noise

was recorded. Fig.3.2 presents the experimental setup employed for the electrical recordings. The mannequin was put in front of a screen to take into account the interferences of a monitor on EEG recording. To avoid any differences between the two datasets I used the same number of trials and samples per trial of simulated data. I referred to this dataset as “mannequin data”.



Figure 3.2 – Experimental setup employed for the simulated electrical recording on a mannequin head by means of a 61 channels EEG cap. The polystyrene mannequin head was posed in front of a screen in order to include the interferences on signals due to the presence of a monitor.

Signal Processing

Both datasets were subjected to the same signal processing procedure, made by the following steps:

- 1) Generation of 20 simulated signals (simulated data) or selection of 20 channels randomly chosen among the 61 used for the recording (mannequin data)
- 2) Functional connectivity estimation performed by means of PDC (details about PDC estimator can be found in Sec.I)
- 3) Extraction of the correspondent binary adjacency matrices by applying a threshold τ achieved in two different ways:
 - by means of shuffling procedure for a significance level of 5% in two conditions: i) not corrected for multiple comparisons and ii) adjusted for multiple comparisons by False Discovery Rate.
 - by fixing the edge density k to predefined values. The levels of such values were chosen equal to those achieved by the shuffling procedure, to avoid

different performances between the two methods due to the selection of a different density of edges.

- 4) Extraction of the graph indexes described above from the adjacency matrices achieved with both methodologies
- 5) Normalization of the indexes achieved at point 4 with those extracted from 100 random graphs generated by maintaining the same number of connections of the correspondent adjacency matrix, to normalize the values to the model dimension.

Analysis of Variance

The signal processing procedure (point 1 to 5 of the previous paragraph) has been repeated 50 times in order to increase the power of the statistical test (ANOVA) computed for comparing the two different modalities used for the extraction of the adjacency matrices.

We computed a two-way ANOVA with each graph index as dependent variable. The main factors were:

- the method used for extracting adjacency matrices (METHOD), with two levels:
 - Shuffling procedure
 - Fixed Edge Density procedure
- the edge density (EDGE) corresponding to two cases:
 - Case 1: percentage of edges survived to the shuffling procedure for a significance level of 5% not corrected. This percentage resulted from the application of the Shuffling procedure and was consequently imposed also to the Fixed Edges procedure, to avoid different performances due to different densities
 - Case 2: percentage of edges survived to the shuffling procedure for a significance level of 5% corrected by FDR. Same procedure described above.

The ANOVA was applied to both simulated and mannequin data.

Results

Simulated data

To describe how I selected the edge density to be used in the two approaches, I reported in Fig.3.3 the histograms describing the distribution of the edge density characterizing the adjacency matrices extracted during different iterations of functional connectivity estimation

process on simulated data. The situation described in the two panels represents the levels Case1 (Fig. 3.3a) and Case2 (Fig 3.3b) used for the ANOVA analysis. In particular the average edge density resulting from the Shuffling procedure applied to simulated (random uncorrelated) data was 7% for the not corrected case and 4% for the FDR corrected case. These results should be interpreted taking in mind that being the simulated data a null model for connectivity estimation, all the survived links can be seen as false positives. This meant that the application of the shuffling procedure together with the FDR correction for multiple comparisons allowed to reduce the number of false positive below the significance threshold imposed in the statistical test (5% of connections could be due to chance).

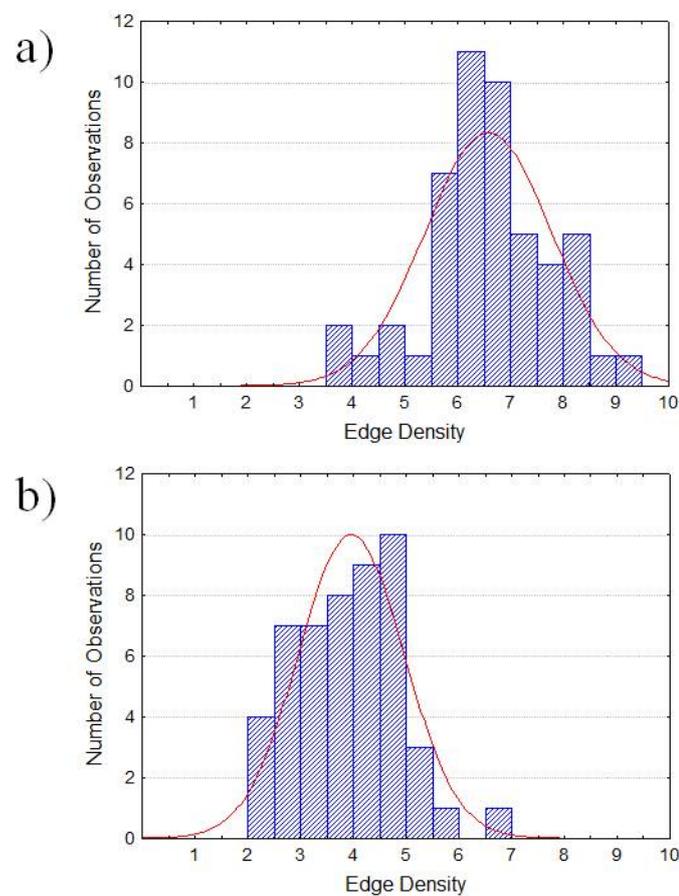


Figure 3.3 - Distribution of the edge density characterizing the adjacency matrices extracted during the different iterations of the connectivity estimation process on simulated data, in two different cases: Case 1 (a) → percentage of edges survived to shuffling procedure for a significance level of 5% not corrected for multiple comparison; Case 2 (b) → percentage of edges survived to shuffling procedure for a significance level of 5% corrected for multiple comparisons by means of FDR.

This first result confirmed the importance of statistical validation process combined with the correction for multiple comparisons. In fact, only the application of the shuffling procedure in the FDR case allowed to discard spurious links (obtained in this case on random, uncorrelated signals) at the correct level (below 5%). The edge densities obtained for the shuffling

procedure, reported in Fig.3.3, were used also in the Fixed Edges method, to avoid different performance of the two methods to be due to the different number of connections. In particular, in the Fixed Edges method, if the imposed edge density is k , the threshold was chosen as the value which allowed to keep the k higher connections of the graph.

The two approaches were statistically compared by means of an ANOVA performed imposing, as described in the Methods paragraph, each derived graph index as dependent variable. The indexes were normalized with the values obtained from 100 random graphs generated by keeping the number of connections of the correspondent adjacency matrix. This process was repeated 50 times in order to increase the robustness of the statistical analysis.

The ANOVA analysis was computed considering the Small-Worldness index as dependent variable. The within main factors were the methods used for adjacency matrices extraction (METHOD) and the edge density of the achieved adjacency matrix (EDGE). The main factor METHOD was composed by two levels: Shuffling Procedure, Fixed Edge Density Method. The main factor EDGE was composed by two levels: Case 1 (edge density associated to significance level 5%, not corrected for multiple comparisons) and Case 2 (edge density associated to significance level 5%, FDR corrected). Results revealed a statistical influence of the main factors METHOD ($p < 0.00001$, $F = 34.87$) and METHOD x EDGE ($p < 0.00001$, $F = 13.46$) on the Small-Worldness index computed on connectivity networks inferred from simulated data.

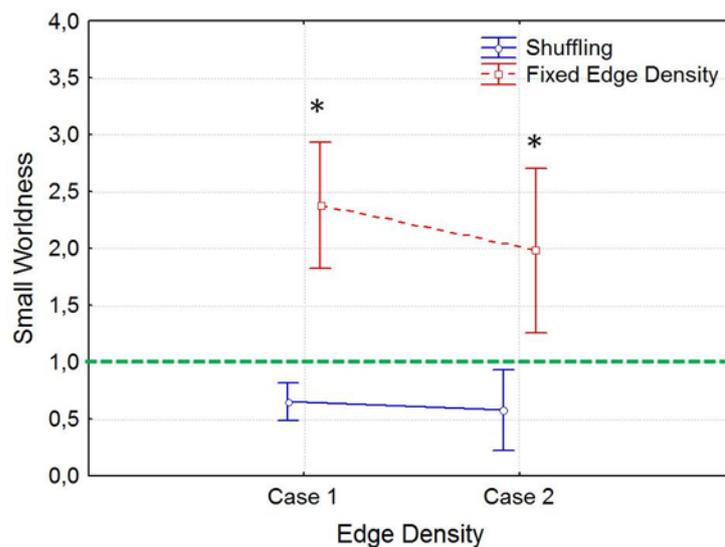


Figure 3.4 - Results of ANOVA performed on the Small-World Index computed on networks inferred from simulated data, using METHOD x EDGE as within main factor. The diagram shows the mean value for the Small-World index computed on the adjacency matrices extracted by means of the Shuffling procedure (blue line) and Fixed Edge Density Method (red line) in Case 1 (edge density as described in Fig. 3.3a) and Case 2 (edge density as described in Fig. 3.3b). The bars represent their relative 95% confidence intervals. The green dotted line represents the threshold above which a network is said to be “small-world”. The symbol (*) indicates a statistical difference between shuffling procedure and

fixed edge density method, highlighted by Tukey's post hoc test ($p < 0.05$).

In Fig. 3.4 I reported results of the ANOVA performed on the Small-World index considering METHOD x EDGE as main factor. The diagram shows the mean value for the Small-Worldness computed on adjacency matrices extracted by means of Shuffling procedure (blue line) and Fixed Edge Density (red line), from the connectivity patterns estimated on simulated data. The bar represented their relative 95% confidence interval. Considering that the edge density is equal by construction for the two methods, the diagram shows significant differences between the two methods in the description of the network in terms of Small-Worldness, confirmed by the post-hoc analysis computed by means of Tukey's test (* symbol in Fig.3.4).

Such results highlighted the importance of statistically validating connectivity patterns and of using the statistical threshold for extracting the adjacency matrix. In fact, the use of the method based on a fixed edge density revealed "small-world" properties of the network obtained from uncorrelated signals, for both density values. On the contrary, the application of the shuffling procedure allowed to correctly identify the absence of small-worldness in the network.

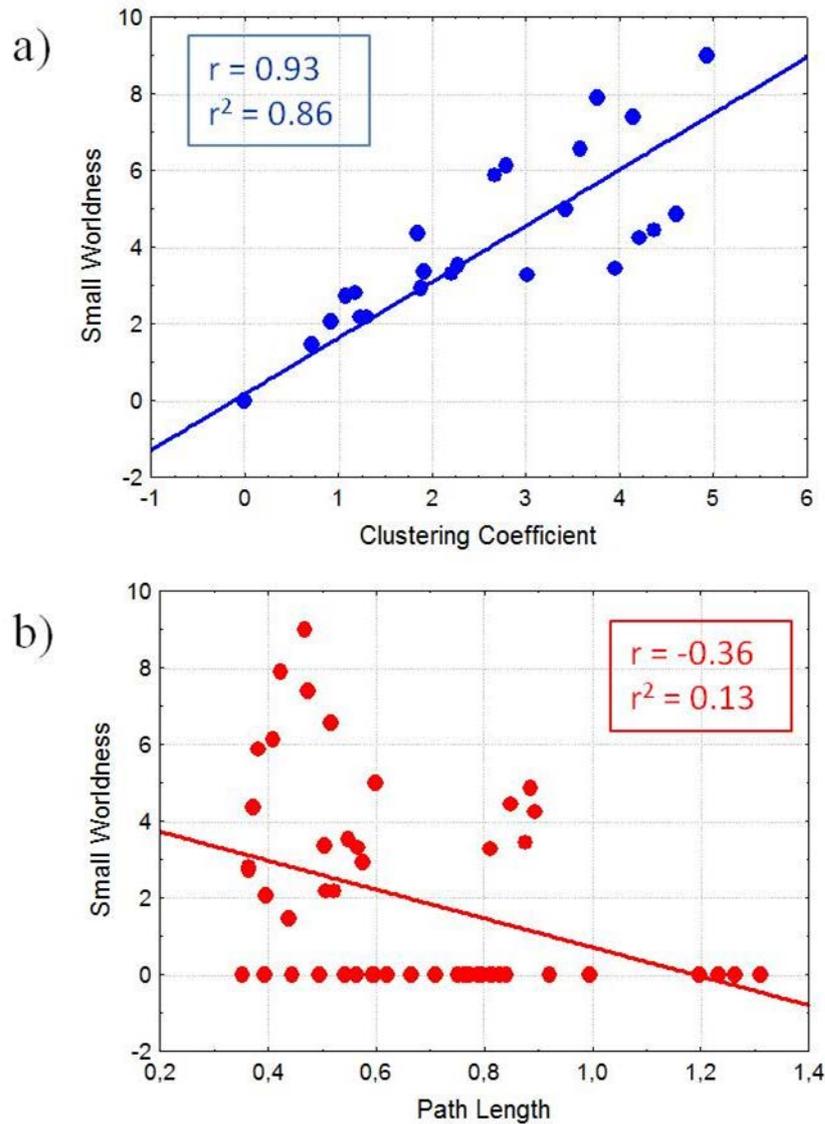


Figure 3.5 - Scatterplot of Small-World index vs Clustering Coefficient (panel a) and Small-World vs Path Length (panel b) for each iteration of the adjacency matrix extraction process computed by means of fixed edge density method for edge densities correspondent to those achieved in Case 2 (as from Fig.3.3b). The solid line represents the linear fitting computed on the data. The associated values of correlation (r) and r-square (r^2) were reported in the box.

To understand if the erroneous attribution of small-worldness to the networks achieved by means of the fixed edge density method is mainly due to the clustering coefficient or to the characteristic path length, correlations between the small-worldness index and these two indexes were computed for the two different edge densities. The results achieved in the Case 2 (edge density as Fig. 3.3b) were shown in Fig.3.5. The diagram showed the scatter plot of Small-World index vs Clustering Coefficient (Fig 3.5a) and Small-World index vs Path Length (Fig. 3.5b) for each iteration of the adjacency matrix extraction process computed by means of the fixed edge density method, in the case of edge density correspondent to those achieved in Case 2 (edge density as Fig. 3.3b). The line in the figure represents the linear

fitting computed on the data. In the box, the associated values of correlation (r) and r-square (r^2) were reported. From these results, it can be inferred that the small-world of networks achieved by means of fixed edge density method in Case 2 can be mainly due to the clustering coefficient, with a correlation of 0.93 and a r-square of 0.86. A minor dependence of small-worldness from path length index is highlighted by low values of correlation coefficient (-0.36) and r-square (0.13). The same effect can be described for Case 1 (small-worldness vs Clustering $r = 0.91$, $r^2 = 0.82$; small-worldness vs Path Length $r = -0.38$, $r^2 = 0.15$).

Mannequin data

The simulated dataset used as null model for functional connectivity estimations represents an ideal case, because it does not take into account the spatial correlation between neighboring electrodes which always occurs during an EEG recording. For this reason, I used a second dataset, composed by signals acquired simultaneously from a mannequin head equipped with a cap positioned over a humidified towel, which, with its absence of physiological signals but with its correlation between neighboring electrodes, represents the null model for connectivity inferred from signals acquired during an EEG experiment. In the second dataset I randomly selected 20 channels among the 61 acquired (same number of signals used for simulated data) and subjected them to functional connectivity estimation process. Then the correspondent adjacency matrix was extracted by means of the two considered methods and some graph indexes, such as Small-Worldness, Path Length and Clustering Coefficient were computed. The indexes were normalized with the values obtained from 100 random graphs generated by keeping the number of connections of the correspondent adjacency matrix. This process was repeated 50 times in order to increase the robustness of the following statistical analysis.

The shuffling procedure was applied for a significance level of 5%, both in the not corrected case and in the case of FDR correction. In Fig. 3.6 I reported two histograms describing the distribution of the edge density characterizing the adjacency matrices extracted during the different iterations of functional connectivity estimation process on mannequin data, for the uncorrected case (Case 1, Fig.3.6a) and for the FDR corrected case (Case 2, Fig.3.6b). In particular the average edge density was 22% for the not corrected case and 16% for the case corrected by means of FDR. This result showed the effect on connectivity measures due to the spatial correlation of neighboring electrodes. In fact the statistical validation process combined with the correction for multiple comparisons could not completely discard spurious links due to random fluctuations of the signals (residual edge density above 5%). The same

edge densities, reported in Fig.3.6, were used in the second method in order to avoid differences between the two methods due to the different number of connections.

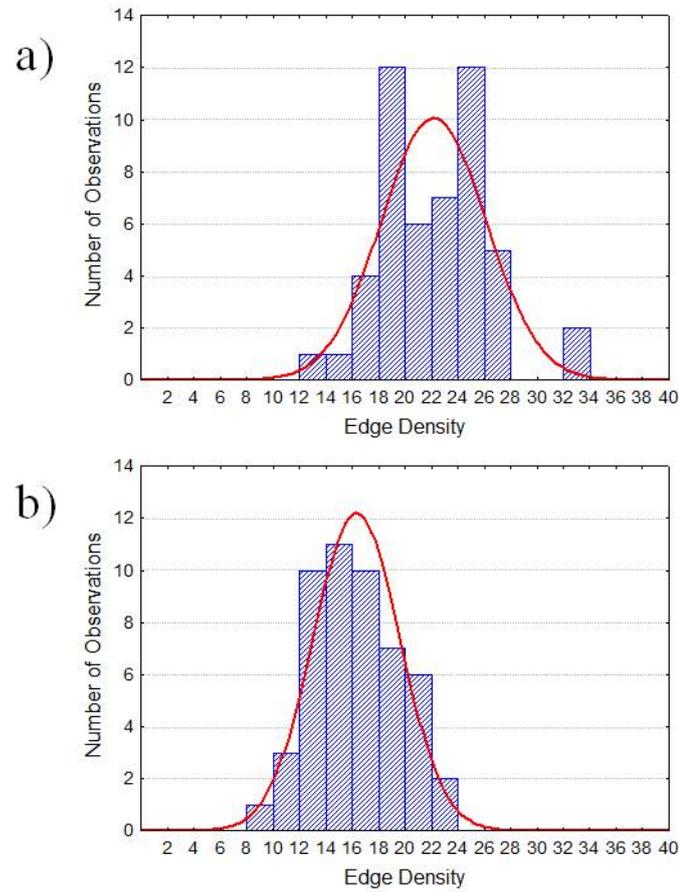


Figure 3.6 - Distribution of the edge density characterizing the adjacency matrices extracted during the different iterations of connectivity estimation process on mannequin data in two different cases: Case 1 (a) → percentage of edges survived to shuffling procedure for a significance level of 5%, not corrected for multiple comparisons; Case 2 (b) → percentage of edges survived to shuffling procedure for a significance level of 5%, FDR corrected.

The same statistical analysis described in the previous paragraph for simulated data, was computed on graph indexes extracted from mannequin networks. In particular the ANOVA was computed considering the Small-World index as dependent variable and the methods used for adjacency matrices extraction (METHOD) and the edge density of the achieved adjacency matrix (EDGE) as within main factors. The main factor METHOD was composed by two levels: Shuffling Procedure and Fixed Edge Density Method. The main factor EDGE was composed by two levels: Case 1 (edge density as in Fig.3.6a) and Case 2 (edge density as in Fig.3.6b). Results revealed statistical influence of the main factors METHOD ($p=0.00001$, $F=23.42$), EDGE ($p<0.00001$, $F=104.47$) and METHOD x EDGE ($p<0.00021$, $F=15.99$) on the Small-World index computed on connectivity networks inferred from mannequin data.

In Fig. 3.7 I reported results of the ANOVA performed on the Small-World index considering METHOD x EDGE as main factor. The diagram shows the mean value for the Small-World computed on adjacency matrices extracted, by means of Shuffling procedure (blue line) and Fixed Edge Density (red line), from the connectivity patterns estimated on mannequin data. The bar represented their relative 95% confidence interval. The Small World index is above 1 for both methodologies, with statistically higher values for Fixed Edge Density in respect to Shuffling procedure in Case 2 as confirmed by the post-hoc analysis computed by means of Tukey's.

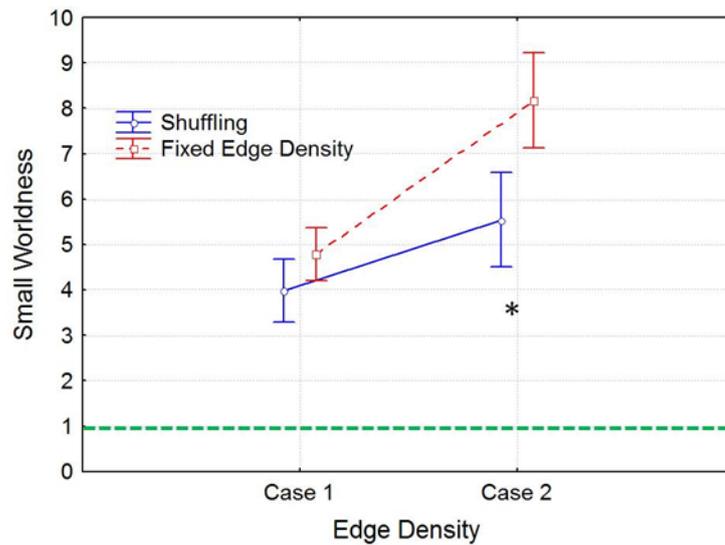


Figure 3.7 - Results of ANOVA performed on the Small-World Index computed on networks inferred from mannequin data, using METHOD and EDGE as within main factors. The diagram shows the mean value for the Small-Worldness computed on the adjacency matrices extracted by means of Shuffling procedure (blue line) and Fixed Edge Density Method (red line) in two cases, Case 1 (edge density as Fig. 3.6a) and Case 2 (edge density as Fig. 3.6b). The bar represents their relative 95% confidence interval. The green dotted line represents the threshold above which a network is said to be “small world”. The symbol (*) indicates a statistical difference between shuffling procedure and fixed edge density method, highlighted by Tukey’s post hoc test ($p < 0.05$).

In order to understand which indexes, between the clustering coefficient and the characteristic path length, mainly contributed to the small-worldness of the networks achieved by means of shuffling procedure and fixed edge density method, correlations between the small-worldness index and these two indexes was computed for the two edge density cases. The results achieved in the case of edge density correspondent to Case 2 (edge density as Fig.3.6b) were showed in Fig.3.8.

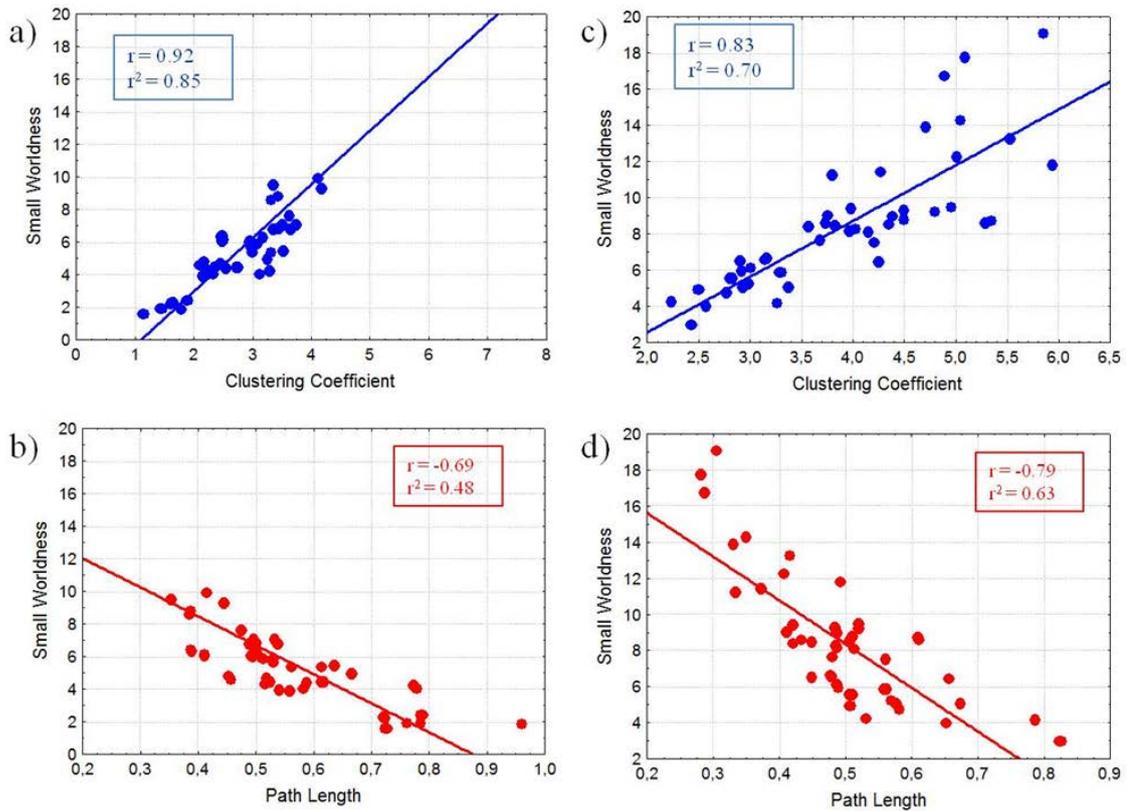


Figure 3.8 - Scatterplot of Small-World index vs Clustering Coefficient (panels a and c) and Small-World index vs Path Length (panel b and c) for each iteration of the adjacency matrix extraction process computed by means of shuffling procedures (first column) and fixed edge density method (second column) for edge densities correspondent to those achieved in Case 2 (edge density as Fig.3.6b). The line represents the linear fitting computed on the data. The associated values of correlation(r) and r-square (r^2) were reported in the boxes.

The diagram showed the scatterplot of Small World index vs Clustering Coefficient (panels a and c) and Small-World index vs Path Length (panels b and d) for each iteration of the adjacency matrix extraction process computed by means of shuffling procedure (first column) and fixed edge density method (second column) in the case of edge density correspondent to those achieved in Case 2 (edge density as Fig. 3.6b). The solid lines in the figure represent the linear fitting computed on the data. In the box the associated values of correlation(r) and r-square (r^2) were reported. The small-worldness of networks achieved by means of shuffling procedure in Case 2 can be due, at the same time, to the clustering coefficient, with a correlation of 0.92 and a r-square of 0.85 and to the path length with a correlation coefficient of -0.69 and a r-square of 0.48. Same consideration could be done for fixed edge density method (Small-Worldness vs Clustering $r = 0.83$, $r^2 = 0.70$; Small-Worldness vs Path Length $r = -0.79$, $r^2 = 0.63$). The same effect could be described for Case 1 (Shuffling procedure: Small-Worldness vs Clustering $r = 0.92$, $r^2 = 0.83$; Small-Worldness vs Path Length $r = -0.79$, $r^2 = 0.63$; Fixed Edge Density: Small-Worldness vs Clustering $r = 0.91$, $r^2 = 0.83$; Small-Worldness vs Path Length $r = -0.87$, $r^2 = 0.76$).

Discussion

The strong dependence of graph measures from the number of nodes, the edge density and the degree of the networks under analysis should lead to reflect on the modalities used for adjacency matrix extraction (Van Wijk et al., 2010). Different methodologies are currently used for this purpose, some of them based on the definition of fixed thresholds (Van den Heuvel et al., 2009; Bassett et al., 2006), other based on fixed average degree (Rubinov and Sporns, 2010; Wang et al., 2009; de Haan et al., 2009; Ferri et al., 2008; He et al., 2008), other on fixed edge density (Wang et al., 2009). The choice of a threshold in order to fix the number of edges or the degree allows to avoid size and density effects in the comparison of networks inferred from two different conditions, but can affect the structure of the network by enforcing not significant links and ignoring significant connections (Van Wijk et al., 2010). To understand the effects on the structural properties of a network due to the method applied for adjacency matrix extraction I computed a statistical comparison between one of the methods most extensively used in graph theory applications for extracting adjacency matrices from brain connectivity patterns (i.e. the method based on fixing the edge density) with an approach based on the statistical validation of achieved connectivity patterns by means of a shuffling procedure. The comparison was performed on two different datasets, one composed by random and uncorrelated simulated data, modeling the null case for the connectivity estimates, and another one composed by signals acquired on a mannequin head in order to take into account the spatial correlation between neighboring electrodes (Nunez et al., 1997). The results presented in this section allow to discuss about some open problems which affect the application of graph measures to the functional connectivity estimates.

The first issue addressed in the present study was the necessity to statistically validate the connectivity measures in order to discard the spurious links due to random fluctuations of the signals considered simultaneously in the multivariate (Takahashi et al., 2007; Schelter et al., 2006; Kaminski et al., 2001) or bivariate model (Amjad et al., 1997; Rosenberg et al., 1989). In this paper I confirmed the importance of the statistical validation combined with the corrections for multiple comparisons in multivariate estimates (Toppi et al., 2011) by showing the edge densities survived to the shuffling procedure on the simulated data (Fig.3.3). Being the simulated data a null model for connectivity estimation, all the survived links can be seen as false positives. The application of the shuffling procedure for a significance level of 5% not corrected produces 7% of false positives. Only applying the statistical correction of FDR the false positives went down the threshold of 5%. Unfortunately the application of shuffling

procedure to connectivity networks inferred on mannequin data led to a high number of false positives (22% in the not corrected case and 16% in the case of FDR correction). This could be explained by taking into account that some of the survival links are due to real correlations between neighboring electrodes induced by the registration on a wet towel, but which can occur also in real EEG recordings (Haueisen et al., 2002; van den Broek et al., 1998).

A second issue to be considered as relevant in graph theory concerned the modality in which the adjacency matrix is extracted from the connectivity network. As already said in the previous paragraphs, the threshold choice is crucial for the computation of graph measures because it affects the topographical properties of real networks. In the present study I made a comparison between one of the methods extensively used in graph theory applications for extracting adjacency matrices from the connectivity patterns (i.e. the method based on fixing the edge density) and an approach based on the statistical validation of achieved connectivity patterns by means of a shuffling procedure, to describe the effects of the modalities for adjacency matrix extraction on the “small world” properties of the network. The results achieved on simulated data highlighted small world properties of the analyzed networks even in random, uncorrelated data, when the fixed edge density method was applied. Such small-worldness is mainly correlated with an increase of the clustering coefficient and disappeared when shuffling procedure was used. The fixed edge density criterion led to an erroneous diagnosis of small-worldness for the connectivity patterns estimated on simulated data, independently from the edge density chosen. In fact, the simulated data, being uncorrelated, should produce connectivity patterns without any topographical properties of small-worldness. These results led to two conclusions. The first is that the shuffling procedure does not just preserve the strongest connections, as demonstrated by different results obtained by means of fixed edge density which is based on this criterion. It means that the significance of a link is not merely related to its strength. The second conclusion is that the choice of an empirical threshold can affect so much the topography of the network that an erroneous definition of small-worldness could result. Thus, a statistical validation, combined with multiple comparisons adjustments, to be applied on connectivity networks, is necessary to define the significance of each edge within the adjacency matrix, in order to extract graph measures able to describe the real properties of the considered network.

The results achieved on mannequin data showed small-world properties of the networks extracted by applying both methodologies. In this case, the shuffling procedure could not prevent the description of mannequin networks as small world networks, even applying the

corrections for multiple comparisons, but the entity of small-worldness is lower than those achieved by means of fixed edge density method. In both cases the Small-Worldness is equally correlated with an increase of the clustering coefficient and with a decrease of the path length. This effect could be explained with the existence of real correlations between electrodes, which can occur in real EEG data, due to volume conduction effect and to the location of the reference (Gerhard et al., 2011; Bialonski et al., 2010; Nunez et al., 1997). These considerations led to a possible redefinition of the meaning of the Small-World Index. In fact, it cannot be considered as an absolute measure, because its value contains some of the real correlations due to neighboring electrodes. A possible solution is to consider only variations of this measure between two conditions within the same subject, or between two subjects in the same conditions, in order to discard all the effects due to the position of the electrodes on the scalp. Another way to mitigate such effect is to apply the connectivity estimation process on the data obtained by methods allowing to reduce the spatial correlation between electrodes, such as all the approaches for the reconstruction of cortical sources from high resolution EEG recordings (Brancucci et al., 2004; Babiloni et al., 2004, 2003, 2001). Such methodologies allow to focus the activations of cerebral sources by means of a high number of sensors, realistic head models and the solution of the associated linear inverse problem (Astolfi et al., 2007, 2005, 2004). It must be also noted that other methods used for reducing spatial correlation at the scalp level, such as Blind Source Separation and Superficial Laplacian (Grave de Peralta Menendez and Gonzalez Andino, 1999), cannot be applied on the data before being subjected to connectivity analysis. In fact their mathematical algorithms introduce correlation in the data, which would, in turn, produces spurious results.

In conclusion the present study allowed to highlight some erroneous results which can be obtained by the application of commonly used approaches for the extraction of adjacency matrix from connectivity patterns. Only the statistical validation of investigated connectivity patterns allowed to reconstruct the real topography of the network and to extract reliable indexes able to describe its salient properties without errors due to the instability of the methodologies applied.

Conclusion

During the last ten years the use of effective connectivity for the description of some neurological phenomena at the basis of cerebral processes has gained a lot of success in different neuroscience applications. In fact, the estimation of the communication flows between different parts of the brain provides further information on the investigated cerebral processes in respect to those achieved by means of traditional brain mapping tools.

The use of effective connectivity in complex neuro-scientific fields led to the necessity to find quantitative descriptors of the main properties of the investigated networks with the aim to synthesize the huge amount of information derived from the application of such advanced methodologies. For this reason, recently, the graph theory approach has been applied to brain networks analysis in order to extract indexes describing salient properties of the achieved connectivity patterns.

As always happens, the translation of a methodology from a mathematical context to a more practical application could not be realized without the adaptation of all the methodologies used. Such adaptation required the refining of the existent approaches and the development of new methods able to solve the open problems in the field.

Notwithstanding the introduction of graph theory approach in connectivity field has started several years ago, a consistent procedure for the extraction of salient properties of investigated connectivity patterns has not been developed yet. In literature different approaches were reported without any demonstration of their validity. The instability of such approaches could lead to misleading results in the extracted graph indexes. Such aspect is still underestimate, if we consider that graph theory indexes are used in the majority of papers reporting studies based on effective connectivity appeared in the latest years.

The necessity of stable and reliable indexes characterizing global and local features of connectivity patterns led to the definition of an approach, based on the use of intrinsic significance of connectivity links, for deriving the adjacency matrix. This approach guarantees that the adjacency matrix obtained is composed only by significant edges and thus reflects the real properties of the investigated network.

The simulation study reported in this Section demonstrated how the procedure for the threshold selection, as commonly used in literature, leads to erroneous descriptions of the properties of networks under investigation. The use of statistical threshold, also derived by

applying corrections for multiple comparisons, mitigated such phenomenon providing correct characterization of connectivity patterns.

The development of such procedure allowed to significantly improve the reliability of the graph theory indexes extraction, thus opening the way to clinical applications in which the reliability of indexes is mandatory.

The study revealed also an important issue regarding the small-world properties of a network. In fact the small-world index has been used since ten years ago in different neuroscience application without taking into account that it is influenced by the correlation of neighboring electrodes used in EEG recordings. This discovery changes the meaning associated to this particular property of real networks. In fact, small-worldness can thus be defined only as variation between two experimental conditions but not as absolute value to be compared with random graphs.

The establishment of a consistent procedure for the characterization of local and global properties of the networks is only one step in the development of this new field. Next steps will be towards the definition of new indexes and in their statistical validation with respect to the chance level.

Section III References

- Amjad, A.M., Halliday, D.M., Rosenberg, J.R., Conway, B.A., 1997. An extended difference of coherence test for comparing and combining several independent coherence estimates: theory and application to the study of motor units and physiological tremor. *J. Neurosci. Methods* 73, 69–79.
- Astolfi, L., Cincotti, F., Babiloni, C., Carducci, F., Basilisco, A., Rossini, P.M., Salinari, S., Mattia, D., Cerutti, S., Dayan, D.B., Ding, L., Ni, Y., He, B., Babiloni, F., 2005. Estimation of the cortical connectivity by high-resolution EEG and structural equation modeling: simulations and application to finger tapping data. *IEEE Trans Biomed Eng* 52, 757–768.
- Astolfi, L., Cincotti, F., Mattia, D., Salinari, S., Babiloni, C., Basilisco, A., Rossini, P.M., Ding, L., Ni, Y., He, B., Marciani, M.G., Babiloni, F., 2004. 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, 1457–1470.
- Astolfi, L., De Vico Fallani, F., Cincotti, F., Mattia, D., Marciani, M.G., Bufalari, S., Salinari, S., Colosimo, A., Ding, L., Edgar, J.C., Heller, W., Miller, G.A., He, B., Babiloni, F., 2007. Imaging functional brain connectivity patterns from high-resolution EEG and fMRI via graph theory. *Psychophysiology* 44, 880–893.
- Babiloni, C., Babiloni, F., Carducci, F., Cincotti, F., Rosciarelli, F., Rossini, P., Arendt-Nielsen, L., Chen, A., 2001. Mapping of early and late human somatosensory evoked brain potentials to phasic galvanic painful stimulation. *Hum Brain Mapp* 12, 168–179.
- Babiloni, C., Babiloni, F., Carducci, F., Cincotti, F., Vecchio, F., Cola, B., Rossi, S., Miniussi, C., Rossini, P.M., 2004. Functional frontoparietal connectivity during short-term memory as revealed by high-resolution EEG coherence analysis. *Behav. Neurosci.* 118, 687–697.
- Babiloni, C., Brancucci, A., Babiloni, F., Capotosto, P., Carducci, F., Cincotti, F., Arendt-Nielsen, L., Chen, A.C.N., Rossini, P.M., 2003. Anticipatory cortical responses during the expectancy of a predictable painful stimulation. A high-resolution electroencephalography study. *Eur. J. Neurosci.* 18, 1692–1700.
- Baccalá, L.A., Sameshima, K., 2001. Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics* 84, 463–474.

- Bassett, D.S., Meyer-Lindenberg, A., Achard, S., Duke, T., Bullmore, E., 2006. Adaptive reconfiguration of fractal small-world human brain functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19518–19523.
- Bialonski, S., Horstmann, M.-T., Lehnertz, K., 2010. From brain to earth and climate systems: small-world interaction networks or not? *Chaos* 20, 013134.
- Brancucci, A., Babiloni, C., Babiloni, F., Galderisi, S., Mucci, A., Tecchio, F., Zappasodi, F., Pizzella, V., Romani, G.L., Rossini, P.M., 2004. Inhibition of auditory cortical responses to ipsilateral stimuli during dichotic listening: evidence from magnetoencephalography. *Eur. J. Neurosci.* 19, 2329–2336.
- Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198.
- Chen, Z.J., He, Y., Rosa-Neto, P., Germann, J., Evans, A.C., 2008. Revealing Modular Architecture of Human Brain Structural Networks by Using Cortical Thickness from MRI. *Cereb Cortex* 18, 2374–2381.
- De Haan, W., Pijnenburg, Y.A.L., Strijers, R.L.M., Van der Made, Y., Van der Flier, W.M., Scheltens, P., Stam, C.J., 2009. Functional neural network analysis in frontotemporal dementia and Alzheimer's disease using EEG and graph theory. *BMC Neurosci* 10, 101.
- De Vico Fallani, F., Toppi, J., Di Lanzo, C., Vecchiato, G., Astolfi, L., Borghini, G., Mattia, D., Cincotti, F., Babiloni, F., 2012. Redundancy in Functional Brain Connectivity from EEG Recordings. *International Journal of Bifurcation and Chaos* 22, 1250158.
- Fagiolo, G., 2007. Clustering in complex directed networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 76, 026107.
- Ferrarini, L., Veer, I.M., Baerends, E., Van Tol, M.-J., Renken, R.J., Van der Wee, N.J.A., Veltman, D.J., Aleman, A., Zitman, F.G., Penninx, B.W.J.H., Van Buchem, M.A., Reiber, J.H.C., Rombouts, S.A.R.B., Milles, J., 2009. Hierarchical functional modularity in the resting-state human brain. *Hum Brain Mapp* 30, 2220–2231.
- Ferri, R., Rundo, F., Bruni, O., Terzano, M.G., Stam, C.J., 2008. The functional connectivity of different EEG bands moves towards small-world network organization during sleep. *Clin Neurophysiol* 119, 2026–2036.
- Freeman, L.C., 1978. Centrality in social networks conceptual clarification. *Social Networks* 1, 215–239.

- Gerhard, F., Pipa, G., Lima, B., Neuenschwander, S., Gerstner, W., 2011. Extraction of Network Topology From Multi-Electrode Recordings: Is there a Small-World Effect? *Front Comput Neurosci* 5, 4.
- Grave de Peralta Menendez, R., Gonzalez Andino, S.L., 1999. Backus and Gilbert method for vector fields. *Hum Brain Mapp* 7, 161–165.
- Grosse-Wentrup, M., Schölkopf, B., 2012. High gamma-power predicts performance in sensorimotor-rhythm brain-computer interfaces. *J Neural Eng* 9, 046001.
- Hauelsen, J., Tuch, D.S., Ramon, C., Schimpf, P.H., Wedeen, V.J., George, J.S., Belliveau, J.W., 2002. The Influence of Brain Tissue Anisotropy on Human EEG and MEG. *NeuroImage* 15, 159–166.
- He, Y., Chen, Z., Evans, A., 2008. Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer’s disease. *J. Neurosci.* 28, 4756–4766.
- Humphries, M.D., Gurney, K., 2008. Network “small-world-ness”: a quantitative method for determining canonical network equivalence. *PLoS ONE* 3, e0002051.
- Kaminski, M., Ding, M., Truccolo, W.A., Bressler, S.L., 2001. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological Cybernetics* 85, 145–157.
- Kintali, S., 2008. Betweenness Centrality : Algorithms and Lower Bounds. arXiv:0809.1906.
- Latora, V., Marchiori, M., 2001. Efficient behavior of small-world networks. *Phys. Rev. Lett.* 87, 198701.
- Nunez, P.L., Srinivasan, R., Westdorp, A.F., Wijesinghe, R.S., Tucker, D.M., Silberstein, R.B., Cadusch, P.J., 1997. EEG coherency. I: Statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalogr Clin Neurophysiol* 103, 499–515.
- Rosenberg, J.R., Amjad, A.M., Breeze, P., Brillinger, D.R., Halliday, D.M., 1989. The Fourier approach to the identification of functional coupling between neuronal spike trains. *Prog. Biophys. Mol. Biol.* 53, 1–31.
- Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069.

- Salvador, R., Suckling, J., Coleman, M.R., Pickard, J.D., Menon, D., Bullmore, E., 2005. Neurophysiological architecture of functional magnetic resonance images of human brain. *Cereb. Cortex* 15, 1332–1342.
- Schelter, B., Winterhalder, M., Eichler, M., Peifer, M., Hellwig, B., Guschlbauer, B., Lücking, C.H., Dahlhaus, R., Timmer, J., 2006. Testing for directed influences among neural signals using partial directed coherence. *Journal of Neuroscience Methods* 152, 210–219.
- Sporns, O., Chialvo, D.R., Kaiser, M., Hilgetag, C.C., 2004. Organization, development and function of complex brain networks. *Trends in Cognitive Sciences* 8, 418–425.
- Stam, C.J., 2004. Functional connectivity patterns of human magnetoencephalographic recordings: a “small-world” network? *Neurosci. Lett.* 355, 25–28.
- Stam, C.J., Reijneveld, J.C., 2007. Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomed Phys* 1, 3.
- Strogatz, S.H., 2001. Exploring complex networks. *Nature* 410, 268–276.
- Takahashi, D.Y., Baccalà, L.A., Sameshima, K., 2007. Connectivity Inference between Neural Structures via Partial Directed Coherence. *Journal of Applied Statistics* 34, 1259–1273.
- Toppi, J., Babiloni, F., Vecchiato, G., Cincotti, F., De Vico Fallani, F., Mattia, D., Salinari, S., Astolfi, L., 2011. 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.
- Toppi, J., Petti, M., De Vico Fallani, F., Vecchiato, G., Maglione, A.G., Cincotti, F., Salinari, S., Mattia, D., Babiloni, F., Astolfi, L., 2012. Describing relevant indexes from the resting state electrophysiological networks. *Conf Proc IEEE Eng Med Biol Soc 2012*, 2547–2550.
- Van den Broek, S., Reinders, F., Donderwinkel, M., Peters, M., 1998. Volume conduction effects in EEG and MEG. *Electroencephalography and Clinical Neurophysiology* 106, 522–534.
- Van den Heuvel, M.P., Stam, C.J., Kahn, R.S., Hulshoff Pol, H.E., 2009. Efficiency of functional brain networks and intellectual performance. *J. Neurosci.* 29, 7619–7624.
- Van Wijk, B.C.M., Stam, C.J., Daffertshofer, A., 2010. Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory. *PLoS One* 5.

- Wang, J., Wang, L., Zang, Y., Yang, H., Tang, H., Gong, Q., Chen, Z., Zhu, C., He, Y., 2009. Parcellation-dependent small-world brain functional networks: a resting-state fMRI study. *Hum Brain Mapp* 30, 1511–1523.
- Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of “small-world” networks. *Nature* 393, 440–442.
- Zhu, C., Guo, X., Jin, Z., Sun, J., Qiu, Y., Zhu, Y., Tong, S., 2011. Influences of brain development and ageing on cortical interactive networks. *Clin Neurophysiol* 122, 278–283.

Section IV

Functional Connectivity for the Study of Resting State and Cognitive Processes in Human

INTRODUCTION

DESCRIBING RELEVANT INDEXES FROM THE RESTING STATE ELECTROPHYSIOLOGICAL NETWORKS

Introduction

Material and Methods

Experimental Design

Pre-Processing and Functional Connectivity Analysis

Results

Discussion

FUNCTIONAL CONNECTIVITY FOR INVESTIGATING THE BASIS OF COGNITIVE PROCESSES

Introduction

Material and Methods

Experimental Design

Visual Oddball Task

Sternberg Task

Pre-Processing of EEG traces

Results

EEG Visual Oddball Task

EEG Sternberg Task

Discussion

CONCLUSION

SECTION IV REFERENCES

Introduction

In the previous sections of this PhD thesis several methodological issues were studied and possible solutions to the existent limitations were provided. All the described approaches showed high performances if tested by means of surrogate data simulating different conditions for EEG signals. Once demonstrated their performances by means of simulation studies, their application to real data is necessary in order to quantify and verify their action field. For this reason in the present section of this PhD thesis I proposed two clinical applications in which the real possibilities of investigated methods could be tested. In particular, I reported the results achieved in the application of effective connectivity for the

study of resting state networks and complex cognitive functions in human. The two studies had the double aim of testing the new methodologies on real EEG data acquired during complex tasks from one side and of studying the cerebral processes at the basis of human conditions not completely investigated from the other side.

In particular, I selected two applications which, for the limitations in the available methodologies, could be investigated only partially in the past, and with results not always consistent among the population. By contrast, applying the methods developed in this PhD work, the results obtained showed a significant consistency across the population, as will be showed by the next paragraphs.

Describing Relevant Indexes from the Resting State Electrophysiological Networks

Introduction

Recent studies in fMRI field highlighted the existence of a “Default Mode network” (DMN) characterizing the brain functions during the rest condition (Greicius et al., 2003; Raichle et al., 2001). The DMN concept was defined as a consistent pattern of deactivation of some regions (precuneus/posterior cingulate cortex (PCC), medial prefrontal cortex (MPFC) and medial, lateral and inferior parietal cortex) which occurs during the initiation of task-related activity. Such network is active during the resting state activity in which an individual is awake and alert, but not actively involved in an attention demanding or goal-directed task (Raichle et al., 2001). It has been demonstrated that the deactivation is correlated with the attentive load required for the task execution. In fact, the more demanding the task, the stronger the deactivation appears to be (Singh and Fawcett, 2008; McKiernan et al., 2006). Moreover, some DMN abnormalities could be also put in relation with a number of different mental disorders (Broyd et al., 2009).

Several electroencephalographic (EEG) studies tried to describe some electrophysiological properties of DMN by correlating the spectral activity in a specific band with a specific group of brain areas belonging to the highlighted network (Chen et al., 2008; Mantini et al., 2007). Despite these efforts, an electrophysiological correlate of the DMN discovered in fMRI field has not yet been found. Its neuroelectrical description could be useful for describing the general properties of a healthy population to be used in the characterization of neurological diseases and in the description of the recovery from a pathological state after specific motor/cognitive trainings.

The aim of this study was to use advanced techniques for functional connectivity estimation for the description of the electrophysiological properties of resting state condition in human. The idea was to extract some salient indexes borrowed from graph theoretical approach for characterizing the connectivity networks elicited by a population of 55 healthy subjects during the rest condition. Once evaluating their uniformity among the population, the values achieved for such indexes were included in a normative database to be used as reference for pathological conditions.

Material and Methods

Experimental Design

55 healthy subjects (age: 27.5 ± 5 years; 25 female) participated in the study. The experiment consisted in 2 minutes of recording session during which the subjects were asked to stay in a rest condition for one minute with their eyes closed and one minute with their eyes opened, without moving or performing any mental activities. A 64-channel system with a sampling frequency of 200 Hz (BrainAmp, Brainproducts GmbH, Germany) was used to record EEG data.

Pre-Processing and Functional Connectivity Analysis

EEG signals were band-pass filtered (1-45 Hz + 50 Hz Notch filter) and ocular artifacts were removed by means of Independent Component Analysis. EEG traces were segmented in epochs of 1s each in order to increase the robustness of methodologies applied in the following and residual artifacts were removed. A subset of 12 channels (spatially distributed on the scalp) among the 64 channels used for the recording (Fp1, Fp2, F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, O2) were selected and functional connectivity was estimated by means of PDC estimator. The achieved connectivity patterns were statistically validated by means of asymptotic statistic method for a significance level of 5% FDR corrected, whose higher performances have been already demonstrated in Sec.I (Toppi et al., 2011). Validated connectivity patterns were then averaged in six frequency bands defined according to the Individual Alpha Frequency (IAF) (Klimesch, 1999). The achieved adjacency matrices were used for extracting several graph indexes.

Thus in order to characterize the features of the networks elicited during the rest condition I computed a statistical study as described in the follow:

- 1) Extraction of several graph indexes from the resting state networks achieved for each subject
- 2) Generation of 100 random graphs for each achieved network. Random graphs were generated by keeping the number of edges of the corresponding real network.
- 3) The same graph indexes used in resting networks were extracted from random networks and averaged across the 100 repetitions.
- 4) Application of a statistical test, repeated measures ANOVA, for comparing the indexes extracted from real and random networks under different conditions.

The comparison between real and random networks allowed to characterize the salient properties of real networks, discarding all what was given by chance or multiple tests.

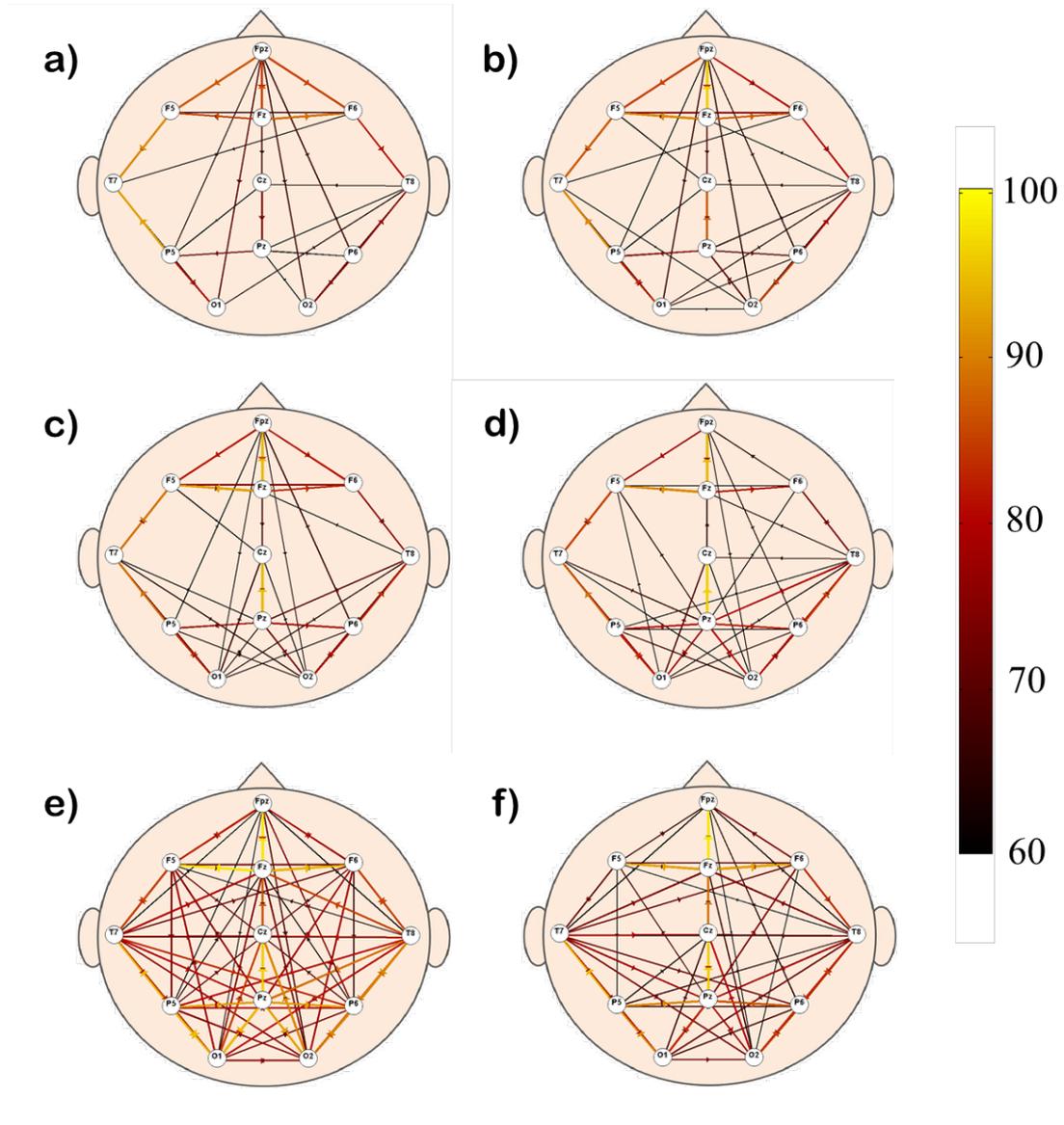


Figure 4.1 - Grand Average of functional connectivity patterns elicited during the eyes-closed condition by 55 healthy subjects in 6 frequency bands defined according to IAF: a) Theta Band (IAF-6, IAF-4), b) Lower1-Alpha Band (IAF-4, IAF-2), c) Lower2-Alpha Band (IAF-2, IAF), d) Upper Alpha Band (IAF, IAF+2), e) Beta Band (IAF+2, IAF+15), f) Gamma Band (IAF+15, IAF+30). Connectivity patterns are represented on a scalp model seen from above with the nose pointing the upper part of the page. Color and size of each arrow code the percentage of subjects who elicited the correspondent connection.

Results

After pre-processing, EEG traces related to the resting condition of 55 healthy subjects were subjected to functional connectivity estimation by means of PDC. The optimal order of the correspondent MVAR model was estimated by means of Akaike Information Criterion (AIC) for each subject. The achieved estimations were averaged within six frequency bands defined

according to IAF ($IAF = 10.25 \pm 0.86$). Fig. 4.1 showed the Grand Average of functional connectivity patterns elicited during the eyes-closed condition by 55 healthy subjects in 6 frequency bands: a) Theta Band (IAF-6, IAF-4), b) Lower1-Alpha Band (IAF-4, IAF-2), c) Lower2-Alpha Band (IAF-2, IAF), d) Upper Alpha Band (IAF, IAF+2), e) Beta Band (IAF+2, IAF+15), f) Gamma Band (IAF+15, IAF+30). Connectivity patterns are represented on a scalp model seen from above with the nose pointing the upper part of the page. Color and size of each arrow code the percentage of subjects who elicited the correspondent connection. The figure highlighted the existence of a consistent pattern elicited by more than 60% of the analyzed population. Such pattern was characterized by a sub-network between electrodes in frontal regions in theta, lower1 and lower2 alpha, by a role of Cz as source of information in all bands and by a sub-network between electrodes located in parietal areas in upper alpha, beta and gamma bands.

In order to extract salient properties for describing the achieved connectivity patterns, I computed graph theory indexes on their associated adjacency matrices. The adjacency matrices were extracted by means of statistical validation of connectivity patterns as demonstrated in Sec.III. In fact, the asymptotic statistic method allowed to building an adjacency matrix whose entries were 1 if the connection resulted significantly different from the null case and 0 if not. Once extracted the correspondent adjacency matrix for each subject and for each band, several indexes, such as global and local efficiencies, symmetries and influences between the two hemispheres or between the anterior and posterior parts of the scalp, were computed.

GRAPH INDEXES	TYPE		TYPE x BAND	
	F	p	F	p
Global efficiency	43.8	< 0.00001	8.5	< 0.00001
Local efficiency	153.9	< 0.00001	16.02	< 0.00001
$S_{\text{Left-Right}}$	1.51	0.224	0.18	0.97
$I_{\text{Left-Right}}$	1.85	0.18	1.28	0.27
$S_{\text{Anterior-Posterior}}$	11.73	< 0.00001	10.76	< 0.00001
$I_{\text{Anterior-Posterior}}$	2.31	0.13	13.67	< 0.00001

Table 4.1 - Results of the two-way ANOVA performed considering as main factors the type of graph (TYPE: real, random) and the band (BAND: theta, lower1-alpha, lower2-alpha, upper alpha, beta, gamma) and as dependent variables all the extracted graph indexes

In order to validate the values achieved for the graph indexes, 100 random graphs with the same number of connections were generated for each adjacency matrix. On each random graph indexes described above were computed and averaged among the iterations. Then a statistical comparison between indexes values achieved on real and random graphs was computed.

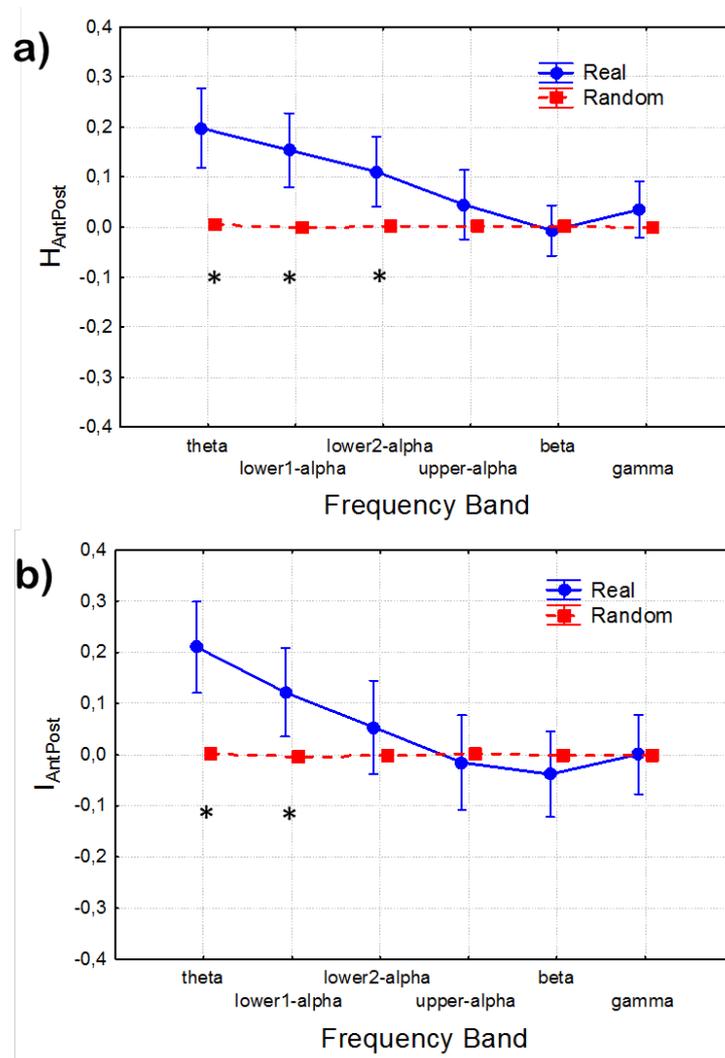


Figure 4.2. Results of ANOVA performed on the indexes Symmetry (a) and Influence (b) between anterior and posterior parts of the scalp computed on resting and random network: plot of means with respect to the factor TYPE x BAND. ANOVA shows a high statistical significance ($F=10.76, p<0.0001$) for the case a) and ($F=13.67, p<0.0001$) for the case b) respectively. The bar on each point represents the 95% confidence interval of the mean errors computed across the subjects. The symbol * codes a significant difference between resting and random networks revealed by Duncan's post-hoc test.

In particular, for each index, a two way ANOVA was computed considering as main factors the type of graph (TYPE: real, random) and the band (BAND: theta, lower1-alpha, lower2-alpha, upper alpha, beta, gamma). The results of each ANOVA were reported in Tab.4.1. In the first column the dependent variables are reported. In the other two columns the results (F and p) related to the effect of main factors TYPE and TYPE x BAND were showed. The table showed a significant influence of factors TYPE and TYPE x BAND on global and local

efficiencies. In particular, the Duncan's post-hoc analysis revealed that the global efficiency of the network elicited during the rest condition is lower than those achieved with random graph for all bands. Instead a higher local efficiency was showed in resting network in respect to random patterns for all bands. No effects of factors TYPE and TYPE x BAND on symmetry and influence between the two hemispheres resulted from the ANOVA. Both indexes remained around zero for both resting and random networks, indicating a complete symmetry between the two hemispheres. Moreover, Tab.4.1 showed statistical influence of the factors TYPE and TYPE x BAND on asymmetry index computed between anterior and posterior parts of the scalp and of the factor TYPE x BAND on the influence index achieved on the same scalp regions.

Fig.4.2 showed the influence of the different levels of the main factor TYPE x BAND on the indexes Symmetry (panel a) and Influence (panel b) computed between the anterior and posterior parts of the brain. The bar on each point represents the 95% confidence interval of the mean errors computed across the subjects. In panel a Duncan's post hoc analysis (* symbol) revealed differences between resting and random networks in theta, lower1-alpha, beta and gamma bands. In particular, in theta, lower1-alpha and lower2-alpha bands a significant higher number of connections in the anterior part of the scalp resulted in respect to the posterior part ($S > 0$). Complete asymmetry was showed in upper-alpha, beta and gamma bands ($S \cong 0$). In panel b Duncan's post hoc analysis revealed differences in influence index between resting and random networks in theta and lower1-alpha bands. The positive values achieved for such index revealed a high influence of the anterior part on the posterior part of the scalp for all these bands.

In order to verify the homogeneity of the extracted graph indexes among the population I applied the cross-validation process reported in Fig.4.3. In particular, I randomly selected a subject among the population and I compared by means of z-score each index of his networks with the corresponding average computed on the remaining population. I repeated the process for all the subjects and for each index I counted all the times in which no significant differences were highlighted between the considered subject and the population. The percentages achieved with this process were showed in the bar diagram of Fig.4.4. For each index is reported the percentage of subjects resulted no significantly different from the population according to the considered index. All the indexes are consistent among the population because no significant differences resulted in more than 90% of population. The homogeneity of these indexes among the population allowed to use them for the creation of a

database of parameters describing the healthy resting condition to be compared with pathological conditions.

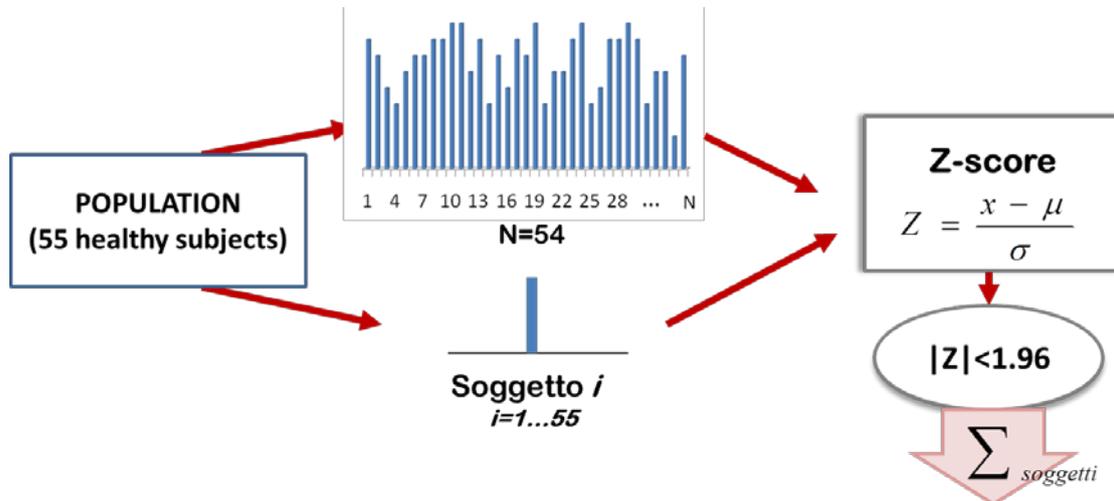


Figure 4.3 – Flowchart of the Cross-validation process used for verifying the homogeneity of the extracted indexes

Discussion

The application of advanced methods for functional connectivity estimation to EEG signals and the use of methods for statistically validate the achieved connectivity patterns also taking into account corrections for multiple comparisons allowed to reconstruct the electrophysiological properties of DMN. The use of PDC, its relative statistical validation and the graph theory approach led to the definition of a consistent neuroelectrical network elicited by a population of 55 healthy subjects during the rest condition. Such network, characteristic of more than 70% of the population, involved mainly the frontal part of the scalp in the lower bands (theta, lower1 and lower2 alpha) and the parietal part of the scalp in the higher bands (beta, gamma) as highlighted by means of grand average connectivity patterns. This result was confirmed by the statistical comparison between resting and random networks in terms of influences and symmetries between anterior and posterior part of the scalp. The same statistical study stated also the existence of a perfect symmetry between the two hemispheres during the rest condition. The ANOVA study highlighted also the small-world properties of the network elicited during the rest condition, characterized by a high local efficiency and a low global efficiency in respect to random graph.

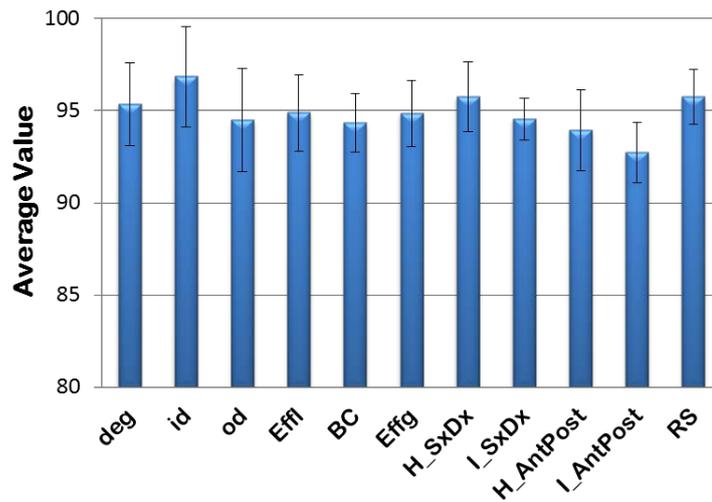


Figure 4.4 – Bar diagram reporting for each index the percentage of subjects resulted no significantly different from the population according to the considered index

The cross-validation study verified the consistency of the results of statistical analysis, in particular the values achieved for all the indexes computed on resting networks are uniform among the population, thus they could be used to build a normative database to be considered in the identification of pathological conditions.

Moreover, the methodological steps performed for the analysis of DMN and the new graph indexes, influence and asymmetry, defined for the purposes of this work, have been revealed as valid procedure for the description of the neuroelectrical properties of any neurological condition, not only the resting state.

Functional Connectivity for investigating the basis of cognitive processes

Introduction

Several studies demonstrated that the processes at the basis of attentive or memory processes involve not isolate and specific cerebral areas but groups of brain areas strictly connected each other. The communication between such areas is characterized by a specific timing and is subjected to a temporal evolution strictly linked to the explicated cognitive function.

For this reason, the study of complex cerebral mechanisms such as those at the basis of cognitive process required methodologies able to describe phenomena evolving in time and which globally involve the brain in terms of functional networks. The methodologies available since few years ago did not allow to follow the temporal evolution of cerebral networks with sufficient accuracy also taking into account all the sources involved in the processes (Möller et al., 2001). The methodological advancements achieved during my PhD provided an important contribution to study the complex mechanisms at the basis of cognitive processes such as attention and memory. In fact a consistent and reliable approach for the estimation of time-varying connectivity patterns with high accuracy and speed in following temporal evolutions of phenomenon was provided. Such method is also able to estimate time-varying patterns including in the process all the cerebral sources without applying any a-priori selection which could affect the results (Toppi et al., 2012).

Advanced methodologies for time-varying connectivity estimation and for the extraction of salient indexes describing the most important features of the investigated networks were used for the study of cerebral mechanisms at the basis of attention and memory. The identification of the indexes describing such cognitive functions were extracted during a pilot testing study conducted in normal healthy ageing (N=17 subjects). The study provided high density EEG data measured during the execution of two well-known tasks targeting different cognitive functions of interest (selective attention and declarative memory). On the basis of scientific evidence derived from literature, I defined and computed descriptors of the cognitive functions based on connectivity networks elicited by the specific tasks and stable across the population analyzed. A statistical analysis proved that such indexes are significantly related to the workload imposed to the subjects or to other indicators of the cognitive performances and

that they are thus suitable to provide measurable physiological indexes of the cognitive performances.

Material and Methods

Experimental Design

17 healthy elderly subjects (between 40 and 60 years old; 8 males) were enrolled in the study. EEG signals were recorded by a BrainAmp Standard amplifier (Brain Products GmbH, Munich, Germany) from 60 electrodes positions according to the extended 10–20 electrode placement system against a linked mastoid reference. Vertical and horizontal EOG signals were recorded with three electrodes in total, two placed on the outer canthi of the eyes and one on the nasion. Electrode impedances were kept below 5 kOhms for the EEG recording and below 10 kOhms for the EOG recording. EEG signals were digitized at 500 Hz and filtered with a 0.01 Hz high-pass and a 100 Hz low-pass.

After 2 minutes of eyes-open and eyes-closed resting EEG recording, each subject performed two cognitive tasks: visual oddball and Sternberg tasks. The two paradigms are built to elicit specific cognitive components of attention and memory functions and therefore different functional networks.

Details about the stimulation and timing for each task are provided in the following paragraphs.

Visual Oddball Task

A classic visual oddball task was administered to elicit attentive processes in the subjects involved in the study (Squires et al., 1975). In the oddball paradigm, trains of visual stimuli (rare stimuli in a stream of frequent stimuli), are used to assess the neural reactions to unpredictable but recognizable events (target, i.e. the rare stimulus). It was shown that the consequent evoked potential across the parieto-central area, called P300, that is usually around 300 ms, is larger after the target stimulus than after the frequent non-target stimuli. An oddball paradigm is often used to study effects of stimulus novelty and significance on information processing (Ferrari et al, 2010). P300 has been shown to be an attention-dependent cognitive component, but it also reflects broad recognition and memory-updating processes.

In my study, a sequence of frequent letter “O” and rare (target) letter “X” was presented according to a randomized order in the middle of a black screen. The probability associated to

frequent and infrequent stimuli was 0.8 and 0.2 respectively. Each stimulus remained in the middle of the screen for 120 ms. The inter-stimulus interval varied randomly between 1 and 1.5 sec. The participant had to press a button with his/her right hand as fast as possible whenever a target appears. In sum, 60 targets and 240 standards were presented, for a duration of the whole task approximately around 7.5 minutes. A description of the timing used for the oddball task is reported in Fig.4.5.

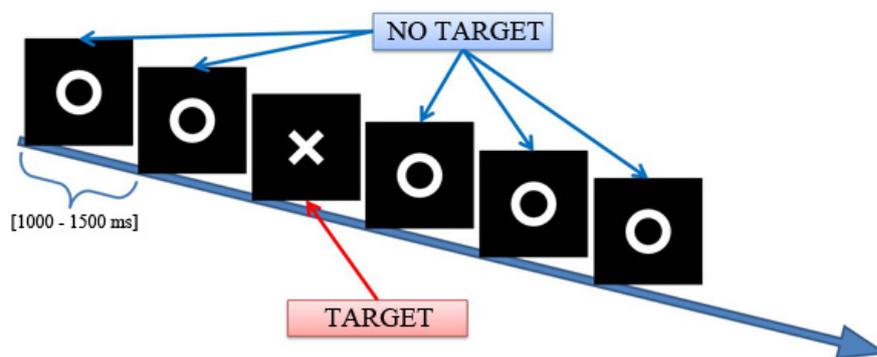


Figure 4.5 – Timing of the visual oddball experiment

Sternberg Task

The Sternberg task is a paradigm used for eliciting all the steps typical of a memory process (Sternberg, 1966). The subject is firstly given a short period for memorizing a series of numeric digits (encoding phase), secondly, he/she has to retain the memorized information for a certain period (storage phase) and then he/she has to retrieve it in a short time interval (retrieval phase). In my study, subjects were asked to remember a set of unique digits (between 0 and 9), and then a probe stimulus in the form of a digit was presented. The subjects were instructed to answer, as quickly as possible, whether the probe was in the previous set of digits or not. The size of the initial set of digits determined the workload on the subject (4 digits → easy, low workload; 6 digits → difficult, high workload).

Each trial starts with the presentation of a fixation cross in the middle of the screen, for 2 seconds. Afterwards, a “memory set” of 4 (e.g. 5682) or 6 digits (e.g. 146372), is presented for 1 second to allow memorization (encoding phase). The presentation of the digits series is followed by another fixation cross window presented for 2 seconds (storage period). Then a single probe digit is presented for 250 ms (retrieval phase) followed by a fixation cross presented for 1250 ms. Afterwards, the question “yes or no?” appears at the screen for a

maximum duration of 1500 ms, and the participant is required to give an answer. If the probe was a member of the preceding set, the participant has to press the left button (Target Condition). If the probe was not a member of the preceding memory set, the participant has to press the right button (No Target Condition). The duration of the presentation of the question “yes or no?” is response-ended, hence, as soon as the participant gives an answer, the next trial starts. The probability of Target condition is 0.5. The digits contained in each trial are presented in a randomized order.

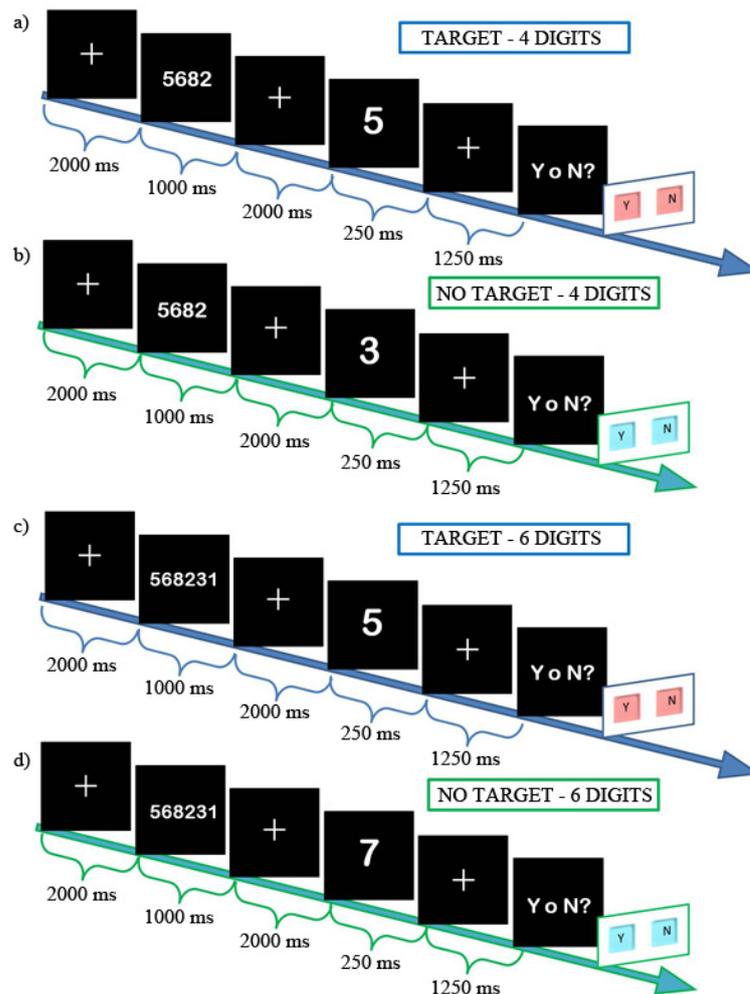


Figure 4.6 – Timing of the Sternberg experiment for the conditions: a) Target_4digits; b) NoTarget_4digits; c) Target_6digits; d) NoTarget_6digits.

The conditions 4/6 digits and Target/NoTarget are randomized within the recording session. 36 trials for each condition were administered. A detailed description of the timing of the experiment for the four different conditions is reported in Fig.4.6.

Pre-Processing of EEG traces

EEG signals were down sampled to 100 Hz and band-pass filtered in the range [1-45] Hz to optimize the following connectivity analysis. Independent Component Analysis (ICA) was used for removing ocular artifacts. EEG traces were segmented according to the specific timing of each paradigm and classified according to different conditions (Oddball: Target and No Target; Sternberg: Target_4digits, No Target_4digits, Target_6digits, No Target_6digits). Residual artifacts were then removed by means of a semi-automatic procedure based on a threshold criterion. Only the artifacts-free trials were considered in the subsequent effective connectivity analysis.

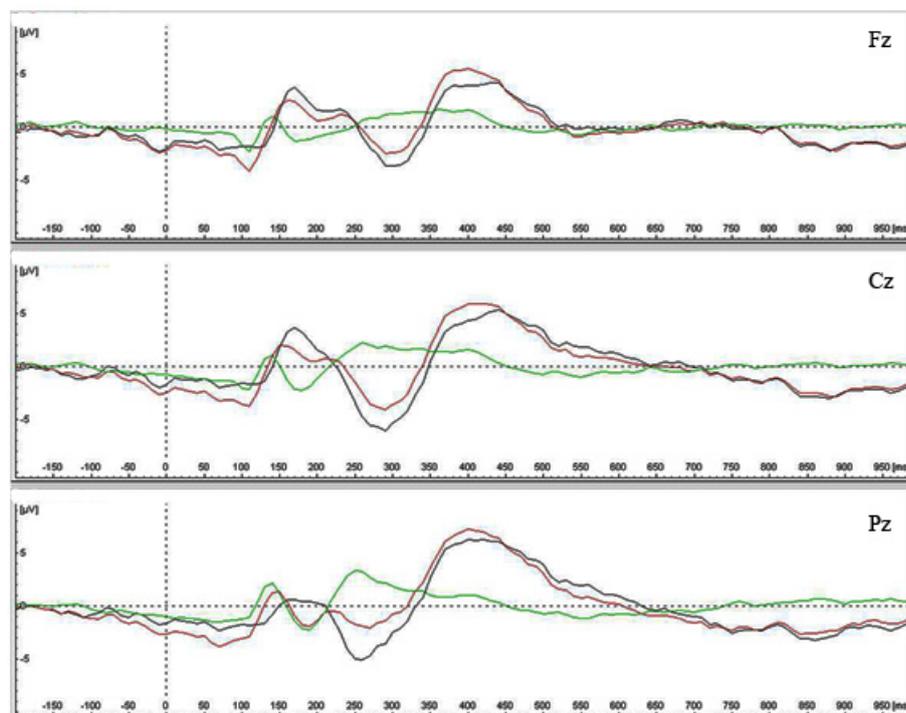


Figure 4.7 – Grand Average across subjects of the EEG signals recorded during the oddball paradigm related to the time interval [-200 : 1000] ms according to the stimulus onset in correspondence of electrodes Fz (first row), Cz (second row), Pz (third row). I reported in red the trace corresponding to Target condition, in green the trace corresponding to No Target condition and in black the difference between them.

Results

EEG Visual Oddball Task

Before applying all the functional connectivity methodologies on data acquired during oddball paradigm, I performed a classical ERP analysis on the data recorded during the oddball task, in order to verify that the paradigm had elicited the expected cognitive components. In particular, a grand average of the EEG traces across the population was computed separately

for both Target and NoTarget conditions and a comparison between them was performed. The results of this procedure were reported in Fig.4.7 for the electrodes F_z , C_z and P_z . In particular, I reported the grand average traces in the time interval [-200: 1000] ms for the Target condition (in green), the No Target condition (in red) and the difference between them (in black). The Grand Average of the difference between Target and No Target conditions highlighted the presence of the P300 event-related brain potential, a positive wave located around 400 ms after the stimulus onset, as expected from an oddball experiment. Such wave reflects the engagement of specific components of attention (alert, selective attention) and of a broad recognition and memory-updating processes. This is in agreement with the task, in which the subjects had to be ready to press the button once the target appeared on the screen and they had to recognize the Target and distinguish it from the No Target (Polich, 2007).

These results are confirmed also at single subject level in 16 out of the 17 subjects included in the study. In Tab.4.2 I reported the amplitude and latency values for the P300 wave elicited in correspondence of F_z (red columns), C_z (green columns) and P_z (blue columns) electrodes. No significant differences resulted between the three scalp locations in terms of amplitude or latency of the P300 wave, as confirmed by two one-way ANOVAs computed considering as main factor the LOCATION of the electrodes (F_z , C_z , P_z) and as dependent variables the P300 amplitude and latency values.

Once checked the involvement of each subject in the task execution and the subsequent elicitation of the attentive components typical of an oddball experiment, I applied a time-varying approach for functional connectivity estimation to extract some descriptors of the neuronal behaviour at the basis of P300 event-related potential. The method GLKF, described in Sec. II was applied to all the data corresponding to Target and No Target conditions in the window [0 1000] ms according to the stimulus onset.

The connectivity patterns estimated for each time sample were averaged in five time intervals ([0:200] ms; [200:400] ms; [400:600] ms; [600:800] ms; [800:1000] ms) and in four frequency bands defined according to the Individual Alpha Frequency (IAF), determined from the Fast Fourier Transform spectra over posterior leads (parietal, parieto-occipital, and occipital). Individually defined bands considered were: Theta (IAF-6/IAF-2), Alpha (IAF-2/IAF+2), Beta (IAF+2/IAF+14) and Gamma (IAF+15/IAF+30). A statistical comparison between Target and No Target conditions was computed in order to discard all the effects due to the environment or to the stimulation used for administering the paradigm. Then a Grand Average across the population was computed and the results achieved in Theta and Alpha

bands were reported in Fig.4.8 and Fig.4.9, respectively. I limited the presentation of the results to only these two bands because they were demonstrated to be highly involved in this type of tasks.

Subjects	Amplitude			Latency		
	F_z	C_z	P_z	F_z	C_z	P_z
01	7.15	5.04	5	370	370	440
02	4.22	4.68	4.47	440	480	480
03	6.81	8.2	7.94	360	380	380
05	8.89	11.77	11.64	430	450	400
06	4.36	4.84	4.37	480	470	470
07	3.99	3.66	5.12	510	500	360
08	7.41	10.31	10.82	380	440	440
09	3.47	8.86	10.17	420	430	430
10	16.22	14.55	13.73	440	400	430
11	5.69	7	5.86	500	500	490
12	2.18	4.025	5.26	430	430	430
13	4.29	11.43	11.8	410	450	430
14	4.1	5.92	7.13	490	490	370
15	13.46	12.07	10.42	370	370	370
16	6.14	3.9	5.79	440	400	400
17	9.76	10.58	10.67	410	410	380
MEAN	6.758	7.927	8.137	430	435.625	418.75
STD	3.785	3.509	3.125	47.046	44.567	40.640

Table 4.2 – Table listing the amplitude and latency values for P300 elicited in correspondence of F_z (red), C_z (green) and P_z (blue) electrodes, for each of the 17 subjects involved in the study

In panel (a) of both figures I reported the Grand Average of functional connectivity patterns elicited during Oddball paradigm execution.

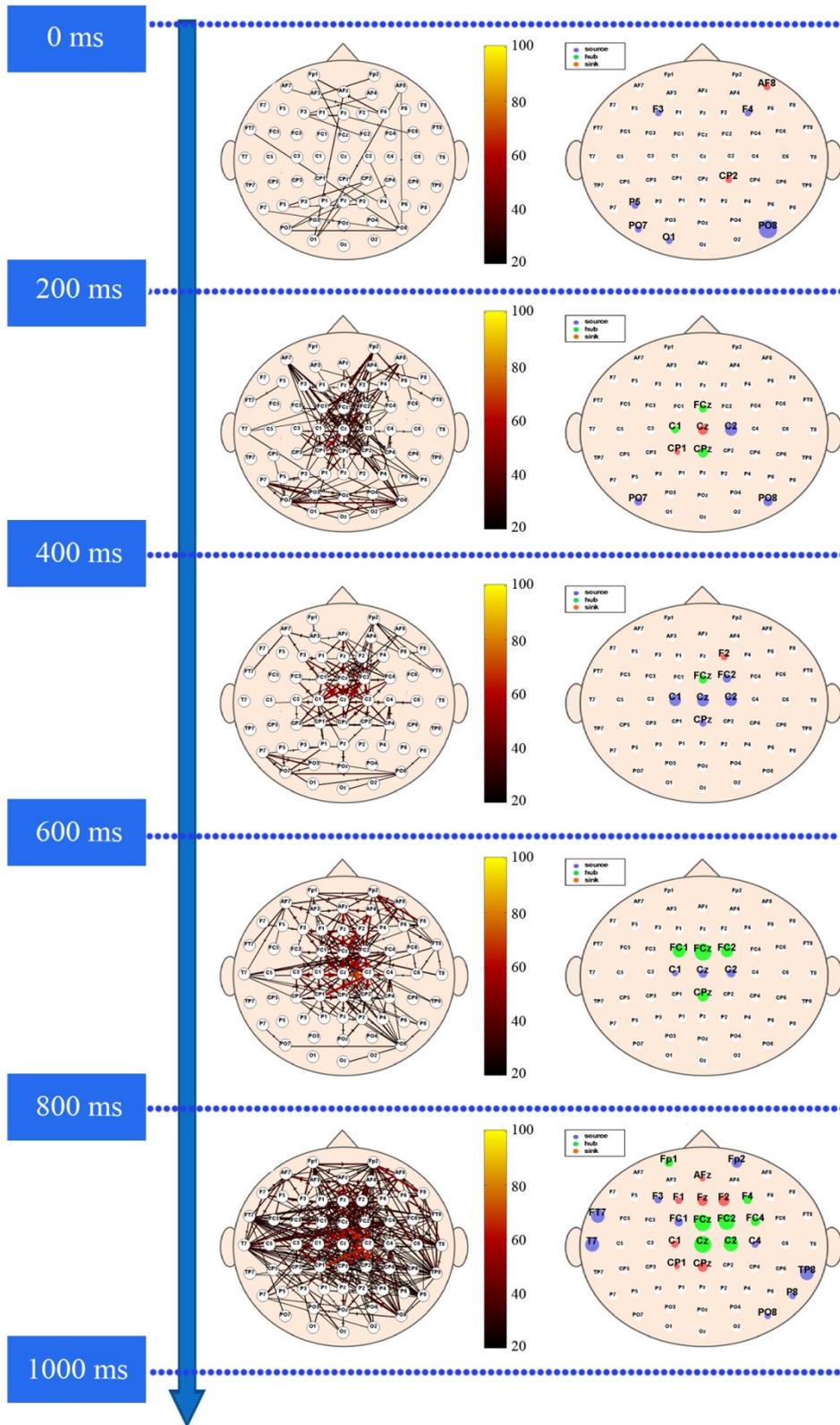


Figure 4.8 – (a) Grand Average of functional connectivity patterns elicited during Oddball paradigm in Theta band. Each network is related to a specific time interval defined according to the stimulus onset: (0:200) ms (first column), (200:400) ms (second column), (400:600) ms (third column), (600:800) ms (fourth column), (800:1000) ms (fifth column). Connectivity patterns are represented on a 2D scalp model seen from above with the nose pointing to the top of the page. Colour and size of the arrows code the percentage of population eliciting the considered connection. (b) Graphical representation of the role of each electrode involved in the connectivity patterns reported in (a). The colour and the diameter of each sphere code for the role (hub in green; sink in red; source in blue) and the magnitude of the degree of the correspondent electrode, respectively.

Each network is related to a specific time interval defined according to the stimulus onset (0 time): (0 ; 200) ms (first column), (200 ; 400) ms (second column), (400 ; 600) ms (third column), (600 ; 800) ms (fourth column), (800 ; 1000) ms (fifth column). Connectivity patterns are represented on a 2D scalp model, seen from above, with the nose pointing to the top of the page. Each electrode is modelled by means of a white circle and the causal connections between different electrodes are represented by means of arrows. Colour and size of each arrow code the percentage of population eliciting the considered connection. All the represented connections were elicited by at least 20% of the considered population.

In panel (b) of Fig.4.8 and Fig.4.9 I reported a graphical representation of the role of each electrode involved in the connectivity patterns reported in (a). Each electrode is represented by means of a sphere. The colour of each sphere codes for the role of the electrode within the network: green if the electrode is a *hub* in the network, meaning that it manages the information flow in both directions (in and out); red if the electrode is a *sink* in the network, mainly receiving information from the rest of the network and not spreading it further; blue if the electrode is a *source* of information, that is a driving element in the network. If the sphere is white, the electrode is not involved in the network. The diameter of each sphere codes for the magnitude of the degree index computed for the correspondent electrode.

Figures 4.8(a) and 4.9(a) showed quite stable connectivity patterns among the population. In fact, all the connections were elicited by at least 50% of the considered population (as represented by their dark to light red colour). The estimated connectivity pattern evolved along the task (zero time: the stimulus onset) in both Theta and Alpha bands. The areas mainly involved in the network during the interval [200–600] ms, which resulted to be the interval of interest for the P300 from the grand-average analysis, belong to the central and the centro-parietal part of the scalp. Figures 4.8(b) and 4.9(b) revealed a high involvement of the central electrodes (C1, Cz, C3, FCz, CPz) in the same interval. In particular, in Theta a role of sink is shown for Cz in the interval [200:400] ms, while it changes to a source in the interval [400:600] ms (Fig 4.8(b)). In Alpha, the same electrode keeps the role of source for both intervals as shown in Fig 4.9(b). The role of Cz in attentive paradigms is in line with the literature. (Polich, 2007; Duncan et al., 2009).

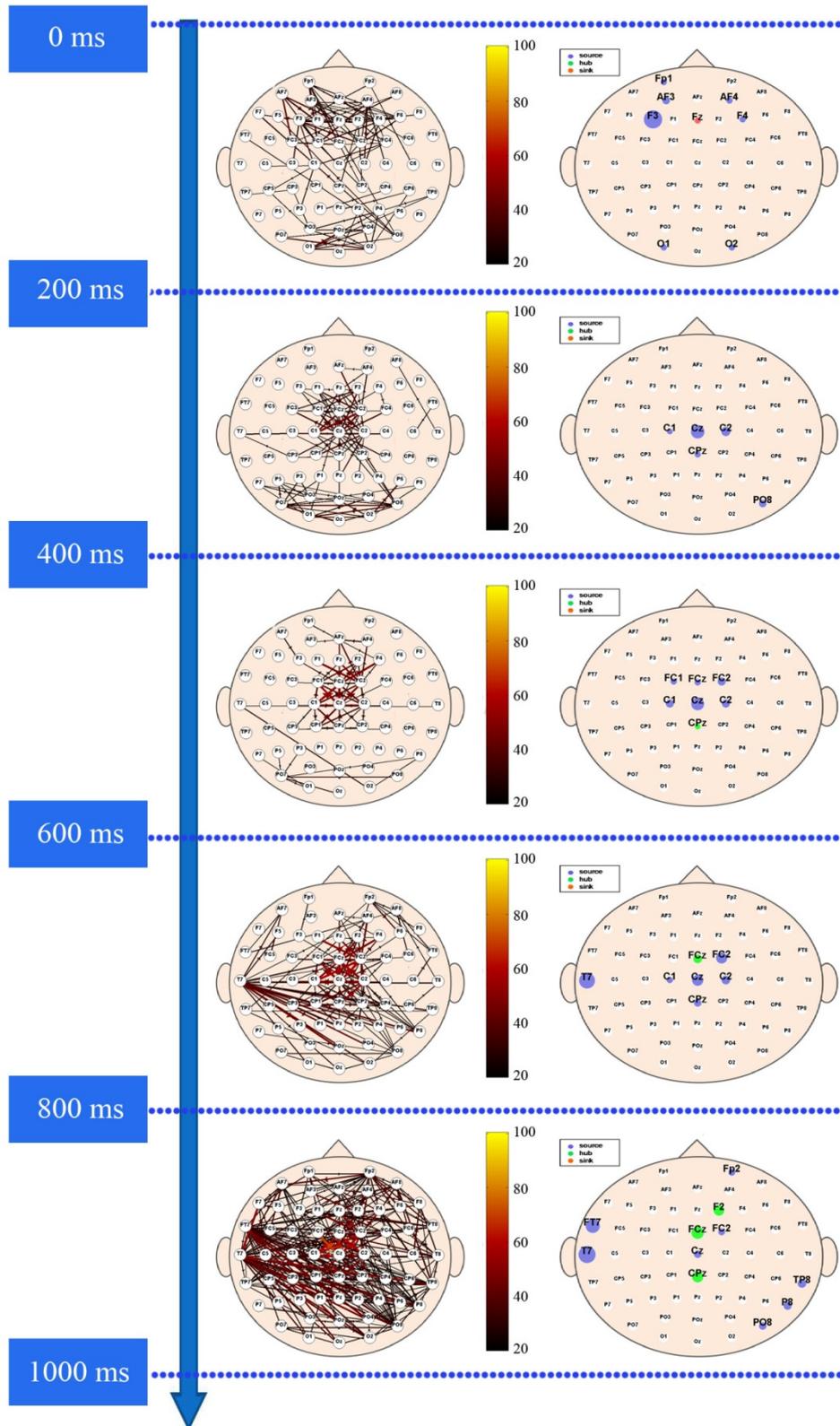


Figure 4.9 – (a) Grand Average of functional connectivity patterns elicited during Oddball paradigm in Alpha band. Each network is related to a specific time interval defined according to the stimulus onset: (0:200) ms (first column), (200:400) ms (second column), (400:600) ms (third column), (600:800) ms (fourth column), (800:1000) ms (fifth column). Connectivity patterns are represented on a 2D scalp model seen from above with the nose pointing to the top of the page. Colour and size of the arrows code the percentage of population eliciting the considered connection. (b) Graphical representation of the role of each electrode involved in the connectivity patterns reported in (a). The colour and the diameter of each sphere code for the role (hub in green; sink in blue) and the magnitude of the degree of the correspondent electrode, respectively.

Given the relevant role of Cz in the network in the theta and alpha bands, I wanted to verify if the Cz degree index can be a possible quantifiable descriptor of the attentive processes at the basis of the execution of the task. To this purpose, I correlated the degree of Cz electrode with amplitude and latency values of the P300 measured at Cz location. This choice is motivated by the fact that amplitude and latency values of P300 can be considered as describing the level of attention showed by the subjects during the execution of an oddball paradigm. In fact, several studies demonstrated that P300 amplitude is proportional to the amount of attentive resources engaged during the cognitive process, while P300 latency is proportional to time required for their allocation (Polich, 2007).

P300 Amplitude in Cz and Cz-degree					
	0-200 ms	200-400 ms	400-600 ms	600-800 ms	800-1000 ms
Theta	0,227	0,694	0,546	0,272	0,386
Alpha	0,242	0,640	0,470	0,181	0,038
Beta	-0,026	0,293	0,334	0,129	-0,026
Gamma	-0,171	0,472	0,204	-0,189	-0,379
P300 Latency in Cz and Cz-degree					
	0-200 ms	200-400 ms	400-600 ms	600-800 ms	800-1000 ms
Theta	0,213	-0,580	-0,628	-0,374	-0,329
Alpha	0,034	-0,659	-0,708	-0,524	-0,513
Beta	0,354	-0,365	-0,082	-0,056	-0,104
Gamma	-0,018	-0,133	-0,034	0,257	0,282

Table 4.3 – Correlation values between the degree of Cz electrode and P300 amplitude (first part of the table, in blue) and latency (second part of the table, in green) elicited in correspondence of Cz electrode, for each band and for each time interval. The values highlighted in red correspond to statistically significant correlation values.

In Tab.4.3 I reported the results of these correlations computed for each frequency band and for each time interval considered in the analysis. I highlighted in red the significant values of correlations achieved ($p < 0.05$). The degree of Cz electrode correlated positively with the amplitude value of P300 in the intervals [200:400] ms and [400:600] ms defined according to

the stimulus onset in Theta band and with the interval [0:200] ms in Alpha band. A significant negative correlation between the degree of Cz and the latency value of P300 measured at Cz location resulted in Theta and Alpha bands for both intervals [200:400] ms and [400:600] ms. The correlation resulted only in the intervals around the P300 onset and in the frequency bands of interest for the task. The different sign of the correlation is in agreement with the neurological processes at the basis of P300, according to which the higher the engagement of attentive resources the higher is the P300 amplitude, and the lower the P300 latency. This result confirmed the hypothesis that the degree of Cz can be a descriptor of the attentive engagement of a subject during an oddball paradigm.

EEG Sternberg Task

The data acquired during the Sternberg task were subjected to a time-varying connectivity estimation, to capture the relevant properties of the brain networks at the basis of memory processes.

The EEG traces recorded during Target_4digits, Target_6digits, NoTarget_4digits and NoTarget_6digits conditions were segmented in the window [0:6000] ms (zero time corresponds to the presentation of the digits string to be remembered) and subjected to time-varying functional connectivity estimation by means of GLKF. The connectivity patterns estimated for each time sample were averaged in three time intervals: [0:1000] ms (encoding phase); [1000:3000] ms (storage phase) and [3000:6000] ms (retrieval phase) and in four frequency bands, defined according to the Individual Alpha Frequency (IAF), determined from the Fast Fourier Transform spectra over posterior leads (parietal, parieto-occipital, and occipital). Individually defined bands considered were: Theta (IAF-6/IAF-2), Alpha (IAF-2/IAF+2), Beta (IAF+2/IAF+14) and Gamma (IAF+15/IAF+30). To discard all the effects due to the environment or to the stimulation used for administering the paradigm, a statistical comparison between each condition and the corresponding baseline was computed for a significance level of 5% FDR corrected for multiple comparisons. The baseline period was the time interval [-1000:0] ms defined according to the onset of the screen containing the digits series. During this interval, the subject had to look at a fixation cross.

After the statistical validation of connectivity patterns, a Grand Average across the population was computed and the results achieved in Alpha band for Encoding and Storage periods are reported in Fig.4.10 and Fig.4.14, respectively. In both figures I reported the Grand Average related to Target (first row) and No Target (second row) conditions for the cases 4 digits (first column) and 6 digits (second column). Connectivity patterns are represented on a 2D scalp

model seen from above, with the nose pointing to the top of the page. Each electrode is represented by means of a white circle and the causal connections between different electrodes are represented by arrows linking the electrodes. Colour and size of each arrow code the percentage of population eliciting the considered connection. All the represented connections were elicited by at least 30% of the considered population in the Encoding phase and in at least 40% of the population in the Storage phase. In Fig.4.11 and Fig.4.15 I reported a graphical representation of the role of each electrode involved in the connectivity patterns reported in Fig.4.10 and Fig.4.14, respectively. Each electrode is represented by means of a sphere. The colour of each sphere codes for the role of the electrode within the network: green if the electrode is a hub in the network, meaning that it manages the information flow in both direction (in and out); red if the electrode is a sink in the network, blue if the electrode is a source for the network (same representation than in the previous figures). If the sphere is white, the electrode has no significant role in the network. The diameter of each sphere codes for the magnitude of the degree index computed for the corresponding electrode.

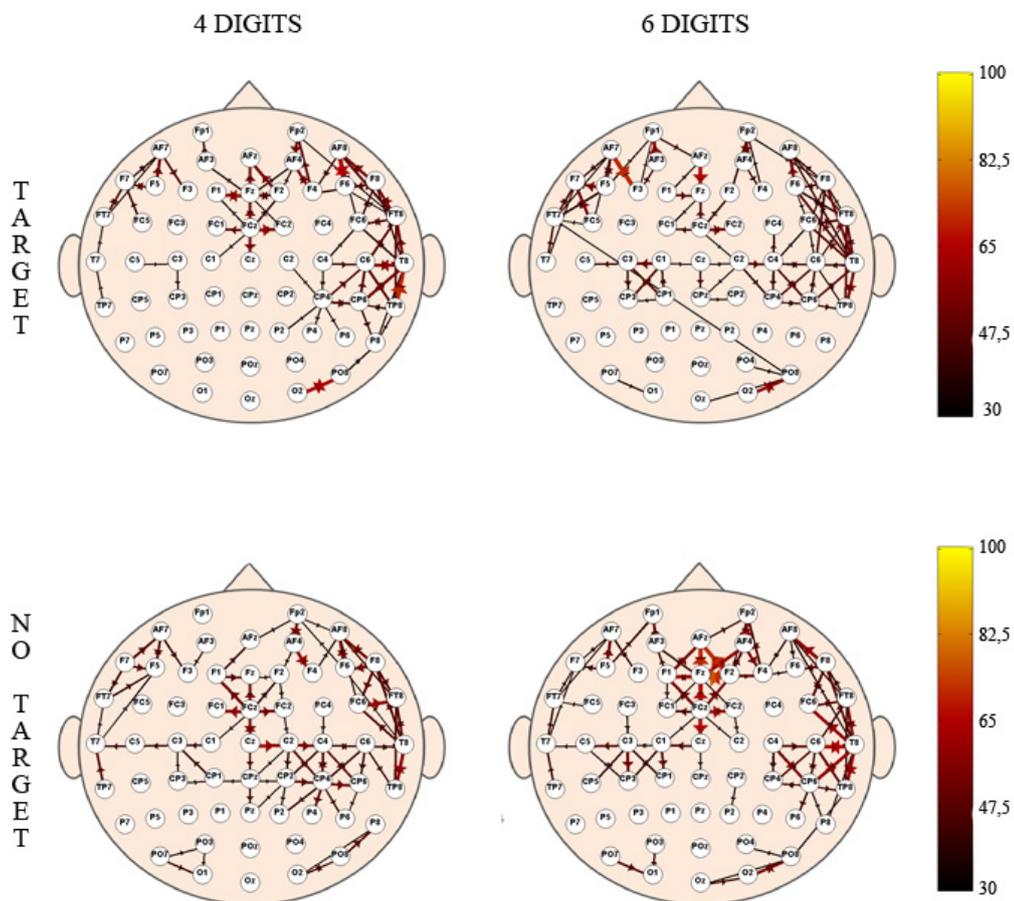


Figure 4.10 - Grand Average of functional connectivity patterns elicited during the Encoding phase of Sternberg paradigm in Alpha band. The patterns are referred to the conditions Target (first row), No Target (second row) for the cases 4 digits (first column) and 6 digits (second column). Connectivity patterns are represented on a 2D scalp model, seen

from above with the nose pointing to the top of the page. Colour and size of the arrows code the percentage of population eliciting the considered connection.

From a visual inspection of the grand average connectivity pattern associated to the Encoding phase in Alpha band, reported in Fig.4.10, no macroscopic differences resulted between Target and No Target condition or between 4 and 6 digits cases. In all the cases the network is characterized by an involvement of the frontal and temporal areas with a small prevalence of the right hemisphere. These results were confirmed by the analysis of the role of each electrode, reported in Fig.4.11. Fronto-central electrodes are involved as source of information (in blue), while the right fronto-temporal electrodes mainly act as hub for all the conditions (in green).

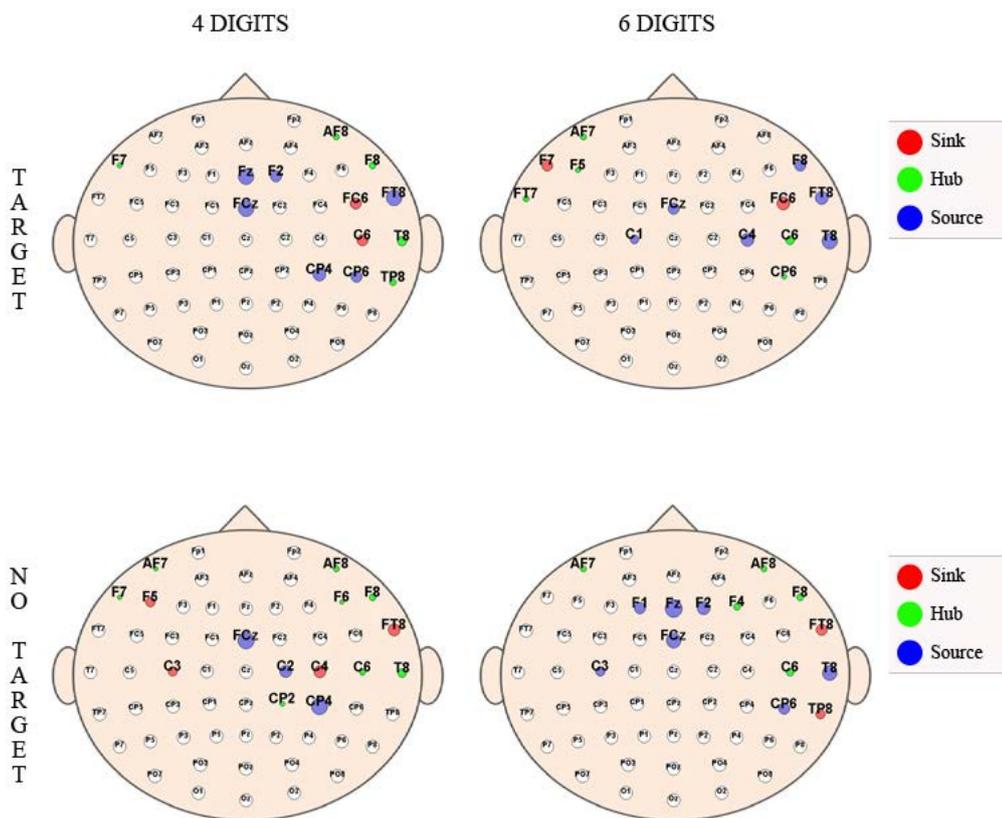


Figure 4.11 - Graphical representation of the role of each electrode involved in the connectivity patterns elicited during the Encoding phase of Sternberg paradigm, in Alpha band. The colour and the diameter of each sphere code for the role (hub in green; sink in red; source in blue) and the magnitude of the degree of the correspondent electrode, respectively.

These results highlight that, in this case, local indexes cannot be considered as descriptors of the process at the basis of encoding phase. In fact, there's no electrode whose role results to be correlated to the level of difficulty imposed during the memorization phase, which is

necessary to assess that the index is not generally linked to the task execution, but is strictly (and in a quantifiable way) related to the cognitive function targeted in the experiment.

We then computed global indexes, able to describe more general properties of the network, to be correlated with the modulation of the workload of the memory process.

In particular, I performed three-way Analysis of Variance (ANOVAs), with the local efficiency, the global efficiency and the small-worldness computed on the connectivity networks in the different conditions as dependent variables, and with TARGET (Yes or No), DIGITS (4 or 6), MEMORY-PHASES (Encoding, Storage, Retrieval) as within-main factors.

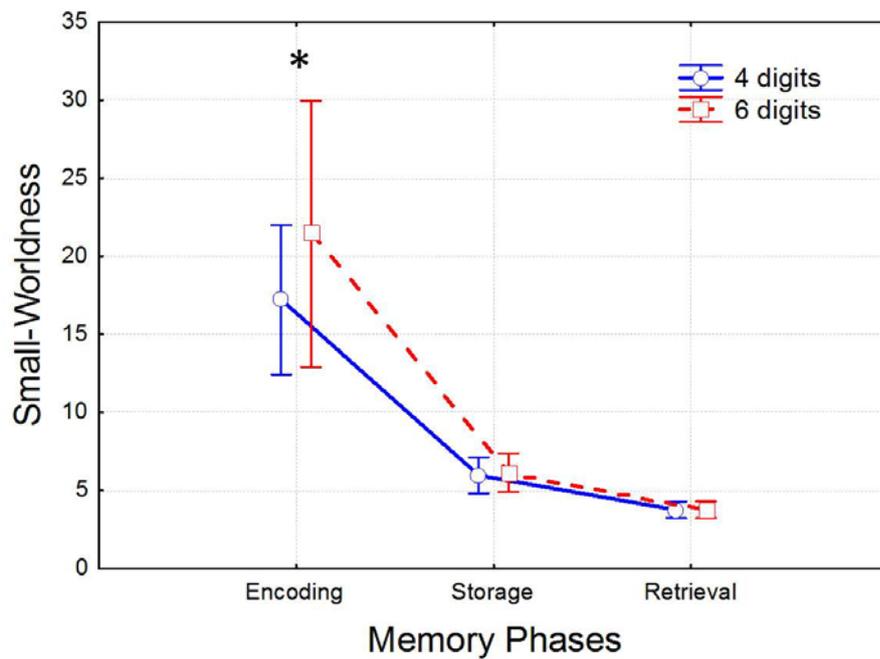


Figure 4.12 - Results of three-way ANOVA performed on the Small-Worldness computed in Alpha Band: plot of means with respect to the interaction between DIGITS (4 or 6) and MEMORY-PHASE (Encoding, Storage or Retrieval) main factors. The symbol () indicates statistical differences between 4 and 6 digits conditions as confirmed by Duncan's post hoc test ($p < 0.05$). The bar on each point represents the 95% confidence interval of the mean errors computed across the subjects*

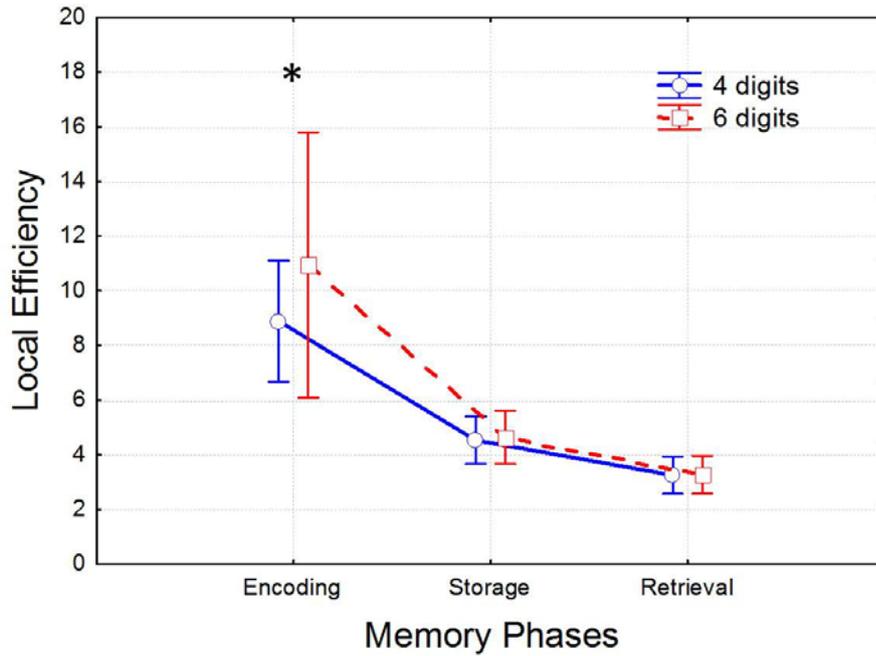


Figure 4.13 - Results of three-way ANOVA performed on the Local Efficiency computed in Alpha Band: plot of means with respect to the interaction between DIGITS (4 or 6) and MEMORY-PHASE (Encoding, Storage or Retrieval) main factors. The symbol () indicates statistical differences between 4 and 6 digits conditions as confirmed by Duncan's post hoc test ($p < 0.05$). The bar on each point represents the 95% confidence interval of the mean errors computed across the subjects*

Statistically significant differences between 4 and 6 digits conditions resulted for the Small-Worldness and for the Local Efficiency computed in Alpha band in the encoding phase, as reported in Fig.4.12 and Fig.4.13. The figures reported the plot of means with respect to the interaction between DIGITS (4 or 6) and MEMORY-PHASE (Encoding, Storage or Retrieval) main factors. The symbol (*) indicates statistical differences between 4 and 6 digits conditions, as confirmed by Duncan's pairwise comparisons ($p < 0.05$). The two figures showed a significant increase of Small-Worldness and Local Efficiency associated to an increase of the workload (from 4 to 6 digits) in the encoding phase, meaning that the network needs to be better organized when the encoding is more difficult, due to the higher number of digits to be remembered. This result indicates that the two indexes of the network organization can be considered as descriptors of the processes specifically at the basis of encoding phase.

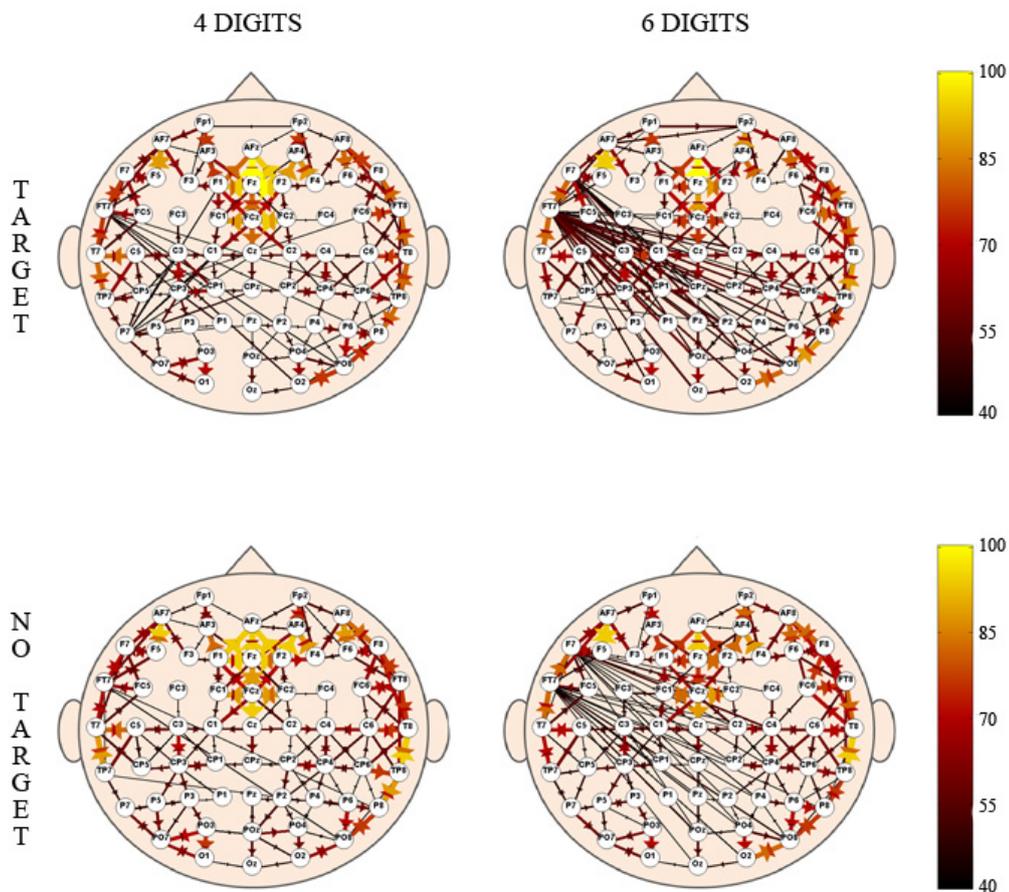


Figure 4.14 - Grand Average of functional connectivity patterns elicited during the Storage phase of Sternberg paradigm in Alpha band. The patterns are referred to the conditions Target (first row), No Target (second row) for the cases 4 digits (first column) and 6 digits (second column). Connectivity patterns are represented on a 2D scalp model seen from above with the nose pointing to the top of the page. Colour and size of the arrows code the percentage of population eliciting the considered connection.

In Fig. 4.14 I reported the grand average of connectivity patterns elicited during the Storage phase of Sternberg paradigm, in Alpha Band. In all the four conditions the pattern involved mainly the fronto-central area of the scalp in correspondence of Fz and FCz electrodes. No macroscopic differences resulted between Target and No Target condition for both 4 and 6 digits cases. In both Target and No Target conditions 4 and 6 digits cases differed for the involvement of left fronto-temporal area when the difficulty associated to the storage process increased. These results were confirmed by the correspondent analysis of the electrodes role, reported in Fig.4.15. The role of FT7 as source of information increased from 4 to 6 digits case in both Target and No Target conditions, while the role of other electrodes remained mostly unchanged. The flows spreading from FT7 was mainly directed to the right occipito-parietal areas of the scalp. The involvement of the left parietal lobe in difficult memory tasks is in agreement with literature. In fact it has already been demonstrated that increases in

alpha-band coherence between left temporal and right parietal sites is associated with an increase of the memory load (Payne and Kounios, 2009).

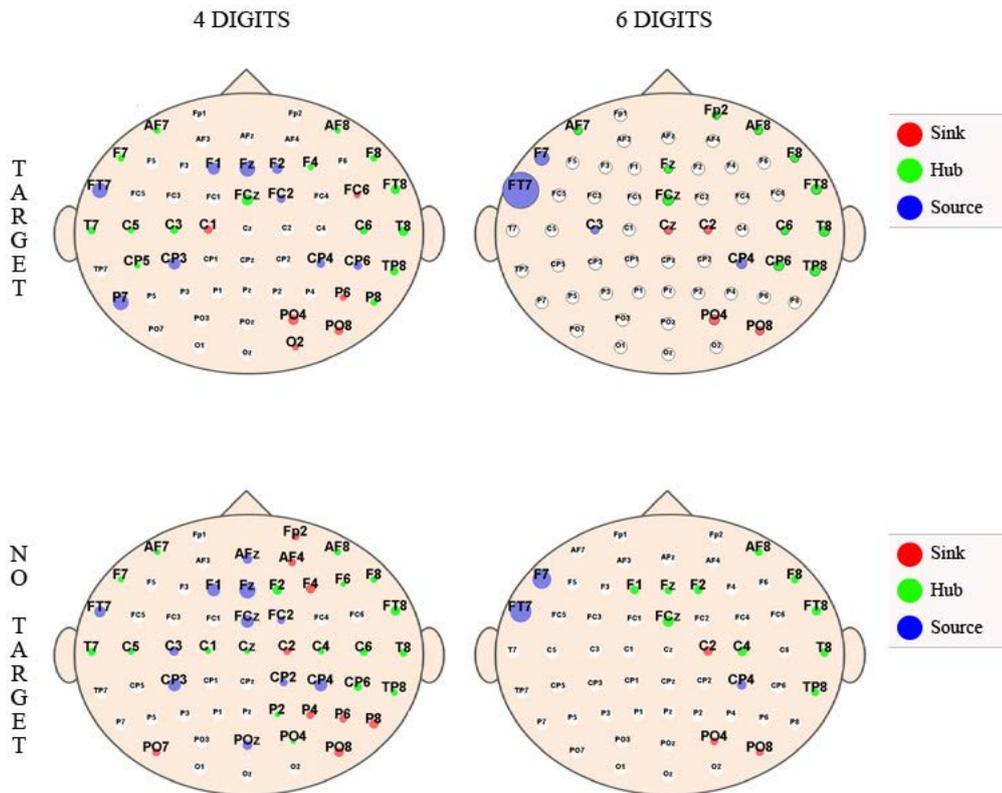


Figure 4.15 - Graphical representation of the role of each electrode involved in the connectivity patterns elicited during the Storage phase of Sternberg paradigm in Alpha band. The colour and the diameter of each sphere code for the role (hub in green; sink in red; source in blue) and the magnitude of the degree of the correspondent electrode respectively.

To evaluate the role of FT7 in the network as a possible descriptor of the processes at the basis of the storage process, I computed a three-way ANOVA with the degree of FT7 in alpha band as dependent variable and TARGET (Yes or No), DIGITS (4 or 6) and MEMORY-PHASES (Encoding, Storage, Retrieval) as within-main factors. The results of the ANOVA revealed statistically significant differences between 4 and 6 digits for the degree of FT7 computed in Alpha band, in the Storage phase, as reported in Fig.4.16. The figure reports the plot of means with respect to the interaction between DIGITS (4 or 6) and MEMORY-PHASE (Encoding, Storage or Retrieval) main factors. The symbol (*) indicates statistically significant differences between the 4 and 6 digits conditions, as confirmed by Duncan’s post hoc test ($p < 0.05$). The figure reveals a significant increase of FT7 degree associated to an increase of the workload (from 4 to 6 digits) in the storage phase, confirming from a statistical point of view what highlighted by the grand average. This index is specific of the storage

phase and can be modulated in relation to the difficulty level of the storage process requested to the subject.

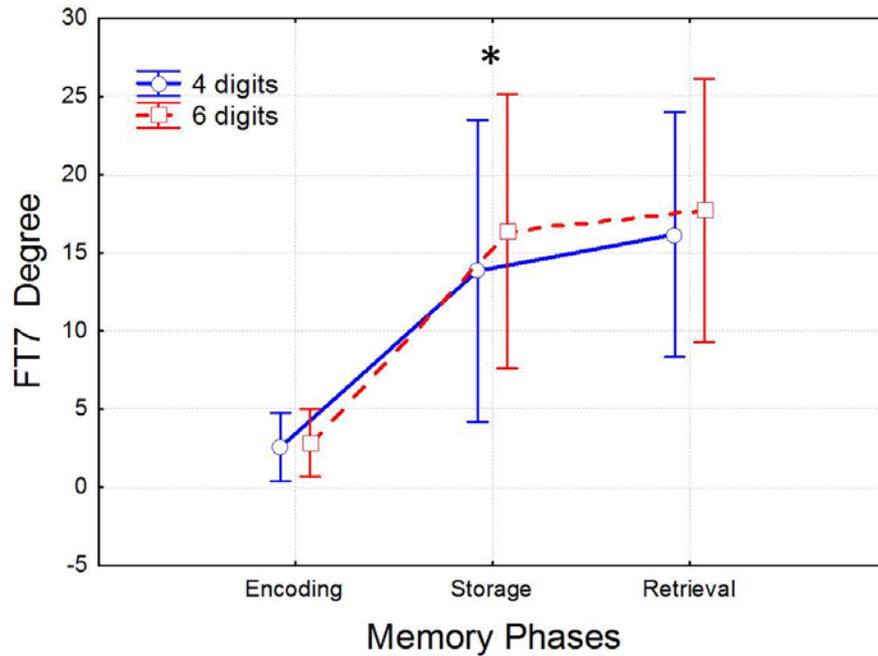


Figure 4.16 - Results of three-way ANOVA performed on the degree of FT7 computed in Alpha Band: plot of means with respect to the interaction between DIGITS (4 or 6) and MEMORY-PHASE (Encoding, Storage or Retrieval) main factors. The symbol () indicates statistical differences between 4 and 6 digits conditions as confirmed by Duncan's post hoc test ($p < 0.05$). The bar on each point represents the 95% confidence interval of the mean errors computed across the subjects.*

Discussion

The main objective of this study was to identify a set of quantifiable indexes derived from EEG-based connectivity and suitable to describe specific cognitive functions. To this purpose, I set up a study involving a group of 17 normal ageing healthy subjects (age between 40 and 60) during the execution of cognitive tasks targeting the cognitive functions under investigation. The targeted specific cognitive functions were: alert, selective attention and broad recognition (oddball) and declarative memory (Sternberg). A body of techniques at the state of the art in the estimation of effective connectivity in the time-frequency domain, together with well-established indexes for the interpretation of the connectivity networks, as well as some new indexes defined ad hoc for the study were applied for this purpose.

The obtained results demonstrated that the methodological advancement allowed to extract indexes able to describe the investigated cognitive functions in a selective manner. Moreover, statistical analysis showed that they are modulated by the workload imposed for the specific target cognitive function.

In the oddball paradigm the degree of Cz, in theta and alpha band, resulted as a valid descriptor of the attentive and memory-updating processes elicited. Its positive correlation with the amplitude of P300 event-related potential and its negative correlation with the P300 latency measured at Cz location suggested a direct correlation with the engagement of the attentive resources. Such correlations resulted only in the intervals characterized by P300 waves.

As for the Sternberg paradigm, I found selective descriptors for the encoding and the storage phases. In particular, the small-worldness and the local efficiency of the connectivity network, in alpha band, showed a significant increase with the workload requested during the memorization phase. A possible interpretation is that the execution of more difficult tasks (memorization of 4 to 6 digits) requires a better organization within the networks, corresponding to a significant increase of the two indexes. Their modulation in relation to the difficulty level of the task shows a relation with the targeted cognitive function. The degree of connectivity of FT7 electrode, in alpha band, proved as well to be a possible index of the success of the storage phase. In fact, its value is statistically related to the difficulty required for the maintenance of the information memorized during the encoding phase.

Conclusion of Section IV

The methodological advancement provided during my PhD allowed to study complex mechanisms at the basis of resting condition and cognitive functions in human. In particular, the consistency of the results and their reliability among the population could be achieved only thanks to the use of advanced methods in part developed and mostly refined during my PhD thesis. The application of asymptotic statistic method as statistical validation approach to be used at the same time to validate the estimated connectivity patterns and to extract the correspondent adjacency matrix allowed to achieve stable and reliable indexes to be used in the investigation of resting condition and in the identification of normative parameters describing the healthy state. The advancement in time-varying connectivity estimation allowed to describe cerebral networks at the basis of attentive and memory processes and to extract quantifiable indexes characterizing the investigated networks also correlating with the workload required for the required cognitive task. These results are completely new and can be seen as a confirmation of the quality of the methodology developed.

Section IV References

- Broyd, S.J., Demanuele, C., Debener, S., Helps, S.K., James, C.J., Sonuga-Barke, E.J.S., 2009. Default-mode brain dysfunction in mental disorders: a systematic review. *Neurosci Biobehav Rev* 33, 279–296.
- Chen, A.C.N., Feng, W., Zhao, H., Yin, Y., Wang, P., 2008. EEG default mode network in the human brain: spectral regional field powers. *Neuroimage* 41, 561–574.
- Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V., 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 253–258.
- Klimesch, W., 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews* 29, 169–195.
- Mantini, D., Perrucci, M.G., Del Gratta, C., Romani, G.L., Corbetta, M., 2007. Electrophysiological signatures of resting state networks in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13170–13175.
- McKiernan, K.A., D’Angelo, B.R., Kaufman, J.N., Binder, J.R., 2006. Interrupting the “stream of consciousness”: an fMRI investigation. *Neuroimage* 29, 1185–1191.
- Möller, E., Schack, B., Arnold, M., Witte, H., 2001. Instantaneous multivariate EEG coherence analysis by means of adaptive high-dimensional autoregressive models. *J. Neurosci. Methods* 105, 143–158.
- Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* 118, 2128–2148.
- Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682.
- Singh, K.D., Fawcett, I.P., 2008. Transient and linearly graded deactivation of the human default-mode network by a visual detection task. *Neuroimage* 41, 100–112.
- Squires, N.K., Squires, K.C., Hillyard, S.A., 1975. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalogr Clin Neurophysiol* 38, 387–401.
- Sternberg, S., 1966. High-speed scanning in human memory. *Science* 153, 652–654.

Toppi, J., Babiloni, F., Vecchiato, G., Cincotti, F., De Vico Fallani, F., Mattia, D., Salinari, S., Astolfi, L., 2011. 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.

Toppi, J., Babiloni, F., Vecchiato, G., Cincotti, F., De Vico Fallani, F., Mattia, D., Salinari, S., Astolfi, L., 2012. Towards the Time Varying Estimation of Complex Brain Connectivity Networks by means of General Linear Kalman Filter. *Conf Proc IEEE Eng Med Biol Soc*.

General Conclusion

The work of these three years of PhD course returns a methodology for the estimation of effective connectivity, for its validation and for the extraction of indexes describing in a stable and reliable way the properties of the brain circuits underlying different cognitive states in humans, which overcomes most of the limitation affecting the approach commonly encountered in literature. In particular:

1. The issue of the statistical validation of the connectivity patterns, also including corrections for multiple comparisons, was addressed with a rigorous simulation study, demonstrating that the asymptotic statistic recently introduced in literature can be applied in a range of conditions usually encountered when dealing with EEG experiments, with better performances with respect to the previously used shuffling surrogate data procedure.
2. The procedure for the extraction of adjacency matrices from connectivity patterns here suggested was proved able to consistently and reliably describe the real properties of the brain networks, avoiding the common errors produced by the method previously used in literature. In fact, the application of qualitative approaches for extracting adjacency matrices, commonly used in literature, was shown to affect the topology of the investigated patterns leading to their erroneous description.
3. The General Linear Kalman Filter method was proved able to accurately describe the temporal evolution of connectivity patterns. Moreover, due to its numerical stability, it showed good performances when dealing with high dimensional models, thus allowing to include an unprecedented number of signals (representing the activity in different cortical sites) in the estimation process. This means no need to select specific brain regions to be included in the model and a reduction of the “hidden source” effect on the connectivity model produced.
4. The developed methodology was successfully tested on real data related to the resting state properties of the brain networks and to cognitive tasks. The results, in terms of brain networks and indexes describing their properties, showed a consistency across the population investigated which is unprecedented in literature. In particular, the indexes obtained were not only able to characterize the specific

function examined, but showed a significant correlation with factors like the workload imposed to the subjects, their age, the difficulty of the task. Such results open the way to clinical applications of the connectivity estimations, which may include the patients' stratification, an aid to the diagnosis of specific levels of cognitive deficits, and the description of phenomena of cortical plasticity resulting from a spontaneous or trained-induced recovery from different pathologies including movement or cognitive impairment.

Notwithstanding the achievements highlighted in this thesis, some considerations about their limits should be provided and addressed in the prosecution of these studies. In particular:

- 1) The use of linear MVAR model for both generating the simulated datasets and for estimating the effective connectivity patterns represents the first limitation in the studies presented in this thesis. In fact in the evaluation of the performances achieved by applying the different investigated methodologies (both stationary and time-varying) only the measurement errors due to the data variability were included. However, the errors related to the choice of a linear model for generating the data, which were not considered in the studies, could negatively affect and thus reduce the performances obtained.
- 2) One of the leitmotifs of the present work is the use of statistics as tool for guaranteeing the consistency, reliability and reproducibility of the achieved results. The tendency to transform a statistical significance in a conceptual/functional significance is a concept to be softened especially in the neuroscience applications. In fact, once highlighted significances in the investigated parameters it is necessary to reflect on their role and function in the investigated context.

Next steps of this research may include a better understanding of the properties of the brain networks that can be captured by the Granger-based approach to connectivity estimation, by testing the methodology here described to simulated datasets generated to include non-linear interactions between the activity in different brain sites. By deepening the model and introducing more physiological parameters into it, this approach could then move further steps toward the comprehension of the physiological phenomena at the basis of the organized brain behavior underlying the most complex cognitive tasks in humans.