

# A Statistical Proposal for Selecting a Data-dependent Threshold in Neurobiology

P. FINOTELLI<sup>1</sup>, P. DULIO<sup>1</sup>, G. VAROTTO<sup>2</sup>, F. ROTONDI<sup>2</sup>, F. PANZICA<sup>2</sup>

<sup>1</sup> Department of Mathematics "F. Brioschi", Politecnico di Milano, Milan, Italy;

<sup>2</sup> Neurophysiology and Diagnostic Epileptology Operative Unit, "C. Besta" Neurological Institute IRCCS Foundation, Milan, Italy

## ABSTRACT

*In this paper we propose a new methodology for introducing thresholds in the analysis of neurobiological databases. Often, in Neuroscience, absolute thresholds are adopted. This is done by cutting the data below (or above) predetermined values of the involved parameters, without an analysis of the distribution of the collected data concerning the phenomenon under investigation. Despite an absolute threshold could be rigorously defined in terms of physic parameters, it can be influenced by many different subjective aspects, including cognitive processes, and individual adaptation to the external stimuli. A possible related risk is that, mainly in experiments also depending on personal reactions, a significant portion of meaningful data, relevant for that specific task, could be neglected. In order to reduce these deviations, we are proposing to adopt a task-dependent approach, based on the comparison between the collected data and some database concerning a different task, assumed as a baseline. After giving the necessary theoretical back-ground, we test our methodology on real EEG data involving two subjects in a musical task. In addition to some natural results, new and unexpected neurological links can be emphasized and discussed.*

## Key words

*Brain networks • Functional connectivity • Graphs • EEG data • Musical task • Threshold*

## Introduction

The construction of a brain network from neurobiological data is really a critical point. It is usually represented in terms of a graph, consisting of nodes and links between pairs of nodes, called edges of the graph. In large-scale analysis, nodes relates to brain regions, while links represent statistical correlations between such regions. Differently, small-scale analysis generally focuses on a specific neural system, either a sub-system of interest (retina, cortical column, etc.) or a small but entire nervous system (e.g., *Caenorhabditis ele-*

*gans*, fruit fly, or grasshopper).

Generally, the number of data obtained performing any kind of experiment (M/EEG, fMRI, PET, etc.) is huge, for example 100 nodes could lead to nearly 10,000 data. As a consequence, some threshold must be introduced, in order to emphasize the main interesting outcomes. The choice of suitable thresholds is really important since these should be able to separate meaningful links from weak and non-significant links that might represent spurious connections, particularly in functional or effective networks.

The problem of thresholding data in Neuroscience can be considered from various points of view. For example, one could concern the selection of suitable thresholds during the process of data acquisition. Information is usually encoded in sequences of spikes, obtained when the membrane potential of the neurons (difference in electrical potential between interior and exterior) crosses a spiking threshold. It varies with firing history and input properties, and different methodologies must be exploited for detecting the main meaningful links between the neural activity and the recorded signals (see for example Lewicki et al., 2009).

A different procedure could depend on the performed experiment, where, sometimes, absolute thresholds might be introduced, based on a predetermined percentage of time that a given stimulus is detected (Waiten, 2007, pp. 120-121; Colman, 2009).

On the other hand, an important aspect of thresholding relates to the analysis of the collected data, and this is the topic we wish to deal with in the present paper.

For a neural network consisting of  $n$  nodes, a connectivity matrix  $A=[a_{ij}]$  is considered, where  $a_{ij}$  represents the detected correlation between the nodes  $i$  and  $j$ , where  $i, j$  belong to  $\{1, \dots, N\}$ . After connectivity matrices have been formed, a new thresholding methodology must be applied to their entries, in order to extract properly the functional weights of the corresponding cerebral areas.

Our proposal is to adopt a task-dependent approach, based on the comparison between the collected data and some database concerning a different task, assumed as a baseline. Basically, our methodology can be adapted to a wealth of data sets which are typically gathered in human subjects by non-invasive methods, such as M/EEG (Rutter et al., 2013) or fMRI (Friston K.J., 1994), but, in principle, there are no restrictions in applying such a methodology to other in-vivo imaging techniques (Friston et al., 1993).

As an applicative illustration, we have tested our proposal on real EEG data (released by the Carlo Besta Neurological Institute, Italy), concerning two different subjects (in the following referred to as Subject A and Subject B), undergoing the same

musical experiment. More precisely, we investigated the changes in the functional connectivity produced by transition from a Basal to a Pleasant condition, due to musical experiments. The Basal test consisted in a kind of resting state (which can be assumed as a sort of baseline), and was determined when no kind of music was proposed to the volunteers. The Pleasant test was obtained while listening a piece of music that, according to previous studies (Koelsch et al., 2006), is considered as pleasant.

In Section 2, we outline the general proposed approach, to be adopted in comparing different homogeneous (same subjects) databases.

Section 3 shows how the methodology can be applied, giving all details, global comments and remarks. In Subsections 3.5.2, 3.5.1 we provide an example of graph theoretical approach to real data. The results are shown in Section 4 ‘Analysis and discussion of the results’, while in Section 5 ‘Conclusions’ we resume our conclusions and outline possible extensions of our work.

Table I. Brain Atlas

Node	Cerebral Area	Node	Cerebral Area
1	Left Amygdala	46	Right Inferior Occipital Gyrus
2	Right Amygdala	47	Left Middle Occipital Gyrus
3	Left Angular Gyrus	48	Right Middle Occipital Gyrus
4	Right Angular Gyrus	49	Left Superior Occipital Gyrus
5	Left Calcarine Sulcus (or Left Calcarine fissure)	50	Right Superior Occipital Gyrus
6	Right Calcarine Sulcus (or Right Calcarine Fissure)	51	Left Olfactory Cortex
7	Left Caudate Nucleus	52	Right Olfactory Cortex
8	Right Caudate Nucleus	53	Left Globus Pallidus
9	Left Anterior Cingulum	54	Right Globus Pallidus
10	Right Anterior Cingulum	55	Left Parahippocampal Gyrus
11	Left Middle Cingulum.	56	Right Parahippocampal Gyrus
12	Right Middle Cingulum.	57	Left Paracentral Lobule
13	Left Posterior Cingulum	58	Right Paracentral Lobule
14	Right Posterior Cingulum	59	Left Inferior Parietal Lobule
15	Left Cuneus	60	Right Inferior Parietal Lobule
16	Right Cuneus	61	Left Superior Parietal Lobule
17	Left Inferior Frontal Gyrus, Opercular Part	62	Right Superior Parietal Lobule
18	Right Inferior Frontal Gyrus, Opercular Part	63	Left Postcentral Gyrus
19	Left Inferior Frontal Gyrus, Orbital Part	64	Right Postcentral Gyrus
20	Inferior Frontal Gyrus, Orbital Part	65	Left Precentral Gyrus
21	Left Inferior Frontal Gyrus, Triangular Part	66	Right Precentral Gyrus
22	Right Inferior Frontal Gyrus, Triangular Part	67	Left Precuneus
23	Left Medial Orbitofrontal Cortex	68	Right Precuneus
24	Right Medial Orbitofrontal Cortex	69	Left Putamen
25	Left Middle Frontal Gyrus	70	Right Putamen
26	Left Middle Frontal Gyrus, Orbital Part	71	Left Gyrus Rectus
27	Right Middle Frontal Gyrus, Orbital Part	72	Right Gyrus Rectus

28	Right Middle Frontal Gyrus	73	Left Rolandic Operculum
29	Left Superior Frontal Gyrus	74	Right Rolandic Operculum
30	Left Medial Frontal Gyrus	75	Left Supplementary Motor Area
31	Right Medial Frontal Gyrus	76	Right Supplementary Motor Area
32	Left Superior Frontal Gyrus, Orbital Part	77	Left Supramarginal Gyrus
33	Right Superior Frontal Gyrus, Orbital Part	78	Right Supramarginal Gyrus
34	Right Superior Frontal Gyrus	79	Left Inferior Temporal Gyrus
35	Left Fusiform Gyrus	80	Right Inferior Temporal Gyrus
36	Right Fusiform Gyrus	81	Left Middle Temporal Gyrus
37	Left Heschl's Gyrus or Left Transverse Temporal Gyrus	82	Right Middle Temporal Gyrus
38	Right Heschl's Gyrus or Right Transverse Temporal Gyrus	83	Left Middle Temporal Pole
39	Left Hippocampus	84	Right Middle Temporal Pole
40	Right Hippocampus	85	Left Superior Temporal Pole
41	Left Insula	86	Right Superior Temporal Pole
42	Right Insula	87	Left Superior Temporal Gyrus
43	Left Lingual Gyrus	88	Right Superior Temporal Gyrus
44	Right Lingual Gyrus	89	Left Thalamus
45	Left Inferior Occipital Gyrus		

## A mathematical proposal to data-depending thresholds

In this section we wish to outline how, in our opinion, a good data-depending threshold should be selected. We assume to perform a neurobiological experiment on a number of volunteers. For each volunteer,  $s$  different tests are performed, each one having duration  $T$ . In order to process data the temporal length  $T$  is split in  $N$  sub-intervals, also referred to as temporal slots, or epochs. Data are collected in matrices whose sizes depend on the number of nodes (denoted by  $n$ ) and on the number of temporal sub-intervals. Hence, we deal with a collection of  $N$  matrices of size  $n \times n$ , namely, every temporal slot is characterized by functional data collected in the form of a matrix with  $n$  rows and  $n$  columns. This holds for each one of the  $s$ -performed test. As well-known (Friston, 1994; Kaiser, 2011; Rubinov and Sporns, 2010; Sporns, 2010), each entry encodes statistical dependence between two voxels, each one corresponding to neural elements, which can be cerebral areas or part of them. Functional brain connectivity forms a full symmetric matrix, where the range of the entries is the  $[0,1]$  real interval (generally, we do not consider negative entries since anti-correlation is not taken into account). We also assume zero entries on the main diagonal, which represents the absence of self-loops in the neural brain network, meaning that node auto-correlation is excluded.

In this paper we deal with functional connectivity and Graph Theory applied to Neuroscience. The following subsections focus on the description of the brain connectivity and on application of the Theory of Graphs to Neuroscience. We refer the readership to (Finotelli and Dulio, 2015) for a more detailed description.

### 2.1. A bird's eye view on the brain connectivity

There are three kinds of brain connectivity: structural, functional and effective connectivity. These refer to a pattern of links between distinct units (individual neurons, neuronal populations, or brain regions) within a nervous system.

- Structural Connectivity (SC): a description of the anatomical connections between network nodes (i.e., brain regions, neurons); for example, reconstructed anatomical projections derived from diffusion MRI, directed anatomical pathways derived from neural tract tracing, or synaptic connections between individual neurons.
- Functional Connectivity (FC): it has been defined as “the temporal correlation of a neurophysiological index measured in different brain areas.” FC represents the coupling between dynamic activity recorded from separate brain areas. Experimentally FC analysis has been used to describe temporal correlations across multiple spatial scales in PET imaging, M/EEG and fMRI. These findings have been used to identify co-activating brain regions as functional networks.
- Effective Connectivity (EC): particular kind of FC which is related to causal relationships between neural units. Similarly to FC, EC is also a measurement of neural interactions in neuroimaging, more specifically it considers simultaneous interaction of several neural elements to explicitly quantify the effect one element has on another. Using dynamic causal modeling (DCM) (see Frackowiack et al., 2004) to estimate

EC, Mechelli et al. (2003) found that early sensory areas changed their effects on category-specific areas in relation to the stimuli, whereas higher-order association regions did not show such changes.

## 2.2. A brief introduction to Theory of Graphs in Neuroscience

There is a wealth of paper and books on these topics (for a survey, see for example Finotelli and Dulio, 2015). The brain is a highly interconnected network; more precisely, it has a complicated SC and FC among its basic constituents, i.e. neurons, synapses and brain regions. The human brain is a complex network whose operations depend on how its neurons and/or cerebral areas are linked together.

A graph is a mathematical representation of a network. Usually, it is denoted by  $G(V,E)$ , and consists of two sets. The first,  $V$ , is the set of vertices (or nodes), and the second,  $E$ , is the set of edges, consisting of pairs of elements from the vertex set. We can distinguish four classes of graphs: directed, undirected, weighted and binary. This relates to the structure of the edges. An edge is directed if its starting and ending vertices are specified, otherwise the edge is said to be undirected. Whenever it is possible to associate a number (for example a probability) to an edge, then this number is called weight of the edge. Differently, if no weight is considered, then the graph is binary. A binary graph can be represented by a binary  $|V| \times |V|$  matrix  $A$ , where  $a_{ij}=1$  if an edge links the vertices  $i, j$ , and  $a_{ij}=0$  otherwise. Historically, Graph Theory was born in the 18<sup>th</sup> century, thanks to Leonhard Euler. He solved the challenging problem of finding a walk through the Prussian city of Königsberg, crossing each one of its seven bridges precisely once.

In Neuroscience, Graph Theory helps in investigating the functional and the structural connectivity among the cerebral areas. In case of functional connectivity, edges represent statistical relationships between couple of different cerebral areas (labeled with nodes), while, for structural connectivity, edges are defined by anatomical connections, the white fibers (rather

than statistical relationships).

Thought, at present, the relationship between SC and FC is not completely clear, it is possible to think of structural brain networks as the physical substrate which underpins the functional brain networks, at least at the resting state (Grecius et al., 2009).

A model of the brain should be sensitive to factors which may alter the underlying connectivity structure or cognitive functionality. Complex brain networks show sensitivity to: behavioral variability (Bassett and Bullmore, 2009), cognitive ability (Van den Heuvel et al., 2009), genetic information (Schmitt et al., 2008), aging (Meunier et al., 2009), gender (Gong et al., 2009), drug treatments (Achard et al., 2007), neurological diseases (for example Alzheimer's (Stam et al., 2009), schizophrenia (Lynall et al., 2010) and a clinical state such as acute depression (Leistedt et al., 2009). Complex network theory is also particularly appealing in the study of clinical neuroscience where disease and other clinical states have been characterized by hypo-connectivity, hyper-connectivity, and non-connectivity profiles.

A very important role in the graph theoretical analysis is played by the metrics. Metrics can be divided into two categories: topological and physical. Some important topological metrics are, for instance, the degree of a node, the presence of cycles, the clustering coefficient, the local and global efficiency, the modularity. Some usually employed physical metrics are the mean connection distance, the tract length and the Rent's exponent. For further details we refer the reader to Sporns (2010) or other similar books.

In the authors' opinion, a basic example of application of Graph Theory to Neuroscience is the Default Mode Network (DMN). The DMN is characterized by the activation of specific cerebral areas during the cognitive rest. Basically, such areas are: The medial prefrontal cortex, medial temporal lobes, posterior cingulate cortex, and the precuneus. These areas, which are strongly mutually connected, provide the nodes making up the vertex set of the graph representing the DMN (DiPasquale et al., 2016; Grecius et al., 2009; Raichele, 2015).



### 2.3. Statistical approach to threshold selection for the analysis of analogies and discrepancies between two tests

The first step in the analysis consists in importing the available row data of each volunteer in terms of manageable files. This can be done by means of some technical computing software (in our approach we have employed Matlab<sup>®</sup>). In order to save space the storage of data can be confined in the upper (or lower) triangular part of the matrix, which leads to  $N \times s$  different triangular  $n \times n$ -sized matrices  $A_i^{(s)}$ . Moreover, each entry should be rounded at some decimal digit, depending on the demanded precision. For each fixed epoch  $t$ , and for each test  $s$ , the entry  $a_{hk}$  of  $A_i^{(s)}$  denotes the functional connectivity between the nodes  $h$  and  $k$ , at the fixed time  $t$ , and during the test  $s$ .

Once the  $s$  tests have been encoded in the collections  $\{A_i^{(s)} : i \in \{1, \dots, N\}\}$ , the comparison between any two different collection should be considered, in order to point out, for a same volunteer, the main analogies and discrepancies resulting from the different tests. A natural approach is to manage the matter by computing the difference matrices of corresponding elements. However, a main problem in this analysis could be caused by the fact that each one of the  $s$  available tests has no temporal correlation with the others, so that it is not clear which elements of two different collections should be considered as correspondent. This means that, for a fixed index  $i \in \{1, \dots, N\}$ , the matrix  $A_i^{(s)}$  in general refers to different temporal slots for different values of  $s$  (performed tests). For instance, the matrix  $A_3^{(1)}$ , belonging to the first test and collected at the third epoch, is not necessarily related to the matrix  $A_3^{(2)}$ , collected at the same third epoch but concerning the second test. As a consequence, a precise picture of the comparison between a pair of tests requires a kind of randomization, namely, we must compute all the possible differences between pairs of matrices of the corresponding collections. In detail, for each test  $\bar{s}$ , and for each fixed temporal slot  $\bar{i} \in \{1, \dots, N\}$ , the matrix  $A_{\bar{i}}^{(\bar{s})}$  should be cross-confronted with all the other  $(s-1) \times N$  ma-

trices, that is, we must consider  $A_i^{(\bar{s})} - A_i^{(s)}$  with  $\bar{s} \neq s$  and  $i \in \{1, \dots, N\}$ .

For any pair  $\bar{s}, \tilde{s}$ , such a randomization creates a cell of order  $N \times N$  made up of matrices whose size is  $n \times n$ .

Since multiple tasks should be compared pairwise, in what follows we can focus, without loss of generality, just on the main case when two different tests  $s = \bar{s}$  and  $s = \tilde{s}$  are performed. If we assume one of them, say  $\bar{s}$ , as a reference baseline, then the comparison procedure detailed below can be interpreted as a general methodology of thresholding the data collected during the test  $\tilde{s}$ , on the basis of their analogies and discrepancies with respect to the baseline.

#### 2.3.1. Analogies and discrepancies between a pair of databases

Assume that two tests  $s = \bar{s}$  and  $s = \tilde{s}$  have been proposed to some volunteers. For each volunteer, we need to consider the matrices  $A_i^{(\bar{s})} - A_j^{(\tilde{s})}$ ,  $i, j \in \{1, \dots, N\}$ , with  $i$  fixed time by time.

In order to point out analogies and discrepancies between the two tests we focus on the entries having, respectively, smaller and greater absolute values, because they denote the greater agreement and difference in correlations. Since we are dealing with the functional connectivity, the entries of each matrix can assume values in  $[0, 1]$ , so that the range of the entries of each matrix  $A_i^{(\bar{s})} - A_j^{(\tilde{s})}$  is the  $[-1, 1]$  real interval. Therefore one should be mainly interested to center the focus on a right neighbor of -1, on a left neighbor of +1, and on a neighbor of 0. We call neighbors of interest (NOI) such neighbors. The width of the NOI depends on the distribution of the entries of each matrix. The right neighbor of -1 represents the situation where the test  $\tilde{s}$  dominates  $\bar{s}$ , while the left neighbor of +1 refers to the opposite situation. The neighbor of 0 gives information about the domain where the tests  $\bar{s}$  and  $\tilde{s}$  can be assumed as invariant. With the purpose of analyzing the entries of the matrices  $A_i^{(\bar{s})} - A_j^{(\tilde{s})}$ , we organize each column, or equivalently each temporal slot  $i$ , in a single  $n \times (n \times N)$ -sized matrix  $C_i = [c_{hl}]_i$  (where  $h \in \{1, \dots, n\}$ ,  $l \in \{1, \dots, n \times N\}$ , and

$i \in \{1, \dots, N\}$ ), obtained by connecting the rows of all the matrices  $A_i^{(\bar{s})} - A_j^{(\bar{s})}$  of size  $n \times n$ ,  $j \in \{1, \dots, N\}$  (see Equation (1) below).

In summary, for each column, that is related to a temporal slot, a single matrix is obtained, made up by placing side by side  $N$  matrices of size  $n \times n$ . For example, for the first temporal slot  $i=1$ , the matrices  $A_1^{(\bar{s})} - A_1^{(\bar{s})}$ ,  $A_1^{(\bar{s})} - A_2^{(\bar{s})}$ ,  $\dots$ ,  $A_1^{(\bar{s})} - A_N^{(\bar{s})}$  must be placed side by side, in order to get one single matrix  $C_1 = [c_{hl}]_1$ , where  $h \in \{1, \dots, n\}$ ,  $l \in \{1, \dots, n \times N\}$ . Similarly for the other temporal slots. At the end of this process  $N$  different matrices  $C_1 = [c_{hl}]_1$ ,  $C_2 = [c_{hl}]_2, \dots, C_N = [c_{hl}]_N$  are generated, that is, for all  $i \in \{1, \dots, N\}$ , we have

$$C_i = [A_i^{(\bar{s})} - A_1^{(\bar{s})} | A_i^{(\bar{s})} - A_2^{(\bar{s})} | \dots | A_i^{(\bar{s})} - A_N^{(\bar{s})}]. \quad (1)$$

Once such matrices are obtained, an analysis on their entries is demanded. The entries of every matrix  $C_i$  are collected in a single array by concatenating the  $N$  rows. This means that  $N$  arrays, each one composed by  $n \times (n \times N)$  elements, are generated. This operation is needed to start the statistical analysis.

**Remark.** A very special case is obtained when  $\bar{s}$  concerns the resting state. Actually, it is the default neurobiological baseline, therefore no test  $\bar{s}$  must be performed, and, consequently, no randomization is required. In this case the matrices  $C_i$  are simply replaced by  $A_i^{(\bar{s})}$ , for all  $i \in \{1, \dots, N\}$ , so we have just one matrix for each epoch. In case different volunteers are considered, then the entries of  $C_i$  are assumed to be the mean values of the corresponding entries in the matrices associated to each volunteer, at the same epoch  $i \in \{1, \dots, N\}$ . We have already applied our method to fMRI real data coming from resting state experiments on 133 different right-handed subjects, males and females, whose age spanned between 6 and 79 years, see the poster by Dipasquale et al. (2015). See also Section 4 ‘Analysis and discussion of the results’ for further comments.

### 2.3.2. First thresholding step

We wish now to detail all the steps leading to the matrices that we propose to work with in performing a neurobiological analysis. Since we are assuming as a reference database, any threshold introduced on the distribution the differences represents an intrinsic thresholding procedure concerning the database  $\bar{s}$ , normalized with respect to the baseline.

As a first step, the distribution of the entries of every array  $C_i$  (consisting of the differences between the databases) should be computed, and increasingly ordered, so obtaining an ordered distribution that can be visually represented in the form of a histogram (see Figure 1). The number of bins can be arbitrarily selected. This is just related to the resolution visualization of the ordered distribution, and has no influence on the further thresholding steps. Since the range of the entries does not necessarily span the whole interval  $[-1, 1]$ , using increasing numbers of bins might allow a better visualization of the tails.

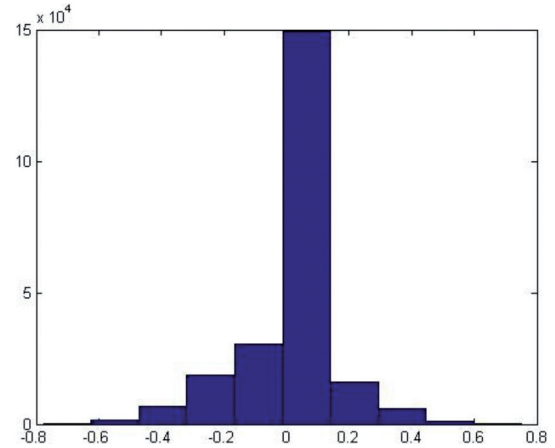


Fig. 1 - A typical histogram representing the frequency of the entries of matrix  $C_i$ , i.e. the distribution of the values of the entries of  $C_i$ ,  $i=1, \dots, N$ .

In particular, if one is interested in the analysis of the discrepancies between the two databases, namely, on the links where  $\bar{s}$  is strongly different from, then the focus is on the tails of the distribution. Differently, if one is interested in the analogies between the databases, namely on the links where  $\bar{s}$  mainly matches, then the focus is on the center of the distribution. Therefore, starting from the ordered distribu-

tion, and according to the specific analysis to be done, we wish to extract a NOI of suitable width. This is obtained by selecting a reasonable percentage of data to be assumed as meaningful for the experiment in the required NOI. To this, we proceed as follows.

- Increasingly ordering  $C_i$ , we get a distribution  $D_i$ , from which we extract the sub-distribution  $D_i^+$ , consisting of the strictly positive entries, and the sub-distribution  $D_i^-$ , consisting of the strictly negative entries.
- Define a quantile subdivision for  $D_i^+$  and  $D_i^-$ . We recall that  $q$ -quantiles partition a distribution into  $q$  subsets, each one containing the  $(100/q)\%$  of the data of the distribution. Therefore, for each integer  $k$  such that  $1 \leq k < q$ , the  $(100k/q)\%$  of the data cumulates below the  $k$ -th quantile, and, of course, the  $(100(q-k)/q)\%$  above. The choice of the quantile subdivision reflects in the width with which we can investigate the NOI. Since we are usually working with a great number of data, a percentile choice, namely  $q=100$ , is mainly suggested.
- Let's now focus on the positive distribution  $D_i^+$ , and select a suitable quantile  $k$  for the analysis of its right tail, corresponding to the NOI of  $+1$ . The length of the tail depends on  $k$ , and consists of the percentage  $\bar{p} = (100(q-k)/q)\%$  of the highest values of  $D_i$ . Of course, the length also depends on  $q$ . The greater  $q$ , the shortest is the tail, and consequently the stronger discrepancies between  $\bar{s}$  and  $\bar{s}$  are detected.
- The same is done for the negative distribution  $D_i^-$ , which leads to the construction of its left tail, corresponding to the NOI of  $-1$ , and formed by the percentage  $\bar{p} = (100(q-k)/q)\%$  of the entries below a quantile  $k'$  suitably selected for the analysis of lowest values of  $D_i$ .
- Concerning the choice of  $k, k'$ , we can investigate on the symmetry of the dis-

tributions  $D_i$ , for instance by comparing  $D_i$  with a normal distribution. This can be done by means of a Q-Q Plot, namely, by plotting the quantiles of  $D_i$  against the quantiles of a normal distribution. In case the answer is positive, namely if the Q-Q Plot is almost linear, then we can assume  $k'=q-k$ , so that the same percentage is employed for selecting the lowest data for the left tail and the highest data for the right tail. The reference quantile  $k$  can be meaningfully selected according to some standard choice, such as the 95-th percentile, or even the 99-th percentile for emphasizing very strong tails.

- Differently, if  $D_i$  does not show symmetric properties, then a convenient choice of  $k, k'$  demands a deeper analysis of the sub-distributions  $D_i^-$  and  $D_i^+$ . In particular,  $k, k'$  should be fixed according to the difference in cardinality of the positive and the negative sub-distributions.
- The same Q-Q Plot comparison can be performed, separately, for the distributions  $D_i^-$  and  $D_i^+$ , in order to check their possible symmetry. In case, we define the NOI of 0 as formed by the union of a negative interval, a zero interval, and a positive interval. The negative interval is the right tail of  $D_i^-$ , consisting of the  $\bar{p} = (100(q-k)/q)\%$  of its highest values. The positive interval is the left tail of  $D_i^+$ , consisting of the  $\bar{p} = (100(q-k)/q)\%$  of its lowest values. The zero interval includes all the links corresponding to  $D_i \setminus (D_i^- \cup D_i^+)$ .
- Differently, if the values of  $D_i^-$  or of  $D_i^+$  (or of both) are not symmetrically distributed, then the suitable quantiles involved in the construction of the right tail of  $D_i^-$ , and of the left tail of  $D_i^+$ , should be selected by means of a deeper analysis of the distributions far away from the corresponding mean values.

Once the reference quantile positions  $k, k'$  have been fixed, their specific values  $\xi_i$  must be computed for each  $i \in \{1, \dots, N\}$ , which reflects in the introduction of a different threshold for each one of the matrices  $C_1, \dots, C_N$ . Concerning the neighbor of -1, and denoting by  $\xi_i(L)$  the value of the quantile  $k'$  at epoch  $i$  ( $L$  stands for left tail), then we have  $\bar{C}_i = [C_i]_L = [\bar{c}_{hl}^i]$ , where

$$\bar{c}_{hl}^i = \begin{cases} 1 & \text{if } c_{hl}^i \leq \xi_i(L), \\ 0 & \text{otherwise.} \end{cases}$$

This turns  $C_i = [c_{hl}^i]$  into a binary matrix  $\bar{C}_i = [\bar{c}_{hl}^i]$ , and consequently related to an undirected graph (Salvador et al., 2005). Analogously, for a neighborhood of +1, denoting by  $\xi_i(R)$  the value of the quantile  $k$  at epoch  $i \in \{1, \dots, N\}$  ( $R$  stands for right tail), we get the binary matrix  $\tilde{C}_i = [C_i]_R^* = [\tilde{c}_{hl}^i]$ , where

$$\tilde{c}_{hl}^i = \begin{cases} 1 & \text{if } c_{hl}^i \geq \xi_i(R), \\ 0 & \text{otherwise.} \end{cases}$$

In case of a neighbor of 0, both  $\xi_i(L)$  and  $\xi_i(R)$  must be employed, and the interval  $(\xi_i(L), \xi_i(R))$  is considered.

This thresholding procedure keeps into account the real meaning of the data for the volunteer, and consequently it seems of special interest for subject-dependent neurobiological analysis. Also, it is easily adaptable to different evaluations, simply changing the reference quantile, and the corresponding percentage of data forming the size of the NOI. Basically, this may be interpreted as a control of the degree of sparsity of the available data, and it is in accordance to what is usually expected (Rubinov and Sporns, 2010; Sporns, 2010).

**Remark.** In the case of the resting state, the tails refer to the neighbors of 0 and of 1 (since no task  $\bar{s}$  is performed). In particular, one is mainly interested in the neighbor of 1, where the highest functional connectivity is registered.

### 2.3.3. Second thresholding step

For each  $i \in \{1, \dots, N\}$ , the corresponding  $N$  left thresholded  $n \times n$ -sized binary matrices  $\bar{C}_i$  should be summed up to form a matrix  $S_i(L)$  that condenses the information in the right neighborhood of 1. Each entry of  $S_i(L)$  ranges from 0 to  $N$ , which denotes the number of slots of  $T/N$  seconds where the corresponding pair of nodes are correlated. Analogously, a matrix  $S_i(R)$ , where the information in the left neighborhood of 1 are condensed, is obtained.

From these matrices several types of graphs can be derived that could be assumed as representative of the transition between a test  $\bar{s}$  and  $\tilde{s}$ . These could be obtained by fixing a new threshold  $\lambda$ ,  $\lambda \in \{1, \dots, N\}$ , corresponding to the minimum number of slots of  $T/N$  seconds that we assumed as indicative of a permanent difference between the  $\bar{s}$  and the  $\tilde{s}$  states. Then we binary threshold each entry  $\alpha$  of  $S_i(L)$  by setting  $\alpha=1$  or  $\alpha=0$  according as  $\alpha \geq \lambda$  or  $\alpha < \lambda$ . Of course  $\lambda=N$  denotes an absolute permanent difference, and represents a theoretical threshold. More realistically, we should relax this threshold to some intermediate value. To this, it is advisable to compute the maximum entry  $M_i$  of each matrix  $S_i(L)$ , and investigate the cases where  $\lambda$  concerns some percentage of  $M_i$ , such as  $\lambda = 4/5M_i$ , or  $\lambda = 1/2M_i$ , corresponding, respectively, to keep all the entries that appear in the 80%, or in the 50% of the slots.

According to the selected value of  $\lambda$ , the representative graph changes. The greater  $\lambda$ , the smaller is the number of extracted links, so that their importance in the neurobiological interpretation increases. On the other side, higher values of  $\lambda$  could miss some important links that, despite their lower weights, could be worth being considered. To this, we suggest to investigate different values of  $\lambda$ , so providing, for a same  $\bar{p}$ , representative graphs having increasing structural richness, and where weaker links, previously excluded, progressively can appear.



### 2.3.4. Graph theoretical approach

Theory of Graphs (Bollobás, 1985) is used to model the brain seen as a complex network (Bullmore and Bassett, 2009; Bullmore and Sporns, 2009; Finotelli and Dulio, 2015; Rubinov and Sporns, 2010; Sporns, 2010) both for healthy subjects and in case of brain disorders (Fornito et al., 2015; Kazaei et al., 2015). This kind of analysis bases on the weights associated to the various edges. In Finotelli and Dulio (2015b), for example, a model to evaluate weights in functional connectivity has been proposed and commented. The selection of the meaningful edges follows from the adopted thresholding procedure. In this respect, after the two steps thresholds have been introduced, our proposal leads to a data-depending edge selection forming a graph  $G_{\bar{p}, \lambda}$ , which describes the brain network associated to the performed test, in the selected NOI, and depending on the considered parameters,  $\bar{p}$  and  $\lambda$ , employed as thresholds.

## Experimental procedure

As an applicative example we consider an acoustical experiment, where two volunteers were subjected to a pleasant music and a resting state.

A similar example was treated in Wilkins et al. (2014) and in Wu et al. (2013). In such papers the effect of music on the brain is considered. From (Wilkins et al. (2014), it comes out that regardless of the acoustical characteristic, functional brain connectivity depends on whether the music is liked or disliked. Interestingly, listening to music that is liked affects functional connectivity in regions involved in self-referential thought and memory encoding, such as the default mode network and the hippocampus.

Moreover, Wu et al. (2013) reveals that music perception induces an increased synchronization of cortical regions and amore random network configuration. The difference in network architecture is not observed between musical stimuli. These findings imply that music perception leads to a more efficient, but less economical, structure, and requires more information processing as well

as cognitive effort.

### 3.1. Subjects and EEG recording protocol

The study involved 2 healthy subjects, the subject A is a male of 39, while B is a female whose age is 27, without any professional musical training. On the other hand, due to differences in age, sex and stimulus reactivity, we expect to find differences as well.

The present study was approved by the Ethics Committee of Carlo Besta Neurological Institute Milan, Italy. Written informed consent was obtained from the subjects.

Hd-EEG (High density EEG) were recorded in a dimly lit room using using a Waveguard 128-channel electrode with active cable-shielding technology (Advanced NeuroTechnology (ANT), Enschede, Netherlands), and acquired at a sampling rate of 512 Hz (ASA Lab, Advanced NeuroTechnology (ANT), Enschede, Netherlands). Each channel was referenced to the average of all channels.

The EEG session started with a 5 minutes eyes-closed resting-state recording. Thereafter, subjects were presented with one minute length of a musical piece, *J. S. Bach: Rejouissance BWV 1069*, alternate with silent condition. This piece was defined as 'pleasant' stimulus, according to previous fMRI and EEG studies (see Sammler et al., 2004 and Koelsch et al., 2006). (sound sample of the stimulus is available at [http://www.stefan-koelsch.de/Music\\_Emotion1](http://www.stefan-koelsch.de/Music_Emotion1)).

Before the analysis, EEG recordings were filtered using a digital 1-120 Hz (12 db/octave) band-pass filter followed by a 50 Hz notch filter to suppress the noise of the electrical powerline. Moreover, the EEG data were normalized by subtracting the mean value and dividing the result by the standard deviation. The EEG section during the pleasant musical stimulus was considered for the analysis, and compared with 1 minutes of EEG recorded at basal condition, free of any physiologic (e.g. eye-blink) and non-physiologic artifacts. 1 min of artifact-free EEG recorded at resting condition<sup>1</sup>.

<sup>1</sup>We remind the reader that the basal condition is a sort of "baseline". It consists of a minute of recording EEG signal while the subject is sit relaxed with her his eyes closed. The

### 3.2. Beamforming source reconstruction

For the construction of the beamformer spatial filter's weights, a time window of 120 seconds (one minute resting-state and one minute during the "pleasant" condition) was used for the computation of the data covariance matrix (Hillebrand et al., 2012). For each recording the forward problem was solved and expressed as a leadfield matrix (i.e. the matrix that contains information on the geometry and conductivity of the model and represents the linear relationship between source-space and measurements). The leadfield matrix was based on a volume conduction model created from a segmentation of the MNI template MRI into brain, skull and scalp compartments having conductivity values of 33 mS/m, 0.41 mS/m and 33 mS/m, respectively. A template electrode set according to the international 10-05 system was used (Oostenveld et al., 2001).

Sensor data was projected into source space by means of a linearly constrained minimum variance beamformer (Van Veen et al., 1997). The LCMV beamformer acts as a spatial filter whose weights are computed to maximize the sensitivity to signals originating from a specific location of interest, whilst reducing the signals from other sources. The beamformer weights for a source at a location of interest are fully determined by the data covariance matrix and the forward solution (leadfield) (Van Veen et al., 1997). The spatial filters obtained for each subject were multiplied with the EEG sensor time series to derive the time course of each voxel (or virtual electrode), defined on a regular 5mm spaced grid covering the entire brain volume (in total 7,780 virtual electrodes were considered).

We clustered the entire set of virtual electrodes in regions of interests (ROIs) according to the AAL atlas by Tzourio-Mazoyer (2002), and assigned to each electrode the label of the corresponding AAL region. For this study 89 regions of the AAL atlas (out of 116) were considered (see Table 1 for a list of the labels used). It should be noted that the number of virtual electrodes was different across ROIs.

According to this anatomical clustering, we reduced the dataset in source space by implementing a strategy described in Hillebrand et al. (2012). For each subject, the spectrum for each virtual electrode time-series was computed, and divided into the 5 classical EEG bands (delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (30-48 Hz)). For each ROI and frequency band separately, we selected the voxel with maximum relative power in the theta band<sup>2</sup>, and used the time-series for this voxel for further connectivity analysis. As a result, we obtained a set of 89 virtual electrodes time series (120s long) for each subject. We decided to focus on the theta band because previous studies demonstrated its relevance in the cognitive processing of musical stimuli (Sammler et al., 2007; Flores-Gutiérrez et al., 2007).

As a final step, the virtual electrodes time courses of each condition (basal and pleasant) were divided into 29 non-overlapped, artifact-free, 2 s length epochs. All the previous computations were performed by custom-made software in Matlab<sup>®</sup> and functions of the Fieldtrip toolbox (Oostenveld et al., 2011).

### 3.3. Connectivity Analysis

The adjacency matrix were estimated by applying the phase locking value (PLV) connectivity method to the source reconstructed signals. The phase locking value (PLV) was first introduced by Lachaux et al. (1999), as a simple and practical method to evaluate frequency-specific phase synchronization between couples of neurophysiological signals.

The PLV formulation was inspired by the first experimental evidences showing that the large-

---

<sup>2</sup> The signal of each source (or virtual electrode) was analyzed by the FFT (Fast Fourier Transform), then, it was calculated the relative power in theta band, which corresponds to the integral of the spectrum between 4 and 8 Hz normalized with respect to the total power (i.e. the integral of the spectrum along the band). At this point, each ROI of the atlas (see Table 1) contains many virtual electrodes (any ROI is associated to nearly hundred virtual electrodes), each with its own value of relative power in theta band. In order to select a single virtual electrode (with its signal), representative of a given ROI, it was chosen the one that showed the maximum power in theta band within that given ROI.

---

subject do not receives any stimulation.

scale integration mechanisms of the brain are supported by the ability of different neuronal populations to oscillate at different frequency ranges and to exchange information by entering into phase locking (see for details Lachaux et al., 1999).

The original process to compute the PLV was aimed at quantifying the stability of the phase differences between couples of signals across multiple trials of the same process. In order to extract the instantaneous phases of each signal at a desired frequency range, we initially band-pass filtered the EEG signals for each experimental condition in 5 different frequency bands:  $\delta$ (0.5-4 Hz),  $\theta$ (4-8 Hz),  $\alpha$ (8-13 Hz),  $\beta$ (13-30 Hz),  $\gamma$ (30-80 Hz). Then the Hilbert transform was used to obtain the analytic phase for a frequency band at a time point  $t$  as:

$$\varphi_x(t) = \arctan\left(\frac{u_x(t)}{v_x(t)}\right), \quad (2)$$

where  $v$  is the real part and  $u$  the imaginary part of the analytic signal, and subscripts  $x$  and  $y$  indicated the two EEG time series being analyzed. The PLV is then defined as

$$PLV = \frac{1}{N} \left| \sum_{t=1}^N e^{i(\phi_x(t) - \phi_y(t))} \right| \quad (3)$$

where  $N$  represents the number of time points in a time window and  $\|\cdot\|$  the complex modulus.

If the phase difference varies little the PLV value is close to 1, otherwise if the phases are uniformly spread on the unit circle the PLV, averaged across a sufficient number of epochs, will be close to 0.

For each temporal epoch, PLV was calculated between each pair of signals, thus obtaining for each subject 3 dimensional Adjacencies Matrices of size: [number of channel x number of channel x number of epochs].

**Remark.** Using individual MRI scans could, in principle, lead to better results. However, we carefully adopted the strategies suggested by previous studies in case the individual MRIs are not available, see for example Michel et al.

(2004). In detail, as suggested by Acar et al. (2013) we used a 4-layer template BEM, accurately warped to a 128 co-registered electrode set. An example of other studies using this technique is shown in Bathelt et al. (2013); in this study the authors demonstrated that, although using a template MRI, important and significant insights about the network organization of the brain can be gained. The selection of a spatial subsample of voxels is a necessary step to reach a physiologically relevant trade-off between computational efficiency (connectivity between tens of thousands voxels would be highly inefficient and could seriously lead to the detection of spurious correlations) and a functionally relevant spatial sampling of the human brain. As demonstrated in Hillebrand et al. (2012) choosing the voxel with maximum relative power in a frequency band lead to meaningful results and is computationally efficient. In fact, in Brookes et al. (2004) the authors demonstrated that the time-series estimated at local maxima best describe the underlying source activity. As possible drawback of this method, it has been demonstrated that, in dataset with large artifact, these artifacts could reflect in power estimation and thus bias the voxel selection. Therefore, in this study, we carefully selected epochs without artifacts that could potentially affect the theta band. Nevertheless, other possible voxel selection techniques have been proposed, described in Hillebrand et al. (2012). In particular, the technique of averaging signals within a ROI before Fourier transformation, could lead to spurious connectivity results since it would generate a signal with disrupted phase characteristics; plus, interference and cancellations of theta band oscillations might occur, thus leading to a representative signal which would not reflect the activity of the ROI at all.

### 3.4. Basal-Pleasant phase transition. Data analysis.

The purpose of our analysis is to apply the task-dependent thresholding procedure for reconstructing a neural graph whose edges represent the connections, common to the two volunteers, which are “steadily different” during the phase

transition Basal-Pleasant. In what follows we give all the details concerning Subject A, while for Subject B we just provide the corresponding results.

3.4.1. Processing the data

The first step in our analysis consisted in importing the available 29x2 different upper-triangular 89x89-sized matrices  $A_t^{(s)}$  as Matlab® files. In this example we treat the phase transition between two phases: the Basal (the database  $\bar{s}$ ) and the Pleasant (the database  $\tilde{s}$ ) condition. As explained in Subsection 2.3.2, for each fixed  $t$  and  $s$ , the entry  $a_{hk} \in [0,1]$  of  $A_t^{(s)}$  denotes the functional connectivity between the nodes  $h$  and  $k$ , at the fixed time  $t$ , and in the fixed state  $s \in \{\bar{s}, \tilde{s}\}$ . Also, due to functional symmetry,  $a_{hk} = a_{kh}$  for all  $h, k \in \{1, \dots, 89\}$ , so that each matrix  $A_t^{(s)}$ ,  $1 \leq t \leq 29$ ,  $s \in \{\bar{s}, \tilde{s}\}$ , was available in its upper-triangular form, where  $a_{hk} = 0$  for all  $h \geq k$ . Zero entries on the main diagonal mean that the autocorrelation of nodes is excluded. Moreover, each entry was rounded at its fourth decimal digit.

The comparison between the Basal and the Pleasant test has been managed by computing the differences of the elements belonging to the corresponding databases.

For a clearer presentation, in the following we denote  $A_t^{(\bar{s})}$  by  $BA_t$ , and  $A_t^{(\tilde{s})}$  by  $PL_t$

3.4.2. Data analysis

For any fixed temporal index  $i \in \{1, \dots, 29\}$ , we have computed  $BA_i - PL_j$  for all  $j \in \{1, \dots, 29\}$ , so collecting a table as represented in Figure 2.

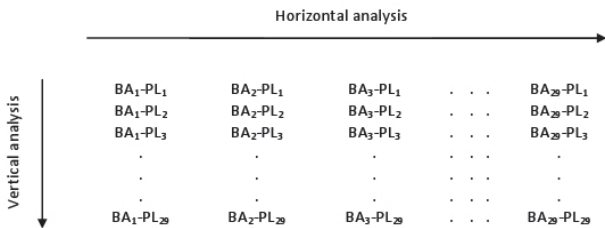


Fig. 2 – Vertical and horizontal analysis

This table can be explored both in the horizon-

tal and in the vertical directions. Each column resumes the comparison of a fixed Basal temporal slot with all the 29 Pleasant temporal slots. The converse holds for the rows.

According to our target, we want to characterize the correlations between pairs of nodes that are more sensitive to the transition from the Basal to the Pleasant test. This means that we need to look for entries having greater absolute values, because these denote the greater differences in correlation. Since the range of the entries is  $[-1, 1]$ , the NOI are a right neighbor of  $-1$ , and a left neighbor of  $+1$ , the width of these neighbors depending on the percentage  $\bar{p}$  of data to be considered.

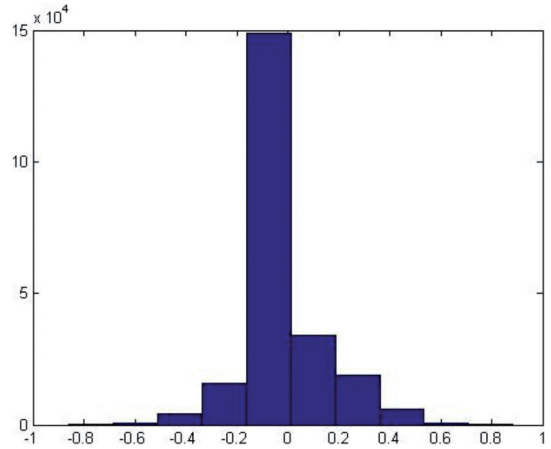


Fig. 3 – Histogram showing the distribution of the values of the entries of the set of matrices  $BA_1 - PL_j$ , with  $j = 1, \dots, 29$

To this purpose, we organized each column in a single 89x2581-sized matrix  $C_i = [c_{hi}]_i$   $i \in \{1, \dots, 29\}$ ,  $h \in \{1, \dots, 89\}$ ,  $l \in \{1, \dots, 2581\}$ , obtained by connecting the rows of all the 29 matrices  $BA_i - PL_j$  ( $j = 1, \dots, 29$ ) of size 89x89 forming the column. Then we computed the distribution of the entries in each matrix, which result in histograms organized in ten different bins, and representing the corresponding ordered data distributions  $D_i$ ,  $i \in \{1, \dots, 29\}$ . See for example Figure 3 for the case  $C_i = BA_1 - PL_j$ ,  $j = 1, \dots, 29$ , concerning the first epoch. By means of a Q-Q Plot we get an almost linear answer at each epoch, so that we can assume that the values of each  $D_i$  are symmetrically distributed. This allows us to constructed the NOI of  $-1$  and of  $+1$  with the same percentage of data, the lowest  $\bar{p}$  and the



highest  $\bar{p}$ , respectively. To this, we have considered a percentile subdivision, and assumed the reference percentile at  $k=95$ , so that, due to symmetry  $k'=5$ , and consequently  $\bar{p}$  consists of the 5% lowest (left tail) and of the 5% highest (right tail) entries of the distributions. We refer to this percentage of data as the *meaningful data* of our analysis.

### 3.4.3. First thresholding step

From the ordered distributions  $D_i$   $i \in \{1, \dots, 29\}$ , we have extracted the sub-distributions  $D_i^-$  and  $D_i^+$  consisting of the negative, and of the positive entries, respectively. We computed a Q-Q Plot against a normal distribution, which always returned a positive answer for all  $i \in \{1, \dots, 29\}$ . Further, also the distributions  $D_i$  proves to be quite good symmetrically approximated for all  $i \in \{1, \dots, 29\}$ , which suggested to work with reference quantiles  $k$ ,  $k'$  such that  $k'=q-k$ . In particular, we adopted a percentile subdivision, namely  $q=100$ , and assumed  $k$  equal to the 95-th percentile, so that  $k'$  corresponds to the 5-th percentile. As a consequence, we have constructed both the left tail of  $D_i^-$  (NOI of -1) and the right tail of  $D_i^+$  (NOI of +1), by using the same percentage  $\bar{p}=5\%$  of the extreme data (negative or positive), for all  $i \in \{1, \dots, 29\}$ . In what follows we give all details of the procedure concerning the right neighbor of -1. It is the neighbor where the Pleasant condition prevails on the Basal condition, namely, it represents the links that are overstimulated by the musical task. The left neighbor of +1 has been treated analogously, and just the corresponding results are shown and commented.

In order to extract the meaningful data, we have computed, at each epoch  $i \in \{1, \dots, 29\}$ , the value of the 5th percentile, representing the threshold  $\xi_i(L)$  determining the right neighborhood of -1. In Table 2 we report the complete list of the first thresholds. Then, for each  $i \in \{1, \dots, 29\}$ , we turned each  $C_i = BA_i - PL_j$  in a thresholded binary matrix  $\bar{C}_i = [C_i]_L = [\bar{c}_{hl}^i]$ , as explained in Sub-section 2.3

The corresponding 29 left thresholded 89x89-

sized binary matrices had been later summed up, to form a matrix  $S_i(L)$  that condensed the information in the right neighborhood of -1. Each entry of  $S_i(L)$  could assume a value ranging from 0 to 29, which denotes the number of slots of 2 seconds where the corresponding pair of nodes was correlated.

### 3.4.4. Second thresholding step

After forming the above thresholded matrices we looked for new thresholds, corresponding to the minimum number of slots of 2 seconds to be assumed as indicative of a permanent difference between the Basal and the Pleasant states. To this, we first computed the effective maximum  $M_i$  of each matrix  $S_i(L)$  (and the same for  $S_i(R)$ ), and later we have investigated the cases where  $\lambda=M_i$ ,  $\lambda=4/5M_i$  (corresponding to the 80% of the slots), and  $\lambda=1/2M_i$  (namely, the 50% of the slots). The results obtained in the case  $i=1$  are shown in the example below.

**Example.** Let us consider the 29 thresholded binary matrices  $\bar{C}_i = [C_i]_L = [BA_i - PL_j]_L$ , (the subscript  $t$  stands for “thresholded”, while the superscript specify that the analysis focused on the right interval of -1). Summing up the same entries in all these matrices we get the integer valued 89x89-sized matrix  $S_1(L)$ . We found that the maximum entry of such a matrix is  $M_1=29$  (note that, a priori, the maximum entry could be less than 29). This means that, when matching the first Basal matrix  $BA_1$  with all the 29 Pleasant matrices  $PL_j$ , some links exist which are always meaningful all over the 29 comparisons (see Figure 4). Therefore, by assuming  $\lambda=29$ , we set any entry  $\alpha$  of  $S_1(L)$  equal to 1 or 0, according as  $\alpha=29$  or  $\alpha < 29$ . The resulting binary matrix represent a graph  $G_{5\%,29}^1$  that we assume as representative of the right neighbor of -1 for what concerns the comparison of the first slot of the Basal test with all the slots of the Pleasant test. By analyzing Figure 4 we note that  $G_{5\%,29}^1$  consists of only one link (the dark-red colored pixel). It represents the connection between the nodes 17 and 21 corre-

sponding, respectively, to the left inferior frontal gyrus, opercular part and the left inferior frontal gyrus, triangular part (see the atlas in Table 1).

Table II. The values of 95-th percentile in the different 29 epochs for Subject A and Subject B.

Epoch	Subject A		Subject B	
	Neighbor -1	Neighbor 1	Neighbor -1	Neighbor 1
	Percentile	Percentile	Percentile	Percentile
1	-0.4540	0.4200	-0.3980	0.4370
2	-0.4180	0.3950	-0.4110	0.3760
3	-0.4360	0.3790	-0.4080	0.4000
4	-0.4350	0.4440	-0.4000	0.4220
5	-0.4310	0.4570	-0.4100	0.4340
6	-0.4480	0.4860	-0.4070	0.4060
7	-0.4730	0.3910	-0.4160	0.4240
8	-0.4470	0.4370	-0.4130	0.4130
9	-0.3990	0.4950	-0.4600	0.4020
10	-0.4470	0.4370	-0.4310	0.4300
11	-0.4420	0.4520	-0.4190	0.3840
12	-0.4800	0.4550	-0.4090	0.4240
13	-0.4460	0.4460	-0.4582	0.4390
14	-0.4190	0.4770	-0.4260	0.4810
15	-0.4570	0.3930	-0.4170	0.3960
16	-0.4500	0.4870	-0.4113	0.3960
17	-0.4530	0.4930	-0.4320	0.3850
18	-0.4530	0.4430	-0.4550	0.4302
19	-0.4630	0.4720	-0.4120	0.3710
20	-0.4460	0.4710	-0.4300	0.4430
21	-0.4700	0.4750	-0.4240	0.3890
22	-0.4370	0.4550	-0.3900	0.4270
23	-0.4570	0.4890	-0.4290	0.4224
24	-0.4450	0.4740	-0.4490	0.4360
25	-0.4690	0.4370	-0.4190	0.4140
26	-0.4550	0.4820	-0.4370	0.3920
27	-0.4430	0.4180	-0.4430	0.3960
28	-0.4530	0.5190	-0.4230	0.3950
29	-0.4430	0.4390	-0.4130	0.3860

By repeating the procedure for all  $i \in \{1, \dots, 29\}$ , we get the graphs  $G_{5\%,29}^i$ , and, from them, the graph  $G_{5\%,29}$  that summarize the analysis in the right neighbor of -1. More precisely, the graph  $G_{5\%,29}$  consists of all links which appear in every graph  $G_{5\%,29}^i$ , for all  $i \in \{1, \dots, 29\}$ .

By weakening  $\lambda$  to the (integer rounding of the) 80% or 50% of the maximum value (respectively 23 and 14 for  $M_i=29$ ), and by repeating the procedure, we get the graphs  $G_{5\%,23}$  and  $G_{5\%,14}$ . The three graphs  $G_{5\%,29}$ ,  $G_{5\%,23}$  and  $G_{5\%,14}$  together, summarize the analysis of the right neighbor of -1 for different values of the data-dependent thresholding parameters.

The same procedure has been carried out on the left neighbor of +1, and repeated on Subject B,

both in the neighbor of -1 and of +1.

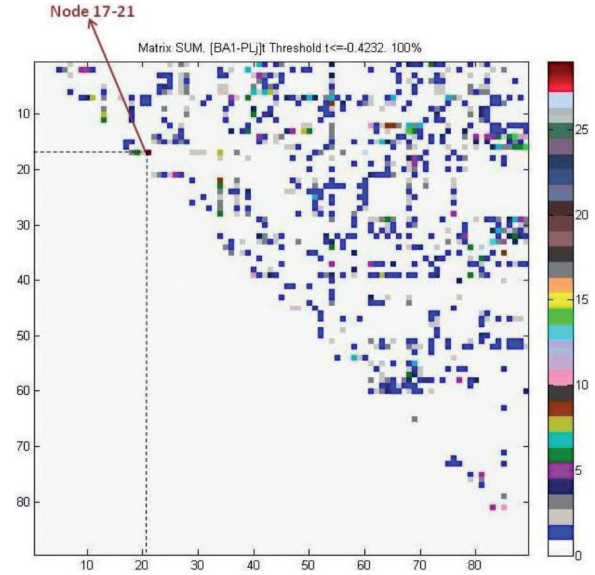


Fig. 4 - The example consider matrix  $S_1(L)$  without a thresholding process. We remind the reader that the colors of the pixels are relative to the presence of the associated link throughout the epochs. For example the dark-red pixel represents the link always active along the 29 epochs. When the domain of analysis is 100%, i.e. matrix not thresholded, there is a number of colors, each representing a value ranging from 0 to 29.

### 3.5. Interpretation of the results

The analysis of the two databases for Subject A and Subject B provides several interesting information. First, for any subject it is possible to understand which are the most recurring links throughout all the temporal slots (horizontal analysis). Second, the results of this analysis can be shown in terms of graphs, representing the involved cerebral network. By using different values of  $\lambda$  we pull out information about the links involving the “stronger” nodes, always present in the vertical analysis (see Figure 2), or links emerging after a relaxed thresholding process (cf. Subsection 3.4.3). We apply this procedure for both subjects, and for the two neighbors of -1 and of +1. In Table 3 and Table 4 we have reported the main results for Subject A, concerning the strongest thresholding choice  $\lambda=M_i$  in the NOI of -1, and the choice  $\lambda=4/5M_i$  in the NOI of +1. Analogously, Table 5 and Table 6 show the outputs for Subject B, where we have considered  $\lambda=1/2M_i$  for the NOI of -1, and  $\lambda=4/5M_i$  for the NOI of +1. The corre-

sponding graphs are shown in Figure 5, Figure 6, Figure 7, and Figure 8, where the thickness of a link connecting two nodes  $h$  and  $k$  relates to the following frequency-based weight formula

$$W_{hk} = \begin{cases} l \cdot 1, & \text{if } l \neq 0 \\ m \cdot 0.8, & \text{if } l = 0, m \neq 0 \\ n \cdot 0.5, & \text{if } l = m = 0, n \neq 0, \end{cases}$$

being  $l$ ,  $m$ ,  $n$  the number of temporal slots in which the link  $h-k$  appears, respectively, in the graphs  $G^{i_{5\%,29}}$ ,  $G^{i_{5\%,23}}$  and  $G^{i_{5\%,14}}$ ,  $i \in \{1, \dots, 29\}$ . The coefficients 1, 0.8, 0.5 relate to the corresponding percentage of  $M_i$  which determine the second threshold. This provides weights condensing information coming from the highest value of  $\lambda$  that allows the link to appear.

**Example.** Let us consider Subject B. In the analysis focusing on the neighborhood of +1, the link 17-25 does not appear in the graph  $G^{i_{5\%,29}}$  (corresponding to  $\lambda=M_i$ ) for any  $i \in \{1, \dots, 29\}$ , so  $l=0$ . Differently, there exist four indices  $i \in \{1, \dots, 29\}$  such that  $G^{i_{5\%,23}}$  (corresponding to  $80\%M_i$ ) contains the link, so  $m=4$ . Therefore, the link 17-25 gets weight  $W_{17-25}=4 \cdot (0.8)=3.2$ . In Figure 8 the link is shown, together with all the other edges of the emergent graph  $G_{5\%,23}$ , which allows an immediate comparison among their thickness, and consequently among their importance in the performed analysis.

### 3.5.1. Subject A: conclusions

Table 3 and Table 4 provide an overview of the results concerning the horizontal analysis over the 29 epochs.

The corresponding emerging graphs are shown in Figure 5 and in Figure 6, respectively.

Table III. NOI of -1 for Subject A.		
Subject A, NOI of -1, $\lambda=M_i$		
Link	Cerebral Areas	Occurrences
16-18	Right cuneus - Right inferior frontal gyrus, opercular part	14
13-52	Left posterior cingulum - Left globus pallidus	13
82-88	Right middle temporal gyrus - Right superior temporal gyrus	8

74-82	Right rolandic operculum - Right middle temporal gyrus	6
80-82	Right inferior temporal gyrus - Right middle temporal gyrus	5
73-88	Left rolandic operculum - Right superior temporal gyrus	4
37-88	Left Heschl's gyrus or Left transverse temporal gyrus - Right superior temporal gyrus	3
74-89	Right rolandic operculum - Left thalamus	2
13-37	Left posterior cingulum - Left Heschl's gyrus or Left transverse temporal gyrus	2
9-39	Left anterior cingulum - Right Heschl's gyrus or Right transverse temporal gyrus	1
51-76	Left olfactory cortex - Right supplementary motor area	1
60-65	Right inferior parietal lobule - Left precentral gyrus	1
73-74	Left rolandic operculum - Right rolandic operculum	1
76-88	Right supplementary motor area - Right superior temporal gyrus	1

Table IV: NOI of +1 for Subject A.		
Subject A, NOI of +1, $\lambda=4/5M_i$		
Link	Cerebral Areas	Occurrences
2-88	Right amygdala - Right superior temporal gyrus	17
85-88	Left superior temporal pole - Right superior temporal gyrus	14
1-88	Left amygdala - Right superior temporal gyrus	12
88-36	Right superior temporal gyrus- Right fusiform gyrus	12
88-83	Right superior temporal gyrus- Left middle temporal pole	4
88-62	Right superior temporal gyrus- Right Superior parietal lobule	3
88-3	Right superior temporal gyrus- Left Angular Gyrus	2
88-33	Right superior temporal gyrus- Right superior frontal gyrus, orbital part	2
88-50	Right superior temporal gyrus- Right superior occipital gyrus	2
1-13	Left amygdala- Left Posterior cingulum	1
1-17	Left amygdala- Left inferior frontal gyrus, opercular part	1
88-10	Right superior temporal gyrus- Right anterior cingulum	1
88-18	Right superior temporal gyrus- Right inferior frontal gyrus, opercular part	1
88-40	Right superior temporal gyrus- Right Hippocampus	1
88-49	Right superior temporal gyrus- Left superior occipital gyrus	1
88-57	Right superior temporal gyrus- Left paracentral lobule	1
88-61	Right superior temporal gyrus- Left Superior parietal lobule	1
88-81	Right superior temporal gyrus- Left middle temporal gyrus	1

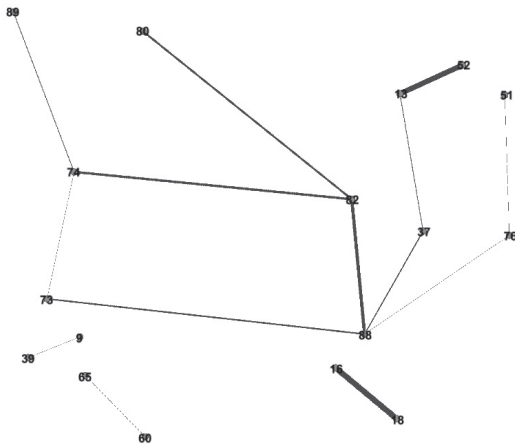


Fig. 5 – The graph  $G_{5\%,29}$  for Subject A when the Pleasant phase is dominant on the Basal phase (NOI of -1).

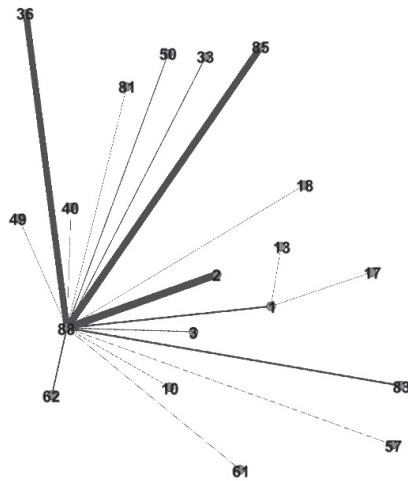


Fig. 6 – The graph  $G_{5\%,23}$  for Subject A when the Basal phase dominates the Pleasant phase (NOI of +1).

3.5.2. Subject B: conclusions

Table 5 and Table 6 provide an overview of the results concerning the horizontal analysis over the 29 epochs.

The corresponding emerging graphs are shown in Figure 7 and in Figure 8, respectively.

Table V: NOI of -1 for Subject B.		
Subject B, NOI of -1, $\lambda=1/2M_i$		
Link	Cerebral Areas	Occurrences
17-60	Left inferior frontal gyrus, opercular part - Right inferior parietal lobule	17
17-20	Left inferior frontal gyrus, opercular	16

	part - Inferior frontal gyrus, orbital part	
17-21	Left inferior frontal gyrus, opercular part - Left inferior frontal gyrus, triangular part	15
17-18	Left inferior frontal gyrus, opercular part - Right inferior frontal gyrus, opercular part	14
17-52	Left inferior frontal gyrus, opercular part - Right olfactory cortex	9
17-56	Left inferior frontal gyrus, opercular part - Right parahippocampal gyrus	6
17-37	Left inferior frontal gyrus, opercular part - Left Heschl's gyrus (or Left transverse temporal gyrus)	5
17-29	Left inferior frontal gyrus, opercular part - Left superior frontal gyrus	1
21-18	Left inferior frontal gyrus, triangular part - Right inferior frontal gyrus, opercular part	1
21-19	Left inferior frontal gyrus, triangular part - Left inferior frontal gyrus, orbital part	1
60-30	Right inferior parietal lobule - Left medial frontal gyrus	1

Table VI: NOI of +1 for Subject B.		
Subject B, NOI of +1, $\lambda=4/5M_i$		
Link	Cerebral Areas	Occurrences
17-22	Left inferior frontal gyrus, opercular part - Right inferior frontal gyrus, triangular part	14
17-26	Left inferior frontal gyrus, opercular part - Right inferior frontal gyrus, triangular part	7
17-25	Left inferior frontal gyrus, opercular part - Left middle frontal gyrus	4
17-23	Left inferior frontal gyrus, opercular part - Left medial orbitofrontal cortex	2
17-61	Left inferior frontal gyrus, opercular part - Left superior parietal lobule	2
17-65	Left inferior frontal gyrus, opercular part - Left precentral gyrus	2
26-47	Left middle frontal gyrus, orbital part - Left middle occipital gyrus	2
26-50	Left middle frontal gyrus, orbital part - Right superior occipital gyrus	2
26-75	Left middle frontal gyrus, orbital part - Left supplementary motor area	2
17-27	Left inferior frontal gyrus, opercular part - Right middle frontal gyrus, orbital part	1
22-88	Right inferior frontal gyrus, triangular part - Right superior temporal gyrus	1
26-36	Left middle frontal gyrus, orbital part - Right fusiform gyrus	1
26-43	Left middle frontal gyrus, orbital part - Left lingual gyrus	1
26-45	Left middle frontal gyrus, orbital part - Left inferior occipital gyrus	1
26-46	Left middle frontal gyrus, orbital	1



	part - Inferior occipital gyrus	
26-49	Left middle frontal gyrus, orbital part - Left superior occipital gyrus	1
26-78	Left middle frontal gyrus, orbital part - Right supramarginal gyrus	1
26-84	Left middle frontal gyrus, orbital part - Right middle temporal pole	1
26-86	Left middle frontal gyrus, orbital part - Right superior temporal pole	1
26-88	Left middle frontal gyrus, orbital part - Right superior temporal gyrus	1

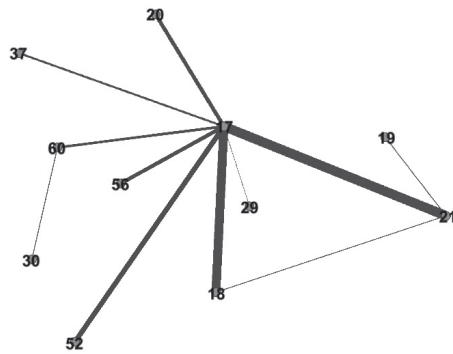


Fig. 7 – The graph  $G_{5\%,14}$  for Subject B when the Pleasant phase is dominant on the Basal phase (NOI of -1).

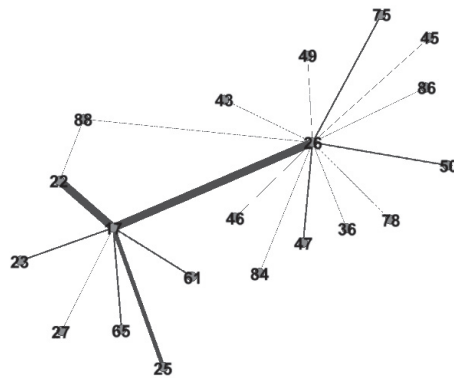


Fig. 8 – The graph  $G_{5\%,23}$  for Subject B when the Basal test dominates the Pleasant test (NOI of +1).

## Analysis and discussion of the results

We are aware that an analysis on two subjects cannot lead to definitive conclusions. Actually, our application has been confined to two sub-

jects, just to show how the whole theoretical procedure must be implemented (and also because no other data were available), without pretending to infer general neurological results. However, as we have previously remarked, our two steps thresholding approach has been recently tested on real fMRI data concerning a sample of 133 different right-handed subjects, of different sex and age, with the main purpose of investigating their functional connectivity changing over the lifespan during the resting state. Results have not been published yet, so we cannot give here deep details. However, we can provide a brief overview. Subjects have been grouped in 28 different age-dependent classes (corresponding to the different epochs of the present paper). To each group, a matrix has been associated, so obtaining a list of 28 different matrices, which are the analogous of the previous  $C_i$  matrices (we recall that, in the resting state, no task  $\tilde{s}$  is performed). Both the thresholding steps have been considered as outlined above. In particular, we have focused on the neighbor of +1, where the highest functional connectivity is registered. As a reference percentile we have assumed  $k=90$ , so considering  $\bar{p} = 10\%$  of the highest data, and  $\lambda=0.5M_i$ . The resulting representative graphs perfectly agree with the available literature about the Default Mode Network (DMN), and also provide some new interesting insights. For example we have found that, both in males and females, the precuneus plays a dominant role in DMN as speculated in [24], [57]. Moreover, our outcomes showed that cingulate cortex, paracingulate gyrus and frontal pole act as important players. In some seminal papers [9], [27], [58] all the cited areas have been identified as the core of the DMN in humans. This seems to represent a further solid validation of our proposal.

### 4.1. Discussion on two subjects' analysis

We wish to point out that, since our approach is data depending, the related results are also data depending, namely, they have a direct correlation with the subjects contributing to the databases, independently of their number. Of course, the higher the number of subjects, the

greater is the extendibility of the results to general neurobiological rules.

Keeping in mind this warning on generalization, a few speculations on the outputs obtained from our acoustical analysis on the two subjects are worth to be outlined.

In what follows, Subsection 4.1.1 and 4.1.2, we describe an analysis in terms of graph theoretical parameters (node degree, weight of the links, presence of cycles,...). To this purpose we focus on the most representative graphs for each subject, meaning that they are not necessarily characterized by the same value of second threshold. Further, in Subsection 4.1.3, we try to detail a possible neurobiological explanation of the outcomes. These show analogies, and also some differences between the two subjects. Analogies denote that indeed, a similar functional behavior has been determined by the musical task. On the other hand the differences could lie in the gender, the age and the emotional answers of the two volunteers.

We recall that the neighbor of -1 corresponds to the links mainly stimulated, with respect to the Basal phase (to be interpreted as the baseline), from the musical test. Conversely, the neighbor of +1 matches links that are “less sensitive” to the test.

#### 4.1.1. Graph Theoretical analysis in the neighbor of -1

The graph theoretical analysis shows the following results:

Subject A. Representative graph:  $G_{5\%,29}$ .

1. The average degree of the graph is  $d=1.529$ . Its floor and ceil approximations are  $\underline{d}=1$  and  $\bar{d}=2$ , respectively.
2. Node 88 has degree  $\text{deg}(88)=4$ , so higher than the average. This suggests that the node could act as a hub.
3. There is a cycle: 73-74-82-88-73. Note that the weights of the links making up the cycle range from 1 (the weight of the link 73-74) to 7 (the weight of the link 88-82) through 4 (the weight of 73-88) and 5 (the weight of 74-82).
4. The links with higher weight, which are called, from now on, the strongest links,

are: 16-18 (14 times out of 29); 13-52 (13); 82-88 (8). We remind the reader that the weight of a link represents its frequency throughout the 29 investigated epochs.

5. The graph consists of four connected components. Four are trivial, since are just single edges, while the fifth is richer.

Subject B. Representative graph:  $G_{5\%,14}$ .

1. The average degree of the graph is  $d=2$ , so that also  $\underline{d}=\bar{d}=2$ .
2. Node 17 could be a hub, since it has a high degree (8) in comparison to the average degree.
3. There is a cycle, whose vertices are nodes 17, 18, 21. Interestingly, links 17-18 and 17-21 have high weights, since they appear, respectively, 14 and 15 times.
4. The strongest links are: 17-60 (17 times), 17-21 (15), and 17-18 (14).
5. The graph is connected.

Main analogies.

1. Presence of possible hubs (nodes 17 and 88, respectively).
2. The hubs have precisely the same floor normalized degree, namely  $\text{deg}(v)/\underline{d}=4$  for both  $v=17$  and  $v=88$ .
3.  $\bar{d}=2$ .
4. Presence of cycles.

Main differences.

1. The number of involved cerebral areas.
2. The number of links contributing to the cycle.
3. The connection of the graphs.

#### 4.1.2. Graph Theoretical analysis in the neighbor of +1

Let us look at the analysis in the neighbor of +1, i.e. when the Basal test dominates on the Pleasant one.

Subject A. Representative graph:  $G_{5\%,23}$ .

1. The average degree of the graph is

$d=1.895$ , so that  $\underline{d}=1$  and  $\bar{d}=2$ , respectively.

2. The temporal lobe, specifically the right superior temporal gyrus (node 88) is strongly suspected to be a hub, being  $\text{deg}(88)=16$ , a very high value if compared to 1.895.
3. There are no cycles.
4. The strongest links are: 88-2 (17 out 23); 88-85 (14); 88-1 (12); 88-36 (12).
5. The graph is connected.

Subject B. Representative graph:  $G_{5\%,23}$ .

1. The average degree is  $d=2$ , so that  $\underline{d}=\bar{d}=2$ .
2. Node 17 could be a hub, since it has a degree equal to 7. Node 26 is strongly believed to be a hub, its degree is 13.
3. Nodes 17 and 26 are connected each other. Hence, due to their central role we can speculate that they form a rich-club organization<sup>3</sup>. In addition, they are a sort of connector hubs.
4. There is a cycle: 17-22-88-26-17. In such a cycle the links 17-22, 17-26 have a high weight.
5. The strongest links are: 17-22; 17-25; 17-26. 6. The graph is connected.

Main analogies.

1. Presence of strong hubs.
2. Connection of the graphs.
3. Significant correspondence of the involved cerebral areas (see Subsection 4.1.3).
4.  $\bar{d}=2$

Main differences.

1. The number of hubs.
2. The floor normalized degree of the hubs.
3. The presence of rich club organization.

We remind that Subject A is a male of 39, while Subject B is a female whose age is 27. Between males and females, there are differences, both in the structural and in the functional connectivity. In addition, the human primary cortex is located in the region of transverse gyrus of Heschl (HG) in the superior temporal plane. It is interesting to remark that it is possible to have more than one HG on one or both sides in approximately 20% of hemispheres, and that the proportion of HG that is occupied by cytoarchitecturally defined primary cortex can vary between 20% and 80% (Frackowiak et al., 2004). These remarks help in understanding why the graphs for Subject A and Subject B are not identical.

#### 4.1.3. Possible neurobiological explanations of the outcomes

Before giving a few possible neurobiological explanations concerning the obtained output graphs, we guess it is useful to recall the main steps of the acoustical process.

When a subject listens to music, or hears someone speaking, the brain must process the incoming sounds, which may depend on, or be affected by, other information, such as the tone, the frequencies, the presence of noise. The first activation is inside the ear, where sounds are converted to vibrations in the middle ear, and then to electrical impulses in the inner ear. The second step occurs at a neurobiological level, when the previous electrical impulses are related to different cerebral areas for their interpretation. Basically, we can select four major players in the acoustical process: the Acoustic Nerve, the Cochlear Nucleus, the Auditory Cortex and, at a higher stage, the Prefrontal Cortex. As expected, we found nodes strongly involved in the processing of acoustical information, in particular 37 (left Heschl's gyrus), 85 (left superior temporal pole) and 88 (right superior temporal gyrus). These nodes are, in some way, related to, or even included in the Primary Auditory Cortex. We remind the reader that the Primary Auditory Cortex is part of the Temporal Lobe. Anatomically, such a cortex is located in the Superior Temporal plane, within the Lateral Fissure, and comprising parts of

<sup>3</sup> The rich club organization (RCO) is characterized by a tendency for high-degree nodes to be more densely connected among themselves than nodes of a lower degree. This metrics provides important information, for example a high rich-club coefficient implies that the hubs are well connected, and global connectivity is resilient to any one hub being removed.

Heschl's Gyrus and the Superior Temporal Gyrus, including planum polare and planum temporale (roughly Brodmann areas 41, 42, and partially 22). Functionally, the Auditory cortex processes auditory information (both in humans and in animals).

Nevertheless, in our outputs are also represented cerebral areas that are not directly related to the auditory process, for example, nodes 49, 50, 52 and 60.

We wish to detail the possible neurological explanations, both in the neighbor of -1 and of +1.

#### *Neighbor of -1.*

By analyzing the outcomes shown in Table 3 and Table 5, it is not surprising to find cerebral areas such as the left and the right inferior frontal gyrus, opercular part (respectively, nodes 17 and 18), the left and the right inferior frontal gyrus, triangular part (nodes 21 and 22), the left superior temporal lobe (node 85) and the right superior temporal gyrus (nodes and 88). All of them, in some way, are related to the processing of auditory information.

If we focus on the representative graphs, for Subject A and Subject B, in the neighbor of -1, we can see that both of them contain the following nodes: 18, 52 and 60. It is not a surprise to find node 18, which is directly involved in the acoustical process. Differently, the presence of nodes 52 and 60 is really interesting. Actually, node 52, the right olfactory cortex<sup>4</sup>, is related to the olfactory process, while node 60, the inferior parietal lobule, is generally involved in the perception of emotions in facial stimuli. The occurrence of these nodes in the representative graphs points out that the acoustical task induced, in both subjects, a functional activation of the same sensorial areas, probably related to the emotional answers to the listened piece of music.

Regarding the strongest links we note that, for Subject A, they are: 16-18, 13-52, 82-88, the last being directly involved in the acoustical process.

For Subject B the strongest links concern the

frontal gyrus, 17-21, 17-60, and, in particular 17-21 is supposed to be involved in the process of acoustical information at a higher stage.

In conclusion, we can infer that, in the neighbor of -1, there is a statistical interaction between areas which are associated to the acoustical process with areas associated to sensorial stimuli.

#### *Neighbor of +1.*

If we consider Table 4 and Table 6, it is not surprising to find nodes that are not directly related to auditory functions, since, in this case, we are investigating cerebral regions that are under-stimulated with respect to the Basal condition. These represent the left and the right amygdala (nodes 1 and 2), the left posterior cingulum (node 13), the left cuneus (node 16), the left globus pallidus (node 52), the right inferior parietal lobule (node 60). All of them seem to be negatively involved in the change of the functional connectivity during the music task. A possible interesting explanation is that these areas are supposed to be functionally linked to emotions, memory, meditation (e.g. posterior cingulum), and consequently are not immediately stimulated in the elaboration of sounds. For example, the cuneus is involved in the basic visual processing, which are modulated by extra-retinal effects, like attention, working memory and reward expectation (see [4]). Also, it is known that the right amygdala plays a role in non-conscious processing of emotion, while the left amygdala is involved in the processing of conscious emotion [34]. Finally, the inferior parietal lobule is involved in the perception of emotion, in facial stimuli, and in the interpretation of sensory information (Radua et al., 2010).

If we focus on the nodes shared by the representative graph of Subject A and Subject B we can find: 17, 49, 50, 61, 88.

Interestingly, nodes 49 (left superior occipital gyrus), 50 (right superior occipital gyrus) and 61 (left superior parietal lobule), are, respectively, important for the visual process (49 and 50) and for the spatial orientation (61). In addition the (left) superior parietal lobule receives a great deal of visual input and it is involved with

<sup>4</sup>The olfactory cortex which includes the piriform cortex (posterior orbitofrontal cortex), amygdala, olfactory tubercle, and parahippocampal gyrus.



other functions of the parietal lobe in general, such as two points spatial discrimination and goal-related activity. From an anatomical point of view, the superior parietal lobule plays a big role since it has major white matter pathway connections with several important areas such as the Cingulum, SLF I<sup>5</sup>.

The presence of node 17 and of node 88, that, as detailed above, are indeed related to the acoustical process, can be explained by the structure of the output graphs. In fact, these nodes are connected with several other cerebral areas that appear as meaningful in the neighbor of +1, so that the occurrence of 17 and 88, rather than to acoustical reasons, should be related to their role of hub in the involved brain network.

In conclusion, in the neighbor of 1, the analysis shows that the network involved in the acoustical process is dominated by other networks which involve areas unrelated to such a process. The functional connectivity appears to be relaxed in cerebral areas which are not traditionally associated with acoustical processes, which agrees, despite of the small number of subjects undergoing the experiment, with what is known by the available literature (see for instance Wilkins et al. (2014) and Wu et al. (2013))

## Conclusions

In this paper we proposed a general methodology for introducing thresholds in the analysis of neurobiological databases. Such a methodology could be applied to data of different nature such as M/EEG, fMRI and PET.

We have proposed to favor a task-dependent approach to threshold selection, based on a statistic analysis of the distribution of the collected data. This is done by gathering together data coming from different tests in form of triangular matrices,

and then by comparing all the possible pairs of matrices. One of the tests is assumed as a baseline, with respect to which the second test is considered as one aims in looking for analogies, or for discrepancies, between the two conditions. In each case, a data-dependent thresholding procedure is carried out by a two steps method, based on parameters  $\bar{p}$  and  $\lambda$ . The first one allows to select the width of the NOI. The second threshold provides a fraction of the data coming from the previously selected NOI, in order to extract the stronger meaningful links. As an output a graph is provided, that represents the neurological results of the performed test with respect to the baseline condition. The greater  $\lambda$ , the smaller is the number of extracted links, so that their importance in the neurobiological interpretation increases. We also suggest to investigate different values of  $\lambda$ , in order to show, for a same threshold  $\bar{p}$ , representative graphs having increasing structural richness.

We have applied the proposed methodology to real EEG data generated in acoustical experiments, where we compared the changes in functional connectivity from a Basal condition (which is essentially a resting state condition), to a Pleasant condition, where the volunteers listened to music that they classified as pleasant.

In our analysis we have considered three different representative values of  $\lambda$ , corresponding to cut off a link when its occurrence is below the 50%, 80% and 100% of the maximum collected frequency. We wish also to point out that this approach works independently of the values of  $M_i$ , so that we can provide meaningful representative graphs even when  $M_i$  is small. Differently, in a database independent thresholding procedure, low values of  $M_i$  could be discarded, which would prevent, since the beginning, the corresponding neurobiological analysis.

We looked for links steadily different, namely, links whose values are always different throughout the experiment duration. Therefore, we focused on the NOI concerning neighbors of -1 and of +1, finding out some natural outcomes, as well as some interesting unexpected results.

A further acoustical analysis could be developed in a neighbor of 0, in order to investigate the links always equal throughout the temporal slots.

<sup>5</sup> The superior longitudinal fasciculus (SLF) is composed of four distinct components: SLF I, SLF II, SLF III, and arcuate fascicle (AF). In humans, these four components are bundled together although they are functionally separate. Notably, in non-human primates, the SLF and AF are anatomically separate and have separate trajectories.

By repeating the task with a significant number of subjects it could be possible to reinforce and generalize the results provided by our analysis. Moreover, by varying the choices of the threshold parameters  $\bar{p}$  and  $\lambda$ , one could also investigate the existence of motifs (Sporns et al., 2004) or leading a clustering analysis to understand the existence of provincial, central hubs and modules (Crossley et al., 2014; Sporns, 2010; Sporns et al., 2007; Van den Heuvel and Sporns, 2013). Such a methodology could be useful to highlight cerebral connections that generally are not expected, also in view of identifying possible regions responsible for psychiatric disorders (Bassett et al., 2008; Bassett and Bullmore, 2009; Fornito et al., 2012). In addition, our methodology can be applied both in case of a resting state, as well as when a task-dependent experiment is considered.

We emphasize that all the data processed, and the results obtained in the different slots are available. To this, we invite everybody who might be interested to write to the corresponding author.

## References

- Acar Z.A. and Makeig S. Effects of Forward Model Errors on EEG Source Localization. *Brain Topogr.*, **26**: 378–396, 2013.
- Achard S. and Bullmore E.T. Efficiency and Cost of Economical Brain Functional Networks. *PLoS Comput. Biol.*, **3**: 174–18, 2007.
- Acheson D.J. and Hagoort P. Stimulating the brain's language network: syntactic ambiguity resolution after TMS to the inferior frontal gyrus and middle temporal gyrus. *J. Cogn. Neurosci.*, **25**: 1664–77, 2013.
- Bassett D.S., Bullmore E., Verchinski B.A., Mattay V.S., Weinberger D.R., Meyer-Lindenberg A. Hierarchical Organization of Human Cortical Networks in Health and Schizophrenia. *J. Neurosci.*, **28**: 9239–9248, 2008.
- Bassett D.S. and Bullmore E.T. Human brain networks in health and disease. *Curr. Opin. Neurol.*, **22**: 340–347, 2009.
- Bathelt J., O'Reilly, H., Clayden, J.D., Cross, J.H., deHaan M. Functional brain network organisation of children between 2 and 5 years derived from reconstructed activity of cortical sources of high-density EEG recordings. *Neuroimage*, **82**: 595–604, 2013.
- Brookes M.J., Gibson A.M., Hall S.D., Furlong P.L., Barnes G.R., Hillebrand A., Singh K.D., Holli-day I.E., Francis S.T., Morris P.G. A general linear model for MEG beamformer imaging. *Neuroimage*, **23**: 936–46, 2004.
- Buckner R.L., Andrews-Hanna J.R., Schacter D.L. The brain's default network. *Ann. NY Acad. Sci.*, **1124**: 138. 2008.
- Bullmore E.T., Bassett D.S. Brain graphs: graphical models of the human brain connectome. *Annu. Rev. Clin. Psychol.*, **7**: 113–140, 2011.
- Bollobás B. *Random graphs*. London: Academic, 1985.
- Bullmore E.T. and Sporn O. Complex Brain Networks Graph Theoretical Analysis Of Structural and Functional Systems. *Nat. Rev. Neurosci.*, **10**: 186–198, 2009.
- Colman A.M. *A Dictionary of Psychology*. Oxford University Press, 2014.
- Crossley N. A., Mechelli A., Scott J., Carletti F., Fox P. T., McGuire P., Bullmore E. T. The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain*, **137**: 2382–95, 2014.
- Dipasquale O., Finotelli P., Costantini I., Baselli G., Baglio F., Dulio P., Cercignani M. Exploring Resting-State Functional Connectivity Invariants across the Life Span using a Novel Graph Model, ISMRM 24th Annual Meeting & Exhibition, Singapore, 2016.
- Friederici A.D., Ruschemeyer S., Hahne A., Fiebach C.J. The Role of Left Inferior Frontal and Superior Temporal Cortex in Sentence Comprehension: Localizing Syntactic and Semantic Processes. *Cereb. Cortex*, **13**: 170–177, 2003.
- Finotelli, P., Dulio, P. Graph Theoretical Analysis of Brain. An Overview. *Scienze e Ricerche*, **9**: 89–96, 2015.
- Finotelli P., Dulio P. A Mathematical Model for Evaluating the Functional Connectivity Strongness in Healthy People. *Arch. Ital. Biol.*, **153**: 279–300, 2015.
- Flores-Gutiérrez E.O., Diaz J.L., Barrios F.A., Favila-Humara R., Guevara M.A., del Rio-Portilla Y., Corsi-Cabrera M. Metabolic and electric brain patterns during pleasant and unpleasant emotions induced by music masterpieces. *Int. J. Psychophysiol.*, **65**: 69–84, 2007
- Fornito A., Zalesky A., Pantelis C., Bullmore E.T. Schizophrenia, neuroimaging and connectomics. *Neuroimage*, **62**: 2296–2314, 2012.
- Fornito A., Zalesky A., Breakspear M. The connectomics of brain disorders. *Nature Reviews Neuroscience*, **16**: 159–172, 2015.
- Frackowiak R.S.J., Friston K.J., Frith C.D., Dolan R.J., Price C.J., Zeki S., Ashburner J.T., Penny, W.D. *Human Brain Function*. 2<sup>nd</sup> edition, San Diego, Academic Press, 2004.
- Friston K.J., Frith C.D., Liddle P.F., Frackowiak R.S.J. Functional connectivity: the principal component

- analysis of large (PET) data sets. *J. Cereb. Blood Flow Metab.*, **13**: 5-14, 1993.
- Fransson P., Marrelec G. The precuneus posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *NeuroImage*, **42**: 1178-1184, 2008.
- Friston K.J. Functional and Effective Connectivity in Neuroimaging: A Synthesis. *Hum. Brain Mapp.*, **2**: 56-78, 1994.
- Friston K.J., Harrison L., Penny W. Dynamic causal modelling. *Neuroimage*, **19**: 1273-1302, 2003.
- Greicius M.D., Krasnow B., Reiss A.L., Menon V. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.*, **100**: 253-258, 2003.
- Gong G., Rosa-Neto P., Carbonell F., Chen Z.J., He Y., and Evans A.C. Age- and gender-related differences in the cortical anatomical network. *J. Neurosci.*, **29**: 15684-15693, 2009.
- Greicius M.D., Supekar K., Menon V., Dougherty R.F. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb. Cortex*, **19**: 72-78, 2009.
- Hillebrand A., Barnes G.R., Bosboom J.L., Berendse H.W., Stam C.J. Frequency-dependent functional connectivity within resting-state networks: An atlas-based MEG beamformer solution. *Neuroimage*, **59**: 3909-3921, 2012.
- Kaiser M. A tutorial in connectome analysis: topological and spatial features of brain networks. *Neuroimage*, **57**: 892-907, 2011.
- Khazaei A., Ebrahimzadeh A., Babajani-Feremi A. Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory. *Clin. Neurophysiol.*, **126**: 2134-2141, 2015.
- Koelsch S., Fritz T., von Cramon, D. Y., Muller, K., & Friederici, A.D. Investigating emotion with music: An fMRI study. *Hum. Brain Mapp.*, **27**: 239-250, 2006.
- Lanteau L., Khalifa S., Rgis J., Marquis P., Chauvel P., Bartolomei F. Emotion induction after direct intracerebral stimulations of human amygdala. *Cereb. Cortex*, **17**: 1307-13, 2007.
- Leistedt S.J.J., Coumans N., Dumont M., Lanquart J.P., Stam C.J., Linkowski, P. Altered sleep brain functional connectivity in acutely depressed patients. *Hum. Brain Mapp.*, **30**: 2207-19, 2009
- Lachaux J.P., Rodriguez E., Martinerie J., Varela F.J. Measuring phase synchrony in brain signals. *Hum. Brain Mapp.*, **8**: 194-208, 1999.
- Lewicki, M.S. A review of methods for spike sorting: the detection and classification of neural action potentials, Network. *Comput. Neural Syst.*, **9** R53-R78, 1998.
- Lynall M.E., Bassett D.S., Kerwin R., McKenna P.J., Kitzbichler M., Muller U. and Bullmore E.T. Functional Connectivity and Brain Networks in Schizophrenia. *J. Neurosci.*, **30**: 9477-9487, 2010.
- Mechelli A., Price C.J., Noppeney U., Friston K.J. A dynamic causal modeling study on category effects: bottom-up or top-down mediation? *J. Cogn. Neurosci.*, **15**: 925-934, 2003.
- Meunier, D., Lambiotte, R., Fornito, A., Ersche, K.D. and Bullmore E.T. Hierarchical Modularity in Human Brain Functional Networks. *Front. Neuroinformatics*, **3**: 37, 2009.
- Michel C.M., Murray M.M., Lantz G., Gonzalez S., Spinelli L., Grave de Peralta R. EEG source imaging, *Clin Neurophysiol.*, **115**: 2195-2222, 2004.
- Oostenveld R. and Praamstra P. The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.*, **112**: 713-9, 2001.
- Oostenveld R., Fries P., Maris E., Schoffelen J.M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput. Intel. Neurosci.*, **2011**: 1-9, 2011.
- Radua J., Phillips M.L., Russell T., Lawrence N., Marshall N.; Kalidindi S.; El-Hage W., McDonald C., Gianpietro G., Brammer M.J., David A.S., Surguladze A.S. Neural response to specific components of fearful faces in healthy and schizophrenic adults. *NeuroImage*, **49**: 939-946, 2010.
- Raichle, M.E. The brain's default mode network. *Annu. Rev. Neurosci.*, **38**: 433-47, 2015.
- Rubinov M., Sporns O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, **52**: 1059-69, 2010.
- Rutter L., Nadar S.R., Holroyd T., Carver F.W., Apud J., Weinberger D.R., Coppola R. Graph theoretical analysis of resting magnetoencephalographic functional connectivity networks. *Front. Comput. Neurosci.*, **7**: 93, 2013.
- Salvador R., Suckling J., Schwarzbauer C., Bullmore E.T. Undirected graphs of frequency-dependent functional connectivity in whole brain networks. *Philos Trans R Soc Lond B Biol Sci.*, **360**: 937-46, 2005.
- Sammler D., Grigutsch M., Koelsch S. Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*, **44**: 293-304, 2007.
- Schmitt J.E., Lenroot R.K., Wallace G.L., Ordaz S., Taylor K.N., Kabani N., Greenstein D., Lerch J.P., Kendler K.S., Neale M.C. and Giedd J.N. Identification of Genetically Mediated Cortical Networks: A Multivariate Study of Pediatric Twins and Siblings. *Cereb. Cortex*, **18**: 1737-1747, 2008.

- Sporns O. *Networks of the Brain*. MIT Press, 2010.
- Sporns O., Honey C. J., and Kötter R. Identification and classification of hubs in brain networks. *PLoS ONE*, **2**, 2007.
- Sporns O., Kötter R. Motifs in Brain Networks. *PLoS Biology*, **2**: e369.
- Stam C.J., de Haan W., Daffertshofer A., Jones B.F., Manshanden I., van Cappellen van Walsum, A.M., Montez T., Verbunt J.P., de Munck J.C., van Dijk B.W., Berendse H.W., Scheltens P. Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. *Brain*, **132**: 213-24, 2009.
- Tzourio-Mazoyer N., Landeau B., Papathanassiou D., Crivello F., Etard O., Delcroix N., Mazoyer B. and Joliot M. Automated Anatomical Labeling of activations in SPM using a Macroscopic Anatomical Parcellation of the MNI MRI single-subject brain. *NeuroImage*, **15**: 273-289, 2002.
- Van Veen, Barry D., et al.. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.*, **44**: 867-880, 1997.
- Utevsky A.V., Smith D.V. and Huettel S.A. Precuneus Is a Functional Core of the Default-Mode Network. *The Journal of Neuroscience*, **34**: 932-940, 2014..
- Waiten W. *Psychology: Themes and variation*. Wadsworth Publishing; 7th edition, 2007.
- Van den Heuvel M.P., Mandl R.C.W., Kahn R.S., Hulshoff Pol E.H. Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Hum. Brain Mapp.*, **30**: 3127-3141, 2009.
- Van den Heuvel M., Stam C., Kahn, R.S., Hulshoff-Pol, H.E.. Efficiency of Functional Brain Networks and Intellectual Performance. *J. Neurosci.*, **29**: 7619-7624, 2009.
- Van den Heuvel and Olaf Sporns Network hubs in the human brain. *Trends Cogn. Sci.*, **17**: 683-96, 2013.
- Wilkins R. W., Hodges D. A., Laurienti P. J., Steen M. and Burdette J. H. Network science and the effects of music preference on functional brain connectivity: from Beethoven to Eminem. *Sci. Rep.*, **4**: 6130, 2014.
- Wu J., Zhang J., Ding X., Li R., Zhou C. The effects of music on brain functional networks: A network analysis. *Neuroscience*, **250**: 49-59, 2013.