Visualization and exploration of multichannel EEG coherence networks

Ji, Chengtao

Publication date:
2018

4 COMPARISON OF BRAIN CONNECTIVITY NETWORKS USING LOCAL STRUCTURE ANALYSIS

ABSTRACT

Object: In this study, we introduce a quantitative method for comparing brain connectivity, which accounts for the connectivity, spatial character and local structure simultaneously. *Methods:* In our approach, we first detect the local structure, in terms of functional units, of EEG coherence networks. The EEG coherence networks are then compared by employing the earth mover's distance (EMD) which calculates the cost of transforming the distribution of functional units in one network into another one. In this work, first a toy example is provided to assess its performance with an existing method, and then an application example with real EEG data is provided to evaluate the variability of brain connectivity for older and younger participants. *Results:* The proposed method has a better performance compared with the existing one-to-one matching method. The variability between EEG coherence networks obtained from different participants is assessed, and the results show that there is a higher variability for older participants compared to the younger participants. *Conclusion:* We proposed a simple quantitative method for comparing brain connectivity data. *Significance:* The proposed method has the potential to be used in comparisons of brain connectivity which are not limited to EEG coherence networks.

4.1 INTRODUCTION

Nowadays, many neuroimaging techniques are able to measure brain activity, such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET), electroencephalography (EEG), or magnetoencephalography (MEG). Brain connectivity datasets derived from these neuroimaging techniques are usually represented as networks in which nodes represent brain regions and links represent anatomical tracts or functional associations [107]. Of particular interest is the *comparison* of brain networks, which is useful for exploring connectivity relationships in individual subjects, or between groups of subjects under different conditions or with different characteristics such as patients and controls. Quantitative comparisons of brain networks between participant populations allow researchers to study how much participants differ from others. This comparison can be further used to reveal presumed connectivity abnormalities in neurological and psychiatric disorders.

Networks, often referred to as *graphs* in the mathematics literature, are widely used to model complex objects and their relations. In this study, we use the terms *network* and *graph* interchangeably. In graph theory, researchers have developed methods to characterize graph properties [107]. These methods are either based on local vertex and edge properties, such as degree and strength, or based on global features of the graph, such as density and modularity [15, 23]. However, such methods are not sufficient to compare brain networks. For example, two graphs with connections between different brain regions can have identical topological properties. In the context of brain networks, information about the absolute and relative spatial location of the vertices (here: electrodes) is a crucial factor for the analysis of brain connectivity. Another way of comparing graphs is based on graph matching, which is a method to find a correspondence between the nodes of different graphs (graph isomorphism) based on the attributes of nodes and edges. Generally, this problem can be cast as a quadratic assignment problem. However, such graph matching approaches are very challenging since they are NP-hard. Thus, approximate methods are used to find suboptimal solutions to compare corresponding networks, for example, using the features that describe or summarize the original networks.

In this chapter, we introduce an approach for comparing brain functional networks using their local structure, and we focus on the analysis of EEG coherence networks. An EEG coherence network is modelled as a collection of vertices, representing electrodes, and a collection of links, representing coherences between pairs of signals recorded by the corresponding electrodes. Among the many available brain mapping techniques, the EEG method provides high-resolution temporal information about brain activity [82].

EEG records the electrical potential of the brain from electrodes attached to the scalp of a subject at multiple positions. Synchronous electrical activity in brain regions is generally assumed to imply functional integration. Many methods have been proposed to measure the synchrony between pairs of brain regions, and these measures are often closely correlated [96]. EEG coherence is one of these measures, which is calculated between pairs of electrode signals as a function of frequency [53, 86].

In this chapter, we propose an extension of the earth mover's distance (EMD) to compare EEG coherence networks quantitatively by the distribution of data-driven regions of interest (ROIs), called *functional unit* (FU) maps. An FU of an EEG coherence network is a collection of nodes which are spatially connected and where each pair of nodes is significantly connected because of volume conduction effects; see [20, 74]. The FUs of the network are displayed in a so-called *FU map*, in which the distribution of FUs can be easily observed; see Section 4.3.1.1 for precise definitions. Therefore, the spatial structure and connectivity features of the network are well represented by the FU-map distribution. Given two EEG coherence networks, the EMD concept is used to assess their (dis)similarity based on their FU representations. The dissimilarity is calculated as the total cost that is needed to turn the FU representation of one network into that of the other one [108]. The EMD method to measure the network dissimilarity also considers the cross-FU information instead of using a simple one-to-one mapping among the FUs of two coherence networks.

The performance of the proposed method is compared with an existing method using synthetic coherence networks. A case study is then presented where the method is applied to brain networks obtained from an oddball experiment for analyzing inter-subject variability.

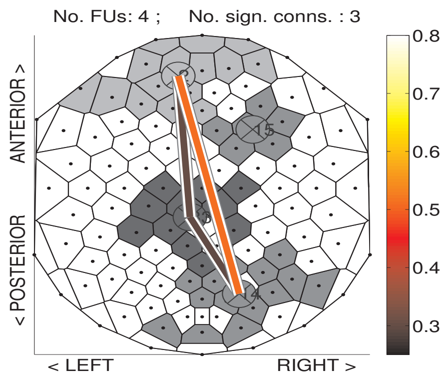


Figure 4.1: *Example of an FU map [18] as obtained during an oddball task. Spatial groups of similarly colored (in gray scale) cells correspond to FUs with a size of at least four, while white cells are part of smaller FUs. Circles overlaid on the cells represent the barycenters of the FUs and are connected by lines whose color reflects the average coherence between all electrodes of the respective FUs (see color bar).*

4.2 RELATED WORK

A large number of measures has been introduced that can be used for network comparison [107]. An individual network measure may characterize one or several aspects of local and global network connectivity, and these methods can therefore be divided into two categories [87]: one is related to local properties, such as vertex degree, strength and centrality; the other one is related to the global features of the graph, such as density and modularity. However, such methods are not suitable to compare brain networks since they ignore the spatial information; for example, two graphs with connections between different brain regions can have identical graph properties.

Another way of comparing graphs is by *graph matching*, where the problem is to find a correspondence between nodes of different graphs

[21]. From the literature, two main categories can be identified: *exact* and *inexact* graph matching [79]. Exact matching aims for a graph isomorphism, which is characterized by the fact that the mapping between the nodes of two graphs must be edge-preserving and, in the case of labelled graphs (which have labels associated with each node or edge), the labels of the vertices and edges should also be preserved. However, the majority of graphs in real-world applications are not isomorphic. For this case, one has to resort to inexact or error-tolerant graph matching. Among this, the most widely used methods are based on the graph edit distance (GED), which counts the costs that are involved in transforming one graph to another one [45]. Many formulations of the graph matching problem are cast as an assignment problem. However, graph matching is an NP-complete problem which has no known polynomial-time solution. Therefore, some approximate methods with polynomial time requirements are often accepted to find suboptimal solutions.

Inspired by image retrieval methods, feature-based graph similarity models have been proposed to compare graphs. They first derive structure information described by features from the graphs, then measure the similarity of graphs based on these features. In [85], the authors transform the structure information into a triple of features: leadership, bonding, and diversity. Leadership is used to measure how the edge connectivity of a graph is dominated by a single vertex; bonding is used to measure triadic closure in a graph (meaning that if among three nodes a, b, c, the nodes a and b as well as the nodes a and c are connected, then also b and c must be connected); and diversity is a measure based on the number of edges that share no common end points, and hence are disjoint. In terms of connectivity information along paths in graphs, Wichterich *et al.* proposed a method where the feature representation encodes which kinds of vertices are connected within a graph and how frequently this coupling occurs [121]. Then, EMD is applied to calculate the dissimilarity between graphs based on their derived features. However, no spatial information is taken into account in these methods for comparing networks.

4.3 METHODS

To compare EEG coherence networks quantitatively, the data model and corresponding FU representation are presented in the Section 4.3.1. The dissimilarity of two coherence networks based on FU representations and EMD is described in the Section 4.3.2. Finally, the performance of the method is studied in the Section 4.3.3.

4.3.1 *Data Model*

A network or graph is denoted by *G=(V,E)* with a node set $V = \{v_1, \ldots, v_n\}$ and an edge set $E \in V \times V$, where $n = |V|$ is the number of nodes. For the case of an EEG coherence network, the node or vertex set is equal to the set of electrode positions (see Figure 4.2), and the edges represent coherence between any pair of signals recorded by corresponding electrodes. Generally, the number and position of the electrodes on the scalp is fixed for all subjects.

For a 2D visualization of the nodes, planar projections are used for the 3D electrode locations on the surface of the head for preserving the spatial structure, and nodes are usually mapped to a top view of the head. To determine the spatial relationships between electrodes, a Voronoi diagram is employed, which partitions the plane into regions of the same nearest vertex [20]. The nodes are referred to as (Voronoi) centers, and the region boundaries as (Voronoi) polygons. The area enclosed by a polygon is called a (Voronoi) cell. Vertices, nodes, and electrodes are used interchangeably in this chapter.

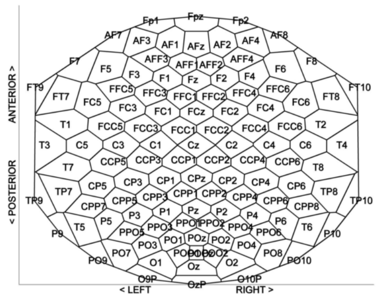


Figure 4.2: *Example of a Voronoi diagram with electrode labels in the corresponding cells (top view of the head, nose at the top). To each electrode a 'Voronoi cell' is associated, consisting of all points that are nearest to that electrode.*

4.3.1.1 *FU Representation*

Under the assumption that multiple electrodes can record the same signal source, a spatially connected set of electrodes recording similar signals is considered as a data-driven region of interest (ROI). Ten Caat *et al.* have proposed methods to detect such ROIs, referred to as *functional units* (FUs) [20]. An FU of an EEG coherence network is a spatially connected maximal clique (a clique is a vertex set in which any pair of vertices is connected by an edge). The FUs of a coherence network can be detected by the maximal clique based (MCB) method. In an EEG coherence network, a larger FU is assumed to correspond to stronger source signals and is considered to be more interesting. The FUs of the network are displayed in a so-called *FU map*, where the barycenter of each FU is the average location of the vertices in this FU (see Figure 4.1).

The FU map is used to illustrate the distribution of FUs, capturing the topology and spatial structure of the coherence networks visually. It partitions the whole electrode set into several FUs. Each FU is a collection of connected polygons (Voronoi cells). The shape of FUs is not regular and the number of electrodes in an FU is also not fixed since the local connectivity is distinct for different regions; for example, in Figure 4.1 we illustrate that the number of electrodes in the FUs can be different. Different FUs represent distinct regions with associated particular connectivity properties.

Given an EEG coherence network $G = (V, E)$, an FU is a set of significantly connected nodes $C = \{v_1, v_2, ..., v_m\} \subseteq V$, where $m = |C|$ is the number of nodes included in C. Since each electrode has a unique x- and y-coordinate, the FU can also be represented by a collection of 2D coordinates: $C = \{(x_{v_1}, y_{v_1}), (x_{v_2}, y_{v_2}), ..., (x_{v_m}, y_{v_m})\}$, where x_{v_i} and y_{v_i} are the x- and y-coordinate of the i-th electrode in C, respectively. The barycenter m_C of C is calculated as $m_C = \frac{1}{|C|}(\sum_{v_i \in C} x_{v_i}, \sum_{v_i \in C} y_{v_i})$.

The FU map F of an EEG coherence network G is a set $F = \{C_1, C_2, ..., C_l\}$ of FUs, where l is the number of FUs in G.

4.3.1.2 *Distance Between FUs*

In FU maps, using only the barycenter of FUs to compute the distance between FUs is not sufficient since the shape of FUs is not regular. For example, any two FUs could have the same barycenter even though their components are totally different. Here, we define the distance $fd(C_1, C_2)$ between FUs C_1 and C_2 as the sum of the spatial distance $D(C_1, C_2)$ and the weighted Jaccard distance $J(C_1, C_2)$:

$$fd(C_1, C_2) = D(C_1, C_2) + \lambda J(C_1, C_2). \tag{4.1}$$

Here, the spatial distance $D(C_1, C_2)$ is defined as the 2D Euclidean distance between the barycenters of C_1 and C_2, and the Jaccard distance is defined as one minus the cardinality of their intersection $|C_1 \cap C_2|$ over the cardinality of their union $|C_1 \cup C2|$: $J(C_1, C_2) = 1 - \frac{|C_1 \cap C_2|}{|C_1 \cup C2|}$. Note that this 2D Euclidean distance is normalized to the interval [0, 1] by scaling it to the maximum possible distance $MaxDist$ in an FU map: $D(C_1, C_2) = \frac{||m_{C_1} - m_{C_2}||}{MaxDist}$. The parameter λ in Eq. (4.1) defines the relative importance of the overlap between FUs; when $\lambda = 0$ the distance between FUs only takes Euclidean distance into account.

4.3.2 *Distance between Coherence Networks*

Functional units, encoding the topology and geometry of local connectivity, capture the local structure of EEG coherence networks. The FU representation of a network describes the distribution of FUs and thus the spatial structure of the original network. Therefore, the comparison of coherence networks can be done in terms of their FU representations. We here propose to define the dissimilarity between two EEG coherence networks as the earth mover's distance (EMD) between the corresponding FU representations using Eq. (4.1) as the ground distance between FUs. For this purpose, we first introduce the concept of EMD, and then model the feature signatures of EEG coherence networks used as input for the EMD.

4.3.2.1 *The Earth Mover's Distance*

The EMD has been used originally to measure the similarity of two distributions. Intuitively, one distribution can be seen as a mass of earth properly spread in space, the other as a collection of holes in that same space. The EMD then measures the minimum work needed to fill the holes with earth. Here, a unit of work corresponds to transporting a unit of earth by a unit of ground distance.

In [108], the Earth Mover's Distance (EMD) was used to compute the similarity of signatures derived from images. Here a *signature* is a representation $\{\mathbf{m}_j, w_j\}$ of a set of clusters, where each cluster is represented by its mean $\mathbf{m}_j$ (the cluster center) in feature space and by the number w_j of points (pixels) in that cluster. The EMD can be computed by solving the well-known transportation problem, which is concerned with finding the minimum cost of transporting a single commodity from a given number of sources to a given number of destinations [28]. The main idea is that a destination can receive its demand from more than one source, and a source can provide commodity for more than one destination. An illustration of the transportation model is shown in Figure 4.3.

4.3.2.2 *Coherence Network Distance*

Here, we extend the EMD method to compare coherence networks with associated FU representations which are compatible with the feature signatures in the EMD.

Given an EEG coherence network $G = (V, E^G)$, the corresponding signature can be represented as $P^G = \{(C_1^G, w_1^G), (C_2^G, w_2^G), ..., (C_m^G, w_m^G)\}$, where C_i^G is the i-th FU of the EEG coherence network G and $w_i^G = \frac{|C_i^G|}{|V|}$ is the relative weight of the corresponding FU C_i^G; $|C_i^G|$ is the number of electrodes in FU C_i^G, and $|V|$ is the number of electrodes for coherence

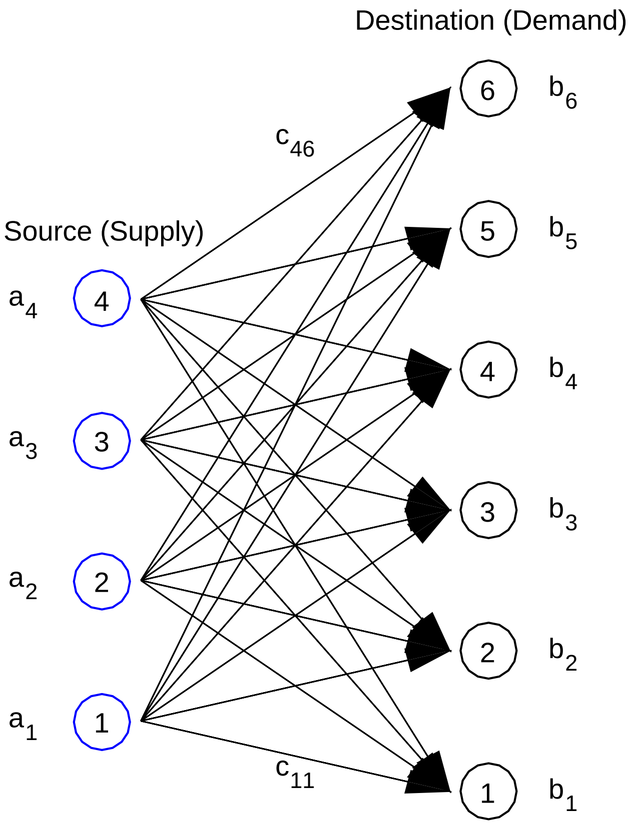


Figure 4.3: *Transportation model with 4 sources and 6 destinations. Each source or destination is represented by a circle; a blue circle represents a source while a black circle represents a destination. The route between a source and a destination is represented by an edge joining the two circles. The amount of supply available at source i is a_i, and the demand required at destination j is b_j. The cost of transporting one unit between source i and destination j is c_{ij}.*

network G. Note that $|V|$ is the same for every EEG coherence network considered in this chapter since the number of electrodes is fixed.

The difference between two networks G_1 and G_2 is then defined as the minimal cost for transforming the signature P^{G_1} into the signature P^{G_2} where Eq. (4.1) determines the cost of transforming a unit of mass from an FU of the first signature to an FU of the second signature. Linear constraints on the movement of mass describe the set of feasible combinations of transformations. This can be formalized as the follows. Given two signatures P^{G_1} and P^{G_2} and a ground distance $fd(C_i, C_j)$, the EMD between P^{G_1} and P^{G_2} is defined as the minimum over all feasible transformations $f \in |P^{G_1}| \times |P^{G_2}|$, where $|P^{G_1}|$ and $|P^{G_2}|$ are the number of clusters in P^{G_1} and P^{G_2}, respectively, and $f(i,j)$ is the flow between $C_i^{G_1}$ and $C_j^{G_2}$:

$$\mathrm{EMD}(P^{G_1}, P^{G_2}) = \min_f \Big\{ \frac{1}{\widetilde{w}} \sum_{i=1}^{m} \sum_{j=1}^{n} fd(C_i^{G_1}, C_j^{G_2}) f(i,j) \Big\}, \tag{4.2}$$

subject to the following constraints:

$$f(i,j) \geq 0 \qquad 1 \leq i \leq m, 1 \leq j \leq n \tag{4.3}$$

$$\sum_{j=1}^{n} f(i,j) \leq w_i^{G_1} \qquad 1 \leq i \leq m \tag{4.4}$$

$$\sum_{i=1}^{m} f(i,j) \leq w_j^{G_2} \qquad 1 \leq j \leq n \tag{4.5}$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} f(i,j) := \widetilde{w} = min\Big\{ \sum_{i=1}^{m} w_i^{G_1}, \sum_{j=1}^{n} w_j^{G_2} \Big\}. \tag{4.6}$$

Constraint 4.3 ensures that the mass of FU is only moved from P^{G_1} to P^{G_2} and not vice versa. Constraints 4.4 and 4.5 ensure that no more mass of FUs is removed from or moved to the FUs than their weights. Constraint 4.6 ensures that in total as much mass as possible is moved. The normalization factor $\widetilde{w}$ (the minimum of the total weights) can be determined before solving Eq. 4.2. The minimization problem in Eq. 4.2 can be solved by the linear programming method (*linprog() function)* available in one of the MATLAB toolboxes.

4.3.3 *Performance*

The performance of the new method is first illustrated on simple synthetic coherence networks. In this toy example, we generated six FU maps (see Figure 4.4). Because larger FUs are more interesting as they capture dominant features of coherence networks, we only analyze FUs whose size is larger than four in this synthetic test set.

In the example, FUs in FU maps a, b, and c are all located at the frontal part of the brain, which means that the frontal part shows more coherent activity for these three participants. Compared to FU maps a, b and c, the FUs of FU map e are only located left and right frontally. In contrast, FUs in FU map d are all located posteriorly, which means that for FU map d the more coherent activity is located at the back of the head. For FU map f, there are two FUs, one mid-frontally and one right-posteriorly. In addition, FU *2* in FU map c is split into FUs *2* and *3* in FU map b, and FU *2* of FU map b is identical to FU *1* of FU map f.

4.3.3.1 *The Influence of λ on EMD*

One issue in computing the EMD is what value of λ (in Eq. (4.1)) to choose. To investigate how the choice of λ impacts the final earth mover's distance, values of λ in the range $[0, 1]$ were considered, in steps of 0.1, and the results of the EMD between the FU maps in Figure 4.4 are shown in Figure 4.5.

From the results it can be observed that with increasing values of λ, which means that the Jaccard distance becomes more and more important in computing the distance between FUs, the EMD between every pair of FU maps is also increasing. This can be easily understood since the Jaccard distance appearing in the right-hand-side of Eq. (4.1) is non-negative. However, the gain in the EMDs for different pairs of FU maps is different. In Figure 4.5, the top three similar pairs of FU maps are (b, c), (a, c), (a, b), and the last seven similar pairs are (a, f), (d, f), (e, f), (d, e), (b, d), (c, d), (a, d), no matter how λ varies. This is obvious, since $D(C_i, C_j)$ and $J(C_i, C_j)$ are constant for every FU pair C_i, C_j.

However, the order of the EMD-values $\mathrm{EMD}(b, e)$, $\mathrm{EMD}(c, e)$, $\mathrm{EMD}(a, e)$, $\mathrm{EMD}(b, f)$, and $\mathrm{EMD}(c, f)$ changes with increasing λ: the increase rate of $\mathrm{EMD}(b, e)$, $\mathrm{EMD}(c, e)$, $\mathrm{EMD}(a, e)$ is similar while $\mathrm{EMD}(b, f)$ has the smallest increase rate. This can be understood as follows. When $\lambda = 0$, only the Euclidean distance between the barycenters of FUs matters for calculating the EMD between FU maps, and the most similar pair of the five is (b, e), followed by (c, e), (a, e), (b, f), and (c, f). This is because FUs of FU maps a, b, c and e are all located frontally while the FUs of FU map f are located posteriorly; hence the distance between FU maps a, b, c and e is less than that between FU maps a, b, c and f. However, when $\lambda = 0.4$, the pair (b, f) has the least EMD among these five EMDs even though one FU of FU map f is far from the FUs of FU map b, followed by (b, e), (c, e), (c, f), and (a, e). This is because when $\lambda > 0$ the overlap between FUs is also taken into account. Since the overlaps (Jaccard distances) between (b, e), (c, e) and (a, e) are larger than those between (b, f) and (c, f), the gains in the corresponding EMDs for in-

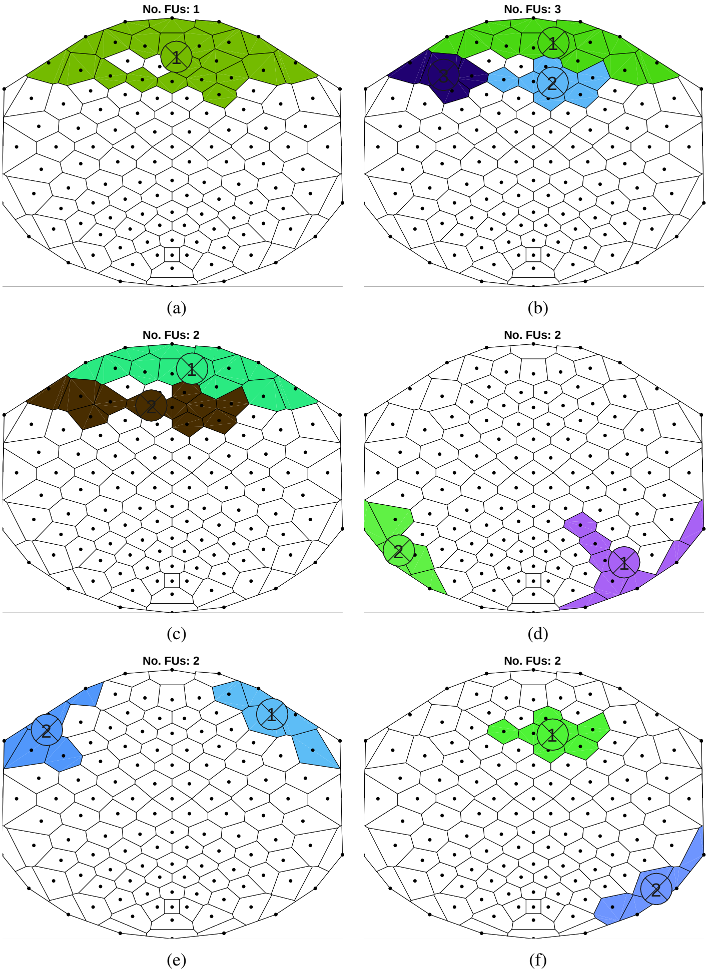


Figure 4.4: *Six simulated FU maps a, b, c, d, e, and f. For each FU map, every FU has a unique colour for distinguishing different FUs. The number within a circle is located at the barycenter of the corresponding FU for reference.*

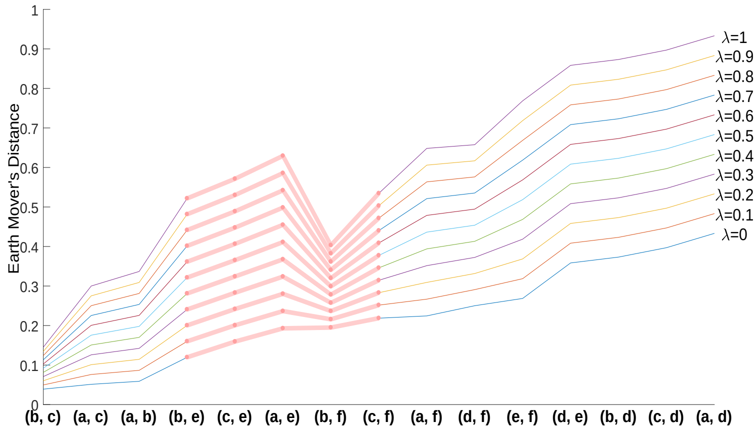


Figure 4.5: *EMD distance between the FU maps shown in Figure 4.4 for values of λ in the range [0, 1].*

creasing λ are larger as well, since λ is simply a factor multiplying the Jaccard distance in the formula for the EMD.

In summary, the values of λ in the range [0, 1] do influence the relative order of the earth mover's distance between FU maps. How to determine the value of λ depends on the situation. When the spatial distance is more important, then λ can be set to a small value. In contrast, when we consider the overlap of FUs to be more important, then λ should be set to a larger value.

4.3.3.2 *Comparison with the Inexact Graph Matching Method*

To demonstrate that the EMD distance measure produces intuitively plausible results, we compare our results with the existing method proposed by [26] which is based on inexact graph matching.

Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, the method can be represented in terms of a bipartite graph $G = (V, E)$, in which the vertex set V can be partitioned into two sets $A = V_1$ and $B = V_2$ such that no edge in E has both endpoints in the same set. The weight of the edges in V represents the cost of transforming one node of A to one node of B. A matching $M \subseteq E$ is a collection of edges such that every vertex of V is incident to at most one edge of M. The method then consist of finding a one-to-one correspondence between nodes of the two graphs such that they "look most similar" when the vertices are labelled according to such a correspondence. If a vertex v has no edge of M incident to it then v is said to be unmatched. In some cases, the dummy vertex concept is used when two graphs have distinct numbers of vertices. An optimal matching M^* is the cheapest matching (lowest total cost) with

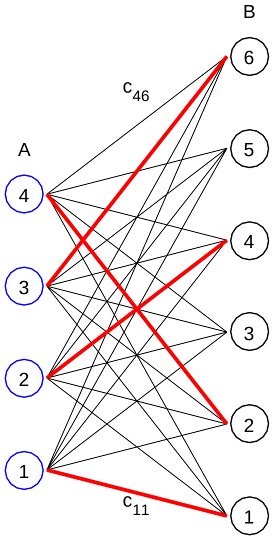


Figure 4.6: *Illustration of the assignment problem. There are four vertices in set A (blue circles), and six in B (black circles). Each node is represented by a circle. The cost for transforming one node in set A to another one in B is c_{ij}. The red edges are an optimal matching M^*. There are no matched nodes in A for 3 and 5 of set B.*

the minimum sum of edge weights [26]. The formulation of this problem can be cast as an assignment problem which is illustrated in Figure 4.6.

We applied both the proposed EMD method and the inexact graph matching method to the FU maps of Figure 4.4. We set $\lambda = 1$ in Eq. (4.1), so that the Jaccard distance and the Euclidean distance are equally important. Tables 1 and 2 give the earth mover's distance and dissimilarity calculated by the inexact graph matching method between the FU maps shown in Figure 4.4. Note that for both for the dissimilarity and the EMD, the lower the values, the more similar the corresponding FU maps are.

It can be observed from Table 1 that in the proposed method the FU map *d* is the least similar with the others. It has a high EMD compared to the other FU maps, since the FUs of FU map *d* are located left and right-posteriorly, while the FUs of the other FU maps are located frontally, except for FU map *f* in which one FU is located right-posteriorly and one is located mid-frontally. However, in the inexact graph matching method (Table 2), FU map *b* is the least similar to the others since *b* has three FUs while each of the remaining FU maps only has two FUs, except for FU map *a* which only has one FU. This method only considers one-to-one matching, which overestimates the dissimilarity between FU maps with a different number of FUs, in which neighbouring FUs are not considered if they are not matched, even though they have

Table 1: *Earth Mover's Distance between the FU maps shown in Figure 4.4 with $\lambda = 1$. The EMDs in ascending order are as follows:* $\mathrm{EMD}(b,c)$, $\mathrm{EMD}(a,c)$, $\mathrm{EMD}(a,b)$, $\mathrm{EMD}(b,f)$, $\mathrm{EMD}(b,e)$, $\mathrm{EMD}(c,f)$, $\mathrm{EMD}(c,e)$, $\mathrm{EMD}(a,e)$, $\mathrm{EMD}(a,f)$, $\mathrm{EMD}(d,f)$, $\mathrm{EMD}(e,f)$, $\mathrm{EMD}(d,e)$, $\mathrm{EMD}(b,d)$, $\mathrm{EMD}(c,d)$, $\mathrm{EMD}(a,d)$.

	a	b	c	d	e	f
a	0	0.3369	0.3000	0.9332	0.6294	0.6483
b		0	0.1451	0.8732	0.5216	0.4033
c			0	0.8970	0.5713	0.5346
d				0	0.8584	0.6576
e					0	0.7686
f						0

some overlap or are spatially close. In contrast, for the EMD method information across FUs is taken into account to compute the distance.

Since one-to-one matching only considers the correspondence between matched FUs and does not use across-FU information, it partially ignores the (spatial) structure of EEG coherence networks. However, in brain connectivity networks the location and number of nodes are fixed for every network, so that the final detected FUs in distinct FU maps are not absolutely different even if they are not matched in the one-to-one matching method. As an example, the two most similar FU maps are (c, e) as detected by the method of Crippa *et al.* and (b, e) as calculated by the EMD method. However, from Figure 4.4, FUs *2* and *3* of FU map b when merged together exactly coincide with FU *2* of FU map c.

Table 2: *Dissimilarity, proposed by [26], between the FU maps shown in Figure 4.4. The dissimilarities in ascending order are as follows:* $dis(c,e)$, $dis(c,f)$, $dis(d,f)$, $dis(a,c)$, $dis(b,c)$, $dis(a,f)$, $dis(a,e)$, $dis(e,f)$, $dis(d,e)$, $dis(c,d)$, $dis(a,d)$, $dis(b,f)$, $dis(b,e)$, $dis(a,b)$, $dis(b,d)$.

	a	b	c	d	e	f
a	0	2.2449	1.2449	1.9269	1.5741	1.4327
b		0	1.2487	2.7070	1.9972	1.9679
c			0	1.8255	1.1426	1.2166
d				0	1.7168	1.2419
e					0	1.6368
f						0

Finally, compared to the assignment problem for graph matching, solving the transportation problem has several advantages. For example, the cost of moving "earth" reflects the notion of nearness properly: items from neighbouring bins now contribute similar costs, and the EMD allows for partial matching in a very natural way. For example, in the proposed method (a, c) is more similar than (c, e) but in the inexact graph matching method the opposite is the case.

4.4 CASE STUDY

We applied the proposed method on EEG coherence networks obtained from 12 subjects (6 young and 6 old) performing a so-called oddball experiment. The median age for the young participants was 29 years (range 25-34 years), and the median age for the old participants was 63.5 years (range 55-67 years).

4.4.1 *Experimental Setup*

Brain responses were recorded during an auditory oddball detection experiment, in which all participants were instructed to count target tones. After the experiment, each participant had to report the number of perceived target tones. For a detailed description of the experiment, please refer to [86].

In the present study we do not consider ongoing EEG but the event-related potential (ERP) which is an EEG recording of the brain response to a sensory stimulus. To calculate the coherence for an ERP with L repetitive stimuli, the EEG data can be separated into L segments of 1 second each, sampled at 256 Hz. Coherences were calculated between pairs of electrode signals. A *significance threshold* for the estimated coherence is given by [53]:

$$\theta = 1 - p^{1/(L-1)}, \tag{4.7}$$

where p is a probability value associated with a confidence level α, such that $p = 1 - \alpha$.

Throughout this section, we use $p = 0.01$, and $L = 13$ segments. In addition, we set $\lambda = 1$ in Eq. (4.1).

4.4.2 *Experimental Results*

We computed the EMD between all FU maps of Figure 4.7 to investigate the inter-subject variability in three frequency bands. Results are presented in Figure 4.8.

The results in general show that similarities between FU maps decrease with increasing frequency band. The EMDs between young participants usually have smaller values across frequency bands than old participants, which means the FU maps of young participants are more similar than those of old participants. High EMD values usually occur between young and old participants. The highest EMD value (0.5333)

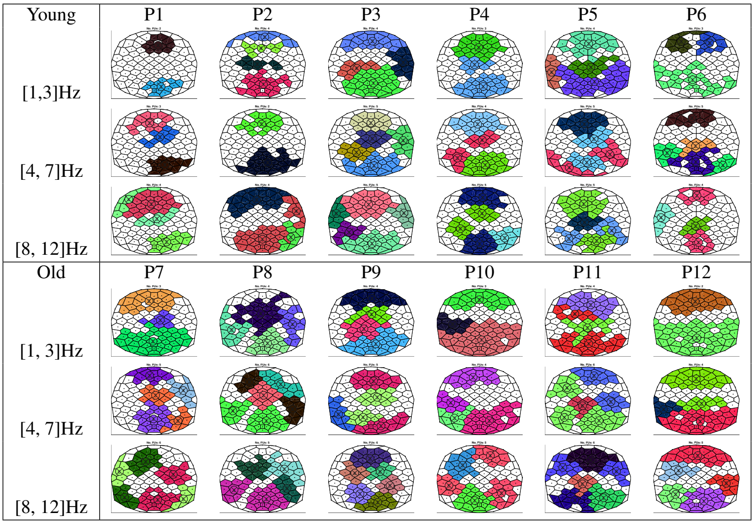


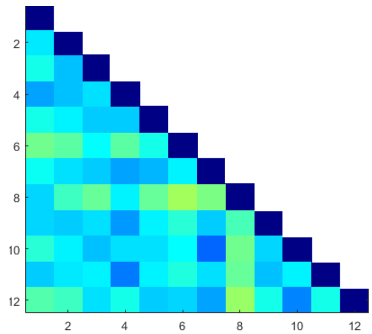
Figure 4.7: *FU-maps for the young and old participants computed for three frequency bands.*

occurs between participants P6 and P8 for the frequency band [1, 3]Hz (Figure 4.8(a)). In addition, the results show a higher inter-group variability than within-group inter-subject variability of young and old participants across frequency bands. The inter-group variability between FU maps also increases with increasing frequency band.

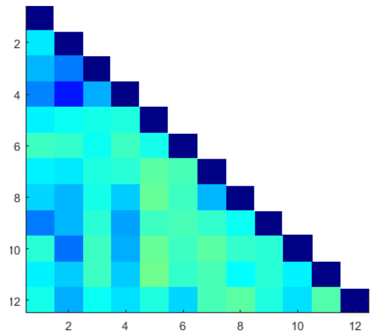
Young participants have smaller EMDs for the frequency band [4, 7]Hz than for the other two frequency bands. It also can be seen from Figure 4.8(b) that the values of EMDs among P1, P2, P3, and P4 are small, which means their FU maps are more similar. The smallest EMD value (0.1454) occurs between participants P2 and P4 for the frequency band [4, 7]Hz (Figure 4.8(b)). In addition, the inter-subject variability within the young group for the frequency band [8, 12]Hz is higher than for the other two frequency bands (Figure 4.8(c)). Participant P6 has the highest EMDs compared to other young participants across the frequency bands.

For the old participants, EMDs between FU maps also increase with increasing frequency. There is a lower inter-subject variability for the frequency band [1, 3]Hz than for the other two frequency bands (Figure 4.8(a)). Participant P8 generally has higher EMD values to other old participants across frequency bands, especially for the frequency band [1, 3]Hz.

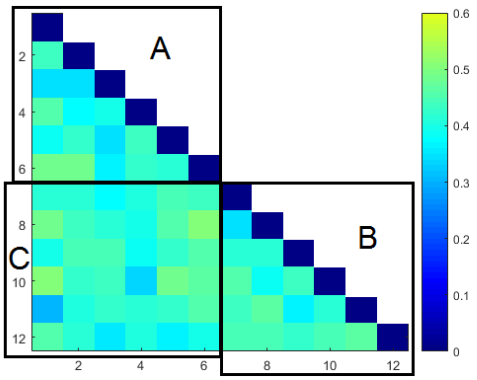
In general we observe more variability in measures obtained from older than from younger participants, due to the different ageing trajectories people experience, i.e., some people age faster than others.



(a) EMDs between FU maps for [1, 3]Hz



(b) EMDs between FU maps for [4, 7]Hz



(c) EMDs between FU maps for [8, 12]Hz

Figure 4.8: *Inter-subject variability based on earth mover's distance between FU maps of Figure 4.7: (a) EMD between FU maps for [1, 3]Hz; (b) EMD between FU maps for [4, 7]Hz; (c) EMD between FU maps for [8, 12]Hz. The color of cells encodes the value of EMDs for the corresponding FU maps. The number at the left and bottom for each triangle is the index of the participants: 1 to 6 for young participants, and 7 to 12 for old participants. Square A of (c) represents the inter-subject variability for the young participants while B is for the old participants, and C is the inter-group variability.*

4.5 CONCLUSIONS AND FUTURE WORK

The key contribution of this chapter is the introduction of a method for comparing EEG coherence networks using local structure. It compares EEG coherence networks using the earth mover's distance (EMD) between their FU representations. This FU representation of a coherence network accounts for connectivity and physical location of the vertices. It thus captures the local spatial structure of coherence networks. The performance of the proposed method on synthetic examples showed a higher ability than the method of [26] which is based on inexact graph matching. Also, our new method showed a high capacity to detect inter-subject variability among functional brain networks obtained from a so-called oddball experiment in which the participants performed a task involving the recognition of target tones.

Currently, our method has a number of limitations. First, the method is proposed as a preliminary step towards a complete quantitative comparison, and its real benefits, including the statistical significance of the network comparisons, still have to be assessed. Second, regarding the application to brain networks, the distance between FUs was assumed to be a combination of spatial (Euclidean) distance and Jaccard distance (accounting for overlap). However, this distance does not fit perfectly since the coherence between electrodes of the same FU is varying. Therefore, other ground distances that take the detailed connectivity within an FU into account may present opportunities to further improve the technique.

Another area for future work is to apply the method for comparison between groups of brain connectivity graphs obtained under different stimulus conditions. The proposed method might help to assess brain activity categorization.

Finally, our method is not limited to EEG coherence networks, but can be extended to other functional networks where the preservation of spatial information is important.