



**HAL**  
open science

# Graph Alignment Exploiting the Spatial Organisation Improves the Similarity of Brain Networks

Anna Calissano, Théodore Papadopoulo, Xavier Pennec, Samuel Deslauriers-Gauthier

► **To cite this version:**

Anna Calissano, Théodore Papadopoulo, Xavier Pennec, Samuel Deslauriers-Gauthier. Graph Alignment Exploiting the Spatial Organisation Improves the Similarity of Brain Networks. *Human Brain Mapping*, 2024, 45 (1), pp.e26554. 10.1002/hbm.26554 . hal-03910761v2

**HAL Id: hal-03910761**

**<https://inria.hal.science/hal-03910761v2>**

Submitted on 22 Dec 2023 (v2), last revised 6 Sep 2024 (v3)




**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Graph Alignment Exploiting the Spatial Organisation Improves the Similarity of Brain Networks

Anna Calissano <sup>1</sup> | Theodore Papadopoulo <sup>1</sup> |  
Xavier Pennec <sup>1</sup> | Samuel Deslauriers-Gauthier <sup>1</sup>

<sup>1</sup>Inria Center at Université Côte d'Azur, Valbonne, 06902, France

## Correspondence

Anna Calissano PhD, Inria Center at Université Côte d'Azur, Valbonne, 06902, France  
Email: [anna.calissano@inria.fr](mailto:anna.calissano@inria.fr)

## Funding information

European Research Council (ERC) (grant agreement G- Statistics No. 786854)

## Conflict of Interest

There are no relevant financial or non-financial competing interests to report.

## Data Availability

The data that support the findings of this study are openly available from the Human Connectome Project at <https://www.humanconnectome.org/>.

Every brain is unique, having its structural and functional organisation shaped by both genetic and environmental factors over the course of its development. Brain image studies tend to produce results by averaging across a group of subjects, under the common assumption that it is possible to subdivide the cortex into homogeneous areas while maintaining a correspondence across subjects. We investigate this assumption: can the structural properties of a specific region of an atlas be assumed to be the same across subjects? This question is addressed by looking at the network representation of the brain, with nodes corresponding to brain regions and edges to their structural relationships. Using an unsupervised graph matching strategy, we align the structural connectomes of a set of healthy subjects, considering parcellations of different granularity, to understand the connectivity misalignment between regions. Firstly, we compare the obtained permutations with four different algorithm initializations: Spatial Adjacency, Identity, Barycenter and Random. Our results suggest that applying an alignment strategy improves the similarity across subjects when the number of parcels is above 100 and when using Spatial Adjacency and Identity initialization (the most

plausible priors). Secondly, we characterize the obtained permutations, revealing that the majority of permutations happens between neighbours parcels. Lastly, we study the spatial distribution of the permutations. By visualizing the results on the cortex, we observe no clear spatial patterns on the permutations and all the regions across the context are mostly permuted with first and second order neighbours.

#### KEYWORDS

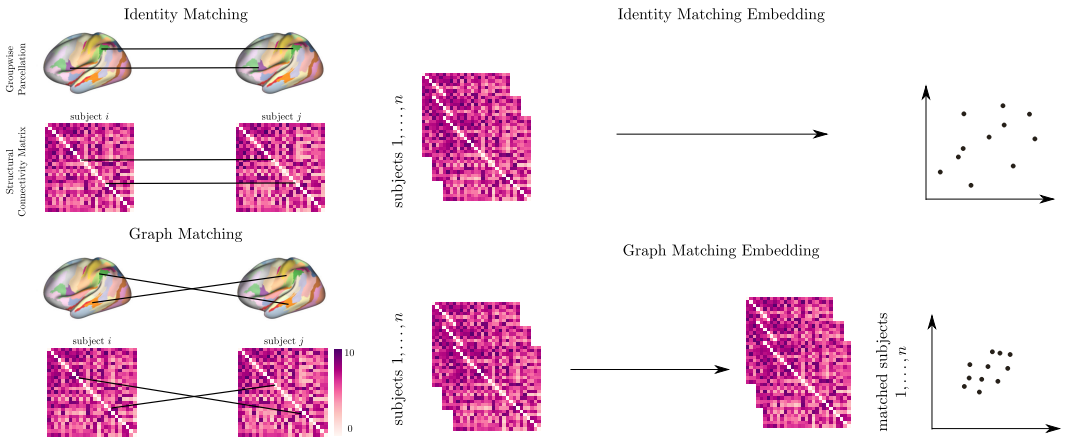
Structural Parcellation, Network Alignment, Graph Matching, Graph Space, Unlabelled Graphs, Healthy Subjects, Group-wise Parcellation

## 1 | INTRODUCTION

Every brain is unique, having its structural and functional organisation shaped by both genetic and environmental factors over the course of its development (Thompson et al., 2001). This uniqueness directly translates into inter-subject variability in the location and extent of functional brain areas, the network organisation of structural connectivity, and the structure–function relationship (Vigneau et al., 2006; Sporns et al., 2005; Mueller et al., 2013; Tate et al., 2014). In imaging studies which tend to produce results by averaging across a group of subjects, this variability introduces temporal and spatial variance in the localisation of specialized brain areas and their network connectivity. In particular, this variability undermines the common assumption that it is possible to subdivide the cortex into homogeneous areas while maintaining a correspondence across subjects. In other words, can we assume the structural and functional properties of a specific region of an atlas to be the same across subjects? This question is of particular interest to studies that represent the brain as a network, where nodes correspond to brain regions and edges to their structural or functional relationship (Sporns et al., 2005; Preti et al., 2017). If there is a spatial variability across subject, then the nodes of the network may not be matched, and the networks of distinct subjects should not be compared directly.

A possible solution to this problem is to make no assumption on the correspondence between the nodes of the network and to try to derive it using a data driven approach. This solution acts at the connectivity matrix level and jointly addresses: a possible spatial misalignment between the parcels that occurred during the registration phase; and a possible structural connectivity difference across subjects. Aligning the nodes of two networks based on their connectivity properties is referred as *graph matching* or *network alignment* (Conte et al., 2004). Given two networks, network alignment corresponds to finding a permutation of the nodes of one graph which minimizes a cost function - usually a dissimilarity measure between the graphs. As the problem is NP-hard, many different techniques have been proposed in the literature, from exact matching to inexact matching (see Conte et al. 2004; Emmert-Streib et al. 2016 for exhaustive reviews on the topic). Notice that network alignment differs from the fingerprint matching and the parcel matching problem (Mars et al., 2016; Gallardo et al., 2018a) as it only relies on the intrinsic properties of the network, for example edge weights, instead of node specific features. The network alignment problem gained attention in the context of neuroscience, aiming at understanding the heterogeneity inter-subjects brain connectivity networks. However, current approaches were designed for the alignment of general networks, for example social networks, and do not typically impose spatial constraints on the node permutations. In the context of brain networks,

arbitrary permutation may result in the large spatial displacement of brain areas which is highly improbable. Therefore, while local permutation explaining inter-subject variability is expected, permutation resulting in large displacement should be discouraged (e.g., permuting regions in the occipital and frontal lobes). Therefore, a graph alignment strategy promoting or enforcing local permutation may be more appropriate to brain network alignment. In Figure 1, we illustrate the difference between the identity matching, where the regions across subjects are considered as the same nodes, and the graph matching, where we search for a better matching of regions across subjects. Both different similarity measures as well as different matching strategies can be defined, creating a wide range of approaches (Frigo et al., 2021; Deslauriers-Gauthier et al., 2020; Rasero et al., 2017; Frost et al., 2014; Takerkart et al., 2014; Becker et al., 2018; Weinstein et al., 2021; Ravindra et al., 2020).



**FIGURE 1** Left: Conceptual visualization of identity matching and graph matching in the context of structural brain connectivity. Right: Conceptual visualization of graph embedding in the context of identity matching and graph matching.

In this work, we aimed at exploring the stability of group-wise parcellations when performing inter-subject analysis by looking at network alignment performances and characteristics on a set of control subjects. Given a set of structural connectivity matrices derived from an extrinsic connectivity parcellation (Gallardo et al., 2018b), we study the network alignment results. Among the different alignment strategies, we choose the Fast Approximate Quadratic (FAQ) algorithm (Vogelstein et al., 2015), as it represents a state of the art in the network alignment literature. It is also known to be well tailored for the analysis of brain connectivity networks (Glaser and Kording, 2016; Vogelstein et al., 2015). The analysis is conducted on the different priors of the matching procedure, by imposing different initial conditions to the graph matching algorithm, and at different parcels granularities. The proposed approach allows for the definition of the spatial initialization as a suitable initialization for structural connectivity data and the identification of the most stable and most unstable brain area in brain matching across subjects.

Our main contribution consists in showing with a purely data driven approach how inter-subject variability of structural connectivity graphs can be decreased using graph matching and with the correct initialization. We also characterized the area of the brain that are more often permuted, showing where the misalignment often resides.

## 2 | MATERIALS AND METHODS

### 2.1 | Brain Graph Matching

A weighted graph is a mathematical object  $G_i = (N, E_i, w_i)$  with a set of nodes  $N = \{1, \dots, n\}$ , a set of edges  $E = \{(j, k) \in \{0, 1\}, j, k \in N\}$ , and a weight function  $w_i : E \rightarrow \mathbb{R}$ . In matrix term, a graph can be represented as a weighted adjacency matrix  $G_i \in \mathbb{R}^{n \times n}$ . In the context of brain networks, the number of nodes  $n$  corresponds to the number of parcels per hemisphere and  $w$  often encodes the connectivity, either structural or functional, between parcels. Given the matrix representation, different metrics can be defined to compare graphs (Donnat and Holmes, 2018; Frigo et al., 2021; Bassett and Bullmore, 2017; Pandit et al., 2013). **In this work, we are using the Frobenious norm:**

$$d_F(G_1, G_2) = \|G_1 - G_2\|^2 = \sum_{i=1}^n \sum_{j=1}^n (G_1(i, j) - G_2(i, j))^2$$

Note that the Frobenious norm is not invariant under node permutations, it thus allows for an analysis of the suggested nodes' permutations. In addition, it allows for a quick graph matching procedure (Vogelstein et al., 2015) and it has been broadly used in the context of structural connectivity comparison. Graph matching corresponds in finding an injective map from the nodes of the graph  $G_i$  to the nodes of the graph  $G_j$ . In geometrical terms, the nodes (parcels in this context) permutation is expressed using the matrix representation of the permutation group  $\mathcal{P} = \{P \in \{0, 1\}^{n \times n} : P^T 1 = P 1 = 1\}$ . Given an adjacency matrix  $G_i$ , it can be permuted into potentially  $n!$  different adjacency matrices by applying the permutation action. The results is an equivalence class of graphs  $[G_i] = \{P^T G_i P : P \in \mathcal{P}\}$ , obtained by permuting the rows and the columns of the adjacency matrices by matrix multiplication. If we are using the metric as the alignment cost function and the permutation matrix representation, aligning two networks  $G_i$  and  $G_j$  corresponds in finding the permutation matrix  $P_{ij}$  minimizing:

$$d([G_i], [G_j]) = \min_{P \in \mathcal{P}} d_F(G_i, P^T G_j P) \quad (1)$$

This minimization problem corresponds to the graph matching problem. In addition,  $d([G_i], [G_j])$  is an induced metric on the quotient space  $\mathbb{R}^{n \times n} / \mathcal{P}$  obtained by applying the permutation action to the space of adjacency matrices. This space is often called graph space (see for details Jain and Obermayer 2009; Calissano et al. 2023). Note that the graph matching is only the simplest operation which can be performed in graph space, as the theoretical framework for the analysis of graphs is richer, including for example Fréchet mean (Jain and Obermayer, 2010), principal components (Calissano et al., 2023), regression (Calissano et al., 2022; Severn et al., 2020), and testing (Ginestet et al., 2017).

In the context of brain network analysis, graph matching represents a way to understand if there is a misalignment between subjects, and provides a way to correct it. A misalignment is detected when the optimal permutation matrix obtained by equation (1) is not the identity. In this situation, the similarity between subjects can be improved by permuting the labels of one subject, implying that the initial labels do not match. The optimal permutation matrix  $P$  to apply to the labels of subject  $j$  is also produced by Equation (1) and can thus be applied as preprocessing step before further analysis or comparison of the networks.

#### 2.1.1 | FAQ algorithm for graph matching

Graph Matching is an NP-hard problem which has been largely explored in the literature (see Conte et al. (2004) for a review). Among the different inexact graph matching algorithms, we selected the Fast Quadratic Assignment

method proposed by Vogelstein et al. (2015) as it represents the state of the art method and it has been used in the context of human and insect brain matching (Frigo et al., 2021; Winding et al., 2023). The FAQ matching is a fast method for inexact graph matching based on the Frobenious norm between adjacency matrices and it allows to select an initialization (a prior) to the matching procedure, which is important in this context where not all the regions permutations are meaningful.

The idea behind the methodology is to relax the domain of permutation matrices from the discrete set  $\mathcal{P}$  to the continuous convex hull of doubly stochastic matrices  $\mathcal{D} = \{P \in \mathbb{R}^{n \times n} : P^T \mathbf{1} = P \mathbf{1} = \mathbf{1}, P \geq 0\}$ . Starting from the minimization of the Frobenious norm between two adjacency matrices  $G_i, G_j \in \mathbb{R}^{n \times n}$ , the problem is manipulated into the following minimization:

$$\text{minimizing } f(P) = -\text{trace}(G_i P G_j^T P^T) \quad (2)$$

$$\text{subject to } P \in \mathcal{D}$$

The local solution is found by gradient descend in the space  $\mathcal{D}$ . To obtain a permutation matrix as output, the solution is then projected onto  $\mathcal{P}$ . As many optimization problems, the FAQ algorithm relies on an initialization  $P_0$ , which could be interpreted as a prior knowledge on the parcels' matching. Given the complexity of the problem, this prior knowledge has a significant impact on the quality of the final solution. In FAQ, such initialization matrices has to be a doubly stochastic matrix (Gagniuc, 2017). Given a permutation matrix, we obtain a doubly stochastic matrix using the Sinkhorn Knopp Algorithm (Sinkhorn and Knopp, 1967). We are proposing four different initializations:

1. **Barycenter:** starting from a non informative doubly stochastic matrix where all entries have the same value. This corresponds to picking all the regions' permutations as equally probable;
2. **Random:** starting from a random doubly stochastic matrix. This encode guessing some permutation at random;
3. **Identity:** starting from the identity matrix. This initialization suggests that the current labels are correct;
4. **Spatial Adjacency:** starting from the spatial structure of the regions. This matrix is derived from the parcellation of the cortical surface where adjacent regions are considered as neighbors. Such initialization suggests a permutation with the spatial neighbours.

## 2.2 | Structural brain networks

Our dataset consists of 100 healthy young adults from the Human Connectome project, (U100 subject group) (Van Essen et al., 2012; Glasser et al., 2013). For each subject, a five-tissue-type image (Smith et al., 2012) was obtained using the Freesurfer pipeline (Fischl, 2012) invoked through Mrtrix3 (Tournier et al., 2019). A response function was estimated for the white matter, gray matter, and cerebrospinal fluid using a maximal spherical harmonic order of 8 for all tissues (Jeurissen et al., 2014). The fiber orientation distribution functions (fODFs) were then computed using the multi-shell multi-tissue constrained spherical deconvolution algorithm (Jeurissen et al., 2014). Finally, the fODFs were used as input for probabilistic anatomically constrained tractography performed with the iFOD2 algorithm (Smith et al., 2012) seeding from the gray matter - white matter interface and obtaining a total of five million streamlines per subject. The structural brain connectivity network are built from the parcellation introduced in Gallardo et al. (2018b), as we are interested in group-wise parcellation. This parcellation is built based on 138 tractograms built from co-registered cortical surfaces from Human Connectome Project (Van Essen et al., 2012). The resulting parcellation is

defined on the average FreeSurfer surface (MSMAll) with 32k vertices. A clustering procedure offers different regions granularity by selecting a threshold and sequentially aggregating regions with similar connectivity structure. Thus, we study the parcellation at different granularity from 50 up to 1000 nodes, to assess the threshold where the original alignment across subjects is inexact.

## 2.3 | Experiments Set Up

The aim is to study how graph matching behaves when comparing healthy subjects' brain structural graph. By analysing the graph matching results at different parcellations and with different initialization, we can infer which areas of the brain most suffer from misalignment errors in the parcellation strategy. To evaluate the quality of the alignment, we use the ratio between  $F_{opt}$ , the objective function obtained FAQ graph matching, and  $F_{id}$  the objective function obtained with the identity matching (i.e., leaving the graphs with the current nodes labels).

$$F_{opt}^{ij}(P) = \text{trace}(G_i P G_j^T P^T) \quad F_{id}^{ij} = \text{trace}(G_i I^T G_j I) \quad (3)$$

Given  $\hat{P} \in \mathcal{P}$ , the permutation matrix estimated via FAQ and  $I \in \mathbb{R}^{n \times n}$  the identity matrix of the original alignment,

$$\frac{F_{opt}^{ij}(\hat{P})}{F_{id}^{ij}}$$

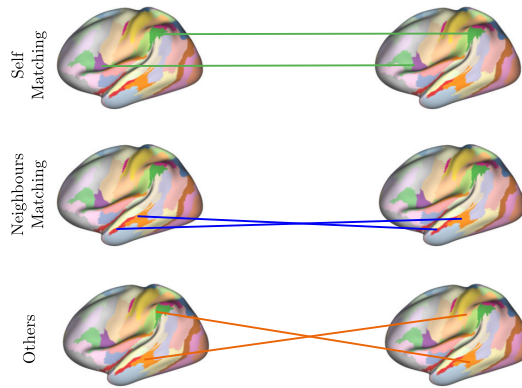
If this ratio is below 1, the observations are better aligned with the identity matching. If the ration is above 1, the observations are better aligned if a permutation is applied.  $F_{opt}^{ij}$  and  $F_{id}^{ij}$  refers to the alignment between subject  $i$  and subject  $j$ . The choice of pairwise matching rather than template matching is driven by two main factors: The computation of an average is non trivial in the case of node unlabelled graphs (see Kolaczyk et al. (2020); Calissano et al. (2023) for further details); The pairwise matching also offers an higher number of matching and permutation matrices, for a more detailed and robust understanding of the misalignment between subjects.

### 2.3.1 | Comparing different initialization

For all the parcellations and hemispheres, we are aligning all the graphs  $G_i$  with all the other graphs  $G_j$ ,  $i, j = 1, \dots, 100$  computing via FAQ the optimal permutation matrices  $\hat{P}_{ij}$ . The aim of this first experiment is to see how the initialization parameter changes the performance of the graph matching with respect to the identity matching.

### 2.3.2 | Comparing different permutations

Once the graph matching is performed, we are interested in understanding how the regions are matched by looking at the optimal permutation matrices  $\hat{P}_{ij}$ . We divide the type of permutations into three categories: Self Matching - when the parcel is permuted with itself; Neighbours Matching - when the parcel is permuted with a neighbouring parcel; and Others - when the permuted regions are not direct neighbors. In Figure 2, we provide a conceptual visualization of the type of permutation between regions.



**FIGURE 2** Types of regions matching: (Top) Self Matching representing regions which are permuted with themselves; (Middle) Neighbours Matching representing regions permuted with their spatial neighbours on the brain; (Bottom) Others representing permutations with other regions rather than themselves or neighbours.

### 2.3.3 | Spatial distribution of permutations

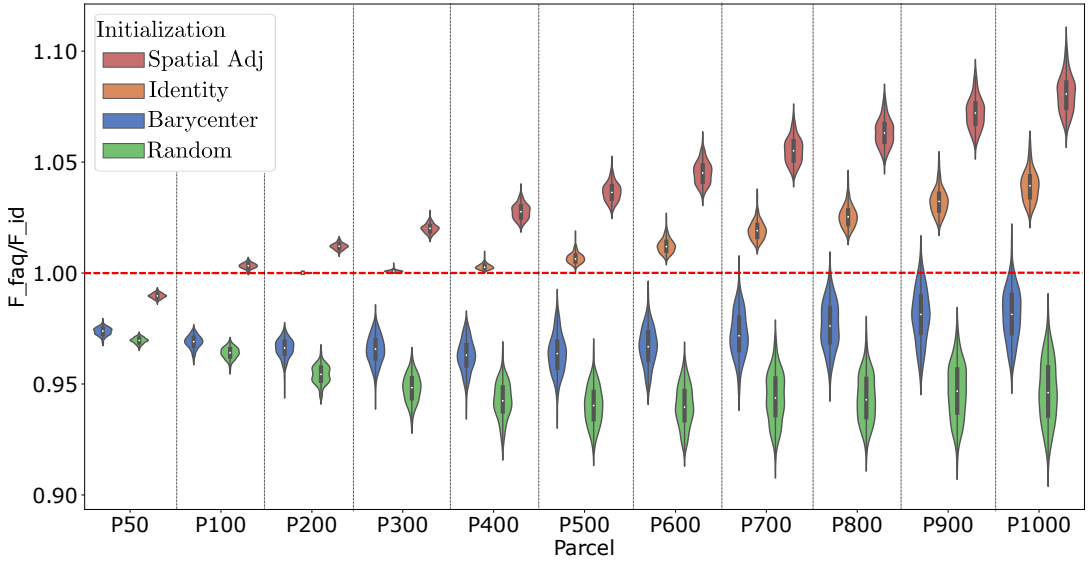
The matching results can be also illustrated on cortical surface to visualize the areas which are more often matched. Specifically, for each region we computed its permutation rate, differentiating over the spatial distance of the proposed permutation. We denote  $r_p^n$  the rate of permutation with an  $n^{\text{th}}$  neighbor. The symbol  $r_p^0$  therefore corresponds to the rate of "no permutation" (or self permutation),  $r_p^1$  the rate of permutation with a direct neighbor, etc. This visualization provides insights on the spatial misalignment across subjects and if certain regions are more permuted than other. It is also possible to visualise individual matching results by computing the rate of permutation of one region with all others across all subjects. This provides a region by region view of the stability of regions and the most likely permutations.

## 3 | RESULTS

### 3.1 | Comparing different initialization strategies

In Figure 3, we illustrate the distribution of the ratio for every parcellation and every initialization. The barycentric and random initialization are systematically outperformed by the original regions matching, indicating that it is preferable to perform no matching. Contrarily, the identity and the spatial initialization outperform the original regions matching for every atlas with a number of regions at or above 100, suggesting inter-subjects alignment. Furthermore, the improvement increases with the number of regions. For example, at 100 parcels with the barycenter initialization the percentage of permuted parcels is 21% and it doubles at 1000 parcels, reaching 58.5%. The results are in line with the expectations, as the most significant priors on the brain regions matching are the identity and the spatial rather than the random and the barycenter, as it is more plausible to have either none or only local misalignment between regions in the brain. The plot shows how the alignment with identity or spatial initialization can decrease the misalignment error inter-subjects.





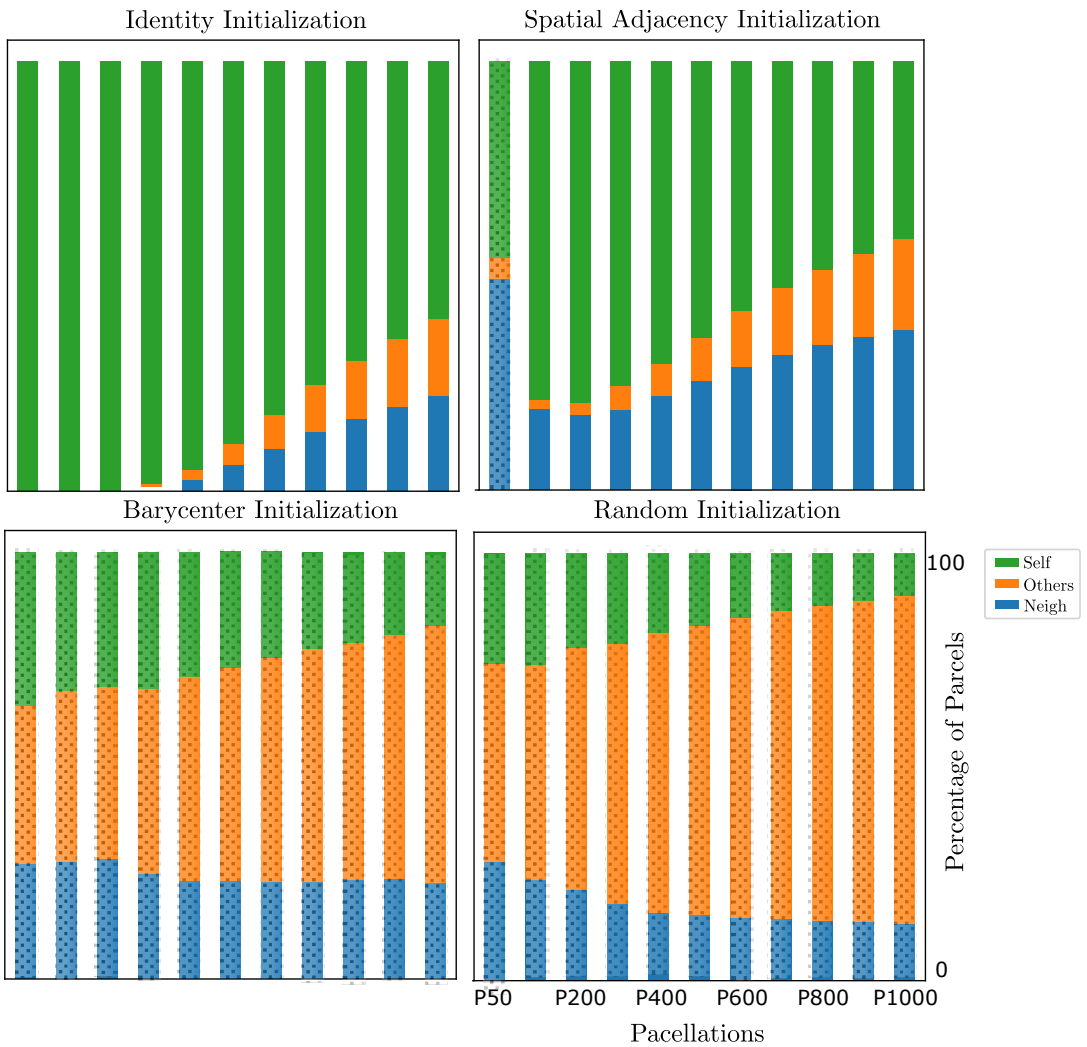
**FIGURE 3** For each parcellation  $\{P50, \dots, P1000\}$ , we compute all the subject-to-subject matching. We plot the distribution of the ratio between the graph matching objective function and the identity matching objective function. When the ratio is above 1, the graph matching is preferable over the identity matching. Different algorithm initializations are considered (Spatial Adjacency, Identity, Barycenter and Random).

### 3.2 | Comparing different permutations

The percentage of regions belonging to Self Matching, Neighbours Matching, and Others is illustrated in Figure 4. When the graph matching was outperformed on average by the identity matching, indicating a poor matching, we report the results with a dashed line. For all four initialization strategies the percentage of permuted nodes increases as the number of regions increases. For the identity and the spatial initializations, the percentage of Self Matched nodes is higher than the percentage of matched nodes. Among the matched nodes, the percentage of nodes matched with Neighbours is predominant with respect to regions permuted with others. These are plausible results, as we expect the regions to be permuted with neighbours due to local misalignment during the parcellation procedure. In addition, the identity initialization shows a higher percentage of self-permuted regions while the spatial initialization shows a higher percentage of neighbours-matched regions.

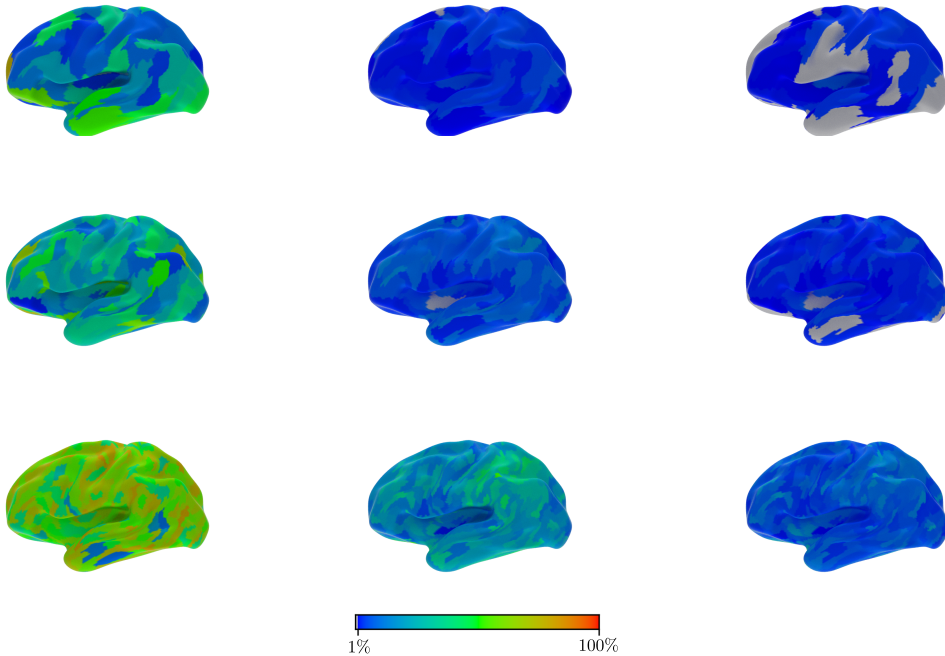
### 3.3 | Spatial organisation of the permutations

To understand the spatial distribution of the different permutations (self, other, neighbours), we represent the permutations obtained with the Adjacency Initialization on the cortex. We chose this initialization as it is the best performing one according to the previous results. Figure 5 illustrates the permutation rates  $r_p^0$ ,  $r_p^1$ , and  $r_p^2$ , for atlases with 100, 300, and 1000 regions. In agreement with the results of Figure 4, the majority of permutations are local, with their rate quickly decreasing with increased pair distance. No spatial organisation is noticeable, indicating that all regions have the same permutation distribution. Examples of permutations for specific regions are illustrated in Figure 6 for similarly located regions across the atlases with 100, 300, and 1000 regions. We observe a high locality of permutation, with the rate of permutation rapidly decreasing with increased distance from the region of interest. As the number



**FIGURE 4** Type of permutations. For each parcellation, the proportion of *self* permuted parcels, parcels permuted with *neighbours* and parcels permuted with *other* parcels. The lines are dotted when the graph matching is under-performing the identity matching. By looking at both the Identity Initialization and the Spatial Adjacency Initialization we notice an increase in the number of permuted parcels when the number of parcels increase, maintaining an higher proportion of parcels permuted with neighbours.

Any Permutation

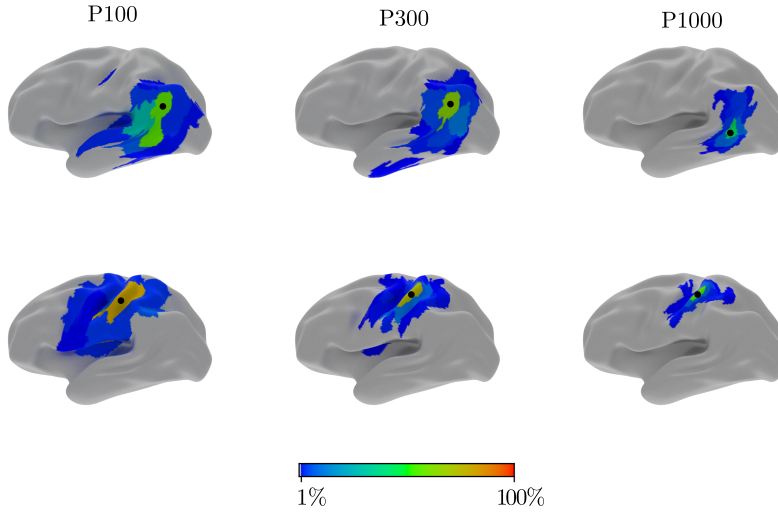
Permutation with  
region not 1<sup>st</sup> neighboursPermutation with  
region not 1<sup>st</sup> and 2<sup>nd</sup> neighbours

**FIGURE 5** Surface visualization of the permutation type. For each regions, we considered the rate at which a regions was permuted with any other region (i.e. not a self permutation), with a region that is not a 1st degree neighbor, and with a region that is not a 1st or 2nd degree neighbor. Warmer and colder colors indicate a higher and lower permutation rate, respectively. Gray indicates a permutation rate below 1%. The three rows correspond to a brain parcellation with 100, 300, 1000 regions.

of regions of the atlas is increased, the permutations become more local and stable around the region of interest. The increased locality for atlases with more parcels is expected as the number of parcels increases the parcel size is reduced, leading to permutations with short distances. This is essentially a consequence of the neighborhood being defined via parcel adjacency instead of cortical distance.

## 4 | DISCUSSION

Studies representing the brain as a network often assume a correspondence between the nodes across the subjects. However, due to inter-subject variability affecting the structural and functional organisation of the cortex, the labels of the nodes may not be matched across subjects, leading to incorrect comparisons. In this work, we proposed a network alignment strategy that exploits the spatial organisation of the cortical surface and thus promotes local permutation of node labels. When applied to the structural connectivity matrix of healthy young adults, our proposed alignment reduced inter-subject variance.



**FIGURE 6** Given the brain region marked with the black dot, we plot the frequency of permutation with other regions for three atlas granularity (100, 300, 1000 regions). The frequencies are computed considering the whole set of subject-to-subject permutations obtained with the spatial initialization. The considered region is permuted with gray regions less than 1% of the times and with the other regions as indicated by the legend: warmer colors indicate high frequency of permutation, colder colors indicate low frequency of permutation.

In more detail, we considered the structural brain networks of 100 healthy subjects from the HCP, obtained from a state-of-the-art diffusion MRI tractography pipeline. The nodes of the network corresponded to the regions of a groupwise structural parcellation of the cortex with the number of regions varied from 50 to 1000. Among the different available network alignment techniques, we selected the FAQ matching as it allowed us to include priors on the spatial organisation of the cortex. We choose a spatial adjacency matrix as initialization of the algorithm, encoding the spatial relations of the parcels. Compared to other initialization strategies, using the adjacency matrix as a starting point systematically increased the similarity between subjects for atlases with more than 50 regions, as illustrated in Figure 3. On the other hand, using a random or barycentric initialisation lead to a decrease of inter-subject similarity. Note that our baseline for similarity is the identity matrix, which corresponds to the assumption that the atlas is matched across subject made in current studies. Our results suggest that, for a healthy cohort, network alignment should be performed before any comparison of the networks across subjects. The whole framework is implemented in a notebook in the `geomstats` python package<sup>1</sup>.

The importance of exploiting the spatial organisation of the cortex when performing network alignment is further explored by looking at the individual permutations. Specifically, we considered the distance between permuted regions, in terms of the neighborhood. As illustrated in Figure 4, most permutation are either with the same label (no permutation  $r_p^0$ ) or with a direct neighbor ( $r_p^1$ ). As the number of regions increase, there is an increase in the rate of long range permutation ( $r_p^n$  with  $n > 1$ ), reaching 21% for 1000 regions. This phenomenon can be explained by the size of the regions of the atlas which decreases as the number of parcel increases. In other words,  $r_p^1$  and  $r_p^2$  may correspond to the same cortical distance on two different atlases. These results suggest that a better notion of

<sup>1</sup><https://github.com/geomstats/geomstats>

neighborhood could be estimated with the distance between cortical regions rather than their adjacency in terms of cortical mesh. Such approach would require the selection of an arbitrary distance threshold.

To understand the spatial distribution, the permutation frequencies for each region of the brain were illustrated in Figure 5. Overall, no specific patterns were identified and all regions presented a similar permutation distribution. From the inter-subject variability point of view, this supports the idea that the variance is not region specific and is instead uniform over the cortex. As illustrated in Figure 6, the permutation patterns were also consistent over different resolutions of the atlas. Regions were in general permuted with close neighbors, again highlighting the importance of the neighborhood information in the graph matching problem. While our proposed strategy favors local permutations, it does not forbid long range ones and they were observed in some cases. Because we considered structural connectivity, the permutation of these nodes can indicate that they fulfil similar roles in terms of structural network organisation, or at least that they have similar connectivity to the rest of the brain. This could occur, for example, in tightly interconnected subnetworks with few connections to other subnetworks.

Previous studies Tian et al. (2022); Chamberland et al. (2017) have reported that interindividual variability is heterogeneous across the cortex, contrasting with the relatively homogeneous permutation maps of our results. Indeed, because our study relies on very similar data and processing, it is reasonable to expect our connectivity matrices to have the same variability profile. However, our goal in performing graph alignment is to improve the global similarity of the matrices and our approach does not seek to specifically correct for regions with high variability. In addition, the permutations identified by the graph alignment do not all have the same impact on the variability. Some permutations may greatly reduce it (e.g. permutations in regions of high variability) while others will have a small impact. Our results in Figure 3 indicate that the variability is indeed systematically reduced when using the spatial initialization (a ratio above 1 corresponds to reduced variability). However, due to noise, approximation in the atlas, and general model mismatch, not all variability can be compensated by the graph alignment and some interindividual subject will remain and may still be heterogeneous. In brief, the maps representing the permutation rate in Figure 5 should not be interpreted as interindividual variability as the relationship between the two is not trivial. Finally, our approach could be optimized to promote not only neighborhood permutations, but also favor regions of high variability susceptible to correction by permutation.

In summary, our results show that brain network alignment increases the inter-subject similarity of structural connectivity, considered through the structural connectome. Notice that the underlying uncertainty might be due to different factors, including the group-wise parcellation misalignment, and the subjects specific differences. A comparison of the registration error in the groupwise parcellation and the graph matching results might decouple the source of uncertainty. In the context of healthy cohorts, the reduction in variability is advantageous as we expect all subjects to have normal, and therefore similar, connectivity. Extending these results to pathology is challenging as the inter-subject variance may be a feature of the disease and should not be corrected explicitly. A possible solution would be to apply network alignment using a different alignment targets per class, as is common in registration. Finally, our results considered only structural connectivity and may not generalize to functional connectivity or scenarios where both are considered simultaneously, e.g. structure-function mapping. In such cases, the alignment strategy should be modified to consider all available data to ensure a consistent permutation across the different networks.

## 5 | CONCLUSION

Considering a set of structural connectivity matrices of healthy subjects, this study shows how the cross-subjects variability can be decreased by applying graph matching techniques. We compare cross-subjects similarity using

different initialization of the FAQ graph matching algorithm. The improvement is significant when the algorithm is initialized using the identity or spatial structure of the regions and when the number of parcels increases. By looking at the results, the suggested matching are mostly between neighbours, showing local misalignments due to group-wise parcellation strategy. These preliminary results highlight the potential of using graph matching prior to group connectivity studies. Applying graph matching between subjects using graph matching technique improve the cross-subjects similarity.

## references

- Bassett, D. S. and Bullmore, E. T. (2017) Small-world brain networks revisited. *The Neuroscientist*, **23**, 499–516. PMID: 27655008.
- Becker, C. O., Pequito, S., Pappas, G. J., Miller, M. B., Grafton, S. T., Bassett, D. S. and Preciado, V. M. (2018) Spectral mapping of brain functional connectivity from diffusion imaging. *Scientific reports*, **8**, 1–15.
- Calissano, A., Feragen, A. and Vantini, S. (2022) Graph-valued regression: Prediction of unlabelled networks in a non-euclidean graph space. *Journal of Multivariate Analysis*, **190**, 104950.
- (2023) Populations of unlabelled networks: Graph space geometry and generalized geodesic principal components. *Biometrika*, asad024.
- Chamberland, M., Girard, G., Bernier, M., Fortin, D., Descoteaux, M. and Whittingstall, K. (2017) On the origin of individual functional connectivity variability: the role of white matter architecture. *Brain connectivity*, **7**, 491–503.
- Conte, D., Foggia, P., Sansone, C. and Vento, M. (2004) Thirty years of graph matching in pattern recognition. *International journal of pattern recognition and artificial intelligence*, **18**, 265–298.
- Deslauriers-Gauthier, S., Zucchelli, M., Frigo, M. and Deriche, R. (2020) A unified framework for multimodal structure–function mapping based on eigenmodes. *Medical Image Analysis*, **66**, 101799.
- Donnat, C. and Holmes, S. (2018) Tracking network dynamics: A survey using graph distances. *The Annals of Applied Statistics*, **12**, 971–1012.
- Emmert-Streib, F., Dehmer, M. and Shi, Y. (2016) Fifty years of graph matching, network alignment and network comparison. *Information Sciences*, **346**, 180–197.
- Fischl, B. (2012) FreeSurfer. *Neuroimage*, **62**, 774–781.
- Frigo, M., Cruciani, E., Coudert, D., Deriche, R., Natale, E. and Deslauriers-Gauthier, S. (2021) Network alignment and similarity reveal atlas-based topological differences in structural connectomes. *Network Neuroscience*, **5**, 711–733.
- Frost, M. A., Esposito, F. and Goebel, R. (2014) Improved correspondence of resting-state networks after macroanatomical alignment. *Human brain mapping*, **35**, 673–682.
- Gagniac, P. A. (2017) *Markov chains: from theory to implementation and experimentation*. John Wiley & Sons.
- Gallardo, G., Gayraud, N. T., Deriche, R., Clerc, M., Deslauriers-Gauthier, S. and Wassermann, D. (2018a) Solving the cross-subject parcel matching problem using optimal transport. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 836–843. Springer.
- Gallardo, G., Wells, W., Deriche, R. and Wassermann, D. (2018b) Groupwise structural parcellation of the whole cortex: A logistic random effects model based approach. **170**, 307–320.
- Ginestet, C. E., Li, J., Balachandran, P., Rosenberg, S. and Kolaczyk, E. D. (2017) Hypothesis testing for network data in functional neuroimaging. *The Annals of Applied Statistics*, **11**, 725–750.

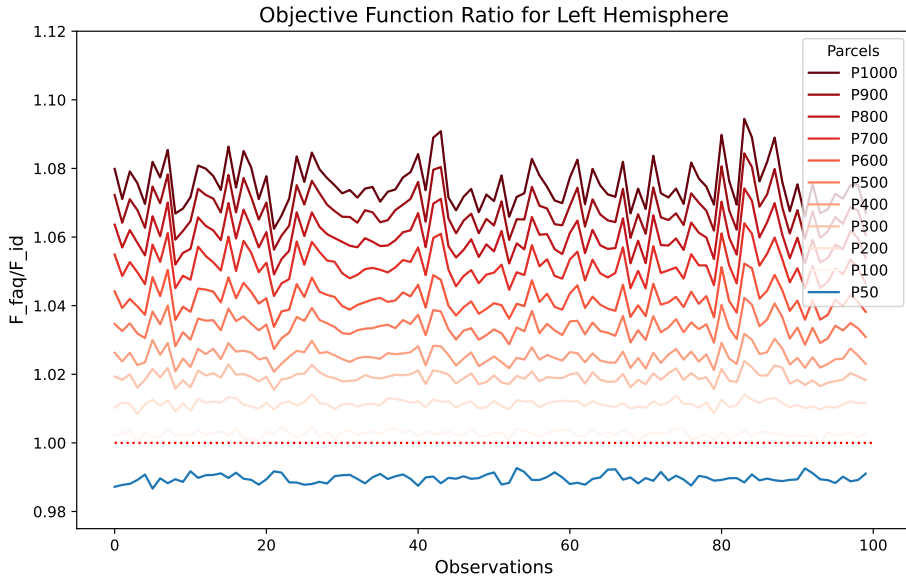
- Glaser, J. I. and Kording, K. P. (2016) The development and analysis of integrated neuroscience data. *Frontiers in computational neuroscience*, **11**.
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C., Jenkinson, M. et al. (2013) The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage*, **80**, 105–124.
- Jain, B. J. and Obermayer, K. (2009) Structure spaces. *Journal of Machine Learning Research*, **10**, 2667–2714.
- (2010) Large sample statistics in the domain of graphs. In *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*, 690–697. Springer.
- Jeurissen, B., Tournier, J.-D., Dhollander, T., Connelly, A. and Sijbers, J. (2014) Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *Neuroimage*, **103**, 411–426.
- Kolaczyk, E. D., Lin, L., Rosenberg, S., Walters, J., Xu, J. et al. (2020) Averages of unlabeled networks: Geometric characterization and asymptotic behavior. *The Annals of Statistics*, **48**, 514–538.
- Mars, R. B., Verhagen, L., Gladwin, T. E., Neubert, F.-X., Sallet, J. and Rushworth, M. F. (2016) Comparing brains by matching connectivity profiles. *Neuroscience & Biobehavioral Reviews*, **60**, 90–97.
- Mueller, S., Wang, D., Fox, M. D., Yeo, B. T., Sepulcre, J., Sabuncu, M. R., Shafee, R., Lu, J. and Liu, H. (2013) Individual variability in functional connectivity architecture of the human brain. *Neuron*, **77**, 586–595.
- Pandit, A. S., Expert, P., Lambiotte, R., Bonnelle, V., Leech, R., Turkheimer, F. E. and Sharp, D. J. (2013) Traumatic brain injury impairs small-world topology. *Neurology*, **80**, 1826–1833.
- Preti, M. G., Bolton, T. A. and Van De Ville, D. (2017) The dynamic functional connectome: State-of-the-art and perspectives. *NeuroImage*, **160**, 41–54. Functional Architecture of the Brain.
- Rasero, J., Pellicoro, M., Angelini, L., Cortes, J. M., Marinazzo, D. and Stramaglia, S. (2017) Consensus clustering approach to group brain connectivity matrices. *Network Neuroscience*, **1**, 242–253.
- Ravindra, V., Nassar, H., Gleich, D. F. and Grama, A. (2020) Rigid graph alignment. In *Complex Networks and Their Applications VIII: Volume 1 Proceedings of the Eighth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2019 8*, 621–632. Springer.
- Severn, K. E., Dryden, I. L. and Preston, S. P. (2020) Non-parametric regression for networks. *arXiv preprint arXiv:2010.00050*.
- Sinkhorn, R. and Knopp, P. (1967) Concerning nonnegative matrices and doubly stochastic matrices. *Pacific Journal of Mathematics*, **21**, 343–348.
- Smith, R. E., Tournier, J.-D., Calamante, F. and Connelly, A. (2012) Anatomically-constrained tractography: improved diffusion MRI streamlines tractography through effective use of anatomical information. *Neuroimage*, **62**, 1924–1938.
- Sporns, O., Tononi, G. and Kötter, R. (2005) The human connectome: a structural description of the human brain. *PLoS computational biology*, **1**, e42.
- Takerkart, S., Auzias, G., Thirion, B. and Ralaivola, L. (2014) Graph-based inter-subject pattern analysis of fmri data. *PLoS one*, **9**, e104586.
- Tate, M. C., Herbet, G., Moritz-Gasser, S., Tate, J. E. and Duffau, H. (2014) Probabilistic map of critical functional regions of the human cerebral cortex: Broca's area revisited. *Brain : a journal of neurology*, **137**, 2773–2782.
- Thompson, P. M., Cannon, T. D., Narr, K. L., van Erp, T., Poutanen, V. P., Huttunen, M., Lönnqvist, J., Standertskjöld-Nordenstam, C. G., Kaprio, J., K., M., Dail, R., Zoumalan, C. I. and Toga, A. W. (2001) Genetic influences on brain structure. *Nature neuroscience*, **4**, 1253–1258.

- Tian, D., Zeng, Z., Sun, X., Tong, Q., Li, H., He, H., Gao, J.-H., He, Y. and Xia, M. (2022) A deep learning-based multisite neuroimage harmonization framework established with a traveling-subject dataset. *NeuroImage*, **257**, 119297.
- Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H. and Connelly, A. (2019) MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, **202**, 116137.
- Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S. W. et al. (2012) The human connectome project: a data acquisition perspective. *NeuroImage*, **62**, 2222–2231.
- Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B. and Tzourio-Mazoyer, N. (2006) Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage*, **30**, 1414–1432.
- Vogelstein, J. T., Conroy, J. M., Lyzinski, V., Podrazik, L. J., Kratzer, S. G., Harley, E. T., Fishkind, D. E., Vogelstein, R. J. and Priebe, C. E. (2015) Fast approximate quadratic programming for graph matching. **10**, e0121002.
- Weinstein, S. M., Vandekar, S. N., Adebimpe, A., Tapera, T. M., Robert-Fitzgerald, T., Gur, R. C., Gur, R. E., Raznahan, A., Satterthwaite, T. D., Alexander-Bloch, A. F. et al. (2021) A simple permutation-based test of intermodal correspondence. *Human brain mapping*, **42**, 5175–5187.
- Winding, M., Pedigo, B. D., Barnes, C. L., Patsolic, H. G., Park, Y., Kazimiers, T., Fushiki, A., Andrade, I. V., Khandelwal, A., Valdes-Aleman, J., Li, F., Randel, N., Barsotti, E., Correia, A., Fetter, R. D., Hartenstein, V., Priebe, C. E., Vogelstein, J. T., Cardona, A. and Zlatic, M. (2023) The connectome of an insect brain. *Science*, **379**, eadd9330.

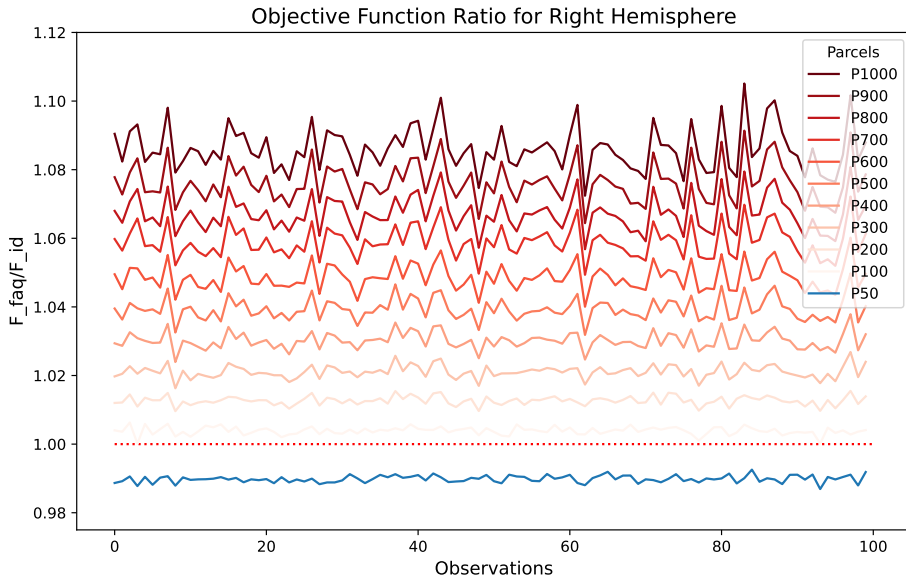
## 6 | APPENDIX

We decompose the results across subjects, to visualize if there is a high variability across subject in the graph matching results. In Figure 7 and 8, we look at all the permutations obtained across subjects, to check if there are any subject anomalies. As the subjects are control subjects, we expect a consistent behaviour for each parcellation. As shown in the figure, the pattern is indeed consistent across subjects and by increasing the number of regions the FAQ suggests a matching procedure rather than leaving the nodes labels stable.





**FIGURE 7** Left Hemisphere: Visualization of the ration between objective functions for each parcellations and each subjects.



**FIGURE 8** Right Hemisphere: Visualization of the ration between objective functions for each parcellations and each subjects.