Rigid Graph Alignment

Vikram Ravindra(✉), Huda Nassar, David F. Gleich, and Ananth Grama

Purdue University, West Lafayette, USA
ravindrv@purdue.edu

Abstract. An increasingly important class of networks is derived from physical systems that have a spatial basis. Specifically, nodes in the network have spatial coordinates associated with them, and conserved edges in two networks being aligned have correlated distance measures. An example of such a network is the human brain connectome – a network of co-activity of different regions of the brain, as observed in a functional MRI (fMRI). Here, the problem of identifying conserved patterns corresponds to the alignment of connectomes. In this context, one may structurally align the brains through co-registration to a common coordinate system. Alternately, one may align the networks, ignoring the structural basis of co-activity. In this paper, we formulate a novel problem – *rigid graph alignment*, which simultaneously aligns the network, as well as the underlying structure. We formally specify the problem and present a method based on expectation maximization, which alternately aligns the network and the structure via rigid body transformations. We demonstrate that our method significantly improves the quality of network alignment in synthetic graphs. We also apply rigid graph alignment to functional brain networks derived from 20 subjects drawn from the Human Connectome Project (HCP), and show over a two-fold increase in quality of alignment. Our results are broadly applicable to other applications and abstracted networks that can be embedded in metric spaces – e.g., through spectral embeddings.

Keywords: Graph alignment · Structural alignment

1 Introduction and Motivation

Graphs are commonly used to represent variety of complex systems such as financial transactions, social and communication networks, chemical reaction pathways, and biomolecular interactions. In an important subclass of such systems, the relative positions of nodes are fixed. While the network, or parts thereof, can be transformed from their preferred (relaxed) positions, their "potential energy" increases in the process, motivating the nodes to return to their relaxed state, i.e., the layout with the minimum energy. A natural way to think about such graphs is to visualize the edges as springs of different lengths. We refer to such graphs as *rigid graphs*.

H. Cherifi et al. (Eds.): COMPLEX NETWORKS 2019, SCI 881, pp. 621–632, 2020.
https://doi.org/10.1007/978-3-030-36687-2_52

In this paper, we focus on the problem of aligning a pair of given rigid graphs. Informally stated, the goal is to find node/edge correspondences across a pair of rigid graphs; e.g., across two given biomolecules or two brain connectomes. Traditional approaches to related problems use either structural embeddings that consider edge lengths (i.e., identifying structural transformations – translation, rotation, dilation, that maximize overlap of graphs) or connectivity (i.e., identifying node label permutations that maximize overlap in adjacency matrices) to find correspondences between the two graphs, but not both. However, in the case of rigid graphs, both of these sources of information are rich and complementary. We demonstrate that integrating these two sources into a single framework leads to markedly better correspondences. We formulate the problem of *rigid graph alignment*, and propose a solution technique for this problem in which the rigidity of nodes as well as their connectivity are used to infer significantly better alignments that either structural or topological alignment alone.

We motivate rigid graph alignment in two distinct application contexts. The first one uses a growing and important database of 3D brain images from the Human Connectome Project. One of the major goals of functional brain studies is to understand spontaneous firings of neurons in absence of any stimulus – the "resting-state" activity. Resting functional state Magnetic Resonance Image (rs-fMRI) is useful in creation of a rough baseline network. A more interesting reason to study these images is that they record cognitive processes of individuals in absence of any activity. It is hypothesized that some of these cognitive processes are unique to individuals [11]. Thus, in principle, it must be possible to uniquely identify an individual from a population, using just their resting state functional brain image. This can be done by aligning the images structurally (through registration), topologically (through network alignment) or, as we demonstrate in this paper, using both structural and topological information, yielding superior results. We further motivate rigid graph alignment in the context of aligning large biomolecules drawn from existing databases of known drug molecules [3]. This problem is of significance, since it underlies biomolecular interactions, identification of drug targets, and design of drug molecules. Specifically, binding pockets in molecules correspond to local atomic structures (active sites) – two molecules that share active sites are likely to bind with the same ligand. This process of matching active sites can be formulated as a rigid graph alignment problem. This is an active area of research, with commonly used methods relying on traditional graph or structural alignment, bootstrapped using techniques such as geometric hashing. These methods do not utilize constraints imposed by the bond lengths and bond angles into the graph alignment process, or the global structure of the molecules. Our novel formulation of rigid graph alignment factors the bond structure through rigid graph alignment, while simultaneously accounting for bond characteristics, using rigid body transformations.

2 Problem Formulation

2.1 Problem Definition

We define the rigid graph alignment problem by first reviewing existing graph and structure alignment formulations, and use these to motivate our new problem.

Network Alignment Review. The literature on network alignment is vast – precluding a comprehensive review. As a representative of a broad class of techniques, we discuss the quadratic programming formulation of a network alignment objective [1]. Let $A = G(V_A, E_A)$ and $B = G(V_B, E_B)$ be two undirected graphs that we aim to align. For simplicity, we set $|V_A| = |V_B| = n$. Let $\mathbf{L} \in \mathbb{R}^{n \times n}$ be a matrix that denotes prior knowledge on likelihood of alignment (or known similarity) between vertices V_A and V_B. If $\mathbf{L}_{ij} > 0$, prior knowledge implies that node i in V_A potentially aligns with node j in V_B, otherwise, node i in V_A cannot align with node j in V_B. The aim of the alignment problem is to find a matching M from V_A to V_B using only weights of $\mathbf{L}$ and adjacency information. We define $\mathbf{X}$ to be a matrix with the same dimensions as L to encode matching $\mathcal{M}$:

$$\mathbf{X}_{ij} = \begin{cases} 1 & \text{if } i \in V_A \text{ is matched with } j \in V_B \text{ under } \mathcal{M} \\ 0 & \text{otherwise} \end{cases}$$

Under such a mapping, we say that an edge $(u, v) \in E_A$ *overlaps* with an edge $(u', v') \in E_B$ if (u, u') and (v, v') are in $\mathcal{M}$.

The goal in this graph alignment formulation is to find a matching $\mathcal{M}$ that maximizes a linear combination of *overlap* and matching *weight*. Note that, if $\mathbf{X}$ is constrained by a set of matching constraints,

$$\forall i \in [|V_A|], \quad \textstyle\sum_j \mathbf{X}_{ij} \leq 1, \qquad \forall j \in [|V_B|], \quad \textstyle\sum_i \mathbf{X}_{ij} \leq 1, \tag{1}$$

then

$$A \bullet \mathbf{X}B\mathbf{X}^T = \textstyle\sum_{ij} A_{ij}(\mathbf{X}B\mathbf{X}^T)_{ij} = \text{overlap of matching } \mathcal{M} \tag{2}$$

where we use the adjacency matrices A and B for the graphs and $\bullet$ to represent a matrix inner product. The corresponding quadratic program is given by:

$$\begin{aligned} \max_{\mathbf{X}} \quad & \alpha\mathbf{L} \bullet \mathbf{X} + \beta A \bullet \mathbf{X}B\mathbf{X}^T \\ \text{s.t.} \quad & \textstyle\sum_i \mathbf{X}_{ij} \leq 1 \quad \forall j = 1 \ldots |V_B|, \\ & \textstyle\sum_j \mathbf{X}_{ij} \leq 1 \quad \forall i = 1 \ldots |V_A|, \quad \mathbf{X}_{ij} \in \{0, 1\} \end{aligned} \tag{3}$$

Here, α, β are non-negative constants that allow for tradeoff between matching weights from the prior and the number of overlapping edges.

Review of Structural Alignment. We now summarize the Orthogonal Procrustes Problem, which aims to compute structural alignments. The problem, originally solved by Schönemann [19] aims to find an orthogonal transformation that

reduces the distance between two matrices in the Frobenius norm. It is often used to find rigid body transformations that describe relationships between two objects. Formally, for any two matrices Y and Z, the problem minimizes:

$$\begin{aligned} \min_{\mathbf{\Omega}} \quad & ||Y - Z\mathbf{\Omega}||_F^2 \\ \text{s.t.} \quad & \mathbf{\Omega}^T\mathbf{\Omega} = I \end{aligned} \tag{4}$$

In our setting, the matrices Y and Z correspond to node coordinates $C_A, C_B \in \mathbb{R}^{n\times d}$. Our objective is to measure the orthogonal transformation between the sets of corresponding nodes of the two graphs. In other words, we aim to find rotation $\hat{\mathbf{R}} \in \mathbb{R}^{d\times d}$ and translation $\hat{t} \in \mathbb{R}^{1\times d}$ such that

$$C_A = C_B\hat{\mathbf{R}} + \mathbb{1}\hat{t} \tag{5}$$

where, $\hat{\mathbf{R}}^T\hat{\mathbf{R}} = I$ and $\mathbb{1}$ is the ones vector of length n. For convenience, we pad C_A and C_B by a vector of ones (for convenience in notation, we don't explicitly show this detail), and combine the translation vector and the rotation matrix into $\mathbf{\Omega}$ as

$$\mathbf{\Omega} = \begin{bmatrix} \hat{\mathbf{R}} & \mathbf{0} \\ \hat{t} & 1 \end{bmatrix} \tag{6}$$

We rephrase the problem as:

$$\min_{\Omega} \quad ||C_A - C_B\mathbf{\Omega}||_F^2 \tag{7}$$

The algorithm due to Kabsch [13] was among the first solutions aimed at solving the problem for 2D and 3D coordinates. We use a related SVD-based method, due to Sabata et al. [18], since it is shown to stable by Eggert et al. [8].

$$\mu_A = \frac{1}{n}\sum_{i=1}^{n} C_{A_i} \qquad \overline{C}_A = C_A - \mathbb{1}\mu_A \tag{8}$$

$$\mu_B = \frac{1}{n}\sum_{i=1}^{n} C_{B_i} \qquad \overline{C}_B = C_B - \mathbb{1}\mu_B \tag{9}$$

Here, C_{A_i} and C_{B_i} refer to the coordinates of the i-th nodes of graphs A and B. Hence, $\mu_A, \mu_B \in \mathbb{R}^{n\times 1}$. Define $\mathbf{H} = \overline{C}_A^T\overline{C}_B$, then the estimated rotation matrix is $\hat{\mathbf{R}}$ is given by:

$$\hat{\mathbf{R}} = VU^T, \tag{10}$$

where V and U are orthonormal matrices that are obtained from SVD of $\mathbf{H}$.

The optimal translation $\hat{\mathbf{t}}$ is defined by:

$$\hat{\mathbf{t}} = \mu_A - \hat{\mathbf{R}}\,\mu_B \tag{11}$$

We note that we can estimate scaling between the two graphs using more generalized formulations of the orthogonal Procrustes problems. Instead, we assume

that the coordinates are drawn from the same units (say meters). This assumption also circumvents the problem of having potentially different coordinate systems.

Rigid Graph Alignment. We now define our problem of rigid graph alignment. Let $C_A \in \mathbb{R}^{n\times d}$ and $C_B \in \mathbb{R}^{n\times d}$ represent coordinates of vertices from V_A and V_B, respectively. Combining the network alignment objective function and rigidity metric, as described in Eqs. 3 and 7, our objective function for *rigid graph alignment* can be written as:

$$\begin{aligned} \mathbf{F} = \max_{\mathbf{X},\mathbf{\Omega}} \quad & \alpha\mathbf{L}\bullet\mathbf{X} + \beta A\bullet\mathbf{X}B\mathbf{X}^T - \gamma||C_A - \mathbf{X}C_B\mathbf{\Omega}||_F^2 \\ \text{s.t.} \quad & \textstyle\sum_i \mathbf{X}_{ij} \le 1 \quad \forall j = 1\ldots|V_B|, \\ & \textstyle\sum_j \mathbf{X}_{ij} \le 1 \quad \forall i = 1\ldots|V_A|, \quad \mathbf{X}_{ij} \in \{0,1\} \end{aligned} \tag{12}$$

Here, we write $\mathbf{X}C_B$ to denote the conformally permuted set of coordinates of graph B after alignment. The first term in this equation, $\mathbf{L}\bullet\mathbf{X}$, corresponds to the consistency between the prior $\mathbf{L}$ and the mapping of vertices across the two graphs $\mathbf{X}$. The second term, $A\bullet\mathbf{X}B\mathbf{X}^T$, corresponds to the alignment of the two networks, and the third term, $||C_A-\mathbf{X}C_B\mathbf{\Omega}||_F^2$, to the structural (mis)alignment. In *rigid graph alignment*, the prior L is driven by spatial constraints. The weights α, β and γ are parameters for the user to adjust relative importance of the prior, graph matching, and structural alignment. It can be seen that when $\alpha = 1, \beta = 0, \gamma = 0$, the problem reduces to the maximum matching problem; when $\alpha = 0, \beta = 1, \gamma = 0$, it is the solution to the problem of maximizing overlap; and when $\alpha = 0, \beta = 0, \gamma = 1$, the problem reduces to rigid body transformation.

The structural alignment error can be minimized by solving for $\mathbf{\Omega}$ using generalized Procrustes method on the ordered set of vertices. To find a good network alignment solution, we need to find the optimal $\mathbf{X}$. For ease of analysis, we view $\mathbf{X}$ as a permutation matrix. Then, the structural error term can be expressed as $C_A - \mathbf{X}C_B\mathbf{\Omega}_F^2 = C_A\bullet C_A + C_B\mathbf{\Omega}\bullet C_B\mathbf{\Omega} - 2C_A\mathbf{\Omega}^TC_B^T\bullet\mathbf{X}$

For a given transformation matrix $\mathbf{\Omega}$, we can rewrite the objective function as

$$\begin{aligned} \mathbf{F} = \max_{\mathbf{X}} \quad & \alpha\underbrace{\mathbf{L}\bullet\mathbf{X}}_{\text{Update Prior}} + \beta\overbrace{A\bullet\mathbf{X}B\mathbf{X}^T}^{\text{Network Alignment}} + \gamma\underbrace{C_A\mathbf{\Omega}^TC_B^T\bullet\mathbf{X}}_{\text{Structural Alignment}} \\ \text{s.t.} \quad & \textstyle\sum_i \mathbf{X}_{ij} = 1 \quad \forall j = 1\ldots|V_B|, \\ & \textstyle\sum_j \mathbf{X}_{ij} = 1 \quad \forall i = 1\ldots|V_A|, \quad \mathbf{X}_{ij} \in \{0,1\} \end{aligned} \tag{13}$$

This form of the objective function suggests that the optimal $\mathbf{X}$ is one that maximizes the network alignment and structural alignment, while respecting the prior. Furthermore, it suggests that the prior $\mathbf{L}$ should be proportional to the similarity between the coordinates, i.e., $C_A\mathbf{\Omega}^TC_B^T$. However, this definition poses the potential problem of a dense prior, which can significantly increase the runtime of network alignment. To circumvent this issue, we constrain the number

of non-zeros in the prior using a distance measure or number of neighbours. One such simple constraint is given by:

$$\mathbf{L}_{i,j} = \begin{cases} \exp(-\left\|C_{A_i} - C_{B_j}\right\|_2^2) & \left\|C_{A_i} - C_{B_j}\right\|_2^2 \leq d_i^k \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

The intuition for an rigid graph alignment technique is as follows: let $\mathbf{P}_{ii'}$ denote the likelihood of matching node i of graph A with i' of graph B. Since we use structural conformity as prior information to guide the network alignment, we say that $\mathbf{P}_{ii'} \propto \exp(-\left\|C_{A_i} - C_{B_{i'}}\right\|_2^2)$. We can similarly define $\mathbf{P}_{jj'}$. Assuming that the probabilities are independent, the expected edge overlap can be written as the product:

$$\mathbb{E}_o = \sum_{i,j,i',j'} \mathbf{P}_{ii'}\mathbf{P}_{jj'}A_{ij}B_{i'j'} \tag{15}$$

$$= \sum_{i,j,i',j'} \exp(-\left\|C_{A_i} - C_{B_{i'}}\right\|_2^2)\exp(-\left\|C_{A_j} - C_{B_{j'}}\right\|_2^2)A_{ij}B_{i'j'} \tag{16}$$

We can see that the expected edge overlap increases as matched nodes move closer to each other. Hence, increased structural alignment results in increased edge overlap. In an analogous manner, accurate matches lead to more accurate transformation matrices, which minimizes the residual error after structural alignment. Premised on the observation that the two objectives depend on, and reinforce each other, we propose an algorithm that optimizes the two terms, alternately.

2.2 Rigid Graph Matching Algorithm

Our approach to rigid graph alignment splits the problem into two tasks and iterates to convergence: (i) align the graphs restricting the prior to pairs of nodes between graphs that are in spatial proximity (using definitions for $\mathbf{L}$ mentioned earlier) – i.e., maximizing the second term of the objective function, and (ii) align the coordinates using the current estimate of the alignment; i.e., maximizing the third term of the objective function. This requires an initial alignment to begin, which we discuss at the end of the section.

The vast majority of network alignment methods take as input α, β, L, A, B (or equivalent inputs). Hence, our goal in step (i) is to leverage the coordinates to estimate a matrix L that constrains the set of alignments considered by the network alignment component. We devise a routine, *get_prior* that creates the matrix $\mathbf{L}$ based on the distance between vertices (using the current transformation Ω). Our implementation of this routine assigns every node i in A with k of its nearest neighbours as possible matches. Each candidate node is assigned a weight inversely proportional to the distance between that node and node i (Eq. 14). In terms of the objective function, note that this choice of L ensures that the term $||C_A - \mathbf{X}C_B\mathbf{\Omega}||_F^2$ stays approximately the same after we use a network alignment method to optimize the permutation.

Step (ii) involves the use of a procedure for structural alignment. We used an SVD-based method to compute orthogonal transformation for structural alignment, as described in Sect. 2.1.

To recap, in step (i), we maximize the term $\alpha \mathbf{L} \bullet \mathbf{X} + \beta A \bullet \mathbf{X} B \mathbf{X}^T$ with the term $\gamma ||C_A - \mathbf{X} C_B \mathbf{\Omega}||_F^2$ approximately fixed, whereas in step (ii), we maximize $-\gamma ||C_A - \mathbf{X} C_B \mathbf{\Omega}||_F^2$ with the graph alignment term fixed.

The resulting *rigid graph alignment* procedure is given in Algorithm 1.

Algorithm 1. Rigid Graph Alignment

Input: Graphs $A(V_A, E_A)$ and $B(V_B, E_B)$, Coordinates C_A and C_B, α, β, γ
Output: Aligned graphs A and B
repeat
  $\mathbf{L} = \text{get\_prior}(C_A, C_B)$
  $\mathbf{X} = \text{align}(A, B, \mathbf{L})$
  $B = \mathbf{X} B \mathbf{X}^T$
  $\mathbf{\Omega} = \text{transform\_coordinates}(C_A, C_B, \mathbf{X})$
  $C_B = \mathbf{X} C_B \mathbf{\Omega}$
until converged

The process is continued until convergence is achieved, in terms of preferred metrics such as edge/node overlap or in the solution to the Orthogonal Procrustes problem. The algorithm proposed (really a meta-algorithm) works with any pair of aligners (graph and structural) that use a prior. In our experiments, we present results where we used *netalignmbp* [1] as the graph aligner and the SVD-based structural aligner.

In the first iteration, when we have no prior knowledge of any correspondence between nodes, we use the following approach to bootstrap Algorithm 1: we compute average pairwise distance for all nodes and populate the prior matrix with nodes that have similar distance profiles. We draw histograms for the distances of each node in both graphs. Then, we use correlation to measure the similarity of distance profiles of pairs of nodes across the two graphs. The prior matrix $\mathbf{L}$ is populated on the basis of this measure. We note that distance profiles are invariant to rotations and translations. Furthermore, correlation measures such as Pearson Correlation are invariant to scaling. In practice, we find that this heuristic works well both in the synthetic datasets, as well as the HCP dataset.

2.3 Analyzing the Human Functional Connectome

We present experimental results in the context of an important application in the analyses of brain connectomes. We demonstrate that our methods are capable of significant improvement in alignment quality over state of the art graph alignment techniques that do not consider structural rigidity of the graph.

Dataset. We construct human brain connectomes from the Young Adult Database. The Young Adult Database was created as part of the Washington

University-University of Minnesota project of the Human Connectome Project (HCP) consortium, described in van Essen et al. [10]. We use resting-state fMRIs from a subset of 20 subjects. The data for each subject is collected in two sessions, which are separated by a few days. Each session lasted 30 min: 15 min of left-to-right (LR) encoding, followed by 15 min of right-to-left (RL) encoding. Each voxel is isotropic, with dimensions 2 mm × 2 mm × 2 mm, and images were acquired once every 720 ms, as explained in Smith et al. [21].

Preprocessing Steps. The preprocessing steps remove spatial and temporal artifacts from the images. Head motion of subjects during acquisition is inevitable, and a potential source of errors. We use the motion correction tool from FMRI Standard Library (FSL), called Motion Correction FSL's Linear Image Registration Tool (MCFLIRT) [12]), and register each (volumetric) image in a session to the first time slice. Motion correction ensures consistent labeling of voxels between images in a session. This is followed by skull stripping, using FSL's Brain Extraction Tool (BET) [22]. The images are then resampled to voxels of size $4mm$. This resampling is done to create networks whose dimensions are accessible to network alignment using state of the art methods. Non-brain voxels are masked out, along with voxels of low variance. The remaining, relevant voxels are vectorized, to create a $voxel \times time$ matrix. It has been shown that spontaneous firings observed in resting state functional MRI is best captured in low frequency fluctuations < 0.1 Hz (Murphy et al. [15]), which is why we use a bandpass filter with limits of 0.001 Hz to 0.08 Hz.

A $voxel \times voxel$ similarity matrix is created by computing Pearson correlation between each pair of time series. We retain the top 5 percentile of correlation values and construct networks with voxels as nodes and high correlation valued edges between nodes. Prior to the creation of the similarity matrix, we regress out the global signal. Since we only use strongly positive values in subsequent steps we sidestep the issue of artificially induced negative correlations.

Rigid Graph Alignment Yields Higher Edge Overlap. In our first set of results, we show that Rigid Graph Alignment improves on a commonly used network alignment metric – edge overlap. For two adjacency matrices A and B, edge overlap is defined as $A \bullet \mathbf{X} B \mathbf{X}^T$. We run Rigid Graph Alignment (Algorithm 1) for intra-subject (same subject across two sessions) and inter-subject (across subjects) analysis on 20 subjects, with equal weights given to the prior (α), edge overlap (β), and structural alignment (γ), as discussed in Sect. 2.1 for a maximum of 20 iterations and a convergence threshold of 0.1% in edge overlap. We find that the edge overlap at the end of first iteration is $20.18 \pm 4.2\%$, whereas the edge overlap after *rigid network alignment* is $53.05 \pm 12.5\%$. The improvement of edge overlap is due to increasingly accurate priors in the later iterations of the algorithm. The increase per iteration in typical intra-subject alignment is shown in Fig. 1. We observe that the scores are largely stagnant in the first few iterations, which is attributed to the fact that the initial prior is not as informative as priors in subsequent iterations. As stated before, to bootstrap the algorithm, we use similarity of distance profiles to populate the initial prior matrix because the "drawings" of graphs could look very different. In the later

iterations, successive transformations result in corresponding nodes of the two graphs being placed close together.

We also characterize the statistical significance of our performance improvement. To do this, we transform the coordinates by random transformation matrices (i.e., we randomly reorient the brains before alignment) in 100 trials. In each case, we found that the edge overlap by rigid graph alignment was higher than edge overlap by regular graph alignment. This implies that our performance improvements are significant and robust.

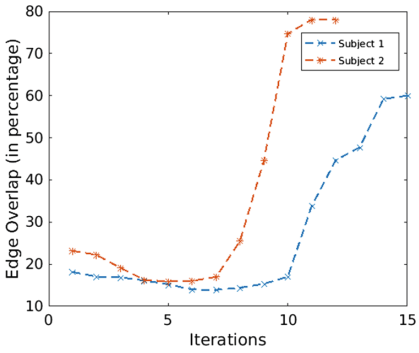
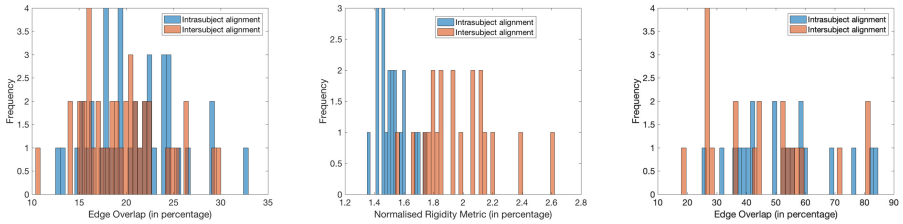


Fig. 1. Increase in edge overlap score while aligning functional brain networks of two subjects with themselves (i.e., across sessions) in resting state functional MRI. These results show significant improvement in alignment quality over state of the art alignment techniques. Note that the alignment results after the first iteration correspond to that of state of the art network alignment technique. Subsequent iterations demonstrate improvements from our method.

Residual Error in Structural Transformation as a Metric for Network Alignment. The problem of identifiability of the connectome, or the so-called *brain fingerprint* involves finding patterns in brain images (either structural or functional) that are unique to an individual. In principle, two networks belonging to the same subject should align "better" than two networks belonging to different subjects. In our context, we need to devise a scoring measure for alignments that can suitably distinguish intra-subject and inter-subject alignment.

Figure 2a shows the histogram of edge overlap at the end of the first iteration. The edge overlap for intra-subject networks is $20.73 \pm 4.45\%$, whereas for inter-subject networks, it is $19.62 \pm 4.28\%$. This suggests that edge overlap is not a strong discriminatory measure for the brain signature. In fact, the improvement in edge overlap as a consequence of our *rigid graph alignment* algorithm is observed both for inter-subject and intra-subject alignments, as can be seen in Fig. 2c. In such applications, where the position of the vertices are of importance, we show that the residual error in structural transformation, which we call "rigidity metric" (Eq. 7) is a better indicator of quality of alignment.

Figure 2b shows the histogram of quality of inter-subject and intra-subject alignments. For intra-subject alignments, the *rigidity metric*, normalized by the



(a) (b) (c)

Fig. 2. Histogram of (a) edge overlap values, in both inter-subject and intra-subject alignments after first iteration. The similar overlap values show that the edge overlap is a not good metric for detecting brain fingerprints, (b) rigidity values for inter-subject and intra-subject alignments. The values are significantly better separated than edge overlap, suggesting that rigidity metric is better than edge overlap in the context of brain fingerprints, and (c) edge overlap, after rigid graph alignment. The overlap has increased in both intra-subject and inter-subject alignments

number of vertices was found to be $1.52 \pm 0.01\%$, whereas for inter-subject alignments, it was $1.99 \pm 0.025\%$. This demonstrates significantly higher distinguishability from the rigidity metric.

3 Related Literature

Network alignment has been an active area of research over several decades [9]. Early formulations of network alignment, in the form of exact graph matching (subgraph isomorphisms) in small chemical networks were analyzed by Sussenguth [23].

For larger graphs, inexact techniques are used for matching graphs, due to the high computational cost of isomorphism-based methods. These formulations largely fall into two classes – local aligners and global aligners. Local aligners (e.g., AlignNemo [6]) aim to find localized regions in graphs that align well. A localized network alignment approach was used by Lässig et al. [2] to align gene regulation networks in *E. coli*. The optimization criteria in this class of methods rewards local matches and does not penalize mismatches over entire graphs. Global aligners, on the other hand, use optimization functions that reward matches and penalize mismatches over the entire graph. In [20], Singh et al. describe IsoRank, a pairwise global network alignment method for protein interaction networks. IsoRank is based on the intuition that a vertex in the first network should be matched with a vertex in the second network if and only if the neighbors of the two vertices are also well matched.

Memory-efficient heuristic methods such as GRAAL [14] and GHOST [16] have been proposed, which require cubic runtime. Many alignment techniques incorporate prior knowledge of potential matches to restrict the search space of potential matches to a smaller subset of nodes, and to guide the search process. Examples of such algorithms include the previously described Isorank [20], where sequence similarity of proteins is often used as a prior for matching networks

of interacting proteins. A message passing algorithm *netalignmbp*, proposed by Bayati et al. [1] is based on making greedy decisions constrained by a given prior, assuming that there are no cycles in the graph. In this paper, we focus on global graph alignment techniques that incorporate prior knowledge to yield a one-to-one mapping of nodes. We use *netalignmbp* with an informative prior as our network alignment substrate. However, our algorithm can be used in conjunction with any network alignment technique that allows the use of a prior.

Structural Alignment. One of the commonly used approaches to *structural alignment* uses the *iterative closest point* (ICP) method of Besl et al. [4], and its variants [5]. The generic steps of ICP described in [17] for matching two structures involve selection of source points from both bodies. ICP then matches the sets of points and computes weights corresponding to the matches appropriately. Points that do not match are rejected and the transformation that minimizes error is computed. This process is repeated to convergence. Our algorithm can be thought of as loosely following this framework, except that we use network alignment as the process to find suitable matching. The transformation itself, as stated in Eq. 4, is called the Orthogonal Procrustes Problem [19]. Kabsch's algorithm [13] provides an efficient solution to this problem in three dimensions. In fact, the graph matching problem as stated by [24] can be thought of as a two-sided Procrustes problem (Conroy et al. [7]), who approximate graph matching when vertices of the two graphs can be parcellated into groups.

Acknowledgement. The authors are supported by the National Science Foundation grants CCF-1149756 and IIS-1546488.

References

1. Bayati, M., Gleich, D.F., Saberi, A., Wang, Y.: Message-passing algorithms for sparse network alignment. ACM Trans. Knowl. Discov. Data **7**(1), 3:1–3:31 (2013). http://doi.acm.org/10.1145/2435209.2435212
2. Berg, J., Lässig, M.: Local graph alignment and motif search in biological networks. Proc. Natl. Acad. Sci. **101**(41), 14689–14694 (2004). https://www.pnas.org/content/101/41/14689
3. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E.: The protein data bank. Nucleic Acids Res. **28**(1), 235–242 (2000)
4. Besl, P., McKay, H.: A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**, 239–256 (1992)
5. Bouaziz, S., Tagliasacchi, A., Pauly, M.: Sparse iterative closest point. In: Proceedings of the Eleventh Eurographics/ACMSIGGRAPH Symposium on Geometry Processing, SGP 2013, pp. 113–123. Eurographics Association, Aire-la-Ville (2013). http://dx.doi.org/10.1111/cgf.12178
6. Ciriello, G., Mina, M., Guzzi, P.H., Cannataro, M., Guerra, C.: Alignnemo: a local network alignment method to integrate homology and topology. PLOS ONE **7**(6), 1–14 (2012)
7. Conroy, B.R., Ramadge, P.J.: The grouped two-sided orthogonal procrustes problem. In: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3688–3691, May 2011

8. Eggert, D., Lorusso, A., Fisher, R.: Estimating 3-D rigid body transformations: a comparison of four major algorithms. Mach. Vis. Appl. **9**(5), 272–290 (1997)
9. Emmert-Streib, F., Dehmer, M., Shi, Y.: Fifty years of graph matching, network alignment and network comparison. Inf. Sci. **346**(C), 180–197 (2016). https://doi.org/10.1016/j.ins.2016.01.074
10. Essen, D.C.V., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K.: The WU-Minn human connectome project: an overview. NeuroImage **80**, 62–79 (2013)
11. Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T.: Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nature Neurosci. **18**(11), 1664–1671 (2015)
12. Jenkinson, M., Bannister, P., Brady, M., Smith, S.: Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage **17**(2), 825–841 (2002)
13. Kabsch, W.: A solution for the best rotation to relate two sets of vectors. Acta Crystallogr. Sect. A **32**(5), 922–923 (1976)
14. Kuchaiev, O., Milenković, T., Memišević, V., Hayes, W., Pržulj, N.: Topological network alignment uncovers biological function and phylogeny. J. Roy. Soc. Interface (2010). http://rsif.royalsocietypublishing.org/content/early/2010/03/24/rsif.2010.0063
15. Murphy, K., Birn, R.M., Handwerker, D.A., Jones, T.B., Bandettini, P.A.: The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? NeuroImage **44**(3), 893–905 (2009)
16. Patro, R., Kingsford, C.: Global network alignment using multiscale spectral signatures. Bioinformatics **28**(23), 3105–3114 (2012)
17. Rusinkiewicz, S., Levoy, M.: Efficient variants of the ICP algorithm. In: Proceedings of the 3DIM 2001, October 2001
18. Sabata, B., Aggarwal, J.: Estimation of motion from a pair of range images: a review. CVGIP: Image Underst. **54**(3), 309–324 (1991). http://www.sciencedirect.com/science/article/pii/104996609190032K
19. Schönemann, P.H.: A generalized solution of the orthogonal procrustes problem. Psychometrika **31**(1), 1–10 (1966)
20. Singh, R., Xu, J., Berger, B.: Pairwise global alignment of protein interaction networks by matching neighborhood topology. In: Proceedings of the 11th Annual International Conference on Research in Computational Molecular Biology, RECOMB 2007, pp. 16–31. Springer, Heidelberg (2007)
21. Smith, S.M., Beckmann, C.F., Andersson, J., Auerbach, E.J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D.A., Griffanti, L., Harms, M.P., Kelly, M., Laumann, T., Miller, K.L., Moeller, S., Petersen, S., Power, J., Salimi-Khorshidi, G., Snyder, A.Z., Vu, A.T., Woolrich, M.W., Xu, J., Yacoub, E., Uğurbil, K., Essen, D.C.V., Glasser, M.F.: Resting-state fMRI in the human connectome project. NeuroImage **80**, 144–168 (2013)
22. Smith, S.M.: Fast robust automated brain extraction. Hum. Brain Mapp. **17**(3), 143–155 (2002)
23. Sussenguth, E.H.: A graph-theoretic algorithm for matching chemical structures. J. Chem. Documentation **5**(1), 36–43 (1965)
24. Umeyama, S.: An eigendecomposition approach to weighted graph matching problems. IEEE Trans. Pattern Anal. Mach. Intell. **10**(5), 695–703 (1988)