Multi-Objective Graph Matching via Signal Filtering

Daniel L. Sussman*, Lin Li†
*Department of Mathematics and Statistics, Boston University
sussman@bu.edu
†Artificial Intelligence Technology and Systems, MIT Lincoln Laboratory
lin.li@ll.mit.edu

***Abstract*—In this white paper we propose a new method which exploits tools from graph signal processing to solve the graph matching problem, the problem of estimating the correspondence between the vertex sets of two graphs. We recast the graph matching problem as matching multiple similarity matrices where the similarities are computed between filtered signals unique to each node. Using appropriate graph filters, these similarity matrices can emphasize long or short range behavior and the method will implicitly search for similarities between the graphs and at multiple scales. Our method shows substantial improvements over standard methods which use the raw adjacency matrices, especially in low-information environments.**

I. Introduction

Many social, economic and biological systems can be better understood when interpreted as graphs, comprised of individual components that interact with each other, resulting in diverse local and global behaviors. When multiple graphs arise from overlapping set of nodes, the challenge of associating nodes across these graphs is a key problem in numerous application domains. The problem of graph matching was initially posed in the context of image and shape analysis, as well as for determining similarities in chemical structures [1], [2], [3]. With the rapid expansion of social network data [4], [5], [6], matching entities (represented by nodes in the graph) across multiple graphs has become a fundamental problem in social network analysis. For example, the knowledge of the mapping between nodes across multiple social media sites provides integration of information to create a more complete picture of user and community activity and trends.

Given two graphs, as represented by their adjacency matrices $\mathbf{A}, \mathbf{B} \in \{0,1\}^{N\times N}$, the graph matching problem is most commonly formulated as solving or approximating

$$\mathbf{P} = \underset{\mathbf{P}\in\Pi}{\operatorname{argmin}} \|\mathbf{A} - \mathbf{P}\mathbf{B}\mathbf{P}^T\|_F^2 = \underset{\mathbf{P}\in\Pi}{\operatorname{argmin}} \operatorname{Tr}(\mathbf{A}\mathbf{P}\mathbf{B}\mathbf{P}^T)$$

where Π denotes the set of $N\times N$ permutation matrices. However, graphs are often constructed from noisy and incomplete data at different times, and the notion of how edges are formed or defined might vary in different graphs. Matching graphs based on the edge discrepancies only considers the immediate adjacencies, but neglects the longer-range structural properties which may be better preserved across networks.

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited. This material is based upon work supported by the Under Secretary of Defense for Research and Engineering under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Under Secretary of Defense for Research and Engineering.

On the other hand, some authors have proposed methods that embed the nodes as vectors in a low-dimensional space using the eigenvectors of the graph Laplacian and solve the matching problem in the embedded space [7], [8]. The node embedding captures longer-range structures of the nodes but these approaches only implicitly accounts for immediate adjacencies. Depending on the topologies of the graphs to be matched, we hypothesize that graph matching can benefit from considering both the immediate adjacency structures and the longer-range structures. The goal of this paper is to design a general framework that exploits tools and theories in graph signal processing to capture both the local and long-range structure properties across graphs, and reformulates the graph matching problem as one of matching similarities among pulse response signals to multiple graph filters.

The contribution of this paper is three-fold. First, this work takes the advantage of graph filtering to obtain multiple matrix representations of a graph that captures the long-range structural similarities between nodes at different scales. To our knowledge, this is the first explicit use of graph signal processing to solve the graph matching problem. Specifically, we use a low-pass diffusion filter to generate a diffusion process around each node. The larger the time scale in the heat diffusion operator, the larger the heat is to spread to its neighborhood, capturing long-range structural information. The proposed approach is also applicable to other types of filters [9], [10], depending on the topology of the graphs to be matched. Importantly, the graph filters can be tailored to emphasize similarities of longer-range connectivity structures via low-pass filters or to emphasize fined-grain correlations between edges. Second, the proposed framework generalizes the existing spectral graph matching methods. We provide a cohesive interpretation of these spectral methods using graph filtering. Third, this work provides a novel way of using multi-objective optimization to perform matching on multiple pairs of graph representations. The advantage of using multi-objective optimization is that graph representations at different scales are treated independently and are all explicitly accounted for.

II. Graph Matching via Multi-Scale Diffusion

A. Filtering for Graph Matching

To solve the graph matching problem from the perspective of graph signal processing, we reformulate the problem as matching similarity matrices across graphs where the similarities are computed between signals associated with each node.

The signals are computed as the response of filters applied to pulse signals at each node in the graph.

Consider a simple graph with N vertices and adjacency matrix $\mathbf{A}$. The normalized Laplacian is defined as $\mathbf{L}_A = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$, with diagonalization $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}$ where $\mathbf{U}$ denotes the eigenvector matrix of $\mathbf{L}$ and $\mathbf{\Lambda}$ is a diagonal matrix containing the corresponding eigenvalues $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{N-1}$. For a frequency response function $h : \mathbb{R}^+ \mapsto \mathbb{R}^+$, the graph filter is given by the matrix $Uf(\Lambda)U^T$ where $h(\Lambda)$ denotes applying h to the diagonal elements of Λ [9], [10].

Given a pair of frequency response function $h, g : \mathbb{R}^+ \mapsto \mathbb{R}^+$ we create pairwise similarity matrices as follows. We associate unit pulses $\mathbf{e}_i$ to each node $i \in V$, where $e_{ii} = 1$ and $e_{ij} = 0$ for $j \neq i$. For graph $\mathbf{A}$, consider the pulse responses given by the signals $x_i^{(h)} = \mathbf{U}_A h(\mathbf{\Lambda}_A)\mathbf{U}_A^T e_i$ where $\mathbf{L}_A = \mathbf{U}_A\mathbf{\Lambda}_A\mathbf{U}_A^T$. Similarly, for graph $\mathbf{B}$ and frequency response function g, we denote the pulse response signals as $y^{(g)} = \mathbf{U}_A g(\mathbf{\Lambda}_B)\mathbf{U}_B^T e_i$, where $\mathbf{L}_B = \mathbf{U}_B\mathbf{\Lambda}_B\mathbf{U}_B^T$.

Finally, we compute the similarities

$$\mathbf{A}_{ij}^{(h)} = \frac{\mathbf{x}_i^{(h)}(t)^T\mathbf{x}_j^{(h)}(t)}{\|\mathbf{x}_i^{(h)}(t)^T\|_2\|\mathbf{x}_j^{(h)}(t)\|_2}$$

and similarly, $\mathbf{B}_{ij}^{(g)} = \frac{\mathbf{y}_i^{(g)}(t)^T\mathbf{y}_j^{(g)}(t)}{\|\mathbf{y}_i^{(g)}(t)^T\|_2\|\mathbf{y}_j^{(g)}(t)\|_2}$, so that the entries of $\mathbf{A}^{(h)}, \mathbf{B}^{(g)} \in \mathbb{R}^{n\times n}$ correspond to the cosine similarity between the pulse response signals associated to each node.

We can now reformulate the graph matching problem as matching these similarity matrices. Given a pair of graphs, $\mathbf{A}, \mathbf{B}$, and frequency response functions, h, g, the objective function to maximize is given by $f_{h,g}(\mathbf{P}) = \frac{1}{2}\text{trace}(\mathbf{A}^{(h)}\mathbf{P}\mathbf{B}^{(g}\mathbf{P}^T)$ which has gradient $\mathbf{A}^{(h)}\mathbf{P}\mathbf{B}^{(g)}$. Note this is still a quadratic assignment problem and approximate quadratic assignment solvers can be employed. However, by selecting appropriate frequency response functions either local or non-local aspects of the graph structure can be emphasized.

Example 1 (Enron Graph)**.** In Figure 1, we show the resulting similarity matrices for three different frequency responses as applied to one of the Enron graphs (see Section III). The leftmost plot shows the normalized Laplacian and the remaining three plots, from left to right, show the similarity matrices associated with the following frequency response functions: $e^{-4\lambda}$ for the low-pass filter, $\lambda^{10}e^{-5\lambda}$ for the band-pass filter, and $e^{4\lambda}$ the highpass filter. Evidently, the lowpass filter highlights substantial similarities among groups of nodes where as the bandpass and highpass filters emphasize some of these same communities but also highlights differences between and among these groups and emphasizes finer grained structures. The ordering of the nodes was chosen by clustering the nodes using the greedy modularity maximization approach of [11], followed by degree sorting.

B. Multiple Data-Dependent Filters

Choosing appropriate filters can have a substantial impact on the overall performance of the matching. To allay the impact of a specific filter choice, we propose using multiple filters along with a multi-objective matching approach to improve performance. To motivate the proposed approach, let us first review the existing spectral graph matching and discuss the similarities of these methods to the graph filtering approach.

While graph matching typically proceeds by matching the adjacency matrices, some authors have matched graphs using the eigenvectors of the Laplacian or adjacency matrices [7], [8], [12]. Specifically, [7] considers the problem minimizing the Frobenius norm $\text{argmin}_{\mathbf{P}\in\Pi}\|\mathbf{L}_A^k - \mathbf{P}\mathbf{L}_B^k\mathbf{P}^T\|_F^2$ where $\mathbf{L}^k$ denotes the truncated eigen-decomposition of the Laplacian $\mathbf{L}$ corresponding to the k smallest eigenvalues. When graphs are near-isomorphic, the formulation is equivalent to applying a low-pass filter of a given threshold to the graphs. On the other hand, [12] proposed a matching algorithm that proceeds by comparing eigenvectors of the adjacency matrix. Specifically, the proposed algorithm solves the linear assignment problem for the matrix $\mathbf{U}\mathbf{V}^T$ where $\mathbf{U}, \mathbf{V} \in O(n)$ are the orthogonal matrices of eigenvectors for the adjacency matrices. In a recent work, [13] proposed a similar algorithm but modified the procedure to perform the linear assignment problem on a different matrix $\sum_{i,j}\frac{1}{(\lambda_i-\mu_j)^2+\eta^2}U_iU_i^T\mathbf{J}V_jV_j^T$, for $\eta > 0$, where λ_i, μ_j are the non-increasing eigenvalues of $\mathbf{A}, \mathbf{B}$, respectively. [13] provide theoretical justifications that this computationally efficient algorithm can accurately match (highly) correlated Erdős-Rényi random graphs. While motivated by very different tools, there are substantial similarities to graph filtering. In particular, by considering the frequency response functions $h_j(\lambda) = \frac{\eta}{(\lambda-\mu_j)^2+\eta^2}$ and $g_j(\lambda) = \delta_{\lambda,\mu_j}$, the approach of [13] can be viewed as matching using N different bandpass filters selected based on the eigenvalues of one of the adjacency matrices. Note that as Erdős-Rényi graphs are nearly regular, the normalized Laplacian has a very similar spectrum to the adjacency matrix, just shifted and scaled. Hence, using bandpass filters in our approach will yield very similar results to their approach. However we consider an iterative QAP solver whereas their approach uses a one-step LAP solver, so that the first step of our procedure corresponds closely to their procedure.

In practice, we adapt the filters to the given graphs based on the spectral properties of the respective normalized Laplacians. In the simulations, we use low-pass filters with frequency response functions of the form $f_t(\lambda) = \exp(-t\lambda)$. After applying these lowpass filters, the resulting signals correspond to a heat diffusion along edges of the graph from an initial point of a unit pulse at a single node. We denote the resulting similarity graphs as *heat graphs*, an example of which is shown in the second panel of Figure 2. For each graph, we select two or more values of t which are spaced logarithmically between $t_1 = \min\{2, \log(10)/\lambda_n\}$ and $t_2 = \log(10)/\lambda_2$. These choices were motivated by the recommendations found in [14] and we found they offer substantial improvements of using the adjacency matrix alone.

C. Multi-objective Matching

The original graph matching problem has a single objective, to maximize the number of common edges $f(\mathbf{P}) = \text{Tr}(\mathbf{A}\mathbf{P}\mathbf{B}\mathbf{P})^T$ between $\mathbf{A}$ and $\mathbf{P}\mathbf{B}\mathbf{P}^T$, where $\mathbf{P}$ denotes a

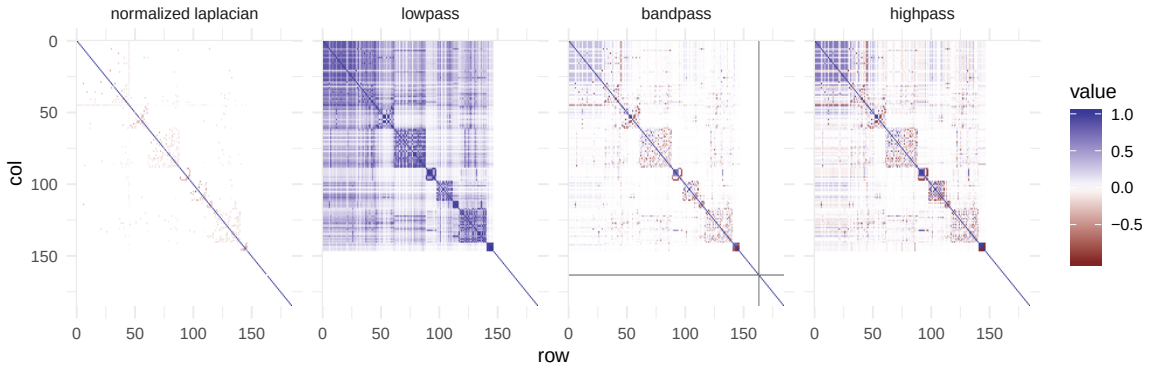

Figure 1: Examples of similarity functions for different filters for a graph based on the Enron dataset.

permutation matrix. Our proposal can incorporate objectives for each filter, as well as the original adjacency matrix, and we use a multi-objective matching approach to search for a Pareto optimal solution.

Let $h_1, \ldots, h_M$ and $g_1, \ldots, g_M$ denote the chosen frequency response functions. Let us index the heat graphs as $\mathbf{A}_m = \mathbf{A}^{(h_m)}$ and $\mathbf{B}_m = \mathbf{B}^{(g_m)}$ for $m = 1, \cdots, M$ where t_m and $\bar{t}_m$ denote the time scales used to generate the diffusion signatures on $\mathcal{G}_A$ and $\mathcal{G}_B$, respectively. Let $\mathbf{A}_0 = \mathbf{A}$ and $\mathbf{B}_0 = \mathbf{B}_0$ be the original adjacency matrices. Leveraging both the adjacency matrices and the heat graphs to find the latent node alignment can be achieved by the following multi-objective optimization [15] :

$$\max_{\mathbf{P} \in \mathcal{D}} (f_0(\mathbf{P}), f_1(\mathbf{P}), \cdots, f_M(\mathbf{P})) \tag{1}$$

where $f_m(\mathbf{P}) := \mathrm{Tr}(\mathbf{A}_m^T \mathbf{P} \mathbf{B}_m \mathbf{P}^T)$ denotes the objective function that measures the edge agreement between graphs represented by the corresponding matrices $\mathbf{A}_m$ and $\mathbf{B}_m$ [16].

To search for the latent alignment that maximizes the M objective functions simultaneously, we use the multiple-gradient descent algorithm [15]. This algorithm generalizes the classical steepest-descent method to multi-objective problems. At each iteration, it computes the gradients of the individual objectives and combines the ascent directions to minimize the combined norm of the gradients. The initial ascent directions are found using the Hungarian algorithm, yielding M permutation matrices. Finding the final ascent direction and step size involves standard optimization techniques.

III. Experimental Results

To illustrate the efficacy of the proposed multi-objective graph matching approach, we compare the matching results found by the proposed method and the method based on minimizing edge discrepancy on a set of real-world graphs. The performance comparison is performed by repeating the process of selecting randomly S number of seeds (known matches) and using each algorithm to match graphs. The performance is measured by the average match ratio, which is defined as the number of correctly matched nodes divided by the number unknown matches $N - S$. For all graphs, we use only $M = 2$ similarities matrices generated by the diffusion low-pass filter in addition to the normalized adjacency matrices.

The graphs include Enron email graphs, two sets of brain graphs and wikipedia graphs. The Enron email dataset [17] consists of $N = 184$ employees of the Enron Corporation. The graphs are unweighted. Edges represent a vertex sending at least one email to another vertex. The Enron email graphs used in this paper are graphs of the consecutive weeks $t = 130$ and $t = 131$ [18]. For the brain graphs, we consider a pair of brain networks derived from diffusion MRI data [19]. The datasets consists of test-retest pairs. The raw image data is converted into a connectome with N regions of interest (ROI) as vertices. The two sets of graphs correspond to two different brain atlases [20]: the Desikan atlas with $N = 70$ and the Talairach atlas with $N = 1105$. The edges are weighted by the estimated number of neural fiber bundle connecting the regions. The Wikipedia graphs consist of articles in both English and French [18]. There are $N = 1382$ English articles and edges represent the intra-language links from one article to another. The corresponding French articles were collected through the inter-language links.

Figures 2-5 show the mean and the standard deviation of the matching accuracy for the 4 sets of real-world graphs. Each dot in the plots represents the average matching accuracy averaged with 100 trials using randomly selected seeds. Observe that the proposed graph matching using multi-scale diffusion features improves the matching performance, especially when the number of seeds is small. In addition, the proposed method is overall less sensitive to the number of seeds.

We also examined the effect of the number of graph filters used to generate the similarity matrices in the multi-objective optimization. We observe that when the percent of seeds is low, increasing the number of graph filters has the greatest impact on the performance improvement. As the number of seeds increases, the effect of the additional heat kernel graphs on improving the matching performance decreases. For the Enron dataset, $M = 2$ seems to be the optimal number, beyond which the matching accuracy does not increase as much. Results on the effects of using other filters on different graphs will be reported in the full paper.

IV. Conclusion

We propose a novel graph matching approach that exploits tools from graph signal processing to capture multiple scales

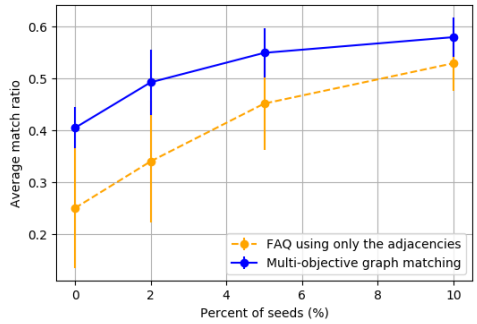

Figure 2: Matching Enron email graphs $t = 130, 131$, plotting average match ratio against the percentage of seeds

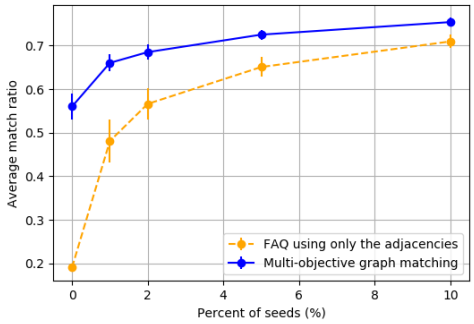

Figure 4: Matching Talairach brain graphs plotting average match ratio against the percentage of seeds

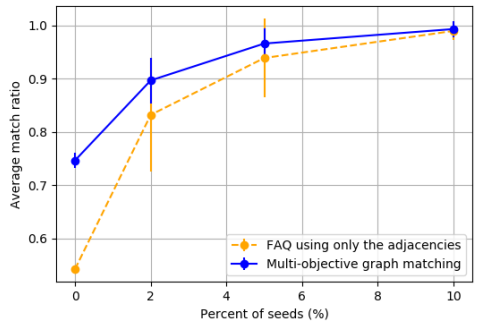

Figure 3: Matching Desikan brain graphs plotting average match ratio against the percentage of seeds

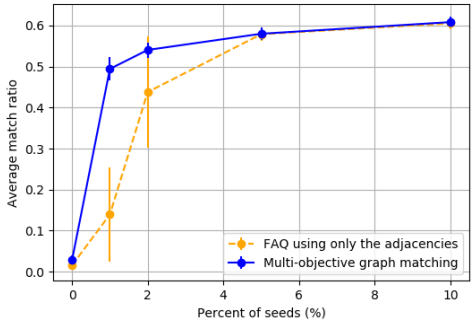

Figure 5: Matching English and French Wikipedia graphs, plotting average match ratio against the percentage of seeds

of graph structure and uses multi-objective optimization approaches to find accurate matchings. Our approach demonstrates substantial improvements over approaches where only the adjacency matrix is used. This improvement is especially high when little or none of the true correspondence is known *a priori*.

References

[1] Emmert-Streib, F., Dehmer, M., Shi, Y.: Fifty years of graph matching, network alignment and network comparison. Information sciences **346–347**, 180–197 (2016)

[2] Conte, D., Foggia, P., Sansone, C., Vento, M.: Thirty years of graph matching in pattern recognition. International Journal of Pattern Recognition and Artificial Intelligence **18**(03), 265–298 (2004)

[3] Foggia, P., Percannella, G., Vento, M.: Graph matching and learning in pattern recognition in the last 10 years. International Journal of Pattern Recognition and Artificial Intelligence **28**(01), 1450001 (2014)

[4] Zhang, J., Yu, P.S.: Multiple anonymized social networks alignment. In: 2015 IEEE International Conference on Data Mining, pp. 599–608. ieeexplore.ieee.org, ??? (2015)

[5] Campbell, W.M., Li, L., Dagli, C., Acevedo-Aviles, J., Geyer, K., Campbell, J.P., Priebe, C.: Cross-domain entity resolution in social media. arXiv preprint arXiv:1608.01386 (2016)

[6] Zhao, W., Tan, S., Guan, Z., Zhang, B., Gong, M., Cao, Z., Wang, Q.: Learning to map social network users by unified manifold alignment on hypergraph. IEEE transactions on neural networks and learning systems **29**(12), 5834–5846 (2018)

[7] Knossow, D., Sharma, A., Mateus, D., Horaud, R.: Inexact matching of large and sparse graphs using laplacian eigenvectors. In: Graph-Based Representations in Pattern Recognition, pp. 144–153. Springer, Berlin Heidelberg (2009)

[8] Mateus, D., Horaud, R., Knossow, D., Cuzzolin, F., Boyer, E.: Articulated shape matching using laplacian eigenfunctions and unsupervised point registration. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8. IEEE, Anchorage, Alaska (2008)

[9] Hammond, D.K., Vandergheynst, P., Gribonval, R.: Wavelets on graphs via spectral graph theory. Applied and Computational Harmonic Analysis **30**(2), 129–150 (2011)

[10] Shuman, D.I., Narang, S.K., Frossard, P., Ortega, A., Vandergheynst, P.: The emerging field of signal processing on graphs: Extending High-Dimensional data analysis to networks and other irregular domains (2012). 1211.0053

[11] Clauset, A., Newman, M.E.J., Moore, C.: Finding community structure in very large networks. Physical review. E, Statistical, nonlinear, and soft matter physics **70**(6 Pt 2), 066111 (2004)

[12] Umeyama, S.: An eigendecomposition approach to weighted graph matching problems. IEEE transactions on pattern analysis and machine intelligence **10**(5), 695–703 (1988)

[13] Fan, Z., Mao, C., Wu, Y., Xu, J.: Spectral graph matching and regularized quadratic relaxations II: Erdős-rényi graphs and universality (2019). 1907.08883

[14] Sun, J., Ovsjanikov, M., Guibas, L.: A concise and provably informative multi-scale signature based on heat diffusion. In: Computer Graphics Forum, vol. 28, pp. 1383–1392 (2009). Wiley Online Library

[15] Désidéri, J.-A.: Multiple-gradient descent algorithm (mgda) for multiobjective optimization. Comptes Rendus Mathematique **350**(5-6), 313–318 (2012)

[16] Vogelstein, J.T., Conroy, J.M., Lyzinski, V., Podrazik, L.J., Kratzer, S.G., Harley, E.T., Fishkind, D.E., Vogelstein, R.J., Priebe, C.E.: Fast approximate quadratic programming for graph matching. PLOS one **10**(4), 0121002 (2015)

[17] Shetty, J., Adibi, J.: The enron email dataset database schema and brief statistical report. Information sciences institute technical report, University of Southern California **4**(1), 120–128 (2004)

[18] Fishkind, D.E., Adali, S., Patsolic, H.G., Meng, L., Singh, D., Lyzinski, V., Priebe, C.E.: Seeded graph matching. Pattern recognition **87**, 203–215 (2019)

[19] Kiar, G., Bridgeford, E., Roncal, W.G., Chandrashekhar, V., Mhembere, D., Ryman, S., Zuo, X.-N., Marguiles, D.S., Craddock, R.C., Priebe, C.E., et al.: A high-throughput pipeline identifies robust connectomes but troublesome variability. bioRxiv, 188706 (2018)

[20] Lancaster, J.: The talairach daemon, a database server for talairach atlas labels. Neuroimage **5**(4), 238–242 (1997)