A mapping between structural and functional brain networks

J. Meier[1], P. Tewarie[2], A. Hillebrand[3], L. Douw[4,5], B.W. van Dijk[3], S.M. Stufflebeam[5,6] and P. Van Mieghem[1]*

1 Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft, The Netherlands;

2 Department of Neurology, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, The Netherlands

3 Department of Clinical Neurophysiology and Magnetoencephalography Center, Neuroscience Campus Amsterdam, VU University Medical Centre, Amsterdam, The Netherlands

4 Department of Anatomy and Neurosciences, VU University Medical Center, Amsterdam, The Netherlands

5 Department of Radiology, Athinoula A. Martinos Center for Biomedical Imaging / Massachusetts General Hospital, Boston MA

6 Department of Radiology, Harvard Medical School, Boston MA

Running Title: *A mapping between structural and functional brain networks*

Keywords: structural networks, functional networks, structural connectivity, functional connectivity, brain networks, matrix mapping

*__Corresponding author:__

Jil Meier

Delft University of Technology

Faculty of Electrical Engineering, Mathematics and Computer Science

P.O. Box 5031, 2600 GA Delft

The Netherlands

E-Mail: j.m.meier@tudelft.nl

Tel.: +31 (0)6 33 30 76 15

Abstract

The relationship between structural and functional brain networks is still highly debated. Most previous studies have used a single functional imaging modality to analyze this relationship. In the present work, we use multimodal data, from functional MRI, magnetoencephalography and diffusion tensor imaging, and assume that there exists a mapping between the connectivity matrices of the resting-state functional and structural networks. We investigate this mapping employing group-averaged as well as individual data. We indeed find a significantly high goodness of fit level for this structure-function mapping. Our analysis suggests that a functional connection is shaped by all walks up to the diameter in the structural network in both modality cases. When analyzing the inverse mapping, from function to structure, longer walks in the functional network also seem to possess minor influence on the structural connection strengths. Even though similar overall properties for the structure-function mapping are found for different functional modalities, our results indicate that the structure-function relation is modality-dependent.

No words in abstract: 159

1) Introduction

Applying network science has become a common practice in neuroscience to understand functional interactions in the healthy brain and to identify abnormalities in brain disorders (Stam, 2014). The collection of these functional connections is often referred to as the functional network and is facilitated by the underlying structural network, i.e. the set of physical connections between neuronal populations. At the same time, functional connections influence modulations of these physical connections by long-term potentiation, plasticity or neuromodulation. In recent years, there has been an increasing interest to understand the emergence of functional brain networks given the constraints of the underlying structural network (Deco et al., 2012; Senden et al., 2014; Abdelnour et al., 2014; Honey et al., 2009). However, the mutual relationship between the structural and functional networks remains highly debated (Robinson, 2012; Robinson et al., 2014; Deco et al., 2014).

Empirical studies have revealed an overlap between structural and (resting-state) functional connections, i.e. the presence of both a structural and functional connection between two brain regions (van den Heuvel et al., 2009; Skudlarski et al., 2008; Hermundstad et al., 2013). However, this overlap is imperfect as functional interactions between brain regions exist in absence of direct structural connections, and also indirect structural connections with the length of two links cannot fully account for these functional connections either (Honey et al., 2009). Moreover, the overlap between structural and functional connections also depends on the time scale considered, where functional connections estimated from larger time windows strongly overlap with the underlying structural connections, for smaller time windows there can be a structural-functional network discrepancy due to distributed delays

between neuronal populations that cause transient phase (de-)synchronization (Messé et al., 2014; Honey et al., 2007; Ton et al., 2014).

On larger time scales, several properties of the underlying structural network have been shown to play an essential role in shaping the functional networks, such as the Euclidian distance between two brain regions (Alexander-Bloch et al., 2013). However, taking into account Euclidean distance alone is insufficient to explain the emergence of long-range functional connections (Vértes et al., 2012). Two recent studies showed that such long-range functional connections may be explained by the product of the degree of two nodes in the structural network, indicating the crucial role of structural hubs for explaining long-range functional connections (Tewarie et al., 2014; Stam et al., 2015). Moreover, Goñi and colleagues (Goñi et al., 2014) demonstrated that shortest paths in the structural network and perturbations from these paths are strong predictors for functional connections as these paths are favorable because of metabolic efficiency and fast communication.

Given these dependencies between structural and functional networks, the challenge is to integrate these different interdependencies into a single framework, for which we may need a more abstract representation. For example, a significant overlap in the connectivity profile of structural and functional networks suggests that part of the functional network connectivity matrix is a linear mapping from its structural counterpart. In addition, functional connections can also be accounted for by several other higher order features of the structural network as outlined above, which refer to non-linear relationships (see (Tewarie et al., 2014) for an example of such non-linearity). Based on the presence of these linear and non-linear features of the relationship between structural and functional networks, we go one step further by assuming that there is a mathematical function that maps the adjacency matrix of

the structural network onto that of the (resting-state) functional network and vice versa (see Figure 1b and Eq. (1) below). If we further assume that our mathematical function is analytic (Whittaker et al., 1996; Titchmarsh, 1939), then the map between structural and functional network can be expressed by a weighted sum of the matrix powers as explained in Section 2.2. Our method consists of a data-driven approach, from which the successive coefficients of this matrix mapping are determined. The major advantage of our method is that an a-priori specific form of a function is not needed. Another implication of such a function is the possible existence of an inverse function, i.e. a mapping from functional networks back to structural networks.

Most previous studies have found relationships between structural and functional networks using a single functional neuroimaging modality (Honey et al., 2009; Damoiseaux and Greicius, 2009), often using functional MRI (fMRI). As the fMRI response is an indirect measure for neuronal activity and contains non-neuronal signals, a structure-function dependency based on this modality could deviate from the same dependency derived from neuroimaging modalities that directly measure neuronal activity and connectivity. In contrast to fMRI, magnetoencephalography (MEG) measures neuronal activity and connectivity directly with excellent temporal resolution. However, given the increasing interest in multimodal imaging approaches there is a need to understand the modality dependency of the structure-function relationship in a single framework. A data-driven approach in the form of a matrix function may be helpful when investigating the modality dependency of the structural-functional network relationship: different modality-dependent coefficients may point to different specific functions for each modality. The relevance of elucidating the modality dependency of a mathematical function can be extended to the clinical field where we could answer questions such

as: which modality would be the most sensitive for picking up functional network changes given disease-specific structural network damage?

The aim of the present study is to analyze the structural-functional network relationship through a mathematical function in a multimodal framework. We use two datasets containing multimodal imaging data ranging from diffusion tensor imaging (DTI) data to MEG and fMRI data. We extend our analysis by also considering the relationship between structural and functional networks at the subject level in a third data set and finally discuss how that relationship can be interpreted neurobiologically.

2) Materials and Methods

2.1) Participants and Data Acquisition

In total, we use three data sets, which all have been used in different previous studies. The first two data sets are group-averaged data sets, obtained from two different centers, but analyzed together in one mapping.

i. A group-averaged structural imaging data set, i.e. a DTI network from 80 healthy subjects in 78 cortical automated anatomical labeling (AAL) brain areas (Gong et al., 2009).
ii. Two group-averaged data sets with functional imaging data, i.e. resting-state MEG and fMRI signals in the same 78 AAL cortical areas, one with 17 and another with 21 healthy subjects (Tewarie et al., 2014; Tewarie et al., 2015).
iii. An individual data set from 11 healthy subjects structural and functional imaging data, i.e. with DTI, resting-state MEG and fMRI time-courses in 219 brain areas (Douw et al., 2015).

For the group-averaged structural connectivity matrix, we use a literature-based structural network (data set (i)) (Gong et al., 2009). In every subject, cortical regions in the AAL atlas were considered to be connected if the end points of two white matter tracts were located in these regions (Gong et al., 2009). Then, a group-averaged structural connectivity matrix was obtained by testing each possible connection for its significance using a non-parametric sign test.

For the group-averaged functional imaging data set (data set (ii)), we use data obtained from our own imaging center. We employ the first data set with 17 healthy controls for our main analysis and the second data set from 21 healthy controls only for validation (Tewarie et al., 2014; Tewarie et al., 2015). The study was approved by the institutional ethics review board of the VUmc and all subjects gave written informed consent prior to participation. Both fMRI and MEG data sets underwent to some extent different pipelines (Tewarie et al., 2014; Tewarie et al., 2015) and are obtained from two different MEG scanners (CTF and Elekta). Detailed information about data acquisition and post-processing can be found in the previous papers. In short, for both MEG and fMRI cortical networks were constructed using the same cortical AAL regions as for the structural network consisting of 78 cortical regions (Gong et al., 2009). The Pearson correlation coefficient was computed between time signals to construct functional networks for fMRI for each subject (the absolute value was taken to avoid negative matrix elements). For MEG, a beamformer approach was used to reconstruct neuronal activity in AAL regions. Subsequently, the phase lag index (PLI), a measure for phase-synchronization, was computed between time series to reconstruct a functional connectivity matrix for each subject in the *alpha2* frequency band (10-13 Hz) (Stam et al., 2007). The present study can be considered as a continuation from previous work where we found a strong relationship between

structural and functional networks in the *alpha2* band and therefore we limited our analysis to this frequency band, although the fit could be generalized (Tewarie et al., 2014). Similar to the structural connectivity matrix, we averaged functional connectivity matrices across subjects for fMRI and MEG separately to obtain one group-averaged functional connectivity matrix for each modality. The averaging over multiple subjects was pursued in the attempt of reducing noise.

For the individual data set (data set (iii)), eleven healthy participants were included, exclusion criteria being psychiatric or neurological disease and use of medication influencing the central nervous system. This study was approved by MGHs institutional review board, and was performed in accordance with the Declaration of Helsinki. All participants gave written informed consent before participation. Pre-processing methodology of the DTI and fMRI data has been described in detail before (Douw et al., 2015). In short, a surface-based atlas approach was used for connectivity analysis of the fMRI and DTI data, using a parcellation scheme with 219 cortical surface parcels (Daducci et al., 2012; Gerhard et al., 2011). In addition, for every entry of the fMRI-based adjacency matrix the absolute value was taken to avoid negative matrix elements. MEG eyes-open resting-state data were collected in a magnetically shielded room with a 306-channel whole-head system (Elekta-Neuromag, Helsinki, Finland) and a sampling rate at 1037 Hz. Vertical and horizontal electro-oculograms were acquired simultaneously for off-line eye-movement artifact rejection. Head positions relative to the MEG sensors were recorded from four head-position indicator coils attached to the scalp. Landmark points of the head were digitized using a 3-D digitizer (Polhemus FASTRAK). MEG data underwent a number of pre-processing steps: (1) bad channel and bad epoch rejection, (2) eye-movement artifact removal via Signal

Space Projection (SSP), (3) downsampling with a decimate factor of 8 (to reduce computational expense). To compute the physical forward solution (lead fields), a single-layer boundary element method was applied to model the brain volume conduction, following an established procedure (Hämäläinen and Sarvas, 1987). The lead field of freely-oriented dipoles was then evaluated at each location. In solving the inverse problem, current density at each source location was approximated by a minimum 2-norm estimate in the same six frequency bands as was used for the second dataset (Hämäläinen and Ilmoniemi, 1994), with noise covariance computed from empty-room recordings on the same day (also band-pass filtered). For each subject, the cortical surface defined by the boundary between the gray and the white matter was reconstructed using FreeSurfer (Fischl et al., 1999), after which time series from the abovementioned 219 cortical surface parcels were reconstructed. The PLI was used as a connectivity measure on these time series (Stam et al., 2007). An average connectivity matrix per participant was calculated over all epochs.

2.2) Mathematical background

We will refer to matrix A as the binary adjacency matrix of the structural network for the group-averaged data (data set (i)) and to matrix W as one of the possible representations of the functional networks, W_{MEG} for MEG functional networks and W_{fMRI} for fMRI functional networks. Both A and W are NxN symmetric matrices, where N equals the number of cortical regions (N=78 for data set (i) and (ii); N=219 for data set (iii)). For both group-averaged and individual data, the matrix W has real elements w_{ij} between 0 and 1. In the case of the individual data, the structural network is described by a weighted adjacency matrix V with real elements between 0 and 1.

As mentioned before, we assume that there exists a function f such that

$$W = f(A) \tag{1}$$

or W = f(V) in the case of a weighted structural connectivity matrix V (see also Figure 1). Under quite mild conditions (Markushevich, 1985), the inverse f^{-1} of the function f exists such that

$$A = f^{-1}(W) . \tag{2}$$

If f(z) is a function of the complex number z and analytic in a disk with radius R around z_0, then f(z) possesses a Taylor series in the complex plane $\mathbb{C}$ that converges for all points z that lie in a disk with radius R around the point $z_0 \in \mathbb{C}$,

$$f(z) = \sum_{k=0}^{\infty} f_k(z_0)(z - z_0)^k \qquad \text{with } f_k(z_0) = \frac{1}{k!}\frac{d^k f(z)}{dz^k}|_{z=z_0}, \tag{3}$$

where $|z-z_0|<R$ and R is called the radius of convergence (Whittaker et al., 1996; Titchmarsh, 1939). It can be shown (Van Mieghem, 2011; Higham, 2008) that, if f(z) is analytic around z_0 and, hence, possesses a Taylor series (3), then for all matrices A, the matrix function f(A) also satisfies this Taylor series, provided each eigenvalue λ of A obeys $| \lambda - z_0 | < R$. Caley-Hamilton's famous theorem (Van Mieghem, 2011) states that any square matrix A satisfies its own characteristic polynomial, which implies that we can write $A^N = p_{N-1}(A)$, where $p_n(z)$ is a polynomial of degree n in z. Iteratively using the Caley-Hamilton theorem to the powers of $k \geq N$ in Eq. (1),

$$f(A) = \sum_{k=0}^{N-1} f_k(z_0)(A - z_0 I)^k + \sum_{k=N}^{\infty} f_k(z_0)(A - z_0 I)^k$$

shows that $\sum_{k=N}^{\infty} f_k(z_0)(A - z_0 I)^k$ can be written as a polynomial of order at most N-1 in A. In summary, any analytic function f, defined by (3), of a matrix A is a polynomial in A of degree at most N-1 (N is the number of nodes, here cortical regions, in the network),

$$f(A) = \sum_{k=0}^{N-1} c_k[f]A^k,$$

(4)

where c_k[f] are coefficients depending on the function f (provided each eigenvalue λ of A lies within the disk, i.e. obeys | λ - z_0 | < R). Because all the analyzed matrices have only zeros on the diagonal, their *trace* is 0. Since the *trace* equals the sum of the eigenvalues of a matrix (Van Mieghem, 2011), the average of the eigenvalues of the empirical matrices here is zero, which suggests us to choose z_0 = 0.

2.3) Mathematical Methodology

The first term $c_0[f] \cdot I$ in (4), which is the product of the constant coefficient c_0[f] and the identity matrix I, provides an offset to adjust the diagonal elements of our fitted matrix. In order to obtain a better goodness of fit, we introduce an offset also for all non-diagonal elements of our matrix. We define this offset as the error matrix E = c * J, where J = u *u^T is the all-one-matrix, $c \in \mathbb{R}$ and u is the all-one vector, u= (1, …, 1)^T. The constant error matrix E can be justified as a first approximation of the part that we do not know yet about the mapping between the structural and functional brain network. Thus, our fitting function is defined as

$$f_{(K)}(A) = \sum_{k=0}^{K} c_k[f]A^k + E,$$

(5)

where K ≤ N-1 is the maximal fitted exponent (N is the dimension of matrix A). We use the non-linear regression algorithm in MATLAB (using the routine *nlinfit.m* version *R2015a*) to estimate the coefficients in (5) by iterative least-squares estimation (for details see SI-H).

Denoting $\widetilde{W}$:= $f_{(K)}$(A), we evaluate the goodness of fit of our mappings using the Frobenius norm (Van Mieghem, 2014, p. 549). In particular, we compute the sum of squared errors (*SSE*), slightly modified as

$$SSE \coloneqq \sum_{i=1}^{N}\sum_{j=1}^{i}(w_{ij} - \widetilde{w}_{ij})^2, \tag{6}$$

where N=78 regions in the case of the group-averaged data and N=219 in the case of the individual data. Here, we only sum the elements of the lower triangular and the diagonal, because all our matrices are symmetric. Since the sum of squared errors is proportional to the number of fitted elements and to compare the different data sets with each other, we introduce a normalized version of *SSE* where we divide *SSE* by the degrees of freedom, which is in our case the number of fitted elements minus one

$$SSE_{norm} \coloneqq \frac{\sum_{i=1}^{N}\sum_{j=1}^{i}\left(w_{ij} - \widetilde{w}_{ij}\right)^2}{df_{top}}, \tag{7}$$

where df_{top}=N*(N-1)/2+N-1, N number of regions. Similarly, we can define the goodness of fit measure from (7) for the function f^{-1}: W → A by interchanging W and A in the description above. When we map all entries of one matrix onto the entries of another matrix, we implement our matrix mapping in the so-called topological domain (at the level of the whole adjacency matrix). The same mapping can also be analyzed in the spectral domain, i.e. at the level of the eigenvalues of the matrices (see SI-A.3).

3) Results

Mapping structural networks to functional networks

Firstly, we estimated the coefficients in Eq. (5) for the mapping from structural networks to functional networks at the group level (see SI Table 2 for K = 6). For both modalities, we can observe that the SSE_{norm} becomes lower, i.e. the fit becomes better, for increasing number of terms (Figure 2). Similarly, with an increasing number of fitted coefficients in Eq. (5), the patterns of the fitted functional connectivity matrices resemble better the empirical fMRI and MEG connectivity matrices (Figure 3, for a complete list of the ROIs see SI B.1). However, for the group-averaged data, there seems to be a limit for the number of terms, since including terms of 6^{th} order and higher did not significantly improve the estimation anymore for both MEG and fMRI under the 5% significance level. For these group-averaged networks, the best fit was reached for the mapping $f : A \rightarrow W_{MEG}$. We obtained significantly different values of the estimated coefficients for the two different modalities under the 5% significance level (see SI Figure 21, 95% confidence intervals did not overlap), indicating a modality-dependent mapping. For the mapping $f: A \rightarrow W_{fMRI}$, estimated coefficient values showed a clear decrease when going from lower to higher order terms, indicating that lower order terms in the expansion (5) contribute more to the estimation of the fMRI network (SI Figure 21). For the mapping $f: A \rightarrow W_{MEG}$, this steep decline in coefficients for higher order terms was not observed (see SI Figure 21). The SSE_{norm} for the data set of individual healthy controls was slightly higher (i.e. worse) than for the group-averaged matrices (Figure 4). Similar to the group level results, the mapping from structural to MEG networks provided better fits than from structural to fMRI networks also at the individual level.

We repeated the same analysis where either the structural or functional connectivity matrices were substituted by a reshuffled version of the empirical matrix (for details see SI-G). The results of this analysis are also displayed in Figure 2, showing a higher SSE_{norm} for all reshuffled cases compared to the original matrices, that is, the empirical results differed significantly ($p < .001$) from the reshuffled results. In addition, we observe that the decline in SSE_{norm} was in most cases for the reshuffled matrices rather narrow in comparison with the empirical matrices (Figure 2). Thus, the observed relationship between structure and function can hardly be reproduced by any reshuffled versions of the matrices. For individual networks, the average performance of the reshuffled matrices was also worse than the empirical original results (Figure 4). We tested the empirical results versus their reshuffles for significant difference with a Mann-Whitney-Wilcoxon (MWW) test and displayed the p-values in Table 1. From this test results, we can conclude that the mapping $f : V \rightarrow W_{fMRI}$ was able to outperform its random reshuffle for all subjects (see Table 1). But the goodness of fit for the mapping $f : V \rightarrow W_{MEG}$ was for 5 out of 11 subjects not better than the random reshuffles, indicating that the relation between the two matrices is less unique than for the anatomical matrix and the fMRI matrices. In order to cross-validate our mapping, we ran the same analysis on a second group-averaged data set (with similar processing pipeline) and found overlapping confidence intervals for the estimated coefficient values (Figures 5 and 6).

Mapping functional networks to structural networks

By reversing the role of A and W and following the same approach as before, we obtained goodness of fit values for the inverse mapping. More specifically, for the group-averaged data, we acquired better fits when starting from W_{fMRI} than from

W_{MEG} (see Figures 7 and 8). Similar to the mapping from structural to functional networks, the estimated coefficients were significantly different under the 5% significance level for the two modalities for the group-averaged data pointing towards a modality-dependent mapping (see SI Figure 22, 95% confidence intervals did not overlap). An overview of the estimated coefficients for this data set is given in SI Table 2. Furthermore, similar to the mapping f, no significant improvement of the goodness of fit level was found by including terms of a higher order than 5 for f^{-1}: $W_{MEG} \rightarrow A$. Even including W^5_{fMRI} in the mapping f^{-1}: $W_{fMRI} \rightarrow A$ hardly improved the fit (no significant improvement under the 5% significance level). Applying the same approach for the individual data, we were able to reach a lower overall error, thus a better fit, for f^{-1} than for f and the differences in modalities with respect to the residuals were very small for f^{-1} (see Figure 9).

To have a benchmark for the overall residuals, we again repeated the same analysis with reshuffled matrices. Similar to f, the function f^{-1} outperformed the random reshuffles for group-averaged networks (see Figure 7, p-value of 0% for MWW-test). On the subject level, the function f^{-1} obtained significantly better results for the empirical matrices than their random reshuffles for most of the individuals under the 5% significance level (two outliers for the p-values of the MWW-test for f^{-1}: $W_{fMRI} \rightarrow A$, see Table 1 and Figure 9). Again, the same analysis using the second group-averaged data set for MEG revealed only for the estimated coefficients $c_1[f]$ and $c_2[f]$ from (5) significant differences between the first and the second data set (for K = 5, Figures 10 and 11). For fMRI, a significant difference could only be determined for $c_1[f]$ but not for the other estimated coefficients from (5), which again cross-validates our mapping between different data sets.

Moreover, the whole analysis was repeated multiple times to check for the stability of the estimated coefficients, which resulted in exactly the same coefficients every time, underlining the robustness of our results. We also analyzed in more detail which connections were well predicted by our approach and which were estimated less accurately (see SI Figures 27-34). A corresponding analysis in the spectral domain (see SI-B.2 for the results) illustrated that the estimated coefficient values were similar to those in the topology domain for the function f but not for f^{-1} (see SI Figures 17-20). The dissimilarities between the spectral and topology domain are most probably due to eigenvector perturbations between the different analyzed empirical matrices. These eigenvector perturbations can probably be traced back to noisy measurements (see SI-F).

4) Discussion

In this study, we have analyzed the mutual dependency of structural and (resting-state) functional networks in a multimodal framework by assuming that there exists a mathematical function that allows for a mapping between the two networks. This function was then analyzed without assuming a priori any specifics and by estimating the coefficients for the mappings in both directions (i.e. structural to functional and functional to structural networks). Our analysis convincingly implicated that our assumption of a mapping between the two networks was justified because we reached overall good fits outperforming random reshuffles and resulting in similar matrix patterns. However, our results also indicated that the mapping was modality-dependent as the coefficients for mappings with MEG- or fMRI-based networks significantly differed.

The existence of such a mathematical function points towards the fact that the functional connectivity of the brain can be described by a combination of the underlying structural connections. Because of the stability of the estimated coefficients and their cross-validation across different data sets, such a mathematical function could potentially be used to predict structure from function or vice versa in future studies. Also, once we can use this mathematical framework to predict 'healthy' functional connectivity, we can compare the matrix to the actual measured functional network of the patient and identify possible malicious connections indicating disease.

4.1) Neurobiological interpretation

If we consider the case of a binary structural adjacency matrix, then the matrix element $(A^k)_{ij}$ equals the number of walks of length k between node i and node j. Each term $c_k[f]A^k$ can be considered as the contribution of walks with hopcount k to the functional network (see SI Figure 23). Here, hopcount is defined as the number of intermediate links between two nodes in a walk (length of the walk). Our approach confirms the ideas postulated by Robinson and co-workers that a functional connection can be regarded as a sum of all possible walks between two regions (Robinson, 2012; Robinson et al., 2014). Additionally, our approach returns the coefficients $c_k[f]$, which can be interpreted as the influence of all walks with hopcount k (see SI Table 2 and Figure 23). In contrast to a path, a walk can traverse the same node more than once. Potential loops in walks are also in line with the belief that re-entry loops can act as a resonating system to enhance a signal that needs to be spread over a long distance (Goñi et al., 2014).

In contrast to most previous studies, we followed a multimodal approach analyzing the mapping for MEG and fMRI data. As opposed to studies that assumed a specific function beforehand, we followed a data-driven approach by fitting coefficients of the general expression (5). More precisely, fMRI networks seemed to be shaped by walks of lower hopcount in the structural network since the coefficients were higher for these configurations (see SI Figure 21). In contrast, for MEG networks all walks from the underlying structural network up to hopcount 5 appeared to contribute more or less equally to the resulting fitted functional network matrix (see SI Figure 21). Overall, we found that estimations from structural networks were more accurate when predicting MEG networks on both individual and group level than when predicting fMRI networks. However, when the functional network was used to predict the structural one, we saw only small differences at the individual level between the modalities but at the group level the fitting using fMRI matrices performed better. These observations together with the significantly different coefficients for MEG and fMRI confirm the modality dependency of the mapping. If ρ denotes the diameter of the network, defined as the hopcount of the longest shortest path in a graph (Van Mieghem, 2011), our analysis for both fMRI and MEG suggests that the diameter of the unweighted structural network ($\rho = 6$) is directly related to the number of terms $K = 5$ in (5) that are sufficient for the best fit of the mapping from structural to functional networks. Hence, a functional connection between two regions seems only to be shaped by walks in the structural network that are shorter than the diameter of this structural network. The important role of the diameter in this fitting procedure can also be mathematically justified (see also SI-A.2).

Besides the possibility of predicting the functional network using the structural network, our analysis also has practical implications on how communication

processes shape brain activity. Bullmore and Sporns proposed the hypothesis that the brain is optimized for efficiency and robustness (Bullmore and Sporns, 2012). Our findings seem to be in line with this idea since the brain seems to use not only (structural) shortest paths (most efficient from a network perspective) for communication but is also transmitting information through less efficient paths or walks. Thus, there seems to be some kind of degeneracy in the brain (Price and Friston, 2002). From a network science perspective, spreading information not only through the shortest path makes the (healthy) brain function more robust against link breakage. However, there seems to be an upper bound for the length of the paths that the brain uses for communication, which corresponds to the diameter of the structural brain network. Walks that are longer than the diameter are highly inefficient for communication. The diameter therefore seems to symbolize the trade-off between efficiency and robustness (Bullmore and Sporns, 2012). It is this degeneracy and robustness that could keep two regions functionally connected when the direct structural connection is damaged in disease. In multiple sclerosis, the structural network gets damaged due to lesions and diffused white matter damage. With this theory we could predict which detours are likely to be taken for functional connections in order to uphold (sub)-optimal network efficiency. Thus, based on the damaged structural network we could be able to make predictions on how this damaged structural network might map onto a functional network. These practical implications seem to agree with several studies that have shown that the averaged path length is higher in diseases than in the healthy brain (Stam, 2014).

Our mathematical approach incorporates previous models on the relationship between structural and functional networks into one single model. For example, a previous study found that the shortest paths and detours along these paths in the

structural network were the strongest predictors for functional connections (Goñi et al., 2014). This result agrees with our finding of the structural-functional network mapping being dependent on the combination of walks with small hopcounts (corresponding to the shortest paths in the network) and detours from these shortest paths. Also the suggestion that network diffusion has the ability to predict functional connections (Abdelnour et al., 2014) is in line with our work. Network diffusion indicates that information is not merely transmitted through the shortest paths, but also through less efficient paths. Furthermore, our mathematical function also includes the predictive value of common neighbors for functional connections (Vértes et al., 2012). The term $c_2[f]A^2$ in (5) corresponds to the weighted number of walks between any pair of nodes with hopcount 2, i.e. walks from any node i to a node j via a common neighbor. In a previous study, Tewarie and coworkers (Tewarie et al., 2014) demonstrated that the degree product between nodes in the structural network together with the Euclidean distance has the ability to predict the functional connections between these nodes. We observed here that our approach with the sum of structural matrices A^k in (5) is correlated not only with the degree product (SI Figure 24) but also with the complete previous model (including Euclidean distance, SI Figure 25).

Predicting the structural network from the functional network has received relatively little attention (Robinson et al., 2014; Abdelnour et al., 2014; Deco et al., 2014; Robinson, 2012). We assumed that the structural network is a weighted sum of powers of the functional network matrix W. However, unlike the structure-to-function mapping f, the interpretation of this mathematical function is less straightforward: If we define the weight of a walk as the product of all weights along this walk, then the matrix entry $(W^k)_{ij}$ represents the summed weights of all possible

walks with hopcount k between node i and node j. Similar to the function f, we find for f^{-1} that higher powers of W do not contribute substantially to the goodness of fit of our mapping. In contrast to the powers of a binary matrix, W^k does not only contain the number of walks with hopcount k but also incorporates information about their weight structure. Still, we can conclude that longer walks in the functional network seem to influence the structural brain network less. Practically, this result not only helps us to reconstruct the structural connections when we have only the functional connectivity matrices, but it also indicates that a direct structural connection between two brain regions seems to be influenced not only by their direct functional connectivity but also by the (functional) communication within a small hopcount neighborhood of those two regions.

Using an additional data set of individual healthy controls (data set (iii)), we found that our mapping can also be generalized to the individual level. For the individual mappings, we also found that nearly all mappings were able to outperform their reshuffled benchmark except for some outliers (see Table 1). Furthermore, we compared the results of the group-averaged data and the individual data (each of these containing data from multiple modalities). In the case of the mapping from structural to functional networks, the performance when using individual fits was similar to that obtained when using the group-averaged matrices (see Figures 2 and 4). However, for the inverse mapping, the individual mappings provided a much better fit than the group-averaged mappings. These results could potentially be explained by the following factors: (1) there exists an even stronger relationship between function and structure at the individual level, (2) the use of weighted structural connectivity matrices (instead of the binary group-averaged structural connectivity matrix), which are more representative of the underlying fiber bundle

structure or (3) the fact that the structural and functional information were gathered from the same group for data set (ii) (in contrast, the group-averaged structural and functional connectivity matrices were based on two different sets of healthy controls).

4.2) Technical implications

Our approach may provide important information about the DTI-obtained structural network that is generally missed due to methodological issues with crossing versus kissing fibers which usually affect inter-hemispheric connections. Given the functional networks, a mapping to the structural network could also allow to distinguish between genuine and false positive connections, which are inherently present in DTI data (Thomas et al., 2014). For example, in the structural networks estimated from MEG and fMRI networks we observed more homologous inter-hemispheric connections than in the actual empirical structural network (see the off-diagonal in SI Figure 14). In addition, for MEG functional connectivity metrics, there are well known methodological issues with volume conduction, signal leakage and field spread. By using our approach and trying out different functional connectivity metrics, one could aim to find the common properties of these mappings, i.e. those that are invariant of the functional metric that was used.

4.3) Methodological considerations

Firstly, we investigated the relationship between the structural network and static patterns of (resting-state) functional connectivity, as functional connectivity was estimated over epochs of several seconds. Therefore, our approach does not consider the dynamical aspects of functional connectivity. It is well known that functional networks obtained from smaller time windows correspond less to the

structural network (Messé et al., 2014; Honey et al., 2007; Ton et al., 2014) and therefore our approach could be less applicable to these smaller time scales.

Secondly, the mapping employed in this study can certainly be influenced by the choice of the parcellation of brain regions. However, as long as the ratio between genuine (functional or structural) connections and noise in the matrices remains similar between parcellation atlases, we do not expect it to have a significant impact on the goodness of fit of our mapping. Despite the well-known limitations of the AAL atlas, it still provides a commonly used framework in neuroimaging studies. By using it, the results from our study are directly relevant for this existing body of work. We also provided a suggestion of how to overcome the dimension differences of the matrices of different parcellations mathematically in SI-I.

Thirdly, our mapping can be influenced by noise in the matrices, such as the presence of false positives in the structural connectivity matrix. However, by randomly adding some connections on top of the existing connections to the structural network and redoing the analysis, we observed that the fluctuation in goodness of fit was relatively small (see SI Figure 26).

Fourthly, we have chosen the *alpha2* band because of high SNR for this frequency band. The mapping between structure and function may be different in terms of coefficients for the other frequency bands because we face there to some extent a different structure in the matrices. To explore the mapping for different frequency bands is a goal for future studies. Since the PLI probably underestimates the connectivity strengths (Stam et al., 2007), future research should apply our methods on other connectivity measures as well which will probably lead to different mappings in terms of different coefficients. Previous studies have used the amplitude envelope correlation to study MEG/fMRI similarity (Brookes et al., 2011). This metric

may be used in future studies to analyze structural versus functional network mappings but this is beyond the scope of the present study.

5) Conclusion

In the present study, we have demonstrated that, irrespective of the functional imaging modality, the relationship between structural and functional networks can be described by a mapping. Such a mathematical function can predict resting-state functional networks from the structural network and vice-versa. This mathematical function can be described by a weighted sum of matrix powers which represent in the binary case the number of walks up to a certain hopcount in the network. Thus, according to our analysis, a functional connection seems to be shaped by shorter walks up to the diameter in the underlying structural network. This result provides a general framework that incorporates previously published models on the relationship between structural and (stationary) functional networks. Also when analyzing the mapping from functional to structural networks, longer walks in the functional brain network appear not to have a big influence on the structural connections. We found different coefficients for MEG and fMRI for our mapping, which point towards a modality dependency for the structure-function relationship. Furthermore, this mathematical function could help to reduce noise and artifacts for the empirical estimation of structural and functional networks. We were also able to extend this mapping relationship to the subject level. For future work, differences in individual mappings between patients and healthy controls may provide insights in the disrupted relationship between the structural and functional brain networks in various diseases.

Acknowledgments

This work was partially supported by a private sponsorship to the VUmc MS center Amsterdam. The VUmc MS center Amsterdam is sponsored through a program grant by the Dutch MS Research Foundation (grant number 09-358d). We thank Menno Schoonheim for acquisition and post-processing of fMRI data. We also thank Gaolang Gong for providing structural network data. We thank Cornelis J. Stam for his useful comments and input that improved the paper.

Author Disclosure Statement

No competing financial interests exist.

References

Abdelnour, F., Voss, H. U., and Raj, A. (2014). Network diffusion accurately models the relationship between structural and functional brain connectivity networks. *NeuroImage*, 90:335-347.

Alexander-Bloch, A. F., Vértes, P. E., Stidd, R., Lalonde, F., Clasen, L., Rapoport, J., Giedd, J., Bullmore, E. T., and Gogtay, N. (2013). The anatomical distance of functional connections predicts brain network topology in health and schizophrenia. *Cerebral cortex*, 23(1):127-138.

Buck, R. (1978). *Advanced Calculus*. McGraw-Hill Book Company, New York.

Bullmore, E., and Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*(5), 336-349.

Brookes, M. J., Woolrich, M., Luckhoo, H., Price, D., Hale, J. R., Stephenson, M. C., Barnes, G. R., Smith, S. M. and Morris, P. G. (2011). Investigating the electrophysiological basis of resting state networks using magnetoencephalography. *Proceedings of the National Academy of Sciences*, *108*(40), 16783-16788.

Daducci, A., Gerhard, S., Griffa, A., Lemkaddem, A., Cammoun, L., Gigandet, X., Meuli, R., Hagmann, P., and Thiran, J.-P. (2012). The connectome mapper: an open-source processing pipeline to map connectomes with MRI. *PloS one*, 7(12):e48121.

Damoiseaux, J. S. and Greicius, M. D. (2009). Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity. *Brain Structure and Function*, 213(6):525-533.

Deco, G., McIntosh, A. R., Shen, K., Hutchison, R. M., Menon, R. S., Everling, S., Hagmann, P., and Jirsa, V. K. (2014). Identification of optimal structural connectivity using functional connectivity and neural modeling. *The Journal of Neuroscience*, 34(23):7910-7916.

Deco, G., Senden, M., and Jirsa, V. (2012). How anatomy shapes dynamics: a semi-analytical study of the brain at rest by a simple spin model. *Frontiers in computational neuroscience*, 6.

Douw, L., DeSalvo, M. N., Tanaka, N., Cole, A. J., Liu, H., Reinsberger, C., and Stufflebeam, S. M. (2015). Dissociated multimodal hubs and seizures in temporal lobe epilepsy. *Annals of Clinical and Translational Neurology*, 2(4):338-352.

Fischl, B., Sereno, M. I., and Dale, A. M. (1999). Cortical surface-based analysis: II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195-207.

Gerhard, S., Daducci, A., Lemkaddem, A., Meuli, R., Thiran, J.-P., and Hagmann, P. (2011). The connectome viewer toolkit: an open source framework to manage, analyze, and visualize connectomes. *Frontiers in Neuroinformatics*, 5(3).

Golub, G. H. and Loan, C. F. V. (1996). *Matrix Computations*. The John Hopkins University Press, Baltimore, third edition.

Gong, G., He, Y., Concha, L., Lebel, C., Gross, D. W., Evans, A. C., and Beaulieu, C. (2009). Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cerebral cortex*, 19(3):524-536.

Goñi, J., van den Heuvel, M. P., Avena-Koenigsberger, A., de Mendizabal, N. V., Betzel, R. F., Griffa, A., Hagmann, P., Corominas-Murtra, B., Thiran, J.-P., and Sporns, O. (2014). Resting-brain functional connectivity predicted by analytic measures of network communication. *Proceedings of the National Academy of Sciences*, 111(2):833-838.

Hämäläinen, M. S. and Ilmoniemi, R. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. Medical & Biological Engineering & Computing, 32(1):35{42.

Hämäläinen, M. S. and Sarvas, J. (1987). Feasibility of the homogeneous head model in the interpretation of neuromagnetic fields. *Physics in Medicine and Biology*, 32(1):91.

Hermundstad, A. M., Bassett, D. S., Brown, K. S., Amino_, E. M., Clewett, D., Freeman, S., Frithsen, A., Johnson, A., Tipper, C. M., Miller, M. B., Grafton, S., and Carlson, J. (2013). Structural foundations of resting-state and task-based functional connectivity in the human brain. *Proceedings of the National Academy of Sciences*, 110(15):6169-6174.

Higham, N. J. (2008). *Functions of matrices: theory and computation*. Siam.

Honey, C., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.-P., Meuli, R., and Hagmann, P. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences*, 106(6):2035-2040.

Honey, C. J., Kötter, R., Breakspear, M., and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proceedings of the National Academy of Sciences, 104(24):10240-10245.

Markushevich, A. I. (1985). *Theory of functions of a complex variable*, volume I - III. Chelsea Publishing Company, New York.

Messé, A., Rudrauf, D., Benali, H., and Marrelec, G. (2014). Relating structure and function in the human brain: Relative contributions of anatomy, stationary dynamics, and non-stationarities. *PLoS computational biology*, 10(3):e1003530.

Meyer, C. D. (2000). *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia.

Price, C. J., and Friston, K. J. (2002). Degeneracy and cognitive anatomy. *Trends in cognitive sciences*, *6*(10), 416-421.

Robinson, P. (2012). Interrelating anatomical, effective, and functional brain connectivity using propagators and neural field theory. *Physical Review E*, 85(1):011912.

Robinson, P., Sarkar, S., Pandejee, G., and Henderson, J. (2014). Determination of effective brain connectivity from functional connectivity with application to resting state connectivities. *Physical Review E*, 90(1):012707.

Senden, M., Deco, G., de Reus, M. A., Goebel, R., and van den Heuvel, M. P. (2014). Rich club organization supports a diverse set of functional network configurations. *NeuroImage*, 96:174-182.

Skudlarski, P., Jagannathan, K., Calhoun, V. D., Hampson, M., Skudlarska, B. A., and Pearlson, G. (2008). Measuring brain connectivity: diffusion tensor imaging validates resting state temporal correlations. *NeuroImage*, 43(3):554{561.

Stam, C. J. (2014). Modern network science of neurological disorders. *Nature Reviews Neuroscience*, 15(10):683-695.

Stam, C. J., Hillebrand, A., van Dellen, E., Meier, J., Tewarie, P., van Straaten, E., and Van Mieghem, P. (2015). The relation between structural and functional connectivity patterns in complex brain networks. International Journal of Psychophysiology.

Stam, C. J., Nolte, G., and Daffertshofer, A. (2007). Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Human Brain Mapping*, 28(11):1178-1193.

Tewarie, P., Hillebrand, A., van Dellen, E., Schoonheim, M., Barkhof, F., Polman, C., Beaulieu, C., Gong, G., van Dijk, B., and Stam, C. (2014). Structural degree predicts functional network connectivity: A multimodal resting-state fMRI and MEG study. *NeuroImage*, 97:296-307.

Tewarie, P., Schoonheim, M. M., Schouten, D. I., Polman, C. H., Balk, L. J., Uitdehaag, B. M., Geurts, J. J., Hillebrand, A., Barkhof, F., and Stam, C. J. (2015). Functional brain networks: linking thalamic atrophy to clinical disability in multiple sclerosis, a multimodal fMRI and MEG study. *Human Brain Mapping*, 36(2):603-618.

Thomas, C., Frank, Q. Y., Irfanoglu, M. O., Modi, P., Saleem, K. S., Leopold, D. A., and Pierpaoli, C. (2014). Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. *Proceedings of the National Academy of Sciences*, 111(46):16574-16579.

Titchmarsh, E. C. (1939). *The theory of functions*.

Ton, R., Deco, G. and Daffertshofer, A. (2014). Structure-Function Discrepancy: Inhomogeneity and Delays in Synchronized Neural Networks. *PLOS Computational Biology,* e1003736.

van den Heuvel, M. P., Mandl, R. C., Kahn, R. S., Pol, H., and Hilleke, E. (2009). Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Human Brain Mapping*, 30(10):3127-3141.

Van Mieghem, P. (1996). The asymptotic behavior of queueing systems: Large deviations theory and dominant pole approximation. *Queueing Systems*, 23(1-4):27-55.

Van Mieghem, P. (2011). *Graph Spectra for Complex Networks*. Cambridge University Press.

Van Mieghem, P. (2014). *Performance Analysis of Complex Networks and Systems*. Cambridge University Press.

Vértes, P. E., Alexander-Bloch, A. F., Gogtay, N., Giedd, J. N., Rapoport, J. L., and Bullmore, E. T. (2012). Simple models of human brain functional networks. *Proceedings of the National Academy of Sciences*, 109(15):5868-5873.

Whittaker, E. T. and Watson, G. N. (1996). *A Course of Modern Analysis*. Cambridge University Press, Cambridge, UK, cambridge mathematical library edition.

Figures

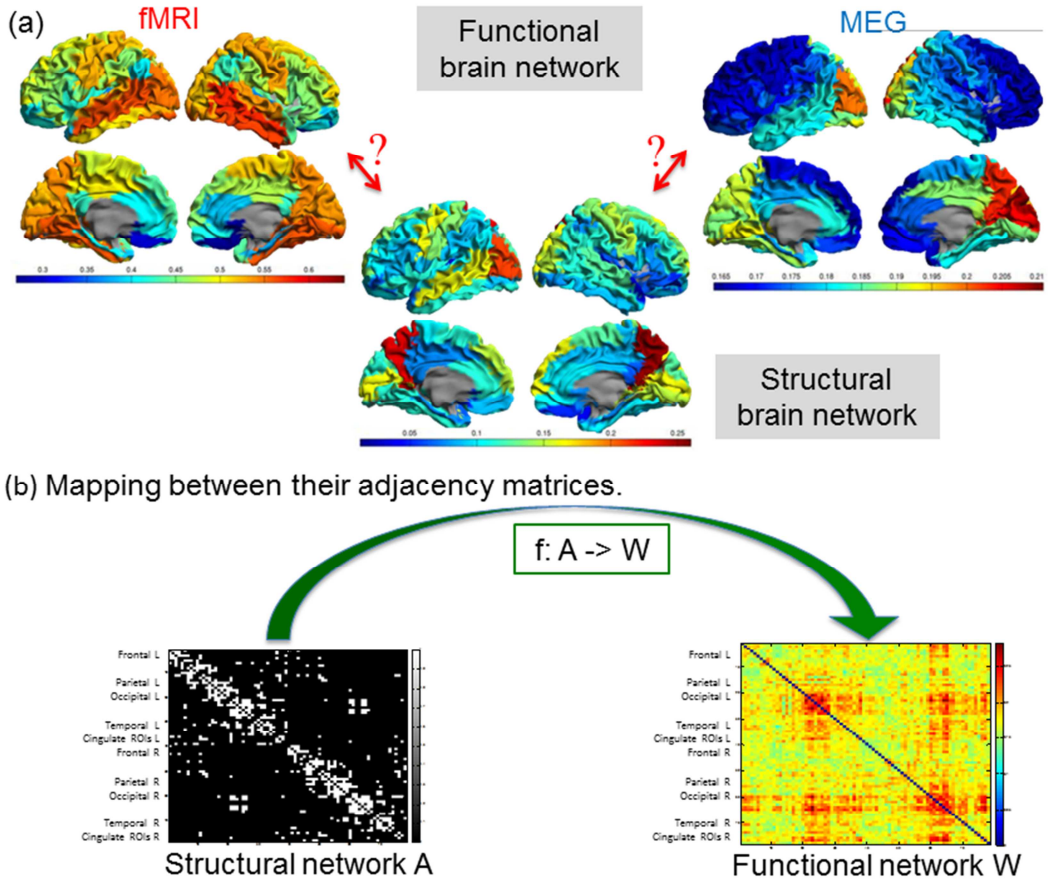


Figure 1: *(a) Visualization of the structural and functional brain network (for fMRI and MEG) for the group-averaged data set, the colors of the different regions represent here their node strength (i.e. the sum of their surrounding link weights). (b) Visualization of the mapping between their adjacency matrices.*

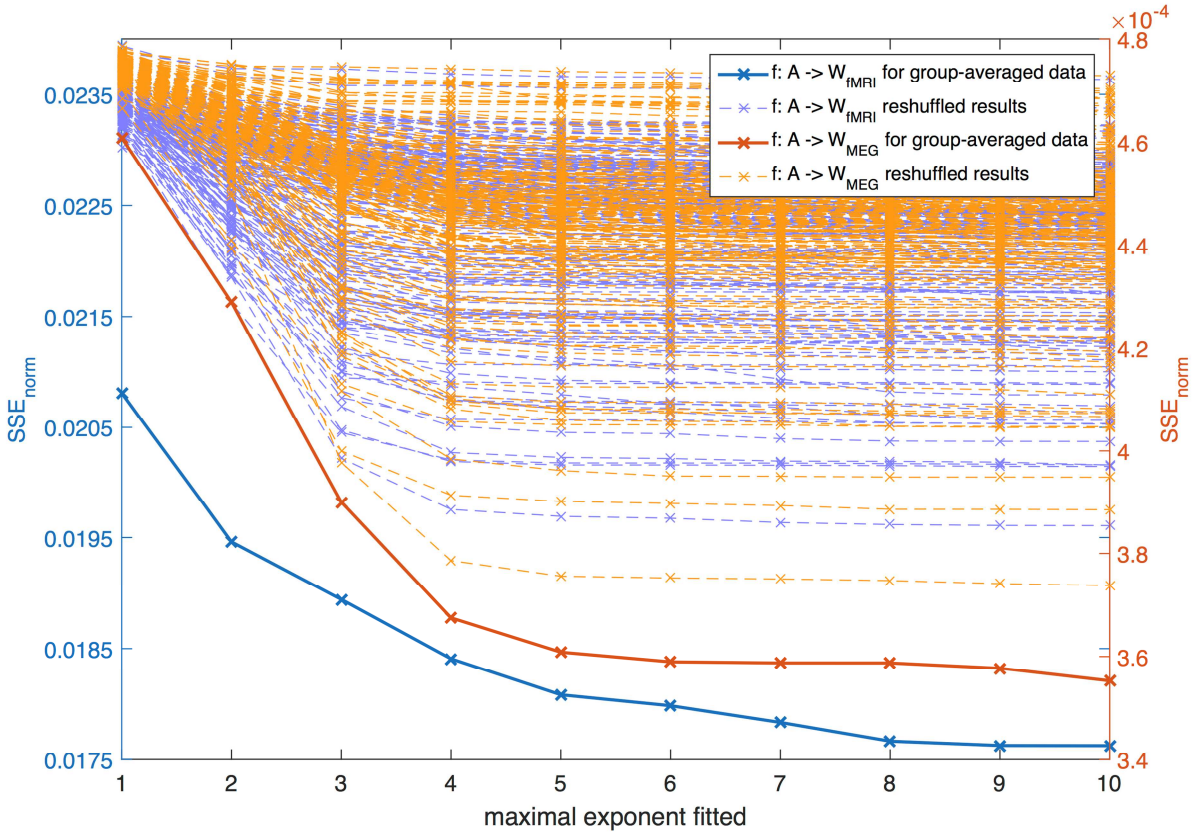


Figure 2: *Visualization of the fitted matrices for different maximal fitted exponents K (abbreviation: maxexp) for the function f : A → W_{fMRI} and f : A → W_{MEG} vs. the empirical matrices for the group-averaged data set.*

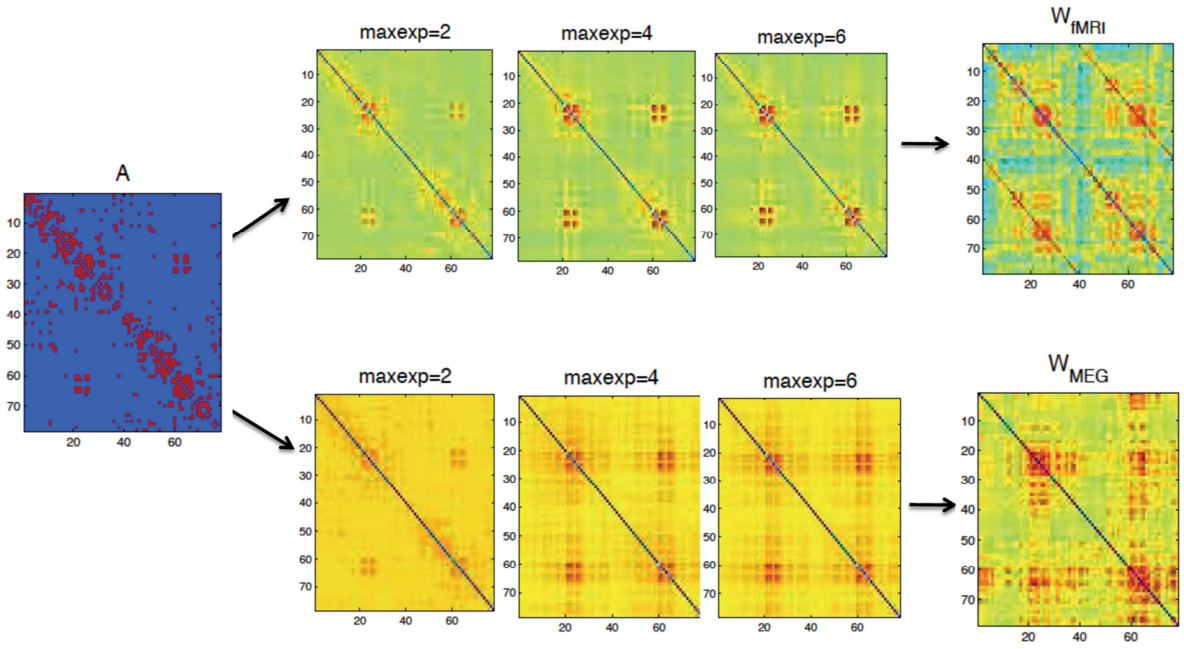


Figure 3: *SSE_{norm} for the group-averaged data set for different maximally fitted exponents K displayed together with the results of the reshuffled matrices. For each mapping we ran the same analysis with 100 reshuffled versions of the matrix A and with 100 reshuffled versions of matrix W.*

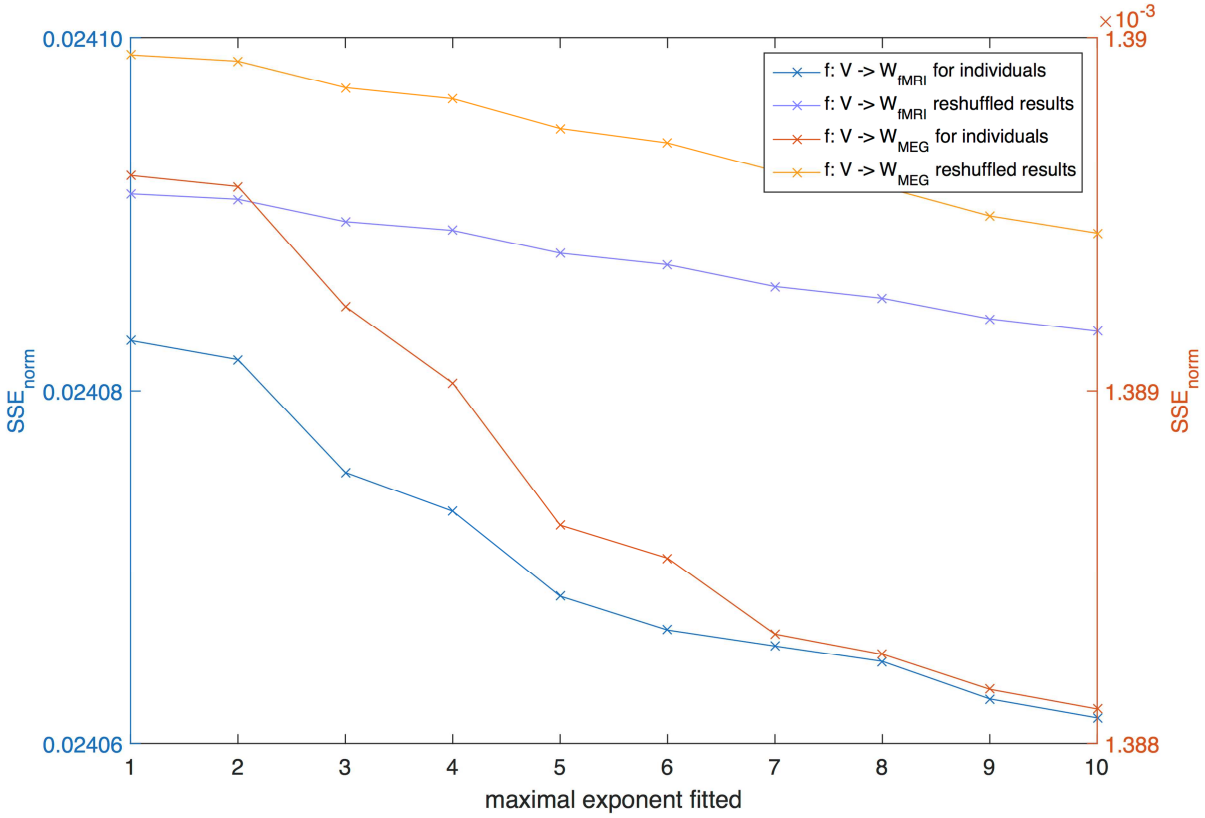


Figure 4: *SSE_{norm} for the individual data set for different maximally fitted exponents K (after averaging over all 11 individual SSE_{norm} results) displayed together with the averaged result of the reshuffled matrices. For each mapping we ran the same analysis with 100 reshuffled versions of the matrix V and with 100 reshuffled versions of matrix W.*

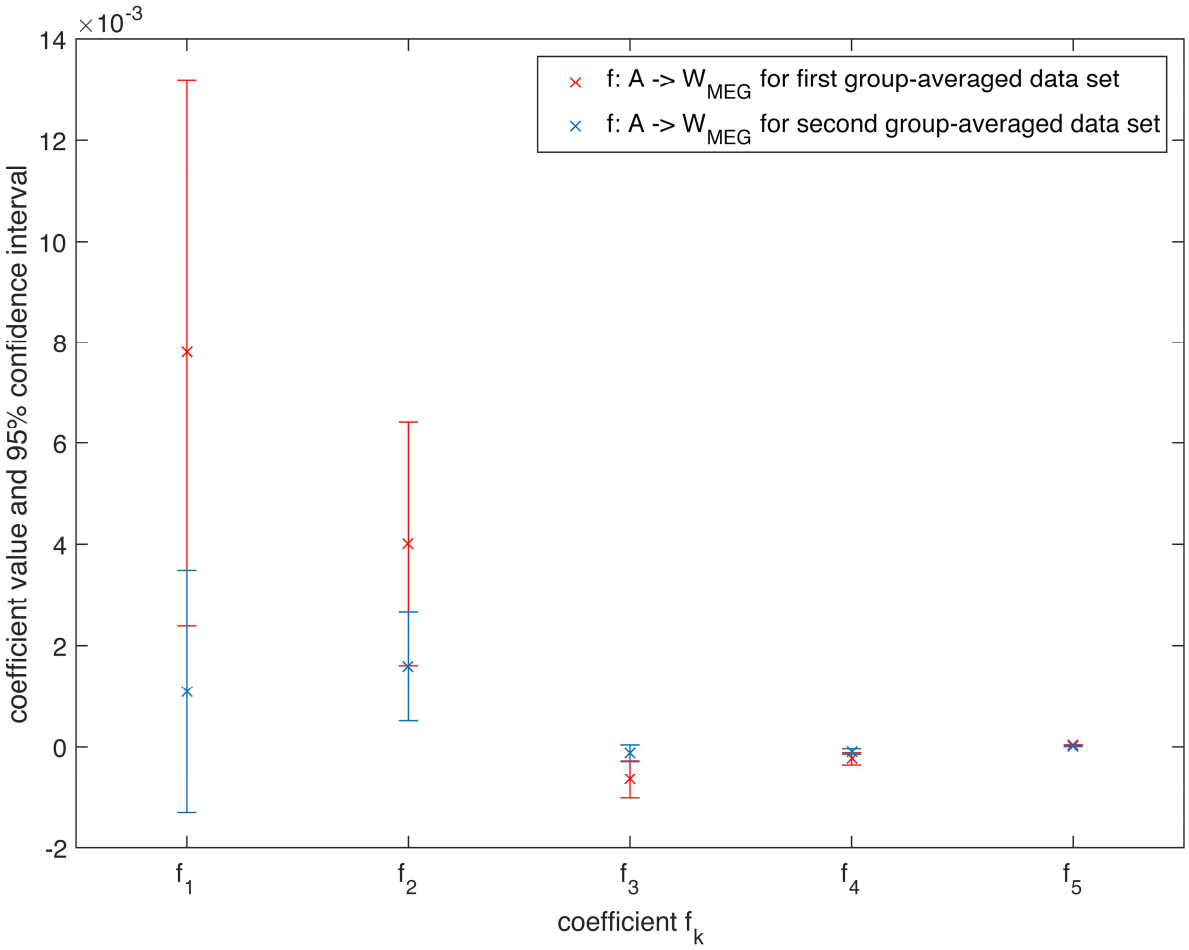


Figure 5: *Estimated coefficients for the mapping f : A → W_{MEG} for K=5 together with their 95% confidence interval for the first group-averaged data set and a second group-averaged data set.*

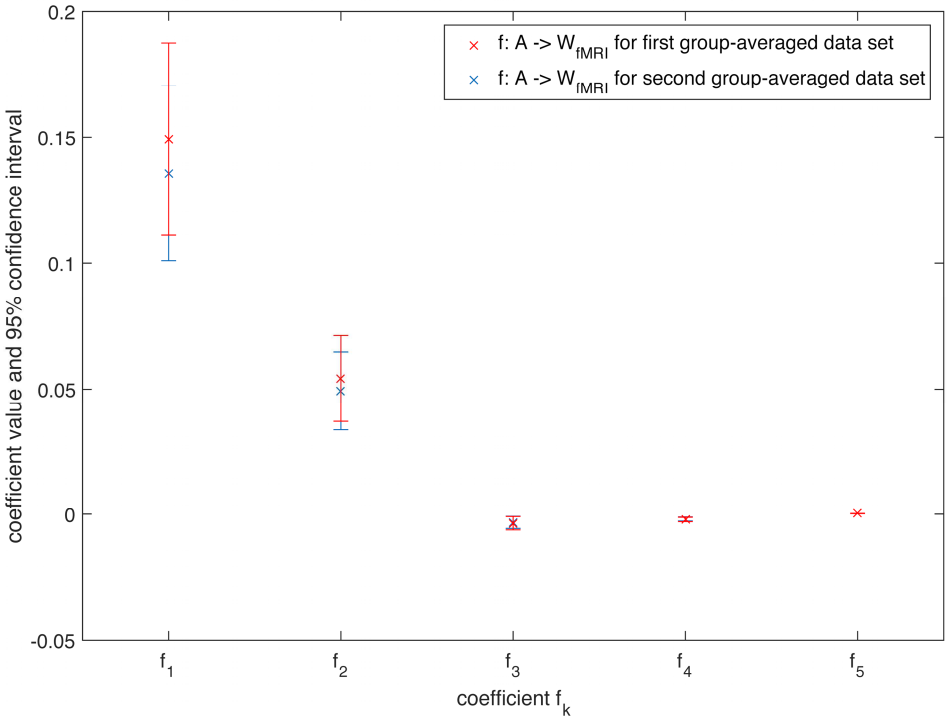


Figure 6: *Estimated coefficients for the mapping f : A → W_{fMRI} for K=5 together with their 95% confidence interval for the first group-averaged data set and a second group-averaged data set.*

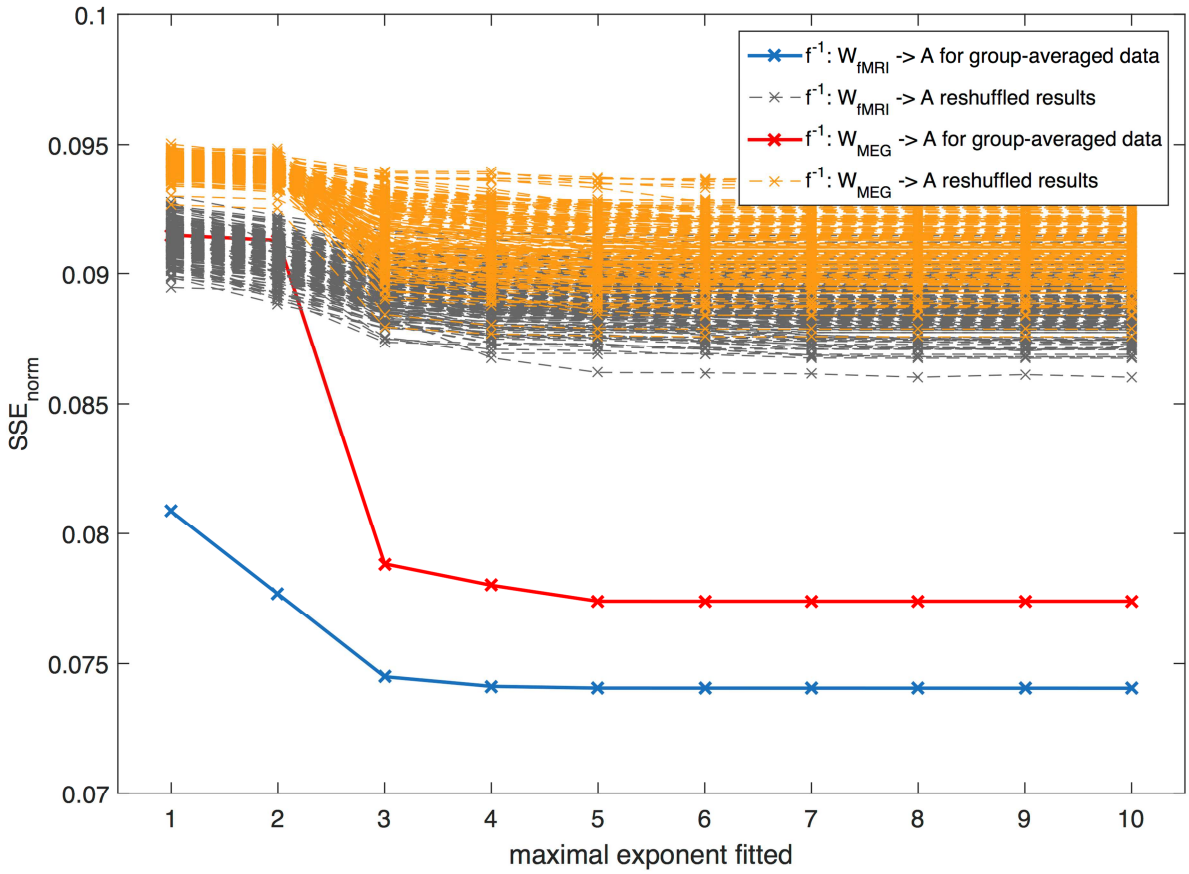


Figure 7: *Visualization of the fitted matrices for different maximal fitted exponents K (abbreviation: maxexp) for the function* $f^{-1} : W_{fMRI} \rightarrow A$ *and* $f^{-1} : W_{MEG} \rightarrow A$ *vs. the empirical matrices for the group-averaged data set.*

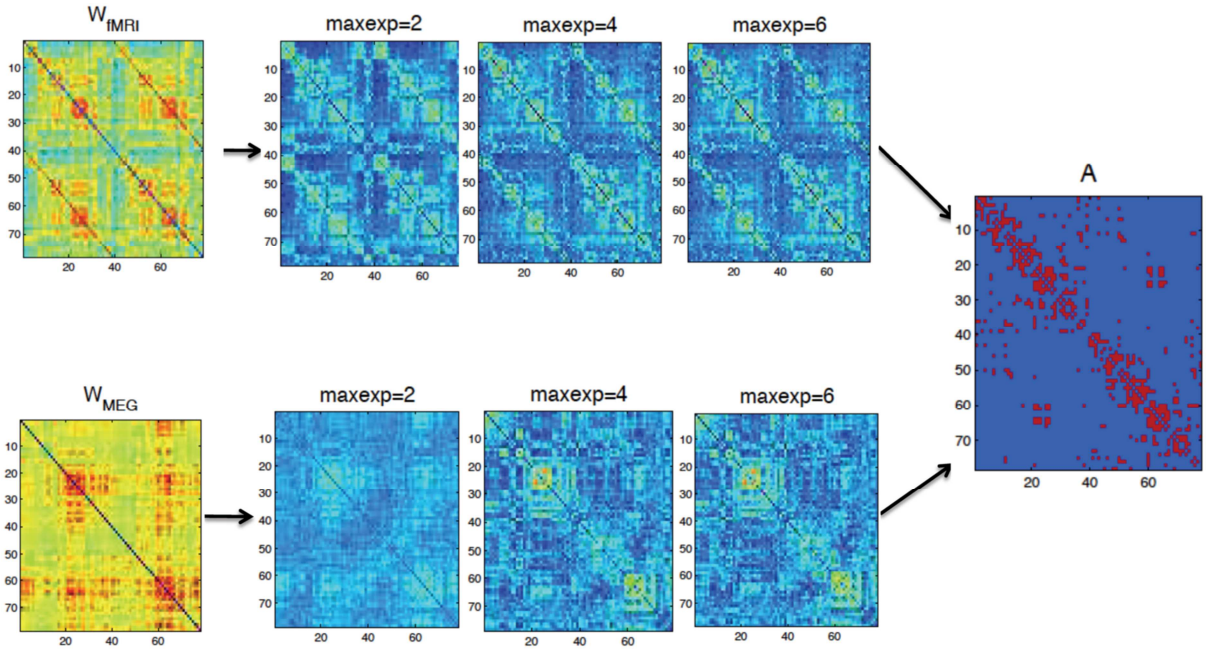


Figure 8: *SSE_{norm} for the group-averaged data set for different maximally fitted exponents K displayed together with the results of the reshuffled matrices. For each mapping we ran the same analysis with 100 reshuffled versions of the matrix A and with 100 reshuffled versions of matrix W.*

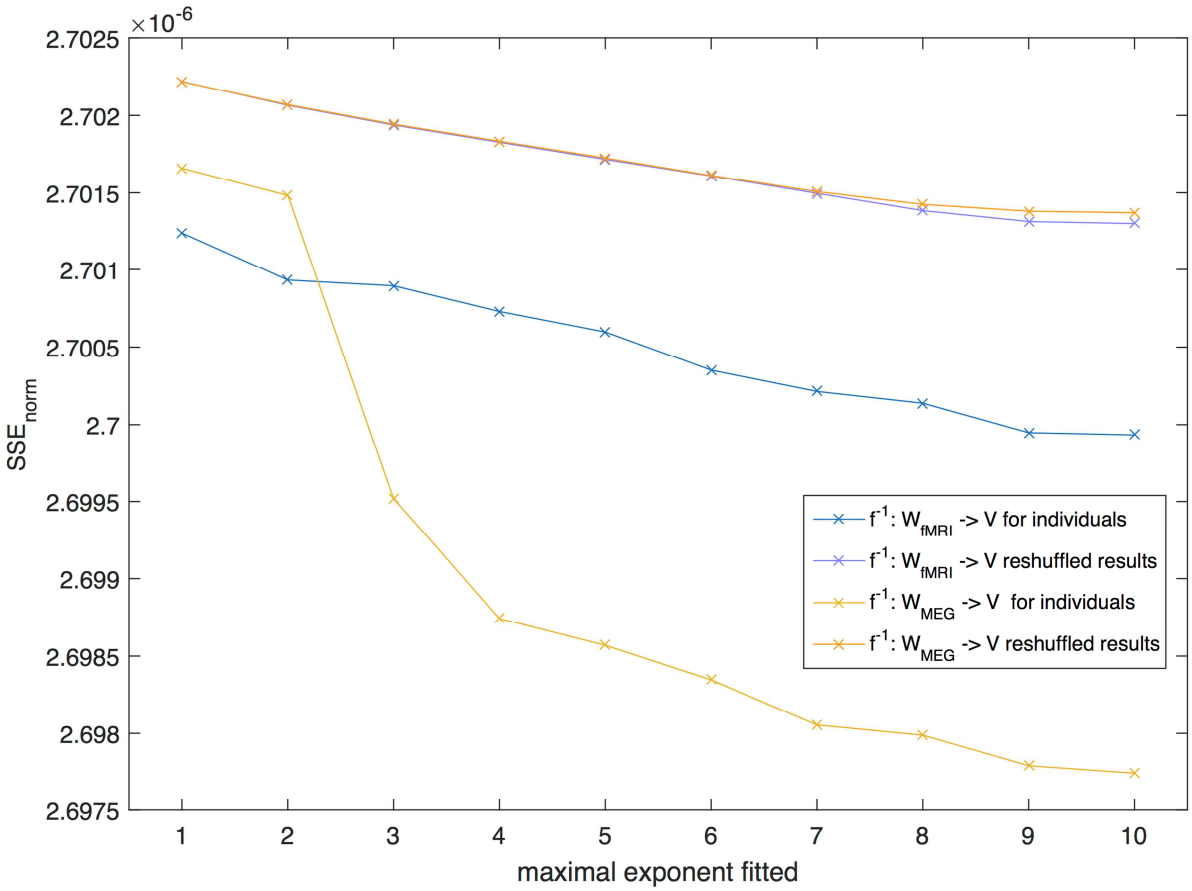


Figure 9: *SSE_{norm} for the individual data set for different maximally fitted exponents K (after averaging over all 11 individual SSE_{norm} results) displayed together with the averaged result of the reshuffled matrices. For each mapping we ran the same analysis with 100 reshuffled versions of the matrix V and with 100 reshuffled versions of matrix W.*

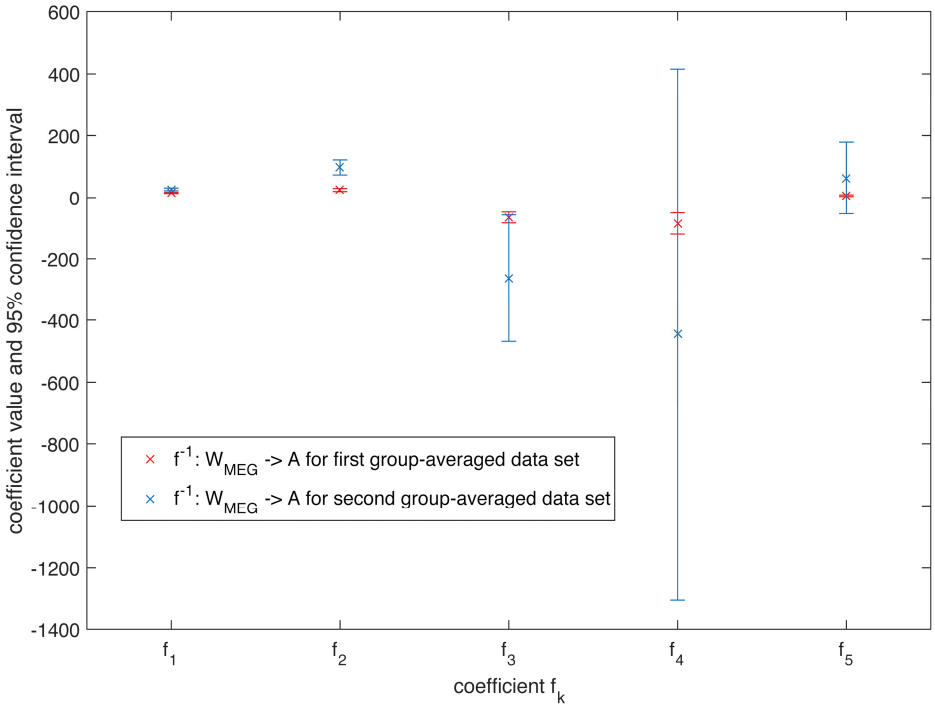


Figure 10: *Estimated coefficients for the mapping $f^{-1} : W_{MEG} \rightarrow A$ for K=5 together with their 95% confidence interval for the first group-averaged data set and a second group-averaged data set.*

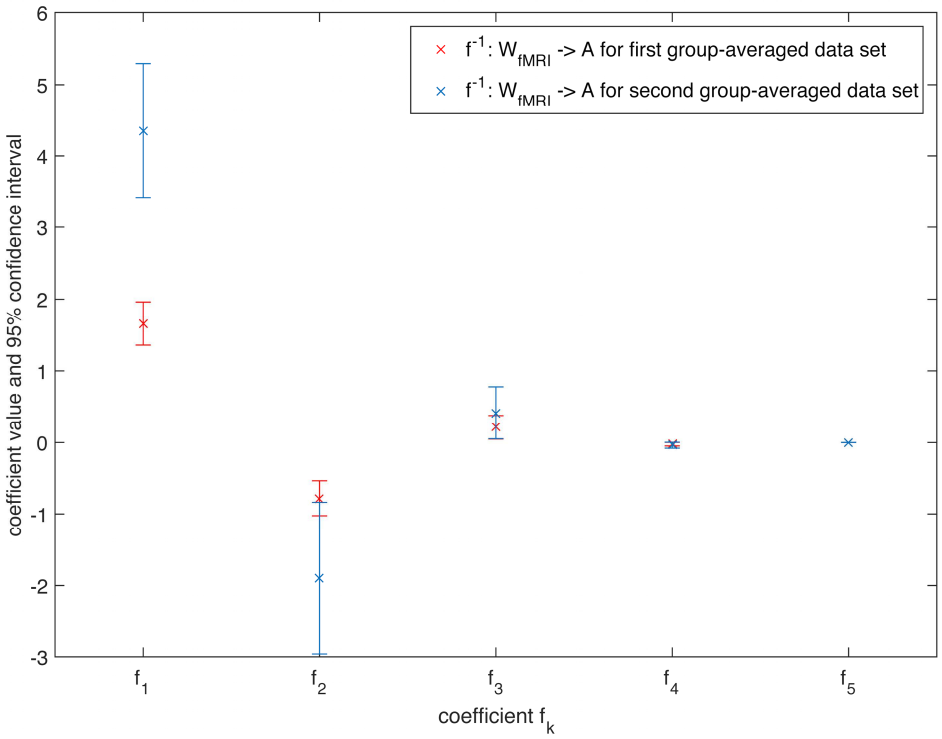


Figure 11: *Estimated coefficients for the mapping $f^{-1} : W_{fMRI} \rightarrow A$ for K=5 together with their 95% confidence interval for the first group-averaged data set and a second group-averaged data set.*

Table 1: p-values for the comparison between SSE_{norm} for the empirical and reshuffled matrices.

Mapping	p1	p2	p3	p4	p5	p6
$f^{-1} : W_{fMRI} \rightarrow V$	.887	<.001*	<.001*	<.001*	.002*	.003*
$f : V \rightarrow W_{fMRI}$	.001*	<.001*	<.001*	<.001*	.011*	<.001*
$f^{-1} : W_{MEG} \rightarrow V$	.001*	<.001*	<.001*	<.001*	<.001*	<.001*
$f : V \rightarrow W_{MEG}$	.339	<.001*	.018*	<.001*	<.001*	.827

	p7	p8	p9	p10	p11
$f^{-1} : W_{fMRI} \rightarrow V$	<.001*	.390	<.001*	<.001*	<.001*
$f : V \rightarrow W_{fMRI}$	<.001*	<.001*	<.001*	<.001*	<.001*
$f^{-1} : W_{MEG} \rightarrow V$	<.001*	<.001*	.001*	<.001*	.002*
$f : V \rightarrow W_{MEG}$	<.001*	.975	.815	<.001*	.130

Table 1: *p-values for the comparison between SSE_{norm} for the empirical and reshuffled matrices. The matrix V denotes the structural network matrices for the individual data and the different columns are for the different 11 analyzed persons (p1 till p11). Note that in most cases a significantly better goodness-of-fit was obtained for the empirical matrices than for the reshuffled matrices (p<.05, indicated with *).*

A mapping between structural and functional brain networks: **Supplementary Information**

J. Meier[*], P. Tewarie[†], A. Hillebrand[‡], L. Douw[§,¶], B.W. van Dijk[‡], S.M. Stufflebeam[¶,‖], and P. Van Mieghem[*]

Contents

[*]Faculty of Electrical Engineering, Mathematics and Computer Science, P.O Box 5031, 2600 GA Delft, The Netherlands; *email*: J.M.Meier@tudelft.nl (corresponding author), P.F.A.VanMieghem@tudelft.nl

[†]Department of Neurology, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, The Netherlands; *email*: p.tewarie@vumc.nl

[‡]Department of Clinical Neurophysiology and Magnetoencephalography Center, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, The Netherlands; *email*: a.hillebrand@vumc.nl, bw.dijk@vumc.nl

[§]Department of Anatomy and Neurosciences, VU University Medical Center, Amsterdam, The Netherlands, *email*: douw@nmr.mgh.harvard.edu

[¶]Department of Radiology, Athinoula A. Martinos Center for Biomedical Imaging / Massachusetts General Hospital, Boston MA.

[‖]Department of Radiology, Harvard Medical School, Boston MA, *email*: sms@nmr.mgh.harvard.edu

A Theory

We provide here some mathematical background of our method.

A.1 Matrix functions

If $f(z)$ is a function of the complex number z and analytic in a disk with radius R around z_0, then it can be shown (see Section 2.2) that for all matrices A, the matrix function $f(A)$ can be expressed as a polynomial of degree at most $N-1$,

$$f(A) = \sum_{k=0}^{N-1} c_k[f] A^k$$

where the coefficients $c_k[f]$ can be specified as

$$c_k[f] = \frac{1}{k!} \sum_{m=1}^{N} \frac{f(\lambda_m)}{\prod_{j=1;j\neq m}^{n} (\lambda_m - \lambda_j)} \frac{d^k}{dx^k} \prod_{j=1;j\neq m}^{N} (x - \lambda_j) \Bigg|_{x=0}$$

Only if f is a polynomial of degree m at most $N-1$, we find, for all $0 \leq k \leq N-1$, that

$$f_k(0) = c_k[f],$$

where $f_k(0)$ is the k-th coefficient of the Taylor series of f for the development point $z_0 = 0$.

A.2 Role of the diameter

An explanation for the importance of the diameter could be as follows: The matrix powers I, A, A^2, A^3,... are all linearly independent of each other up to A^ρ (Van Mieghem, 2011). For higher powers, we cannot be sure of their dependency. Our analysis shows that using higher powers of the structural connectivity matrix than its diameter ($\rho = 6$) does not improve the goodness of fit of our estimation. Furthermore, in the binary matrix A, the sum $\sum_{k=1}^{m} A^k$ can have zero entries for all $m < \rho$. Reducing the number of zero entries can also be a reason why the goodness of fit increases gradually until adding A^ρ and then converges.

A.3 Analysis in the spectral domain

If a mapping postulated in the previous section is valid, then such a mapping should also hold in the spectral domain (Van Mieghem, 2011). Provided A is symmetric such that $A = X\Lambda X^T$, where the matrix X contains the eigenvectors of A in the columns and $\Lambda = \mathrm{diag}(\lambda_k)$, $1 \leq k \leq N$, with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$ the real eigenvalues of A, then there exists an alternative to compute $f(A)$. Since eigenvectors are orthogonal, the matrix X is an orthogonal matrix satisfying $X^T X = I$ and $XX^T = I$, where the latter follows from the fact that $X^{-1} = X^T$ and the fact that a matrix and its

inverse commute. We assume here that all eigenvalues are different, as is the case in most real-world networks (Van Mieghem, 2011). Then,

$$f(A) = Xf(\Lambda)X^T = \sum_{k=0}^{N} f(\lambda_k)\, x_k x_k^T \tag{1}$$

where x_k is the eigenvector of A belonging to the eigenvalue λ_k. Using the spectral form (1) in our assumption (1) reveals that

$$W = Xf(\Lambda)X^T, \tag{2}$$

Since W is symmetric, the spectral decomposition equals

$$W = Y\Upsilon Y^T \tag{3}$$

where Y is the orthogonal matrix containing the eigenvectors $y_1, y_2, \ldots, y_N$ belonging to the eigenvalues $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_N$ and $\Upsilon = \text{diag}(\mu_k), 1 \leq k \leq N$ is the diagonal matrix of the eigenvalues of W. Using the properties of orthogonal matrices, we find from (2) the diagonal matrix $f(\Lambda) = \text{diag}(f(\lambda_k))$ as

$$f(\Lambda) = X^T W X.$$

With (3), we obtain

$$f(\Lambda) = \left(X^T Y\right)\Upsilon\left(Y^T X\right). \tag{4}$$

Since both $f(\Lambda)$ and Υ are diagonal matrices and $\left(X^T Y\right) = \left(Y^T X\right)^{-1}$, there must hold that

$$f(\lambda_k) = \mu_k \tag{5}$$

for each $1 \leq k \leq N$. Furthermore, from (4) follows that

$$\left(Y^T X\right) f(\Lambda) = \Upsilon\left(Y^T X\right). \tag{6}$$

From (5) and the fact that we multiply the matrix $Y^T X$ with the same diagonal matrix from both sides, we can conclude that (under the condition that all eigenvalues are different) $Y^T X = I$ and therefore $X = Y$.

In summary, equation (1) implies that both the structural matrix A and the functional matrix W must have the same eigenvectors and that the function f maps the eigenvalues (ordered) of A onto those of W (also ordered, see (5)). Moreover, (5) shows that $f(x)$ is non-decreasing in x.

If the empirical matrices have different eigenvectors, then that difference may be due to noise of the measurement. If the difference cannot be explained by noise perturbation, our assumption in (1) that there exists an analytic function f needs to be revisited and a more general form of our fitting function can be considered with matrix coefficients instead of scalar ones

$$F(A) = \sum_{k=0}^{\infty} F_k (A - z_0 I)^k \tag{7}$$

where F_k are $N \times N$ matrices, which reduces when $N = 1$ again to our previous assumption (3).

B Data Analysis

B.1 Visualization of best fits

In Figures 12, 13 and 14, we visualized the best fitted matrices of the group-averaged data set on a bigger scale to better localize the biggest discrepancies between our best fits and the empirical findings. The numbers on the axes of the matrices, 1 to 78 refer to specific brain regions. You can find on the next page a complete list of the regions of interest (ROIs) that we display mostly with numbers:

1	Rectus-L	40	Rectus-R
2	Olfactory-L	41	Olfactory-R
3	Frontal-Sup-Orb-L	42	Frontal-Sup-Orb-R
4	Frontal-Med-Orb-L	43	Frontal-Med-Orb-R
5	Frontal-Mid-Orb-L	44	Frontal-Mid-Orb-R
6	Frontal-Inf-Orb-L	45	Frontal-Inf-Orb-R
7	Frontal-Sup-L	46	Frontal-Sup-R
8	Frontal-Mid-L	47	Frontal-Mid-R
9	Frontal-Inf-Oper-L	48	Frontal-Inf-Oper-R
10	Frontal-Inf-Tri-L	49	Frontal-Inf-Tri-R
11	Frontal-Sup-Medial-L	50	Frontal-Sup-Medial-R
12	Supp-Motor-Area-L	51	Supp-Motor-Area-R
13	Paracentral-Lobule-L	52	Paracentral-Lobule-R
14	Precentral-L	53	Precentral-R
15	Rolandic-Oper-L	54	Rolandic-Oper-R
16	Postcentral-L	55	Postcentral-R
17	Parietal-Sup-L	56	Parietal-Sup-R
18	Parietal-Inf-L	57	Parietal-Inf-R
19	SupraMarginal-L	58	SupraMarginal-R
20	Angular-L	59	Angular-R
21	Precuneus-L	60	Precuneus-R
22	Occipital-Sup-L	61	Occipital-Sup-R
23	Occipital-Mid-L	62	Occipital-Mid-R
24	Occipital-Inf-L	63	Occipital-Inf-R
25	Calcarine-L	64	Calcarine-R
26	Cuneus-L	65	Cuneus-R
27	Lingual-L	66	Lingual-R
28	Fusiform-L	67	Fusiform-R
29	Heschl-L	68	Heschl-R
30	Temporal-Sup-L	69	Temporal-Sup-R
31	Temporal-Mid-L	70	Temporal-Mid-R
32	Temporal-Inf-L	71	Temporal-Inf-R
33	Temporal-Pole-Sup-L	72	Temporal-Pole-Sup-R
34	Temporal-Pole-Mid-L	73	Temporal-Pole-Mid-R
35	ParaHippocampal-L	74	ParaHippocampal-R
36	Cingulum-Ant-L	75	Cingulum-Ant-R
37	Cingulum-Mid-L	76	Cingulum-Mid-R
38	Cingulum-Post-L	77	Cingulum-Post-R
39	Insula-L	78	Insula-R

We can see that the patterns of the fitted matrices seem to be similar to the empirical ones and that the value range is overlapping. Therefore, we can conclude that also only from visual inspections of the fitted matrices our mapping seems to be convincingly accurate.

B.2 Spectral data analysis

After plotting the eigenvalue couples (λ_k, μ_k) in a scatter plot, we can obtain the function f in the spectral domain (see (5)). Polynomial functions were fitted to all possible combinations of scatter plots (i.e. for the combinations structure-function (MEG/fMRI) and vice versa) by minimizing the sum of squared errors. An example of such a fit is depicted in Figure 15.

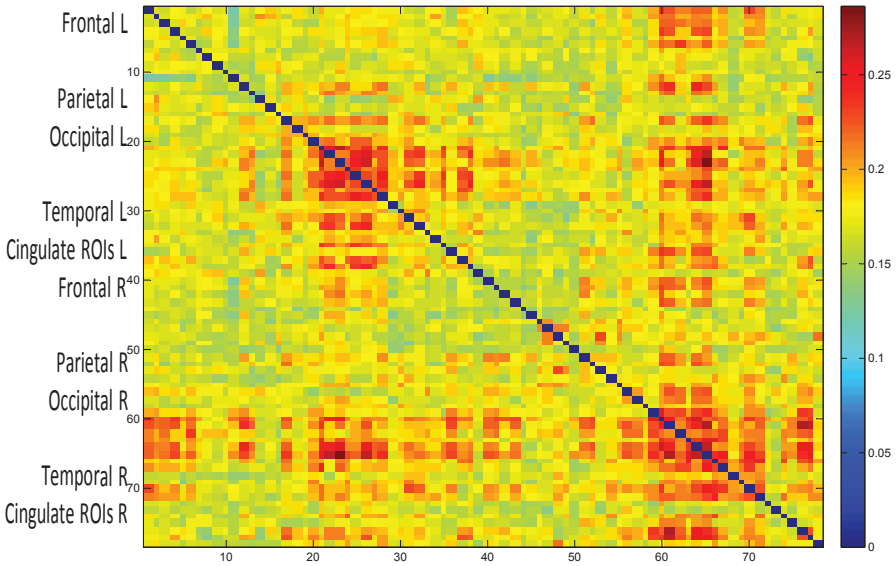
For the goodness of fit in the spectral domain, we computed the adjusted R^2 value for the different mappings (see Figure 16). Overall, we reached already for $K \geq 4$ with all combinations of matrices an adjusted R^2 value of higher than 0.9 indicating a good fit of our mapping. We followed the same approach here as we did for the topology domain, and reversed A and W to repeat the spectral analysis for the function $f^{-1} : W \to A$. Results of this analysis can also be found in Figure 16. For this spectral approach, the adjusted R^2 value did not improve much after adding the same number of terms as was used for the mathematical function in the topology domain (compare Figures 16 with Figure 2). Thus, a functional expression with $K = 5$ was again sufficient for the analyzed mappings. This conclusion held for both modalities (fMRI and MEG) and for both functions $f : A \to W$ and $f^{-1} : W \to A$.

Since we conducted a similar analysis in the topology and in the spectral domain respectively for the functions f and f^{-1}, we compared the estimated coefficient values for the spectral and topology domain. For a correct comparison, we must omit the error matrix E in (5) in the topology domain. A plot of the estimated coefficient values and confidence intervals (obtained by the least squared parameter estimation in MATLAB) is illustrated in SI Figures 17 and 18 for the mapping of structural to functional matrices and in SI Figures 19 and 20 for the other direction. For the mapping f, we only faced small differences between the coefficients (see SI Figures 17 and 18) and in most cases their confidence intervals were overlapping. But for the other direction, function f^{-1}, we observed quite different estimated coefficients. In the case of $f^{-1} : W_{MEG} \to A$ we obtained large confidence intervals when many (> 5) coefficients were fitted (see SI Figure 19), implying insecure estimations of their exact value. This result is in agreement with the finding that 5 coefficients were sufficient to describe the mapping between W and A (Figure 2), and that these extra coefficients did not contribute relevant information to the mapping. The discrepancies between the estimated coefficient values in the topology and spectral domain could be originating from the different eigenvectors of the 3 analyzed empirical matrices. These eigenvector perturbations can potentially be caused by noise in the different measurement techniques.

C Comparison of fitted coefficient values for different modalities

In Table 1, we displayed the different estimated values for the coefficients of our mapping for a maximal fitted exponent of $K = 6$ and $z_0 = 0$ for the group-averaged data set. In addition, in Figures 21 and 22 we show the different coefficient values (without the offset estimates) for the two modalities for both

(a) W_{MEG}



(b) $\widetilde{W}_{MEG}$

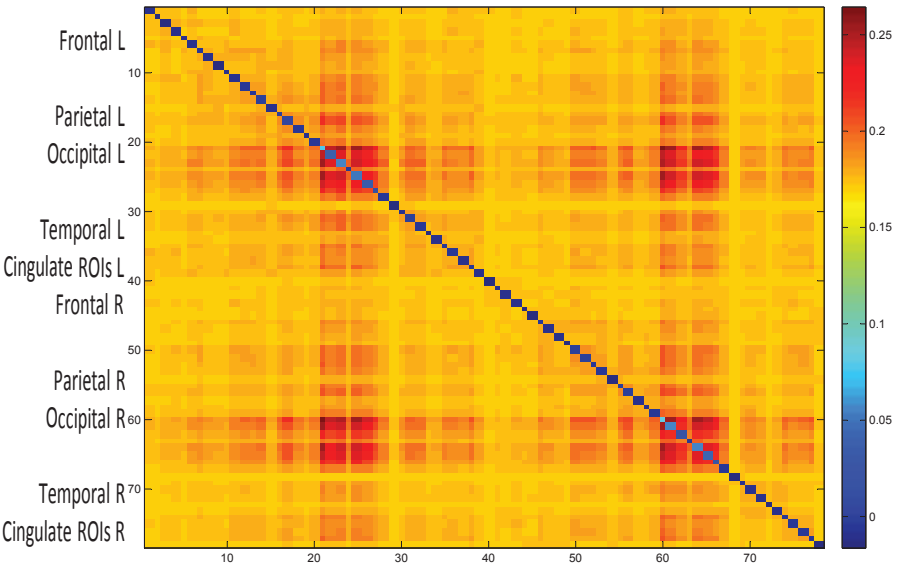
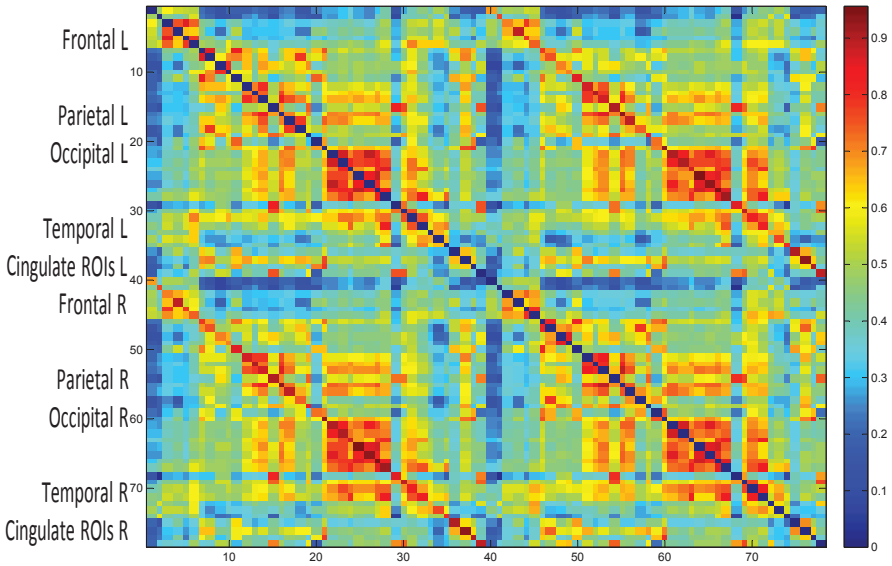


Figure 12: Visualization of the best fits for the function f, which was $\widetilde{W}_{MEG}$ (for $K = 6$ with an error matrix E), under the empirical adjacency matrix W_{MEG} for the group-averaged data set.

(a) W_{fMRI}



(b) $\widetilde{W}_{fMRI}$

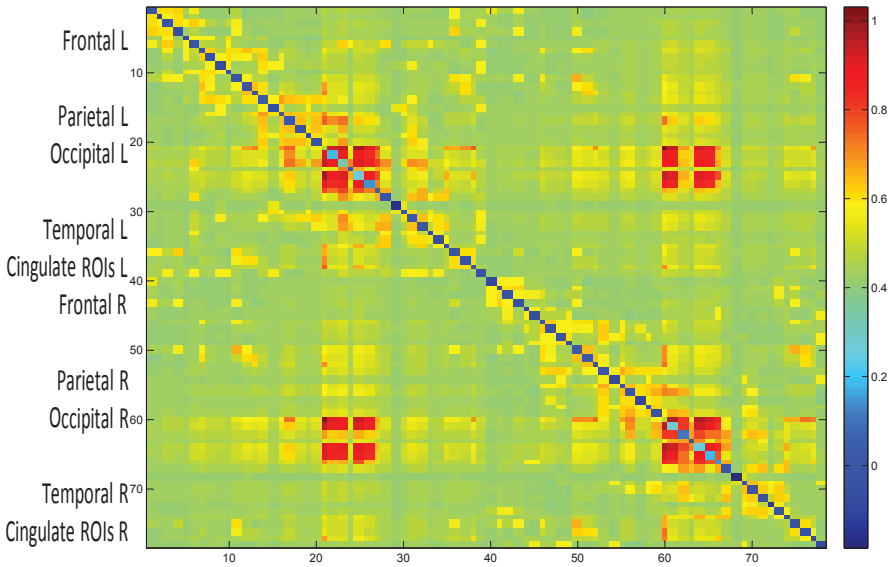
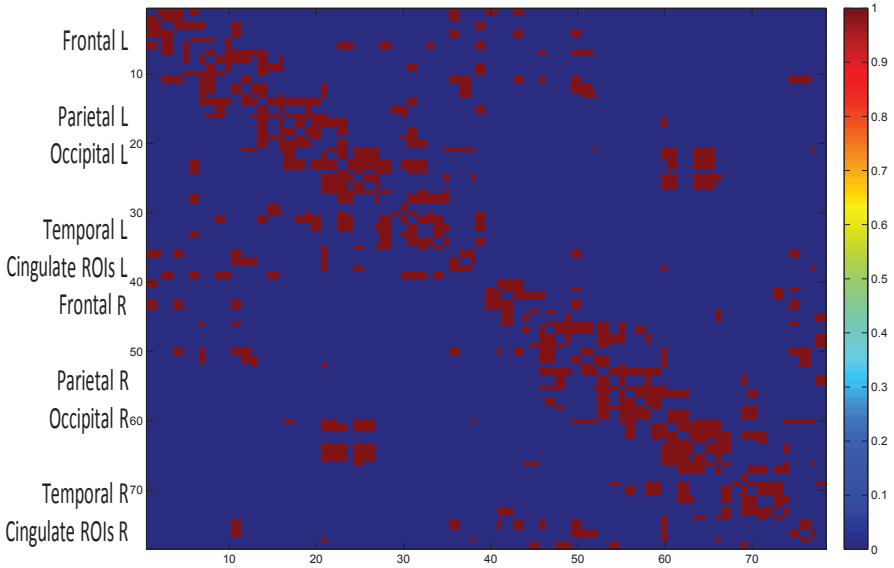


Figure 13: Visualization of the best fits for the function f, which was $\widetilde{W}_{fMRI}$ (for $K = 6$ with an error matrix E), under the empirical adjacency matrix W_{fMRI} for the group-averaged data set.

(a) A



(b) $\widetilde{A}$

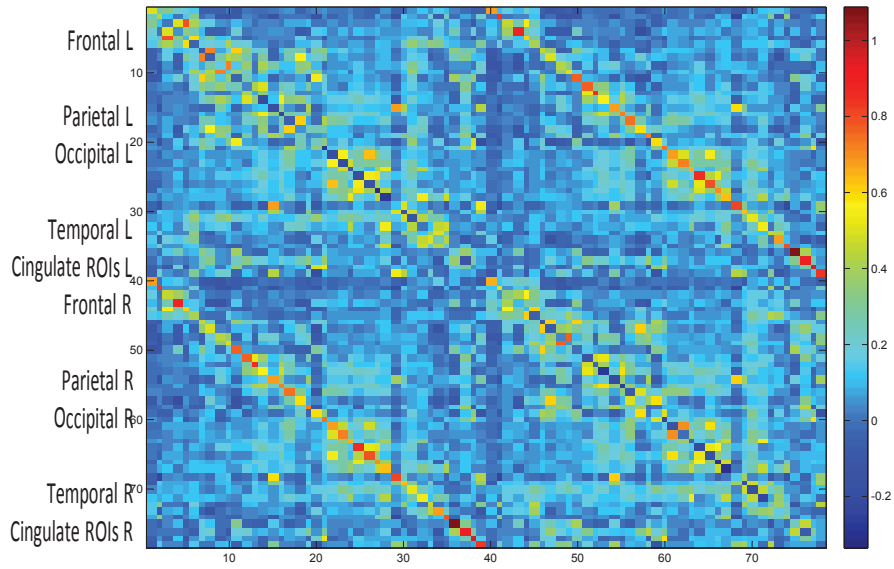


Figure 14: Visualization of the best fits for the function f^{-1}, which was $\widetilde{A}$ (again for $K = 6$ with an error matrix E) under the empirical adjacency matrix A for the group-averaged data set.

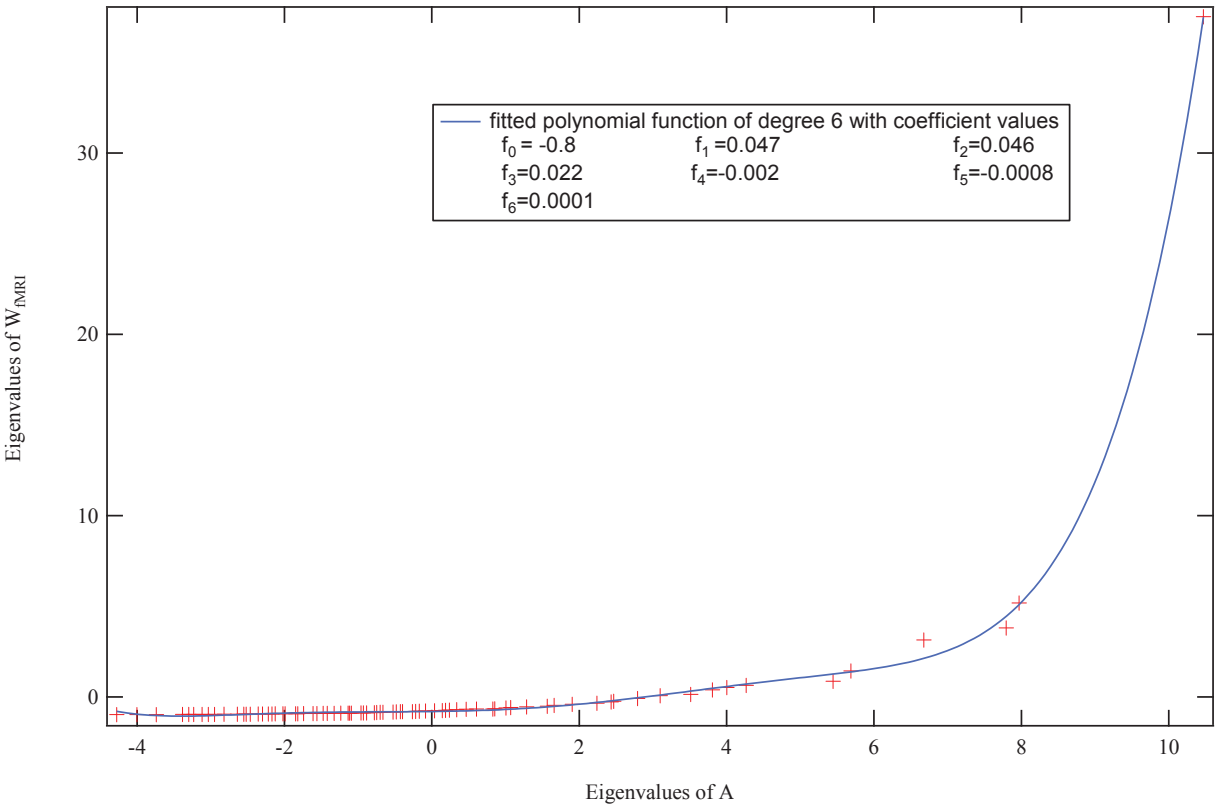


Figure 15: Scatter plot of the eigenvalues of the structural matrix A against the eigenvalues of W_{fMRI} with a least-squared fitted polynomial function with a maximal fitted exponent of $K = 6$ (including the intercept f_0) for the group-averaged data set.

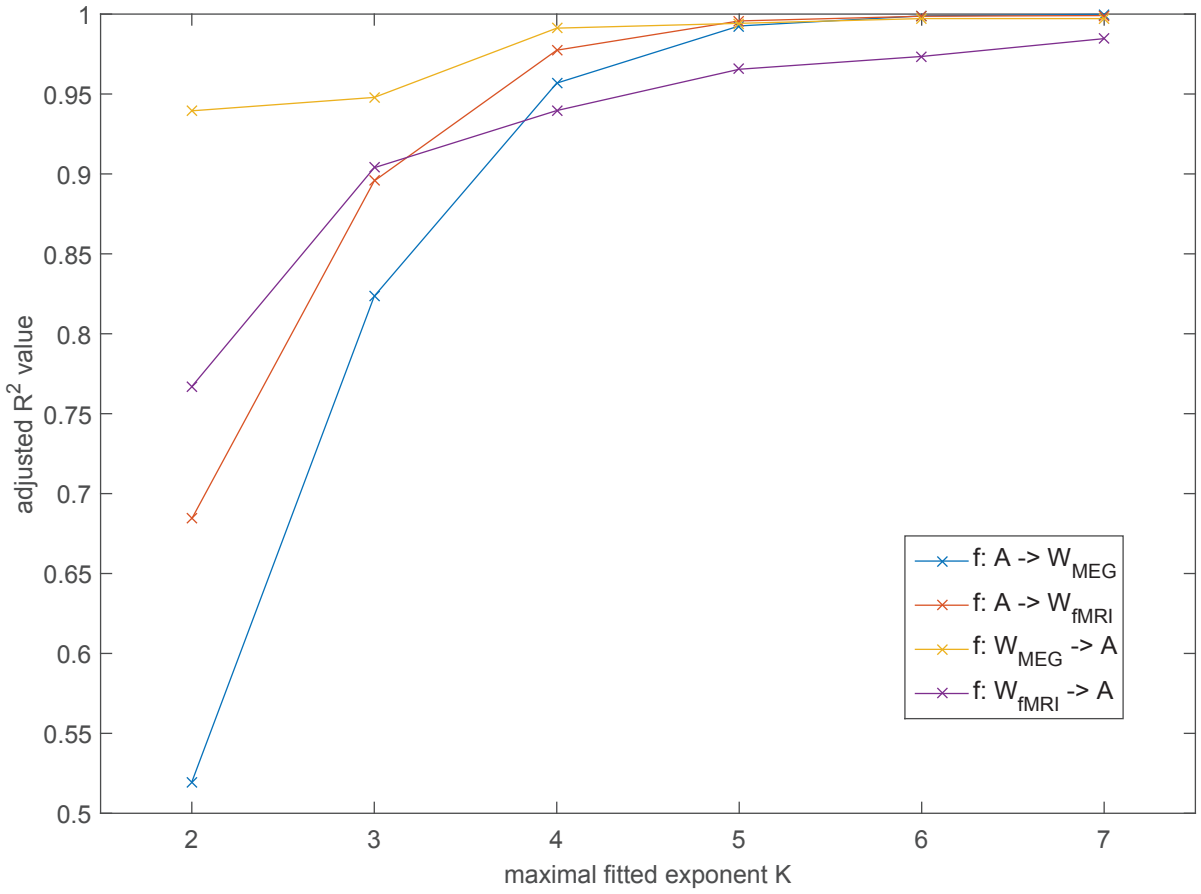


Figure 16: Adjusted R^2 value of the different mappings in the spectral domain for different maximal fitted exponents K for the group-averaged data set for $z_0 = 0$.

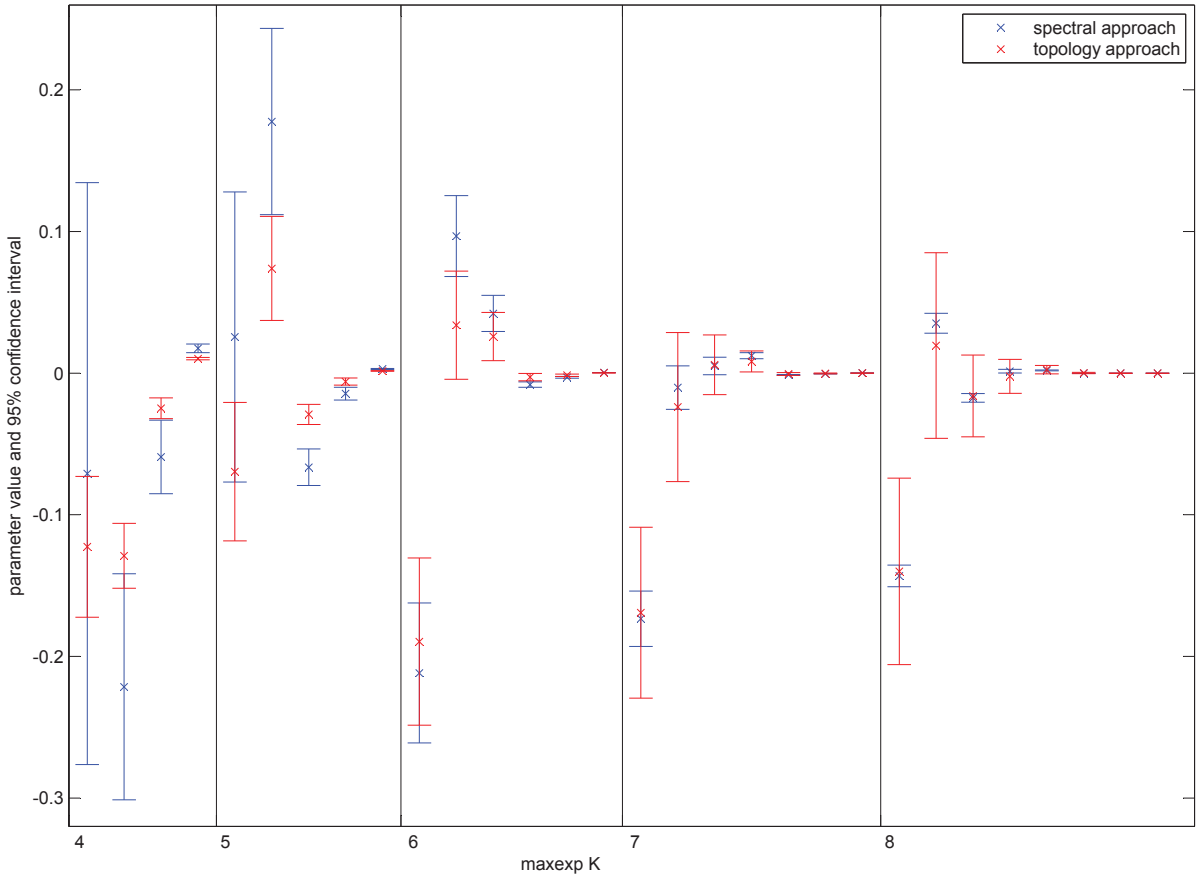


Figure 17: Plot of the estimated coefficient values for different maximal exponents K with their 95% confidence interval as an errorbar for the mapping $f : A \rightarrow W_{MEG}$ for the group-averaged data set. The spectral and topology approach are marked in blue and red, respectively. Note that, the confidence intervals of the coefficient values overlap and the distance between the estimated coefficient values is becoming smaller when more coefficients are used.

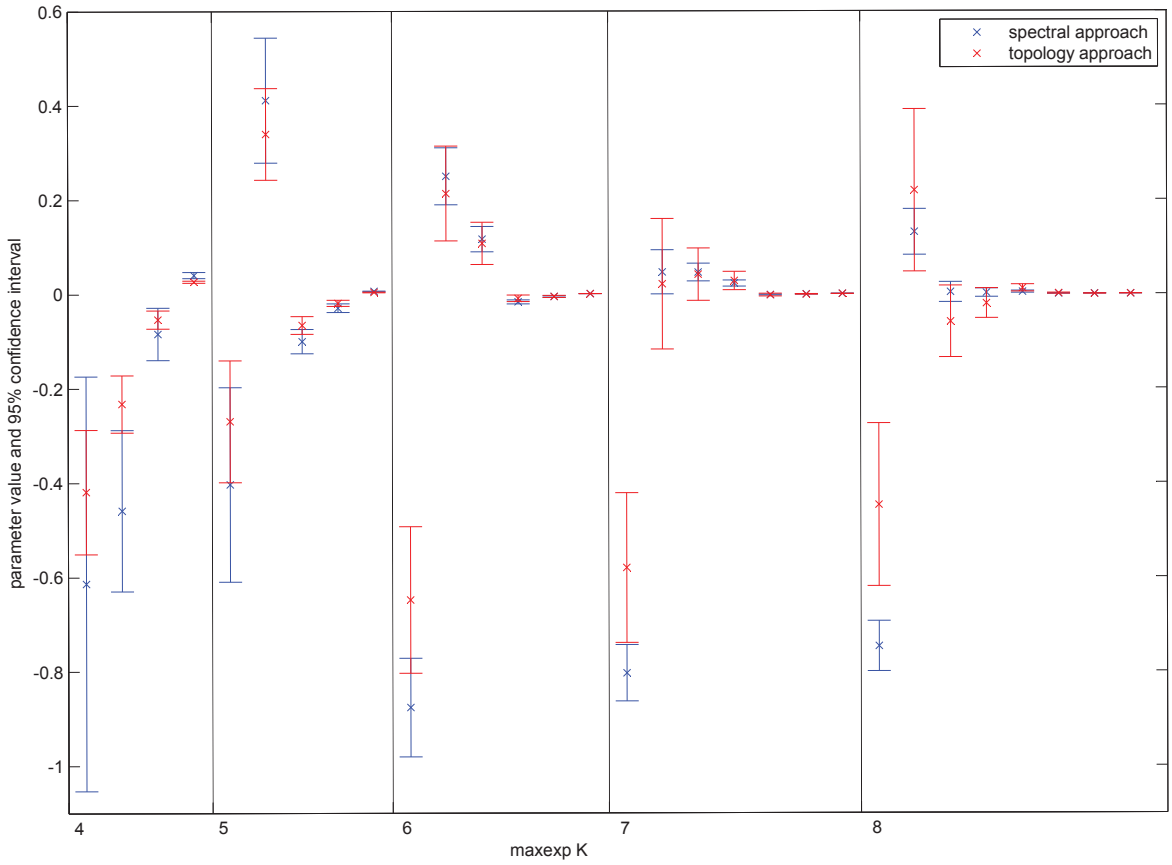


Figure 18: Plot of the estimated coefficient values for different maximal exponents K with their 95% confidence interval as an errorbar for the mapping $f : A \to W_{fMRI}$ for the group-averaged data set. The spectral and topology approach are marked in blue and red, respectively. Note that, in nearly all cases, the confidence intervals of the coefficient values overlap pointing towards similar estimated coefficients for the spectral and topology domain.

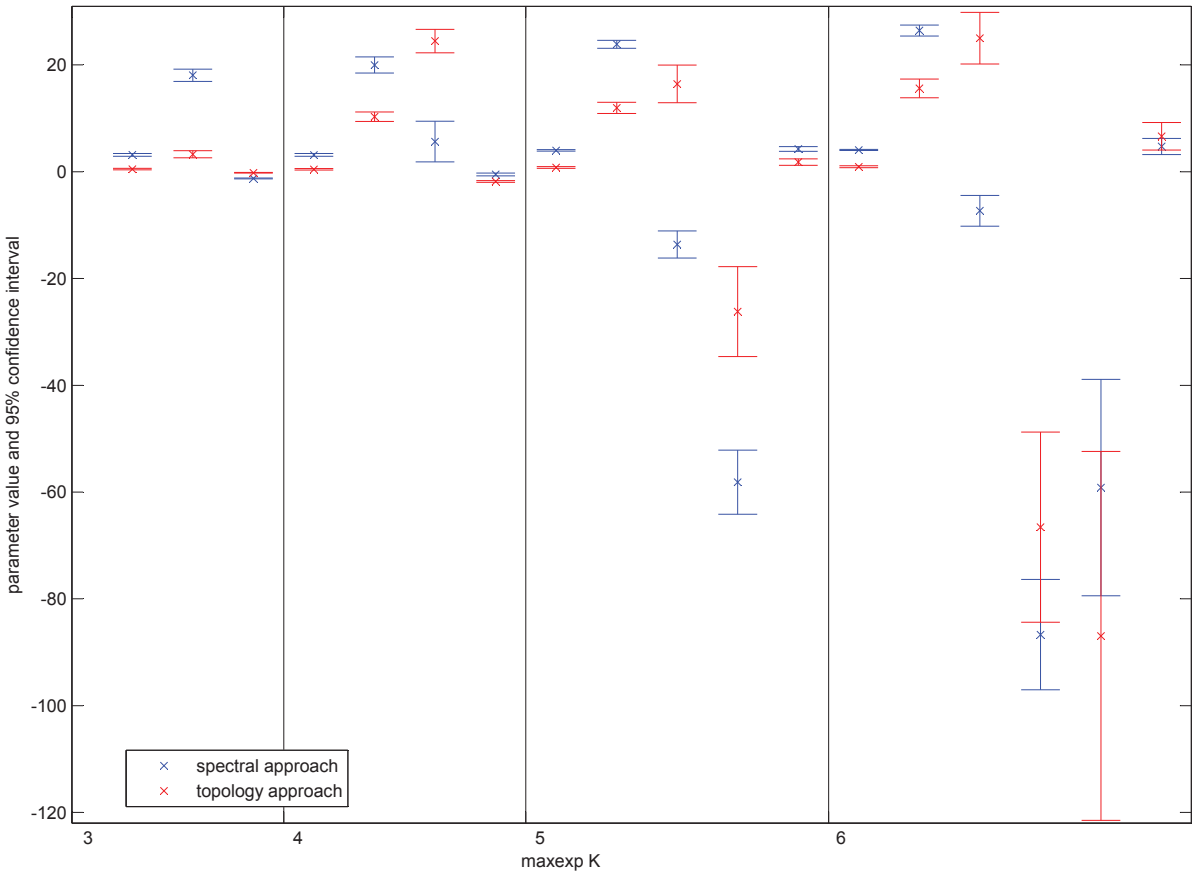


Figure 19: Plot of the estimated coefficient values for different maximal exponents K with their 95% confidence interval as an errorbar for the mapping $f^{-1} : W_{MEG} \to A$ for the group-averaged data set. The spectral and topology approach are marked in blue and red, respectively.

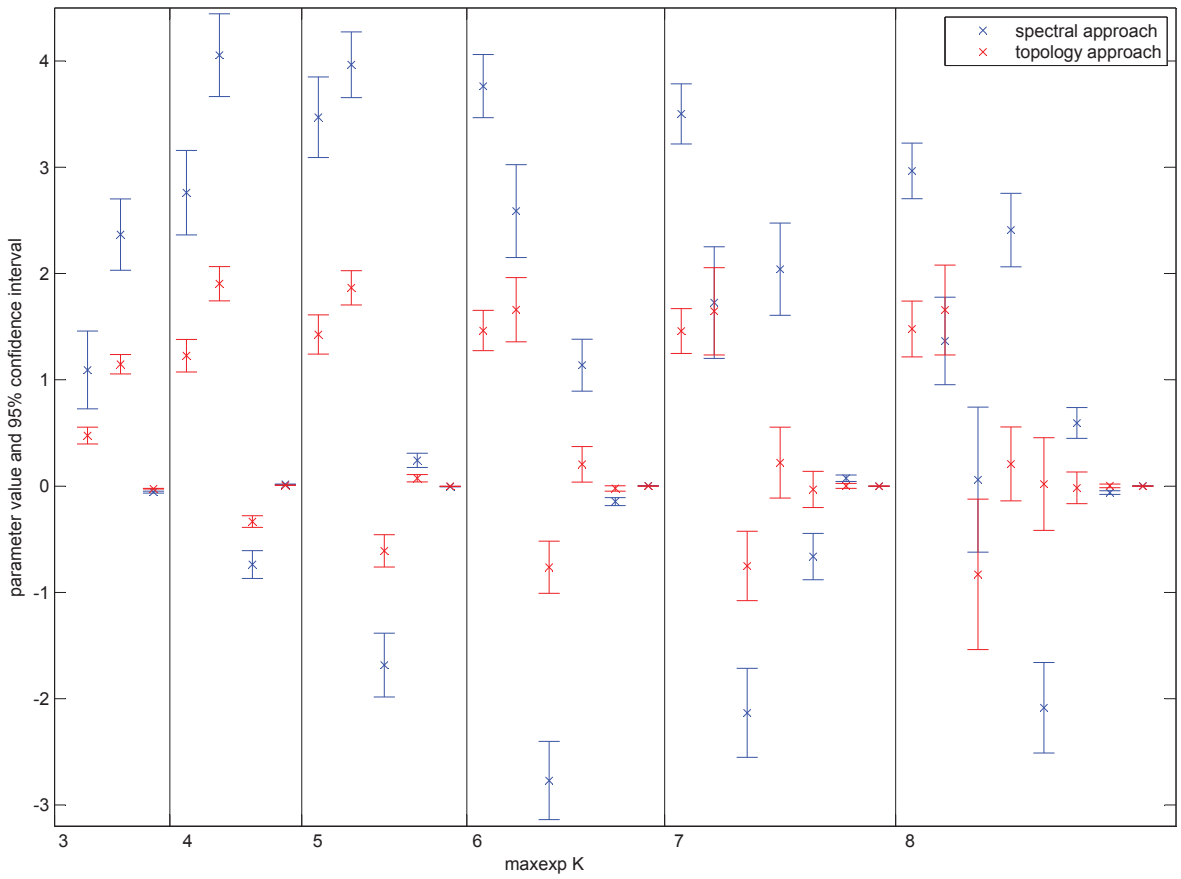


Figure 20: Plot of the estimated coefficient values for different maximal exponents K with their 95% confidence interval as an errorbar for the mapping $f^{-1} : W_{fMRI} \to A$ for the group-averaged data set. The spectral and topology approach are marked in blue and red, respectively.

mappings together with their 95% confidence intervals, from structure to function and from function to structure. In both displayed figures, the 95% confidence intervals for the estimated coefficients of the different modalities do not overlap indicating significantly different coefficients using the 5% significance level. This table and these confidence intervals show clearly the differences between the different analyzed mappings pointing towards a modality-dependent mapping.

Table 1: Estimated coefficient values for a maximal exponent of $K = 6$ and $z_0 = 0$ for the group-averaged data set.

parameter in front of	f: A -> W_{MEG}	f: A -> W_{fMRI}	parameter in front of	f^{-1}: W_{MEG} -> A	f^{-1}: W_{fMRI} -> A
I	-0.186244926	-0.622088682	I	0.931541504	1.472348531
E	0.170142739	0.417817587	E	-0.226229939	-0.098880043
A	-0.002689163	0.073378816	W	15.61376486	1.633796259
A^2	0.000397359	0.027965465	W^2	25.77141943	-0.760008801
A^3	0.00136165	0.010963972	W^3	-64.46444579	0.241736267
A^4	-7.34E-05	-0.000932557	W^4	-92.90312812	-0.03937664
A^5	-5.42E-05	-0.000401534	W^5	-6.670779246	0.002673357
A^6	5.47E-06	3.96E-05	W^6	0.993949647	-4.75E-05

D Interpretation with walks

In Figure 23, we visualized our mapping in terms of the number of walks of the structural brain network (without the error matrix E).

E Comparison with a previous study

In Figures 24 and 25, we compare our results with a previous study (Tewarie et al., 2014). Tewarie and coworkers (Tewarie et al., 2014) demonstrated that the degree product between nodes in the structural network together with the Euclidean distance has the ability to predict the functional connections between these nodes. If we merely focus on the degree product, we observe that our approach with the sum of structural matrices A^k (see Eq. (5)) is correlated with the degree product (SI Figure 24). The correlation between those two measures is indeed positive with a Spearman correlation of $R = 0.57$ (p-value < 0.001). There are two clouds in the scatterplot: the upper cloud corresponds to direct connections whereas the lower cloud corresponds to all possible indirect connections, consisting of all walks larger than one. If we investigate the relationship between the previous model (including degree product and Euclidean distance as predictors for functional connectivity) and the mapping approach from this paper, the Spearman correlation R becomes higher ($R = 0.64$, p-value < 0.001, see Figure 25). This result raises the question whether the Euclidean distance as a separate term in a model for explaining functional connections is required (Alexander-Bloch et al., 2013). In our approach, we only incorporated topological distance, which means the distance with respect to intermediary nodes and links in the structural network, and not Euclidean distance; however, these findings might suggest that topological distance and Euclidean distance between nodes are related.

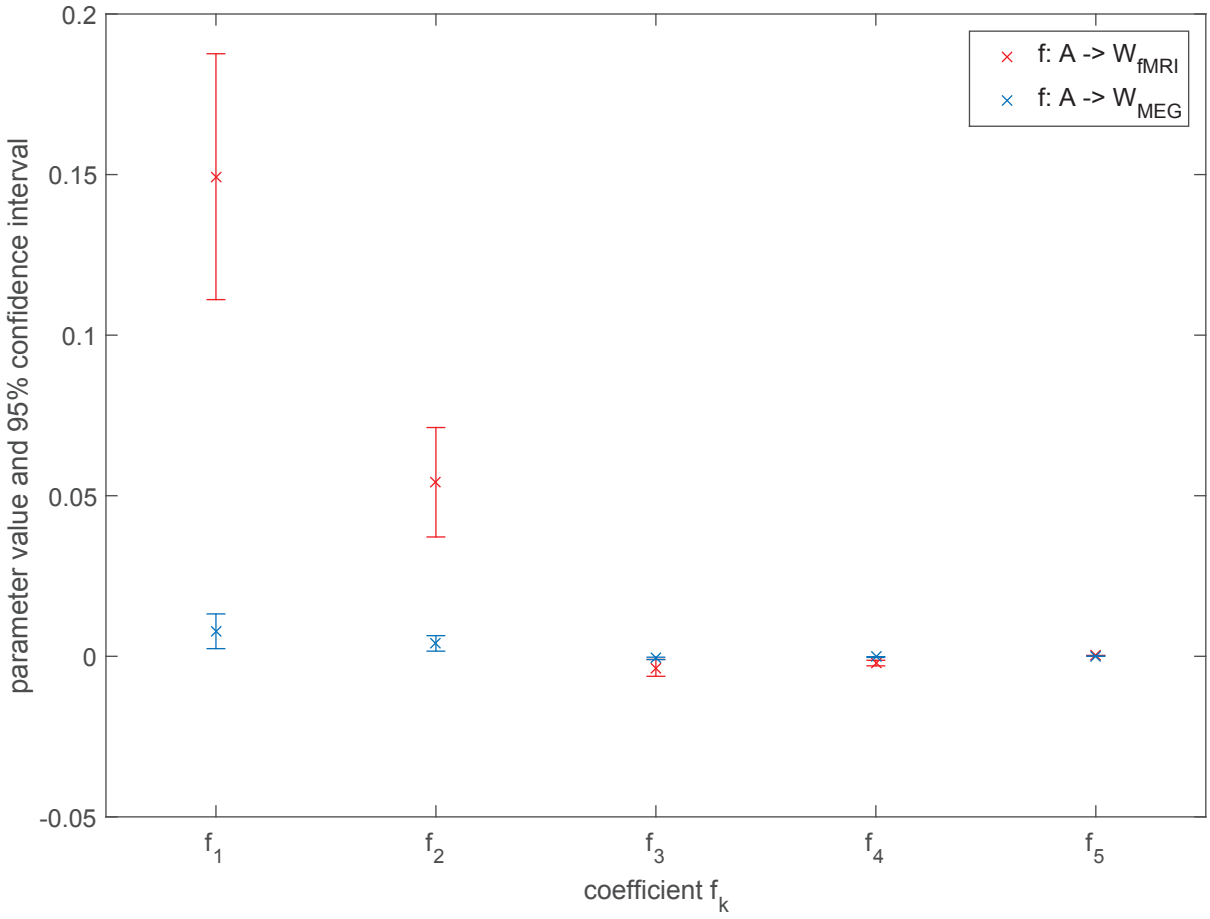


Figure 21: Plot of the estimated coefficient values for maximal fitted exponent $K = 5$ with their 95% confidence interval as an errorbar for the mapping from structural to functional networks for the group-averaged data set. Note that the displayed intervals do not overlap, thus we face here significantly different estimated values.

F Error analysis

The equation (1) assumes that W and A are known exactly. In reality, all types of error mask the true structure so that we actually measure

$$\widetilde{W} = W + \varepsilon_W R_W$$

where R_W is a realization of a random matrix with unit norm and ε_W is the maximum amplitude of the error. Similarly,

$$\widetilde{A} = A + \varepsilon_A R_A$$

and the assumption becomes

$$\widetilde{W} = f\left(\widetilde{A}\right)$$

or

$$W + \varepsilon_W R_W = f\left(A + \varepsilon_A R_A\right)$$

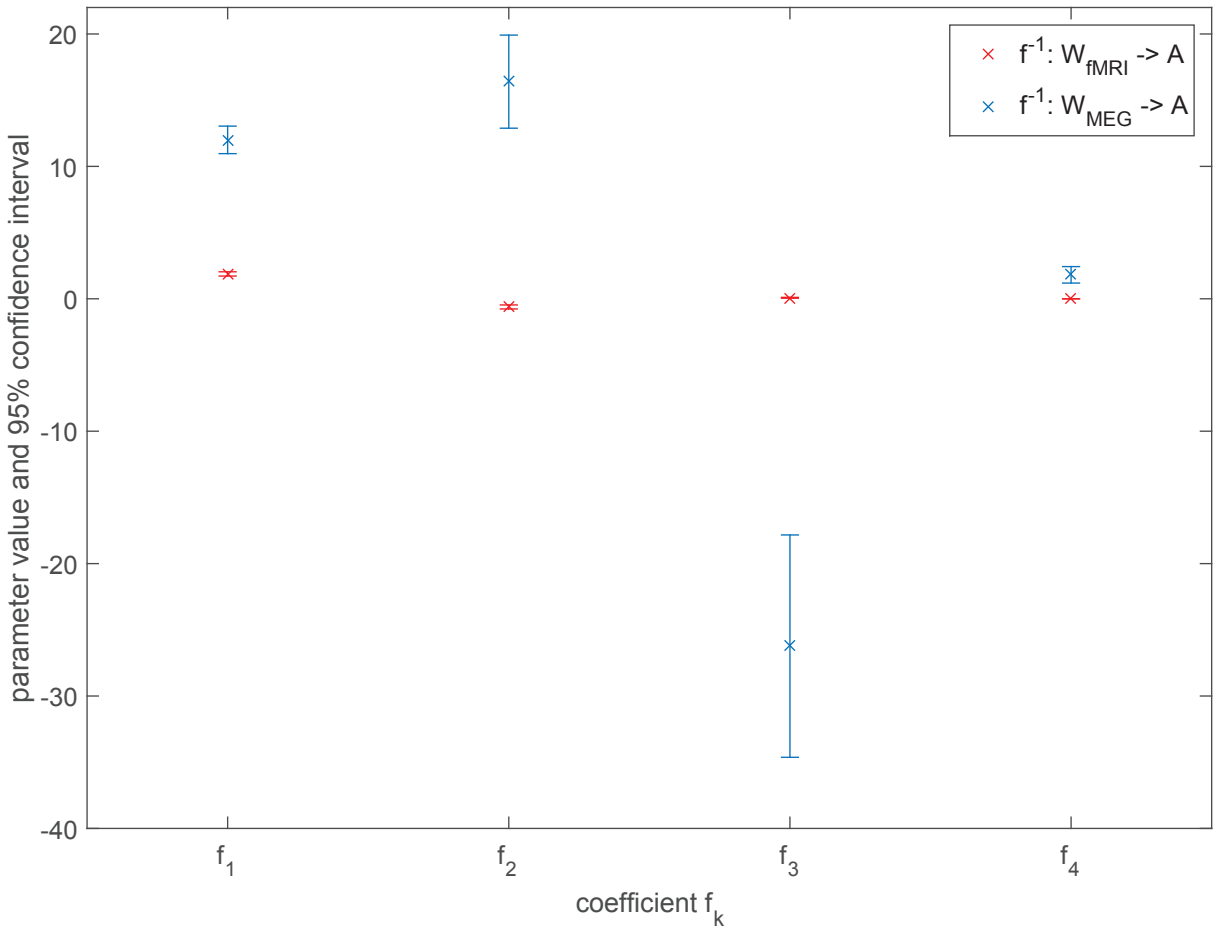


Figure 22: Plot of the estimated coefficient values for maximal exponents $K = 4$ with their 95% confidence interval as an errorbar for the mapping from functional to structural networks for the group-averaged data set. Note that the displayed intervals do not overlap, thus we face here significantly different estimated values.

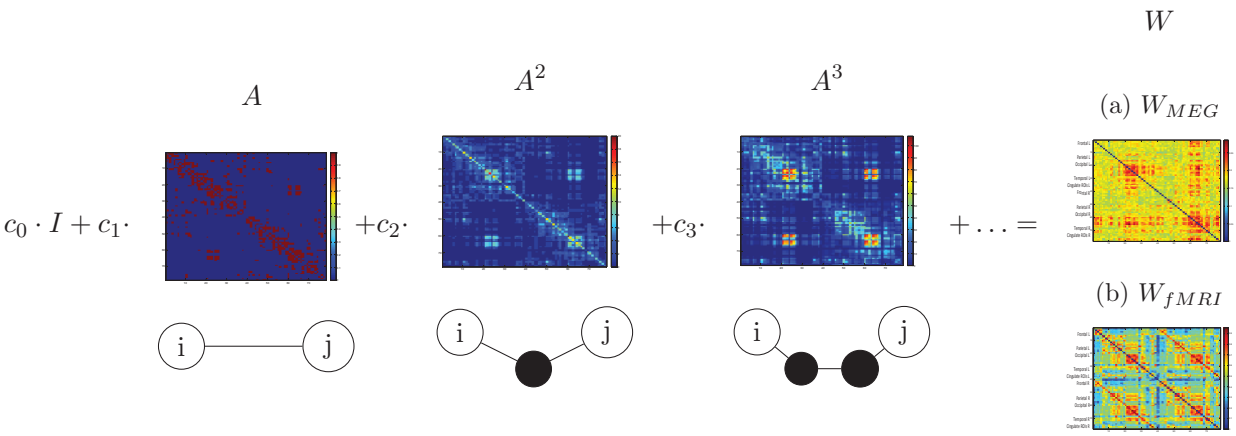


Figure 23: Visualization of a simplified version of our model (from (4) where $c_k = c_k[f]$): Walks of different length between node i and j, first the direct connection, then with one intermediate node, with two and so on, adding up to the functional connectivity matrix.

Using the Taylor expansion (3),

$$f\left(A+\varepsilon_A R_A\right)=\sum_{k=0}^{\infty} f_k\left(A\right)\varepsilon_A^k\left(R_A\right)^k$$

up to first order (assuming that ε_A is sufficiently small!), then

$$W+\varepsilon_W R_W = f\left(A\right)+f_1\left(A\right)\varepsilon_A R_A + O\left(\varepsilon_A^2\right)$$

Invoking the assumption (1) shows a relation between the different types of errors

$$\varepsilon_W R_W = f_1\left(A\right)\varepsilon_A R_A + O\left(\varepsilon_A^2\right)$$

Given that the assumption (1) is correct and that A is known exactly, we could derive a method to improve the measurements $\widetilde{W}$ based on Section SI-A.3, which suggests that all eigenvectors of W are fixed and known (i.e. X is the same as for A), so that $\widetilde{W}$ needs to be modified to incorporate this property. This analysis is a suggestion for future work.

To investigate the influence of false positives in the structural matrix, we randomly added connections (1 % new connections) in the structural matrix of the group-averaged data set and redid the analysis (power series in topology domain with 6 terms). The results can be found in the boxplots of the goodness of fit (SSE_{norm}) in SI Figure 26. If we define the change in SSE_{norm} due to noise as the noise influence NI whereas $NI(SSE_{norm}) := standard\ deviation(SSE_{norm})/mean(SSE_{norm})$, then we can calculated

$$\begin{aligned}
NI(SSE_{norm})(W_{fMRI} \to A) &= 0.0174\\
NI(SSE_{norm})(W_{MEG} \to A) &= 0.0169\\
NI(SSE_{norm})(A \to W_{fMRI}) &= 0.0124\\
NI(SSE_{norm})(A \to W_{MEG}) &= 0.0158.
\end{aligned}$$

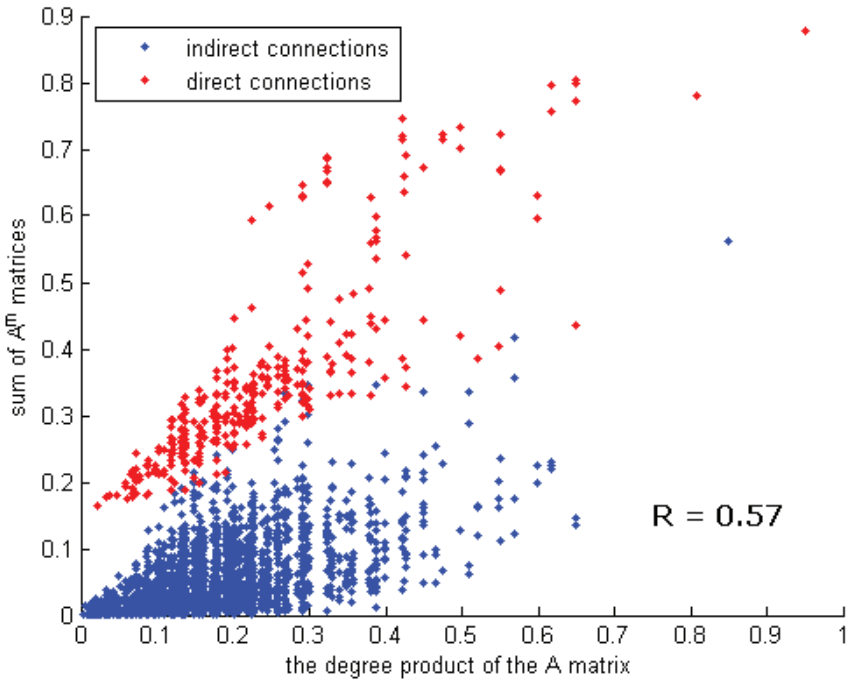


Figure 24: Scatter plot of the structural degree product against the sum of the powers of the structural matrix A (from power $k = 1$ up to $k = 6$).

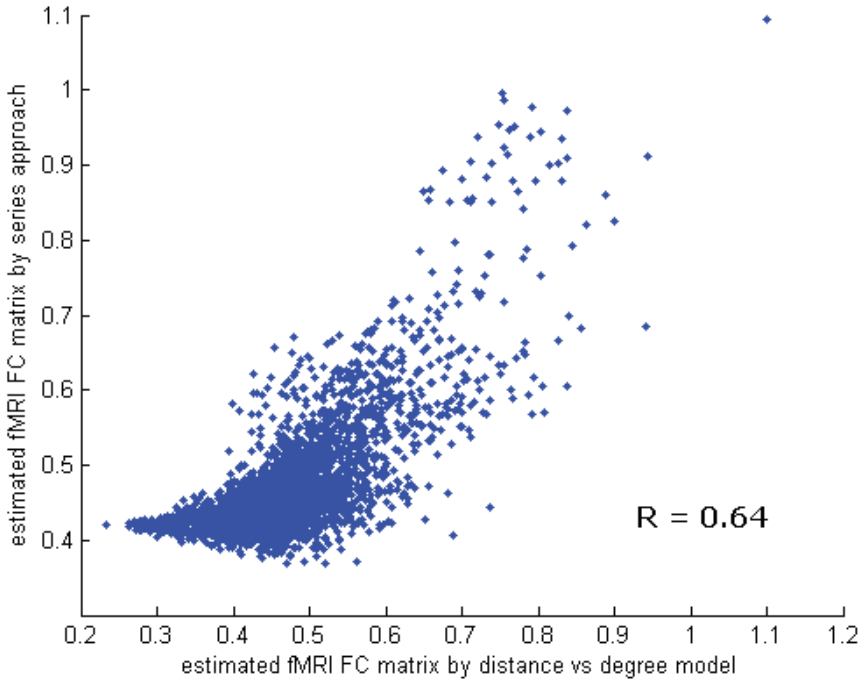
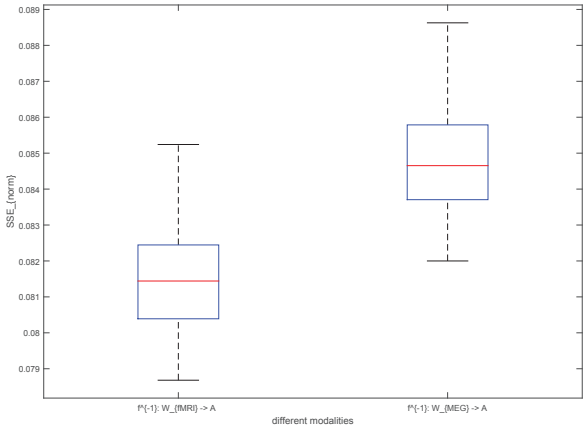
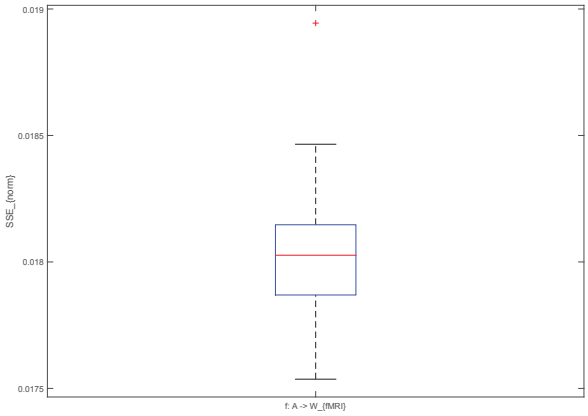


Figure 25: Scatter plot of the estimated fMRI correlation matrix from the distance and degree model described in a previous study (Tewarie et al., 2014) against the estimated values using the mapping approach on the structural matrix A.

(a) $f^{-1} : W_{fMRI} \to A$ and $f^{-1} : W_{MEG} \to A$



(b) $f : A \to W_{fMRI}$



(c) $f : A \to W_{MEG}$

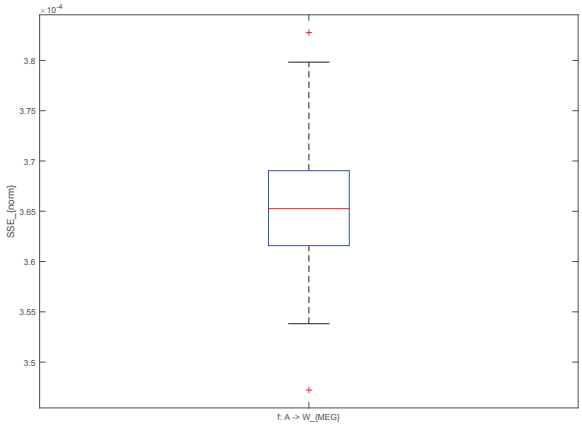


Figure 26: Boxplot of the variations in SSE_{norm} in the topology domain for $K = 6$ with group-averaged data for $z_0 = 0$. To investigate the influence of false positives in the structural connectivity matrix, we randomly added connections (1% new connections) in the structural matrix and redid the analysis.

It can be observed from the above that the change in goodness of fit is small in the presence of little noise (in the order of 1%), thus the mapping does not seem to be sensitive to small noise fluctuations.

If we analyze the region-to-region variability, we find that the inter-hemispheric connections were often quite different between the empirical matrix and its fitted version (see the secondary diagonal in Figures 27 - 30 where we displayed the absolute error). This result confirms our interpretation about the more homologous inter-hemispheric connections in the fitted networks than in the empirically observed networks. Furthermore, we were also interested in which regions benefitted more from an increasing number of coefficients (darker regions in SI Figures 31-34). Those regions that benefitted from an increasing number of fitted coefficients were possibly most influenced by longer walks in the underlying structural network. For the estimated structural and functional networks by our mapping we observe as a result quite a diverse homogenous spreading of the benefitting regions over the entire group of regions except for the diagonal and secondary diagonal. Thus, we can again conclude that the inter-hemispheric connections are probably benefitting most from our mapping approach.

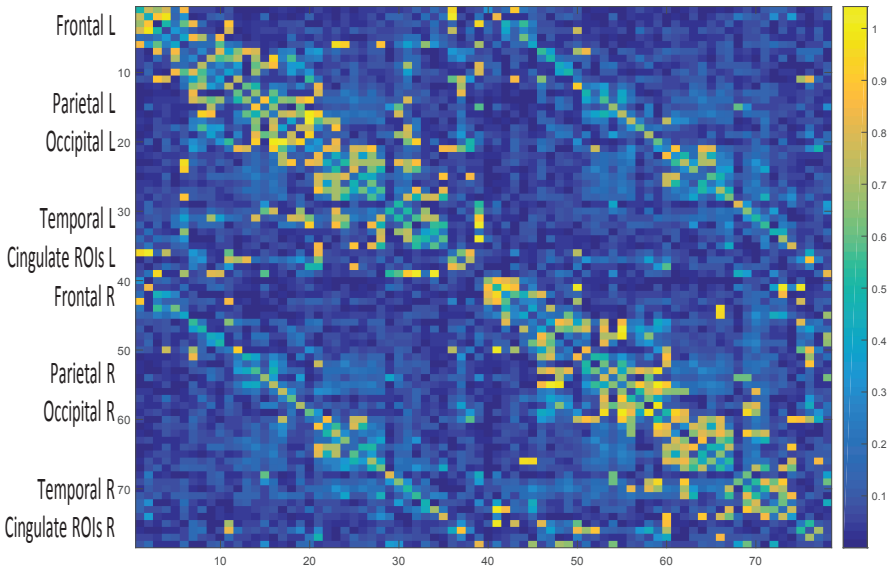


Figure 27: Absolute distance matrix between the best fit for $\widetilde{A} = f(W_{fMRI})$ ($K = 6$, $z_0 = 0$) and the empirical matrix A for the group-averaged data set.

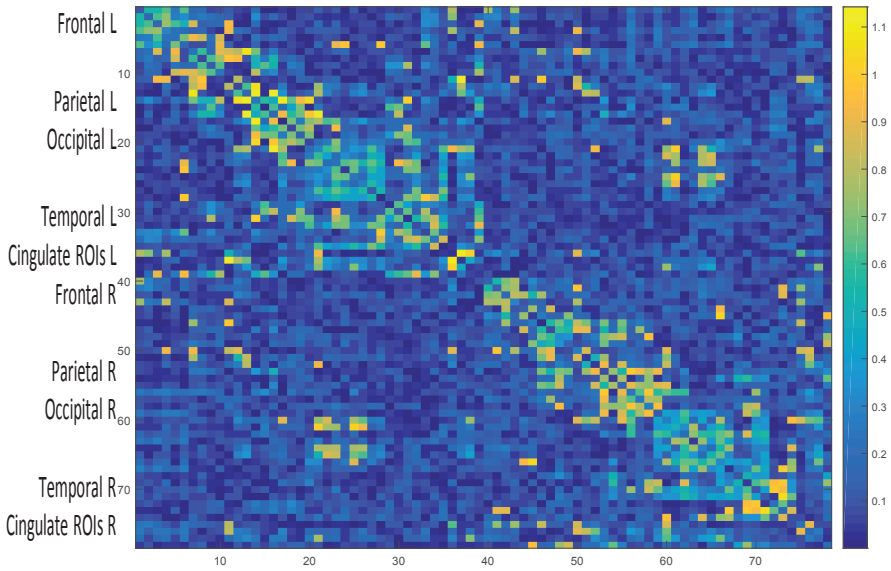


Figure 28: Absolute distance matrix between the best fit for $\widetilde{A} = f(W_{MEG})$ ($K = 6$, $z_0 = 0$) and the empirical matrix A for the group-averaged data set.

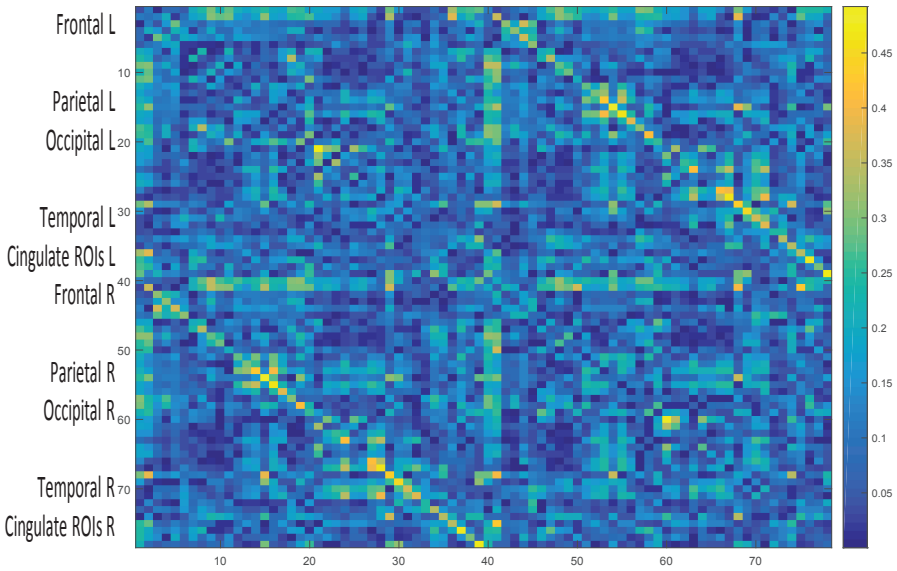


Figure 29: Absolute distance matrix between the best fit for $\widetilde{W}_{fMRI} = f(A)$ ($K = 6$, $z_0 = 0$) and the empirical matrix W_{fMRI} for the group-averaged data set.

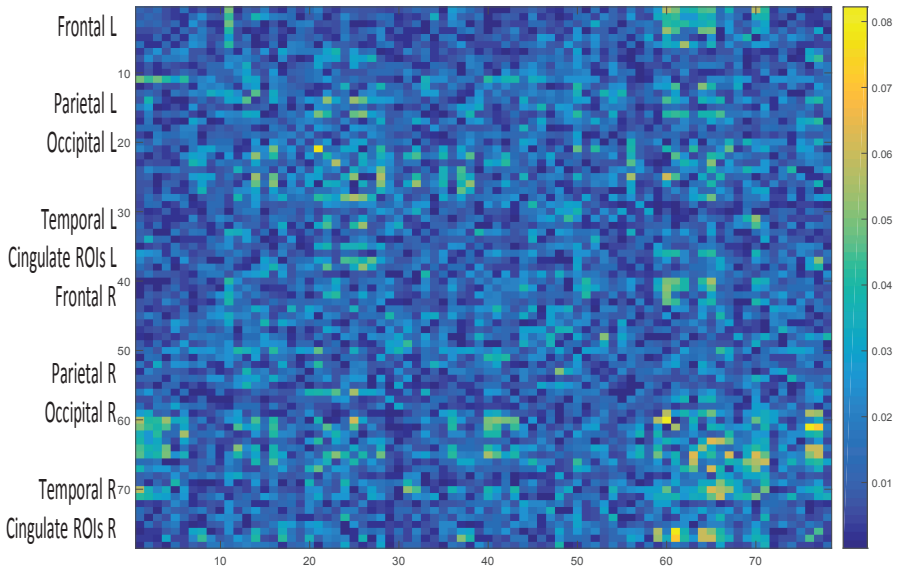


Figure 30: Absolute distance matrix between the best fit for $\widetilde{W}_{MEG} = f(A)$ ($K = 6$, $z_0 = 0$) and the empirical matrix W_{MEG} for the group-averaged data set.

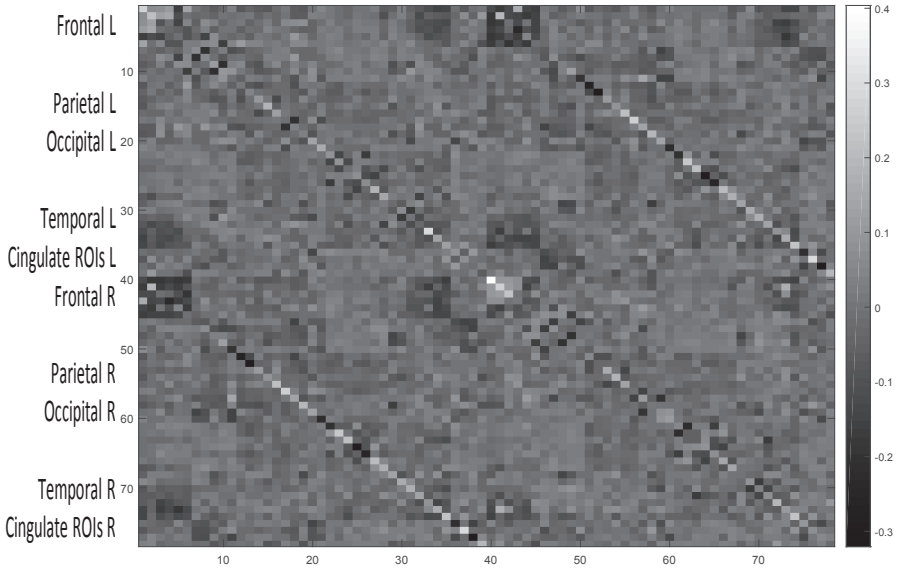


Figure 31: Absolute differences in the error for 2 in comparison with the error for 6 fitted coefficients ($K = 6$, $z_0 = 0$) for the mapping $f^{-1} : W_{fMRI} \to A$ for the group-averaged data set. Note that darker areas correspond here to regions that benefitted more from a higher number of coefficients.

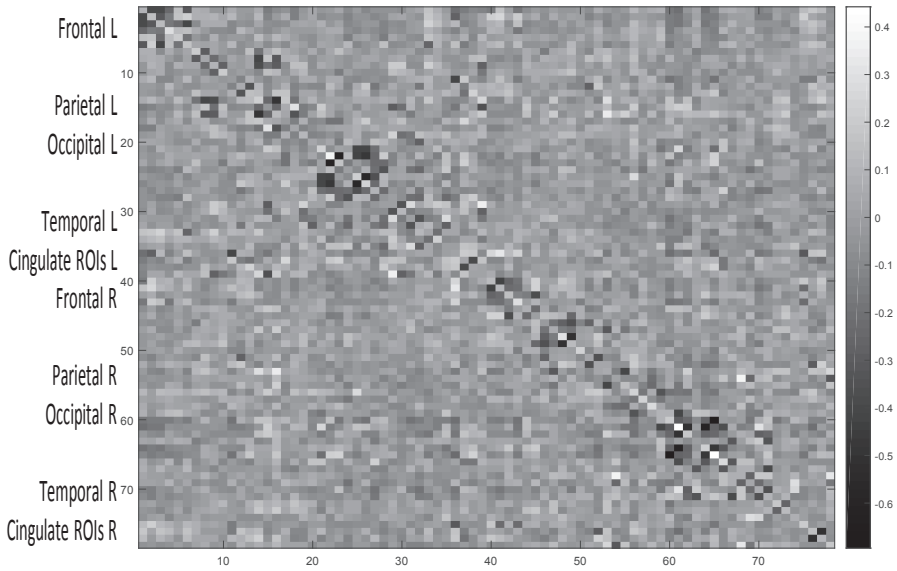


Figure 32: Absolute differences in the error for 2 in comparison with the error for 6 fitted coefficients ($K = 6$, $z_0 = 0$) for the mapping $f^{-1} : W_{MEG} \to A$ for the group-averaged data set. Note that darker areas correspond here to regions that benefitted more from a higher number of coefficients.

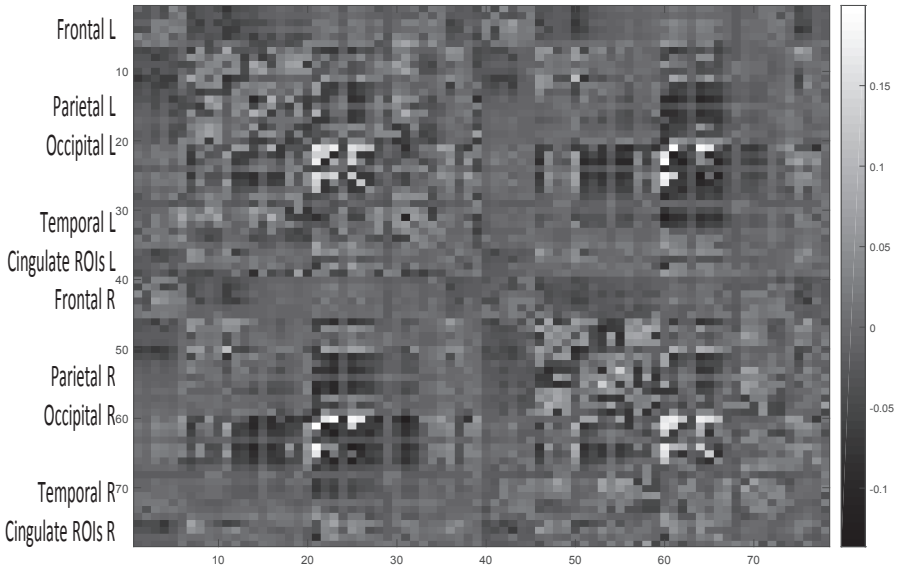


Figure 33: Absolute differences in the error for 2 in comparison with the error for 6 fitted coefficients ($K = 6$, $z_0 = 0$) for the mapping $f : A \to W_{fMRI}$ for the group-averaged data set. Note that darker areas correspond here to regions that benefitted more from a higher number of coefficients.

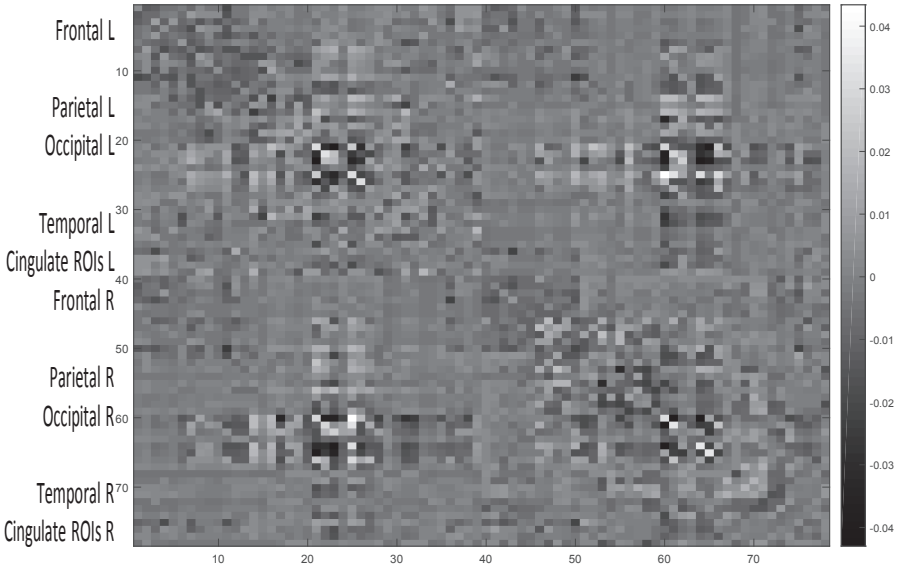


Figure 34: Absolute differences in the error for 2 in comparison with the error for 6 fitted coefficients ($K = 6$, $z_0 = 0$) for the mapping $f : A \to W_{MEG}$ for the group-averaged data set. Note that darker areas correspond here to regions that benefitted more from a higher number of coefficients.

G Reshuffled matrices

We used one reshuffling technique on all matrices: we selected two matrix entries at random and then exchanged their entries in the matrix (their link weights in the network). We repeated this step 5000 times to obtain a reshuffled version of every matrix, which is again symmetric. Thereby we preserved the distribution of the weights. After generating 100 reshuffled matrices, we compared the goodness of fit distribution of the mappings using reshuffled matrices to our empirical results (see SI Figures 35 and 36). More precisely, for each mapping from any matrix M to N, we first replaced the underlying matrix M with a reshuffled version of this adjacency matrix and ran the same mapping analysis on it. Then we replaced the image matrix N with a randomly reshuffled version of itself (keeping the underlying matrix M as the original empirical matrix) and ran the fitting algorithm on that combination (thus always one empirical matrix with one reshuffled version of the other matrix together).

For the group-averaged data set, the fit errors of 100 reshuffled matrices were larger compared to those obtained using the experimental data for all analyzed mappings (see SI Figures 35 and 36). Therefore, we can conclude that the mapping from structure to function (and vice versa) fails when reshuffled connections are used. Therefore, the empirical matrices seem to possess a special structure making the relationship between the structural and functional brain network closer than when reshuffled version of those matrices are used.

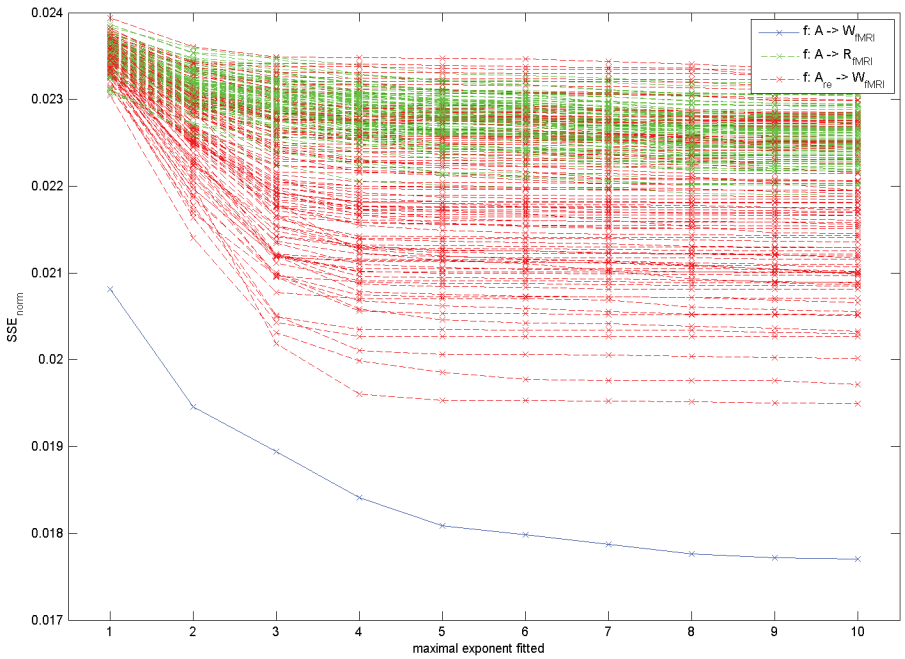
For the data set of individual healthy controls, we also used 100 reshuffled matrices of their structural and functional networks, respectively, and displayed the percentage of those matrices that achieved better goodness of fit than our empirical data (see SI Figures 37 and 38).

When using 5 coefficients or more, we obtained good results for the mapping $f : V \to W_{fMRI}$ for nearly all subjects. Only subject number 5 seems to be an exception which could be due to some measurement errors or specific individual attributes of that subject. We also identified two individuals as outliers whose goodness of fit level did not outperform the random reshuffles for the mapping $f^{-1} : W_{fMRI} \to V$. The function f^{-1} starting from MEG networks obtained good results for $K > 2$. The only mapping that was not able to outperform the reshuffled matrices as a benchmark for most of the subjects was $f : V \to W_{MEG}$ in the individual healthy control data. To sum up, for the individual structure-function relationships, in most of the cases the mapping performs worse when we apply it to random reshuffles indicating a high goodness of fit level for the mapping between the original empirical matrices.

H Details of the fitting procedure

In order to use the non-linear regression algorithm in MATLAB (using the routine *nlinfit.m* version *R2015a*) to determine the coefficients in (5) by iterative least squares estimation, we need to adapt our data first. Because all involved matrices are symmetric, we only need to fit the lower triangular matrices and the diagonal to get our fitting results. Thus, we first write all matrices in a vectorized form only containing their lower triangular and diagonal entries. For any matrix M of dimension $N \times N$ this vector will be denoted by $ltd(M)$. To be able to use the standard equation $Y = X \cdot \beta$ (with X design matrix, β parameter vector and Y image matrix) for a linear model, we need to define the

(a) $f : A \to W_{fMRI}$



(b) $f : A \to W_{MEG}$

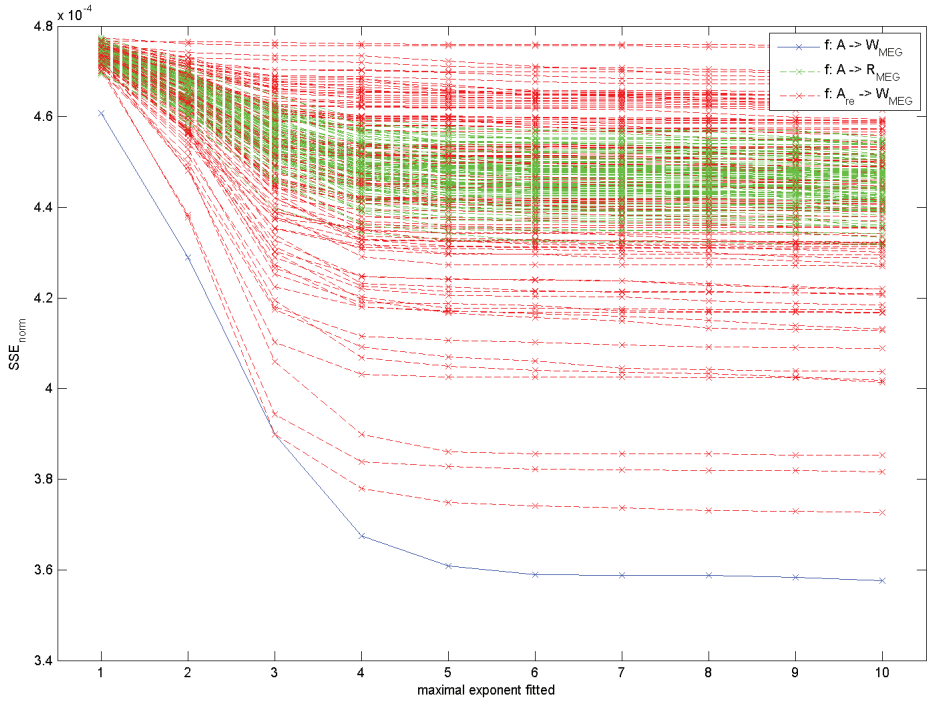
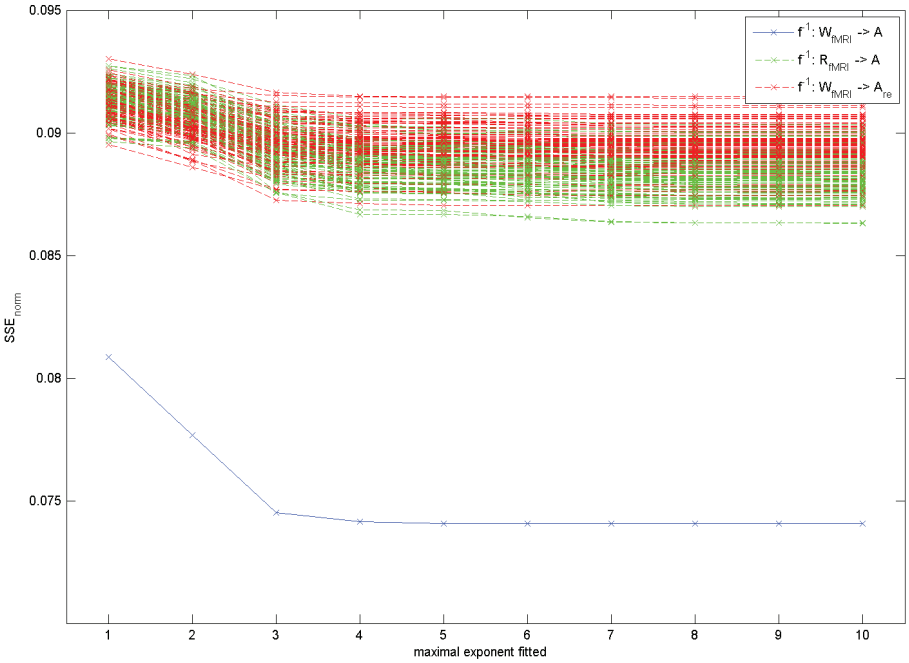


Figure 35: Plot of the normalized sum of squared errors (SSE_{norm}) of the function f in the topology domain for different maximal fitted exponents K and always in combination with a randomly reshuffled matrix R (R_{MEG}, R_{fMRI} and A_{re} denoting the reshuffled versions of W_{MEG}, W_{fMRI} and A, respectively) averaged over a range of z_0 values from -3 till 3 (always including an error matrix E) for the group-averaged data set.

(a) $f^{-1} : W_{fMRI} \rightarrow A$



(b) $f^{-1} : W_{MEG} \rightarrow A$

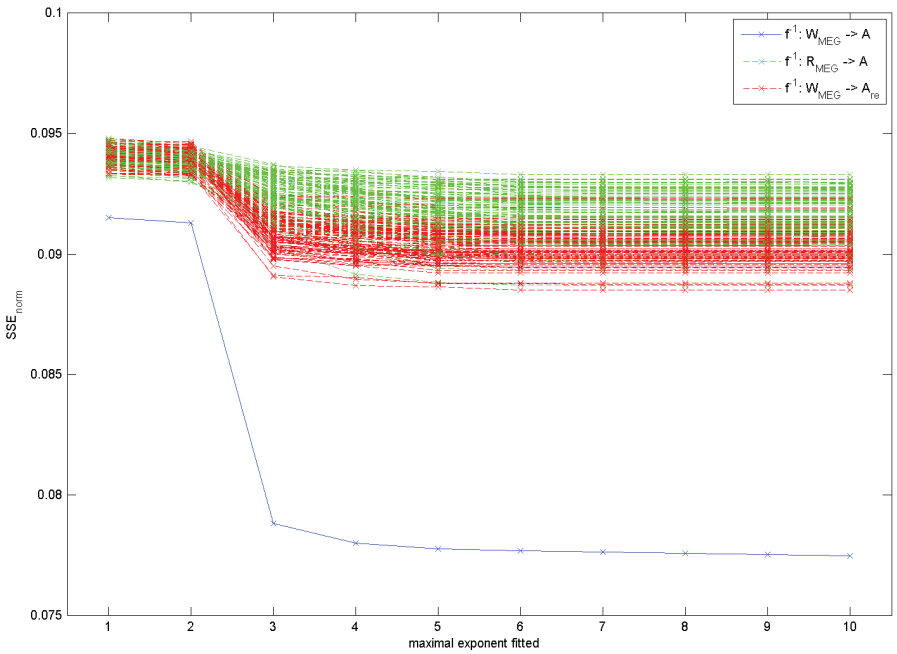
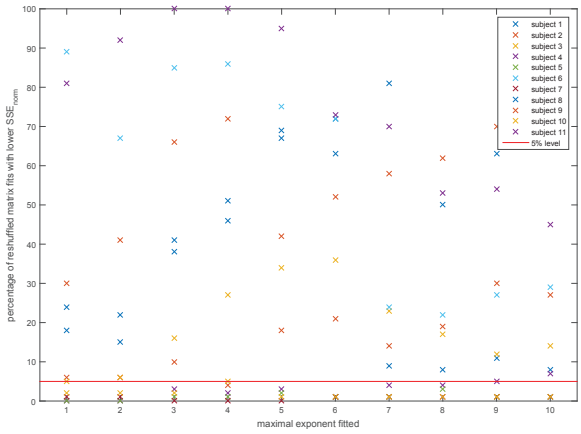
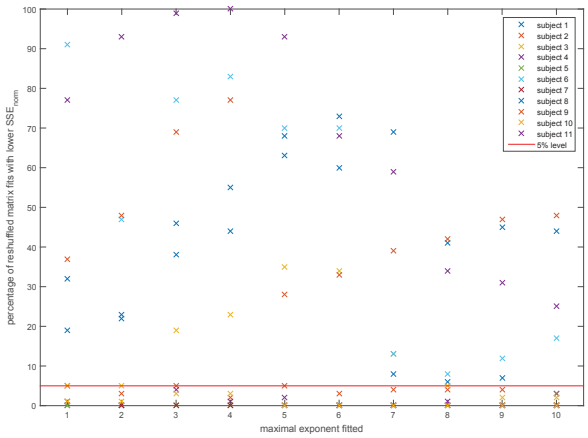


Figure 36: Plot of the normalized sum of squared errors (SSE_{norm}) of the function f^{-1} in the topology domain for different maximal fitted exponents K and always in combination with a randomly reshuffled matrix R (R_{MEG}, R_{fMRI} and A_{re} denoting the reshuffled versions of W_{MEG}, W_{fMRI} and A, respectively) averaged over a range of z_0 values from -3 till 3 (always including an error matrix E) for the group-averaged data set.

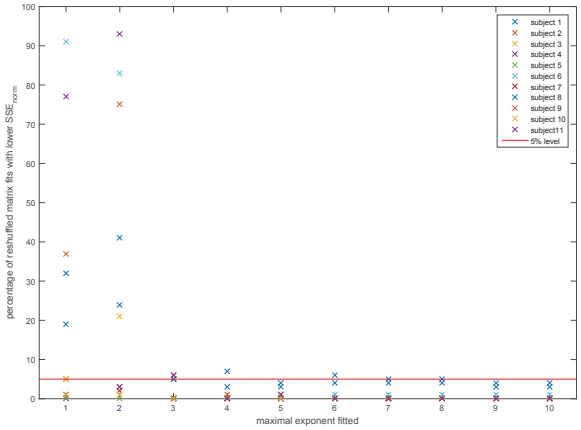
(a) $f : V_{re} \to W_{MEG}$



(b) $f : V \to R_{MEG}$



(c) $f^{-1} : R_{MEG} \to V$



(d) $f^{-1} : W_{MEG} \to V_{re}$

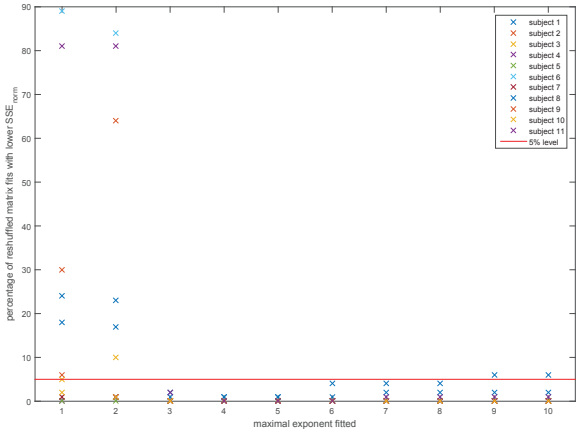
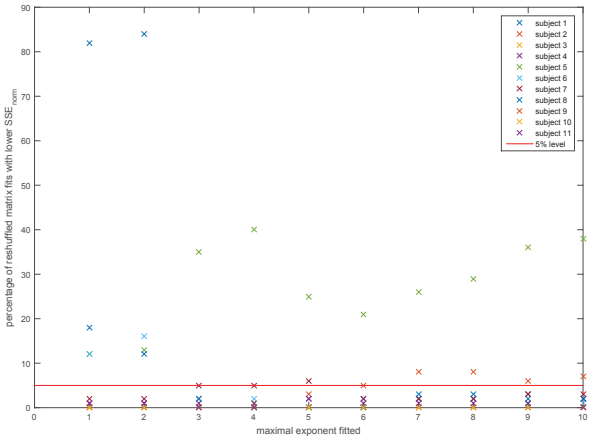
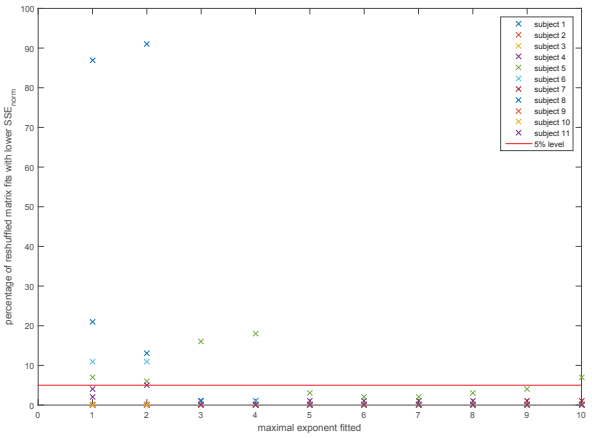


Figure 37: Plot of the percentages of reshuffled matrices that resulted in a lower normalized sum of squared errors (SSE_{norm}) in the topology domain for different maximal fitted exponents K with individual healthy controls results for all mapping including MEG (V denoting the weighted structural matrix and V_{re} denoting its randomly reshuffled version) for $z_0 = 0$ (with an error matrix E).

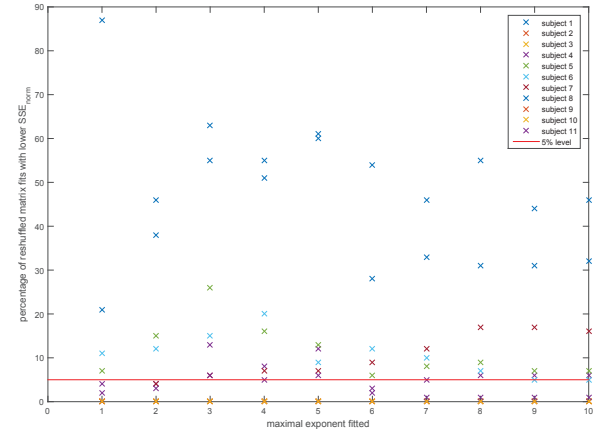
(a) $f : V_{re} \to W_{fMRI}$



(b) $f : V \to R_{fMRI}$



(c) $f^{-1} : R_{fMRI} \to V$



(d) $f^{-1} : W_{fMRI} \to V_{re}$

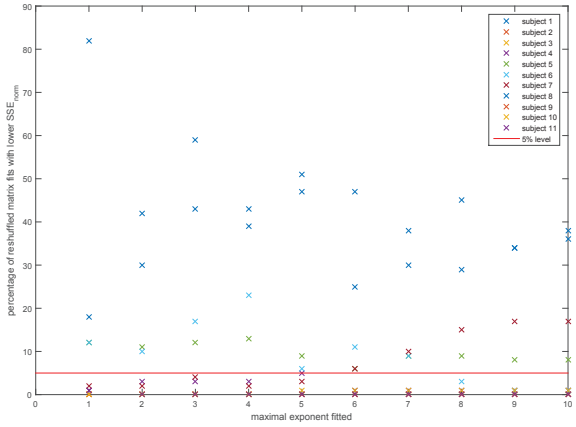


Figure 38: Plot of the percentages of reshuffled matrices that resulted in a lower normalized sum of squared errors (SSE_{norm}) in the topology domain for different maximal fitted exponents K with individual healthy controls results for all mappings including fMRI (V denoting the weighted structural matrix and V_{re} denoting its randomly reshuffled version) for $z_0 = 0$ (with an error matrix E).

variables for our case. In the case of $f(A) = W$, the response Y is just the image matrix W written as a vector containing the lower triangular and diagonal entries, $Y = ltd(W)$. The design matrix X is in the case of K as the maximal exponent

$$X = \Big(ltd(J) \quad ltd(I) \quad ltd(A) \quad ltd(A^2) \quad ... \quad ltd(A^K) \Big)$$

Therefore, the parameter vector β has the length $(K+2)$, where the first entry will be the coefficient c in front of the all-one matrix J and the second one is the coefficient for the identity matrix I and the others are in front of the matrix powers of A. Because the matrix powers of A are exploding in magnitude quickly, we normalize all the matrices beforehand dividing every entry by the absolute maximum entry of each matrix, which has the consequence that all matrices now have values between 0 and 1. Then, the *nlinfit.m* algorithm can be applied to our data using the underlying function *myfun.m* displayed here

```
function F = myfun(beta,xdata)

        F = xdata*beta;
end
```

where **xdata** refers to our design matrix X and **beta** is the parameter vector β. Because our model resembles a GLM, we could also use the pseudo-inverse of our design matrix **xdata** (*pinv(**xdata**)*) and multiply it with the vectorized matrix Y in order to obtain the same estimated coefficients. In order to obtain the coefficient values for the original powers of the A matrix, we have to denormalize the estimated values by dividing the estimated coefficients each by the absolute normalization value from before.

I Dimension differences

In practice, the $m \times m$ measured matrix W^* may be of a different dimension than the $N \times N$ matrix W. If $m \geq N$, then we can transform the measured matrix W^* to W as follows. Since W^* is symmetric, the spectral decomposition is

$$W^* = Y^* \Upsilon^* Y^{*T}$$

where the diagonal matrix $\Upsilon^* = \mathrm{diag}\big(\mu_1^*, \ldots, \mu_N^*, \mu_{N+1}^*, \ldots, \mu_m^*\big)$ with the real eigenvalues ordered as $|\mu_1^*| \geq |\mu_2^*| \geq \ldots \geq |\mu_m^*|$. The ordering here is different than the usual ordering in Section A.3, because eigenvalues of W may be negative (in principle; although those of a correlation matrix are non-negative). Next, we let $\mu_k = \mu_k^*$ for $1 \leq k \leq N$ and $\mu_k = 0$ for $k > N$ and

$$Y^* = \begin{bmatrix} (Y_{11})_{N \times N} & (Y_{12})_{N \times (m-N)} \\ (Y_{21})_{(m-N) \times N} & (Y_{22})_{(m-N) \times (m-N)} \end{bmatrix}$$

so that

$$\widetilde{W^*} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} \Upsilon & O \\ O & O \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}^T = \begin{bmatrix} Y_{11}\Upsilon Y_{11}^T & Y_{11}\Upsilon Y_{21}^T \\ Y_{21}\Upsilon Y_{11}^T & Y_{21}\Upsilon Y_{21}^T \end{bmatrix}$$

from which we choose $W = Y_{11}\Upsilon Y_{11}^T$. This method is well-known in the theory of singular value decompositions (see e.g. (Golub and Loan, 1996)) and provides the best $N \times N$ (in the mean-square sense) approximation of an $m \times m$ matrix.

References

Alexander-Bloch, A. F., Vértes, P. E., Stidd, R., Lalonde, F., Clasen, L., Rapoport, J., Giedd, J., Bullmore, E. T., and Gogtay, N. (2013). The anatomical distance of functional connections predicts brain network topology in health and schizophrenia. *Cerebral cortex*, 23(1):127–138.

Golub, G. H. and Loan, C. F. V. (1996). *Matrix Computations*. The John Hopkins University Press, Baltimore, third edition.

Tewarie, P., Hillebrand, A., van Dellen, E., Schoonheim, M., Barkhof, F., Polman, C., Beaulieu, C., Gong, G., van Dijk, B., and Stam, C. (2014). Structural degree predicts functional network connectivity: A multimodal resting-state fMRI and MEG study. *NeuroImage*, 97:296–307.

Van Mieghem, P. (2011). *Graph Spectra for Complex Networks*. Cambridge University Press.