

## **A Riemannian approach to predicting brain function from the structural connectome**

Oualid Benkarim<sup>1\*</sup>, Casey Paquola<sup>1,2</sup>, Bo-yong Park<sup>1,3,4</sup>, Jessica Royer<sup>1</sup>, Raúl Rodríguez-Cruces<sup>1</sup>, Reinder Vos de Wael<sup>1</sup>, Bratislav Misic<sup>1</sup>, Gemma Piella<sup>5</sup>, Boris C. Bernhardt<sup>1\*</sup>

<sup>1</sup>*McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada;* <sup>2</sup>*Institute of Neuroscience and Medicine (INM-1), Forschungszentrum Jülich, Jülich, Germany;* <sup>3</sup>*Department of Data Science, Inha University, South Korea;* <sup>4</sup>*Center for Neuroscience Imaging Research, Institute for Basic Science, South Korea;* <sup>5</sup>*BCN MedTech, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain*

\* Correspondence to [oualid.benkarim@mcgill.ca](mailto:oualid.benkarim@mcgill.ca) or [boris.bernhardt@mcgill.ca](mailto:boris.bernhardt@mcgill.ca)

## ABSTRACT

Ongoing brain function is largely determined by the underlying wiring of the brain, but the specific rules governing this relationship remain unknown. Emerging literature has suggested that functional interactions between brain regions emerge from the structural connections through mono- as well as polysynaptic mechanisms. Here, we propose a novel approach based on diffusion maps and Riemannian optimization to emulate this dynamic mechanism in the form of random walks on the structural connectome and predict functional interactions as a weighted combination of these random walks. Our proposed approach was evaluated in two different cohorts of healthy adults (Human Connectome Project, HCP; Microstructure-Informed Connectomics, MICs). Our approach outperformed existing approaches and showed that performance plateaus approximately around the third random walk. At macroscale, we found that the largest number of walks was required in nodes of the default mode and frontoparietal networks, underscoring an increasing relevance of polysynaptic communication mechanisms in transmodal cortical networks compared to primary and unimodal systems.

**Keywords:** functional connectivity, structural connectome, diffusion maps, manifold optimization

## INTRODUCTION

Neuroscience has increasingly embraced a paradigm shift, away from focusing on single regions towards conceptualizations that emphasize the analysis of the brain as a complex, interconnected network (Avena-Koenigsberger *et al.*, 2018; Bassett *et al.*, 2017; Bullmore & Sporns, 2009; Fornito *et al.*, 2013; Sporns *et al.*, 2005). With progress in multimodal imaging and modelling, in particular advances in diffusion MRI acquisition and tractographic reconstructions, it is now possible to visualize the structural connectome (SC), as a representation of structural wiring with increasing biological validity (Hagmann *et al.*, 2007). Mapping the SC has critical appeal to network neuroscience, as it is generally considered to provide the physical wiring diagram of the brain that shapes and constrains ongoing brain function and dynamics, which in turn are thought to underlie the emergence of cognitive functions and behavior (Bressler & Menon, 2010).

Although a fundamental goal of systems neuroscience is to identify how structure gives rise to ongoing brain function, establishing a direct link between the SC and the functional connectome (FC) as a representation of ongoing signal interactions remains an open problem. While neural signals can be immediately transmitted between anatomically connected locations in the SC, the principles governing the flow of information between different unconnected regions of the brain remain elusive. In both humans and non-human primates, the strength of a structural connection linking two brain regions has shown to be a relatively robust predictor of the strength of their functional interaction (Honey *et al.*, 2009; Shen *et al.*, 2012; Skudlarski *et al.*, 2008). Nonetheless, the predictive accuracy of the SC is far from perfect, especially when aiming to explain the functional connectivity between anatomically unconnected regions, which may be mediated by polysynaptic communication paths (Damoiseaux & Greicius, 2009; Goñi *et al.*, 2014; Honey *et al.*, 2009). Moreover, the strength of the structure-function relationship has generally been recognized to vary across the brain, with a tight coupling of structural and functional connectivity profiles in unimodal sensory regions that is gradually relaxed in higher-order transmodal association areas, notably regions of the default mode and frontoparietal networks (Baum *et al.*, 2020; Vázquez-Rodríguez *et al.*, 2019).

Several approaches have been proposed to explain the mapping between structural and functional networks, including statistical associative techniques (Mišić *et al.*, 2016), biophysical models (Breakspear, 2017; Deco *et al.*, 2013; Honey *et al.*, 2009; Robinson, 2012; Wang *et al.*, 2019), structural connectome harmonics (Abdelnour *et al.*, 2014, 2018; Becker *et al.*, 2018; Rosenthal *et al.*, 2018), network communication models (Avena-Koenigsberger *et al.*, 2018; Bazinet *et al.*, 2021; Goñi *et al.*, 2014; Mišić *et al.*, 2015), and deep learning methods (Rosenthal *et al.*, 2018; Sarwar *et al.*, 2021). Among these approaches, those based on the eigenvectors on the SC and network communication have attracted mounting interest recently (Abdelnour *et al.*, 2018; Atasoy *et al.*, 2016; Becker *et al.*, 2018; Gabay *et al.*, 2018; Surampudi *et al.*, 2018; Tewarie *et al.*, 2020; Wang *et al.*, 2017). These approaches incorporate polysynaptic communication mechanisms through more than one structural connection to account for the flow of information between not only directly connected regions but also intermediary pathways (Atasoy *et al.*, 2016; Seguin *et al.*, 2020; Suárez *et al.*, 2020). By working on the structural embeddings, network communication can be modelled in a straightforward manner based on random walks on the SC (*i.e.*, signal diffusion through the entire SC) or as combinations of the structural eigenvectors (Tewarie *et al.*, 2020). As such, these techniques allow for the modelling of both mono- as well as polysynaptic communication mechanisms to incorporate increasingly high-order structural interactions, which may

ultimately reconstruct a dense FC from a relatively sparse SC representation (Honey et al., 2009; Suárez et al., 2020).

Here, we propose a novel approach to predict FC from SC and better understand structure-function coupling in the human brain. The proposed approach is formulated as a kernel fusion method, where each kernel can be depicted as a putative intermediate state while the brain is propagating information through the static white matter fibers. These multi-scale diffusion kernels are implemented as random walks on the structural eigenspace, by leveraging a diffusion maps framework to identify low dimensional components describing variance in SC (Coifman et al., 2005; Coifman & Lafon, 2006). Since the functional diffusion coordinates we attempt to synthesize may have a different orientation than their analogous structural coordinates, we need a transformation to align their low-dimensional representations in manifold space. To do so, we formulate our task as a Riemannian optimization problem over the product manifold of rotations (Absil et al., 2009; Hu et al., 2020), where each rotation is used to identify the optimal paths of a specific length, and subsequently build the corresponding intermediate diffusion kernels. The proposed approach is illustrated in **Figure 1**. The workflow was evaluated on the prediction of FC from SC at the individual level rather than group level. Results are reported for two different datasets, namely: 326 unrelated subjects from the Human Connectome project (Van Essen et al., 2013) and 50 unrelated subjects from the Microstructure-Informed Connectomics dataset (Royer et al., 2021). We used the proposed approach to: *i*) study the relationship between the length of the random walks (*i.e.*, length of indirect paths between brain regions) and prediction accuracy (quantified via Pearson's correlation coefficient), *ii*) test the contribution and identify the weighting schemes to define the SC that best explain the observed brain function, *iii*) perform region- and network-specific analyses of FC prediction as a function of path length, and *iv*) compare the prediction performance of the proposed approach with several state-of-the-art methods.

## MATERIALS AND METHODS

### Datasets

We evaluated our proposed approach on diffusion magnetic resonance imaging (dMRI) and resting-state functional MRI (rs-fMRI) data provided by the Human Connectome Project (HCP) repository (Van Essen *et al.*, 2013), and on the Microstructure-Informed Connectomics (MICs) dataset (Royer *et al.*, 2021).

All MRI data used in this study were publicly available and anonymized. For HCP, we used data from the minimally processed S900 release. Participants who did not complete full imaging data and who had family relationships were excluded, resulting in a total of 326 participants (mean  $\pm$  SD age =  $28.56 \pm 3.73$  years; 55% females). Participant recruitment procedures and informed consent forms, including consent to share de-identified data, were previously approved by the Washington University Institutional Review Board as part of the HCP. For MICs, the data consist of 50 healthy volunteers ( $29.82 \pm 5.73$  years; 21 females) scanned between April 2018 and September 2020. All participants denied a history of neurological illness. The MICs dataset was approved by the Ethics Committee of the Montreal Neurological Institute and Hospital. Written informed consent, including a statement for openly sharing all data in anonymized form, was obtained from all participants.

### MRI acquisition

HCP participants were scanned using a Siemens Skyra 3T at Washington University. The T1-weighted (T1w) images were acquired using a magnetization-prepared rapid gradient echo (MPRAGE) sequence (repetition time (TR) = 2,400 ms; echo time (TE) = 2.14 ms; field of view (FOV) =  $224 \times 224$  mm<sup>2</sup>; voxel size = 0.7 mm<sup>3</sup>; and number of slices = 256). The T2-weighted (T2w) structural data were obtained with the T2-SPACE sequence, with an identical geometry as the T1w data but different TR (3,200 ms) and TE (565 ms). The dMRI data were acquired with the spin-echo echo-planar imaging (EPI) sequence (TR = 5,520 ms; TE = 89.5 ms; FOV =  $210 \times 180$  mm<sup>2</sup>; voxel size = 1.25 mm<sup>3</sup>; b-value = three different shells *i.e.*, 1,000, 2,000, and 3,000 s/mm<sup>2</sup>; number of diffusion directions = 270; and number of b0 images = 18). The rs-fMRI data were collected using a gradient-echo EPI sequence (TR = 720 ms; TE = 33.1 ms; FOV =  $208 \times 180$  mm<sup>2</sup>; voxel size = 2 mm<sup>3</sup>; number of slices = 72; and number of volumes = 1,200). During the rs-fMRI scan, participants were instructed to keep their eyes open looking at a fixation cross. Two sessions of rs-fMRI data were acquired; each of them contained data of left-to-right and right-to-left phase-encoded directions, providing up to four time series per participant.

For MICs, participants were scanned at the McConnell Brain Imaging Centre of the Montreal Neurological Institute and Hospital on a 3T Siemens Magnetom Prisma-Fit equipped with a 64-channel head coil. Participants underwent a T1w structural scan, followed by multi-shell dMRI and rs-fMRI. In addition, a pair of spin-echo images was acquired for distortion correction of individual rs-fMRI scans. Two T1w scans with identical parameters were acquired with a 3D-MPRAGE sequence (TR = 2300 ms, TE = 3.14 ms, TI = 900 ms, flip angle = 9°, iPAT = 2, partial Fourier = 6/8, voxel size = 0.8 mm<sup>3</sup>, matrix =  $320 \times 320$ , and number of slices = 224). Both T1w scans were visually inspected to ensure minimal head motion before they were submitted to further processing. A spin-echo EPI sequence with multi-band acceleration was used to obtain dMRI data, consisting of three shells with b-values 300, 700, and 2000s/mm<sup>2</sup> and 10, 40, and 90 diffusion weighting directions per shell, respectively (TR = 3500 ms, TE = 64.40 ms, voxel size = 1.6 mm<sup>3</sup>, flip angle = 90°, refocusing flip angle = 180°, FOV =  $224 \times 224$  mm<sup>2</sup>,

slice thickness = 1.6 mm, mb factor = 3, echo spacing = 0.76 ms, number of b0 images = 3). One rs-fMRI scan was acquired using multiband accelerated 2D-BOLD EPI (TR = 600 ms, TE = 30 ms, voxel size = 3mm<sup>3</sup>, flip angle = 52°, FOV = 240×240 mm<sup>2</sup>, slice thickness = 3 mm, mb factor = 6, echo spacing = 0.54 ms). Participants were instructed to keep their eyes open, look at a fixation cross, and not fall asleep. A complete list of acquisition parameters can be found in the detailed imaging protocol provided by (Royer *et al.*, 2021).

### *Data preprocessing*

HCP data underwent the initiative's minimal preprocessing pipelines (Glasser *et al.*, 2013). In brief, structural MRI data underwent gradient nonlinearity and b0 distortion correction, followed by co-registration between the T1w and T2w data using a rigid-body transformation. Bias field correction was performed by capitalizing on the inverse intensities from the T1- and T2-weighting. Processed data were nonlinearly registered to MNI152 space and the white and pial surfaces were generated by following the boundaries between different tissues (Dale *et al.*, 1999; Fischl, 2012; Fischl *et al.*, 1999a, 1999b). The white and pial surfaces were averaged to generate a mid-thickness surface, which was used to generate the inflated surface. The spherical surface was registered to the Conte69 template with 164k vertices (Van Essen *et al.*, 2012) using MSMAll (Glasser *et al.*, 2016; Robinson *et al.*, 2014) and downsampled to a 32k vertex mesh. The dMRI data underwent b0 intensity normalization, and EPI distortions were corrected by leveraging reversed phase-encoded directions. The dMRI data was also corrected for eddy current distortions and head motion. The rs-fMRI data preprocessing involved corrections for EPI distortions and head motion, and fMRI data were registered to the T1w data and subsequently to MNI152 space. Magnetic field bias correction, skull removal, and intensity normalization were performed. Noise components attributed to head movement, white matter, cardiac pulsation, arterial, and large vein related contributions were automatically removed using FIX (Salimi-Khorshidi *et al.*, 2014). The minimal preprocessing with FIX-denoising pipeline of the HCP performs a high-pass filtering with a cutoff of 2,000 s full width at half maximum (FWHM) (Glasser *et al.*, 2013). Preprocessed time series were mapped to standard grayordinate space, with a cortical ribbon-constrained volume-to-surface mapping algorithm. The total mean of the time series of each left-to-right/right-to-left phase-encoded data was subtracted to adjust the discontinuity between the two datasets and they were concatenated to form a single time series data.

For the MICs dataset, T1w images were anonymized and de-identified by defacing all structural volumes. Each T1w scan was deobliqued and reoriented. T1w scans were then linearly co-registered and averaged, automatically corrected for intensity nonuniformity (Tustison *et al.*, 2010), and intensity normalized. Resulting images were skull-stripped, and subcortical structures were segmented using FSL FIRST (Jenkinson *et al.*, 2012). Cortical surface segmentations were generated from native T1w scans using FreeSurfer 6.0 (Dale *et al.*, 1999; Fischl, 2012; Fischl *et al.*, 1999a, 1999b). The dMRI data were pre-processed using MRtrix (Tournier *et al.*, 2012, 2019). The dMRI data underwent b0 intensity normalization, and were corrected for susceptibility distortion, head motion, and eddy currents. Required anatomical features for tractography processing were co-registered to native dMRI space using affine transformation tools implemented in Advanced Neuroimaging Tools (ANTs) (Tustison & Avants, 2013). Diffusion processing and tractography were performed in native dMRI space. For rs-fMRI, images were pre-processed using AFNI (Cox, 1996) and FSL (Jenkinson *et al.*, 2012). The first five volumes were discarded to ensure magnetic field saturation. Images were reoriented, and motion as well as distortion

corrected. Nuisance variable signal was removed using ICA-FIX (Salimi-Khorshidi *et al.*, 2014) and by performing spike regression. Volume timeseries were averaged for registration to native FreeSurfer space using boundary-based registration (Greve & Fischl, 2009). Native timeseries were mapped to individual surface models using a boundary-based registration and smoothed using a Gaussian kernel (FWHM=10mm, smoothing performed on native midsurface mesh) using workbench. The preprocessing of the MICs dataset was performed with micapipe (<https://micapipe.readthedocs.io>).

### Functional and structural connectome generation

To estimate FC matrices, individual rs-fMRI timeseries mapped to individual surface models were averaged within parcels defined by the cortical parcellation scheme. Cortical timeseries were sampled from each vertex of the native FreeSurfer space cortical surface segmentation, and averaged within surface parcels. Individual functional connectomes were generated by cross-correlating all nodal timeseries.

SC representations were generated from preprocessed dMRI data using MRtrix (Tournier *et al.*, 2012, 2019). Different tissue types of cortical and subcortical grey matter, white matter, and cerebrospinal fluid were segmented using T1-weighted image for anatomical constrained tractography (Smith *et al.*, 2012). Multi-shell and multi-tissue response functions were estimated (Christiaens *et al.*, 2015) and constrained spherical-deconvolution and intensity normalization were performed (Jeurissen *et al.*, 2014). The initial tractogram was generated with 40 million streamlines, with a maximum tract length of 250 and a fractional anisotropy cutoff of 0.06. Spherical-deconvolution informed filtering of tractograms (SIFT2) was applied to reconstruct whole brain streamlines weighted by cross-section multipliers (Smith *et al.*, 2015). To build a structural connectome, the reconstructed cross-section streamlines were mapped to selected cortical parcellation schemes.

### Proposed framework

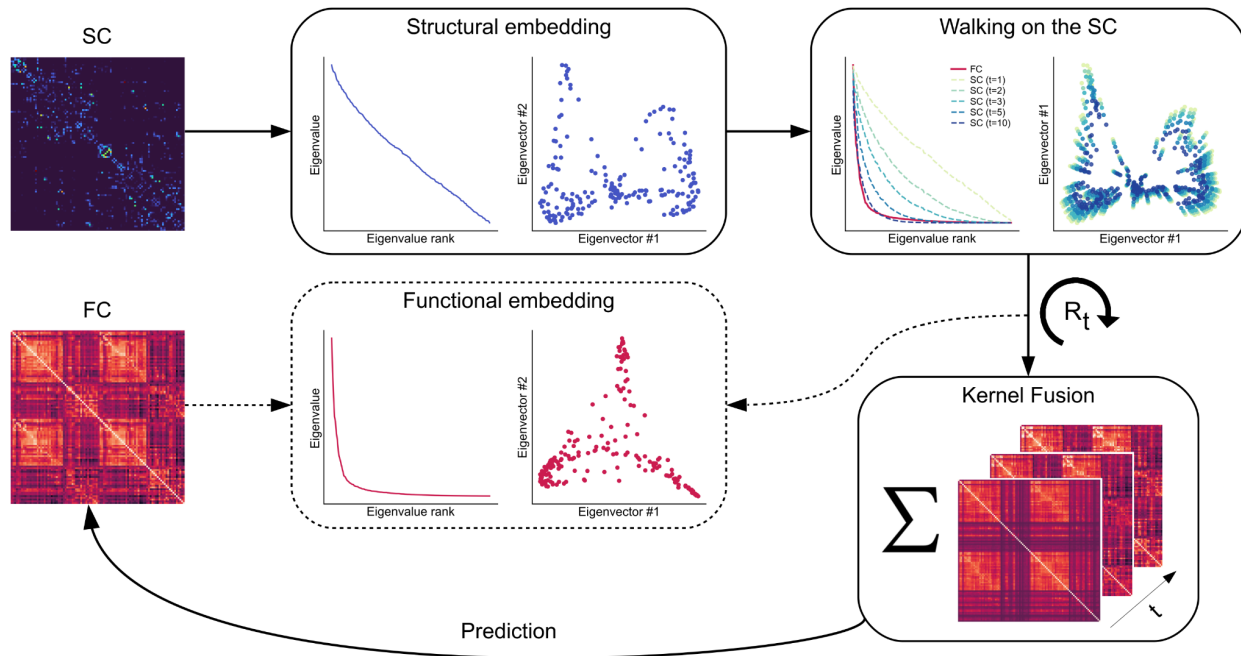
Let  $S \in \mathcal{R}^{n \times n}$  be the connectivity matrix representing a given SC, where each entry  $S(i, j)$  is the weight of the edge connecting the  $i$ -th and  $j$ -th cortical locations, computed as the total number of streamlines connecting both locations, such that  $S(i, j) = S(j, i)$  and  $S(i, j) \geq 0, \forall i, j = 1, \dots, n$ . Our purpose is to predict the FC, the correlations of resting-state functional signals, from its corresponding SC (i.e.,  $S$ ). To do so, we leverage diffusion maps (Coifman & Lafon, 2006). We first proceed by normalizing the SC matrix to define the diffusion operator  $P$ :

$$P(i, j) = \frac{S(i, j)}{q(i)}, \quad (1)$$

where  $q(i) = \sum_k S(i, k)$  denotes the degree in the connectome, such that  $\sum_k P(i, k) = 1$ . Now  $P_{ij}$  can be viewed as the probability for a random walker on the SC  $S$  to make a step from the  $i$ -th to  $j$ -th cortical locations. As  $P$  is not symmetric, we can further define a symmetric operator  $\Delta$ :

$$\Delta(i, j) = \frac{\sqrt{q(i)}}{\sqrt{q(j)}} P(i, j) = \frac{S(i, j)}{\sqrt{q(i)} \sqrt{q(j)}}. \quad (2)$$





**Figure 1. Schematic of the proposed approach.** Given a pair of structural (SC) and functional (FC) connectivity matrices, first we use diffusion maps to obtain the structural embedding of the SC. The functional embedding (dashed) is only shown for visualization purposes, to illustrate the differences that might exist between the structural and functional spectra and also their eigenvectors. By increasing the diffusion time  $t$  (see Walking on the SC), we can see that the structural eigenvalues approximate the functional eigenvalues shown in red. For each diffusion time, we can obtain a different representation of the structural diffusion coordinates. The larger the diffusion time is, the closer the brain regions are to each other in the structural embedding, and hence their pairwise connectivity is increased. Then, a rotation matrix is used for each diffusion time to obtain a kernel that represents the predicted FC at time  $t$ . This can be seen as finding a rotation matrix to align to structural embedding to the functional one, see dashed line. The kernels are then fused to obtain the final predicted FC matrix.

In matrix notation, we have that  $\Delta = Q^{-1/2} S Q^{-1/2}$ , where  $Q$  denotes the degree matrix of  $S$  (a diagonal matrix such that  $Q(i, i) = q(i) = \sum_k S(i, k)$ ). Using spectral theory, it can be shown that  $\Delta$  has the following eigendecomposition:

$$\Delta(i, j) = \sum_{k \geq 0} \lambda_k^2 \psi_k(i) \psi_k(j), \quad (3)$$

where  $1 = \lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \dots$  is the eigenspectrum and  $\{\psi_k\}$  the corresponding eigenvectors of  $\Delta$ . This operator shares the same spectrum with  $P$  and its eigenvectors are orthogonal, unlike those of  $P$  (Coifman & Hirn, 2014). Now, a walk of length  $t$  in the SC can be represented by the diffusion maps  $\Psi_t$  as follows:

$$\Psi_t(i) = (\lambda_0^t \psi_0(i), \lambda_1^t \psi_1(i), \dots)^T, \quad (4)$$



where  $T$  stands for transpose. A desirable property of the diffusion map  $\Psi_t$  is that it embeds the data into a Euclidean space in which the Euclidean distance is equal to the diffusion distance  $D_t$ :

$$D_t^2(i, j) = \sum_{k \geq 0} \lambda_k^{2t} (\psi(i) - \psi(j))^2 = \|\Psi_t(i) - \Psi_t(j)\|^2, \quad (5)$$

where  $\|\cdot\|$  denotes the  $l_2$ -norm. With this metric, we can capture the connectivity of two cortical locations for each length of the walks  $t$  in the SC. Note that as  $t$  grows, the diffusion distance between the cortical locations will decrease and will be mainly driven by the first diffusion coordinates, i.e., those coordinates corresponding to the largest eigenvalues. This will allow us to approximate the distances based solely on the dominant eigenvectors and reduce the dimensionality of the diffusion maps.

Let  $F \in \mathcal{R}^{n \times n}$  be the FC matrix, built using time-series correlation analysis of an fMRI scan from the same subject. Since diffusion maps define a Euclidean space, we propose to predict the FC using a kernel fusion approach:

$$\begin{aligned} & \underset{\alpha \in \mathcal{R}^m}{\operatorname{argmin}} \left\| F - \sum_{t=1}^m \alpha_t K_t \right\|_F^2 + \mu_1 \|\alpha\|_2^2 \\ \text{s.t.} \quad & \alpha_t \geq 0, \forall t = 1, \dots, m, \end{aligned} \quad (6)$$

where  $K_t(i, j) = \exp(-\gamma \|\Psi_t(i) - \Psi_t(j)\|^2) = \exp(-\gamma D_t^2(i, j))$  is a radial basis function (RBF) kernel built from the diffusion maps  $\Psi_t$  for each diffusion time  $t$ , and  $\gamma$  is the kernel bandwidth;  $m$  is the total number of walks considered,  $\alpha_t \geq 0$  is the coefficient corresponding to the RBF kernel  $K_t$ ,  $\mu_1$  is a trade-off parameter, and  $\|\cdot\|_F$  is the Frobenius norm. Given that  $K_t(i, j) \in [0, 1]$  and  $F(i, j) \in [-1, 1]$ , in our setting we scale the FC matrices to the range  $[0, 1]$  for training and undo this operation after prediction. Since we will assess performance using Pearson's correlation coefficient, which is invariant under positive linear transformations, this scaling has no effect on the results.

Although we can approximate the functional eigenvalues with an increasing number of walks in the structural embedding, as shown in **Figure 1**, the structural and functional embeddings do not share the same diffusion coordinates. We therefore propose to find a transformation of the structural embedding to more faithfully reconstruct the FC. Let  $\Upsilon_t \in \mathcal{R}^{n \times p}$  be a matrix representing the diffusion coordinates of the SC at time  $t$  and  $p \leq n$ , we aim to find a rotation matrix  $\Omega \in \mathcal{R}^{n \times n}$  to transform the diffusion coordinates as follows:

$$\begin{aligned} & \underset{\substack{\Omega \in \mathcal{R}^{n \times n} \\ \alpha \in \mathcal{R}^m}}{\operatorname{argmin}} \left\| F - \sum_{t=1}^m \alpha_t k(\Omega \Upsilon_t) \right\|_F^2 + \mu_1 \|\alpha\|_2^2 \\ \text{s.t.} \quad & \Omega \Omega^T = \Omega^T \Omega = I_n, \det(\Omega) = 1, \\ & \alpha_t \geq 0, \forall t = 1, \dots, m, \end{aligned} \quad (7)$$

where  $I_n$  is an  $n \times n$  identity matrix,  $\det(\cdot)$  stands for matrix determinant, and  $\mathfrak{I}$  and  $\Gamma_t = \Omega \Upsilon_t$ , with  $\Gamma_t$  denoting the rotated structural diffusion map, and  $\Gamma_t(i)$  the  $i$ -th row of  $\Gamma_t$ . It is worth noting that the

scaling of the eigenvectors is different for each diffusion time  $t$ , and because of the spectrum decay, we will be having fewer and fewer diffusion coordinates contributing to the computation of the RBF kernels. This will be producing a different kernel  $K_t$  for each diffusion time  $t$ . Based on this observation, we extend our approach to include a rotation for each diffusion time:

$$\begin{aligned} & \underset{\substack{\Omega_1, \dots, \Omega_m \in \mathcal{R}^{n \times n} \\ \alpha \in \mathcal{R}^m}}{\operatorname{argmin}} \quad \|F - \sum_{t=1}^m \alpha_t k(\Omega_t \Upsilon_t)\|_F^2 + \mu_1 \|\alpha\|_2^2 + \frac{\mu_2}{m-1} \sum_{t=1}^{m-1} d(\Omega_t, \Omega_{t+1}) \\ \text{s.t.} \quad & \Omega_t \Omega_t^T = \Omega_t^T \Omega_t = I_n, \det(\Omega_t) = 1, \\ & \alpha_t \geq 0, \forall t = 1, \dots, m, \end{aligned} \quad (8)$$

where we include a third term and its trade-off parameter  $\mu_2$  to avoid overfitting, with  $d(\Omega_t, \Omega_{t+1})$  denoting the distance between two consecutive rotation matrices. When  $\mu_2 = 0$ , no restrictions are imposed on the rotation matrices, whereas when  $\mu_2 > 0$ , this term enforces the rotations of consecutive diffusion times to be similar to each other. In the extreme case where  $\mu_2$  tends to infinity, the problem amounts to finding one single rotation shared by all diffusion times (*as shown in Eq. (7)*). Biologically, one can think of these rotation matrices as identifying the optimal paths through which to propagate information between different regions of the brain. Since we have one rotation matrix for each diffusion time, each rotation will identify paths of a specific length. That is, the rotation corresponding to diffusion time  $t$  will attempt to find the optimal paths of length  $t$  that connect a pair of brain regions.

To infer the latent variables in our problem, we employ an alternating optimization technique. We minimize the cost function in Eq. (8) for each output variable, while holding the estimates of the other unknowns constant. Note that, for brevity, in our formulation we only considered a single subject, but the optimization is performed for multiple subjects. To determine the set of rotations, we recast our cost function to a Riemannian manifold optimization problem (Absil *et al.*, 2009; Hu *et al.*, 2020). Riemannian optimization translates a constrained optimization problem into an unconstrained optimization problem where the constraints are implicitly defined by the search space. Particularly, for our problem, the target manifold is the Special Orthogonal group  $\mathcal{SO}(n)$ . For more than one rotation, we can define the target manifold as the product manifold of rotations,  $(\Omega_1, \Omega_2, \dots, \Omega_m) \in \mathcal{SO}(n)^m$ :

$$\underset{\Omega_1, \dots, \Omega_m \in \mathcal{SO}(n)}{\operatorname{argmin}} \quad \|F - \sum_{t=1}^m \alpha_t k(\Omega_t \Psi_t)\|_F^2 + \frac{\mu_2}{m-1} \sum_{t=1}^{m-1} \|\log(\Omega_t^T \Omega_{t+1})\|_F, \quad (9)$$

where the distance between rotation matrices  $d(\Omega_t, \Omega_{t+1}) = \|\log(\Omega_t^T \Omega_{t+1})\|_F$  is suitably chosen to be the geodesic distance on  $\mathcal{SO}(n)$  and  $\log(\cdot)$  denotes the matrix logarithm (Boumal *et al.*, 2014). Given that kernels corresponding to consecutive numbers of walks are similar, this term imposes that their rotation matrices must lie close to each other on the rotation manifold. The problem in Eq. (9) can be solved using Riemannian optimization algorithms. In our case, we use the Riemannian conjugate gradient algorithm (Absil *et al.*, 2009) as implemented in Pymanopt (Townsend *et al.*, 2016).

Once we have estimated the rotation matrices, the next step is solving for  $\alpha$ , which amounts to ridge regression with non-negativity constraints on the coefficients. Let  $X = (\operatorname{uvec}(K_1), \operatorname{uvec}(K_2), \dots, \operatorname{uvec}(K_m))$  and  $y = \operatorname{uvec}(F)$ , where  $\operatorname{uvec}(\cdot)$  returns the upper

triangular part of a symmetric matrix, without the elements in the main diagonal. This optimization problem amounts to:

$$\underset{\alpha \in \mathcal{R}^m}{\operatorname{argmin}} \alpha^T (X^T X + \mu_1 I) \alpha - 2y^T X \alpha + y^T y \quad \text{s.t.} \quad \alpha \geq 0, \quad (10)$$

where the third term does not affect the optimization. This problem can be solved with conventional quadratic programming. The proposed approach is implemented in Python and the code is publicly available at ([https://github.com/MICA-MNI/micaopen/sf\\_prediction](https://github.com/MICA-MNI/micaopen/sf_prediction)).

### Experimental settings

The performance of the proposed approach in predicting FC from SC was assessed in the HCP and MICs datasets. For MICs, prediction accuracy was based on a 3-fold cross-validation (CV) strategy based on all 50 individuals in the dataset. For HCP (n=326), the data was randomly split in 3 subsets: 50 individuals were selected for CV, 250 for holdout, and the remaining 26 individuals were used for parameter tuning. Following previous studies, prediction performance was reported in terms of Pearson's correlation coefficient based on the upper triangular parts (excluding the main diagonal) of both empirical and predicted FC matrices. Regarding parameter tuning, we used an independent subset (26 subjects) from the HCP dataset to find the optimal values for the hyperparameters (i.e.,  $\mu_1$ ,  $\mu_2$ , and  $\gamma$  of the RBF kernels) of our proposed approach. Note that this subset was not used to report any results. The best values for both  $\mu_1$  and  $\mu_2$  were chosen from a grid of 9 equidistant points in logarithmic scale in the interval [1e-4, 1e4]. The optimal value for  $\mu_1$  was found to be 100. For the Riemannian regularization, our initial benchmarking showed a negligible improvement for  $\mu_2 \leq 0.001$ , and we therefore chose to report our results for  $\mu_2 = 0$ . For the RBF kernel,  $\gamma$  was chosen to be the standard deviation of the diffusion distances for each random walk. These optimal hyperparameter values were then used to assess performance in both the HCP and MICs datasets.

Our proposed approach was tested using SC and FC matrices built based on two different cortical atlases: *i*) a parcellation derived using functional MRI data (Schaefer et al., 2018), and *ii*) a structurally-defined parcellation based on a more fine-grained clustering of the well-established Desikan Killiany parcellation (Desikan et al., 2006; Vos de Wael et al., 2020). To assess the robustness of the proposed approach across different spatial scales, our experiments were repeated for two different parcellation granularities, using parcellations with 100 and 200 cortical regions. We also explored different weighting schemes to define the SC matrices in explaining ongoing brain function. We analyzed three versions of the SC: *i*) binary SC, where each entry in the SC matrix is set to 1 only if there is at least one streamline connecting the corresponding regions of the brain, *ii*) length-based SC, which is built using an RBF kernel based on the mean length of the streamlines connecting pairs of brain regions, and *iii*) count-based SC, which is defined based in the number of streamlines connecting each pairs of brain regions. Unless otherwise stated, our experiments are based on SC matrices defined using the latter weighting scheme (i.e., count-based). We also investigated the contribution of diffusion time at network level, and the role it plays in strengthening the structure-function coupling and its relationship with the principal functional gradient (Margulies et al., 2016). The principal functional gradient was estimated using BrainSpace (Vos de Wael et al., 2020).

We compared the performance of the proposed approach with state-of-the-art methods. Among those based on the structural eigenvectors, we compared our approach to the Multiple Kernel Learning (MKL) (Surampudi et al., 2018) and Spectral (Becker et al., 2018) approaches. Briefly, in the MKL approach, multiple diffusion kernels on the SC are linearly combined using LASSO (Tibshirani, 1996) to predict FC. The Spectral approach, on the other hand, learns a shared functional embedding and a mapping from the functional to the structural eigenvalues for each individual, which are then used to build the predicted FC matrices. Note that, in the Spectral approach, the individual structural eigenvectors are not taken into consideration to predict FC. Outside the eigenvector-based category, we included results using the single Laplacian-based diffusion kernel (SDK) proposed in (Abdelnour et al., 2014), and the series expansion approach (NLSA) in (Meier et al., 2016). SDK defines a single diffusion kernel at a specific scale (or diffusion time) from the symmetric normalized Laplacian matrix of the SC. In our work, the optimal scale was chosen so that it provided the best performance in the training data. NLSA, on the other hand, predicts FC as a truncated Taylor series expansion of the SC (Tewarie et al., 2020). We used code provided by the authors for both MKL (<https://github.com/govindasurampudi/MKL>) and Spectral (<https://brainopt.github.io/spectral-mapping>). Finally, our approach used multiple kernels along with their corresponding rotation matrices. To elucidate the contribution of each of these components, we included two additional versions of our method to the comparison table: 1) "SingleLength" only used one single kernel or random walk for prediction, and 2) "SharedRot" used a single rotation matrix shared by all the random walks.

## RESULTS

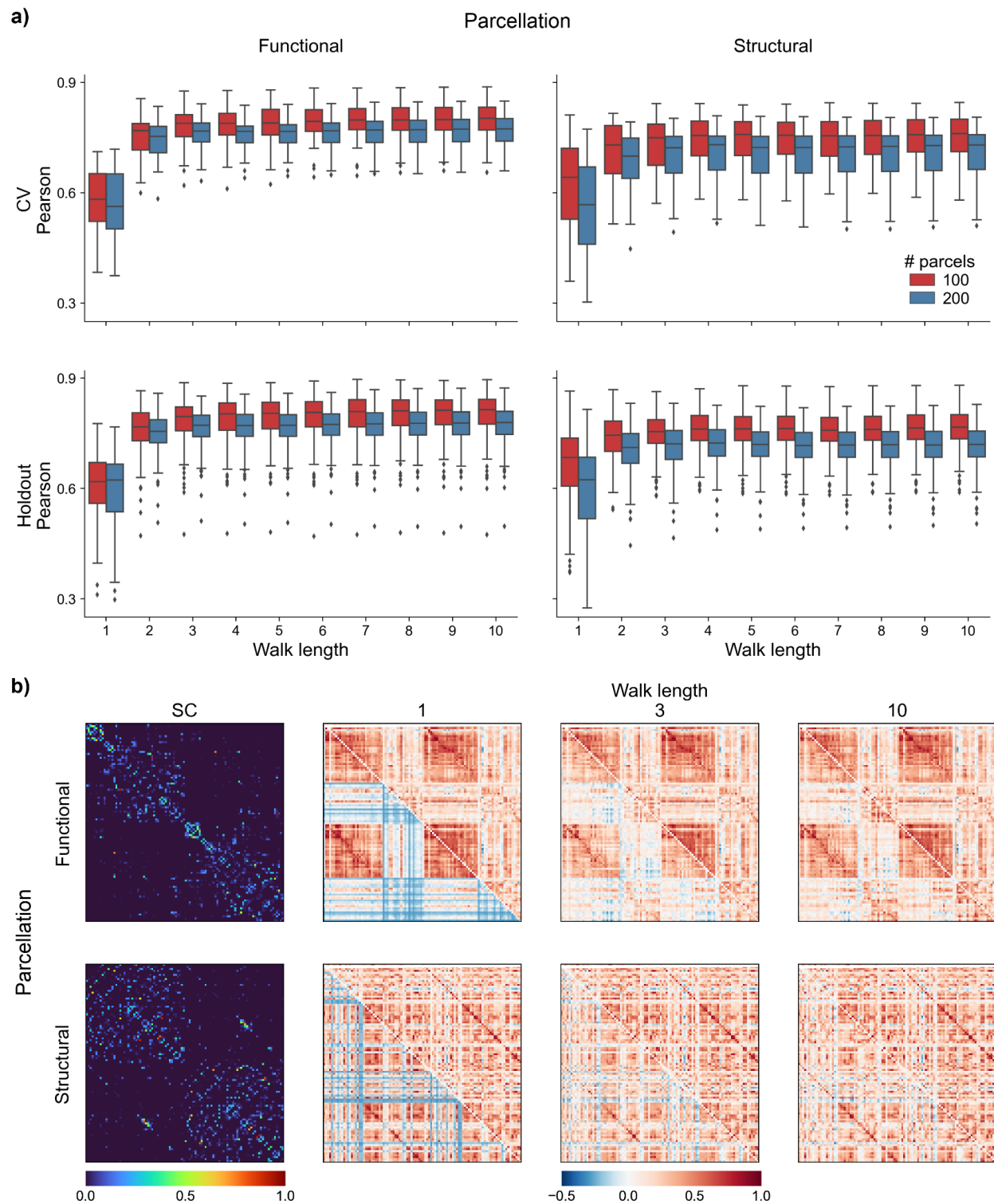
### *Walking on the structural connectome increases prediction accuracy*

To faithfully reconstruct the FC from the sparse SC, we first investigated the role of diffusion time (or path length) in the prediction accuracy of our proposed approach. We explored the contribution of random walks of different lengths, ranging from 1 to 10, such that for a given maximum length of 5 for example, we considered random walks of any length below or equal to 5. In this way, we are not only including the static direct connections present in the SC, but also allow our method to incorporate information about the SC at multiple scales, *i.e.*, after propagating information between indirectly connected brain regions.

The prediction accuracy of our proposed approach was assessed in two different datasets, namely: HCP and MICs. For HCP, **Figure 2a** shows boxplots with the correlation between the empirical and predicted FC matrices using structural random walks with maximum lengths ranging from 1 to 10. Results are reported in both cross-validation (using a 3-fold CV based on 50 randomly chosen subjects) and holdout data (250 subjects) subsets, and using different cortical parcellation atlases (defined using structural and functional information) and number of regions (100 and 200 parcels). These results showed that the highest performance (in terms of Pearson's correlation coefficient) was achieved with approximately a maximum path length of 3-4. After that, performance plateaued and the change in prediction accuracy, although positive, was in most scenarios negligible. Overall, these trends were consistent in both cross-validation and holdout datasets, and across the different parcellations and numbers of regions. In the holdout set of HCP, for example, prediction accuracy improved substantially when considering random walks of length up to 3 (0.174/0.165 increase in mean Pearson's correlation coefficient for 100/200 regions functional parcellations, and 0.085/0.112 for the analogous structural parcellations), whereas with the incorporation of additional walks (those with path lengths from 4 to 10), we only found very small improvements in performance (0.017/0.007 and 0.013/0.002 average increase for the 100/200-node functional and structural parcellations, respectively). Moreover, when comparing the type and size of the parcellations, prediction accuracy was consistently lower in structural than in functional parcellations, and dropped when increasing the number of parcels used to build the connectivity matrices. Results on the MICs dataset, based on a 3-fold cross-validation, are shown in **Figure S1** of the supplementary materials. Similar to HCP, highest performance was achieved with random walks of length 3 or shorter, with prediction accuracy decaying when the number of cortical parcels increased. However, no clear trend was found when comparing performance based on the structural and functional parcellation (see *Comparison with state-of-the-art methods* section).

To further illustrate how brain function emanates from the SC as the length of the random walks increases (*i.e.*, as we increasingly incorporate indirect connections between cortical regions), **Figure 2b** displays the original SC matrices and the predicted FC matrices corresponding to the individuals that achieved the median Pearson's correlation coefficient in the holdout subset of the HCP dataset. The predicted FC matrices corresponding to random walks of length 1, 3 and 10 are displayed in the lower triangular parts, with their respective empirical FC shown in the upper triangular parts. Results shown in **Figure 2b** are based on both the structural and functional parcellations with 100 regions. For the 200-node parcellations, results are shown in **Figure S2**. With both parcellation types and sizes, we found some similarities emerge with random walks of length 1, but these considerably increased as larger walks (indirect paths) were incorporated. This points to communication between different cortical regions through indirect paths of lengths larger than one, which highlights the importance of polysynaptic mechanisms in the emergence of the brain's function and, at the same time, shows that only a small number of hops (path lengths  $\leq 3$ ) in the

SC may be required to accurately predict FC, as shown by the results in **Figure 1a**. Furthermore, we also found that functionally-defined cortical parcellations were more suitable than structurally-derived ones in the prediction of FC from SC.



**Figure 2. Functional connectivity prediction accuracy using random walks of different lengths in the structural connectome. a)** Boxplots of Pearson's correlation coefficient between empirical and estimated FC matrices using



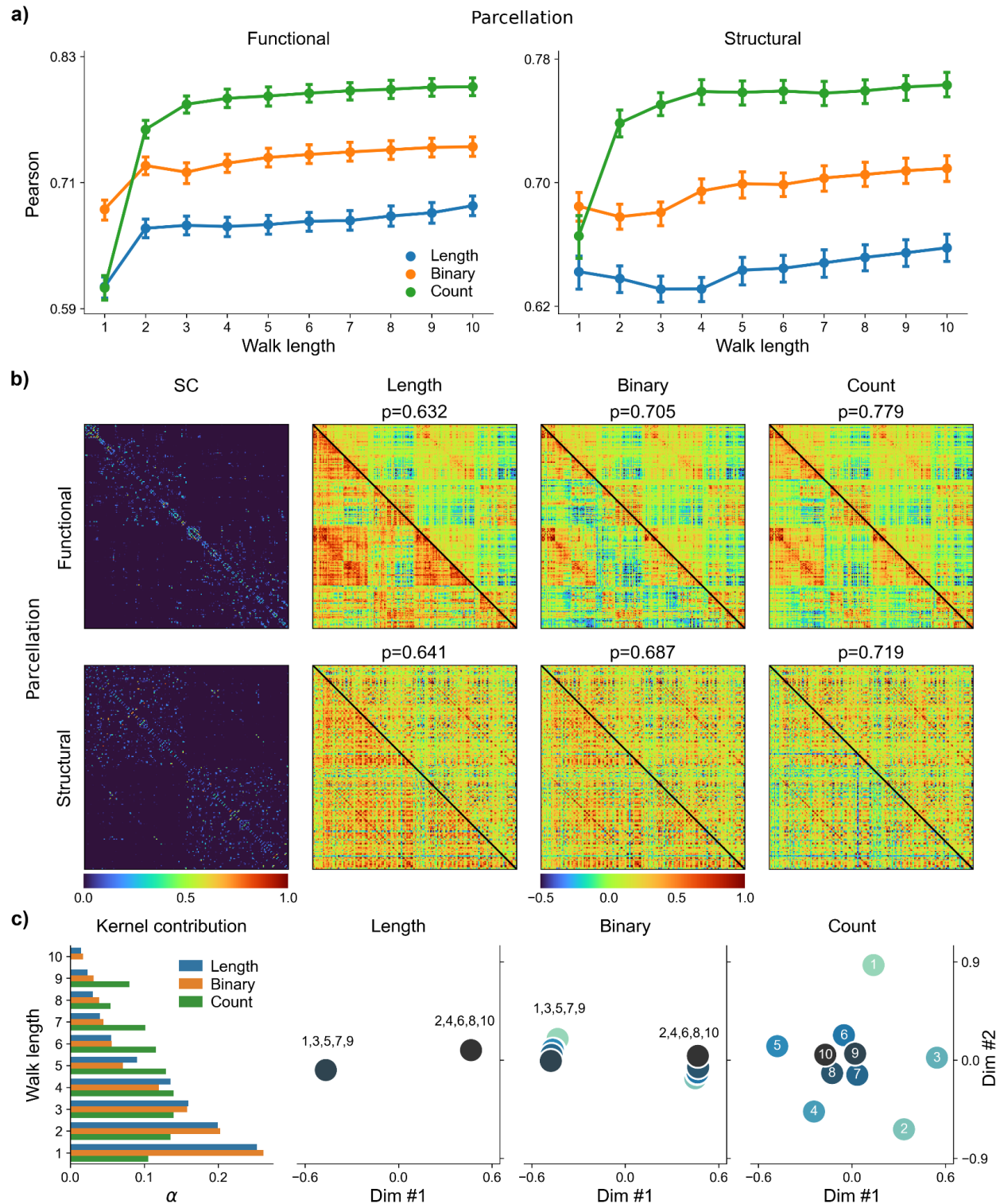
random walks of maximum length ranging from 1 to 10. Results are reported in cross-validation (3-fold CV based on 50 subjects) and holdout (250 subjects) data of HCP using functionally- (*left*) and structurally-derived cortical parcellations (*right*) with 100 and 200 regions. Boxes denote the interquartile range (IQR) between the first and third quartiles, and the line inside denotes the median. Whiskers extend to points that lie within 1.5 IQRs of the lower and upper quartiles, and the black diamonds denote outliers. **b**) Estimated functional connectivity matrices corresponding to the subjects that achieved the median Pearson's correlation coefficient based on functional (*top*) and structural (*bottom*) parcellations with 100 parcels. From left to right: structural connectome and estimated FC matrices for random walks of length 1, 3 and 10. Empirical and estimated functional connectivity matrices are shown in upper and lower triangular parts, respectively.

### Comparison of structural connectome characteristics

The structural connectivity matrices used so far in our experiments were derived from fiber density estimates, such that each entry in the matrix denotes the number of streamlines connecting the specific pair of brain regions. Here, we sought to investigate the contribution of different weighting schemes to define SC. First, we used different versions of the structural connectivity matrices: *i*) the original SC matrices, where edges carried information about the streamline count, *ii*) binary SC, where we preserved the same edges but ignored edge information (with the weights of all edges set to 1), and *iii*) length-based SC, with the same edges but the weights were based on the inverse of the average length of the streamlines. We repeated the experiments to learn the set of rotations and kernel coefficients for each version of the structural connectome.

As shown in **Figure 3a**, based on results on the holdout data of HCP and 200-node parcellations, when using SC based on streamline count we achieved the best accuracy in predicting FC, with a considerable improvement over binary SC, which in turn outperformed length-based SC. More importantly, diffusion time did not seem to have the same contribution when using binary or length-based SC matrices for prediction. With the functional parcellation, a path length of 1 with the binary SC provided better FC predictions than the count-based SC, however, as the lengths of the paths increased, the boost in performance only occurred with count-based SC. With the structural parcellation, this was even more evident, with prediction accuracy remaining almost constant as diffusion time increased when using binary and length-based SC. Following the diffusion maps framework, with increasing diffusion time, pairs of brain regions will increase the strengths of their connections according to the strengths of their immediate direct connections and those of their neighbors. For the purposes of brain function, this may suggest that the more important the 'structural' connection between a pair of regions is, the more fibers the brain invests in its construction. Examples of predicted FC matrices using binary and length-based SC are displayed in **Figure 3b**. These examples correspond to the subjects that achieved the mean Pearson's correlation coefficient in the holdout set of HCP for the 200-node functional and structural parcellations. As we can see, the proposed approach produced the least accurate FC matrices when using length-based SC matrices. The fact that better prediction accuracies were achieved by binary SC matrices, which may be an oversimplified representation of the connectome, indicates that the length of a given connection (quantified as the mean length along all interconnecting streamlines) may not be as informative as its density.





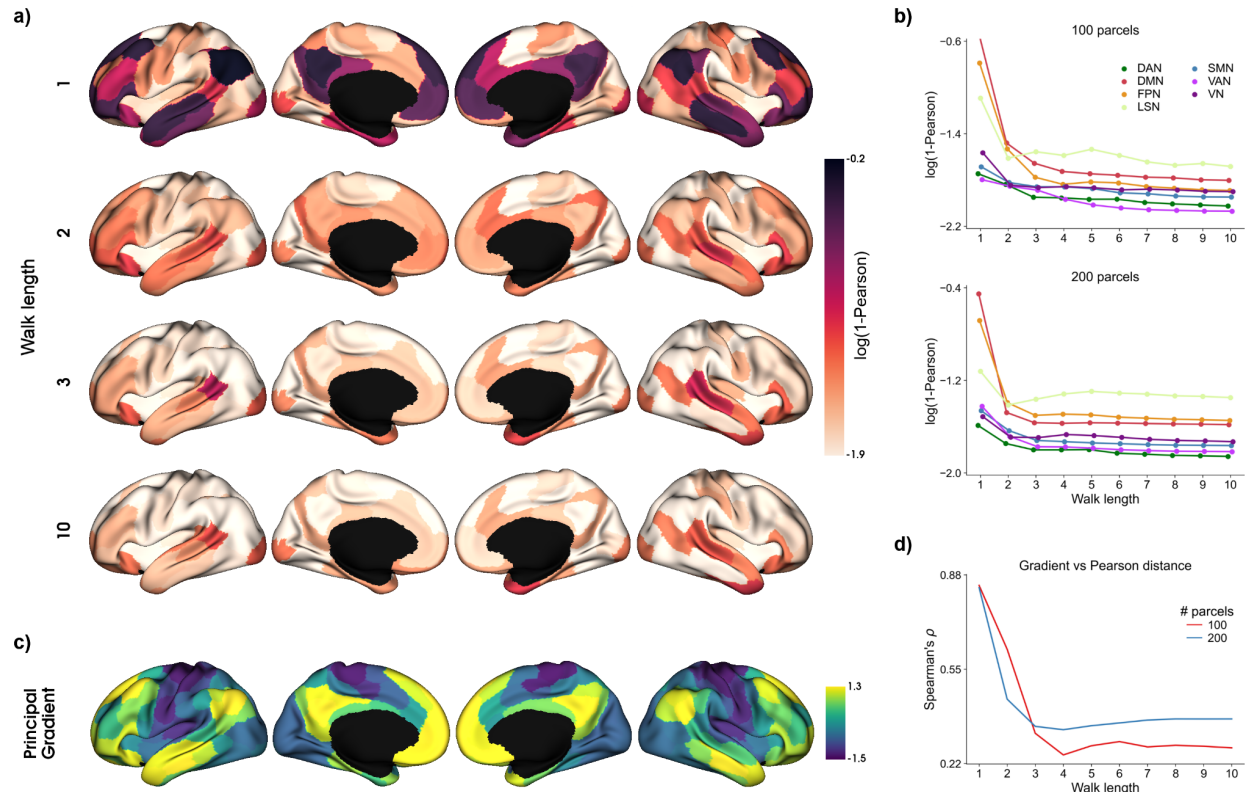
**Figure 3. Comparison of prediction performance and model parameters based on binary, length- and count-based SC matrices.** **a)** Mean Pearson correlation (and 95% confidence interval) between empirical and estimated FC matrices based on SC matrices built using streamline length, streamline count, and the binarized connectome. Results are reported for random walks of length 1 to 10 on the SC for both the 200-node functional (*left*) and structural (*right*) parcellations. **b)** Estimated FC matrices corresponding to the subjects that achieved the median Pearson's correlation

coefficient based on functional (*top*) and structural (*bottom*) parcellations with 200 parcels. From left to right: count-based SC matrices, and empirical and estimated FC matrices based on streamline length, binary SC, and streamline count. The displayed estimated connectivity matrices correspond to using random walks of length  $\leq 10$ . **c)** Optimal kernel coefficients (normalized),  $\alpha$ , and rotations matrices obtained when trained with binary, length- and count-based SC matrices using random walks of length  $\leq 10$  and the 200-node functional parcellation. Rotation scatterplots are based on the two first dimensions of multidimensional scaling (using the geodesic distance between rotation matrices). Results are reported in the holdout data of the HCP dataset.

Finally, in **Figure 3c** we can find the optimal kernel coefficients and rotations matrices learnt by our proposed approach when trained with each of the three different versions of the SC matrices. When using count-based SC, the kernels that most contributed to the final prediction corresponded to diffusion times 2-5, followed by 6 and 1. Kernels with higher diffusion times ( $t > 6$ ) had the least contribution. On the other hand, with both binary and length-based SC, kernel contributions decayed monotonically with increasing diffusion time. The most relevant differences, however, were found when comparing the rotation matrices learnt from the different SC versions. The rotation matrices, displayed using multidimensional scaling (based on the geodesic distance between rotations, as defined in Eq. (9)), were clustered in two groups of very similar rotations when using binary and length-based SC. On the other hand, with count-based SC matrices, we found different rotations for each diffusion time until 5, while rotations were very similar for diffusion times of 6 and greater. These results indicate that diffusion time (*i.e.*, considering increasingly larger paths) is of substantial added value to the prediction of FC when using count-based SC, as opposed to binary and length-based SC data.

#### Region- and network-wise analysis of predicted functional connectivity

In this section, we scrutinized our results to further investigate the role of diffusion time at the regional and network-level. **Figure 4a** shows spatial maps of prediction error (measured as  $\log(1-Pearson)$ ) for different diffusion times (*i.e.*, 1, 2, 3 and 10). Here we display the average prediction error in the holdout data of HCP for each cortical parcel (*i.e.*, derived from row-wise correlations). There were clear improvements across the whole cortex in the quality of the predictions as increasingly longer paths were considered, although they became subtler with higher diffusion times. The cortical regions where our approach produced the largest prediction errors (irrespective of diffusion time) were confined bilaterally to the lateral temporal lobe and frontal cortices. At the network level, derived using a previous cortical decomposition into seven intrinsic functional networks (Yeo *et al.*, 2011), prediction error decreased substantially from walk lengths of 1 to 3 (see **Figure 4b**). With higher walk lengths, we can only observe minor changes in prediction error across all networks. These findings were consistent across different parcellation granularities. Moreover, from these results we could identify two different groups of networks according to their prediction errors. FC prediction was more accurate in visual, somatomotor and both attention networks than in the default mode, frontoparietal and limbic networks. The latter were the networks that most benefited from incorporating indirect paths to the prediction, indicating that polysynaptic connections may have an important contribution to functional connectivity patterns, especially in these transmodal *i.e.*, heteromodal and paralimbic systems. This was further confirmed when analyzing the relationships between the principal functional gradient (see **Figure 4c**) and the spatial maps of prediction error. As shown in **Figure 4d**, these maps showed very high correlations with the principal functional gradient when using short paths (Spearman's  $r=0.843/0.835$  in 100/200-node functional parcellation when diffusion time=1), but considerably decrease as diffusion time increased (Spearman's  $r=0.250/0.338$  when diffusion time=4).



**Figure 4. Prediction accuracy at the network level.** **a)** Lateral and medial views of left and right hemispheres showing spatial cortical maps of average FC prediction error in terms of log Pearson distance (defined as  $1 - \text{Pearson}$ ) based on the 200-node functional parcellation and the 250 subjects in the holdout data from HCP. Prediction error maps are shown for random walks of length 1, 2, 3 and 10. Each parcel depicts the mean prediction errors across all its edges, with darker colors denoting higher prediction errors. **b)** Network-wise average prediction errors estimated between empirical functional connectivity profiles and the predicted profiles for walks of lengths ranging from 1 to 10. Results are shown for each of the 7 functional Yeo networks based on the 100- and 200-node functional parcellations. **c)** Principal functional gradient derived from the empirical FC in the holdout set of HCP. **d)** Spearman's correlation of the prediction error maps produced from each walk length and the principal functional gradient. Results are shown for both 100- and 200-node functional parcellations. Abbreviations: dorsal attention (DAN), frontoparietal (FPN), default mode (DMN), visual (VN), limbic (LSN), somatomotor (SMN), and ventral attention (VAN) networks.

### Comparison with state-of-the-art methods

We compared the prediction accuracy (average Pearson's correlation coefficient) of our proposed approach to several state-of-the-art methods and to other 2 "reduced" versions of our approach (*SingleLength*, *SharedRot*). Benchmarking results are reported in **Table 1** for the experiments in the HCP dataset, and **Table S1** for MICs. These results correspond to using random walks of length 10. In both tables, we found a clear dichotomy in performance between approaches based on the eigenvectors of the SC and the rest (*i.e.*, SDK and NLSA). The latter achieved the lowest prediction accuracies, with SDK slightly outperforming NLSA in most experimental settings. It is worth noting that the SDK approach only uses one single diffusion kernel. **Table 1** further reports the performances of the MKL and Spectral approaches, which achieved higher performances than SDK and NLSA, with the Spectral approach producing better predictions than MKL in most scenarios (except when using the 100-node structural parcellation in MICS and the

holdout dataset in HCP). Nonetheless, all these approaches were considerably outperformed by our proposed approach in all experimental scenarios, and in both HCP and MICs datasets.

We further assessed the contribution of the different components of our proposed method. As shown in **Table 1** and **Table S1** for HCP and MICs, respectively, the reduced versions of our approach achieved lower prediction accuracies than the original version. When only considering random walks with a single length (*SingleLength*), predictions were more accurate than when using a single rotation shared by all random walks (*SharedRot*). These results underscore the importance of considering random walks of multiple lengths and emphasizes the role of using a different rotation for each length.

Regarding the experimental settings, the functionally-derived parcellation showed to be more beneficial for the prediction of FC than the structural parcellation used in our work. Although the structural parcellation consistently outperformed the functional one with SDK and NLSA, there was a substantial improvement with the functional parcellation over the structural one when considering the best performing approaches (Spectral, *SingleLength*, *SharedRot*, Proposed). Finally, we found a clear drop in prediction accuracy across all the methods we evaluated as the number of parcels increased from 100 to 200, regardless of parcellation type (functional or structural) and dataset.

		Functional		Structural	
		CV	Holdout	CV	Holdout
	<b>SDK</b>	0.238±0.029	0.234±0.027	0.343±0.057	0.347±0.069
	<b>NLSA</b>	0.210±0.037	0.216±0.036	0.351±0.050	0.348±0.061
	<b>MKL</b>	0.713±0.101	0.726±0.106	0.664±0.110	0.705±0.121
<b>100</b>	<b>Spectral</b>	0.779±0.059	0.761±0.082	0.703±0.090	0.669±0.105
	<b>SingleLength</b>	0.757±0.050	0.767±0.051	0.731±0.052	0.748±0.050
	<b>SharedRot</b>	0.733±0.050	0.751±0.048	0.701±0.068	0.728±0.056
	<b>Proposed</b>	<b>0.802±0.054</b>	<b>0.804±0.060</b>	<b>0.751±0.065</b>	<b>0.764±0.054</b>
	<b>SDK</b>	0.241±0.027	0.237±0.025	0.245±0.043	0.243±0.035
	<b>NLSA</b>	0.211±0.039	0.213±0.030	0.241±0.039	0.240±0.043
	<b>MKL</b>	0.702±0.105	0.701±0.099	0.612±0.124	0.627±0.132
<b>200</b>	<b>Spectral</b>	0.741±0.073	0.725±0.081	0.681±0.083	0.650±0.145
	<b>SingleLength</b>	0.727±0.048	0.738±0.044	0.692±0.063	0.700±0.054
	<b>SharedRot</b>	0.701±0.056	0.720±0.047	0.539±0.112	0.571±0.102
	<b>Proposed</b>	<b>0.775±0.049</b>	<b>0.776±0.052</b>	<b>0.715±0.071</b>	<b>0.726±0.055</b>

**Table 1. Comparison of functional connectivity prediction accuracy with state-of-the-art methods.** Performance is reported using the average Pearson’s correlation coefficient (and standard deviation) between the upper triangular

parts of empirical and predicted FC matrices (excluding the main diagonal) using SDK, NLSA, MKL, Spectral and the 3 versions of our proposed approach: *i*) consider random walks of a specific length (SingleLength), *ii*) one single rotation shared by all random walks (SharedRot), and *iii*) one rotation for each length of the random walks (Proposed). Comparisons of prediction accuracy were carried out for 100- and 200-node functional and structural parcellations. Results are reported for a 3-fold cross validation (CV) and in the holdout data in HCP. For NLSA, Spectral and all the versions of our proposed approach, reported performances correspond to using random walks of length 10, and 16 for MKL. For SDK, the diffusion time was chosen to be the one that achieved the best performance in the training set. For findings in the MICs dataset, see Table S1.



## DISCUSSION

This work presented a Riemannian approach to predict functional connectivity (FC) from the underlying structural connectome (SC) at the individual participant level. The proposed approach leveraged the diffusion maps framework to model the exchange of information through the fibers between different brain regions and learn intermediate kernels capturing this information flow. By capitalizing on manifold optimization, we did not only consider the relationships between structural and functional spectra, but also incorporated a mapping (rotation matrices) between the eigenvectors of both domains. With the proposed approach, we were able to find a robust mapping between the structural and functional embeddings, as shown by our results across different datasets, cortical parcellation atlases and parcellation granularities. Furthermore, our approach allowed us to investigate and understand how brain function gradually emanates from structure as information is propagated through increasingly longer (*i.e.*, multiple hops) paths across the structural backbone.

FC does not simply reflect the static and direct wiring of the SC, but it also captures higher-order interactions between potentially only indirectly connected areas (Honey *et al.*, 2009). We hypothesized that by accounting for polysynaptic communication mechanisms, we could explain, to a large extent, the FC observed between pairs of brain regions that lack a direct structural connection. Polysynaptic signaling was modelled by controlling the diffusion time parameter of the diffusion maps framework (Coifman *et al.*, 2005; Coifman & Lafon, 2006). By increasing the diffusion time for a given structural embedding, we were able to generate diffusion coordinates that increasingly captured the interactions between brain regions that were only indirectly connected by structural links. In other words, we were gradually incorporating indirect paths of longer and longer lengths (paths of length 2, 3, 4, and so forth), hence allowing information to flow through increasingly higher-order connections. These interactions were then represented using a radial basis function kernel for each diffusion time. Furthermore, these kernels can be interpreted as intermediate states of the brain while neural information is propagated through the structural fibers. Finally, the different kernels were then fused to provide the predicted FC. Our results showed that, although prediction accuracy increased almost monotonically with increasing diffusion time, there was a clear leveling off around random walks of lengths 3-4, after which only negligible improvements were found. Prior work in the prediction of FC from SC also highlighted shorter structural walks as the strongest contributors to the resulting FC (Becker *et al.*, 2018). This is in line with the economically-optimized configuration of the brain (Bullmore & Sporns, 2012). In addition to minimizing the axonal wiring costs to build the connectome, the expensive metabolic costs spent on information processing also encourage the transfer of information through more economic pathways (Achard & Bullmore, 2007; Avena-Koenigsberger *et al.*, 2014; Bullmore & Sporns, 2012; Laughlin *et al.*, 1998), and hence might be favouring shorter polysynaptic paths.

With diffusion maps, information flow is driven by the transition probabilities of the diffusion operator, which are in turn derived from the edge weights of the original SC matrices. In the current work, we sought to explore the role of different weighting schemes in explaining ongoing brain function: *i)* binary weights only indicating presence or absence of connections, *ii)* weights based on mean streamline length, and *iii)* weights denoting fiber density (conventional scheme). Note that we did not change the connections of the structural graph (*i.e.*, add and/or remove edges), only their weights. These weights play a crucial role in guiding the random walker-based information propagation process through the SC. In the binary weighting scheme, the probability of transitioning from one brain region to another is evenly distributed among all its adjacent regions (*i.e.*, those with a direct structural connection), whereas in the

other weighting schemes, the likelihood of following a particular (direct) connection depends on its length or fiber density. With inadequate weight assignments, we might be misleading the diffusion process and deviating from the true putative flow of information, which can be especially harmful for neural function synchronization of brain regions relying on polysynaptic communication mechanisms. Indeed, this is what we observed with the binary and length-based weighting schemes, which were substantially outperformed when using weights based on fiber density. More importantly, increasing diffusion time showed no consistent contribution to prediction accuracy, and only minor changes were found from one diffusion time to another. Moreover, the length-based weighting scheme attained the lowest prediction accuracy, suggesting that the length of the streamlines might not be the most important indicator of their degree of participation in polysynaptic communication. According to our results, this role appears to be better explained by the density of the streamlines. Regardless of the weighting scheme, our approach relies on diffusion-based communication processes (Coifman *et al.*, 2005; Coifman & Lafon, 2006; Masuda *et al.*, 2017), in which the random walker is only driven by local information. This communication strategy lies on a continuous spectrum of communication processes that ranges from unbiased random walks (such as our proposed approach) at one extreme, through biased random walks that incorporate both local properties and information about the global topology of the structural network, to shortest paths walkers that only consider global information at the other extreme (Avena-Koenigsberger *et al.*, 2014, 2019). Network communication models incorporating both local and global network properties may therefore provide additional information for understanding the correspondence of brain structure and function and hence enhance the predictive power of the proposed approach. Given that signal propagation is strongly influenced by the number of streamlines and also (to a lesser extent) by their lengths (Hermundstad *et al.*, 2013), future work may consider incorporating global information about these properties into the diffusion process to favor transmission of neural signals through shorter and more reliable pathways (Fornito *et al.*, 2016; Goñi *et al.*, 2014).

With the purpose of supporting the dynamic emergence of coherent neural activity patterns, the anatomical substrate of the brain establishes the routing network architecture that facilitates communication between disparate cerebral regions. These structural communication channels have an important say in how brain function is shaped. It is well known that the structure-function relationship is spatially-varying across the cortex, with strong coupling in primary sensory areas that gradually weakens as we move in the sensory-fugal direction (Baum *et al.*, 2020; Suárez *et al.*, 2020; Sydnor *et al.*, 2021; Vázquez-Rodríguez *et al.*, 2019), following a cortical hierarchy of functional and structural organization (Margulies *et al.*, 2016; Paquola *et al.*, 2019; Park *et al.*, 2020). In addition to corroborating these observations, our results showed a tightening of the structure-function coupling in transmodal cortices as diffusion time increased, which, in turn, also translated into a divergence of this coupling from the principal functional gradient. As mentioned above, most of these changes occurred during the first 3 or 4 diffusion timesteps. Although we found some improvements in prediction accuracy with increasing diffusion time in unimodal areas, diffusion time had disproportionately more impact in explaining brain function in transmodal regions. These findings suggest that brain function emerges through polysynaptic communication mechanisms in transmodal cortices, while shorter communication pathways (*e.g.*, monosynaptic) are needed in unimodal regions. In a recent study characterizing the directionality of neural information propagation from undirected structural connectome data (Seguin *et al.*, 2019), unimodal and transmodal cortices were found to be at extremes of the send-receive asymmetry spectrum, with unimodal cortices being more likely to be senders and transmodal cortices more likely to be receivers. This communication asymmetry may therefore account for the larger diffusion times required



by transmodal cortices, which are thought to integrate multiple streams of information originating from the unimodal sensory areas that require fewer diffusion timesteps.

Most work studying structure-function coupling in the brain investigated this relationship based on group-level SC and FC matrices (Goñi *et al.*, 2014; Seguin *et al.*, 2020; Suárez *et al.*, 2020). Predictions at the group level, however, do not take into account inter-subject variability, *e.g.*, using group consensus structural matrices that only consider connections when they are present in at least one-fourth of the participants (Avena-Koenigsberger *et al.*, 2014; Rosenthal *et al.*, 2018). At the individual level, this may translate into less accurate predictions. For example, (Sarwar *et al.*, 2021) using deep learning achieved very high prediction accuracies at the group level ( $r=0.900$ ) that substantially dropped when considering individual predictions ( $r=0.550$ ). In this work, our proposed approach was able to accurately predict brain function at the individual level in both HCP and MICs datasets. Across the different parcellation schemes considered, we observed a consistent increase in prediction accuracy when reducing the spatial scale of the parcellations. The type of information used to create the parcellations also had an important impact on out-of-sample performance, with the functionally-derived parcellations typically producing better predictions than the structurally-defined parcellations. In cross-validation, the highest prediction accuracy ( $r=0.802$ ) was achieved when using the coarsest functional parcellation scheme, whereas the lowest performance ( $r=0.715$ ) was obtained with the most granular (*i.e.*, 200 regions) structural parcellation, both in HCP and MICs. In comparison with the state of the art, our approach attained the best prediction accuracies, regardless of the type of parcellation and dataset. The proposed approach showed solid improvements over the currently best performing method (*i.e.*, Spectral), with percentage increases ranging from 2.95% to 14.20% in HCP (across both cross-validation and holdout datasets), and from 11.77% to 18.15% in the MICs dataset. In addition, the ablation analysis of our method highlighted the importance of including different walk lengths and considering multiple rotations (one rotation per length) for providing accurate prediction. In the MICs dataset, for the 200-node structural parcellation, for example, the mean Pearson's correlation coefficient dropped by 4.63% when only considering a single diffusion time (and hence a single kernel), with an even more substantial drop of 7.30% when using one single rotation shared by all diffusion times. In HCP, prediction accuracy decreased even further with these two reduced versions of our proposed approach. On the other hand, according to our results, the Riemannian regularization (encouraging consecutive diffusion times to use similar rotation matrices) that we initially included in our approach did not show an important contribution to the prediction of FC. This regularization was used to encourage consecutive diffusion times to use similar rotation matrices, however, our results showed that the kernels that contributed the most to the prediction of brain function had very dissimilar rotation matrices, whereas the remaining kernels that barely improved the quality of our predictions had similar rotations. This dichotomy may explain the little contribution of the Riemannian regularization term to the final predictions.

In conclusion, our results show that only a few kernels are necessary to reliably reconstruct functional connectivity from a model of the structural connectome. Moreover, our results underscored that visual, somatomotor and attention networks require generally shorter communication paths than transmodal systems such as the default mode, frontoparietal and limbic networks. The requirement of larger diffusion times in these networks highlights the reliance on more polysynaptic communication mechanisms as we go up the putative cortical hierarchy. Finally, the proposed approach produced highly competitive predictions vis-a-vis current state-of-the-art methods, and this performance improvement was observed across different experimental settings (*i.e.*, across different datasets, parcellation schemes, and

parcellation granularities). Overall, our findings support a likely contribution of polysynaptic signaling in macroscale brain function, especially in transmodal cortices and thus outline potential mechanisms underlying gradients of structure-function coupling in human cortical networks.

## ACKNOWLEDGEMENTS

Oualid Benkarim was funded by a Healthy Brains for Healthy Lives (HBHL) postdoctoral fellowship and the Quebec Autism Research Training (QART) program. Casey Paquola was funded through a postdoctoral fellowship of the Fonds de la Recherche due Quebec - Santé (FRQ-S). Bo-yong Park was funded by the National Research Foundation of Korea (NRF-2021R1F1A1052303), Institute for Information and Communications Technology Planning and Evaluation (IITP) funded by the Korea Government (MSIT) (2020-0-01389, Artificial Intelligence Convergence Research Center, Inha University; 2021-0-02068, Artificial Intelligence Innovation Hub), and Institute for Basic Science (IBS-R015-D1). Boris Bernhardt acknowledges research support from the National Science and Engineering Research Council of Canada (NSERC Discovery-1304413), the Canadian Institutes of Health Research (CIHR FDN-154298, PJT-174995), SickKids Foundation (NI17-039), Azrieli Center for Autism Research (ACAR-TACC), BrainCanada (Azrieli Future Leaders), and the Tier-2 Canada Research Chairs program. Jessica Royer was funded by a CIHR fellowship. Reinder Vos de Wael was funded by a studentship from the Savoy Foundation. We would also like to acknowledge support from the Helmholtz Foundation and the Healthy Brains for Healthy Lives initiative.

## AUTHOR CONTRIBUTIONS

O.B. and B.C.B. designed the experiments, analyzed the data, and wrote the manuscript. C.P., and B.P. aided with the experiments. J.R., R.R-C., R.V., B.M. and G.P. reviewed the manuscript. O.B. and B.C.B. are the corresponding authors of this work and have responsibility for the integrity of the data analysis.

## COMPETING INTERESTS

The authors declare no conflicts of interest.

## CODE AND DATA AVAILABILITY

The code for the proposed framework is publicly available at [https://github.com/MICA-MNI/micaopen/sf\\_prediction](https://github.com/MICA-MNI/micaopen/sf_prediction). The HCP and MICs dMRI and rs-fMRI data are publicly available at <https://www.humanconnectome.org/> and <https://portal.conp.ca/dataset?id=projects/mica-mics>.

## REFERENCES

- Abdelnour, F., Dayan, M., Devinsky, O., Thesen, T., & Raj, A. (2018). Functional brain connectivity is predictable from anatomic network's Laplacian eigen-structure. *NeuroImage*, *172*, 728–739. <https://doi.org/10.1016/j.neuroimage.2018.02.016>
- Abdelnour, F., Voss, H. U., & Raj, A. (2014). Network diffusion accurately models the relationship between structural and functional brain connectivity networks. *NeuroImage*, *90*, 335–347. <https://doi.org/10.1016/j.neuroimage.2013.12.039>
- Absil, P.-A., Mahony, R., & Sepulchre, R. (2009). *Optimization algorithms on matrix manifolds*. Princeton University Press.
- Achard, S., & Bullmore, E. (2007). Efficiency and Cost of Economical Brain Functional Networks. *PLoS Computational Biology*, *3*(2), e17. <https://doi.org/10.1371/journal.pcbi.0030017>
- Atasoy, S., Donnelly, I., & Pearson, J. (2016). Human brain networks function in connectome-specific harmonic waves. *Nature Communications*, *7*(1), 10340. <https://doi.org/10.1038/ncomms10340>
- Avena-Koenigsberger, A., Goñi, J., Betzel, R. F., van den Heuvel, M. P., Griffa, A., Hagmann, P., Thiran, J.-P., & Sporns, O. (2014). Using Pareto optimality to explore the topology and dynamics of the human connectome. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1653), 20130530. <https://doi.org/10.1098/rstb.2013.0530>
- Avena-Koenigsberger, A., Misić, B., & Sporns, O. (2018). Communication dynamics in complex brain networks. *Nature Reviews Neuroscience*, *19*(1), 17–33.
- Avena-Koenigsberger, A., Yan, X., Kolchinsky, A., Heuvel, M. P. van den, Hagmann, P., & Sporns, O. (2019). A spectrum of routing strategies for brain networks. *PLOS Computational Biology*, *15*(3), e1006833. <https://doi.org/10.1371/journal.pcbi.1006833>
- Bassett, D. S., Khambhati, A. N., & Grafton, S. T. (2017). Emerging frontiers of neuroengineering: A network science of brain connectivity. *Annual Review of Biomedical Engineering*, *19*, 327–352.
- Baum, G. L., Cui, Z., Roalf, D. R., Ciric, R., Betzel, R. F., Larsen, B., Cieslak, M., Cook, P. A., Xia, C. H., Moore, T. M., Ruparel, K., Oathes, D. J., Alexander-Bloch, A. F., Shinohara, R. T., Raznahan, A., Gur, R. E., Gur, R. C., Bassett, D. S., & Satterthwaite, T. D. (2020). Development of structure–function coupling in human brain networks during youth. *Proceedings of the National Academy of Sciences*, *117*(1), 771–778. <https://doi.org/10.1073/pnas.1912034117>
- Bazinet, V., Vos de Wael, R., Hagmann, P., Bernhardt, B. C., & Misić, B. (2021). Multiscale communication in cortico-cortical networks. *NeuroImage*, *243*, 118546. <https://doi.org/10.1016/j.neuroimage.2021.118546>
- Becker, C. O., Pequito, S., Pappas, G. J., Miller, M. B., Grafton, S. T., Bassett, D. S., & Preciado, V. M. (2018). Spectral mapping of brain functional connectivity from diffusion imaging. *Scientific Reports*, *8*(1), 1–15.
- Boumal, N., Singer, A., Absil, P.-A., & Blondel, V. D. (2014). Cramér–Rao bounds for synchronization of rotations. *Information and Inference: A Journal of the IMA*, *3*(1), 1–39.
- Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nature Neuroscience*, *20*(3), 340–352. <https://doi.org/10.1038/nn.4497>
- Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences*, *14*(6), 277–290. <https://doi.org/10.1016/j.tics.2010.04.004>
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186–198.
- Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*(5), 336–349. <https://doi.org/10.1038/nrn3214>
- Christiaens, D., Reisert, M., Dhollander, T., Sunaert, S., Suetens, P., & Maes, F. (2015). Global tractography of multi-shell diffusion-weighted imaging data using a multi-tissue model. *NeuroImage*, *123*, 89–101. <https://doi.org/10.1016/j.neuroimage.2015.08.008>
- Coifman, R. R., & Hirn, M. J. (2014). Diffusion maps for changing data. *Applied and Computational Harmonic Analysis*, *36*, 79–107.

- Coifman, R. R., & Lafon, S. (2006). Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1), 5–30.
- Coifman, R. R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F., & Zucker, S. W. (2005). Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the National Academy of Sciences*, 102(21), 7426–7431. <https://doi.org/10.1073/pnas.0500334102>
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, an International Journal*, 29(3), 162–173. <https://doi.org/10.1006/cbmr.1996.0014>
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage*, 9(2), 179–194.
- Damoiseaux, J. S., & 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. <https://doi.org/10.1007/s00429-009-0208-6>
- Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G. L., Hagmann, P., & Corbetta, M. (2013). Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *Journal of Neuroscience*, 33(27), 11239–11252.
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3), 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021>
- Fischl, B. (2012). FreeSurfer. *NeuroImage*, 62(2), 774–781.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999a). Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2), 195–207.
- Fischl, B., Sereno, M. I., Tootell, R. B. H., & Dale, A. M. (1999b). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8(4), 272–284. [https://doi.org/10.1002/\(SICI\)1097-0193\(1999\)8:4<272::AID-HBM10>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<272::AID-HBM10>3.0.CO;2-4)
- Fornito, A., Zalesky, A., & Breakspear, M. (2013). Graph analysis of the human connectome: Promise, progress, and pitfalls. *Mapping the Connectome*, 80, 426–444. <https://doi.org/10.1016/j.neuroimage.2013.04.087>
- Fornito, A., Zalesky, A., & Bullmore, E. T. (2016). *Fundamentals of Brain Network Analysis*. Academic Press. <https://doi.org/10.1016/B978-0-12-407908-3.09999-4>
- Gabay, N. C., Babaie-Janvier, T., & Robinson, P. A. (2018). Dynamics of cortical activity eigenmodes including standing, traveling, and rotating waves. *Physical Review E*, 98(4), 042413. <https://doi.org/10.1103/PhysRevE.98.042413>
- Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., & Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615), 171–178. <https://doi.org/10.1038/nature18933>
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., & Polimeni, J. R. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage*, 80, 105–124.
- 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., & 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.
- Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1), 63–72. <https://doi.org/10.1016/j.neuroimage.2009.06.060>
- Hagmann, P., Kurant, M., Gigandet, X., Thiran, P., Wedeen, V. J., Meuli, R., & Thiran, J.-P. (2007). Mapping Human Whole-Brain Structural Networks with Diffusion MRI. *PLoS ONE*, 2(7), e597. <https://doi.org/10.1371/journal.pone.0000597>
- Hermundstad, A. M., Bassett, D. S., Brown, K. S., Aminoff, E. M., Clewett, D., Freeman, S., Frithsen, A., Johnson, A., Tipper, C. M., Miller, M. B., Grafton, S. T., & Carlson, J. M. (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. <https://doi.org/10.1073/pnas.1219562110>
- Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.-P., Meuli, R., & Hagmann, P. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences*, *106*(6), 2035–2040.
- Hu, J., Liu, X., Wen, Z.-W., & Yuan, Y.-X. (2020). A Brief Introduction to Manifold Optimization. *Journal of the Operations Research Society of China*, *8*(2), 199–248. <https://doi.org/10.1007/s40305-020-00295-9>
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. <https://doi.org/10.1016/j.neuroimage.2011.09.015>
- Jeurissen, B., Tournier, J.-D., Dhollander, T., Connelly, A., & Sijbers, J. (2014). Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage*, *103*, 411–426. <https://doi.org/10.1016/j.neuroimage.2014.07.061>
- Laughlin, S. B., de Ruyter van Steveninck, R. R., & Anderson, J. C. (1998). The metabolic cost of neural information. *Nature Neuroscience*, *1*(1), 36–41. <https://doi.org/10.1038/236>
- Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., Bezgin, G., Eickhoff, S. B., Castellanos, F. X., & Petrides, M. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences*, *113*(44), 12574–12579.
- Masuda, N., Porter, M. A., & Lambiotte, R. (2017). Random walks and diffusion on networks. *Physics Reports*, *716–717*, 1–58. <https://doi.org/10.1016/j.physrep.2017.07.007>
- Meier, J., Tewarie, P., Hillebrand, A., Douw, L., van Dijk, B. W., Stufflebeam, S. M., & Van Mieghem, P. (2016). A mapping between structural and functional brain networks. *Brain Connectivity*, *6*(4), 298–311.
- Mišić, B., Betzel, R. F., De Reus, M. A., Van Den Heuvel, M. P., Berman, M. G., McIntosh, A. R., & Sporns, O. (2016). Network-level structure-function relationships in human neocortex. *Cerebral Cortex*, *26*(7), 3285–3296.
- Mišić, B., Betzel, R. F., Nematzadeh, A., Goñi, J., Griffa, A., Hagmann, P., Flammini, A., Ahn, Y.-Y., & Sporns, O. (2015). Cooperative and Competitive Spreading Dynamics on the Human Connectome. *Neuron*, *86*(6), 1518–1529. <https://doi.org/10.1016/j.neuron.2015.05.035>
- Paquola, C., Vos De Wael, R., Wagstyl, K., Bethlehem, R. A. I., Hong, S.-J., Seidlitz, J., Bullmore, E. T., Evans, A. C., Misić, B., Margulies, D. S., Smallwood, J., & Bernhardt, B. C. (2019). Microstructural and functional gradients are increasingly dissociated in transmodal cortices. *PLOS Biology*, *17*(5), 1–28.
- Park, B., Vos de Wael, R., Paquola, C., Larivière, S., Benkarim, O., Royer, J., Tavakol, S., Cruces, R. R., Li, Q., Valk, S. L., Margulies, D. S., Misić, B., Bzdok, D., Smallwood, J., & Bernhardt, B. C. (2020). Signal diffusion along connectome gradients and inter-hub routing differentially contribute to dynamic human brain function. *Neuroimage*, *224*, 117429.
- Robinson, E. C., Jbabdi, S., Glasser, M. F., Andersson, J., Burgess, G. C., Harms, M. P., Smith, S. M., Van Essen, D. C., & Jenkinson, M. (2014). MSM: A new flexible framework for Multimodal Surface Matching. *NeuroImage*, *100*, 414–426. <https://doi.org/10.1016/j.neuroimage.2014.05.069>
- Robinson, P. A. (2012). Interrelating anatomical, effective, and functional brain connectivity using propagators and neural field theory. *Physical Review E*, *85*(1), 011912.
- Rosenthal, G., Váša, F., Griffa, A., Hagmann, P., Amico, E., Goñi, J., Avidan, G., & Sporns, O. (2018). Mapping higher-order relations between brain structure and function with embedded vector representations of connectomes. *Nature Communications*, *9*(1), 2178. <https://doi.org/10.1038/s41467-018-04614-w>
- Royer, J., Rodríguez-Cruces, R., Tavakol, S., Larivière, S., Herholz, P., Li, Q., Wael, R. V. de, Paquola, C., Benkarim, O., Park, B., Lowe, A. J., Margulies, D., Smallwood, J., Bernasconi, A., Bernasconi, N., Frauscher, B., & Bernhardt, B. C. (2021). *An Open MRI Dataset for Multiscale Neuroscience* (p. 2021.08.04.454795). <https://doi.org/10.1101/2021.08.04.454795>
- Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of

- classifiers. *NeuroImage*, *90*, 449–468.
- Sarwar, T., Tian, Y., Yeo, B. T. T., Ramamohanarao, K., & Zalesky, A. (2021). Structure-function coupling in the human connectome: A machine learning approach. *NeuroImage*, *226*, 117609. <https://doi.org/10.1016/j.neuroimage.2020.117609>
- Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral Cortex*, *28*(9), 3095–3114.
- Seguin, C., Razi, A., & Zalesky, A. (2019). Inferring neural signalling directionality from undirected structural connectomes. *Nature Communications*, *10*(1), 4289. <https://doi.org/10.1038/s41467-019-12201-w>
- Seguin, C., Tian, Y., & Zalesky, A. (2020). Network communication models improve the behavioral and functional predictive utility of the human structural connectome. *Network Neuroscience*, *4*(4), 980–1006. [https://doi.org/10.1162/netn\\_a\\_00161](https://doi.org/10.1162/netn_a_00161)
- Shen, K., Bezdin, G., Hutchison, R. M., Gati, J. S., Menon, R. S., Everling, S., & McIntosh, A. R. (2012). Information Processing Architecture of Functionally Defined Clusters in the Macaque Cortex. *The Journal of Neuroscience*, *32*(48), 17465–17476. <https://doi.org/10.1523/JNEUROSCI.2709-12.2012>
- Skudlarski, P., Jagannathan, K., Calhoun, V. D., Hampson, M., Skudlarska, B. A., & Pearlson, G. (2008). Measuring Brain Connectivity: Diffusion Tensor Imaging Validates Resting State Temporal Correlations. *NeuroImage*, *43*(3), 554–561. <https://doi.org/10.1016/j.neuroimage.2008.07.063>
- Smith, R. E., Tournier, J.-D., Calamante, F., & Connelly, A. (2012). Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage*, *62*(3), 1924–1938. <https://doi.org/10.1016/j.neuroimage.2012.06.005>
- Smith, R. E., Tournier, J.-D., Calamante, F., & Connelly, A. (2015). SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. *NeuroImage*, *119*, 338–351. <https://doi.org/10.1016/j.neuroimage.2015.06.092>
- Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology*, *1*(4), e42.
- Suárez, L. E., Markello, R. D., Betzel, R. F., & Misic, B. (2020). Linking Structure and Function in Macroscale Brain Networks. *Trends in Cognitive Sciences*, *24*(4), 302–315. <https://doi.org/10.1016/j.tics.2020.01.008>
- Surampudi, S. G., Naik, S., Surampudi, R. B., Jirsa, V. K., Sharma, A., & Roy, D. (2018). Multiple kernel learning model for relating structural and functional connectivity in the brain. *Scientific Reports*, *8*(1), 1–14.
- Sydnor, V. J., Larsen, B., Bassett, D. S., Alexander-Bloch, A., Fair, D. A., Liston, C., Mackey, A. P., Milham, M. P., Pines, A., Roalf, D. R., Seidlitz, J., Xu, T., Raznahan, A., & Satterthwaite, T. D. (2021). Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron*, *109*(18), 2820–2846. <https://doi.org/10.1016/j.neuron.2021.06.016>
- Tewarie, P., Prasse, B., Meier, J. M., Santos, F. A., Douw, L., Schoonheim, M. M., Stam, C. J., Van Mieghem, P., & Hillebrand, A. (2020). Mapping functional brain networks from the structural connectome: Relating the series expansion and eigenmode approaches. *NeuroImage*, *216*, 116805.
- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*(1), 267–288.
- Tournier, J.-D., Calamante, F., & Connelly, A. (2012). MRtrix: Diffusion tractography in crossing fiber regions. *International Journal of Imaging Systems and Technology*, *22*(1), 53–66. <https://doi.org/10.1002/ima.22005>
- Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H., & Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, *202*, 116137. <https://doi.org/10.1016/j.neuroimage.2019.116137>
- Townsend, J., Koep, N., & Weichwald, S. (2016). Pymanopt: A python toolbox for optimization on manifolds using automatic differentiation. *ArXiv Preprint ArXiv:1603.03236*.
- Tustison, N. J., & Avants, B. B. (2013). Explicit B-spline regularization in diffeomorphic image registration. *Frontiers in Neuroinformatics*, *7*, 39. <https://doi.org/10.3389/fninf.2013.00039>
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK:



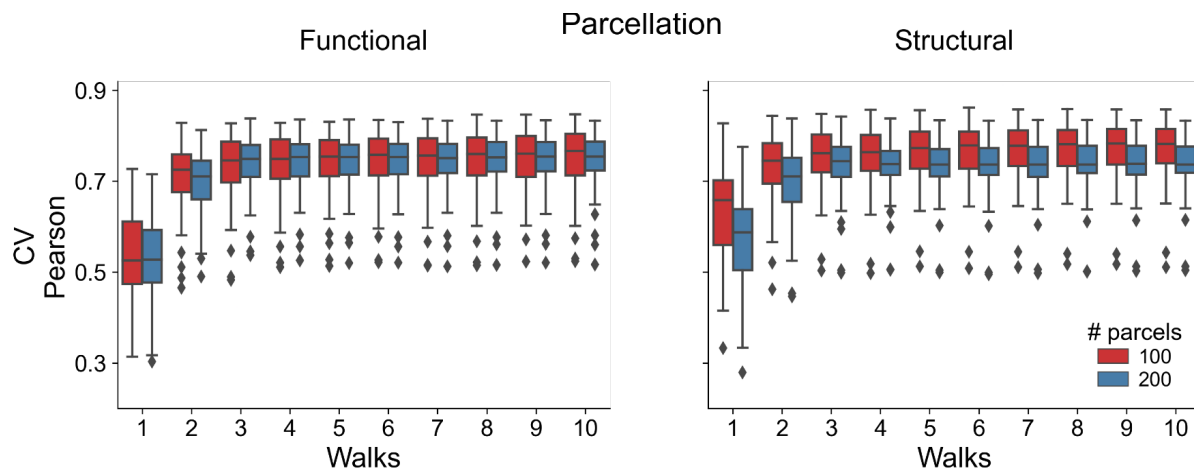
- Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6), 1310–1320.  
<https://doi.org/10.1109/tmi.2010.2046908>
- Van Essen, D. C., Glasser, M. F., Dierker, D. L., Harwell, J., & Coalson, T. (2012). Parcellations and Hemispheric Asymmetries of Human Cerebral Cortex Analyzed on Surface-Based Atlases. *Cerebral Cortex*, 22(10), 2241–2262. <https://doi.org/10.1093/cercor/bhr291>
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & Consortium, W.-M. H. (2013). The WU-Minn human connectome project: An overview. *Neuroimage*, 80, 62–79.
- Vázquez-Rodríguez, B., Suárez, L. E., Markello, R. D., Shafiei, G., Paquola, C., Hagmann, P., Heuvel, M. P. van den, Bernhardt, B. C., Spreng, R. N., & Misisic, B. (2019). Gradients of structure–function tethering across neocortex. *Proceedings of the National Academy of Sciences*, 116(42), 21219–21227.  
<https://doi.org/10.1073/pnas.1903403116>
- Vos de Wael, R., Benkarim, O., Paquola, C., Larivière, S., Royer, J., Tavakol, S., Xu, T., Hong, S. J., Langa, G., Valk, S., Misisic, B., Milham, M., Margulies, D., Smallwood, J., & Bernhardt, B. C. (2020). BrainSpace: A toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. *Communications Biology*, 3.
- Wang, M. B., Owen, J. P., Mukherjee, P., & Raj, A. (2017). Brain network eigenmodes provide a robust and compact representation of the structural connectome in health and disease. *PLOS Computational Biology*, 13(6), e1005550. <https://doi.org/10.1371/journal.pcbi.1005550>
- Wang, P., Kong, R., Kong, X., Liégeois, R., Orban, C., Deco, G., Heuvel, M. P. van den, & Yeo, B. T. T. (2019). Inversion of a large-scale circuit model reveals a cortical hierarchy in the dynamic resting human brain. *Science Advances*. <https://www.science.org/doi/abs/10.1126/sciadv.aat7854>
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., & Polimeni, J. R. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106, 1125–1165.

## A Riemannian approach to predicting brain function from the structural connectome

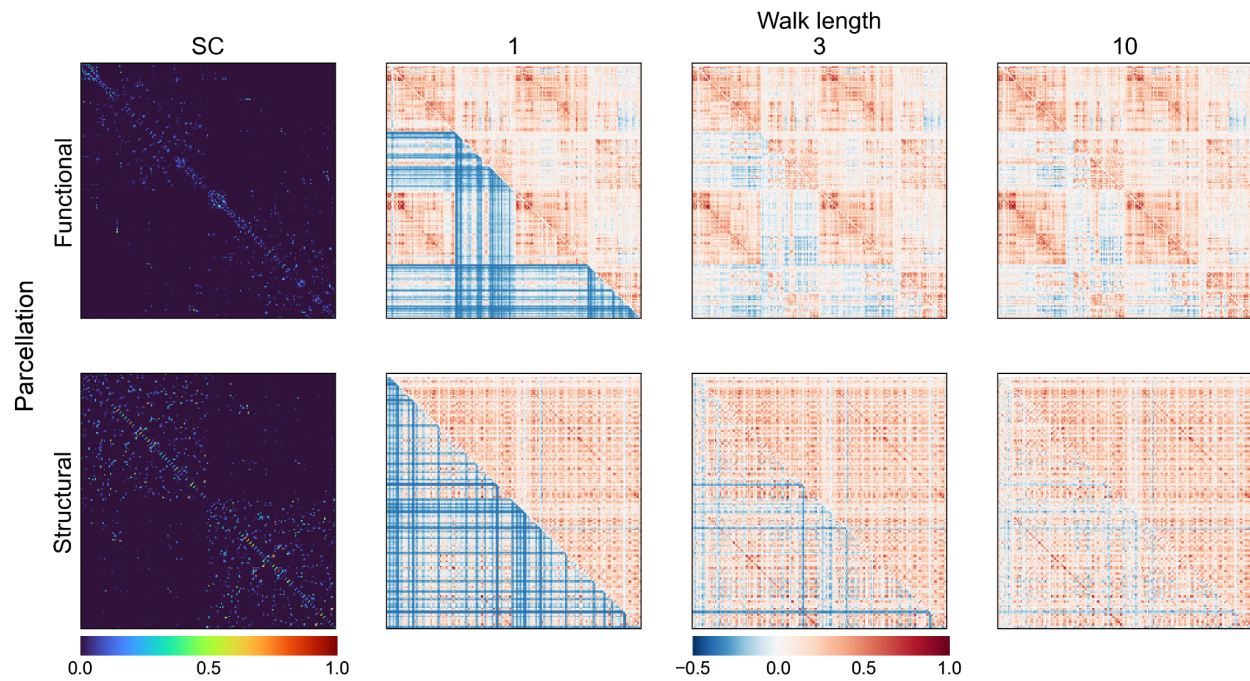
### (Supplementary material)

Oualid Benkarim<sup>1\*</sup>, Casey Paquola<sup>1,2</sup>, Bo-yong Park<sup>1,3,4</sup>, Jessica Royer<sup>1</sup>, Raúl Rodríguez-Cruces<sup>1</sup>, Reinder Vos de Wael<sup>1</sup>, Bratislav Misic<sup>1</sup>, Gemma Piella<sup>5</sup>, Boris C. Bernhardt<sup>1\*</sup>

<sup>1</sup>McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada; <sup>2</sup>Institute of Neuroscience and Medicine (INM-1), Forschungszentrum Jülich, Jülich, Germany; <sup>3</sup>Department of Data Science, Inha University, South Korea; <sup>4</sup>Center for Neuroscience Imaging Research, Institute for Basic Science, South Korea; <sup>5</sup>BCN MedTech, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain



**Figure S1.** MICs dataset. Boxplots of correlation between empirical and estimated FC matrices using 1 to 10 random walks. Results are reported in cross-validation (top) and holdout (bottom) using functionally- (left) and structurally-derived cortical parcellation (right) with 100 and 200 parcels.



**Figure S2.** Functional connectivity prediction using random walks of different lengths in the structural connectome in the HCP dataset. Estimated functional connectivity matrices corresponding to the subject that achieved the median Pearson's correlation coefficient based on functional (top) and structural (bottom) parcellations with 200 parcels. From left to right: structural connectome and estimated FC matrices for random walks of length 1, 3 and 10. Empirical and estimated functional connectivity matrices are shown in upper and lower triangular parts, respectively.

	<b>Functional</b>	<b>Structural</b>
<b>SDK</b>	0.320±0.045	0.427±0.067
<b>NLSA</b>	0.299±0.043	0.416±0.065
<b>MKL</b>	0.632±0.088	0.678±0.090
<b>100 Spectral</b>	0.671±0.085	0.650±0.081
<b>SingleLength</b>	0.723±0.073	0.741±0.075
<b>SharedRot</b>	0.711±0.070	0.734±0.070
<b>Proposed</b>	<b>0.750±0.075</b>	<b>0.768±0.072</b>
<b>SDK</b>	0.317±0.042	0.338±0.048
<b>NLSA</b>	0.303±0.042	0.326±0.059
<b>MKL</b>	0.601±0.081	0.614±0.090
<b>200 Spectral</b>	0.655±0.076	0.624±0.079
<b>SingleLength</b>	0.727±0.067	0.706±0.065
<b>SharedRot</b>	0.693±0.069	0.685±0.075
<b>Proposed</b>	<b>0.742±0.069</b>	<b>0.735±0.067</b>

**Table S1.** Mean correlation and standard deviation using SDK, NLSA, MKL, Spectral and the 3 versions of our proposed approach. Performance is reported for 100 and 200 parcels using 3-fold cross validation (CV) in HCP and MICS datasets, and holdout in HCP. For Spectral and our approach, the performance reported corresponds to using 10 random walks, and 16 for MKL.