



Comparing models of information transfer in the structural brain network and their relationship to functional connectivity: diffusion versus shortest path routing

Josh Neudorf¹ · Shaylyn Kress¹ · Ron Borowsky¹

Received: 6 July 2022 / Accepted: 16 January 2023 / Published online: 1 February 2023
© The Author(s) 2023

Abstract

The relationship between structural and functional connectivity in the human brain is a core question in network neuroscience, and a topic of paramount importance to our ability to meaningfully describe and predict functional outcomes. Graph theory has been used to produce measures based on the structural connectivity network that are related to functional connectivity. These measures are commonly based on either the shortest path routing model or the diffusion model, which carry distinct assumptions about how information is transferred through the network. Unlike shortest path routing, which assumes the most efficient path is always known, the diffusion model makes no such assumption, and lets information diffuse in parallel based on the number of connections to other regions. Past research has also developed hybrid measures that use concepts from both models, which have better predicted functional connectivity from structural connectivity than the shortest path length alone. We examined the extent to which each of these models can account for the structure–function relationship of interest using graph theory measures that are exclusively based on each model. This analysis was performed on multiple parcellations of the Human Connectome Project using multiple approaches, which all converged on the same finding. We found that the diffusion model accounts for much more variance in functional connectivity than the shortest path routing model, suggesting that the diffusion model is better suited to describing the structure–function relationship in the human brain at the macroscale.

Keywords Structure–function relationship · DTI · rsfMRI · Graph theory · Diffusion model · Shortest path routing model

Introduction

Graph theory analyses of structural brain connectivity have been vital to providing breakthroughs in our understanding of how the underlying structure of the brain can influence the patterns of coordinated functional activity (see Avena-Koenigsberger et al. 2018 for a review; see also Goñi et al. 2014; Neudorf et al. 2020a, b, 2022). Defining this relationship between structural and functional connectivity using advanced techniques including graph theory has recently been highlighted as an important frontier in neuroscience (Suárez et al. 2020). When it comes to choosing graph

theory measures of connectivity, important assumptions must be made about how information is transferred through the structural network, and the effectiveness of these measures for predicting functional connectivity is dependent on the accuracy of these assumptions about the human brain. Two primary graph theory models of information transfer in the brain include shortest path routing and diffusion. The shortest path routing model relies on the calculation of the shortest path to the destination region. This model is straightforward to calculate and underlies many useful graph theory measures that have been helpful in describing brain networks and networks in general (e.g., characteristic path length, Watts and Strogatz 1998; global efficiency as an indicator of small-worldness, Latora and Marchiori 2001; nodal and local efficiency, Latora and Marchiori 2001; van den Heuvel and Sporns 2013; etc.). One problem with the shortest path routing model when it comes to brain networks is that it assumes each region has whole-brain level knowledge about the most efficient path to use (Avena-Koenigsberger

✉ Ron Borowsky
ron.borowsky@usask.ca

¹ Cognitive Neuroscience Lab, Department of Psychology and Health Studies, University of Saskatchewan, 9 Campus Dr., Saskatoon, SK S7N 5A5, Canada

et al. 2019; Seguin et al. 2018, 2022; Zamani Esfahlani et al. 2022).

An alternative graph theory model has been proposed that does not assume whole-brain knowledge about the shortest path but instead assumes that information diffuses along random paths in the network influenced by the relative weighting of each path. Under these model assumptions, information propagates through the network as a “random walker” that is constrained by the structural architecture. Furthermore, information can be transferred in parallel, whereas shortest path routing describes information traveling along a single path to the destination (Fornito et al. 2016a).

Novel graph theory metrics combining both diffusion and shortest path routing models have been developed for use in brain research and applied to the task of predicting functional connectivity from the underlying structural connectivity (Goñi et al. 2014). *Search information* was developed as a measure of how many distractor paths may lead a random walker away from the shortest path, while *path transitivity* measures how likely a random walker on a detour will end up back on the shortest path. While these measures were more successful than the shortest path length alone at predicting functional connectivity from structural connectivity, there are other graph theory measures of connectivity that consider the full range of possible paths based on a diffusion model, rather than hinging on what happens around the shortest path during information transfer. One such measure of diffusion efficiency is the *mean first passage time* (Wang and Pei 2008), which calculates the number of steps it takes a random walker on average to travel from region A to region B. This measure has been used to show that biological brain networks typically display a balance between diffusion efficiency and global efficiency (sensitive to shortest path length; Goñi et al. 2013). Another measure relying on the diffusion model of information transfer is *communicability* (Estrada and Hatano 2008), which takes into consideration all possible walks from region A to region B. Walks with less edges, n , are weighted much higher than those with more, with walks weighted by the factor $1/n!$. *Communicability* is described as reflecting the capacity for a network to transfer information in parallel assuming a diffusion model of information transfer (Fornito et al. 2016a, b). This measure has been useful in distinguishing patients from controls, including stroke (Crofts et al. 2011) and multiple sclerosis (Li et al. 2013).

Considering past success with the hybrid measures combining the diffusion and shortest path routing models of information transfer (Goñi et al. 2014), this research will apply the exclusively diffusion-based measures of mean first passage time and communicability as well as the shortest path routing measure of shortest path length to structural connectivity, and the results will be used to predict functional connectivity to determine to what extent these

measures are able to account for variance in functional connectivity. Crucially, this research will extend past research that has examined the ability of multiple graph theory communication measures to predict functional connectivity from structural connectivity (Betzel et al. 2022; Vázquez-Rodríguez et al. 2019; Zamani Esfahlani et al. 2022) and benchmarking the ability for different communication measures to predict functional connectivity (Seguin et al. 2018, 2020, 2022), by directly comparing two commonly used models (diffusion and shortest path routing) using multiple linear regression analyses, partial least squares regression, and principal components analysis to determine which graph theory model is most important in this relationship. Research suggests that brain networks (at both the macroscale and microscale) typically demonstrate a balance of diffusion efficiency and global efficiency (Goñi et al. 2013), while also suggesting that this balance may lean more towards dominance of diffusion efficiency in human brains, in which case we expect that the diffusion measures examined here will be more relevant than shortest path length to the structure–function relationship in the brain.

Methods

Dataset

MRI data for 998 subjects from the Human Connectome Project (HCP; Van Essen et al. 2013) were used including diffusion tensor imaging (DTI) and resting state functional magnetic resonance imaging (rsfMRI). We used the preprocessed version of the rsfMRI data. This data has been preprocessed using FSL FIX (Salimi-Khorshidi et al. 2014). The DTI data used was also preprocessed. The HCP pipelines for preprocessing are described by Glasser et al. (2013). The Automated Anatomical Labelling 90 region atlas (AAL; Tzourio-Mazoyer et al. 2002) was used as well as the Brainnetome 246 region atlas (Fan et al. 2016). Activation at each rsfMRI acquisition was used to calculate the mean activation for the atlas regions. The rsfMRI sessions were standardized using a z -score for the regions for each session separately. The activation in these regions was then submitted to bandpass filtering (separately for each session) allowing only frequencies within 0.01 Hz and 0.1 Hz (see Hallquist et al. 2013).

Connectivity measures

To calculate the functional connectivity measures for each combination of regions, we calculated the Pearson correlation coefficient using all of the 4800 acquisitions. To calculate the structural connectivity measures, DSI Studio (<http://dsi-studio.labsolver.org>) was used with quantitative

anisotropy (Yeh et al. 2013) as the termination index to calculate the streamline count. Generalized q -sampling (Yeh et al. 2010) was used, and tracking used 1 million fibers, 75° maximum angular deviation, and a 20 mm minimum and 500 mm maximum fiber length. To calculate the structural connectivity matrix containing the number of streamlines for each cell, a whole brain seed was used. The connectivity values for structural and functional connectivity were averaged using the mean for all subjects. The weighted structural connectivity density (the sum of connection weights divided by the total possible connection weights, where each weight has a maximum of 1.0) was 0.016 for the AAL atlas and 0.004 for the Brainnetome atlas.

Graph theory structural connectivity measures of mean first passage time (Wang and Pei 2008) and communicability (Estrada and Hatano 2008) were calculated as diffusion model measures (also discussed in Fornito et al. 2016a, b). Mean first passage time was calculated as

$$MFPT_{ij} = \sum_{n=1}^N \left[(I - U_j)^{-1} \right]_{ni}, \quad (1)$$

where I is the identity matrix, i is the starting node, j is the destination node, and N is the number of regions in the network, and

$$U_j = WS^{-1}, \quad (2)$$

but with the j th row set to zero so that a random walker is unable to enter j . W is the weighted structural connectivity adjacency matrix, and

$$S = \begin{bmatrix} s_1 & 0 & 0 \\ 0 & \ddots & \vdots \\ 0 & \dots & s_N \end{bmatrix}, \quad (3)$$

with s_n representing the strength (weighted number of connections) of region n . Communicability was calculated as

$$Com_{ij} = \left[e^{S^{-1/2}WS^{-1/2}} \right]_{ij}, \quad (4)$$

where $e^{S^{-1/2}WS^{-1/2}}$ is the matrix exponential of $S^{-1/2}WS^{-1/2}$, the reduced structural connectivity adjacency matrix (see Crofts et al. 2011). Long walks are weighted more weakly (by a factor of $n!$ where n is the number of steps) in this formula as the series expansion equates to

$$Com_{ij} = \sum_{n=0}^{\infty} \frac{[(S^{-1/2}WS^{-1/2})^n]_{ij}}{n!}. \quad (5)$$

Shortest path length was calculated as the shortest path routing model measure using the NetworkX python library (Hagberg et al. 2008; function `shortest_path_length`, using the Dijkstra algorithm described by Dijkstra 1959, and given

the inverse value of structural connectivity edges so that the edges represent resistance in the network).

Permutation testing was performed using 500 null models created from the structural connectivity following the generalized Maslov–Sneppen (Maslov and Sneppen 2002) rewiring algorithm developed by Rubinov and Sporns (2011) for use with weighted networks to control for node strength and degree while randomizing the connection weights. Permutation p -values were calculated as the number of null models resulting in the same or better variance accounted for in the models as a proportion of the total number of null models.

Results

AAL

Linear regression models

Linear regression models were computed for each log-transformed independent variable (mean first passage time, communicability, and shortest path length) with functional connectivity as the dependent variable using the `lm` function from the `lme4` library (Bates et al. 2015) in R (R Core Team 2018). Mean first passage time demonstrated an inverse relationship with functional connectivity, whereby a high mean first passage time was associated with poorer functional connectivity, as expected, $R(4003) = -0.376$, $p < 0.001$ (null model permutation $p = 0.044$; see Fig. 1A). Communicability demonstrated a positive relationship with functional connectivity, whereby high communicability was associated with better functional connectivity, as expected, $R(4003) = 0.316$, $p < 0.001$ (null model permutation $p = 0.014$; see Fig. 1B). Shortest path length demonstrated an inverse relationship with functional connectivity, whereby a high shortest path length was associated with poorer functional connectivity, as expected, $R(4003) = -0.379$, $p < 0.001$ (null model permutation $p < 0.002$; see Fig. 1C). The magnitude of the structure–function relationship for each of these measures was relatively comparable, so additional multiple linear regression approaches were also taken to determine which measures are primarily driving the relationship between structural and functional connectivity.

Multiple linear regression models

Multiple linear regression models were then investigated starting with mean first passage time and shortest path length included in the model as independent variables, with functional connectivity as the dependent variable. These models were again calculated in R using the `lm` function from the `lme4` library, as well as `spcor` from the `ppcor` library to calculate the semi-partial correlation (Kim 2015) and `vif`

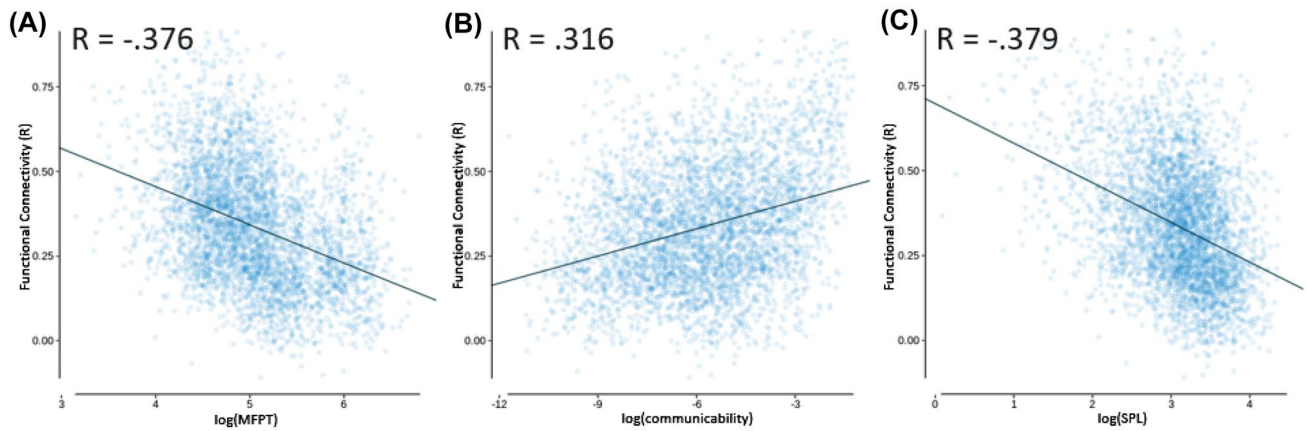


Fig. 1 AAL linear regression models with functional connectivity as the independent variable and dependent variables of: **A** mean first passage time (MFPT; log transformed), $R_{\text{adj}} = -0.376$; **B** communi-

cability (COM; log transformed), $R_{\text{adj}} = 0.316$; **C** and shortest path length (SPL; log transformed), $R_{\text{adj}} = -0.379$

Table 1 AAL independent variable correlation matrix

| Variable | MFPT | COM | SPL |
|----------|--------|--------|--------|
| MFPT | 1 | -0.236 | 0.841 |
| COM | -0.236 | 1 | -0.491 |
| SPL | 0.841 | -0.491 | 1 |

Variables include mean first passage time (MFPT), communicability (COM), and shortest path length (SPL)

from the *car* library to calculate the variance inflation factor (Fox and Weisberg 2019). Prior to this, the correlation matrix of these measures was examined, which indicated that there were no extreme correlations (e.g., greater than 0.9) between the independent variables, with the highest value being $R = 0.841$ between mean first passage time and shortest path length (see Table 1). This correlation is theoretically interesting though, as it indicates there is a high level of redundancy between mean first passage time and shortest path length, suggesting that information in the diffusion model naturally follows paths that are similarly efficient when compared to the shortest path. This potential for decentralized information transfer strategies to take advantage of the shortest paths in the network has been noted in past research (Avena-Koenigsberger et al. 2017; Goñi et al. 2014; Seguin et al. 2018; Vézquez-Rodríguez et al. 2020). As seen in Table 2, mean first passage time and shortest path

length both produced significant effects, with the shortest path length having a slightly larger semi-partial correlation (see Fig. 2A for predicted vs. empirical functional connectivity). However, when adding communicability to the model as seen in Table 3, the overall variance accounted for increased, and the semi-partial correlation of shortest path length was greatly reduced (though still significant), while the diffusion-based measures of mean first passage time and communicability had a much larger combined magnitude of semi-partial correlation. This model accounted for more variance than any of the measures independently ($R^2 = 0.165$; see Fig. 2B for predicted vs. empirical functional connectivity). It should be noted that the variance inflation factor (VIF) for the shortest path length in model 2 was greater than 5 ($VIF = 5.567$), indicating that multicollinearity between the independent variables may have affected the variance of the shortest path length coefficient. To address this, we also examined these variables using partial least squares regression, which is robust against multicollinearity.

Partial least squares regression

A partial least squares regression analysis was conducted with a dependent variable of functional connectivity and independent variables of mean first passage time, communicability, and shortest path length, using the *pls* function from the *pls* library in R (Mevik and Wehrens 2007). The

Table 2 AAL multiple linear model 1, with dependent variable functional connectivity. $R^2 = 0.157$, $R_{\text{adj}}^2 = 0.157$ (null model permutation $p = 0.012$)

| Effects | SPC | Estimate | Std. Error | <i>t</i> -value | <i>p</i> -value | VIF |
|-----------|--------|----------|------------|-----------------|-----------------|-------|
| Intercept | | 0.845 | 0.023 | 36.577 | <0.001 | |
| log(MFPT) | -0.117 | -0.060 | 0.007 | -8.051 | <0.001 | 2.927 |
| log(SPL) | -0.127 | -0.067 | 0.008 | -8.745 | <0.001 | 2.927 |

SPC semi-partial correlation, VIF variance inflation factor

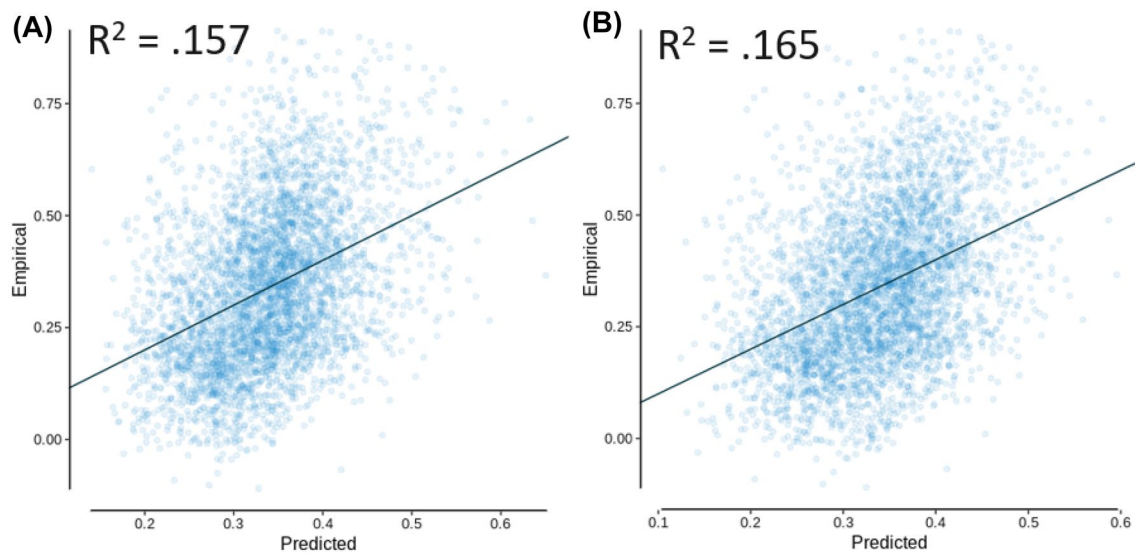


Fig. 2 AAL multiple linear regression models with empirical functional connectivity as a function of the predicted functional connectivity, for **A** model 1 and **B** model 2

Table 3 AAL multiple linear model 2, with dependent variable functional connectivity. $R^2=0.165$, $R_{adj}^2=0.165$ (null model permutation $p=0.016$)

| Effects | SPC | Estimate | Std. Error | <i>t</i> -value | <i>p</i> -value | VIF |
|-----------|--------|----------|------------|-----------------|-----------------|-------|
| Intercept | | 0.851 | 0.023 | 36.960 | <0.001 | |
| log(MFPT) | −0.138 | −0.075 | 0.008 | −9.592 | <0.001 | 3.229 |
| log(COM) | 0.089 | 0.012 | 0.002 | 6.184 | <0.001 | 2.493 |
| log(SPL) | −0.031 | −0.022 | 0.010 | −2.112 | 0.035 | 5.567 |

SPC semi-partial correlation, VIF variance inflation factor

independent variables were log transformed and standardized to have a mean of 0 and a standard deviation of 1. To validate the model and check for overfitting a k-fold cross-validation scheme was used with 10 folds. The number of components to include was decided when additional components no longer substantially decreased the root mean squared error of prediction. With 2 components included, the root mean squared error of prediction reached its minimum of 0.161, so 2 components were used. Cross-validation determined that the model was able to account for 16.3% ($R^2=0.163$, $R_{adj}^2=0.162$) of the variance in functional connectivity of novel validation samples, while the model accounted for 16.5% ($R^2=0.165$, $R_{adj}^2=0.164$) of the variance in functional connectivity when predicting the data for all connections (null model permutation $p=0.012$). These cross-validation results indicate that over-fitting is minimal. Finally, by investigating the coefficients for each of the independent variables, the pattern of results seen in the multiple linear regression model can be confirmed. For diffusion measures, mean first passage time had a coefficient of -0.037 , communicability had a coefficient of 0.018 , and shortest path length had a coefficient of -0.025 . Note that the negative relationship between structure and function

for mean first passage time and shortest path length was expected, as a higher value for these structural measures indicates weaker connectivity, while a positive relationship was expected for communicability as higher values indicate stronger connectivity. These coefficients support what was observed for the multiple linear regression, indicating that the effects of the diffusion model-based measures were greater in combined magnitude than that of the shortest path length.

Principal components analysis

The principal components analysis for functional connectivity, mean first passage time, communicability, and shortest path length shown in Fig. 3 demonstrates the unique component space occupied by each measure. This analysis was conducted using the *prcomp* function of the core *stats* library in *R*. To aid the interpretation of the principal component loadings of each variable, mean first passage time and shortest path length were multiplied by -1 so that larger values indicate better connectivity for all measures. In particular, Principal Component 1 seems to be sensitive to the variance in common between functional connectivity and the graph

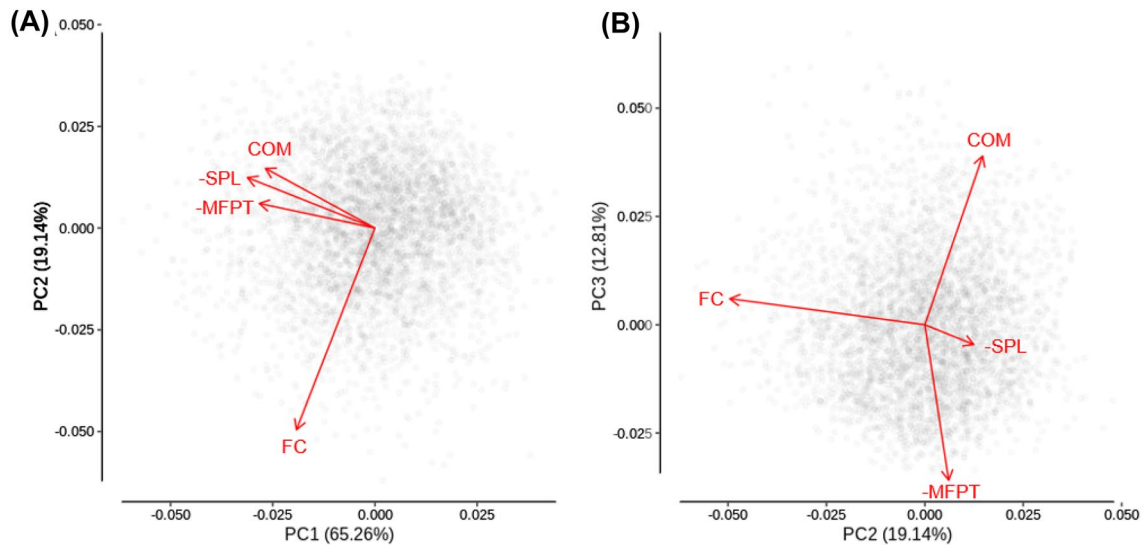


Fig. 3 AAL principal components analysis with data points and variable loadings as vectors. Variables considered were functional connectivity (FC), mean first passage time (MFPT; log transformed and

multiplied by -1 so that more positive values indicate better connectivity), communicability (COM; log transformed), and shortest path length (SPL; log transformed and multiplied by -1)

Table 4 AAL principal components analysis loadings for all 3 principal components

| Variable | PC1 | PC2 | PC3 |
|--------------------------|--------|--------|--------|
| FC | -0.359 | -0.926 | 0.112 |
| $-1 * \log(\text{MFPT})$ | -0.529 | 0.113 | -0.671 |
| $\log(\text{COM})$ | -0.501 | 0.274 | 0.728 |
| $-1 * \log(\text{SPL})$ | -0.583 | 0.232 | -0.086 |

Variables considered were functional connectivity (FC), mean first passage time (MFPT; log transformed and multiplied by -1 so that more positive values indicate better connectivity), communicability (COM; log transformed), and shortest path length (SPL; log transformed and multiplied by -1 as with MFPT)

theory measures, as these all have loadings in the same direction (Fig. 3A and Table 4). Conversely, functional connectivity loads strongly onto Principal Component 2, while the graph theory measures load weakly and in the opposite direction, suggesting that this component identifies variance in functional connectivity that is not well accounted for by the graph theory measures (Fig. 3A and Table 4). Finally, functional connectivity and shortest path length load very weakly onto Principal Component 3, while the loadings for mean first passage time and communicability are strong and in opposite directions, suggesting that this component speaks to the unique position in the component space of these diffusion model measures (Fig. 3B and Table 4).

Table 5 Brainnetome linear regression analyses

| Variable | R | R_{adj} | p -value | permutation p -value |
|---------------------|--------|------------------|------------|------------------------|
| $\log(\text{MFPT})$ | -0.232 | -0.231 | <0.001 | 0.014 |
| $\log(\text{COM})$ | 0.270 | 0.270 | <0.001 | <0.002 |
| $\log(\text{SPL})$ | -0.250 | -0.250 | <0.001 | <0.002 |

Dependent variable was functional connectivity, and independent variables considered were mean first passage time (MFPT; log transformed), communicability (COM; log transformed), and shortest path length (SPL; log transformed)

Brainnetome

Linear regression models

Using the Brainnetome atlas, there was again an inverse relationship between mean first passage time and functional connectivity, a positive relationship between communicability and functional connectivity, and an inverse relationship between shortest path length and functional connectivity. Again, the magnitude of the structure–function relationship for each of these measures was relatively comparable (see Table 5 and Fig. 4).

Multiple linear regression models

Multiple linear regression models were also investigated for the Brainnetome atlas, demonstrating that the semi-partial correlation for the shortest path length was much less than the combined magnitude for the diffusion-based measures of

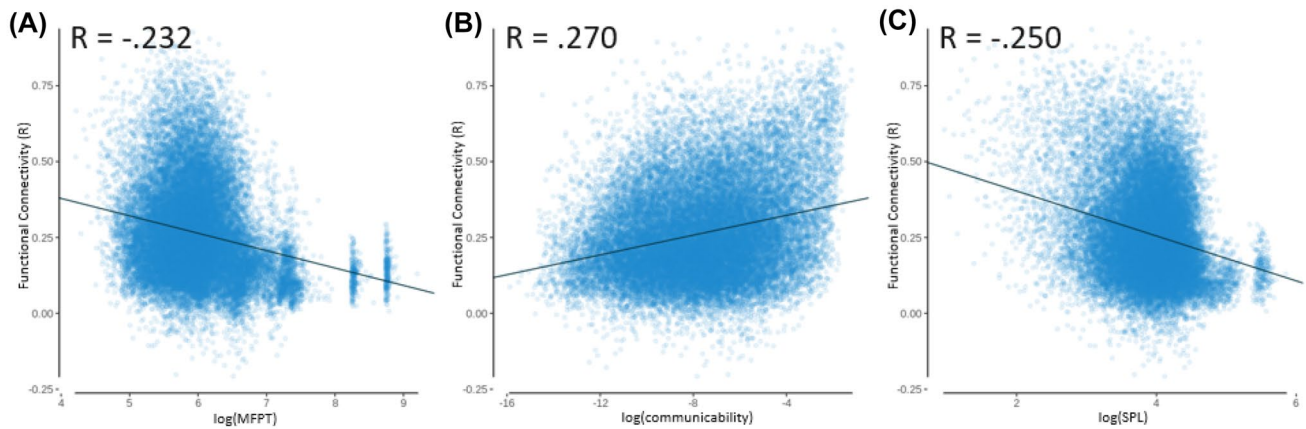


Fig. 4 Brainnetome linear regression models with functional connectivity as the independent variable and dependent variables of: **A** mean first passage time (MFPT; log transformed), $R_{adj} = -0.231$; **B** communicability (COM; log transformed), $R_{adj} = 0.270$; **C** and shortest path length (SPL; log transformed), $R_{adj} = -0.250$. With outlier clusters

removed from **A** by excluding cases with $\log(\text{MFPT}) > 8$ and **C** by excluding cases with $\log(\text{SPL}) > 5.33$ the correlation remains significant, with $R = -0.207$ for $\log(\text{MFPT})$ and $R = -0.224$ for $\log(\text{SPL})$. These outlier clusters are due to more isolated regions of the atlas that take more steps to reach than most other regions

Table 6 Brainnetome independent variable correlation matrix

| Variable | MFPT | COM | SPL |
|----------|--------|--------|--------|
| MFPT | 1 | -0.087 | 0.824 |
| COM | -0.087 | 1 | -0.342 |
| SPL | 0.824 | -0.342 | 1 |

Variables include mean first passage time (MFPT), communicability (COM), and shortest path length (SPL)

mean first passage time and communicability (see Tables 6, 7 and 8 and Fig. 5).

Partial least squares regression

The partial least squares regression again demonstrated that the combined magnitude of the coefficients for the diffusion model measures of mean first passage time and communicability were much greater than the shortest path routing model measure of shortest path length (see Table 9).

Principal components analysis

The principal components analysis was again conducted, but this time for the Brainnetome atlas. This analysis replicated

Table 7 Brainnetome multiple linear model 1, with dependent variable functional connectivity. $R^2 = 0.067$, $R_{adj}^2 = 0.067$ (null model permutation $p = 0.002$)

| Effects | SPC | Estimate | Std. Error | t-value | p-value | VIF |
|-----------|--------|----------|------------|---------|---------|-------|
| Intercept | | 0.614 | 0.008 | 73.783 | <0.001 | |
| log(MFPT) | -0.067 | -0.025 | 0.002 | -12.024 | <0.001 | 2.275 |
| log(SPL) | -0.116 | -0.052 | 0.002 | -20.796 | <0.001 | 2.275 |

SPC semi-partial correlation, VIF variance inflation factor

Table 8 Brainnetome multiple linear model 2, with dependent variable functional connectivity. $R^2 = 0.090$, $R_{adj}^2 = 0.090$ (null model permutation $p < 0.002$)

| Effects | SPC | Estimate | Std. Error | t-value | p-value | VIF |
|-----------|--------|----------|------------------------|---------|---------|-------|
| Intercept | | 0.558 | 0.008 | 65.877 | <0.001 | |
| log(MFPT) | -0.112 | -0.044 | 0.002 | -20.383 | <0.001 | 2.534 |
| log(COM) | 0.152 | 0.015 | 5.442×10^{-4} | 27.643 | <0.001 | 2.702 |
| log(SPL) | 0.034 | 0.022 | 0.004 | 6.107 | <0.001 | 4.969 |

SPC semi-partial correlation, VIF variance inflation factor

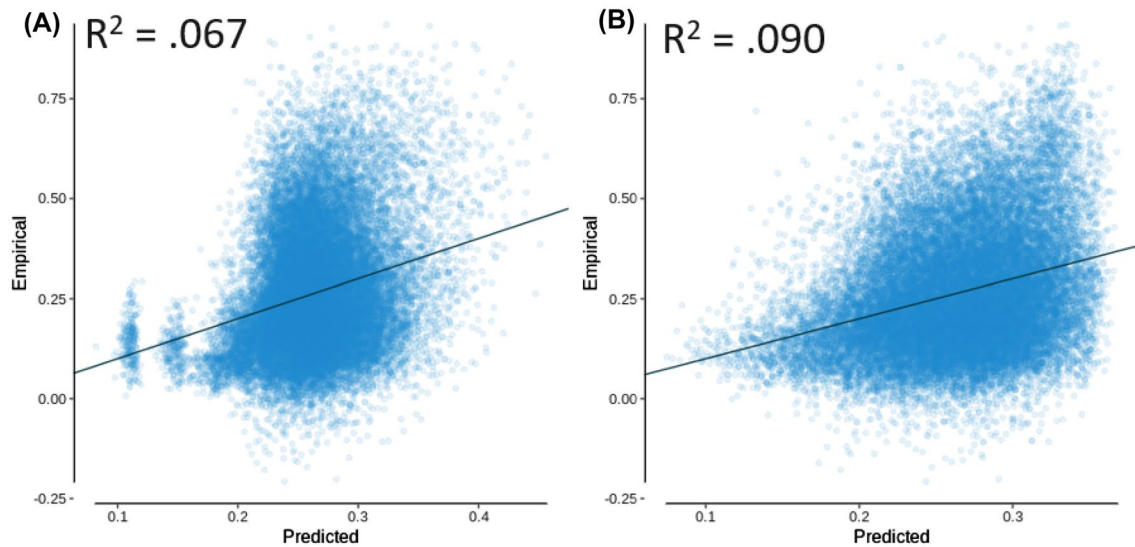


Fig. 5 Brainnetome multiple linear regression models with empirical functional connectivity as a function of the predicted functional connectivity, for **A** model 1 and **B** model 2

Table 9 Brainnetome partial least squares (PLS) regression

| Variable | Coefficient |
|-----------|-------------|
| log(MFPT) | -0.015 |
| log(COM) | 0.035 |
| log(SPL) | -0.003 |

Dependent variable was functional connectivity, and independent variables considered were mean first passage time (MFPT; log transformed), communicability (COM; log transformed), and shortest path length (SPL; log transformed). Variance accounted for was $R^2=0.087$, $R_{adj}^2=0.087$, null model permutation $p < 0.002$

the pattern of results seen for the AAL atlas (see Fig. 6 and Table 10).

Supplementary analyses

Individual-level analyses

Individual-level analyses were performed that mirrored the mean-level analyses, demonstrating that the variance accounted for was reduced but significant in all cases, with the same pattern of results as in the mean-level analyses (see Supplementary Tables 1 through 10). These analyses utilized the upper triangle of each connectivity matrix, reshaped to a single dimensional array for each individual and then concatenated across all individuals into a single large array for each of the mean first passage time, communicability,

shortest path length, and functional connectivity measures. These arrays were used as variables in the statistical analyses in the same way as for the mean-level data.

Split-half analyses

Half of the data was used as a training set to train the models and the other half was used as a test set to test the models on novel data to demonstrate the predictability of out-of-sample data. In all cases, the test sets were predicted with comparable accuracy (see Supplementary Tables 11 through 18).

PCA null models

Null models were used to produce a null distribution for loadings of each of the 4 variables on the 3 Principal Components, and demonstrated that there was a significant distance between the null and empirical PC loadings (see Supplementary Tables 19 and 20).

Discussion

This investigation and comparison of graph theory structural connectivity measures based on two different theories of how information passes from one region to another in the brain has highlighted the importance of the diffusion model relative to the more straightforward but less biologically plausible shortest path routing model. Diffusion measures of mean first passage time and communicability as well as the shortest path routing measure of shortest path length were calculated from the brain structural connectivity, and

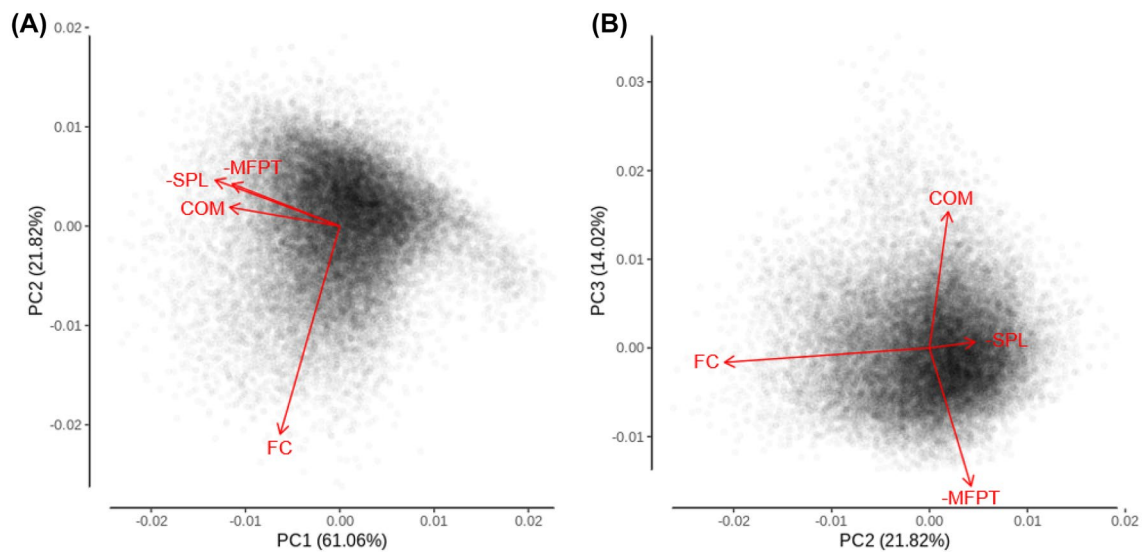


Fig. 6 Brainnetome principal components analysis with data points and variable loadings as vectors. Variables considered were functional connectivity (FC), mean first passage time (MFPT; log transformed and multiplied by -1 so that more positive values indicate

better connectivity), communicability (COM; log transformed), and shortest path length (SPL; log transformed and multiplied by -1 as with MFPT)

Table 10 Brainnetome principal components analysis loadings for all 3 principal components

| Variable | PC1 | PC2 | PC3 |
|-------------------------------|--------|--------|--------|
| FC | -0.288 | -0.954 | -0.074 |
| $-1 \times \log(\text{MFPT})$ | -0.520 | 0.194 | -0.710 |
| $\log(\text{COM})$ | -0.532 | 0.087 | 0.700 |
| $-1 \times \log(\text{SPL})$ | -0.603 | 0.211 | 0.030 |

Variables considered were functional connectivity (FC), mean first passage time (MFPT; log transformed and multiplied by -1 so that more positive values indicate better connectivity), communicability (COM; log transformed), and shortest path length (SPL; log transformed and multiplied by -1 as with MFPT)

these measures were related to functional connectivity. In isolation, each of these measures were comparable in the level to which they were related to functional connectivity. When analysed as hybrid models with each of these measures included, more variance was accounted for than any measure on its own, supporting past research suggesting that a hybrid/balance of both the diffusion and shortest path routing models is appropriate for describing the structure–function relationship (Goñi et al. 2013, 2014). However, when analysed together in multiple linear regression, partial least squares regression, and principal components analysis, it was clear that the diffusion model and the ideas it can express capture more of the variance in functional connectivity than shortest path routing, suggesting that the diffusion model may be closer to describing how information travels from one region to another in the structural connectivity

network to produce the observed patterns of functional connectivity. These findings suggest that shortest path routing may be somewhat redundant when diffusion models are able to travel along similarly efficient routes in the brain and that diffusion models are able to additionally tap into aspects of functional connectivity that are not considered by shortest path routing. These results are not surprising, given that diffusion models are more biologically plausible than shortest path routing as there is no evidence to suggest that regions have global network knowledge about what path would be the shortest (Avena-Koenigsberger et al. 2019; Seguin et al. 2018, 2022; Zamani Esfahlani et al. 2022). Past research has demonstrated that both diffusion models and shortest path routing models are important frameworks for understanding the architecture of brain networks (Goñi et al. 2013, 2014), but this work has taken an important next step in distinguishing the greater relative ability for the diffusion measures to accurately predict function from the underlying structural connectivity.

Limitations and future directions

While diffusion and shortest path routing models are the most commonly discussed in network neuroscience, other models have been suggested that may add a unique perspective on how information travels in the brain. One such proposed theory for how information may transfer is *greedy navigation*, in which the Euclidean (three dimensional) or geodesic (two-dimensional flattened surface of cortex) distance between regions is used to travel through the network

to whichever region is closest to the target region. This model has been investigated in simulated networks, which have shown that greedy navigation is able to successfully send information between regions without getting stuck (i.e., ending up at a node that is closer than all neighbours to the destination but unconnected to the destination) when there is a balance between clustered connections of proximal nodes and a power-law like distribution whereby clusters are connected by a small number of highly connected hub nodes (Boguñá et al. 2009). The pattern of navigation under these conditions is such that the path tends to travel to a nearby hub node, travel a long distance to another hub node, then travel to a low-degree node in a cluster close to the destination. This pattern has also been demonstrated in the macaque brain network (Harriger et al. 2012) and in *C. elegans* (Towlson et al. 2013). Research using this theory in the human brain has demonstrated the prediction of functional connectivity from structural connectivity (Seguin et al. 2018, 2020). Future research should implement a model of *greedy navigation* to replicate that the path length of a greedy navigator predicts functional connectivity, and additionally investigate whether the pattern of paths resembles that seen in simulated models and animal models.

Internet and computer analogies have also been proposed for application to brain networks, with encouraging results (Graham and Rockmore 2011; Mišić et al. 2014). Information flow in computer networks (such as the internet) is limited by bandwidth, which represents the amount of information that can travel in a certain amount of time (e.g., bits/s). Information can be transmitted via packet switching, which breaks messages into packets that are labeled with the intended destination. These packets traverse the network efficiently by utilizing connections at separate times and buffering in a node if a connection is full, with the downside that if the node buffer is full then the information is lost. The structural connectivity of the macaque has been used to simulate a message-switched variant of this model of information flow, showing that compared to other networks there was more message loss, lower throughput, but faster transit times (Mišić et al. 2014). This suggests that under the model assumptions speed would be optimized in the macaque brain over maintaining the integrity of each individual signal. Future research should work on also applying a packet switching simulation model using the human brain to uncover whether similar signatures of information flow are seen between human and macaque networks, and whether the time taken for information to travel between regions is predictive of functional connectivity measures.

As another example of further work to be done in the field, brain network analysis of the cat brain has found that regions with similar connectivity profiles (connect to the same or similar regions) tended to correspond with groups of regions performing similar tasks (Zamora-López et al.

2010). One measure of the similarity between connectivity profiles of regions is *cosine similarity*, which has been used to demonstrate that functionally similar regions in the macaque brain also had high cosine similarity, and that regions with high cosine similarity were also likely to be connected and located close together (Song et al. 2014). This measure has recently been investigated along with many other measures in the context of the human brain (Zamani Esfahlani et al. 2022), and more research should be done to investigate whether human brain regions with similar connectivity profiles are more likely to be connected, to what extent this pattern differs between functionally similar clusters and core hub regions, and to what extent the structural cosine similarity is able to predict the functional connectivity and functional cosine similarity.

Finally, a recent line of research has investigated the potential of hybrid models of information transfer that allow for regional heterogeneity in the navigation strategy applied. This is an extension of research that has shown differences in structure–function coupling when looking at unimodal vs. transmodal regions, and the development of these patterns with age (Baum et al. 2020). These models allow for different regions to apply unique strategies of information transfer, and have shown promising findings (Avena-Koenigsberger et al. 2019; Vázquez-Rodríguez et al. 2019; Zamani Esfahlani et al. 2022). This is a promising line of research that we expect will continue improving our understanding of how information is propagated throughout the structural connectivity network of the brain.

Conclusion

Although the models described here do not claim to be a fully accurate and comprehensive description of how information is transferred in the brain, especially considering that the nature of the structure–function relationship will inevitably vary as the scale of interest goes from the macroscale to the microscale, by examining the diffusion and shortest path routing models together, this research has demonstrated that diffusion models are better suited to describing the relationship between structural and functional connectivity at the macroscale. In the future, alternative models discussed here could be examined together, to contribute to a fuller picture of which aspects of these theoretical models are able to best approximate the ground truth of information transfer in the human brain network.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00429-023-02613-2>.

Acknowledgements This research was supported by the Natural Sciences and Engineering Research Council of Canada through Alexander Graham Bell Canada Graduate Scholarships to Josh Neudorf and Shaylyn Kress, and Discovery Grant 18968-2013-25 to the senior author Ron Borowsky. The authors affirm that there are no conflicts of interest to disclose. Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Author contributions J.N. performed the analyses and wrote the main manuscript text. All authors contributed to the conceptualization of the analysis, analysis decisions, and review of the manuscript.

Availability of data and material All data are made publicly available by the Human Connectome Project (Van Essen et al. 2013).

Code availability The code used to apply the equations is available upon reasonable request.

Declarations

Competing interests The authors declare no competing interests.

Conflicts of interest The authors report no conflicts of interest.

Ethics approval, consent to participate, and consent for publication Please see Van Essen et al. (2013) for a description of ethics and consent.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Avena-Koenigsberger A, Mišić B, Hawkins RXD, Griffa A, Hagmann P, Goñi J, Sporns O (2017) Path ensembles and a tradeoff between communication efficiency and resilience in the human connectome. *Brain Struct Funct* 222(1):603–618. <https://doi.org/10.1007/s00429-016-1238-5>
- Avena-Koenigsberger A, Misis B, Sporns O (2018) Communication dynamics in complex brain networks. *Nat Rev Neurosci* 19(1):1. <https://doi.org/10.1038/nrn.2017.149>
- Avena-Koenigsberger A, Yan X, Kolchinsky A, van den Heuvel MP, Hagmann P, Sporns O (2019) A spectrum of routing strategies for brain networks. *PLOS Comput Biol* 15(3):3. <https://doi.org/10.1371/journal.pcbi.1006833>
- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67(1):1–48. <https://doi.org/10.18637/jss.v067.i01>
- Baum GL, Cui Z, Roalf DR, Ciric R, Betzel RF, Larsen B, Cieslak M, Cook PA, Xia CH, Moore TM, Ruparel K, Oathes DJ, Alexander-Bloch AF, Shinohara RT, Raznahan A, Gur RE, Gur RC, Bassett DS, Satterthwaite TD (2020) Development of structure-function coupling in human brain networks during youth. *Proc Natl Acad Sci USA* 117(1):771–778. <https://doi.org/10.1073/pnas.1912034117>
- Betzel RF, Faskowitz J, Mišić B, Sporns O, Seguin C (2022) Multi-policy models of interregional communication in the human connectome. *bioRxiv*. <https://doi.org/10.1101/2022.05.08.490752>
- Boguñá M, Krioukov D, Claffy KC (2009) Navigability of complex networks. *Nat Phys* 5(1):1. <https://doi.org/10.1038/nphys1130>
- Crofts JJ, Higham DJ, Bosnell R, Jbabdi S, Matthews PM, Behrens TEJ, Johansen-Berg H (2011) Network analysis detects changes in the contralesional hemisphere following stroke. *Neuroimage* 54(1):161–169. <https://doi.org/10.1016/j.neuroimage.2010.08.032>
- Dijkstra EW (1959) A note on two problems in connexion with graphs. *Numer Math* 1(1):269–271
- Estrada E, Hatano N (2008) Communicability in complex networks. *Phys Rev E* 77(3):036111. <https://doi.org/10.1103/PhysRevE.77.036111>
- Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, Yang Z, Chu C, Xie S, Laird AR, Fox PT, Eickhoff SB, Yu C, Jiang T (2016) The human brainnetome atlas: a new brain atlas based on connective architecture. *Cereb Cortex* 26(8):3508–3526. <https://doi.org/10.1093/cercor/bhw157>
- Fornito A, Zalesky A, Bullmore ET (2016a) Fundamentals of brain network analysis. Elsevier/Academic Press
- Fornito A, Zalesky A, Bullmore ET (eds) (2016b) Chapter 7—paths, diffusion, and navigation. In: Fundamentals of brain network analysis, pp 207–255. Academic Press. <https://doi.org/10.1016/B978-0-12-407908-3.00007-8>
- Fox J, Weisberg S (2019) An R Companion to Applied Regression (Third). Sage. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>. Accessed 15 Dec 2022
- Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M (2013) The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 80:105–124. <https://doi.org/10.1016/j.neuroimage.2013.04.127>
- Goñi J, Avena-Koenigsberger A, Velez de Mendizabal N, van den Heuvel MP, Betzel RF, Sporns O (2013) Exploring the morphospace of communication efficiency in complex networks. *PLoS ONE* 8(3):3. <https://doi.org/10.1371/journal.pone.0058070>
- Goñi J, van den Heuvel MP, Avena-Koenigsberger A, Velez de Mendizabal N, Betzel RF, Griffa A, Hagmann P, Corominas-Murtra B, Thiran J-P, Sporns O (2014) Resting-brain functional connectivity predicted by analytic measures of network communication. *Proc Natl Acad Sci USA* 111(2):2. <https://doi.org/10.1073/pnas.1315529111>
- Graham D, Rockmore D (2011) The packet switching brain. *J Cogn Neurosci* 23(2):267–276. <https://doi.org/10.1162/jocn.2010.21477>
- Hagberg A, Schult, D, Swart P (2008) Exploring network structure, dynamics, and function using networkx. In: Varoquaux G, Vaught T, Millman J (eds) Proceedings of the 7th Python in science conference (SciPy 2008), pp 11–15
- Hallquist MN, Hwang K, Luna B (2013) The nuisance of nuisance regression: Spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. *Neuroimage* 82:208–225. <https://doi.org/10.1016/j.neuroimage.2013.05.116>
- Harriger L, van den Heuvel MP, Sporns O (2012) Rich club organization of macaque cerebral cortex and its role in network

- communication. PLoS ONE 7(9):e46497. <https://doi.org/10.1371/journal.pone.0046497>
- Kim S (2015) ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Commun Stat Appl Methods* 22(6):665–674. <https://doi.org/10.5351/CSAM.2015.22.6.665>
- Latora V, Marchiori M (2001) Efficient behavior of small-world networks. *Phys Rev Lett* 87(19):198701. <https://doi.org/10.1103/PhysRevLett.87.198701>
- Li Y, Jewells V, Kim M, Chen Y, Moon A, Armao D, Troiani L, Markovic-Plese S, Lin W, Shen D (2013) Diffusion tensor imaging based network analysis detects alterations of neuroconnectivity in patients with clinically early relapsing-remitting multiple sclerosis. *Hum Brain Mapp* 34(12):3376–3391. <https://doi.org/10.1002/hbm.22158>
- Maslov S, Sneppen K (2002) Specificity and stability in topology of protein networks. *Science* 296(5569):910–913. <https://doi.org/10.1126/science.1065103>
- Mevik B-H, Wehrens R (2007) The pls package: principal component and partial least squares regression in R. *J Stat Softw* 18:1–23. <https://doi.org/10.18637/jss.v018.i02>
- Mišić B, Sporns O, McIntosh AR (2014) Communication efficiency and congestion of signal traffic in large-scale brain networks. *PLoS Comput Biol* 10(1):e1003427. <https://doi.org/10.1371/journal.pcbi.1003427>
- Neudorf J, Ekstrand C, Kress S, Borowsky R (2020a) Brain structural connectivity predicts brain functional complexity: diffusion tensor imaging derived centrality accounts for variance in fractal properties of functional magnetic resonance imaging signal. *Neuroscience* 438:1–8. <https://doi.org/10.1016/j.neuroscience.2020.04.048>
- Neudorf J, Kress S, Gould L, Gibb K, Mickleborough M, Borowsky R (2020b) Language lateralization differences between left and right temporal lobe epilepsy as measured by overt word reading fMRI activation and DTI structural connectivity. *Epilepsy Behav* 112:107467. <https://doi.org/10.1016/j.yebeh.2020.107467>
- Neudorf J, Kress S, Borowsky R (2022) Structure can predict function in the human brain: a graph neural network deep learning model of functional connectivity and centrality based on structural connectivity. *Brain Struct Funct* 227(1):331–343. <https://doi.org/10.1007/s00429-021-02403-8>
- R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/>. Accessed 15 Dec 2022
- Rubinov M, Sporns O (2011) Weight-conserving characterization of complex functional brain networks. *Neuroimage* 56(4):2068–2079. <https://doi.org/10.1016/j.neuroimage.2011.03.069>
- Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM (2014) Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* 90:449–468. <https://doi.org/10.1016/j.neuroimage.2013.11.046>
- Seguin C, van den Heuvel MP, Zalesky A (2018) Navigation of brain networks. *Proc Natl Acad Sci* 115(24):6297–6302. <https://doi.org/10.1073/pnas.1801351115>
- Seguin C, Tian Y, Zalesky A (2020) Network communication models improve the behavioral and functional predictive utility of the human structural connectome. *Netw Neurosci* 4(4):980–1006. https://doi.org/10.1162/netn_a_00161
- Seguin C, Mansour LS, Sporns O, Zalesky A, Calamante F (2022) Network communication models narrow the gap between the modular organization of structural and functional brain networks. *Neuroimage* 257:119323. <https://doi.org/10.1016/j.neuroimage.2022.119323>
- Song HF, Kennedy H, Wang X-J (2014) Spatial embedding of structural similarity in the cerebral cortex. *Proc Natl Acad Sci USA* 111(46):16580–16585. <https://doi.org/10.1073/pnas.1414153111>
- Suárez LE, Markello RD, Betzel RF, Misisic B (2020) Linking structure and function in macroscale brain networks. *Trends Cogn Sci* 24(4):302–315. <https://doi.org/10.1016/j.tics.2020.01.008>
- Towlson EK, Vértes PE, Ahnert SE, Schafer WR, Bullmore ET (2013) The rich club of the *C. elegans* neuronal connectome. *J Neurosci* 33(15):6380–6387. <https://doi.org/10.1523/JNEUROSCI.3784-12.2013>
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15(1):273–289. <https://doi.org/10.1006/nimg.2001.0978>
- van den Heuvel MP, Sporns O (2013) Network hubs in the human brain. *Trends Cogn Sci* 17(12):683–696. <https://doi.org/10.1016/j.tics.2013.09.012>
- Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K (2013) The WU-Minn Human Connectome Project: an overview. *Neuroimage* 80:62–79. <https://doi.org/10.1016/j.neuroimage.2013.05.041>
- Vázquez-Rodríguez B, Suárez LE, Markello RD, Shafiei G, Paquola C, Hagmann P, van den Heuvel MP, Bernhardt BC, Spreng RN, Misisic B (2019) Gradients of structure-function tethering across neocortex. *Proc Natl Acad Sci USA* 116(42):21219–21227. <https://doi.org/10.1073/pnas.1903403116>
- Vázquez-Rodríguez B, Liu Z-Q, Hagmann P, Misisic B (2020) Signal propagation via cortical hierarchies. *Netw Neurosci* 4(4):1072–1090. https://doi.org/10.1162/netn_a_00153
- Wang S-P, Pei W-J (2008) First passage time of multiple Brownian particles on networks with applications. *Phys A* 387(18):4699–4708. <https://doi.org/10.1016/j.physa.2008.03.032>
- Watts DJ, Strogatz SH (1998) Collective dynamics of “small-world” networks. *Nature* 393(6684):440–442. <https://doi.org/10.1038/30918>
- Yeh F, Wedeen VJ, Tseng WI (2010) Generalized q-Sampling Imaging. *IEEE Trans Med Imaging* 29(9):1626–1635. <https://doi.org/10.1109/TMI.2010.2045126>
- Yeh F-C, Verstynen TD, Wang Y, Fernández-Miranda JC, Tseng W-YI (2013) Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0080713>
- Zamani Esfahlani F, Faskowitz J, Slack J, Mišić B, Betzel RF (2022) Local structure-function relationships in human brain networks across the lifespan. *Nat Commun* 13(1):1. <https://doi.org/10.1038/s41467-022-29770-y>
- Zamora-López G, Zhou C, Kurths J (2010) Cortical hubs form a module for multisensory integration on top of the hierarchy of cortical networks. *Front Neuroinform*. <https://doi.org/10.3389/neuro.11.001.2010>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.