Structurally Constrained Effective Brain Connectivity

Alessandro Crimi[1,7], Luca Dodero[1,2], Fabio Sambataro[3], Vittorio Murino[1,4], Diego Sona[1,5]

How function arises from structure is of interest in many fields from proteomics to neuroscience. In particular, among the brain research community the fusion of structure and function data can shed new lights on underlying operational network principles in the brain. Targeting this issue, the manuscript proposes a constrained autoregressive model generating "effective" connectivity given structural and functional information. In practice, an initial structural connectivity representation is altered according to functional data, by minimizing the reconstruction error of an autoregressive model constrained by the structural prior. The proposed model has been tested in a community detection framework, where the brain is partitioned using the effective networks across multiple subjects. The model is further validated in a case-control experiment, which aims at differentiating healthy subjects from young patients affected by autism spectrum disorder. Results showed that using effective connectivity resulted in clusters that better describe the functional interactions between different regions while maintaining the structural organization, and a better discrimination in the case-control classification task.

The investigation of how structure sculpts function is present in several fields. Whereby a protein function can be predicted by its sequence and structure is a common issue in proteomics [1, 2]. Moreover, network representation can bring across data of different types and domains [3]. Those are present at many scales from cell-specific metabolic or regulatory pathways inside neurons to interactions between systems of cortical areas which is the focus of this paper. In neuroscience, it is now well established that functional segregation is a bearing principle of brain organization. However, the human brain is a complex network characterized by spatially interconnected regions that can be causally or concurrently activated during specific tasks or at rest. As a consequence, integration of segregated regions is emerging as the most probable structural organization explaining the complexity of brain function. This principle is however difficult to prove and the easiest approach is to investigate it through functional connectivity analysis, which is usually based on the determination of functional correlations. In this context, we propose a model that sculpts structural data from functional information.

Current functional connectivity analyses do not capitalize on structural information, thus preventing the interpretation of their results in terms of effective connectivity. Indeed, understanding the relationships between the functional activity in different brain regions and the structural network highlighted using tractography can convey useful information about brain underlying principles. While in Vincent et al. [4] and Fukushima et al. [5] it has been shown there is a significant overlap between neuroanatomical connections and correlations of brain functional signals, it is yet to be understood how whole-brain network interacts during a specific task or at rest. Among other networks, the default mode network (DMN) [6], which is a set of brain regions more active during rest relative to a goal directed task, has emerged for its possible role in allocating attention, self-referential processing, and memory. A predictive framework based on multiple sparse linear regression has been used to predict functional series from structural data [7, 8], and a model called the "virtual brain" has been designed to simulate brain activity in injured and healthy subjects [9]. With the aim to highlight the relationships between function and structural connectivity, a different approach relies on a Bayesian framework to estimate the functional connectivity using a structural graph as a prior [10]. More recently, some approaches modeling nodal activity as Ornstein-Uhlenbeck processes have been proposed. More specifically, those approaches obtain causal structures combining those processes to l_1 minimization [11], prior knowledge related to rich-clubs [12], or differences between task-based and resting-state sequences [13]. However, results generating functional connectivity from structural information have been so far challenging, due to the fact that certain high correlations appear between brain regions not directly linked by structural connections. Another approach based on causal graphs has proposed to use the notion of an external intervention that actively perturbs the system defined by the brain region connectome [14].

An alternative schema adopted to infer the functional di-

[1]Pattern Analysis and Computer Vision, Istituto Italiano di Tecnologia, Italy
[2]Department of Radiology and Medical Informatics, University of Geneva, Switzerland
[3]Department of Experimental and Clinical Medical Sciences, University of Udine, Italy
[4]Department of Computer Science, University of Verona, Italy
[5]Neuroinformatics Laboratory, Fondazione Bruno Kessler, Trento, Italy
[6]Neuropathology department, University Hospital of Zurich, Switzerland

rected connections from brain activity is based on methods investigating the functional causality like the Dynamical Causal Model (DCM) [15] and the Granger causality [16]. DCM uses an explicit model of regional neural dynamics to capture changes in regional activations and inter-regional coupling in response to stimulus or task demands. Statistical inference is used to estimate parameters for directed influences between neuronal elements. While this is a powerful method to study effective connectivity, its main limitation is the combinatorial complexity on the number of modeled regions and connections, which limits its applicability to only few regions. Granger causality also measures the causal influence and the flow of information between regions. Despite the slow dynamics and the regions variability of hemodynamic response making it a controversial method for functional magnetic resonance images (fMRIs), it has been also used to identify the dynamics of Blood-Oxygen-Level Dependent (BOLD) signal flow between brain regions [17, 18, 19]. Extensions as non-linear formulation of the Granger causality using Kernels [20], and sparsity [21, 22] have been proposed.

Although spurious influences might be obtained because the hemodynamic activity is known to have differential delays with respect to the underlying neuronal activity, the impact of such variable hemodynamic delays it has been shown to be negligible when the sampling interval is as long as 2s [23]. Moreover, indirect connections could be addressed by conditional Granger causality analysis [24]. Li et al. proposed to extract BOLD signals from the two ends of an individual white matter fiber derived from tractography [18], although the model can explain causality between fiber-connected regions, it is subjected to noise and other imprecision derived from single fiber tractography and it performs univariate analysis looking at one connection at a time, while the method proposed in this manuscript is a multivariate approach working on whole brain structural connectome. The method has been extended to handle the multivariate case by using a Lasso regression [25], though this requires the optimization of the regularization parameters via cross-validation. Indeed, tractography might be imprecise and introduce false positive connections [26]. To refine the fiber tracking by keeping only the fibers which are used during tasks or resting-state might be a way to reduce false positive, and the proposed method aims at doing so by using an autoregressive model constrained in a novel manner.

A multivariate autoregressive model (MAR) is a random process specifying that the variables depend linearly on their own previous values and on a stochastic term. Thank to this, Granger causality does not suffer an excessive computational complexity. Nonetheless, in case of a large number of regions involved in the analysis, it is affected by cancellation issues and high sensitivity to noise. This makes it impossible to perform a whole brain analysis involving many brain regions. Furthermore, Granger causality approach has not yet considered structural connectivity in a multivariate manner.

Taking inspiration from these limitations, this paper presents an extension of the MAR model introducing a constrained MAR (CMAR), which uses the structural connectivity as a prior to bound the search space during parameter fitting. This fusion of structural connectivity and functional time-series aims at representing an effective brain connectivity, addressing a whole brain analysis thanks to the sparse representation of the connectivity matrix. The validation of the proposed method has been performed first by looking in a reduced subset of areas related to the DMN, testing the quality of results in a typical brain community detection framework, using a recently developed technique based on group-wise spectral clustering [27], and a classification task aiming at discriminate autistic and typically developing children [28]. This investigation allowed to validate the hypothesis that clustering the brain using the effective connectivity matrix retrieved with CMAR, while preserving the structural partitioning, it also optimizes the graph cut minimizing the loss of functional interactions. The classification task between autistic and typically developing subjects also improved. This shows the benefits of using effective connectivity to study the modular organization of the human brain in both healthy and diseased populations.

Structurally Constrained Autoregressive model

More in details, a multivariate autoregressive model of order n (MAR(n)) is a stochastic process defining r variables $\mathbf{y}(t)$ as linearly dependent on their own previous values and on a stochastic term:

$$\mathbf{y}(t) = \sum_{i=1}^{n} \mathbf{A}_i \cdot \mathbf{y}(t-i) + \epsilon, \tag{1}$$

where the r-by-r coefficient matrices $\mathbf{A}_i$ are the model parameters and ϵ is the additive Gaussian noise. The weight matrices $\mathbf{A}_i$ describe the linear dependencies between the r time-series $[y_1, \ldots, y_r]$, therefore, they can be intended as functional connectivity matrix at different time lags. The aim of the proposed method is to infer a functional connectivity justified by the structural connectivity, so that the functional activity of the brain can be described only using the direct physical connections between the brain regions.

To obtain this result, we resort to a set of constraints on the above MAR model, which force the MAR model to only fit the parameters associated to existing structural connections. In this way the signal in a brain region is described by a linear combination of the function of structurally connected regions only. Moreover, to avoid negative correlations between the regions only a non-negative reconstruction is allowed. Formally these two constrains can be expressed as:

$$\begin{cases} \mathbf{A}_i \odot \mathbf{B} = \mathbf{A}_i \\ \mathbf{A}_i \geq 0 \end{cases} \tag{2}$$

where $\odot$ denotes the Hadamard or element-wise product, and $\mathbf{B}$ is an indicator matrix of the structural connectivity $\mathbf{A}_{\text{init}}$:

$$B_{uv} = \begin{cases} 0 & \text{if the } uv \text{ element of } \mathbf{A}_{\text{init}} \text{ is } 0 \\ 1 & \text{otherwise.} \end{cases}$$

This $\mathbf{B}$ matrix is used as a structural bias into the model, constraining each $\mathbf{A}_i$ to have zero elements where no structural connectivity appears.

Given an initial structural connectivity matrix $\mathbf{A}_{\text{init}}$ and the functional signal in a column-vector $\mathbf{y}$ for all T time-frames, an effective connectivity matrix can be determined minimizing a reconstruction error of the MAR model in eq. (1) subject to the structural constraint in eq. (2) as follows:

$$E = \frac{1}{2}\sum_{t=n}^{T}\left\|\mathbf{y}(t) - \sum_{i=1}^{n}(\mathbf{A}_i \odot \mathbf{B}) \cdot \mathbf{y}(t-i)\right\|_2^2 . \tag{3}$$

The effective connectivity matrix is therefore the matrix $\mathbf{A}_i$ which minimize the reconstruction error above. The details of the minimization process are described in the supplementary *Data and Method section*. The choice of the model order n can be defined by Akaike information criterion [29], or empirically in a cross-validation settings. Within the experiments we have adopted this latter approach to choose the order of the CMAR model.

The reconstruction error shown per subject in Figure 2(f) shows, as expected, that the lowest error is given by considering the whole network in the CMAR computation. However, when removing some connections according to the clustering results, the communities determined from the effective connectivity matrix show to be more self explanatory in terms of functional activity then the communities obtained from the structural connectivity only. Moreover, Figures 2 (a), (b) and supplementary Figures A.2 depicts an example of initial structural matrix and final effective connectivity matrix where each row and column represents a region of interest (ROI). Showing that some connections are canceled out, while others are proportionally strengthened. The difference in values is related to the convergence of the algorithm to match the functional series.

We further validate the method with three experiments. A proof of concept with using only DMN regions, an effective brain community detection task, and a discrimination task between subjects in the autism spectrum disorder and typically developing subjects. In all our experiments we used the Power atlas defining 264 putative functional ROIs [30]. A functional atlas has been chosen given the need of mapping functional series.

Default Mode Network analysis

To first validate this model, we analyzed a subset of areas given by the functional DMN and the connections that relate each other. Some DMN areas are not directly structurally connected, though some activity is visible simultaneously [6]. To see if the structural connections relating the functional areas of DMN are maintained after we filter the connections with our CMAR model, we conducted an experiment using the structural and functional information of the main DMN areas (frontal middle gyrus, and posterior cingulate cortex, and fusiform gyrus; all both left and right). The proposed model has been used on a whole dataset by using these areas showing that they persist after the autoregressive model and still identify areas which are physically connected and used functionally as shown in the example in Figure 1. The depicted values are an average connectivity matrix for the whole dataset (a) and an example of tractography among those ROIs for one subject (c). This experiment was carried out by using the diffusion tensor and fMRI volumes of 200 healthy subjects from the Nathan Kline Institute-Rockland dataset [31]. Moreover, the results were compared to those obtained by using the method most similar to the proposed one: the predictive framework based on multiple sparse linear regression of Deligiani et al. [7]. It was noted that the method produced similar results if the whole dataset of 200 subjects was used, and the structural support was set to the maximum (to allow cancellation of sparsity only to 1% [7]). If only 10 random samples were used, the proposed method was still producing the same results, while the Deligiani et al. framework was canceling out some connections (as shown in Figure 1(b)). With only 5 random samples, the proposed framework was still working properly - as it is supposed to produce results with just a pair of structural/functional volumes - but the Deligiani et al. framework was not able to terminate the randomized Lasso by using the implementation of the authors themselves, despite the Ledoit-Wolf regularization during the estimation of the precision matrices [7].

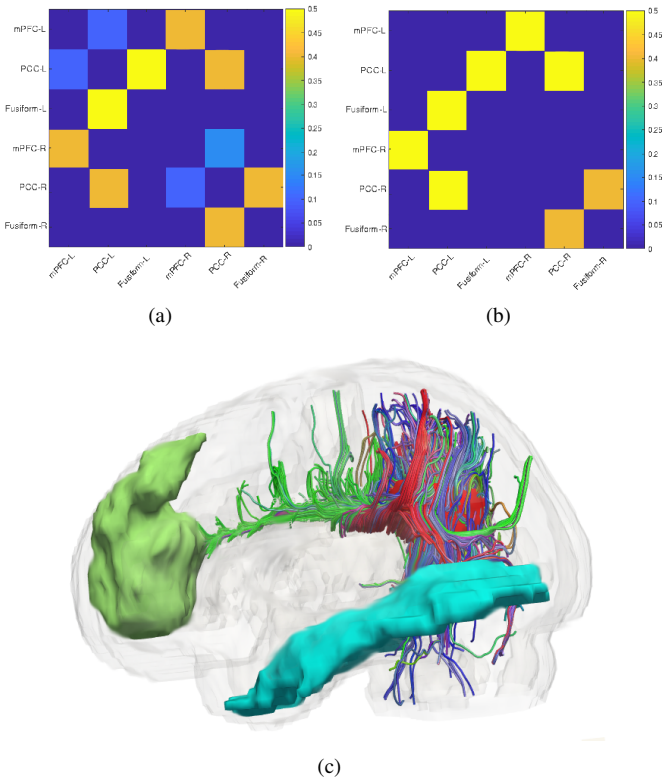


Figure 1: (a) Average effective connectivity matrix resulting after the filtering performed by the MAR model using only 10 random samples (similar results were obtained using the whole dataset), and (b) average effective connectivity matrix computed with the framework described in [7]. Here the connectivity values are shown according to the colorbar depicted. The reported acronyms are 'mPFC' = medial PreFrontal Cortex, 'PCC' = Posterior Cingulate Cortex, 'L'=Left, and 'R'=Right. This reflects structural and functional expected connections depicted in (c): Sagittal example of structural tracts connecting the functional regions of interest selected as related to the DMN. Those ROIs are obtained as a subset of the DMN ROIs of the Power atlas according to the coordinates specified in [32]. The PCCs are depicted as red, mPFCs are depicted as green and the fusiform gyri in cyan. Commisural fibers are omitted.

Effective Brain community detection

Clustering group of nodes within brain networks is relevant since it can show coherent meaningful behavior sub-graphs across multiple samples. A characteristic that exists in other several real world systems including social, biological, and political systems [33]. In neuroscience the related question is on how neural units cluster into densely interconnected groups which provide coordinated activities such as perception, action, and adaptive behaviors [34], moreover the principle of modularity characterize the fundamental organization of human brain functional connectivity during learning [35]. As sparsely interconnected modules allow faster adaptation or evolution of the system in response to changing environmental conditions, the identification of effectively interconnected community simplify the analysis of such modules. In our experiments, the identification of the "effectively" used structural connections has clearly an impact on the subsequent clustering, which is reflected in Figure 2. We resorted to a joint Laplacian diagonalization algorithm [27] to find those effective communities (see supplementary Data and Method section for detail). Laplacian representations have been already used to infer resting-state network from spacially defined oscillations on the human cortex [36]. Indeed, by analyzing the group-wise eigenvalues resulting from the joint Laplacians diagonalization, it has been noted a spectral gap at the 4th and 8th eigenvalues for both structural and effective connectivity matrices, in agreement with previous studies on other datasets [37]. The value $k = 8$ has been chosen to clusters the brain as it better explains the brain known communities. The resulting clustering of the brain regions based on the structural connectome (Figure 2(c)) and on the effective connectome (Figure 2(d)) are slightly different while preserving the overall organization. For this experiment we used the diffusion tensor and fMRI volumes of 200 healthy subjects from the Nathan Kline Institute-Rockland dataset [31].

Evidence is also obtained when analyzing the cluster functional separation, defined as the average ratio between the intra- and inter-cluster cross-correlation (for details see the section Data and Methods). This index has been computed for both structural and effective clustering result. Figure 2(f) shows that cluster functional separation with clusters determined using our CMAR approach is significantly higher when compared with the structural clusters ($p < 0.001$), demonstrating that the effective clusters are also underpinned by the functional connectivity. More specifically, we also devised an analysis to assess whether the clusters obtained form the autoregressive filtered data are more meaningful in relation to the fMRI time-series then the cluster obtained from the structural information. We carried out a block-wise definition of the effective connectivity matrices where one block at a time, defined by the brain regions belonging to a cluster, is used in a CMAR model involving only the relative fMRI series. Then the reconstruction error of the fitted CMAR models for each cluster have been summed up over all clusters and compared each other. The underlying intuition is that partitioning the brain using an effective connectivity information would remove those structural connections which are also meaningless from a functional perspective, at least in the analyzed experimental data. This is the overall message conveyed in Figure 2.

Classification with Effective Connectivity in Autism

Autism spectrum disorder (ASD) is a developmental disorder characterized by repetitive, restricted behavior and deficits in communication and social interactions [38]. Many aspects of ASD are still unclear, and the discrimination from typically developing (TD) subjects by using medical imaging is still a challenging task [39]. Nevertheless, connectome-based classifiers for ASD and TD individuals have been recently introduced also characterizing local and global graph theoretical metrics of structural and functional networks [40]. A recent study has also compared functional connections of individual affected by ASD, Schizophrenia and attention deficit disorders versus TD [39].

To test the descriptive power of the effective connectivity computed using the proposed approach we carried out a classification task using a linear support vector machine (SVM) to discriminate between ASD and TD individuals from the ABIDE-II dataset [38]. We used both the original structural connectomes and the effective connectomes obtained with first and second order MAR models, evaluating how the embedding of functional information on structural connectomes produces a graph better describing the intrinsic characteristic of both functional and structural brain interactions. As features we used the vectorized matrices, and for the case of MAR(2) the two weight matrices were vectorized and concatenated.

We used both the structural and the effective matrices in a leave-one-out classification pipeline to verify on a case-control study whether the embedding of functional information in the structural connectivity can bring some information on the overall brain organization, hence, on the pathology characterization.

We investigated the most significant connections obtained from the SVM discrimination. Those are the weights larger than the 95-th percentile or smaller than the 5-st percentile of a random weight distribution representing the null hypothesis, as depicted in Figure 3(d-f). The null hypothesis SVM weights are obtained by performing 1000 random permutations of the labels of the two groups. Notice that the most discriminating connections of the MAR model are are very different from those determined with structural connectomes. Figure 3 (d-f) show the whole spectrum for the averaged SVM weights respectively for the structural, effective MAR(1) and functional connectomes. It is noticeable that the weights for the MAR cases are larger than those related to the original structural connections, and the weights related to the functional data are more numerous. The detected connections using the structural data comprised spread ipsilateral connections in agreement with previous studies [41, 42]. The detected functional connections are also confirming former results obtained by using another dataset [39]. For instance, the connection between the left frontal pole and the right central gyrus. As well as the left inferior frontal gyrus and the right inferior temporal gyrus. The "effective" detected connections partially resemble the detected structural and the functional connections. Interestingly, these detected features

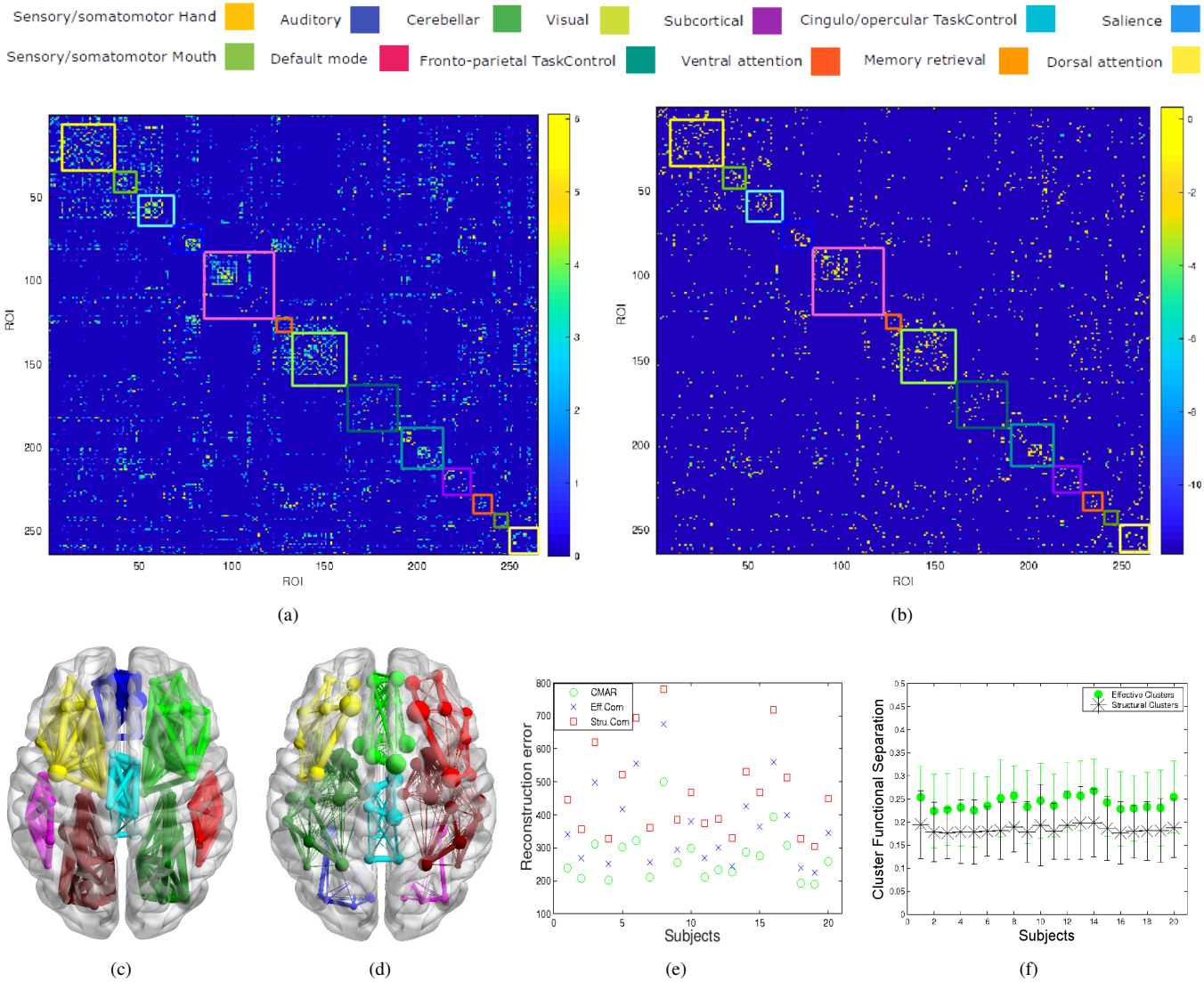


Figure 2: Example of initial structural connectivity matrix (a) and final effective connectivity matrix (b) for the same subject. Each row and column represents one of the 264 ROIs in which the brain is partitioned. Observing the resulting connectivity matrices, as expected, the effective connectivity resulted to be composed of a subset of the structural connections though proportionally stronger. An alternative representations of the graph connections is given in the supplementary material using chord diagrams. Axial view of joint spectral clustering using k=8 is depicted in (c) for the original structural joint eigenspace, and (d) on the joint eigenspace given by effective connectivity matrices. The ROIs highlighed in the connectivity matrices follow the color-code listed on top, where the initial areas are considered as "uncertain" and therefore not enclosed by squares [30]. It has to be noted that the color-code not necessarily match the identified clusters in (c) and (d) as these are the result of clustering. The difference is probably related to the fact that some ROIs are very small compared to the others and therefore merged. The resulting connectivity matrices as the one in (b) and the related clustering (d) are not only sparser than the original, they also lead to better reconstruction of the functional data as shown in the other figures: Subfigure (e) depicts the reconstruction error of CMAR model (only 20 random subjects shown) after converging to the effective connectivity (green circles) or according to block-wise MAR based on the structural communities (red squares) and the effective communities (blue crosses). The lower the better. While subfigure (f) depicts the functional segregation of clusters using the effective communities (green dots) or structural communities (black stars). The higher the better.

are even more in agreement with the discriminant features identified in previous studies. As the ectopic connectivity of precentral and striatal and occipital regions [43], the left superior parietal and right inferior frontal gyrus, and rectus and cuneus region [39]. The classification accuracy between ASD and TD subject by using the identified connections was respectively for the structural, functional, effective connectivity using MAR(1) and MAR(2) 60.57%, 69.1% 70.21% and 72.34%. It was noted that the detected functional connections were larger in number and both negative and positive to compensate the signs given by correlations and anti-correlation.

Conclusions

Despite the temporal resolution limitation of BOLD signals, a model such as these defined by the Granger causality can make explicit the definition of causality in time by the use of MAR models which is easier to validate [17]. This represent a

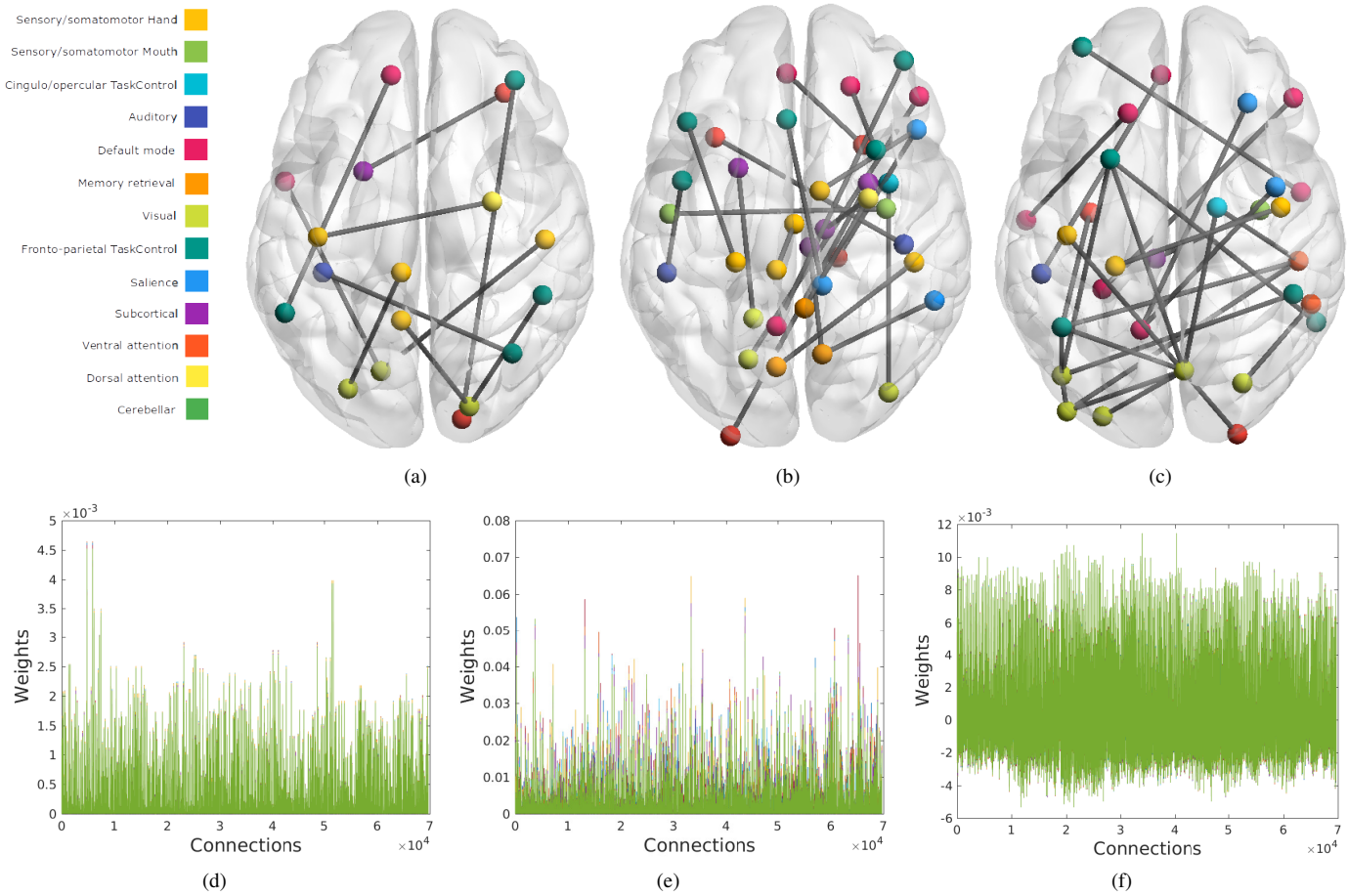


Figure 3: Axial views of most discriminant connections according to the SVM weights (a) original connectomes, (b) MAR(1) effective connectomes, and (c) functional connectomes. (d), (e) and (f) show the average weights respectively for the structural, the MAR(1) effective and functional connectomes. The highlighted structural connections in (a) are mostly ipsilateral and few commisural connections. Most of the detected functional connections (c) are already known [39] as the link between the left frontal pole and the right central gyrus, and the left inferior frontal gyrus and the right inferior temporal gyrus. The "effective" connections in (b) also highlight known connections as the left superior parietal and right Inferior frontal gyrus, and rectus and cuneus region [39].

valid alternative to models based on DCM which are computationally limited.

The effective connectivity inferred by the proposed constrained MAR model highlights a different brain architecture underpinned by both structural and functional connectivity, which is in-line with current neuroscience principles [44]. The DMN experiment showed a proof of concept of the validity of the method where ground-truth is available. Namely, the effective influence can be identified among specific brain regions with which they interact, given by effective connectivity rather than just functional co-activation or structural connection [45].

Thanks to this, the method can lead to new insights into understanding brain effective connections in healthy and pathological subjects. Moreover the proposed constrained model is not limited to fMRI and diffusion volumes, but it can be also applied to different domains where structural and time varying data is applied as two-phonton Ca^{2+} imaging [46], electroencephalography, dynamic contrasts imaging, and metabolic positron emission tomography [47].

[1] David Lee, Oliver Redfern, and Christine Orengo, "Predicting protein function from sequence and structure," *Nature Reviews Molecular Cell Biology*, vol. 8, no. 12, pp. 995–1005, 2007.

[2] James D Watson, Roman A Laskowski, and Janet M Thornton, "Predicting protein function from sequence and structural data," *Current opinion in structural biology*, vol. 15, no. 3, pp. 275–284, 2005.

[3] Seung Wook Oh, Julie A Harris, Lydia Ng, Brent Winslow, Nicholas Cain, Stefan Mihalas, Quanxin Wang, Chris Lau, Leonard Kuan, Alex M Henry, et al., "A mesoscale connectome of the mouse brain," *Nature*, vol. 508, no. 7495, pp. 207–214, 2014.

[4] JL Vincent, GH Patel, MD Fox, AZ Snyder, JT Baker, DC Van Essen, JM Zempel, LH Snyder, M Corbetta, and ME Raichle, "Intrinsic functional architecture in the anaesthetized monkey brain," *Nature*, vol. 447, no. 7140, pp. 83–86, 2007.

[5] Makoto Fukushima, Richard F Betzel, Ye He, Martijn van den Heuvel, Xi-Nian Zuo, and Olaf Sporns, "Structure–function relationships during segregated and integrated network states of human brain functional connectivity," *Brain Structure and Function*, vol. 223, pp. 1091–1106, 2018.

[6] Michael D Greicius, Ben Krasnow, Allan L Reiss, and Vinod Menon, "Functional connectivity in the resting brain: a network analysis of the default mode hypothesis," *Proceedings of the National Academy of Sciences*, vol. 100, no. 1, pp. 253–258, 2003.

[7] Fani Deligianni, Gaël Varoquaux, Bertrand Thirion, David Sharp, Christian Ledig, Robert Leech, and Daniel Rueckert, "A framework for inter-subject prediction of functional connectivity from structural networks," *Medical Imaging, IEEE Transactions on*, vol. 32, pp. 2200–2214, 2013.

[8] Bernard Ng, Gaël Varoquaux, Jean-Baptiste Poline, and Bertrand Thirion, "A novel sparse graphical approach for multimodal brain connectivity inference," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pp. 707–714. Springer, 2012.

[9] Viktor Jirsa, Olaf Sporns, Michael Breakspear, Gustavo Deco, and Anthony Randal McIntosh, "Towards the virtual brain: network modeling of the intact and the damaged brain," *Archives italiennes de biologie*, vol. 148, no. 3, pp. 189–205, 2010.

[10] Max Hinne, Luca Ambrogioni, Ronald J Janssen, Tom Heskes, and Marcel AJ van Gerven, "Structurally-informed bayesian functional connectivity analysis," *Neuroimage*, vol. 86, pp. 294–305, 2014.

[11] Schiefer J, Niederbhl A, Pernice V, Lennartz C, Hennig J, LeVan P, and Rotter S., "From correlation to causation: Estimating effective connectivity from zero-lag covariances of brain signals," *PLoS Comput Biol*, vol. 14, no. 3, pp. e1006056, 2018.

[12] Mario Senden, Niels Reuter, Martijn P van den Heuvel, Rainer Goebel, Gustavo Deco, and Matthieu Gilson, "Task-related effective connectivity reveals that the cortical rich club gates cortex-wide communication," *Human brain mapping*, vol. 39, no. 3, pp. 1246–1262, 2018.

[13] Matthieu Gilson, Gustavo Deco, Karl Friston, Patric Hagmann, Dante Mantini, Viviana Betti, Gian Luca Romani, and Maurizio Corbetta, "Effective connectivity inferred from fMRI transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions," *NeuroImage*, 2017.

[14] Daniel Chicharro and Stefano Panzeri, "Algorithms of causal inference for the analysis of effective connectivity among brain regions," *Frontiers in neuroinformatics*, vol. 8, pp. 64, 2014.

[15] Karl J Friston, "Functional and effective connectivity: a review," *Brain connectivity*, vol. 1, no. 1, pp. 13–36, 2011.

[16] Clive WJ Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, pp. 424–438, 1969.

[17] Rainer Goebel, Alard Roebroeck, Dae-Shik Kim, and Elia Formisano, "Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and Granger causality mapping," *Magnetic resonance imaging*, vol. 21, no. 10, pp. 1251–1261, 2003.

[18] Xiang Li, Kaiming Li, Lei Guo, Chulwoo Lim, and Tianming Liu, "Fiber-centered Granger causality analysis," in *MICCAI 2011*, pp. 251–259. Springer, 2011.

[19] W Liao, J Ding, D Marinazzo, et al., "Small-world directed networks in the human brain: multivariate Granger causality analysis of resting-state fMRI," *NeuroImage*, vol. 54, pp. 2683–2694, 2011.

[20] Daniele Marinazzo, Wei Liao, Huafu Chen, and Sebastiano Stramaglia, "Nonlinear connectivity by granger causality," *Neuroimage*, vol. 58, no. 2, pp. 330–338, 2011.

[21] Rahul Garg, Guillermo A Cecchi, et al., "Full-brain auto-regressive modeling (FARM) using fMRI," *Neuroimage*, vol. 58, pp. 416–441, 2011.

[22] Pedro A Valdés-Sosa, Jose M Sánchez-Bornot, Agustín Lage-Castellanos, Mayrim Vega-Hernández, Jorge Bosch-Bayard, Lester Melie-García, and Erick Canales-Rodríguez, "Estimating brain functional connectivity with sparse multivariate autoregression," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, no. 1457, pp. 969–981, 2005.

[23] Gopikrishna Deshpande, Xiaoping Hu, Randall Stilla, and Krish Sathian, "Effective connectivity during haptic perception: a study using Granger causality analysis of functional magnetic resonance imaging data," *Neuroimage*, vol. 40, no. 4, pp. 1807–1814, 2008.

[24] Qing Jiao, Guangming Lu, Zhiqiang Zhang, Yuan Zhong, Zhengge Wang, Yongxin Guo, Kai Li, Mingzhou Ding, and Yijun Liu, "Granger causal influence predicts BOLD activity levels in the default mode network," *Human brain mapping*, vol. 32, no. 1, pp. 154–161, 2011.

[25] Dajiang Zhu, Xiang Li, Xi Jiang, Hanbo Chen, Dinggang Shen, and Tianming Liu, "Exploring high-order functional interactions via structurally-weighted lasso models," in *Information Processing in Medical Imaging*. Springer, 2013, pp. 13–24.

[26] Hanbo Chen, Tao Liu, Yu Zhao, Tuo Zhang, Yujie Li, Meng Li, Hongmiao Zhang, Hui Kuang, Lei Guo, Joe Z Tsien, et al., "Optimization of large-scale mouse brain connectome via joint evaluation of dti and neuron tracing data," *NeuroImage*, vol. 115, pp. 202–213, 2015.

[27] Luca Dodero, Alessandro Gozzi, Adam Liska, Vittorio Murino, and Diego Sona, "Group-wise functional community detection through joint Laplacian diagonalization," in *MICCAI 2014*, pp. 708–715. Springer, 2014.

[28] Alessandro Crimi, Luca Dodero, Vittorio Murino, and Diego Sona, "Case-control discrimination through effective brain connectivity," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2017.

[29] Hirotugu Akaike, "A new look at the statistical model identification," *IEEE transactions on automatic control*, vol. 19, pp. 716–723, 1974.

[30] Jonathan Power, Alexander Cohen, Steven Nelson, Gagan Wig, Kelly Barnes, Jessica Church, Alecia Vogel, Timothy Laumann, Fran Miezin, Bradley Schlaggar, et al., "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.

[31] KB Nooner et al., "The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry. front. neurosci. 6, 152," 2012.

[32] WR Shirer, S Ryali, E Rykhlevskaia, V Menon, and MD Greicius, "Decoding subject-driven cognitive states with whole-brain connectivity patterns," *Cerebral cortex*, vol. 22, no. 1, pp. 158–165, 2012.

[33] Santo Fortunato, "Community detection in graphs," *Physics reports*, vol. 486, no. 3, pp. 75–174, 2010.

[34] David Meunier, Renaud Lambiotte, and Edward T Bullmore, "Modular and hierarchically modular organization of brain networks," *Frontiers in neuroscience*, vol. 4, 2010.

[35] Danielle S Bassett, Nicholas F Wymbs, Mason A Porter, Peter J Mucha, Jean M Carlson, and Scott T Grafton, "Dynamic reconfiguration of human brain networks during learning," *Proceedings of the National Academy of Sciences*, vol. 108, no. 18, pp. 7641–7646, 2011.

[36] Selen Atasoy, Isaac Donnelly, and Joel Pearson, "Human brain networks function in connectome-specific harmonic waves," *Nature communications*, vol. 7, pp. 10340, 2016.

[37] Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J Honey, Van J Wedeen, and Olaf Sporns, "Mapping the structural core of human cerebral cortex," *PLoS Biol*, vol. 6, no. 7, pp. e159, 2008.

[38] I Fishman, M Datko, Y Cabrera, RA Carper, and RA Müller, "Reduced integration and differentiation of the imitation network in autism: A combined functional connectivity magnetic resonance imaging and diffusion-weighted imaging study," *Annals of neurology*, vol. 78, no. 6, pp. 958–969, 2015.

[39] N Yahata, J Morimoto, R Hashimoto, G Lisi, K Shibata, Y Kawakubo, et al., "A small number of abnormal brain connections predicts adult autism spectrum disorder," *Nature communications*, vol. 7, 2016.

[40] JD Rudie, JA Brown, D Beck-Pancer, LM Hernandez, EL Dennis, PM Thompson, SY Bookheimer, and M Dapretto, "Altered functional and structural brain network organization in autism," *NeuroImage: clinical*, vol. 2, pp. 79–94, 2013.

[41] Hai Li, Zhong Xue, Timothy M Ellmore, Richard E Frye, and Stephen TC Wong, "Network-based analysis reveals stronger local diffusion-based connectivity and different correlations with oral language skills in brains of children with high functioning autism spectrum disorders," *Human brain mapping*, vol. 35, no. 2, pp. 396–413, 2014.

[42] Fan Zhang, Peter Savadjiev, Weidong Cai, Yang Song, Yogesh Rathi, Birkan Tunç, Drew Parker, Tina Kapur, Robert T Schultz, Nikos Makris, et al., "Whole brain white matter connectivity analysis using machine learning: an application to autism," *NeuroImage*, 2017.

[43] Sarah K Noonan, Frank Haist, and Ralph-Axel Müller, "Aberrant functional connectivity in autism: evidence from low-frequency BOLD signal fluctuations," *Brain research*, vol. 1262, pp. 48–63, 2009.

[44] Richard E Passingham, Klaas E Stephan, and Rolf Kötter, "The anatomical basis of functional localization in the cortex," *Nature Reviews Neuroscience*, vol. 3, no. 8, pp. 606–616, 2002.

[45] Devarajan Sridharan, Daniel J Levitin, and Vinod Menon, "A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks," *Proceedings of the National Academy of Sciences*, vol. 105, no. 34, pp. 12569–12574, 2008.

[46] Alireza Sheikhattar, Sina Miran, Ji Liu, Jonathan B Fritz, Shihab A Shamma, Patrick O Kanold, and Behtash Babadi, "Extracting neuronal functional network dynamics via adaptive granger causality analysis," *Proceedings of the National Academy of Sciences*, p. 201718154, 2018.

[47] Harald Hampel, Gordon Wilcock, Sandrine Andrieu, Paul Aisen, Kaj Blennow, K Broich, Maria Carrillo, Nick C Fox, Giovanni B Frisoni, Maria Isaac, et al., "Biomarkers for alzheimer's disease therapeutic trials," *Progress in neurobiology*, vol. 95, no. 4, pp. 579–593, 2011.

[48] Ziad S Saad et al., "Correcting brain-wide correlation differences in resting-state fMRI," *Brain connectivity*, vol. 3, no. 4, pp. 339–352, 2013.
[49] Kevin Murphy and Michael D Fox, "Towards a consensus regarding global signal regression for resting state functional connectivity mri," *Neuroimage*, vol. 154, pp. 169–173, 2017.
[50] Jonathan D Power, Kelly A Barnes, Abraham Z Snyder, Bradley L Schlaggar, and Steven E Petersen, "Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion," *Neuroimage*, vol. 59, no. 3, pp. 2142–2154, 2012.
[51] Evan M Gordon, Timothy O Laumann, Babatunde Adeyemo, Jeremy F Huckins, William M Kelley, and Steven E Petersen, "Generation and evaluation of a cortical area parcellation from resting-state correlations," *Cerebral cortex*, p. bhu239, 2014.
[52] Eleftherios Garyfallidis, Matthew Brett, Bagrat Amirbekian, Ariel Rokem, Stefan Van Der Walt, Maxime Descoteaux, Ian Nimmo-Smith, and Dipy Contributors, "Dipy, a library for the analysis of diffusion MRI data," *Frontiers in neuroinformatics*, vol. 8, 2014.
[53] Tiago P Peixoto, "Model selection and hypothesis testing for large-scale network models with overlapping groups," *Physical Review X*, vol. 5, no. 1, pp. 011033, 2015.
[54] Alessandro Crimi, Luca Dodero, Vittorio Murino, and Diego Sona, "Effective brain connectivity through a constrained autoregressive model," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp. 140–147.
[55] Fragkiskos D Malliaros and Michalis Vazirgiannis, "Clustering and community detection in directed networks: A survey," *Physics Reports*, vol. 533, no. 4, pp. 95–142, 2013.

ONLINE CONTENT

Data and Methods

Nathan Kline Institute-Rockland dataset

To estimate the connectome communities, a large dataset obtained by the F1000 connectomes initiative was used. This comprised 200 right-handed healthy subjects from the Nathan Kline Institute-Rockland dataset [31], and publicly available at the url *http://fcon_1000.projects.nitrc.org/*. For each subject, resting state fMRI (rs-fMRI), diffusion tensor imaging (DTI) and T1 have been acquired by using a Siemens (Munich, Germany) Magnetom scanner and co-registered. FMRI data were acquired using a 3 Tesla scanner, with TR/TE times as 1.4s/30ms, flip angle 65°, and isotropic voxel size of 2 mm; resulting in resting-state time series 10 minutes long, where subjects were asked to keep the eyes open. DTI volumes were acquired with a 1.5 Tesla scanner by using 35 gradient directions and TR/TE 2.4ms/85ms, flip angle 90°, and isotropic voxel-size of 2.5 mm. The T1 weighted MRI data were acquired with the same scanner, using as TR/TE times 1.1s/4.38ms, flip angle 15°, and isotropic voxel-size of 1 mm.

Autism Brain Imaging Data Exchange II dataset

To perform the case-control classification task, the experiments have been performed on the San Diego State University cohort of the ABIDE-II dataset publicly available [38] at the url *http://fcon_1000.projects.nitrc.org/indi/abide/abide_II.html*. Excluding subjects with excessive motion or other imaging artifacts, the final dataset included 26 ASD and 21 TD subjects. For each subject, rs-fMRI, DTI and T1-weighted have been acquired and co-registered. Imaging data were acquired on a GE (Milwaukee, WI) 3T MR750 scanner. For a detailed description of the experimental protocols refer to [38]. Briefly, rs-fMRI volumes were acquired using a single-shot gradient-recalled, echo-planar pulse sequence, in one 6:10- minute eyes-open scan consisting of 185 whole brain volumes at TR/TE = 2s/30ms, flip angle 90°, and isotropic 3.4 mm voxel-size. The DTI volumes were acquired with a dual spin echo excitation using TR/TE = 8.5s/84.9ms, flip angle 90°, and $1.88\times1.88\times2mm^3$ voxel-size. Total diffusion-weighted scan time was about 9 minutes. T1-weighted inversion recovery spoiled gradient echo sequence were acquired at TR/TE = 8.1s/3,172ms, flip angle 8°, and isotropic 1 mm voxel-size.

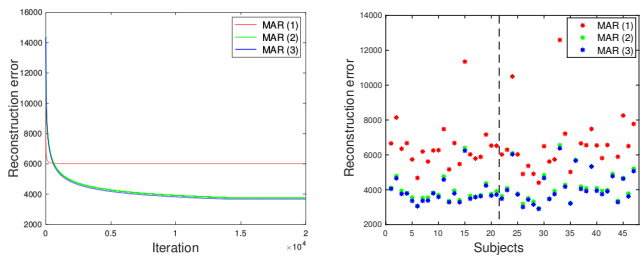
Pre-processing and Connectome construction

FMRI data have been pre-processed according to a standard pipeline: motion correction, mean intensity subtraction, pass-band filtering with cutoff frequencies of [0.005-0.1 Hz] and skull removal. To account for potential noise from physiological processes such as cardiac and respiratory fluctuations, nine covariates of no interest have been identified for inclusion in our analyses [48]. Spatial smoothing has been performed only while averaging the voxels for the ROIs. Global signal has not being regressed as in the given settings it would just add spurious anticorrelations [49]. To further reduce the effects of motion, compensation for frame-wise displacement has been carried out [50]. For both dataset, the DTI data have been cleaned

by eddy current interference, and the skull has been removed. Linear registration has been applied between the Power atlas [30] and the T1 reference volume by using linear registration with 12 degrees of freedom. The Power atlas defines 264 functional ROIs. Each ROI is a 5mm sphere around a voxel of peak significant activity during tasks. This atlas is considered one of the most accurate for functional mapping [51].

Tractographies for all subjects have been generated processing DTI data with the Python library Dipy [52]. In particular, a deterministic algorithm called Euler Delta Crossings has been used stemming from 2,000,000 seed-points and stopping when the fractional anisotropy was smaller than < 0.1. Tracts shorter than 30 mm or in which a sharp angle occurred have been discarded. The final result yielded to about 250,000 fibers. To construct the connectome, the graph nodes have been determined using the 264 regions in the Power atlas. Specifically, the structural connectome has been built counting the number of tracts connecting two regions, for any pair of regions. The same regions have been used to compute the functional activity just averaging the voxel activity in each area.

The aim of the proposed method is to enforces the association between structural brain connectivity and functional brain activation. To obtain this result we resort to a multivariate autoregressive model properly modified in order to allow for the estimation of the temporal brain activation biased by the structural connectivity.



(a) (b)

Figure A.1: (a) Example of convergence of the reconstruction error for one subject using different orders of the MAR model. (b) Final reconstruction errors for different orders. The stars at the left and right of the line depict respectively the errors for the control (TD) and case (ASD) subjects from the classification experiment.

Structurally constrained Autoregressive model

A multivariate autoregressive model of order n (MAR(n)) is a stochastic process defining r variables $\mathbf{y}(t)$ as linearly dependent on their own previous values and on a stochastic term:$\mathbf{y}(t) = \sum_{i=1}^{n} \mathbf{A}_i \cdot \mathbf{y}(t-i) + \epsilon$, where the r-by-r coefficient matrices $\mathbf{A}_i$ are the model parameters and ϵ is the additive Gaussian noise. The weight matrices $\mathbf{A}_i$ describe the linear dependencies between the r timeseries $[y_1, \ldots, y_r]$, therefore, they can be intended as functional connectivity matrix at different time lags. The proposed method aims at merging functional and structural connectivity, leaving on the structural connections used by functional series.

Given an initial structural connectivity matrix $\mathbf{A}_{\text{init}}$ and the functional signal in a column-vector $\mathbf{y}$ for all r regions of interest and for all T time-frames, an effective connectivity matrix can be determined minimizing the reconstruction error in equation (3). More specifically, to fit the model parameters we used a gradient descent approach. The direction of the gradient can be computed for all matrices $\mathbf{A}_j$ of a specified order as follows:

$$\frac{\partial \mathbf{E}}{\partial \mathbf{A_j}} = -\sum_{t=n}^{T} \left[\mathbf{y}(t) - \sum_{i=1}^{n} (\mathbf{A}_i \odot \mathbf{B}) \cdot \mathbf{y}(t-i) \right] \cdot \mathbf{y}'(t-j) \odot \mathbf{B}. \quad \text{(A.1)}$$

and the update rule is therefore $\mathbf{A}_j = \left(\mathbf{A}_j + \eta \cdot \partial \mathbf{E} / \partial \mathbf{A}_j\right)$, where η is the learning rate.

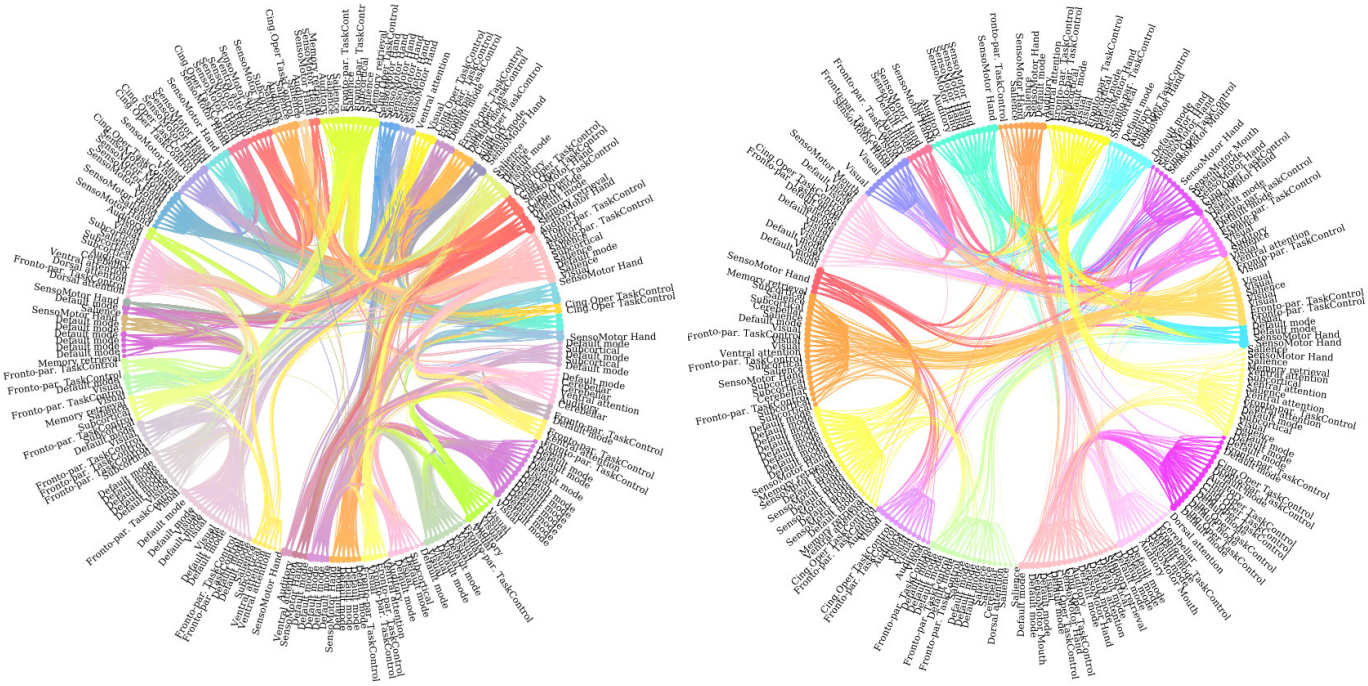
We generated the effective connectivity matrix for each subject as outlined in Section *Structurally constrained Autoregressive model*, ranging over different MAR orders. The learning rate parameter η was set empirically to optimize the convergence rate [54]. Figure A.2 depicts the initial structural and resulting effective connectivity as chord diagrams which are an alternative representation of graphs to adjacency matrices. Figure A.1 highlights the beneficial effect of using more time points in the MAR model for the autism dataset [38]. As expected, the higher the order of MAR model the better was the reconstruction error of the signal, though at some point the improvements were not worthwhile. In particular, it is visible that the reconstruction error of MAR(2) is always significantly smaller than MAR(1). However, in all our experiments, going beyond order 2 did not improve the reconstruction error compared to MAR(2). This is probably related to the type of signal. Indeed BOLD signal in used fMRI recordings has a sampling rate that compared to the underlying brain activity makes it useless to go beyond two time steps.

Effective Brain Community Detection

To assess whether the proposed method produces an effective connectivity information characterizing the structural connectivity enriched with functional information, a community detection analysis has been performed using a group-wise graph clustering algorithm recently proposed in [27] both on the set of structural connectivity matrices and on the effective connectivity matrices.

Given a set of connectivity matrices $\mathcal{W} = \{\mathbf{W}_i\}$ representing undirected weighted graphs with positive weights, the normalized graph Laplacian is built as $\mathbf{L_i} = \mathbf{D_i}^{-\frac{1}{2}}(\mathbf{D_i} - \mathbf{W}_i)\mathbf{D_i}^{-\frac{1}{2}}$, where $\mathbf{D_i}$ is the diagonal degree matrix of $\mathbf{W}_i$. However, in general, the connectivity matrices resulting from the above CMAR model computed for each subject are asymmetric (i.e., edges are directed), thus, they have been converted to undirected graphs aiming at maintaining the properties of the original graphs estimated from CMAR. To this aim, a symmetrization based on random walk was applied [55].

More specifically, given a directed graph $\mathbf{M}$, the transition matrix of the random walk can be defined as $\mathbf{P} = \mathbf{D_{out}^{-1}}\mathbf{M}$, where $\mathbf{D_{out}}$ is a diagonal matrix built using nodes' out-degree. The symmetric graph can be therefore defined as $\mathbf{M_{sym}} = \frac{1}{2}(\mathbf{\Pi P} + \mathbf{P'\Pi})$, where $\mathbf{\Pi}$ is a the diagonal matrix which defines



(a) (b)

Figure A.2: (a) Example of initial structural and (b) resulting effective connectivity depicted as chord diagrams highlighting the connections between ROIs. Here connections between ROIs are represented by chords bent according to visual algorithm described in [53] where the color-code does not represent ROIs but connections between them. The ROIs defined as uncertain in [30] are not labeled.

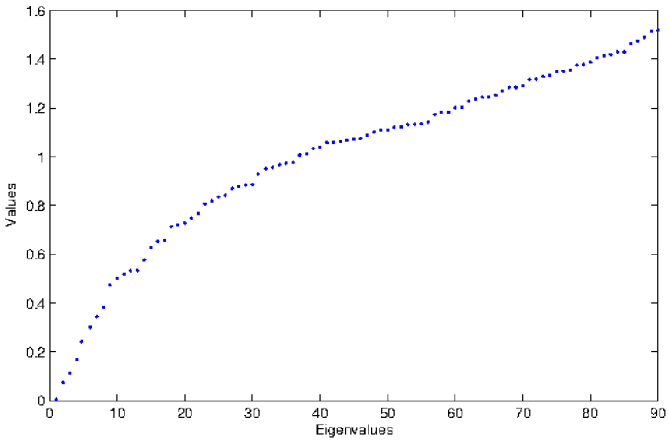


Figure A.3: Eigengap that allows to identify the present clusters for the aforementioned experiment. It is visible that the largest gap is present at the 8th eigenvalue.

the probability of a walker to stay in each node in a stationary distribution, defined as $\mathbf{\Pi} = \frac{d_{\text{out}}}{m}$, where d_{out} is the vector of the out-degree of each node and m is the number of nodes. Thanks to this new representation, the pipeline described by the authors in [27] can be applied, generating the Normalized Graph Laplacians for each subject, performing the joint diagonalization of multiple Laplacians to find a unique eigenspace and, finally, applying spectral clustering on the smallest joint eigenvectors.

In order to decide the number of clusters, as usual in spectral clustering, we can look at the spectral gap of the mean approximated eigenvalues. Figure A.3 depicts mean approximated eigenvalues where a gap between the 8th and 9th values is visible.

The cluster functional separation (CFS) was defined as the average ratio between the intra- and inter-cluster cross-correlation. as follows:

$$CFS = \frac{1}{k}\sum_{s=1}^{k} \frac{\sum_{i<j\in C_s} w_{ij}}{\sum_{i<j\in C_s} w_{ij} + \sum_{i\in C_s}\sum_{j\in C_t\neq C_s} w_{ij}} \tag{A.2}$$

where w_{ij} is the functional cross-correlation of the time-series for nodes i and j.