

# Effective Brain Connectivity Through a Constrained Autoregressive Model

Alessandro Crimi<sup>1</sup>(✉), Luca Dodero<sup>1</sup>, Vittorio Murino<sup>1,2</sup>, and Diego Sona<sup>1,3</sup>

<sup>1</sup> Pattern Analysis and Computer Vision, Istituto Italiano di Tecnologia, Genoa, Italy  
{alessandro.crimi, luca.dodero, vittorio.murino, diego.sona}@iit.it

<sup>2</sup> Department of Computer Science, University of Verona, Verona, Italy

<sup>3</sup> Neuroinformatics Laboratory, Fondazione Bruno Kessler, Trento, Italy

**Abstract.** Integration of functional and structural brain connectivity is a topic receiving growing attention in the research community. Their fusion can, in fact, shed new light on brain functions. Targeting this issue, the manuscript proposes a constrained autoregressive model allowing to generate an “effective” connectivity matrix that model the structural connectivity integrating the functional activity. 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 network across multiple subjects. Results showed that using the effective connectivity the resulting clusters better describe the functional interactions of different regions while maintaining the structural organization.

**Keywords:** Autoregressive model · Spectral clustering · Connectome · Effective connectivity · Community detection · Brain imaging

## 1 Introduction

It is now well established in neuroscience 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 organization explaining the complexity of brain function. This principle is however difficult to prove and a possible approach is to investigate it through functional connectivity analysis, which is usually based on the determination of functional correlations.

The main limitation of functional connectivity analysis is that structural information is not used at all, preventing any interpretation 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 functions. While in [19]

it has been shown there is a significant overlap between neuroanatomical connections and correlations of fMRI signals, it is yet to be understood how whole-brain network interactions relate during specific tasks or at rest, where fundamental signals have been suggested to play a key role [4].

It has also been shown, for example, that the intensity of functional activity can be predicted by the strength of structural connections, suggesting that it is possible to predict resting-state functional activity from structural information [11]. On the same line, a predictive framework based on multiple sparse linear regression has been recently used to predict functional series from structural data [2], and a model called the “virtual brain” has been designed to simulate brain activity in injured and healthy subjects [12]. With the aim of highlighting the relationships between functional and structural connectivity, a different approach relies on a Bayesian framework to estimate the functional connectivity using a structural graph as a prior [10] has been proposed. 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.

An alternative schema adopted to infer the functional directed connections from brain activity is based on methods investigating the functional causality like the Dynamical Causal Model (DCM) [5] and the Granger Causality (GC) [8]. 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 being this 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. GC 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 fMRI data, it has been also used to identify the dynamics of Blood-Oxygen-Level Dependent (BOLD) signal flow between brain regions [7, 14]. The advantage of GC is that it is based on a Multivariate Autoregressive model (MAR), a random process specifying that the variables depend linearly on their own previous values and on a stochastic term. Thank to this, GC 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. Moreover, this approach usually does not consider the structural connectivity, unless considering one fiber/connection at a time with univariate analysis [13].

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. Besides, the advantages of the proposed method are that only the structural connections justified by the functionality survive, and it can refine possible false positive connections [1]. The validation of the proposed method has been performed testing the quality of results in a typical brain community detection framework, using a recently developed technique based on group-wise spectral clustering [3]. 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.

## 2 Method

The aim of the proposed method is to reinforce 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.

### 2.1 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  $\epsilon$ :  $\mathbf{y}(t) = \sum_{i=1}^n \mathbf{A}_i \cdot \mathbf{y}(t-i) + \epsilon$ , where the coefficient matrices  $\mathbf{A}_i$  are the model parameters. These can be estimated by minimizing the discrepancy between the observed and the implied covariance or constrained by prior knowledge, though this is cumbersome for detailed brain-wise tasks.

Given a structural connectivity matrix  $\mathbf{A}_{\text{init}}$ , the functional signal in a column-vector  $\mathbf{y}$  for all  $r$  regions of interest and all  $T$  frames, an improved connectivity matrix can be determined minimizing a reconstruction error of a MAR(1) model. Indeed, to fit the model parameters  $\mathbf{A}$  we used a gradient descent approach minimizing the reconstruction error in a least square fashion:

$$E = \frac{1}{2} \sum_{t=0}^{T-1} \|\mathbf{A} \cdot \mathbf{y}_t - \mathbf{y}_{t+1}\|_2^2, \quad (1)$$

with the direction of the gradient computed regarding  $\mathbf{A}$  as follows:

$$\nabla \mathbf{E} = \sum_{t=0}^{T-1} (\mathbf{A} \cdot \mathbf{y}_t - \mathbf{y}_{t+1}) \cdot \mathbf{y}'_t, \quad (2)$$

where  $()'$  denotes the transpose. To introduce the structural bias into the model, the parameters fitting is constrained by the structural information in the update rule:

$$\mathbf{A}_{\text{new}} = (\mathbf{A}_{\text{old}} + \eta \nabla \mathbf{E}) \odot \mathbf{B}, \quad (3)$$

where  $\eta$  is the learning rate,  $\odot$  is the Hadamard or element-wise product, and  $\mathbf{B}$  is a matrix of the same size of  $\mathbf{A}_{\text{init}}$  with each element  $b_{ij}$  defined as

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

In this manner, the lack of connections among certain regions in the initial matrix  $\mathbf{A}_{\text{init}}$  is reinforced at each iteration keeping the connection originally set to zero when the gradient descent has introduced some non-null values in them. Thus,  $\mathbf{B}$  and  $\mathbf{A}_{\text{init}}$  encode the prior structural connectivity that reinforce the relationship between functional and structural data. During the gradient descent, it is possible that some values of the matrix  $\mathbf{A}_{\text{new}}$  become negative which might be meaningless in terms of causality. Therefore, at each iteration those values are set to zero to correct the descent. Setting to zero negative values at each iteration in the matrix  $\mathbf{A}_{\text{new}}$  is a common way to enforce non-negativity during the learning. Clearly, this has an effect on accuracy as the solution is sub-optimal with respect to an unconstrained fitting. On the other side, the need for non-negative coefficients is motivated by the fact that negative causalities cannot be interpreted.

The approach can be easily generalized to higher order MAR, where a different matrix  $\mathbf{A}_i$  has to be optimized for each order. Choosing the right model-order is a trade-off between optimizing the variance and the model complexity. In our experiments to have a direct validation of the method, we limited the model to the first order.

## 2.2 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 [3] 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 [15].

More specifically, given a directed graph  $\mathbf{M}$ , the transition matrix of the random walk can be defined as  $\mathbf{P} = \mathbf{D}_{\text{out}}^{-1}\mathbf{M}$ , where  $\mathbf{D}_{\text{out}}$  is a diagonal matrix built using nodes' out-degree. The symmetric graph can be therefore defined as  $\mathbf{M}_{\text{sym}} = \frac{1}{2}(\mathbf{P}\mathbf{P} + \mathbf{P}'\mathbf{P})$ , where  $\mathbf{P}$  is a the diagonal matrix which defines the probability of a walker to stay in each node in a stationary distribution, defined as  $\mathbf{P} = \frac{d_{\text{out}}}{m}$ , where  $d_{\text{out}}$  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 [3] 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.

### 3 Data and Experimental Settings

All experiments have been performed on 20 right-handed healthy subjects from the Nathan Kline Institute-Rockland dataset [16]. For each subject, fMRI, DTI and T1 have been acquired and co-registered. FMRI data were acquired using a 3 T scanner, with TR/TE times as 1.4 s/30 ms, flip angle  $65^\circ$ , and isotropic voxel size of 2 mm; resulting in resting-state time series 10 min long, where subjects were asked to keep the eyes open. DTI volumes were acquired with a 1.5 Tesla scanner and isotropic voxel-size of 2.5 mm using 35 gradient directions. The T1 weighted MRI data were acquired with the same scanner, using as TR/TE times 1.1 s/4.38 ms, flip angle  $15^\circ$ , and isotropic voxel-size of 1 mm.

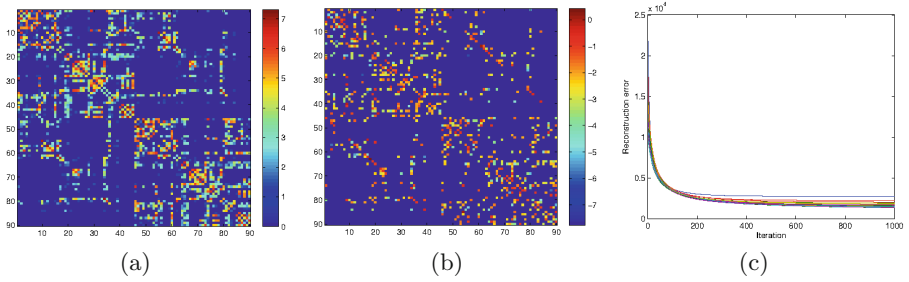
#### 3.1 Pre-processing and Connectome Construction

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 [18]. To further reduce the effects of motion, compensation for frame-wise displacement has been carried out [17]. Eddy current correction and skull stripping have been carried out as the pre-processing for the DTI data. Linear registration has been applied between the AAL atlas and the T1 reference volume by using linear registration with 12 degrees of freedom.

Tractographies for all subjects have been generated processing DTI data with the Python library Dipy [6]. 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 90 regions in the AAL 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 averaged functional time series from the voxels in each region.

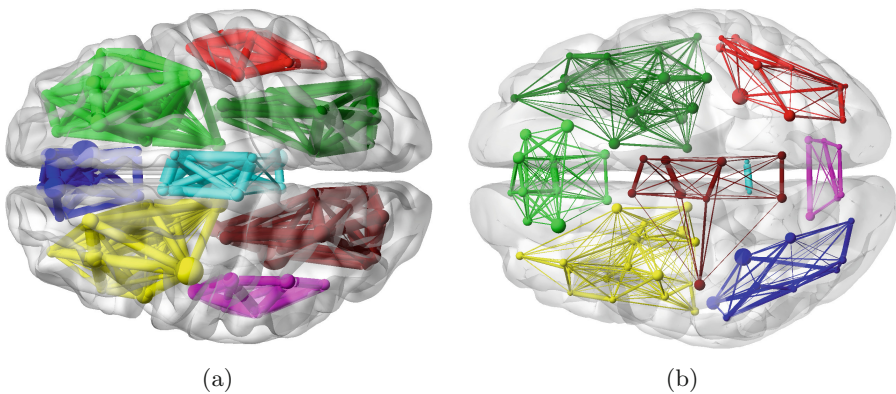
## 4 Results and Discussions

Figure 1 depicts an example from one subject of (a) the original structural connectivity matrix and (b) the effective connectivity resulting from autoregressive

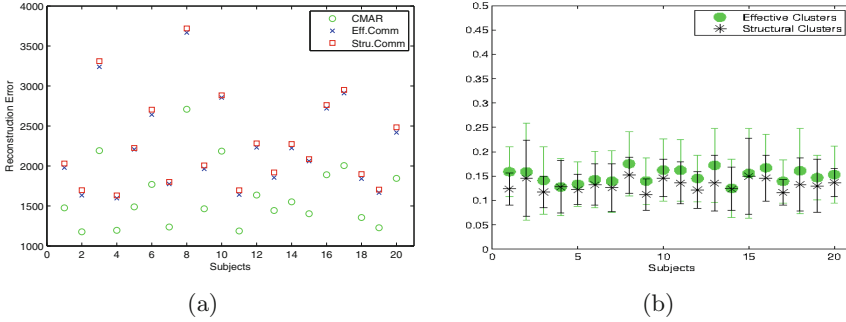


**Fig. 1.** Example of adjacency matrix for one subject plotted in logscale: (a) initial matrix obtained from the tractography; (b) effective matrix obtained with the proposed autoregressive model. The convergence process is shown in subfigure (c) as a decrease of error defined in Eq. (1), where each color line is a different subject.

model filtering. The figures highlight that some structural connections which are not “used” by the resting-state functional data are canceled out. This has clearly an effect on the subsequent clustering, which is shown in Fig. 2. 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 [9]. The value  $k = 8$  has been chosen to cluster the brain as it better explains the brain known communities. The resulting clustering of the brain regions based on the structural connectome (Fig. 2(a)) and on the effective connectome (Fig. 2(b)) are slightly different while preserving the overall organization. Regarding convergence, the gradient descent finds a sub-optimal solution by definition. The zeroing step makes convergence more cumbersome but does not prevent from reaching a minimum as shown in Fig. 1(c).



**Fig. 2.** Axial view of joint spectral clustering using  $k=8$  on (a) the original structural joint eigenspace, and (b) on the joint eigenspace given by effective connectivity matrices.



**Fig. 3.** (a) Reconstruction error of CMAR model 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. (b) Functional segregation of clusters using the effective communities (green dots) or structural communities (black stars). The higher the better.

We also devised an analysis to assess whether the clusters obtained from the autoregressive filtered data are more meaningful in relation to the fMRI time-series than 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.

The reconstruction error per subject in Fig. 3(a) 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 than the communities obtained from the structural connectivity only. Similar evidence is obtained when analyzing the cluster functional separation (CFS), defined as the average ratio between the intra- and inter-cluster cross-correlation as follows:

$$CFS = \frac{1}{k} \sum_{s=1}^k \frac{\sum_{i < j \in C_s} w_{ij}}{\sum_{i < j \in C_s} w_{ij} + \sum_{i \in C_s} \sum_{j \in C_t \neq C_s} w_{ij}} \quad (4)$$

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

This index has been computed for both structural and effective clustering result. Figure 3(b) shows that CFS 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. Although, larger datasets experiments are required.



## 5 Conclusions

The effective connectivity inferred by the proposed CMAR model highlights a different brain architecture underpinned by both structural and functional connectivity. Thanks to this, the method can lead to new insights into understanding brain effective connections in healthy and pathological subjects.

## References

1. Chen, H., et al.: Optimization of large-scale mouse brain connectome via joint evaluation of DTI and neuron tracing data. *NeuroImage* **115**, 202–213 (2015)
2. Deligianni, F., et al.: A framework for inter-subject prediction of functional connectivity from structural networks. *IEEE TMI* **32**(12), 2200–2214 (2013)
3. Doderer, L., Gozzi, A., Liska, A., Murino, V., Sona, D.: Group-wise functional community detection through joint Laplacian diagonalization. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) *MICCAI 2014*. LNCS, vol. 8674, pp. 708–715. Springer, Heidelberg (2014). doi:[10.1007/978-3-319-10470-6\\_88](https://doi.org/10.1007/978-3-319-10470-6_88)
4. Fox, M.D., et al.: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *PNAS* **102**(27), 9673–9678 (2005)
5. Friston, K.J.: Functional and effective connectivity: a review. *Brain Connect.* **1**(1), 13–36 (2011)
6. Garyfallidis, E., et al.: Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinformatics* **8**, 8 (2014)
7. Goebel, R., et al.: Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magn. Reson. Imaging* **21**(10), 1251–1261 (2003)
8. Granger, C.: Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**(3), 424–438 (1969)
9. Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., et al.: Mapping the structural core of human cerebral cortex. *PLoS Biol.* **6**(7), e159 (2008)
10. Hinne, M., Ambrogioni, L., Janssen, R.J., Heskes, T., et al.: Structurally-informed Bayesian functional connectivity analysis. *NeuroImage* **86**, 294–305 (2014)
11. Honey, C., et al.: Predicting human resting-state functional connectivity from structural connectivity. *PNAS* **106**(6), 2035–2040 (2009)
12. Jirsa, V., et al.: Towards the virtual brain: network modeling of the intact and the damaged brain. *Arch. Ital. Biol.* **148**(3), 189–205 (2010)
13. Li, X., Li, K., Guo, L., Lim, C., Liu, T.: Fiber-centered granger causality analysis. In: Fichtinger, G., Martel, A., Peters, T. (eds.) *MICCAI 2011*. LNCS, vol. 6892, pp. 251–259. Springer, Heidelberg (2011). doi:[10.1007/978-3-642-23629-7\\_31](https://doi.org/10.1007/978-3-642-23629-7_31)
14. Liao, W., et al.: Small-world directed networks in the human brain: multivariate Granger causality analysis of resting-state fMRI. *NeuroImage* **54**, 2683–2694 (2011)
15. Malliaros, F.D., Vazirgiannis, M.: Clustering and community detection in directed networks: a survey. *Phys. Rep.* **533**(4), 95–142 (2013)
16. Nooner, K., et al.: The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry. *Front. Neurosci.* **6**, 152 (2012)
17. Power, J.D., et al.: Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* **59**(3), 2142–2154 (2012)
18. Saad, Z.S., et al.: Correcting brain-wide correlation differences in resting-state fMRI. *Brain Connect.* **3**(4), 339–352 (2013)
19. Vincent, J., et al.: Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* **447**(7140), 83–86 (2007)