ELSEVIER

Using partial correlation to enhance structural equation modeling of functional MRI data

Guillaume Marrelec[a,b,c,*], Barry Horwitz[d], Jieun Kim[d], Mélanie Pélégrini-Issac[a,b],
Habib Benali[a,b,c], Julien Doyon[a,b,c]

[a]Inserm, U678, F-75013 Paris, France
[b]Université Pierre et Marie Curie, Faculté de médecine Pitié-Salpêtrière, F-75013 Paris, France
[c]Université de Montréal, UNF, MIC, Montréal, Canada H3W 1W5
[d]Brain Imaging and Modeling Section, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, MD 20892, USA

Abstract

In functional magnetic resonance imaging (fMRI) data analysis, effective connectivity investigates the influence that brain regions exert on one another. Structural equation modeling (SEM) has been the main approach to examine effective connectivity. In this paper, we propose a method that, given a set of regions, performs partial correlation analysis. This method provides an approach to effective connectivity that is data driven, in the sense that it does not require any prior information regarding the anatomical or functional connections. To demonstrate the practical relevance of partial correlation analysis for effective connectivity investigation, we reanalyzed data previously published [Bullmore, Horwitz, Honey, Brammer, Williams, Sharma, 2000. How good is good enough in path analysis of fMRI data? NeuroImage 11, 289–301]. Specifically, we show that partial correlation analysis can serve several purposes. In a pre-processing step, it can hint at which effective connections are structuring the interactions and which have little influence on the pattern of connectivity. As a post-processing step, it can be used both as a simple and visual way to check the validity of SEM optimization algorithms and to show which assumptions made by the model are valid, and which ones should be further modified to better fit the data.
© 2007 Elsevier Inc. All rights reserved.

Keywords: fMRI; Functional brain interactivity; Effective connectivity; Partial correlation; Structural equation modeling

1. Introduction

Numerous studies have shown that the analysis of blood oxygen level dependent (BOLD) signals obtained through functional magnetic resonance imaging (fMRI) provides accurate maps of *where* functional processes may be instantiated in the brain [1,2]. More recently, some researchers have also suggested that fMRI data contain evidence of *how* the brain orchestrates the interactions between regions to implement cognitive functions. A new field has consequently emerged in fMRI data analysis, whose goal is to investigate effective connectivity, i.e., the influence that regions exert on one another [3–5].

To date, structural equation modeling (SEM) has been the main approach to examine effective connectivity [6–9].

Starting from a set of D regions, a structural model is determined a priori between these regions, based on existing anatomical and functional information regarding the network under investigation. The model defines the time course $z_i(t)$ of each region i, $i=1,\ldots,D$, as a linear function of other regions' time courses,

$$z_i(t) = \sum_{j \neq i} \lambda_{ij} z_j(t) + \epsilon_i(t).$$

In this expression, each path coefficient λ_{ij} quantifies the strength that region j exerts on region i. Some coefficients are constrained to 0, while others remain free to vary. Setting a structural model is equivalent to defining a directed graph, i.e., (i) identifying nodes, each of which stands for a brain region; (ii) setting arrows between these nodes, where an arrow $j{\rightarrow}i$ is present if the corresponding coefficient λ_{ij} is not constrained to 0; and (iii) determining the intensity λ_{ij} of each arrow $j{\rightarrow}i$. Arrows of a graph are

* Corresponding author. Inserm U678, CHU Pitié-Salpêtrière, 91 boulevard de l'Hôpital, 75634 Paris Cedex 13, France. Tel.: +33 01 53 82 84 28; fax: +33 01 53 82 84 48.
 E-mail address: marrelec@imed.jussieu.fr (G. Marrelec).

usually thought to represent anatomical connections that may be functionally expressed during the task under consideration. In this setting, there is thus an anatomical connection underlying each functional interaction.

Most SEM techniques require prior completion of Steps (i) and (ii) described above. They then estimate all unconstrained path coefficients from the data. The present study focuses on Step (ii), which remains a real challenge. A common rule is to resort to known *anatomical information*: consideration of existing structural pathways makes it possible sometimes to narrow the set of potential connections between regions. However, this procedure does not usually suffice to provide a fully specified structural model; the available information must then be completed with further *assumptions*. Such hypotheses usually originate from functional considerations, but no rule exists as to how they should be incorporated into the modeling or their bearing on the accuracy and robustness of the method. Indeed, most studies assume that the model resulting from Steps (i) and (ii) proposed for SEM analysis is true and exclusively focus on estimating the path coefficients. Very few methods confer any kind of feedback relative to the validity of the various assumptions introduced into the structural model during Step (ii). The researcher is hence left with little clue with respect to which of his or her structural/functional hypotheses are supported by the data and the influence of these hypotheses on the conclusiveness of the answers provided to his or her questions by the data.

Recent developments have attempted to devise methods to select the most appropriate structural model within a set of models [9]. Such approaches must deal with very complex issues. Most importantly, the complexity of effective connectivity roughly increases exponentially with the number of regions. Working with D regions in SEM can bring a huge number of graphs — potentially up to $4^{D(D-1)/2}$. For a network of merely $D=5$ regions, this amounts to searching a space of $4^{10} \approx 10^6$ potential graphs. Consequently, algorithms that perform model comparison have to resort to deterministic, local and stepwise procedures to search the space of models. In contrast, directed graphs are very complex structures with global constraints that cannot be determined locally. For instance, whether a graph is acyclic cannot be determined by observation of one arrow at a time but requires consideration of the whole graph. Whether an "optimized" structural model obtained through SEM is indeed optimal (and not merely a local extremum) remains an open issue. So is the question of how probable other potential graphs are. Given a structural assumption (e.g., stating that region a has an effect onto region b), it would be of interest to be able to determine how much evidence from the data supports it regardless of the rest of the model. Bullmore et al. [9] proposed a method to provide a structural model that "best fits" the data (according to some criterion) as the result of a blind search, as well as assessing whether a given structural model, selected on

anatomical and functional considerations prior to data analysis, could be discarded as being significantly different from this "best-fit" model. The algorithm calculated a global measure of discrepancy between the best-fit model and the theoretically proposed graph, providing some kind of feedback with regard to the relevance of the set of assumptions used to construct the theoretically preferred model. However, the validity of local structural patterns of effective connectivity cannot be assessed separately. Providing such a local assessment would prove to be a convenient tool to improve a theoretical model in the face of the actual data: Which connections of the proposed structural model should be kept? Which ones should be removed? To our knowledge, no approach yet allows for such a detailed feedback.

Here, we propose a method for solving this SEM problem by using partial correlation. Upon examination of a network of D regions, partial correlation is considered as the conditional correlation between any two regions with respect to the set of $D-2$ remaining regions. We propose to apply partial correlation analysis to remove mutual dependencies or common influences from other brain areas. Partial correlation has already been used in fMRI data analysis to investigate functional connectivity [10–12]. Such a measure provides a convenient summary of conditional independencies and has striking graphical properties [13] that make it relatively easy to compare to structural models. Also, in Marrelec et al. [14], theoretical considerations led us to hypothesize that partial correlation might be a relevant measure of effective connectivity. In this paper, we go one step further in this direction and provide a direct comparison of partial correlation and SEM analyses. Specifically, we advocate that partial correlation analysis can be used for two purposes: (i) to make educated guesses regarding potential structural connections that should be included into, or could be removed from, a structural model prior to an SEM analysis; and (ii) once an SEM model has been selected, to provide convincing feedback concerning the validity of the local hypotheses used to construct the model, as well as the robustness of the SEM algorithm. In order to illustrate the main features of partial correlation analysis, we resort to a dataset on which SEM analysis has already been performed and published [9]. This dataset, which investigates semantic decision and subvocal rehearsal, was examined with regard to effective connectivity and is thought to be a good benchmark on which to assess the validity of the method proposed here.

The outline of this paper is as follows: in the next section, we introduce the dataset and the structural model used. Section 3 demonstrates that it is possible to infer the partial correlation matrix from the data and shows the bearing of this matrix on setting the structural model. Finally, Section 4 shows that partial correlation analysis can also be applied as a post-processing analysis to check the robustness of SEM algorithms and to make local comparisons between the information incorporated into the

structural model and the structure actually contained in the data. Further issues are raised in the discussion.

2. Study presentation

2.1. General background

The starting point for the present study originated from the study of Bullmore et al. [9] in which group fMRI data of a task requiring semantic decision and subvocal rehearsal were analyzed using classical activation detection methods.

Here, we only introduce details that are relevant for our purpose. Further topics, including issues raised by the group analysis, are developed in the discussion.

Based on the activation maps generated and previous anatomical knowledge, the following $D = 5$ left hemispheric cortical regions of interest were selected: the ventral extrastriate cortex (VEC), the prefrontal cortex (PFC), the supplementary motor area (SMA), the inferior frontal gyrus (IFG) and the inferior parietal lobule (IPL). Thus, the anatomical network considered is the following:

$$\mathcal{R} = \{\text{VEC, PFC, SMA, IFG, IPL}\}.$$

Each region was then associated to a time course for a total of five time courses of length $T = 96$ time samples. The sample correlation matrix corresponding to these time courses, given in Bullmore et al. [9], is reported in Table 1. The time courses were a group average over the subjects and the correlation matrix corresponds to the correlations of the averaged time series.

2.2. SEM analysis

Bullmore et al. [9] first proposed a plausible structural model based on anatomical and functional considerations; the only structural connection for which they could find no evidence in the literature was the VEC–SMA connection. Because removing only two arrows (VEC→SMA and SMA→VEC) resulted in a number of path coefficients that were too high and prevented SEM analysis, further hypotheses were introduced. All the subsequent decisions to remove or keep other links were based on purely functional assumptions. The resulting model, henceforth referred to as the "theoretically preferred model" (or "tp"), is represented in Fig. 1, left. With the use of SEM, the six path coefficients of this model ($\lambda_{15}, \lambda_{21}, \lambda_{32}, \lambda_{43}, \lambda_{51}$ and λ_{54}) were then estimated (see Table 2).

A procedure implemented in the LISREL proprietary software package[1] was then run. Using the sole correlation matrix given in Table 1, it calculated a best-fit model from the data, henceforth referred to as such (or "bf "). The algorithm both proposed a structural model and fitted its path coefficients. The resulting model, schematized in Fig. 1, right, shared some structural similarities with the theoretically preferred model, such as the VEC→PFC, PFC→SMA and IFG→IPL pathways. By contrast, other features also differed. Some connections that were present in the theoretically preferred model were not selected in the best-fit model, such as VEC→IPL and SMA→IFG, while connections that were absent in the theoretically preferred model appeared in the best-fit model, such as PFC→IFG and SMA→IPL. Accordingly, the best-fit model also had six path parameters: $\mu_{15}, \mu_{21}, \mu_{32}, \mu_{42}, \mu_{53}$ and μ_{54} were then estimated (see Table 2).

The theoretically preferred and the best-fit models differed both structurally (different set of arrows) and numerically (different path coefficients for arrows that are common to both models). While the former was based on, and incorporated, cognitive evidence, the latter was purely data driven. Consequently, its relation with, as well as its interpretation in the light of, existing cognitive knowledge was not so obvious. As a result, a method was proposed to assess whether the theoretically preferred model really differed from the best-fit model. Despite the fact that both models were different, Bullmore et al. [9] eventually concluded that the data did not contain enough evidence to enable one to discard the theoretically preferred model as being significantly different from the best-fit model.

3. Pre-SEM analysis

To demonstrate how partial correlation analysis can provide cogent information for SEM, we now return to one step before SEM analysis is performed. Specifically, we assume that only Step (i) of the analysis, i.e., region selection, has successfully been completed, leading to the set of $D = 5$ aforementioned regions. We propose a data-driven pre-processing step, based on partial correlations, that extracts cogent evidence regarding the structure underlying the data based upon the correlation matrix of Table 1.

3.1. Partial correlation

Let $\boldsymbol{z} = (\boldsymbol{z}_t)_{t=1,\ldots,T}$ be the BOLD fMRI time courses of the five regions in $\mathcal{R}$, with each $\boldsymbol{z}_t$ further assumed to be a realization of a five-dimensional Gaussian variable $\boldsymbol{y} = (y_i)_{i=1,\ldots,D}$ of (population) mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Partial correlation between two regions i and j,

Table 1
Sample correlation matrix of the real data set examined in Ref. [9]

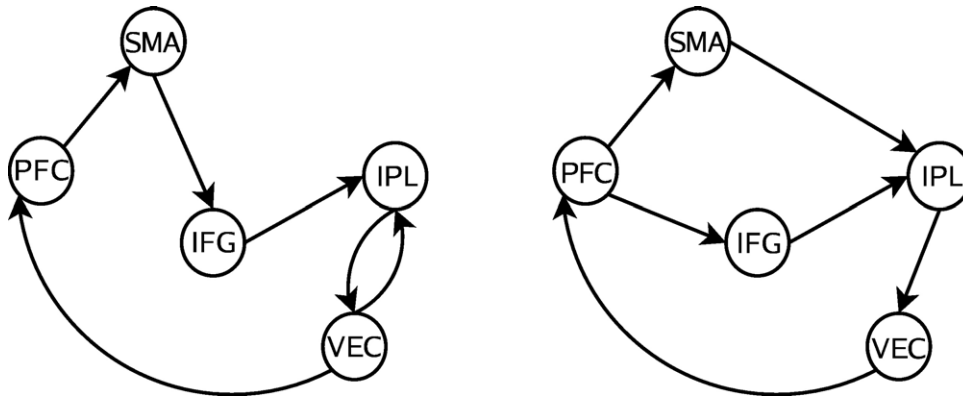|	(1) VEC	(2) PFC	(3) SMA	(4) IFG	(5) IPL
(1) VEC	1				
(2) PFC	0.661	1			
(3) SMA	0.525	0.660	1		
(4) IFG	0.486	0.507	0.437	1	
(5) IPL	0.731	0.630	0.558	0.517	1

---

[1] http://www.ssicentral.com/lisrel/.

Fig. 1. Theoretically preferred (left) and best-fit (right) models (from Ref. [9]).

denoted by Π_{ij}, is here defined as the correlation between these two regions conditioned on the set $\mathcal{R}\setminus\{i,j\}$ of remaining regions, i.e., [13]

$$\Pi_{ij} = \text{Corr}[y_i, y_j | \boldsymbol{y}_{\mathcal{R}\setminus\{i,j\}}], \qquad (1)$$

The $D(D-1)/2$ partial correlation coefficients form the D-by-D partial correlation matrix $\boldsymbol{\Pi}=(\Pi_{ij})$, which can be readily calculated from the concentration, or precision, matrix $\boldsymbol{\Sigma}^{-1}=\boldsymbol{\Upsilon}=(\Upsilon_{ij})$ as [13]

$$\Pi_{ij} = -\frac{\Upsilon_{ij}}{\sqrt{\Upsilon_{ii}\Upsilon_{jj}}} \qquad (2)$$

for two distinct regions i and j, and $\Pi_{ii}=1$.

Estimating $\boldsymbol{\Pi}$ from the data can be readily performed using a numeric sampling scheme (Appendix A; see also Ref. [15] or [12]). The sample obtained can then be used to approximate estimates of the partial correlation coefficients. The values of these coefficients can also be tested against the null hypothesis (H_0) that $\Pi_{ij}= 0$, the sampling scheme then giving approximations of the coefficient significance

levels (see Appendix A). Classically defining the marginal correlation matrix as $\boldsymbol{\Omega}=(\Omega_{ij})$ with

$$\Omega_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\cdot\Sigma_{jj}}}, \qquad (3)$$

the posterior distribution of this matrix, $\text{Pr}(\boldsymbol{\Omega}|\boldsymbol{z})$, can incidentally also be approximated with the same procedure.

Note that the sampling scheme proposed in Appendix A requires the sample covariance matrix, whereas Bullmore et al. [9] only refer to the sample correlation matrix (reproduced here in Table 1). However, this fact has no effect on our analysis, since we are not so much interested in the covariance matrix as in the marginal and partial correlation matrices, i.e., matrices that are normalized. On this account, working with the sample covariance matrix or with the sample correlation matrix leads to the same result.

3.2. Results

We used the sampling scheme of Appendix A to compare each marginal and partial correlation coefficient to 0 and to test whether they significantly differed from 0. More specifically, we tested each coefficient Ω_{ij} of $\boldsymbol{\Omega}$ against the null hypothesis (H_0) $\Omega_{ij}=0$, and similarly with Π_{ij}. The significance levels for the marginal correlation coefficients were all smaller than 0.001; the significance levels corresponding to the partial correlation coefficients, denoted α_{ij}, are shown in Table 3. In Fig. 2, we represented the corresponding log-significance levels, $-\log(\alpha_{ij})$, in the form of a graph.

Table 2
Estimated path coefficients for the theoretically preferred (top) and the best-fit (bottom) models (from Ref. [9])

	(1) VEC	(2) PFC	(3) SMA	(4) IFG	(5) IPL
λ					
(1) VEC	0	0	0	0	0.80
(2) PFC	0.59	0	0	0	0
(3) SMA	0	0.60	0	0	0
(4) IFG	0	0	0.31	0	0
(5) IPL	−0.16	0	0	0.52	0
μ					
(1) VEC	0	0	0	0	0.61
(2) PFC	0.50	0	0	0	0
(3) SMA	0	0.58	0	0	0
(4) IFG	0	0.43	0	0	0
(5) IPL	0	0	0.27	0.58	0

Table 3
Inference

	(1) VEC	(2) PFC	(3) SMA	(4) IFG
(1) VEC				
(2) PFC	0.002			
(3) SMA	0.409	<0.001		
(4) IFG	0.188	0.055	0.192	
(5) IPL	<0.001	0.100	0.045	0.033

Significance levels α_{ij} associated with the partial correlation matrix $\boldsymbol{\Pi}$ given the data $\boldsymbol{z}$.
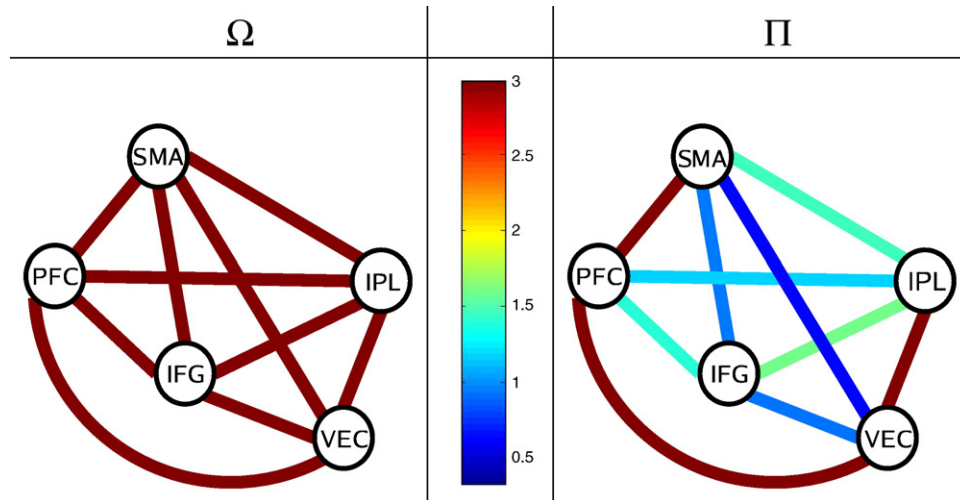
Fig. 2. Inference. Significance graphs associated with the marginal (left) and partial (right) correlation matrices. The scale is logarithmic, e.g., $\alpha = 0.01 = 10^{-2}$ is associated to a log-significance of 2. Maximum (red) is 3 and above; minimum (blue) is $-\log_{10}(0.5) \approx 0.301$.

Examination of the results clearly illustrates the inadequacy of marginal correlation as a way to quantify effective connectivity: all marginal correlations significantly differ from 0. This is in part due to the transitive property of correlation; a correlation between two variables can be induced, e.g., by a common input. This is confirmed by our example. A fully connected model does not make much sense for the data considered. Any relevant measure of effective connectivity should therefore have at least one or several values that do not significantly differ from 0. For instance, since no evidence of an anatomical connection between VEC and SMA was found, no effective connectivity is expected between the two regions. In contrast, all marginal correlation coefficients (Fig. 2, left) significantly differ from 0 ($P = .001$). In particular, the data still exhibit a marginal correlation between VEC and SMA, Ω_{13}, that is significant.

By contrast, the graph of partial correlation log-significances is much more informative (see Table 3). Three links significantly differ from 0 at a significance level of $P = .01$: VEC–IPL, PFC–SMA and VEC–PFC. Two links are under or around the $P = .05$ threshold limit: IFG–IPL, SMA–IPL and PFC–IFG. Finally, four partial correlation coefficients are above the limit: PFC–IPL, VEC–IFG, SMA–IFG and VEC–SMA. Interestingly, the partial correlation associated with the only absent anatomical connection that was previously reported, VEC–SMA, Π_{13}, does not significantly differ from 0 ($P < .05$). Indeed, it is the coefficient with the lowest significance level.

Second, the structure of significant and nonsignificant partial correlations can be compared to the information introduced into the theoretically preferred model. The three most significant partial correlation coefficients, VEC–IPL, VEC–PFC, and PFC–SMA, are represented in the form of directed paths (VEC↔IPL, VEC↔PFC and PFC→SMA, respectively). Among the significant links ($P \approx .05$), one is represented in the structural model (IFG→IPL), two are not (PFC–IFG and SMA–IPL). The VEC–SMA, VEC–IFG and

PFC–IPL connections, which have low partial correlations (Π_{13}, Π_{14} and Π_{25}, respectively), have consistently not been kept in the structural model. One nonsignificant link has been represented in the structural model (SMA→IFG). Both resemblances and dissimilarities will be further interpreted later.

4. Post-SEM analysis

In this section, we assume that all SEM analyses described in Ref. [9] and summarized in Section 2 have been conducted. Partial correlation analysis can now serve two goals. The results previously found can be further interpreted in the light of the new structural models. These models can also be used to generate surrogate data where the expected partial correlation structure can be compared to the real dataset to test for the validity of the estimated models.

4.1. Partial correlation vs. SEM inference

The results of the pre-processing step, detailed in Section 3.2, can also be compared a posteriori to the results of the SEM analysis. Indeed, there exists a striking resemblance between the graphs of partial correlation log-significance and the SEM analyses as presented in Ref. [9]. When both the theoretically preferred and the best-fit models agreed, good agreement was found with partial correlations as well: a lack of connection VEC–SMA, strong connections IPL–VEC, VEC–PFC, and PFC–SMA. On the other hand, when both structural models disagreed (e.g., regarding how information flows from PFC to IPL), partial correlation analysis showed that both models introduced different structural connections (e.g., SMA→IFG for the theoretically preferred model; PFC→IFG and SMA→IPL for the best-fit model) that were associated with rather low partial correlation values and, hence, did not strongly structure the data.

The inadequacy of using marginal correlation can be further exemplified as follows. The correlation between PFC

Table 4
Surrogate data

	(1) VEC	(2) PFC	(3) SMA	(4) IFG
tp				
(1) VEC				
(2) PFC	<0.001			
(3) SMA	0.497	<0.001		
(4) IFG	0.270	0.492	0.0056	
(5) IPL	<0.001	0.498	0.497	<0.001
bf				
(1) VEC				
(2) PFC	<0.001			
(3) SMA	0.499	<0.001		
(4) IFG	0.498	<0.001	0.249	
(5) IPL	<0.001	0.499	0.0151	0.0119

Levels of significance α_{ij} associated to the sample partial correlation matrix for the theoretically preferred (top) and best-fit (bottom) models.

and IPL was quite high (0.630±0.063). Yet, in both the theoretically preferred and the best-fit structural models, this functional connection can be explained by indirect effective connections: through PFC→SMA→IFG→IPL in the theoretically preferred model; through PFC→SMA→IPL and PFC→IFG→IPL in the best-fit model. Consistently, the corresponding partial correlation was lower (0.100±0.135)

and not significantly different from 0 ($P=.05$), confirming the unlikeliness of an effective PFC–IPL connection.

4.2. Surrogate data

Once an SEM analysis has been completed, an important objective is to assess the validity of the resulting effective connectivities. Indeed, since most algorithms used are deterministic, local and stepwise procedures, many potential factors might have biased the analysis. It would consequently be of interest to provide a simple method that could test whether the SEM obtained after a complex processing procedure is indeed relevant to the data under investigation. To this end, we propose the following procedure. Since SEM analysis has been fully completed, the structural models have been fully determined and it is possible to calculate the corresponding covariance matrices, Σ_{tp} and Σ_{bf} in our case. Given a structural model and its coefficients, the distribution of a sample covariance matrix can hence easily be calculated and simulated using a sampling scheme similar to that used for the inference process (see Appendix B). The distribution of the sample marginal and partial correlation matrices, $\boldsymbol{R}$ and $\boldsymbol{P}$, respectively, can also be approximated. For both the theoretically preferred and the best-fit models, the
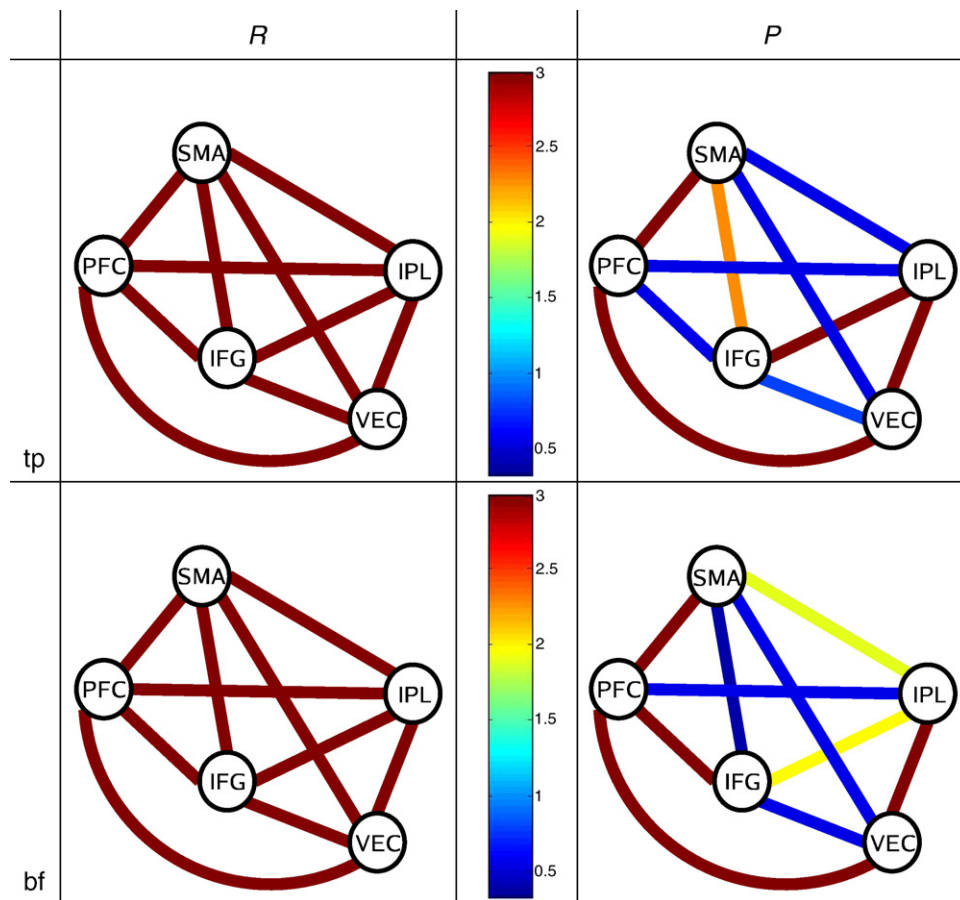
Fig. 3. Surrogate data. Graphs representing the sample marginal, $\boldsymbol{R}$, and partial, $\boldsymbol{P}$, correlation coefficients for the theoretically preferred (top) and the best-fit (bottom) models.

significances of the partial correlations are given in Table 4 (all marginal correlation coefficients are significantly different from 0). The log-significance graphs have also been represented in Fig. 3 for both **R** and **P**.

As in Section 3, the significance of the sample marginal correlation coefficients does not convey much information, as they all significantly differ from 0. The structure of partial correlation log-significances, on the other hand, strikingly resembles that of the structural model that was used to generate them. As a matter of fact, each partial correlation structure, as represented in Fig. 3, matches the skeleton of the corresponding structural model, obtained by transforming the arrows into unoriented edges: VEC–PFC, VEC–IPL, PFC–SMA, SMA–IFG and IFG–IPL for the theoretically preferred model; VEC–PFC, VEC–IPL, PFC–SMA, SMA–IPL, PFC–IFG and IFG–IPL for the best-fit model. This remarkable feature, which was hinted at in Marrelec et al. [14] from a theoretical perspective, is here demonstrated on simulated data. Note that the significance of effective connections taken into account in each of the two graphs (e.g., PFC→SMA for the theoretically preferred model) largely differs from the significance of effective connections that were not modeled (e.g., between PFC and IFG in the theoretically preferred model). For each simulation setting (i.e., each structural graph), a $P=.05$ significance threshold unambiguously retrieved the skeleton of the structural graph. Varying the threshold from 0.25 to about 0.015 did not change the results.

5. Discussion and perspectives

In this paper, we proposed a method that, given a set of regions, performs a partial correlation analysis. This method provides a way to approach effective connectivity that is data driven, in the sense that it does not require any prior information regarding the anatomical or functional connections. The results reveal that partial correlation analysis can serve several purposes. In a pre-processing step for a subsequent SEM analysis, inferring the partial correlation structure from the data can give insight into the effective connections that are structuring the interactions (these can consequently be incorporated into the SEM) and those that have little influence onto the pattern of connectivity (allowing us to remove them and gain degrees of freedom). As a post-processing step, generation of synthetic data using the path coefficient estimates obtained by any SEM optimization algorithm provides a simple visual way to check the validity of the algorithm used; it also shows which assumptions made by the SEM model are valid and those that should be modified to better fit the data.

The relationships between SEM and conditional correlation have been the topic of much research and involve graph theoretic concepts like morality and d-separation [13,16,17]. Theoretical considerations led us to hypothesize that partial and, more generally, conditional, correlation coefficients could extract the (undirected) structure of

effective connectivity from the data [14]. The analysis developed in this paper strongly supports this assumption. Indeed, while we demonstrated that a lack of partial correlation between two regions can potentially be related to a lack of underlying anatomical connection, the example used suggests that a strong and significant partial correlation can be interpreted as the presence of an effective connection. Whether this behavior is a general property of fMRI data or only incidental remains to be investigated. Nonetheless, we believe that partial correlation will prove essential to effective connectivity investigation, for it can compensate for some of the most important drawbacks from which SEM analysis suffers (i.e., difficulty to provide a structural model a priori and lack of control over the SEM algorithms and results).

Partial correlation analysis, as introduced in this paper, has a very strong bearing on both the analysis and the consistency of SEM. In this perspective, it stands as an efficient way to provide feedback regarding the relevance of a model given a set of data. A good way to perform joint SEM and partial correlation analysis would be as follows: a blind partial correlation analysis could be performed from the data, based on the data-driven inferential process detailed in this paper. The resulting structure could then be compared to previous knowledge from the literature and checked for any discrepancies. If no discrepancies are observed, a first structural model could then be proposed, integrating both prior information from the literature and paths with significant partial correlations. Orientation of the connections (i.e., transformation of the unoriented links into arrows) should still rely solely on prior knowledge. A classical SEM analysis could then be conducted with the model obtained. After estimation of the corresponding path coefficients, simulations could be run to compare the partial correlation structure expected from such a structural model to the partial correlation structure actually observed in the data. This comparison would exhibit both connections on which there is a good agreement (i.e., well modeled) and connections for which there exists a discrepancy between the simulated expected value and the value observed in the data. Such a comparison could give some insight regarding potential improvements that could be made to allow for a better fit. In case of a discrepancy between the extracted pattern of partial correlation and previous knowledge, further research must be conducted in order to provide rules to properly constrain the partial correlation matrix with existing information.

At this stage, how SEM and partial correlation deal with group analysis deserves a comment. SEM software by and large only accepts a single correlation matrix as input, thereby not explicitly considering any potential group variability — only the "average" subject can be examined. To compensate for this flaw, various pre-processing schemes try to artificially incorporate some variability induced by the group. For instance, Bullmore et al. [9] first calculated one spatially averaged time series for each of the five regions and each of the 20 subjects. For each of the five regions,

they then performed PCA of the 20 corresponding subject time series and kept the first eigenseries as representative of an average pattern of response to the experimental design over all subjects in each region. This step led to a total of five times series (one for each region), from which they calculated a correlation matrix (Table 1). It is this correlation matrix that was subsequently used for path analysis in Bullmore et al. [9] and for partial correlation investigation in this manuscript. While it is fairly easy to model some group variability in the method that we expounded here (see, e.g., Ref. [12]), how such results can be compared to an SEM analysis remains to be investigated.

The relevance of partial correlation to investigate SEM comes up as a cogent demonstration that this technique provides a pertinent first step to bridge the gap that has endured between functional and effective connectivity. Marginal correlation is often used in functional connectivity analyses (e.g., Refs. [18–23]). However, it has been suspected of being a rather poor measure of effective connectivity, based on the prediction that two regions that were indirectly connected or driven by the same stimulus could still exhibit a significant marginal correlation (e.g., Refs. [14,15]). Our results, both from Marrelec et al. [12] and this article, tend to experimentally confirm this theoretical assumption. By contrast, the same research tends to demonstrate that partial correlation, while also being data driven, sheds some light on the effective connectivity structure of a brain network.

A final key issue that needs mentioning is that partial correlation is unable to recover connection directionality from the data. Note first that inferring direct interactions and directionality are two separate issues. The best proof of this assertion is that, while some data-driven methods, mainly based on Granger causality, try to cope with directionality (e.g., Refs. [24,25]), none has been able to differentiate between direct and mediated interactions. Directionality represents an information that is complementary to interaction mediation and still has to be assumed a priori. Nonetheless, our method still allows one to drastically limit the number of potential graphs by putting strong constraints onto the skeleton of any potential structural model.

Acknowledgments

Appendix A. Inferring the marginal and partial correlation matrices

With the use of standard Bayesian theory, it can be shown that the covariance matrix Σ given the data z follows an inverse Wishart distribution with $T-1$ degrees of freedom and scale matrix $U = S^{-1}$, where

$$S = \sum_{t=1}^{T} (z_t - \overline{z_t})(z_t - \overline{z_t})^{t}$$

is proportional to the sample covariance matrix, and $\overline{z}_t$ is the temporal mean [26]. Calculation of the posterior probability density function (pdf) of partial correlation cannot be performed in close form from this distribution. To approximate the distribution of the marginal and partial correlation matrices, we can nevertheless resort to the following sampling scheme [12,15]). For sample l,

(1) sample $\Sigma^{[l]}$ according to its inverse Wishart distribution [26] (Appendix A);
(2) calculate $\Upsilon^{[l]} = (\Sigma^{[l]})^{-1}$, and $\Pi^{[l]}$ from $\Upsilon^{[l]}$ according to Eq. (2); $\Omega^{[l]}$ from Eq. (3).

Once a large number L (10,000 in this article) of samples has been drawn following this process, the marginal pdf of a given quantity can be approximated by the frequency histogram obtained from the sample. Likewise, all statistics and estimators can be approximated by their sample counterparts. For instance,

$$\mathrm{E}\left[\Pi_{ij}|z\right] \approx M_{ij} = \frac{1}{L} \sum_{l=1}^{L} \Pi_{ij}^{[l]}, \tag{A.1}$$

$$\mathrm{Var}\left[\Pi_{ij}|z\right] \approx X_{ij} = \frac{1}{L} \sum_{l=1}^{L} \left(\Pi_{ij}^{[l]} - M_{ij}\right)^2. \tag{A.2}$$

Significance tests can also be approximated in the same way. For instance, testing against the null hypothesis (H_0) $\Pi_{ij} = \Pi_{ij,0}$ can be associated with the following significances: $\mathrm{Pr}(\Pi_{ij} > \Pi_{ij,0})$ if $M_{ij} > \Pi_{ij,0}$ or $\mathrm{Pr}(\Pi_{ij} < \Pi_{ij,0})$ if $M_{ij} < \Pi_{ij,0}$. These quantities can, in turn, be approximated from the sample by

$$1 - \alpha_{ij} = \mathrm{Pr}\left(\Pi_{ij} > \Pi_{ij,0}\right) \approx \frac{1}{L} \# \left\{ \Pi_{ij}^{[l]} > \Pi_{ij,0} \right\}$$

and

$$1 - \alpha_{ij} = \mathrm{Pr}\left(\Pi_{ij} < \Pi_{ij,0}\right) \approx \frac{1}{L} \# \left\{ \Pi_{ij}^{[l]} < \Pi_{ij,0} \right\},$$

respectively, where $\#S$ stands for the cardinal of set S. For instance, setting $\Pi_{ij,0}$ to 0 makes it possible to test the null hypothesis $\Pi_{ij} = 0$.

Appendix B. Generation of surrogate data

A structural model can be defined in matrix form as

$$y = Ky + e, \tag{B.1}$$

$K = (K_{ij})$ contains the path coefficients. Both the theoretically preferred and the best-fit models have different matrices, denoted by K_{tp} and K_{bf}, respectively, that are

functions of the path coefficients. We also assume that the noise e is composed of spatially and temporally independent Gaussian variables with diagonal covariance matrix:

$$\text{Var}[\boldsymbol{e}] = \boldsymbol{V} = \begin{pmatrix} V_1^2 & & 0 \\ & \ddots & \\ 0 & & V_D^2 \end{pmatrix}.$$

The first equation can be rewritten as $\boldsymbol{y} = (\boldsymbol{I}-\boldsymbol{K})^{-1}\boldsymbol{e}$, where $\boldsymbol{I}$ stands for the D-dimensional unit matrix; it is straightforward to show that $\boldsymbol{y}$ is also Gaussian distributed with covariance matrix [27]

$$\boldsymbol{\Sigma}_m = (\boldsymbol{I} - \boldsymbol{K}_m)^{-1} \boldsymbol{V} \left[(\boldsymbol{I} - \boldsymbol{K}_m)^{-1} \right]^{\text{t}}, \tag{B.2}$$

with $m \in \{\text{tp}, \text{bf}\}$. Since $\boldsymbol{K}_m$ is a function of the path coefficients, so is $\boldsymbol{\Sigma}_m$. The likelihood of the data given $\boldsymbol{\Sigma}_m$ then reads:

$$\Pr(\boldsymbol{z} | \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_m) = \prod_{t=1}^{T} \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_m; \boldsymbol{z}_t).$$

In this expression, $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{z}_t)$ stands for the multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The value of the sample covariance matrix $\boldsymbol{S}$ is then given by [27]

$$\boldsymbol{S} | \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_m \sim \text{Wishart}\left(T-1, \frac{1}{T-1}\boldsymbol{\Sigma}_m \right).$$

To approximate the distribution of the sample marginal and partial correlation matrices $\boldsymbol{R}$ and $\boldsymbol{P}$, we resorted to a sampling scheme similar to that used for the inference process and detailed in Appendix A.

References

[1] Frackowiak RSJ, Friston KJ, Frith CD, Dolan RJ, Price CJ, Zeki S, et al, editors. Human brain function. 2nd ed. Paris: Elsevier Academic Press; 2004.

[2] Huettel SA, Song AW, McCarthy G. Functional magnetic resonance imaging. Sunderland: Sinauer; 2004.

[3] Horwitz B. Data analysis paradigms for metabolic-flow data: combining neural modeling and functional neuroimaging. Hum Brain Mapp 1994;2:112–22.

[4] Friston KJ. Functional and effective connectivity in neuroimaging: a synthesis. Hum Brain Mapp 1994;2:56–78.

[5] Horwitz B, Tagamets M-A, McIntosh AR. Neural modeling, functional brain imaging and cognition. Trends Cogn Sci 1999;3:91–8.

[6] McIntosh AR, Gonzalez-Lima F. Structural equation modeling and its application to network analysis of functional brain imaging. Hum Brain Mapp 1994;2:2–22.

[7] McIntosh AR, Grady CL, Ungerleider LG, Haxby JV, Rapoport SI, Horwitz B. Network analysis of cortical visual pathways mapped with PET. J Neurosci 1994;14:655–66.

[8] Büchel C, Coull JT, Friston KJ. The predictive value of changes in effective connectivity for human learning. Science 1999;283:1538–41.

[9] Bullmore E, Horwitz B, Honey G, Brammer M, Williams S, Sharma T. How good is good enough in path analysis of fMRI data? Neuroimage 2000;11:289–301.

[10] Salvador R, Suckling J, Coleman M, Pickard JD, Menon D, Bullmore E. Neurophysiological architecture of functional magnetic resonance images of human brain. Cereb Cortex 2005;34:387–413.

[11] Salvador R, Suckling J, Schwarzbauer C, Bullmore E. Undirected graphs of frequency-dependent functional connectivity in whole brain networks. Philos Trans R Soc Lond B Biol Sci 2005;360:937–46.

[12] Marrelec G, Krainik A, Duffau H, Pélégrini-Issac M, Lehéricy S, Doyon J, et al. Partial correlation for functional brain interactivity investigation in functional MRI. Neuroimage 2006;32:228–37.

[13] Whittaker J. Graphical models in applied multivariate statistics. Chichester: J. Wiley and Sons; 1990.

[14] Marrelec G, Daunizeau J, Pélégrini-Issac M, Doyon J, Benali H. Conditional correlation as a measure of mediated interactivity in fMRI and MEG/EEG. IEEE Trans Signal Process 2005;53:3503–16.

[15] Marrelec G, Doyon J, Pélégrini-Issac M, Benali H. Heading for data-driven measures of effective connectivity in functional MRI. Proceedings of the International Joint Conference on Neural Networks. 2005. p. 1528–33.

[16] Lauritzen SL. Graphical models. Oxford: Oxford University Press; 1996.

[17] Pearl J. Causality: models, reasoning, and inference. Cambridge: Cambridge University Press; 2001.

[18] Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echoplanar MRI. Magn Reson Med 1995;34:537–41.

[19] Biswal BB, Kylen JV, Hyde JS. Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps. NMR Biomed 1997;10:165–70.

[20] Cordes D, Haughton VM, Arfanakis K, Wendt GJ, Turski PA, Moritz CH, et al. Mapping functionally related regions of brain with functional connectivity MR imaging. Am J Neuroradiol 2000;21:1636–44.

[21] Greicius MD, Krasnow B, Reiss AL, Menon V. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci USA 2003;100:253–8.

[22] Dodel S, Golestani N, Pallier C, ElKouby V, Le Bihan D, Poline J-B. Condition-dependent functional connectivity: syntax network in bilinguals. Philos Trans R Soc Lond B Biol Sci 2005;360:921–35.

[23] Bellec P, Perlbarg V, Jbabdi S, Pélégrini-Issac M, Anton J-L, Doyon J, et al. Identification of large-scale networks in the brain using fMRI. Neuroimage 2006;29:1231–43.

[24] Goebel R, Roebroeck A, Kim D-S, Formisano E. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magn Reson Imaging 2003;21:1251–61.

[25] Roebroeck A, Formisano E, Goebel R. Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage 2005; 25:230–42.

[26] Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian data analysis. Texts in statistical science. London: Chapman & Hall; 1998.

[27] Anderson TW. An introduction to multivariate statistical analysis. Wiley publications in statistics. New York: John Wiley and Sons; 1958.