

Research Article

A Theoretical Investigation of the Relationship between Structural Equation Modeling and Partial Correlation in Functional MRI Effective Connectivity

Guillaume Marrelec^{1,2} and Habib Benali^{1,2}

¹Inserm, U678, Laboratoire d'Imagerie Fonctionnelle, F-75013 Paris, France

²UPMC Univ Paris 06, UMR-S 678, Laboratoire d'Imagerie Fonctionnelle, F-75013 Paris, France

Correspondence should be addressed to Guillaume Marrelec, marrelec@imed.jussieu.fr

Received 31 October 2008; Accepted 15 May 2009

Recommended by Fabio Babiloni

An important field of blood oxygen level dependent (BOLD) functional magnetic resonance imaging (fMRI) is the investigation of effective connectivity, that is, the actions that a given set of regions exert on one another. We recently proposed a data-driven method based on the partial correlation matrix that could provide some insight regarding the pattern of functional interaction between brain regions as represented by structural equation modeling (SEM). So far, the efficiency of this approach was mostly based on empirical evidence. In this paper, we provide theoretical fundamentals explaining why and in what measure structural equation modeling and partial correlations are related. This gives better insight regarding what parts of SEM can be retrieved by partial correlation analysis and what remains inaccessible. We illustrate the different results with real data.

Copyright © 2009 G. Marrelec and H. Benali. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Blood oxygen level dependent (BOLD) functional magnetic resonance imaging (fMRI) is an imaging technique that allows to dynamically and noninvasively follow metabolic and hemodynamic consequences of brain activity [1, 2]. Since Biswal et al. [3], an increasing number of studies have suggested that fMRI data could be used to explore how brain regions interact to perform functional tasks. A key concept in investigation of functional brain interactions is effective connectivity, which has been defined as the influence that regions exert on one another [4].

Path analysis, or structural equation modeling (SEM), has been the major way to examine effective connectivity in fMRI [5–7]. Starting from a set of D regions, a model is set a priori that expresses the time course $z_i(t)$ of each region as a linear function of the time course of other regions

$$z_i(t) = \sum_{j \neq i} \lambda_{ij} z_j(t) + e_i(t), \quad (1)$$

with some coefficients λ_{ij} being constrained to 0, the others are free to vary. λ_{ij} quantifies the strength that region j

exerts on region i . Setting an SEM is equivalent to defining a directed graph, where each node stands for a region, a given arrow $j \rightarrow i$ is present if and only if the corresponding coefficient λ_{ij} is not constrained to zero, and, finally, λ_{ij} represents the intensity of arrow $j \rightarrow i$. Once the structural model is completely set, the unconstrained coefficients λ_{ij} are estimated. To this aim, the model covariance matrix Σ , which is a function of the parameters, is compared to the sample covariance matrix S using a discrepancy function that is minimized [8, 9]. In fMRI data analysis, the following maximum likelihood function is often used [7]:

$$l(\Sigma) = \text{tr}(\Sigma^{-1}) - \ln |\Sigma^{-1} S| - D, \quad (2)$$

where $\text{tr}(\cdot)$ stands for the standard matrix trace function. The major flaw of this approach is that it requires the prior definition of a structural model, that is, of regions and arrows, each arrow requiring itself information regarding connection and direction. By contrast, information regarding the functional interactions present within the network of interest is likely to be scarce, since it is often the very reason why an fMRI study of effective connectivity is carried

out. This is all the more problematic that the approach does not really provide any clear way to challenge the model or to provide information relative to where or how the model under investigation could be improved.

We recently proposed a novel approach to gain insight on effective connectivity. We first showed that, unlike marginal (i.e., regular) correlation, conditional correlation could account for many patterns of interaction as modeled by SEM [10, 11]. We then proposed to focus on a specific set of conditional correlations, namely partial correlations [12]. Given a set of D regions, denoted by \mathbb{R} , and a variable y_i associated to each region i (of which $z_i(t)$ mentioned in (1) is a realization), the method estimates the partial correlation of any region pair (i, j) given the set of $D - 2$ remaining regions,

$$\text{Corr}[y_i, y_j \mid \mathbf{y}_{\mathbb{R} \setminus \{i, j\}}]. \quad (3)$$

On both real [13] and synthetic data [14], it was observed that a large partial correlation value between two regions was often associated with the presence of an effective connectivity between these regions. However, the reason for such a behavior remained unclear. In the present paper, we further delve into the relationship between SEM and partial correlation in order to better understand why and in what measure partial correlation can extract information that is relevant for effective connectivity analysis. To this aim, we provide a theoretical relationship between SEM and partial correlation through the computation of the inverse covariance matrix (also-called concentration or precision matrix). To illustrate the results so obtained, we use a dataset on which SEM analysis has already been performed and published [7].

2. From SEM to Partial Correlation

2.1. Bullmore et al. [7] SEM Study. We here quickly recall the essentials of a previous study on which our investigation of partial correlation relies. For more detail, we refer the reader to Bullmore et al. [7]. The study focused on $D = 5$ left hemispheric cortical regions of interest: 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). Each region was associated to a time course for a total of five time courses of length $T = 96$ time samples. The sample marginal and partial correlation matrices corresponding to these time courses are 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.

A plausible structural model, henceforth referred to as the “theoretically preferred model” (or “TP”), was proposed and is represented in Figure 1(a). Using the correlation matrix of Table 1, a procedure implemented in the LISREL proprietary software package (<http://www.ssicentral.com/lisrel/>) computed a so-called “best fit” model from the data, henceforth referred to as such (or “BF”) and represented in Figure 1(b). While similar in some ways, the two models had different features:

TABLE 1: Sample marginal correlation coefficients of the real data set examined in Bullmore et al. [7].

		(1)	(2)	(3)	(4)	(5)
		VEC	PFC	SMA	IFG	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

- (i) VEC \rightarrow IPL and SMA \rightarrow IFG were present in the theoretically preferred model but were not selected in the best fit model;
- (ii) PFC \rightarrow IFG and SMA \rightarrow IPL were absent in the theoretically preferred model but appeared in the best fit model.

We now go back to a different perspective. Indeed, the structure of any SEM entails specific constraints on the covariance matrix, as well as other matrices characteristic of the process, such as the concentration matrix and the marginal and partial correlation matrices.

2.2. SEM Modeling. Generally speaking, a structural model can be defined in matrix form as

$$\mathbf{y} = \mathbf{K}\mathbf{y} + \mathbf{e}, \quad (4)$$

where \mathbf{y} is the D -dimensional variable characterizing the state of each region and \mathbf{e} is a temporally independent and identically distributed (i.i.d.) Gaussian noise with diagonal covariance matrix. $\mathbf{K} = (K_{ij})_{i,j=1,\dots,D}$ contains the path coefficients. The N time samples $(\mathbf{z}(t_n))_{n=1,\dots,N}$, where $\mathbf{z}(t_n)$ is the signal measured in each of the D regions at time t_n , are supposed to be N i.i.d. realizations of \mathbf{y} . The matrices corresponding to the theoretically preferred and the best fit models are, respectively, given by (see also Figure 1)

$$\mathbf{K}_{\text{TP}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \lambda_{15} \\ \lambda_{21} & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_{32} & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_{43} & 0 & 0 & 0 \\ \lambda_{51} & 0 & 0 & \lambda_{54} & 0 & 0 \end{pmatrix}, \quad (5)$$

$$\mathbf{K}_{\text{BF}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \mu_{15} \\ \mu_{21} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu_{32} & 0 & 0 & 0 & 0 \\ 0 & \mu_{42} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mu_{53} & \mu_{54} & 0 & 0 \end{pmatrix}.$$

2.3. SEM and Covariance. Classically, we further assume that the noise \mathbf{e} of (4) is composed of spatially and temporally

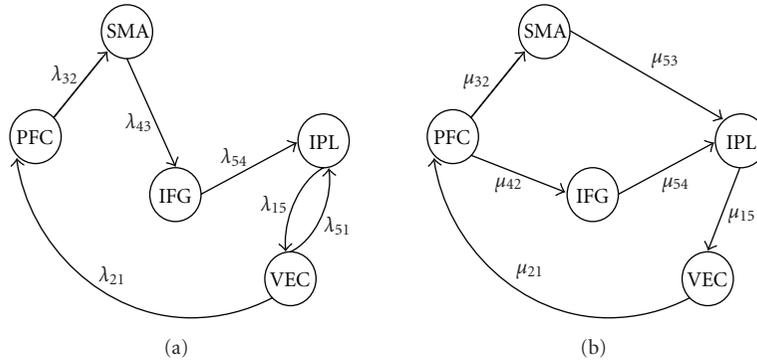


FIGURE 1: Structural models and path coefficients corresponding to the theoretically preferred (a) and best fit (b) models (from [7]).

independent Gaussian variables with diagonal covariance matrix:

$$\text{Var}[\mathbf{e}] = \mathbf{V} = \begin{pmatrix} V_1 & & 0 \\ & \ddots & \\ 0 & & V_D \end{pmatrix}. \quad (6)$$

Since (4) rereads $\mathbf{y} = (\mathbf{I} - \mathbf{K})^{-1}\mathbf{e}$, where \mathbf{I} stands for the D -dimensional unit matrix, it is straightforward to show that \mathbf{y} is also Gaussian distributed with covariance matrix [15]

$$\mathbf{\Sigma} = (\mathbf{I} - \mathbf{K})^{-1}\mathbf{V}[(\mathbf{I} - \mathbf{K})^{-1}]^t, \quad (7)$$

where “ t ” stands for matrix transposition. Since \mathbf{K} is a function of the path coefficients, so is $\mathbf{\Sigma}$. This relationship is central to SEM analysis, for most methods rely on comparing the covariance matrix $\mathbf{\Sigma}$ implied by a structural model to the data sample covariance matrix using normal theory maximum likelihood—leading to the discrepancy function of (2)—, generalized least squares, or ordinary least squares [8, 9]. Note that, in (2), $\mathbf{\Sigma}$ only appears through its inverse $\mathbf{Y} = \mathbf{\Sigma}^{-1}$. \mathbf{Y} is called the concentration, or precision, matrix and it is on this matrix that we will focus to get a better understanding of the data structure.

2.4. SEM and Concentration. Indeed, \mathbf{Y} has intriguing structural properties when related to a structural model. Using (7), this matrix is given by

$$\mathbf{Y} = (\mathbf{I} - \mathbf{K})^t\mathbf{V}^{-1}(\mathbf{I} - \mathbf{K}). \quad (8)$$

\mathbf{V} being a diagonal matrix, the expression for each element Y_{ij} of the concentration matrix can easily be expanded as

$$Y_{ij} = \sum_l \frac{(\delta_{ii} - K_{li})(\delta_{lj} - K_{lj})}{V_l}. \quad (9)$$

Given that $K_{ii} = 0$, the previous equation yields

$$Y_{ii} = \frac{1}{V_i} + \sum_{l \neq i} \frac{K_{li}^2}{V_l}, \quad (10)$$

and, for $i \neq j$,

$$Y_{ij} = -\frac{K_{ij}}{V_i} - \frac{K_{ji}}{V_j} + \sum_{l \notin \{i,j\}} \frac{K_{li}K_{lj}}{V_l}. \quad (11)$$

Equation (11) can be used to compute the concentration coefficients corresponding to the TP and BF structural models. For instance, we have for the TP model

$$\begin{aligned} Y_{12} &= -\frac{\lambda_{21}}{V_2}, \\ Y_{13} &= 0, \\ Y_{14} &= \frac{\lambda_{51}\lambda_{54}}{V_5}. \end{aligned} \quad (12)$$

From this example, we see that two cases can arise. In the first case (e.g., Y_{13}), the value of the concentration coefficient is equal to zero, not because of the specific numerical values that have been assigned to the path coefficients, but because of the structure of the SEM itself. In the second case (e.g., Y_{12} or Y_{14}), the concentration coefficient is equal to zero only if the path coefficients are set to certain values (e.g., $\lambda_{21} = 0$ for Y_{12} ; $\lambda_{51} = 0$ or $\lambda_{54} = 0$ for Y_{15}). For our purpose, the exact values taken by the nonzero Y_{ij} are of minor importance; we rather focus on the elements that, such as Y_{13} , are structurally equal to zero, that is, that are equal to zero independently of the values taken by the path coefficients. More generally, it can be shown using (11) that Y_{ij} is identically equal to zero regardless of the numerical values of the path coefficients if and only if the three terms of the right-hand side of (11) are equal to zero, that is,

- (C₁) $K_{ij} = 0$ and $K_{ji} = 0$: neither region i nor region j has an effect on each other;
- (C₂) $K_{li}K_{lj} = 0$: regions i and j do not jointly influence region l , for all $l \neq i, j$.

In other words, $Y_{ij} = 0$ if and only if there are no such structures as $i \rightarrow j$, $i \leftarrow j$, or $i \rightarrow l \leftarrow j$ for any l in the structural graph: according to (C₁), there is no structural connection between i and j and, according to (C₂), regions i and j do not jointly influence a third region l . When a pair of

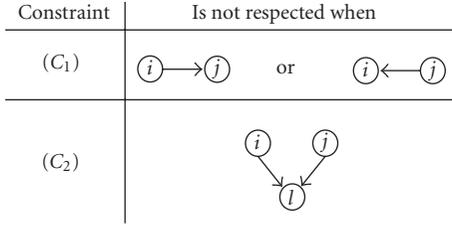


FIGURE 2: Structures that render either constraint (C₁) or (C₂) invalid for the pair i - j , thereby leading to $Y_{ij} \neq 0$ or, equivalently, $\Pi_{ij} \neq 0$.

regions is not directly connected in the structural model or both regions do not jointly point to any common region, the coefficient of partial correlation between these two regions is expected to be structurally equal to zero. On the other hand, if either condition is not satisfied, the corresponding coefficient of partial correlation is not structurally equal to zero (see Figure 2). Turning our attention back to the TP model, we see that, while regions VEC and SMA satisfy both (C₁) and (C₂) (implying $\Pi_{13} = 0$), regions VEC and PFC do not satisfy (C₁) (since we have $VEC \rightarrow PFC$) and regions VEC and IFG do not satisfy (C₂) (since we have $VEC \rightarrow IPL \leftarrow IFG$). As a matter of fact, all cases can be found in both the TP and the BF models, as shown in Tables 2 and 3. Using the aforementioned rule, we are able to retrieve the following structural constraints for partial correlation:

- (i) for the TP model: $Y_{13} = Y_{24} = Y_{25} = Y_{35} = 0$;
- (ii) for the BF model: $Y_{13} = Y_{14} = Y_{25} = 0$.

2.5. SEM and Partial Correlation. As correlation matrices are often easier to interpret than covariance matrices, we can decide to examine partial correlation matrices rather than concentration matrices. The partial correlation coefficient between two regions i and j , denoted by Π_{ij} , is here defined as a particular conditional correlation coefficient; it is the correlation between these two regions conditioned on the set of remaining regions, that is,

$$\Pi_{ij} = \text{Corr}[y_i, y_j \mid \mathbf{y}_{\mathbb{R} \setminus \{i,j\}}]. \quad (13)$$

There are hence $D(D-1)/2$ partial correlation coefficients (10 in our example) that form the D -by- D partial correlation matrix $\mathbf{\Pi} = (\Pi_{ij})$. $\mathbf{\Pi}$ can readily be calculated from \mathbf{Y} through the following relationship [17]:

$$\Pi_{ij} = -\frac{Y_{ij}}{\sqrt{Y_{ii} \cdot Y_{jj}}} \quad (14)$$

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

$$Y_{ij} = 0 \iff \Pi_{ij} = 0, \quad (15)$$

and what has been said about the relationship between the structural model and the structural zeros of the concentration matrix, namely conditions (C₁) and (C₂), also holds for

the partial correlation matrix. Furthermore, since the partial correlation coefficients are correlation coefficients, they are not influenced by any scale effect and remain between -1 and 1 ; for this reason, they are much easier to analyze and interpret than elements of the concentration matrix.

3. Validating Partial Correlation Structures

As we saw, a structural model has unique implications in terms of the structural pattern of partial correlation that can be expected from the data. Since the partial correlation matrix is a quantity that can be inferred from the data, we can use partial correlation analysis as a way to validate a structural model by comparing what is expected and what is observed.

3.1. Local Hypotheses. The approach consists of translating the structural hypotheses in terms of partial correlation. Indeed, according to Tables 2 and 3, the two structural models entail different hypotheses in term of partial correlation. For the theoretically preferred model, we have

$$\begin{aligned} \Pi_{13} &= 0 & (H_{TP1}), \\ \Pi_{24} &= 0 & (H_{TP2}), \\ \Pi_{25} &= 0 & (H_{TP3}), \\ \Pi_{35} &= 0 & (H_{TP4}), \end{aligned} \quad (16)$$

and, for the best fit model,

$$\begin{aligned} \Pi_{13} &= 0 & (H_{BF1}), \\ \Pi_{14} &= 0 & (H_{BF2}), \\ \Pi_{25} &= 0 & (H_{BF3}). \end{aligned} \quad (17)$$

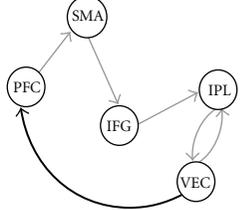
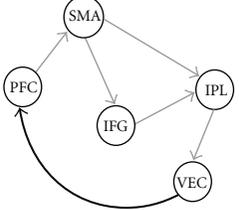
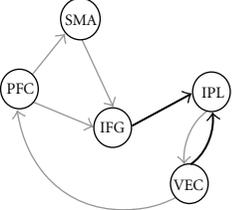
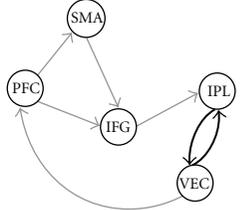
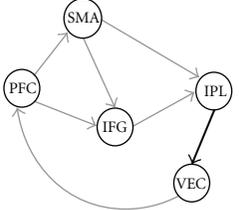
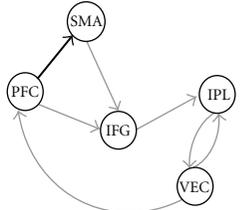
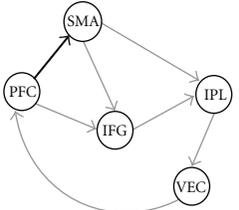
While some hypotheses are identical for both models, ($H_{TP1} = H_{BF1}$) and ($H_{TP3} = H_{BF3}$), others have no equivalent in the other model, such as (H_{TP1}), (H_{TP4}), and (H_{BF2}). The objective is then to infer the validity of these hypotheses with regard to the data.

3.2. Inference. Assessing the validity of the various hypotheses can be done by first estimating the partial correlation matrix. Inference of $\mathbf{\Pi}$ can be performed in a Bayesian framework using a numerical sampling scheme ([11, 13], see also the appendix). While direct computation of $p(\mathbf{\Pi} \mid \mathbf{z})$ is rather complex, this technique provides a simple approximation by sampling L (e.g., $L = 5000$) matrices $(\mathbf{\Pi}^{[l]})_{l=1, \dots, L}$ from $p(\mathbf{\Pi} \mid \mathbf{z})$. We then quantify the relevance of all hypotheses as follows. First, the probability of a coefficient Π_{ij} to be higher than 0 can be approximated by

$$p_{ij}^+ = \Pr(\Pi_{ij} > 0) \approx \frac{1}{L} \#\{l : \Pi_{ij}^{[l]} > 0\}, \quad (18)$$

where “#” stands for the cardinal of a set (i.e., its size). The probability p_{ij}^- of a coefficient to be lower than 0 could be approximated in a similar way, but only one of these two

TABLE 2: Partial correlation constraints in the TP and BF models (1/2). For each link between regions and each model, examination of whether (C_1) and (C_2) are satisfied.

Link $i-j$	TP model			BF model		
	(C_1) satisfied	(C_2) satisfied	Υ_{ij}	(C_1) satisfied	(C_2) satisfied	Υ_{ij}
VEC-PFC	no: 	yes	$\neq 0$	no: 	yes	$\neq 0$
VEC-SMA	yes	yes	$= 0$	yes	yes	$= 0$
VEC-IFG	yes	no: 	$\neq 0$	yes	yes	$= 0$
VEC-IPL	no: 	yes	$\neq 0$	no: 	yes	$\neq 0$
PFC-SMA	no: 	yes	$\neq 0$	no: 	yes	$\neq 0$

quantities need to be computed, since we have $p_{ij}^+ + p_{ij}^- = 1$. From there, the bearing of having $\Pi_{ij} > 0$ can be quantified by the log-odd ratio

$$e_{ij} = 10 \log_{10} \frac{p_{ij}^+}{p_{ij}^-} = 10 \log_{10} \frac{p_{ij}^+}{1 - p_{ij}^+}. \quad (19)$$

If e_{ij} is large and positive, we are more inclined to accept $\Pi_{ij} > 0$, while, if it is large and negative, we are more inclined to accept $\Pi_{ij} < 0$. Usually, a value of 10 dB can be considered as good evidence in favor of the hypothesis (see Table 4 for some relationships between p_{ij}^+ and e_{ij}). We finally take $|e_{ij}|$ as a measure of how Π_{ij} differs from zero and, hence, as a way to quantify the deviation of the data from hypothesis $\Pi_{ij} = 0$: values close to zero indicate a coefficient close to zero, while large values suggest a large coefficient value.

Since we here focus on the partial correlation constraints entailed by the structural models, (16) and (17), we only need the corresponding log odd ratios, summarized in Table 5. If all these hypotheses were true, then we would expect the absolute values of all log odd ratios to be lower than 10 dB. While this is the case for the three hypotheses related to the BF model, it is not the case for two of the four hypotheses related to the TP model: according to these results, (H_{TP2}) and (H_{TP4}) are rather unlikely to be true.

4. Discussion and Perspectives

In this paper, we further examined how partial correlation could be used to investigate effective connectivity in fMRI. We introduced theoretical fundamentals explaining why and in what measure the structure of the partial correlation matrix

TABLE 3: Partial correlation constraints in the TP and BF models (2/2). For each link between regions and each model, examination of whether (C_1) and (C_2) are satisfied.

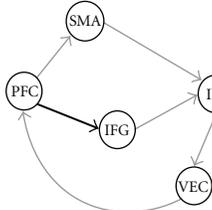
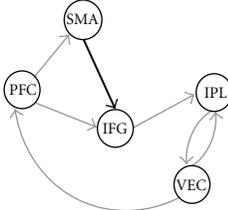
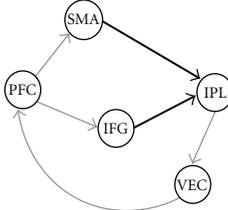
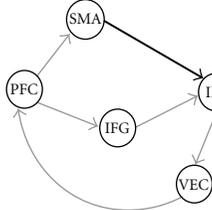
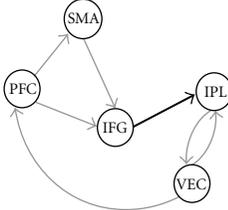
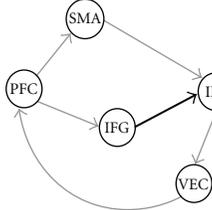
Link $i-j$	TP model			BF model		
	(C_1) satisfied	(C_2) satisfied	Y_{ij}	(C_1) satisfied	(C_2) satisfied	Y_{ij}
PFC-IFG	yes	yes	$= 0$	no: 	yes	$\neq 0$
PFC-IPL	yes	yes	$= 0$	yes	yes	$= 0$
SMA-IFG	no: 	yes	$\neq 0$	yes	no: 	$\neq 0$
SMA-IPL	yes	yes	$= 0$	no: 	yes	$\neq 0$
IFG-IPL	no: 	yes	$\neq 0$	no: 	yes	$\neq 0$

TABLE 4: Evidence e_{ij} and probability p_{ij}^+ (from [16]). For evidences of 3 and 6, p_{ij}^+ is only approximately equal to the fraction.

e_{ij} (dB)	p_{ij}^+
0	$1/2 = 0.50$
3	$2/3 \approx 0.67$
6	$4/5 = 0.80$
10	$10/11 \approx 0.91$
20	$100/101 \approx 0.99$
30	$1000/1001 \approx 0.999$
40	$10000/10001 \approx 0.9999$

can be related to a structural model. More precisely, we showed that, given a structural model, the partial correlation Π_{ij} between i and j is structurally equal to zero if and only if (C_1) neither region i nor region j has an effect on each

other, and (C_2) regions i and j do not jointly influence a third region l ; in other words, if and only if none of the following patterns are observed: $i \leftarrow j$, $i \rightarrow j$, or $i \rightarrow l \leftarrow j$ for any l . From there, the definition of a structural model entails a unique set of constraints that can be tested from the data, supporting or invalidating the plausibility of the corresponding structural model.

When examining the global relevance of partial correlation analysis to the investigation of effective connectivity, we must jointly consider two complementary effects, namely, the theoretical relationship between structural models and partial correlation matrices on the one hand and, on the other hand, the quality of the inference process. From a purely theoretical standpoint, this result shows that partial correlation analysis comes up as a combination of two effects. First, constraints (C_1) and (C_2) imply that

$$\Pi_{ij} = 0 \implies \neg(i \leftarrow j), \neg(i \rightarrow j), \quad (20)$$

TABLE 5: Real data. Relevance of hypotheses related to the TP and the BF models, respectively. Log odd ratios above a threshold of 10 dB are represented in bold.

Structural model	Constituting hypotheses	Structural constraints	$ e_{ij} $
TP	(H_{TP1})	$\Pi_{13} = 0$	1.6 dB
	(H_{TP2})	$\Pi_{24} = 0$	12.4 dB
	(H_{TP3})	$\Pi_{25} = 0$	9.7 dB
	(H_{TP4})	$\Pi_{35} = 0$	13.1 dB
BF	(H_{BF1})	$\Pi_{13} = 0$	1.6 dB
	(H_{BF2})	$\Pi_{14} = 0$	6.4 dB
	(H_{BF3})	$\Pi_{25} = 0$	9.7 dB

where \neg stands for the negation. In other words, a zero partial correlation between i and j implies the absence of a direct link between these two regions. Were there only (C_1) , this implication would be an equivalence and having $\Pi_{ij} \neq 0$ would imply a direct link between i and j . However, this is not true in general and, more specifically, for any pair of regions for which constraint (C_2) is satisfied. Such pairs are not connected but still have a nonzero partial correlation coefficient. As a consequence, all that can be said is that the set of set of pairs of regions with a zero partial correlations is a subset of the sets of pairs not directly connected in the structural model or, equivalently, that the set of pairs of regions connected in the structural model is a subset of the set of pairs of regions with a nonzero partial correlations. These features can easily be related to basic graph theoretic concepts. Condition (C_1) states that regions i and j are not neighbors; condition (C_2) states that i and j satisfy the so-called Wermuth condition [17]. As a consequence, the partial correlation constraints imposed by a structural model can be read off the graph obtained by adding undirected edges to eliminate all Wermuth configurations (for condition (C_2)) and transforming all arrows into undirected edges (for condition (C_1)). Such a graph is called the moral graph associated with the structural model. Depending on how many variables share common parents, the moral graph can be more or less close to the structural graph. For instance, in each of the two models used in this paper, condition (C_2) was only met once. Whether this is a general feature of fMRI data or only a characteristic induced by the structure selected remains to be cleared.

Another theoretical issue that needs to be tackled is the fact that having a partial correlation that is not constrained to 0 (e.g., Π_{14} for the theoretically preferred model) does not preclude its value to be equal to zero, due to a numerical coincidence. Indeed, (11) shows that specific values of \mathbf{K} and \mathbf{V} could be selected to induce $\Upsilon_{ij} = 0$ and, consequently, also $\Pi_{ij} = 0$. Even though this event is possible, it should be considered as rather unlikely, unless there is an underlying constraint at stake that forces the coefficient values to respect a certain relationship.

Another, more important issue deals with inference and how confident we can be in the partial correlation estimates and, critically, in the tests that their values are different from zero. The major difference between partial correlation and marginal correlation is that the former is obtained

by removing the effect of $D - 2$ regions as evidenced by (14). Importantly, the partialization process tends to decrease the value of correlation regardless of the exact relationship between the two variables and the conditioning set. Consequently, the values of partial correlation coefficients usually tend to be lower than their marginal counterparts; this is an observation that we have made consistently, and with only few exceptions. Also, as a rule of thumb, the posterior variance associated with a (marginal or partial) correlation coefficient (e.g., $\text{Var}[\Pi_{ij} \mid \mathbf{y}]$ for partial correlation) is roughly a decreasing function of the absolute value of its posterior mean (e.g., $E[\Pi_{ij} \mid \mathbf{y}]$ for partial correlation). For instance, it is asymptotically $(1 - \Pi_{ij}^2)/(N - 1)$ (which is indeed a decreasing function of Π_{ij}) for partial correlation and a similar result hold for marginal correlation [15]. A lower mean value therefore also implies a higher variance and, essentially, a bigger difficulty to discriminate a nonzero value from zero.

Altogether, these various factors, both theoretical and inferential, have different consequences on the relationship between the inferred pattern of partial correlation and the underlying structural model. Although we have observed a rather good agreement between expected and inferred patterns so far, in the lack of gold standard, these consequences must be further investigated.

Still, one of the main reasons why partial correlation analysis might become an important tool for the investigation of effective connectivity is that it is, to our knowledge, the only fully exploratory approach. Its key feature is its ability to retrieve local patterns of interaction. Indeed, while the method developed for the estimation of structural parameters, for example, (2), globally assesses the goodness of fit of the whole model and accordingly provides a general measure of it, partial correlation analysis provides a rather local assessment of effective connectivity, since the fact that two regions have a nonzero partial correlation depends on their connection with each other and of a potential connection with a common third region. For instance, in our example, while Bullmore et al. [7] concluded that the data did not contain enough evidence to prefer the BF model over the TP model (global statement), we showed that the TP model entails two partial correlation constraints ($\Pi_{24} = 0$ and $\Pi_{35} = 0$) that are rather unlikely to be true in the data (local statements). According to this result, we should discard the BF model or, at least, exert great caution

when using it. Furthermore, if one only had the theoretically preferred model and were testing it, the large log odd ratios corresponding to hypotheses $\Pi_{24} = 0$ and $\Pi_{35} = 0$ would hint that the corresponding constraints might not hold and that there might be a direct connection between regions PFC and IFG on the one hand and, on the other hand, between regions SMA and IPL.

In this paper, we determined whether certain coefficients could be considered as different from zero or not in a Bayesian framework. This led us to the use of the evidence e_{ij} of (19). While increasingly used, evidence admittedly remains rather uncommon in the brain imaging literature, where significance is often asserted with respect to a significance threshold, or P -value. It would therefore be tempting to propose a direct connection between P -values and evidence or, at least, interpret results of our Bayesian approach in terms of significance and P -value (see, e.g., [12]). Unfortunately, doing so is both inaccurate and misleading, because of the strong difference between Bayesian probability intervals and their frequentist counterparts, confidence intervals. Under the null hypothesis (H_0): $\Pi_{ij} = 0$, thresholding a statistic $\hat{\Pi}_{ij}$ at 10% in a frequentist framework (corresponding to a statistic of $P_{10\%}$) implies that, assuming that (H_0) is true, there is only 10% to obtain data with a statistic above the threshold, that is,

$$p(\hat{\Pi}_{ij} > P_{10\%} \mid H_0) = 0.10. \quad (21)$$

In this case, there is no mention whatsoever of any alternative hypothesis: we only assess how typical the data under consideration are. By contrast, thresholding a Bayesian probability at 10% means that we only consider cases where the alternative hypothesis (H_1) of (H_0): $\Pi_{ij} \leq 0$ has a probability of more than 0.9, that is,

$$p(H_1 \mid P_{ij}) > 0.90. \quad (22)$$

While a frequentist threshold of 10% might appear permissive, a Bayesian threshold of 10% is already conservative, since it implies that (H_1) is about 10 times more probable than (H_0). For more details on this topic, the reader can refer to Jaynes [18].

A last question is the possibility to apply partial correlation to other imaging modalities, such as electroencephalography (EEG) and magnetoencephalography (MEG). While the issue of removing the effect of other regions when considering the interactions between two regions remains relevant, whether partial correlation as defined here can provide a cogent solution remains to be investigated. One of the major properties of the fMRI signal is that, due to the convolution with the hemodynamic response, the temporal information that it conveys is usually considered as less relevant than in EEG or MEG. This is one of the major reasons why most EEG or MEG analyses are performed in the frequency domain. Of interest would therefore be to use partial correlation in this frequency domain. This analysis could be performed over time windows that are narrow enough to assume stationarity of the signal. How such an approach could be related to partial coherence [19, 20] remains to be clarified.

Appendix

Numerical Sampling Scheme

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

$$\mathbf{S} = \sum_{t=1}^T (z_t - \bar{z}_t)(z_t - \bar{z}_t)^t \quad (A.1)$$

is proportional to the sample covariance matrix, and \bar{z}_t is the temporal mean [21]. Calculation of the posterior probability density function (pdf) of the partial correlation matrix, $p(\Pi \mid \mathbf{z})$ cannot be performed in close form from this distribution. To approximate this distribution, we can nevertheless resort to the following sampling scheme [11, 13]. For sample l ,

- (1) sample $\Sigma^{[l]}$ according to its inverse Wishart distribution ([21], Appendix A);
- (2) calculate $\mathbf{Y}^{[l]} = (\Sigma^{[l]})^{-1}$, and $\Pi^{[l]}$ from $\mathbf{Y}^{[l]}$ according to (14).

Once a large number L of samples have 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,

$$E[\Pi_{ij} \mid \mathbf{z}] \approx M_{ij} = \frac{1}{L} \sum_{l=1}^L \Pi_{ij}^{[l]}, \quad (A.2)$$

$$\text{Var}[\Pi_{ij} \mid \mathbf{z}] \approx X_{ij} = \frac{1}{L} \sum_{l=1}^L (\Pi_{ij}^{[l]} - M_{ij})^2.$$

One can also compute evidence as explained in the main text.

References

- [1] W. Chen and S. Ogawa, "Principles of BOLD functional MRI," in *Functional MRI*, C. Moonen and P. Bandettini, Eds., pp. 103–113, Springer, Berlin, Germany, 1999.
- [2] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*, Sinauer, Sunderland, UK, 2004.
- [3] B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI," *Magnetic Resonance in Medicine*, vol. 34, no. 4, pp. 537–541, 1995.
- [4] K. J. Friston, "Functional and effective connectivity in neuroimaging: a synthesis," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 56–78, 1994.
- [5] A. R. McIntosh and F. Gonzalez-Lima, "Structural equation modeling and its application to network analysis in functional brain imaging," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 2–22, 1994.
- [6] F. Gonzalez-Lima and A. R. McIntosh, "Analysis of neural network interactions related to associative learning using structural equation modeling," *Mathematics and Computers in Simulation*, vol. 40, no. 1-2, pp. 115–140, 1995.

- [7] E. Bullmore, B. Horwitz, G. Honey, M. Brammer, S. Williams, and T. Sharma, "How good is good enough in path analysis of fMRI data?" *NeuroImage*, vol. 11, no. 4, pp. 289–301, 2000.
- [8] K. A. Bollen, *Structural Equation with Latent Variables*, Wiley-Interscience, New York, NY, USA, 1989.
- [9] R. Cudeck, K. J. Klebe, and S. J. Henly, "A simple Gauss-Newton procedure for covariance structure analysis with high-level computer languages," *Psychometrika*, vol. 58, no. 2, pp. 211–232, 1993.
- [10] G. Marrelec, J. Daunizeau, M. Péligrini-Issac, J. Doyon, and H. Benali, "Conditional correlation as a measure of mediated interactivity in fMRI and MEG/EEG," *IEEE Transactions on Signal Processing*, vol. 53, no. 9, pp. 3503–3516, 2005.
- [11] G. Marrelec, J. Doyon, M. Péligrini-Issac, and H. Benali, "Heading for data-driven measures of effective connectivity in functional MRI," in *Proceedings of the International Joint Conference on Neural Networks*, pp. 1528–1533, 2005.
- [12] G. Marrelec, A. Krainik, H. Duffau, et al., "Partial correlation for functional brain interactivity investigation in functional MRI," *NeuroImage*, vol. 32, no. 1, pp. 228–237, 2006.
- [13] G. Marrelec, B. Horwitz, J. Kim, M. Péligrini-Issac, H. Benali, and J. Doyon, "Using partial correlation to enhance structural equation modeling of functional MRI data," *Magnetic Resonance Imaging*, vol. 25, no. 8, pp. 1181–1189, 2007.
- [14] G. Marrelec, J. Kim, J. Doyon, and B. Horwitz, "Large-scale neural model validation of partial correlation analysis for effective connectivity investigation in functional MRI," *Human Brain Mapping*, vol. 30, no. 3, pp. 941–950, 2009.
- [15] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley Publications in Statistics, John Wiley & Sons, New York, NY, USA, 1958.
- [16] E. T. Jaynes, *Probability Theory: The Logic of Science. Vol. I: Principles and Elementary Applications*, Cambridge University Press, Cambridge, UK, 2003.
- [17] J. Whittaker, *Graphical Models in Applied Multivariate Statistics*, John Wiley & Sons, Chichester, UK, 1990.
- [18] E. T. Jaynes, "Confidence intervals vs. Bayesian intervals," in *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, W. L. Harper and C. A. Hooker, Eds., pp. 175–257, D. Reidel, Dordrecht, The Netherlands, 1976.
- [19] F. H. Lopes da Silva, J. E. Vos, J. Mooibroek, and A. V. Rotterdam, "Relative contributions of intracortical and thalamo-cortical processes in the generation of alpha rhythms, revealed by partial coherence analysis," *Electroencephalography and Clinical Neurophysiology*, vol. 50, no. 5-6, pp. 449–456, 1980.
- [20] J. R. Rosenberg, D. M. Halliday, P. Breeze, and B. A. Conway, "Identification of patterns of neuronal connectivity—partial spectra, partial coherence, and neuronal interactions," *Journal of Neuroscience Methods*, vol. 83, no. 1, pp. 57–72, 1998.
- [21] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis. Texts in Statistical Science*, Chapman & Hall, London, UK, 1998.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

