Investigation of two neural mass models for DCM-based effective connectivity inference in temporal epilepsy

Wentao Xiang, Ahmad Karfoul, Chunfeng Yang, Huazhong Shu, Régine Le Bouquin Jeannès

Investigation of two neural mass models for DCM-based effective connectivity inference in temporal epilepsy

Wentao Xiang[a,b,c], Ahmad Karfoul[b,c], Chunfeng Yang[c,d], Huazhong Shu[c,d], Régine Le Bouquin Jeannès[b,c,*]

[a] *Key Laboratory of Clinical and Medical Engineering, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, 211166, China*
[b] *Univ Rennes, Inserm, LTSI, UMR 1099, Rennes, F-35000, France*
[c] *Univ Rennes, Inserm, SEU, LIA - Centre de Recherche en Information Biomédicale Sino-français (CRIBs), F-35000, Rennes, France*
[d] *Laboratory of Image Science and Technology (LIST), School of Computer Science and Engineering, Southeast University, Nanjing, 210096, China*

Abstract

Background and objective: Recently, spectral Dynamic Causal Modelling (DCM) has been used increasingly to infer effective connectivity from epileptic intracranial electroencephalographic (iEEG) signals. In this context, the Physiology-Based Model (PBM), a neural mass model, is used as a generative model. However, previous studies have highlighted out the inability of PBM to properly describe iEEG signals with specific power spectral densities (PSDs). More precisely, PSDs that have multiple peaks around β and γ rhythms (*i.e.* spectral characteristics at seizure onset) are concerned.
Methods: To cope with this limitation, an alternative neural mass model, called the complete PBM (cPBM), is investigated. The spectral DCM and two recent variants are used to evaluate the relevance of cPBM over PBM.
Results: The study is conducted on both simulated signals and real epileptic iEEG recordings. Our results confirm that, compared to PBM, cPBM shows (i) more ability to model the desired PSDs and (ii) lower numerical complexity whatever the method.
Conclusions: Thanks to its intrinsic and extrinsic connectivity parameters as well as the input coming into the fast inhibitory subpopulation, the cPBM provides a more expressive model of PSDs, leading to a better understanding of epileptic patterns and DCM-based effective connectivity inference.

*Corresponding author. Tel.: +33 2 23 23 69 19. Fax: +33 2 23 23 69 17.
Email address: regine.le-bouquin-jeannes@univ-rennes1.fr (Régine Le Bouquin Jeannès)

Highlights

- The complete Physiology-Based Model (cPBM) allows to generate multiple Power Spectral Densities (PSDs) peaks and highly contributes to a better understanding of epileptic patterns than PBM.
- To the best of our knowledge, cPBM has never been introduced in the context of DCM. Bayesian model comparison was used to establish the identifiability of models of effective connectivity in the setting of DCM with cPBM.
- cPBM offers lower computational complexity better estimation accuracy compared to PBM.

Keywords: Epilepsy; Neural mass model; Effective connectivity; Power spectral density; Spectral dynamic causal modelling

1. Introduction

Epilepsy is a neurological disease that affects approximately 1% of the world population and is characterized by the repetition of seizures whose frequency and duration are variable [1]. The treatment for epileptic patients generally begins with medication. Recent studies have shown that up to 70% of children and adults with epilepsy can be successfully treated with Anti-Epileptic Drugs (AEDs) [2, 3]. However, there are still about 30% of epileptic patients who cannot be adequately controlled with AEDs, which means that (i) seizures are not (completely) suppressed and/or (ii) patients suffer from strong side-effects due to AEDs [4]. These patients are often diagnosed as drug-resistant or refractory epileptic patients [5, 6]. For these drug-resistant epilepsies, a surgical intervention can be considered as an alternative solution to remove the Epileptogenic Zone (EZ). This zone has been defined as the area of cortex primarily responsible for the generation of clinical seizures [7]. The standard pre-surgery evaluation includes clinical review, brain imaging with Magnetic Resonance Imaging (MRI), ElectroEncephaloGraphy (EEG) or intracranial EEG (iEEG) signals recordings, neuropsychological and psychiatric assessments [8]. Thanks to its high temporal resolution, the iEEG technique permits to capture very fast neuronal dynamics related to the epileptic seizure. This technique is considered as gold standard for EZ identification and allows a more precise brain exploration in spite of its invasiveness. In fact, seizure activity is not limited to the EZ but also may be modulated – or be mediated by – distributed dynamics in other involved regions. To better understand the organization of epileptic seizures in terms of onset and propagation, one may resort to effective connectivity [9] (*i.e.* causal effects of one neural system over another).

A variety of approaches have been proposed to infer effective connectivity. These methods can be divided into two categories, model-free and model-based

ones. In [10], authors stressed that effective connectivity analysis tends to be more hypothesis-driven (model-based) than data-driven (model-free). In this paper, we mainly focus on model-based approaches, which rely on two aspects: the generative model (to describe the observed signals) and the model comparison (to seek the best model structure among a set of predefined model structures based on the model evidence). Recently, among model-based techniques, Dynamic Causal Modelling (DCM) [11] is increasingly used in quantifying effective connectivity in the scope of epilepsy. Epileptic seizures always show some particular spectral characteristics, *i.e.* fast activities also referred as rapid discharges corresponding to power spectral densities (PSDs) with several peaks around β and γ frequency bands ($15 \sim 50$ Hz) at the seizure onset [12]. Therefore, researchers have been tempted to use spectral DCM [13, 14] to infer causality among brain regions in the context of epilepsy [15, 16, 17, 18] to cite a few, under the assumption of the generative model (*i.e.* neural mass model)—Canonical MicroCircuit (CMC) model proposed in [19, 20] or Physiology-Based Model (PBM) proposed in [12, 21, 22, 23].

Each population in the CMC model comprises four subpopulations: (i) superficial pyramidal, (ii) spiny stellate, (iii) inhibitory interneuron and (iv) deep pyramidal. More precisely, in [15, 24] authors used spectral DCM in the CMC model to infer causal coupling between brain areas in order to evaluate the relative contribution of changes in intrinsic (excitatory and inhibitory) connectivity and endogenous input. This study was conducted using EEG recordings acquired from two female patients with recurrent partial seizures. Authors concluded that changes in excitatory/inhibitory connections were able to explain the observed seizure activity and that these changes were reproduced over seizures. Later, authors in [25] proposed a Bayesian belief updating scheme to accelerate the convergence of the spectral DCM. The efficiency of the proposed approach in terms of model inversion compared to the standard DCM was reported based on EEG seizure recordings. In addition, authors in [16] used a Bayesian model reduction-based scheme in [26] to identify the key synaptic (*e.g.* synaptic time constant in deep pyramidal cells) and connectivity parameters

(*e.g.* self-inhibitory connection in inhibitory subpopulation) that underlie seizure onset in a rat model with induced epilepsy. These different studies [15, 24, 16], stated that the inhibitory subpopulation plays an important role in explaining the seizure activity under CMC model.

As emphasized in [21], the onset of seizures is often characterized by the appearance of a fast activity in mesial structures, but this activity cannot be modeled in Jansen's model [27]. Authors in [21] proposed another extension of Jansen's model, known as PBM. This model takes into account: (i) the respective role of (somatic) fast and (dendritic) slow inhibitory interneurons and (ii) the possible inhibition-related functional reorganizations that take place in the epileptic tissue. The PBM is particularly suitable for studying neuronal dynamics (especially fast activity) during epileptic seizures. With such classification, authors in [12, 23] argued that the PBM can be used to simulate iEEG signals in the hippocampus during the transition from interictal to fast ictal activity. As also stated in [28], the fast inhibitory interneurons acting on pyramidal cells can be viewed as a prominent mechanism in seizure triggering. PBM was considered to infer effective connectivity in the context of epilepsy using spectral DCM [17] and, to deal with the local maximization issue of the free energy, two strategies have been proposed in [18], the L-DCM where a local adjustment strategy scheme was utilized and the D-DCM where a deterministic annealing scheme was employed. However, PBM is difficult to fit PSDs with multiple peaks around β and γ rhythms and presents a high computational complexity due to the calculation of equilibrium point in Jacobian matrix [18] related to the predicted PSD.

Later, Zavaglia et al. [29] as well as Ursino et al. [30] proposed a complete Physiology-Based Model, abbreviated hereafter by cPBM, to simulate the multiple/multimodal PSD in cortical regions for simple motor tasks by adding a new negative self-loop as well as a new afferent input to the fast inhibitory interneurons. Also, the cPBM comprises a new extrinsic connectivity parameter modelling direct connections from pyramidal neurons in one population to fast inhibitory interneurons in another population. As indicated in [29, 30, 31], the

cPBM is much more suitable to produce two distinct resonance peaks and may be more consistent with the physiological reality than the PBM. To the best of our knowledge, cPBM has never been introduced in the context of epileptic seizures and DCM-based framework. In addition, contrary to PBM, cPBM sounds to be more attractive from a computational point of view as the system identification is performed through a linearization of the generative model and output equations around the null equilibrium point [30]. Therefore, thanks to the abovementioned properties that cPBM enjoys, this model is considered in the current study.

Following [18] and thanks to the ability of cPBM to simulate intracranial EEG signals having PSDs with multiple peaks, the motivation of this study is to investigate a potential relationship between the estimation of effective connectivity and the underlying generative model. Therefore, the rest of this paper is organized as follows. A detailed description of the cPBM model is given in section 2. Section 3 presents a full explanation of the cPBM model in the framework of spectral DCM, along with two recently proposed DCM algorithms, L-DCM and D-DCM algorithms [18]. Section 4 is dedicated to numerical experiments where a comparative study between DCM, L-DCM and D-DCM algorithms in the cPBM framework was firstly conducted using simulated iEEG data. Secondly, the three algorithms were compared in both PBM and cPBM frameworks using real iEEG signals. Finally, section 5 presents discussion.

2. The neural mass model

2.1. Complete Physiology-Based Model

Similarly to the PBM [21, 18], each population in the cPBM consists of four neural subpopulations: pyramidal neurons (P_p), excitatory interneurons (P_e), slow inhibitory interneurons (P_s) and fast inhibitory interneurons (P_f), respectively. Each subpopulation represents a group of neurons of the same type, which approximately share the same membrane potential and can be lumped together. These subpopulations are interconnected with excitatory

$\{C_{ll}^{ep}, C_{ll}^{pe}, C_{ll}^{sp}, C_{ll}^{fp}\}$ and inhibitory $\{C_{ll}^{ps}, C_{ll}^{fs}, C_{ll}^{pf}, C_{ll}^{ff}\}$ intrinsic connections, as shown in Fig. 1. The differential equations governing cPBM [30] describe the dynamic of a set of hidden states (*e.g.* $\boldsymbol{x}_l(t) = [x_{1,l}(t), \cdots, x_{14,l}(t)]^\intercal$) for a given population $pop_l, l \neq l'$ (l' is related to another population, $pop_{l'}$):

$$
\begin{aligned}
\dot{x}_{1,l}(t) &= x_{2,l}(t) \\
\dot{x}_{2,l}(t) &= G_{e,l}\omega_{e,l}S(x_{3,l}(t)) - 2\omega_{e,l}x_{2,l}(t) - \omega_{e,l}^2 x_{1,l}(t) \\
\dot{x}_{3,l}(t) &= C_{ll}^{pe}x_{5,l}(t) - C_{ll}^{ps}x_{8,l}(t) - C_{ll}^{pf}x_{11,l}(t) \\
\dot{x}_{4,l}(t) &= x_{5,l}(t) \\
\dot{x}_{5,l}(t) &= G_{e,l}\omega_{e,l}(S(x_{6,l}(t)) + K_{ll'}^{pp}S(x_{3,l'}(t-\tau))/C_{ll}^{pe} \\
&\quad + u_l^p(t)/C_{ll}^{pe}) - 2\omega_{e,l}x_{5,l}(t) - \omega_{e,l}^2 x_{4,l}(t) \\
\dot{x}_{6,l}(t) &= C_{ll}^{ep}x_{2,l}(t) \\
\dot{x}_{7,l}(t) &= x_{8,l}(t) \\
\dot{x}_{8,l}(t) &= G_{s,l}\omega_{s,l}S(x_{9,l}(t)) - 2\omega_{s,l}x_{8,l}(t) - \omega_{s,l}^2 x_{7,l}(t) \\
\dot{x}_{9,l}(t) &= C_{ll}^{sp}x_{2,l}(t) \\
\dot{x}_{10,l}(t) &= x_{11,l}(t) \\
\dot{x}_{11,l}(t) &= G_{f,l}\omega_{f,l}S(x_{14,l}(t)) - 2\omega_{f,l}x_{11,l}(t) - \omega_{f,l}^2 x_{10,l}(t) \\
\dot{x}_{12,l}(t) &= x_{13,l}(t) \\
\dot{x}_{13,l}(t) &= G_{e,l}\omega_{e,l}(u_l^f(t) + K_{ll'}^{fp}S(x_{3,l'}(t-\tau))) \\
&\quad - 2\omega_{e,l}x_{13,l}(t) - \omega_{e,l}^2 x_{12,l}(t) \\
\dot{x}_{14,l}(t) &= C_{ll}^{fp}x_{2,l}(t) - C_{ll}^{fs}x_{8,l}(t) - C_{ll}^{ff}x_{11,l}(t) + x_{13,l}(t)
\end{aligned}
\tag{1}
$$

Comparing PBM with cPBM as shown in Fig. 1, we can note some important differences as summarized in Table 1 [30, 22]. The values of the parameters specifying Eq. (1) are given in Table 2. According to Fig. 1, the two extrinsic random inputs, $u_l^p(t) = \sigma_l^p w_l^p(t)$ and $u_l^f(t) = \sigma_l^f w_l^f(t)$ in population pop_l, are independent and stand for white noises where $w_l^p(t)$ and $w_l^f(t)$ follow Gaussian probability distributions $\mathcal{N}(0, \alpha_l^p)$ and $\mathcal{N}(0, \alpha_l^f)$ respectively. To accommodate the amplitude of $u_l^p(t)$ and $u_l^f(t)$, we set the diffusion parameters σ_l^p and σ_l^f to

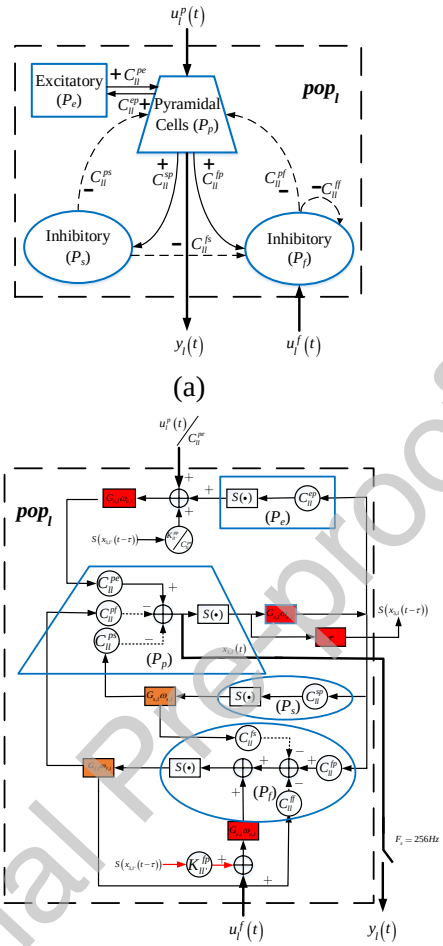


Figure 1. The population pop_l in cPBM [30] with **(a)** interactions between four neuronal subpopulations (*e.g.* pyramidal neurons (P_p), excitatory interneurons (P_e), slow inhibitory interneurons (P_s) and fast inhibitory interneurons (P_f), respectively). **(b)** block diagram for each hidden state as described in Eq. (1), where excitatory and inhibitory intrinsic connections $\{C_{ll}^{ep}, C_{ll}^{pe}, C_{ll}^{sp}, C_{ll}^{fp}\}$ and $\{C_{ll}^{ps}, C_{ll}^{fs}, C_{ll}^{pf}, C_{ll}^{ff}\}$, respectively. $S(\bullet)$ is the sigmoid function, $u_l^p(t)$ and $u_l^f(t)$ are the inputs, $y_l(t)$ is the output.

$20\sqrt{2}$ and 10 as shown in Table 2.

The output of cPBM for the l-th population is given by:

$$y_l(t) = x_{3,l}(t) \tag{2}$$

Table 1. Comparison of PBM and cPBM. Note that $u_l^f(t) = \sigma_l^f w_l^f(t)$ and $w_l^f(t) \sim \mathcal{N}\left(0, \alpha_l^f\right)$.

Case	PBM	cPBM
Sigmoid ($S(x(t))$)	$\frac{2e_0}{1+\exp(r_0(V_0-x(t)))}$	$\frac{2e_0}{1+\exp(r_0(V_0-x(t)))} - e_0$
Self-loop in P_f (C_{ll}^{ff})	×	√
Input in P_f ($u_l^f(t)$)	×	√
Extrinsic connection from P_p to P_f ($K_{ll'}^{fp}$)	×	√
Delay (τ)	×	√

where $y_l(t)$ reflects the mean membrane potential in pyramidal cells which induces the measured iEEG signal, recorded in an invasive way with high temporal resolution.

Table 2 lists the parameters priors used in Eq. (1) and Eq. (2), such prior distributions are introduced for model inversion (*e.g.* parameters estimation described in section 3.1.2) based on a Bayesian approach. For simplicity, these parameters can be divided into two categories depending whether the variance $\sigma_{\theta_i}^2$ is equal to zero or not: (i) fixed parameters (for which the prior variances are equal to 0), *i.e.* $\Phi_1 = \{C_{ll}^{ep}, ..., C_{ll}^{ff}, \sigma_l^p, \sigma_l^f, e_0, r, v_0, \tau\}$, and (ii) free parameters *i.e.* $\Phi_2 = \{\omega_{e,l}, \omega_{s,l}, \omega_{f,l}, G_{e,l}, G_{s,l}, G_{f,l}, \alpha_l^p, \alpha_l^f, K_{ll'}^{pp}, K_{ll'}^{fp}\}$, which, instead of being estimated directly, are estimated through the vector $\boldsymbol{\theta}_l = \left[\theta_{\omega_{e,l}}, \theta_{\omega_{s,l}}, \theta_{\omega_{f,l}}, \theta_{G_{e,l}}, \theta_{G_{s,l}}, \theta_{G_{f,l}}, \theta_{\alpha_l^p}, \theta_{\alpha_l^f}, \theta_{K_{ll'}^{pp}}, \theta_{K_{ll'}^{fp}}\right]^\intercal$. The i-th component, θ_i, of vector $\boldsymbol{\theta}_l$ is defined such that the parameter $P_i = V_i \exp(\theta_i)$ with $P_i \in \Phi_2$ (see Table 2). In fact, the free parameters Φ_2 are expected to vary during a transition from background brain activity to an epileptic seizure activity.

Table 2. Model parameters for $pop_l, l \neq l'$, corresponding to cPBM, where the values in red are extracted from [22], the values in blue from [30], and the values in black correspond to our proposal. Prior distributions of $P_i = V_i \exp(\theta_i)$, where V_i is a fixed default parameter value and $\theta_i \sim p(\theta_i) = \mathcal{N}\left(\mu_i, \sigma^2_{\theta_i}\right)$. If $\sigma^2_{\theta_i} = 0$, P_i is considered as constant.

Parameter (P_i)	Description (Unit)	Value (V_i)	Prior ($p(\theta_i)$)
	Synaptic time parameter (s)		
$\omega_{e,l}$	Excitatory	1/75	$\mathcal{N}(0, 1/32)$
$\omega_{s,l}$	Dendritic slow inhibitory	1/30	$\mathcal{N}(0, 1/32)$
$\omega_{f,l}$	Somatic fast inhibitory	1/100	$\mathcal{N}(0, 1/32)$
	Intrinsic connection within population		
C_{ll}^{ep}	$P_p \to P_e$	54	$\mathcal{N}(0, 0)$
C_{ll}^{pe}	$P_e \to P_p$	54	$\mathcal{N}(0, 0)$
C_{ll}^{sp}	$P_p \to P_s$	54	$\mathcal{N}(0, 0)$
C_{ll}^{ps}	$P_s \to P_p$	67.5	$\mathcal{N}(0, 0)$
C_{ll}^{fp}	$P_p \to P_f$	54	$\mathcal{N}(0, 0)$
C_{ll}^{fs}	$P_s \to P_f$	27	$\mathcal{N}(0, 0)$
C_{ll}^{pf}	$P_f \to P_p$	54	$\mathcal{N}(0, 0)$
C_{ll}^{ff}	$P_f \to P_f$	27	$\mathcal{N}(0, 0)$
	Extrinsic connection between populations		
$K_{ll'}^{pp}$	$P_p(pop_{l'}) \to P_p(pop_l)$	54	$\mathcal{N}(0, 1/4)$
$K_{ll'}^{fp}$	$P_p(pop_{l'}) \to P_f(pop_l)$	27	$\mathcal{N}(0, 1/4)$
	Input White Gaussian noise (s^{-1})		
σ_l^p	Diffusion	$20\sqrt{2}$	$\mathcal{N}(0, 0)$
σ_l^f	Diffusion	10	$\mathcal{N}(0, 0)$
α_l^p	Variance	1	$\mathcal{N}(0, 1/128)$
α_l^f	Variance	1	$\mathcal{N}(0, 1/128)$
	Sigmoid		
e_0	Sigmoid saturation (s^{-1})	2.5	$\mathcal{N}(0, 0)$
v_0	Activation threshold (mV)	0	$\mathcal{N}(0, 0)$
r	Sigmoid steepness (mV^{-1})	0.56	$\mathcal{N}(0, 0)$
	Time delay (ms)		
τ		10	$\mathcal{N}(0, 0)$
	Gain (mV)		
$G_{e,l}$	Synaptic excitation	5	$\mathcal{N}(0, 1/16)$
$G_{s,l}$	Dendritic slow inhibition	3	$\mathcal{N}(0, 1/16)$
$G_{f,l}$	Somatic fast inhibition	20	$\mathcal{N}(0, 1/16)$

2.2. Model structure and family

Based on the vectors of extrinsic coupling parameters $\boldsymbol{K}_{jk} = \left[K_{jk}^{pp}, K_{jk}^{fp}\right]^{\intercal}$ in pop_j and $\boldsymbol{K}_{kj} = \left[K_{kj}^{pp}, K_{kj}^{fp}\right]^{\intercal}$ in pop_k for cPBM (see Fig. 2 **(a)**), a model space with 16 model structures, M$m, m \in \{1, 2, ..., 16\}$, (see Fig. 2 **(b)**), can be defined. In fact, this model space describes all possible directional coupling configurations between the two populations. $\boldsymbol{\theta} = \text{vec}\left([\boldsymbol{\theta}_j, \boldsymbol{\theta}_k]^{\intercal}\right)$ denotes the vector of all parameters to be estimated under a given model structure Mm. When a coupling parameter (*e.g* $K_{ll'}^{pp}$ or $K_{ll'}^{fp}$, $\{l, l'\} \in \{j, k\}^2$, $l \neq l'$) is zero, the corresponding $\theta_{K_{ll'}^{pp}}$ or $\theta_{K_{ll'}^{fp}}$ is dropped from the vector of parameters $\boldsymbol{\theta}_l$. For instance, in the case of M5, the parameter $\theta_{K_{jk}^{pp}}$ is omitted in the vector $\boldsymbol{\theta}_j$, and the parameters $\theta_{K_{kj}^{pp}}$ and $\theta_{K_{kj}^{fp}}$ are omitted in the vector $\boldsymbol{\theta}_k$ since K_{jk}^{pp}, K_{kj}^{pp} and K_{kj}^{fp} are equal to zero. Furthermore, we can classify the 16 model structures into four families, corresponding to 4 types of causal architectures (see Fig. 2 **(c)**).

Therefore, for a given pair of observed signals, identifying the best model structure or family in the model space is the issue we have to deal with. Such issue can be solved by the DCM-based methods (see section 3).

3. Methods

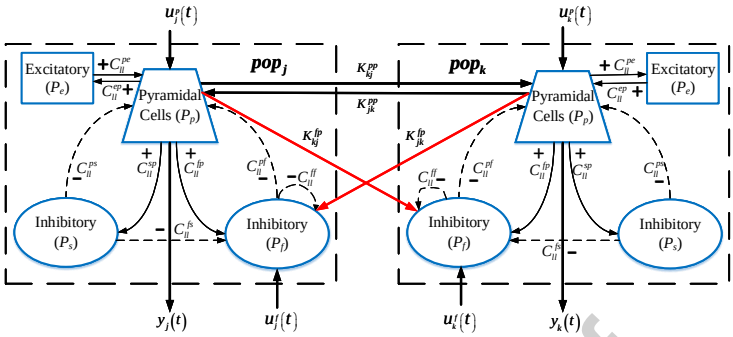
3.1. Spectral DCM

3.1.1. The predicted PSD

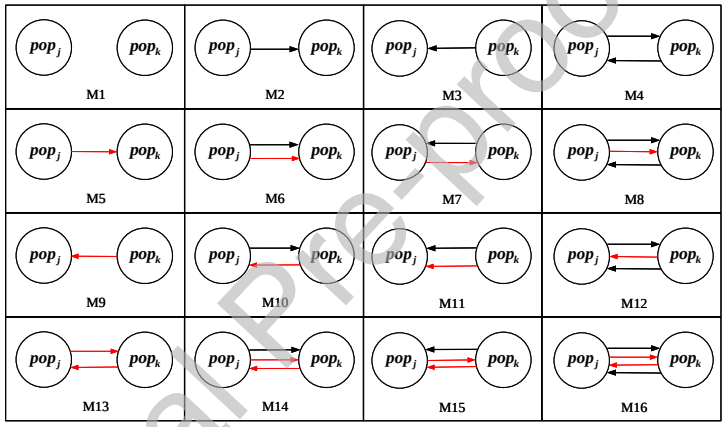
The cPBM for two potentially coupled populations represented by Eq. (1) and (2) can be described in a set of $(2 \times n_s)$ first-order delay differential equations ($n_s = 14$ is the number of state variables) and a two-component output vector $\boldsymbol{y}(t) = [y_j(t), y_k(t)]^{\intercal}$, as follows:

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \boldsymbol{f}\left(\boldsymbol{x}(t, \boldsymbol{\Psi}), \boldsymbol{\theta}\right) + \boldsymbol{D}\boldsymbol{u}(t) \\ \boldsymbol{y}(t) = \boldsymbol{Q}\boldsymbol{x}(t) \end{cases} \tag{3}$$

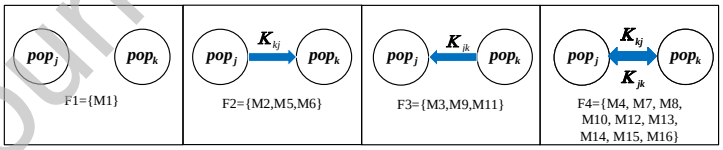
where $\boldsymbol{x}(t) = \text{vec}([\boldsymbol{x}_j(t), \boldsymbol{x}_k(t)]^{\intercal}) \in \mathbb{R}^{2n_s \times 1}$, $\boldsymbol{x}_l(t) = [x_{1,l}(t), \cdots, x_{n_s,l}(t)]^{\intercal} \in \mathbb{R}^{n_s \times 1}$ is the state vector associated to pop_l, $l \in \{j, k\}$. The nonlinear function $\boldsymbol{f}\left(\boldsymbol{x}(t, \boldsymbol{\Psi}), \boldsymbol{\theta}\right)$ is given by $\boldsymbol{f}\left(\boldsymbol{x}(t, \boldsymbol{\Psi}), \boldsymbol{\theta}\right) = \text{vec}([\boldsymbol{f}_j\left(\boldsymbol{x}(t, \boldsymbol{\Psi}), \boldsymbol{\theta}_j\right), \boldsymbol{f}_k\left(\boldsymbol{x}(t, \boldsymbol{\Psi}), \boldsymbol{\theta}_k\right)]^{\intercal}) \in \mathbb{R}^{2n_s \times 1}$. The



(a)



(b)



(c)

Figure 2. The effective connectivity between pop_j and pop_k in cPBM. **(a)** two potentially coupled populations; **(b)** the model space including the 16 model structures derived from possible extrinsic coupling(s) between the two populations; **(c)** the family space including four families and corresponding model structures ($\boldsymbol{K}_{jk} = \left[K_{jk}^{pp}, K_{jk}^{fp}\right]^{\mathsf{T}}$, $\boldsymbol{K}_{kj} = \left[K_{kj}^{pp}, K_{kj}^{fp}\right]^{\mathsf{T}}$).

components of $\boldsymbol{f}_l(\boldsymbol{x}(t,\boldsymbol{\Psi}),\boldsymbol{\theta}_l) \in \mathbb{R}^{n_s\times 1}$, $l \in \{j,k\}$, are the restrictions of the right terms of Eq. (1) on the restricted domains induced respectively by $u_l^p(t)=0$ and $u_l^f(t)=0$. Note that the first equation in Eq. (3) is a vector representation of the following component-wise differential equation:

$$\dot{x_i}(t) = f_i(\boldsymbol{x}(t,\boldsymbol{\Psi}(i,:)),\boldsymbol{\theta}) + (\boldsymbol{D}\boldsymbol{u})_i(t) \tag{4}$$

where $\forall i \in \{1,\cdots,2n_s\}$, $x_i(t)$, $f_i(\boldsymbol{x}(t,\boldsymbol{\Psi}(i,:)),\boldsymbol{\theta})$ and $(\boldsymbol{D}\boldsymbol{u})_i(t)$ denote the i-th component of the vectors $\boldsymbol{x}(t)$, $\boldsymbol{f}(\boldsymbol{x}(t,\boldsymbol{\Psi}),\boldsymbol{\theta})$ and $\boldsymbol{D}\boldsymbol{u}(t)$, respectively. For the sake of convenience, $\boldsymbol{\Psi}(i,:)$ denotes the i-th row of the delay matrix $\boldsymbol{\Psi} \in \mathbb{R}^{2n_s\times 2n_s}$ and $x_n(t-\Psi_{i,n})$ stands for the n-th component of the $2n_s$-th dimensional vector $\boldsymbol{x}(t,\boldsymbol{\Psi}(i,:))$. According to Eq. (1) and Eq. (4), the entries of $\boldsymbol{\Psi}$ are zeros excepted the four entries $\Psi_{9,6},\Psi_{25,6},\Psi_{10,5}$ and $\Psi_{26,5}$ which are equal to τ (*e.g.* $\Psi_{9,6}=\tau$ indicates that $x_6(t-\Psi_{9,6})$ is induced by $\dot{x}_9(t)$ in Eq. (4)). As far as the input matrix $\boldsymbol{D} \in \mathbb{R}^{2n_s\times 4}$ is concerned, all its entries are zeros excepted for $D_{9,1},D_{10,2},D_{25,3}$ and $D_{26,4}$ which are equal to $G_{e,j}w_{e,j}/C_{jj}^{pe}, G_{e,k}w_{e,k}/C_{kk}^{pe}$, $G_{e,j}w_{e,j}$ and $G_{e,k}w_{e,k}$, respectively. The input vector $\boldsymbol{u}(t)$ is equal to $\left[u_j^p(t),u_k^p(t),u_j^f(t),u_k^f(t)\right]^\intercal$. The output matrix $\boldsymbol{Q} \in \mathbb{R}^{2\times 2n_s}$ is an all-zero matrix excepted for two entries: $Q_{1,5}=Q_{2,6}=1$.

As we need an analytical expression (at least approximate) of the output PSD, a linearization of Eq. (3) around the equilibrium state, following [32, 13], using first-order Taylor expansion is given as follows:

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \tilde{\mathfrak{F}}(\boldsymbol{\theta})\,\boldsymbol{x}(t) + \boldsymbol{D}_1\boldsymbol{w}(t) \\ \boldsymbol{y}(t) = \boldsymbol{Q}\boldsymbol{x}(t) \end{cases} \tag{5}$$

where $\tilde{\mathfrak{F}}(\boldsymbol{\theta}) \simeq (\boldsymbol{I}+\boldsymbol{\Psi}\circ\mathfrak{F}(\boldsymbol{\theta}))^{-1}\mathfrak{F}(\boldsymbol{\theta})$ [33] ($\circ$ is Hadamard product) and $\boldsymbol{w}(t) = \left[w_j^p(t),w_k^p(t),w_j^f(t),w_k^f(t)\right]^\intercal$. Note that, $\mathfrak{F}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{\theta})}{\partial \boldsymbol{x}(t)}|_{\boldsymbol{x}(t)=\boldsymbol{0}} \in \mathbb{R}^{2n_s\times 2n_s}$ is the Jacobian matrix of $\boldsymbol{f}(\cdot,\boldsymbol{\theta})$, computed at $\boldsymbol{x}(t)=\boldsymbol{0}$ (*i.e.* the equilibrium point of Eq. (3) is equal to the null vector since the sigmoid function and $\boldsymbol{w}(t)$ are centered and the mean of is equal to zero as indicated in [30]). $\boldsymbol{D}_1 = (\boldsymbol{I}+\boldsymbol{\Psi}\circ\mathfrak{F}(\boldsymbol{x}_0,\boldsymbol{\theta}))^{-1}\boldsymbol{D}\mathrm{diag}(\boldsymbol{\sigma})$ with $\mathrm{diag}(\boldsymbol{\sigma})$ denoting a square diagonal matrix

whose diagonal corresponds to the elements of the vector $\boldsymbol{\sigma} = \left[\sigma_j^p, \sigma_k^p, \sigma_j^f, \sigma_k^f\right]^{\mathsf{T}}$. Then, the Fourier transform to both sides of Eq. (5) is applied:

$$\begin{cases} i2\pi\nu\,\breve{\boldsymbol{x}}\,(\nu) = \tilde{\mathfrak{F}}\,(\boldsymbol{\theta})\,\breve{\boldsymbol{x}}\,(\nu) + \boldsymbol{D}_1\breve{\boldsymbol{w}}\,(\nu) \\ \breve{\boldsymbol{y}}\,(\nu) = \boldsymbol{Q}\breve{\boldsymbol{x}}\,(\nu) = \boldsymbol{H}(\nu,\boldsymbol{\theta})\breve{\boldsymbol{w}}\,(\nu) \end{cases} \tag{6}$$

where $\boldsymbol{H}(\nu,\boldsymbol{\theta}) = \boldsymbol{Q}\Big(i2\pi\nu\boldsymbol{I} - \tilde{\mathfrak{F}}\,(\boldsymbol{\theta})\Big)^{-1}\boldsymbol{D}_1$, ν denotes the frequency variable and $\boldsymbol{H}(\nu,\boldsymbol{\theta}) \in \mathbb{R}^{2\times 4}$ is the transfer function ($\boldsymbol{I}$ is the identity matrix). The analytical expression of the PSD obtained from the parameters (named predicted PSD) is denoted by $\boldsymbol{G}\,(\nu,\boldsymbol{\theta}) = \begin{pmatrix} G_{jj}(\nu,\boldsymbol{\theta}) & G_{jk}(\nu,\boldsymbol{\theta}) \\ G_{kj}(\nu,\boldsymbol{\theta}) & G_{kk}(\nu,\boldsymbol{\theta}) \end{pmatrix} \in \mathbb{R}^{2\times 2}$ and computed from Eq. (6), as follows [13]:

$$\boldsymbol{G}\,(\nu,\boldsymbol{\theta}) = \boldsymbol{H}\,(\nu,\boldsymbol{\theta})\,\boldsymbol{G}_{\boldsymbol{w}}\,(\nu,\boldsymbol{\theta})\,\boldsymbol{H}^{\mathsf{H}}\,(\nu,\boldsymbol{\theta}) \tag{7}$$

where

$$\boldsymbol{G}_{\boldsymbol{w}}\,(\nu,\boldsymbol{\theta}) = \begin{pmatrix} G^{pp}_{w,jj}\,(\nu,\boldsymbol{\theta}) & G^{pp}_{w,jk}\,(\nu,\boldsymbol{\theta}) & G^{pf}_{w,jj}\,(\nu,\boldsymbol{\theta}) & G^{pf}_{w,jk}\,(\nu,\boldsymbol{\theta}) \\ G^{pp}_{w,kj}\,(\nu,\boldsymbol{\theta}) & G^{pp}_{w,kk}\,(\nu,\boldsymbol{\theta}) & G^{pf}_{w,kj}\,(\nu,\boldsymbol{\theta}) & G^{pf}_{w,kk}\,(\nu,\boldsymbol{\theta}) \\ G^{fp}_{w,jj}\,(\nu,\boldsymbol{\theta}) & G^{fp}_{w,jk}\,(\nu,\boldsymbol{\theta}) & G^{ff}_{w,jj}\,(\nu,\boldsymbol{\theta}) & G^{ff}_{w,jk}\,(\nu,\boldsymbol{\theta}) \\ G^{fp}_{w,kj}\,(\nu,\boldsymbol{\theta}) & G^{fp}_{w,kk}\,(\nu,\boldsymbol{\theta}) & G^{ff}_{w,kj}\,(\nu,\boldsymbol{\theta}) & G^{ff}_{w,kk}\,(\nu,\boldsymbol{\theta}) \end{pmatrix} = \begin{pmatrix} \alpha_j^p & 0 & 0 & 0 \\ 0 & \alpha_k^p & 0 & 0 \\ 0 & 0 & \alpha_j^f & 0 \\ 0 & 0 & 0 & \alpha_k^f \end{pmatrix} \tag{8}$$

denotes the PSD matrix associated with the white Gaussian noise vector $\boldsymbol{w}(t)$. Since the inputs $w_j^p(t), w_k^p(t), w_j^f(t)$ as well as $w_k^f(t)$ in pop_j and pop_k are assumed to be independent in/between populations, $\boldsymbol{G}_{\boldsymbol{w}}\,(\nu,\boldsymbol{\theta})$ is diagonal with diagonal entries, $G^{\xi\xi}_{w,ll}\,(\nu,\boldsymbol{\theta}) = \alpha_l^{\xi}$, $\xi \in \{p,f\}$ and $l \in \{j,k\}$, denoting the auto PSD of the input $w_l^{\xi}(t)$.

3.1.2. Model inversion

Concerning the sample PSD matrix, obtained from the observed iEEG signals, it is computed using the SPM spectral toolbox with a 12-pole AR process [13]. The sample PSD, denoted by $\tilde{\boldsymbol{G}}\,(\nu) = \begin{pmatrix} \tilde{G}_{jj}(\nu) & \tilde{G}_{jk}(\nu) \\ \tilde{G}_{kj}(\nu) & \tilde{G}_{kk}(\nu) \end{pmatrix} \in \mathbb{R}^{2\times 2}$ at each frequency ν, can also be considered as a noisy version of $\boldsymbol{G}\,(\nu,\boldsymbol{\theta})$:

$$\tilde{\boldsymbol{G}}\,(\nu) = \boldsymbol{G}\,(\nu,\boldsymbol{\theta}) + \boldsymbol{E}\,(\nu) \tag{9}$$

where $\boldsymbol{E}(\nu) = \begin{pmatrix} E_{jj}(\nu) & E_{jk}(\nu) \\ E_{kj}(\nu) & E_{kk}(\nu) \end{pmatrix}$ is the noise matrix. Let us consider N frequency bins $(\nu_i, i = 1, 2, ..., N)$ of $\tilde{\boldsymbol{G}}(\nu)$ stacked in a column vector $\tilde{\boldsymbol{g}}$, then following Eq. (9), $\tilde{\boldsymbol{g}}$ can be written:

$$\tilde{\boldsymbol{g}} = \boldsymbol{g}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon} \tag{10}$$

where

$$\begin{aligned}
&\tilde{\boldsymbol{g}} = [\tilde{\boldsymbol{g}}_{jj}, \tilde{\boldsymbol{g}}_{jk}, \tilde{\boldsymbol{g}}_{kj}, \tilde{\boldsymbol{g}}_{kk}]^\intercal \in \mathbb{R}^{4N\times 1} \\
&\tilde{\boldsymbol{g}}_{ll'} = [\tilde{G}_{ll'}(\nu_1), ..., \tilde{G}_{ll'}(\nu_N)]^\intercal \in \mathbb{R}^{N\times 1}, \{l, l'\} \in \{j, k\} \\
&\boldsymbol{g}(\boldsymbol{\theta}) = [\boldsymbol{g}_{jj}(\boldsymbol{\theta}), \boldsymbol{g}_{jk}(\boldsymbol{\theta}), \boldsymbol{g}_{kj}(\boldsymbol{\theta}), \boldsymbol{g}_{kk}(\boldsymbol{\theta})]^\intercal \in \mathbb{R}^{4N\times 1} \\
&\boldsymbol{g}_{ll'}(\boldsymbol{\theta}) = [G_{ll'}(\nu_1, \boldsymbol{\theta}), ..., G_{ll'}(\nu_N, \boldsymbol{\theta})]^\intercal \in \mathbb{R}^{N\times 1} \\
&\boldsymbol{\varepsilon} = [\boldsymbol{\varepsilon}_{jj}, \boldsymbol{\varepsilon}_{jk}, \boldsymbol{\varepsilon}_{kj}, \boldsymbol{\varepsilon}_{kk}]^\intercal \in \mathbb{R}^{4N\times 1} \\
&\boldsymbol{\varepsilon}_{ll'} = [E_{ll'}(\nu_1), ..., E_{ll'}(\nu_N)]^\intercal \in \mathbb{R}^{N\times 1}
\end{aligned}$$

$\boldsymbol{\varepsilon}$ is the vector of prediction error with $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}})$, where the posterior covariance $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}$ is related to a hyper-parameter vector $\boldsymbol{\lambda}$ under Gaussian assumptions [34]. According to Eq. (10), the vector of parameters $\boldsymbol{\theta}$ for a given model structure Mm can be obtained by a Bayesian model inversion step using the model evidence $p(\tilde{\boldsymbol{g}}|\mathrm{M}m)$ [34, 35]. The logarithm of model evidence, $\ln p(\tilde{\boldsymbol{g}}|\mathrm{M}m)$, can be expressed as:

$$\ln p(\tilde{\boldsymbol{g}}|\mathrm{M}m) = F_m + KL(q(\boldsymbol{\theta}), p(\boldsymbol{\theta}|\tilde{\boldsymbol{g}}, \mathrm{M}m)) \tag{11}$$

where KL is the Kullback-Leibler divergence, $p(\boldsymbol{\theta}|\tilde{\boldsymbol{g}}, \mathrm{M}m)$ is the posterior distribution of the vector $\boldsymbol{\theta}$, $q(\boldsymbol{\theta}) \sim \mathcal{N}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\right)$ and F_m is the well-called free energy (of the model structure Mm). Since KL is non negative, then $\forall q(\boldsymbol{\theta})$: $\ln p(\tilde{\boldsymbol{g}}|\mathrm{M}m) \geq F_m$, the equality being verified if $KL(q(\boldsymbol{\theta}), p(\boldsymbol{\theta}|\tilde{\boldsymbol{g}}, \mathrm{M}m)) = 0$, *i.e.* if $q(\boldsymbol{\theta}) = p(\boldsymbol{\theta}|\tilde{\boldsymbol{g}}, \mathrm{M}m)$ (almost everywhere), a maximum value of F_m can be always seen as a lower bound of $\ln p(\tilde{\boldsymbol{g}}|\mathrm{M}m)$. The definition of free energy is given as follows [34]:

$$F_m = \int q(\boldsymbol{\theta}) \ln p(\tilde{\boldsymbol{g}}, \boldsymbol{\theta}|\mathrm{M}m)\, d\boldsymbol{\theta} - \int q(\boldsymbol{\theta}) \ln q(\boldsymbol{\theta})\, d\boldsymbol{\theta} \tag{12}$$

Maximizing the free energy, F_m, is usually done in an iterative way using a variational scheme [34], named as the variational EM algorithm[1].

3.1.3. Model and family selection

Once the above model inversion step is achieved, a model selection step is performed. This step consists in selecting, in the predefined model space, the best model structure underlying the observed data. Recall that an optimal model structure is the one defined in terms of the maximum free energy compared to the free energies of the other considered model structures. Note that causality identification (*i.e.* effective connectivity identification) in cPBM can also be based on the family identification. Now, this causality identification depends on the vectors of coupling parameters $\boldsymbol{K}_{kj}$ and $\boldsymbol{K}_{jk}$. As an example, the family F2 (*i.e.* $\boldsymbol{K}_{jk} = \left[K_{jk}^{pp}, K_{jk}^{fp}\right]^{\intercal} = \mathbf{0}$) includes three model structures, *i.e.* M2, M5, and M6 (see Fig. 2 **(c)**), in cPBM with one or two element(s) of the vector $\boldsymbol{K}_{kj} = \left[K_{kj}^{pp}, K_{kj}^{fp}\right]^{\intercal}$ that are non-zero and positive. To decide the most plausible family in cPBM, the family posterior probability of each family Ff, $p\left(\tilde{\boldsymbol{g}}|\mathrm{F}f\right)$ is computed. This posterior probability is equal to the sum of model evidence $p\left(\tilde{\boldsymbol{g}}|\mathrm{M}m\right)$ times the prior $p\left(\mathrm{M}m|\mathrm{F}f\right)$ [36, 37]:

$$p\left(\tilde{\boldsymbol{g}}|\mathrm{F}f\right) = \sum_{\mathrm{M}m \in \mathrm{F}f} p\left(\tilde{\boldsymbol{g}}|\mathrm{M}m\right) p\left(\mathrm{M}m|\mathrm{F}f\right) \tag{13}$$

where the prior $p\left(\mathrm{M}m|\mathrm{F}f\right)$ is given as follows (see Fig. 2 **(c)**):

$$\begin{aligned}
&p\left(\mathrm{M1}|\mathrm{F1}\right) = 1, \mathrm{M1} \in \mathrm{F1} \\
&p\left(\mathrm{M}m|\mathrm{F2}\right) = \tfrac{1}{3}, \forall \mathrm{M}m \in \mathrm{F2} = \{\mathrm{M2}, \mathrm{M5}, \mathrm{M6}\} \\
&p\left(\mathrm{M}m|\mathrm{F3}\right) = \tfrac{1}{3}, \forall \mathrm{M}m \in \mathrm{F3} = \{\mathrm{M3}, \mathrm{M9}, \mathrm{M11}\} \\
&p\left(\mathrm{M}m|\mathrm{F4}\right) = \tfrac{1}{9}, \forall \mathrm{M}m \in \mathrm{F4} = \{\mathrm{M4}, \mathrm{M7}, \mathrm{M8}, \mathrm{M10}, \mathrm{M12}, \mathrm{M13}, \mathrm{M14}, \mathrm{M15}, \mathrm{M16}\}
\end{aligned} \tag{14}$$

[1]Although we have referred to the variational scheme as a variational EM algorithm, technically it is known as Variational Laplace. Variational Laplace is a generic form of variational Bayes under the Laplace assumption (*i.e.* under the assumption that the prior and posterior are Gaussian). This enables one to estimate the posterior expectation and covariance of both the parameters $\boldsymbol{\theta}$ and hyper-parameter vector $\boldsymbol{\lambda}$. Please see [34] for details.

Similar to the logarithm of model evidence, the family free energy, FF_f index which can be defined as $\ln p\left(\tilde{\boldsymbol{g}}|\mathrm{F}f\right)$, is given by:

$$\begin{aligned} FF_f = \ln p\left(\tilde{\boldsymbol{g}}|\mathrm{F}f\right) &= \ln \sum_{\mathrm{M}m\in\mathrm{F}f} p\left(\tilde{\boldsymbol{g}}|\mathrm{M}m\right) p\left(\mathrm{M}m|\mathrm{F}f\right) \\ &\simeq \ln \sum_{\mathrm{M}m\in\mathrm{F}f} \exp(F_m) + \ln p\left(\mathrm{M}m|\mathrm{F}f\right) \end{aligned} \tag{15}$$

where $p\left(\tilde{\boldsymbol{g}}|\mathrm{M}m\right) \simeq \exp(F_m)$ (see Eq. (11)). Likewise, the best family in cPBM is identified as the one with the maximum family free energy. The pseudo-code summarizing spectral DCM algorithm is given in **Algorithm 1**. Note that, the variational EM algorithm stops when the convergence condition is met for eight successive iterations. The convergence condition is fulfilled when the difference in the estimation of F_m exhibits a value smaller than or equal to a predefined threshold (2) or a maximum number of iterations (60) is reached.

3.2. Sub-estimation problem

As shown in Fig. 3 **(a)**, the variational EM algorithm in DCM is highly sensitive to the initialization, so it may lead to a sub-estimation of the free energy and consequently may lead to a mis-estimation of the optimal model structure/family.

Algorithm 1 (Spectral) DCM algorithm

(a) Considering the current model structure Mm based on $\boldsymbol{\theta}$
(b) Initialization $\boldsymbol{\theta}_0 = \mathrm{vec}\left([\boldsymbol{\theta}_{0,j}, \boldsymbol{\theta}_{0,k}]^{\mathsf{T}}\right)$, $\boldsymbol{\lambda}_0$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}$, $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}_0}$
(c) Perform the following EM steps:
Repeat until convergence or a maximum number of iterations
E-step: $(\hat{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}) \leftarrow \underset{fixed\ \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}}{argmax}\ F_m(\hat{\boldsymbol{\theta}}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}})$
M-step: $(\boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}) \leftarrow \underset{fixed\ \hat{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}}{argmax}\ F_m(\hat{\boldsymbol{\theta}}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}})$
End
$F_m = F_m(\hat{\boldsymbol{\theta}}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}})$
(d) Turn to next model structure Mm, repeat from step (a)
(e) Model/Family selection to identify the optimal model structure/family

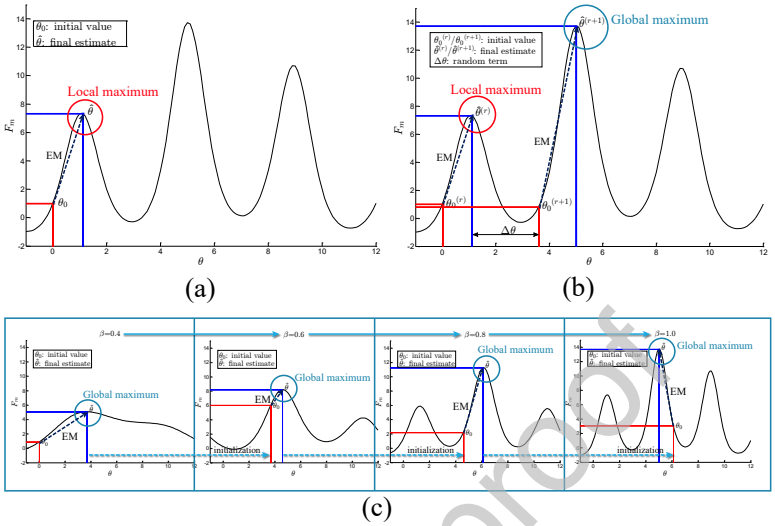


Figure 3. Scheme of three algorithms to maximize the free energy F_m, under considering Mm with **(a)** DCM, **(b)** L-DCM and **(c)** D-DCM ($\beta = 0.4$, $\beta = 0.6$, $\beta = 0.8$ and $\beta = 1$) respectively.

3.3. L-DCM

To deal with the sub-estimation problem in DCM, the L-DCM algorithm employing a local adjustment strategy was proposed in [18], and the scheme of L-DCM to maximize the free energy is shown in Fig. 3 **(b)**. As indicated in [18], the sub-estimation, when two populations are considered, relies on a threshold, $Th_{ll} = \sum_{\nu=\nu_l^a}^{\nu_l^b} \gamma^2 \tilde{g}_{ll}(\nu)^2$ and a prediction error, $Er_{ll} = \sum_{\nu=\nu_l^a}^{\nu_l^b} \left(g_{ll}\left(\nu, \hat{\theta}\right) - \tilde{g}_{ll}\left(\nu\right)\right)^2$, $l \in \{j, k\}$, for each population, where $\tilde{g}_{ll}\left(\nu\right)$ and $g_{ll}\left(\nu, \hat{\boldsymbol{\theta}}\right)$ stand for the sample PSD computed using the SPM spectral toolbox with a 12-pole AR process [13] and the estimation of the predicted PSD obtained by substituting the model parameters estimate $\hat{\boldsymbol{\theta}}$ using the variational EM algorithm, respectively. According to [18], we set $\gamma = 0.1$, $\nu_l^a = \nu_l^m - 10$ and $\nu_l^b = \nu_l^m + 10$ with ν_l^m standing for the frequency bin corresponding to the maximum value of $\tilde{g}_{ll}\left(\nu\right)$. A good estimation of the PSD $g_{ll}\left(\nu, \hat{\boldsymbol{\theta}}\right)$ is obtained when $Er_{ll} \leq Th_{ll}$, $l \in \{j, k\}$. However, in case this inequality is not fulfilled, three possible situations

can be figured out: situation ①: $Er_{jj} \leq Th_{jj}$, $Er_{kk} > Th_{kk}$; situation ②: $Er_{jj} > Th_{jj}$, $Er_{kk} \leq Th_{kk}$; situation ③: $Er_{jj} > Th_{jj}$, $Er_{kk} > Th_{kk}$. The L-DCM algorithm proposed in [18] refined the variational EM estimation results by performing few supplementary well-conditioned variational EM sweeps with a local adjustment strategy for the initialization. The initial guess for the $(r+1)$-th EM sweep $\boldsymbol{\theta}_0^{(r+1)} = \text{vec}\left(\left[\boldsymbol{\theta}_{0,j}^{(r+1)}, \boldsymbol{\theta}_{0,k}^{(r+1)}\right]^{\mathsf{T}}\right)$ is then locally adjusted, $\boldsymbol{\theta}_0^{(r+1)} = \hat{\boldsymbol{\theta}}^{(r)} + \Delta\boldsymbol{\theta}$, where $\hat{\boldsymbol{\theta}}^{(r)} = \text{vec}\left(\left[\hat{\boldsymbol{\theta}}_j^{(r)}, \hat{\boldsymbol{\theta}}_k^{(r)}\right]^{\mathsf{T}}\right)$ and $\Delta\boldsymbol{\theta}$ is a random vector dropping from zero-mean normal distribution with a predefined variance. By taking into account (i) the encountered situation (*i.e.* ① or ② or ③) and (ii) the model structure (*i.e.* Mm), $\Delta\boldsymbol{\theta}$ is given as follows:

$$\Delta\boldsymbol{\theta} = \begin{cases} \text{vec}\left([\mathbf{0}, \Delta\boldsymbol{\theta}_k]^{\mathsf{T}}\right), \forall \text{M}m \in \{\text{F1}, \text{F2}\} \text{ under } ① \\ \text{vec}\left([\Delta\boldsymbol{\theta}_j, \mathbf{0}]^{\mathsf{T}}\right), \forall \text{M}m \in \{\text{F1}, \text{F3}\} \text{ under } ② \\ \text{vec}\left([\Delta\boldsymbol{\theta}_j, \Delta\boldsymbol{\theta}_k]^{\mathsf{T}}\right), \text{otherwise} \end{cases} \tag{16}$$

where $\Delta\boldsymbol{\theta}_j \sim \mathcal{N}\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$ and $\Delta\boldsymbol{\theta}_k \sim \mathcal{N}\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$, $\boldsymbol{I}$ is the identity matrix and σ^2 is a parameter that was empirically set to 0.02. The pseudo-code of L-DCM algorithm can be seen in **Algorithm 2**.

Algorithm 2 L-DCM algorithm

(a) Considering the current model structure Mm based on $\boldsymbol{\theta}$
(b) Initialization $r = 1$, $\boldsymbol{\theta}_0^{(r)} = \text{vec}\left([\boldsymbol{\theta}_{0,j}, \boldsymbol{\theta}_{0,k}]^{\mathsf{T}}\right)$, $\boldsymbol{\lambda}_0$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}$, $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}_0}$
(c) Perform the following EM steps:
Repeat until convergence or a maximum number of iterations
E-step: $(\hat{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}) \leftarrow \underset{fixed\ \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}}{argmax}\ F_m(\hat{\boldsymbol{\theta}}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}})$
M-step: $(\boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}) \leftarrow \underset{fixed\ \hat{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}}{argmax}\ F_m(\hat{\boldsymbol{\theta}}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}})$
End
$F_m^{(r)} = F_m(\hat{\boldsymbol{\theta}}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}})$
(d) If $r <= 6$ and situation ① or ② or ③ occurs,
Repeat from step (c) by $\boldsymbol{\theta}_0^{(r+1)} = \hat{\boldsymbol{\theta}}^{(r)} + \Delta\boldsymbol{\theta}$ as a new initial value and $r = r + 1$;
Otherwise stop
(e) $F_m = \max\{F_m^{(r)}\}, r = 1, 2, ..., 6$
(f) Turn to next model structure Mm, repeat from step (a)
(g) Model/Family selection to identify the optimal model structure/family

3.4. D-DCM

Another way to deal with the sub-estimation issue in spectral DCM relies essentially on a deterministic annealing scheme [38, 39], leading to a new algorithm called D-DCM hereafter. More precisely, the deterministic annealing scheme, inspired from the thermodynamic theory, is based on tracking the optimum of the objective function under consideration from a high temperature wherein the objective function is smoothed (*i.e.* it may have one global optimum, see Fig. 3 **(c)** $\beta = 0.4$) to low temperature wherein the objective function returns to its initial form (see Fig. 3 **(c)** $\beta = 1$). By introducing a temperature related parameter, denoted by β (temperature $T = \frac{1}{\beta}$), the objective function (*e.g.* the free energy) F_m in Eq. (12) can be rewritten as:

$$F_m = \int q(\boldsymbol{\theta}) \ln p(\tilde{\boldsymbol{g}}, \boldsymbol{\theta} | \mathrm{M}m)\, d\boldsymbol{\theta} - \frac{1}{\beta} \int q(\boldsymbol{\theta}) \ln q(\boldsymbol{\theta})\, d\boldsymbol{\theta} \tag{17}$$

The variational EM algorithm [34] is used to perform such maximization in spectral DCM framework. In this way, one guarantees a good initialization through decreasing temperature values (*e.g.* β varying from very small values up to 1) as indicated in Fig. 3 **(c)**. The pseudo-code of the D-DCM algorithm is given in **Algorithm 3**.

Algorithm 3 D-DCM algorithm

(a) Considering the current model structure Mm based on $\boldsymbol{\theta}$
(b) Set $\beta \leftarrow \beta_{\text{int}}$ $(0 < \beta_{\text{int}} < 1)$
(c) Initialization $\boldsymbol{\theta}_0 = \text{vec}\left([\boldsymbol{\theta}_{0,j}, \boldsymbol{\theta}_{0,k}]^{\mathsf{T}}\right)$, $\boldsymbol{\lambda}_0$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}$, $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}_0}$
(d) Perform the following EM steps:
Repeat until convergence or a maximum number of iterations
E-step: $(\hat{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}) \leftarrow \underset{fixed\ \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}}{argmax}\ F_m(\hat{\boldsymbol{\theta}}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}, \beta)$
M-step: $(\boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}) \leftarrow \underset{fixed\ \hat{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}}{argmax}\ F_m(\hat{\boldsymbol{\theta}}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}, \beta)$
End
$F_m = F_m(\hat{\boldsymbol{\theta}}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}, \beta)$
(e) $\beta_{new} = \beta + const$
(f) If $\beta_{new} <= 1$
Repeat from step (d) with $\boldsymbol{\theta}_0 = \hat{\boldsymbol{\theta}}$ from β_{new} as a new initial value;
Otherwise stop
(g) Turn to next model structure Mm, repeat from step (a)
(h) Model/Family selection to identify the optimal model structure/family

4. Results

4.1. Simulations

In a previous study, we have performed a comparative evaluation between the L-DCM, D-DCM and DCM algorithms in inferring effective connectivity among epileptic neural populations at the seizure onset using PBM [18]. Hereafter, the simulated iEEG data are generated using the cPBM. Two connectivity configurations, the unidirectional and the bidirectional influence, between the two neural populations are considered. Several model structures can fulfill unidirectional and bidirectional coupling configurations between two populations (see Fig. 2 **(b)**). Therefore, model structures sharing the same coupling pattern (unidirectional/bidirectional) are considered to belong to the same family as shown in Fig. 2 **(c)**. Furthermore, two families are also considered: (i) the family F2 composed of model structures M2, M5 and M6 reflecting unidirectional coupling from pop_j to pop_k (see section 4.1.1) and (ii) the family F4 composed of model structures M4, M7, M8, M10, M12, M13, M14, M15 and M16 reflecting bidirectional coupling between the two populations. For simplicity, only M8 and M16 were used to generate data. These models were taken to be representative of F4 (see section 4.1.2).

4.1.1. Unidirectional influence

As mentioned previously, the family F2 = $\{\mathrm{M2}, \mathrm{M5}, \mathrm{M6}\}$ is considered here. Figs. 4 **(a-i)**, **(b-i)** and **(c-i)** show the three different model structures (corresponding to M2, M5 and M6 respectively) and their related vector $\boldsymbol{\theta}$. Note that, for each of model structure in F2, we simulated 100 trials of 10-second length signals by integrating delay differential equations [33] and using Runge-Kutta 4-th order method to solve Eq. (3) at a sampling frequency of 256 Hz and the values of the parameters in $\boldsymbol{\theta}$ are chosen in such a way that the model output, $\boldsymbol{y}(t)$, is equivalent to an epileptic iEEG signal with a peak around 25 Hz in its PSD observed at the seizure onset [12, 18]. Fig. 4 **(a-c)** depicts the temporal signature of the simulated iEEG and the sample auto/cross PSDs for M2, M5 and M6, respectively.

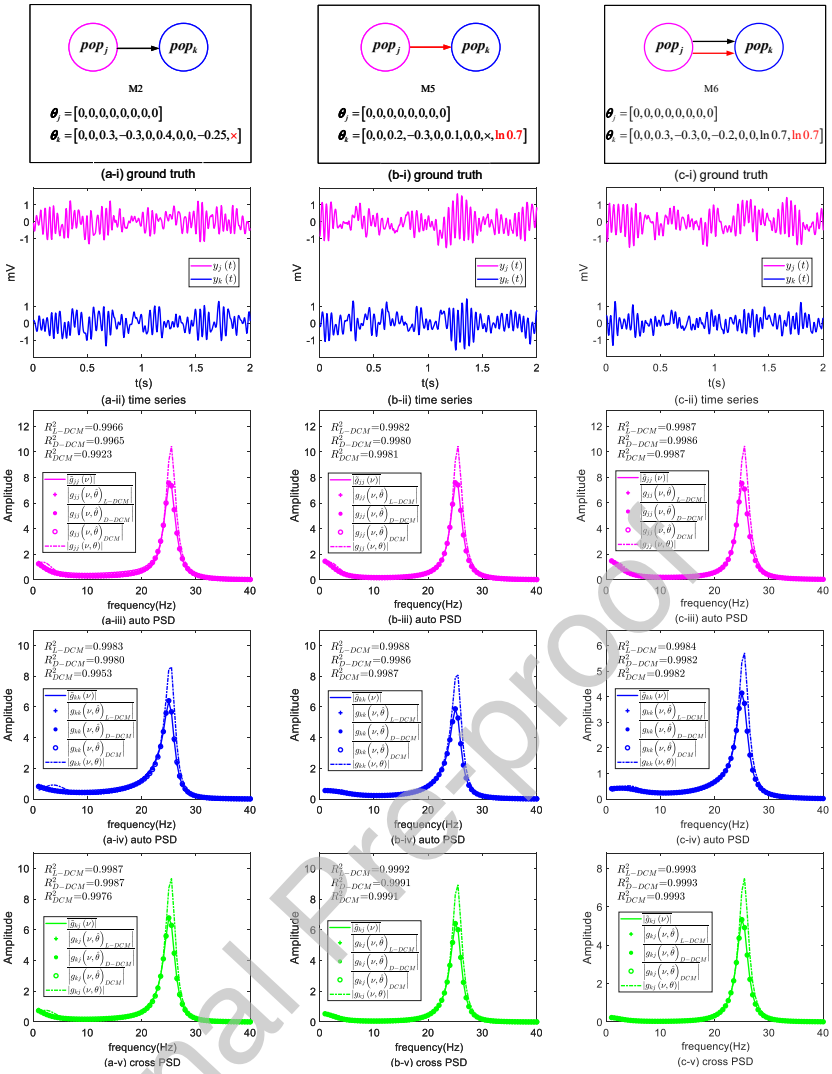


Figure 4. **(a-i)**, **(b-i)**, **(c-i)**: Three ground truths M2, M5 and M6 for unidirectional influence, the black arrows correspond to excitatory connections, whereas the red arrows correspond to extrinsic connections that target inhibitory populations, the corresponding vector $\boldsymbol{\theta} = \text{vec}\left([\boldsymbol{\theta}_j, \boldsymbol{\theta}_k]^\intercal\right)$, where $\boldsymbol{\theta}_j = \left[\theta_{\omega_{e,j}}, \theta_{\omega_{s,j}}, \theta_{\omega_{f,j}}, \theta_{G_{e,j}}, \theta_{G_{s,j}}, \theta_{G_{f,j}}, \theta_{\alpha_j^p}, \theta_{\alpha_j^f}\right]^\intercal$ and $\boldsymbol{\theta}_k = \left[\theta_{\omega_{e,k}}, \theta_{\omega_{s,k}}, \theta_{\omega_{f,k}}, \theta_{G_{e,k}}, \theta_{G_{s,k}}, \theta_{G_{f,k}}, \theta_{\alpha_k^p}, \theta_{\alpha_k^f}, \theta_{K_{kj}^{pp}}, \theta_{K_{kj}^{fp}}\right]^\intercal$. Simulated time signals displaying epileptic activities (*i.e.* **(a-ii)**, **(b-ii)**, **(c-ii)**), corresponding auto PSD amplitudes (*i.e.* **(a-iii, a-iv)**,**(b-iii, b-iv)**,**(c-iii, c-iv)**) and cross PSD amplitudes (*i.e.* **(a-v)**,**(b-v)**,**(c-v)**) for the three ground truths M2, M5 and M6, respectively. The sample PSD, denoted by $\tilde{g}_n(\nu)$, $n \in \{jj, kk, kj\}$, is computed from the simulated EEG signals based on the non-linear model Eq. (1) and the SPM spectral toolbox [13], $g_n(\nu, \boldsymbol{\theta})$ is the predicted PSD with the correct $\boldsymbol{\theta}$ based on the linearized model Eq. (5), $g_n\left(\nu, \hat{\boldsymbol{\theta}}\right)_{L-DCM}$, $g_n\left(\nu, \hat{\boldsymbol{\theta}}\right)_{D-DCM}$ and $g_n\left(\nu, \hat{\boldsymbol{\theta}}\right)_{DCM}$ represent the estimated PSDs with $\hat{\boldsymbol{\theta}}$ (under the optimized model structure) using L-DCM, D-DCM and DCM respectively. The coefficients of determination R^2_{L-DCM}, R^2_{D-DCM} and R^2_{DCM} stand for the measures of the goodness of fit between sample and estimated PSDs averaged over 100 trials, where $R^2{}_\kappa = 1 - \frac{\sum_\nu \left(\tilde{g}_n(\nu) - g_n\left(\nu, \hat{\boldsymbol{\theta}}\right)_\kappa\right)^2}{var(\tilde{g}_n(\nu))}$, $\kappa \in \{L-DCM, D-DCM, DCM\}$.

Table 3. The four most plausible model structures and model identification using simulated iEEG signals in the context of unidirectional influence in cPBM, the values in bold indicate the best performance under the true model structure.

Ground truth	Method	The four most plausible model structures (Model identification count over 100 trials)			
M2	L-DCM	M2 (38/100)	M6 (35/100)	M16 (13/100)	M10 (13/100)
	D-DCM	M2 (**48/100**)	M6 (35/100)	M10 (7/100)	M16 (7/100)
	DCM	M16 (59/100)	M2 (4/100)	M6 (14/100)	M10 (22/100)
M5	L-DCM	M5 (57/100)	M6 (14/100)	M7 (19/100)	M8 (8/100)
	D-DCM	M5 (71/100)	M6 (9/100)	M7 (17/100)	M8 (0/100)
	DCM	M5 (**76/100**)	M6 (6/100)	M7 (14/100)	M8 (4/100)
M6	L-DCM	M6 (87/100)	M2 (0/100)	M8 (12/100)	M10 (0/100)
	D-DCM	M6 (**99/100**)	M2 (0/100)	M10 (0/100)	M12 (0/100)
	DCM	M6 (97/100)	M2 (0/100)	M10 (0/100)	M16 (0/100)

Model selection

Table 3 shows the identification rates[2] computed over 100 trials for the three methods in identifying model structures M2, M5 and M6 when these models were successively considered as a ground truth. For each algorithm and each simulated model structure, the identification rates of only the four most plausible model structures among the sixteen structures depicted in Fig. 2 **(b)** are reported in this table. Regarding the ground truth M2, both D-DCM and L-DCM algorithms are able to identify the true model structure (*e.g.* M2) with identification rates of 48% and 38% respectively. This fact does not hold true for the DCM algorithm

[2]The identification rate of Mm refers to the degrees of preference of Mm computed over 100 trials. The identifiability of different models was assessed using the Bayesian model comparison at the group level where the fixed effects analysis was conducted [36]. More precisely, for each model in a set of plausible sixteen model structures, the median (less prone to outliers than the mean) of the maximized free energy was computed over 100 trials and then used to define the confusion matrices (more details are given in **Appendix: A**).

where the M16 model structure with an identification rate of 59% is rather retained than the M2 model structure. As far as ground truth M5 is concerned, all algorithms are able to identify the true model structure (*e.g.* M5). However, the DCM algorithm shows a slight superiority (76%) especially compared to the D-DCM algorithm (71%). When the true model structure is M6, D-DCM, DCM and L-DCM algorithms succeed in identifying M6 with identification rates of 99%, 97% and 85%, respectively.

To sum up, false detection in the context of model identification is generally highly related to a spurious estimation of one or more coupling links between neural populations under study (more details are presented in section **Support materials S1**). For instance, when M2 is the ground truth, estimating M6 instead of M2 for DCM, L-DCM and D-DCM approaches is mainly due to the fact that estimation errors during the optimization process could lead to some small non-null value of the coupling parameter K_{jk}^{fp} (the red line in Fig. 2 **(a)**). To deal with such problem, spurious connections can be removed through a thresholding step (more details are presented in **Appendix: B**). On the other hand, in the case of false detection, one can interestingly note that, even if a wrong model structure is identified (*e.g.* M6 instead of M2 in our example), this structure generally belongs to the same family as the right model structure. This shows that the algorithms are still able to quantify the direction of coupling between populations under study (*e.g.* M6 and M2 belong to the same family F2, see Fig. 2 **(c)**) which is the main goal of effective connectivity. Therefore, resorting to family identification analysis would provide a new insight regarding the estimation of effective connectivity between the neural populations.

The performance of the three algorithms is also evaluated using the coefficients of determination R^2_{L-DCM}, R^2_{D-DCM} and R^2_{DCM}, which stand for the measures of the goodness of fit, computed between the sample and estimated PSDs, as depicted in Figure 4 **(a-c)**. This figure shows comparable behaviors of the three algorithms with slight superiority of the L-DCM one, since the L-DCM algorithm applies a local adjustment strategy based on the estimation errors [18].

Family selection

As just mentioned, an analysis of family identification could be preferable to circumvent the false detection encountered in the model identification due to spurious estimation of coupling parameters. This idea comes from the observation that some wrong identified model structures belong to the same family as the target model structure. Since the main goal of effective connectivity is to point out the influence of one neural population on another one, it could be interesting to quantify, in addition to model identification, the ability of the algorithms to identify the correct family. Table 4 shows both the family identification rate and the execution time (in minutes) over the 100 trials for L-DCM, D-DCM and DCM methods for the previous ground truths M2, M5 and M6, respectively. According to this table, the D-DCM algorithm shows higher performance in identifying the true family (*e.g.* F2) than the DCM and L-DCM algorithms for M2 and M6.

Table 4. Family selection over 100 trials using simulated iEEG signals in the context of unidirectional influence in cPBM, the values in bold indicate the best performance.

Ground truth	Method	Family identification count over 100 trials (Averaged run time over 100 trials (mn))			
		F1	F2	F3	F4
M2	L-DCM	0/100 (0.4)	73/100 (1.9)	0/100 (1.0)	27/100 (3.4)
	D-DCM	0/100 (0.2)	**83/100** (1.0)	0/100 (0.9)	17/100 (2.9)
	DCM	0/100 (0.1)	18/100 (0.6)	0/100 (0.6)	82/100 (2.0)
M5	L-DCM	0/100 (0.1)	71/100 (0.9)	0/100 (0.4)	29/100 (2.1)
	D-DCM	0/100 (0.1)	80/100 (0.5)	0/100 (0.6)	20/100 (2.2)
	DCM	0/100 (0.1)	**82/100** (0.3)	0/100 (0.3)	18/100 (1.7)
M6	L-DCM	0/100 (0.2)	87/100 (0.8)	0/100 (0.7)	13/100 (2.3)
	D-DCM	0/100 (0.1)	**99/100** (0.5)	0/100 (0.8)	1/100 (2.5)
	DCM	0/100 (0.1)	97/100 (0.4)	0/100 (0.5)	3/100 (1.8)

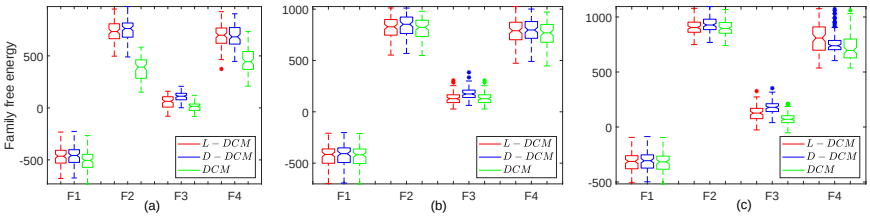


Figure 5. Boxplots of the family free energy ($FF_f, f = 1, 2, 3, 4$) values over 100 trials using L-DCM, D-DCM and DCM for the four families (unidirectional influence) in cPBM. **(a-c)**: the three ground truths (see Figure 4 **(a-i)**, **(b-i)**, **(c-i)**) corresponding to M2, M5 and M6, respectively.

Indeed, for M2, the D-DCM algorithm provides an identification rate of 83% compared to 73% and 18% for the L-DCM and DCM approaches respectively. Furthermore, the D-DCM algorithm provides, for M6, an identification rate of 99% compared to 97% and 87% for the DCM and L-DCM approaches respectively. Regarding ground truth M5, all algorithms succeed in identifying the true family (*e.g.* F2 family) with a slight preference for DCM. As far as the computation time is concerned, both L-DCM and D-DCM algorithms show, as expected, higher computation time compared to DCM in Table 4. In fact, this higher computation time is essentially due to the additional EM sweeps in the L-DCM (see Fig. 3 **(b)**) and to the adopted deterministic annealing scheme in D-DCM (see Fig. 3 **(c)**).

This result is assessed also in terms of the family free energy as depicted in Fig. 5. According to Fig. 5 **(a)** for ground truth M2, both L-DCM and D-DCM approaches show higher family free energies compared to the DCM algorithm. More particularly, for the true family (*e.g.* F2), the D-DCM and L-DCM algorithms provide a higher family evidence than the DCM one. This figure together with Table 4 highlight the fact that the DCM algorithm suffers from higher false family identification rate. Regarding the correct model M5 and M6 in Figs. 5 **(b-c)**, all methods are able to identify the true family F2 with comparable family identification rates.

4.1.2. Bidirectional influence

Family F4 comprises all model structures reflecting bidirectional coupling between pop_j and pop_k, as shown in Fig. 2 **(c)**. The performance of L-DCM, D-DCM and DCM in identifying F4 is evaluated in three different ground truths presented in Fig. 6 **(a-i)**, **(b-i)** and **(c-i)**, where the model structure M8 and two variants of the model structure M16 were used. More precisely, symmetric extrinsic couplings (*e.g.* $K^{pp}_{jk} = K^{pp}_{kj}$, $K^{fp}_{jk} = K^{fp}_{kj}$, see Fig. 6 **(b-i)** for the values of the parameters) between the two populations were used giving rise to $M16_s$ and asymmetric extrinsic couplings (*e.g.*$K^{pp}_{jk} \neq K^{pp}_{kj}$, $K^{fp}_{jk} \neq K^{fp}_{kj}$, see Fig. 6 **(c-i)** for the values of the parameters) were used giving rise to $M16_{as}$. For each ground truth, 100 trials of 10-second length signals at a sampling frequency of 256 Hz were simulated. The vector of parameters $\boldsymbol{\theta}$ in model structure M8 and $M16_s$ is tuned such that the model output, $\boldsymbol{y}(t)$, is a fast sinusoidal activity with only one peak around 25 Hz (Fig. 6 **(a-b)**) similar to the epileptic fast activity [12, 18]. Regarding the symmetric nature of $M16_s$, the vectors of parameters $\boldsymbol{\theta}_j$ and $\boldsymbol{\theta}_k$ associated to pop_j and pop_k, respectively, are effectively equal. However, these two vectors are not identical in the case of ground truth $M16_{as}$. In Fig. 6 **(c)** we represent, under $M16_{as}$ configuration, the epileptic-like activity $\boldsymbol{y}(t)$, the auto (cross) PSDs of the two populations displaying two peaks (around 20 Hz and 28 Hz), similar to possible epileptic activity at the seizure onset.

Model selection

Table 5 gives the four most plausible model structures and the model identification rate over 100 Monte Carlo trials for the bidirectional influence under the three ground truths M8, $M16_s$ and $M16_{as}$, respectively. According to this table, all techniques globally succeed in identifying the correct model structure. Furthermore, L-DCM outperforms D-DCM and DCM in the ground truth M8 (85% identification rate *vs* 80% and 79% respectively) whereas the D-DCM approach outperforms DCM and L-DCM in the ground truth $M16_s$ (96% identification rate *vs* 94% and 80% respectively). Finally, all methods can identify M16 with a 100% identification rate in the ground truth $M16_{as}$. More

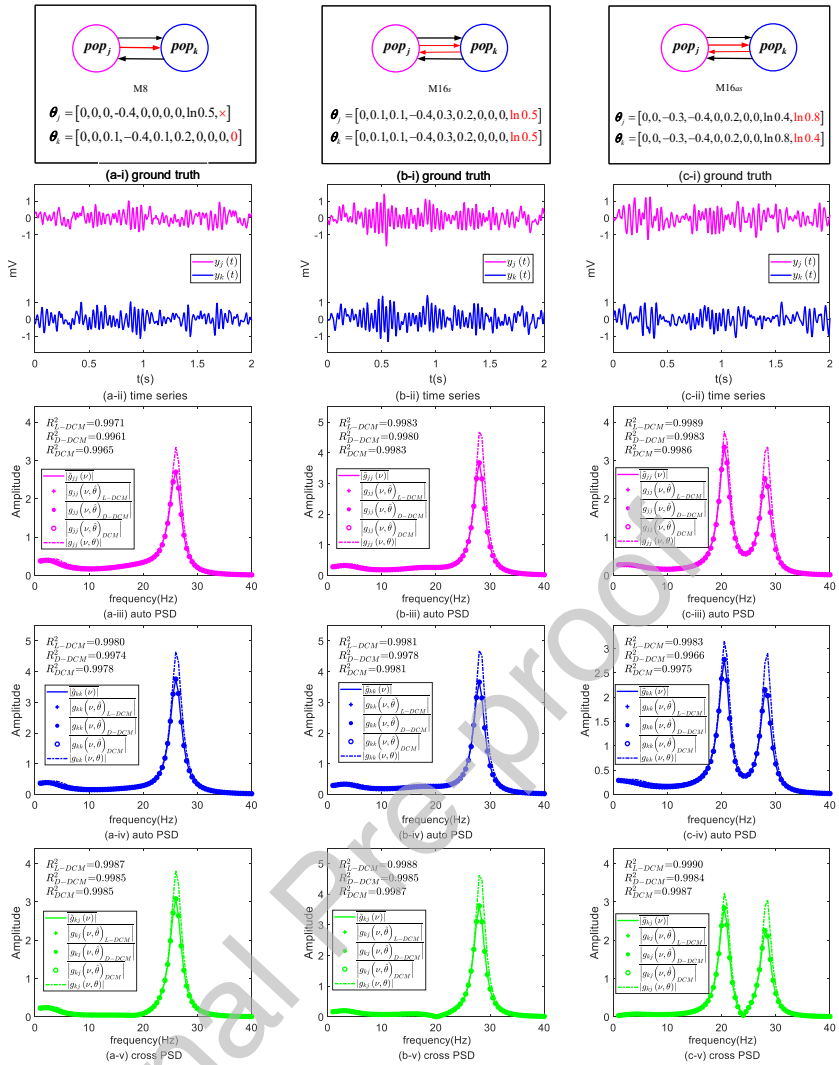


Figure 6. **(a-i)**, **(b-i)**, **(c-i)**: Three ground truths M8, M16$_s$ and M16$_{as}$ for bidirectional influence, the black arrows correspond to excitatory connections, whereas the red arrows correspond to extrinsic connections that target inhibitory populations, the corresponding vector $\boldsymbol{\theta} = \text{vec}\left([\boldsymbol{\theta}_j, \boldsymbol{\theta}_k]^\intercal\right)$, where $\boldsymbol{\theta}_j = [\theta_{\omega_{e,j}}, \theta_{\omega_{s,j}}, \theta_{\omega_{f,j}}, \theta_{G_{e,j}}, \theta_{G_{s,j}}, \theta_{G_{f,j}}, \theta_{\alpha_j^p}, \theta_{\alpha_j^f}, \theta_{K_{jk}^{pp}}, \theta_{K_{jk}^{fp}}]^\intercal$ and $\theta_k = [\theta_{\omega_{e,k}}, \theta_{\omega_{s,k}}, \theta_{\omega_{f,k}}, \theta_{G_{e,k}}, \theta_{G_{s,k}}, \theta_{G_{f,k}}, \theta_{\alpha_k^p}, \theta_{\alpha_k^f}, \theta_{K_{kj}^{pp}}, \theta_{K_{kj}^{fp}}]^\intercal$. Simulated time signals displaying epileptic activities (*i.e.* **(a-ii)**, **(b-ii)**, **(c-ii)**), corresponding auto PSD amplitudes (*i.e.* **(a-iii, a-iv)**,**(b-iii, b-iv)**,**(c-iii, c-iv)**) and cross PSD amplitudes (*i.e.* **(a-v)**,**(b-v)**,**(c-v)**) for the three ground truths M8, M16$_s$ and M16$_{as}$, respectively. The sample PSD, denoted by $\tilde{g}_n(\nu)$, $n \in \{jj, kk, kj\}$, is computed from the simulated EEG signals based on the non-linear model Eq. (1) and the SPM spectral toolbox [13], $g_n(\nu, \boldsymbol{\theta})$ is the predicted PSD with the correct $\boldsymbol{\theta}$ based on the linearized model Eq. (5), $g_n\left(\nu, \hat{\boldsymbol{\theta}}\right)_{L-DCM}$, $g_n\left(\nu, \hat{\boldsymbol{\theta}}\right)_{D-DCM}$ and $g_n\left(\nu, \hat{\boldsymbol{\theta}}\right)_{DCM}$ represent the estimated PSDs with $\hat{\boldsymbol{\theta}}$ (under the optimized model structure) using L-DCM, D-DCM and DCM respectively. The coefficients of determination R^2_{L-DCM}, R^2_{D-DCM} and R^2_{DCM} stand for the measures of the goodness of fit between sample and estimated PSDs averaged over 100 trials, where $R^2{}_\kappa = 1 - \frac{\sum_\nu \left(\tilde{g}_n(\nu) - g_n\left(\nu, \hat{\boldsymbol{\theta}}\right)_\kappa\right)^2}{var(\tilde{g}_n(\nu))}$, $\kappa \in \{L-DCM, D-DCM, DCM\}$.

Table 5. The four most plausible model structures and model identification using simulated iEEG signals in the context of bidirectional influence in cPBM. $M16_s$: symmetrical structure. $M16_{as}$: asymmetrical structure, the values in bold indicate the best performance under the true model structure.

Ground truth	Method	The four most plausible model structures (Model identification count over 100 trials)			
M8	L-DCM	M8 (**85/100**)	M16 (14/100)	M12 (1/100)	M2 (0/100)
	D-DCM	M8 (80/100)	M16 (19/100)	M12 (1/100)	M6 (0/100)
	DCM	M8 (79/100)	M16 (21/100)	M12 (0/100)	M6 (0/100)
$M16_s$	L-DCM	M16 (80/100)	M8 (11/100)	M12 (9/100)	M6 (0/100)
	D-DCM	M16 (**96/100**)	M8 (2/100)	M12 (2/100)	M6 (0/100)
	DCM	M16 (94/100)	M12 (1/100)	M8 (5/100)	M6 (0/100)
$M16_{as}$	L-DCM	M16 (**100/100**)	M10 (0/100)	M14 (0/100)	M15 (0/100)
	D-DCM	M16 (**100/100**)	M10 (0/100)	M4 (0/100)	M14 (0/100)
	DCM	M16 (**100/100**)	M10 (0/100)	M14 (0/100)	M15 (0/100)

details are presented in section **Support materials S2**. Moreover, in terms of PSD estimation, all methods show again good estimation accuracy which is clearly assessed by the obtained coefficients of determination depicted in Fig. 6 **(a-c)**.

Family selection

Regarding the family selection and the averaged execution time for each family over 100 trials, they are displayed in Table 6. All techniques perfectly identify the correct family. Furthermore, Fig. 7 depicts the boxplots of family free energy over the 100 trials using the three methods. Clearly, the maximal median of the family free energy is globally obtained for F4 whatever the method and the ground truth. Furthermore, D-DCM show relatively higher boxplots of family free energy than L-DCM and DCM for all ground truths and families. Let

Table 6. Family selection over 100 trials using simulated iEEG signals in the context of bidirectional influence in cPBM, the values in bold indicate the best performance.

Ground truth	Method	Family identification count over 100 trials (Averaged run time over 100 trials (mn))			
		F1	F2	F3	F4
M8	L-DCM	0/100 (0.2)	0/100 (0.7)	0/100 (0.8)	**100/100** (2.4)
	D-DCM	0/100 (0.2)	0/100 (0.7)	0/100 (0.8)	**100/100** (2.6)
	DCM	0/100 (0.1)	0/100 (0.4)	0/100 (0.5)	**100/100** (1.8)
$M16_s$	L-DCM	0/100 (0.3)	0/100 (0.8)	0/100 (0.8)	**100/100** (2.8)
	D-DCM	0/100 (0.2)	0/100 (0.7)	0/100 (0.7)	**100/100** (2.4)
	DCM	0/100 (0.1)	0/100 (0.5)	0/100 (0.5)	**100/100** (1.7)
$M16_{as}$	L-DCM	0/100 (0.1)	0/100 (0.7)	0/100 (0.8)	**100/100** (2.4)
	D-DCM	0/100 (0.1)	0/100 (0.6)	0/100 (0.6)	**100/100** (2.2)
	DCM	0/100 (0.1)	0/100 (0.4)	0/100 (0.5)	**100/100** (1.6)

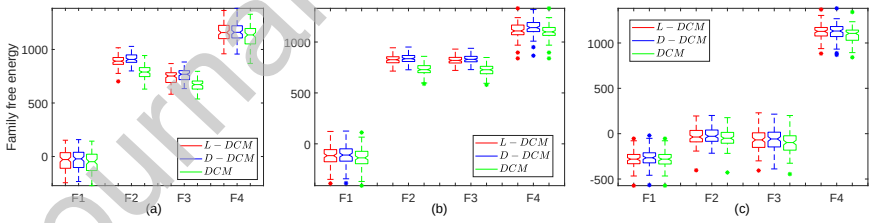


Figure 7. Boxplots of the family free energy ($FF_f, f = 1, 2, 3, 4$) values over 100 trials using L-DCM, D-DCM and DCM for the four families (bidirectional influence) in cPBM. **(a-c)**: the three ground truths (see Figure 6 **(a-i)**, **(b-i)**, **(c-i)**) corresponding to M8, $M16_s$ and $M16_{as}$, respectively.

us note that the boxplots of FF_2 and FF_3 are almost the same due to symmetric model structure (*i.e.* $\boldsymbol{\theta}_j = \boldsymbol{\theta}_k$ for $M16_s$) for all methods as shown in Fig. 7 (b).

4.2. Real iEEG signals

Our purpose is to detect brain regions responsible for the seizure onset and effective connectivity consists in inferring the directional flow between brain regions. In this study, iEEG signals were recorded in an epileptic patient in the Centre Hospitalier Universitaire in Rennes. Two particular channels were selected by the clinician: the Pp4 channel, whose activity was confirmed to be part of the seizure onset zone and the Dp1 channel, corresponding to a less epileptogenic region triggered in the seizure. It is worth mentioning that the bipolar signal derived from channel Pp4 (obtained as the difference between Pp4 and Pp5) was associated with the output $y_j(t)$ of population pop_j whereas the bipolar signal derived from channel Dp1 (obtained as the difference between Dp1 and Dp2) was assigned with the output $y_k(t)$ of population pop_k. Effective connectivity based on PBM, where the simulations were already investigated in [18], and cPBM between these two bipolar signals was evaluated for the first part of the ictal phase ($22 \sim 46$s) where channel Pp4 showed a fast onset activity (see Fig. 8 **(a)**). The segment for each scenario was partitioned using 2s length sliding windows without any overlap. Each windowed signal segment of 2s length was normalized and used to calculate the sample PSD, $\tilde{G}(\nu)$. A unidirectional connectivity from Pp4 to Dp1 in the ictal phase could be considered as ground truth according to the clinician. Therefore, the model structure M2 in PBM or family F2 (*i.e.* M2, M5, M6) in cPBM stands for the ground truth based on clinical expertise, as indicated in Fig. 8 **(b)**.

Regarding the model/family selection, under the cPBM model, L-DCM, D-DCM and DCM methods identify M5 as the optimal model structure underlying the observed iEEG signals as shown in **Support materials S3**. In terms of the maximized free energy and the family free energy depicted in Fig. 9 **(a-b)**, all methods point to a causal effect from pop_j to pop_k (M2 for Fig. 9 **(a)** and F2 for Fig. 9 **(b)**) whatever the NMM (*e.g.* PBM or cPBM).

Regarding the coefficients of determination between the sample and estimated PSDs and considering the estimated PSD of pop_j as given in Fig. 8 **(c-i) (d-i)**, under the cPBM framework, a higher estimation quality is obtained for the

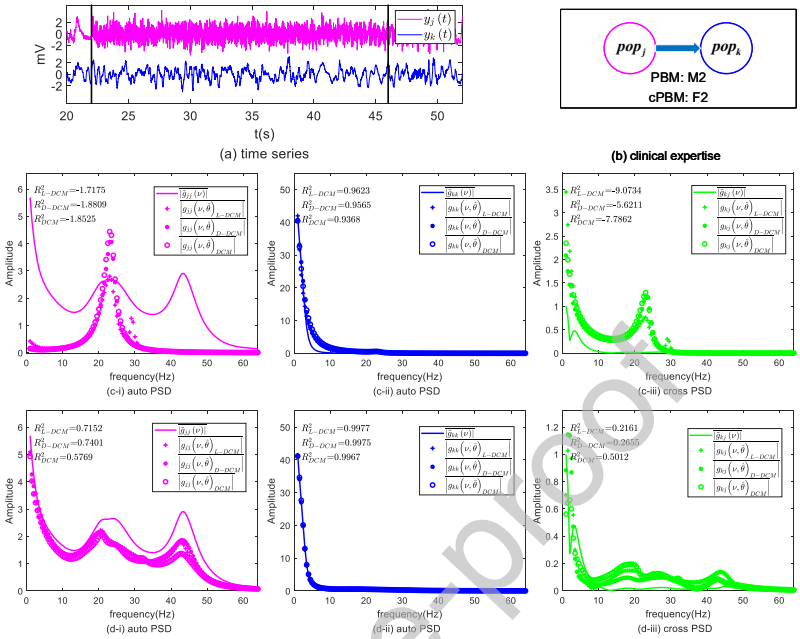


Figure 8. **(a)**: Two iEEG signals with a seizure onset up to 24s in channels $y_j(t)$ and $y_k(t)$ from 22s to 46s in the ictal phase. Each channel corresponds to a bipolar iEEG signal. **(b)**: Prior information provided by the clinical expertise. Results corresponding to PBM **(c-i, c-ii, c-iii)** and cPBM **(d-i, d-ii, d-iii)** in terms of the sample and estimated auto/cross PSD amplitudes together with their coefficients of determination R^2.

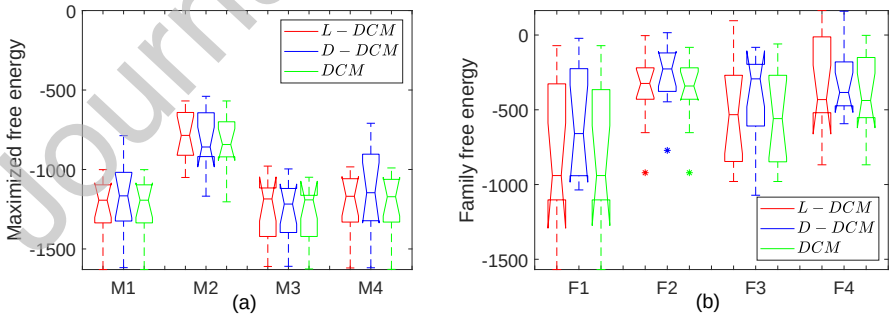


Figure 9. **(a)**: Boxplots of the maximized free energy for the four model structures in PBM. **(b)**: Boxplots of the family free energy for the four families in cPBM.

three algorithms (R^2_{D-DCM} =0.7401, R^2_{L-DCM} =0.7152 and R^2_{DCM} =0.5769) compared to the case of PBM framework (R^2_{D-DCM} = -1.8809, R^2_{L-DCM} =-1.7175 and R^2_{DCM} =-1.8525). This is due to the fact that, under PBM framework, no method is able to reproduce a PSD with two peaks (around 22 Hz and 44 Hz) corresponding to the sample PSD computed from the observed iEEG signals. On the contrary, the cPBM allows to recover two peaks in the PSD whatever the method, this PSD being close to the sample one. We can also observe that, in the case of cPBM, the D-DCM method provides the best performance in the population pop_j, as shown in Fig. 8 **(d-i)**. Regarding the estimated PSD of pop_k, we also note from Fig. 8 **(c-ii)(d-ii)** that all methods show good estimation accuracy under cPBM (R^2_{L-DCM} = 0.9977, R^2_{D-DCM} =0.9975 and R^2_{DCM} =0.9967), compared to PBM (R^2_{L-DCM} = 0.9623, R^2_{D-DCM} =0.9565 and R^2_{D-DCM} =0.9368).

Regarding the execution time shown in Table 7, all methods show, as expected, an execution time that is considerably smaller under the cPBM compared to the PBM. Around 65%, 61% and 41% of averaged execution time is saved for L-DCM, D-DCM and DCM, under the assumption of cPBM compared to the case of PBM. This is due to the fact that, contrary to PBM, the equilibrium point for $\Im(\boldsymbol{\theta})$ in Eq. (5) in cPBM to derive the estimated PSD during the optimization process does not need to be computed. This is since both the sigmoid function and the population inputs are centered in cPBM. Consequently, cPBM provides a faster model inversion than PBM even if the number of model parameters and model structures are higher in cPBM compared to PBM.

Table 7. Averaged execution time for all model structures over all windows (mn) for PBM and cPBM, NMM: neural mass model.

Method / NMM	L-DCM	D-DCM	DCM
PBM	10.49	12.25	5.73
cPBM	3.70	4.75	3.36

5. Discussion

5.1. Conclusions

We attempted to contribute to the analysis of brain effective connectivity in the context of drug-resistant epilepsy in the framework of neural mass models. This work focused on the model structures used to simulate intracranial epileptic activities similar to the ones observed at the seizure onset, and more particularly on the fast activity around 25Hz (sinusoidal dynamics) [12].

Regarding the simulated iEEG signals using cPBM, the D-DCM algorithm showed some superiority in terms of identification when an extrinsic coupling from the pyramidal cells to the slow inhibitory interneurons exists (case of unidirectional influence) between two populations. In the case of bidirectional coupling, the three techniques we presented, performed well. The D-DCM and L-DCM algorithms outperformed DCM in terms of maximizing the free energy in cPBM, with a slight superiority in favour of D-DCM over L-DCM.

Regarding the real epileptic signals, the analysis was conducted on a couple of iEEG channels. Results showed that the potential causal influence was properly identified using either PBM or cPBM. Compared to PBM, cPBM offered a lower computational complexity since computing the equilibrium point was not any more required. This leaded to a faster model inversion even if the number of model parameters and model structures was relatively high. Moreover, the cPBM provided a more expressive model of PSDs and a better estimation accuracy compared to PBM as already stated in our obtained results on real epileptic iEEG signals.

5.2. Limitations

Clearly, there are still some issues that we have not addressed in this paper. Firstly, the D-DCM approach can be improved by incorporating a local adjustment initialization strategy as the one in L-DCM. Secondly, our study only concerned the causal coupling between two neural populations and it would be interesting to extend the L-DCM and D-DCM methods to the

case of several coupled neural populations or larger network implementations since potentially more than two distant populations can be involved in the seizure initiation. Let us note that the scalability of the cPBM model for larger network implementations may be probably solved by the Baysian model reduction [40]. In addition to the direct coupling between several coupled neural populations, identifying indirect causal coupling could also improve effective connectivity inference. Finally, since epileptic seizures are patient-dependent, the performance of the L-DCM and D-DCM algorithms should be assessed by an extensive study using a larger dataset. A comparison to other methods, such as Granger Causality [41], transfer entropy [42] and DCM approaches could also be investigated. However, these different types of approaches do not address the same causality concept and are still disputed.

Declaration of Ethical Approval

All experimental procedures have been conducted at the CHU (Centre Hospitalier Universitaire), Rennes, France, following the ethical and regulatory standards. The patient has signed a consent and has been informed that his iEEG data would be used for clinical research and might serve for publication.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work has been supported by Natural Science Foundation of China under Grant (62001240, 31400842, 61876037), the National Key Research and Development Program of China (2017YFB1303200), the National Key R&D Program of China (2017YFC0107900), the Short-term Recruitment Program of Foreign Experts (WQ20163200398), the CRIBs, the Leading-edge Technology and Basic Research Program of Jiangsu Province (BK20192004) and the China

Scholarship Council. The authors thank Jean-Jacques Bellanger, Isabelle Merlet, and Fabrice Wendling for their fruitful discussions and Rennes CHU for providing the database.

Appendix A: confusion matrices

Bayesian model comparison at the group level [37] can be also performed using the confusion matrix where the evidence of the sixteen model structures that has been used to invert the different datasets generated using the six model structures (*e.g.* M2, M5, M6, M8, $M16_s$ and $M16_{as}$) is reported. To this end, three confusion matrices, one per method, were generated and illustrated in Fig. 10. Each entry of these matrices is the median of the maximized free energy computed over 100 trials. From this figure, we observed that the three methods can select the correct model structure except that the DCM method which selected M16 whereas the ground truth was M2.

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16
M2	-465	**712**^{1}	48	-637	-470	698^{2}	51	-615	-467	640^{4}	32	-780	-162	83	-616	651^{3}
M5	-416	464	-453	-504	**828**^{1}	815^{2}	783^{3}	733^{4}	125	99	128	117	498	-718	-718	572
M6	-313	797^{2}	84	-563	409	**904**^{1}	377	789^{3}	-336	698^{4}	117	637	-11	-654	63	646
M8	-28	886^{4}	747	-241	-33	881	670	**1146**^{1}	-33	859	745	951^{3}	147	171	467	1081^{2}
$M16_s$	-116	821	815	-37	58	825^{4}	775	1050^{2}	-114	781	818	1047^{3}	118	497	505	**1103**^{1}
$M16_{as}$	-281	-171	-199	-20	-39	-105	-5	-126	-73	846^{2}	-80	-228	-372	55^{3}	15^{4}	**1128**^{1}

(a)

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16
M2	-458	**736**^{1}	112	-614	-463	706^{2}	120	-580	-460	619^{3}	111	-555	135	140	-226	615^{4}
M5	-408	455	-408	-304	**852**^{1}	845^{2}	781^{3}	687^{4}	172	163	173	161	-382	-676	-676	498
M6	-306	764^{2}	178	-380	483	**927**^{1}	442	610	-308	731^{3}	163	661^{4}	241	-602	261	625
M8	-22	900	768	-208	-23	907^{4}	752	**1149**^{1}	-28	903	763	1014^{3}	501	-13	545	1131^{2}
$M16_s$	-110	826	823	34	30	836^{4}	794	1041^{2}	-107	795	830	1033^{3}	483	567	563	**1142**^{1}
$M16_{as}$	-266	-139	-179	193^{3}	-46	-92	5	-80	-63	538^{2}	-66	-74	-257	128^{4}	88	**1132**^{1}

(b)

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16
M2	-505	371^{2}	2	-944	-591	365^{3}	-56	-817	-590	358^{4}	7	-842	-259	-479	-629	**436**^{1}
M5	-419	340	-480	-543	**823**^{1}	796^{2}	759^{3}	732^{4}	125	99	128	116	-429	-718	-718	405
M6	-316	751^{2}	70	-570	409	**896**^{1}	377	535	-386	661^{3}	-184	592	-22	-654	45	628^{4}
M8	-48	783	672	-360	-101	787^{4}	658	**1106**^{1}	-116	758	658	934^{3}	144	170	288	1080^{2}
$M16_s$	-138	728	720	-37	-10	726^{4}	713	996^{3}	-196	715	724	1003^{2}	117	230	224	**1097**^{1}
$M16_{as}$	-281	-179	-202	-34	-74	-126	-159	-302	-106	584^{2}	-103	-228	-377	55^{3}	13^{4}	**1106**^{1}

(c)

Figure 10. The confusion matrices for **(a)** L-DCM **(b)** D-DCM and **(c)** DCM, respectively. Each value in these matrices corresponds to the median of the maximized free energy over 100 trials. Entries in green and bold refer to the identified model structure for methods L-DCM, D-DCM and DCM, respectively. The superscript number indicates the four most plausible model structures for each method and each ground truth as given in Tables 3 and 5.

Appendix B: thresholding for removing spurious connections

To deal with potential false detections induced by the small non-null value(s) of the coupling parameter(s), a thresholding based post-processing step is adopted. This means that, if one/some of the estimated parameter(s) $exp(\hat{\theta}_{K_{kj}^{pp}})$, $exp(\hat{\theta}_{K_{kj}^{pp}})$, $exp(\hat{\theta}_{K_{kj}^{pp}})$ and $exp(\hat{\theta}_{K_{kj}^{pp}})$ is/are less than a pre-defined threshold parameter ϕ, it/they is/are set to 0. This can potentially lead to define another model structure and consequently modify its identification rate as shown in Tables 8 and 9, where several threshold values (e.g. $\phi \in \{0, 0.05, 0.1, 0.15\}$) are considered. From the above two tables, we can observe that: (1) in the case of unidirectional flow, the model identification rate of the correct model structure is increased when the value of the threshold ϕ is increased as well regardless the considered method. It results in a decreased false detection rate. Regarding the case of bidirectional flow, no significant improvement in the model identification rate can be observed

Table 8. The identification rates of the identified model structure (according to Fig. 10) using simulated iEEG signals under six ground truths over 100 trials. The values in bold indicate the best performance.

Ground truth	Method	Mm	Model identification count over 100 trials			
			$\phi = 0$	$\phi = 0.05$	$\phi = 0.1$	$\phi = 0.15$
M2	L-DCM	M2	38/100	51/100	62/100	75/100
	D-DCM	M2	**48/100**	**60/100**	**67/100**	**81/100**
	DCM	M16	59/100	59/100	58/100	44/100
M5	L-DCM	M5	57/100	66/100	79/100	91/100
	D-DCM	M5	71/100	73/100	**83/100**	91/100
	DCM	M5	**76/100**	**76/100**	81/100	**92/100**
M6	L-DCM	M6	87/100	91/100	95/100	99/100
	D-DCM	M6	**99/100**	**99/100**	**100/100**	**100/100**
	DCM	M6	97/100	97/100	98/100	**100/100**
M8	L-DCM	M8	**85/100**	**85/100**	**85/100**	**86/100**
	D-DCM	M8	80/100	80/100	81/100	81/100
	DCM	M8	79/100	79/100	79/100	82/100
M16$_s$	L-DCM	M16	80/100	80/100	80/100	80/100
	D-DCM	M16	**96/100**	**96/100**	**96/100**	**96/100**
	DCM	M16	94/100	94/100	94/100	94/100
M16$_{as}$	L-DCM	M16	**100/100**	**100/100**	**100/100**	**100/100**
	D-DCM	M16	**100/100**	**100/100**	**100/100**	**100/100**
	DCM	M16	**100/100**	**100/100**	**100/100**	**100/100**

Table 9. The identification rates of the second plausible model structure (according to Fig. 10) using simulated iEEG signals under six ground truths over 100 trials. The values in bold indicate the best performance.

Ground truth	Method	Mm	Model identification count over 100 trials			
			$\phi = 0$	$\phi = 0.05$	$\phi = 0.1$	$\phi = 0.15$
M2	L-DCM	M6	**35/100**	33/100	24/100	11/100
	D-DCM	M6	**35/100**	**30/100**	**23/100**	**10/100**
	DCM	M2	4/100	17/100	31/100	37/100
M5	L-DCM	M6	14/100	11/100	**6/100**	5/100
	D-DCM	M6	9/100	8/100	7/100	**4/100**
	DCM	M6	**6/100**	**6/100**	**6/100**	5/100
M6	L-DCM	M2	**0/100**	**0/100**	**0/100**	**0/100**
	D-DCM	M2	**0/100**	**0/100**	**0/100**	**0/100**
	DCM	M2	**0/100**	**0/100**	**0/100**	**0/100**
M8	L-DCM	M16	**14/100**	**14/100**	**14/100**	**13/100**
	D-DCM	M16	19/100	19/100	18/100	17/100
	DCM	M16	21/100	21/100	21/100	19/100
M16$_s$	L-DCM	M8	11/100	11/100	11/100	11/100
	D-DCM	M8	2/100	2/100	2/100	2/100
	DCM	M12	**1/100**	**1/100**	**1/100**	**1/100**
M16$_{as}$	L-DCM	M10	**0/100**	**0/100**	**0/100**	**0/100**
	D-DCM	M10	**0/100**	**0/100**	**0/100**	**0/100**
	DCM	M10	**0/100**	**0/100**	**0/100**	**0/100**

whatever the threshold value. (2) The D-DCM algorithm shows a higher model identification rate compared to the DCM and the L-DCM ones in the cases where M2, M6 and M16s are considered as ground truths and also whatever the threshold value. Besides, regarding the case where M8 is the ground truth, the superiority of L-DCM over DCM and D-DCM are also to be noticed whatever the value of ϕ.

References

[1] I. E. Scheffer, S. Berkovic, G. Capovilla, M. B. Connolly, J. French, L. Guilhoto, E. Hirsch, S. Jain, G. W. Mathern, S. L. Moshé, et al., Ilae classification of the epilepsies: position paper of the ilae commission for classification and terminology, Epilepsia 58 (4) (2017) 512–521.

[2] M. J. Brodie, P. Kwan, Staged approach to epilepsy management, Neurology 58 (8 suppl 5) (2002) S2–S8.

[3] T. Wu, D. Chen, Q. Chen, R. Zhang, W. Zhang, Y. Li, L. Zhang, H. Liu, S. Wan, T. Jiang, et al., Automatic lateralization of temporal lobe epilepsy based on MEG network features using support vector machines, Complexity 4325096 (2018).

[4] P. van Mierlo, M. Papadopoulou, E. Carrette, P. Boon, S. Vandenberghe, K. Vonck, D. Marinazzo, Functional brain connectivity from EEG in epilepsy: Seizure prediction and epileptogenic focus localization, Progress in Neurobiology 121 (2014) 19–35.

[5] A. Zigouris, G. A. Alexiou, M. S. Themistocleous, S. Voulgaris, Aed strategy after refractory epilepsy surgery., Current pharmaceutical design 23 (42) (2017) 6505–6507.

[6] J. Engel Jr, Outcome with respect to epileptic seizures., Surgical Treatment of the Epilepsies (1993) 609–621.

[7] D. R. Nair, A. Mohamed, R. Burgess, H. Lders, A critical review of the different conceptual hypotheses framing human focal epilepsy, Epileptic Disorders 6 (2) (2004) 77–83.

[8] J. S. Duncan, G. P. Winston, M. J. Koepp, S. Ourselin, Brain imaging in the assessment for epilepsy surgery, The Lancet Neurology 15 (4) (2016) 420–433.

[9] M. Rubinov, O. Sporns, Complex network measures of brain connectivity: uses and interpretations, Neuroimage 52 (3) (2010) 1059–1069.

[10] W. D. Penny, K. E. Stephan, A. Mechelli, K. J. Friston, Modelling functional integration: a comparison of structural equation and dynamic causal models, Neuroimage 23 (2004) S264–S274.

[11] K. J. Friston, L. Harrison, W. Penny, Dynamic causal modelling, Neuroimage 19 (4) (2003) 1273–1302.

[12] F. Wendling, P. Chauvel, A. Biraben, F. Bartolomei, From intracerebral EEG signals to brain connectivity: identification of epileptogenic networks in partial epilepsy, Frontiers in Systems Neuroscience 4 (154) (2010).

[13] R. J. Moran, K. E. Stephan, T. Seidenbecher, H.-C. Pape, R. J. Dolan, K. J. Friston, Dynamic causal models of steady-state responses, Neuroimage 44 (3) (2009) 796–811.

[14] K. J. Friston, A. Bastos, V. Litvak, K. E. Stephan, P. Fries, R. J. Moran, DCM for complex-valued data: cross-spectra, coherence and phase-delays, Neuroimage 59 (1) (2012) 439–455.

[15] G. K. Cooray, B. Sengupta, P. Douglas, M. Englund, R. Wickstrom, K. Friston, Characterising seizures in anti-NMDA-receptor encephalitis with dynamic causal modelling, Neuroimage 118 (2015) 508–519.

[16] M. Papadopoulou, G. Cooray, R. Rosch, R. Moran, D. Marinazzo, K. Friston, Dynamic causal modelling of seizure activity in a rat model, Neuroimage 146 (2017) 518–532.

[17] W. Xiang, C. Yang, J.-J. Bellanger, H. Shu, R. Le Bouquin Jeannès, Inferring effective connectivity in epilepsy using dynamic causal modeling, IRBM 36 (6) (2015) 335–344.

[18] W. Xiang, A. Karfoul, H. Shu, R. Le Bouquin Jeannès, A local adjustment strategy for the initialization of dynamic causal modelling to infer effective

connectivity in brain epileptic structures, Computers in Biology and Medicine 84 (2017) 30–44.

[19] A. M. Bastos, W. M. Usrey, R. A. Adams, G. R. Mangun, P. Fries, K. J. Friston, Canonical microcircuits for predictive coding, Neuron 76 (4) (2012) 695–711.

[20] H. Wei, A. Jafarian, P. Zeidman, V. Litvak, A. Razi, D. Hu, K. J. Friston, Bayesian fusion and multimodal dcm for eeg and fmri, NeuroImage 211 (2020) 116595.

[21] F. Wendling, F. Bartolomei, J. Bellanger, P. Chauvel, Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition, European Journal of Neuroscience 15 (9) (2002) 1499–1508.

[22] F. Wendling, A. Hernandez, J.-J. Bellanger, P. Chauvel, F. Bartolomei, Interictal to ictal transition in human temporal lobe epilepsy: insights from a computational model of intracerebral EEG, Journal of Clinical Neurophysiology 22 (5) (2005).

[23] F. Shayegh, S. Sadri, R. Amirfattahi, K. Ansari-Asl, A model-based method for computation of correlation dimension, lyapunov exponents and synchronization from depth-eeg signals, Computer methods and programs in biomedicine 113 (1) (2014) 323–337.

[24] M. Papadopoulou, M. Leite, P. van Mierlo, K. Vonck, L. Lemieux, K. Friston, D. Marinazzo, Tracking slow modulations in synaptic gain using dynamic causal modelling: validation in epilepsy, Neuroimage 107 (2015) 117–126.

[25] G. K. Cooray, B. Sengupta, P. K. Douglas, K. Friston, Dynamic causal modelling of electrographic seizure activity using Bayesian belief updating, Neuroimage 125 (2016) 1142–1154.

[26] K. J. Friston, V. Litvak, A. Oswal, A. Razi, K. E. Stephan, B. C. van Wijk, G. Ziegler, P. Zeidman, Bayesian model reduction and empirical bayes for group DCM studies, Neuroimage 128 (2016) 413–431.

[27] B. H. Jansen, V. G. Rit, Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns, Biological Cybernetics 73 (4) (1995) 357–366.

[28] O. Grinenko, J. Li, J. C. Mosher, I. Z. Wang, J. C. Bulacio, J. Gonzalez-Martinez, D. Nair, I. Najm, R. M. Leahy, P. Chauvel, A fingerprint of the epileptogenic zone in human epilepsies, Brain 141 (1) (2017) 117–131.

[29] M. Zavaglia, F. Cona, M. Ursino, A neural mass model to simulate different rhythms in a cortical region, Computational Intelligence and Neuroscience 2010 (2010).

[30] M. Ursino, F. Cona, M. Zavaglia, The generation of rhythms within a cortical region: analysis of a neural mass model, Neuroimage 52 (3) (2010) 1080–1094.

[31] F. Cona, M. Zavaglia, M. Massimini, M. Rosanova, M. Ursino, A neural mass model of interconnected regions simulates rhythm propagation observed via TMS-EEG, Neuroimage 57 (3) (2011) 1045–1058.

[32] R. J. Moran, S. J. Kiebel, K. Stephan, R. Reilly, J. Daunizeau, K. J. Friston, A neural mass model of spectral responses in electrophysiology, Neuroimage 37 (3) (2007) 706–720.

[33] O. David, S. J. Kiebel, L. M. Harrison, J. Mattout, J. M. Kilner, K. J. Friston, Dynamic causal modeling of evoked responses in EEG and MEG, Neuroimage 30 (4) (2006) 1255–1272.

[34] K. Friston, J. Mattout, N. Trujillo-Barreto, J. Ashburner, W. Penny, Variational free energy and the Laplace approximation, Neuroimage 34 (1) (2007) 220–234.

[35] J. Daunizeau, K. J. Friston, S. J. Kiebel, Variational Bayesian identification and prediction of stochastic nonlinear dynamic causal models, Physica D: Nonlinear Phenomena 238 (21) (2009) 2089–2118.

[36] K. E. Stephan, W. D. Penny, J. Daunizeau, R. J. Moran, K. J. Friston, Bayesian model selection for group studies, Neuroimage 46 (4) (2009) 1004–1017.

[37] W. D. Penny, K. E. Stephan, J. Daunizeau, M. J. Rosa, K. J. Friston, T. M. Schofield, A. P. Leff, Comparing families of dynamic causal models, PLoS computational biology 6 (3) (2010) e1000709.

[38] N. Ueda, R. Nakano, Deterministic annealing EM algorithm, Neural Networks 11 (2) (1998) 271–282.

[39] K. Katahira, K. Watanabe, M. Okada, Deterministic annealing variant of variational Bayes method, Journal of Physics: Conference Series 95 (1) (2008).

[40] K. Friston, T. Parr, P. Zeidman, Bayesian model reduction, arXiv preprint arXiv:1805.07092 (2018).

[41] A. K. Seth, A MATLAB toolbox for Granger causal connectivity analysis, Journal of neuroscience methods 186 (2) (2010) 262–273.

[42] C. Yang, R. Le Bouquin Jeannès, J.-J. Bellanger, H. Shu, A new strategy for model order identification and its application to transfer entropy for EEG signals analysis, IEEE Transactions on Biomedical Engineering 60 (5) (2012) 1318–1327.

Support materials

S1: Bayesian model comparison for unidirectional influence

These results for Table 3 can be interpreted in terms of the median (less prone to outliers than the mean) of the maximized free energy depicted in Figs. 11, 12 and 13. For instance, regarding the L-DCM and D-DCM algorithms, a relatively higher false model identification rate occurred when the ground truth is M2, as expected. This is mainly due to the fact that the difference between the median of the maximized free energy of M6, M10 or M16 models and the one of the M2 one is significantly small (see Fig. 11). Now, the DCM shows higher false model identification rate compared to the L-DCM and D-DCM methods. This can be interpreted by the highest median of the maximized free energy obtained for the M16 model. Regarding the correct model M5, as the medians of the maximized free energy of models M5, M6, M7 and M8 are to some extent comparable (see Fig. 12), with slight superiority for M5, some false model identification can occur for the three algorithms. As far as the ground truth M6 is concerned, Table 3 clearly shows very high performance in terms of detection rate, especially for D-DCM (99%) and DCM (97%). These results can also be interpreted in terms of the median of the maximized free energy depicted in Fig. 13. According to this figure, the median of the maximized free energy obtained for identifying the correct model, M6, for D-DCM and DCM algorithms, is significantly higher than the ones related to the other model structures. Consequently, a very low false detection rate for these algorithms is expected.

As relatively high false detection rate is to some extent problematic, investigating the model evidence, $p(\tilde{\boldsymbol{g}}|\mathrm{M}m)$, would help to complete the whole scene regarding the ability of each algorithm in identifying the model structure. This model evidence is nothing else than the model posterior probability under flat prior and Mm [40]. This quantity is computed via the softmax function of the median of the maximized free energy computed over the 100 conducted trials, that means: $p(\tilde{\boldsymbol{g}}|\mathrm{M}m) = \frac{\exp(\tilde{F}_m)}{\sum_{i=1}^{16}\exp(\tilde{F}_i)}$ where $\tilde{F}_i$ denotes the median of the maximized free energy associated to the model structure Mi. It is clear

from Figs. 11, 12 and 13 that the model evidence measure allows for a global quantification of the ability of the different algorithms in identifying the correct model structure except in ground truth M2 where M16 is identified using DCM.

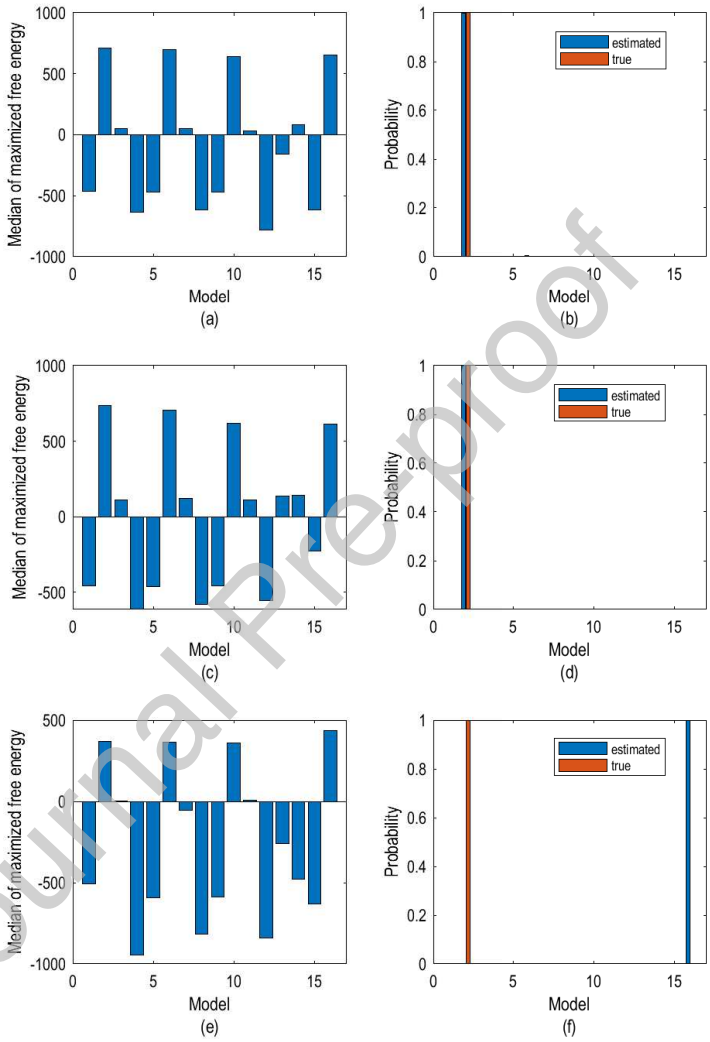


Figure 11. Bayesian model comparison in terms of the median of maximized free energy over 100 trials and the model posterior probability using **(a-b)** L-DCM, **(c-d)** D-DCM and **(e-f)** DCM for the ground truth M2 in the context of unidirectional influence.

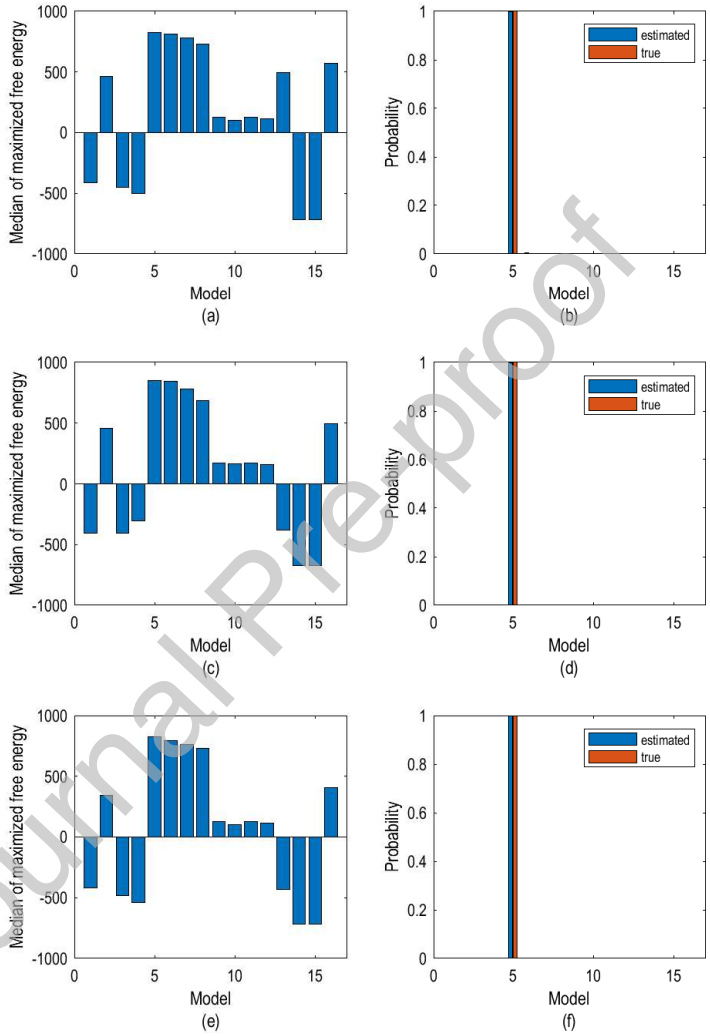


Figure 12. Bayesian model comparison in terms of the median of maximized free energy over 100 trials and the model posterior probability using **(a-b)** L-DCM, **(c-d)** D-DCM and **(e-f)** DCM for the ground truth M5 in the context of unidirectional influence.

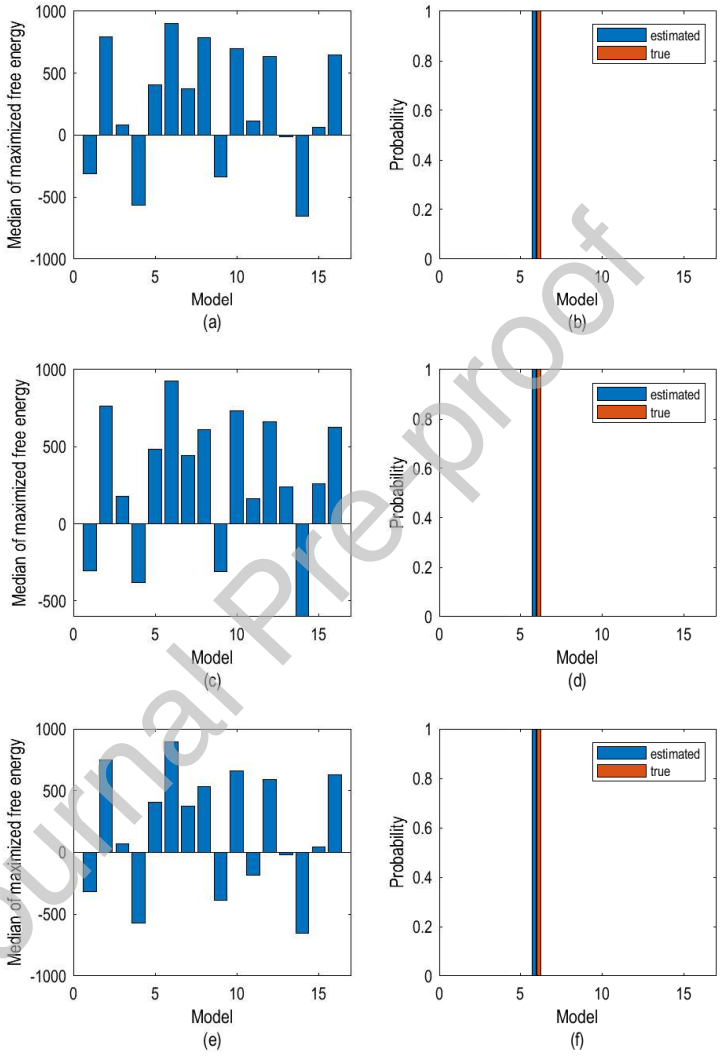


Figure 13. Bayesian model comparison in terms of the median of maximized free energy over 100 trials and the model posterior probability using **(a-b)** L-DCM, **(c-d)** D-DCM and **(e-f)** DCM for the ground truth M6 in the context of unidirectional influence.

S2: Bayesian model comparison for bidirectional influence

The results for Table 5 can be furthermore interpretable in Figs. 14, 15 and 16, which show the results of model selection based on the median of the maximized free energy over 100 Monte Carlo trials corresponding to ground truths M8, $M16_s$ and $M16_{as}$ respectively. Clearly, the correct model structure can be identified with almost 100% model posterior probability by all techniques as shown in these figures. As shown in Fig. 14, on the one hand, the gap on the median of the maximized free energy between the optimal model structure M8 and the model structure M16 (the model structure giving the second more important free energy among the 16 model structures) is small (< 30 for D-DCM and DCM) leading to the false detection of model structure (M16). On the other hand, this gap is slightly higher for L-DCM than for D-DCM and DCM, leading to a reduced false detection rate of M16 for L-DCM (14% *vs* 19% and 21% respectively), as shown in Table 5. A quite comparable behaviour of the three algorithms can be observed in the ground truth $M16_s$ (see Fig. 15), such that the gap on the median of the maximized free energy between the optimal model structure M16 and the model structure M8 (resp. M12) (the two model structures giving the second and third more important free energies among the 16 model structures) is small, this gap between M16 and M8 (resp. M12) is slightly higher for D-DCM than for DCM and L-DCM, leading to a reduced false detection rate of M8 (resp. M12) for D-DCM, DCM and L-DCM *i.e.* 2% *vs* 5% and 11% respectively (resp. 2% *vs* 1% and 9% respectively) as displayed in Table 5. Furthermore, as depicted in Fig. 16 where model structure $M16_{as}$ stands for the ground truth, the gap on the median of the maximized free energy between the optimal model structure M16 and the model structure M10 (the model structure giving the second more important free energy among the 16 model structures in this situation) is significant (higher than 270) whatever the method. The optimal model structure (M16) has a greater model posterior probability than any other model structure, so that 100% model identification rate is obtained by all techniques in this case (see Table 5).

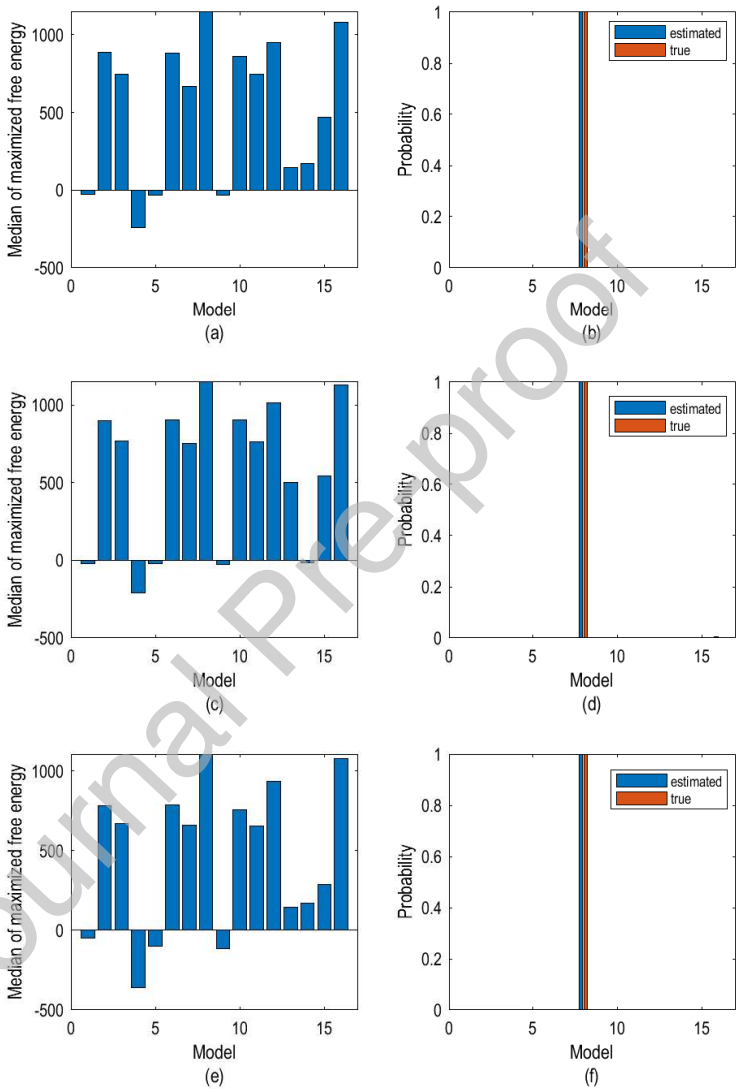


Figure 14. Bayesian model comparison in terms of the median of maximized free energy over 100 trials and the model posterior probability using **(a-b)** L-DCM, **(c-d)** D-DCM and **(e-f)** DCM for the ground truth M8 in the context of bidirectional influence.

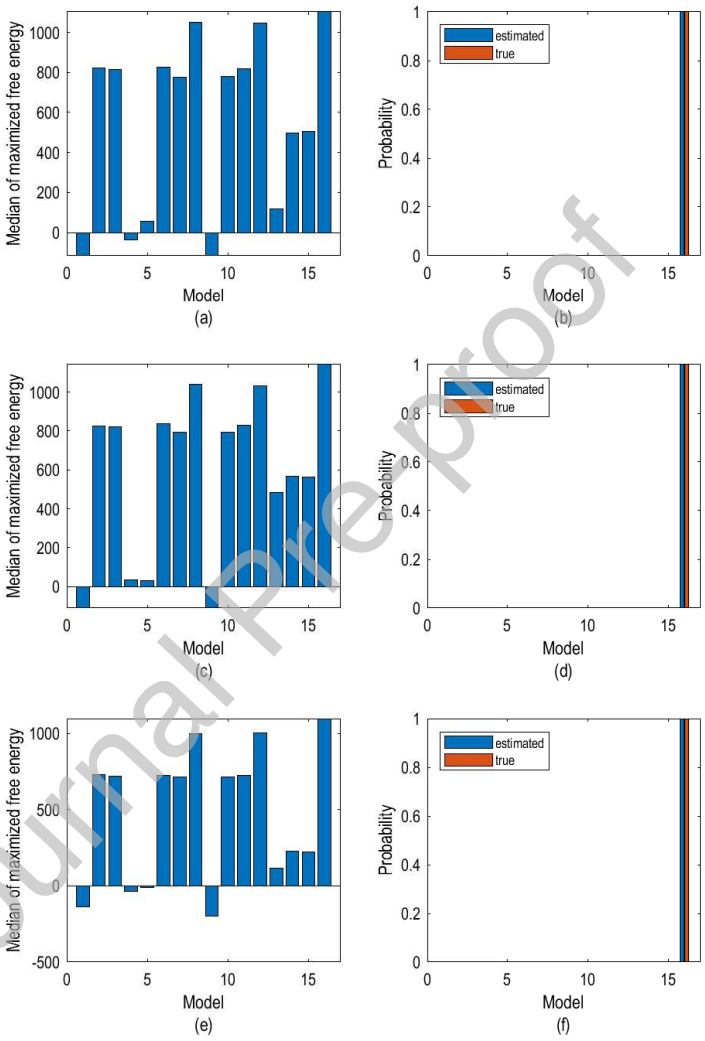


Figure 15. Bayesian model comparison in terms of the median of maximized free energy over 100 trials and the model posterior probability using **(a-b)** L-DCM, **(c-d)** D-DCM and **(e-f)** DCM for the ground truth $M16_s$ in the context of bidirectional influence.

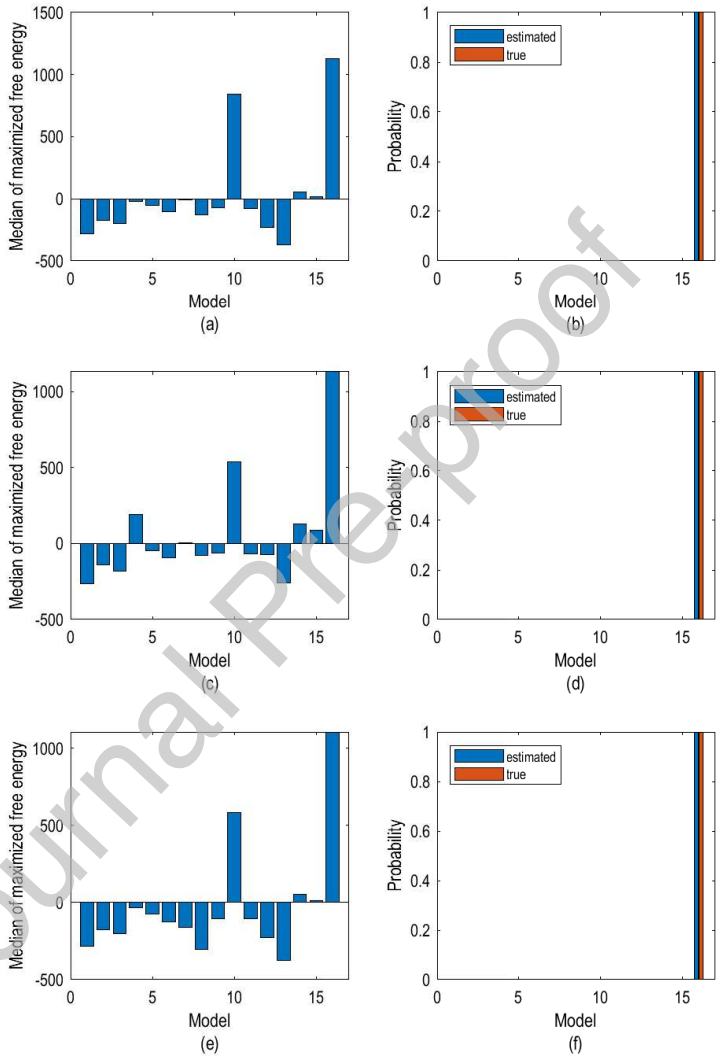


Figure 16. Bayesian model comparison in terms of the median of maximized free energy over 100 trials and the model posterior probability using **(a-b)** L-DCM, **(c-d)** D-DCM and **(e-f)** DCM for the ground truth M16$_s$ is the ground truth) in the context of bidirectional influence.

S3: Bayesian model comparison for real signals

Model selection in terms of the median of the maximized free energy over 12 trials in ictal phase is given in Fig 17 for cPBM. All methods provide almost 100% model evidence to identify M5 as depicted in Fig. 17 based on the median of maximized free energy.

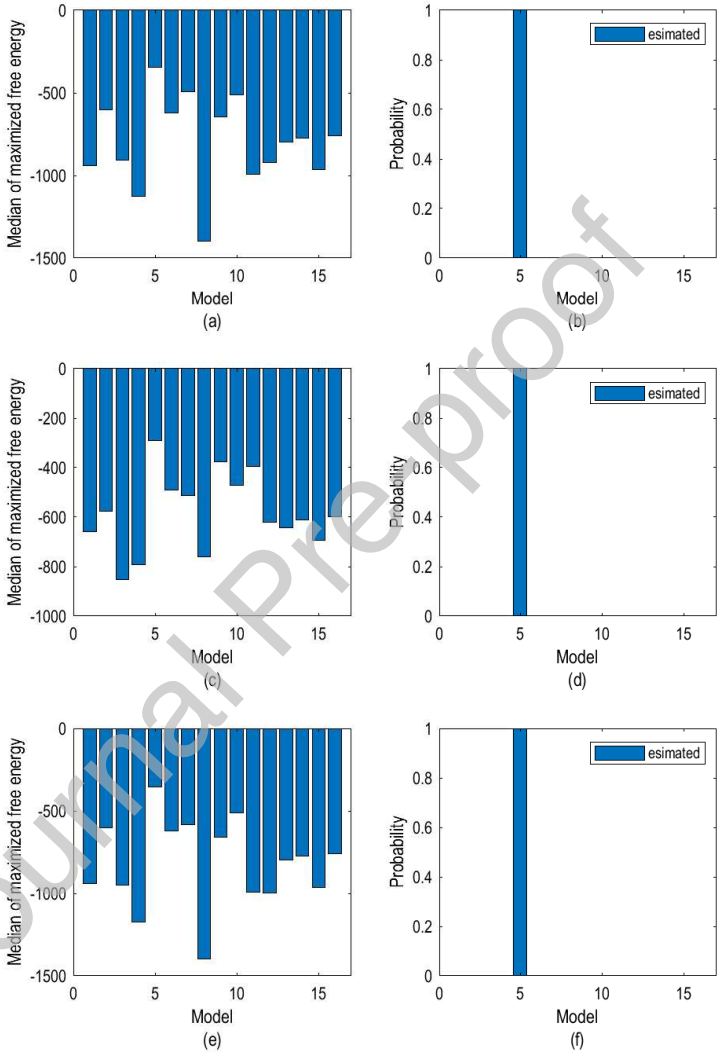


Figure 17. Bayesian model comparison in terms of the median of maximized free energy over 12 trials and the model posterior probability using **(a-b)** L-DCM, **(c-d)** D-DCM and **(e-f)** DCM for the Ictal phase in the context of real signals.