UNIVERSITÀ
DI PAVIA

FACOLTA' DI INGEGNERIA

DIPARTIMENTO DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

CORSO DI LAUREA MAGISTRALE IN BIOINGEGNERIA

TESI DI LAUREA

MODEL SELECTION STRATEGY IN DYNAMIC CAUSAL MODELING FOR FUNCTIONAL MAGNETIC RESONANCE IMAGING: APPLICATION TO AN ACTION OBSERVATION TASK

STRATEGIA DI SELEZIONE DEL MODELLO NEL DYNAMIC CAUSAL MODELING PER LA RISONANZA MAGNETICA FUNZIONALE: APPLICAZIONE A UN TASK DI OSSERVAZIONE DELL'AZIONE.

Candidato:

Udo-Chukwu Ifunanya Chukwu Oruche

Relatore:

Prof.ssa Claudia A. M. Gandini Wheeler-Kingshott

Correlatori:

Dott.ssa Roberta Maria Lorenzi

Dott.ssa Gökçe Korkmaz

A.A. 2024/2025

INDEX

1. PRINCIPLE OF MAGNETIC RESONANCE IMAGING 11
   1.1 PHYSICAL SIGNAL TO IMAGE 11
   1.2 ACQUISITION 16
   1.3 IMAGE RECONSTRUCTION 20
   1.4 TISSUE CONTRAST 23
2. FUNCTIONAL MAGNETIC RESONANCE IMAGING 26
   2.1 ACQUISITION OF FUNCTIONAL MAGNETIC RESONANCE IMAGING 28
   2.2 PRE-PROCESSING of fMRI DATA 29
   2.3 GENERAL LINEAR MODEL 30
3. BRAIN CONNECTIVITY 34
   3.1 STRUCTURAL CONNECTIVITY 34
   3.2 FUNCTIONAL CONNECTIVITY 34
   3.3 EFFECTIVE CONNECTIVITY 35
4. DYNAMIC CAUSAL MODELLING 36
   4.1 NEURONAL MODEL 37
   4.2 HEMODYNAMIC MODEL 38
   4.3 BAYESIAN INFERENCE 40
   4.4 BAYESIAN MODEL SELECTION 44
   4.5 BAYESIAN MODEL COMPARISON 46
   4.6 PARAMETRIC EMPIRICAL BAYES 47
   4.7 AIM OF STUDY 48
5. MATERIALS AND METHODS 50
   5.1 PARTICIPANTS 50
      5.1.1 MRI ACQUISITION PROTOCOL 50
      5.1.2 EXPERIMENTAL PARADIGM 51
   5.2 PREPROCESSING ANALYSIS 53
   5.3 VOI EXTRACTION FOR DCM ANALYSIS 55
   5.4 EFFECTIVE CONNECTIVITY ESTIMATION USING DCM 59
   5.5 GROUP-LEVEL PEB ANALYSIS 64
6. RESULTS 65
   6.1 MODEL SELECTION 65
   6.2 GROUP-LEVEL CONNECTIVITY RESULTS 68

7. DISCUSSION 73
8. CONCLUSIONS 75
BIBLIOGRAPHY 76

LIST OF FIGURES

Figure 1.1: Behaviour of nuclei as magnetic dipoles. In (a), a rotating charged particle is shown, where rotation generates a magnetic moment. In (b), a bar magnet analogy is shown, with north and south poles, where the magnetic field lines go from north to south (Brown et al., 2014). 12

Figure 1.2: [Precession of the magnetic moment around B0. In this figure, the vertical arrow represents B0, while the tilted arrow represents the magnetic moment. The circular trajectory represents the precessional motion. (Brown et al.,2014). 13

Figure 1.3: Evolution of the longitudinal magnetization (Mz(t)) over time. The curve shows how Mz approaches its equilibrium value, reflecting the energy exchange with the surrounding environment as spins return to a lower energy state (Coriasco et al., 2014) 14

Figure 1.4: Evolution of the Mxy(t) over time. The curve shows an oscillating signal whose amplitude decreases over time, producing the FID. This decay is due to the loss of phase coherence between spins (Coriasco et al., 2014) 15

Figure 1.5: RF pulses and signal evolution over time. In the top row, the 90° and 180° RF pulses are shown, while the bottom row represents the signal evolution over time. After the 90° pulse, spins are in the transverse plane; therefore, the signal starts oscillating (FID). However, it decays very rapidly due to both spin-spin interactions and magnetic field inhomogeneities. As a result, the observed decay corresponds to T2*. The 180* pulse is applied after time t1. It flips the spins and reverses the dephasing caused by magnetic field inhomogeneities (Coriasco et al 2014), 16

Figure 1.6: SE sequence. The temporal diagram shows a 90° RF pulse. Subsequently, dephasing occurs. Then, a 180° pulse is applied, after which rephasing occurs. Finally, an echo is formed and the signal is acquired. (Brown et al., 2014). 17

Figure 1.7: GRE sequence. The temporal diagram shows that an RF pulse is applied, after which dephasing occurs. Subsequently, gradient reversal is applied, and then the echo is generated, and the signal is acquired (Brown et al., 2014). 18

Figure 1.8: EPI sequence. The temporal diagram shows a single RF excitation. Subsequently, rapid gradient switching is applied. Then, multiple echoes are acquired, resulting in a fast acquisition. (Brown et al., 2014). 19

Figure 1.9: Example image illustrating the effect of TR and TE on MRI image contrast. (Coriasco et al., 2014). 24

Figure 2.1: BOLD SIGNAL AND CBF. The top graph represents the change in blood flow over time, while the bottom graph shows the corresponding BOLD signal. 27

Figure 2.2: HRF. Immediately after the stimulus, a slight decrease in signal is observed (initial dip). After a few seconds, the signal increases and reaches a peak. Following the peak, the signal decreases below baseline (post-stimulus undershoot). This reflects the vascular response to neural activity (Buxton et al., 2004). 28

Figure 5.1: Squeeze-ball device and visual feedback system. The squeeze-ball is MRI-compatible, ensuring safe use within the scanner environment. The visual display

shows both the target force level and the actual force produced, represented by a bar. The colour green indicates correct performance, whereas red indicates a mismatch between target and applied force (Alahmadi et al., 2016). 52

Figure 5.2; Experimental paradigm. Panel A corresponds to the AE condition, where the subject performs the squeeze-ball task with visual feedback. Panel B corresponds to the AO condition, in which the subject observes the action without performing any movement. Panel C corresponds to the AOBAR condition, where the subject observes the task with an additional visual cue. (Casiraghi et al., 2019) 53

Figure 5.3: Example of VOI extraction in AO. This figure shows the time series extracted for a representative subject across all regions in the network. Each subplot corresponds to a VOI. The signals represent the temporal evolution of the BOLD signal and show oscillations over time, reflecting indirect neural activity. Each region exhibits its own dynamics, resulting in differences across regions. These differences form the basis for connectivity analysis. The extracted time series are used as inputs for DCM to estimate interactions between regions. 58

Figure 5.4: Example of VOI extraction in AO-BAR. This figure shows the time series extracted for a representative subject across all regions in the network. Each subplot corresponds to a VOI. The signals represent the temporal evolution of the BOLD signal and show oscillations over time, reflecting indirect neural activity. Each region exhibits its own dynamics, resulting in differences across regions. These differences form the basis for connectivity analysis. The extracted time series are used as inputs for DCM to estimate interactions between regions. 59

Figure 5.5: Schematic illustration of the full effective connectivity model of cortico–cerebellar loops. The model represents the full architecture obtained by connecting all regions. 61

Figure 5.6: Schematic illustration of the reduced effective connectivity models of cortico–cerebellar loops. Each model corresponds to a specific hypothesis. Models S1.1–S1.5 represent reduced architectures obtained by removing specific connections. The models differ in the presence or absence of connections between visual, motor, and cerebellar regions. Model 4 corresponds to S1.4. (Lorenzi et al., 2025) 63

Figure 6.1: AO. The bars represent the exceedance probability of the reduced models 66

Figure 6.2: AO-BAR. The bars represent the exceedance probability of the reduced models 67

Figure 6.3: Heatmap of supported connections for AO (Pp > 0.95). Each value equal to 1 indicates a strongly supported connection, while 0 indicates no strong support 68

Figure 6.4: Heatmap of supported connections for AO-BAR (Pp > 0.95). No connections exceed the threshold; therefore, no strong group-level connections are identified 69

Figure 6.5: Network representation of supported effective connectivity (Pp > 0.95) for the AO condition. Arrows represent the direction of connectivity. The external input is applied to V1. 70

Figure 6.6: Network representation of effective connectivity for the AO-BAR condition. No connections exceed the threshold (Pp > 0.95). The external input is applied to V1. 72

ABSTRACT (ITALIAN)

Le tecniche di Risonanza Magnetica (MRI) consentono di studiare il cervello umano da diverse prospettive, che includono l'anatomia (ad esempio la caratterizzazione morfologica delle diverse regioni cerebrali), la microstruttura (ad esempio le connessioni anatomiche tra regioni) e la funzione (ad esempio le connessioni funzionali tra regioni). La risonanza magnetica funzionale (fMRI) misura il flusso sanguigno come marcatore indiretto dell'attività neurale. Durante l'esecuzione di un compito, infatti, l'attività neurale aumenta, comportando un maggiore consumo di ossigeno. A questo segue una risposta vascolare che compensa in eccesso la richiesta di ossigeno, determinando variazioni nel rapporto tra ossiemoglobina e deossiemoglobina, che a loro volta influenzano il segnale BOLD. Pertanto, le variazioni del flusso sanguigno riflettono la domanda metabolica sottostante piuttosto che l'attività neurale in modo diretto.

La correlazione tra i segnali BOLD di coppie di regioni definisce la connettività funzionale, ossia l'interazione tra regioni cerebrali distribuite, fondamentale per comprendere il funzionamento delle reti cerebrali. Tuttavia, sebbene la connettività funzionale sia una caratteristica importante per descrivere la dinamica cerebrale, essa non fornisce informazioni causali. Infatti, misura la correlazione tra segnali regionali senza considerare i meccanismi che guidano tali interazioni. Le interazioni causali all'interno di una rete cerebrale possono essere studiate attraverso la connettività effettiva. Il Dynamic Causal Modelling (DCM) è un framework consolidato che modella le interazioni causali tra regioni cerebrali, stimando sia la direzione sia l'influenza di tali interazioni attraverso un modello generativo.

L'obiettivo di questa tesi è stato quello di analizzare la connettività effettiva all'interno di una rete visuomotoria, definita come una rete che integra l'elaborazione visiva e la pianificazione motoria, comprendente la Corteccia Visiva, la Corteccia Motoria Primaria, l'Area Supplementare Motoria/Corteccia Premotoria, la Corteccia Cingolata, il Lobulo Parietale Superiore e il Cervelletto. In particolare, sono state confrontate due condizioni sperimentali: Action Observation (AO), in cui i soggetti osservano un'azione, e Action Observation con stimolo visivo aggiuntivo (AO-BAR), in cui l'osservazione è accompagnata da un segnale visivo supplementare utilizzato per guidare il compito.

L’analisi è stata condotta su due dataset fMRI specifici per le condizioni AO e AO-BAR ed è stata implementata in MATLAB utilizzando il toolbox SPM12. La pipeline ha incluso fasi di preprocessing, seguite dall’estrazione delle regioni della rete visuomotoria e delle corrispondenti serie temporali rappresentative. Successivamente, la modellazione della connettività è stata effettuata tramite DCM, e sono stati confrontati diversi modelli di organizzazione della rete, includendo sia modelli completamente connessi sia modelli ridotti basati su ipotesi derivate dalla letteratura (ad esempio Lorenzi et al., 2025), a livello sia del singolo soggetto sia del gruppo.

I risultati hanno mostrato che la condizione AO presenta un pattern di connettività effettiva stabile e ben supportato, con connessioni multiple, ad esempio dalla Corteccia Visiva verso le regioni motorie e il Cervelletto, riflettendo l’integrazione tra elaborazione visiva, pianificazione motoria e coordinazione all’interno della rete visuomotoria. Al contrario, la condizione AO-BAR non ha mostrato connessioni fortemente supportate e presenta una maggiore variabilità tra soggetti. Il Modello 4, che rappresenta un’architettura distribuita con interazioni tra regioni visive, motorie e cerebellari, è stato identificato come il modello più plausibile per il dataset AO.

Questi risultati suggeriscono che la rete visuomotoria è attivata in modo coerente durante l’osservazione di un’azione, mentre l’aggiunta di un segnale visivo non migliora la connettività e può invece aumentare il carico cognitivo, riducendo la stabilità della rete. Complessivamente, i risultati evidenziano come la connettività dipenda fortemente dal design del compito e come anche piccole modifiche sperimentali possano influenzare significativamente l’organizzazione delle reti cerebrali.

ABSTRACT (ENGLISH)

Magnetic Resonance Imaging (MRI) techniques enable the investigation of the human brain under multiple perspectives, ranging from its anatomy (e.g., the morphologic characterisation of different brain regions), microstructure (e.g., the anatomical connections between different regions) and its functions (e.g., the functional connections between different regions). Functional MRI (fMRI) measures blood flow as an indirect marker of neural activity. During the performance of a task, neural activity increases, leading to a higher consumption of oxygen. This is followed by a vascular response that overcompensates the oxygen demand, resulting in changes in the balance between oxyhemoglobin and deoxyhemoglobin, which in turn affects the BOLD signal. Therefore, changes in blood flow reflect the underlying metabolic demand rather than neural activity directly.

The correlation between the BOLD signals of pairs of regions defines functional connectivity, i.e., the interaction between distributed brain regions, which is essential to understand how functional brain networks operate. Despite functional connectivity being a fundamental feature to understand brain dynamics, it lacks causal information. Indeed, functional connectivity measures the correlation between regional signals without considering the mechanisms that drive these interactions. Causal interactions in a brain network can be measured through effective connectivity. Dynamic Causal Modelling (DCM) is an established framework that models causal interactions between brain regions, estimating both the direction and the influence of these interactions using a generative model.

The aim of this thesis was to examine effective connectivity within a visuomotor network, defined as a network integrating visual processing and motor planning, including the Visual Cortex, Primary Motor Cortex, Supplementary Motor Area/Premotor Cortex, Cingulate Cortex, Superior Parietal Lobule and the Cerebellum. In particular, two experimental conditions were compared: Action Observation (AO), where subjects observe an action, and Action Observation with visual cue (AO-BAR), where observation is combined with an additional visual cue. This additional visual information is used to guide the task.

The analysis was performed using two fMRI datasets specific for AO and AO-BAR. The pipeline included preprocessing steps, followed by the extraction of regions embedded in the visuomotor network and the estimation of representative time series. Connectivity modelling was then performed using DCM, and different models of network organisation, including both fully connected and hypothesis-driven reduced models informed by prior literature (e.g. Lorenzi et al., 2025), were compared at both subject and group levels.

The results showed that the AO condition exhibited a stable and well-supported effective connectivity pattern, with multiple connections such as from the Visual Cortex to motor regions and to the Cerebellum, reflecting the integration between visual processing, motor planning and coordination within the visuomotor network. In contrast, the AO-BAR condition did not show strongly supported connections and presented greater variability across subjects. Model 4, representing a distributed architecture with interactions between visual, motor and cerebellar regions, was identified as the most plausible architecture for the AO dataset.

These findings suggest that the visuomotor network is coherently engaged during action observation, whereas the addition of a visual cue does not improve connectivity and may instead increase cognitive demand, reducing the stability of the network. Overall, the results highlight how connectivity strongly depends on task design, and how even small experimental changes can significantly affect brain network organisation.

1. PRINCIPLE OF MAGNETIC RESONANCE IMAGING

Medical imaging includes a wide range of techniques that can be classified based on the type of radiation they use. A fundamental distinction is made between techniques that rely on ionizing radiation and those that use non-ionizing electromagnetic fields (Brown et al., 2014; McRobbie et al., 2017).

Techniques based on ionizing radiation, such as X-ray imaging and computed tomography (CT), use high-energy radiation that can penetrate biological tissues. While these methods are highly effective for diagnostic purposes, their use is associated with potential risks due to radiation exposure (Brown et al., 2014).

In contrast, Magnetic Resonance Imaging (MRI) is a non-invasive technique that does not use ionizing radiation. Instead, it is based on the interaction between magnetic fields and radio frequency pulses to generate detailed images of internal body structures. This makes MRI a safer alternative for repeated measurements and particularly suitable for studying tissues, including the brain (McRobbie et al., 2017).

1.1 PHYSICAL SIGNAL TO IMAGE

MRI is based on the principles of nuclear magnetic resonance (NMR). In this section, the physical mechanisms underlying MRI signal generation are introduced (Brown et al., 2014; McRobbie et al., 2017).

In NMR, atomic nuclei with non-zero spin behave as small magnetic dipoles, possessing angular momentum associated with a magnetic moment (Brown et al., 2014). When placed in an external magnetic field, these magnetic moments interact with the applied field. This behaviour is illustrated in Figure 1.1.

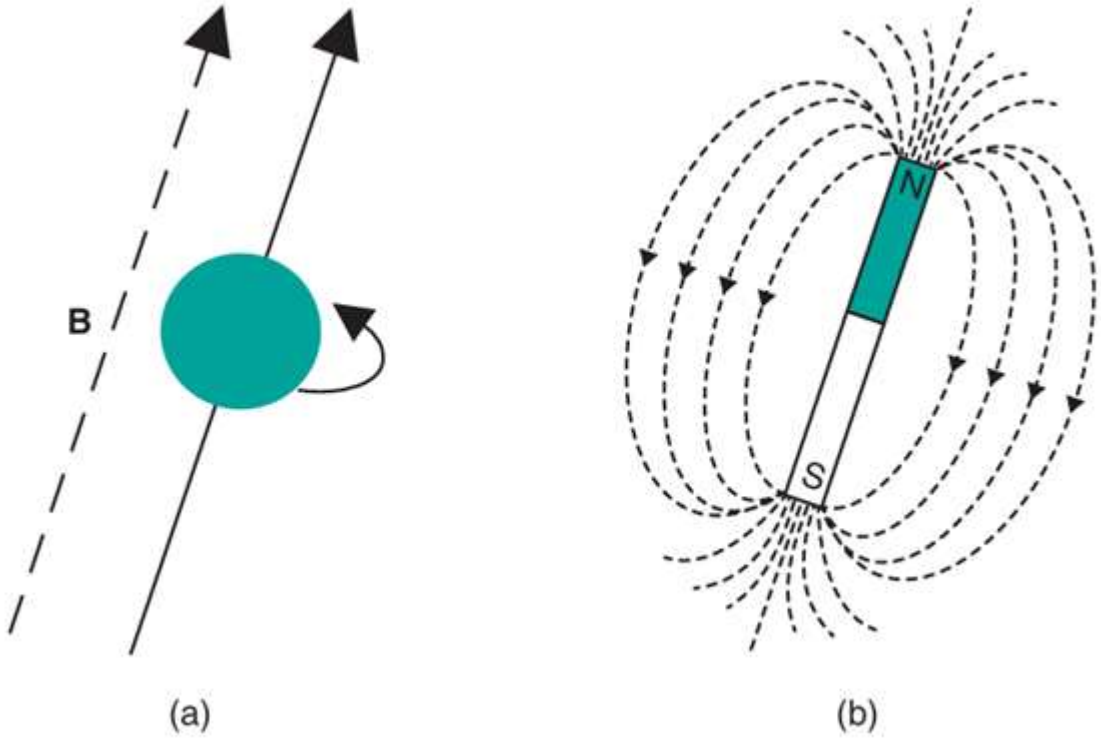

Figure 1.1: Behaviour of nuclei as magnetic dipoles. In (a), a rotating charged particle is shown, where rotation generates a magnetic moment. In (b), a bar magnet analogy is shown, with north and south poles, where the magnetic field lines go from north to south (Brown et al., 2014).

In biological tissues, hydrogen (^{1}H) nuclei are particularly abundant due to the presence of water and fat. MRI primarily exploits ^{1}H because of its high gyromagnetic ratio and strong signal sensitivity. (McRobbie et al., 2017). As a result, the measured MRI signal mainly arises from ^{1}H nuclei.

MRI does not detect individual spins, but rather the collective behaviour of a large ensemble of spins. To observe this macroscopic signal, a strong magnetic field, denoted as B0, is applied (Brown et al., 2014)

Under the influence of B0, nuclear spins precess around the direction of magnetic field, conventionally defined as the z-axis (McRobbie et al.,2017). At equilibrium spins can adopt two orientations: parallel or antiparallel. The parallel configuration corresponds to a lower energy state, while the antiparallel configuration corresponds to a higher energy state. Because slightly more spins occupy the lower energy state,

a net magnetization vector (M0) emerges along the direction of B0. (Brown et al., 2014). This behaviour is illustrated in Figure 1.2.

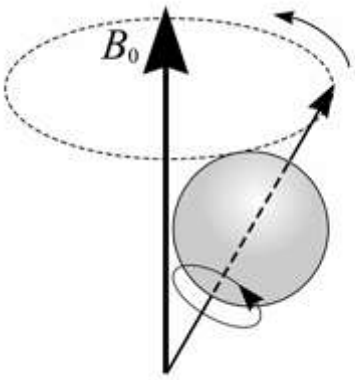

Figure 1.2: [Precession of the magnetic moment around B0. In this figure, the vertical arrow represents B0, while the tilted arrow represents the magnetic moment. The circular trajectory represents the precessional motion. (Brown et al.,2014).

The frequency of precession of nuclear spins is described by the Larmor equation:

$$\omega_0 = \gamma B_0 \tag{1.1}$$

Where γ is the gyromagnetic ratio, a constant specific to each type of nucleus (McRobbie et al., 2017)

At equilibrium, no measurable MR signal is observed, as the transverse components of magnetization cancel out. The net magnetization is therefore aligned along the longitudinal axis.

To generate a measurable signal, a radiofrequency (RF) pulse is applied. This pulse perturbs the system, rotating the net magnetization away from the longitudinal axis into the transverse plane. Once a transverse component is created, it becomes detectable by the MRI system (Brown et al. 2014).

The RF pulse is applied for a short duration. After excitation, the system gradually returns to equilibrium through a process known as relaxation (McRobbie et al., 2017).

Two main types of relaxation can be observed:

a. Longitudinal relaxation (T1)

T1 describes the recovery of magnetization along the direction of B0. This characteristic time constant of this process is the longitudinal relaxation time, T1:

$$M_Z(\mathrm{t}) = M_0\left(1 - \mathrm{e}^{-\frac{\mathrm{t}}{\mathrm{T1}}}\right) \tag{1.2}$$

Where Mz(t) represents the magnetization along the z-axis at time t, M0 is the equilibrium magnetization, and T1 is the time required for longitudinal magnetization to recover approximately 63% of its equilibrium value. This process is also referred to as spin-lattice relaxation (Brown et al., 2014) and it is illustrated on Figure 1.3.

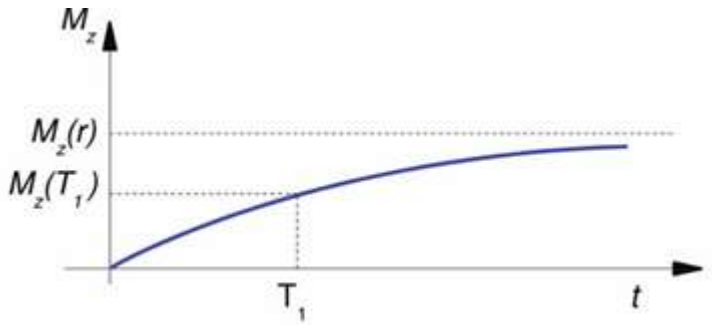

Figure 1.3: Evolution of the longitudinal magnetization (Mz(t)) over time. The curve shows how Mz approaches its equilibrium value, reflecting the energy exchange with the surrounding environment as spins return to a lower energy state (Coriasco et al., 2014)

b. Transverse relaxation (T2)

T2 describes the decay of magnetization in the transverse plane:

$$M_{xy}(\mathrm{t}) = M_0\left(\mathrm{e}^{-\frac{\mathrm{t}}{\mathrm{T2}}}\right) \tag{1.3}$$

Where Mxy(t) is the transverse magnetization at time t, Mxy(0) is the initial value, and T2 is the time constant describing the decay of transverse magnetization to 37% of its initial value. This process is also known as spin-spin relaxation (McRobbie et al.,2017)

The day of transverse magnetization following RF excitation is observed as the Free Induction Decay (FID). (McRobbie et al., 2017). This process is illustrated in Figure 1.4.

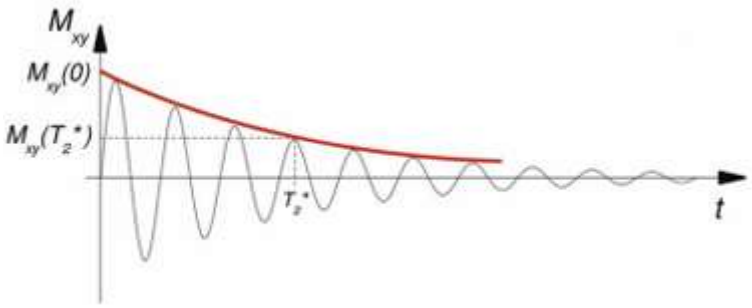

Figure 1.4: Evolution of the Mxy(t) over time. The curve shows an oscillating signal whose amplitude decreases over time, producing the FID. This decay is due to the loss of phase coherence between spins (Coriasco et al., 2014)

In practice, transverse relaxation is influenced not only by spin-spin interactions but also by magnetic field inhomogeneities. Real MRI systems do not have a perfectly uniform B0, and local susceptibility differences introduce additional dephasing effects (Haacke et al., 1999).

These combined effects are described by effective transverse relaxation time, T2*, defined as:

$$\frac{1}{\mathrm{T2}*} = \frac{1}{\mathrm{T2}} + \frac{1}{\mathrm{T2}'} \tag{1.4}$$

Where T2* represents additional dephasing due to field inhomogeneities. As a result, T2* is always shorter than T2 (Haacke et al., 1999).

The difference is illustrated in Figure 1.5.

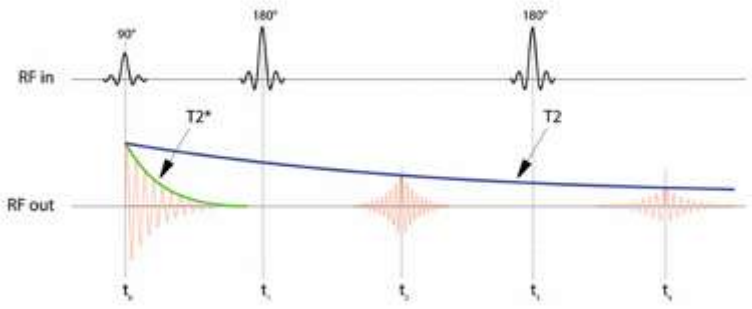

Figure 1.5: RF pulses and signal evolution over time. In the top row, the 90° and 180° RF pulses are shown, while the bottom row represents the signal evolution over time. After the 90° pulse, spins are in the transverse plane; therefore, the signal starts oscillating (FID). However, it decays very rapidly due to both spin-spin interactions and magnetic field inhomogeneities. As a result, the observed decay corresponds to T2*. The 180* pulse is applied after time t1. It flips the spins and reverses the dephasing caused by magnetic field inhomogeneities (Coriasco et al 2014),

1.2 ACQUISITION

A pulse sequence defines the timing, order, and combinations of RF pulses and magnetic field gradients used to acquire MR data (Brown et al., 2014; McRobbie et al., 2017).

Several types of pulse sequences are commonly used, including Spin Echo (SE), Gradient Echo (GRE), and Echo Planar Imaging (EPI).

SE sequences involve at least two RF pulses: initial excitation pulse followed by one or more 180° refocusing pulses. The refocusing pulse compensates for dephasing cause by static magnetic field inhomogeneities, allowing the formation of a SE. As a result, SE sequences produce a signal that reflects true T2 contrast (Brown et al., 2014). This process is illustrated in figure 1.6.

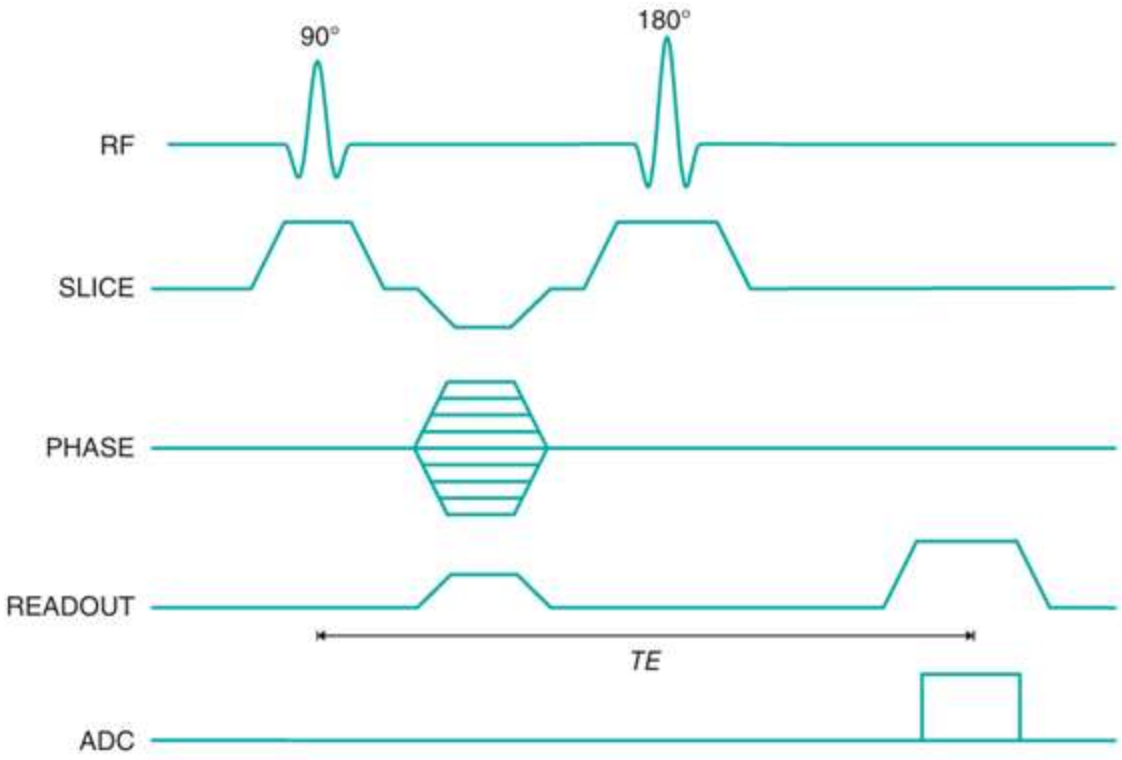

Figure 1.6: SE sequence. The temporal diagram shows a 90° RF pulse. Subsequently, dephasing occurs. Then, a 180° pulse is applied, after which rephasing occurs. Finally, an echo is formed and the signal is acquired. (Brown et al., 2014).

GRE sequences, in contrast, do not use a 180° refocusing pulse. Instead, the echo is generated by reversing the polarity of the magnetic field gradients. This reversal compensates for dephasing and produces a GRE. Because GRE sequences do not correct for magnetic field inhomogeneities, they are sensitive to T2* effects (McRobbie et al., 2017). This process is illustrated in figure 1.7.

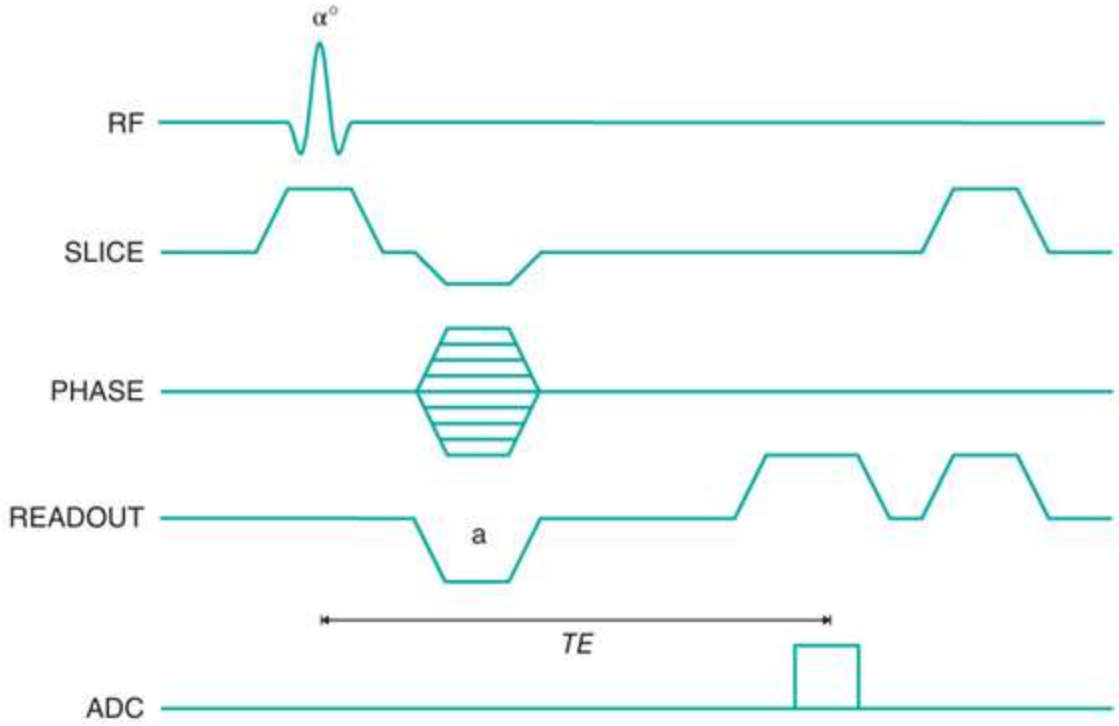

Figure 1.7: GRE sequence. The temporal diagram shows that an RF pulse is applied, after which dephasing occurs. Subsequently, gradient reversal is applied, and then the echo is generated, and the signal is acquired (Brown et al., 2014).

EPI is a fast-imaging technique characterized by rapid gradient switching. Following a single RF excitation, multiple lines of k-space are acquired through a series of gradient reversals. This allows very fast image acquisition, making EPI particularly suitable for functional MRI (fMRI). However, EPI is also highly sensitive to magnetic susceptibility (Turner, 1994). This process is illustrated in figure 1.8.

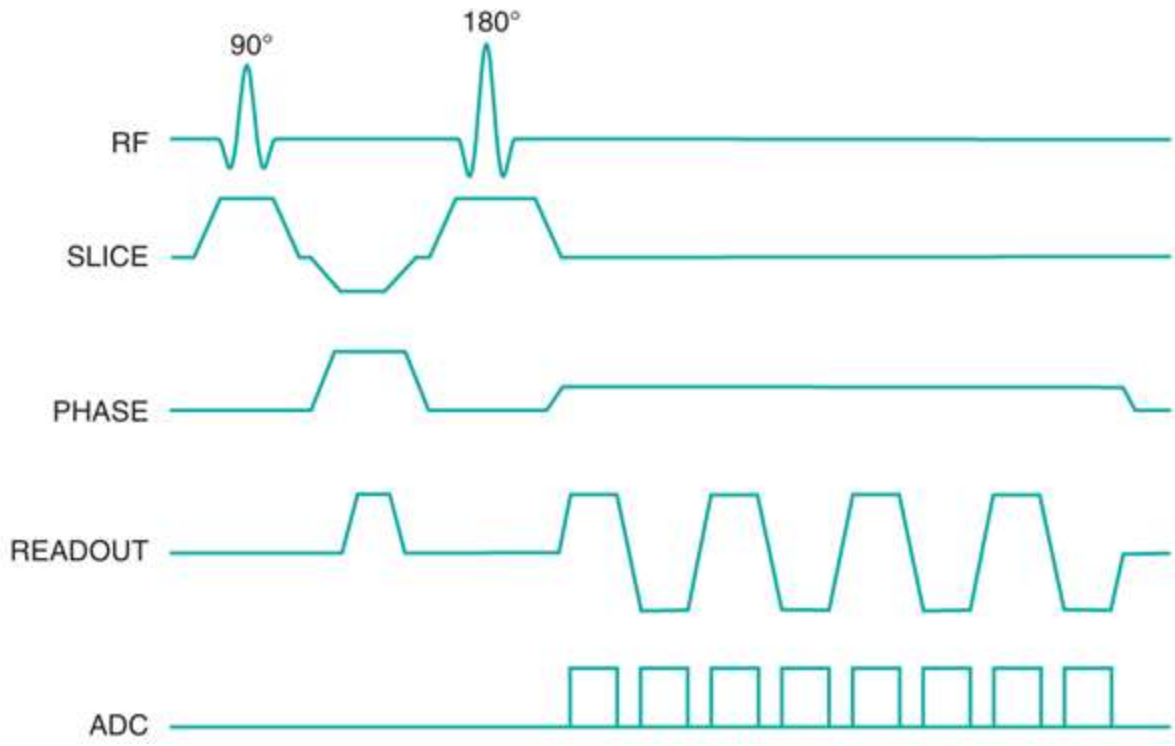

Figure 1.8: EPI sequence. The temporal diagram shows a single RF excitation. Subsequently, rapid gradient switching is applied. Then, multiple echoes are acquired, resulting in a fast acquisition. (Brown et al., 2014).

To acquire an MR image, spatial information must be encoded into a measured signal through a controlled sequence of RF pulses and magnetic field gradients. This process determines how the object is sampled in space and how the corresponding data populate k-space.

k-space is the spatial-frequency domain in which the MR signal is encoded through phase- and frequency-encoding gradients. It is filled progressively by repeated executions of the pulse sequence, with each repetition sampling a distinct k-space trajectory or line, until the full k-space matrix required for accurate spatial reconstruction of the image is acquired. In both spin-echo (SE) and gradient-echo (GRE) sequences, each repetition samples a distinct line or trajectory of k-space until the full matrix required for image reconstruction is acquired. (Brown et al., 2014).

Slice selection restricts the acquisition to a specific anatomical region. This is achieved by applying an RF pulse in the presence of a slice-select gradient, typically along the z-axis. Because the Larmor frequency varies linearly with position under a gradient,

only spins within a specific frequency range are excited. The RF bandwidth and gradient strength therefore determine the thickness and location of the slice (Brown et al., 2014).

In practice, multiple slices are acquired sequentially to cover the entire brain. To reduce interference between adjacent slices (slice cross-talk), interleaved acquisition schemes are often used, where non-adjacent slices are acquired in successive repetitions.

Following slice selection, spatial encoding within the slice is achieved through phase encoding. A gradient is briefly applied along one direction (e.g., the y-axis), introducing a position-dependent phase shift in the transverse magnetization.

By repeating the acquisition with different phase-encoding gradients, spatial information is encoded in k-space. (Buxton. 2013).

Spatial encoding along the orthogonal direction is achieved through frequency encoding. During signal readout, a gradient is applied (e.g. along x-axis), causing spins at different locations to precess at different frequencies. This allows spatial localisation based on frequency differences in the received signal (Brown et al., 2014).

The combination of slice selection, phase encoding, and frequency encoding enables full spatial reconstruction of the image.

Each repetition of the acquisition sequence contributes data to a specific region of k-space. By systematically varying the phase-encoding gradient and acquiring frequency-encoding signals, k-space is progressively filled, forming the basis for image reconstruction.

1.3 IMAGE RECONSTRUCTION

The MRI signal arises from the combined contribution of a large number of nuclear spins, resulting in a time-domain signal in which spatial information is not directly accessible (Brown et al. 2014; Liang & Lauterbur, 2000).

To encode spatial information, magnetic field gradients are applied, introducing a controlled linear variation in the static magnetic field B0 (Brown et al., 2014). These gradients allow the signal to become position dependent.

To reconstruct an image from the acquired signal, a mathematical transformation is required. Specifically, the Fourier transform is used to convert the measured signal from the time domain into the frequency domain. The acquired signal can therefore be interpreted as the Fourier transform of the spatial distribution of spin density (Liang & Lauterbur, 2000).

The spatially encoded data are stored in k-space, a mathematical representation that describes how spatial frequencies contribute to the final image (Moratal et al., 2008). The coordinates of k-space, typically denoted kx and ky, are related to the time integral of the applied magnetic field gradients, scaled by γ.

Data acquired near the centre of k-space mainly influence image contrast and overall signal intensity, while data in the outer regions contribute to spatial resolution and edge definition.

In practice, continuous sampling of k-space is not feasible, and only a finite number of points can be acquired. The quality of the reconstructed image therefore depends on how completely and accurately k-space is sampled. (Liang & Lauterbur, 2000)

The MRI signal arises from the combined contribution of a large number of nuclear spins, resulting in a time-domain signal in which spatial information is not directly accessible (Brown et al., 2014; Liang & Lauterbur, 2000).

To encode spatial information, magnetic field gradients are applied, introducing a controlled linear variation in the static magnetic field B0 (Brown et al., 2014). These gradients allow the signal to become position dependent.

To reconstruct an image from the acquired signal, a mathematical transformation is required. The Fourier transform is used to convert the measured signal from the time domain into the frequency domain. The acquired signal can therefore be interpreted as the Fourier transform of the spatial distribution of spin density (Liang & Lauterbur, 2000).

The spatially encoded data are stored in k-space, a mathematical representation that describes how spatial frequencies contribute to the final image (Moratal et al., 2008). The coordinates of k-space, typically denoted as kx and ky, are related to the time integral of the applied magnetic field gradients, scaled by γ.

Data acquired near the centre of k-space mainly influence image contrast and overall signal intensity, while data in the outer regions contribute to spatial resolution and edge definition.

In practice, continuous sampling of k-space is not feasible, and only a finite number of points can be acquired. The quality of the reconstructed image therefore depends on how completely and accurately k-space is sampled (Liang & Lauterbur, 2000).

The Nyquist criterion defines the minimum sampling frequency required to acquire a signal without loss of information. This frequency must be at least twice the highest frequency present in the signal (Bernstein et al., 2004).

If this condition is not satisfied, aliasing occurs. In this case, high-frequency components are incorrectly represented and overlap with lower-frequency components, resulting in signal distortion (Hashemi et al., 2010).

In the context of MRI, the signal is sampled in k-space. However, if the sampling is insufficient, i.e. undersampling, this leads to aliasing. Consequently, in the reconstructed image, structures overlap and appear as wrap-around artifacts (Bernstein et al., 2004).

To avoid this phenomenon, k-space must be sampled sufficiently. However, this results in an increase in acquisition time (Bernstein et al., 2004).

A trade-off therefore exists between acquisition time and spatial resolution: higher resolution requires denser sampling of k-space, which increases acquisition time (Bernstein et al., 2004).

Finally, the reconstructed image is obtained by applying the inverse Fourier transform to the sampled k-space data (Brown et al., 2014).

1.4 TISSUE CONTRAST

The contrast observed in MRI images arises from intrinsic tissue properties and acquisition parameters. In particular, the relaxation times T1, T2 and T2* reflect the interactions between nuclear spins and their surrounding molecular environment (Brown et al., 2014; McRobbie et al., 2017).

These intrinsic properties cannot be directly controlled by the operator. However, acquisition parameters such as repetition time (TR), and echo time (TE) can be adjusted to influence how the signal is sampled.

TR is defined as the time interval between successive RF excitation pulses applied to the same slice. TR determines the extent to which longitudinal magnetization recovers before the next excitation, thereby affecting signal amplitude and image contrast (McRobbie et al, 2017).

TE is the time interval between the RF excitation pulse and the acquisition of MR signal. TE determines how much transverse magnetization decay had occurred at the time of signal measurement, influencing the detected signal intensity (Brown et al., 2014).

Although TR and TE can be adjusted independently, their combination defines the temporal structure of the acquisition sequence and strongly influences image contrast.

Image weighting arises from the interaction between intrinsic tissue properties (T1, T2, T2*) and the selected acquisition parameters (TE, TR). By appropriately choosing these parameters, it is possible to emphasize specific tissue characteristics depending on the imaging objectives.

In T1-weighted imaging, short TR values are used to limit longitudinal recovery between excitations enhancing differences in T1 relaxation times and providing high-resolution anatomical contrast.

In T2-weighted imaging, long TR values are combined with long TE values to minimize T1 effects and emphasize differences in transverse relaxation. This type of contrast is

particularly useful for detecting fluid-related and pathological changes (McRobbie et al., 2017).

T2*-weighted imaging is sensitive to magnetic field inhomogeneities and susceptibility effects. In this case, TE is selected to enhance sensitivity to T2* decay, resulting in contrast that reflects local variations in magnetic properties (Haacke et al., 1999).

In proton density (PD)-weighted imaging, the influence of T1 and T2 relaxation is minimized by using long TR and short TE values. This allows the signal intensity to primarily reflect the concentration of 1H nuclei within the tissue.

These differences are illustrated in Figure 1.9.

These contrast mechanisms form the basis for functional MRI (fMRI), where T2*-weighted imaging is used to detect changes in the following chapter.

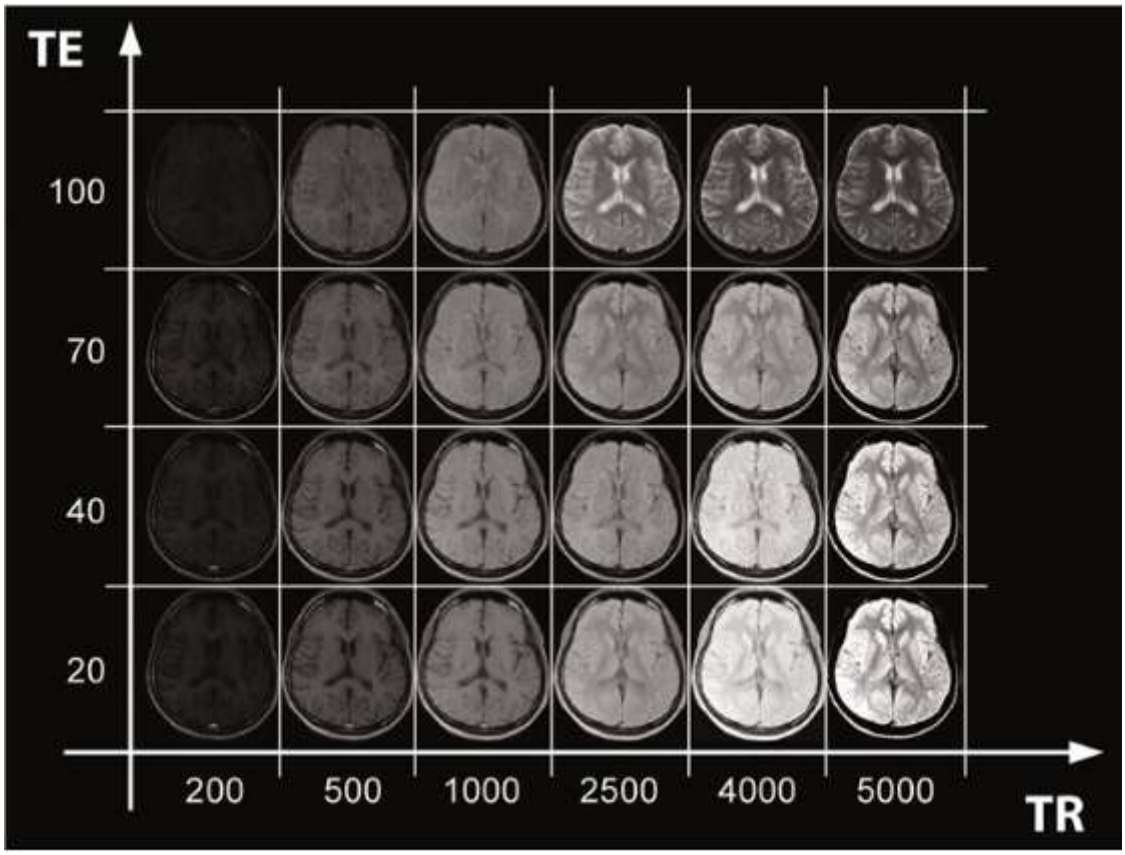

Figure 1.9: Example image illustrating the effect of TR and TE on MRI image contrast. (Coriasco et al., 2014).

Figure 1.9 demonstrates how varying TR and TE influences image contrast. As TE increases, transverse magnetization decay becomes more pronounced, leading to a reduction in signal intensity. As TR increases, longitudinal magnetization recovery is greater, resulting in an increase in signal intensity. Different combinations of TR and TE emphasize different contrast weightings, allowing specific tissue properties to be highlighted.

2. FUNCTIONAL MAGNETIC RESONANCE IMAGING

Functional MRI (fMRI) is an MRI technique used to investigate the relationship between neural activity and cerebral blood flow (CBF). The primary objective of fMRI is to detect changes in metabolic demand and oxygen consumption associated with neuronal activation (Buxton et al., 2004; Logothetis et al., 2001).

When neurons become active, synaptic activity increases, which elevates metabolic demand and leads to enhanced consumption of oxygen and glucose. This process, known as neurovascular coupling, triggers a localized increase in CBF, and consequently, an increase in cerebral blood volume (CBV) (Buxton et al., 2004; Logothetis et al., 2001).

An important component of blood is haemoglobin, an iron-containing protein. Depending on its oxygenation state, haemoglobin exists in two forms: oxyhaemoglobin and deoxyhaemoglobin. Deoxyhaemoglobin is paramagnetic, whereas oxyhaemoglobin is diamagnetic (Pauling & Coryell, 1936).

Deoxyhaemoglobin introduces microscopic magnetic susceptibility gradients, which accelerate T2* decay (Ogawa et al., 1990). When CBF increases, the concentration of deoxyhaemoglobin decreases. This reduction leads to a decrease in susceptibility gradients, a slowing of T2* decay, and consequently an increase in MR signal intensity.

This phenomenon is known as the Blood Oxygenation Level Dependent (BOLD) effect and it is illustrated in Figure 2.1.

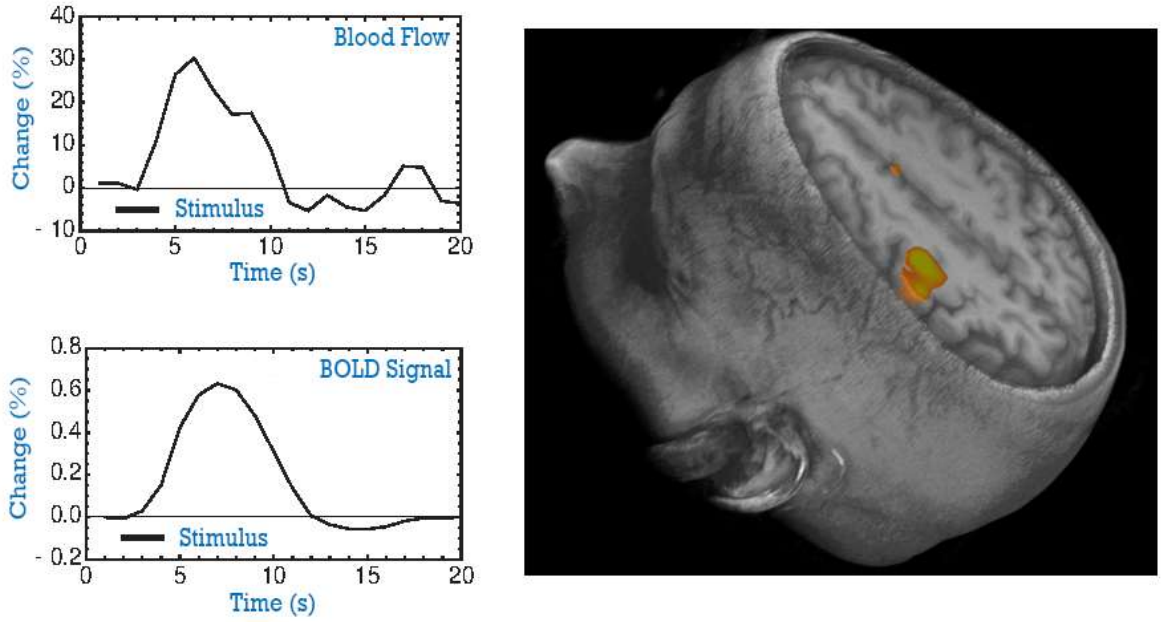

Figure 2.1: BOLD SIGNAL AND CBF. The top graph represents the change in blood flow over time, while the bottom graph shows the corresponding BOLD signal.

It represents an indirect measure of neural activity, reflecting changes in T2*-weighted signal due to variations in blood oxygenation (Ogawa et al., 1990; Buxton et al., 2004).

As discussed in Chapter 1, MRI is sensitive to magnetic properties. Since haemoglobin changes its magnetic behaviour depending on oxygen binding, neural activity produces measurable variations in the MR signal, which are captured in fMRI BOLD contrast.

This temporal evolution of the signal is known as the Hemodynamic Response Function (HRF).

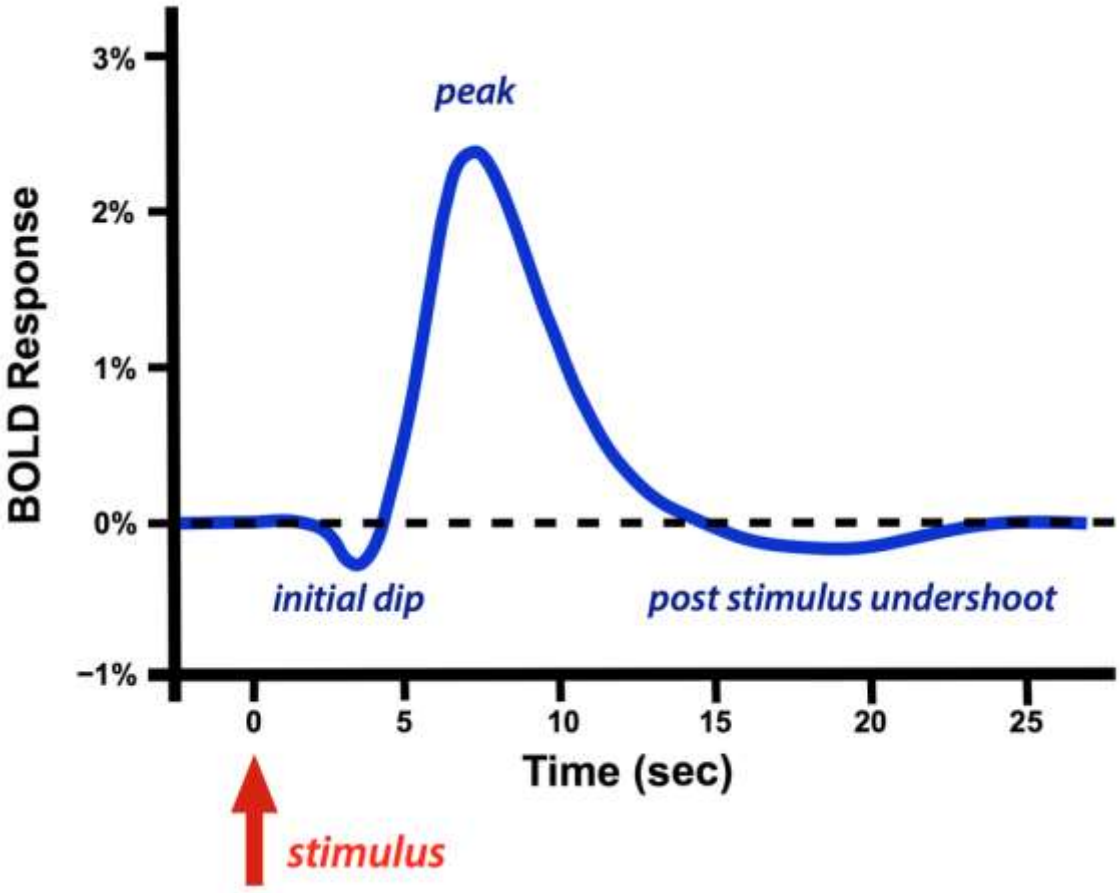

Figure 2.2: HRF. Immediately after the stimulus, a slight decrease in signal is observed (initial dip). After a few seconds, the signal increases and reaches a peak. Following the peak, the signal decreases below baseline (post-stimulus undershoot). This reflects the vascular response to neural activity (Buxton et al., 2004).

The BOLD response presents an initial delay of several seconds, reaches a peak and is followed by a post-stimulus undershoot due to vascular dynamics (Buxton et al., 2004).

2.1 ACQUISITION OF FUNCTIONAL MAGNETIC RESONANCE IMAGING

As described in the previous section, fMRI detects BOLD contrast, which reflects changes in blood oxygenation levels.

The BOLD response evolves over a time scale of seconds, so the acquisition of fMRI must reflect this temporal characteristic. For this reason, EPI is used. As introduced in

Section 1.3, GRE EPI allows rapid acquisition of multiple k-space lines following a single RF excitation, enabling fast whole-brain imaging (Turner, 1994; Mansfield, 1977).

Another important aspect of fMRI concerns the acquisition method. MRI data are acquired slice by slice due to spatial encoding constraints, and multiple slices are combined to form a whole-brain volume. Slice acquisition can be performed sequentially or interleaved manner, depending on the acquisition protocol (Huettel et al., 2014).

From the perspective of fMRI, the TR corresponds to the time required to acquire one whole-brain volume. A shorter TR improves temporal resolution but may reduce spatial coverage or signal-to-noise ratio (SNR) (Huettel et al., 2014).

Finally, fMRI data are acquired as a series of repeated volumes over time. This repeated acquisition allows tracking of BOLD signal changes and enables statistical modelling of task-related activity (Poldrack et al., 2011).

2.2 PRE-PROCESSING OF FMRI DATA

fMRI data are inherently complex and contain a mixture of meaningful neural signals and unwanted variability. Factors such as head motion, physiological fluctuations, and scanner-related imperfections introduce artifacts that can obscure true brain activity. To ensure that subsequent analyses are valid and the data to reduce these confounding influences.

Pre-processing consists of a series of steps designed to improve data quality and prepare the data quality and prepare the data for statistical modelling (Poldrack et al., 2011).

One of the first steps is slice timing correction (STC). Since slices are acquire at slightly different time points within each TR, interpolation is used to temporarily align them, ensuring that all slices correspond to the same acquisition time (Poldrack et al., 2001).

Another important step is motion correction, also known as realignment. Head motion during scanning can introduce successive images. To address this, each image in the fMRI time series is aligned to a reference image using image registration techniques, assuming a rigid body transformation (Jenkinson et al., 2002).

Spatial normalization is then applied to transform individual brain images into a standard anatomical space. This step enables comparisons across subjects by ensuring that corresponding brain regions are aligned across datasets (Ashburner & Friston, 1999).

Spatial smoothing is performed by convolving the images with a three-dimensional (3D) Gaussian kernel. This step increases the SNR and helps account for anatomical variability across subjects (Poldrack et al., 2011).

Finally, temporal filtering is applied to remove low-frequency drift and other unwanted temporal fluctuations, improving the quality of the signal for subsequent statistical analysis (Worsley & Friston, 1995).

2.3 GENERAL LINEAR MODEL

The fMRI signal is complex because it contains a mixture of different sources, such as physiological fluctuations, head motion, and scanner-related noise. Therefore, a model is required to separate these contributions and explain the signal. This model is known as the General Linear Model (GLM) (Friston et al., 1994; Poldrack et al., 2011).

The signal can be described as composed of two parts: an explained component and an unexplained component. The explained part corresponds to the experimental task, while the unexplained part corresponds to noise.

To apply the GLM correctly and obtain reliable results in fMRI, the experimental paradigm must be accurately represented in terms of its temporal structure. This is done by describing the experiment as a series of conditions, each corresponding to a specific task or stimulus. Each condition is characterised by two key elements: onsets and durations. The onset specifies the time point at which a task or stimulus begins, while the duration indicates how long the task or stimulus lasts. Together, these

parameters define the neural events that are hypothesized to evoke changes in the BOLD signal.

These neural events are then convolved with a canonical Hemodynamic Response Function (HRF), which models how neural activity is translated into a measurable BOLD signal (Glover, 1999; Friston et al., 1998). This process generates regressors that represent the expected BOLD response over time for each condition.

By including these regressors in the design matrix, the GLM estimates the contribution of each experimental condition to the observed BOLD signal while accounting for noise and confounding effects. Correct specification of onsets and durations is therefore critical, as inaccurate timing may lead to poorly modelled responses and reduced statistical sensitivity.

These neural events are then convolved with a canonical HRF,), which models how neural activity is translated into a measurable BOLD signal (Glover, 1999; Friston et al., 1998). This process generates regressors that represent the expected BOLD response over time for each condition.

To obtain a signal comparable to the measured one, the regressors are convolved with the HRF. The resulting convolved regressors represent predictors of the BOLD signal. The HRF is important because the BOLD signal is temporally delayed and dispersed; therefore, it introduces both the delay and the characteristic shape of the response.

Regressors can include task-related components as well as physiological contributions, often referred to as confounds. All these regressors are combined to form the design matrix (X). The columns of X represent the regressors, while the rows represent time points corresponding to the TR.

The GLM is described by the equation:

$$Y = X\beta + \varepsilon \tag{2.1}$$

Where Y represents the observed data, X is the design matrix, β are the parameter estimates, and ε represents the unexplained variance. The GLM estimates the β parameters to explain the observed signal, allowing the identification of task-related activity (Friston et al., 1994).

Once the parameters β are estimated, it is necessary to test specific hypotheses about the experimental conditions. The β parameters alone are not sufficient, as they represent numerical estimates that require interpretation in relation to the experimental hypothesis. This is done using contrasts, which are linear combinations of β (Friston et al., 1994).

In fMRI analysis, contrasts are used to test specific hypotheses about brain activity by comparing the contributions of different conditions or regressors in the GLM. There are two main types of contrasts: t-contrasts and F-contrasts.

A t-contrast evaluates a specific hypothesis regarding a single condition or a comparison between conditions. It is directional, meaning it tests whether the effect of one condition is greater or smaller than another. For example, a t-contrast comparing condition A to condition B tests the hypothesis “A > B,” identifying brain regions where activity is significantly higher during condition A than condition B. t-contrasts produce a t-statistic at each voxel, which is used to assess statistical significance across the brain.

The F-contrast, instead, evaluates multiple parameters simultaneously to determine whether a set of conditions has a significant effect on the data. It is non-directional, as it tests whether any of the specified effects explain a significant portion of the signal variance, without indicating the direction of the effect.

The result of applying the contrasts are statistical maps. These maps are images in which each voxel is associated with a statistical value, such as a t-value or F-value. They are obtained by applying the contrast at each voxel, resulting in a statistical value for every location.

A statistical threshold is then applied to the resulting parameter estimates to identify only those effects that are unlikely to have occurred by chance. This procedure

ensures that only statistically significant activations are highlighted, reducing the likelihood of false-positive findings. As a result, only activations associated with the task are retained. These statistical maps therefore represent the spatial distribution of brain activity related to the experimental conditions (Friston et al., 1994; Poldrack et al., 2011).

3. BRAIN CONNECTIVITY

The brain can be described as a complex network of interacting regions. Connectivity refers to the relationship between spatially separated brain areas.

Brain connectivity can be broadly classified into three main categories: structural connectivity, functional connectivity, and effective connectivity (Friston, 2011; Sporns, 2013).

3.1 STRUCTURAL CONNECTIVITY

Structural connectivity refers to the physical information-carrying connections between neural populations, typically represented by white matter tracts connecting predefined grey matter regions (Sporns et al., 2005). It is typically obtained by processing diffusion-weighted MRI data (DWI), and it represents anatomical pathways, thereby providing structural constraints (Basser et al., 1994; Le Bihan, 2003).

The key principle is that each voxel is acquired multiple times, with each acquisition being sensitive to a specific diffusion direction. From this set of images, the spatial profile of preferred diffusion directions can be estimated (Sporns et al., 2003).

This allows the inference of the presence and orientation of white matter fibres within a voxel and enables reconstruction of fibre pathways across the brain. By combining this information spatially, tractography techniques generate a representation of white matter tracts connecting different brain regions (Mori & van Zijl, 2007).

The resulting tractogram provides a detailed visualisation of structural connections, as well as quantitative information about connectivity between brain regions, often summarised in a structural connectivity matrix.

3.2 FUNCTIONAL CONNECTIVITY

Functional connectivity is usually inferred based on correlations among measurements of neuronal activity. It is defined as statistical dependencies among spatially remote neurophysiological signals (Friston, 1994).

It can be quantified using measures of statistical relationships between signals measured in different brain regions, such as correlation or coherence. These analyses usually entail identifying predominant patterns of correlation, for example using principal component analysis (PCA) or independent component analysis (ICA), or

assessing whether a specific correlation between two regions is statistically significant (Friston, 2011; Smith et al., 2011).

3.3 EFFECTIVE CONNECTIVITY

Effective connectivity refers to the influence that one neural system exerts over another (Friston, 2011). It is dynamic and activity-dependent, and it relies on an explicit model of interactions between brain regions. In this sense, effective connectivity corresponds to the parameters of a model that aims to explain observed statistical dependencies, such as functional connectivity (Friston, 2011).

Therefore, it reflects directed causal interactions rather than simple correlations. Effective connectivity is typically used to test hypotheses about coupling architecture by comparing different models in terms of their statistical evidence given the observed data (Friston et al., 2003; Stephan et al., 2010).

4. DYNAMIC CAUSAL MODELLING

As discussed in Section 2.3, the GLM identifies task-related activations; therefore, based on GLM alone it is not possible to infer interactions between brain regions. In fact, activation is local rather than relational (Friston et al., 1994; Poldrack et al., 2011).

Another aspect, as explained in Chapter 3, is that functional connectivity is based on correlations between time series, but it does not explain the direction or causality of interactions between brain regions. Therefore, correlation does not imply causal influence (Friston, 1994; Friston, 2011).

The brain can be considered a dynamic system, meaning that interactions between regions evolve over time. For this reason, functional connectivity alone is not sufficient to fully describe these interactions, as the state of one region can influence another over time (Friston et al., 2003).

To understand these interactions, there is a need for a generative model, i.e. a model that produces the observed data and explains the underlying mechanisms. In this sense, it explains not only the observed data but also how these data are generated (Friston et al., 2003; Stephan et al., 2010).

A model that satisfies these requirements is Dynamic Causal Modelling (DCM). It is a generative and hypothesis-driven model, meaning that it is based on predefined hypotheses about the interactions between brain regions. In this framework, DCM aims to explain how neuronal activity evolves over time and how interactions between regions give rise to the observed data (Friston et al., 2003; Stephan et al., 2009).

In the context of fMRI, DCM explains the measured BOLD signal in terms of underlying neuronal activity, allowing inference on hidden neural processes (Stephan et al., 2009).

In this framework, the system is described in terms of hidden neural states, which are not directly observable, and a forward model that maps these states to the measured BOLD signal (Friston et al., 2003; Stephan et al., 2010).

DCM is not a single model; it is composed of two interacting levels. The first level, known as the neuronal model, describes the interactions between brain regions and how activity evolves over time. The second level, known as the hemodynamic model, links neuronal activity to the BOLD signal and explains how the measured signal is generated (Friston et al., 2003).

These components will be described in detail in the following sections (Friston et al., 2003; Stephan et al., 2010).

4.1 NEURONAL MODEL

Neural states change over time and are mathematically represented as dx/dt. They depend on the current state (x) and inputs (u), and this relationship is described by a state equation, which models how neural activity evolves over time (Friston et al., 2003; Stephan et al., 2010).

The system describing the brain is nonlinear due to its complex structure and the nature of interactions between regions. However, solving such a system analytically is difficult; therefore, a Taylor expansion is used as an approximation, making the system tractable (Friston et al., 2003).

This provides a local approximation of the nonlinear function, corresponding to a linearization around a baseline. The baseline is defined by the resting state x = 0 and the absence of input u = 0 (Friston et al., 2003).

Brain regions are represented as state vectors. Considering multiple interacting regions (nodes), the system can be expressed in vector form, and interactions between regions can be represented using matrices.

After applying the approximation, the neural state equation becomes:

$$\dot{x} = \left(A + \sum u_i B_j\right) x + Cu \tag{4.1}$$

This equation is known as the bilinear approximation, as it is linear in x and u separately (Friston et al., 2003).

The connectivity matrix A represents the intrinsic coupling between regions in the absence of input, describing how one region influences another at baseline (Friston et al., 2003; Stephan et al., 2010).

The matrices B_j represent changes in coupling induced by external inputs. They encode how inputs modulate the interactions between regions, meaning that connections are not fixed but can be influenced by experimental conditions (Friston et al., 2003).

The matrix C represents the direct influence of external inputs on neuronal activity, specifying where inputs enter the system and drive regional responses (Friston et al., 2003).

The difference between B and C is that B describes how inputs modulate connections between regions, whereas C defines where inputs directly affect regional activity.

The set of matrices {A, B_j, C} defines the connectivity parameters of the model and characterizes the functional architecture and interactions among brain regions at the neuronal level. The goal of DCM is to estimate these parameters to infer the strength and direction of interactions and to test hypotheses about network connectivity (Friston et al., 2003; Stephan et al., 2010).

4.2 HEMODYNAMIC MODEL

As discussed in Chapter 2, fMRI is an indirect technique to measure neural activity, and it produces the BOLD signal. Therefore, to establish a link between neural activity and the BOLD signal, a model is required. This model is the hemodynamic model, which links neuronal dynamics to the measured signal (Buxton et al., 1998; Friston et al., 2003).

This model is necessary because the neuronal model provides neural activity x(t), but fMRI does not directly observe x(t); instead, it measures the BOLD signal. Therefore,

a transformation is required that links neural activity to the BOLD signal through vascular responses. This transformation is neither linear nor instantaneous and is described by the Balloon model (Buxton et al., 1998).

To describe this transformation, the Balloon model is used to consider the physiological cascade: neural activity triggers a vasodilatory signal, which causes an increase in blood flow. Consequently, there is a change in blood volume, followed by changes in deoxyhaemoglobin.

These processes evolve over time (in seconds), and as a result, they generate the BOLD signal (Buxton et al., 1998).

For each brain region, four state variables are considered in the Balloon model: vasodilatory signal s(t), blood inflow f(t), blood volume v(t), and deoxyhaemoglobin content q(t). These variables are hidden states, as they are not directly observable, and they depend on the neuronal activity of that region. They are described by a system of differential equations (Friston et al., 2003).

The BOLD signal is not directly equal to flow or volume; instead, it is a nonlinear function of v(t) and q(t). As a result, the final output is y(t) (Buxton et al., 1998).

Therefore, the neuronal model and the hemodynamic model together form the full DCM framework (Friston et al., 2003; Stephan et al., 2010).

Considering the hemodynamic parameters, they exhibit slow dynamics and tend to be similar across regions; consequently, they are less informative for the study of connectivity.

The main objective is to understand connectivity, which is primarily determined by neuronal parameters, while the hemodynamic model plays a secondary role in the interpretation.

In practice, hemodynamic parameters are often fixed or less interpreted, whereas the focus is on neuronal parameters, as they provide information about connectivity between brain regions (Zeidman et al., 2019).

4.3 BAYESIAN INFERENCE

The problem of solving the model is that the model considered is a forward model, meaning that the starting point is the parameters and the output are observations. The starting point is the observations, and the goal is to estimate the parameters; therefore, it is necessary to introduce model inversion, with the aim of identifying parameter values such that the forward model generates the observed data (Friston et al., 2003; Stephan et al., 2010).

To perform this inversion, Bayesian inference is used. It updates beliefs about parameters after observing data, i.e. it combines prior knowledge and observed data (Friston et al., 2003).

The Bayes' rule is:

$$\mathrm{p}(\theta|\mathrm{y},\mathrm{m}) = \frac{\mathrm{p}(\mathrm{y}|\theta,\mathrm{m})\ \mathrm{p}(\theta|\mathrm{m})}{p(y|m)} \tag{4.2}$$

The prior p(θ | m) represents beliefs about parameters before observing data. Therefore, it constrains the model and prevents unrealistic solutions. In DCM, priors encode biological plausibility (Friston et al., 2003; Stephan et al., 2010).

The likelihood p(y | θ, m) is the probability of observing data given parameters and model. Therefore, it measures how well the model explains the data. In DCM, the generative model predicts the BOLD signal, which is then compared with the observed BOLD signal (Friston et al., 2003).

The posterior p(θ | y, m) is the updated belief after observing data, i.e. it combines prior and likelihood. Therefore, this represents the parameter estimation in DCM (Friston et al., 2003).

The model evidence p(y | m) is the probability of observing data under a model, i.e. it integrates over all possible parameter values:

The model evidence p(y | m) is the probability of observing data under a model, i.e. it integrates over all possible parameter values:

$$p(y \mid m) = \int p(y \mid \theta, m)\, p(\theta \mid m) d\theta$$

This means that model evidence measures how well a model explains the data and, at the same time, penalizes overly complex models, providing a balance between accuracy and complexity (Stephan et al., 2010).

Overall, the goal is to fit the data well (high likelihood) while maintaining model simplicity (respecting priors). Consequently, model evidence allows comparison between models, which leads to Bayesian Model Selection (BMS) (Stephan et al., 2010; Penny et al., 2004).

As seen before, model evidence is defined as an integral over the parameter space. However, the parameter space is highly dimensional; therefore, it is not analytically tractable (Friston et al., 2003; Stephan et al., 2010).

Instead of computing model evidence directly, an approximation of the posterior distribution q(θ) is introduced, which is chosen to be tractable (Friston et al., 2003). The goal is to make

$$q(\theta) \approx p(\theta \mid m) \tag{4.4}$$

To quantify the difference between these two distributions, the Kullback–Leibler (KL) divergence is used:

$$KL(q(\theta) \,||\, p(\theta \mid y, m)) \tag{4.5}$$

It measures the distance between the approximate posterior q(θ) and the true posterior p(θ | y, m). By definition, KL ≥ 0, and KL = 0 only when the two distributions are identical (Friston et al., 2003).

The aim of Variational Bayes is to minimize KL(q || p), thus making q(θ) as close as possible to the true posterior. However, instead of minimizing KL directly, Variational Bayes maximizes the Free Energy (F) (Friston et al., 2003; Penny et al., 2004).

This follows from the relation:

$$\log p(y \mid m) = F + KL(q \,||\, p) \tag{4.6}$$

Since KL ≥ 0, it follows that F ≤ log p(y | m); therefore, F represents a lower bound of the model evidence.

Maximizing F is therefore equivalent to minimizing the KL divergence, improving the approximation of the posterior and, at the same time, approximating the model evidence.

In practice, F can be expressed as:

$$F = \text{accuracy} - \text{complexity} \tag{4.7}$$

Accuracy measures how well the model predicts the data and corresponds to the likelihood term. Complexity measures the divergence between posterior and prior distributions, and it is therefore related to the KL divergence. It penalizes overly complex models and prevents overfitting (Stephan et al., 2010).

A good model must therefore fit the data well (high accuracy) while remaining simple (low complexity). F provides a balance between these two aspects.

In DCM, F is used to perform model inversion, where inversion is formulated as an optimization problem. The goal is to find the parameters θ that maximize F, corresponding to the best trade-off between accuracy and complexity, and therefore the best explanation of the data (Friston et al., 2003).

Instead of computing the posterior distribution exactly, inference is performed by maximising the variational free energy F with respect to q(θ), yielding an approximate posterior distribution while simultaneously providing an approximation to the model evidence.

Maximizing F does not yield a single parameter estimate, but rather a posterior distribution q(θ). Therefore, model parameters are treated probabilistically.

However, a distribution cannot be directly interpreted, so it is necessary to summarize the posterior using representative quantities.

The posterior mean (Ep) represents the expected value of the parameters and corresponds to the central estimate, i.e. the most likely parameter value.

The posterior covariance (Cp) represents the uncertainty of the parameters. The variance reflects the confidence in each estimate, while the covariance captures relationships between parameters (Friston et al., 2003).

Together, Ep and Cp provide a complete description of the posterior distribution.

Another important quantity is the posterior probability (Pp). It represents the probability that a parameter differs from zero and therefore measures the strength of evidence. A high value of Pp indicates strong evidence for a connection, whereas a low value indicates weak or no evidence (Stephan et al., 2010).

DCM uses a specific Variational Bayesian method, namely Variational Laplace. This method assumes that the posterior q(θ) follows a Gaussian distribution, which is fully described by its mean and covariance (Friston et al., 2003).

The procedure starts from the prior, then updates the posterior while maximizing Free Energy. This process is repeated iteratively until convergence is reached.

Therefore, model inversion in DCM consists of estimating the posterior distribution and maximizing F, resulting in both parameter estimates and an approximation of model evidence (Friston et al., 2003; Stephan et al., 2010).

In summary, F plays a central role in DCM by enabling parameter estimation and model evaluation within a unified Bayesian framework. It provides estimates of posterior distribution of model parameters including estimation through the posterior quantities (Ep, which represent the most likely parameter estimates, the posterior covariance, Cp, which reflects the uncertainty or variability around these estimates, and Pp), which quantify the confidence that a given parameter differs from zero, providing an approximation of model evidence, and forms the basis for model comparison.

4.4 BAYESIAN MODEL SELECTION

In the previous section, F was defined, where each model is associated with a score.

The goal is to compare multiple models, notated as m, (m_1, m_2, …, m_n) in order to determine which, one best explains the data.

Model comparison is performed using model evidence, which is approximated by F. Therefore, the best model corresponds to the one with the highest evidence (Friston et al., 2003; Penny et al., 2004).

Model evidence can be converted into posterior model probabilities p(m | y), which represent the probability that a given model generated the observed data (Stephan et al., 2009).

Based on these probabilities, model selection is performed through BMS (Penny et al., 2004; Stephan et al., 2009).

BMS is a framework for model comparison. In practice, data are collected from multiple subjects; therefore, it requires group-level inference. This introduces assumptions about how evidence is combined across subjects, leading to Fixed Effects (FFX) and Random Effects (RFX).

FFX assumes that the same model generates the data for all subjects, with no between-subject variability in model structure (Penny et al., 2004).

Mathematically, FFX is described by combining evidence across subjects:

$$\log p(y \mid m) = \Sigma \log p(y_i \mid m)$$

(4.7)

where the evidence is summed across subjects.

This implies that a single model must explain the data from all subjects simultaneously, which constitutes a strong constraint. Consequently, FFX inference is highly sensitive to outliers, as a single subject with low model evidence can disproportionately penalize the model at the group level. For this reason, FFX may be overly restrictive for neuroimaging data, as it assumes an identical network architecture across all subjects (Stephan et al., 2009).

RFX, in contrast, assumes that different subjects may use different models, and therefore the model is treated as a random variable (Stephan et al., 2009).

This approach introduces a hierarchical structure with two levels: at the subject level, each subject has its own model evidence; at the group level, a distribution over models is estimated.

In practice, RFX estimates the probability of each model across the population, often referred to as model frequency.

The main metrics used to characterize this distribution are the posterior model probability and the exceedance probability (denoted as xp)[, which represents the probability that a given model is more likely than any other model in the model space (Stephan et al., 2009).

Based on these metrics, it is possible to identify the most prevalent model in the population.

The main advantage of RFX is that it is robust to inter-subject variability and is not dominated by a single subject. For this reason, RFX is widely used in fMRI studies, particularly in DCM, where brain networks are expected to vary across individuals (Stephan et al., 2009).

4.5 BAYESIAN MODEL COMPARISON

As discussed in the previous section, each model has its own evidence value, and BMS compares models based on this evidence. However, one model may have the highest evidence. The problem is that BMS selects a model, but this selection is not absolute, as other models may still be plausible (Penny et al., 2004; Stephan et al., 2009).

In fact, the data may favour one model, but not eliminate all alternatives, so choosing only one model discards uncertainty across models (Stephan et al., 2009). Consequently, inference may become overconfident, particularly when models have similar evidence (Stephan et al., 2009).

For each model, parameters are estimated by Bayesian inversion. This estimation combines prior assumptions and observed data, resulting in a posterior estimate of the parameters. This corresponds to Maximum A Posteriori (MAP) estimation (Friston et al., 2003; Penny et al., 2004).

The first solution is Bayesian Model Averaging (BMA). BMA does not retain only one model; instead, it considers all plausible models in the model space. Practically, each model has a posterior probability (Pp) and associated parameter estimates. Models with higher Pp contribute more to the final estimate. Therefore, parameter estimates are combined across models using these weights, resulting in weighted parameter estimates and posterior expectations (Ep) (Penny et al., 2004; Friston et al., 2003).

BMA incorporates model uncertainty directly into parameter inference, providing robust estimates and avoiding over-reliance on a single winning model (Penny et al., 2004).

Another possible solution is Bayesian Model Reduction (BMR). The full model may be complex due to the inclusion of all possible connections, and some parameters may be unnecessary. These unnecessary parameters increase model complexity and may lead to overfitting (Friston et al., 2016).

BMR starts from the full model and systematically removes parameters to generate reduced models. These reduced models are nested within the full model, where some

connections are switched off (i.e. set to zero), resulting in simpler models (Friston et al., 2016).

One important advantage of BMR is that reduced models do not need to be estimated from scratch, as BMR uses the inversion of the full model. This makes the procedure computationally efficient (Friston et al., 2016). The reduced models are then compared using F, allowing identification of models that achieve a better balance between accuracy and complexity (Friston et al., 2016).

Although BMA and BMR may appear unrelated, there is a clear relationship between them. BMR defines or refines the model space, while BMA averages parameter estimates across models. Therefore, BMR and BMA are complementary approaches for addressing model uncertainty (Friston et al., 2016; Penny et al., 2004)

4.6 PARAMETRIC EMPIRICAL BAYES

Since the parameters differ across subjects, simple averaging is not sufficient, as it ignores variability and uncertainty. Therefore, a structured model is required (Friston et al., 2016; Zeidman et al., 2019).

Parametric Empirical Bayes (PEB) is a hierarchical Bayesian model that operates on DCM parameters and aims to explain variability across subjects (Friston et al., 2016; Friston et al., 2015). It begins from estimated parameters, which are then explained by a group model and can be expressed as:

$$\theta_i = X\beta + \varepsilon_i \tag{4.8}$$

where X encodes the group design, β represents group-level effects, and ε_i represents between-subject variability (Friston et al., 2015; Zeidman et al., 2019).

In this framework, priors are not fixed but are estimated from the data. These are known as empirical priors, meaning that group data inform the priors and improve the estimation of group effects (Friston et al., 2015; Friston et al., 2016).

The estimated parameters of PEB are Ep, representing group-level parameters, Cp representing the uncertainty of parameters, and Pp representing the probability that a parameter is different from zero (Zeidman et al., 2019). Therefore, Ep reflects the strength of connections, while Pp provides evidence for their presence.

PEB starts from a full model in which all candidate connections between regions are included and all parameters are active (Friston et al., 2016).

4.7 AIM OF STUDY

The problem faced in this study is that effective connectivity is not directly observable. Another issue is that multiple models can explain the same data. Therefore, there is a need to compare models to determine which model best explains the data.

The task of this study is visuomotor, and the network is a cortico-cerebellar loop. The regions of this network are: Visual Cortex (V1), Primary Motor Cortex (M1), Supplementary Motor Area/Premotor Cortex (SMAPMC), Cingulate Cortex (CC), Superior Parietal Lobule (SPL) and Cerebellum (CRBL) (Lorenzi et al., 2025).

These regions subserve complementary functions within the network: V1 is involved in visual processing and receives visual input (Hubel & Wiesel, 1968; Wandell, 1995), SPL supports visuomotor integration and spatial transformations (Andersen & Buneo, 2002), SMAPMC is associated with motor planning and action representation (Rizzolatti et al., 1996), M1 is responsible for motor execution and force production (Porter & Lemon, 1993), and the cerebellum contributes to coordination, prediction, and error correction (Ito, 2008).

In this study there are two conditions: Action Observation (AO) and Action Observation with visual cue (AO-BAR), where a visual cue is present compared to AO where it is absent. Therefore, the study compares the presence versus absence of visual input modulation.

V1 is fundamental because it represents the visual driving input of the network. It activates the network and is influenced by the task condition.

The aim of this study is to identify the most plausible architecture, to test the presence or absence of connections and to understand the propagation of network activity.

The method used is DCM, which models effective connectivity. The model space is composed of a full model and reduced models corresponding to biological hypotheses. The models are then compared using BMS, and PEB is subsequently applied.

From a scientific point of view, the aim is not to determine the model winner but to infer the causal organisation of the network, to understand visuo-motor dynamics and to interpret the interactions between regions. Therefore, starting from modelling, the comparison of models allows inference, and the results can be interpreted in a neuroscientific context.

5. MATERIALS AND METHODS

5.1 PARTICIPANTS

The subjects considered in this analysis were healthy volunteers with no history of neurological or psychiatric disorders, and with normal or corrected-to-normal vision. All participants were right-handed, and handedness was assessed using the Edinburgh Handedness Inventory (Alahmadi et al. 2016).

The data used in this study are based on previously published works and include Action Execution (AE), AO and AO-BAR (Alahmadi et al., 2016; Casiraghi et al., 2019; Lorenzi et al., 2025). Although AE exists, in this analysis, only AO and AO-BAR were used

The AO dataset included 9 healthy volunteers (age: 26 ± 8 years), of whom 3 were male. For the AO-BAR dataset, 12 subjects were initially recruited (age: 26 ± 3.5 years), of whom 7 were female

5.1.1 MRI ACQUISITION PROTOCOL

MRI data were acquired using a 3 Tesla (T) Philips Achieva scanner (Philips Healthcare, Best, The Netherlands) equipped with a 32-channel head coil. Functional images were acquired using a T2*-weighted GRE-EPI sequence to measure the BOLD signal. EPI enables rapid acquisition, while T2* weighting provides sensitivity to haemodynamic changes (McRobbie et al., 2017; Huettel et al., 2004).

The acquisition parameters were repetition time (TR) = 2500 ms (temporal resolution), echo time (TE) = 35 ms (T2* sensitivity), flip angle = 90° (signal excitation), voxel size = 3 × 3 × 3 mm^3 (spatial resolution), and 200 volumes per run (length of the time series). Three functional runs were acquired. Whole-brain coverage was obtained, including both cortical and cerebellar regions.

Structural images were acquired using a 3D T1-weighted Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence with the following parameters: TR = 6.9 ms, TE = 3.1 ms, inversion time (TI) = 824 ms, flip angle = 8°, and voxel size = 1 × 1 × 1 mm^3.

Data were acquired using an event-related design. In this design, stimuli are presented as discrete trials, each consisting of a task period followed by a rest period. This temporal structure allows the separation of individual events and enables modelling of the BOLD response using the GLM (Friston et al., 1994; Poldrack et al., 2011).

5.1.2 EXPERIMENTAL PARADIGM

The experimental paradigm consists of a visuomotor task based on a squeeze-ball (power grip) paradigm performed with the right hand. This controlled motor task is designed to investigate visuomotor processing by engaging both action execution and action observation mechanisms (Casiraghi et al., 2019; Alahmadi et al., 2016). The paradigm includes three conditions: AE, AO and AO-BAR, which share the same task structure but differ in the level of engagement.

Force production is controlled parametrically and defined relative to the Maximum Voluntary Contraction (MVC), which is measured prior to scanning to normalize force across subjects. Five force levels are used (20%, 30%, 40%, 50%, and 60% MVC), allowing comparison across participants and introducing a parametric manipulation of motor output.

The experiment follows an event-related design. In this framework, stimuli are presented as discrete trials, each consisting of a task period (3 s) followed by a rest interval ranging between 2 and 12 s. A total of 75 trials are performed, balanced across force levels and experimental conditions, and presented in a randomized order. This design allows the separation of individual events and enables estimation of the BOLD response at the single-trial level (Friston et al., 1994; Poldrack et al., 2011).

Subjects receive visual feedback during the task, consisting of a display showing both the target force level and the actual force produced, typically represented by a coloured bar (green/red). This continuous feedback establishes a coupling between visual and motor systems, which is critical for visuomotor integration.

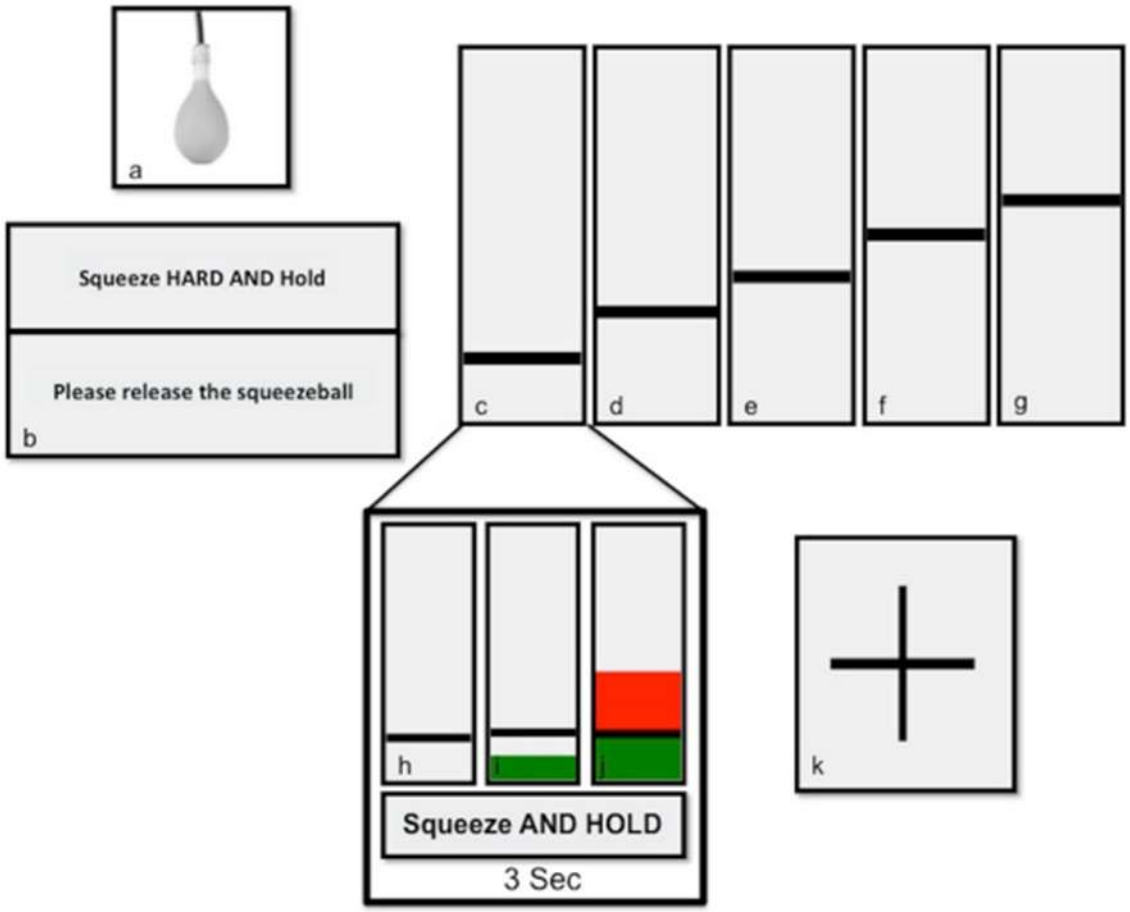

Figure 5.1: Squeeze-ball device and visual feedback system. The squeeze-ball is MRI-compatible, ensuring safe use within the scanner environment. The visual display shows both the target force level and the actual force produced, represented by a bar. The colour green indicates correct performance, whereas red indicates a mismatch between target and applied force (Alahmadi et al., 2016).

Participants undergo a training phase prior to scanning, including observation, practice outside the scanner, and practice inside the scanner. This procedure ensures familiarity with the task and reduces inter-subject variability.

The three experimental conditions differ as follows. In the AE condition, subjects actively perform the squeeze-ball task, producing force according to the predefined %MVC levels. In the AO condition, subjects observe an actor performing the same task without receiving explicit force information, thereby isolating the observation component. In the AOBAR condition, subjects observe the task while receiving an additional visual cue indicating the force level, introducing explicit visual guidance.

This progression from AE to AO to AOBAR allows the investigation of the role of visual information in modulating visuomotor processing.

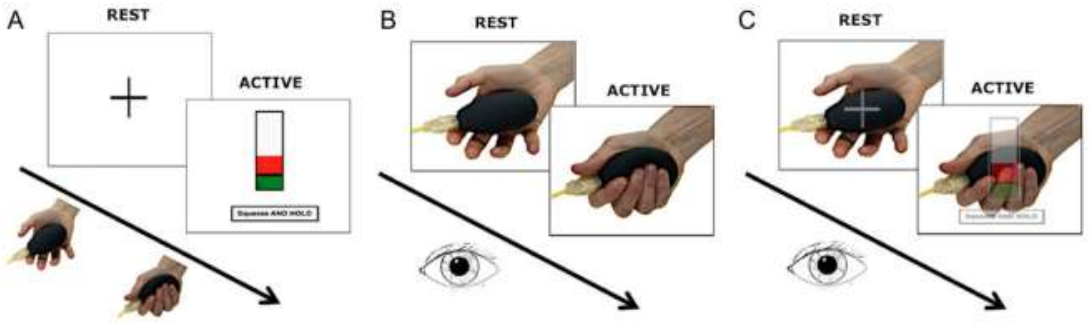

Figure 5.2; Experimental paradigm. Panel A corresponds to the AE condition, where the subject performs the squeeze-ball task with visual feedback. Panel B corresponds to the AO condition, in which the subject observes the action without performing any movement. Panel C corresponds to the AOBAR condition, where the subject observes the task with an additional visual cue. (Casiraghi et al., 2019)

A parametric modulator modulates the amplitude of the BOLD signal, which depends on a continuous variable. In this case, this variable is grip force (GF).

From a neurophysiological perspective, this task engages a distributed visuomotor network, including visual, motor, and associative regions involved in sensorimotor integration (Casiraghi et al., 2019; Lorenzi et al., 2025).

5.2 PREPROCESSING ANALYSIS

The analysis is restricted to the AO and AOBAR datasets, as it focuses on action observation and the effect of visual cues on visuomotor processing.

The goal of preprocessing is to prepare anatomical and functional data by removing noise and artefacts and ensuring spatial and temporal alignment. The data for both datasets consist of anatomical images (T1-weighted) and functional images (EPI).

The goal of anatomical preprocessing is to prepare the structural image for alignment. Intensity normalization was applied to correct for intensity inhomogeneity, ensuring comparable signal across voxels. Subsequently, brain extraction was performed to

remove non-brain tissues (e.g., skull and scalp), resulting in a brain-only anatomical image (Poldrack et al., 2011).

The goal of functional preprocessing is to clean the fMRI signal and correct artefacts. First, denoising using MP-PCA was applied to remove noise based on redundancy in the data (Veraart et al., 2016).

Since slices are acquired at different time points, slice timing correction was applied to temporally align all slices based on the TR and acquisition order, resulting in a time-corrected 4D image (Poldrack et al., 2011).

Head motion across volumes was then corrected by realigning all volumes to the mean image, resulting in motion-corrected images (Poldrack et al., 2011).

Functional and anatomical images differ in space; therefore, co-registration was performed to align functional images to the anatomical space. Subsequently, normalization to a standard anatomical space, specifically the Montreal Neurological Institute (MNI) template, was applied, allowing group-level comparisons (Ashburner & Friston, 1999; Mazziotta et al., 2001).

Cerebrospinal fluid (CSF) represents a non-neuronal signal; therefore, a mask was applied to remove these components and focus the analysis on brain tissue (Poldrack et al., 2011).

To remove confounding signals (nuisance regressors), including motion, physiological noise, and low-frequency drift, a GLM-based regression approach was used to model and remove unwanted components (Friston et al., 1994; Poldrack et al., 2011).

Finally, data quality was assessed using temporal signal-to-noise ratio (tSNR) and signal-to-fluctuation-noise ratio (SFNR). tSNR reflects temporal stability, while SFNR quantifies the relationship between signal and noise, where higher values indicate better data quality (Welvaert & Rosseel, 2013).

The output of preprocessing is cleaned and aligned fMRI data, ready for subsequent statistical analysis.

5.3 VOI EXTRACTION FOR DCM ANALYSIS

The goal of VOI (Volume of Interest) extraction is to extract subject-specific time series and to define regions for DCM analysis.

This step is fundamental because DCM does not use whole-brain data but focuses on predefined regions (nodes). For each region, an fMRI time series is required, representing the temporal evolution of the BOLD signal. This signal is summarized using the first eigenvariate, which reduces the contribution of multiple voxels into a single representative signal.

To extract VOIs, a GLM is first estimated at the subject level, as it models the BOLD signal in relation to the task (Friston et al., 1994). The outputs include the SPM.mat file, beta images, and residual maps.

A contrast is then defined; in this analysis, the contrast “action vs rest” is used to identify task-related activation, highlighting voxels that are more active during the task.

Subsequently, a group-level analysis is performed using a one-sample t-test to identify robust activation across subjects. The output of this step is a group statistical map, which serves as the basis for VOI localization.

Once the group peaks are identified, these are used as coordinates for VOIs.

A functional constraint is applied to determine where activation is present (Lorenzi et al., 2025).

Then, an anatomical constraint is applied. To achieve this, atlas-based masks are used. In particular, for cortical regions, BAM (Brainnetome Atlas Mask) masks are applied, while for the cerebellum the SUIT (Spatially Unbiased Infratentorial Template) atlas is used. This ensures anatomically informed and more precise localization of the regions (Lorenzi et al., 2025).

Finally, the geometrical constraint defines the shape and dimensions of the VOIs, and a statistical threshold is applied (Lorenzi et al., 2025).

Region	Center	Geometry	Dimension	Threshold	Mask	Voxels
V1_15115_RL	5.3 92 -1.4	box	90 x 90 x 90	0.3	BAM	246
SPL_138_L	-18.3 -72.8 37.5	sphere	15 mm	0.3	BAM	183
CC_117_L	-2.1 -2.6 30.8	sphere	15 mm	0.3	BAM_	48
M1_131_L	-39 -8.9 53.4	sphere	15 mm	0.3	BAM	345
SMAPMC_122_L	-53.9 6.7 15	sphere	15 mm	0.3	BAM	314
CRBL_R	27.4 63.1 23	sphere	15 mm	0.3	SUIT	345

The results of the group-level analysis are summarized in Tables 5.1 and 5.2.

Table 5.1: Group-level peak coordinates and VOI definitions for the AO-BAR dataset.

Region	Center	Geometry	Dimension	Threshold	Mask	Voxels
V1_15115_RL	$\begin{matrix}9.1\\-91\\0.4\end{matrix}$	box	90 x 90 x 90	0.3	BAM	241
SPL_138_L	$\begin{matrix}-32.6\\-46.8\\59.2\end{matrix}$	sphere	15 mm	0.3	BAM	711
CC_117_L	$\begin{matrix}-10.8\\22.8\\28.6\end{matrix}$	sphere	15 mm	0.3	BAM	89
M1_131_L	$\begin{matrix}-33.7\\-9.4\\57.7\end{matrix}$	sphere	15 mm	0.3	BAM	376
SMAPMC_122_L	$\begin{matrix}-54.8\\7.3\\23.9\end{matrix}$	sphere	15 mm	0.3	BAM	376
CRBL_R	$\begin{matrix}13.3\\-65.4\\-21.8\end{matrix}$	sphere	15 mm	0.3	SUIT	612

Table 5.2: Group-level peak coordinates and VOI definitions for the AO dataset.

Using these group-defined regions, subject-specific VOIs are then extracted, ensuring that the same regions are used across subjects while preserving individual variability.

In this analysis, not all subjects were retained. Selection was based on the presence of activation in all predefined regions. Therefore, for the AO dataset, six subjects were included, while for the AO-BAR dataset only two subjects met this criterion.

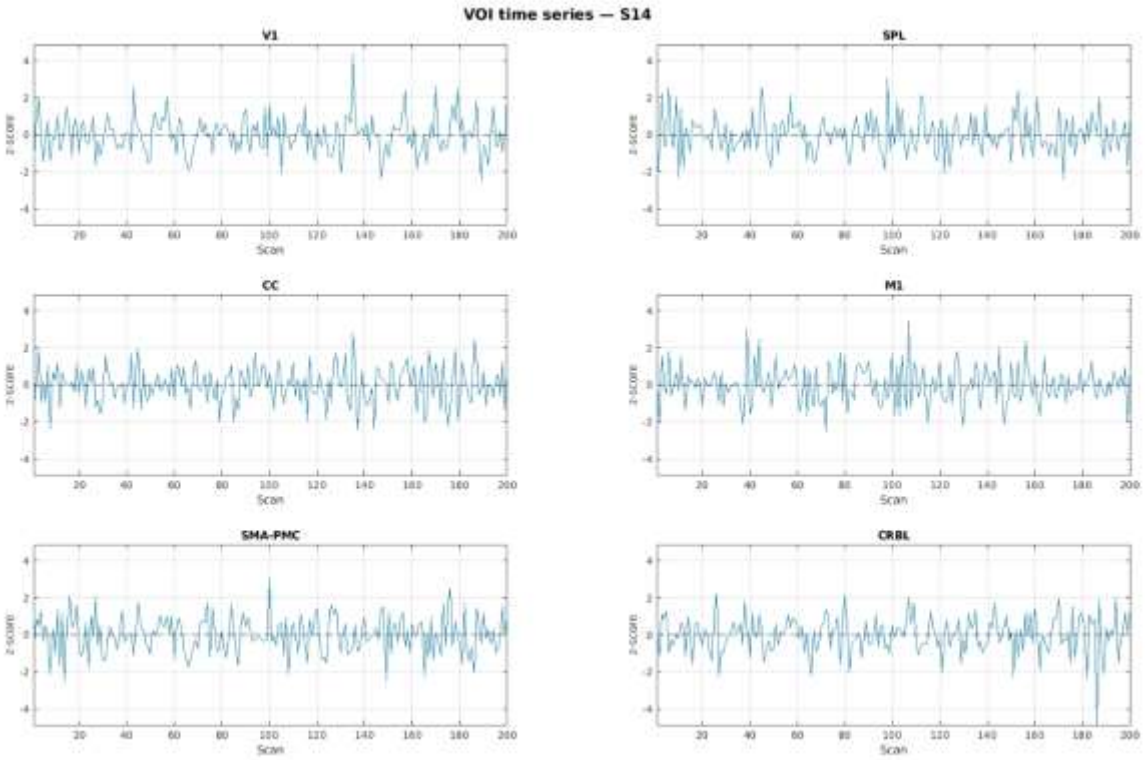

Figure 5.3: Example of VOI extraction in AO. This figure shows the time series extracted for a representative subject across all regions in the network. Each subplot corresponds to a VOI. The signals represent the temporal evolution of the BOLD signal and show oscillations over time, reflecting indirect neural activity. Each region exhibits its own dynamics, resulting in differences across regions. These differences form the basis for connectivity analysis. The extracted time series are used as inputs for DCM to estimate interactions between regions.

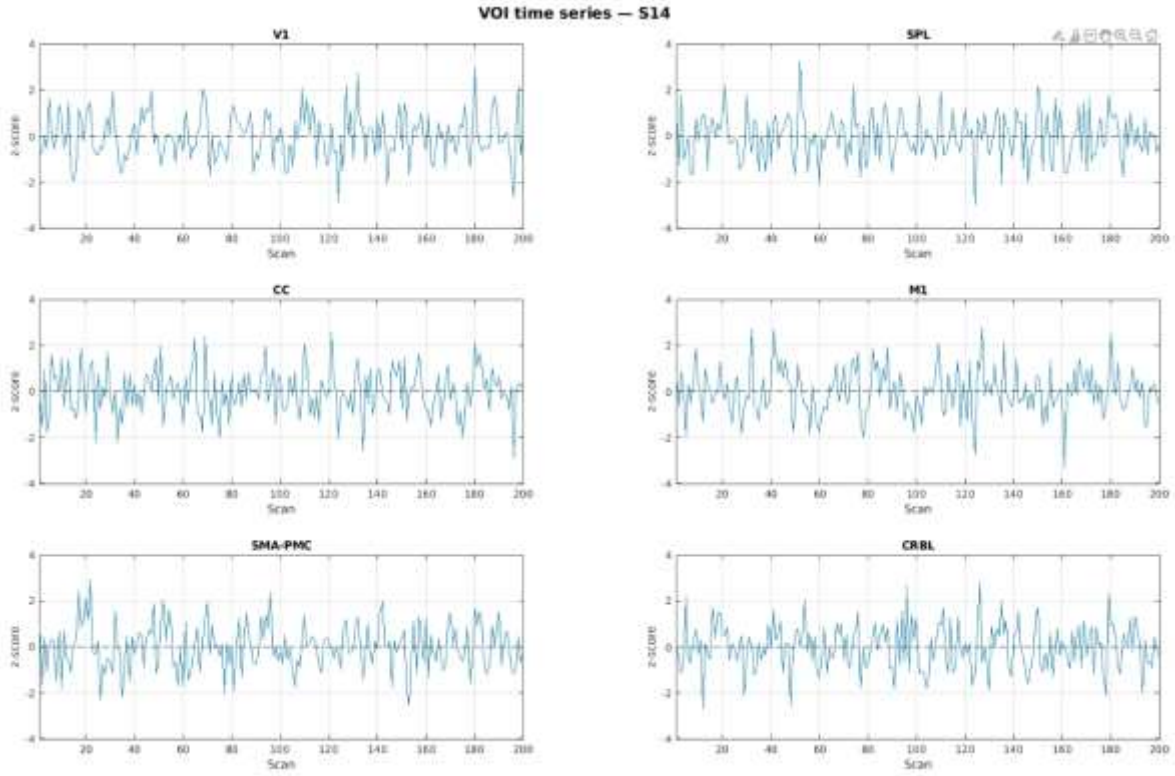

Figure 5.4: Example of VOI extraction in AO-BAR. This figure shows the time series extracted for a representative subject across all regions in the network. Each subplot corresponds to a VOI. The signals represent the temporal evolution of the BOLD signal and show oscillations over time, reflecting indirect neural activity. Each region exhibits its own dynamics, resulting in differences across regions. These differences form the basis for connectivity analysis. The extracted time series are used as inputs for DCM to estimate interactions between regions.

5.4 EFFECTIVE CONNECTIVITY ESTIMATION USING DCM

The goal of this analysis is to estimate effective connectivity, meaning directed interactions and causal influences between brain regions. DCM takes as input the VOI time series, where one time series is extracted per region, as described in the previous section. These signals represent the observed BOLD responses used by the model.

Only subjects with all VOIs available were included in the analysis in order to ensure that all nodes of the network were present. This is necessary because DCM requires a complete set of regions; missing regions would result in an incomplete model specification.

For model specification, the regions correspond to those described in Section 5.1.2 and define the nodes of the network. These anatomical regions are translated into a model structure representing the interactions between brain areas.

Regarding connectivity matrices, six models were defined, corresponding to different hypotheses on network architecture. The A-matrix (intrinsic connectivity) included one fully connected model, in which all regions are interconnected, and five reduced models (namely S1.1–S1.5) obtained by switching off specific connections based on prior hypotheses and literature (Lorenzi et al., 2025).

The full model (S1.0) represents the most complex architecture, as all possible interactions between regions are allowed. In contrast, the reduced models impose constraints on connectivity, enabling the testing of specific hypotheses about the network organization. The model structure is illustrated in Figure 5.5.

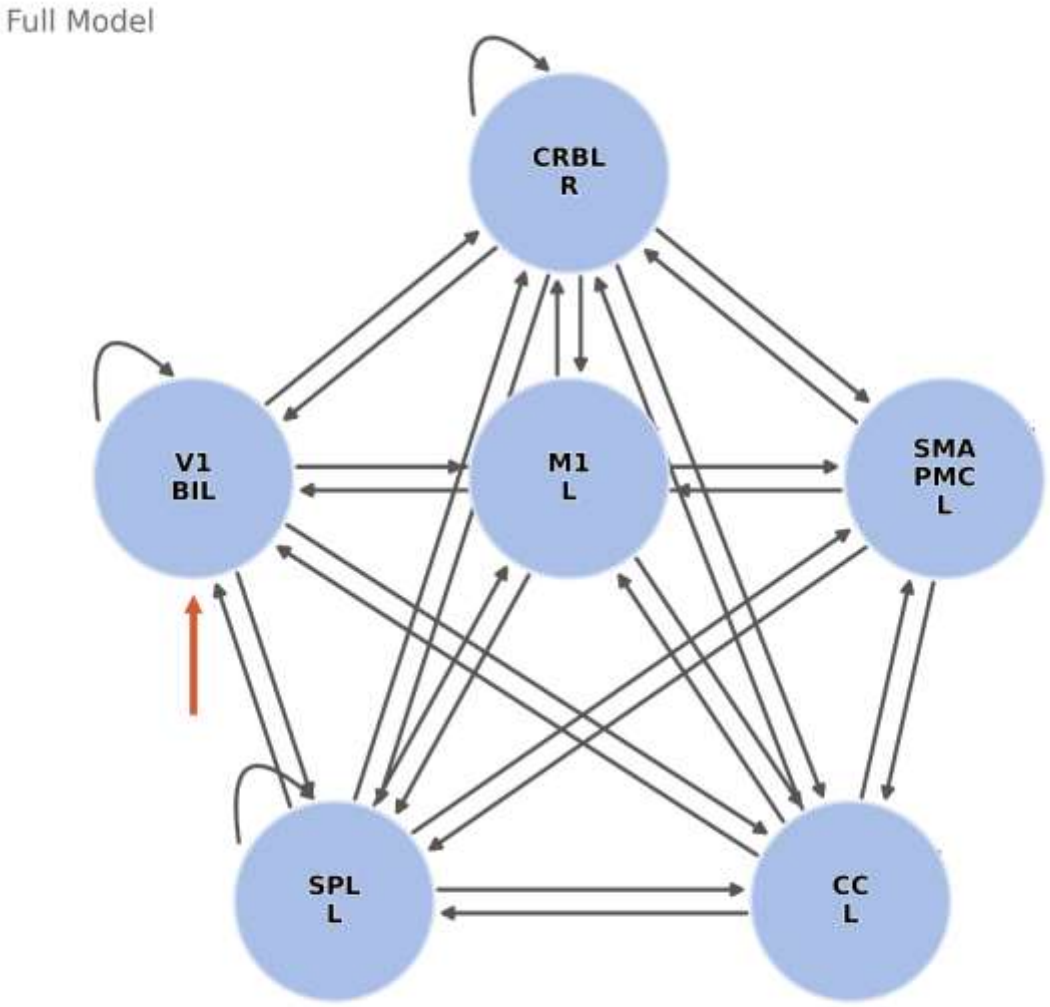

Figure 5.5: Schematic illustration of the full effective connectivity model of cortico–cerebellar loops. The model represents the full architecture obtained by connecting all regions.

Model 1 (S1.1) is a restricted model characterized by mainly bidirectional interactions. The network is distributed and includes interactions between visual and cerebellar regions, with self-connections in key regions such as V1. The dominant flow is mixed, including both feedforward and feedback pathways. This model represents a simplified architecture in which visuomotor, and cerebellar interactions are preserved while less relevant connections are excluded.

Model 2 (S1.2) is a restricted model characterized by mainly unidirectional interactions. The network is distributed and includes cerebellar and motor regions. The dominant flow is primarily feedforward, with limited feedback interactions. This model tests how visual input from V1 propagates towards cerebellar and motor regions.

Model 3 (S1.3) is a restricted model with mainly bidirectional interactions, together with some unidirectional pathways. The network is distributed and includes self-connections in key regions such as V1, M1, and CRBL. The dominant flow is mixed. This model tests how V1 influences the motor system indirectly, without a direct connection between V1 and the cerebellum.

Model 4 (S1.4) is a more complex restricted model with predominantly bidirectional interactions. The network is distributed and includes direct connections between V1 and multiple regions, including CRBL, M1, and SMA-PMC. The dominant flow is mixed. This model tests a more complex architecture in which visual input directly influences multiple components of the visuomotor network.

Model 5 (S1.5) is a restricted model with both bidirectional and unidirectional interactions. The network is distributed, with self-connections in V1, M1, and CRBL. The dominant flow is mixed. This model tests a configuration in which V1 primarily interacts with the motor system, with reduced involvement of cerebellar pathways.

The reduced models are illustrated in Figure 5.6.

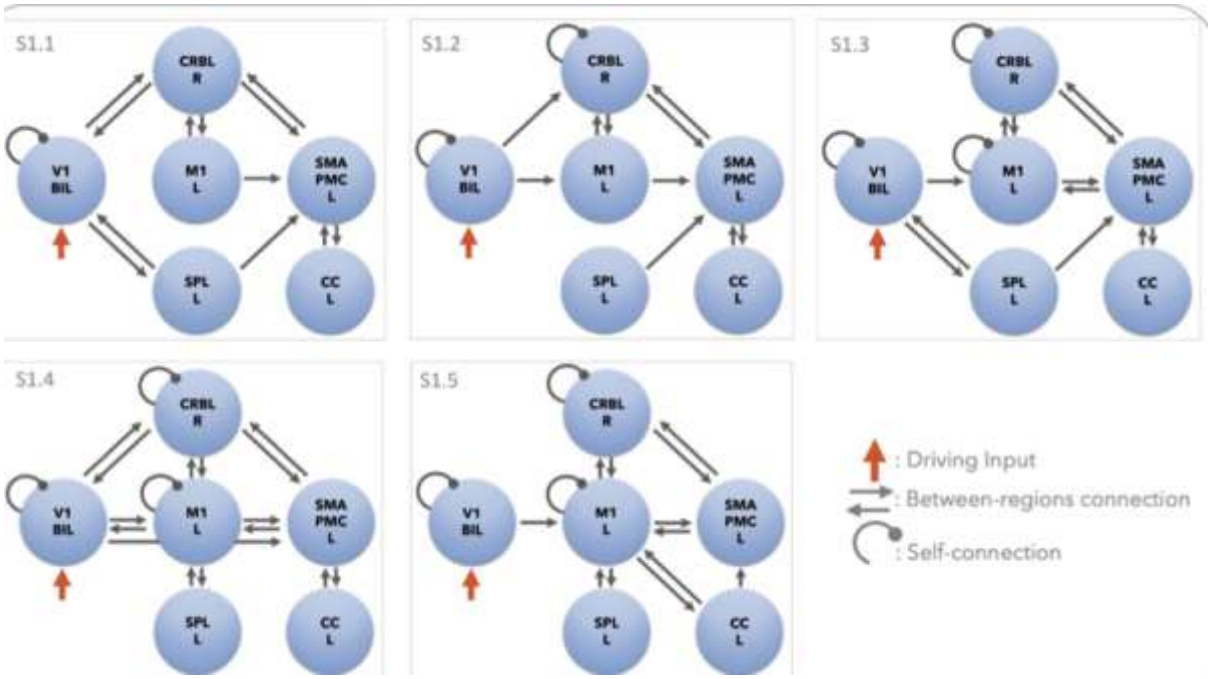

Figure 5.6: Schematic illustration of the reduced effective connectivity models of cortico–cerebellar loops. Each model corresponds to a specific hypothesis. Models S1.1–S1.5 represent reduced architectures obtained by removing specific connections. The models differ in the presence or absence of connections between visual, motor, and cerebellar regions. Model 4 corresponds to S1.4. (Lorenzi et al., 2025)

The B-matrix (modulatory effects) was not included and was therefore set to zero, as no task-dependent modulation was tested. The C-matrix (driving input) was specified such that the external input entered only the V1 region, representing visual input and indicating that the stimulus enters the network through the visual cortex (Friston et al., 2003).

The model space is composed of the six models described above, allowing the comparison of alternative hypotheses on effective connectivity.

Model inversion was performed using Bayesian estimation through Variational Laplace (Friston et al., 2003; Stephan et al., 2010). The goal of this step is to estimate the model parameters that best explain the observed data. The outputs of model

inversion include the posterior Ep and the F, which approximates the model evidence and is used for subsequent model comparison.

5.5 GROUP-LEVEL PEB ANALYSIS

The goal of the group-level PEB analysis is to estimate effective connectivity at the group level, focusing on the A-matrix.

As input, a General Linear Model of connectivity (GCM) was used, containing subject-specific DCMs. The AO and AOBAR datasets were analysed separately. As described previously, only subjects with all VOIs available were included in order to ensure a complete network.

The second-level design matrix included only the group mean and was implemented as a column of ones. This allows the estimation of the average connectivity across subjects.

PEB was estimated at the group level in order to model effects on intrinsic connectivity. Following PEB estimation, BMR was applied, followed by BMA. This procedure allows the identification of reduced models and the averaging of parameters across plausible models.

6. RESULTS

6.1 MODEL SELECTION

The model space is defined by a full model and reduced models.

To perform the analysis, a subset of subjects was used. Due to the requirement that all subjects must have all VOIs available, some subjects were excluded. Therefore, for the AO dataset, 6 subjects were included, while for the AO-BAR dataset, only 6 subjects met this criterion.

However, the comparison is performed only on the reduced models. RFX-BMS was applied to five reduced models, each representing different connectivity hypotheses, and was performed based on model evidence (F). The results are expressed in terms of exceedance probability (xp), which represents the probability that a model is the most likely among the set. Figures 6.1 and 6.2 show the exceedance probabilities across the reduced models for the AO and AO-BAR datasets, respectively.

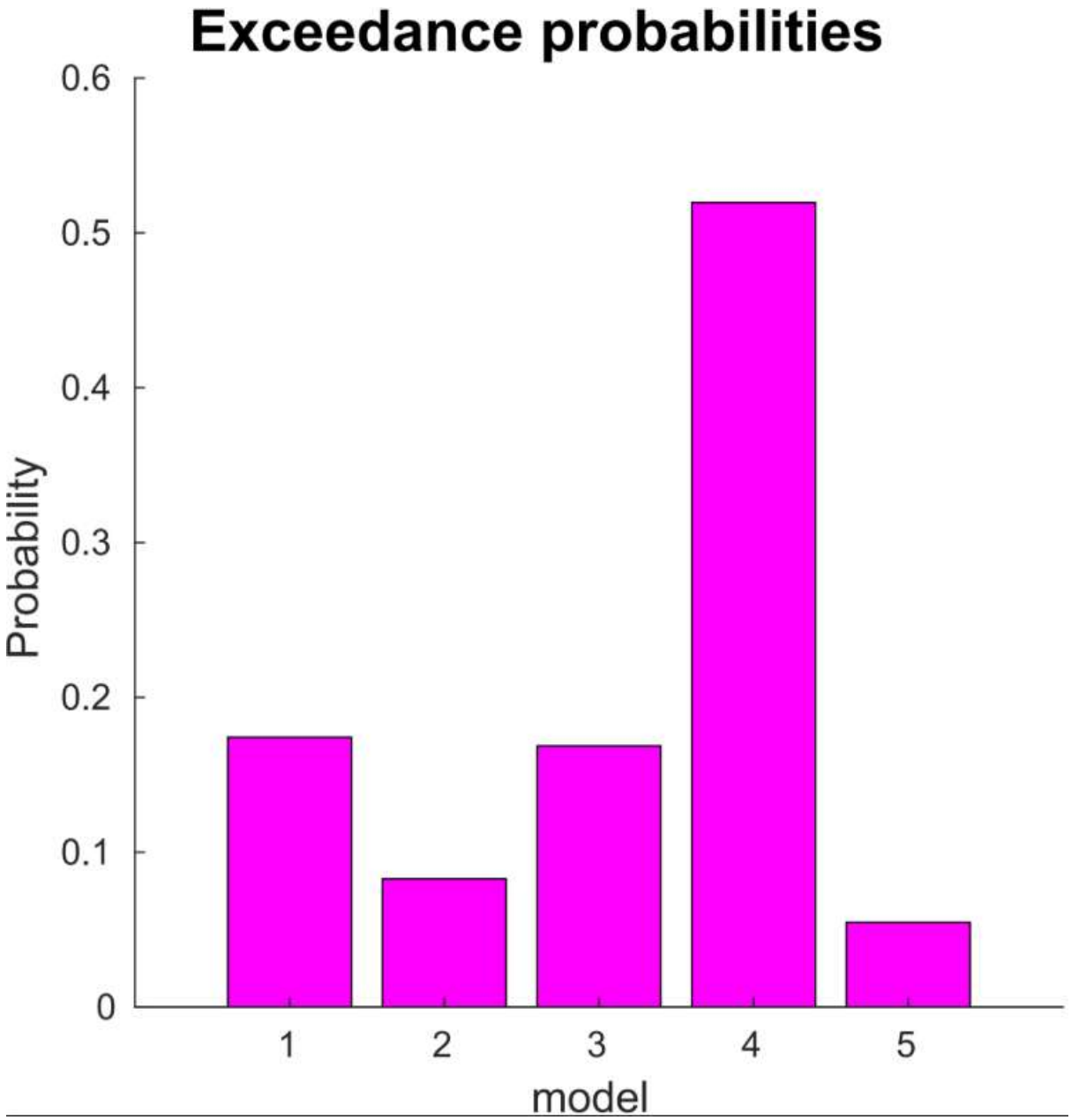

Figure 6.1: AO. The bars represent the exceedance probability of the reduced models

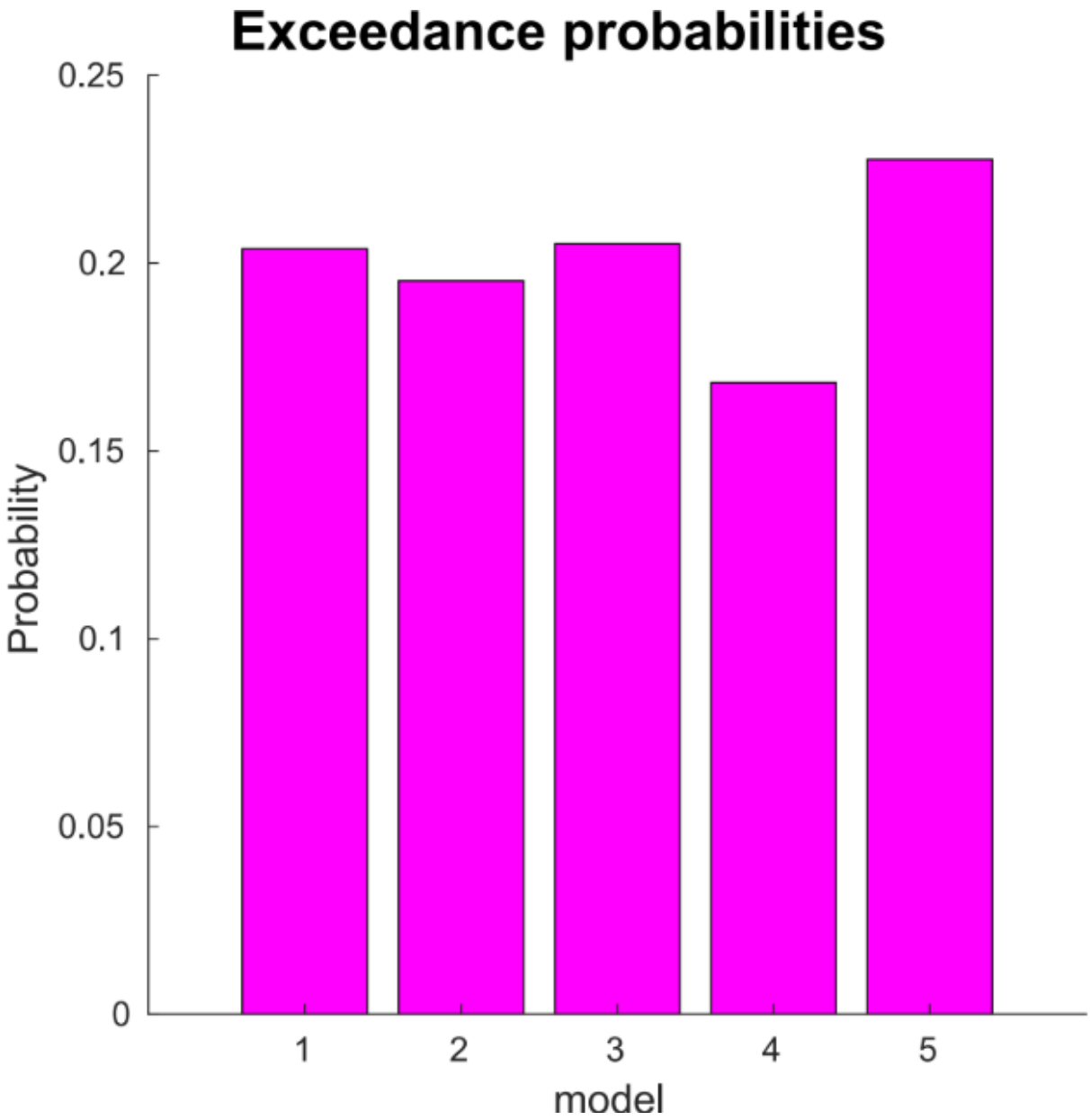

Figure 6.2: AO-BAR. The bars represent the exceedance probability of the reduced models

As shown in Figure 6.1, for the AO dataset, Model 4 has the highest exceedance probability (xp ≈ 0.5) and is therefore identified as the winning model.

As shown in Figure 6.2, for the AOBAR dataset, Model 5 has the highest exceedance probability (xp ≈ 0.2–0.25). However, the values are relatively low, indicating weak evidence for model selection and suggesting that no model clearly dominates.

6.2 GROUP-LEVEL CONNECTIVITY RESULTS

The results were visualised as heatmaps, including Ep.A and Pp.A. In addition, connection support was evaluated using posterior probability thresholds, with Pp > 0.95 indicating strong evidence for a connection. The output of the analysis consists of group-level effective connectivity estimates, posterior probabilities, and the set of supported connections. The colours in the heatmap correspond to Ep. The symbol * corresponds to supported connections (Pp > 0.95), and the order of the regions follows the ordering used in the DCM specification. This output is illustrated in Figures 6.3 and 6.4.

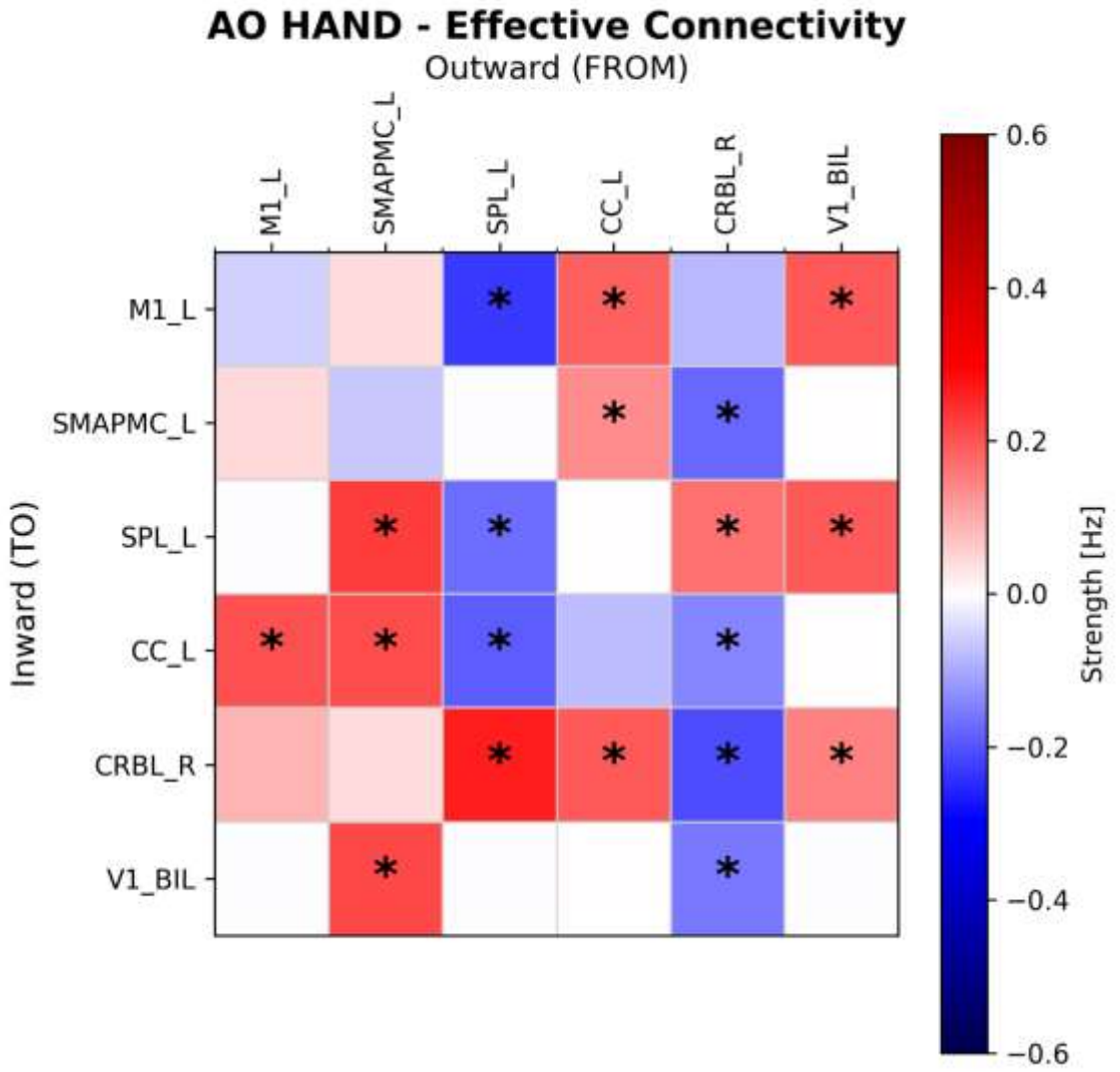

Figure 6.3: Heatmap of supported connections for AO (Pp > 0.95). Each value equal to 1 indicates a strongly supported connection, while 0 indicates no strong support

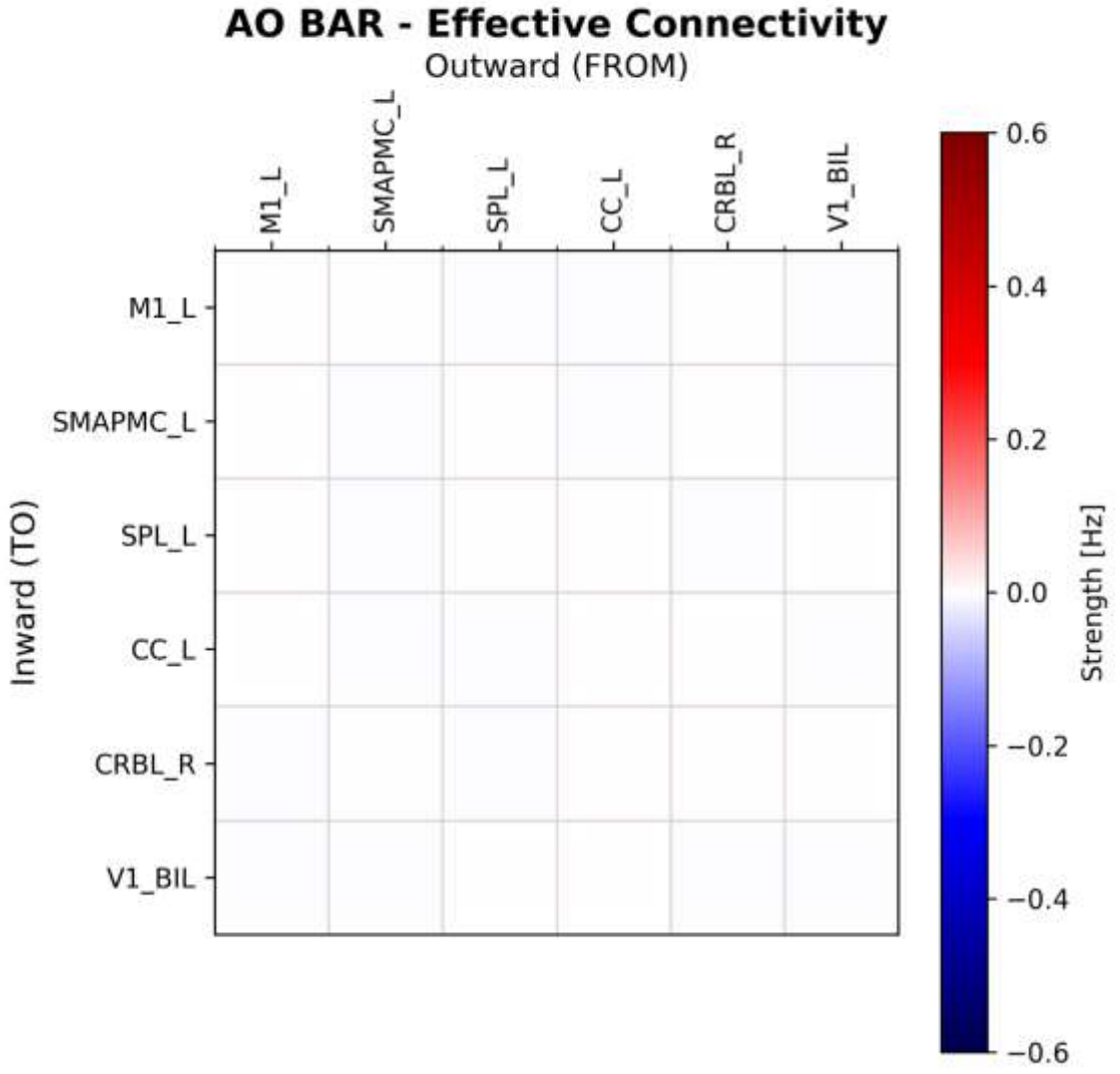

Figure 6.4: Heatmap of supported connections for AO-BAR (Pp > 0.95). No connections exceed the threshold; therefore, no strong group-level connections are identified

Figures 6.3 demonstrates that the AO condition presents supported connections, indicating a structured and distributed network. In contrast, figure 6.4 shows that the AO-BAR condition illustrates instability, with no strongly supported connections.

To better visualise the AO network, the heatmap was transformed into a network representation (Figure 6.5), where connections between regions are illustrated. The driving input was included to indicate where the network activity originates.

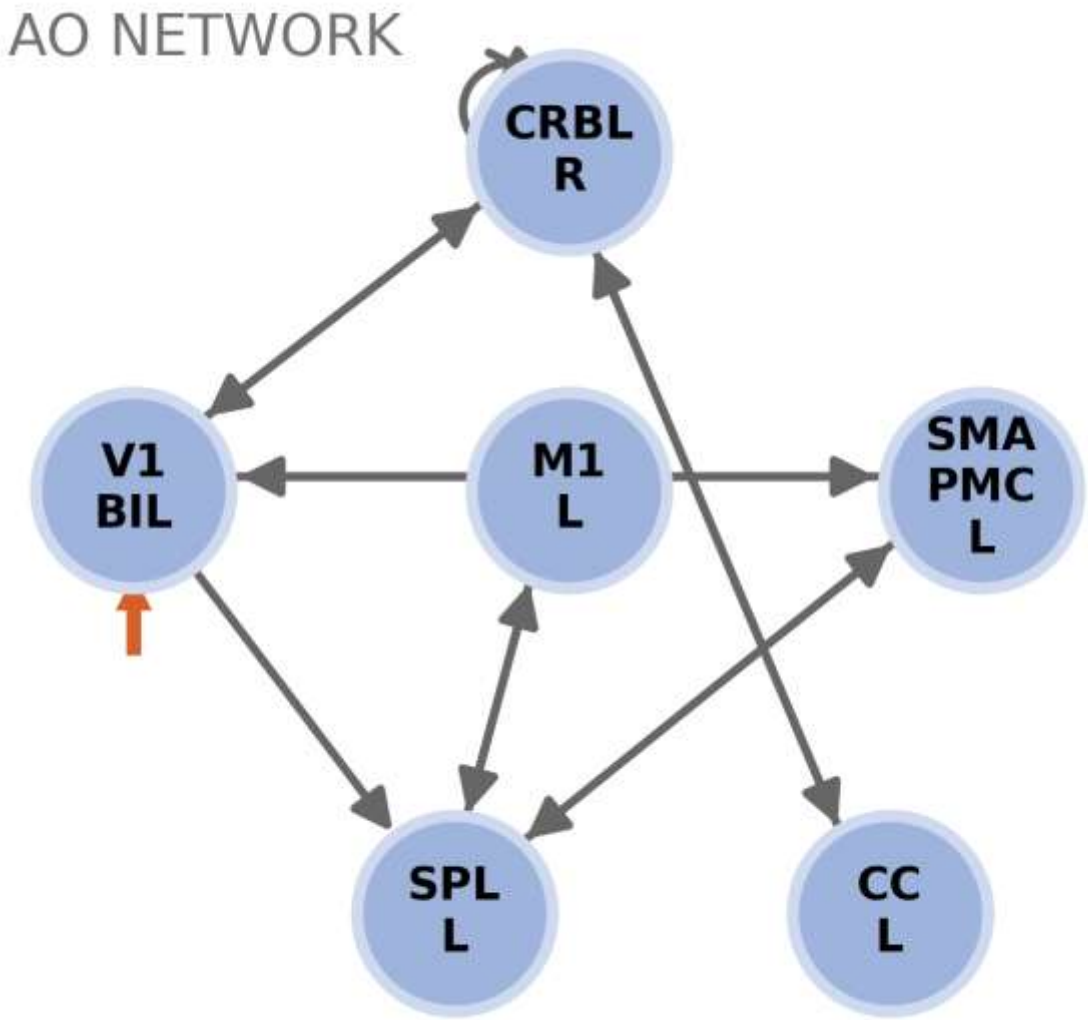

Figure 6.5: Network representation of supported effective connectivity (Pp > 0.95) for the AO condition. Arrows represent the direction of connectivity. The external input is applied to V1.

The AO network shows multiple supported connections, indicating that the AO condition is structured and distributed.

In the AO condition, there are connections from V1 to other regions, e.g. V1 → M1 and V1 → SPL. In addition, there are motor interactions between regions such as M1,

SMAPMC, SPL, and CC. Furthermore, the cerebellum (CRBL) is connected to motor regions, e.g. CRBL → SMAPMC.

Comparing this network with Model 4 (S1.4), the similarities include interactions between V1 and other regions, motor network interactions, and cerebellar involvement. However, a notable difference is the presence of additional connections, such as CRBL → motor regions and feedback connections towards V1.

Therefore, Model 4 can be considered the most plausible hypothesis describing this dataset, although it does not fully capture all the connections observed in the AO network.

In contrast, for the AO-BAR condition, no supported connections were identified due to higher variability. This is illustrated in Figure 6.6.

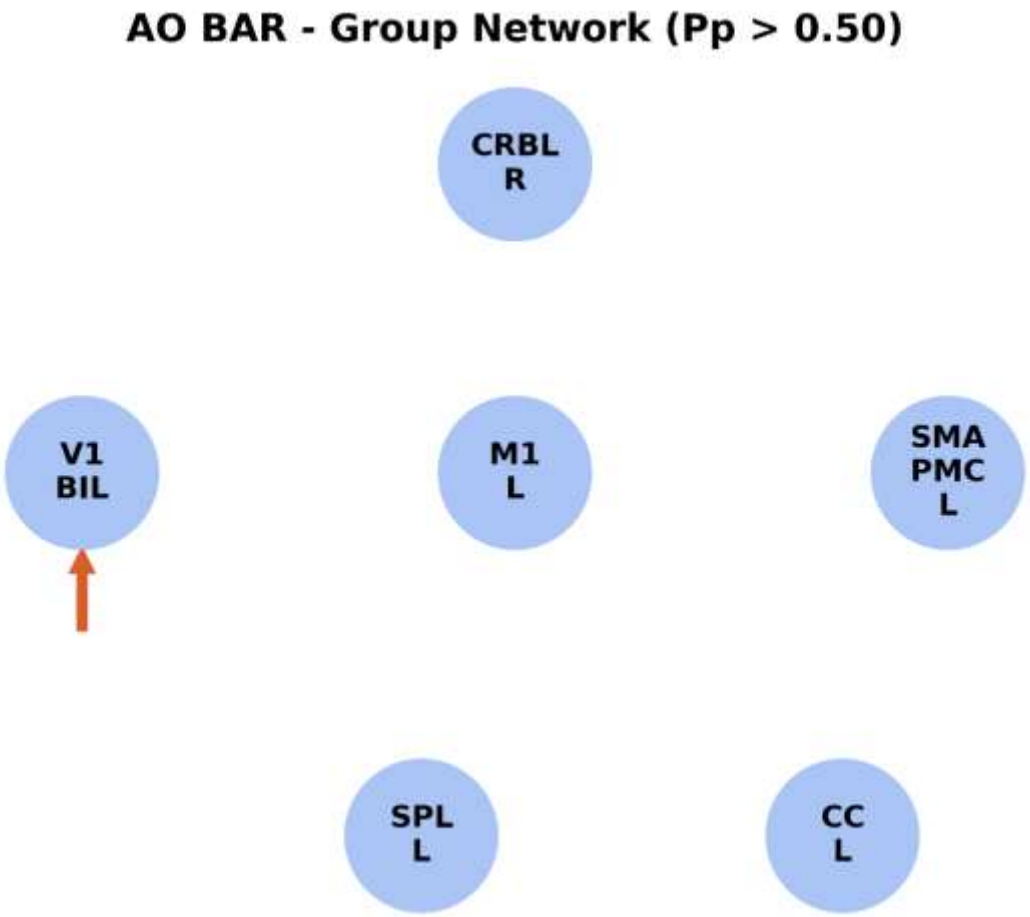

Figure 6.6: Network representation of effective connectivity for the AO-BAR condition. No connections exceed the threshold (Pp > 0.95). The external input is applied to V1.

7. DISCUSSION

The aim of this analysis is to determine effective connectivity, i.e. causal interactions during observation. The methods used to achieve this are DCM, BMS and PEB. The two datasets, AO and AO-BAR, were then compared. The network investigated is the visuomotor network within cortico–cerebellar loops.

The highlight of this analysis is Model 4. It represents a distributed network characterised by interactions between visual, motor and cerebellar regions. It is dominated by bidirectional connections, indicating a continuous exchange of information. This suggests that the model captures visuomotor integration, where visual input is processed, transformed into motor planning, and then into an internal simulation of the movement.

More specifically, Model 4 reflects a visuo–motor loop organisation, in which V1 provides the visual input that is propagated to motor regions and to the cerebellum (CRBL). The cerebellum is not only connected but plays an active role in prediction and coordination. The motor-related regions (SMAPMC, M1, SPL and CC) interact with each other to build the motor plan. This organisation is not linear but loop-based, with bidirectional interactions (e.g. V1 ↔ CRBL ↔ SMAPMC ↔ CC), meaning that information is continuously updated and integrated rather than simply transmitted (Lorenzi et al., 2025).

In this framework, the cerebellum can be interpreted as a forward controller, as it uses incoming information to predict and adjust motor processes. Therefore, it is not an accessory region but contributes to controlling the dynamics of the network. This explains why Model 4 is identified as the best model in the AO condition, as AO requires a complete integration of visual, motor and predictive processes.

In the AO network, a distributed pattern with stable connections is observed. In particular, V1 plays a key role as it represents the driving input of the network. Within the motor system, there are interactions between regions, and the cerebellum contributes to coordination and prediction. This indicates that the system is coherently engaged in action simulation.

Model 4 represents a general theoretical structure, while the AO network can be considered a filtered version of this structure, as only strongly supported connections (Pp > 0.95) are retained. Therefore, they are not identical, but the presence of overlapping connections suggests that the model is meaningful, while the absence of others reflects the refinement introduced by PEB.

Accordingly, the AO network corresponds to the winning model identified by BMS, and the connections supported by PEB provide further evidence for this structure. This shows that BMS and PEB results are coherent.

Comparing AO with AO-BAR shows that AO is more stable, with supported connections and clear visuo-motor integration, whereas AO-BAR shows high variability and no strong connections. One possible explanation is that AO is a simpler task, leading to a more direct mapping between visual input and motor representation. In contrast, in AO-BAR the presence of a visual cue increases task complexity, introduces divided attention, and may involve additional circuits, resulting in lower consistency across subjects.

Comparing these results with previous studies (Lorenzi et al., 2025; Casiraghi et al., 2019; Alahmadi et al., 2016), the AO condition is consistent with the literature, particularly regarding the organisation of cortico–cerebellar visuomotor loops and the role of the cerebellum in prediction and coordination.

The limitations of this study include the number of regions considered, which may result in an incomplete representation of the network. In addition, the sample size is limited, particularly for AO-BAR, leading to reduced statistical power and higher variability.

Therefore, Model 4 can be considered valid but incomplete, as it describes the general structure of the network, while PEB shows that not all connections are consistently active. In AO, the network is coherent, while in AO-BAR the results are weaker, suggesting that task design plays a crucial role in shaping effective connectivity.

8. CONCLUSIONS

Based on the results, it can be concluded that the AO condition shows a stable and supported network, whereas AO-BAR does not show supported connections.

Although both conditions are based on the same network, they show different behaviour. The presence of a visual cue does not improve performance but increases task complexity.

These results were obtained using DCM, which allowed the investigation of causal interactions and the comparison between conditions.

The main limitations are the sample size, particularly for AO-BAR, where only 2 subjects met the inclusion criterion. Consequently, there is high variability and no strongly supported connections are observed.

Future studies could include a larger number of subjects and the inclusion of additional visual regions.

Overall, these findings indicate that task design influences effective connectivity, and even small changes can modify the network organisation. This is important when considering future work as the condition of the task may impact on results, hence comparisons of studies across centres and using locally-implemented tasks should take this into account.

BIBLIOGRAPHY

Alahmadi, A. A. S., Pardini, M., Samson, R. S., D'Angelo, E., Friston, K. J., & Toosy, A. T. (2016). Differential involvement of cortical and cerebellar areas using dominant and nondominant hands: An fMRI study. Human Brain Mapping, 37(2), 507–520. https://doi.org/10.1002/hbm.23054

Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. Annual Review of Neuroscience, 25, 189–220. https://doi.org/10.1146/annurev.neuro.25.112701.142922

Ashburner, J., & Friston, K. J. (1999). Nonlinear spatial normalization using basis functions. Human Brain Mapping, 7(4), 254–266. https://doi.org/10.1002/(SICI)1097-0193(1999)7:4<254::AID-HBM4>3.0.CO;2-G

Basser, P. J., Mattiello, J., & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. Biophysical Journal, 66(1), 259–267. https://doi.org/10.1016/S0006-3495(94)80775-1

Bernstein, M. A., King, K. F., & Zhou, X. J. (2004). Handbook of MRI Pulse Sequences. Elsevier.

Brown, R. W., Cheng, Y. N., Haacke, E. M., Thompson, M. R., & Venkatesan, R. (2014). Magnetic Resonance Imaging: Physical Principles and Sequence Design. Wiley.

Buxton, R. B., Wong, E. C., & Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation. Magnetic Resonance in Medicine, 39(6), 855–864. https://doi.org/10.1002/mrm.1910390602

Buxton, R. B. (2004). Introduction to Functional Magnetic Resonance Imaging. Cambridge University Press.

Casiraghi, L., Alahmadi, A. A. S., Toosy, A. T., Gandini Wheeler-Kingshott, C. A. M., & D'Angelo, E. (2019). I see your effort: Force-related BOLD effects in an extended action execution–observation network involving the cerebellum. NeuroImage, 184, 63–75. https://doi.org/10.1016/j.neuroimage.2018.09.040

Coriasco, M. (2014). Elementi di risonanza magnetica: Dal protone alle principali applicazioni diagnostiche. Springer.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. Human Brain Mapping, 2(4), 189–210. https://doi.org/10.1002/hbm.460020402

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. NeuroImage, 19(4), 1273–1302. https://doi.org/10.1016/S1053-8119(03)00202-7

Friston, K. J. (2011). Functional and effective connectivity: A review. Brain Connectivity, 1(1), 13–36. https://doi.org/10.1089/brain.2011.0008

Friston, K. J., et al. (2015). A DCM for resting state fMRI. NeuroImage, 94, 396–407. https://doi.org/10.1016/j.neuroimage.2013.12.009

Friston, K. J., Litvak, V., Oswal, A., et al. (2016). Bayesian model reduction and empirical Bayes for group (DCM) studies. NeuroImage, 128, 413–431. https://doi.org/10.1016/j.neuroimage.2015.11.015

Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage, 9(4), 416–429. https://doi.org/10.1006/nimg.1998.0419

Haacke, E. M., et al. (1999). Magnetic Resonance Imaging: Physical Principles and Sequence Design. Wiley.

Huettel, S. A., Song, A. W., & McCarthy, G. (2014). Functional Magnetic Resonance Imaging. Sinauer Associates.

Iacoboni, M., et al. (1999). Cortical mechanisms of human imitation. Science, 286(5449), 2526–2528. https://doi.org/10.1126/science.286.5449.2526

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. Nature Reviews Neuroscience, 9(4), 304–313. https://doi.org/10.1038/nrn2332

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for robust linear registration. NeuroImage, 17(2), 825–841. https://doi.org/10.1006/nimg.2002.1132

Kilner, J. M., et al. (2009). Predictive coding and the mirror neuron system. Cognitive Processing, 10(3), 135–142. https://doi.org/10.1007/s10339-009-0274-7

Liang, Z. P., & Lauterbur, P. C. (2000). Principles of Magnetic Resonance Imaging. IEEE Press.

Logothetis, N. K., et al. (2001). Neurophysiological basis of the fMRI signal. Nature, 412(6843), 150–157. https://doi.org/10.1038/35084005

Lorenzi, R. M., Korkmaz, G., Alahmadi, A. A. S., Monteverdi, A., Casiraghi, L., D'Angelo, E., Palesi, F., & Gandini Wheeler-Kingshott, C. A. M. (2025). Cerebellar control over inter-regional excitatory/inhibitory dynamics discriminates execution from observation of an action. The Cerebellum, 24, 115. https://doi.org/10.1007/s12311-025-01863-6

Mazziotta, J., et al. (2001). A probabilistic atlas of the human brain. Philosophical Transactions of the Royal Society B, 356(1412), 1293–1322. https://doi.org/10.1098/rstb.2001.0915

McRobbie, D. W., et al. (2017). MRI from Picture to Proton. Cambridge University Press.

Mori, S., & van Zijl, P. C. M. (2007). Human white matter atlas. NeuroImage, 40(2), 570–582. https://doi.org/10.1016/j.neuroimage.2007.11.029

Ogawa, S., et al. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. PNAS, 87(24), 9868–9872. https://doi.org/10.1073/pnas.87.24.9868

Pauling, L., & Coryell, C. D. (1936). The magnetic properties of hemoglobin. PNAS, 22(4), 210–216. https://doi.org/10.1073/pnas.22.4.210

Penny, W. D., et al. (2004). Comparing dynamic causal models. NeuroImage, 22(3), 1157–1172. https://doi.org/10.1016/j.neuroimage.2004.03.026

Poldrack, R. A., et al. (2011). Handbook of Functional MRI Data Analysis. Cambridge University Press.

Porter, R., & Lemon, R. (1993). Corticospinal Function and Voluntary Movement. Oxford University Press.

Rizzolatti, G., et al. (1996). Premotor cortex and the recognition of motor actions. Cognitive Brain Research, 3(2), 131–141. https://doi.org/10.1016/0926-6410(95)00038-0

Rizzolatti, G., & Craighero, L. (2004). The mirror neuron system. Annual Review of Neuroscience, 27, 169–192. https://doi.org/10.1146/annurev.neuro.27.070203.144230

Smith, S. M., et al. (2011). Network modelling methods for fMRI. NeuroImage, 54(2), 875–891. https://doi.org/10.1016/j.neuroimage.2010.08.063

Sporns, O. (2013). Structure and function of complex brain networks. Dialogues in Clinical Neuroscience, 15(3), 247–262.

Stephan, K. E., et al. (2009). Bayesian model selection for group studies. NeuroImage, 46(4), 1004–1017. https://doi.org/10.1016/j.neuroimage.2009.03.025

Stephan, K. E., et al. (2010). Ten simple rules for dynamic causal modelling. NeuroImage, 49(4), 3099–3109. https://doi.org/10.1016/j.neuroimage.2009.11.015

Turner, R. (1994). Functional MRI at high field. Journal of Magnetic Resonance Imaging, 4(3), 271–279.

Veraart, J., et al. (2016). Denoising of diffusion MRI using random matrix theory. NeuroImage, 142, 394–406. https://doi.org/10.1016/j.neuroimage.2016.08.016

Wandell, B. A. (1995). Foundations of Vision. Sinauer Associates.

Welvaert, M., & Rosseel, Y. (2013). On the definition of signal-to-noise ratio. PLOS ONE, 8(11), e77089. https://doi.org/10.1371/journal.pone.0077089

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited. NeuroImage, 2(3), 173–181. https://doi.org/10.1006/nimg.1995.1014

Zeidman, P., Jafarian, A., Corbin, N., et al. (2019). A guide to group effective connectivity analysis using dynamic causal modelling for fMRI. NeuroImage, 200, 174–190. https://doi.org/10.1016/j.neuroimage.2019.06.032