# **1. Theoretical Background and Aims**

## 1.1 The benefits of combining EEG and fMRI

Neuronal activity of cognitive or affective processes can be studied from a large variety of measures, thereby revealing unique perspectives on brain activation.

Electroencephalography (EEG) mainly reflects the summation of postsynaptic potentials in pyramid cells with a similar orientation at a cortical level (Luck, 2005). Through sufficient coverage of the head surface with electrodes, synchronised activity of these cells can be recorded at a high temporal resolution. Due to this capacity of observing changes on a scale of milliseconds, EEG is often chosen as a direct link to cortical activity. However, EEG is recorded at a relatively large distance from the cells and considerable portions of the original activity spikes fall off outside a 50 µm radius (Henze, Borhegyi, & Csicsvari, 2000). In addition, shorter spike durations with high-frequency oscillations far above 200 Hz decrease the odds of spike summation. Therefore, the skull prevents higher frequency signals from affecting the EEG and the recorded signal predominantly consists of slower Local Field Potentials (LFP).

Compared to action potentials of single cells and multi-unit activity (MUA), LFP are bound to temporal and spatial summation. For this reason, EEG only represents the summation of surface potentials. Furthermore, despite advances in signal source estimation (e.g. Lei et al., 2011), its spatial resolution is severely limited. The inability to pinpoint neuronal sources and to reconstruct the original flow of current of a given potential on the head surface is referred to as the inverse problem in the EEG literature (Koles, 1998; Christophe Phillips, Rugg, & Friston, 2002). Source estimation analyses, as a tool of uncovering dipoles and brain areas most likely responsible for electric voltage fluctuation at the surface, rely on several assumptions. These include largely homogenous electric conductivity and resistance throughout brain tissues, a mathematical approximation of the orientation and fluctuation of the current as well as a fixed template for the anatomical structure underlying these approximations (Cuffin, 1998; Koles, 1998; Sanei, Chambers, Sanei, & Chambers, 2013; Xu, Xu, & He, 2004). Whereas the lack of knowledge on individual test subjects’ brains can be compensated by letting structural magnetic resonance imaging (sMRI) informing source estimations (Cristophe Phillips, Rugg, & Friston, 2002; Whittingstall, Bartels, Singh, Kwon, & Logothetis, 2010), other assumptions remain unlikely to be met.

While the electromagnetic fields measured in the EEG directly relate to neuronal activity, MRI is taking advantage of differing magnetic properties of nuclei within tissues of the human brain to produce images of different spatial resolutions. Functional magnetic resonance imaging (fMRI) for instance is based on blood oxygenation. By building a strong static magnetic field () during MRI measurements, the nuclei within atoms forming the observed tissue adapt an orientation parallel or anti-parallel to the magnetic field. Through a head coil placed above the subject, a controlled manipulation of the field via radiofrequency pulses causes the nuclei to change their orientation and emit energy. The gradients forming a new magnetic field () enable the successive acquisition of slices, which taken together shape a 3D volume of the subject. Relating the amount of emission to a certain spatial point within the field is a basic mechanism of all MRI techniques. In fMRI specifically the repeated assessment of haemoglobin in the brain and its location in the brain is used as an indicator of brain activation. The hemodynamic signal assessed by fMRI is linked to the oxygen consumption of neuron populations. Results show the flow of oxygenated blood in accordance to the metabolic demands of brain regions (Logothetis & Wandell, 2004). For this reason, the signal used in fMRI contrasts is referred to as blood oxygenation level dependent (BOLD). With the idea in mind that brain regions supporting a cognitive process, consume more resources, such as oxygen, the BOLD is thought to vary in correspondence to increased synaptic current flow. Since the BOLD signal is a correlate of neuronal activity, it is regarded as an indirect measure. Plus, it is confined to a low temporal resolution on a timescale of seconds. In return, functional BOLD contrasts offer a higher spatial resolution compared to other imaging methods, while still operating entirely non-invasively. As such, MRI is a powerful method for studying the spatial dynamics of brain activation and for gaining anatomical information without harming patients or test subjects.

Comparing the two methods, it becomes apparent that EEG and fMRI complement each other. Together they combine next to ideal temporal and spatial resolutions (Debener, Ullsperger, Siegel, & Engel, 2006). Both measures require an in depth understanding about its signals’ physiological properties in order to draw reasonable conclusions from experimental results about brain activity. Their respective limitations often decrease the conclusions’ validity. Instead of relying on a selective view with a single method, simultaneous or parallel recordings provide multifaceted insights into brain activation. In principle, a setup for parallel EEG-fMRI experiments entails that subjects are tested at least twice using one for each method separately. This type of experiment is easier to perform and ensures higher data quality. In short, due to the EEG measuring voltage changes sensitive to powerful magnetic fields as in the MRI and the MRI relying on relative homogeneity of the magnetic field across time as well as the removal of all ferromagnetic and electrically conductive materials, the two methods in their basic configurations inflict measurement artefacts on each other (Allen, Josephs, & Turner, 2000; Bénar et al., 2003; Iannotti, Pittau, Michel, Vulliemoz, & Grouiller, 2014). In spite of these, simultaneous recordings yield the greater potential. While free from artefacts, parallel recordings do not represent identical psychological processes in test subjects. No brain activation at a given time point in a given experiment can be perfectly replicated, due to the nature of individual cognitive processing. Even when performing identical experiments with the same experimental protocol, the timelines of signal changes in the two experiments eventually diverge. Thus, it is impossible to relate for instance single-trial EEG and MRI signals, since they were acquired successively instead of concurrently. Other problems, such as training effects, habituation or fatigue, further add to the limitations of parallel recordings.

Perhaps even more notable than the complementing spatial and temporal resolutions in combined EEG and fMRI is the benefit stemming from their physiological relation. Variation in LFP often bears more similarity to changes in BOLD than to recordings of single cell activity or MUA (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). At the same time, it is irrefutable that EEG and fMRI present brain activity from two very different perspectives. Considering their physiological basis, it seems plausible that modulations across experimental manipulations of BOLD and EEG activity often do not align (Im, Jung, & Fujimaki, 2005; Nunez & Silberstein, 2000). Whereas EEG signals only show the result of multiple activity summations across cortical layers, changes in the BOLD signal across time reflect fluctuations of oxygen concentrations in different brain regions.

The fact that the two signals do not align completely can be regarded as an upside and a downside to concurrent EEG and fMRI recordings. Neurovascular decoupling of neuronal activity and cerebral blood flow impedes the validation of a result found in one method but not in the other. When relating, for instance, Event-related potentials (ERP) to functional contrasts, neurovascular coupling would yield both results to be more meaningful. Plus, information from both sides aid the interpretation and integration of results into the greater theoretical background. Then again, neurovascular decoupling might provide as meaningful information as coupling (Rosa, Daunizeau, & Friston, 2010). For one, decoupling could be merely the result of failed signal detection or could be entirely unrelated to experimental conditions. However, it could also be attributed to pathological characteristics (Schridde et al., 2008).

As a result, approaches for combined EEG-fMRI recordings allow analysing shared and discrete signal variation (see Fig. 1) in these methods (Herrmann & Debener, 2007). Highlighting neurovascular coupling and decoupling promises new insights for the study of physiological foundations of EEG and fMRI as wells as experimental investigations of cognitive processes.

Event-related

Unrelated

**EEG**

**fMRI**

Fig. 1. Illustration of variance proportions in EEG and fMRI signal attributed uniquely to EEG (red) or fMRI (blue) and event-related neurovascular coupling (striped) or event-unrelated coupling (not striped shared area) from Herrmann and Debener (2007).

## 1.2 Approaches for multimodal data fusion

Despite the short time combined EEG and fMRI has been emerging as a research field, there is already a wealth of literature for statistical analyses. Recently, multimodal data fusion has received attention most of all (Multert & Lemieux, 2009).

This approach stands opposed to the isolated (i.e. ERP, frequency analyses, fMRI contrasts, connectivity analyses) and asymmetric types of data analyses, such as EEG-informed fMRI and fMRI-informed EEG. In all of these cases one of the two methods takes precedence over or excludes the other (Huster, Debener, & Eichele, 2012). Further, each of these analyses only uses part of the data. The main advantage of data fusion is that it represents a multivariate approach which takes into account almost all available information.

### 1.2.1 Joint Independent Component Analysis

A popular method for fusing different kinds of medical imaging and EEG data is the joint Independent Component Analysis (jICA) (Calhoun, Adali, & Liu, 2006; Calhoun, Liu, & Adali, 2009; Eichele et al., 2008; Kyathanahally, Franco-Watkins, Zhang, Calhoun, & Deshpande, 2016). As with ICA in single modalities, a generative model with an unknown, linear mixing process of signal components is assumed to underlie the data. The jICA also aims at identifying maximally independent components contributing to the signal by unmixing signal parameters. However, a spatiotemporal decomposition is performed on at least two different data modalities. In terms of ERP and fMRI data, spatiotemporal decomposition refers to the ERP time course and voxel intensity. Moreover, the jICA adds a strong constraint by assuming that sources associated with the two data modalities modulate the same way across subjects. Therefore, only sources with identical linear covariation are extracted from the unmixed data matrix. Correspondingly, beta weights are assigned to pairs of components from both data modalities (see Fig. 2). When extracting complementary components from ERP time courses and fMRI contrasts, each time point in the extracted ERP time course is assigned a combination of the associated fMRI voxels, adding spatial to the temporal data.



Fig. 2. Illustration of jICA with coupled feature matrix of multimodal datasets and in a shared data matrix (left) and in an umixed matrix with shared beta weights (right) from Calhoun and Adali (2009).

The likelihood function used in the jICA is similar to the usual ICA as well. The joint unmixing data matrix W of two datasets and from the same sample of test subjects N is estimated so that the likelihood L(W) is maximal. In the estimated unmixing matrix each dataset has the dimensions N and voxels () or ERP time course (). The basic estimation of L(W) is

,

where the vectors and represent observations for sample n = 1,… , N in the matrices and (Calhoun & Adali, 2009).

With this framework for multimodal data fusion Kyathanahally et al. (2016) were able to detect decision making components underlying simple to more complex Delay Discounting Tasks. Joint components identified in simpler tasks with certain rewards could predict parts of activation patterns found in more complex tasks with reward and punishment uncertainty, indicating that decision making processes might occur in an additive fashion.

Although data fusion attempts to use all the available information, the jICA framework puts limitations to this principle. First of all, the constraint of identical modulation of features restricts the amount of data going into signal components. This constraint is necessary to only focus on shared data sources. It can be relaxed by either choosing different datatypes or assuming correlated instead of identical modulation of the datatypes (i.e. parallel ICA). Moreover, while the authors advise to use the jICA as a second level analysis on single subject data (e.g. Calhoun & Adali, 2009), other approaches have gone further by performing second level data fusion on single trials (Debener et al., 2006; Huster et al., 2011; Murta, Hu, Tierney, Chaudhary, & Walker, 2016).

### 1.2.2 Hierarchical fusion framework

Also taking advantage of ICA techniques, Dong, Gong, Valdes-Sosa, Xia and Luo (2014) have proposed a hierarchical fusion framework using a comparatively new measure of correlation, the Maximal Information Coefficient (MIC) (Reshef et al., 2011), for temporal matching and a Bayesian approach for spatial matching on the trial level. They chose the MIC for temporal matching, because, as in the jICA, most conventional methods for temporal matching assume a linear relation between datatypes. The assumption of linearity is not necessarily false, but non-linear, more complex relations between modalities should be considered as well. For spatial matching of fMRI with EEG they used a network based method for source imaging in the EEG (NESOI) (Lei et al., 2011). In principle, NESOI employs ICA to detect multiple, temporally coherent networks from fMRI as covariance components of a Parametric Empirical Bayesian model (Friston, Penny, Phillips, Kiebel, & Hinton, 2002) for EEG imaging. With the information from spatiotemporal EEG and fMRI, they constructed a three-level, hierarchical model. It included spatiotemporal associations at the first, complementary ICA components from the temporal or spatial matching at the second and unexplained, unmatched components, which likely represented noise, at a third level (see Fig. 3).



Fig. 3. The basic preprocessing (top left) and analysing steps in the hierarchical fusion framework from Dong et al. (2014). At the bottom is the three-level, hierarchical model with components obtained from spatiotemporal matching. Spatiotemporally matched components make up the first level, components matched temporally or spatially the second level and unmatched components the third level.

Not only does this framework use a larger extend of available information, but it also accounts for shared, discrete and unexplained variation in combined EEG and fMRI data. Yet, the validity of its temporal matching is predicated on the generality and equitability assumptions of the MIC made by Reshef et al. (2011). The former refers to the coverage of a range of statistical relations (i.e. linear, exponential or periodic) in the data provided by a sufficient sample size. By the latter, more problematic assumption the authors imply that statistics should assign similar scores to equally noisy relations of different datatypes (Reshef et al., 2011). This assumption still lacks sufficient evidence (Kinney & Atwal, 2014).

### 1.2.3 Phase-amplitude coupling

One final data fusion approach that shall be considered here broadens the range of measures taken into account in combining EEG and fMRI. Next to amplitudes and latencies in ERP data, EEG offers insight into the oscillations of cortical activity. The power of frequency bands identified in the EEG oscillations is associated with various cognitive, motor and perceptual processes. Special relevance to the BOLD signal, however, is ascribed to phase-amplitude coupling (PAC). PAC is a type of cross-frequency coupling in EEG and occurs when high oscillation power fluctuations synchronise with the phase of a slower oscillation (Cohen, 2008). The interaction of the phase of a slower frequency activity with the amplitude of a higher frequency activity has generated interest, because of the unclear association of the BOLD signal and LFP.

As outlined in section 1.1, changes in the LFP show greater similarity to changes in the BOLD signal than single cell activity or MUA. Whether it is the power of the LFP or the firing rate of neurons that most closely relates to the BOLD signal, is subject to many investigations (Logothetis et al., 2001; Nir et al., 2007). Recently, Murta et al. (2016) argued that PAC was of particular interest for the prediction of the BOLD signal, because of the finding that an increase in γ power time-locked to δ phase predicted MUA (Whittingstall & Logothetis, 2009). Further, they pointed out that fluctuations in the power of multiple frequency bands were a better predictor than single power fluctuations (Magri, Schridde, Murayama, Panzeri, & Logothetis, 2012). Hence, PAC might have a more complex relation to the BOLD signal and explain more variability than just the LFP.

For multimodal data fusion, Murta et al. (2016) conducted simultaneous intracranial EEG and fMRI recordings with subjects performing a finger-tapping task. They entered two single trial estimates of PAC strength as regressors, and , into a general linear model (GLM) for modelling variation in the BOLD signal. Other than the general contributions of PAC regressors, they tested if PAC regressors explained variance in addition to single-trial power fluctuations in α, β, and γ band power, which were associated with the finger-tapping task. Next to a negative correlation of PAC strength and BOLD amplitude, they found that PAC regressors indeed had incremental value to a combination of α, β, and γ band power in explaining BOLD variation. This suggests that PAC might relate differently to the BOLD signal compared to LFP. Thus, it might explain BOLD signal variance in addition to more common measures of simultaneous recordings, such as ERP amplitude or frequency band power.

## 1.3 Partial Least Squares Regression for EEG and fMRI