
Biostatistical Modeling and Analysis of Combined fMRI and EEG Measurements

Stefanie Kalus



München 2012

Biostatistical Modeling and Analysis of Combined fMRI and EEG Measurements

Stefanie Kalus

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Stefanie Kalus
aus Bad Brückenau

München, den 17. September 2012

Erstgutachter: Prof. Dr. Ludwig Fahrmeir

Zweitgutachter: Prof. Dr. Thomas Kneib

Tag der mündlichen Prüfung: 31. Oktober 2012

Abstract

The purpose of brain mapping is to advance the understanding of the relationship between structure and function in the human brain. Several techniques—with different advantages and disadvantages—exist for recording neural activity. Functional magnetic resonance imaging (fMRI) has a high spatial resolution, but low temporal resolution. It also suffers from a low-signal-to-noise ratio in event-related experimental designs, which are commonly used to investigate neuronal brain activity. On the other hand, the high temporal resolution of electroencephalography (EEG) recordings allows to capture provoked event-related potentials. Though, 3D maps derived by EEG source reconstruction methods have a low spatial resolution, they provide complementary information about the location of neuronal activity. There is a strong interest in combining data from both modalities to gain a deeper knowledge of brain functioning through advanced statistical modeling.

In this thesis, a new Bayesian method is proposed for enhancing fMRI activation detection by the use of EEG-based spatial prior information in stimulus based experimental paradigms. This method builds upon a newly developed mere fMRI activation detection method. In general, activation detection corresponds to stimulus predictor components having an effect on the fMRI signal trajectory in a voxelwise linear model. We model and analyze stimulus influence by a spatial Bayesian variable selection scheme, and extend existing high-dimensional regression methods by incorporating prior information on binary selection indicators via a latent probit regression. For mere fMRI activation detection, the predictor consists of a spatially-varying intercept only. For EEG-enhanced schemes, an EEG effect is added, which is either chosen to be spatially-varying or constant. Spatially-varying effects are regularized by different Markov random field priors.

Statistical inference in resulting high-dimensional hierarchical models becomes rather challenging from a modeling perspective as well as with regard to numerical issues. In this thesis, inference is based on a Markov Chain Monte Carlo (MCMC) approach relying on global updates of effect maps. Additionally, a faster algorithm is developed based on single-site updates to circumvent the computationally intensive, high-dimensional, sparse Cholesky decompositions.

The proposed algorithms are examined in both simulation studies and real-world applications. Performance is evaluated in terms of convergency properties, the ability to produce interpretable results, and the sensitivity and specificity of corresponding activation classification rules. The main question is whether the use of EEG information can increase the power of fMRI models to detect activated voxels.

In summary, the new algorithms show a substantial increase in sensitivity compared to existing fMRI activation detection methods like classical SPM. Carefully selected EEG-prior information additionally increases sensitivity in activation regions that have been distorted by a low signal-to-noise ratio.

Zusammenfassung

Die Gehirnkartierung verfolgt das Ziel, Hirnarealen Funktionen zuzuweisen und damit die Funktionsweise des Gehirns zu ergründen. Es existieren verschiedene bildgebende Verfahren zur Aufzeichnung neuronaler Aktivität, die unterschiedliche Vor- und Nachteile aufweisen. Funktionelle Magnetresonanztomographie (fMRT) bietet eine hohe räumliche, aber schlechte zeitliche Auflösung. Zusätzlich ist die Erkennung neuronaler Aktivität in gängigen ereigniskorrelierten experimentellen Designs durch ein niedriges Signal-zu-Rauschen Verhältnis erschwert. Ein weiteres Verfahren stellt die Elektroenzephalographie (EEG) dar: Durch deren hohe zeitliche Auflösung können in gemittelten EEG Zeitreihen ereigniskorrelierte Potentiale klar identifiziert werden. 3D EEG-Karten, die über Quellenlokalisationsmethoden berechnet werden, haben jedoch eine niedrige räumliche Auflösung. Trotzdem enthalten sie komplementäre Information über die Lokation neuronaler Aktivität. Es ist von hohem Interesse, die Information aus beiden Verfahren zu kombinieren, um tieferes Wissen über die Funktionsweise des Gehirns zu erlangen, und komplexe statistische Verfahren sind dafür unabdingbar.

In dieser Arbeit wird ein neuer Bayesianischer Ansatz vorgestellt, der in stimulusbasierten experimentellen Designs die Performanz von fMRT-Aktivierungserkennungsverfahren verbessern kann indem er zusätzlich EEG-basierte räumliche Priori-Information einbezieht. Dieser Ansatz basiert auf einem neu entwickelten reinen fMRT-Aktivierungserkennungsverfahren. Bei der Aktivierungserkennung wird voxelspezifisch evaluiert, inwieweit ein experimentell gesetzter Stimulus einen Einfluss auf die gemessene fMRT-Signalzeitreihe hat. Der vorgestellte Ansatz stützt diese Evaluierung auf eine Bayesianische Variablenselektion und erweitert vorhandene hochdimensionale Regressionsverfahren indem die Prioriwahrscheinlichkeiten der binären Indikatorvariablen mit einer latenten Probitregression verknüpft werden. Für reine fMRT-Aktivierungserkennungsverfahren besteht der Prädiktor aus einem räumlich-variierenden Intercept. Für EEG-basierte Verfahren wird ein EEG Effekt hinzugefügt, der entweder räumlich-variiert oder konstant ist. Zur Schätzung dieser räumlichen Effektkarten werden Regularisierungsverfahren verwendet, die auf verschiedenen Markov-Zufallsfeldtypen basieren.

Die Herausforderungen bei der Entwicklung statistischer Inferenzmethoden für hochdimensionale hierarchische Modelle sind groß: Sowohl modellierungstechnische als auch numerische Probleme sind dabei zu lösen. In dieser Arbeit beruht die Inferenz auf einem Markov Chain Monte Carlo (MCMC)-Ansatz, der die Effektkarten simultan aktualisiert. Um die darin verwendeten hochdimensionalen Choleskyzerlegungen zu umgehen, wurde darüber hinaus ein auf Geschwindigkeit hin optimierter Algorithmus entwickelt, der auf elementweisen Updates beruht.

Die Performanz der vorgeschlagenen Algorithmen wird anhand umfassender Simulationsstudien und realer Anwendungen untersucht. Dabei wird die Güte der Verfahren hinsichtlich deren Konvergenzverhalten, Interpretierbarkeit sowie Sensitivität und Spezifität der daraus abgeleiteten Aktivierungsklassifikationsregeln beurteilt. Von besonderem Interesse ist die Frage, inwieweit die Hinzunahme von EEG-Information die Performanz von Aktivierungserkennungsverfahren erhöhen kann.

Zusammenfassend zeigen die neuen Algorithmen eine deutliche Steigerung der Sensitivität gegenüber bestehenden Verfahren wie SPM, ohne an Spezifität einzubüßen. Durch die Hinzunahme von sorgfältig ausgewählter EEG-Information wird eine zusätzliche Sensitivitätssteigerung in verrauschten Aktivierungsregionen erzielt.

Contents

Outline	1
1 Introduction	9
1.1 Neurophysiological basis	9
1.1.1 Neuronal activity	10
1.1.2 Generators of the EEG	12
1.1.3 Generators of the fMRI signal	15
1.1.4 Comparing fMRI and EEG information	19
1.2 The fMRI-EEG data fusion	20
1.2.1 EEG-to-fMRI approaches	23
1.2.2 fMRI-to-EEG approaches	26
1.2.3 Symmetrical approaches	28
1.2.4 A new asymmetrical EEG-to-fMRI approach	30
2 Regression models for fMRI	33
2.1 Description of predictor components	35
2.1.1 Baseline - modeling of low frequency confounds	35
2.1.2 Confounding variables	37
2.1.3 Modeling experimental stimuli	38
2.1.4 Linearization of all model components	44
2.1.5 Matrix notation	46
2.2 Distributional assumptions about the error term	46
3 An EEG-enhanced fMRI model	49
3.1 The fMRI regression model	50
3.2 Bayesian activation detection	51
3.3 Prior specifications	52
3.4 Posterior inference	58
3.4.1 Sampling activation probabilities	59

3.4.2	Binary regression level	61
3.4.3	Algorithmic details	68
3.4.4	Monte Carlo Estimates	70
3.4.5	Classification of voxels	72
3.5	A model extension: A non-negative global EEG effect	77
4	A fast CAR model for the EEG-enhanced fMRI model	79
4.1	Single site updates for spatially-varying coefficients	81
4.2	Calculating the determinant of large sparse update matrices	82
4.2.1	Overview of existing methods	82
4.2.2	Laplace eigenspectrum of grids	84
4.3	An accelerated fMRI-EEG algorithm	86
5	Software implementation	89
5.1	Usage	90
5.1.1	Preliminaries: Configuring an <code>fmrieeg</code> algorithm run	90
5.1.2	Usage of the R-package <code>Rfmrieeg</code>	92
5.1.3	Compilation and usage of the executable <code>CfmrieegMain</code>	98
5.2	Implementational issues	100
5.2.1	Efficient linear algebra subprograms	100
5.2.2	Efficient GMRF updates via Cholesky decompositions	101
5.2.3	Software Design	102
5.3	Technical notes	103
5.3.1	Image data format	106
5.3.2	External dependencies	107
5.3.3	Influences on the running time	107
6	Simulation studies	111
6.1	Simulation studies based on artificial data	111
6.1.1	Design for generating an artificial dataset	112
6.1.2	Notation and common model settings	116
6.1.3	Evaluation of different EEG predictor forms	117
6.1.4	Comparison of different fMRI activation detection algorithms	135
6.1.5	Evaluation of EEG-enhanced detection schemes	144
6.1.6	Summary	152
6.2	Simulation studies based on real-world data	154

7	Application to an oddball study	159
7.1	Material	159
7.2	Methods	162
7.3	Results	169
7.3.1	Comparison of different fMRI activation detection algorithms	169
7.3.2	Evaluation of EEG-enhanced fMRI detection schemes	178
7.4	Summary	197
8	Proof of concept	203
9	Conclusion and Perspective	215
	Appendix	223
A	Derivation of further components of the EEG-enhanced fMRI model	223
A.1	Marginalized conditional distributions of binary indicators	223
A.2	Conditional posterior means of fMRI stage parameters	225
A.3	Marginal prior probability for activation	227
B	Robustness of variance parameter estimation	229
B.1	Robustness of gCAR5 results	229
B.2	Robustness of fCAR5 results	234
B.3	Robustness of iMRF5 results	237
B.4	Summary	238
C	Implementation: Arguments for a call to the <code>fmri_eeeg</code> algorithm	241
D	Supplementary results: Results of shortened fMRI data trajectories	247
	Bibliography	257

Outline

Recording electrophysiological data simultaneously with functional magnetic resonance imaging (in the following referred to as fMRI) targets at combining the advantages of both modalities to gain deeper insight into brain functioning. Recently, extensive experimental and methodological research has been done to achieve this goal (see Mulert and Lemieux, 2010, for a compelling review). Both measurement techniques capture neuronal activity. Scalp electroencephalography (in the following referred to as EEG) measures the summed activity of postsynaptic currents. Hereby, time series of electrical activity are recorded from multiple electrodes placed on the scalp (see Figures 1.2 on p. 13 and 1.3 on p. 14). fMRI is one of the most recently developed and at present most popular form of noninvasive imaging of human brain activity. Just recently, celebrating 20 years of fMRI, an entire special issue was released by the NeuroImage journal (Bandettini, 2012) demonstrating the popularity of the fMRI technique. fMRI also captures neuronal activity but in a more indirect way than EEG: Electrically active neurons are subject to a higher energy demand leading to an increase in blood flow into active brain regions. fMRI is based on the blood oxygenation level dependent (BOLD) effect and mirrors the changes in cerebral perfusion of capillaries with oxygenated blood. The fMRI technique provides a time series of three-dimensional images of the brain, see Figure 1.4 on p. 17. Both modalities, EEG and fMRI, are based on brain processes related to neuronal activity. Therefore, activated brain regions can be located by analyzing both fMRI and EEG data.

EEG and fMRI can be contrasted by advantages and disadvantages associated with each technique. Neuronal activity is directly reflected in EEG recordings. Identifying the exact spatial localization of active neurons, however, is complicated by, on the one hand, physical constraints underlying the current conduction through brain tissue and, on the other hand, dispersion effects on the scalp. The time resolution is high and allows to identify short-lasting or high-frequency events in the submillisecond range. In contrast to EEG, with fMRI a high spatial resolution in the millimeter range is obtained, whereas the time resolution is in practice limited through several technical factors, so that data acquisition of a single whole brain volume takes about 2 seconds.

EEG and fMRI each capture different aspects of neuronal activity, but neither captures all. For being recorded at the scalp surface, electric events in deeper brain regions, which are detected by fMRI, emerge as damped oscillations in the EEG and are hard to localize. On the contrary, signals recorded by EEG need not necessarily be identifiable in the simultaneously recorded fMRI time series. For instance, no explicit BOLD response was identified for some short-lasting EEG signals with high amplitude, as e.g. K-complexes. See Ritter and Villringer (2006) and Daunizeau et al. (2010) for a more detailed discussion of situations where discrepancies between concurrent fMRI and EEG recordings emerge, and Section 1.1.4 for a summary.

The main motivation for developing statistical methods for the fMRI-EEG data fusion lies in compensating shortcomings—related to technical issues as well as the biophysical generation process—of one technique by adding information from the other to gain a deeper insight into brain functioning. A close coupling between measuring EEG and fMRI data seems indispensable for this. Therefore, in the last couple of years, recording EEG simultaneously with fMRI has become firmly established. For this to happen, several technical challenges had to be overcome (Ritter and Villringer, 2006).

Statistical methods that aim at combining EEG and fMRI face the challenge of integrating two data types that differ strongly in their structure as well as their generative process. Basically, so far, three different analysis approaches to multimodal integration have been proposed (Laufs et al., 2008; Daunizeau et al., 2010; Rosa et al., 2010). We distinguish between (i) EEG-to-fMRI approaches, where the fMRI signal is correlated with an EEG-defined event or feature to gain information on the location of electrophysiological phenomena, (ii) fMRI-to-EEG approaches, where the spatial information of the fMRI is used for a (spatiotemporal) source reconstruction of the EEG and (iii) symmetrical approaches, usually referring to the use of a common forward or a generative model that explains both EEG and fMRI data.

In this thesis, this set of approaches is extended by introducing a new type of method: an EEG-informed fMRI activation detection method. It is based on a high-dimensional fMRI regression model, which is also the basis for (i). In preparation for incorporating EEG information, a suitable uninformed fMRI activation detection method was developed. It grounds on ideas of Smith et al. (2003) and Smith and Fahrmeir (2007), who suggested to use a Bayesian variable selection approach and model averaging techniques to assess brain activity. Activity is, thereby, assessed locally at each voxel (a volume element of a three-dimensional image). A voxel is said to be active—in a narrower sense, to be responsive—if its fMRI signal trajectory responds to stimuli. That is, a voxel is active

when a predictor component describing stimulus presentation has a significant effect on the fMRI signal. Alternatively, in terms of Smith et al. (2003), a voxel is said to be active if the inclusion of the stimulus predictor component plays a role in explaining the fMRI signal. To assess voxelwise activity, Smith et al. (2003) introduced voxelwise binary activation parameters in the form of selection indicators for the stimulus component and inspected the size of corresponding selection probability estimates, which serve as activation probability estimates. To account for the correlation between neighboring voxels, estimation of activation probabilities was regularized by an Ising prior. The authors even provided a way to incorporate external prior information. This prior information, however, had to be available in the form of prior probability maps.

In this thesis, we extend the work of Smith et al. (2003) by providing means to include general prior activation information in continuous or binary form. For this, we replace the Ising prior with a latent, spatial probit regression stage that accomplishes the mapping between general variable values and probability estimations. For uninformed fMRI activation detection, the probit regression predictor consists of a spatially-varying intercept only (similar to the work of Smith and Smith, 2006). For EEG-informed fMRI activation detection schemes, spatial EEG information can be included as a further variable—possibly with a spatially-varying effect. To adjust for correlations within spatially-varying effect vectors, two different types of Gaussian Markov random field (GMRF) priors were used resulting in two different modelling choices.

Another extension to the models of Smith et al. (2003) was required to make the activation detection scheme applicable to general experimental designs commonly applied in the fMRI literature. We had to adapt the form of the fMRI regression predictor to incorporate popular event-related stimulus paradigms (Rosen et al., 1998)¹, further covariates and an advantageous choice of highpass filter regressors. For this, we followed modelling choices as used in the software package Statistical Parametric Mapping (SPM), which is in common use as a standard tool for assessing human brain activity. The theoretical framework behind implemented features of SPM is discussed in Friston et al. (2008). Event-related stimulus types of interest can either be set by an experimental procedure (external stimuli) or can be defined by spontaneous electrophysiological events (internal stimuli). Though, we focussed on event-related designs, the model formulation can also incorporate block designs.

Bayesian posterior analysis is based on a Markov Chain Monte Carlo (MCMC)-approach

¹In event-related paradigms single, short-lasting stimuli are presented. This is in contrast to block designs where longer periods of experimental conditions alternate with control conditions

that allows to directly calculate all parameters and features of interest. In particular, the marginal posterior probability for activation at each voxel is of interest, because it can be thresholded to provide an activation map. Our MCMC implementation relies on global updates of spatially-varying effect maps. Additionally, a faster algorithm is developed based on single-site updates circumventing the computationally intensive, high-dimensional, sparse Cholesky decompositions.

In comparison to previous methods, our approach has several significant advantages. It shares several advantages of the Ising model (see Smith et al., 2003). First, our Bayesian formulation allows the explicit modeling of the probability that a voxel is activated, which circumvents either the problematic interpretation of frequentist p -values in classical approaches or the counterintuitive selection of an activation threshold on the level of Bayesian activation amplitude effect estimates (Friston et al., 2002). Second, it incorporates spatial correlations at the level of activation probabilities, which are the parameters of interest, and not just indirectly on the level of activation effects as, for example, in Gössl et al. (2001a), Penny et al. (2005) and Groves et al. (2009). Third, because the posterior distribution incorporates the spatial structure of a GMRF prior, there is no need for spatial adjustments in a postprocessing step.

Above that, our approach possesses a major advantage over the Ising model. Our modelling formulation allows to incorporate very general forms of prior information, which can be used to enhance fMRI activation detection. The intended use of our approach is to incorporate EEG-prior activation information in the form of three-dimensional source reconstruction maps (Michel et al., 2004), though any kind of external prior information can be used as long as it is available as a single three-dimensional map with activation information. The scale of the contained measurements is quite arbitrary: Voxel values can be binary or continuous. We generally assume, however, that larger values indicate activation. Extensions to more than one predictor variable, also of the categorical type, can easily be accomplished. As a EEG-fMRI data fusion technique, our approach is related to (i) for being based on fMRI regression, but does not simply search for brain regions showing a correlation with EEG features. Our approach extends fMRI regression by adding spatial EEG information to activation detection.

Compared to classical SPM and the Ising model, our uninformed fMRI activation schemes prove to have superior performance. They have highly increased sensitivity without losing their high specificity level and possess excellent edge-preserving properties.

Our EEG-enhanced fMRI activation schemes are equally sensitive or at least as sensitive as our uninformed algorithms. This hints at the usefulness of our enhanced approaches and is

a prerequisite for their intended purpose. These procedures are developed to compensate a low signal-to-noise ratio in the fMRI signal, which is often a problem in event-related designs, to (a) make EEG phenomena visible otherwise not detected and (b) bring out activation regions more clearly when activation regions are blurred by noise. Though, three-dimensional maps derived by EEG source reconstruction methods have a low spatial resolution, they provide complementary information on the location of neuronal activity. To make use of EEG information without being biased by inaccurate location information, EEG-enhanced detection schemes need to be robust against prior misspecifications. Our developed model ensures this for adapting to the level of EEG and fMRI congruency within the brain. This is achieved at the expense of a sensitivity increase. However, we identified data settings in which a substantial performance gain could be achieved compensating activation loss due to noise.

Though, we focus on deriving a method for combining fMRI and EEG data, our technique can be used for solving problems arising in other applications as well. Our method is a special kind of spatial classification technique that can be applied in the absence of a training sample, but in the presence of structural spatial information. That is, it targets at estimating a latent binary Markov random field. As noted by Smith and Smith (2006), who discussed a wider scope for the fMRI activation technique proposed in Smith et al. (2003), problems adequate for applying the technique have the following properties. First, the data at each specific spatial location are distributed as a mixture of two distributions representing two underlying classes. Second, the data likelihood assumes spatial independence in the data, conditional on the binary variables and any other parameters. The objective is to provide a binary classification at each site accounting for and making use of spatial correlations. Hence, in exactly the same way as in Smith and Smith (2006), the uninformed technique can, for example, be adopted for identifying cells in data images recorded by confocal fluorescence microscopy. In addition to this, the informed technique can now be used when additional, spatial information on classification is at hand. It seems, for example, conceivable, that our technique can be applied for tumor identification. Tumor tissue can, for example, be identified by typical contrast-curve shapes recorded by dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) (Parker and Padhani, 2003; Padhani et al., 2005). Incorporating prior information in form of, e.g., a positron emission tomography (PET) image can help to quantify location and size of a tumor. Parker and Padhani (2003) note that it seems unlikely that infiltrating tumors will be entirely delineated by DCE-MRI techniques and that additional techniques such as proton spectroscopy and PET may provide additional clarification of the extent of tumor spread. Our proposed

informed classification technique provides a conceptual framework of incorporating such additional information into DCE-MRI tumor detection methods.

Beside the methodological development, the implementation of a software tool was a major part of this thesis. High diligence was put into arriving at a software solution that fulfills a high quality standard. The implementation follows object-orientated programming paradigms and meets requirements as user friendliness and easy extensibility. It incorporates automatic data management routines for data formats that are commonly used in neuroimaging analysis and has a good runtime performance for relying on high-performance linear algebra software libraries. The software is freely available as R package (R Core Team, 2012) `Rfmri EEG`. Alternatively, C++ source code is provided for a binary `Cfmri EEGMain` program. Both kinds of software packages can be obtained from the author.

To facilitate the understanding for the reader unfamiliar with the field of neuroscience, we start with a short overview and introduction on the neurophysiological basis that both EEG and fMRI rely on in Chapter 1. To convey a basic understanding of neuronal activity, we explain what aspects of neuronal activity are measured by each modality. After this, we give more details on published approaches to combine fMRI and EEG and highlight in what way the new method proposed in this thesis supplements these approaches.

The main research objective of this thesis is the development of an EEG-enhanced fMRI regression model. In Chapter 2, we describe the voxelwise fMRI regression framework it is based on in detail and—in particular—we provide information about three types of predictor components we decided to integrate into our model: highpass filter regressors, confounders and regressors that model the contribution of experimental stimuli.

The methodological framework for the EEG-informed fMRI activation detection model, which incorporates the uninformed model, is proposed in Chapter 3. For this, the underlying Bayesian variable selection approach for activation detection is introduced. We then describe in detail the used hierarchical parameter structure for incorporating the flexible latent probit regression. Thereby, hierarchical Bayesian models are based on suitable choices of (conditional) prior distributions. In particular, spatially-varying probit effects are regularized by two different types of GMRF priors. Posterior inference is implemented by a Markov Chain Monte Carlo scheme. Corresponding to the type of GMRF prior and exact predictor choice, we derive different forms of activation detection algorithms. In Chapter 4, model extensions are described to improve runtime performance leading to an accelerated algorithm.

Chapter 5 contains a description of the developed software making the analyzing techniques discussed in this thesis available to researchers and practitioners. An introductory section gives details on the basic usage of the `RfmriEEG` R-package and the binary `CfmriEEGMain` program. After this, we provide some insight into numerical solutions and the software design. Technical notes are included for users that aim at optimizing the runtime of the implemented algorithms.

In Chapter 6, we report the results of extensive simulation studies based on both controlled, artificial as well as realistic data settings. These were conducted, on the one hand, to support model building decisions and to gain insight into the functioning of the algorithms and, on the other hand, to gain knowledge about the comparative performance of the proposed models.

In Chapter 7, we demonstrate the applicability of variants of our model in the analysis of a single-subject analysis of data derived from an auditory oddball experiment. Because the performance gain by adding EEG information appears limited with our specific choice of combined fMRI and EEG data, we demonstrate the usefulness of algorithms when applying them to slightly modified, but realistic, data in Chapter 8. With it, we are able to list data properties that are favorable to a successful application of EEG-enhanced fMRI activation detection algorithms. The promising analysis results of an (unmodified) data combination from the auditory oddball experiment that is chosen in accordance to these characteristics can be found at the end of Chapter 7.

The thesis closes with concluding remarks and perspectives for the future.

Statistical inference in high-dimensional hierarchical models becomes rather challenging from a modeling perspective as well as regarding numerical issues. Before arriving at a solution to EEG-enhanced fMRI activation detection and a description of characteristics of its applicational scope, various modifications to modeling choices as well as different data analysis scenarios had to be considered. Therefore, the work at hand does not simply summarize finished work, but documents the process of finding both a suitable model and a characterization of its applicability. Hence, in a wider scope, this thesis serves as an example for accomplishing a complex statistical modeling task for a challenging practical research question that statisticians face today.

1 Introduction

This thesis focusses on developing statistical models for locating neuronally active brain regions by analyzing both fMRI and EEG data. Though a rather abstract notion of neuronal activity is sufficient to understand most details of the methodological transfer, we give in Section 1.1 a short overview and introduction on the neurophysiological basis that each technique relies on. With a basic understanding of neuronal activity, challenges and problems that arise when combining data from both modalities are easier to comprehend. In Section 1.2 we give some more details on statistical approaches to combine fMRI and EEG published so far and highlight in what way the here proposed new method extends these approaches. Readers interested solely in the technical transfer and application of newly proposed methods may skip this chapter.

1.1 Neurophysiological basis

Brain imaging modalities like EEG or fMRI are able to capture activity of large ensembles of neurons, which are connected to each other to form neural networks. A neuron is, hereby, the basic information-processing unit of the brain whereby neurons are located in the cortex and the subcortical nuclei. A neuron is active if it processes and transmits information by electrical and chemical signaling.

In the following, we describe the basics of neuronal signal processing and point out the connection to biophysical properties that can be recorded by either EEG or fMRI.

Note that the following description of biophysical processes is simplistic and targets an audience unfamiliar with neurophysiology. It is meant to give a basic understanding of the biophysical processes that generate fMRI and EEG signals. More comprehensive descriptions can be found in Zschocke and Hansen (2012) and Huettel et al. (2004) which constitute the basis for this section.

1.1.1 Neuronal activity

Before we describe the process of neuronal signal processing, the basic morphological features of a typical neuron need to be described. A cell diagram with the blueprint of a typical neuron is shown in Figure 1.1¹. A typical neuron is organized in three parts: a soma (cell body), dendrites and an axon. The cell body contains cytoplasm, organelles—such as the Golgi apparatus and mitochondria—and the nucleus that contains the DNA, the cell's genetic information. Dendrites are thin structures that arise from the cell body. They serve to integrate signals coming from other neurons that connect to the dendrites via synapses, i.e. specialized connections that transfer chemical signals. An axon is a special cellular process that branches off from the cell body at a site called the axon hillock. If the neuron's excitation is strong enough, an action potential is generated and travels down the axon, which, at its end, branches out and connects to subsequent neurons via synapses.

In the following, we look at the neuronal signal transmission between two neurons that are connected to each other by a synapse. Parts of the neuron that are located before the synapse and emit the signal, are called presynaptic. In contrary, the parts of the neuron that receive the signal, are called postsynaptic. Typically, these parts are located on another/second neuron, either its dendrites or its soma.

A postsynaptic neuron receives a signal if neurotransmitters, released by presynaptic neurons, diffuse across the synaptic cleft and bind to receptors located on the membrane of its dendrites or cell body (chemical signaling). These receptors open ion channels provoking an influx or efflux of ions that changes the local polarization of the postsynaptic neuron (electrical excitation). During rest, a neuron has a negative resting membrane potential (net potential) of about -80 to -70 mV, as there is an excess of positive ions outside the cell membrane. That is, if we place two electrodes inside and outside the cell body of a neuron an electrical potential difference of about 70-80 mV can be recorded. If an open ion channel provokes an influx of positive ions into the cell, this potential difference decreases and the postsynaptic cell membrane becomes locally depolarized. This process is referred to as an excitatory postsynaptic potential (EPSP). In turn, if ion exchange leads to an increase in potential difference, a local hyperpolarization of the neuronal membrane is observed. This process is referred to as an inhibitory postsynaptic potential (IPSP). Whether an EPSP or IPSP arises at the postsynaptic neuron depends on the type of neurotransmitter released by the presynaptic neuron. A specific presynaptic

¹accessed at http://en.wikipedia.org/wiki/File:Complete_neuron_cell_diagram_en.svg in March 2012

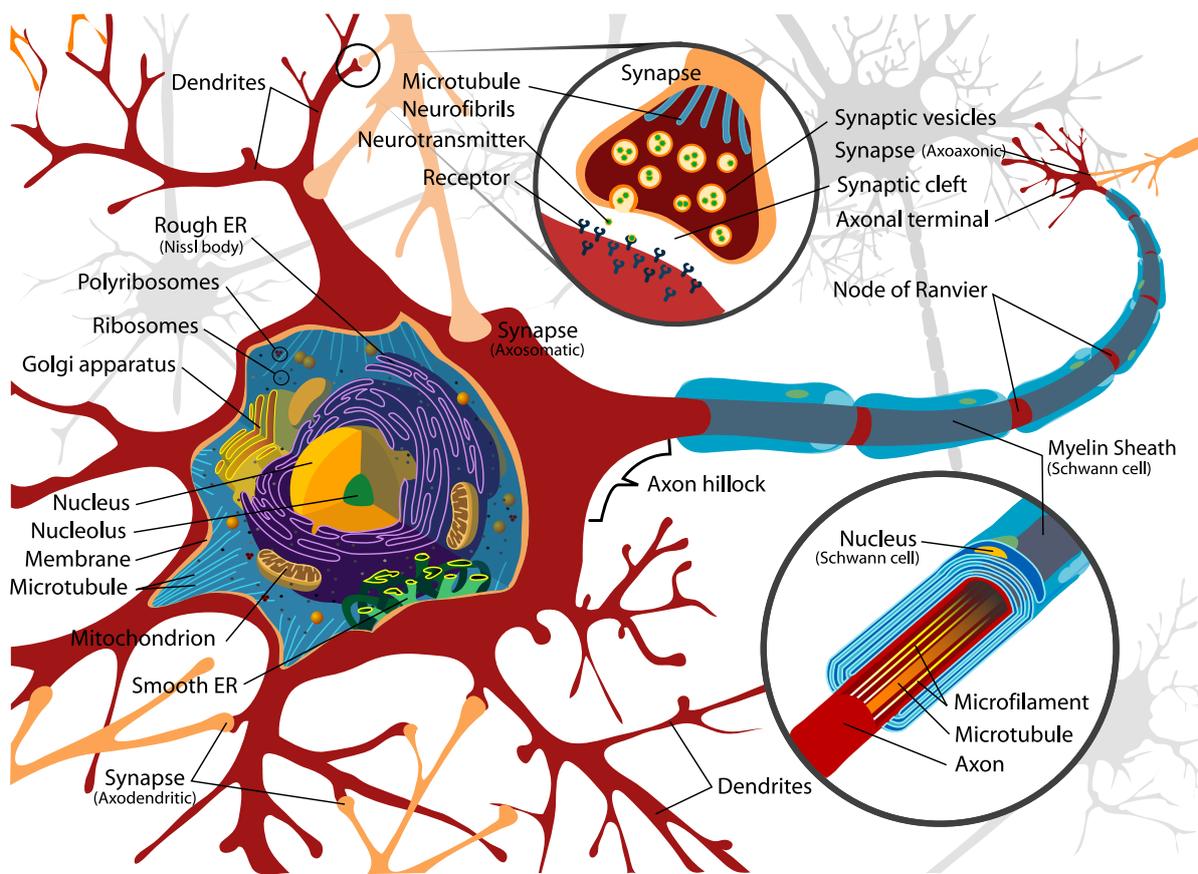


Figure 1.1: Typical structure of a neuron (nerve cell).

neuron always releases the same type of neurotransmitters promoting the same type of postsynaptic potential at postsynaptic neurons.

Once a signal is transmitted across the synaptic cleft to form an electric excitation, it is carried across the soma to the axon hillock, which is located where the axon emerges from the cell body. Individual EPSPs and IPSPs, which have been generated about the same time by signal transmissions on all the neuron's synapses, superimpose each other to influence the net polarization of the axon hillock. IPSPs, hereby, have a hyperpolarizing effect inhibiting a further signal transmission to successive neurons, whereas EPSPs have a depolarizing effect promoting a further signal transmission to successive neurons.

If the resulting sum potential at the axon hillock is larger (more depolarized) than a cell-specific threshold, an action potential (or nerve impulse) in the form of a wave of depolarization is generated, which sweeps down the axon. Eventually the action potential reaches the end of the axon, where its presynaptic terms form synapses with another neuron. Here the wave of depolarization causes a release of neurotransmitters into the synaptic cleft, again initiating either an EPSP or IPSP at the postsynaptic membrane.

Once an action potential has been emitted, it takes some time for the cell to restore its resting membrane potential. During this period no new action potential can be generated. In order to regenerate the negative resting potential, ion pumps come into operation that can transport ions against their concentration gradient to restore it. This operation, however, requires cellular sources of energy. The most important process relevant pump, the sodium-potassium pump, needs adenosine triphosphate (ATP) to work.

1.1.2 Generators of the EEG

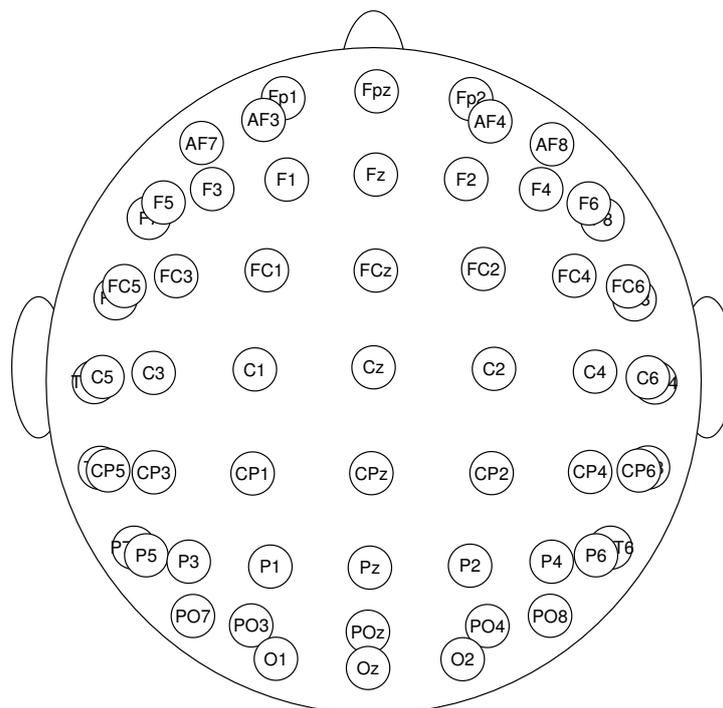
For simplicity, assume that a resting neuron has a constant net potential between extra- and intracellular space across its whole membrane, which goes along with a uniform ion density on both sides of the membrane. If neurotransmitters provoke an influx/efflux of ions by postsynaptic stimulation, this equilibrium is disturbed and a potential difference between the axon hillock and the region of synaptic activity is created, i.e. a postsynaptic potential, leading to a movement of ions. This ionic current flow is carried forward extracellularly, generating a widespread field potential. Field potentials that reach the head's surface can be measured by voltmeters on the scalp. During such recordings, potential differences are recorded from several electrodes fixated on the scalp, with one or several averaged electrodes serving as reference.

The electric potentials generated by single neurons are far too small to be picked up by the EEG. EEG activity therefore always reflects the summation of the synchronous activity of thousands or millions of neurons that have similar spatial orientation—in particular, the surface EEG is sensitive to an orientation perpendicular to the cortical surface. Excitation of cells with similar spatial orientation leads to an alignment of corresponding ionic current flows. If neurons are aligned (approximately) perpendicularly to the scalp surface, these ionic current flows lead to a charge transfer by increasing/decreasing the amount of ions under the scalp surface. Particularly, ensembles of the pyramidal neurons of the cortex share these characteristics, so they are thought to mostly contribute to the EEG signal. Because ionic current flows from deep sources hardly reach the scalp surface, their activity is more difficult to detect than currents near the skull, since voltage fields fall off with the square of distance. This is the reason why activity in the subcortical nuclei of the brain is difficult to be detected by the surface EEG.

In summary, EEG measures the summed postsynaptic activity of a large group of neurons that are located (approximately) perpendicular to the head surface. This postsynaptic



(a) Subject with EEG cap



(b) EEG electrode positions of a 64-channel EEG (top view)

Figure 1.2: Experimental EEG setup.

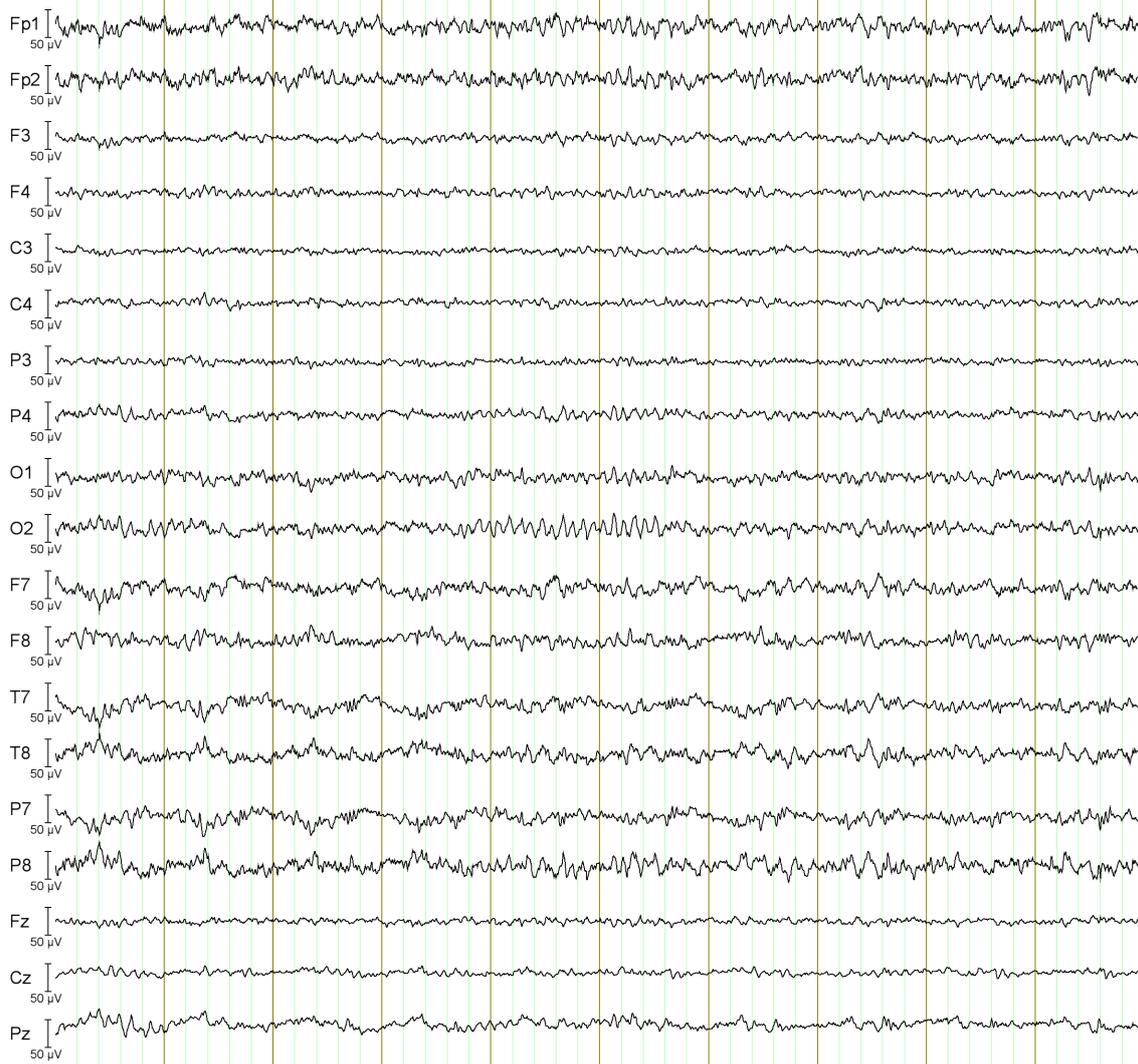


Figure 1.3: Extract (10 s) of EEG signal trajectories of a selected set of electrodes.

activity provokes ionic current flows leading to a charge transfer on the head surface which can be measured by voltmeters. The EEG signal measured at one electrode then consists of the difference in the voltage between that electrode and a reference.

Figure 1.2a depicts a subject with an EEG cap with a small number of electrodes. To enhance the spatial resolution of the EEG data, the number of electrodes can be increased. Figure 1.2b shows the electrode positions of a 64-channel EEG used for recording the EEG dataset analyzed in Chapter 7. The acronyms indicate the location on the head (i.e. the cortical lobe above which the electrode lies): frontal pole (Fp), frontal (F), central (C), temporal (T), parietal (P) or occipital (O). Combinations of two letters indicate intermediate locations. In Figure 1.3, the signal trajectories of a subset of the 64 electrodes are depicted.

The depicted trajectories represent an extract of the full EEG recordings that is processed and analyzed in Chapter 7. They were recorded in an acoustic oddball experiment during wakefulness (Kiehl et al., 2005). Typical alpha activity with a basis frequency of around 10 Hz and occipital emphasis are shown. Alpha waves are regularly recurring oscillating waves, which are a characteristic of the EEG of a normal adult in a fully conscious but relaxed state. It should be noted that deflections caused by single stimuli (e.g. acoustic stimuli) can often not be seen by the naked eye, whereas averaging many events typically unmasks such event-related potential.

1.1.3 Generators of the fMRI signal

In the following, we give a simplified description of neurophysiological processes that are supposed to underlie the fMRI signal. Up to now, the exact connection between neuronal activity and the fMRI signal is not fully understood. However, there is converging evidence that by several subprocesses, referred to as neurovascular coupling, the neuron's electrical activity triggers a vascular response that serves the supply of glucose and oxygen. It is rather these secondary hemodynamic response processes that are captured by the fMRI signal and that give indirect insights into neuronal activity. In any case, it neither reflects mere neuronal activity nor captures all neuronal activity happening (see Section 1.1.4 for more details). Nevertheless, we provide some insight into generators of the fMRI signal enabling the reader to comprehend what is measured. To this end, we follow the explanations from Huettel et al. (2004).

The fMRI signal captures neuronal activity even more indirectly than the EEG. Better to say, fMRI creates images of physiological activity that is only correlated with neuronal activity. As described above, the information-processing activity of neurons increases their metabolic demand, e.g. by operating ion pumps. Energy must be provided to meet this demand. The primary energy source for cells in the human body is adenosine triphosphate (ATP). The majority of ATP is produced in the mitochondria of cells using glucose and oxygen which are both provided by the vascular system. When transported to active regions, blood contains oxygenated blood, i.e. hemoglobin with oxygen molecules bound to it. After the oxygen has been consumed, the venous blood contains deoxygenated hemoglobin, which is the form of hemoglobin without the bound oxygen.

ATP can be produced in two ways: Anaerobic and aerobic glycolysis. Anaerobic glycolysis is fast but inefficient for consuming two ATP molecules while producing only four ATP

molecules. As its name suggests, it does not need oxygen. Contrariwise, aerobic glycolysis generates additional 34 ATP molecules. Overall, aerobic glycolysis is more efficient in terms of energy mount supply, but also more slowly.

When neuronal activity suddenly increases, an increase in blood flow ensures a sufficient supply of glucose and oxygen to satisfy its demand within ATP genesis. The blood-flow change is localized in arterioles within the distance of 2 or 3 mm of where the neural activity is. The brain, thereby, transiently prefers the fast but inefficient anaerobic method for producing ATP to satisfy the local demand—consuming glucose, but less oxygen. This leads to an oversupply of oxygenated hemoglobin and an decrease in deoxygenated hemoglobin, so that the ratio of oxygenated and deoxygenated hemoglobin increases.

Importantly, the fMRI signal depends on the so-called blood oxygenation level dependent (BOLD) effect, which can now be described as follows (Jezzard et al., 2001): Hemoglobin contains iron and, thus, possesses magnetic properties. If oxygen is bound to hemoglobin, the molecule alters from being paramagnetic (having a significant magnetic effect on its environment) to being diamagnetic (having little effect). Inside an MR scanner, paramagnetic deoxygenated hemoglobin generates local magnetic field inhomogeneities that decreases the MR signal, whereas diamagnetic oxygenated hemoglobin does not affect the recorded signal. Such oxygen dependence makes hemoglobin a sensitive magnetic marker to the level of blood oxygenation and thus indirectly to neuronal activity. If the oxygenation level of the blood decreases (or more specifically, if the level of deoxygenated hemoglobin increases), it causes the fMRI signal to decrease. Conversely, if the blood oxygenation level rises, e.g. due to increased neuronal activity, then the fMRI signal increases.

An fMRI image taken at a given time point contains structural information related to the blood oxygenation level of examined tissue types and can, hence, be used to visualize brain anatomy. Conclusions about activation can only be drawn by observing BOLD changes over time. If the signal in a voxel shows a significant correlation with the presentation of a series of stimuli, neuronal activation induced by this stimuli is observed at this voxel. Hence, to gain binary brain maps indicating neuronal responsive voxels, the signal time series of all voxels have to be examined simultaneously. That is, an fMRI dataset consists of a time series of three-dimensional brain images as depicted in Figure 1.4.

In this thesis, we focus on applications for event-related experimental designs (Rosen et al., 1998) where single, short-lasting stimuli are presented to examine the hemodynamic response they provoke. Event-related designs are very flexible. For example, the ability to randomize and mix different types of events as well as randomized inter-stimulus-times

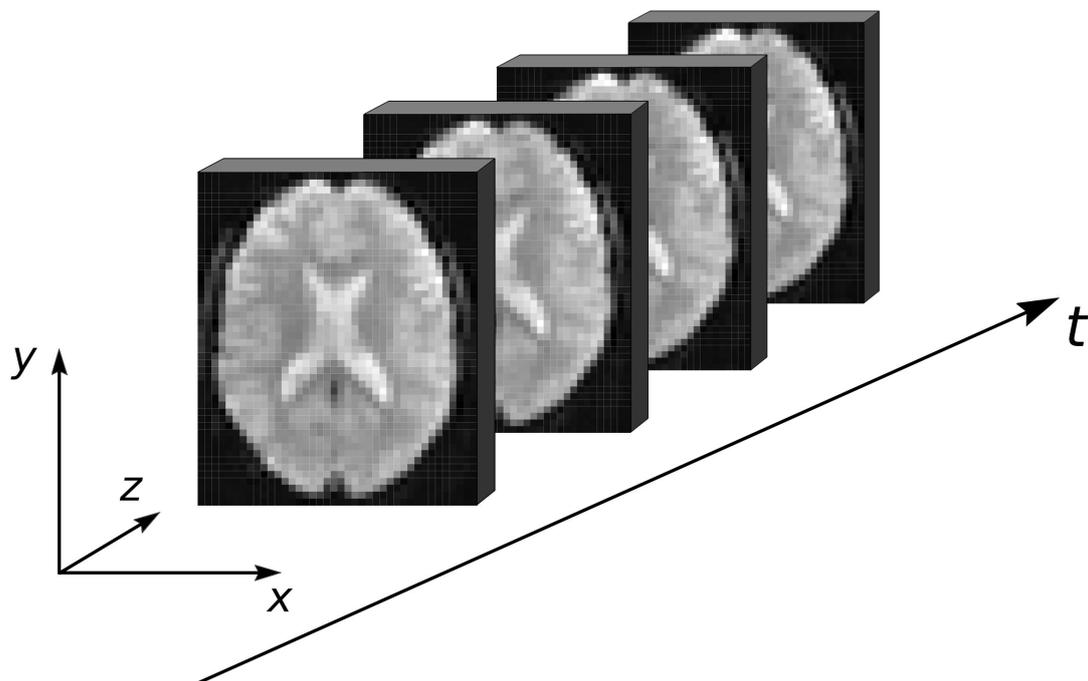
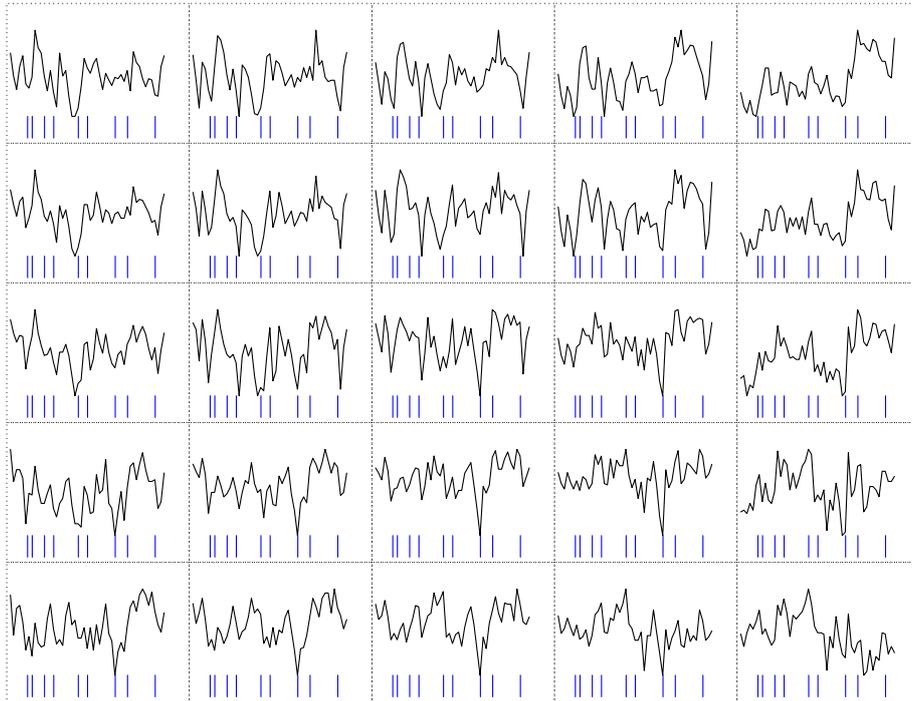
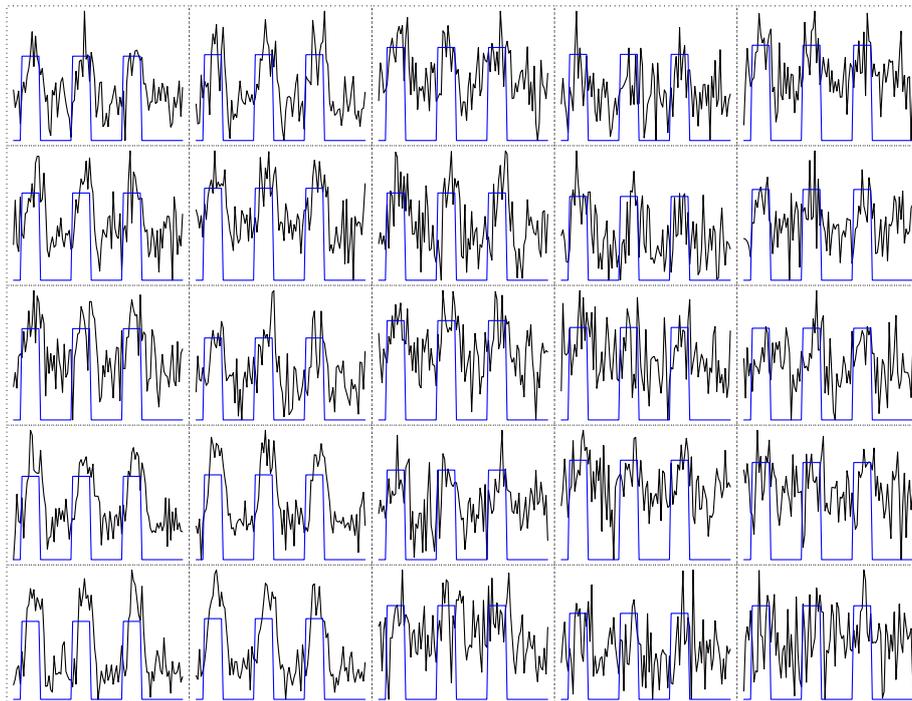


Figure 1.4: fMRI time series.

are among the major advantages of these designs. Such designs are commonly applied in electrophysiological studies and, hence, are suitable for simultaneous EEG and fMRI studies. Figure 1.5a displays the voxel time series recorded at the boundary of the auditory cortex from the acoustic oddball experiment described in Chapter 7 (voxels in the upper left corner are active). However, as can be seen here, the signal-to-noise ratio of these designs is low; that is, the relationship between stimulus presentation and hemodynamic response can hardly be spotted by visual inspection. This is in contrast to block designs where longer periods of experimental conditions alternate with control conditions. In block designs a high signal-to-noise ratio can be achieved leading to high statistical power in detecting activation. Activation can sometimes even be spotted by eye as can be seen in Figure 1.5b, which depicts—for illustration—simulated signal trajectories of a special type of block design, i.e. a boxcar design. This is in analogy to the visualization of stimulus time series of a visual boxcar experiment presented in Gössl (2001), Chap. 1.2. For event-related designs, the recording time and sensitivity of statistical models must be increased to gain equally high efficiency.



(a) 100 s extract of event-related design time series



(b) 200 s extract of boxcar design time series

Figure 1.5: fMRI signal time series (black lines) for a five by five subset of a voxel layer. Presented stimulus time series are shown in blue color.



Figure 1.6: Set-up for a combined fMRI and EEG experiment: A subject with an EEG cap is lying on a table that slides into the large cylinder of the MR scanner before recording starts.

An MR scanner is depicted in Figure 1.6. We refrain from explaining the functioning of an MR scanner, which makes use of the property of nuclear magnetic resonance (NMR) to image nuclei of atoms inside the body. The interested reader is referred to a detailed introduction in Huettel et al. (2004) or Jezzard et al. (2001).

1.1.4 Comparing fMRI and EEG information

In this section, we comment on differences in the informational content that both signal types convey. A more detailed introduction can be found in Daunizeau et al. (2010), which is summarized in the following.

As noted above, an exact biophysical model that explains the connection between neuronal activity and hemodynamics does not exist. Therefore, we do not know what aspect of neural information is reflected in the BOLD signal. While several studies report a significant correlation between the time courses of hemodynamic and electrophysiological signals, other research results have been published showing significant differences between the regions implicated by EEG and fMRI, respectively.

Ritter and Villringer (2006) and Daunizeau et al. (2010) list several physiological reasons that can explain a decoupling between electrophysiological activity (recorded by EEG) and hemodynamic activity (recorded by fMRI). For example, there might be a distance between the neuronal population, whose electrical activity is generating the EEG signal, and the vascular tree, which provides the blood supply settling the energy demand of this activity. Otherwise, we observe BOLD signal changes that are not reflected in the EEG at all. This might, on the one hand, be traced back to other brain processes with a metabolic demand, like neurotransmitter synthesis, causing hemodynamic BOLD changes without EEG correlates. On the other hand, this might be traced back to neuronal activity that is invisible to EEG, e.g. the activity of non-synchronized neuronal networks. Contrariwise, EEG events can be recorded without having a BOLD response. This applies, for instance, for short-lasting electrophysiological activity.

Technical reasons for a mismatch between EEG and fMRI exist as well. Low signal-to-noise (or rather contrast-to-noise) ratios in recordings of one modality may obscure neuronal activity. The statistical approach proposed in this thesis is mainly concerned with detecting masked neuronal activation contained in noisy fMRI signals with the help of EEG. We contrast our approach with other statistical models for the fMRI-EEG data fusion in the next section.

1.2 The fMRI-EEG data fusion

As described in the preceding section, neither EEG nor fMRI techniques capture all neuronal activity—and, above that, there is a mismatch between the signals contained in EEG and fMRI for not being necessarily caused by the same underlying neuronal processes. Therefore, Daunizeau et al. (2010) follows Friston (2005) and suggests to define “neuronal activity” operationally as a state of nodes in a network responding to specific events (e.g. cognitive, sensorimotor or spontaneous changes in brain activity). This definition unfortunately leads to the inclusion of remote correlates of neuronal activity that should be treated as artifacts. In fMRI analysis such an artifact, for example, is activity correlated pulsation of distant arteries or veins like the superior sagittal sinus. In spite of this definition, activation detection results must be carefully interpreted to differentiate between artifacts and neuronal activity in a narrower sense.

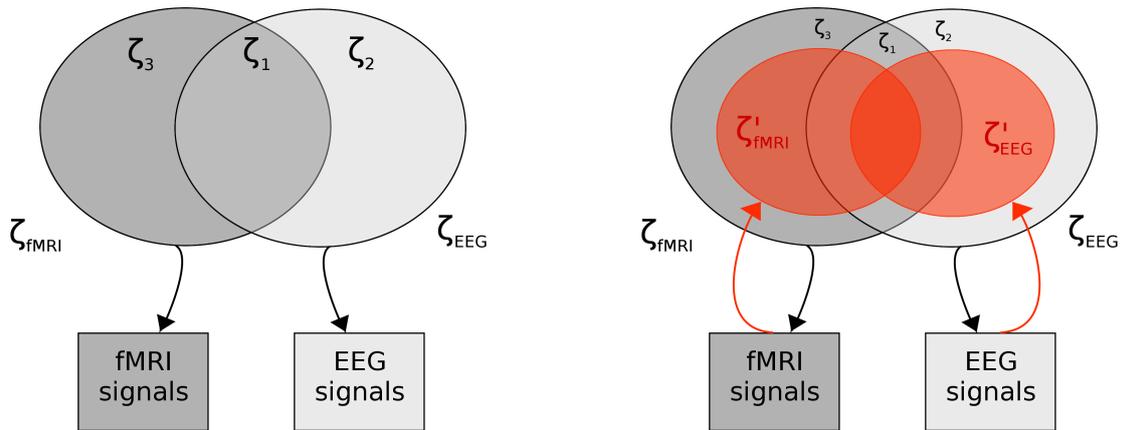
In the following, we adapt a neuronal source model proposed by Daunizeau et al. (2010) to formalize the coupling-uncoupling between bioelectric and hemodynamic responses.

EEG measurements are generated by a subset ζ_{EEG} of neuronal activity, whereas fMRI is generated by another subset ζ_{fMRI} . Both subsets overlap forming the common source ζ_1 of both signals. In the following, we denote ζ_3 as the fMRI source without components contributing to EEG, and ζ_2 vice versa. A diagram of this model can be found in Figure 1.7a. There, the modality specific sets of neuronal generators are depicted as two overlapping gray ellipses.

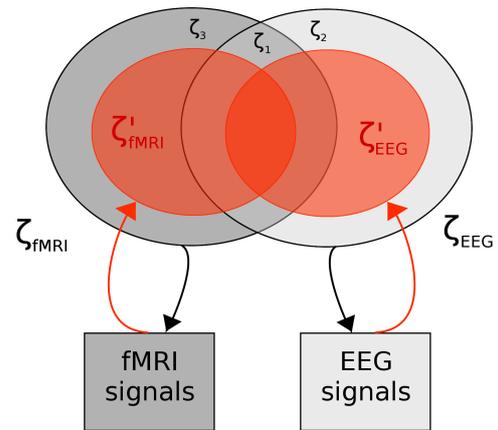
We extend this model and attempt to visualize the contribution of the different inference techniques to identify neuronal sources. First of all, we observe that all models and measuring techniques are subject to error and noise. Hence, EEG analysis only provides information on a subset ζ'_{EEG} of ζ_{EEG} and fMRI analysis only provides information on a subset ζ'_{fMRI} of ζ_{fMRI} . These portions of the true neuronal source are depicted as red ellipses in Figure 1.7b. Black arrows stand for the process of measuring signals deduced from neuronal sources, whereas red arrows stand for the application of modeling and inference techniques to infer neuronal generators. To visualize the contribution of combined fMRI-EEG analyzing techniques and their corresponding benefit compared to unimodal methods, we use the following color codes: Dashed yellow areas indicate neuronal sources that both unimodal and multimodal methods infer, whereas plain yellow areas indicate neuronal sources that multimodal approaches intend to increase information on.

In the following subsections, we give some more details on combined fMRI-EEG analyzing techniques proposed so far in the literature. As mentioned in the introductory section of this chapter, approaches can be subdivided into three types (Laufs et al., 2008; Daunizeau et al., 2010; Rosa et al., 2010): two asymmetrical ones, which add information from one modality into commonly applied analyzing techniques of the other, and symmetrical data-fusion techniques modeling a common generation process. As noted before, we distinguish between (i) EEG-to-fMRI approaches (its use is depicted in Figure 1.7c), (ii) fMRI-to-EEG approaches (Figure 1.7d), and (iii) symmetrical approaches (Figure 1.7e). At the end of each subsection, we give an explanation for which reason the respective combined analyzing technique is classified into the introduced neuronal source model as depicted in Figure 1.7.

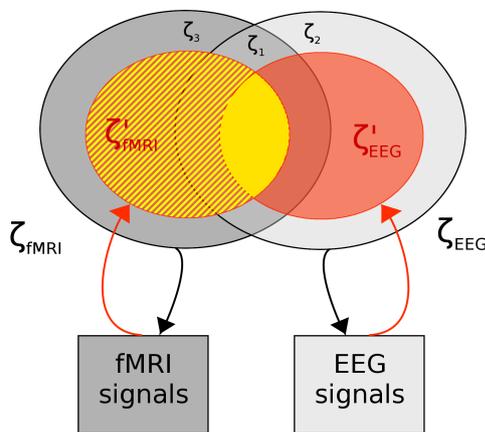
We close this section with a motivation for the EEG-enhanced fMRI activation detection model proposed in this thesis, which can be interpreted as an extension to (i), and show its additional benefit it has when inferring sources of neuronal activity.



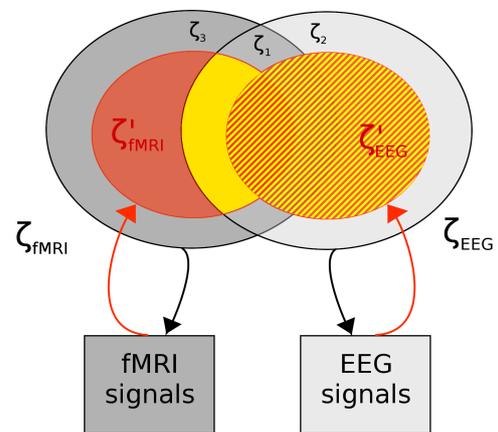
(a) Basic coupling-uncoupling model



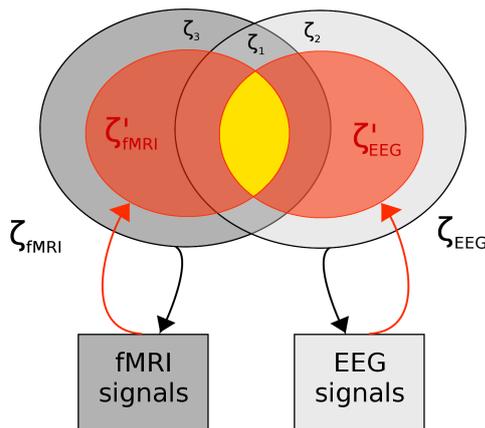
(b) Extension with modeling error and measurement noise



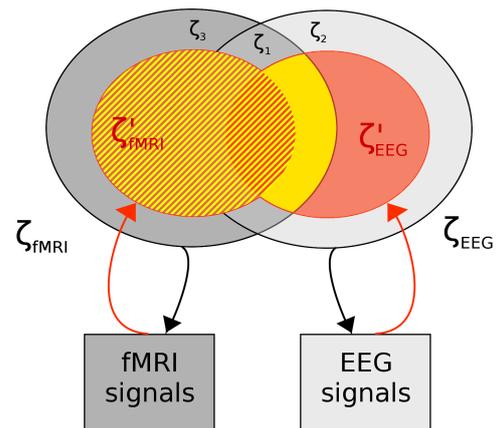
(c) EEG-to-fMRI approaches



(d) fMRI-to-EEG approaches



(e) Symmetrical approaches



(f) EEG-enhanced fMRI

Figure 1.7: Formalization of the EEG/fMRI coupling-uncoupling with visualization of neuronal activity components different data fusion approaches aim at. For more details see the description in the introductory part of Section 1.2.

1.2.1 EEG-to-fMRI approaches

EEG-to-fMRI approaches utilize the well developed methodological framework of (general) linear models (LMs) conventionally applied for fMRI analysis. For mere fMRI analyses without EEG information, they have been introduced by Friston et al. (1994) and Worsley and Friston (1995). Basically, voxelwise linear models are used to decompose the time series $y_{i,t}, t = 1, \dots, T$ of T fMRI recordings at each voxel i ($i = 1, \dots, N$) into

$$y_{i,t} = f_1(i, t) + f_2(i, t) + \dots + f_p(i, t) + \epsilon_{i,t},$$

where $f_1(i, t), \dots, f_p(i, t)$ are p additive regressor components, like baseline, confounders, stimulus components, etc., and $\epsilon_{i,t}$ is a Gaussian noise term. For linear regression analysis components are chosen to be linear in unknown parameters:

$$y_{i,t} = \sum_{j=1}^p \mathbf{X}'_{j,t} \boldsymbol{\beta}_{j,i} + \epsilon_{i,t},$$

where $\mathbf{X}'_{j,t}$ is the design vector at time t for the j -th regressor component (which is the same for all voxels) and $\boldsymbol{\beta}_{j,i}$ is the voxel-specific parameter vector of the j -th component. Classical approaches locate neuronal activation by voxelwise hypothesis tests. Activation induced by a stimulus modeled via component j can, for instance, be assessed by testing $H_0 : \boldsymbol{\beta}_{j,i} = \mathbf{0}$ for all i . If this stimulus component is modeled by a single regressor, its effect can be quantified via voxelwise T -statistics. Voxels with T -statistic values surviving a critical threshold (derived by multiple test correction procedures) indicate the location of activation. In Figure 1.8, an exemplary activation map of testing the effect of an odd stimulus in an auditory oddball experiment is shown (own data, details not shown). For a description of the experimental paradigm, see Kiehl et al. (2005). Data analysis has been conducted using the statistical parametric mapping (SPM) software (<http://www.fil.ion.ucl.ac.uk/spm/software/>), version SPM8. The critical threshold in this example was 4.985, which controls the family wise error (FWE) rate at a significance level of 0.05 in an analysis of 203 577 brain voxels. Only voxels surviving this threshold are depicted in heat-colors. For a more detailed discussion of the linear fMRI regression model used in this thesis and the modeling of predictor components of interest, see Chapter 2.

EEG-to-fMRI approaches extend this model by adding an EEG regressor component to the model. That is, the influence of temporal electrophysiological data on the fMRI signal is

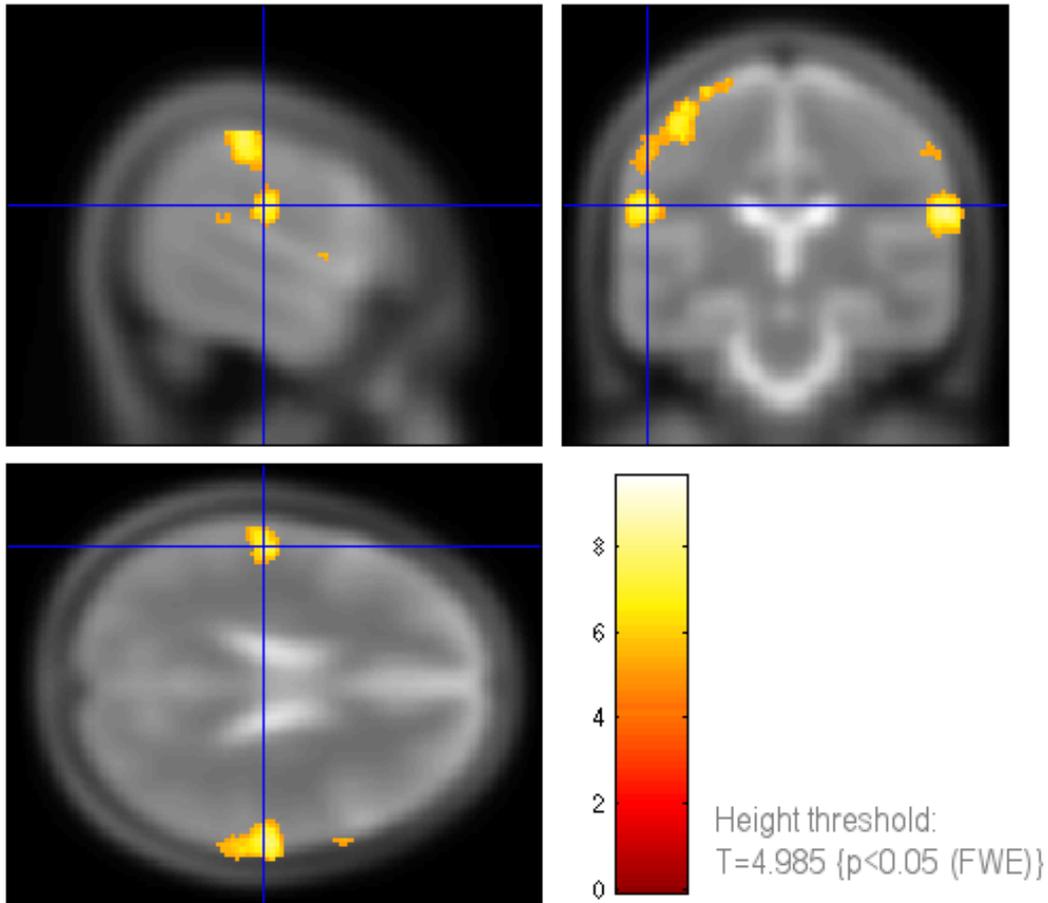


Figure 1.8: Classical activation map for the odd stimulus component of an acoustic oddball experiment. For this demonstration, the brain is cut into sagittal, coronal and transverse planes, which corresponds to a cut in x -, y - and z -direction of a three-dimensional image. Regions in heat colors indicate activated voxels according to a significance test of the odd stimulus effect. Heat-colors indicate thereby the size of the used T -statistic.

examined at each voxel i using

$$y_{i,t} = f_1(i,t) + f_2(i,t) + \dots + f_p(i,t) + f_{EEG}(i,t) + \epsilon_{i,t},$$

Thereby, categorical or continuous variables are derived from EEG-defined features or events, where applicable convolved with a hemodynamic response function, to form $f_{EEG}(i,t)$. If corresponding explanatory variables have a (significant) effect on the fMRI response of a voxel, the neuronal generators of the EEG signal are supposed to be located there. Hence, thresholded T - or F -statistic maps indicate where neuronal generators of

included electrophysiological features are located in the brain. Review articles by, e.g. Salek-Haddadi et al. (2003) and Menon and Crottaz-Herbette (2005), list several applications for this approach. In the following, some examples are given.

Inclusion of EEG regressors in fMRI LMs is, for example, applied in observational studies with stimulus- and task-free conditions like sleep or resting states (Damoiseaux et al., 2006) to study spontaneous physiological as well as pathological brain activity. The inclusion of power trajectories of frequency bands of interest, for instance, gives information on the location of corresponding brain rhythm generators (Laufs et al., 2003; Ritter and Villringer, 2006). Other examples of EEG features that are used to define events or epochs as fMRI predictors, are interictal epileptiform abnormalities or sleep spindles (Gotman et al., 2006; Schabus et al., 2007).

In experimental studies, regressors defined upon event-related EEG potentials (ERPs) are, for example, incorporated to find corresponding fMRI activation foci. For one trial, i.e. experimental session, an ERP is thereby calculated by averaging the EEG signal over all presented stimuli of one stimulus type within a pre-defined time window to ensure a sufficient signal-to-noise ratio. Recently, researchers focused on single-trial analyses (related to random effects analyses) where subject-specific patterns and their variability are examined instead of group profiles (Eichele et al., 2005; Debener et al., 2006; Bagshaw and Warbrick, 2007).

Besides locating EEG-defined events or features, electrophysiological data can be used to define confounds that can be included into fMRI regression to regress out artificial effects or increase power by reducing error noise. Based on EEG data, for example, cardiac and respiratory cycles can be captured that might confound effects of regressors of interest (Glover et al., 2000; Liston et al., 2006).

To classify EEG-to-fMRI approaches into the neuronal source model, introduced in the beginning of this section, we consider the following arguments. Essentially, EEG-to-fMRI approaches have the same task as mere fMRI regression approaches; that is, via the LM framework the fMRI signal is decomposed into predictor components. In EEG-to-fMRI approaches, in particular, the contribution of EEG regressors to the fMRI signal is evaluated to locate neuronal EEG generators by using the high spatial resolution of the fMRI. As such it is not a new analyzing technique developed for the fMRI-EEG data fusion. Hence, we decided to depict the benefit of these approaches as shown in Figure 1.7c. Yellow stripes indicate detected sources found by non-EEG regressors, whereas the plain yellow area indicate the benefit of including EEG regressors to gain information on the common source of fMRI and EEG.

1.2.2 fMRI-to-EEG approaches

fMRI-to-EEG approaches target at enhancing the performance of the EEG source reconstruction problem. EEG source reconstruction techniques try to find the three-dimensional location of neuronal generators of the corresponding (two-dimensional) signals measured on the scalp surface. Two different classes of models have been developed to solve the problem (Michel et al., 2004; Daunizeau et al., 2010): equivalent current dipole (ECD) models (Mosher et al., 1992; Kiebel et al., 2008) and distributed source models (Hämäläinen and Ilmoniemi, 1994).

For ECD models, an assumption must be made beforehand on how many activation foci are to be found. This assumption is generally content-driven. Algorithms then estimate the location and orientation of dipoles via a nonlinear optimization procedure, because the used forward model is nonlinear in dipole locations. The spatial distribution of the activated area involved, however, cannot be inferred. ECD models are no imaging technique in a narrower sense, because they do not calculate a three-dimensional voxel image, but only a limited number of dipole locations. fMRI information can be incorporated as a prior constraint for the position of dipoles, which are supposed to lie in the center of fMRI activation regions (cf. e.g. Wagner and Fuchs, 2001).

Distributed source models do not suffer from mentioned disadvantages of ECD models like yielding position estimates for only a small, specified number of dipoles without information about their spatial extend. In the following, we give some more details on the methodology behind distributed source models based solely on EEG, because we pursue to incorporate these in the combined fMRI-EEG model proposed later in this thesis. Afterwards we give some information on fMRI-constrained distributed source models proposed so far in the literature.

Distributed source models calculate the strength and direction of dipoles on a dense (voxel) grid typically distributed all over the cortical sheet producing a three-dimensional image of current density values. This unknown current density distribution is the solution of an under-determined system of equations, in which the number of parameters p to be estimated exceeds the number of observations N_e by far ($N_e \ll p$). Hence, restrictions have to be introduced to constrain the solution space by prior information. The distributed source model is generally written as a multivariate linear model

$$\mathbf{M} = \mathbf{K}\mathbf{J} + \mathbf{E}$$

with \mathbf{M} being the $N_e \times T$ observation matrix of scalp EEG measurements recorded at time points $t = 1, \dots, T$, \mathbf{K} is a known $N_e \times p$ transition matrix (i.e. a lead field matrix as described in Fuchs et al., 2002) containing the mapping of cortical sources to the scalp surface by an underlying physical model, and \mathbf{J} is a $p \times T$ matrix of unknown current density parameters of p dipoles (located on the pre-specified voxel grid) at each time point t . Note that, the number of parameters p coincides with the number of grid voxels, which we generally denote as N in this thesis. It is assumed that the model is correctly specified except for an additive white noise term \mathbf{E} . Proposed distributed source models use different regularization techniques to yield solutions with certain properties (Mattout et al., 2006). Estimation is generally based on minimizing a quadratic loss function in the form of a penalized least squares criterion (PLS):

$$PLS(\lambda) = \|\mathbf{M} - \mathbf{KJ}\|_{\mathbf{W}_{N_e}}^2 + \lambda \|\mathbf{J}\|_{\mathbf{W}_p}^2$$

where $\|\cdot\|_{\mathbf{W}}^2$ is the L_2 norm with respect to metric \mathbf{W} and λ is a regularization parameter controlling the relative weight of both terms, which should be minimized. Given a normally distributed error, this criterion corresponds to a Bayesian linear regression model with Gaussian error prior, $\mathbf{E} \sim \mathcal{N}(0, \mathbf{C}_{N_e})$ with $\mathbf{C}_{N_e} = (\mathbf{W}'_{N_e} \mathbf{W}_{N_e})^{-1}$, and $\mathbf{J} \sim \mathcal{N}(0, \mathbf{C}_p)$ with $\mathbf{C}_p = (\lambda \mathbf{W}'_p \mathbf{W}_p)^{-1}$ as prior for \mathbf{J} . Choosing $\mathbf{W}_p = \mathbf{I}_p$, with \mathbf{I}_p being the p -dimensional unity matrix, leads to the minimum-norm approach of Hämäläinen and Ilmoniemi (1994). If \mathbf{W}_p equals a spatial Laplacian matrix the smooth LORETA solution is calculated (Pascual-Marqui et al., 1994). Mattout et al. (2006) model \mathbf{C}_{N_e} and \mathbf{C}_p as a linear combination of variance component matrices within a two-stage hierarchical model to adapt variance structure estimates to more correlated correlation structures. The BASTA method of Daunizeau et al. (2006) can as well be derived from a multi-stage hierarchical model. Here, in contrast to Mattout et al. (2006) and related to ECD models, the current density distribution estimation is based on a brain parcelling into homogeneous clusters.

Source reconstruction of raw EEG trajectories is of minor practical interest, because single-stimulus profiles suffer from a low signal-to-noise ratio. More commonly, several EEG responses to a selected stimulus type are averaged in a defined time window, with this average being referred to as event-related potential (ERP). The improved signal-to-noise ratio reveals the typical response profile evoked by this stimulus type. For selected time points of this mean ERP, a source reconstruction method can then be used to infer the location of neuronal generators. In Figure 1.9, the process of a source reconstruction of two selected time points of an averaged ERP signal in response to acoustic stimuli is visualized. Data stem from the acoustic oddball experiment analyzed in Chapter 7.

Figure 1.9 depicts the mean ERP response to the so-called odd tone stimulus type. The shown source distribution solution has been calculated by the sLORETA software—an extension to the LORETA software (<http://www.uzh.ch/keyinst/loreta.htm>).

Instead of prior information influencing solely the structure of the solution, like its smoothness, fMRI information (or other external information) can be included to constrain the solution space (Dale and Sereno, 1993). Generally, fMRI activation information in the form of a binary activation map \mathbf{Z} can be incorporated into the prior variance-covariance matrix of \mathbf{J} by choosing $\mathbf{C}_p \propto f(\mathbf{Z})$ where $f(\cdot)$ is an appropriate transformation function (Daunizeau et al., 2006). The function $f(\mathbf{Z})$ is often chosen as a linear combination of variance-component matrices with one variance-covariance component being proportional to $\lambda_{\mathbf{Z}} \mathbf{D}_{\mathbf{Z}}$ with $\lambda_{\mathbf{Z}}$ being an additional, positive weight parameter and $\mathbf{D}_{\mathbf{Z}}$ the diagonal matrix of the vectorized \mathbf{Z} entries (Liu et al., 1998; Phillips et al., 2002; Daunizeau et al., 2006). Hence, if voxel i is activated, weight $\lambda_{\mathbf{Z}}$ is added to its element-specific variance putting higher weight resp. probability to \mathbf{J}_i values farther away from 0. This approach is related to the concept of spike-and-slab priors well-known in the Bayesian variable selection literature (George and McCulloch, 1993).

To classify fMRI-to-EEG approaches into the introduced neuronal source model, we assert that the high spatial resolution of the fMRI is used to add information to the ill-posed source construction problem to find more reliable source locations. Noise and modeling bias is intended to be reduced to detect neuronal sources that could not be detected on the basis of the EEG alone. Hence, fMRI-to-EEG approaches cover a larger area of neuronal sources. This is visualized in Figure 1.7d. The gain in knowledge about neuronal sources is visualized by the plain yellow area, whereas the dashed yellow area indicates sources that both the unimodal and multimodal method can infer. Having said this, we have to admit that this is a rather optimistic view. Incorporating fMRI information as prior information into EEG source reconstruction models can also introduce bias when some sort of decoupling occurs (as discussed in Section 1.1.4). For more details on this, we refer to Ahlfors and Simpson (2004), Ritter and Villringer (2006) and Daunizeau et al. (2010), and references therein.

1.2.3 Symmetrical approaches

Both asymmetrical approaches suffer from problems arising from the decoupling of EEG and fMRI. EEG-to-fMRI approaches, for instance, only locate EEG generators contained

in the fMRI signal, whereas results from fMRI-constrained EEG source localization can be biased if signals are generated by partly- or non-overlapping sources. Symmetrical approaches are sought that assess brain function by utilizing both electrical and hemodynamic measurements simultaneously and symmetrically. In the following, we give a short overview of this topic. Details can, for example, be found in Daunizeau et al. (2010) and Rosa et al. (2010).

The ultimate goal of symmetric data fusion approaches would be to infer neuronal sources generating both EEG and fMRI signals with high precision by using the advantages of each modality. To achieve this, first, two separate generative (resp. forward) models have to be developed that specify the relationship between neuronal generators and each signal. Second, these two separate models have to be linked to form a combined generative model. If data is at hand, inversion of this forward model can be applied to infer unknown underlying sources, which are treated as unknown model parameters. A probabilistic framework, thereby, is desired for allowing to establish a connection between informational content and uncertainty contained in the parameter estimation.

The development of such data fusion model is still an integral part of ongoing research. There are different approaches to this task: Either commonly applied data analyzing strategies are extended or realistic neurophysiological models describing the connection between neuronal generators and electrical resp. hemodynamic measurements are developed (neurovascular coupling models).

Developing a neurophysiological model is a challenging task. Due to the lack of consensus on the exact nature of the neurovascular coupling process, up to now no generally accepted modeling approach has been established. Further, due to the complexity of the brain processes and derived models, no multimodal neurophysiological model has been proposed containing parameters whose estimation can conjointly be informed by both EEG and fMRI measurements. Existing neurophysiological models have mainly proven useful when different biophysiological plausible hypotheses are compared in the context of building a neurovascular coupling model. Hence, so far they are not ready to be used as data fusion strategies.

Alternatively, joint models have been developed based on established, unimodal analyzing techniques of each data type. These are combined in an equally weighted, symmetrical way to fuse electrophysiological and hemodynamic data. This circumvents modeling the complex neurovascular coupling by solely estimating certain properties detected by those two analyzing techniques the data fusion approach is based on. A fine example for

such an approach is the following: Daunizeau et al. (2007) (and in an extension Luessi et al., 2011)) combine both the fMRI LM and EEG forward model to estimate the common spatial activation profile of electrophysiological and metabolic responses. Inferring the common source of both types of signals, this symmetrical approach can be classified into the neuronal source model scheme as depicted in Figure 1.7e. As noted in Daunizeau et al. (2010), the intersection of EEG and fMRI generators (depicted as plain yellow area) is the neuronal subset that data fusion approach intend to capture. More precisely, a multimodal approach should benefit from the complementary nature of EEG and fMRI by providing different perspectives on the common subspace. The goal of symmetrical approaches is, hence, not to infer other sources than the common neuronal generators. Since no information about ζ_2 (resp. ζ_3) is available from EEG (resp. fMRI), no multimodal procedure will provide a better characterization of this activity subspace than a unimodal EEG (resp. fMRI) analysis.

1.2.4 A new asymmetrical EEG-to-fMRI approach

In this thesis, we propose an asymmetric model that extends established EEG-to-fMRI approaches. That is, we use a spatial regression model for analyzing the fMRI signal and incorporate EEG information to increase our knowledge about the underlying neuronal activation profile. Hereby, we focus on modeling the response to a stimulus in event-related paradigms. Three-dimensional EEG information is used to enhance the detection of activation provoked by the examined type of stimuli.

The proposed approach differs fundamentally from EEG-to-fMRI approaches in the literature by not introducing EEG information in the form of a further predictor variable (as described in Section 1.2.1), but as prior information in the evaluation of whether a voxel can be classified as active, i.e. responds to a stimulus. The aim of these two approaches is essentially different. Whereas current EEG-to-fMRI approaches search for the location of EEG generators, our approach intends to enhance fMRI activation detection by using EEG-prior information. For this, we extend the approach of Smith et al. (2003) and Smith and Fahrmeir (2007) who use a spatial Bayesian variable selection approach to determine whether a regressor modeling the experimental stimulus contributes to explaining voxelwise fMRI signals. Smith et al. (2003) and Smith and Fahrmeir (2007) use a spatial Ising prior (Hurn et al., 2003) to incorporate information of neighbors when determining whether a voxel is active. Instead of this, we now use EEG information in the form of three-dimensional source maps to influence activation classification. The general

motivation is that if a strong EEG signal is present at a specific voxel, the prior probability of a voxel being active in an fMRI model should as well be increased.

In this process, an EEG source map is calculated by means of unimodal source reconstruction methods as briefly described in Section 1.2.2. Hereby, the high temporal resolution of the EEG signal renders it possible to access the source distribution of virtually arbitrary time points within a recording session. For instance, a time point of maximal depolarization after a presented stimulus (i.e. a specific positive ERP component) constitutes a characteristic of special interest. In an application presented in Chapter 7, we use EEG source maps calculated by the sLORETA method (Pascual-Marqui, 2002). Alternatively, other source reconstruction methods like for example BASTA (Daunizeau et al., 2006) can as well be used.

Despite the intention to combine EEG and fMRI data, our model is not restricted to the use of EEG data. Generally, any kind of three-dimensional data based on external knowledge or measurements from other human brain mapping techniques (e.g. structural MRI) can be used that potentially adds information to activation classification. For yielding an increase in sensitivity, it must, however, be ensured that the activation information contained in the fMRI signal and external prior information match. Though, results cannot be biased if they contain conflicting information. The prior is adaptive, because the proposed model incorporates a parameter structure modeling the relationship between data types flexibly. If the external information completely contradicts the fMRI signal, its weight is decreased locally or globally—depending on the exact type of proposed algorithm.

The intended use of the proposed model is the following: First, the model should make EEG phenomena within activation profiles visible when otherwise not detected. Second, activation regions should be brought out more clearly when the corresponding fMRI signal is damped by noise. Overall, the sensitivity of standard fMRI analysis should be increased. This can as well be used to reduce the examination time of a subject, or the number of subjects needed for a clinical question. The proposed model is constructed in a way that the neuronal activity it intends to capture has to be contained in the fMRI signal in some way, but is covered by noise. Additional information in the form of EEG should help to detect it. Hence, our approach can be classified into the neuronal source modeling scheme as depicted in Figure 1.7f. As before, the potential gain in knowledge about neuronal sources is depicted as the plain yellow area.

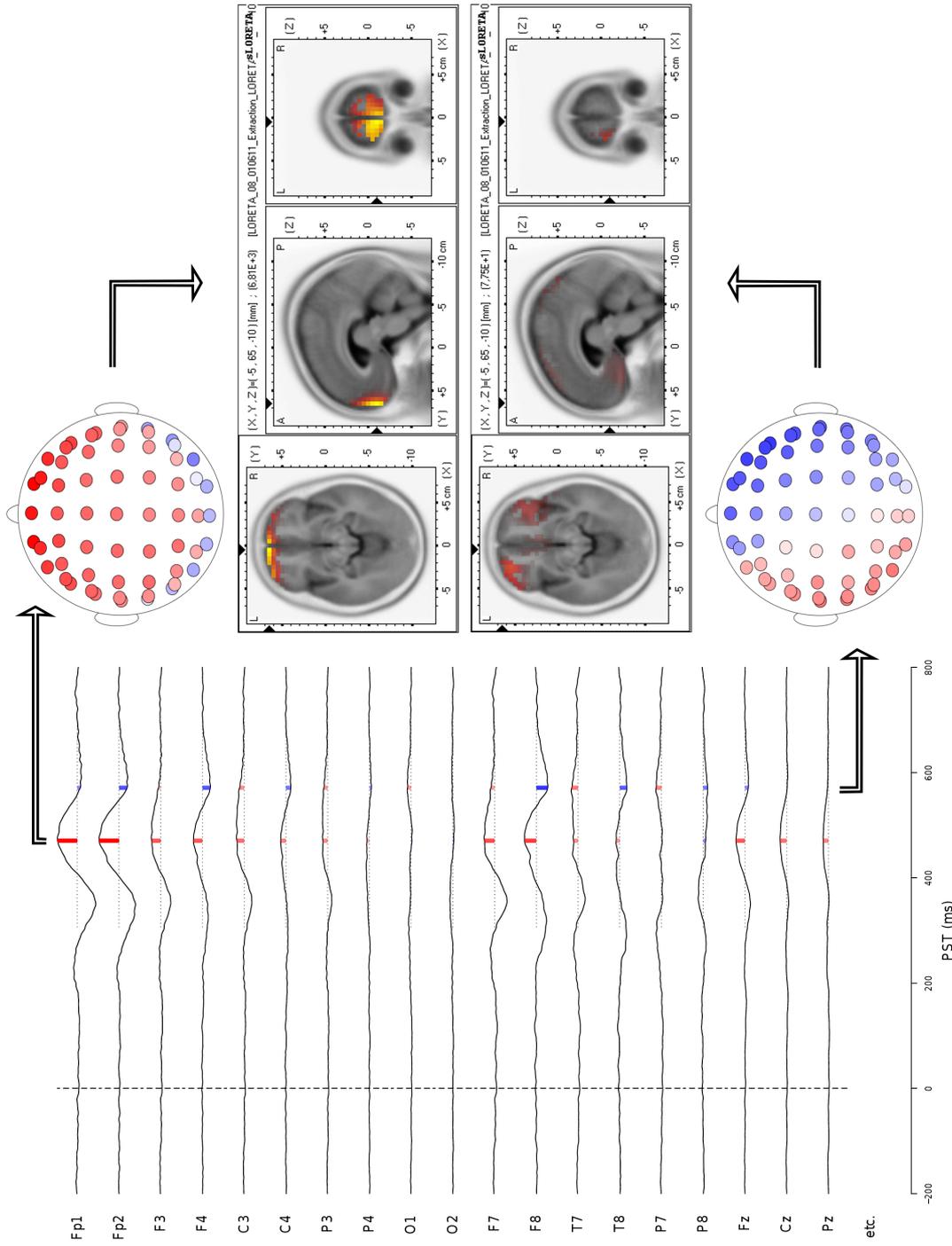


Figure 1.9: EEG source reconstruction process visualized via a 64 channel EEG. The user selects time points of interest in the mean ERP trajectories as depicted on the left (several electrode trajectories are omitted for lack of space). Source reconstruction techniques (here: sLORETA) calculate a 3D solution (center right) based on ERP amplitudes values of all electrodes of each time point and corresponding electrode positions on the scalp surface (depicted bottom resp. top right).

2 Regression models for fMRI

Generally, fMRI studies evaluate brain-wide differences in the response strength to experimental stimuli. Thereby, the main focus lies on the identification of activated voxels. A voxel is said to be activated if its fMRI time series (compare Figure 2.1a) shows a relationship with the presentation rate of experimental stimuli (compare Figure 2.1b). Early studies used voxelwise correlation coefficients (Bandettini et al., 1993) to quantify correspondence between signal time series and stimulus time series. To adjust for confounding effects, Friston et al. (1994) proposed to use voxelwise multiple regression analysis to assess the effect of a stimulus regressor—which is a direct extension to the use of mere correlation coefficients. As Friston et al. (2008) notes (p. 16), with very few exceptions, every fMRI model is based on a (general) linear model formulation. Likewise, our models are based on the following single-session regression model.

An fMRI dataset consists of the fMRI signal time series $y_{i,t}, t = 1, \dots, T$, at voxels $i = 1, \dots, N$. We assume that the fMRI signal at voxel i and time t , $y_{i,t}$, can be decomposed into a linear combination of unknown functions

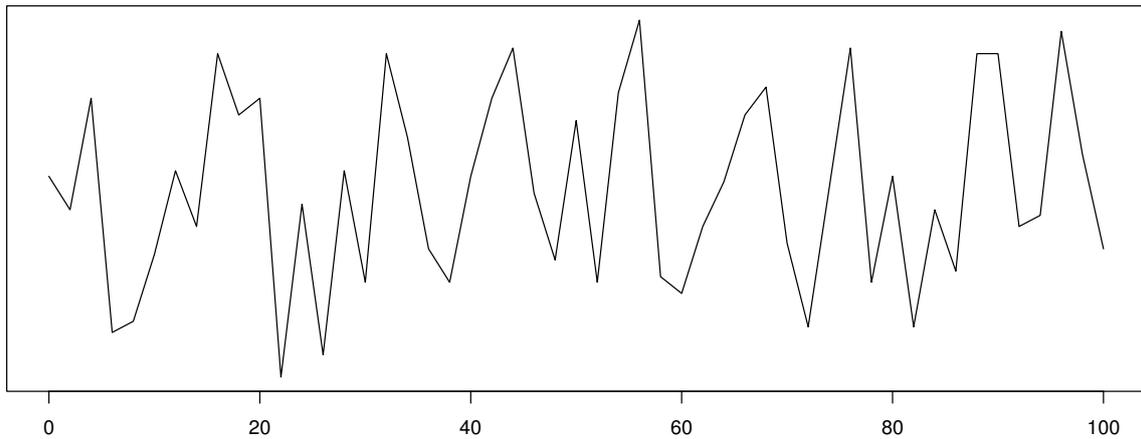
$$y_{i,t} = f_{base}(i, t) + f_{con}(i, t) + f_{stim}(i, t) + \epsilon_{i,t}, \quad (2.1)$$

where

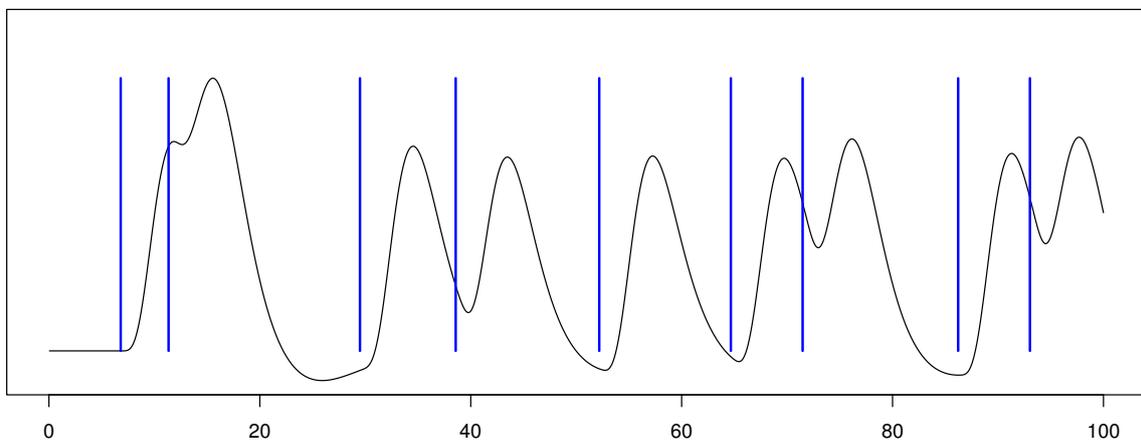
- $f_{base}(i, t)$ is the value of the baseline trend at voxel i and time t ,
- $f_{con}(i, t)$ is the effect of further (confounding) covariables at voxel i and time t ,
- $f_{stim}(i, t)$ is the value of the hemodynamic response to experimental stimuli at voxel i and time t ,

and

$\epsilon_{i,t}$ is the value of a random error term at voxel i and time t .



(a) Observed fMRI signal time series



(b) Stimulus time series (blue stick functions) and expected BOLD response (black line)

Figure 2.1: Observed and expected fMRI signal time series of an activated voxel in an event-related experiment.

The three predictor components can be linearized with respect to unknown parameters so that voxelwise linear models can be obtained. The exact form of each regression component is discussed in depth in the following paragraphs.

Note that, in the following, we distinguish between scan time t and real time t^* : Whereas $t = 1, \dots, T$ enumerates the 3D fMRI images in the order of their recording, real time t^* denotes the time in seconds after recording onset of the first image, i.e. $t^* = TR(t - 1)$. TR is thereby the fixed time of repetition of scans, i.e. the interval between the recording of two images. By relating a whole image to one time point t resp. t^* , we assume that images are slice-time corrected in a preprocessing step. An image slice is thereby a traverse plane of one 3D brain image. For technical reasons, a whole brain volume is not recorded at

once but as a series of measured 2D slices. Slice-time correction procedures are applied to adjust for different slice scanning times.

2.1 Description of predictor components

2.1.1 Baseline - modeling of low frequency confounds

The baseline term $f_{base}(i, t)$ corrects for slow periodic variations and drift inherent to the scanning procedure and serves as a highpass filter. It is modeled as a linear combination of a few, i.e. p_1 , simple basis functions $w_k(t)$, $k = 1, \dots, p_1$, which do not vary over voxels, and voxelspecific weights $\delta_{i,k}$, $k = 1, \dots, p_1$. In the literature, different choices of basis functions $w_k(t)$ can be found. In Smith et al. (2003), Smith and Fahrmeir (2007) and Brezger et al. (2007) they consist of lower frequency terms of a Fourier expansion and/or piecewise continuous polynomials. In SPM, the highpass filter consists of a discrete cosine transform (DCT) set¹ (Friston et al., 2008, p. 123). DCT was introduced by Ahmed et al. (1974) and is defined for scan time points $t = 1, \dots, T$ as

$$w_k(t) = \frac{1}{\sqrt{2}}, \quad k = 1$$

$$w_k(t) = \sqrt{\frac{2}{T}} \cos\left(\frac{\pi(2t-1)(k-1)}{2T}\right), \quad k = 2, \dots, p_1.$$

The DCT basis set constitutes an orthonormal basis of a subspace of a vector space spanned by a Toeplitz matrix, i.e. a diagonal-constant matrix (Strang, 1999). To calculate the number p_1 of DCT basis functions needed for a given highpass cutoff d_{cut} , the argument of the cosine function $\frac{\pi(2t-1)(p_1-1)}{2T}$ at $t = d_{cut}$ can be equated with 2π . This yields $p_1 = \lfloor \frac{4T}{2d_{cut}-1} + 1 \rfloor$. With this value of p_1 , one oscillation of the cosine function is finished approximately by time d_{cut} . Basis function $k = 2, \dots, p_1 - 1$ model lower frequencies. With the constant basis function $w_1(t)$, an intercept term is implicitly contained in the model. Note, if you specify a highpass filter cutoff d_{cut}^* in real time, e.g. the default of 128s, it has to be transformed appropriately to yield a d_{cut} value in scan time: $d_{cut} = d_{cut}^*/TR$. In Figure 2.2 an exemplary set of DCT basis functions is plotted for a highpass filter of 128s against scan time, i.e. one oscillation of the basis function with the highest frequency is finished after 64 scan time points when the repetition time equals 2s. For 200 scans in total, the set consists of 7 basis functions.

¹More precisely, the DCT II set.

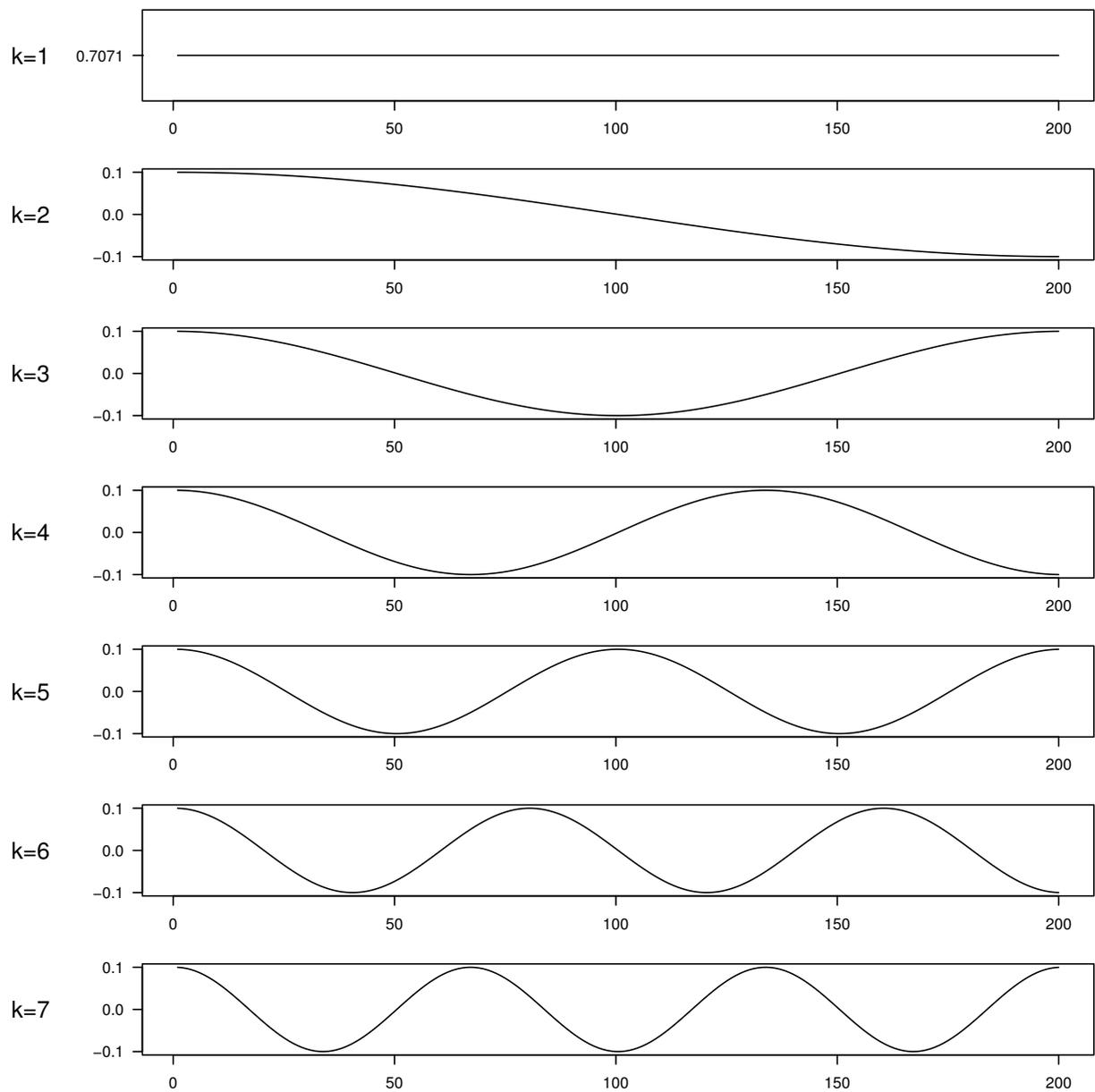


Figure 2.2: DCT basis functions for a highpass filter of 128s in a design with repetition time equal to 2s. The respective basis function values are plotted against scan time.

To calculate the baseline trend at voxel i and time t , the basis functions are evaluated at time t to receive

$$f_{base}(i, t) = \mathbf{w}(t)' \boldsymbol{\delta}_i \quad (2.2)$$

where $\mathbf{w}(t) = (w_1(t), \dots, w_{p_1}(t))'$ and $\boldsymbol{\delta}_i = (\delta_{i,1}, \dots, \delta_{i,p_1})'$.

For our analysis, we model the baseline drift with this DCT basis set. Unlike SPM, we include baseline regressors directly into the model formula, whereas in SPM the highpass filter is applied to the data in a preprocessing step (Frackowiak et al., 2004, p. 123). Both approaches are equivalent, but our choice circumvents a further preprocessing step.

2.1.2 Confounding variables

The second term $f_{con}(i, t)$ accounts for further confounding effects. It is assumed that according information is available in the form of several univariate, global variables with value $c_k(t)$, $k = 1, \dots, p_2$, at time t . This leads to a linearization of $f_{con}(i, t)$ into

$$f_{con}(i, t) = \mathbf{c}(t)' \boldsymbol{\nu}_i \quad (2.3)$$

where $\mathbf{c}(t) = (c_1(t), \dots, c_{p_2}(t))'$ and $\boldsymbol{\nu}_i = (\nu_{i,1}, \dots, \nu_{i,p_2})'$ is the vector of according voxelspecific effects.

Generally, we consider to include the following covariates into the fMRI regression predictor. We include the six rotation and translation parameters from a rigid body transformation (Friston et al., 2008, p. 53) as covariates. In a preprocessing step, fMRI scans are spatially aligned to adjust for head movements. A substantial change in head position, which is reflected in large movement parameters, can distort the fMRI signal. Thus, movement parameters are included to adjust for confounding movement effects.

Additionally, two global signal variables enter the regression. These covariates contain the scanwise fMRI signal means of white matter and liquor voxels, respectively. In contrast to gray matter tissue, white matter tissue and liquor are brain components that cannot contain neuronal activity. A voxel is classified as white matter voxel (resp. liquor voxel) if it survives a 98% probability threshold for being a white matter voxel (resp. liquor voxel) as calculated from a brain segmentation (Friston et al., 2008, Chap. 6). These global covariates seem able to capture effects from cardiac and respiratory cycles, which may confound stimulus effects and are, thus, considered as further covariables.

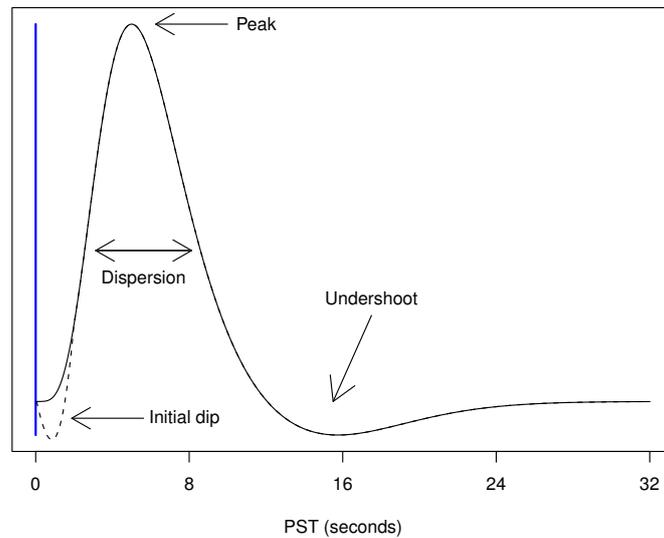


Figure 2.3: Typical (canonical) BOLD response to one impulsive stimulation (blue stick function) plotted with respect to peristimulus time (PST), i.e. time after stimulus onset.

2.1.3 Modeling experimental stimuli

The third term $f_{stim}(i, t)$ includes the transformed signal time series of a given stimulus type. The transformation of the signal time series is necessary for the following reason: If a neuronal response is induced within a brain region, it occurs immediately (within a few hundred milliseconds) after presentation of a stimulus. So there is a close correspondence between stimulus presentation and neuronal activity. The fMRI signal, however, represents aggregated and time delayed neuronal activity. Therefore, the stimulus signal is transformed to the level of the fMRI response to obtain models that are closer to reality and yield a higher degree of explained residual variance.

This mapping is described via the so called hemodynamic response function (HRF), which is based on the typical BOLD response to a single, impulsive stimulation. The canonical BOLD impulse response is depicted in Figure 2.3—in analogy to Friston et al. (2008), p. 178. After a possible initial dip, the response peaks approximately 5 s after stimulation, and is followed by an undershoot that lasts till 32 s after stimulus presentation. This functional form of the response has to be taken into account when calculating stimulus regressor values.

Different proposals exist for transforming the stimulus time series to the level of hemodynamic response. The decision about which modeling approach should be chosen depends

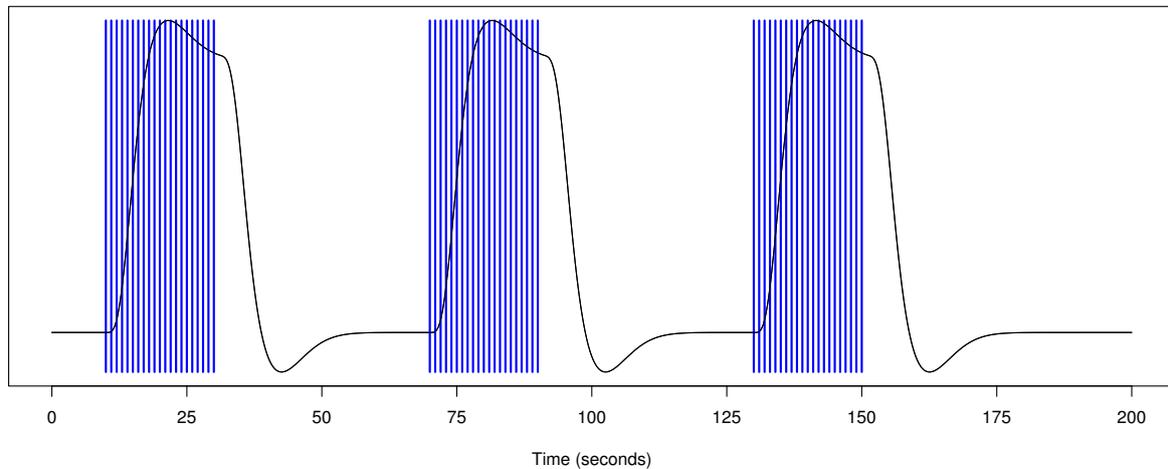


Figure 2.4: Typical BOLD response (black line) in block designs where epochs with stimulus presentations (blue stick functions) alternate with periods of rest.

on several factors, e.g. on the underlying experimental design. In the case of conventional block designs, where periods of rest alternate with stimulus epochs of fixed size, direct approaches exist for estimating the HRF (Genovese, 2000; Gössl et al., 2001b, and references therein). Compare Figure 2.4 for a typical BOLD response in block designs. In direct approaches, the typical stages of the response (e.g. rise, plateau phase, fall, undershoot) are modeled by parameters in combination with a few truncated, informative basis functions.

However, the intended application of analyzing strategies proposed in this thesis are experiments with event-related stimuli. Event-related designs greatly expand the flexibility of the fMRI experiments, since researchers are no longer bound by the constraints of a formal block design (Lazar, 2008). In these studies stimuli are presented individually, separated by a, possibly random, interstimulus interval (ISI). If requested, stimulus presentation can depend on the immediate response of observed subjects, or stimulus types can (randomly) vary from trial to trial. Short ISI of a few seconds lead to an overlap of the BOLD response to single stimuli (compare Figure 2.5), but are comparable to those typically used in behavioral and electrophysiological studies and are generally more efficient from a statistical perspective than studies with long ISI (Friston et al., 2008, p. 178). Hence, the modeling of hemodynamic response has to accommodate these peculiarities.

A more general and even popular approach to HRF modeling is a transformation of the stimulus time series via a convolution model. Applying these models to experiments with event-related stimuli is straightforward, whereas direct approaches become infeasible

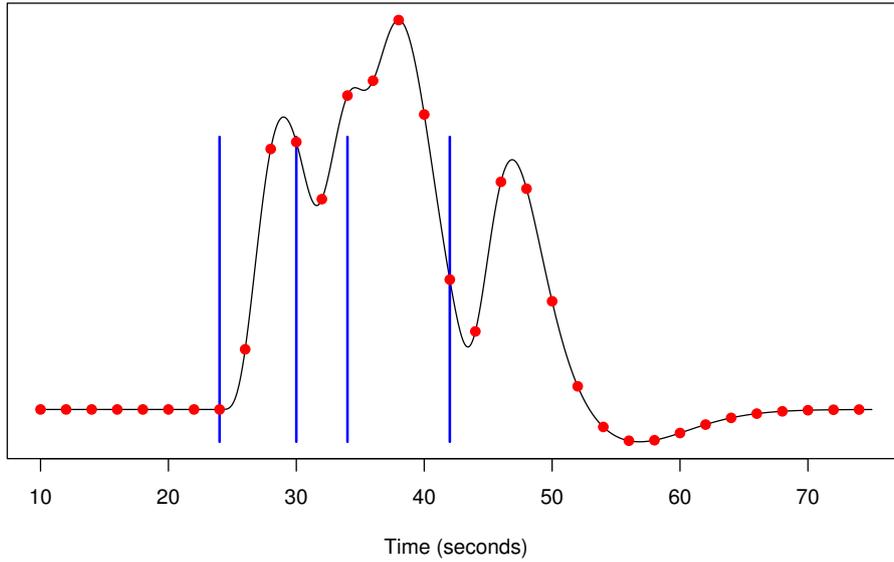


Figure 2.5: Event-related designs with overlapping BOLD response. Stimuli are denoted by blue stick functions, the black function is the modeled BOLD response, red dots denote measurement time points of the fMRI time series.

when stimuli are presented irregularly and are not well separated leading to arbitrarily overlapping responses. Therefore, we use the following convolution model (Josephs et al., 1997; Friston et al., 1998a,b):

$$f_{stim}(i, t) = \int_0^{\tau_{max}} hrf(i, \tau) u(t^* - \tau) \partial\tau$$

where $t^* = TR(t - 1)$. Conforming to the literature, we write down the components of the convolution model in real time notation. The function $u(t^*)$ describes the given time course of neuronal activity and $hrf(i, \tau)$ is the unknown HRF at voxel i . The integral models the neuronal signal aggregation from 0 to τ_{max} time units in the past. Neuronal signals prior to τ_{max} time units from t^* are assumed to not have an effect on the signal at time t^* . The HRF controls the weight of the neuronal signal $u(t^* - \tau)$ at time $t^* - \tau$ that is used to enter the aggregation at voxel i at time t^* .

The exact form of the voxelspecific $hrf(i, \tau)$ is usually unknown. Therefore, we have to estimate it. The approach we follow is a flexible modeling strategy with basis functions B_k

and corresponding voxelspecific weights $\zeta_{i,k}$ (Josephs et al., 1997; Friston et al., 1998a):

$$hrf(i, \tau) = \sum_{k=1}^{p_3} \zeta_{i,k} B_k(\tau).$$

This approach leads to a flexible and data driven estimation of the specific functional form of the hemodynamic response. Usually, a small informed basis set is used making regularization schemes on weight parameters superfluous. However, estimated HRF shapes do not necessarily make sense for all voxels. Therefore, Woolrich et al. (2004a) developed constraints on effect parameters resulting in sensible HRF shapes. For the Bayesian variable selection approach, which we intend to develop, it is not clear if it is feasible to include these constraints. For this, we choose to use an unconstrained estimation scheme.

Different choices of basis sets exist (Henson et al., 2001). We focus on the canonical basis function set (Friston et al., 2008, pp. 181) and a gamma basis function set as proposed in Friston et al. (1998b).

The gamma basis function set consists of gamma density functions (and possibly their time derivatives):

$$f_{a,s}(\tau) = \frac{1}{s^a \Gamma(a)} \tau^{(a-1)} e^{-(\tau/s)}.$$

We prefer a gamma basis function set of 3 gamma density functions with shape parameter $a \in \{4, 8, 16\}$ and scale parameter $s = 1$ as used in Friston et al. (1998b), i.e. $B_1(\tau) = f_{4,1}(\tau)$, $B_2(\tau) = f_{8,1}(\tau)$ and $B_3(\tau) = f_{16,1}(\tau)$.

For the canonical set, we follow the SPM implementation. In SPM, the canonical basis function set consists of maximal 3 basis functions including the canonical HRF, i.e. a sum of two weighted gamma densities, an approximation of its time derivative and of the—so called—dispersion derivative. This set is motivated by a first order Taylor expansion allowing for small variations in the form of the canonical HRF with respect to its onset and dispersion (Friston et al., 2008, pp. 181). In SPM, computation of derivatives are approximated by differential quotients. From the SPM implementation, the canonical HRF is defined as

$$B_1(\tau) = \frac{(\Delta_\tau/\rho_3)^{\rho_1/\rho_3}}{\Gamma(\rho_1/\rho_3)} (\tau/\Delta_\tau - \rho_6)^{\rho_1/\rho_3-1} \exp\left(\frac{-\Delta_\tau}{\rho_3} (\tau/\Delta_\tau - \rho_6)\right) - \frac{1}{\rho_5} \frac{(\Delta_\tau/\rho_4)^{\rho_2/\rho_4}}{\Gamma(\rho_2/\rho_4)} (\tau/\Delta_\tau - \rho_6)^{\rho_2/\rho_4-1} \exp\left(\frac{-\Delta_\tau}{\rho_4} (\tau/\Delta_\tau - \rho_6)\right)$$

where parameter ρ_1 is the delay of the response (relative to onset, default $\rho_1 = 6s$), ρ_2 is the delay of the undershoot (relative to onset, default $\rho_2 = 16s$), ρ_3 is the dispersion of the response (default $\rho_3 = 1s$), ρ_4 is the dispersion of the undershoot (default $\rho_4 = 1s$), ρ_5 is the ratio of the peak response to the peak undershoot (default $\rho_5 = 6$), ρ_6 is an offset for the onset (default $\rho_6 = 0$). The parameter Δ_τ controls the window length, i.e. the time it takes for the HRF response to return to baseline (default $\Delta_\tau = 0.125$ gives a window length of about 32 s). SPM uses the Taylor approximation to adjust for variability in ρ_3 (dispersion) and ρ_6 (time). Instead of using an approximation via differential quotients, we use the exact derivatives in our implementation. For the time derivative define a gamma function $g(\tau; \rho_a, \rho_b, \Delta_\tau, \rho_6)$ as

$$g(\tau; \rho_a, \rho_b, \Delta_\tau, \rho_6) = \frac{(\Delta_\tau/\rho_b)^{\rho_a/\rho_b}}{\Gamma(\rho_a/\rho_b)} (\tau/\Delta_\tau - \rho_6)^{\rho_a/\rho_b - 1} \exp\left(\frac{-\Delta_\tau}{\rho_b} (\tau/\Delta_\tau - \rho_6)\right).$$

Its derivative with respect to ρ_6 is

$$\frac{\partial g(\tau; \rho_a, \rho_b, \Delta_\tau, \rho_6)}{\partial \rho_6} = g(\tau; \rho_a, \rho_b, \Delta_\tau, \rho_6) \frac{1}{(\tau/\Delta_\tau - \rho_6)} \left[\frac{\Delta_\tau}{\rho_b} (\tau/\Delta_\tau - \rho_6) - \frac{\rho_a}{\rho_b} + 1 \right].$$

Hence, the time derivative can be calculated as

$$B_2(\tau) = \frac{\partial B_1(\tau)}{\partial \rho_6} = \frac{\partial g(\tau; \rho_1, \rho_3, \Delta_\tau, \rho_6)}{\partial \rho_6} - \frac{1}{\rho_5} \frac{\partial g(\tau; \rho_2, \rho_4, \Delta_\tau, \rho_6)}{\partial \rho_6}.$$

The dispersion derivative can be derived as

$$\begin{aligned} B_3(\tau) &= \frac{\partial B_1(\tau)}{\partial \rho_3} \\ &= -g(\tau; \rho_1, \rho_3, \Delta_\tau, \rho_6) \\ &\quad \times \frac{1}{\rho_3^2} \left[\rho_1 \left(1 + \log\left(\frac{\Delta_\tau}{\rho_3}\right) - \Psi\left(\frac{\rho_1}{\rho_3}\right) + \log(\tau/\Delta_\tau - \rho_6) \right) - \Delta_\tau (\tau/\Delta_\tau - \rho_6) \right], \end{aligned}$$

where $\Psi(x)$ is the Digamma or Psi function, $\Psi(x) = \frac{\partial}{\partial x} \log \Gamma(x)$.

Within the estimation procedure in SPM all three basis functions are orthogonalized with respect to each other via the Gram-Schmidt process. Due to the fact that regression analysis accounts for non-orthogonality we choose to use them as defined above. Additionally, in contrast to our implementation, SPM normalizes basis functions to sum 1. A visualization of the two alternative basis function sets we use can be found in Figure 2.6.

The time series of neuronal activity $u(t^*)$ is set equal to the signal time series, which is mod-

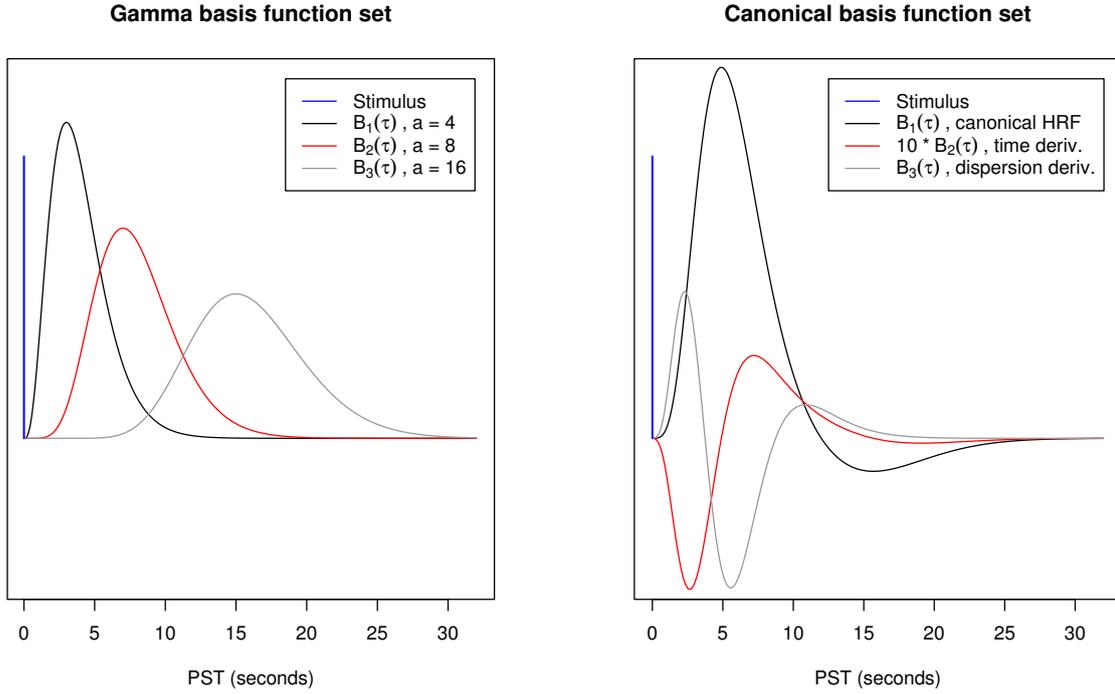


Figure 2.6: Gamma and canonical basis function sets plotted with respect to peristimulus time (PST), i.e. time after stimulus onset. For the 3 gamma basis functions, shape parameters $a=4, 8, 16$ have been used. The second canonical HRF basis function, the time derivative, is multiplied by 10 for better visualization.

eled as follows: Suppose we have event-related stimuli of one type at times $\tau_1, \tau_2, \dots, \tau_M$. A stimulus at time τ_m is modeled via a dirac delta function $\delta(t^* - \tau_m)$, i.e. a stick-function, so that $u(t^*) = \sum_{m=1}^M \delta(t^* - \tau_m)$. Thus, the stimulus predictor for all presented stimuli simplifies as follows

$$\begin{aligned}
 f_{stim}(i, t) &= \sum_{k=1}^{p_3} \zeta_{i,k} \int_0^{\tau_{max}} B_k(\tau) u(t^* - \tau) \partial\tau \\
 &= \sum_{k=1}^{p_3} \zeta_{i,k} \sum_{m=1}^M \int_0^{\tau_{max}} B_k(\tau) \delta(t^* - \tau_m - \tau) \partial\tau \\
 &\approx \sum_{k=1}^{p_3} \zeta_{i,k} \underbrace{\sum_{m=1}^M B_k(t^* - \tau_m)}_{=z_k(t)} \\
 &= \sum_{k=1}^{p_3} \zeta_{i,k} z_k(t)
 \end{aligned}$$

$$= \mathbf{z}(t)\boldsymbol{\zeta}_i, \quad (2.4)$$

where $\mathbf{z}(t) = (z_1(t), \dots, z_{p_3}(t))'$ and $\boldsymbol{\zeta}_i = (\zeta_{i,1}, \dots, \zeta_{i,p_3})'$. The approximation in the third step is dependent on the appropriate choice of basis functions with property $B_k(t^* - \tau_m) \approx 0$ for $(t^* - \tau_m) > \tau_{max}$ and $B_k(t^* - \tau_m) = 0$ for $(t^* - \tau_m) < 0$, which holds for the selected basis function sets. For other basis sets, formulas must slightly be adapted to incorporate the integration limits of 0 to τ_{max} .

In Figure 2.7, the estimation process is visualized via a toy example with well separated stimuli times. An estimation procedure yields weight estimates $\hat{\zeta}_{i,k}$, $k = 1, 2, 3$. Each basis function is multiplied with the according weight and the final HRF estimate is formed by the summed weighted basis functions.

Although having the same theoretical background, our implementation differs from the one in SPM: While we have decided to realize the proposed continuous convolution model, SPM implements a discretized version (Friston et al., 2008, pp. 119). For this, they use both a discrete binary stimulus time series in user-provided fine resolution, e.g. 16 time bins per fMRI scan, and a discrete linear convolution model. We do not expect larger differences between results of both implementations for the BOLD response occupying a rather low-frequency band. As noted in Friston et al. (2008), pp. 119, responses to trials a few hundred milliseconds apart are virtually indistinguishable.

Note, so far $f_{stim}(i, t)$ is defined for one stimulus type presented at times $\tau_1, \tau_2, \dots, \tau_M$. This approach is easily extendable to several stimulus types: For each stimulus type s presented at times $\tau_1^{(s)}, \tau_2^{(s)}, \dots, \tau_M^{(s)}$ a separate predictor $f_{stim}^{(s)}(t)$ is defined as in (2.4). The final predictor is then formed additively by $f_{stim}(t) = \sum_{s=1}^S f_{stim}^{(s)}(t)$ where S is the total number of different stimulus types.

2.1.4 Linearization of all model components

As has been shown in the preceding paragraphs, each model component in (2.1) can be linearized with respect to unknown parameters. Using (2.2)–(2.4), the regression model (2.1) can be rewritten as

$$y_{i,t} = \mathbf{w}(t)'\boldsymbol{\delta}_i + \mathbf{c}(t)'\boldsymbol{\nu}_i + \mathbf{z}(t)'\boldsymbol{\zeta}_i + \epsilon_{i,t}. \quad (2.5)$$

Thus, for inference, we can revert to the vast literature on linear regression models.

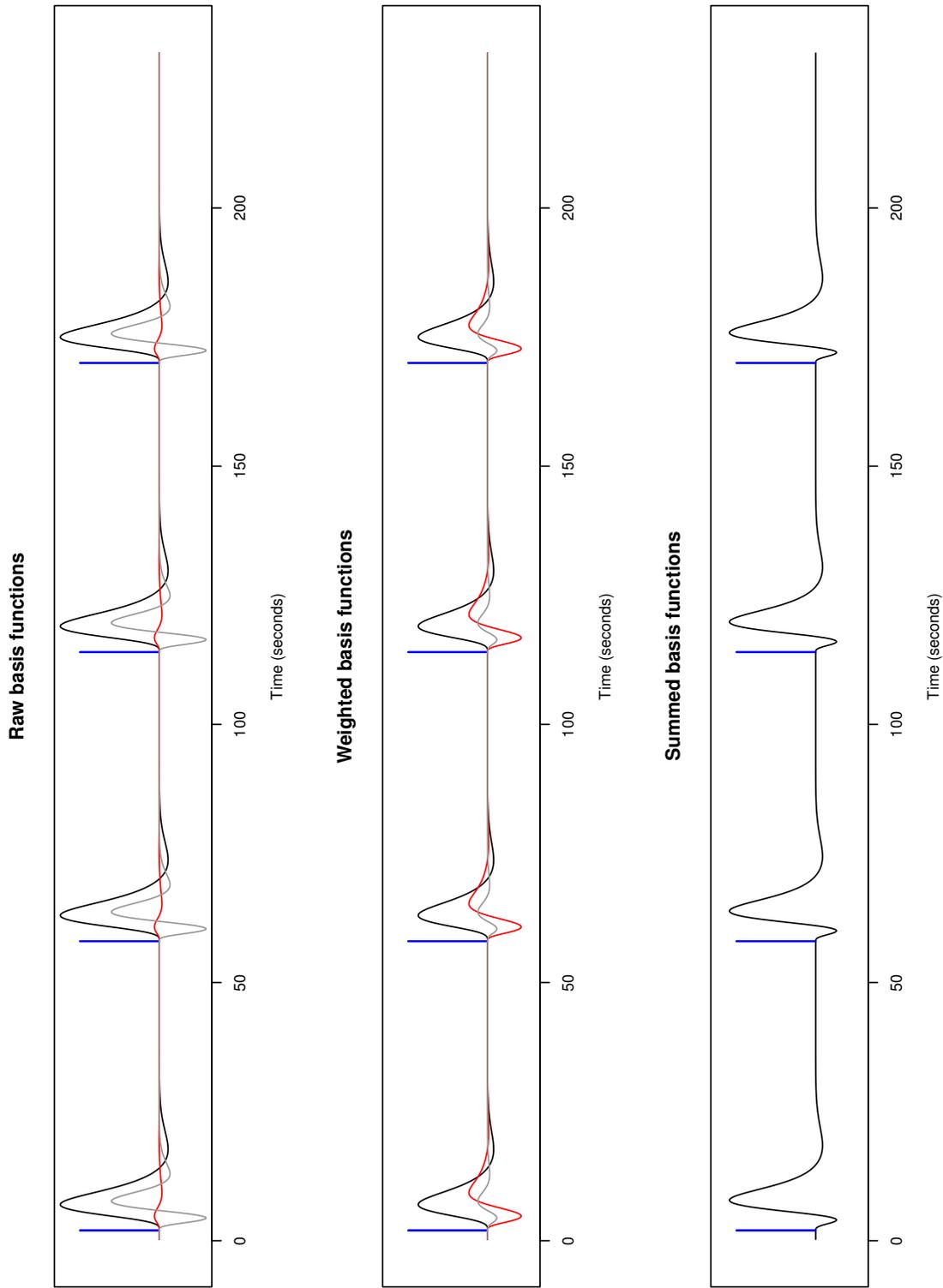


Figure 2.7: Demonstration of the voxel-specific HRF estimation process of the basis functions approach. In the topmost plot, the raw basis functions of the canonical set, which follow a stimulus (blue stick), are depicted. Via an estimation procedure estimates for the weights $\zeta_{i,k}$ are calculated. The correspondingly weighted basis functions are plotted in the middle, whereas the HRF estimation, i.e. the summed, weighted basis functions, are shown in the lower part of the figure.

2.1.5 Matrix notation

Let \mathbf{y}_i be the vector $\mathbf{y}_i = (y_{i,1}, \dots, y_{i,T})'$ of the fMRI signal time series of length T at voxel i . Accordingly, $\boldsymbol{\epsilon}_i$ is defined as the vector of measurement noise $\boldsymbol{\epsilon}_i = (\epsilon_{i,1}, \dots, \epsilon_{i,T})'$ at voxel i . Then regression model (2.5) can be rewritten as

$$\mathbf{y}_i = \mathbf{W}\boldsymbol{\delta}_i + \mathbf{C}\boldsymbol{\nu}_i + \mathbf{Z}\boldsymbol{\zeta}_i + \boldsymbol{\epsilon}_i$$

where

$$\mathbf{W}_{T \times p_1} = \begin{pmatrix} \mathbf{w}(1)' \\ \vdots \\ \mathbf{w}(T)' \end{pmatrix}, \quad \mathbf{C}_{T \times p_2} = \begin{pmatrix} \mathbf{c}(1)' \\ \vdots \\ \mathbf{c}(T)' \end{pmatrix}, \quad \mathbf{Z}_{T \times p_3} = \begin{pmatrix} \mathbf{z}(1)' \\ \vdots \\ \mathbf{z}(T)' \end{pmatrix}.$$

We define the overall $T \times p$ design matrix \mathbf{X} as $\mathbf{X} = [\mathbf{W}|\mathbf{C}|\mathbf{Z}]$ with $p = p_1 + p_2 + p_3$ and an according parameter vector of unknown weights and effects $\boldsymbol{\beta}_i = (\boldsymbol{\delta}'_i, \boldsymbol{\nu}'_i, \boldsymbol{\zeta}'_i)'$. Our further modeling approach is then based on the following voxelwise model (Friston et al., 1995)

$$\mathbf{y}_i = \mathbf{X}\boldsymbol{\beta}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, N.$$

Note that the design matrix does not vary over voxels as in Smith et al. (2003) and Smith and Fahrmeir (2007) where voxel specific HRF estimates and according transformed stimulus regressors were obtained in a preprocessing step to form a voxel dependent design matrix \mathbf{X}_i . Instead, with a basis function approach, differences in hemodynamic response between voxels are now subsumed in the estimation of $\boldsymbol{\beta}_i$, which are already voxel dependent. Despite being formulated for a constant design matrix \mathbf{X} , our modeling approach proposed in Chapter 3 can directly be reformulated for a voxel dependent \mathbf{X}_i .

2.2 Distributional assumptions about the error term

For fMRI being measured on a metric scale, it is assumed that $\boldsymbol{\epsilon}_i$ follows a multivariate normal distribution

$$\boldsymbol{\epsilon}_i \sim \mathcal{N}(0, \sigma_i^2 \mathbf{V})$$

with voxel dependent variance σ_i^2 and correlation matrix \mathbf{V} .

Usually fMRI data exhibit short-range serial or temporal correlations (Friston et al., 2008, p. 121). Long-range correlations are assumed to be removed by the applied highpass filter, cf. Section 2.1.1. If ignoring serial correlations, inappropriate estimates for degrees of freedom are obtained, which enter T - or F -statistics in classical inference approaches. Hence, \mathbf{V} is typically chosen to model serial correlations. A popular choice of \mathbf{V} consists of a $AR(1)$ plus white noise process. Classical estimation for \mathbf{V} can be accomplished by ReML procedures. We refer to Friston et al. (2008) for a detailed discussion of modeling serial correlations in fMRI.

If the baseline captures small-range trend changes as in Gössl et al. (2000, 2001a), an uncorrelated error structure seems adequate. Within their Bayesian modeling framework, the authors have observed that temporal trends are sufficiently captured by their time-varying trend effect, confirming the assumption of white error noise. In their experience no gain in precision is obtained by considering correlated error terms. Note, however, that the application of this small-scale trend component is restricted to the analysis of block experiments. In event-related designs, this trend component might confound stimulus effects.

As a starting point, we derive the model proposed in Chapter 3 on the assumption of independent error terms:

$$\epsilon_i \sim \mathcal{N}_T(\mathbf{0}, \sigma_i^2 \mathbf{I}),$$

where \mathbf{I} is the identity matrix of size T . Though, our chosen highpass filter removes only long-range frequencies, the derivation of a combined fMRI-EEG model is substantially complicated when adding serial correlations. For testing whether the model exhibits a sufficient usefulness, we revert to independent noise.

3 An EEG-enhanced fMRI model

The main focus of an fMRI study lies on the identification of activated voxels. Thereby, a voxel is said to be activated if its fMRI time series correlates with the transformed stimulus time series. In the classical fMRI regression framework, activation detection is implemented by voxelwise statistical hypothesis tests evaluating whether stimulus effect parameters equal zero (see for example Friston et al., 1994, 2002). Applying a significance threshold (derived by spatial multiple test procedures) to corresponding T- or F-maps results in a so-called statistical parametric map (SPM). In analogy to this, activation maps of Bayesian models are derived by thresholding posterior probability maps (PPMs) of effect parameters (see for example Gössl et al., 2000; Friston et al., 2002; Woolrich et al., 2004b). However, neither the classical approach nor the mentioned Bayesian approach constitutes a direct measure of activation. In the classical approach, a voxel is declared to be active when a corresponding test statistic is larger than a critical value. This critical value represents a limit for test statistic values that are acceptable under the null hypothesis of the true effect being zero. PPMs constitute a more direct measure for being defined as the posterior probability that the effect is larger than some threshold. The threshold, though, rests upon a definition from which effect size on a voxel can be declared activated—which can be quite arbitrary.

To overcome these problems, Smith et al. (2003) resp. Smith and Fahrmeir (2007) proposed to use a Bayesian activation detection scheme estimating voxelwise probabilities of activation directly. For this, they applied the theory of Bayesian variable selection (Smith and Kohn, 1996; George and McCulloch, 1997) introducing a binary variable being 1 if the stimulus regressor is selected, i.e. if there is a relationship between stimulus and signal time series, and 0 otherwise. Estimation of these binary indicators was spatially regularized by an Ising prior (Hurn et al., 2003). The posterior probabilities of voxelwise indicators being 1 constitute the posterior activation probabilities.

For our combined fMRI and EEG model, we choose to extend their approach to incorporate EEG information in an fMRI regression model. For this, we exchange the Ising prior with

a prior based on a binary regression model. EEG information is included as predictor with an—optionally—spatially-varying coefficient allowing for adaption to local brain response. Estimation of spatially-varying coefficients is regularized by Gaussian Markov random fields (GMRF). We consider two alternative regularization schemes either based on an intrinsic GMRF (Lang and Brezger, 2004; Rue and Held, 2005) or on a Gaussian conditional autoregression (CAR) (Weir and Pettitt, 2000; Pettitt et al., 2002; Smith and Smith, 2006). With this spatial probit model, a strong EEG effect at voxel i should increase its activation probability.

This chapter is organized as follows. After briefly reviewing our choice for the fMRI regression model for readers who skipped Chapter 2, we introduce the activation detection scheme based on Bayesian variable selection in Section 3.2. Prior specifications for all parameters, including the priors for incorporating EEG information in a probit hierarchy, are discussed in Section 3.3. Posterior inference based on a Markov Chain Monte Carlo (MCMC) scheme is derived in Section 3.4. A minor model extension is described in Section 3.5.

3.1 The fMRI regression model

The design of the predictor we use for modeling the fMRI signal is discussed in depth in Chapter 2. For ease of reading, we shortly summarize its form.

Let \mathbf{y}_i denote the vector $\mathbf{y}_i = (y_{i,t}, t = 1, \dots, T)'$ of the fMRI signal time series at voxel i , $i = 1, \dots, N$. We include three predictor components which can be linearized with respect to unknown effect parameters. Hence, the following multiple regression model results

$$\mathbf{y}_i = \mathbf{W}\boldsymbol{\delta}_i + \mathbf{C}\boldsymbol{\nu}_i + \mathbf{Z}\boldsymbol{\zeta}_i + \boldsymbol{\epsilon}_i,$$

where \mathbf{W} is the $T \times p_1$ design matrix for the baseline trend, \mathbf{C} is the $T \times p_2$ design matrix of covariates and \mathbf{Z} is the $T \times p_3$ design matrix for modeling the hemodynamic response to stimuli. The vector $\boldsymbol{\epsilon}_i$ is the vector of random errors $\boldsymbol{\epsilon}_i = (\epsilon_{i,t}, t = 1, \dots, T)'$. Note that all three design matrices do not vary over voxels, whereas voxel dependent effect estimates allow for adaption to local brain response.

Simplifying the notation, we combine all linear regression parts into model

$$\mathbf{y}_i = \mathbf{X}\boldsymbol{\beta}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, N, \quad (3.1)$$

where $\mathbf{X} = [\mathbf{W}|\mathbf{C}|\mathbf{Z}]$ is the overall $T \times p$ design matrix (with $p = p_1 + p_2 + p_3$) and $\boldsymbol{\beta}_i = (\boldsymbol{\delta}'_i, \boldsymbol{\nu}'_i, \boldsymbol{\zeta}'_i)'$ is an according parameter vector of unknown weights and effects.

The likelihood of our model depends on the assumption $\boldsymbol{\epsilon}_i \sim \mathcal{N}_T(\mathbf{0}, \sigma_i^2 \mathbf{I})$, where \mathbf{I} is the identity matrix of size T .

3.2 Bayesian activation detection

We define a spatial Bayesian model for variable selection by introducing binary indicator variables $\boldsymbol{\gamma}_i = (\gamma_{i,1}, \dots, \gamma_{i,p})'$, such that $\beta_{i,j} = 0$ if $\gamma_{i,j} = 0$ and $\beta_{i,j} \neq 0$ if $\gamma_{i,j} = 1$ (Smith et al., 2003; Smith and Fahrmeir, 2007). Given $\boldsymbol{\gamma}_i$, let $\boldsymbol{\beta}_i(\boldsymbol{\gamma}_i)$ be the vector of nonzero regression coefficients from regression i and let $\mathbf{X}(\boldsymbol{\gamma}_i)$ be the corresponding design matrix; then the regression model (3.1) can be rewritten as

$$\mathbf{y}_i = \mathbf{X}(\boldsymbol{\gamma}_i)\boldsymbol{\beta}_i(\boldsymbol{\gamma}_i) + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, N.$$

The p -dimensional vector $\boldsymbol{\gamma}_i$ implements a variable selection scheme for all p regressors within model equation (3.1), including the baseline trend and further covariates. But, primary interest lies in the selection of stimulus regressors forming $f_{stim}(i, t)$. In this case, it may prove sufficient to confine variable selection to the detection of activated voxels. Consider the special case of modeling the response of one stimulus type with p_3 basis functions in (2.4). Then, $\boldsymbol{\gamma}_i$ can for example be restricted to the form

$$\boldsymbol{\gamma}_i = \left(\underbrace{1, \dots, 1}_{p_1+p_2\text{-times}}, \underbrace{\gamma_i, \dots, \gamma_i}_{p_3\text{-times}} \right)'. \quad (3.2)$$

With $\gamma_{i,1} = \dots = \gamma_{i,p_1+p_2} = 1$, baseline and covariate regressors are always kept in the model. The scalar binary parameter γ_i controls the simultaneous selection of all p_3 basis functions for the stimulus component. This leads us to the following interpretation:

$$\gamma_i = \begin{cases} 1 & \text{if voxel } i \text{ is activated,} \\ 0 & \text{if voxel } i \text{ is not activated.} \end{cases}$$

The vector of scalar binary indicators $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_N)'$ then represents an activation surface. Per definition, it holds for the hemodynamic response that

$$f_{stim}(i, t) \neq 0 \text{ iff } \gamma_i = 1,$$

$$f_{stim}(i, t) = 0 \text{ iff } \gamma_i = 0,$$

or equivalently

$$\zeta_{i,1}, \dots, \zeta_{i,p_3} \neq 0 \text{ iff } \gamma_i = 1,$$

$$\zeta_{i,1}, \dots, \zeta_{i,p_3} = 0 \text{ iff } \gamma_i = 0.$$

3.3 Prior specifications

A Bayesian model requires the specification of prior distributions for all parameters in the model. In a hierarchical setup, these priors can be chosen to depend on further parameter constellations imposing specific structural assumptions. We use the following priors.

From the definition of the variable selection scheme, $\beta_{i,j} = 0 | \gamma_{i,j} = 0$, a prior is only required for the nonzero coefficients $\beta_i(\gamma_i)$. This prior, however, has to be proper. If improper priors are placed on these coefficients, they will be estimated as 0 (Mitchell and Beauchamp, 1988; Smith and Kohn, 1996; Kohn et al., 2001). Following the literature (see Smith et al. (2003) and Smith and Fahrmeir (2007) for applications in the context of fMRI analysis), we select a proper conditional prior for nonzero coefficients by setting it proportional to a fraction of the likelihood, $p(\beta_i(\gamma_i) | \mathbf{y}_i, \sigma_i^2, \gamma_i) \propto p(\mathbf{y}_i | \sigma_i^2, \gamma_i, \beta_i(\gamma_i))^{1/T}$, such that

$$\beta_i(\gamma_i) | \mathbf{y}_i, \sigma_i^2, \gamma_i \sim \mathcal{N} \left(\hat{\beta}_i(\gamma_i), T \sigma_i^2 (\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1} \right), \quad (3.3)$$

where $\hat{\beta}_i(\gamma_i) = (\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1} \mathbf{X}(\gamma_i)' \mathbf{y}_i$. Being a more dispersed version of the likelihood, this prior avoids the problems of employing improper priors without using additional prior information where it does not exist. This prior is related to the g -prior of Zellner (1986) and can be motivated by fitting a linear model based on an “imaginary” sample prior to the actual analysis. It is attractive because it is invariant to location changes and possesses automatic rescaling properties (Kohn et al., 2001). The multiplication factor T of the covariance matrix is chosen to keep $T(\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1}$ approximately constant as the number of scans T increases. The presence of σ_i^2 in the covariance of this prior makes this prior an element of the family of conjugate hierarchical setups (George and McCulloch, 1997). As opposed to non-conjugate priors as applied in George and McCulloch (1993), $\beta_i, i = 1, \dots, N$, and $\sigma_i^2, i = 1, \dots, N$, can be eliminated by routine integration from the full

posterior to yield a marginalized posterior for γ_i . As will be seen in subsequent sections, this feature yields fast and efficient computational estimation routines.

For the variance parameter $\boldsymbol{\sigma}^2 = (\sigma_1^2, \dots, \sigma_N^2)'$, we assume standard independent noninformative priors. That is,

$$p(\boldsymbol{\sigma}^2) = \prod_{i=1}^N p(\sigma_i^2) \propto \prod_{i=1}^N \frac{1}{\sigma_i^2}. \quad (3.4)$$

Following Smith and Fahrmeir (2007), the prior of the binary indicator variables vector $\boldsymbol{\gamma} = (\gamma_{i,j}, i = 1, \dots, N, j = 1, \dots, p)'$ factorizes over regressors. Assume that $\boldsymbol{\gamma}_{(j)} = (\gamma_{1,j}, \dots, \gamma_{N,j})'$, then $p(\boldsymbol{\gamma}) = \prod_{j=1}^p p(\boldsymbol{\gamma}_{(j)})$. Spatial dependencies are then incorporated by an appropriate choice of $p(\boldsymbol{\gamma}_{(j)})$. In Smith et al. (2003) and Smith and Fahrmeir (2007) a binary spatial Markov random field (MRF) prior in form of an Ising prior was used. In this thesis, however, we consider an alternative approach where variable selection, i.e. $p(\gamma_{i,j} = 1)$, is based on additional continuous information like a variable with EEG measurements.

In the following, we restrain our derivations to the special case proposed in Section 3.2 where variable selection is confined to the simultaneous selection of p_3 basis functions forming $f_{stim}(i, j)$, i.e. $\boldsymbol{\gamma}_i = (1, \dots, 1, \gamma_i, \dots, \gamma_i)'$ as in (3.2) with γ_i being a scalar binary indicator. Thus, only a prior for the reduced vector $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_N)'$ must be specified. The reason for this restriction is that external information in form of EEG data can serve as an indicator for activation, but not for baseline trend or covariate selection. For simplicity, no variable selection is applied to these terms. But if variable selection is requested for all regressors, it seems conceivable that Ising priors can be used for the non-stimulus regressors as noted above.

Spatial correlations and prior information in form of voxelwise, scalar EEG information $J_i, i = 1, \dots, N$ is introduced through priors motivated by binary regression models. Conditional on a further parameter vector $\boldsymbol{\theta}$, the binary indicators γ_i follow conditional independent Bernoulli distributions, so that

$$p(\boldsymbol{\gamma}|\boldsymbol{\theta}) = \prod_{i=1}^N p(\gamma_i|\boldsymbol{\theta}) = \prod_{i=1}^N \pi_i(\boldsymbol{\theta})^{\gamma_i} (1 - \pi_i(\boldsymbol{\theta}))^{1-\gamma_i} \quad (3.5)$$

where $\pi_i(\boldsymbol{\theta}) = p(\gamma_i = 1|\boldsymbol{\theta})$. The parameter vector $\boldsymbol{\theta}$ contains further parameters in lower hierarchical levels of our model. Following standard theory for generalized linear models, the parameter $\pi_i(\boldsymbol{\theta})$, which coincides with the expectation of a single Bernoulli distribution,

is linked to a predictor η_i by a strictly monotonic increasing function h with $h(\eta_i) \in [0, 1]$. In our application, we use probit transformations as link between predictor and $\pi_i(\boldsymbol{\theta})$:

$$\pi_i(\boldsymbol{\theta}) = p(\gamma_i = 1 | \eta_i) = \Phi(\eta_i), \quad i = 1, \dots, N.$$

We consider additive predictors of the form

$$\eta_i = f_0(i) + f(J_i), \quad i = 1, \dots, N,$$

where $J_i \in \mathbb{R}^+$ are the voxelwise EEG measurements. Here, $f_0(i)$ is an intercept term that is assumed to consist of either a global effect $\alpha_{0,G}$ or a local effect $\alpha_{0,i}$ or of a combination of both. If requested, $f_0(i)$ can also be removed from the model. That is, $f_0(i)$ can be either one of

$$f_0(i) = 0$$

$$f_0(i) = \alpha_{0,G} \tag{3.6}$$

$$f_0(i) = \alpha_{0,i} \tag{3.7}$$

$$f_0(i) = \alpha_{0,G} + \alpha_{0,i}. \tag{3.8}$$

Whereas the global effect $\alpha_{0,G}$ models the basic inclination of the brain being activated aggregated over all brain regions, the local effect $\alpha_{0,i}, i = 1, \dots, N$ is a spatially-varying intercept adapting to regional differences.

The term $f(J_i)$ of the predictor models the EEG contribution to activation. Like $f_0(i)$, it consists of either only a global effect α_G or a local effect α_i or of a combination of both. If requested, $f(J_i)$ can also be removed from the model formula. That is, $f(J_i)$ can be either one of

$$f(J_i) = 0$$

$$f(J_i) = \alpha_G J_i \tag{3.9}$$

$$f(J_i) = \alpha_i J_i \tag{3.10}$$

$$f(J_i) = \alpha_G J_i + \alpha_i J_i. \tag{3.11}$$

The global EEG effect α_G is an aggregated effect over all brain voxels, while the local effect $\alpha_i, i = 1, \dots, N$ is a spatially-varying EEG effect modeling areal differences in the response to the EEG signal. Inclusion of this flexible structure seems necessary for adapting to brain regions where either EEG measurements contradict or support fMRI evidence. For

	IGMRF			CAR		
	$f_0(i)$	+	$f(J_i)$	$f_0(i)$	+	$f(J_i)$
1	$\alpha_{0,G}$	+	$\alpha_G J_i + \alpha_i J_i$	$\alpha_{0,G}$	+	$\alpha_i J_i$
2	$\alpha_{0,G} + \alpha_{0,i}$	+	$\alpha_G J_i$	$\alpha_{0,i}$	+	$\alpha_G J_i$
3	$\alpha_{0,G} + \alpha_{0,i}$			$\alpha_{0,i}$		
4			$\alpha_G J_i + \alpha_i J_i$			$\alpha_i J_i$
5	$\alpha_{0,G} + \alpha_{0,i}$	+	$\alpha_G J_i + \alpha_i J_i$	$\alpha_{0,i}$	+	$\alpha_i J_i$

Table 3.1: Predictor forms of interest.

instance, small values of the local effect α_i can downweight the influence of a large J_i -value in parts of the brain where fMRI activation is entirely absent. Note that, generally, larger values of J_i are considered to be indicative for activation. For a start, we do not introduce restrictions on $\alpha_i, i = 1, \dots, n$ and α_G , and evaluate whether and which restrictions are sensible after first analyses.

If a local effect is requested, one parameter per voxel is included in the model, i.e. $\alpha_0 = (\alpha_{0,1}, \dots, \alpha_{0,N})'$ for the spatially-varying intercept and $\alpha = (\alpha_1, \dots, \alpha_N)'$ for the spatially-varying EEG coefficient. In this case, the number of parameters for each effect equals the number of latent observations $\gamma_1, \dots, \gamma_N$. Hence, estimation of each local effect has to be regularized by a spatial prior. Suitable choices of priors impose an estimable degree of smoothness reducing the number of effective parameters. We use either an intrinsic Gaussian Markov random field (IGMRF)(Lang and Brezger, 2004; Rue and Held, 2005) or a Gaussian conditional autoregressive (CAR)(Weir and Pettitt, 2000; Pettitt et al., 2002; Smith and Smith, 2006) prior. The choice of spatial prior has an effect on the choice of predictor. At first sight the global effects in linear predictors (3.8) and (3.11) seem to be redundant for being automatically included in the local effect. This is true in models with CAR prior. However, as can be seen below, it has to be included in models with IGMRF prior.

Summarizing, all predictor forms of interest are listed in Table 3.1. We assume that at least one of the two components $f_0(i)$ or $f(J_i)$ contains a spatially-varying coefficient for adapting to local brain response. If an IGMRF prior is used for a local effect, the according global effect is also included in the model. For the CAR model, it suffices to include either a global or a local effect.

In case of an intrinsic Gaussian random field, the following priors are attached to α_0 and α :

$$p(\alpha_0 | \xi_0^2) \propto (\xi_0^2)^{-(N-1)/2} \exp\left(-\frac{1}{2\xi_0^2} \alpha_0' \mathbf{Q} \alpha_0\right) \quad (3.12)$$

$$p(\alpha|\xi^2) \propto (\xi^2)^{-(N-1)/2} \exp\left(-\frac{1}{2\xi^2} \alpha' \mathbf{Q} \alpha\right) \quad (3.13)$$

where \mathbf{Q} is the precision matrix of each IGMRF.

Following the literature for first order IGMRFs on regular lattices (Rue and Held, 2005, Chap. 3), the precision matrix \mathbf{Q} is the Laplacian matrix of the underlying voxel grid, i.e. \mathbf{Q} has elements

$$Q_{ij} = \begin{cases} n_i, & i = j, \\ -1, & i \sim j, \\ 0, & \text{else,} \end{cases}$$

with n_i the number of neighbors of voxel i and $i \sim j$ denoting i and j are neighbors. This choice of \mathbf{Q} has a rank deficiency of 1, which is due to linear constraints saying that the differences in α_i -values resp. $\alpha_{0,i}$ -values sum to zero over all circuits in the lattice graph. Thus, the priors (3.12) and (3.13) are also partially improper, which—usually—does not lead to problems as long as the full posterior of the model is proper (see Fahrmeir and Kneib, 2009, for a discussion of sufficient conditions for propriety in the context of structured additive regression models). The conditional distribution of an α_i - resp. $\alpha_{0,i}$ -value at voxel i is only dependent on the value of neighboring voxels and has the form

$$\alpha_{0,i} | \alpha_{0,j \neq i}, \xi_0^2 \sim \mathcal{N}\left(\frac{1}{n_i} \sum_{j \in \partial_i} \alpha_{0,j}, \frac{\xi_0^2}{n_i}\right),$$

$$\alpha_i | \alpha_{j \neq i}, \xi^2 \sim \mathcal{N}\left(\frac{1}{n_i} \sum_{j \in \partial_i} \alpha_j, \frac{\xi^2}{n_i}\right),$$

where ∂_i denotes the neighborhood of i , $\alpha_{j \neq i}$ is the vector α without α_i and $\alpha_{0,j \neq i}$ is the vector α_0 without $\alpha_{0,i}$. Hence, the prior mean for α_i resp. $\alpha_{0,i}$ is the mean of its neighbors. With this, departures from a smooth surface in α_i and $\alpha_{0,i}$ are penalized. The degree of enforced smoothness is determined by the variance parameters ξ^2 and ξ_0^2 . Complete smoothness, i.e. the case of a constant scaling factor $\alpha_1 = \dots = \alpha_N$ and intercept $\alpha_{0,1} = \dots = \alpha_{0,N}$, is included as a limiting case as $\xi^2 \rightarrow 0$ resp. $\xi_0^2 \rightarrow 0$. Both ξ_0^2 and ξ^2 are not set fixed in advance, but are estimated within the derived procedure and with ξ^2 and ξ_0^2 , the respective degree of smoothness. Prior assumptions about these variance parameters are modeled by inverse gamma (IG) prior distributions

$$\xi_0^2 \sim IG(a_0, b_0),$$

$$\xi^2 \sim IG(a, b),$$

with densities

$$p(\xi_0^2) \propto (\xi_0^2)^{-(a_0+1)} \exp\left(-\frac{b_0}{\xi_0^2}\right)$$

$$p(\xi^2) \propto (\xi^2)^{-(a+1)} \exp\left(-\frac{b}{\xi^2}\right).$$

Either adding/subtracting a constant to the spatially-varying intercept or the spatially-varying EEG coefficient does not change the corresponding prior value in (3.12) or (3.13), because a constant vector lies in the null space of the precision matrix \mathbf{Q} . Thus, implicitly an improper prior is used for the global level. However, we have seen some improvements in the convergency of parameter trajectories within the MCMC sampling procedures if proper priors are used for every parameter. To keep the model identifiable, we choose to add a global intercept and/or a global EEG coefficient and center the according spatially-varying coefficients to mean 0.

If global effects are included, i.e. either in linear predictors (3.8) and (3.11) or in linear predictors (3.6) and (3.9), corresponding prior distributions have to be specified. The global EEG coefficient α_G then is modeled to follow an informative univariate normal distribution $\mathcal{N}(\mu_G, \xi_G^2)$, whereas its variance parameter has an $IG(a_G, b_G)$ prior. The global intercept $\alpha_{0,G}$ then is modeled to follow an informative univariate normal distribution $\mathcal{N}(\mu_{0,G}, \xi_{0,G}^2)$, whereas its variance parameter has an $IG(a_{0,G}, b_{0,G})$ prior.

In case of the CAR model, the prior of spatially-varying coefficients are proper with positive definite precision matrices \mathbf{P}_0 and \mathbf{P} which are defined in analogy to Pettitt et al. (2002). Our choice is based on the good performance of the precision matrix shown in studies of Smith and Smith (2006) on latent binary Markov random fields. The spatially-varying intercept has prior $\alpha_0 | \tau_0^2, \xi_0^2 \sim \mathcal{N}(\mathbf{0}, \xi_0^2 \mathbf{P}_0^{-1})$, where $\mathbf{P}_0 = \mathbf{I} + \tau_0^2 \mathbf{Q}$. Analogously, the prior for the spatially-varying EEG coefficient is set to $\alpha | \tau^2, \xi^2 \sim \mathcal{N}(\mathbf{0}, \xi^2 \mathbf{P}^{-1})$, where $\mathbf{P} = \mathbf{I} + \tau^2 \mathbf{Q}$.

Like for the IGMRF prior, with the CAR prior the α_i - resp. $\alpha_{0,i}$ -value at voxel i is only dependent on the value of neighboring voxels. Being a special case of a general GMRF distribution, the conditional distributions can be derived as normal distributions with the following form (Rue and Held, 2005, p. 22):

$$\alpha_{0,i} | \alpha_{0,j \neq i}, \xi_0^2, \tau_0^2 \sim \mathcal{N} \left(\frac{\tau_0^2}{1 + \tau_0^2 n_i} \sum_{j \in \partial_i} \alpha_{0,j}, \frac{\xi_0^2}{1 + \tau_0^2 n_i} \right), \quad (3.14)$$

$$\alpha_i | \alpha_{j \neq i}, \xi^2, \tau^2 \sim \mathcal{N} \left(\frac{\tau^2}{1 + \tau^2 n_i} \sum_{j \in \partial_i} \alpha_j, \frac{\xi^2}{1 + \tau^2 n_i} \right). \quad (3.15)$$

Note that Pettitt et al. (2002) and Smith and Smith (2006) use a more general definition of \mathbf{Q} based on varying weights and possibly negative covariance parameters. Our choice of equal weights in \mathbf{Q} is justified by a neighborhood definition where all pairs of neighbors are equally far apart from each other, which is for example the case for a 3D neighborhood including the six nearest neighbors of a voxel only. Parameters τ^2 and τ_0^2 are constrained to be positive (this is emphasized by using the square in notation). This goes along with a positive correlation between neighboring sites (Pettitt et al., 2002, see Section 2.2)—which is a sound assumption for our model. A positively truncated $\mathcal{N}(\mu_p, \sigma_p^2)$ prior is both used for τ^2 and τ_0^2 . In our analysis, we set $\mu_p = 0$ and $\sigma_p^2 = 5^2$. Variance parameter ξ^2 (ξ_0^2) either follows an inverse gamma prior with parameters a (a_0) and b (b_0) or is set constant to 1 to yield the model used in Smith and Smith (2006).

3.4 Posterior inference

Bayesian model estimation and inference is based on posterior quantities, i.e. quantities of the posterior multivariate probability distribution of all unknown parameters conditional on the data $\mathbf{y} = (\mathbf{y}'_1, \dots, \mathbf{y}'_N)'$ collected:

$$p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\sigma}^2, \boldsymbol{\theta} | \mathbf{y}) \propto \prod_{i=1}^N [p(\mathbf{y}_i | \gamma_i, \boldsymbol{\beta}_i(\gamma_i), \sigma_i^2) p(\boldsymbol{\beta}_i(\gamma_i) | \gamma_i, \sigma_i^2, \mathbf{y}_i) p(\sigma_i^2)] p(\boldsymbol{\gamma} | \boldsymbol{\theta}) p(\boldsymbol{\theta}),$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \dots, \boldsymbol{\beta}'_N)'$ and $\boldsymbol{\theta}$ contains all parameters in lower hierarchical levels—which depends on the exact model formulation. For ease of exploration, the marginal posteriors, i.e. the probability distribution of each parameter(-vector) given \mathbf{y} , are often examined. In the case of fMRI analysis, the quantities of immediate interest are the marginal posterior probabilities of activation $p(\gamma_i = 1 | \mathbf{y})$ for all voxels $i, i = 1, \dots, N$, if γ_i is the indicator for selecting $f_{stim}(i, t)$ as in (3.2). Closed form calculation of this quantity is not possible because it involves integration out of the binary variables that have support on 2^N possible

indicator combinations—amongst other parameters of lower hierarchical levels of our model. Thus, inference relies on an adequate computational strategy that allows exploring features of the posterior distribution.

An approach to posterior inference, which is commonly applied in the Bayesian variable selection literature, is a Markov Chain Monte Carlo (MCMC) scheme (see Smith and Fahrmeir, 2007, and references therein). For each unknown parameter, MCMC methods are used to generate a Markov chain that converges within a burn-in phase to the desired marginal distribution, i.e. its equilibrium distributions. A particular Markov chain algorithm is the Gibbs Sampler (Gelfand and Smith, 1990). It can be applied if the (full) conditional distribution of each unknown parameter(-subvector) given all other parameters in the model and the data \mathbf{y} are in a known form. Then, a Gibbs Sampler is based on repeated sampling from the full conditional distributions given the actual state of all other model components. If some (or all) full conditional distributions have an unknown form, the Gibbs Sampler can be extended to include Metropolis-Hasting steps. Within these, parameters within a batch are sampled from a suitable proposal density and kept with an accordingly defined acceptance probability, which defines the transition kernel of the targeted Markov chain. For a thorough introduction to MCMC schemes, see Gilks et al. (1996) or Gelman et al. (2004).

In the following subsections, the full conditional distributions of the different model components are derived. In case of the indicator variables γ , the according full conditional can be marginalized with respect to higher level parameters β and σ^2 to yield a more efficient update scheme.

3.4.1 Sampling activation probabilities

As noted above, a quantity of immediate interest is the marginal posterior probability of $p(\gamma_i|\mathbf{y})$ or rather $p(\gamma_i|\mathbf{y})$ in the case of (3.2). To apply the Gibbs sampler and to obtain a sample from these marginals, the full conditional distributions have to be derived.

The chosen Bayesian variable selection scheme belongs to the family of conjugate hierarchical setups (George and McCulloch, 1997), for which the full conditional distribution of γ_i , $p(\gamma_i|\gamma_{j \neq i}, \beta, \sigma^2, \theta, \mathbf{y})$, can analytically be marginalized with respect to β and σ^2 to yield $p(\gamma_i|\gamma_{j \neq i}, \theta, \mathbf{y})$ up to a proportionality constant. This can be used to derive a more efficient computational routine in which a sample of γ_i does not depend on the actual values of β and σ^2 . Note that $\gamma_{j \neq i}$ denotes the vector γ without γ_i .

We will proceed as follows. First, we discuss the form of $p(\gamma_i|\gamma_{j\neq i}, \beta, \sigma^2, \theta, \mathbf{y})$ for a general $\gamma_i = (\gamma_{i,1}, \dots, \gamma_{i,p})$. Then, we turn towards the special case of simultaneous selection of p_3 stimulus basis functions and restrict γ_i to the form proposed in (3.2).

In analogy to Smith and Kohn (1996) and Smith and Fahrmeir (2007), it can be shown that $p(\gamma_i|\gamma_{j\neq i}, \theta, \mathbf{y}) \propto p(\mathbf{y}_i|\gamma_i)p(\gamma_i|\gamma_{j\neq i}, \theta)$, where

$$p(\mathbf{y}_i|\gamma_i) \propto S_i(\gamma_i)^{-T/2}(1+T)^{-q_i/2}.$$

Here, $q_i = \sum_{j=1}^p \gamma_{i,j}$ is the number of nonzero regressors in regression i , and

$$S_i(\gamma_i) = \mathbf{y}'_i \mathbf{y}_i - \mathbf{y}'_i \mathbf{X}(\gamma_i) (\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1} \mathbf{X}(\gamma_i)' \mathbf{y}_i$$

is the sum of squares in regression i corresponding to subset γ_i . The derivation of this distributional form can be found in Appendix A.1.

Since γ_i has support on 2^p different values, it is impractical to generate a random sample for γ_i as a whole. Instead, the full conditional is split up into the different regressor components $p(\gamma_{i,k}|\gamma_{i,l\neq k}, \gamma_{j\neq i}, \theta, \mathbf{y})$ to yield separate Bernoulli distributions.

In the following, we consider the special form of γ_i as given in (3.2). Here, we have only one scalar γ_i per voxel which implements a simultaneous selection of p_3 stimulus basis functions. The sampler for the full conditional can then be derived by the following procedure. First, note that the posterior distribution can be written as a transform $A(\cdot)$ of γ_i given θ :

$$p(\gamma_i|\gamma_{j\neq i}, \theta, \mathbf{y}) \propto p(\mathbf{y}_i|\gamma_i)p(\gamma_i|\gamma_{j\neq i}, \theta) \propto A(\gamma_i|\theta). \quad (3.16)$$

Then, $A(\gamma_i = 1|\theta)$ and $A(\gamma_i = 0|\theta)$ can be calculated. Through normalization, the posterior update probability is $p(\gamma_i = 1|\gamma_{j\neq i}, \theta, \mathbf{y}) = 1/(1+\mathcal{H}(\theta))$, where $\mathcal{H}(\theta) = A(\gamma_i = 0|\theta)/A(\gamma_i = 1|\theta)$.

For the case when $\gamma_i = 1$,

$$A(\gamma_i = 1|\theta) = S_{i1}^{-T/2}(1+T)^{-p/2}p(\gamma_i = 1|\gamma_{j\neq i}, \theta),$$

where $S_{i1} = \mathbf{y}'_i \mathbf{y}_i - \mathbf{y}'_i \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'_i \mathbf{y}_i$ and \mathbf{X} is the full design matrix, i.e. $\mathbf{X} = (\mathbf{W}|\mathbf{C}|\mathbf{Z})$.

For the case when $\gamma_i = 0$,

$$A(\gamma_i = 0|\theta) = S_{i0}^{-T/2}(1+T)^{-(p_1+p_2)/2}p(\gamma_i = 0|\gamma_{j\neq i}, \theta),$$

where $S_{i0} = \mathbf{y}'_i \mathbf{y}_i - \mathbf{y}'_i \mathbf{X}_0 (\mathbf{X}'_0 \mathbf{X}_0)^{-1} \mathbf{X}'_0 \mathbf{y}_i$ and $\mathbf{X}_0 = (\mathbf{W} | \mathbf{C})$. Normalizing the results, we get

$$\mathcal{H}(\boldsymbol{\theta}) = \frac{A(\gamma_i = 0 | \boldsymbol{\theta})}{A(\gamma_i = 1 | \boldsymbol{\theta})} = \exp(l_i) \frac{p(\gamma_i = 0 | \boldsymbol{\gamma}_{j \neq i}, \boldsymbol{\theta})}{p(\gamma_i = 1 | \boldsymbol{\gamma}_{j \neq i}, \boldsymbol{\theta})} \quad (3.17)$$

with

$$l_i = \log \left\{ \left(\frac{S_{i1}}{S_{i0}} \right)^{T/2} (1 + T)^{p_3/2} \right\}.$$

The term l_i is the log-likelihood ratio of the marginalized likelihoods $p(\mathbf{y}_i | \gamma_i)$. A sample γ_i from the marginalized full conditional can now be drawn from a Bernoulli distribution with probability $p(\gamma_i = 1 | \boldsymbol{\gamma}_{j \neq i}, \boldsymbol{\theta}, \mathbf{y}) = 1 / (1 + \mathcal{H}(\boldsymbol{\theta}))$. For an exact specification of $\mathcal{H}(\boldsymbol{\theta})$, the relationship to the parameter vector $\boldsymbol{\theta}$ must be taken into account, which is elaborated in the next section.

3.4.2 Binary regression level

As discussed before, we restrict the EEG-informed variable selection to the form of γ_i in (3.2), so that a scalar γ_i for detecting brain activation due to stimuli is considered, and use according priors proposed in Section 3.3.

The exact specification of $\boldsymbol{\theta}$ and related distributional assumptions depend on the selected binary regression setup for the scalar activation parameters γ_i used in this level of the Bayesian model hierarchy. In particular, this means that the exact model formulation depends on the response function $h(\cdot)$ used to define the relationship between binary variable γ_i and the predictor η_i with transformed EEG information $f(J_i)$ (compare Section 3.3). In our application, we consider a probit link, because it can simply be embedded in a Gibbs sampling scheme.

The parameter $\boldsymbol{\theta}$ then comprises the regression coefficients of the binary linear model stage and parameters of the hierarchical stages below. As discussed in Section 3.3, EEG related regression coefficients comprise either the spatially-varying EEG coefficients $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)'$ or a global EEG coefficient α_G or both in case of an IGMRF prior. Analogously, either a spatially-varying intercept $\boldsymbol{\alpha}_0 = (\alpha_{0,1}, \dots, \alpha_{0,N})'$ is included or a global intercept $\alpha_{0,G}$ or both when an IGMRF prior is used. Additionally, prior distributions of these regression parameters depend on unknown variance parameters.

In the following, $\boldsymbol{\theta}_{-\nu}$ denotes the parameter vector $\boldsymbol{\theta}$ without the parameter(-subvector) ν .

In analogy to a generalized linear model framework, the $\gamma_i, i = 1, \dots, N$, which serve as latent binary observations, are assumed to be conditionally independent given the regression coefficients in η_i . Hence, the prior can be written as

$$\begin{aligned} p(\boldsymbol{\gamma}|\boldsymbol{\theta}) &= \prod_{i=1}^N p(\gamma_i|\eta_i) \\ &= \prod_{i=1}^N h(\eta_i)^{\gamma_i} (1 - h(\eta_i))^{1-\gamma_i} \\ &= \prod_{i=1}^N \Phi(\eta_i)^{\gamma_i} (1 - \Phi(\eta_i))^{1-\gamma_i}, \end{aligned} \quad (3.18)$$

where $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_N)'$. Thus, in the model with probit link, the response function $h(\cdot)$ is equal to the distribution function $\Phi(\cdot)$ of the standard normal distribution. However, our exact model does not rely on (3.18), but on a slight model modification. A convenient posterior sampling scheme can be derived by a data augmentation approach proposed by Albert and Chib (1993). For this, N latent variables U_1, \dots, U_N are introduced, where the U_i follow independent $\mathcal{N}(\eta_i, 1)$. Hence, $\mathbf{U} = (U_1, \dots, U_N) \sim \mathcal{N}_N(\boldsymbol{\eta}, \mathbf{I})$ with $\boldsymbol{\eta} = (\eta_1, \dots, \eta_N)'$. Define $\gamma_i = 1$ if $U_i > 0$ and $\gamma_i = 0$ otherwise. It can be shown that this goes along with the assumption of the γ_i 's being independent Bernoulli random variables with $p(\gamma_i = 1|\eta_i) = \Phi(\eta_i)$. The assumption of a threshold equal to 0 is relaxed by the intercept term $f_0(i)$, which can be interpreted as an automatic threshold estimation for the EEG part $f(J_i)$ of the predictor in U_i . Note that $\boldsymbol{\theta}$ now comprises \mathbf{U} besides regression and related variance parameters.

The relationship of U_i and γ_i in form of a probability distribution can be expressed through a mixture of point masses

$$p(\gamma_i|U_i) = I(U_i > 0)I(\gamma_i = 1) + I(U_i \leq 0)I(\gamma_i = 0), \quad (3.19)$$

where $I(A)$ is an indicator function equal to 1 if A is true and 0 otherwise. Again, the assumption of conditional independence of the γ_i 's given U_i holds, so that the conditional joint distribution of $\boldsymbol{\gamma}$ can be written as $p(\boldsymbol{\gamma}|\mathbf{U}) = \prod_{i=1}^N p(\gamma_i|U_i)$.

The full conditional distributions for the latent variables U_i can be derived as follows. Being conditionally independent given the regression coefficients—consisting of all or a subset of $\alpha_{0,G}$, α_G , α_0 and $\boldsymbol{\alpha}$ —a rather simple form can be derived by

$$p(U_i|\boldsymbol{\theta}_{-U_i}, \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) \propto p(\gamma_i|U_i)p(U_i|\eta_i)$$

$$= [I(U_i > 0)I(\gamma_i = 1) + I(U_i \leq 0)I(\gamma_i = 0)] \\ \times (2\pi)^{-1/2} \exp\left(-\frac{1}{2}(U_i - \eta_i)^2\right).$$

Hence, samples for U_i can be drawn from a truncated normal distribution, i.e.

$$p(U_i|\gamma_i = 1, \boldsymbol{\gamma}_{j \neq i}, \boldsymbol{\theta}_{-U_i}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) \propto (2\pi)^{-1/2} \exp\left(-\frac{1}{2}(U_i - \eta_i)^2\right)$$

restricted on $U_i > 0$ if $\gamma_i = 1$,

$$p(U_i|\gamma_i = 0, \boldsymbol{\gamma}_{j \neq i}, \boldsymbol{\theta}_{-U_i}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) \propto (2\pi)^{-1/2} \exp\left(-\frac{1}{2}(U_i - \eta_i)^2\right)$$

restricted on $U_i \leq 0$ if $\gamma_i = 0$.

For updating γ_i , the exact form of $\mathcal{H}(\boldsymbol{\theta})$ has to be derived. For this, we have to take into account the following. If a realization of U_i is put into $p(\gamma_i|U_i) = I(U_i > 0)I(\gamma_i = 1) + I(U_i \leq 0)I(\gamma_i = 0)$, then $p(\gamma_i = 1|U_i)$ is either 1 or 0. If $p(\gamma_i = 1|U_i)$ and $p(\gamma_i = 0|U_i) = 1 - p(\gamma_i = 1|U_i)$ is then used to calculate $\mathcal{H}(\boldsymbol{\theta})$, a reducible sampling scheme evolves, in which a proposal for γ_i is fully determined by U_i and, thus, by prior beliefs only. Then, $\gamma_i = 1$ for $U_i > 0$, and in the next iteration $U_i > 0$ because $\gamma_i = 1$. Hence, information from the data \mathbf{y} in form of the marginalized likelihood ratio would be entirely suppressed and the Markov chain gets stuck. To avoid this, we suggest to update (U_i, γ_i) as a pair by first sampling from $p(\gamma_i|\boldsymbol{\theta}_{-U_i}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y})$ and then generating from $p(U_i|\gamma_i, \boldsymbol{\theta}_{-U_i}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) = p(U_i|\boldsymbol{\theta}_{-U_i}, \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y})$. The latter coincides with the above full conditional for U_i .

The conditional posterior distribution $p(\gamma_i|\boldsymbol{\theta}_{-U_i}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y})$ is derived by marginalizing $p(\mathbf{y}_i|\gamma_i)p(\gamma_i|\boldsymbol{\gamma}_{j \neq i}, \boldsymbol{\theta})$ in (3.16) with respect to U_i . In particular, for this the prior component of (3.16) is marginalized with respect to U_i . The prior component can be rewritten as

$$p(\gamma_i|\boldsymbol{\gamma}_{j \neq i}, U_i, \mathbf{U}_{j \neq i}, \boldsymbol{\theta}_{-\mathbf{U}}) = p(\gamma_i|U_i, \mathbf{U}_{j \neq i}, \boldsymbol{\theta}_{-\mathbf{U}})$$

using that the elements of $\boldsymbol{\gamma}$ are conditionally independent given U_i . Marginalizing then leads to

$$p(\gamma_i|\mathbf{U}_{j \neq i}, \boldsymbol{\theta}_{-\mathbf{U}}) = \int p(\gamma_i|U_i)p(U_i|\mathbf{U}_{j \neq i}, \boldsymbol{\theta}_{-\mathbf{U}})\partial U_i \\ = \int p(\gamma_i|U_i)p(U_i|\boldsymbol{\theta}_{-\mathbf{U}})\partial U_i$$

$$= \int [I(U_i > 0)I(\gamma_i = 1) + I(U_i < 0)I(\gamma_i = 0)]p(U_i|\boldsymbol{\theta}_{-U})\partial U_i.$$

The conditional density $p(U_i|\boldsymbol{\theta}_{-U})$ is the density of a $\mathcal{N}(\eta_i, 1)$ distribution. Evaluating for $\gamma_i = 1$, we get

$$p(\gamma_i = 1|\mathbf{U}_{j \neq i}, \boldsymbol{\theta}_{-U}) = \int_0^\infty p(U_i|\boldsymbol{\theta}_{-U})\partial U_i = 1 - \Phi(0; \eta_i, 1) = \Phi(\eta_i; 0, 1), \quad (3.20)$$

where $\Phi(x; \eta_i, 1)$ is the value of the distribution function of a normal distribution with mean η_i and variance 1 at quantile $x \in \mathbb{R}$. For $\gamma_i = 0$, we get

$$p(\gamma_i = 0|\mathbf{U}_{j \neq i}, \boldsymbol{\theta}_{-U}) = \int_{-\infty}^0 p(U_i|\boldsymbol{\theta}_{-U})\partial U_i = \Phi(0; \eta_i, 1) = 1 - \Phi(\eta_i; 0, 1).$$

Substituting the marginalized prior distributions into (3.17) yields an improved updating scheme for γ_i as a Bernoulli variate with probability $1/(1 + \mathcal{H}(\boldsymbol{\theta}))$ and

$$\mathcal{H}(\boldsymbol{\theta}) = \exp(l_i) \frac{1 - \Phi(\eta_i)}{\Phi(\eta_i)}.$$

Which regression coefficients and related variance parameters have to be updated depends on the exact specification of the predictor. See Table 3.1 for a listing of suitable predictor forms. However, full conditional distributions for all parameters can be derived in a general form—despite the exact choice of predictor and spatial prior. For a specific choice of predictor and spatial prior, related full conditionals can be selected to update parameters appearing in the hierarchical modeling setup.

We first elaborate on the derivation of full conditionals for the update of the whole spatially-varying EEG coefficient vector. Assume that $\boldsymbol{\Sigma}^-$ is the precision matrix of the multivariate normal spatial prior, i.e. $\boldsymbol{\Sigma}^- = \frac{1}{\xi^2} \mathbf{Q}$ in case of the IGMRF prior and $\boldsymbol{\Sigma}^- = \frac{1}{\xi^2} \mathbf{P}$ in case of the CAR prior. Let $\mathbf{J} = \text{diag}(J_1, \dots, J_N)$ be the design matrix part for the varying EEG coefficient. The full conditional for the parameter vector $\boldsymbol{\alpha}$ can be derived by extracting the components of the full posterior dependent on $\boldsymbol{\alpha}$ and recognizing a multivariate normal distribution:

$$\begin{aligned} p(\boldsymbol{\alpha}|\boldsymbol{\gamma}, \boldsymbol{\theta}_{-\boldsymbol{\alpha}}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) &\propto p(\boldsymbol{\alpha}, \boldsymbol{\theta}_{-\boldsymbol{\alpha}}, \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2|\mathbf{y}) \\ &\propto p(\mathbf{U}|\boldsymbol{\eta})p(\boldsymbol{\alpha}|\xi^2) \\ &\propto \exp\left(-\frac{1}{2}(\mathbf{U} - \boldsymbol{\eta}_{-\boldsymbol{\alpha}} - \mathbf{J}\boldsymbol{\alpha})'(\mathbf{U} - \boldsymbol{\eta}_{-\boldsymbol{\alpha}} - \mathbf{J}\boldsymbol{\alpha})\right) \exp\left(-\frac{1}{2}\boldsymbol{\alpha}'\boldsymbol{\Sigma}^-\boldsymbol{\alpha}\right) \end{aligned}$$

$$\begin{aligned}
&= \exp \left(-\frac{1}{2} [(\mathbf{U} - \boldsymbol{\eta}_{-\alpha})'(\mathbf{U} - \boldsymbol{\eta}_{-\alpha}) - (\mathbf{U} - \boldsymbol{\eta}_{-\alpha})\mathbf{J}\boldsymbol{\alpha} - \boldsymbol{\alpha}'\mathbf{J}'(\mathbf{U} - \boldsymbol{\eta}_{-\alpha}) \right. \\
&\quad \left. + \boldsymbol{\alpha}'\mathbf{J}'\mathbf{J}\boldsymbol{\alpha}] - \frac{1}{2}\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{-}\boldsymbol{\alpha} \right) \\
&\propto \exp \left(-\frac{1}{2} [\boldsymbol{\alpha}'(\mathbf{J}'\mathbf{J} + \boldsymbol{\Sigma}^{-})\boldsymbol{\alpha} - (\mathbf{U} - \boldsymbol{\eta}_{-\alpha})\mathbf{J}'\boldsymbol{\alpha} - \boldsymbol{\alpha}'\mathbf{J}'(\mathbf{U} - \boldsymbol{\eta}_{-\alpha})] \right), \quad (3.21)
\end{aligned}$$

where $\boldsymbol{\eta}_{-\alpha}$ denotes the vector of predictor values without the $\mathbf{J}\boldsymbol{\alpha}$ component, i.e. $\boldsymbol{\eta}_{-\alpha} = \boldsymbol{\eta} - \mathbf{J}\boldsymbol{\alpha}$. Equation (3.21) can be completed to a multivariate normal distribution with mean $(\mathbf{J}'\mathbf{J} + \boldsymbol{\Sigma}^{-})^{-1}\mathbf{J}'(\mathbf{U} - \boldsymbol{\eta}_{-\alpha})$ and variance-covariance matrix $(\mathbf{J}'\mathbf{J} + \boldsymbol{\Sigma}^{-})^{-1}$, which is then the full conditional of $\boldsymbol{\alpha}$, i.e.

$$\boldsymbol{\alpha} | \cdot \sim \mathcal{N} \left((\mathbf{J}'\mathbf{J} + \boldsymbol{\Sigma}^{-})^{-1}\mathbf{J}'(\mathbf{U} - \boldsymbol{\eta}_{-\alpha}), (\mathbf{J}'\mathbf{J} + \boldsymbol{\Sigma}^{-})^{-1} \right). \quad (3.22)$$

Similar arguments lead to a full conditional distribution for α_0 . This random field is updated with a multivariate normal distribution with mean $(\mathbf{I} + \boldsymbol{\Sigma}_0^{-})^{-1}(\mathbf{U} - \boldsymbol{\eta}_{-\alpha_0})$ and variance-covariance matrix $(\mathbf{I} + \boldsymbol{\Sigma}_0^{-})^{-1}$, where $\boldsymbol{\Sigma}_0^{-} = \frac{1}{\xi_0^2}\mathbf{Q}_0$ in case of the IGMRF prior and $\boldsymbol{\Sigma}_0^{-} = \frac{1}{\xi_0^2}\mathbf{P}_0$ in case of the CAR prior. In notation,

$$\alpha_0 | \cdot \sim \mathcal{N} \left((\mathbf{I} + \boldsymbol{\Sigma}_0^{-})^{-1}(\mathbf{U} - \boldsymbol{\eta}_{-\alpha_0}), (\mathbf{I} + \boldsymbol{\Sigma}_0^{-})^{-1} \right). \quad (3.23)$$

For each random field, all components are updated in one block. For the IGMRF prior, the generated $\boldsymbol{\alpha}$ (α_0) is centered to mean 0 immediately after sampling. An efficient sampling scheme from these multivariate normal distributions utilizes the sparse matrix nature of the variance-covariance matrices at hand as described in Rue (2001) and Lang and Brezger (2004). Implementational details on this efficient Cholesky decomposition algorithm can be found in Section 5.2.2.

If a global parameter with one-dimensional normal prior is included in the predictor, the according full conditional distribution is derived as follows. For the global intercept with $\mathcal{N}(\mu_{0,G}, \sigma_{0,G}^2)$ prior, we have,

$$\begin{aligned}
p(\alpha_{0,G} | \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\alpha_{0,G}}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) &\propto p(\mathbf{U} | \boldsymbol{\eta}) p(\alpha_{0,G} | \mu_{0,G}, \xi_{0,G}^2) \\
&\propto \exp \left(-\frac{1}{2} (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{0,G}} - \mathbf{1}\alpha_{0,G})' (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{0,G}} - \mathbf{1}\alpha_{0,G}) \right) \\
&\quad \times \exp \left(-\frac{1}{2\xi_{0,G}^2} (\alpha_{0,G} - \mu_{0,G})^2 \right)
\end{aligned}$$

$$\propto \exp \left(-\frac{1}{2} \left[\alpha_{0,G}^2 \left(\mathbf{1}'\mathbf{1} + \frac{1}{\xi_{0,G}^2} \right) - 2 \left((\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{0,G}})' \mathbf{1} + \frac{\mu_{0,G}}{\xi_{0,G}^2} \right) \alpha_{0,G} \right] \right),$$

which can be completed to a univariate normal distribution with variance $\left(N + \frac{1}{\xi_{0,G}^2}\right)^{-1}$ and mean $\left(N + \frac{1}{\xi_{0,G}^2}\right)^{-1} \left((\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{0,G}})' \mathbf{1} + \frac{\mu_{0,G}}{\xi_{0,G}^2} \right)$, where $\mathbf{1}$ is the N -dimensional vector of ones.

Similarly, we can derive a full conditional for a global EEG coefficient with $\mathcal{N}(\mu_G, \sigma_G^2)$ prior as univariate normal distribution with mean $\left(\mathbf{j}'\mathbf{j} + \frac{1}{\xi_G^2}\right)^{-1} \left((\mathbf{U} - \boldsymbol{\eta}_{-\alpha_G})' \mathbf{j} + \frac{\mu_G}{\xi_G^2} \right)$ and variance $\left(\mathbf{j}'\mathbf{j} + \frac{1}{\xi_G^2}\right)^{-1}$, where $\mathbf{j} = (J_1, \dots, J_N)'$ is the vectorized map of EEG measurements. If both global parameters are included in the predictor, the two global parameters can be updated in one block using a two-dimensional multivariate normal distribution. We have found that doing a simultaneous update improves convergency of the algorithm in the model with IGMRF priors. The full conditional is derived as follows. The assumption of two independent normal priori distributions leads to a spheric two-dimensional prior distribution $\mathcal{N}_2(\underline{\boldsymbol{\mu}}_G, \boldsymbol{\Sigma}_G)$ with $\underline{\boldsymbol{\mu}}_G = (\mu_{0,G}, \mu_G)'$ for $(\alpha_{0,G}, \alpha_G)'$ and

$$\boldsymbol{\Sigma}_G = \begin{pmatrix} \xi_{0,G}^2 & 0 \\ 0 & \xi_G^2 \end{pmatrix}.$$

Let $\underline{\boldsymbol{\alpha}}_G = (\alpha_{0,G}, \alpha_G)$ and $\mathbf{X}_G = (\mathbf{1}, \mathbf{j})$, then

$$\begin{aligned} p(\underline{\boldsymbol{\alpha}}_G | \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\underline{\boldsymbol{\alpha}}_G}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) &\propto p(\mathbf{U} | \boldsymbol{\eta}) p(\underline{\boldsymbol{\alpha}}_G | \underline{\boldsymbol{\mu}}_G, \boldsymbol{\Sigma}_G) \\ &\propto \exp \left(-\frac{1}{2} (\mathbf{U} - \boldsymbol{\eta}_{-\underline{\boldsymbol{\alpha}}_G} - \mathbf{X}_G \underline{\boldsymbol{\alpha}}_G)' (\mathbf{U} - \boldsymbol{\eta}_{-\underline{\boldsymbol{\alpha}}_G} - \mathbf{X}_G \underline{\boldsymbol{\alpha}}_G) \right) \\ &\quad \times \exp \left(-\frac{1}{2} (\underline{\boldsymbol{\alpha}}_G - \underline{\boldsymbol{\mu}}_G)' \boldsymbol{\Sigma}_G^{-1} (\underline{\boldsymbol{\alpha}}_G - \underline{\boldsymbol{\mu}}_G) \right) \\ &\propto \exp \left(-\frac{1}{2} \left[\underline{\boldsymbol{\alpha}}_G' (\mathbf{X}_G' \mathbf{X}_G + \boldsymbol{\Sigma}_G^{-1}) \underline{\boldsymbol{\alpha}}_G - ((\mathbf{U} - \boldsymbol{\eta}_{-\underline{\boldsymbol{\alpha}}_G})' \mathbf{X}_G + \underline{\boldsymbol{\mu}}_G' \boldsymbol{\Sigma}_G^{-1}) \underline{\boldsymbol{\alpha}}_G \right. \right. \\ &\quad \left. \left. - \underline{\boldsymbol{\alpha}}_G' (\mathbf{X}_G' (\mathbf{U} - \boldsymbol{\eta}_{-\underline{\boldsymbol{\alpha}}_G}) + \boldsymbol{\Sigma}_G^{-1} \underline{\boldsymbol{\mu}}_G) \right] \right), \end{aligned}$$

where again $\boldsymbol{\eta}_{-\underline{\boldsymbol{\alpha}}_G}$ is the predictor value without the global part $\mathbf{X}_G \underline{\boldsymbol{\alpha}}_G$, i.e. $\boldsymbol{\eta}_{-\underline{\boldsymbol{\alpha}}_G} = \boldsymbol{\eta} - \mathbf{X}_G \underline{\boldsymbol{\alpha}}_G$. The last equation can be completed to a multivariate normal distribution with mean $(\mathbf{X}_G' \mathbf{X}_G + \boldsymbol{\Sigma}_G^{-1})^{-1} (\mathbf{X}_G' (\mathbf{U} - \boldsymbol{\eta}_{-\underline{\boldsymbol{\alpha}}_G}) + \boldsymbol{\Sigma}_G^{-1} \underline{\boldsymbol{\mu}}_G)$ and variance-covariance matrix $(\mathbf{X}_G' \mathbf{X}_G + \boldsymbol{\Sigma}_G^{-1})^{-1}$. Note that we choose to not estimate $\underline{\boldsymbol{\mu}}_G$ later on but to set its value. The reason for this is that the complexity of the model should not further be increased and,

thus, to prevent an overload of the estimation procedure. We set the value, for instance, to zero or to the estimates from a probit model with global variables only (evaluating their effect on an initial guess of γ).

The choice of a conditional conjugate prior for each variance parameter, i.e. an inverse gamma prior, leads to a simple update scheme for all variances. The full conditional distribution of each variance parameter is again an inverse gamma distribution. In the IGMRF case, for ξ^2 (the variance parameter of the spatially-varying EEG coefficient) the update distribution can be derived by

$$\begin{aligned} p(\xi^2 | \gamma, \boldsymbol{\theta}_{-\xi^2}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) &\propto p(\boldsymbol{\alpha} | \xi^2) p(\xi^2) \\ &\propto (\xi^2)^{-(N-k)/2} \exp\left(-\frac{1}{2\xi^2} \boldsymbol{\alpha}' \mathbf{Q} \boldsymbol{\alpha}\right) \left(\frac{1}{\xi^2}\right)^{(a+1)} \exp\left(-\frac{b}{\xi^2}\right) \\ &= (\xi^2)^{-(a+\frac{1}{2}(N-k)+1)} \exp\left(-\frac{1}{\xi^2} \left[b + \frac{1}{2} \boldsymbol{\alpha}' \mathbf{Q} \boldsymbol{\alpha}\right]\right) \end{aligned}$$

where k is the rank deficiency of the singular matrix \mathbf{Q} . This can be completed to an inverse gamma distribution with shape parameter $a + \frac{1}{2}(N - k)$ and scale parameter $b + \frac{1}{2} \boldsymbol{\alpha}' \mathbf{Q} \boldsymbol{\alpha}$. When using a first-order IGMRF as proposed, $k = 1$. If the spatial CAR prior is used, $k = 0$ and \mathbf{Q} has to be substituted with \mathbf{P} .

In an analogous way, it can be shown that in the IGMRF case the full conditional for ξ_0^2 (the variance parameter of the spatially-varying intercept) is an inverse gamma distribution with shape parameter $a_0 + \frac{1}{2}(N - k)$ and scale parameter $b_0 + \frac{1}{2} \boldsymbol{\alpha}'_0 \mathbf{Q} \boldsymbol{\alpha}_0$. When using a first-order IGMRF as proposed, $k = 1$. Again, if the CAR prior is used, $k = 0$ and \mathbf{Q} has to be replaced with \mathbf{P}_0 .

Furthermore, we get for the variance parameter ξ_G^2 of the prior for a global EEG effect α_G an inverse gamma distribution with shape parameter $a_G + 0.5$ and scale parameter $b_G + 0.5(\alpha_G - \mu_G)^2$. For the variance parameter $\xi_{0,G}^2$ of the prior for a global intercept $\alpha_{0,G}$, we get an inverse gamma distribution with shape parameter $a_{0,G} + 0.5$ and scale parameter $b_{0,G} + 0.5(\alpha_{0,G} - \mu_{0,G})^2$.

If the CAR prior for the random field $\boldsymbol{\alpha}$ or $\boldsymbol{\alpha}_0$ is used, the spatial dependency parameter τ^2 or τ_0^2 has to be updated additionally. Updates are generated using Metropolis steps (Gelman et al., 2004, Chap. 11). A proposal $\tau^{2,new}$ is drawn from a positively truncated $\mathcal{N}(\tau^{2,old}, \sigma_{\tau^2}^2)$ distribution, which is symmetric in $\tau^{2,new}$ and $\tau^{2,old}$. The acceptance

probability is defined by

$$\begin{aligned} \alpha(\tau^{2,new}|\tau^{2,old}) &= \min \left\{ \frac{p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\sigma}^2, \boldsymbol{\theta}_{-\tau^2}, \tau^{2,new}|\mathbf{y})}{p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\sigma}^2, \boldsymbol{\theta}_{-\tau^2}, \tau^{2,old}|\mathbf{y})}, 1 \right\} \\ &= \min \left\{ \frac{|\mathbf{P}^{new}|^{1/2} \exp\left(-\frac{1}{2\xi^2} \boldsymbol{\alpha}' \mathbf{P}^{new} \boldsymbol{\alpha}\right) \exp\left(-\frac{1}{2\sigma_p^2} (\tau^{2,new} - \mu_p)^2\right)}{|\mathbf{P}^{old}|^{1/2} \exp\left(-\frac{1}{2\xi^2} \boldsymbol{\alpha}' \mathbf{P}^{old} \boldsymbol{\alpha}\right) \exp\left(-\frac{1}{2\sigma_p^2} (\tau^{2,old} - \mu_p)^2\right)}, 1 \right\}, \end{aligned} \quad (3.24)$$

where $\mathbf{P}^{new} = \mathbf{I} + \tau^{2,new} \mathbf{Q}$ and $\mathbf{P}^{old} = \mathbf{I} + \tau^{2,old} \mathbf{Q}$. A proposal $\tau_0^{2,new}$ is drawn from a positively truncated $\mathcal{N}(\tau_0^{2,old}, \sigma_{\tau_0}^2)$ distribution. It is accepted with probability

$$\alpha(\tau_0^{2,new}|\tau_0^{2,old}) = \min \left\{ \frac{|\mathbf{P}_0^{new}|^{1/2} \exp\left(-\frac{1}{2\xi_0^2} \boldsymbol{\alpha}_0' \mathbf{P}_0^{new} \boldsymbol{\alpha}_0\right) \exp\left(-\frac{1}{2\sigma_{0,p}^2} (\tau_0^{2,new} - \mu_{0,p})^2\right)}{|\mathbf{P}_0^{old}|^{1/2} \exp\left(-\frac{1}{2\xi_0^2} \boldsymbol{\alpha}_0' \mathbf{P}_0^{old} \boldsymbol{\alpha}_0\right) \exp\left(-\frac{1}{2\sigma_{0,p}^2} (\tau_0^{2,old} - \mu_{0,p})^2\right)}, 1 \right\}. \quad (3.25)$$

Determinant computation is done via a Cholesky decomposition $\mathbf{P} = \mathbf{L}\mathbf{L}'$, resp. $\mathbf{P}_0 = \mathbf{L}_0\mathbf{L}_0'$, as proposed in Rue (2001). The determinant is then computed by $|\mathbf{P}| = |\mathbf{L}||\mathbf{L}'| = |\mathbf{L}|^2 = \left(\prod_{i=1}^N l_{ii}\right)^2$, where l_{ii} are the diagonal elements of \mathbf{L} . Likewise, $|\mathbf{P}_0| = \left(\prod_{i=1}^N l_{0,ii}\right)^2$, where $l_{0,ii}$ are the diagonal elements of \mathbf{L}_0 . For being numerically more stable, we use the log transformation $\log |\mathbf{P}| = 2 \sum_{i=1}^N \log(l_{ii})$ resp. $\log |\mathbf{P}_0| = 2 \sum_{i=1}^N \log(l_{0,ii})$.

Determinant calculation via a Cholesky decomposition, however, is not as efficient as it may seem at the first sight. Determinants must be recalculated for every iterate for depending on the spatial dependency parameter. This slows down the MCMC procedure substantially. We cannot make use of the Cholesky decomposition calculated for the GMRF updates: Update precision matrix and \mathbf{P} resp. \mathbf{P}_0 do not coincide. But for an initial testing phase of the combined fMRI-EEG algorithm, we rely on this simple determinant computation approach.

3.4.3 Algorithmic details

Based on the derived full conditional distributions, we propose the following MCMC algorithm. Due to marginalizing with respect to fMRI level regression parameters, MCMC iterations update parameters of the EEG stage only. The update of $\boldsymbol{\gamma}$, however, additionally depends on the output (a transformation of voxelwise likelihood ratio statistics) of the pure voxelwise fMRI linear model. Note that this is the point where the estimation of EEG-level

parameters is connected to the fMRI level. Therefore, we start with a mass-univariate least squares fit of model (3.1) and then, switch over to the following iterative procedure.

We list an update scheme for a model with all parameters that might appear in the EEG probit stage. If specific predictor components are not contained in a requested model, their update can be omitted in later applications.

1. **Initialization:** Initialize EEG-stage regression parameters to the following values: Initialize α and α_0 with independent zero-mean normal random numbers with variance 1. If global parameters are contained in the model, initialize $\alpha_{0,G}$ and α_G to their prior means. If a CAR prior is used, choose initial values for τ^2 and τ_0^2 . Choose a maximal number of MCMC cycles L and set $l = 1$.

2. **Estimation:** For $l = 1, \dots, L$, draw random samples from the full conditionals:

a. For $\gamma_i^{(l)}, i = 1, \dots, N$, calculate $p(\gamma_i = 1 | \eta_i^{(l)}) = \Phi(\eta_i^{(l)})$ and $p(\gamma_i = 0 | \eta_i^{(l)}) = 1 - \Phi(\eta_i^{(l)})$ and with it

$$\mathcal{H}(\boldsymbol{\theta}) = \exp(l_i) \frac{1 - \Phi(\eta_i^{(l)})}{\Phi(\eta_i^{(l)})}.$$

Draw $\gamma_i^{(l)}$ from a Bernoulli distribution with parameter $1/(1 + \mathcal{H}(\boldsymbol{\theta}))$;

b. For $i = 1, \dots, N$,

$$U_i^{(l)} \sim \mathcal{N}(\eta_i^{(l)}, 1)$$

restricted on $U_i > 0$ if $\gamma_i^{(l-1)} = 1$ or $U_i \leq 0$ if $\gamma_i^{(l-1)} = 0$;

c1. Update of variance parameters in models with IGMRF priors:

- $\xi_{0,G}^{2(l)} \sim IG(a_{0,G} + 0.5, b_{0,G} + 0.5(\alpha_{0,G}^{(l-1)} - \mu_{0,G})^2)$;
- $\xi_G^{2(l)} \sim IG(a_G + 0.5, b_G + 0.5(\alpha_G^{(l-1)} - \mu_G)^2)$;
- $\xi_0^{2(l)} \sim IG(a_0 + 0.5(N - 1), b_0 + 0.5\boldsymbol{\alpha}_0^{(l-1)'} \mathbf{Q} \boldsymbol{\alpha}_0^{(l-1)})$;
- $\xi^{2(l)} \sim IG(a + 0.5(N - 1), b + 0.5\boldsymbol{\alpha}^{(l-1)'} \mathbf{Q} \boldsymbol{\alpha}^{(l-1)})$;

c2. Update of variance parameters in models with CAR priors:

- $\xi_{0,G}^{2(l)} \sim IG(a_{0,G} + 0.5, b_{0,G} + 0.5(\alpha_{0,G}^{(l-1)} - \mu_{0,G})^2)$;

- $\xi_G^{2(l)} \sim IG \left(a_G + 0.5, b_G + 0.5(\alpha_G^{(l-1)} - \mu_G)^2 \right);$
- $\xi_0^{2(l)} \sim IG \left(a_0 + 0.5N, b_0 + 0.5\alpha_0^{(l-1)'} \mathbf{P}_0^{(l-1)} \alpha_0^{(l-1)} \right);$
- $\xi^{2(l)} \sim IG \left(a + 0.5N, b + 0.5\alpha^{(l-1)'} \mathbf{P}^{(l-1)} \alpha^{(l-1)} \right);$
- Draw a proposal $\tau^{2,new}$ from a positively truncated $\mathcal{N}(\tau^{2,old}, \sigma_p^2)$ distribution and accept with probability $\alpha(\tau^{2,new} | \tau^{2,old})$ from (3.24) $\rightarrow \tau^{2(l)}$;
- Draw a proposal $\tau_0^{2,new}$ from a positively truncated $\mathcal{N}(\tau_0^{2,old}, \sigma_{0,p}^2)$ distribution and accept with probability $\alpha(\tau_0^{2,new} | \tau_0^{2,old})$ from (3.25) $\rightarrow \tau_0^{2(l)}$;

d. Update of effect parameters:

- $\alpha^{(l)} \sim \mathcal{N} \left((\mathbf{J}'\mathbf{J} + \Sigma^-)^{-1} \mathbf{J}'(\mathbf{U} - \boldsymbol{\eta}_{-\alpha}^{(l)}), (\mathbf{J}'\mathbf{J} + \Sigma^-)^{-1} \right)$
 $\Sigma^- = \frac{1}{\xi^{2(l)}} \mathbf{Q}$ in case of the IGMRF prior and $\Sigma^- = \frac{1}{\xi^{2(l)}} \mathbf{P}^{(l)}$ in case of the CAR prior; if using an IGMRF prior, center $\alpha^{(l)}$ afterwards by subtracting its realized vector mean and add this mean to $\alpha_G^{(l-1)}$;
- $\alpha_0^{(l)} \sim \mathcal{N} \left((\mathbf{I} + \Sigma_0^-)^{-1} (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_0}^{(l)}), (\mathbf{I} + \Sigma_0^-)^{-1} \right)$
 $\Sigma_0^- = \frac{1}{\xi_0^{2(l)}} \mathbf{Q}$ in case of the IGMRF prior and $\Sigma_0^- = \frac{1}{\xi_0^{2(l)}} \mathbf{P}^{(l)}$ in case of the CAR prior; if using an IGMRF prior, center $\alpha_0^{(l)}$ afterwards by subtracting its realized vector mean and add this mean to $\alpha_{0,G}^{(l-1)}$;
- $\alpha_G^{(l)} \sim \mathcal{N} \left(\left(\mathbf{j}'\mathbf{j} + \frac{1}{\xi_G^{2(l)}} \right)^{-1} \left((\mathbf{U} - \boldsymbol{\eta}_{-\alpha_G}^{(l)})' \mathbf{j} + \frac{\mu_G}{\xi_G^{2(l)}} \right), \left(\mathbf{j}'\mathbf{j} + \frac{1}{\xi_G^{2(l)}} \right)^{-1} \right);$
- $\alpha_{0,G}^{(l)} \sim \mathcal{N} \left(\left(N + \frac{1}{\xi_{0,G}^{2(l)}} \right)^{-1} \left((\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{0,G}}^{(l)})' \mathbf{1} + \frac{\mu_{0,G}}{\xi_{0,G}^{2(l)}} \right), \left(N + \frac{1}{\xi_{0,G}^{2(l)}} \right)^{-1} \right);$
- If both global parameters are contained in the model:
 $\underline{\alpha}_G^{(l)} \sim \mathcal{N} \left(\underline{\mu}_{\alpha_G}^{(l)}, \underline{\Sigma}_{\alpha_G}^{(l)} \right)$ with
 $\underline{\mu}_{\alpha_G}^{(l)} = (\mathbf{X}'_G \mathbf{X}_G + \Sigma_G^{-1(l)})^{-1} (\mathbf{X}'_G (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_G}^{(l)}) + \Sigma_G^{-1(l)} \underline{\mu}_G)$
 $\underline{\Sigma}_{\alpha_G}^{(l)} = (\mathbf{X}'_G \mathbf{X}_G + \Sigma_G^{-1(l)})^{-1};$

3.4.4 Monte Carlo Estimates

We first comment on posterior inference for a vector-valued parameter γ . An estimate for the posterior selection probability $p(\gamma_{i,k} = 1 | \mathbf{y})$ for the k -th regressor at voxel i can be

obtained in two ways. Assume that $\{\gamma^{(l)}; l = 1, \dots, L\}$ and $\{\theta^{(l)}; l = 1, \dots, L\}$ are series of Monte Carlo samples from the posterior distribution. On the one hand, an estimate can be obtained by averaging over the MCMC sample of the corresponding activation probability (Smith and Fahrmeir, 2007), i.e.

$$\hat{p}(\gamma_{i,k} = 1|\mathbf{y}) = \frac{1}{L} \sum_{l=1}^L p(\gamma_{i,k} = 1 | \gamma_{j \neq i}^{(l)}, \theta^{(l)}, \mathbf{y}).$$

On the other hand, the $\{\gamma_{i,k}^{(l)}; l = 1, \dots, L\}$ series is per construction a sample from the desired marginal posterior distribution of $\gamma_{i,k}|\mathbf{y}$, so that the relative frequency $\frac{1}{L} \sum_{l=1}^L \gamma_{i,k}^{(l)}$ is an estimate for $p(\gamma_{i,k} = 1|\mathbf{y})$ as well.

In case of (3.2), when variable selection is confined to a scalar voxelwise activation indicator γ_i , estimators are constructed likewise. Voxelwise estimates of posterior selection probabilities are of special interest, because these can be thresholded to provide an activation map.

Posterior inference for the regression coefficients β_i and variances σ_i^2 , $i = 1, \dots, N$ can be based on mixture estimates of the mean of corresponding marginal posterior distributions. These estimates rely on the theorem of iterated expectations and are in analogy to “model average” estimates of Raftery et al. (1997). In particular, as in Smith and Fahrmeir (2007), it holds

$$E(\beta_i|\mathbf{y}) = \sum_{\gamma_i} E(\beta_i|\gamma_i, \mathbf{y}_i)p(\gamma_i|\mathbf{y}), \quad (3.26)$$

$$E(\sigma_i^2|\mathbf{y}) = \sum_{\gamma_i} E(\sigma_i^2|\gamma_i, \mathbf{y}_i)p(\gamma_i|\mathbf{y}). \quad (3.27)$$

Closed form expressions for estimates of $E(\beta_i|\gamma_i, \mathbf{y}_i)$ and of $E(\sigma_i^2|\gamma_i, \mathbf{y}_i)$ exist for a given value of γ_i , so there is no need to estimate these means from an MCMC sample for these parameters. Specifically, $E(\beta_i|\gamma_i, \mathbf{y}_i) = \hat{\beta}_i(\gamma_i)$ and $E(\sigma_i^2|\gamma_i, \mathbf{y}_i) = S_i(\gamma_i)/(T - 2)$. See Appendix A.2 for the derivation of these components.

If γ_i is vector-valued and can take 2^p different forms, an model average estimate for β_i (Smith and Fahrmeir, 2007) is

$$\hat{\beta}_i = \frac{1}{L} \sum_{l=1}^L \hat{\beta}_i(\gamma_i^{(l)}),$$

which is based on approximation of the marginal mean by

$$E(\boldsymbol{\beta}_i|\mathbf{y}) \approx \frac{1}{L} \sum_{l=1}^L E(\boldsymbol{\beta}_i|\boldsymbol{\gamma}_i^{(l)}, \mathbf{y}_i).$$

Likewise, an model average estimate for σ_i^2 is

$$\hat{\sigma}_i^2 = \frac{1}{L} \sum_{l=1}^L S_i(\boldsymbol{\gamma}_i^{(l)})/(T-2),$$

which is based on

$$E(\sigma_i^2|\mathbf{y}) \approx \frac{1}{L} \sum_{l=1}^L E(\sigma_i^2|\boldsymbol{\gamma}_i^{(l)}, \mathbf{y}_i).$$

If $\boldsymbol{\gamma}_i$ comprises just one scalar γ_i as in (3.2), the computational burden of calculating the conditional means is further reduced by using Equations (3.26) and (3.27) directly:

$$\begin{aligned} E(\boldsymbol{\beta}_i|\mathbf{y}) &\approx \hat{\boldsymbol{\beta}}_i(\gamma_i = 1)\hat{p}(\gamma_i = 1|\mathbf{y}) + \hat{\boldsymbol{\beta}}_i(\gamma_i = 0)\hat{p}(\gamma_i = 0|\mathbf{y}) \\ E(\sigma_i^2|\mathbf{y}) &\approx \frac{S_i(\gamma_i = 1)}{T-2}\hat{p}(\gamma_i = 1|\mathbf{y}) + \frac{S_i(\gamma_i = 0)}{T-2}\hat{p}(\gamma_i = 0|\mathbf{y}) \end{aligned}$$

with $\hat{p}(\gamma_i = 1|\mathbf{y}) = 1 - \hat{p}(\gamma_i = 0|\mathbf{y})$ being an estimate for $p(\gamma_i = 1|\mathbf{y})$ as given above. These two estimates can be evaluated after the last iteration of the Gibbs sampling algorithm.

3.4.5 Classification of voxels

As noted at the beginning of this chapter, the main focus of an fMRI study lies on the identification of activated voxels. To this end, we are interested in binary maps with voxelwise activation classifications. Thereby, for each voxel a classification into active versus non-active has to be made according to a classification rule.

The Bayes decision rules are based on posterior probabilities for classes resp. models. In our case, class probabilities are the voxelwise posterior activation probabilities $p(\gamma_i|\mathbf{y})$ —which can be seen as the continuous output from a discrimination algorithm. Finding an optimal classification rule corresponds to finding an optimal threshold for dichotomizing this output.

In the context of variable selection in Gaussian regression models, Barbieri and Berger (2004) derived a classification rule for an optimal predictive model selection based on squared error loss. They showed that the median probability model, i.e. a model consisting of variables with posterior selection probabilities of at least 0.5 only, is the best predictive model (given some regularity conditions). This implies for our model that stimuli regressors should enter the model for voxel i if the corresponding posterior activation probability $p(\gamma_i|\mathbf{y})$ is larger than a threshold of 0.5 to obtain the best predictive fMRI regression model.

If evaluating classification rules, however, the squared error loss does not appear to be a suitable choice. An alternative criterion for classification algorithms is the 0-1 error loss—counting the classification errors made in total. Nevertheless, it can be shown that also the Bayes decision rule classifying a voxel as active with $p(\gamma_i|\mathbf{y}) > 0.5$ is optimal concerning the 0-1 error loss, i.e. it has the smallest expected error count (Fahrmeir et al., 1996).

This rule puts equal weight on false positives and false negatives. If requested otherwise, an optimal threshold search can be based upon statistical tools from biometrical diagnostics. Then, threshold search can be based on ROC curve estimates, cost functions or other constraints like, e.g., a minimal acceptable specificity or sensitivity value.

Generally, a threshold $\neq 0.5$ can, hence, be seen as a classification rule putting unequal weight on sensitivity and specificity. Therefore, other threshold choices can as well be motivated.

The following approach from Smith et al. (2003) and Smith and Fahrmeir (2007) shows a possibility to find a decision rule being comparable to the results of classical significance tests. Such a decision rule is appealing for enabling us to compare our results to widely-used activation detection algorithms like SPM. The authors follow Raftery (1996b) and note that $-2 \log(p(\gamma_i = 0|\mathbf{y})/p(\gamma_i = 1|\mathbf{y}))$ is on the same scale as a likelihood ratio LR statistic Λ and (for calibration purposes) is distributed approximately $\chi^2(p_3)$ where p_3 is the additional number of regressors the model $\gamma_i = 1$ has compared to model $\gamma_i = 0$ (cf. Chapter 2). This relationship is based on the assumption that the posterior odd $p(\gamma_i = 0|\mathbf{y})/p(\gamma_i = 1|\mathbf{y})$ is equal resp. approximates the Bayes factor (BF)

$$\text{BF}_{01} = \frac{p(\gamma_i = 0|\mathbf{y}) p(\gamma_i = 1)}{p(\gamma_i = 1|\mathbf{y}) p(\gamma_i = 0)},$$

which in turn can be approximated by Λ . Clearly, equality of BF and the posterior odd holds for equal prior model probabilities. For the EEG-enhanced fMRI model, where prior

model probabilities depend on a probit stage with EEG information, it is not obvious whether we can act on the assumption of equal prior probabilities.

To clarify whether equal prior probabilities are given in our model, we have to decide which is the appropriate form of prior model probability. A priori our voxelwise model probabilities depend on parameters of the probit stage: $p(\gamma_i = 1|\boldsymbol{\theta})$. To calculate BF, the unconditional prior probability $p(\gamma_i = 1)$ is needed. Therefore, $p(\gamma_i = 1|\boldsymbol{\theta})$ has to be marginalized with respect to $\boldsymbol{\theta}$. For this purpose, we use the alternative probit formulation with latent Gaussian variables U_i and translate the search for $p(\gamma_i = 1)$ to the search for $p(U_i > 0)$ and calculate

$$p(U) = \int \int p(\mathbf{U}|\boldsymbol{\theta})p(\alpha_{0,G}|\xi_{0,G}^2)p(\alpha_G|\xi_G^2)p(\boldsymbol{\alpha}_0|\xi_0^2)p(\boldsymbol{\alpha}'|\xi^2)\partial(\alpha_{0,G}, \alpha_G, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}') \\ p(\xi_{0,G}^2)p(\xi_G^2)p(\xi_0^2)p(\xi^2)\partial(\xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2).$$

Without loss of generality, we use the full form of predictor. If a sub-predictor is used for the probit stage, corresponding terms can simply be removed. The inner integral can be solved by concatenating effect vectors $\tilde{\boldsymbol{\alpha}} = (\alpha_{0,G}, \alpha_G, \boldsymbol{\alpha}'_0, \boldsymbol{\alpha}')'$ and recognizing a multivariate normal distribution yielding (see Appendix A.3 for details)

$$p(U) = \int C(\xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2) \exp\left(-\frac{1}{2}\mathbf{U}'[\mathbf{I} - \tilde{\mathbf{X}}(\tilde{\mathbf{X}}'\tilde{\mathbf{X}} + \tilde{\mathbf{Q}})^{-1}\tilde{\mathbf{X}}']\mathbf{U}\right) \\ p(\xi_{0,G}^2)p(\xi_G^2)p(\xi_0^2)p(\xi^2)\partial(\xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2), \quad (3.28)$$

where $C(\xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2)$ is a term depending on constants and probit effect variance parameters, $\tilde{\mathbf{X}}$ is the design matrix of the full predictor with effect vector $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\mathbf{Q}} = \text{diag}(1/\xi_{0,G}^2, 1/\xi_G^2, \mathbf{Q}/\xi_0^2, \mathbf{Q}/\xi^2)$ is the combined prior precision matrix of $\tilde{\boldsymbol{\alpha}}$. Note that, for this result, we use prior means equal to zero (resp. zero vectors) for probit effects. The resulting integral cannot be solved analytically. If the exact value of the marginal $p(U_i > 0)$ probability is needed, we can either calculate it by Monte Carlo integration or deduce it from the form of (3.28). From (3.28) we see, in fact, that marginalizing with respect to variance parameters affects the variance-covariance matrix of \mathbf{U} , but not the mean structure. Hence, the marginalized mean of \mathbf{U} is the zero vector. Therefore, we conclude that $p(U_i > 0) = 0.5 \forall i$ for zero prior mean of probit effects and thus, that the BF is the quotient of posterior probabilities.

With this, we can try to apply the approach of Smith and Fahrmeir (2007) and derive a classification threshold that is calibrated to a significance level of 0.05.

The statement “being on the same scale as the likelihood ratio statistic” of Raftery (1996b) can easily be misunderstood. It does not necessarily mean that we can simply equate $-2 \log(\text{BF}_{01})$ with the critical value of Λ , which is a mistake found in the literature. Raftery (1996a) gives approximations of the BF based on the LR statistic. For each given approximation, an additional term has to be taken into account.

Belonging to the model class of (generalized) linear models, we can use the approximations from Raftery (1996a). In the following, we elaborate on approximating BFs for our combined fMRI-EEG model. As can be seen below, these BF approximations can be derived for one voxelspecific regression model and then be applied voxelwise. This is due, on the one hand, to the conditional independence of voxelspecific models in our complete fMRI-EEG model and, on the other hand, to the specific prior and posterior mean structure of fMRI effect parameters. Therefore, we drop in the following the index i for simplicity and denote $\beta(\gamma)$ (as used in Section 3.2) as β_γ .

Let the prior mean of regression effect parameters be $E(\beta_\gamma|\gamma) = \omega_\gamma$ and the corresponding prior covariance matrix be $\text{Var}(\beta_\gamma|\gamma) = \mathbf{W}_\gamma$. Additionally, $\mathbf{G}_\gamma = \mathbf{W}_\gamma^{-1}$ and \mathbf{F}_γ is the observed or expected Fisher information matrix. Then Raftery (1996a) shows that

$$-2 \log \text{BF}_{01} \approx \Lambda + (E_1 - E_0),$$

where $\Lambda = 2 \left(l_1(\hat{\beta}_1) - l_0(\hat{\beta}_0) \right)$ is the standard likelihood ratio test statistic with $l_\gamma(\beta_\gamma) = \log p(\mathbf{y}|\beta_\gamma, \gamma)$ being the log likelihood for model with $\gamma = 1$ or $\gamma = 0$ and in the case of normal priors

$$E_\gamma = \log |\mathbf{G}_\gamma| - (\hat{\beta}_\gamma - \omega_\gamma)' \mathbf{C}_\gamma (\hat{\beta}_\gamma - \omega_\gamma) - \log |\mathbf{F}_\gamma + \mathbf{G}_\gamma|, \quad (3.29)$$

where $\mathbf{C}_\gamma = \mathbf{G}_\gamma (\mathbf{I} - \mathbf{H}_\gamma (2 - \mathbf{F}_\gamma \mathbf{H}_\gamma) \mathbf{G}_\gamma)$ and $\mathbf{H}_\gamma = (\mathbf{F}_\gamma + \mathbf{G}_\gamma)^{-1}$.

From Equation (3.3) and Section 3.4.4, we see that the prior mean and the conditional posterior mean estimate coincide, so that the quadratic form in (3.29) is zero. The inverse prior covariance is $\mathbf{G}_\gamma = \frac{1}{T\sigma^2} \mathbf{X}(\gamma)' \mathbf{X}(\gamma)$ and the (expected and observed) Fisher information matrix is $\mathbf{F}_\gamma = \frac{1}{\sigma^2} \mathbf{X}(\gamma)' \mathbf{X}(\gamma)$. Hence, we have

$$E_\gamma = \log |\mathbf{G}_\gamma| - \log |\mathbf{F}_\gamma + \mathbf{G}_\gamma| = \log |\mathbf{G}_\gamma (\mathbf{F}_\gamma + \mathbf{G}_\gamma)^{-1}| = \log \left| \left(\frac{1}{T+1} \right) \mathbf{I} \right| = -p_\gamma \log(T+1)$$

where p_γ is the number of effects in model with indicator value γ .

Hence, we obtain the approximation

$$-2 \log \text{BF}_{01} \approx \Lambda - p_{\Delta} \log(T + 1),$$

where p_{Δ} is the difference in parameter numbers, i.e. in the notation of our model: $p_{\Delta} = p_3$. This approximation is approximately equal to the *Schwarz criterion* derived in Schwarz (1978).

With this, a threshold for classification can be derived for a critical value κ_{α} of the likelihood ratio test with significance level α as

$$\frac{\exp(0.5(\kappa_{\alpha} - p_{\Delta} \log(T + 1)))}{1 + \exp(0.5(\kappa_{\alpha} - p_{\Delta} \log(T + 1)))}.$$

Inserting realistic values ($\alpha = 0.05$, $T = 300$ and $p_{\Delta} = 3$) into this equation yields, however, a strange threshold below 0.01. Hence, the given approximation leads to an inconsistent model selection criterion always favoring the more complex model.

It is a known issue that there are difficulties in objective Bayesian model selection procedures based on both BFs or their approximations. Stone (1979) shows that, for instance, the information criterion BIC—which is closely related to the Schwarz criterion—can lead to an inconsistent model choice. Berger and Pericchi (2001) discuss difficulties arising in model selection procedures directly based on BFs: Both noninformative and vague priors can lead to arbitrary model choices.

Note that the g -prior (Equation (3.3)) as used for β is vague and the choice of multiplication factor T for the variance is quite arbitrary. As a limiting case ($T \rightarrow \infty$), the g -prior is noninformative. Both properties of the g -prior can be a crucial factor for problems with a model selection based on derived BFs. Being an essential component of the combined fMRI-EEG model—enabling us to marginalize with respect to fMRI regression parameter—the g -prior cannot easily be replaced by other types of priors. Being beyond the scope of this thesis, we decided to abandon the search for an optimal classification threshold at this point and use a conservative threshold of 0.8722, which has been derived and used in Smith and Fahrmeir (2007). This threshold can be calculated by equating $-2 \log(\text{BF}_{01})$ with the critical value $\kappa_{0.05}$ of Λ with test distribution $\chi(1)$ and solving for the posterior activation probability. To evaluate the robustness of corresponding activation maps, we also consider a small set of further thresholds: 0.5, 0.75, 0.95.

3.5 A model extension: A non-negative global EEG effect

First analysis results show that models with predictor type 2 (cf. Table 3.1) may yield a negative global EEG effect. If a spatially-varying EEG effect is used, we allow local EEG effects to be negative for indicating regions with non-congruent EEG and fMRI information. A negative global EEG effect, however, leads to the counterintuitive interpretation that EEG generally prohibits fMRI activation. Hence, we think it is sensible to introduce a restriction on the global EEG coefficient in predictor type 2 models. For being independent of the exact choice of spatially-varying effect priors, the following restriction can be used in both CAR and IGMRF models—as well as in the fast version of the CAR model (cf. Chapter 4).

We constrain the global EEG coefficient α_G to be non-negative by using a log normal prior, i.e. $\alpha_G \sim \mathcal{LN}(\mu_G, \xi_G^2)$. With this the full conditional distribution has the following form:

$$\begin{aligned} p(\alpha_G | \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\alpha_G}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) &\propto p(\alpha_G, \boldsymbol{\theta}_{-\alpha_G}, \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2 | \mathbf{y}) \propto p(\mathbf{U} | \boldsymbol{\eta}) p(\alpha_G | \mu_G, \xi_G^2) \\ &\propto \exp\left(-\frac{1}{2}(\mathbf{U} - \boldsymbol{\eta}_{-\alpha_G} - \mathbf{j}\alpha_G)'(\mathbf{U} - \boldsymbol{\eta}_{-\alpha_G} - \mathbf{j}\alpha_G)\right) \cdot \frac{1}{\alpha_G} \exp\left(-\frac{1}{2\xi_G^2}(\log(\alpha_G) - \mu_G)^2\right) \\ &\propto \frac{1}{\alpha_G} \exp\left(-\frac{1}{2}\left[\alpha_G^2 \mathbf{j}'\mathbf{j} - 2(\mathbf{U} - \boldsymbol{\eta}_{-\alpha_G})'\mathbf{j}\alpha_G + \frac{\log(\alpha_G)^2}{\xi_G^2} - 2\frac{\log(\alpha_G)\mu_G}{\xi_G^2}\right]\right). \end{aligned}$$

This form cannot be completed to a known distribution from which random numbers can easily be drawn. Therefore, we have to change over to a Metropolis-Hasting step. As proposal density, we take a distribution with support on positive α_G values. We decided to use a log-normal distribution with density

$$J(\alpha_{G,new} | \alpha_{G,old}) \propto \frac{1}{\alpha_{G,new}} \exp\left(-\frac{1}{2\xi_{prop,G}^2}(\log(\alpha_{G,new}) - \log(\alpha_{G,old}))\right).$$

This density is not symmetric in $\alpha_{G,new}$ and $\alpha_{G,old}$. Hence, we have to account for this in the Metropolis-Hastings step. The corresponding acceptance probability can be derived as

$$\begin{aligned} \alpha(\alpha_{G,new} | \alpha_{G,old}) &= \min\left\{\frac{p(\boldsymbol{\gamma}, \boldsymbol{\theta}_{-\alpha_G}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \alpha_{G,new} | \mathbf{y}) / J(\alpha_{G,new} | \alpha_{G,old})}{p(\boldsymbol{\gamma}, \boldsymbol{\theta}_{-\alpha_G}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \alpha_{G,old} | \mathbf{y}) / J(\alpha_{G,old} | \alpha_{G,new})}, 1\right\} \\ &= \min\left\{\frac{\exp\left(-\frac{1}{2}\left[\alpha_{G,new}^2 \mathbf{j}'\mathbf{j} - 2(\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{G,new}})'\mathbf{j}\alpha_{G,new} + \frac{\log(\alpha_{G,new})^2}{\xi_G^2} - 2\frac{\log(\alpha_{G,new})\mu_G}{\xi_G^2}\right]\right)}{\exp\left(-\frac{1}{2}\left[\alpha_{G,old}^2 \mathbf{j}'\mathbf{j} - 2(\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{G,old}})'\mathbf{j}\alpha_{G,old} + \frac{\log(\alpha_{G,old})^2}{\xi_G^2} - 2\frac{\log(\alpha_{G,old})\mu_G}{\xi_G^2}\right]\right)}, 1\right\}. \end{aligned}$$

The update for the related variance parameter ξ_G^2 changes as follows:

$$\begin{aligned}
 p(\xi_G^2 | \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\xi_G^2}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) &\propto p(\alpha_G | \mu_g, \xi_G^2) p(\xi_G^2) \\
 &\propto \frac{1}{\sqrt{2\pi\xi_G^2} \alpha_G} \exp\left(-\frac{1}{2\xi_G^2} (\log(\alpha_G) - \mu_G)^2\right) (\xi_G^2)^{-(a_G+1)} \exp\left(-\frac{b_G}{\xi_G^2}\right) \\
 &\propto (\xi_G^2)^{-(a_G+0.5+1)} \exp\left(-\frac{1}{\xi_G^2} ((\log(\alpha_G) - \mu_G)^2/2 + b_G)\right).
 \end{aligned}$$

This can be completed to an inverse gamma distribution with shape parameter $a_G + 1$ and scale parameter $(\log(\alpha_G) - \mu_G)^2/2 + b_G$, which can be used as the full conditional distribution for updating ξ_G^2 in a Gibbs sampling step.

4 A fast CAR model for the EEG-enhanced fMRI model

Depending on the number of analyzed voxels, the computation time for running the model proposed in Chapter 3 can become very long. The most time-consuming steps are the simultaneous updates of the whole spatially-varying EEG coefficient vector and the spatially-varying intercept vector. The calculation of one Cholesky decomposition of a large update matrix can take about 2.5 seconds with a brain cuboid of 65 280 voxels in the application presented in Chapter 7, which leads to a total running time of the MCMC algorithm of about 4.2 h (in the IGMRF prior case updating one spatially-varying coefficient in 6 000 iterations). When using the CAR model, updates of the spatial dependency parameters add likewise to the increase of running time. Acceptance probabilities of the corresponding Metropolis-Hastings step need the evaluation of the determinant of the CAR precision matrix, which we currently also base on a Cholesky decomposition with same time demand. Thus, running time is doubled. Generally, computation time increases superlinearly with the number of voxels.

Computation time can for example be decreased by incorporating an analyzing mask, i.e. a binary image indicating a subset of voxels that should be analyzed. Analysis can for example be restricted to voxels within gray-matter—where activation is assumed to happen. Segmentation procedures (Friston et al., 2008, pp. 81) can provide gray-matter probabilities for voxels in an fMRI dataset. If we restrict analysis to voxels with a gray-matter probability larger than 0.5, the number of voxels in the given example (Chapter 7) reduces to about 16 500. Thus, one iteration in the MCMC update scheme takes about 0.2 seconds decreasing the overall computation time of a full CAR model with EEG information to about 30 min. In Figure 4.1, the resulting analyzing mask is plotted for the discussed example. Note that the resulting mask corresponds to a connected analyzing graph, i.e. all pairs of brain voxels (vertices) are connected by a path. This is necessary for isolated voxel groups being able to distort the estimation process. Nevertheless, the resulting mask exhibits many small gaps in otherwise continuous gray-matter tissues.

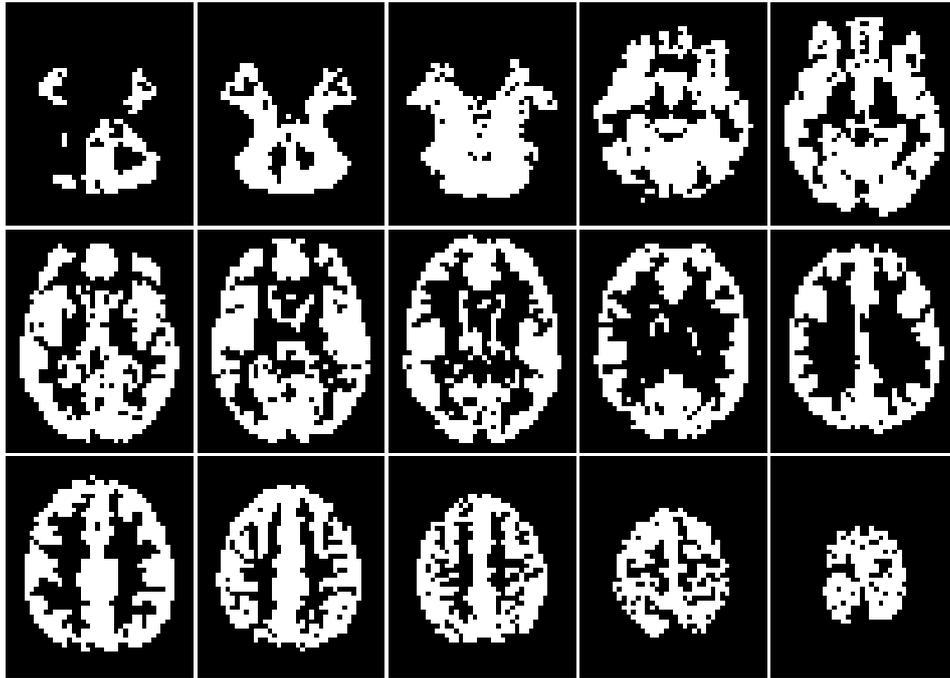


Figure 4.1: Gray-matter analyzing mask for fMRI image layers 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31. White voxels denote voxels with a gray-matter probability larger than 0.5.

Alternatively, an analyzing mask can be derived depending on a specified threshold. This selection mechanism is used by software like SPM. Here, a voxel is selected for analysis when all of its fMRI time series values survive a grand mean dependent threshold. Details of this procedure can be found below (Section 4.3). As before, we deselect voxels not connected to the main brain region. The resulting mask exhibits hardly any discontinuities (cf. Figure 4.2). In total, this threshold mask selects 27 300 voxels increasing the time demand of one MCMC iteration of a full CAR model with EEG information to 1.8 seconds, which, in turn, leads to an overall computation time of about 3 h.

To circumvent the derivation of suitable analyzing masks, which themselves may be prone to estimation error, we propose an alternative strategy to fasten the fMRI-EEG procedure. As noted above, both the random field updates for the spatially-varying coefficients and the determinant calculation in the update of the spatial dependency updates slow down MCMC iterations. To speed-up the algorithm, both steps must be accelerated. Speeding-up random field updates can be accomplished by single-site updates. This idea is discussed in Section 4.1. Different approaches exist for calculating the determinant of large sparse

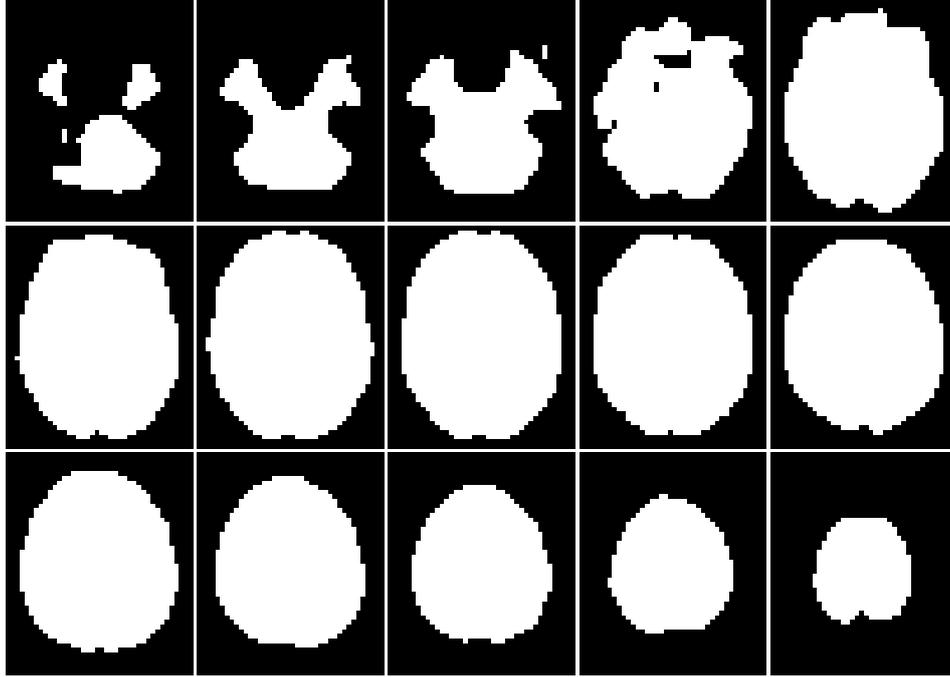


Figure 4.2: Threshold based analyzing mask for fMRI image layers 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31. White voxels denote voxels with fMRI time series values surviving a grand mean dependent threshold.

symmetric positive matrices. We look at different proposals and review the most promising one in Section 4.2. In Section 4.3, we explain how the combined fMRI-EEG algorithm of Section 3.4.3 has to be adapted to be accelerated.

4.1 Single site updates for spatially-varying coefficients

Instead of using a global update of the whole varying coefficient vector based on full conditionals (3.22) and (3.23), we derive full conditionals for the single site updates.

For the EEG coefficient α_i at voxel i , the full conditional can be derived as:

$$\begin{aligned}
 p(\alpha_i | \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\alpha_i}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) &\propto p(\alpha_i, \boldsymbol{\theta}_{-\alpha_i}, \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2 | \mathbf{y}) \\
 &\propto p(\mathbf{U} | \boldsymbol{\eta}) \underbrace{p(\alpha_i | \boldsymbol{\alpha}_{j \neq i}, \xi^2, \tau^2) p(\boldsymbol{\alpha}_{j \neq i} | \xi^2, \tau^2)}_{=p(\boldsymbol{\alpha} | \xi^2, \tau^2)} \propto p(\mathbf{U} | \boldsymbol{\eta}) p(\alpha_i | \boldsymbol{\alpha}_{j \neq i}, \xi^2, \tau^2) \\
 &\propto \exp \left(-\frac{1}{2} (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_i} - \mathbf{X}_{J_i} \alpha_i)' (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_i} - \mathbf{X}_{J_i} \alpha_i) \right) p(\alpha_i | \boldsymbol{\alpha}_{j \neq i}, \xi^2, \tau^2)
 \end{aligned}$$

$$\begin{aligned}
&\propto \exp \left(-\frac{1}{2} \left[\alpha_i^2 J_i^2 - 2\alpha_i \mathbf{X}'_{J_i} (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_i}) + \left(\frac{1 + \tau^2 n_i}{\xi^2} \right) \alpha_i^2 - 2\alpha_i \frac{\tau^2}{\xi^2} \sum_{j \in \partial_i} \alpha_j \right] \right) \\
&= \exp \left(-\frac{1}{2} \left[\alpha_i^2 \left(J_i^2 + \frac{1 + \tau^2 n_i}{\xi^2} \right) - 2\alpha_i \left(\mathbf{X}'_{J_i} (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_i}) + \frac{\tau^2}{\xi^2} \sum_{j \in \partial_i} \alpha_j \right) \right] \right),
\end{aligned}$$

where $\mathbf{X}'_{J_i} = (0, \dots, 0, J_i, 0, \dots, 0)$ is a vector with zeros except element i , which equals J_i . Hence, the full conditional for the i -th EEG coefficient is

$$\alpha_i | \cdot \sim \mathcal{N} \left(\frac{\xi^2}{J_i^2 \xi^2 + 1 + \tau^2 n_i} \left(\mathbf{X}'_{J_i} (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_i}) + \frac{\tau^2}{\xi^2} \sum_{j \in \partial_i} \alpha_j \right), \frac{\xi^2}{J_i^2 \xi^2 + 1 + \tau^2 n_i} \right). \quad (4.1)$$

The full conditional for the intercept at voxel i can be derived as

$$\begin{aligned}
p(\alpha_{0,i} | \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\alpha_{0,i}}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{y}) &\propto p(\mathbf{U} | \boldsymbol{\eta}) p(\alpha_{0,i} | \boldsymbol{\alpha}_{j \neq i}, \xi_0^2, \tau_0^2) \\
&= \exp \left(-\frac{1}{2} (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{0,i}} - \mathbf{X}_{0,i} \alpha_{0,i})' (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{0,i}} - \mathbf{X}_{0,i} \alpha_{0,i}) \right) p(\mathbf{U} | \boldsymbol{\eta}) p(\alpha_{0,i} | \boldsymbol{\alpha}_{j \neq i}, \xi_0^2, \tau_0^2) \\
&\propto \exp \left(-\frac{1}{2} \left[\alpha_{0,i}^2 \left(1 + \frac{1 + \tau_0^2 n_i}{\xi_0^2} \right) - 2\alpha_{0,i} \left(\mathbf{X}'_{0,i} (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{0,i}}) + \frac{\tau_0^2}{\xi_0^2} \sum_{j \in \partial_i} \alpha_{0,j} \right) \right] \right),
\end{aligned}$$

where $\mathbf{X}'_{0,i} = (0, \dots, 0, 1, 0, \dots, 0)$ is the i -th unit vector. Hence, the full conditional for the i -th intercept term is

$$\alpha_{0,i} | \cdot \sim \mathcal{N} \left(\frac{\xi_0^2}{\xi_0^2 + 1 + \tau_0^2 n_i} \left(\mathbf{X}'_{0,i} (\mathbf{U} - \boldsymbol{\eta}_{-\alpha_{0,i}}) + \frac{\tau_0^2}{\xi_0^2} \sum_{j \in \partial_i} \alpha_{0,j} \right), \frac{\xi_0^2}{\xi_0^2 + 1 + \tau_0^2 n_i} \right). \quad (4.2)$$

4.2 Calculating the determinant of large sparse update matrices

4.2.1 Overview of existing methods

The Metropolis step for updating the covariance parameter τ^2 of the CAR model involves the evaluation of the determinant of the precision matrix, $|\mathbf{P}| = |\mathbf{I} + \tau^2 \mathbf{Q}|$, where \mathbf{I} is the N -dimensional identity matrix and \mathbf{Q} is the Laplacian matrix of the underlying graph describing the 3D voxel grid (for more details on this see Section 4.2.2). The determinant $|\mathbf{P}|$ has to be recalculated for every update of τ^2 . This can be done via its Cholesky

decomposition $\mathbf{P} = \mathbf{L}\mathbf{L}'$ as suggested in Rue (2001) using $|\mathbf{P}| = |\mathbf{L}||\mathbf{L}'| = |\mathbf{L}|^2 = \left(\prod_{i=1}^N l_{ii}\right)^2$, where l_{ii} are the diagonal elements of \mathbf{L} . For large datasets, numerical stability is achieved by using $\log |\mathbf{P}| = 2 * \sum_{i=1}^N \log(l_{ii})$. Calculating the Cholesky decomposition in every iteration is very time-consuming with time growing superlinearly with N . Hence, Pettitt et al. (2002) and Smith and Smith (2006) suggest to calculate the eigenvalues $\lambda_1, \dots, \lambda_N$ of $-\mathbf{Q}$ once prior to MCMC iterations to yield an efficient evaluation scheme using

$$|\mathbf{P}| = \prod_{i=1}^N (1 - \tau^2 \lambda_i)$$

or rather

$$\log |\mathbf{P}| = \sum_{i=1}^N \log(1 - \tau^2 \lambda_i).$$

However, calculation of all eigenvalues of the sparse Laplacian matrix \mathbf{Q} , resp. $-\mathbf{Q}$, for general graphs present in brain imaging is problematic for large N . Iterative solutions to the eigenvalue problem aim at calculating only a small subset of eigenvalues. See, for example Bai et al. (2000), Chapter 4.1, for an overview of several available algorithms and parts of the spectrum aimed at. In addition, even recently developed algorithms need a reasonable amount of time for calculating a relatively small number—i.e. in the order of hundreds—of eigenvalues in large scale problems (Arbenz et al., 2005). Hence, this limits their usefulness in speeding up MCMC brain analysis.

There is an exception where the eigenspectrum of the Laplacian matrix of the voxel-grid is analytically tractable. If analysis is refined to a rectangular voxel-slice or voxel-cube, explicit formulas are available. Results are based on the Cartesian sum of path graphs with known spectra and are a standard result of algebraic graph theory. A derivation can, for example, be found in Mohar (2004) and is summarized in Section 4.2.2.

Simpson et al. (2008) mentions different alternative approaches to approximate the determinant appearing in the literature: Methods are based on sparse approximating inverses (Reusken, 2001), Gauss quadrature (Bai et al., 1996) and diagonal approximations (Ipsen and Lee, 2003). Simpson et al. (2008) also states that the method of Bai et al. (1996) is the most popular one. It has to be evaluated which approach is the fastest and can be implemented with acceptable time and effort at the same time. For a moment, we do not pursue one of these approaches.

For a start, we take a closer look at the eigenspectrum approach. There are in principle

two conceivable ways for applying it to brain analysis. If analysis should be restricted to an individual head shape, we can take the Laplacian spectrum of the enveloping brain cuboid to approximate the determinant of \mathbf{P} . On the other hand, we can analyze all voxels of the brain cuboid and calculate the Laplacians exact spectrum to calculate $|\mathbf{P}|$.

It may seem reasonable to remove voxels outside the brain. This enhances the performance of analyzing procedures (1) for reducing the memory needed to store all generated objects, (2) for speeding up computations which have a demand of larger than $O(N)$ and (3) for supporting smoothing procedures, which can be deteriorated by smoothing over non-brain voxels. However, if we approximate $|\mathbf{P}|$ by taking the Laplacian spectrum of an enveloping brain cuboid, we must guarantee that this approximation does not deteriorate MCMC inference.

For our implementation, we prefer to realize the second approach building upon a cubical analyzing mask. With the single-site update proposed in Section 4.1, all MCMC update parts are $O(N)$, so that an increase in voxel numbers does not interfere substantially with the speed-up. When smoothing over non-brain voxels, we rely on the good edge-preserving properties of the CAR prior.

For this purpose, we derive the Laplacian spectrum of a 3D grid in the following section.

4.2.2 Laplace eigenspectrum of grids

The following derivation of Laplacian eigenvalues of q -dimensional grids can be found in Mohar (2004).

Let $G = (V, E)$ be a non-directed graph with finite vertex set $V = 1, 2, \dots, N$ and edge set E of two-element subsets of V . The degree n_i of vertex i is the number of edges incident to vertex i . Let $\mathbf{D} = \text{diag}(n_1, \dots, n_N)$ be the diagonal matrix of vertex degrees. The Laplacian matrix is $\mathbf{Q} = \mathbf{D} - \mathbf{A}$, where \mathbf{A} is the (0,1)-adjacency matrix of the graph. The Laplacian \mathbf{Q} is a real symmetric matrix, which can be shown to be positive semi-definite. It therefore has N non-negative real eigenvalues $\lambda_i = \lambda_i(G)$, $i = 1, \dots, N$. For the row sums being equal to 0, the vector $\mathbf{1} = (1, 1, \dots, 1)'$ is an eigenvector with corresponding eigenvalue 0. By convention, let

$$0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$$

be the eigenvalues in ascending order.

Based on these definitions, a 3D voxel grid in brain analysis is built up as follows. Each analyzed brain voxel denotes a vertex in the grid. Two voxel vertices are connected by an edge if they are declared as neighbors by prior believes—as defined by the choice of adjacency matrix \mathbf{A} that is incorporated in the precision matrix of multivariate prior distributions (e.g. in the IGMRF precision matrix \mathbf{Q} or in the CAR precision matrix \mathbf{P}).

The following theorem is useful for determining the number of connected brain regions in a brain imaging example: The multiplicity of 0 as an eigenvalue of \mathbf{Q} is equal to the number of connected components of G . Hence, 0 is a simple eigenvalue of \mathbf{Q} if and only if the graph G is connected. Applied to brain imaging analysis, a multiplicity of $k > 1$ of eigenvalue 0 implies the brain voxel grid is decomposed into k subgraphs. Especially, if small brain parts become disconnected to the main brain grid, effect estimation based on neighborhood information can be deteriorated.

The Laplacian spectrum of a q -dimensional grid graph builds upon the Cartesian product of path graphs. A path of length $n - 1$ ($n \geq 2$), denoted by P_n , is a graph with n vertices, say x_1, \dots, x_n , and with $n - 1$ edges in which x_i and x_{i+1} are connected by an edge for $i = 1, \dots, n - 1$. The Laplace eigenvalues of P_n are

$$\lambda_i^{(n)} = 4 \sin^2 \left(\frac{\pi(i-1)}{2n} \right), \quad i = 1, \dots, n.$$

The Cartesian product $G_1 \times G_2$ of two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ has vertex set $V(G_1 \times G_2) = V(G_1) \times V(G_2)$, where vertice (i_1, j_1) is adjacent to (i_2, j_2) if and only if $i_1 = i_2$ and $\{j_1, j_2\}$ is an edge in G_2 , or $j_1 = j_2$ and $\{i_1, i_2\}$ is an edge in G_1 .

The Laplace eigenvalues of the Cartesian product $G_1 \times G_2$ are

$$\lambda_i(G_1) + \lambda_j(G_2)$$

for $i = 1, 2, \dots, |V(G_1)|$ and $j = 1, 2, \dots, |V(G_2)|$, where $|V(G_k)|$ is the cardinality of set $V(G_k)$, $k = 1, 2$.

Applied to two path graphs P_n and P_m of length n and m , respectively, the $n \times m$ grid graph $P_n \times P_m$ has Laplace eigenvalues

$$\lambda_{i,j}^{(n \times m)} = \lambda_i^{(n)} + \lambda_j^{(m)} = 4 \sin^2 \left(\frac{\pi(i-1)}{2n} \right) + 4 \sin^2 \left(\frac{\pi(j-1)}{2m} \right),$$

$$i = 1, \dots, n, \quad j = 1, \dots, m,$$

which applies in brain imaging to the spectrum of a Laplacian matrix of a two-dimensional voxel slice of size $n \times m$. Reiterating this procedure to a third dimension gives the Laplacian spectrum of a three-dimensional $n \times m \times k$ grid as

$$\lambda_{i,j,l}^{(n \times m \times k)} = \lambda_i^{(n)} + \lambda_j^{(m)} + \lambda_l^{(k)} = 4 \sin^2 \left(\frac{\pi(i-1)}{2n} \right) + 4 \sin^2 \left(\frac{\pi(j-1)}{2m} \right) + 4 \sin^2 \left(\frac{\pi(l-1)}{2k} \right),$$

$$i = 1, \dots, n, j = 1, \dots, m, l = 1, \dots, k, \quad (4.3)$$

which are the Laplace eigenvalues of $P_n \times P_m \times P_k$, i.e. a graph of a 3D cuboid with side length $n \times m \times k$.

4.3 An accelerated fMRI-EEG algorithm

To implement determinant calculation via the Laplacian spectrum of a brain cuboid, a cubical analyzing mask has to be determined prior to MCMC iterations. For this, we first select all brain voxels and then choose the enveloping cubical mask as analyzing mask.

In our implementation we define brain voxels as voxels with all time series values larger than $1/8$ of the grand mean μ_{gm} . The grand mean is thereby calculate as in SPM as the mean of scanwise global means $\mu_{gm} = \frac{1}{T} \sum_{t=1}^T \mu_{gm,t}$. A global mean $\mu_{gm,t}$ is the mean of all voxels with signal larger than $\mu_t/8$, i.e.

$$\mu_{gm,t} = \frac{1}{n_t} \sum_{i: y_{i,t} > \mu_t/8} y_{i,t},$$

where μ_t is the overall signal mean of scan t and n_t is the number of voxels with $y_{i,t} > \mu_t/8$. Note that this is the default selection scheme for a brain mask in SPM.

Having selected all brain voxels that should be analyzed, we determine the corresponding minimum and maximum indices to define the image cuboid. Let x, y, z denote the three spatial indices, the cuboid dimensions are defined as

$$n = x_{max} - (x_{min} - 1)$$

$$m = y_{max} - (y_{min} - 1)$$

$$k = z_{max} - (z_{min} - 1).$$

Therefore, the corresponding Laplacian spectrum can be calculate via (4.3) as

$$(\lambda_i^{(+)}, i = 1, \dots, N) = (\lambda_{i,j,l}^{(n \times m \times k)}), \quad i = 1, \dots, n, \quad j = 1, \dots, m, \quad l = 1, \dots, k).$$

For calculating the determinant we need, however, the eigenvalues of $-\mathbf{Q}$ which can be simply calculated as

$$\lambda_i = -\lambda_i^{(+)}, \quad i = 1, \dots, N.$$

The MCMC algorithm proposed in Section 3.4.3 can then be accelerated by using the following adoptions. On the one hand, a for-loop over single site updates (4.1) and (4.2) in step **d** replaces global updates. On the other hand, updates of τ^2 and τ_0^2 in **c2** use the efficient determinant calculation via

$$|\mathbf{P}| = \prod_{i=1}^N (1 - \tau^2 \lambda_i)$$

and

$$|\mathbf{P}_0| = \prod_{i=1}^N (1 - \tau_0^2 \lambda_i).$$

5 Software implementation

The algorithms derived from the methodology presented in the previous chapters have been implemented and combined into an user-friendly software tool. These include the EEG-enhanced fMRI detection algorithms with IGMRF and CAR prior with all discussed predictor types (cf. Table 3.1) from Chapter 3. The CAR algorithm can be run as version with global updates or as the accelerated version from Chapter 4.3. Additionally, the fMRI activation detection algorithm based on an Ising prior from Smith et al. (2003) is implemented.

The user can choose between two different software package versions. On the one hand, all routines are callable from R (R Core Team, 2012), a software environment for statistical computing and graphics, and are available as package `Rfmrieeeg`. On the other hand, the analyzing routine can directly be compiled as a binary `CfmrieeegMain` program from its C++ source code (which can then e.g. be executed in a shell). Both software versions build upon the identical C++ algorithm collection. Whereas the main executable `CfmrieeegMain` calls provided algorithms directly within C++, the R package provides an interface so that algorithms can be accessed from within R. Besides this interface—which is the core of the R-package—`Rfmrieeeg` provides additional utilities for setting configuration arguments, output analysis and presentation of results. Both the R-package as well as the C++ source code (which is contained within the R-package source folder `src`) are provided upon request by the author. In the following, we denote parts of the implementation and algorithms common to `Rfmrieeeg` and `CfmrieeegMain` as `fmrieeeg`.

Section 5.1 is written as a manual for readers intending to use the software. We describe how to configure and run the `fmrieeeg` algorithms using a configuration text-file in Section 5.1.1. A call to the `fmrieeeg` algorithms by the corresponding `Rfmrieeeg`-procedure as well as by `CfmrieeegMain` relies on this file. In Section 5.1.2, we give details on installing the R-package `Rfmrieeeg` and demonstrate its functionality by going through the main analysis steps of an exemplary data analysis. In Section 5.1.3, information on the installation process of `CfmrieeegMain` is given as well as a short note on its use.

In Section 5.2, we shed light on the implementation of the C++ algorithms. Efficient numerical solutions are discussed to solve linear algebra problems (Section 5.2.1 and 5.2.2). Notes on the object-oriented software architecture of the C++ program can be found in Section 5.2.3.

Section 5.3 gives some more technical details for users unfamiliar with the used image data format. Additionally, some advice is given on changing compilation settings that might increase the performance of the algorithms.

In the following, we denote the `fmrieeeg` algorithms with CAR prior with global updates of spatially-varying effects from Chapter 3 as gCAR and the fast algorithms from Chapter 4 as fCAR. The model with IGMRF prior is denoted as iMRF. It relies exclusively on simultaneous map updates of map coefficients (as described in Chapter 3). We label algorithm names with numbers 1-5 to differentiate between predictor types, i.e. gCAR1-gCAR5 denotes the gCAR models with predictor 1 to 5. The same applies analogously to the fCAR model and iMRF model: Model names fCAR1-fCAR5 denote the fCAR models with predictor 1 to 5. Model names iMRF1-iMRF5 denote the iMRF models with predictor 1 to 5.

5.1 Usage

5.1.1 Preliminaries: Configuring an `fmrieeeg` algorithm run

The `fmrieeeg` C++ algorithm—the core of both the R package and the binary—can be controlled by the user via a set of parameters. Though R provides flexible means of passing arguments to functions, options are limited when executing a binary in a shell. For ease of usage, we decided to pass arguments by using a configuration file in ASCII text format in the form `'keyword = value'` (see Figure 5.1 for an example). We prefer the file ending `'.inp'` to distinguish it from other text files. The path to this file is the only direct argument the C++ program takes. Hence, we decided that the binary as well as the corresponding R-procedure are started by passing the configuration file path.

The argument values in the configuration file can be adapted to ones needs either by editing a template (made available with the source code) in a text editor or by using configuration functions from the provided R-package `Rfmrieeeg` (see next section). Some arguments are mandatory to be specified, e.g. the path to the actual data, stimulus presentation times, etc.,

```
pathResults = Z:/testsuite
t0 = 0
tr = 2
nStimTypes = 1
stimActDect = 1
nPerStimType = 36
pathStimTimes= Z:/testsuite/stim-times.txt
confounder = FALSE
pathConf =
freqHPF = 128
pathNifti = Z:/testsuite/nifti_fmrits-RS.nii

...
```

Figure 5.1: Example of a configuration ASCII text file.

whereas default values are assigned to remaining arguments if not specified otherwise. We differentiate between common arguments (needed for all model types) and arguments specific to the model choice for the Bayesian activation detection hierarchy. That is, the set of arguments depends on the choice of spatial binary prior, i.e. Ising or probit prior. For the latter, we differentiate between CAR and IGMRF prior and between predictor types. The keyword `speedup` can be used to switch between gCAR and fCAR algorithms. An overview of arguments common to all models can be found in Table C.1 and an overview of model specific arguments can be found in Table C.3 in the Appendix.

A call to the algorithms within R can be conducted by passing a string with the path to the configuration file to the `RfmrieegMain()` function:

```
R> returnObj <- RfmrieegMain("path/to/ConfigurationFile.inp")
```

The `CfmrieegMain` program is executed by specifying the path to the configuration file after the `'-config'` flag:

```
% CfmrieegMain -config path/to/ConfigurationFile.inp
```

In both cases, results are stored in a subfolder of the directory specified with keyword `pathResults` or in the working directory if left unspecified.

5.1.2 Usage of the R-package `Rfmrieeg`

For installing the R-package `Rfmrieeg`, the provided version of the package source¹ must be saved to a local folder. Having launched R and having changed the working directory to the corresponding folder (via `setwd()`), the package can easily be installed using the command:

```
R> install.packages("Rfmrieeg", repos=NULL, type="source")
```

Note that installing from source can take several minutes. If working on windows without an installation of the Rtools collection², the windows binaries in the zip format can be used and installed by using:

```
R> install.packages("Rfmrieeg", repos=NULL)
```

Note that installation from a local repository requires the R-packages `Rniftilib` (Granert, 2012), `lattice` (Plummer et al., 2006) and `coda` (Sarkar, 2008) to be already installed. In a next step, the package needs to be loaded into the current R session:

```
R> library("Rfmrieeg")
Loading required package: coda
Loading required package: lattice
Loading required package: Rniftilib
```

To view the documentation on provided functions, e.g. `RfmrieegMain`, or the package itself use the help command, for example `?RfmrieegMain` or `?Rfmrieeg`.

The package `Rfmrieeg` provides functionality for the entire data analysis process, from configuring the algorithm up to MCMC diagnostics and presentation of results.

In the following, the main functionality of `Rfmrieeg` is demonstrated with a subset of the simulated dataset from Chapter 6.1.1 with small error noise (where an error term ϵ_i^* is a realization from $\mathcal{N}(0, \sigma_i^{2,*} I)$). Subset images have size $27 \times 25 \times 5$. We take a time series of 200 images. This smaller dataset is part of a test suite, which can be used to test whether

¹Current version at time of thesis publication: `Rfmrieeg_1.0.tar.gz`

²These are tools for building R-packages on windows and are available from <http://cran.r-project.org/bin/windows/Rtools/>.

the software is properly installed. This test suite is provided by the author upon request and includes the following data files. The EEG data is contained in a 3D NIfTI-1 image *nifti_J-RS.nii*, the fMRI data is stored in a 4D NIfTI-1 image *nifti_fmrits-RS.nii* and the vector with stimulus times is contained in *stim-times.txt*. For more details on the NIfTI-1 file format see Section 5.3.1.

The following exemplary features for carrying out an analysis are presented: (1) Configuring a model run using the `S4 Config` class, (2) running the algorithm, (3) preparing results with `S4 class Results`, and (4) looking at MCMC diagnostics made available by `S4 class Diagnostics`. The `S4` class system was used for providing the functionality of object-oriented programming concept within R. Each class, thereby, defines class-specific member variables characterizing their objects and according member functions, which act as the intermediaries for retrieving or modifying the member variable values to ensure data integrity. Thus, setting the values of member variables is normally done via mutator methods to control for side-effects. In the following, we denote these functions as setter functions.

After the R session has been started, the user may (1) define the absolute paths to needed data objects, which are later assigned to the configuration settings, and (2) change the working directory to a selected folder (via `setwd()`). If not specified otherwise, results outputted by the procedure are saved to the working directory by default. Depending on whether he/she is working on a unix or windows operating system, the paths have to be set in the according correct format. For instance, use

```
R> pathFMRI.Test <- "Z:/testsuite/nifti_fmrits-RS.nii"
R> pathTimes.Test <- "Z:/testsuite/stim-times.txt"
R> pathEEG.Test <- "Z:/testsuite/nifti_J-RS.nii"
R> setwd("Z:/testsuite/")
```

when working on a windows system or

```
R> pathFMRI.Test <- "/home/kalus/testsuite/nifti_fmrits-RS.nii"
R> pathTimes.Test <- "/home/kalus/testsuite/stim-times.txt"
R> pathEEG.Test <- "/home/kalus/testsuite/nifti_J-RS.nii"
R> setwd("/home/kalus/testsuite/")
```

when working on a unix system. In the following, we report the output from an R session that was run on a windows system. Hence, paths are printed out alike.

Setter function	Relates to:
<code>setConfig.fmri()</code>	fMRI regression stage arguments
<code>setConfig.MCMC()</code>	MCMC algorithm specific arguments
<code>setConfig.pathResults()</code>	Setting the path where a results folder should be generated
<code>setConfig.stimCanonicalBF()</code> or <code>setConfig.stimGammaBF()</code>	Setting the HRF basis function type and related arguments

Table 5.1: Setter functions for setting arguments common to all model types.

CAR	iMRF	Ising
<code>setConfig.params.CAR1()</code>	<code>setConfig.params.iMRF1()</code>	<code>setConfig.paramsIsing()</code>
<code>setConfig.params.CAR2()</code>	<code>setConfig.params.iMRF2()</code>	
<code>setConfig.params.CAR3()</code>	<code>setConfig.params.iMRF3()</code>	
<code>setConfig.params.CAR4()</code>	<code>setConfig.params.iMRF4()</code>	
<code>setConfig.params.CAR5()</code>	<code>setConfig.params.iMRF5()</code>	

Table 5.2: Setter functions for setting the algorithm type of the model run and according arguments.

As noted in the preceding section, the `fmri` algorithm is run via a configuration file. If such a file is not at hand, the `S4` class `Config` can be used to generate one. For this, the class handles up to 59 arguments being employed in different model runs. The arguments of the `fmri` algorithm are hereby the member variables of the `Config` class. We structured the setting of member variables by providing setter functions that subdivide arguments with regard to content. The setter function `setConfig.fmri()`, for instance, takes and sets arguments needed to configure solely the fMRI regression stage of the model. See Table 5.1 for an overview of setters for model parameters common to all model runs and Table 5.2 for setter of algorithm specific arguments. If an algorithm type specific setter is called, arguments are set in such a way that the corresponding algorithm is selected. Additionally, these setters take arguments relevant for the specific model type. Which set of arguments a model setter function takes can be inferred from the according help file and Table C.3 in the Appendix. For a start, the majority of `fmri` arguments can be left to default values, so that configuring an algorithm run is straightforward. For this, initialize a new `Config`-object and specify arguments that do not have a default value (cf. Table C.1 in the Appendix). Note that warnings are printed if mandatory arguments are not yet set.

```
R> configObj <- new("Config")
R> configObj <- setConfig.pathResults(configObj)
[1] "WARNING: Use setConfig.fmri() to specify paths to fmri data"
```

(i.e. `pathNifti`, `pathStimTimes`)."

Note that this call to `setConfig.pathResults()` sets the results path to the working directory. If another directory should be used, the corresponding path has to be passed via the setter argument `pathResults`. If most arguments of the fMRI predictor stage should be kept to their defaults, only the paths to the actual data have to be specified via

```
R> configObj <- setConfig.fmri(configObj, pathNifti=pathFMRI.Test,
+   pathStimTimes=pathTimes.Test)
```

The setter functions listed in Table 5.2 can then be used to request the fitting of a specific algorithm type as well as to set corresponding model options. If no model setter function is called, `RfmriEEGMain` fits per default an Ising model for uninformed fMRI activation detection, which can alternatively be requested by using the function `setConfig.paramsIsing()`. If another model type is asked for, the corresponding setter has to be used. In the following, we would like to fit a gCAR5 model for EEG-enhanced fMRI activation detection and, hence, use

```
configObj <- setConfig.params.CAR5(configObj, pathEEG=pathEEG.Test,
+   speedup=FALSE,
+   tauStart=2,
+   sigma2ProposalTau=0.25,
+   tau0Start=25,
+   sigma2ProposalTau0=3)
```

to set configurations alike and pass over the path to the 3D EEG image file. To improve convergency, spatial dependency hyperparameters are given non-default values. Note that, `speedup=FALSE` selects a gCAR5 model instead of a fCAR5 model, which can be chosen by `speedup=TRUE`. Per default, a gCAR5 model is selected.

As a next step, we write a configuration text-file to the `pathResults` directory (in this example the working directory) by using the function `writeConfig()`. We append the string `"_gCAR5"` to the filename to indicate its use. This function returns the path to the written file, which can then be used to start the algorithm. We assign the returned string to variable `pathConfigFile`.

```
R> pathConfigFile <- writeConfig(configObj, add2Filename="_gCAR5")
[1] "Z:/testsuite/configFile_gCAR5.inp"
```

The algorithm is started via

```
R> returnObj <- RfmrieegMain(pathConfigFile)
```

and returns the path where all results and diagnostic objects are stored:

```
R> returnObj  
[1] "Z:\\testsuite\\results1"
```

Note that R uses either “\\” or “/” as a separator since “\” is a special character.

To access the results, the `Results`-class of the package can be used. For using its functionality, a new `Results`-object has to be generated and initialized by passing over the path to results (e.g. stored in `returnObj`) and the configuration object `configObj`.

```
R> resObj <- new("Results")  
R> resObj <- initResults(resObj, returnObj=returnObj,  
+                       configObj=configObj)
```

The results objects stored in the results folder can easily be accessed and, for example, an activation map can be written to the folder via `calcActivMap()` while returning the path where it can be found. For plotting, it first has to be read by the `Rniftilib`-function `nifti.image.read()` to be loaded again into the current R session. The function `plotNifti()` can then be used to draw the map. The resulting plot is shown in Figure 5.2. Note that `plotNifti()` provides an option to write out a PDF of the plot. Depending on the image dimensions and—optionally—on a calculated pleasant value for the plotting parameter `mfrow` resp. `mfcol`, width and height of the PDF are then automatically chosen in such a way that voxels appear as proper squares. If another device type is preferred, the user can calculate optimal graphic dimensions by using function `calcDeviceHeight()`.

```
R> activPath <- calcActivMap(resObj)  
R> activMap <- nifti.image.read(activPath)  
R> plotNifti(activMap, pdfoutput=TRUE, pdfName="gCAR5active")  
null device
```

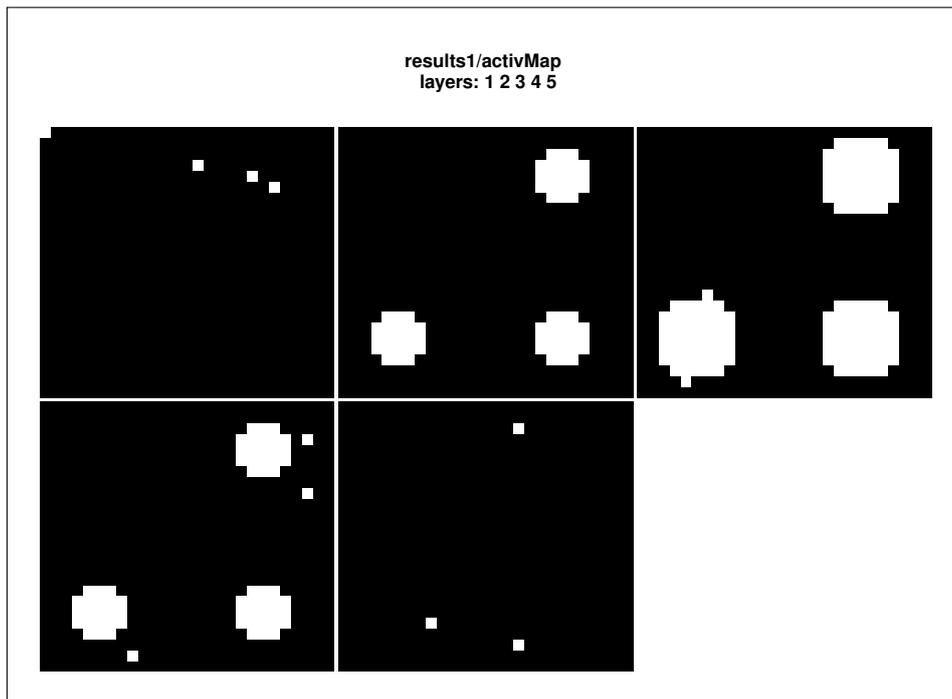


Figure 5.2: Estimated activation map from the exemplary gCAR5 model run.

Besides this, the `Results` class provides functions to plot the parameter map estimates via `plotMapEst()` and to compute a summary of global parameter estimates via `summary()`.

To determine whether the algorithm shows an acceptable convergency behavior the `Diagnostic`-class can be used. Again, a `Diagnostic`-object has to be generated and initialized for this:

```
R> diagObj <- new("Diagnostics")
R> diagObj <- initDiagnostics(diagObj, returnObj=returnObj,
                             configObj=configObj)
```

The function `summary()` prints a six-point summary of global parameter traces.

```
R> summary(diagObj)
This is a summary of global parameter estimates of the following
Rfmrieeg model run.
Used algorithm: gCAR5
Results folder: Z:\testsuite\results1
```

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
xi20	10.540	13.350	14.260	14.310	15.180	18.960
tau20	11.450	13.630	14.620	14.950	16.230	20.110
xi2	3.960	4.693	4.938	4.951	5.171	6.658
tau2	1.552	2.361	2.724	2.769	3.162	4.291

Trace plots of parameter trajectories can be obtained via the `plot()`-function.

```
R> plot(diagObj, what="global", type="trace")
```

If requested, this function plots autocorrelation functions and diagnostics of voxel-specific estimates as well. Exemplary trace plots of global parameter trajectories of the gCAR5 run can be found in Figure 5.3.

5.1.3 Compilation and usage of the executable `CfmrieegMain`

If desired, the C++ algorithm collection can be accessed via the `CfmrieegMain` program without relying on an installation of the R environment. For compiling and building `CfmrieegMain` in a shell follow the subsequent steps.

The source code of the C++ application is contained within folder `src` of the `Rfmrieeg` package including a makefile, which controls the compilation process. The most convenient way to build the `CfmrieegMain` program is, hence, to change directory to folder `src` and use the provided makefile via

```
% make -f makefileCpp
```

The binary `CfmrieegMain` is then generated in this directory from where it can then be executed.

Note that this instruction assumes a unix-type operating system with linear algebra libraries BLAS (Blackford et al., 2002) and LAPACK (Andersen et al., 1999) already installed in the standard library path of your system. These have to be C-libraries that were built upon a C interface (`cblas`) to the original Fortran routines. If available, `libblas.a` (`libblas.so`) and `liblapack.a` (`liblapack.so`) can, e.g., be found in `usr\libs\`. If `CfmrieegMain` should be linked to another BLAS or LAPACK library, the makefile has to be adapted. Further information on this can be found in Section 5.3.3.

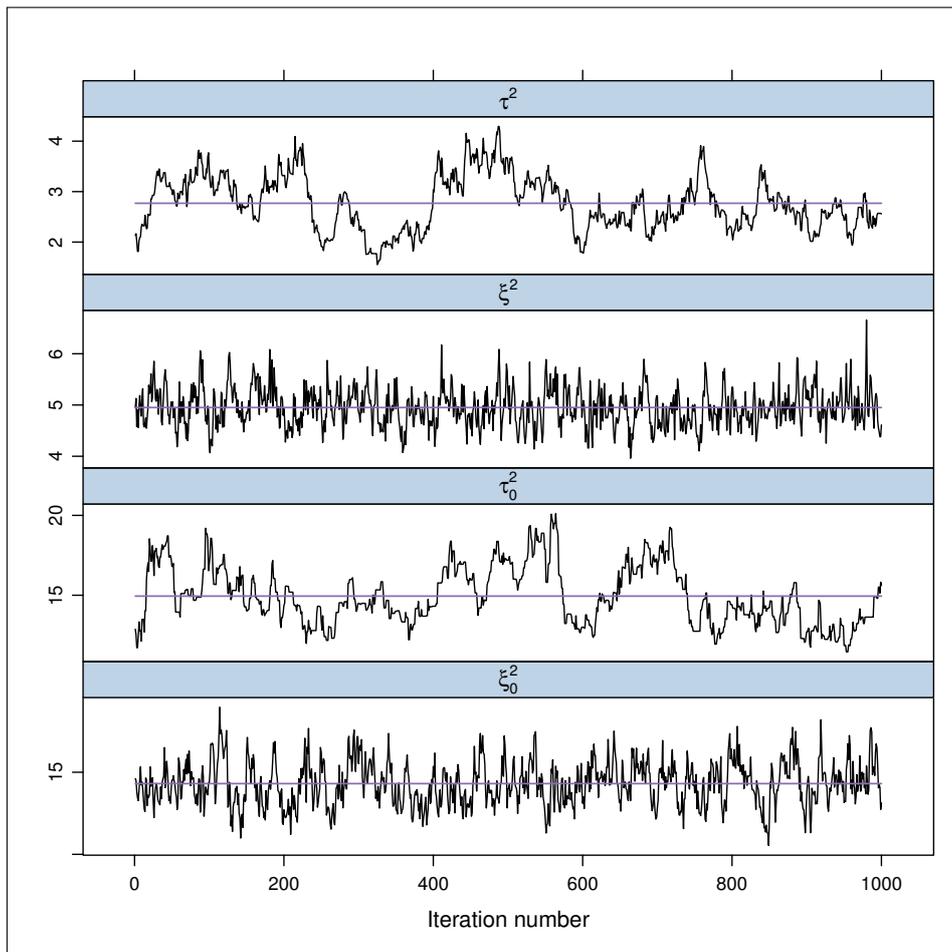


Figure 5.3: Trace plots of global parameters from the exemplary gCAR5 model run.

Standard BLAS and LAPACK routines are generally slow. For high performance, we recommend to install an optimized native BLAS implementation, like ATLAS (Whaley and Petitet, 2005) or GotoBLAS (Goto and Van De Geijn, 2008), which also include optimized LAPACK routines. The makefile can be adapted to incorporate one of these libraries. Further details on this can also be found in Section 5.3.3.

As noted in Section 5.1.1, the `CfmrieegMain` program is executed in a shell by using

```
% CfmrieegMain -config path/to/ConfigurationFile.inp
```

Results are stored in a subfolder of the directory specified with keyword `pathResults` or in the working directory if left unspecified. No extra functionality is provided for output analysis. For this, the user may revert to procedures provided by the R-package `Rfmrieeg`.

5.2 Implementational issues

The R-package `Rfmrieeeg` was developed on basis of a C++ implementation of the `fmrieeeg` algorithm. It provides a standardized installation process and easy to handle analyzing routines. Above that, the R-package includes extra-functionality for configuring the algorithm, presenting results and diagnostics. The heart of the `fmrieeeg` algorithm, however, is implemented in C++ code. The reason for not using R more extensively lies in the formerly insufficient and missing image data handling routines in R at the beginning of this project. At this time, the NIfTI-1 C library (<http://nifti.nimh.nih.gov/>) provided convenient data handling routines for image data in the NIfTI-1 format (and still does)—cf. Section 5.3.1 for more details on this. The `Rniftilib` package has made such routines available in R not before 2009. Additionally, to design a fast MCMC algorithm, it is recommended to implement update loops in a lower-level programming language like C/C++, which are closer to the machine language. Hence, we decided to use C++ for implementation. High diligence was put into implementing an algorithm with a good run-time performance.

To accomplish this, the implementation relies on external, open-source C-libraries that provide state-of-the-art numerical solutions for efficient linear algebra operations. These are described in Section 5.2.1 and 5.2.2.

When starting a new programming project, the coding task has to be structured to obtain comprehensible, easy-extensible code. For this, we decided to use an object-oriented approach that was constructed using the unified modeling language (UML) (see e.g. Kecher, 2009). To provide insight into the underlying software architecture, the corresponding UML diagram of the implementation project is discussed and presented in Section 5.2.3.

5.2.1 Efficient linear algebra subprograms

The algorithm comprises several linear algebra operations like scalar products, vector- and matrix multiplications and a multivariate least squares fit. Where feasible, computations have been vectorized to omit time-consuming loops. To yield a short computation time overall, vector and matrix operations are calculated by external libraries, which have been especially designed to achieve good run-time performances. For basic linear algebra calculations, we use C-libraries building on BLAS (Blackford et al., 2002). For a multivariate least squares fit, an interface to the GELS function of the LAPACK library is used—which relies on a QR matrix factorization (Andersen et al., 1999).

As noted above, standard BLAS and LAPACK routines are generally not the fastest available. Further details on including an optimized library can be found in Section 5.3.3.

For a convenient and efficient C++ interface to (optimized) BLAS and LAPACK routines, the matrix library FLENS (Lehn et al., 2005) is used. FLENS possesses the advantage of no run-time overhead compared to directly calling BLAS and LAPACK and no side-effects like internal creation of temporary objects—in spite of providing easy to handle dynamic vector and matrix objects and related array subsetting functionality.

5.2.2 Efficient GMRF updates via Cholesky decompositions

Updating large random fields in one block is supposed to be beneficial for the convergency behavior of the MCMC procedure. This applies to the update of parameter vectors α and α_0 , i.e. the spatially-varying coefficients of the latent probit stage (cf. Chapter 3). An efficient sampling scheme from corresponding multivariate normal update distributions utilizes the sparse matrix nature of the variance-covariance matrices at hand as described in Rue (2001) and Lang and Brezger (2004). In the following, we briefly summarize the main steps for an update of α . An update for α_0 can be implemented in an analogue way.

Let \mathbf{K} be the precision matrix for the GMRF update of α , i.e. $\mathbf{K} = (\mathbf{J}'\mathbf{J} + \Sigma^-)$. Drawing random numbers from the multivariate normal full conditional distribution $p(\alpha|\cdot)$ with mean $\mathbf{m} = \mathbf{K}^{-1}\mathbf{J}'(\mathbf{U} - \boldsymbol{\eta}_{-\alpha})$ and precision matrix \mathbf{K} is achieved as follows. After calculating its Cholesky decomposition $\mathbf{K} = \mathbf{L}\mathbf{L}'$, we proceed by solving $\mathbf{L}'\alpha = \mathbf{z}$, where \mathbf{z} is a vector of independent standard Gaussians. It follows that $\alpha \sim \mathcal{N}(0, \mathbf{K}^{-1})$. The update mean is calculated by solving $\mathbf{K}\mathbf{m} = \mathbf{J}'(\mathbf{U} - \boldsymbol{\eta}_{-\alpha})$. This is achieved by first solving $\mathbf{L}\boldsymbol{\nu} = \mathbf{J}'(\mathbf{U} - \boldsymbol{\eta}_{-\alpha})$ by forward substitution followed by a backward substitution $\mathbf{L}'\mathbf{m} = \boldsymbol{\nu}$. Adding \mathbf{m} to α yields $\alpha \sim \mathcal{N}(\mathbf{m}, \mathbf{K}^{-1})$.

Note that this procedure is not applied to the original matrices, but to permuted ones. Fill-reducing permutations yield an astonishing increase in speed. Rue (2001) and Lang and Brezger (2004) suggest to use the Cuthill-McKee algorithm (George and Liu, 1981) to reduce the band-width of sparse matrices. We, however, slightly modified their approach by allowing for different reordering algorithms for the underlying spatial grid. Cholesky decompositions are efficiently calculated using the routines of the C-library CHOLMOD (Davis and Hager, 1999). In an analyze step, CHOLMOD automatically determines the best strategy for the dataset at hand to reorder the voxel-IDs of the regular 3D grid to gain a smaller bandwidth of sparse precision matrices $\mathbf{J}'\mathbf{J} + \Sigma^-$ or $\mathbf{I} + \Sigma_0^-$. CHOLMOD

chooses between the natural ordering, AMD (Amestoy et al., 1996, 2004), COLAMD (Davis et al., 2004; Davis, 2005), different versions of a nested dissection algorithm (George, 1973) and—if available—the METIS algorithm (Karypis and Kumar, 1998). For an in-depth discussion of used reordering strategies see the CHOLMOD user manual. A test on an exemplary brain imaging dataset has shown that CHOLMOD finds permutations leading to smaller factor bandwidths than the Cuthill-McKee algorithm. The analyze step for determining the permutation is done once prior to MCMC iterations and thus, all Cholesky decompositions. The reordering, i.e. index permutation, is repeated in every iteration, but with negligible time demand. The most time-consuming step in our procedure is the calculation of Cholesky factors of permuted matrices. Generally, updating large random fields in one block constitutes the most time-consuming step in the `fmrieeg` MCMC procedure.

5.2.3 Software Design

Several conceptual requirements were imposed for implementing the C++ `fmrieeg` algorithm. These include a data-independent, easily extensible implementation with a comprehensible structure. To meet this needs, an object-oriented design was chosen, which is visualized by the UML diagram in Figure 5.4. A class diagram usually includes the classes with their attributes and methods. For sake of a condensed presentation, these features are omitted in the below figure as well as the relationships to external library classes—except `MTRand`, which is the Mersenne Twister class for generating random numbers (Matsumoto and Nishimura, 1998). Three sorts of class relationships are visualized: (1) Every arrow joins the derived class with its base class (e.g. `stimCanonicalBF` is derived from `stimBF`), (2) arrows with open diamonds indicate weak aggregation (e.g. an instance of `XStimulus` contains an `ScanInfo`, `StimulusInfo` and `StimBF` object), (3) dashed arrows indicate a directed dependency between classes (e.g. an `ParamsMIsing` object knows about the `DiagnosticsMIsing` object but not vice versa). Abstract classes are printed in italics. For a more detailed introduction on UML, the interested reader is referred to Kecher (2009).

The software architecture reflects the main two hierarchies of the `fmrieeg` model. On the one hand, there is the fMRI predictor stage. On the other hand, there is the Bayesian activation detection stage with the latent probit resp. Ising model implementation.

The integral part of the fMRI regression stage is the calculation of the fMRI design matrix, which contains a baseline, (optional) confounders and stimulus regressors (cf. Chapter 2).

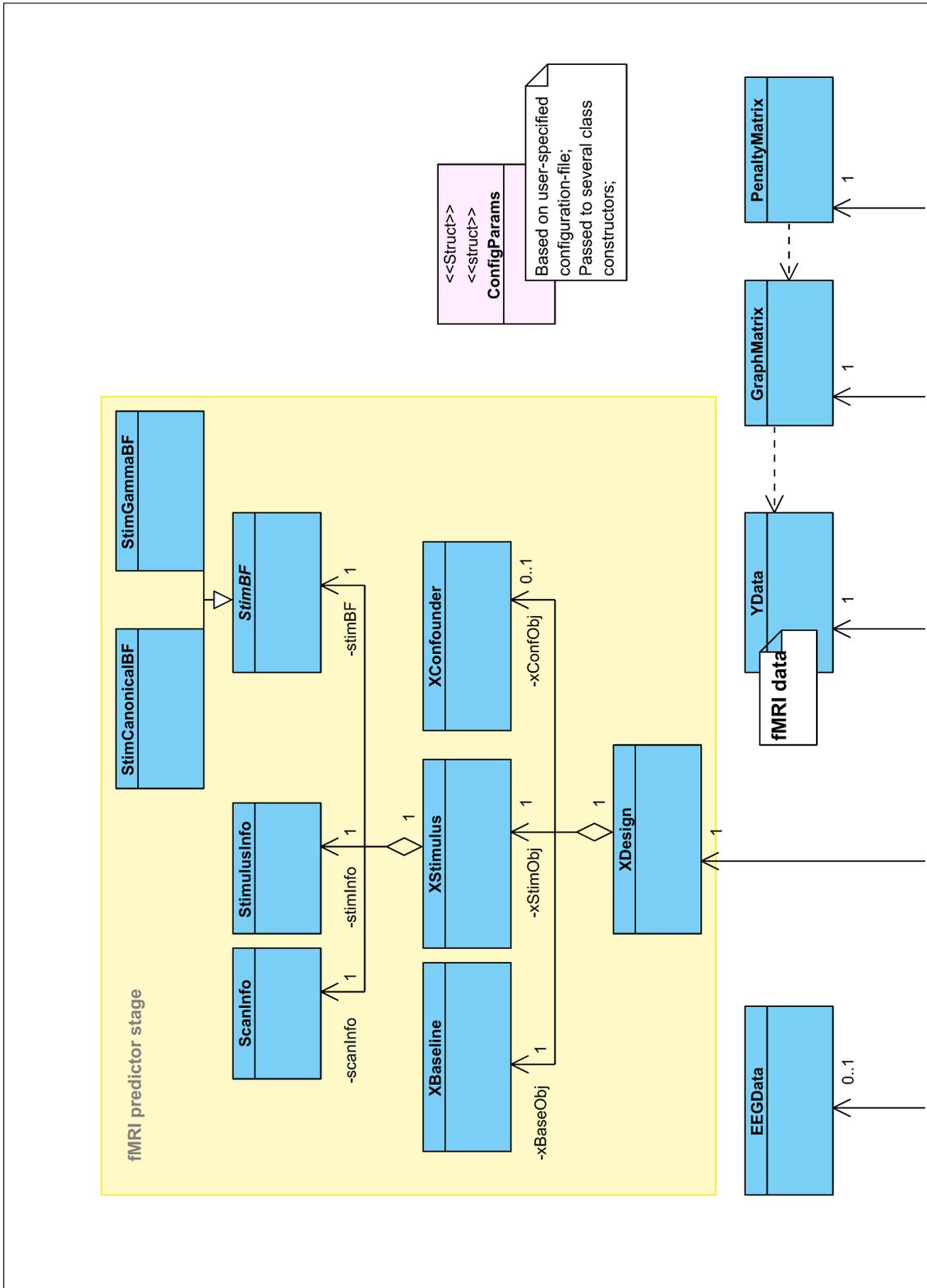
This is reflected by the three classes `XBaseline`, `XConfounder` and `XStimulus`. An instance of each of these three classes is (can be) contained in an `XDesign` object, which combines them to one design matrix. Whereas it is optional to include the `XConfounder` object (which is instantiated by reading a confounder matrix from file), the `XBaseline` object contains a matrix with at least one basis function modeling the intercept.

The `Xstimulus` class provides means for calculating the stimulus regressors of an event-related experimental paradigm. For this, the user has to specify information about the scanning procedure (`ScanInfo`), about the experimental setup (`StimulusInfo`) and the requested basis functions (`StimBF`). To incorporate different basis function types, the class `StimBF` is realized as an abstract class that cannot be instantiated but provides the prototype for a virtual `bfValue()` function calculating basis function values for a given peristimulus time. This function is then implemented by concrete `StimBF` descendants (classes `StimCanonicalBF` and `StimGammaBF`). Extensions to other basis function types can thus be easily incorporated into the implementation by adding new descendant classes. Classes for storing, reading and manipulating data objects are provided by classes `EEGData` and `YData`. The latter provides access to the fMRI data and includes masking routines restricting the program's internal access to a requested subgroup of voxels. Upon the definition of the `YData` class, neighborhood information of the voxel graph is generated and stored in an object of class `GraphMatrix`, which can be used to calculate a sparse Laplacian matrix (class `PenaltyMatrix`).

Having defined all classes with information needed for executing the `fmri_eeeg` update schemes, the model classes can be defined. Again, for easy extensibility we define an abstract model class `Params` which provides an interface to a general MCMC algorithm (implemented in class `MCMC`). Descendants of `Params` implement an `updateParams()` function embedding model-specific MCMC update functions. Optionally, an object from class `Diagnostic` with concrete, model-specific realizations can be generated to store diagnostics for requested model runs.

5.3 Technical notes

In this section, we provide additional information on details that are of interest for users not being familiar with the used data format and for users seeing the need to improve the run-time of the algorithm. For this, we give a short introduction on the NIfTI-1 image data format in Section 5.3.1. Before advice is given on controlling speed-relevant factors in Section 5.3.3, we point out the dependencies on external C++ libraries in Section 5.3.2.



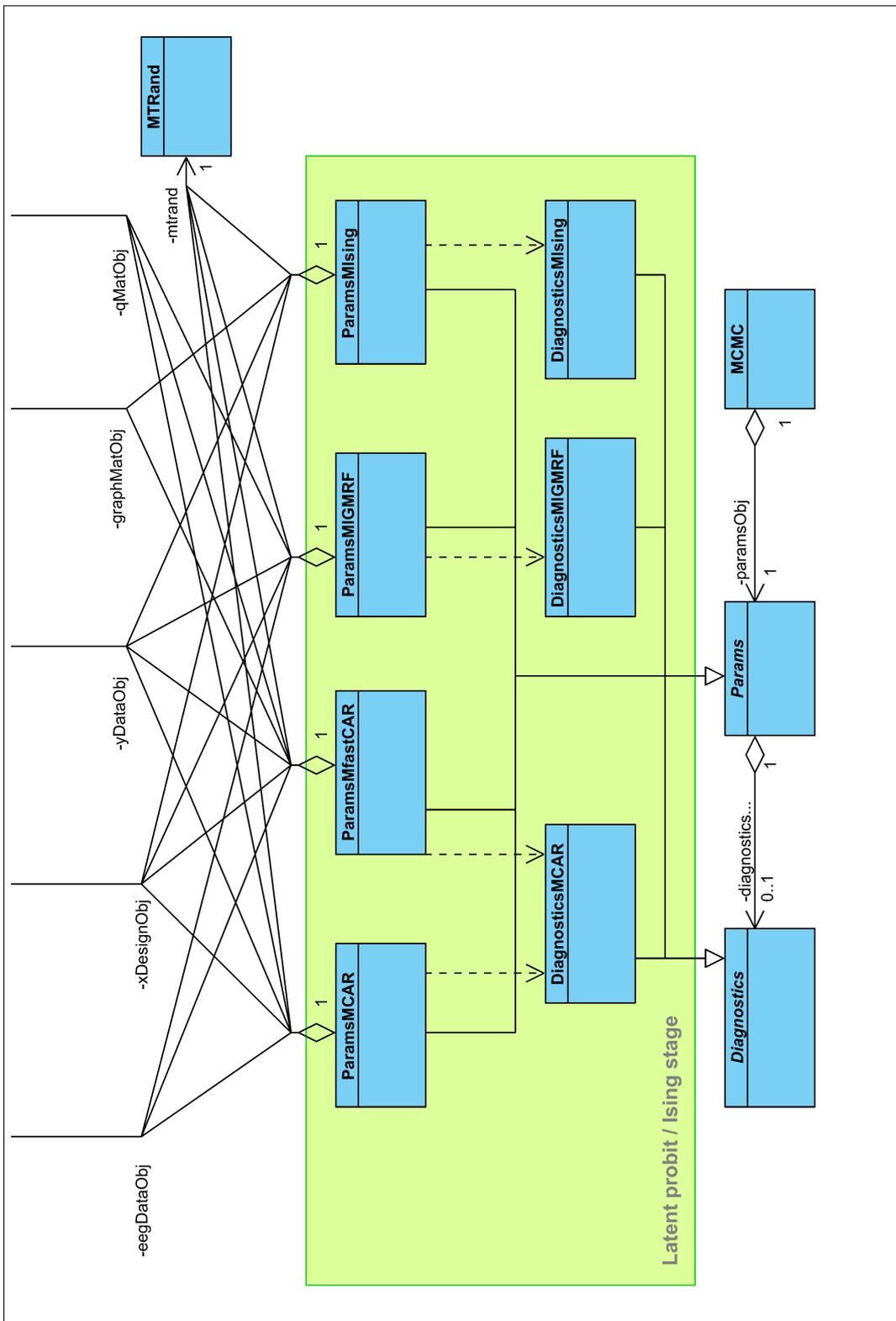


Figure 5.4: UML diagram of the `fmri.eeg` library

5.3.1 Image data format

The NIfTI-1 format (<http://nifti.nimh.nih.gov/>) is widely used in the functional neuroimaging field. Popular software packages (AFNI, BrainVoyager, FSL, SPM) agreed to read and write NIfTI-1 files and provide utilities for this since the end of 2004. Before fMRI data is analyzed via the `fmrii` algorithm, the data is generally preprocessed (see Chapter 7.1 for an example) using one of these software packages and will thus be available in the NIfTI-1 format. Hence, we decided to incorporate data management routines from the NIfTI-1 C library to read and write image data with the `fmrii` package. For users unfamiliar with this data format, the following notes may be helpful.

The NIfTI-1 file format extends the ANALYZE™-7.5 file format to incorporate features a task force of functional imaging software experts agreed on. This task force (Data Format Working Group, DFWG) was commissioned to find a technical solution to the problem of multiple data formats used in fMRI research and created the NIfTI-1 definition as unifying file format. The NIfTI-1 C library is made available by the DFWG on the above given website to provide utilities for working with the NIfTI-1 format.

A NIfTI-1 file is a binary data format consisting of a 348 byte header followed by image data stored in an one-dimensional array. The header comprises fields (variables) describing the structure of the subsequent data (cf. <http://nifti.nimh.nih.gov/nifti-1/documentation/nifti1fields>) including e.g. the dimension of the data and the datatype. A NIfTI-1 file can be available as a single file (file ending `.nii`) or as an `.hdr/.img` file pair where the data of the 348 byte header and the image data are stored separately in a `.hdr` and `.img` file, respectively. For the `fmrii` software it is assumed that the 4D fMRI data is available as a single `.nii` file to simplify reading. Tools for aggregating a time series of single images into one `.nii` file are for example available by an installation of FSL (Smith et al., 2004). We recommend to use the NIfTI-1 file format as well for the EEG data map and the (optional) map containing an analyzing mask—instead of a text-file with row-wise vector entries, which also can be read. To the text-file no extra information, e.g. about the storage order of voxels in the data array, can be attached. If this information about image files is missing, it cannot be ascertained that all image maps intended to be used for analysis are compatible with each other—which constitutes the basis for all data fusion software. Besides the storage order, for instance, image dimension and voxel size must match.

A convenient interface to the NIfTI-1 C library in R is provided by the package `Rniftilib`. After reading a NIfTI-1 image, the data can easily be accessed via the R-specific array

subsetting.

5.3.2 External dependencies

The liberal license of several external libraries used within the `fmrieeeg` code allows us to include them into our software package. For details see the license files distributed with the `Rfmrieeeg` package. We decided to include SuiteSparse sublibraries (CHOLMOD, COLAMD and AMD), FLENS and the NIfTI-1 C library into the package as well as a Mersenne Twister random number generator (Matsumoto and Nishimura, 1998). Linear algebra libraries BLAS and LAPACK are not distributed with our package for their general availability, large size and customizability to system-specific conditions.

Both the R-package `Rfmrieeeg` and the `CfmrieeegMain` program (contained in folder `src` of the `Rfmrieeeg` package) automatically compile these libraries and include needed code sections when being build.

If a user prefers individual installations of external libraries, these can be included into the `CfmrieeegMain` program. For this, we provide a dedicated makefile, where ‘include’ directories and library paths of favored libraries can be specified. This makefile can be found in folder `src\fmrieeegDP` when `CfmrieeegMain` should work with double-precision routines (the default). Alternatively the makefile in `src\fmrieeeg` can be adapted and used to build `CfmrieeegMain` with single-precision routines. The compatibility of other versions than otherwise automatically included libraries, however, cannot be ensured.

The inclusion of external libraries is sensible in some cases. These cases are discussed in the following section—as well as difficulties in incorporating external libraries in an R-package installation.

5.3.3 Influences on the running time

There were several decisions to be made during the process of software development that have influenced the runtime performance of the `fmrieeeg` C++ algorithm. As noted above, high diligence was put into using efficient linear algebra routines building on BLAS and LAPACK. Additionally, when installing the software the user can customize the compilation and installing process to some extend to increase the runtime performance of the algorithm even more. Customizing, however, can only be done within the scope of

the given software application framework. There are less options when configuring the R-package installation. Hence, in the following, we differentiate between the R-package and binary program building.

Modifying the building process of `CfmrieegMain`

The building process of the `CfmrieegMain` program can be modified at two stages: Changes can be applied to the building configuration of (A) the self-contained package of automatically built sublibraries and of (B) a separate `fmrieeg` library compilation with links to already installed external libraries. For (A) configurations can be changed by modifying the makefile 'makefileCpp' contained in directory `src` of the R-package. For (B) the makefile 'makefileCpp' contained in the `fmrieegDP` resp. `fmrieeg` subfolder can be adapted (depending on whether the single- or double precision library should be compiled).

In the following, we discuss three options to increase the runtime performance of `CfmrieegMain`: (1) the use of alternative BLAS/LAPACK library installations; (2) the choice between a single- or double-precision version of `fmrieeg`; (3) usage of own installations of external libraries. Modifications in form of (1) and (2) can be accomplished within the compilation process of the self-contained package (case (A)). Thus, we primarily focus on realizing changes that can be made to (A). In contrast to this, full flexibility, e.g. in the choice of external library installations, can only be handled by case (B). This, however, requires that the user knows how to compile external libraries on her/his own. Because of this, we only comment very briefly on incorporating a specific external library implementing the METIS algorithm (see below).

An optimized BLAS/LAPACK library increases the speed of the `fmrieeg` algorithm in any case. For high performance, we recommend to install an optimized native BLAS implementation like ATLAS (Whaley and Petitet, 2005) or GotoBLAS (Goto and Van De Geijn, 2008), which also include optimized LAPACK routines. Note that GotoBLAS must be compiled in the single-threaded version to not interfere with CHOLMOD routines. To link against an optimized library, adjust the `MY_BLASLAPACK_LIBS` statement in the makefile 'makefileCpp' contained in folder `src` of the R-package. The alternative BLAS/LAPACK versions can then simply be incorporated in the building of `CfmrieegMain`. Acceleration, generally, depends on the size of the image data. For a whole-brain analysis, we expect a substantial decrease in running time. For our installation, we linked the `CfmrieegMain`

program to the GotoBLAS library and achieved an acceleration of factor 4 in a brain image analysis of an fMRI time series with 27 257 voxels.

The `fmriieg` algorithm originally has been implemented using single-precision variables (except for code sections where a higher accuracy is needed for specific calculations). Reasons for this were threefold: We decided to use float variables on the basis that image data at hand did not need high precision data formats for being available as short integers. Hence, on the one hand, float variables offered the advantage of saving working memory and, on the other hand, running time potentially could be decreased. Acceleration, however, could not be observed on a Quad-Core AMD Opteron(tm) Processor 8384 (64 bit system). Nevertheless, the user might revert to the single-precision library when working memory is limited or another platform promises an increase in speed when processing floats instead of doubles. The single precision `CfmriiegMain` program can be compiled by setting configurations in the makefile 'makefileCpp' provided in directory `src` of the R-package alike (see description contained).

A further factor that might influence running time is the inclusion of the METIS C-library (Karypis and Kumar, 1998). The METIS algorithm is capable of calculating an advantageous permutation to reduce the bandwidth of sparse Laplacian matrices used in the time-consuming update of large GMRF maps as described in Section 5.2.2. Depending on the voxel grid at hand, the METIS algorithm can increase the speed of the `fmriieg` algorithm. Because of its license, the METIS C-library is neither distributed with the SuiteSparse library, which includes CHOLMOD, nor with our self-contained package, which is used to build `CfmriiegMain`. Therefore, an individual installation of the SuiteSparse library may be desired to incorporate the METIS C-library routines. To revert to the METIS algorithm, the library must be installed on your platform (METIS is available at <http://glaros.dtc.umn.edu/gkhome/metis/metis/overview>) and SuiteSparse has to be recompiled to find the METIS library (for more information on this, see <http://www.cise.ufl.edu/research/sparse/SuiteSparse/>). To link against external libraries adapt the makefile 'makefileCpp' in the `fmriieg` or `fmriiegDP` subfolder of the R-package folder `src` (depending on whether the single- or double precision library should be compiled). Adjust the `METIS_LPATH` (`METIS_LIB`) and `SUITESPARSE_DIR` statements within the makefile to link against METIS and an own SuiteSparse installation.

Changing installation settings within R

It must be ensured that installation of R-packages work on both windows and unix systems. Although compilation of package-specific source code can be configured using a 'MAKEVARS' ('MAKEVARS.win') makefile included in folder `src`, the standard installation process of an R-package must be accomplishable with the software components provided by the R-package and a standard R installation. Especially, inclusion of external libraries is problematic, because their presence cannot generally be assumed. An How-To on installing general C/C++ libraries on a windows system and including them in an R-package installation process is beyond the scope of this introduction, hence, we give some advice on including external libraries on a unix system.

As noted in the preceding section there may be reasons to use BLAS and LAPACK libraries not included by default. Foremost, the inclusion of an optimized BLAS is desirable, but the user might also want to include a library providing single-precision routines. An installation of the software environment R makes a BLAS and LAPACK library available. These libraries provide double-precision routines only, so that the single-precision `fmrieeg` library cannot be used by default. If another BLAS/LAPACK library should be used, we recommend to proceed as follows.

We assume that an alternative BLAS/LAPACK library (like ATLAS or GotoBLAS) is installed. If one of these should be used within the R-package, it must be ensured that the whole R environment uses this library as well to avoid conflicts. There are different ways of incorporating a shared BLAS library. Probably the most easy one is to symlink a dynamic BLAS library to '`R_HOME/lib/libRblas.so`' as described in the R Installation and Administration-guide (available at <http://cran.r-project.org/doc/manuals/R-admin.pdf>). Accordingly, the library containing LAPACK routines must be symlinked to '`R_HOME/lib/libRlapack.so`'. If an R-package with BLAS/LAPACK routines is now installed within this modified R version, it now uses routines of the specified alternative library/libraries.

If single-precision routines should be used or the external BLAS/LAPACK library is called by a `cblas` interface, corresponding 'MAKEVARS' statements of the `Rfmrieeg` package have to be adjusted as described there.

External libraries like METIS and an accordingly recompiled SuiteSparse can be included by changing the 'MAKEVARS' ('MAKEVARS.win') file to include corresponding 'include'-paths and linking statements. Some changes to the present form of the 'MAKEVARS' have to be applied that are not provided by us, but can be accomplished by the user.

6 Simulation studies

Reasons for conducting the studies in this chapter are threefold: First, we want to base model building decisions upon the results of data arising under controlled conditions. Second, uninformed fMRI activation detection schemes should be compared to algorithms found in the literature to assess their basic performance. Third, promising EEG-enhanced detection schemes should be compared to uninformed schemes to assess whether our models are generally capable of increasing sensitivity by the use of spatial EEG information. Model building decisions are based on convergency properties and the capability of yielding sensible estimates for latent fMRI and/or EEG predictor components. Performance is assessed by the power of the different algorithms to detect activated voxels. Simulation studies are designed in such a way that a true activation map is at hand. Hence, we quantify performance in terms of sensitivity and specificity.

This chapter is divided into two parts: The first section, Section 6.1, contains simulation studies based on an artificial simulation design. This section includes an evaluation of different predictor forms leading to model selection decisions. In Section 6.2, a simulation study is based on the structure of a given real-world dataset.

6.1 Simulation studies based on artificial data

In this section, simulation setup is based on an artificial simulation design. We control position, spread and composition of regions of activation or rather non-activation with each region being very homogenous. In Section 6.1.1, we describe the simulation design that is used to investigate the performance of models in a controlled setting. Notation and parameter settings common to all models used in this study are introduced in Section 6.1.2. In Section 6.1.3, we evaluate different predictor forms of the EEG hierarchy, i.e. the way EEG information is included in the model. Besides this, the type of spatial prior (i.e. CAR versus IGMRF prior) and the accelerated version of the CAR algorithm are examined.

We conclude with a recommendation on which algorithm should be used in further analyses. In Section 6.1.4, we compare the performance of our proposed models based solely on fMRI (so-called null models) to competing activation detection methods. Finally, in Section 6.1.5, we evaluate whether the recommended combined fMRI-EEG algorithms have increased sensitivity compared to their corresponding null models. Performance of different algorithms is then evaluated at different noise levels.

6.1.1 Design for generating an artificial dataset

A complete dataset for our combined fMRI-EEG model consists of a 4D fMRI dataset and a 3D EEG dataset, which both need to be generated. We will proceed by first setting up EEG predictor components including the EEG map. Based on the according predictor map, a binary activation map is derived which is—in turn—the basis for the generation of the fMRI dataset.

We construct a dataset with five brain layers of dimension 47×56 . A binary mask from a real-world brain dataset is used to fill only voxels within the shape of a human head.

First, we calculate the predictor maps of the latent probit stage of the combined fMRI-EEG model:

$$\begin{aligned} \eta_i &= f_0(i) + f(J_i) = \alpha_{0,i} + \alpha_i J_i, \quad i = 1, \dots, N \\ \Leftrightarrow \boldsymbol{\eta} &= \boldsymbol{\alpha}_0 + \mathbf{J}\boldsymbol{\alpha} \end{aligned}$$

where $\boldsymbol{\alpha}_0$ and $\boldsymbol{\alpha}$ are the vectorized parameter maps of spatially-varying coefficients and \mathbf{J} is the diagonal matrix of voxelwise EEG measurements as in Chapter 3.

In our design, we differentiate between four different kinds of activated regions and center them in layer 3. The regions differ in the combination of $\boldsymbol{\alpha}_0$ -, $\boldsymbol{\alpha}$ - and \mathbf{J} -values showing activation. We assumed that activation in $\boldsymbol{\alpha}$ is only present if activation in \mathbf{J} is present. See Table 6.1 for a listing of the considered activation constellations and Figure 6.1 for a visualization of these. An activated region is represented as the peak density values of a three-dimensional multivariate normal density function with centers given by x , y and z coordinates in Table 6.1. The following variance-covariance matrix is used for all multivariate normal distributions

$$\boldsymbol{\Sigma} = \begin{pmatrix} 12 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Region	α_0	\mathbf{J}	α	x	y	z
R_1	0	1	0	15	35	3
R_2	1	0	0	30	35	3
R_3	0	1	1	15	20	3
R_4	1	1	1	30	20	3

Table 6.1: Simulation design: The binary indicators in columns α_0 , \mathbf{J} and α indicate whether activation, i.e. an area of positive values, is present in region 1 to 4. The regions are centered in layer 3 with row and column coordinates given in x and y .

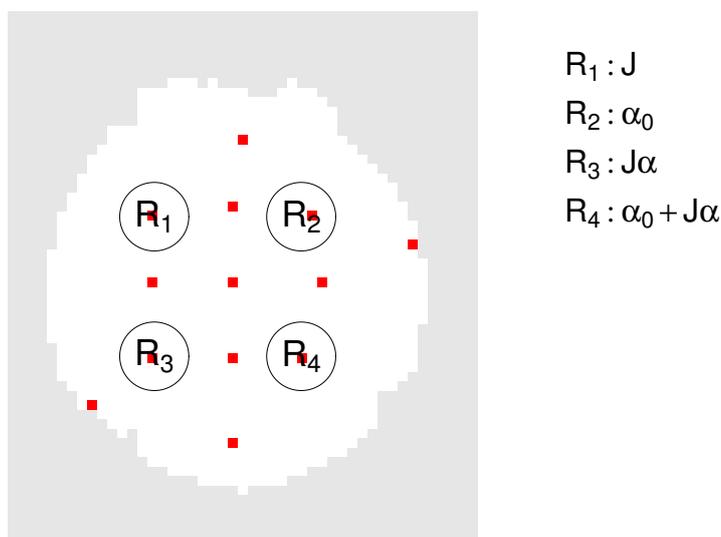


Figure 6.1: Overview of experimentally varied activation regions and selected voxels for diagnostics (labeled as red squares).

For calculating the α_0 -map, two separate multivariate normal density maps corresponding to R_2 and R_4 are calculated. The α_0 -map is then composed as the voxelwise maximum of these two maps. The resulting map is normalized by its maximal value and multiplied by $\Phi(0.975)$ to ensure that voxels in this region are set as active in the true activation map below. To test edge-preserving properties of algorithms, the multivariate normal density values are truncated at $\Phi^{-1}(0.8722) = 1.136852$, which is later used as a threshold for activation classification.

The α -map is calculated in a similar way by setting up separate multivariate normal density maps for region R_3 and R_4 . The whole α -map is then composed as the voxelwise maximum of the two maps. After this the α -map is normalized to a maximal absolute value of $\Phi(0.975)$ and truncated like the α_0 -map.

The EEG map \mathbf{J} is again realized by the voxelwise maxima of multivariate normal densities of EEG activated areas (R_1, R_3 and R_4). The \mathbf{J} -map is also normalized to a maximal value of $\Phi(0.975)$, but not truncated for the EEG source construction solutions being very smooth.

Having set all predictor components of the EEG stage, we calculate $\boldsymbol{\eta} = \boldsymbol{\alpha}_0 + \mathbf{J}\boldsymbol{\alpha}$ and threshold the resulting map with $\Phi^{-1}(0.8722) = 1.136852$ to derive a binary activation map, which is denoted as the true activation map $\boldsymbol{\gamma}^*$ in the following simulation studies.

The $\boldsymbol{\alpha}_0$ -, $\boldsymbol{\alpha}$ -, \mathbf{J} -maps, the corresponding predictor map of the latent probit stage and the derived true activation map $\boldsymbol{\gamma}^*$ are visualized in Figure 6.2.

The fMRI dataset is calculated as follows. Recall, the fMRI voxel time series $y_{i,t}, t = 1, \dots, T$ of voxel i is decomposed into

$$y_{i,t} = f_{base}(i, t) + f_{con}(i, t) + f_{stim}(i, t) + \epsilon_{i,t}, \quad t = 1, \dots, T,$$

which can be linearized and written as multiple regression model (cf. Chapter 2)

$$\mathbf{y}_i = \mathbf{X}\boldsymbol{\beta}_i + \boldsymbol{\epsilon}_i.$$

We select $y_{i_{max},t}, t = 1, \dots, T$, i.e. the voxel time series of voxel i_{max} of a given real-world dataset with maximal likelihood ratio test statistic of the mass-voxelwise tests $H_{0,i} : f_{stim}(i, t) = 0 \forall t, i = 1, \dots, N$. This data trajectory stems from an additional female subject who underwent a run of the same experimental procedure as described in Chapter 7, i.e. an oddball experiment with 60 odd and 540 even stimuli. The total duration of this session, however, was slightly longer: The resulting fMRI dataset consists of $T = 354$ scans.

The new dataset is generated by

$$\mathbf{y}_i = \mathbf{X}^*\boldsymbol{\beta}_i^* + \boldsymbol{\epsilon}_i^*, \quad i = 1, \dots, N$$

where $\boldsymbol{\beta}_i^* = E(\boldsymbol{\beta}_i | \boldsymbol{\gamma} = \boldsymbol{\gamma}^*, \mathbf{y}) = (\mathbf{X}^{*\prime} \mathbf{X}^*)^{-1} \mathbf{X}^{*\prime} \mathbf{y}_{i_{max}}$ is constructed as the posterior mean of a model with activation map $\boldsymbol{\gamma}^*$. The matrix $\mathbf{X}^* = \mathbf{X}(\boldsymbol{\gamma}_i^*)$ is thereby the full design matrix if $\boldsymbol{\gamma}_i^* = 1$ and it is the design matrix without the stimulus regressor component if $\boldsymbol{\gamma}_i^* = 0$. The error term $\boldsymbol{\epsilon}_i^*$ is a realization from $\mathcal{N}(0, \sigma_i^{2,*} \mathbf{I})$ with $\sigma_i^{2,*} = E(\sigma_i^2 | \boldsymbol{\gamma} = \boldsymbol{\gamma}^*, \mathbf{y}) = S_i(\boldsymbol{\gamma}_i^*) / (T - 2)$ where $S_i(\boldsymbol{\gamma}_i^*) = \mathbf{y}_{i_{max}}' \mathbf{y}_{i_{max}} - \mathbf{y}_{i_{max}}' \mathbf{X}^* (\mathbf{X}^{*\prime} \mathbf{X}^*)^{-1} \mathbf{X}^{*\prime} \mathbf{y}_{i_{max}}$. To investigate effects of increased noise, we consider noise levels $\sigma_i^{2,*}, 2\sigma_i^{2,*}, 4\sigma_i^{2,*}, 6\sigma_i^{2,*}, 8\sigma_i^{2,*}, 10\sigma_i^{2,*}, 12\sigma_i^{2,*}$.

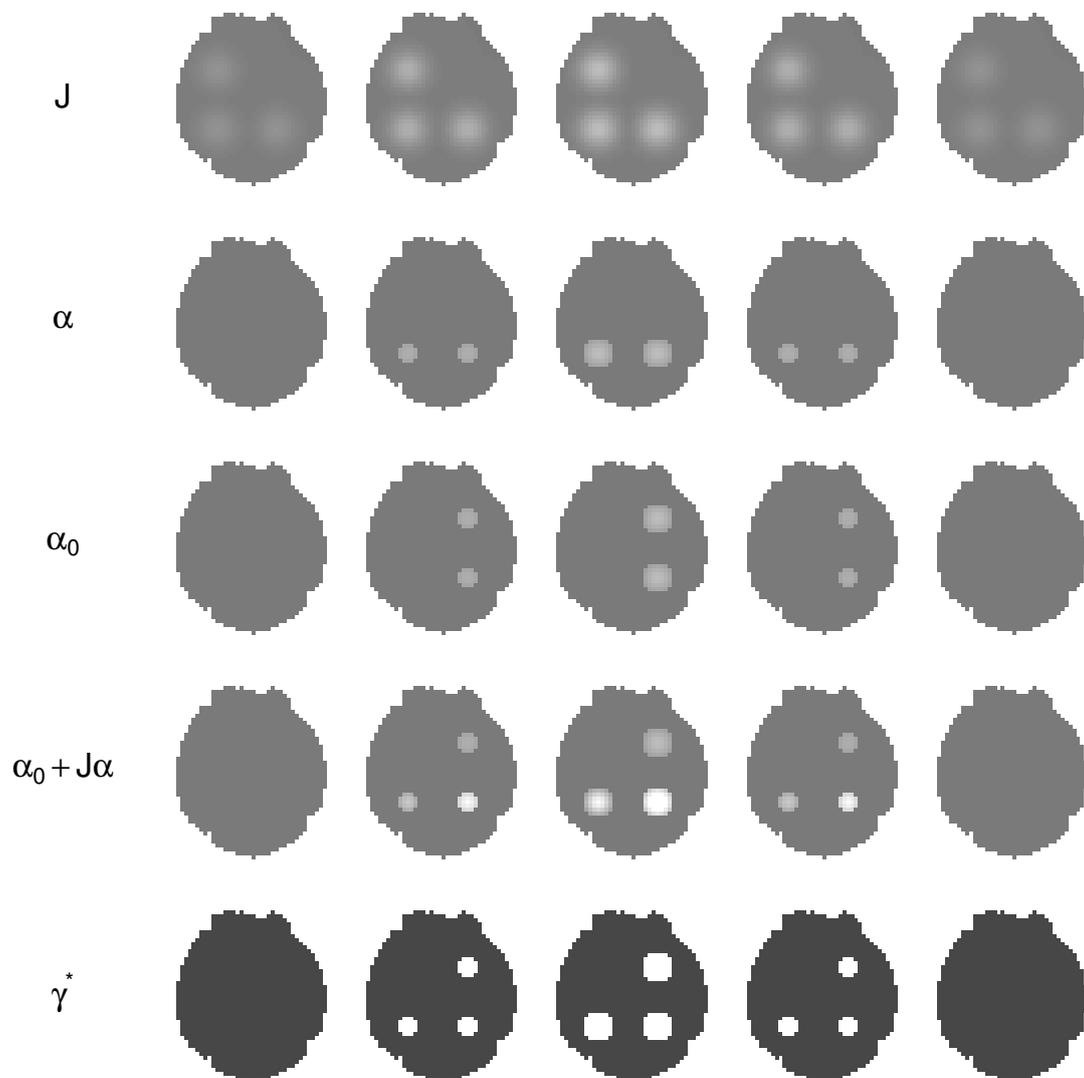


Figure 6.2: Generated maps of the artificial data simulation set up (EEG hierarchy): EEG- (1st row), effect- (2nd, 3rd row) and predictor maps (4th row). In row five, the resulting activation map is plotted. Figure columns 1-5 represent brain layers 1-5.

6.1.2 Notation and common model settings

For our simulation studies, we consider both CAR prior models and IGMRF prior models. For the model with CAR prior, we use the form with global updates of spatially-varying effects from Chapter 3 (denoted as gCAR) as well as the fast algorithm from Chapter 4 (denoted as fCAR). The model with IGMRF prior is denoted as iMRF. It relies exclusively on simultaneous map updates of map coefficients—as described in Chapter 3.

We use the notation gCAR1-gCAR5 to denote the gCAR models with predictor 1 to 5 (cf. Table 3.1). The same applies analogously to the fCAR model and iMRF model: Model names fCAR1-fCAR5 denote the fCAR models with predictor 1 to 5. Model names iMRF1-iMRF5 denote the iMRF models with predictor 1 to 5.

For all models, we use the canonical HRF with three basis functions to model odd stimulus contribution to the fMRI predictor. Regressors for modeling the even stimulus are not included for introducing collinearity issues. We include the DCT basis set implementing regressors for a highpass filter of 128s and a set of movement parameters and a global signal as confounders. This confounder set has been used for analyzing the real-world dataset, which serves as the prototype for this artificial dataset. Details about the calculation of these regressors can be found in Chapter 2.

In case of the CAR prior, the prior mean of the spatial dependency parameter τ^2 (and/or τ_0^2) is set to 0 and its prior variance is chosen to be rather uninformative with value 25. In a testing phase of the gCAR and fCAR algorithms, we have found that convergency of MCMC trajectories is sensitive to both the choice of ξ^2 (ξ_0^2) hyperparameters, i.e. inverse gamma shape and scale parameters of respective priors, and τ^2 (τ_0^2) hyperparameters, i.e. proposal variance and starting value. Hyperparameters of the algorithm are chosen in such a way that convergency to the equilibrium distribution is acceptable within 1 000 iterations (burnin). Convergency is assessed by visual inspection of trace plots and corresponding diagnostic tools. For variance parameters ξ^2 , we use an inverse gamma prior with shape 227 and scale 1 017. This corresponds to an inverse gamma mean of 4.5 and a variance of 0.09. For ξ_0^2 , we use an inverse gamma prior with shape 452 and scale 4 059. This corresponds to an inverse gamma mean of 9 and a variance of 0.18. Suitable hyperparameters for the spatial dependency parameter τ^2 and τ_0^2 differ between fCAR and gCAR. In the gCAR model, we use the following values. For τ^2 , we found a proposal variance of 0.25 and a starting parameter of 2 to be useful. For τ_0^2 , we chose a proposal variance of 3 and a starting parameter of 25. Note that for gCAR4, these values had slightly to be adapted to yield stationarity. In the fCAR model, we use for τ^2 a starting parameter

of 0.05 and a proposal variance of 0.02. For τ_0^2 , we specify a starting parameter of 5 and a proposal variance of 0.1.

In case of the IGMRF prior, according IGMRF variance parameters follow an inverse gamma distribution with shape 204.5 and scale 915.75. This corresponds to an inverse gamma mean of 4.5 and variance 0.1. In a testing phase of the iMRF algorithm, these parameter values led to acceptable sample trajectories showing a tolerably stationary behavior.

If included in the model (either in gCAR, fCAR or iMRF), global effect parameters have prior mean zero and a variance parameter that follows an inverse gamma distribution with shape parameter 3 and scale parameter 1 which corresponds to inverse gamma mean 0.5 and variance 0.25.

The neighborhood, on which the random fields are based on, are defined upon the next 6 (three-dimensional) voxel neighbors.

For each run, we use 6 000 MCMC iterations including a burnin phase of 1 000 iterations. To remove strong autocorrelations, we thin out resulting parameter trajectories with a stepping of 5.

6.1.3 Evaluation of different EEG predictor forms

The prior probability of each voxel being activated is linked to the EEG measurements via a probit link, $p(\gamma_i = 1|\eta_i) = \Phi(\eta_i)$, $i = 1, \dots, N$. The predictor contains an intercept term plus an EEG effect, i.e. $\eta_i = f_0(i) + f(J_i)$. In Table 3.1, all predictor forms of interest are listed. We assume that at least one of the two components $f_0(i)$ or $f(J_i)$ contains a spatially-varying coefficient for adapting to local brain response. If an IGMRF prior is used for a local effect, the according global effect is also included in the model. For the model with CAR priors, it suffices to include either a global or local effect. See Section 3.3 for a detailed discussion of predictor types and prior assumptions.

We conduct the following study to determine which predictor form is most suitable for EEG-informed fMRI activation detection. For this, we run both the CAR prior models (global and fast version) and IGMRF prior models with all possible predictor forms, i.e. we apply gCAR1-gCAR5, fCAR1-fCAR5 and iMRF1-iMRF5 algorithms. Evaluation is then based upon several criterions. We examine convergency properties, the ability to produce reasonable probit effect estimates, and their sensitivity to detect activation. The figures used in this section to display probit effect estimates comprise the following contents:

Column	Expression	Description
1	$\hat{\mathbf{f}}_0(i) = \left(\hat{f}_0(i) \right)$	Intercept map estimate
2	$\hat{\mathbf{f}}(J_i) = \left(\hat{f}(J_i) \right)$	EEG effect map estimate
3	$\hat{\boldsymbol{\eta}} = \left(\hat{f}_0(i) + \hat{f}(J_i) \right)$	Predictor map estimate
4	$\hat{\boldsymbol{\pi}} = (\hat{p}(\gamma_i = 1 \boldsymbol{\theta}, \mathbf{y}))$	Posterior activation probability map estimate
5	$\hat{\boldsymbol{\gamma}} = (\hat{\gamma}_i = 1)$	Activation map estimate (*)

(*) $\hat{\boldsymbol{\gamma}}$ is calculated by binarizing $\hat{\boldsymbol{\pi}}$ with threshold 0.8722

We restrict simulation runs to an artificial dataset with medium noise, i.e. error noise ϵ_i^* is generated from $\mathcal{N}(0, 6\sigma_i^{2,*}\mathbf{I})$. Figure 6.3 depicts the map with voxelwise log-likelihood ratio (LR) statistic values for testing whether the fMRI stimulus component is 0 in the fMRI dataset at hand. This map does not depend on the EEG probit stage and is, hence, the same for all models. It depicts the mere fMRI contribution to activation.

Convergency behavior

For determining whether the MCMC algorithms have converged, we look at the trajectories of global parameters and the local parameter trajectories of 20 selected voxels. These selected voxels have been chosen to lie in the center and on the margin of activation regions and in non-activated areas—which include the brain margin. See Figure 6.1 for the position of selected brain voxels in the medium brain layer. Exemplary trace plots and diagnostics for the gCAR5 model are shown in Figure 6.4. Diagnostics of other model runs can be provided upon request.

After an initial search for suitable hyperparameters, the MCMC procedure did hardly show any convergency problems for fCAR. Though global parameter trajectories of variance and spatial dependency parameters exhibit some autocorrelations, the mean level of posteriori distributions seems to be well identified.

Stronger mixing problems arise in both model types with global updates, the gCAR and iMRF models. The gCAR models, however, were more informative concerning possible reasons for convergency problems. Priors of the gCAR and iMRF models rely on different variance-covariance matrix forms. Thereby the CAR prior relies on an additional spatial dependency parameter τ^2 (resp. τ_0^2). After inspection of convergency diagnostics of the gCAR models, we conclude that identifying the level of spatial dependency is problematic for all predictor types, whereas acceptable sample trajectories are achieved for their variance parameters ξ^2 (resp. ξ_0^2). Spatial dependency trajectories exhibit strong

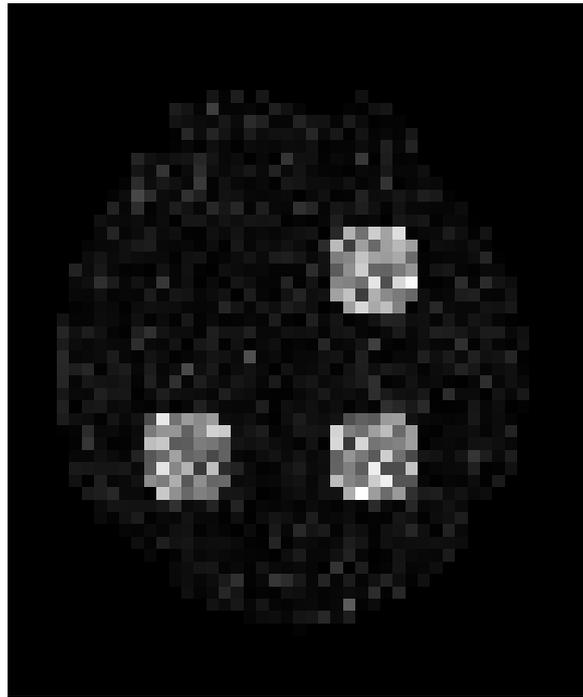


Figure 6.3: This map depicts voxelwise $\log(\text{LR})$ statistic values for testing whether the fMRI stimulus component is 0. These indicate how much activation information is contained in an fMRI signal time series of a voxel. The $\log(\text{LR})$ -map of the central brain layer is depicted, i.e. layer 3.

autocorrelations, which are even detectable by visual inspection. The identification of a mean level, however, seems to work to some extent, i.e. trajectories do not diverge, but samples fluctuate around approximately the same level.

The iMRF variance-covariance parameter ξ^2 (resp. ξ_0^2) subsumes both degree of variance and spatial dependency. Hence, we cannot differentiate between mixing problems in the identification of variance or spatial dependency parameters. Additionally, the centering of spatial-varying effects provides trajectories of global intercept and EEG effect. Surprisingly, mixing issues appear in these global effect trajectories as well. Generally, convergence problems are reflected in occasional departures from a sample mean and higher autocorrelations. In iMRF2, non-stationary behavior of global parameters is reflected in iMRF field samples. The strongest non-stationary behavior is detected in the model with predictor 4 (iMRF4): Especially iMRF4 variance parameter ξ^2 and the global parameters α_G show strong trends. Like in the iMRF2 run, this time non-convergence of global parameters is reflected in the remaining parameter trajectories of the iMRF parameter fields. Conver-

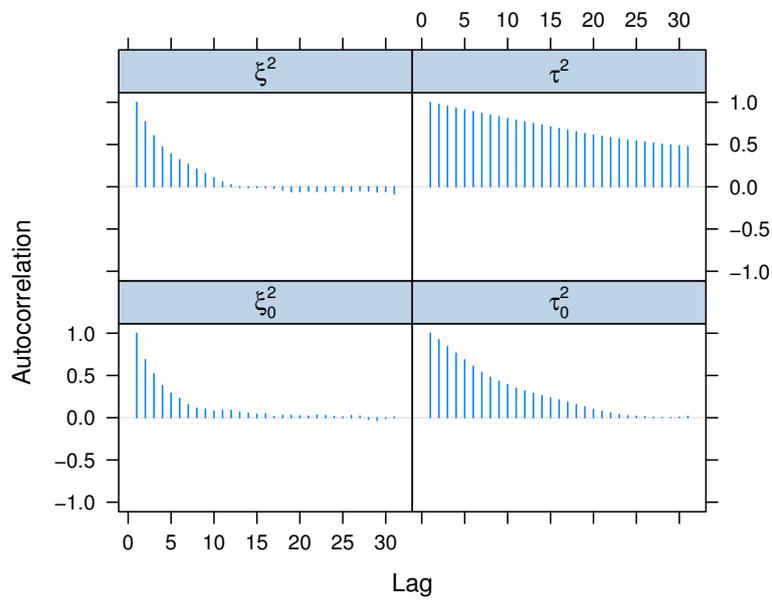
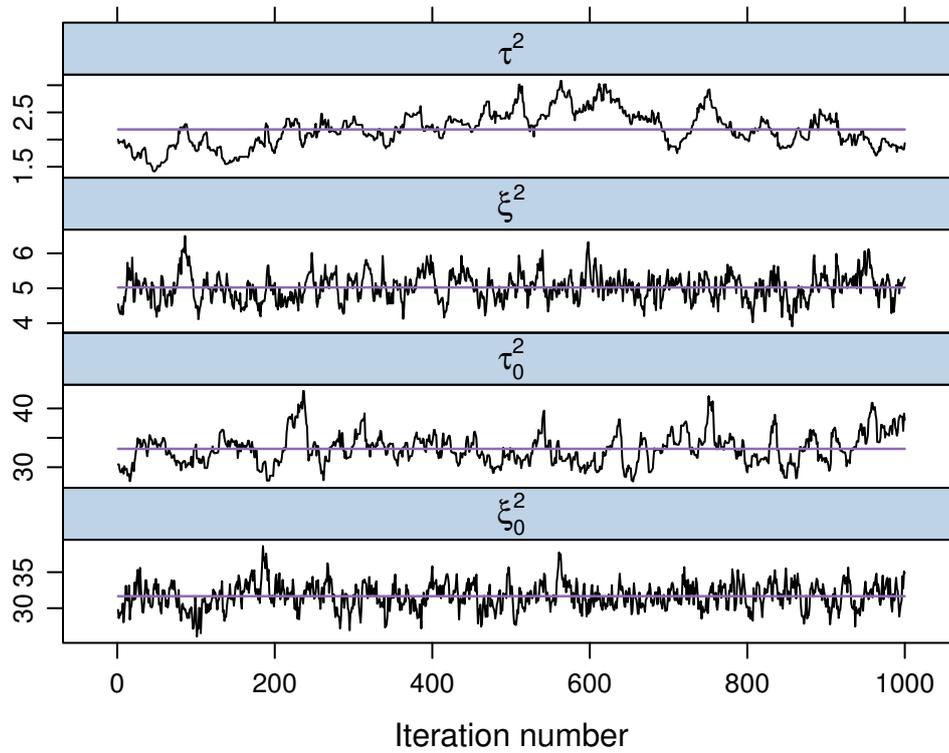


Figure 6.4: Trace plots and corresponding autocorrelation functions of global parameter trajectories of the gCAR5 model.

gency of sample trajectories is improved considerably if both components are allowed to vary spatially with predictor 5 (iMRF5). In summary, we observe that especially the iMRF algorithms have difficulties to identify either the global effect levels and/or the magnitude of variation and spatial dependency within effect maps.

For both the gCAR and iMRF algorithm, it is not out of question whether hyperparameter constellations can be found that improve convergency. However, search for good parameter values is tedious and must be restarted for every new dataset. To evaluate the influence of convergency problems of variance-covariance parameters on parameters of interest, i.e. activation classification, we conducted a robustness study. The results of this study can be found in Appendix B. The main conclusion from this study is that even strong variations in variance parameter values lead to nearly negligible changes in activation maps. Hence, we consider variance-covariance parameters as nuisance and do not try any further to improve their convergency behavior.

Influence of predictor forms in the gCAR model

In Figure 6.5, the results of gCAR models are shown for brain layer 3. Corresponding activation maps of all models are compared to the underlying true activation map from the simulation design in Figure 6.6. Based on the assumption that gCAR5 contains the full model, we compare its posterior activation probabilities with the results of the remaining models via difference maps in Figure 6.7. Sensitivity and specificity estimates are presented in Table 6.2.

We found that estimates for gCAR1 contain a negative estimate for the global intercept, thus, reducing the influence of non-activated regions. Based on its parametrization, the EEG part of the predictor is 0 when EEG measurements are zero. We observe that EEG effects are non-zero as soon as EEG measurements are non-zero. That is, in region R_1 where no fMRI activation is present, EEG shows a slight negative effect, and in regions R_3 and R_4 , congruent fMRI activation is mirrored in positive EEG effects. Otherwise, the local EEG effect is zero. Surprisingly, the activation in R_2 is mirrored in $\hat{\pi}$ although no positive predictor value could be estimated for it. This is likely due to the fact that posterior activation probability estimates depend on a transformation of fMRI LR values (cf. Equation (3.17), p. 61), which is able to compensate for this misspecification in the predictor. Compared to gCAR5, this model obtains smaller activation probability estimates overall, which leads especially in activation regions to a strong decrease in sensitivity. The

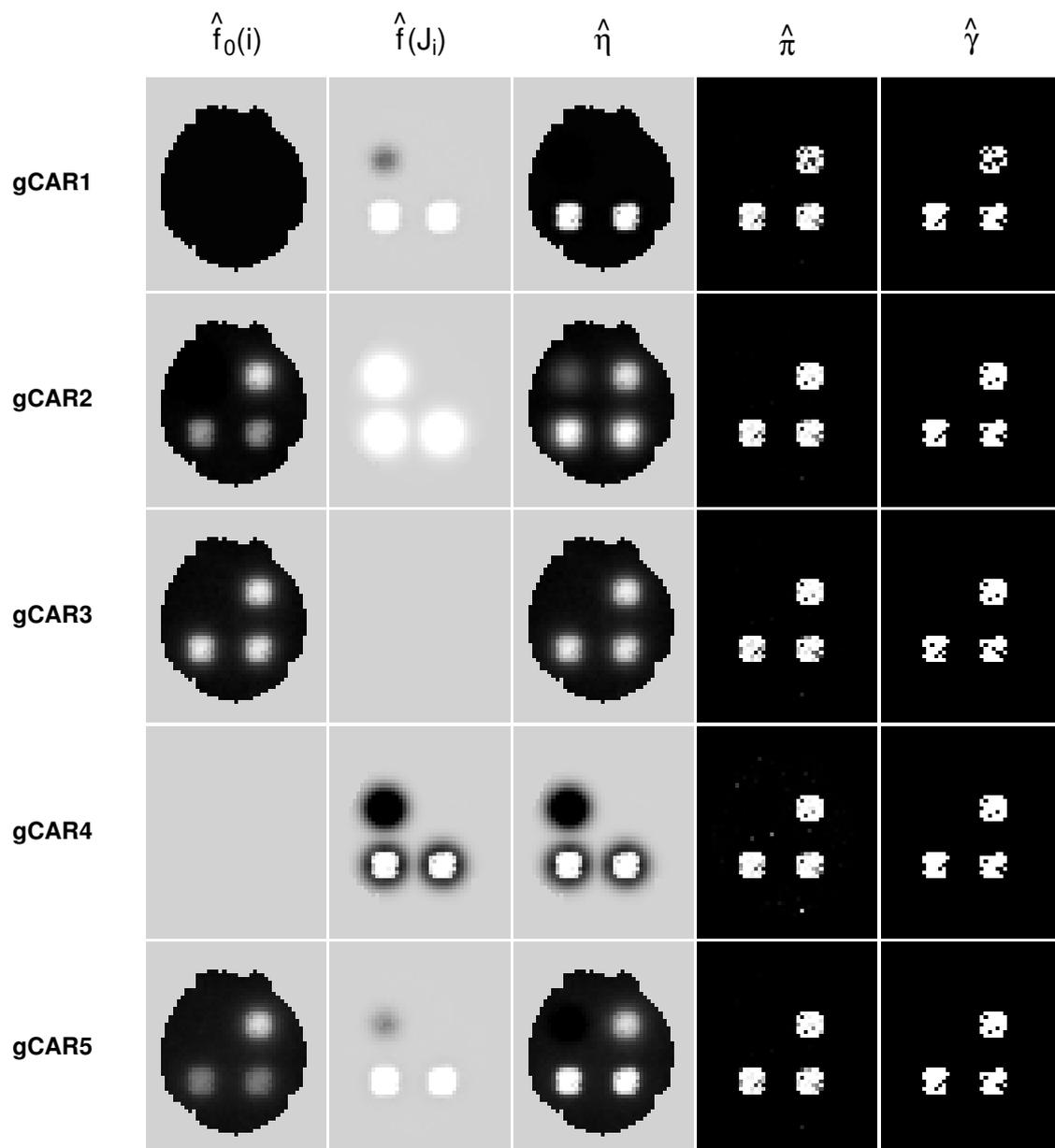


Figure 6.5: Map estimates of the EEG probit stage calculated by the gCAR algorithms for brain layer 3 (cf. legend on p. 118).

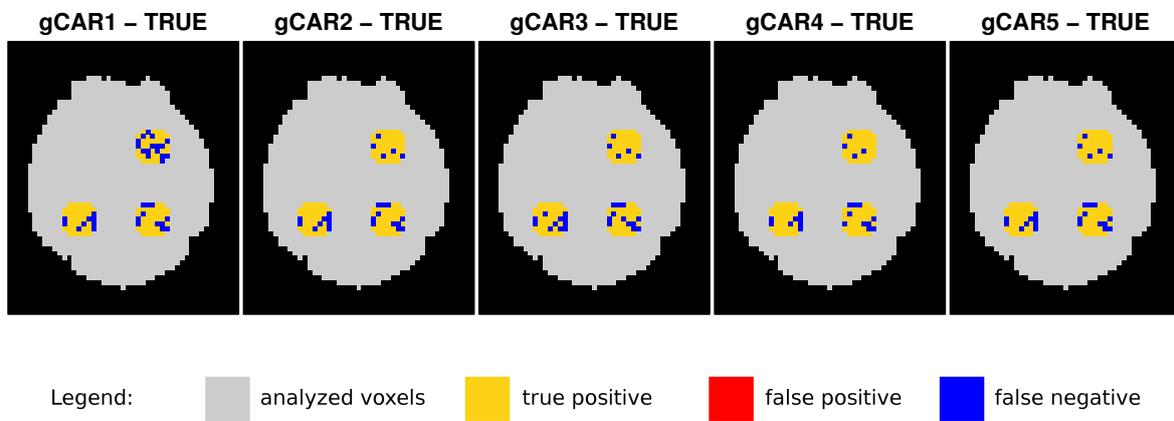


Figure 6.6: Activation maps of gCAR models in comparison to the true activation map as used in the simulation design. The maps contain the results of the central brain layer, i.e. layer 3.

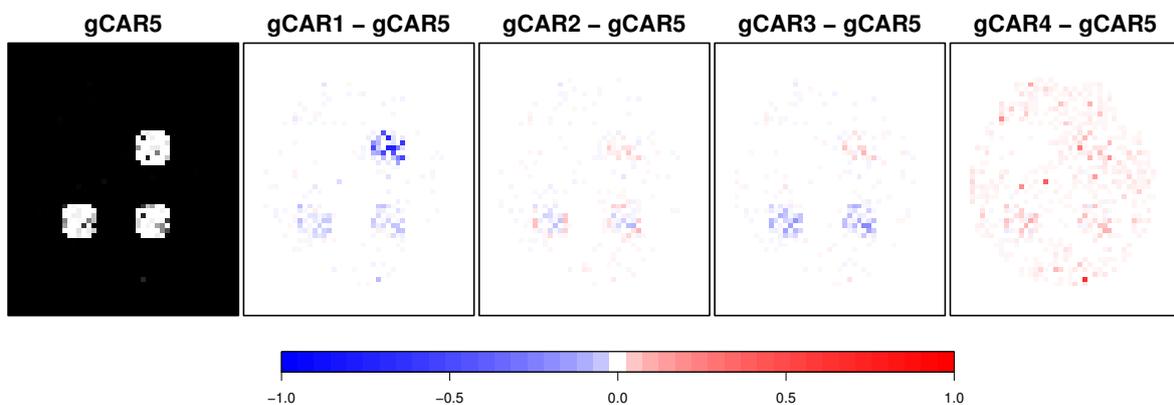


Figure 6.7: Activation probability map estimates calculated by the gCAR algorithms for layer 3. In the leftmost plot, $\hat{\pi}_{gCAR5}$ is plotted. The following plots depict the differences in π -estimation of the remaining models to $\hat{\pi}_{gCAR5}$. Higher probability values of the gCAR5 model are marked in shades of blue, lower probability values are marked in shades of red.

Predictor	gCAR	fCAR	iMRF	Predictor	gCAR	fCAR	iMRF
1	0.7356	0.6743	0.7701	1	1.0000	1.0000	1.0000
2	0.8161	0.5824	0.9042	2	1.0000	0.9998	1.0000
3	0.8046	0.8544	0.8314	3	1.0000	0.9998	1.0000
4	0.8582	0.8544	0.8544	4	1.0000	1.0000	1.0000
5	0.8276	0.8582	0.9080	5	1.0000	0.9997	1.0000

(a) Sensitivity. (b) Specificity.

Table 6.2: Influence of EEG predictor forms on sensitivity and specificity. Note that fCAR specificity is calculated in relation to the used brain mask. Negligible changes occur if the cuboid mask is used.

largest negative differences are obtained for R_2 . Hence, it seems disadvantageous that R_2 cannot be recovered by this type of predictor.

Estimates for gCAR2 contain a spatially-varying intercept and a global EEG coefficient. For the latter, a positive estimate is obtained leading to the desired interpretation that EEG has a positive effect on fMRI activation. Hence, the global EEG effect can be used to increase activation probability where positive EEG measurements are present. A global EEG effect, however, prohibits that the regionally varying relationship between EEG and fMRI is captured: Region R_1 gets the same large positive EEG effect as R_3 and R_4 . The spatially-varying intercept compensates this predictor misspecification to some extent. No false positives occur in R_1 . The sensitivity of gCAR2 is slightly lower than the sensitivity of gCAR5 and larger than the ones of gCAR1 and gCAR3. A further advantage compared to predictor type 1 is that R_2 is now reflected in the spatial intercept estimate.

Estimates for the gCAR3 model contain only the spatially-varying intercept. The corresponding map estimate reflects the fMRI log(LR) statistic map in Figure 6.3 well. The missing EEG component does not lead to a substantial decrease in sensitivity compared to the full model gCAR5. But it apparently yields lower activation probability estimates in regions R_3 and R_4 . We do not want to discuss this phenomenon here. This is postponed to a general comparison of methods on pp. 135.

Estimates for the gCAR4 model contain only the EEG predictor component. Due to the parametrization of this model, the predictor cannot decrease the effect to subzero values (and thus to prior activation probability values below 0.5) in regions with zero EEG measurements. Only non-zero EEG measurements can be used to adapt probit effect estimates to the data. We see that R_3 and R_4 are well recovered by the spatially-varying EEG effect. Good edge-preserving properties of this prior are mirrored in the strong drop the effect field exhibits on the edge of these activation regions. As desired, R_1 obtains

negative effect estimates. Like in the gCAR1 model, region R_2 cannot be recovered in the predictor map estimate. Peculiarities of this predictor type can also be seen in the posterior activation probability map: The brain background obtains higher probability estimates than gCAR5, which might be an undesirable feature of the missing intercept component leading possibly to false positives. Activation regions R_3 and R_4 (and also R_2 , which is hardly silhouetted against brain background) yield larger posterior probabilities leading at last to an increased sensitivity compared to gCAR5.

In the gCAR5 model both model components are allowed to vary spatially. The spatially-varying intercept reflects non-activation in the brain background by negative estimates, but R_2 is still recovered in the intercept estimate. In region R_1 where no fMRI activation is observed, non-activation is supported by a slight negative EEG effect. This even leads to a drop in the total predictor map below brain background. If activation is present in EEG and fMRI, the EEG model components obtain a larger effect compared to the intercept component. Hence, both model components do not contribute equally to the total predictor map. One might be inclined to interpret the two model components as follows: The intercept reflects the mere BOLD contribution to activation, whereas the EEG component reflects effects on activation, which can be traced back to the common source of fMRI and EEG. This interpretation, however, must be used with caution. Whether this interpretation is valid has to be corroborated by a specially designed experiment where both sources can be separated. Whether such a design is realizable is, however, doubtful. Beyond that, the separation of intercept and EEG component might also be connected with identification problems of both components. This might be the reason for the decrease in sensitivity compared to gCAR4. A workaround for possible identification problems of this type does not exist at the moment.

Influence of predictor forms in the fCAR model

We compare results of the fCAR model with the gCAR model. In Figure 6.8, the predictor and activation related result maps of the fCAR model are plotted. The most prominent difference between these two figures appears to be the underlying brain mask. Whereas the gCAR model can be run with various forms of brain masks (like the used head shape), the fCAR model is based on a cubical brain mask including all brain voxels, which is applied for accelerating the update of CAR spatial dependency parameters. Thus with the fCAR model, we also obtain effect estimates for non-brain regions. This, however, has no effect on activation (probability) map estimates. As can be inferred from voxelwise

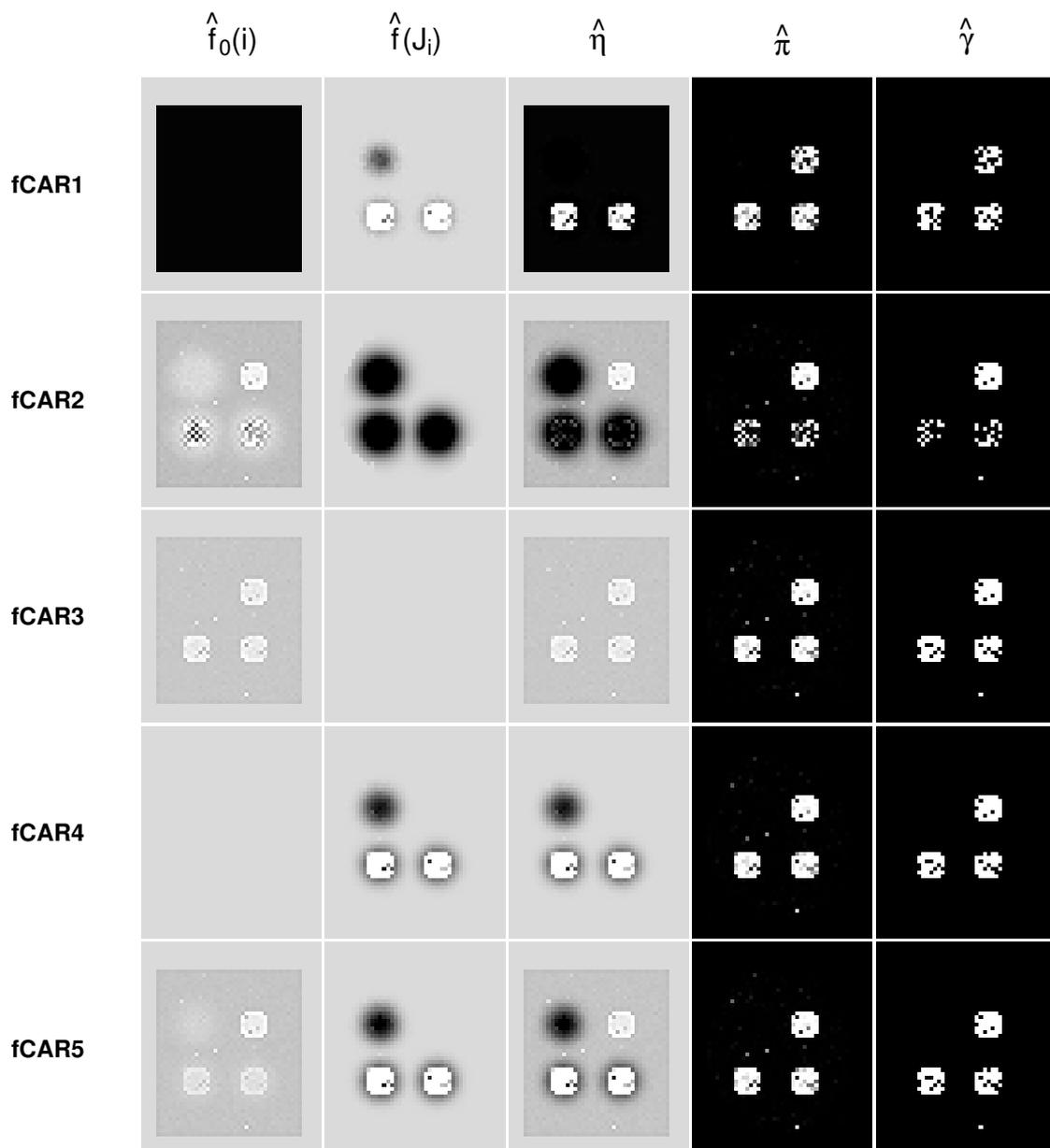


Figure 6.8: Map estimates of the EEG probit stage calculated by the fCAR algorithms for brain layer 3 (cf. legend on p. 118).

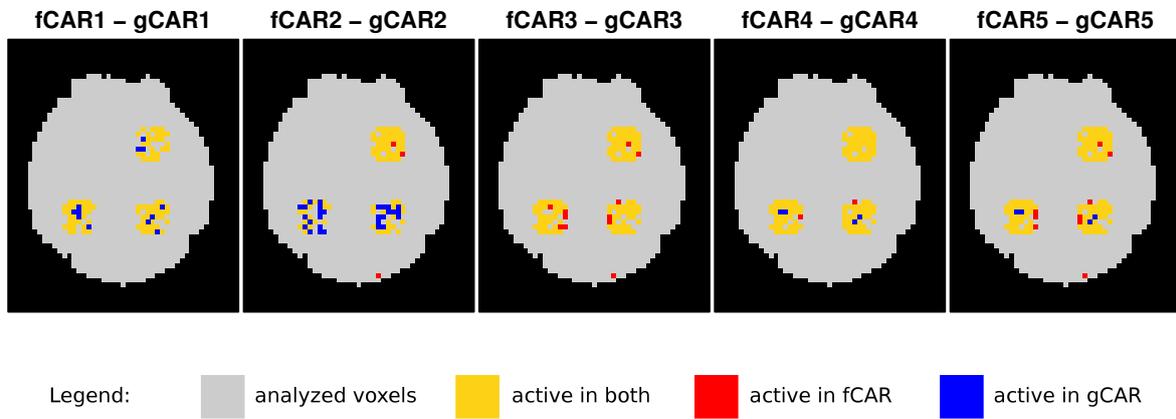


Figure 6.9: Activation maps of fCAR models in comparison to corresponding gCAR activation maps. Separate plots depict differences in classifications for each predictor type (1-5). The maps contain the results of the central brain layer, i.e. layer 3.

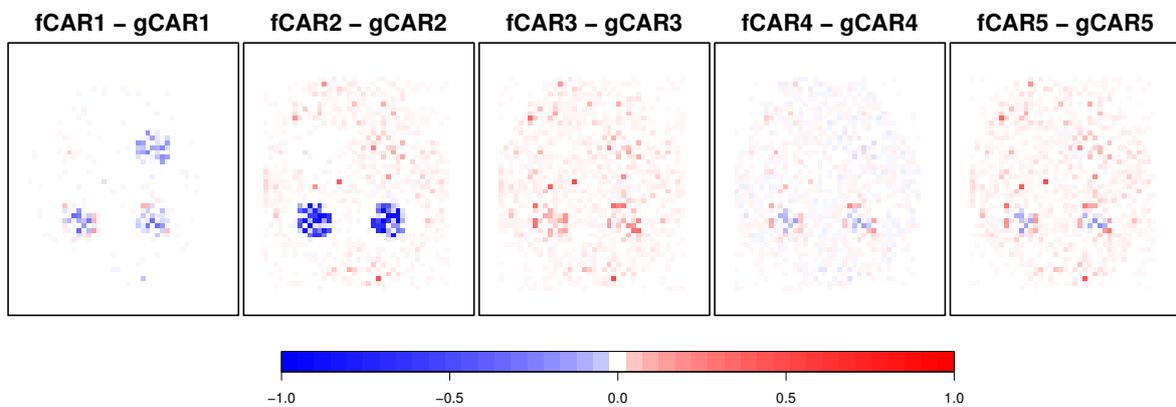


Figure 6.10: Activation probability map estimates calculated by the fCAR algorithms in comparison to corresponding gCAR maps. Separate plots depict the differences in π -estimation for each predictor type (1-5). Higher probability values of the fCAR model are marked in shades of red, lower probability values are marked in shades of blue. The maps contain the results of the central brain layer, i.e. layer 3.

difference maps of activation probabilities (Figure 6.10), non-brain voxels have activation probability values of about 0. Activation difference maps between gCAR and fCAR models can be found in Figure 6.9.

Pred.	Min.	1% Qu.	25% Qu.	Median	Mean	75% Qu.	99% Qu.	Max
1	-0.355	-0.004	0.000	0	-1.618e-03	0.000	0.001	0.311
2	-0.994	-0.009	0.000	0	-5.747e-03	0.001	0.027	0.497
3	-0.003	0.000	0.000	0	6.014e-03	0.002	0.064	0.521
4	-0.146	-0.005	0.000	0	5.854e-04	0.001	0.009	0.385
5	-0.204	-0.001	0.000	0	5.112e-03	0.002	0.051	0.567

Table 6.3: Summary of in activation probability difference values between gCAR and fCAR models

Activation probability differences between gCAR and fCAR models depend on the predictor type. Maximum and minimum difference values per predictor type appear to be large, but these can be recognized as outliers. Most of the difference values lie in the interval $[-0.01; 0.065]$ (cf. Table 6.3 for an overview of summary statistics of difference values).

For predictor type 1, only minor differences are obtained for the brain background, but fCAR1 obtains considerably lower activation probabilities for regions R_2 , R_3 and R_4 —which goes along with a drop in sensitivity.

For predictor type 2, fCAR2 obtains a negative global EEG effect estimate leading to the counterintuitive interpretation that EEG has a negative effect on fMRI activation. In region R_3 and R_4 , the spatially-varying intercept cannot compensate for this. Hence, activation probability values are low and many voxels are falsely classified as non-active. Voxels in the brain background yield larger activation probabilities than the global model.

Results for predictor types 3, 4 and 5 are similar. For each type, the brain background yields higher probability values in the fCAR models than in the gCAR models. This can be seen as an undesirable property of the fast algorithm that is not able to separate activated and non-activated regions as well as the global model. Even false positive voxels are obtained. Within activated regions, we observe the following: fCAR outperforms gCAR when it comes to detect active voxels at edges to non-activated regions. Nevertheless, gCAR appears to be more stable to identify active voxels in the center of activation regions—where fluctuating effects in the fCAR model lead to a drop in sensitivity. If calculated over the whole brain region, fCAR (with predictor 3, 4, 5) obtains higher sensitivity values than the corresponding gCAR model. Unfortunately, it seems that using EEG information does not result in an increase of sensitivity.

It is not obvious why fCAR and gCAR algorithms yield such different results. After all, fCAR and gCAR rely on the same probability model. However, differences might be traced back to differences in variance and spatial dependency parameter estimates (results not shown). These might be provoked by deviating update schemes and analysis masks. The fCAR algorithms yield estimates by also averaging over non-brain areas with possibly lower spatial correlations. This might lead to smaller global estimates for the spatial dependency parameter, which—in turn—can enhance edge-detection, but interferes with exploiting information from neighboring voxels.

Influence of predictor forms in the iMRF model

In Figure 6.11, the results of iMRF models are shown for brain layer 3. Sensitivity and specificity estimates are given in Table 6.2 on p. 124. Activation maps of all iMRF models are compared to the underlying true activation map from the simulation design in Figure 6.12. Based on the assumption that iMRF5 contains the full model, we compare its posterior activation probabilities with the results of the remaining models via difference maps in Figure 6.13.

At first sight, we notice that predictor map estimates of type 1-4 are very similar to the estimates of the gCAR model. Effect estimates on edges of regions appear slightly more blurred than with the gCAR model. That is, the IGMRF prior yields smoother results. This might have two consequences: Either activated voxels on edges are more easily detected than in the gCAR model or more false positive voxels are found that lie directly at the border to activated regions. Concerning the full model, iMRF5 estimates differ slightly from gCAR5. On the one hand, the spatially-varying EEG effect is based on a larger variance estimate, so that non-activated regions neighboring activation sites are covered by the positive EEG effect region. This is partly compensated by the spatially-varying intercept effect. On the other hand, R_1 is assigned a positive EEG effect and no negative effect as in gCAR5. The effect size is not as pronounced as in the actual activation regions R_3 and R_4 and is sufficiently low for not producing false positives. A positive EEG effect region in R_1 may indicate that the model tries to transcribe supporting EEG information into positive effects—which is our intention.

Compared to the gCAR results, the iMRF models obtain higher sensitivity values (except for predictor type 4). Especially for predictor 2 and 5, a very high sensitivity can be achieved, which is even larger than for corresponding fCAR models. A comparison between iMRF5 (resp. iMRF2) and iMRF3 indicates that EEG information adds value to the

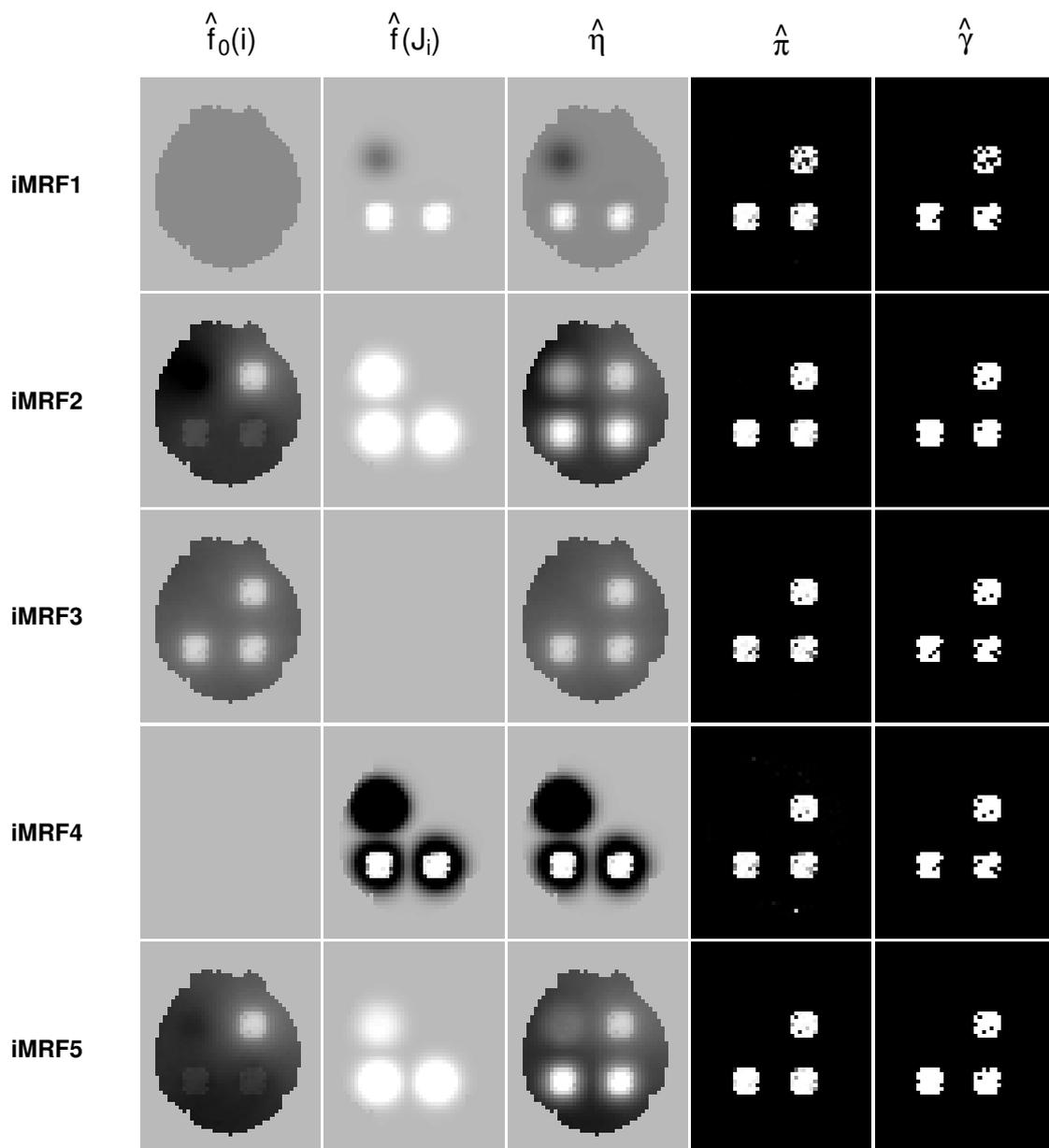


Figure 6.11: Map estimates of the EEG probit stage calculated by the iMRF algorithms for brain layer 3 (cf. legend on p. 118).

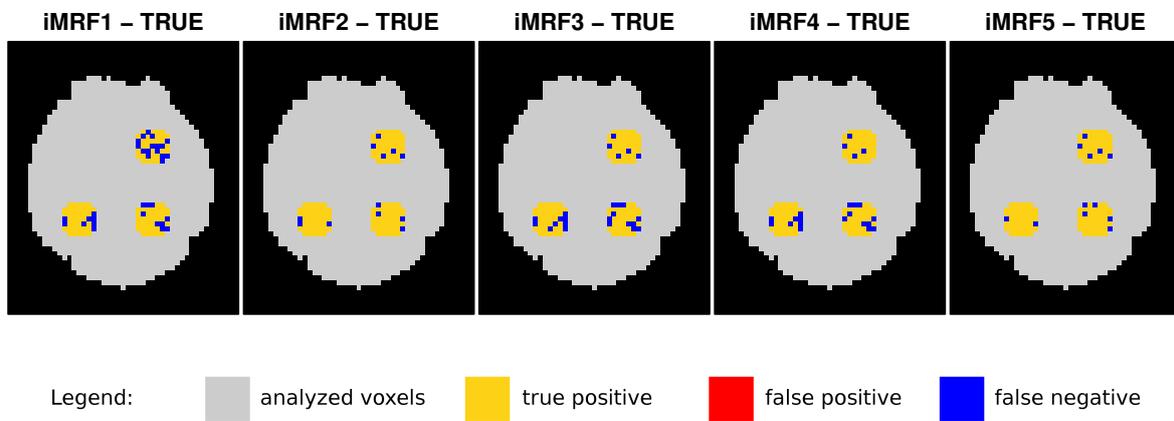


Figure 6.12: Activation maps of iMRF models in comparison to the true activation map as used in the simulation design. The maps contain the results of the central brain layer, i.e. layer 3.

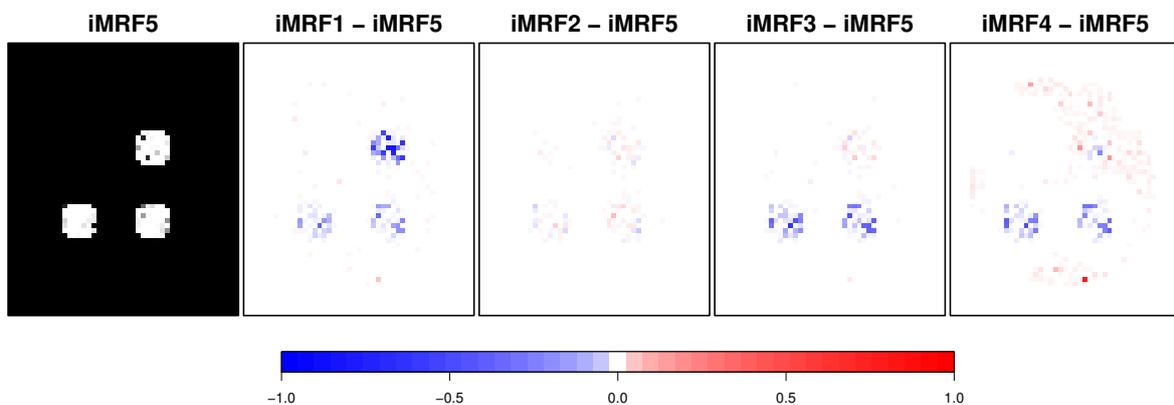


Figure 6.13: Activation probability map estimates calculated by the iMRF algorithms for layer 3. In the leftmost plot, $\hat{\pi}_{iMRF5}$ is plotted. The following plots depict the differences in π -estimation of the remaining models to $\hat{\pi}_{iMRF5}$. Higher probability values of the iMRF5 model are marked in shades of blue, lower probability values are marked in shades of red.

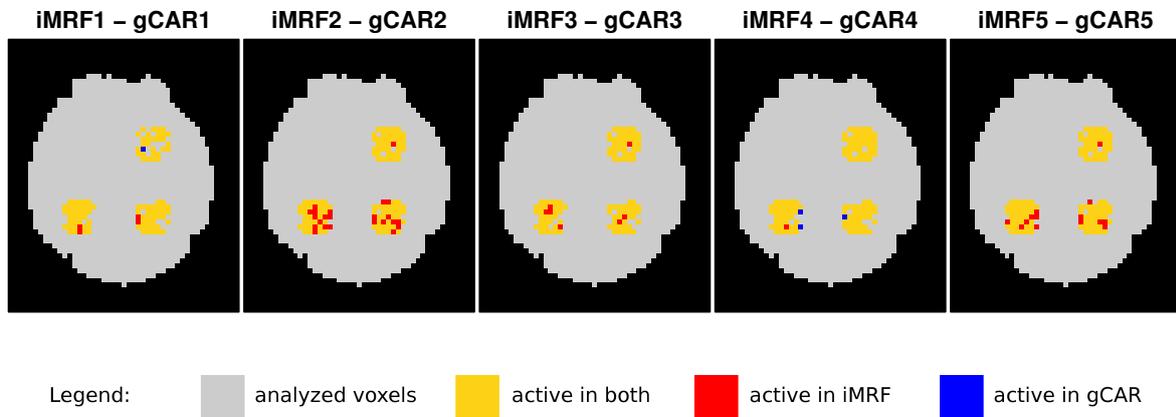


Figure 6.14: Activation maps of iMRF models in comparison to corresponding gCAR activation maps. Separate plots depict differences in classifications for each predictor type (1-5). The maps contain the results of the central brain layer, i.e. layer 3.

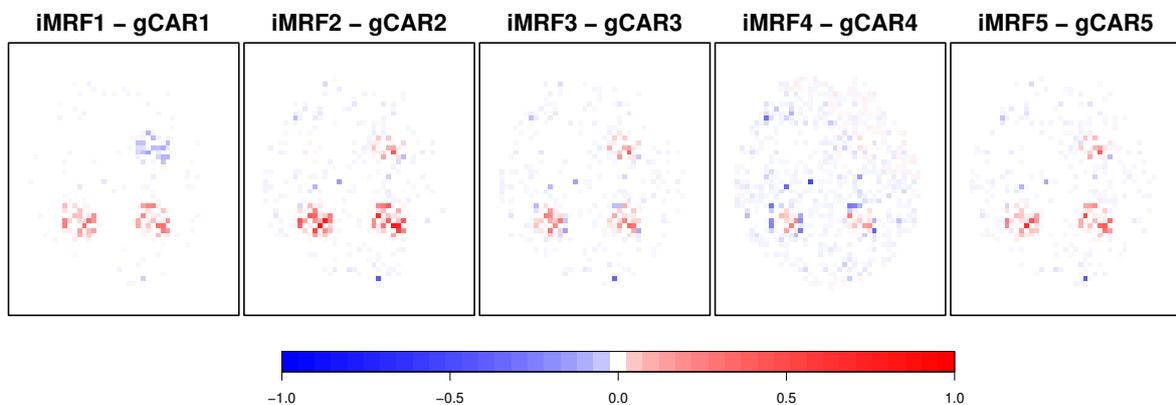


Figure 6.15: Activation probability map estimates calculated by the iMRF algorithms in comparison to corresponding gCAR maps. Separate plots depict the differences in π -estimation for each predictor type (1-5). Higher probability values of the iMRF model are marked in shades of red, lower probability values are marked in shades of blue. The maps contain the results of the central brain layer, i.e. layer 3.

activation detection: The increase in sensitivity by using EEG information is substantial, i.e. in the order of magnitude of 7%.

Figures 6.14 and 6.15 directly depict the differences between iMRF and gCAR results. It seems likely that the higher degree of estimated smoothness leads to an increase in sensitivity in smooth activated regions for the iMRF. At first sight, this seems to be a desirable feature of the iMRF model. However, smoothness is present by simulation design in the dataset at hand. The performance of the iMRF model remains to be tested in settings with rougher activation structures. Enforced smoothness might then lead to an increase in classification errors.

Summary

In summary, the examination of predictor and spatial prior types yields the following conclusions. The activation maps are quite insensitive to misspecifications in the probit predictor stage. Surprisingly, although R_2 cannot be recovered via predictor 1 and 4, reasonable posterior activation probability estimates evolve because the fMRI LR statistic share in the posterior π distribution compensates for this. That is, posterior activation probability estimates also depend on a transformation of fMRI LR values (cf. Equation (3.17), p. 61). This leads to appropriately high activation probabilities for voxels showing fMRI but no EEG activation.

However, the exact form of the predictor does matter. Differences in sensitivity are observed for the examined predictor types with predictor 5 being the most (or second most) sensitive one. Virtually no differences are obtained in specificity values, which might be traced back to the simulation design where brain background is generated with higher white noise.

For the following analyses, we prefer to use the full predictor 5 with both spatially-varying components. The full predictor is able to recover EEG influence on the one hand and fMRI influence on the other hand. Particularly, the spatially-varying intercept makes the fMRI LR contribution visible in probit effect estimates. It is able to capture fMRI activation where zero EEG measurements are observed (as in R_2). The brain background is identified by negative estimates downweighting prior activation probabilities in these regions. The spatially-varying EEG coefficient adapts to local brain response to EEG. If there is no fMRI activation in spite of non-zero EEG measurements, the according EEG effect is reduced (in the iMRF case) or estimated to be zero or subzero (in the gCAR case). In regions with congruent EEG and fMRI activation, we recover a positive EEG effect.

Even though the gCAR4 model obtains superior performance compared to gCAR5, we do not recommend its use. The gCAR4 algorithm seems to have worse convergence properties than the other predictor types and its parametrization yields unreasonable probit effect estimates. However, it might be interesting to examine its performance in a real-world application.

Alternatively, an EEG-enhanced fMRI detection model can be based on predictor type 2. A positive EEG effect then tries to upweight all regions with positive EEG measurements. Predictor type 2 can, hence, be interpreted as a model implementing a less flexible prior: Whether reasonable or not, the prior probability for activation is enforced to be higher in regions with positive EEG measurements. Considered from this point of view, this is the direct way of substituting (non-adaptive) expert knowledge with EEG information. However, if local EEG information is non-congruent to fMRI information, the spatially-varying intercept estimate can compensate for this in this region to some extent. The predictor type 2 models as used here, however, carry one risk: If a negative EEG effect is recovered (e.g. because of the EEG and fMRI activation is mostly not congruent), a counterintuitive situation evolves leading to the interpretation that EEG has a negative effect on (fMRI) activation. Hence, introducing a restriction of non-negative global EEG estimates seems sensible. A model extension for this is derived in Section 3.5.

Based on the results presented in this section, we cannot give a final recommendation whether gCAR or iMRF models should be preferred. In this simulation study, iMRF models showed clearly the best performance (including the desired increase in power when EEG information is used). This, however, can be due to the smooth simulation design where the iMRF model with its better smoothing properties can be advantageous. Rougher datasets may require the better edge-preserving properties of the gCAR algorithm.

The comparison between gCAR and fCAR model appears to be difficult. On the one hand, computation time of the fCAR algorithm is in fact in an order of magnitude shorter. Convergence issues are easier to be solved and higher sensitivity estimates are obtained. On the other hand, results of the fCAR model appear to be less stable than in the gCAR model, which might be due to single-site updates of spatial fields. For example, fCAR2 recovers a negative global EEG effect leading to false positives, brain background is upweighted and centers of activation regions are not as well recovered as might be desired. Additionally, in the fCAR algorithm, we cannot retrieve a gain in power when using EEG information. From this point of view, further evaluation of the fCAR algorithm is necessary to determine whether its use in practice is recommendable.

For the following analyses, we use all three model types: iMRF, gCAR and fCAR. For an EEG-enhanced activation detection algorithm, we use predictor type 5 models and the newly developed predictor type 2 models with an global EEG effect that has been restricted to be non-negative (cf. Section 3.5).

6.1.4 Comparison of different fMRI activation detection algorithms

This part of the simulation study is conducted to assess the comparative performance of the proposed Bayesian activation detection algorithms (with both IGMRF and CAR priors) without EEG information. Existing methods provide no means of using additional information to enhance activation detection—so far. Hence, for fairness we compare only our proposed predictor type 3 models, i.e. gCAR3, fCAR3, iMRF3, with the performance of competing activation detection software. Performance relates to the power of the different algorithms to detect activated voxels and therefore, we quantify performance in terms of sensitivity and specificity.

We compare our proposed models to the following methods: The Ising model from Smith et al. (2003) and the widely used classical SPM algorithm (Friston et al., 1995) as implemented in SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/software/>).

Like our method, the Ising model is a Bayesian activation detection model applying a Bayesian variable selection approach to determine if stimulus components contribute to explaining the fMRI signal. Bayesian variable selection approaches were introduced by Smith and Kohn (1996) and George and McCulloch (1997) for selecting covariates in linear regression models via i.i.d. indicator priors. To account for spatial structures these approaches had to be extended. As indicated by its name, the Ising model relies on an Ising prior to spatially regularize the estimation of a binary field and hence, corresponding activation probabilities—whereas our model uses a probit link to a predictor with spatially-varying effects regularized by the spatial CAR or IGMRF prior. The Ising prior is parameterized by two parameters: A spatial dependency parameter θ , which models the influence intensity a neighboring voxel has on estimation at a given voxel, and the—so-called—external field δ which is depending on the specified prior probability of activation. The external field δ is usually chosen upon prior believes. It can also be spatially-varying if spatially-varying information is available, e.g. in form of a gray-matter probability map. In our artificial dataset no such information exists, so that we decided to use a global δ parameter. In Smith and Fahrmeir (2007), an algorithm was presented

for estimating the spatial dependency parameter θ . However, the authors do not provide software for fitting the model. Therefore, we decided to rely on a simple grid-search to find optimal values of δ and θ in the simulated dataset of medium noise—where noise ϵ_i^* is generated from $\mathcal{N}(0, 6\sigma_i^{2,*} \mathbf{I})$. We found that sensitivity can arbitrarily be increased with δ becoming larger. Hence, in this simulation study we chose δ in such a way (Smith et al., 2003) that it corresponds to a prior activation probability of 0.04146, which is the frequency of active voxels in the true activation map of the simulation design (relative to voxels analyzed). Besides this we found that results are rather insensitive to the choice of θ . Therefore, we set it to a medium level of 0.45, which also has been used in Smith et al. (2003).

As mentioned above, we chose the classical SPM model for comparing our models to a widely used classical activation detection algorithm. Activation maps are derived by thresholding an F -map from the voxelwise statistical F -Tests of HRF effects being zero. A multiple test correction is applied via the implemented random field theory (Worsley et al., 1996). Significance level was set to 0.05. For not being directly comparable (compare Section 3.4.5), we first evaluate the effect of different classification thresholds on the performance of our models and compare, then, the corresponding activation maps (differing in their sensitivity level) to classical SPM results on pp. 142.

We analyze seven different datasets differing in the level of noise: $\sigma_i^{2,*}$, $2\sigma_i^{2,*}$, $4\sigma_i^{2,*}$, $6\sigma_i^{2,*}$, $8\sigma_i^{2,*}$, $10\sigma_i^{2,*}$, $12\sigma_i^{2,*}$.

Comparison of competing Bayesian activation detection algorithms

In this section, we focus on a comparison of Bayesian activation detection algorithms that do not rely on external EEG information, i.e. the proposed predictor type 3 models and the Ising model of Smith et al. (2003). Activation differences to reference model gCAR3 are depicted in Figure 6.16. Activation probability differences to reference model gCAR3 are shown in Figure 6.17. In Table 6.4 sensitivities for the different Bayesian activation schemes are printed for different classification thresholds.

We foremost base interpretations of activation classifications and, hence, interpretation of sensitivity and specificity values on a classification threshold of 0.8722 (cf. Section 3.4.5). Robustness of results with respect to other choices of thresholds are discussed below. For ease of comparability, we choose gCAR3 as a reference model to highlight differences between algorithms.

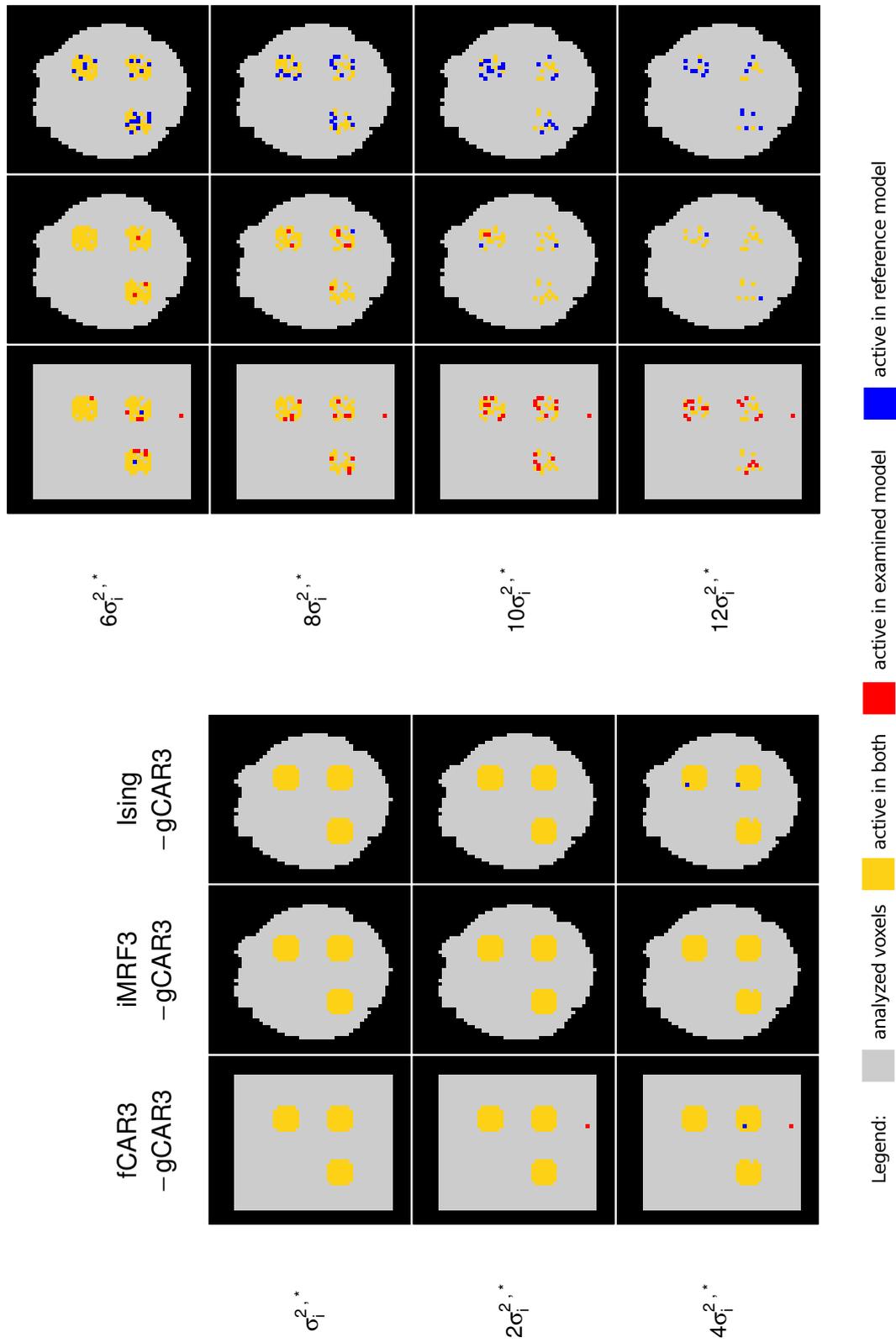


Figure 6.16: Activation difference maps of fCAR3, iMRF3 and Ising models in comparison to the activation map of gCAR3 for different error noise. The maps contain the results of the central brain layer, i.e. layer 3.

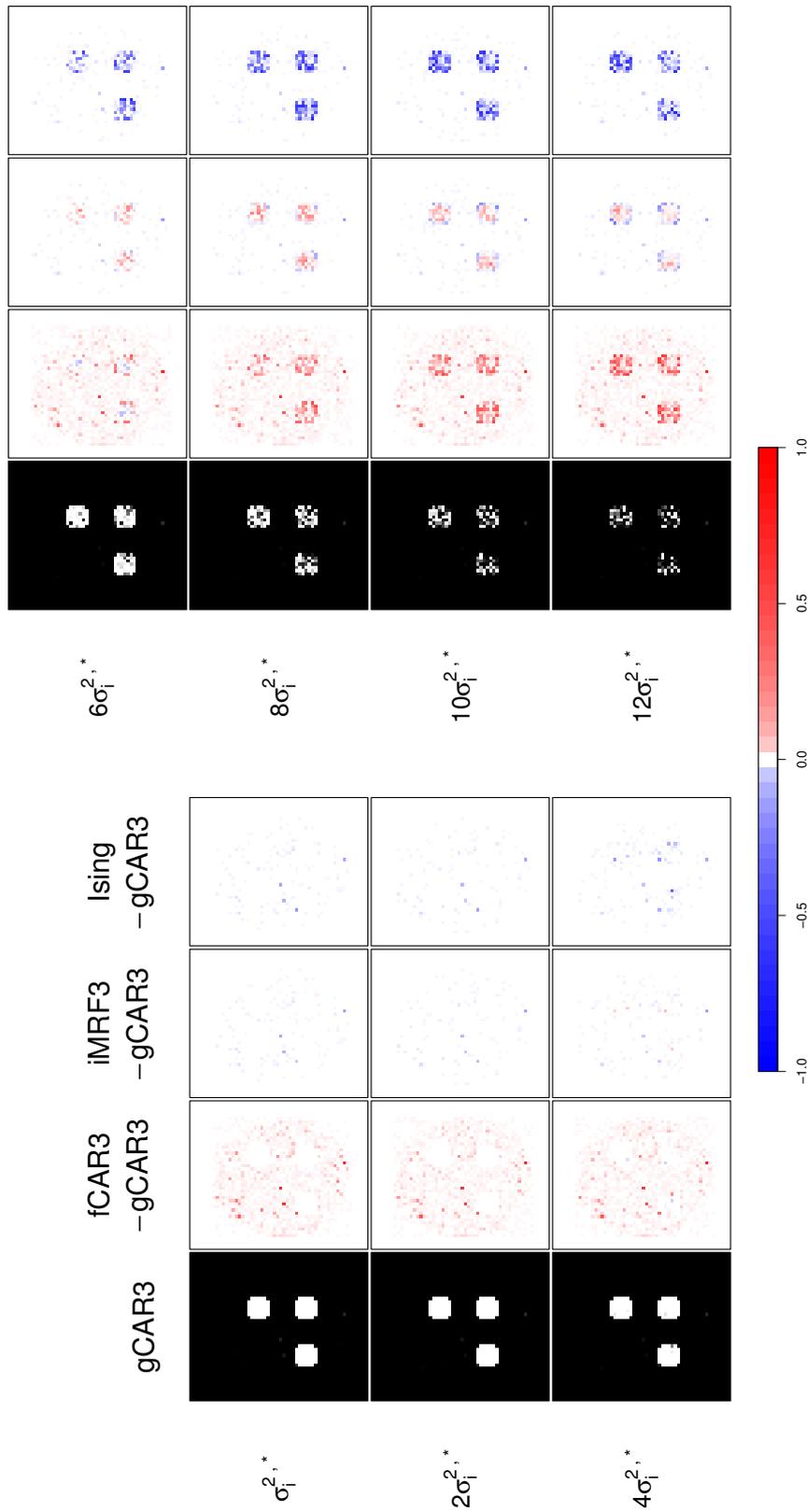


Figure 6.17: Activation probability maps of fCAR3, iMRF3 and Ising models in comparison to the activation probability map of gCAR3 for different error noise. In the leftmost plot, $\hat{\pi}_{gCAR3}$ is plotted. The following plots depict the differences in π -estimation of the fCAR3, iMRF3 and Ising models to $\hat{\pi}_{gCAR3}$. Higher probability values of the gCAR3 model are marked in shades of blue, lower probability values are marked in shades of red. The maps contain the results of the central brain layer, i.e. layer 3.

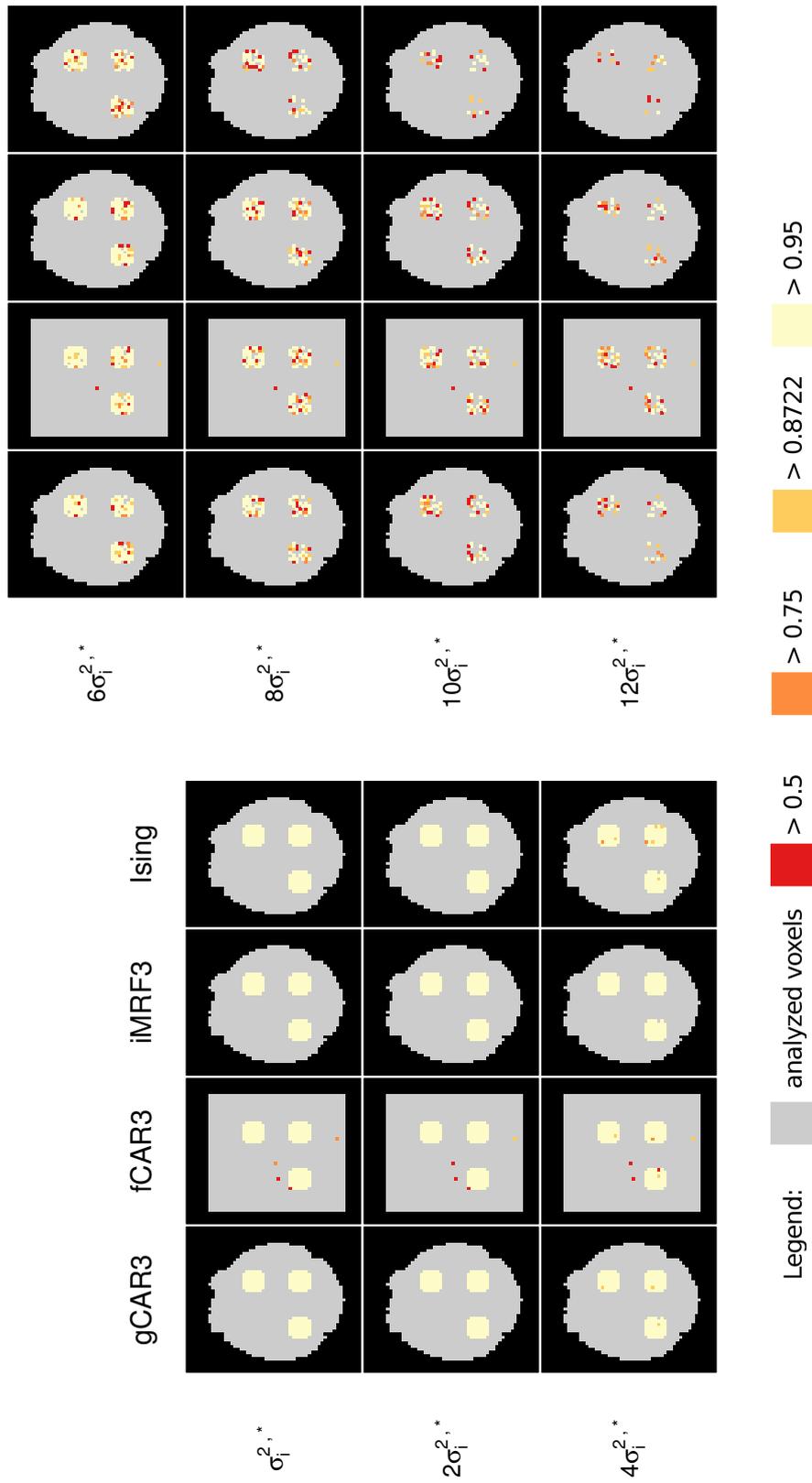


Figure 6.18: Threshold dependent activation maps for the gCAR3, fCAR3, iMRF3 and Ising models. Color codes represent activated voxels with activation probability exceeding thresholds 0.5, 0.75, 0.8722 and 0.95. The brighter the color the higher is the probability threshold the respective voxel’s activation probability survives. The maps contain the results of the central brain layer, i.e. layer 3.

Whereas sensitivity is 1 for all methods in the lowest noise levels, differences increase with noise level. The higher the noise level, the worse is the sensitivity of the Ising model compared to gCAR3. The worse performance of the Ising model can also be inferred from the posterior activation probability difference maps in Figure 6.17. The Ising model obtains lower activation probability estimates than the gCAR3 model over the whole brain region, but especially in activation regions. This might be due to two reasons. First, spatial dependency parameter θ possibly needs to be adjusted in high noise level settings. Second, our choice of external field parameter δ might be suboptimal in this artificial setting. In real data situations, a more meaningful choice of δ (or rather δ) might lead to an increase in sensitivity. The evaluation of this hypothesis is postponed to Section 6.2.

There are slight differences between the iMRF3 and gCAR3 model at medium noise levels: In runs with $4\sigma_i^{2,*}$, $6\sigma_i^{2,*}$, $8\sigma_i^{2,*}$, iMRF3 performs better. In all other runs, sensitivity is approximately at the same level in both models. Activation probability estimates indicate that iMRF3 obtains increased values in activation region centers, whereas gCAR3 achieves higher probabilities at corresponding edges.

The fCAR3 results are substantially different to the gCAR3 results. Like in Section 6.1.3, fCAR3 generally obtains higher probability estimates in the brain background, which can be a source of false positives. In our simulation design, only a small decrease in specificity can be observed, which is likely to be traced back to the peculiarities of our simulation design. At medium and high noise levels, fCAR3 yields an astonishing increase in sensitivity. It outperforms all remaining models in an order of magnitude. Differences between gCAR3 and fCAR3 models are surprising, because they do not differ in the underlying probability model, but only in their algorithmic implementation: fCAR3 relies on single-site updates of spatially-varying coefficients and uses a cubical brain mask for analyses (needed for the fast determinant calculation). We see that posterior activation probability maps of gCAR3 and fCAR3 hardly differ outside the brain region. At first sight, this might indicate that not the brain mask, but only single-site updates are likely to lead to these differences. However, both factors may provoke comparably small fCAR3 spatial dependency estimates (not shown). On the one hand, the large-scale spatial correlation structure may not be caught by single-site updates and, on the other hand, low spatial correlations in non-brain areas might have an diminishing effect on a spatial dependency parameter that is aggregated over all analyzed voxels. A small spatial dependency parameter is connected to good edge-preserving properties, but inhibits advantages going along with higher smoothness. Hence, fCAR3 is capable of capturing coarse underlying true activations, but also might suffer from noisy effect map estimates.

	0.5	0.75	0.8722	0.95
$\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000
$2\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000
$4\sigma_i^{2,*}$	0.9962	0.9962	0.9962	0.9847
$6\sigma_i^{2,*}$	0.9234	0.8659	0.8046	0.7548
$8\sigma_i^{2,*}$	0.6858	0.5977	0.5479	0.4751
$10\sigma_i^{2,*}$	0.4943	0.3831	0.3372	0.2644
$12\sigma_i^{2,*}$	0.3218	0.2490	0.1954	0.1418

(a) gCAR3

	0.5	0.75	0.8722	0.95
$\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000
$2\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000
$4\sigma_i^{2,*}$	1.0000	0.9962	0.9923	0.9847
$6\sigma_i^{2,*}$	0.9387	0.9042	0.8544	0.7701
$8\sigma_i^{2,*}$	0.7625	0.6858	0.6245	0.5441
$10\sigma_i^{2,*}$	0.6284	0.5211	0.4713	0.3985
$12\sigma_i^{2,*}$	0.5057	0.4253	0.3448	0.2720

(b) fCAR3

	0.5	0.75	0.8722	0.95
$\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000
$2\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000
$4\sigma_i^{2,*}$	0.9962	0.9962	0.9962	0.9962
$6\sigma_i^{2,*}$	0.9387	0.8774	0.8314	0.7739
$8\sigma_i^{2,*}$	0.7241	0.6360	0.5862	0.5019
$10\sigma_i^{2,*}$	0.5134	0.4100	0.3410	0.2874
$12\sigma_i^{2,*}$	0.3257	0.2529	0.1916	0.1379

(c) iMRF3

	0.5	0.75	0.8722	0.95
$\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000
$2\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000
$4\sigma_i^{2,*}$	0.9962	0.9923	0.9808	0.9464
$6\sigma_i^{2,*}$	0.8008	0.7318	0.6705	0.6015
$8\sigma_i^{2,*}$	0.4866	0.3908	0.3372	0.2989
$10\sigma_i^{2,*}$	0.2567	0.1877	0.1494	0.1111
$12\sigma_i^{2,*}$	0.1418	0.0881	0.0651	0.0345

(d) Ising

Table 6.4: Sensitivity of algorithms for selected thresholds.

Specificity is perfect for all methods and noise levels—except for the fCAR3 algorithm, which shows a marginal decrease. This is likely to be owed to the simulation design (cf. Table 6.7 for specificity values of our proposed models; specificity values of the Ising model are not shown here, but can be provided on request).

For evaluating the robustness of results, we examine changes in activation classification (Figure 6.18) and corresponding changes in sensitivity (Table 6.4). Clearly more voxels are classified as active with decreasing probability threshold, but differences between algorithms are of similar size for most noise levels. Only for high noise, sensitivity differences of Ising and fCAR3 models compared to gCAR3 decrease with increasing threshold. Specificity estimates do hardly change: For the fCAR3, model a slight drop in specificity can be observed for decreasing threshold value, the remaining models retain a perfect specificity level (results not shown).

Comparison with classical SPM

A comparison of Bayesian activation detection schemes based on posterior activation probability estimations and classical methods is not as straightforward as in the previous section, where algorithms rely on both the same inference and detection technique. Despite the application in Smith et al. (2003) and Smith and Fahrmeir (2007), theoretical considerations yield that the relationship between the classical significance level and a Bayesian probability threshold—these publications rely on—does not hold (cf. Section 3.4.5). Hence, we decided to evaluate the robustness of resulting activation maps in relation to several choices of significance levels (SPM) and probability thresholds (Bayesian procedures).

In Table 6.4, sensitivity estimates of Bayesian algorithms are listed for different probability thresholds. In Table 6.5, sensitivity estimates of the SPM procedure are listed for selected significance levels. Corresponding categorical activation maps are depicted in Figure 6.18 (Bayesian) and 6.19 (SPM).

In the following, we compare the sensitivity values of the least conservative significance level of 0.2 with the most strict probability threshold of the Bayesian activation detection methods, i.e. 0.95. The fCAR3 algorithm outperforms SPM in any case. The gCAR3 and iMRF3 methods perform only in very high noise levels worse. At threshold 0.8722, SPM yields only a better sensitivity value than gCAR3 and iMRF3 at noise level $12\sigma_i^{2,*}$. Hence, we conclude that our proposed Bayesian activation schemes have a superior performance in most of the cases.

	0.2	0.1	0.05	0.01	0.001
$\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000	1.0000
$2\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000	1.0000
$4\sigma_i^{2,*}$	0.9732	0.9732	0.9425	0.9004	0.7510
$6\sigma_i^{2,*}$	0.7241	0.6820	0.6284	0.5249	0.3525
$8\sigma_i^{2,*}$	0.4981	0.4674	0.3946	0.2529	0.1494
$10\sigma_i^{2,*}$	0.3218	0.2529	0.2261	0.1341	0.0421
$12\sigma_i^{2,*}$	0.2107	0.1686	0.1341	0.0575	0.0115

Table 6.5: Sensitivity of classical SPM for selected significance levels.

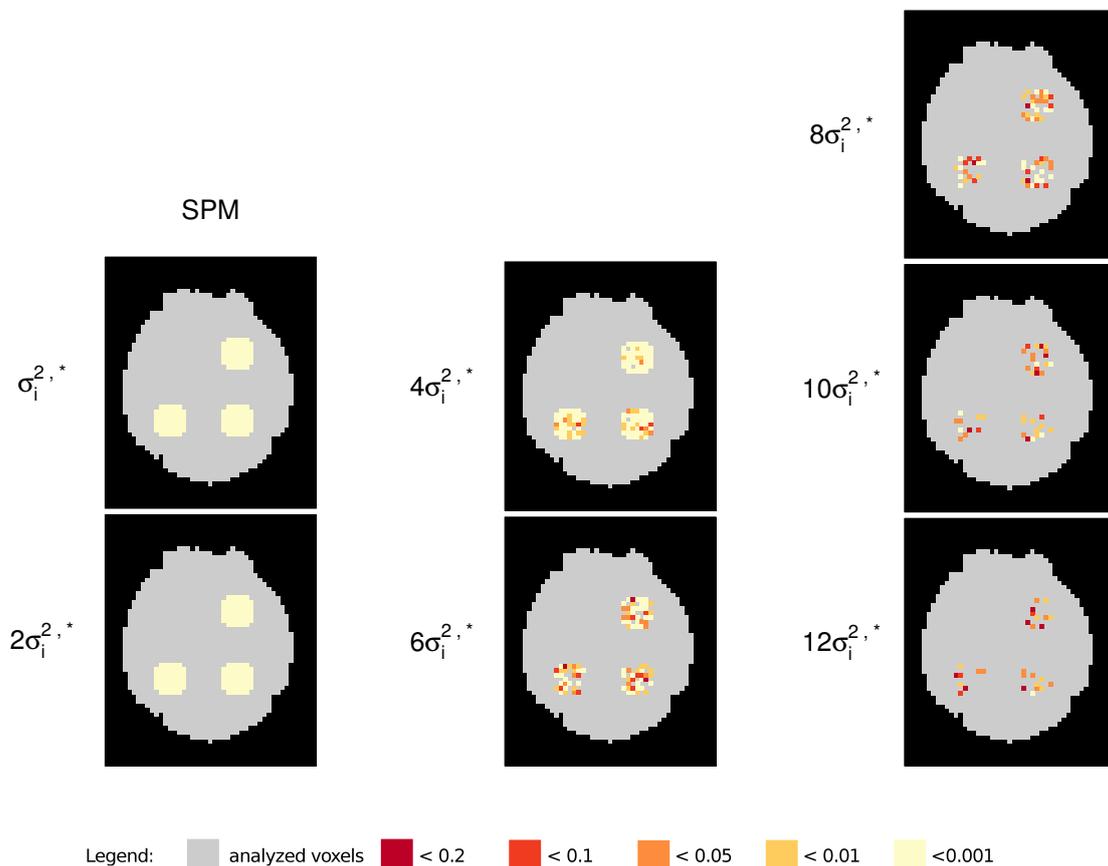


Figure 6.19: Significance level dependent activation maps for SPM. Color codes represent activated voxels which exceed the critical value corresponding to significance level 0.001, 0.01, 0.05, 0.1 and 0.2. The maps contain the results of the central brain layer, i.e. layer 3.

The conclusions from the comparison to the Ising model is not quite as clear-cut as before. SPM still inclines to perform better at higher noise levels. For the Ising model lacking an externally motivated prior information (for determining the external field parameters) in this artificial data setting, we postpone a comparison between these two methods to a subsequent part of this thesis.

6.1.5 Evaluation of EEG-enhanced detection schemes

To evaluate whether the proposed models are able to increase sensitivity by using EEG information, we compare our EEG-enhanced activation detection algorithms with the corresponding algorithms without EEG information.

For the proposed models, we make use of the results from the previous simulation study in Section 6.1.3. We use predictor type 3 models as null models, which do not rely on additional EEG information. Due to its better stability and interpretability, we use predictor 5 to include EEG information. Additionally, we examine predictor type 2 models with a global EEG effect restricted to be non-negative (cf. Section 3.5), which can be interpreted as a model with a less flexible EEG-prior. That is, if a voxel has a positive EEG measurement, activation probability should be forced to be increased—though, the spatially-varying intercept estimate is able to compensate too large EEG effects when necessary.

We use the same settings as in the preceding sections. We analyze the seven artificial datasets differing in the level of noise: $\sigma_i^{2,*}$, $2\sigma_i^{2,*}$, $4\sigma_i^{2,*}$, $6\sigma_i^{2,*}$, $8\sigma_i^{2,*}$, $10\sigma_i^{2,*}$, $12\sigma_i^{2,*}$.

In Figure 6.20, activation difference maps of gCAR2 and gCAR5 models are plotted in comparison to the activation map of the null model gCAR3. In Figure 6.21, activation probability maps of gCAR2 and gCAR5 models are plotted in comparison to the activation probability maps of gCAR3. Activation difference maps indicate minor differences between predictor types. In Table 6.6 and Table 6.7, sensitivity and specificity values are printed, respectively. As can be seen from Table 6.6, sensitivity can only slightly be increased by using gCAR2 and gCAR5 at medium noise level. Model gCAR5 performs somewhat better than gCAR2. At high noise level only gCAR5 has a marginally higher sensitivity than the other two models—which are on a comparable level. Hence, adding EEG information to gCAR does hardly yield a gain in power. However, from Figure 6.21, we see that gCAR2 and gCAR5 in principle do what they are supposed to do: At medium and high noise level activation probabilities are larger in regions with supporting EEG information. This increase, however, cannot be observed for all voxels in activation regions R_3 and

	gCAR3	gCAR2	gCAR5	fCAR3	fCAR2	fCAR5	iMRF3	iMRF2	iMRF5
$\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$2\sigma_i^{2,*}$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$4\sigma_i^{2,*}$	0.9962	0.9962	0.9962	0.9923	0.9923	0.9923	0.9962	0.9962	0.9962
$6\sigma_i^{2,*}$	0.8046	0.8238	0.8276	0.8544	0.8582	0.8582	0.8314	0.8927	0.9080
$8\sigma_i^{2,*}$	0.5479	0.5747	0.5862	0.6245	0.6245	0.6245	0.5862	0.6590	0.6743
$10\sigma_i^{2,*}$	0.3372	0.3410	0.3448	0.4713	0.4713	0.4713	0.3410	0.3985	0.4215
$12\sigma_i^{2,*}$	0.1954	0.1916	0.2184	0.3448	0.3487	0.3410	0.1916	0.2184	0.2299

Table 6.6: Evaluation of EEG enhancement: Sensitivity of Bayesian activation detection schemes in artificial simulation study

	gCAR3	gCAR2	gCAR5	fCAR3	fCAR2	fCAR5	iMRF3	iMRF2	iMRF5
$\sigma_i^{2,*}$	1.0000	1.0000	1.0000	0.9998	0.9998	0.9998	1.0000	1.0000	1.0000
$2\sigma_i^{2,*}$	1.0000	1.0000	1.0000	0.9998	1.0000	0.9998	1.0000	1.0000	1.0000
$4\sigma_i^{2,*}$	1.0000	1.0000	1.0000	0.9998	0.9998	0.9997	1.0000	1.0000	1.0000
$6\sigma_i^{2,*}$	1.0000	1.0000	1.0000	0.9998	0.9998	0.9997	1.0000	1.0000	1.0000
$8\sigma_i^{2,*}$	1.0000	1.0000	1.0000	0.9998	0.9998	0.9998	1.0000	1.0000	1.0000
$10\sigma_i^{2,*}$	1.0000	1.0000	1.0000	0.9997	0.9998	0.9998	1.0000	1.0000	1.0000
$12\sigma_i^{2,*}$	1.0000	1.0000	1.0000	0.9998	0.9998	0.9998	1.0000	1.0000	1.0000

Table 6.7: Evaluation of EEG enhancement: Specificity of Bayesian activation detection schemes in artificial simulation study. Note that fCAR specificity is calculated in relation to the used brain mask. Negligible changes occur if the cuboid mask is used.

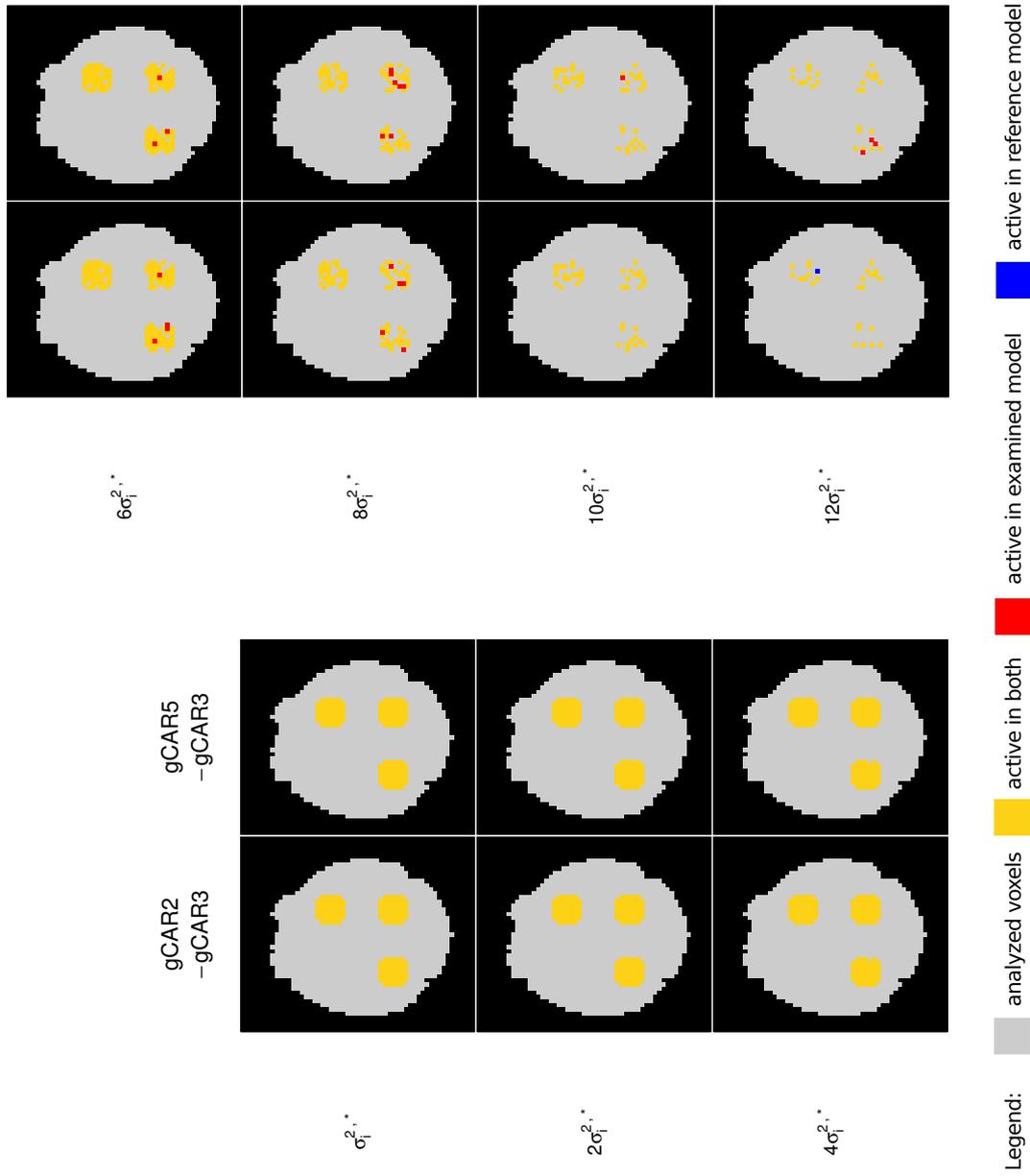


Figure 6.20: Activation difference maps of gCAR2 and gCAR5 models in comparison to the activation map of gCAR3 for different error noise. The maps contain the results of the central brain layer, i.e. layer 3.

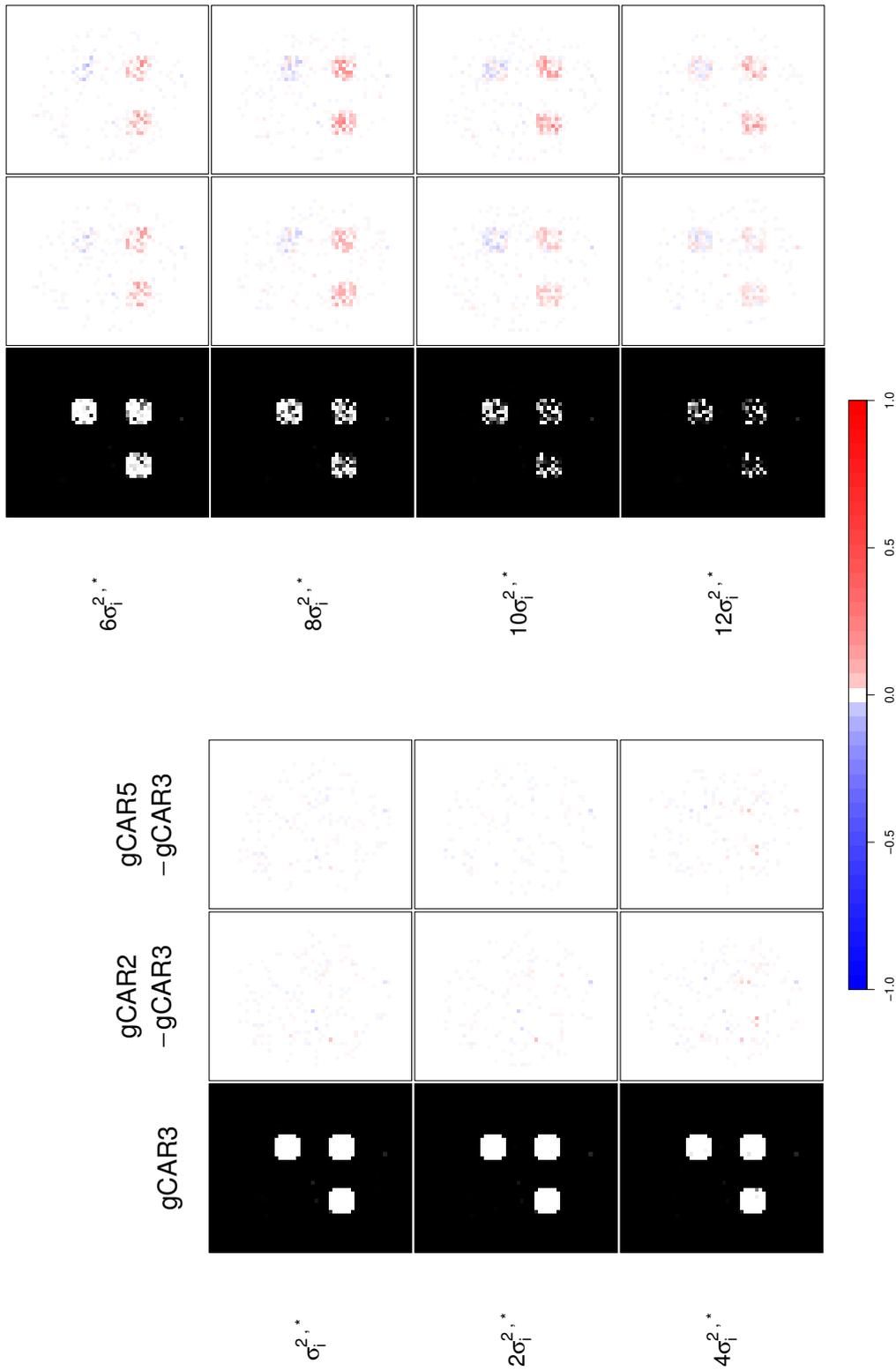


Figure 6.21: Activation probability maps of gCAR2 and gCAR5 models in comparison to the activation probability map of gCAR3 for different error noise. In the leftmost plot, $\hat{\pi}_{gCAR3}$ is plotted. The following plots depict the differences in π -estimation of the gCAR2 and gCAR5 models to $\hat{\pi}_{gCAR3}$. Higher probability values of the gCAR3 model are marked in shades of blue, lower probability values are marked in shades of red. The maps contain the results of the central brain layer, i.e. layer 3.

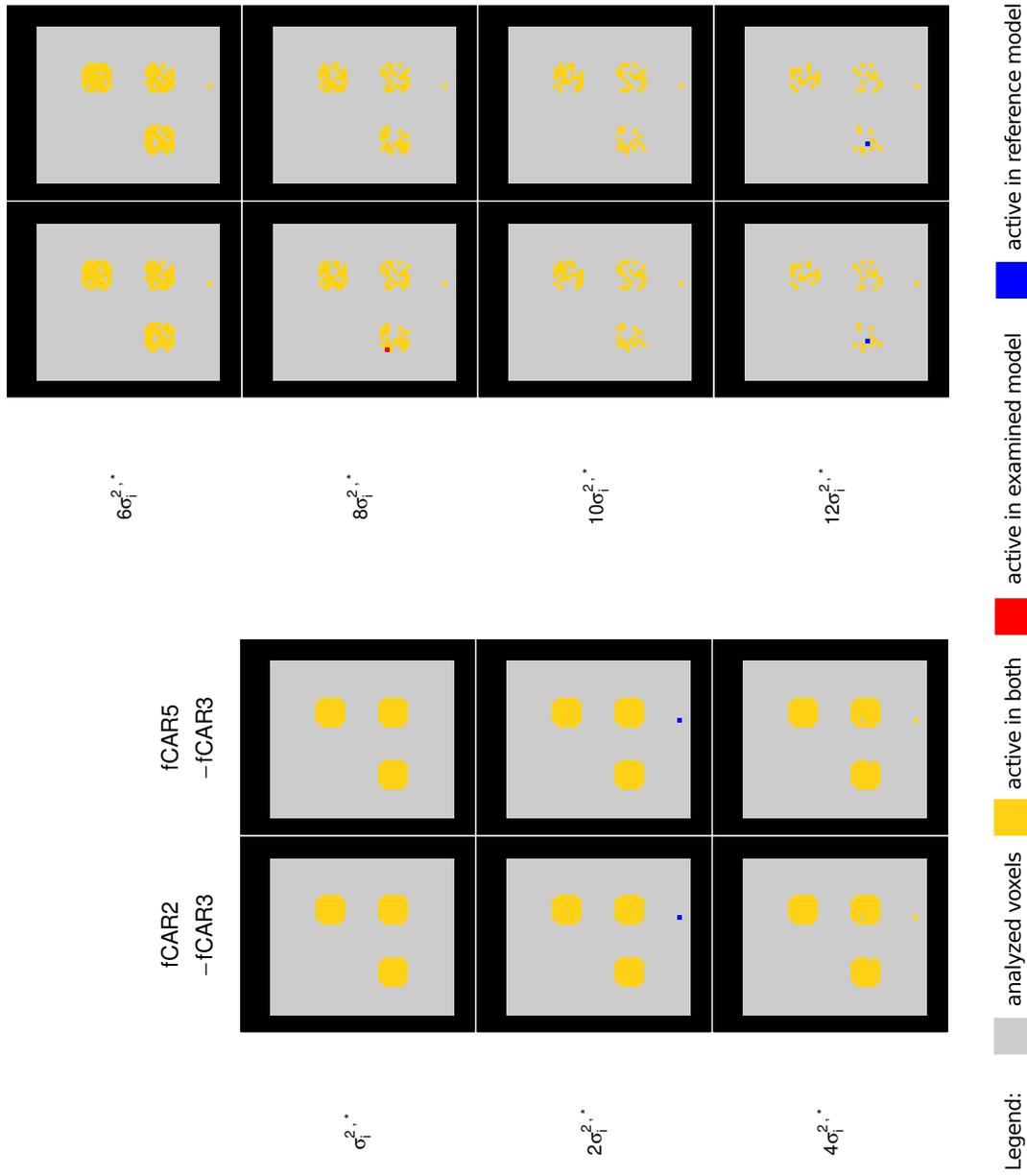


Figure 6.22: Activation difference maps of fCAR2 and fCAR5 models in comparison to the activation map of fCAR3 for different error noise. The maps contain the results of the central brain layer, i.e. layer 3.

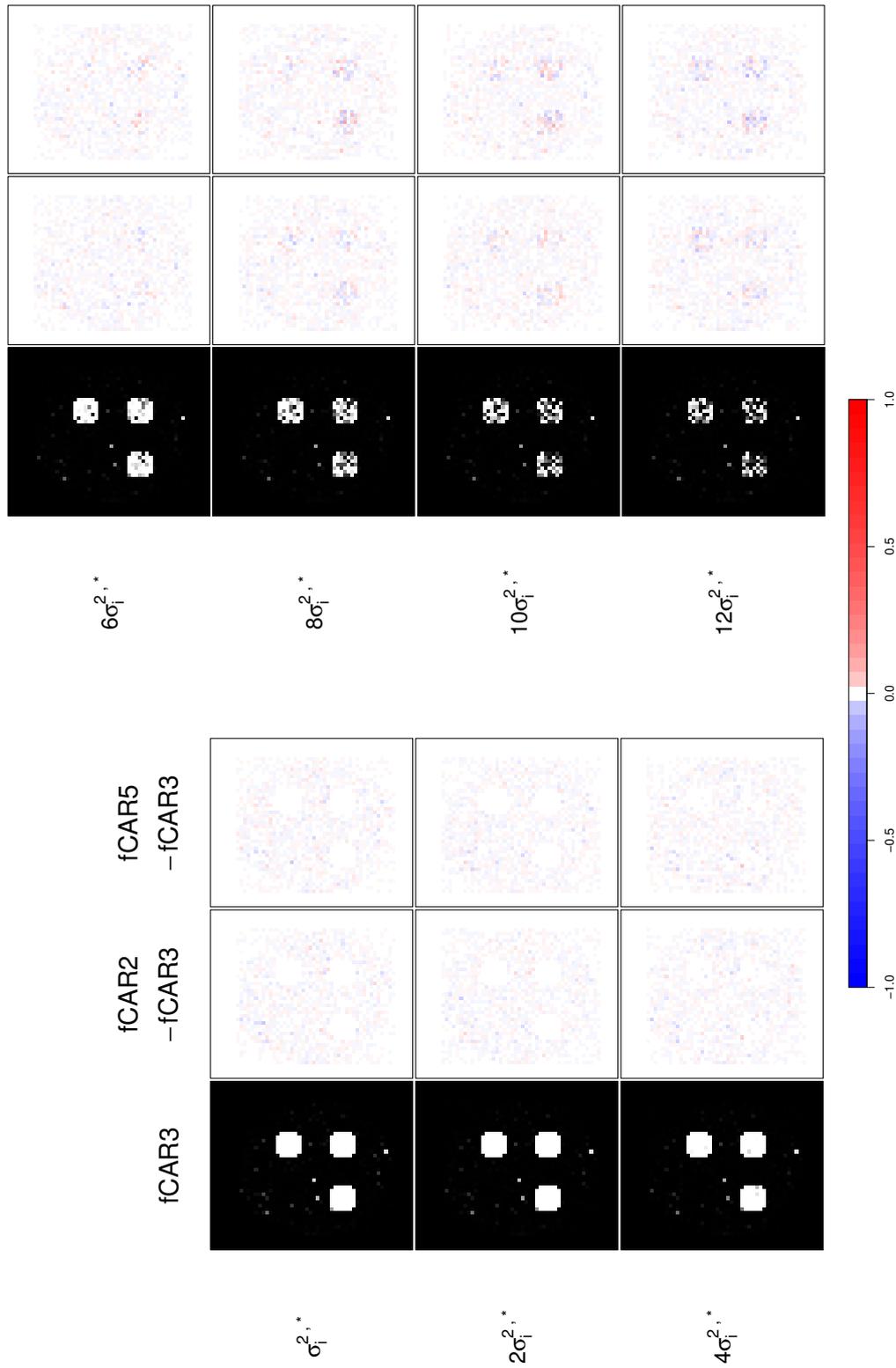


Figure 6.23: Activation probability maps of fCAR2 and fCAR5 models in comparison to the activation probability map of fCAR3 for different error noise. In the leftmost plot, $\hat{\pi}_{fCAR3}$ is plotted. The following plots depict the differences in π -estimation of the fCAR2 and fCAR5 models to $\hat{\pi}_{fCAR3}$. Higher probability values of the fCAR3 model are marked in shades of blue, lower probability values are marked in shades of red. The maps contain the results of the central brain layer, i.e. layer 3.

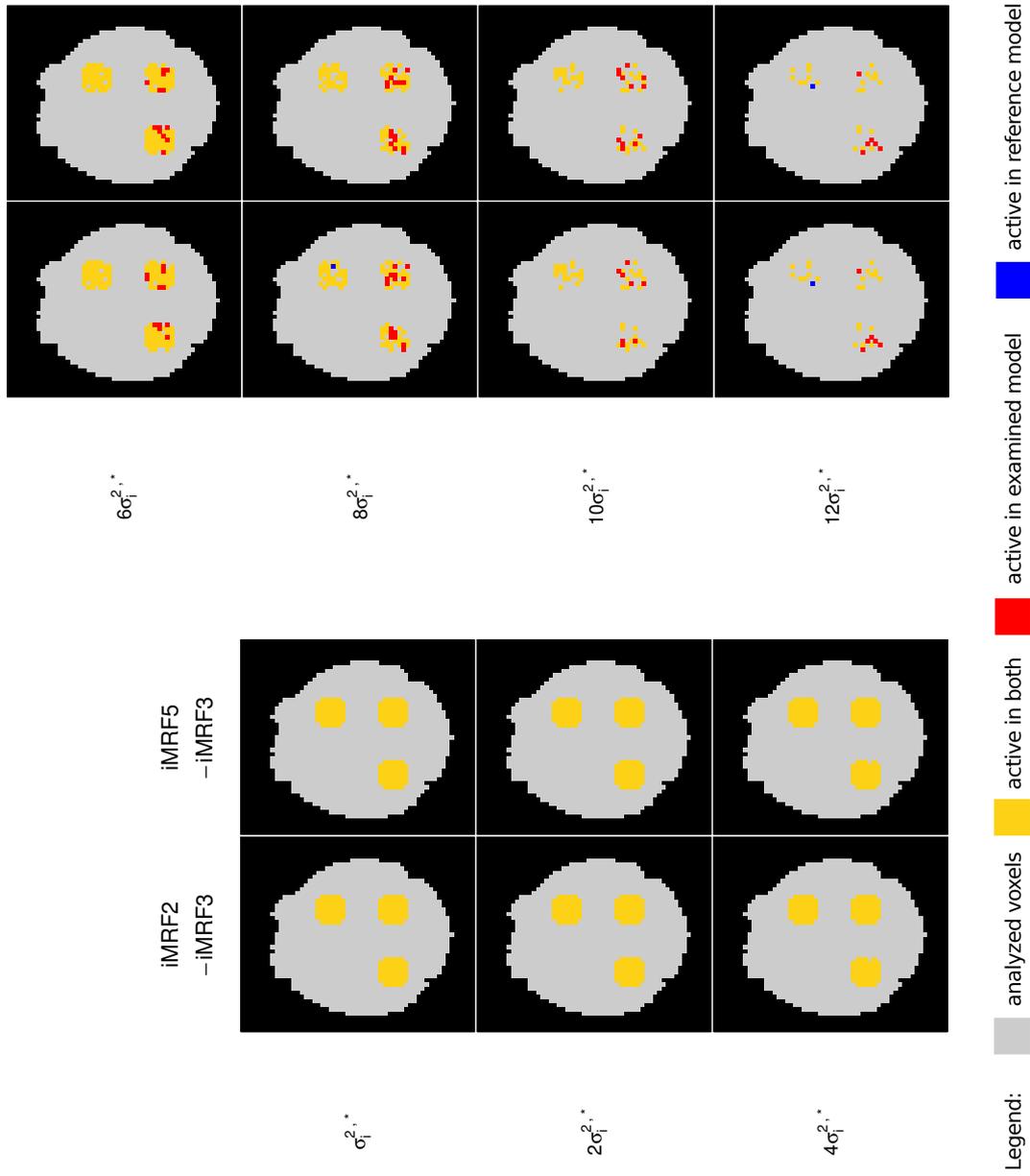


Figure 6.24: Activation difference maps of iMRF2 and iMRF5 models in comparison to the activation map of iMRF3 for different error noise. The maps contain the results of the central brain layer, i.e. layer 3.

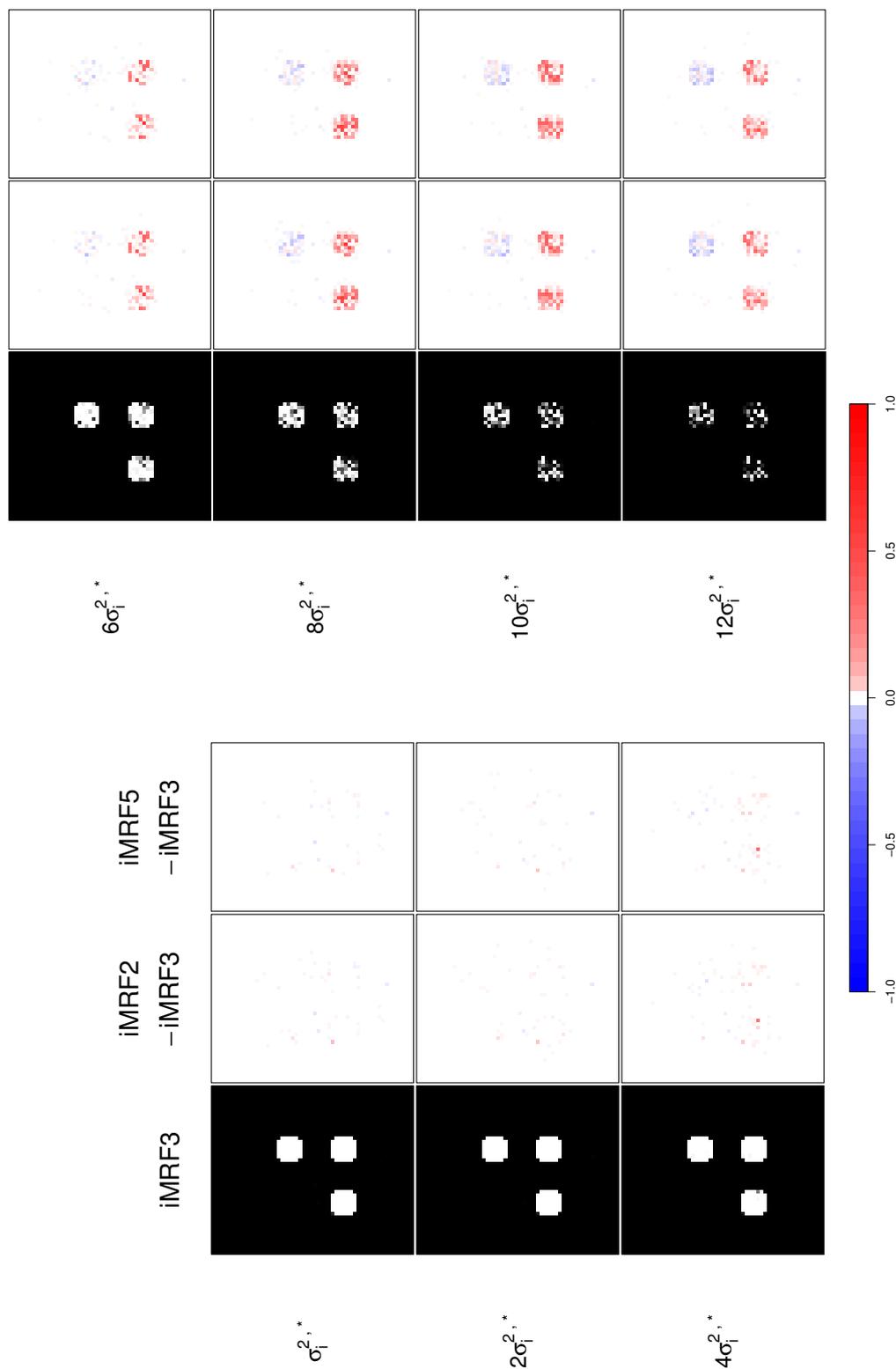


Figure 6.25: Activation probability maps of iMRF2 and iMRF5 models in comparison to the activation probability map of iMRF3 for different error noise. In the leftmost plot, $\hat{\pi}_{iMRF3}$ is plotted. The following plots depict the differences in π -estimation of the iMRF2 and iMRF5 models to $\hat{\pi}_{iMRF3}$. Higher probability values of the iMRF3 model are marked in shades of blue, lower probability values are marked in shades of red. The maps contain the results of the central brain layer, i.e. layer 3.

R_4 —maybe because of the rather good edge-preserving properties of the CAR prior, which may be able to detect also edges within activation regions produced by high noise. Hence, the CAR smoothness prior seems not to be able to increase the number of true positives substantially.

The fCAR2 and fCAR5 models do not yield a gain in sensitivity by using additional EEG information. As can be seen from Figure 6.22 and Table 6.6, sensitivity values of all three predictor types are about the same for all noise levels. In contrast to the gCAR models, not even activation probability estimates are increased in regions with congruent fMRI and EEG activation (cf. Figure 6.23). Nevertheless, all predictor types are able to yield surprisingly large sensitivity values at high noise levels. In these cases, fCAR outperforms gCAR and iMRF by far.

The iMRF models perform similar to gCAR models at low ($\sigma_i^{2,*}$, $2\sigma_i^{2,*}$, $4\sigma_i^{2,*}$) and very high ($12\sigma_i^{2,*}$) noise level. At medium noise levels ($6\sigma_i^{2,*}$, $8\sigma_i^{2,*}$, $10\sigma_i^{2,*}$), an impressive gain in performance can be observed. In these cases, iMRF2 and iMRF5 have superior performance compared to all examined models. From Table 6.6, we see that an increase of about 8% in sensitivity can be achieved by using iMRF5. We note that iMRF2 tends to perform slightly worse than iMRF5. In spite of this, the fCAR models outperform the iMRF models at noise levels $10\sigma_i^{2,*}$ and $12\sigma_i^{2,*}$. Like for gCAR, we see in Figure 6.25, that the EEG-enhanced activation scheme works: Activation probability can definitely be increased in regions with congruent EEG and fMRI activation.

In both gCAR and iMRF models, we see that activation probability estimates in R_2 are highly variable at high noise levels. This is due to the higher noise in the fMRI data in this region. Hence, both model types seem to be less stable to recover activation probability estimates in these cases. This might also be the case for all runs of the fCAR model, because its probability difference maps are also very scattered.

6.1.6 Summary

We pursued different goals with conducting simulation studies based on artificial data. We were able to evaluate the basic properties of our proposed algorithms in a strictly controlled setting.

First, we determined which form of predictor in the EEG probit stage is most suitable for EEG-enhanced activation detection (Section 6.1.3). We recommend to use models with

predictor type 5, i.e. the predictor with both spatially-varying intercept and spatially-varying EEG effect. It is the most flexible predictor type recovering regional-specific EEG influence on the one hand and varying fMRI influence (captured by the intercept) on the other hand. If necessary, it downweights EEG contribution in regions with non-congruent fMRI activation.

Alternatively, we recommend to use predictor type 2 models with a non-negatively restricted global EEG effect—where the EEG effect estimate has the same value overall the brain and does not adapt to regional distinctions. It possesses the ability to increase prior activation probabilities proportional to the observed EEG measurements as soon as a reasonable large, positive EEG effect could be retrieved. If necessary, the spatially-varying intercept can have a region-specific compensating effect to improve the goodness-of-fit.

In Section 6.1.5, we evaluated whether the use of EEG information can increase performance of activation detection algorithms. Depending on the algorithm, we obtained the following results: (1) The sensitivity of the gCAR model can only slightly be increased by adding EEG information. In either case (whether EEG is used or not), gCAR performance is below the other two models. (2) Performance of the fCAR model is extremely high in high noise settings and stands out from the performance of remaining models. Though, incorporating EEG information in fCAR has no effect. (3) In contrast to this, the iMRF models increases performance substantially in medium noise settings. That is, iMRF5 yields an performance increase of up to 8.8% and iMRF2 still achieves a performance increase of up to 7.3%. In these cases, iMRF2 and iMRF5 have superior performance compared to all examined models. Hence, the EEG-enhanced activation scheme works with the iMRF model. That is, in certain settings, activation probability can substantially be increased in regions with congruent EEG and fMRI activation leading to increased sensitivity of the iMRF2 and iMRF5 algorithms.

As a conclusion from this simulation study, we recommend to use the fCAR3 model in high noise settings and the iMRF2/iMRF5 model (with supporting EEG information) in medium noise settings. This recommendation is supported by results from Section 6.1.4, where our proposed models outperformed both the Bayesian activation scheme based on the Ising prior and classical SPM.

Being based on an artificial data setup, performance of our models needs to be evaluated with real-world data. Various factors—which cannot easily be reproduced in the used artificial datasets—could affect performance of our models. On the one hand, we are interested in the effect of less smooth activation structures and the effect of white noise

level in real-world datasets on the performance of our models. We assume that lacking smoothness and high noise can lead to a superior performance of fCAR (i.e. fCAR3). However, it remains to be tested whether fCAR yields high sensitivity values at the expense of specificity. On the other hand, we are interested in the comparative performance of the Ising model with a meaningful derived external field. In Section 6.2, we provide some answer to those questions.

6.2 Simulation studies based on real-world data

In this section, we describe the design and results from a simulation study with the same structure as the application described in Chapter 7. fMRI regression stage and hyperparameter settings are chosen as described there.

For all simulation runs, we take the output from an initial run of the gCAR3 model described in Chapter 7 and label the estimated activation map as γ^* . We calculate the posterior expected values for the fMRI model parameters conditional on this activation map, that is $\beta_i^* = E(\beta_i | \gamma = \gamma^*, \mathbf{y}) = (\mathbf{X}^{*\prime} \mathbf{X}^*)^{-1} \mathbf{X}^{*\prime} \mathbf{y}_i$ and $\sigma_i^{2,*} = E(\sigma_i^2 | \gamma = \gamma^*, \mathbf{y}) = S_i(\gamma_i^*) / (T - 2)$ where $S_i(\gamma_i^*) = \mathbf{y}_i' \mathbf{y}_i - \mathbf{y}_i' \mathbf{X}^* (\mathbf{X}^{*\prime} \mathbf{X}^*)^{-1} \mathbf{X}^{*\prime} \mathbf{y}_i$. The matrix $\mathbf{X}^* = \mathbf{X}(\gamma_i^*)$ is thereby the full design matrix if $\gamma_i^* = 1$ and the design matrix without the stimulus regressor components if $\gamma_i^* = 0$. To obtain a new, simulated fMRI dataset, $\beta_i^*, i = 1, \dots, N$, is plugged into the regression model. Then, voxelwise error vectors generated from $\mathcal{N}(\mathbf{0}, \sigma_i^{2,*} \mathbf{I})$ are added. In contrast to the artificial simulation study described before in Section 6.1, each voxel i contains the recorded fMRI time series \mathbf{y}_i and, thus, the data structure of the original dataset is preserved. To investigate effects of deviations in noise levels, we consider noise levels $0.5\sigma_i^{2,*}, \sigma_i^{2,*}, 2\sigma_i^{2,*}, 4\sigma_i^{2,*}$.

Starting configurations and hyperparameter settings of algorithms are chosen as in Chapter 7. For the evaluation of EEG-enhanced activation schemes, we select a redundant prior information map. As described in Section 7.2, this map is an fMRI activation map estimate from an uninformed activation detection run with extended activation region boundaries. It contains mostly congruent activation information. Hence, potentially adverse results are not due to non-matching EEG information. For each algorithm and noise level, we conduct 50 simulation runs. For each run, we calculate sensitivity and specificity estimates. Then, we report the corresponding minimum, median and maximum over the 50 simulated datasets. Note that we calculate sensitivity and specificity estimates in relation to voxels analyzed. When the fCAR model is examined, estimates can either be calculated

based on the used cubical mask or the threshold mask of the reference model gCAR3. Sensitivity values, however, do not change per definition and changes in specificity values are negligible. Hence, we report the performance values of all models in relation to the used threshold mask (cf. Section 7.1). Summaries of sensitivity and specificity estimates are given in Table 6.8a and Table 6.8b, respectively.

All models obtain (almost) perfect specificity. In contrast to our speculation, we notice that the drop in specificity for the fCAR models (fCAR3, fCAR2, fCAR5) is negligible. Hence, in the following, we evaluate all algorithms in terms of their sensitivity.

We first look at the simulation results from model runs without EEG, i.e. gCAR3, fCAR3, iMRF3 and Ising model results. In contrast to the artificial simulation study, we can now calculate an informative external field map for the Ising model based on anatomical prior information (cf. Chapter 7), enabling a valid comparison between algorithms. As before we set its spatial dependency parameter θ to a medium level of 0.45. Looking at the sensitivity estimates, we see that the proposed models outperform the Ising model in any case. The iMRF3 and gCAR3 models are similarly sensitive in all model runs, whereas the fCAR3 model has increased sensitivity in runs with noise levels $2\sigma_i^{2,*}$ and $4\sigma_i^{2,*}$. At low ($0.5\sigma_i^{2,*}$) and medium ($\sigma_i^{2,*}$) noise level sensitivity of our algorithms is high—demonstrating the reliability of our approaches.

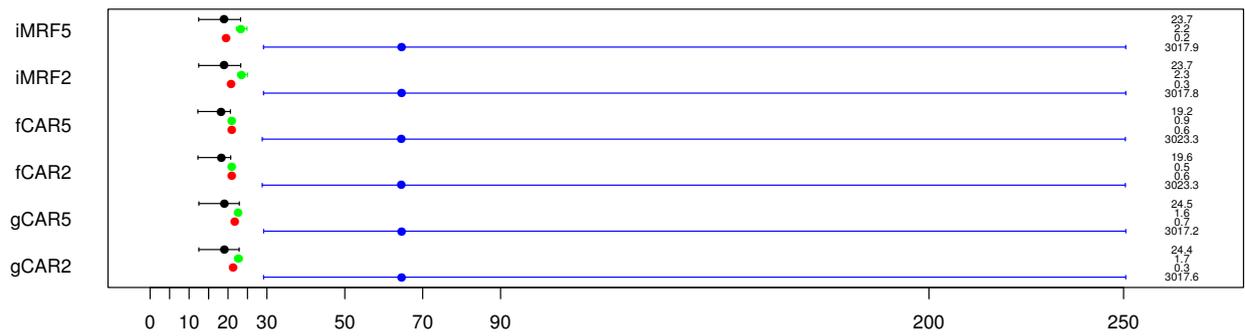
Comparing the results from the EEG-enhanced activation schemes to the corresponding results from schemes without EEG, we see no substantial gain when using EEG information. The algorithms with global updates (gCAR and iMRF) obtain slightly higher sensitivity estimates in higher noise level settings $2\sigma_i^{2,*}$ and $4\sigma_i^{2,*}$ with both predictor types. But no gain in sensitivity can be observed with fCAR2 and fCAR5. Nevertheless, the fCAR algorithm retains its high performance level in any case and outperforms all gCAR and iMRF versions.

To find an explanation why EEG enhancement yields no increased performance, we try to find a characterization of voxels who switch their state from non-active to active (or vice versa). In the following, we look at log(LR) statistics for activation, which can be derived for each voxel as $T \log(S_{i0}/S_{i1})$ where S_{i1} and S_{i0} are the residual sum of squares of the fMRI regression with and without stimulus component as in Section 3.4.1. To characterize different voxel types, we split up true active voxels into four groups depending on their estimated activation state in analyses with and without EEG-prior: voxels staying non-active, voxels changing from non-active to active, voxels changing from active to non-active and voxels staying active. For each simulation run, we calculate 2.5%,

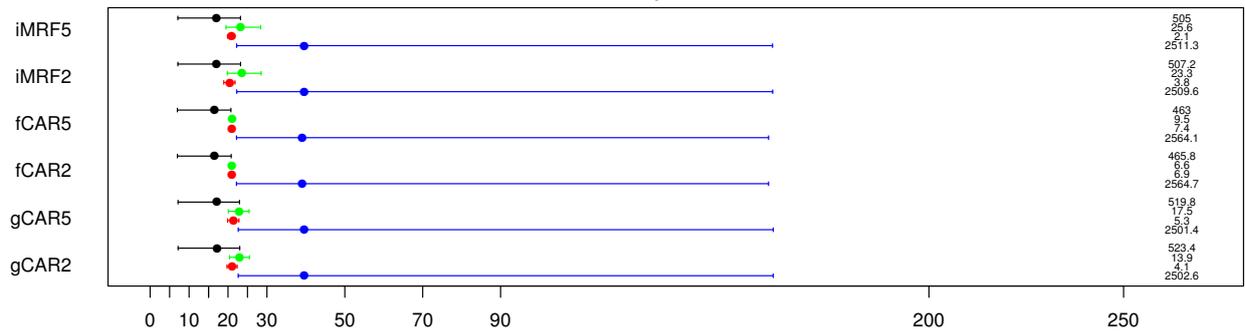
50% and 97.5% quantiles of log-likelihood values within subsets. In Figure 6.26, we depict corresponding average (interval) values over all 50 simulation runs. We notice a good separation between voxels staying active and non-active, respectively, for all algorithms and simulation settings indicating that information contained in the fMRI regression dominates activation classification. A border between these two groups can be drawn at a $\log(\text{LR})$ statistic value of about 20. Voxels changing their activation state are located around this border. Thereby, voxels becoming active generally have higher values than those becoming non-active. The higher the noise in the dataset, the more voxels change their state. In low noise settings ($0.5\sigma_i^{2,*}$), nearly no state changes happen.

As a conclusion, we suppose that mainly voxels having $\log(\text{LR})$ statistic values of about 20 can change their state from non-active to active if suitable EEG-prior information is at hand—in combination with supporting information from neighboring voxels.

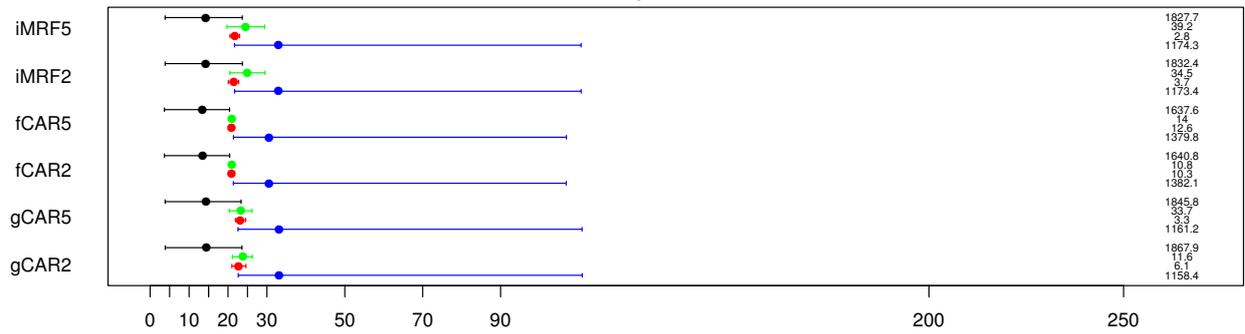
In the next chapter, the application this study is based on is described in detail. We inspect results from single runs of selected fMRI and EEG data combinations and examine estimates of parameter maps beyond the level of binary activation classifications. This should provide deeper insight into the functionality of proposed algorithms in real data settings.



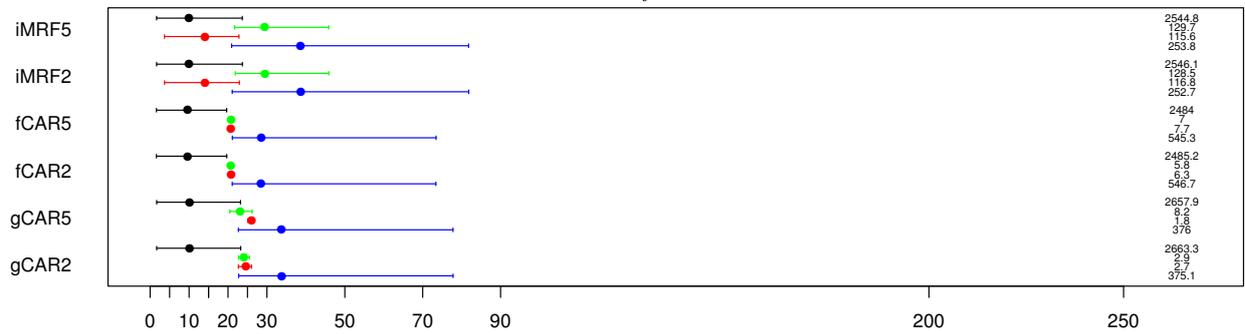
(a) $0.5\sigma_i^{2,*}$



(b) $\sigma_i^{2,*}$



(c) $2\sigma_i^{2,*}$



(d) $4\sigma_i^{2,*}$

Subsets: █ non-active █ non-active to active █ active to non-active █ active

Figure 6.26: Mean of log(LR) statistic intervals (based on 2.5% and 97.5% quantiles) of voxel subsets. Annotation on the right indicates mean subset size.

		$0.5\sigma_i^{2,*}$	$\sigma_i^{2,*}$	$2\sigma_i^{2,*}$	$4\sigma_i^{2,*}$
gCAR3	Median	99.18	82.41	38.22	12.37
	Min	98.75	80.81	36.47	11.20
	Max	99.44	83.31	40.14	13.86
gCAR2	Median	99.21	82.62	38.45	12.39
	Min	98.85	80.91	36.60	11.30
	Max	99.47	83.64	40.24	13.76
gCAR5	Median	99.20	82.65	39.31	12.61
	Min	98.78	81.18	37.39	11.37
	Max	99.47	83.77	41.03	13.99
fCAR3	Median	99.31	84.41	45.76	18.15
	Min	99.05	83.18	44.15	17.05
	Max	99.67	85.97	47.31	19.28
fCAR2	Median	99.33	84.46	45.73	18.18
	Min	99.08	83.21	44.42	17.08
	Max	99.70	85.87	47.24	19.25
fCAR5	Median	99.34	84.54	45.81	18.13
	Min	99.08	83.15	44.38	17.12
	Max	99.70	85.81	47.21	19.22
iMRF3	Median	99.18	82.59	38.70	12.12
	Min	98.78	80.95	36.56	10.81
	Max	99.44	83.74	40.80	13.27
iMRF2	Median	99.23	83.21	39.72	12.52
	Min	98.85	81.41	37.71	11.24
	Max	99.54	84.23	41.85	13.90
iMRF5	Median	99.21	83.33	39.91	12.63
	Min	98.85	81.54	37.81	11.24
	Max	99.54	84.40	42.05	13.86
Ising	Median	63.12	46.55	18.15	5.72
	Min	62.68	45.50	17.08	5.16
	Max	63.67	48.19	19.35	6.24

	$0.5\sigma_i^{2,*}$	$\sigma_i^{2,*}$	$2\sigma_i^{2,*}$	$4\sigma_i^{2,*}$	
gCAR3	Median	100.00	100.00	100.00	100.00
	Min	99.99	99.99	99.99	99.99
	Max	100.00	100.00	100.00	100.00
gCAR2	Median	100.00	100.00	100.00	100.00
	Min	99.99	99.99	99.99	99.99
	Max	100.00	100.00	100.00	100.00
gCAR5	Median	100.00	100.00	100.00	100.00
	Min	99.99	99.99	99.99	99.99
	Max	100.00	100.00	100.00	100.00
fCAR3	Median	99.98	99.98	99.98	99.98
	Min	99.95	99.95	99.95	99.95
	Max	100.00	100.00	100.00	100.00
fCAR2	Median	99.98	99.98	99.98	99.98
	Min	99.94	99.95	99.94	99.95
	Max	100.00	100.00	99.99	100.00
fCAR5	Median	99.98	99.98	99.98	99.98
	Min	99.95	99.95	99.95	99.95
	Max	99.99	100.00	100.00	100.00
iMRF3	Median	100.00	100.00	100.00	100.00
	Min	100.00	99.99	100.00	100.00
	Max	100.00	100.00	100.00	100.00
iMRF2	Median	100.00	100.00	100.00	100.00
	Min	100.00	100.00	100.00	100.00
	Max	100.00	100.00	100.00	100.00
iMRF5	Median	100.00	100.00	100.00	100.00
	Min	100.00	100.00	100.00	100.00
	Max	100.00	100.00	100.00	100.00
Ising	Median	100.00	100.00	100.00	100.00
	Min	100.00	100.00	100.00	100.00
	Max	100.00	100.00	100.00	100.00

(a) Sensitivity

(b) Specificity

Table 6.8: Summary of sensitivity and specificity estimates for real data simulations with four noise levels.

7 Application to an oddball study

Proposed algorithms—both mere fMRI and combined approaches—performed well in artificial simulation studies in Chapter 6. On the one hand, mere fMRI activation detection schemes outperformed other approaches found in the literature. On the other hand, specific choices of EEG-enhanced activation schemes had proven to be useful in various settings. In this chapter, we apply proposed algorithms to real-world data to reveal whether they prove to be useful in practice. The dataset used for testing is based on simultaneously recorded fMRI and EEG time series measurements from an acoustic oddball study. For the assessment of EEG-enhanced activation schemes, we consider different forms of prior EEG information and fMRI data (the latter differing in their level of smoothness as controlled by a preprocessing option). Varying the data scenario aims at gaining deeper insight into the functioning of combined algorithms and at identifying settings where EEG enhancement is able to improve activation detection in realistic settings. In Section 7.1 and 7.2, we describe material and methods used (including preprocessing procedures). For a condensed, but thorough description, we revert to technical terms found in the neuroscience literature. For explanation of these, the interested reader is referred to introductory texts like Zschocke and Hansen (2012) and Huettel et al. (2004). In Section 7.3, we describe the results from both our combined fMRI-EEG algorithm and algorithms based solely on fMRI information.

7.1 Material

Ethics Statement

The study protocol, which was designed in the context of Max Planck Institute of Psychiatry (MPIP) based fMRI/EEG sleep research activities, followed the guidelines of the Declaration of Helsinki and was approved by the local ethical committee (Bayrische Landesärztekammer, Germany, Nr. 01102).

Subjects

Young healthy subjects were recruited by public advertising and gave their written informed consent prior to the study. All participants underwent a careful screening including a clinical interview with a particular focus on previous or current neurological or psychiatric conditions. They further received an anatomical MRI scan to become acquainted to the MR environment and to exclude gross abnormality of normal variants that prevents standard analysis of functional images. Further exclusion criteria were any chronic medical condition, any regular psychotropic medication, consumption of more than 2 cups of coffee per day or more than 5 alcoholic drinks per week, any regular medication, crossing of time zones up to 3 months before the study and any contraindications to MRI. All subjects were right-handed and non-smokers. Eventually, a total of 10 male volunteers (mean \pm SD: 26.0 \pm 2.7 years) underwent the study protocol.

Acoustic oddball paradigm

For acoustic stimulation, we used an active two tone oddball paradigm programmed using the Presentations Software (Neurobehavioral Systems, Albany, USA). The paradigm has been chosen for the following reason (Kiehl et al., 2005). Hemodynamic imaging studies have shown that processing of low probability task-relevant target stimuli (i.e. oddballs) elicit widespread activity in diverse, spatially distributed cortical and subcortical systems of the brain. The nature of this distributed response supports the model that processing of salient and novel stimuli engages many brain regions regardless of whether said regions were necessary for task performance. Due to its remarkable robustness and due to established knowledge on the brain's response to the acoustic oddball stimulation both at the level of the BOLD signal (Kiehl et al., 2005) and the neurophysiological level (evoked potentials) (Volpe et al., 2007), the paradigm is particularly interesting for combining both techniques.

In our paradigm, rare (high-pitched) odd tones (1 500 Hz, duration 50 ms) appeared with 10% probability against the background of frequent (low-pitched) tones (1 000 Hz, duration 50 ms). The interstimulus interval (ISI) was set to an average of 1 000 ms as this period allows to average electrophysiological responses with minimum overlap between the single response. A random jitter was used to prevent habituation that occurs more easily with a strictly regular ISI. While an average of 10% rare tones was defined, the exact order of the tones was randomly assigned, with two additional rules: (1) Rare

tones always had to be separated by at least two frequent tones; (2) The 10% rare tone probability had to be fulfilled in subsets of 20 tones. Acoustic stimuli were delivered by a magnetostatic headphone (MR-Confon, Magdeburg, Germany). In addition, subjects wore foam ear plugs for safety reasons. To adjust loudness of the acoustic stimuli, a preparation scan was performed during which subjects had to repeatedly decide whether or not they perceived the tones as loud as the fMRI scanner sound. This resulted in a defined level of subjectively identical loudness of tones and scanner noise. For the eventual experiment, tones were delivered 3 dB louder than the individualized threshold to guarantee perception. The main acoustic frequency component of the fMRI sequence was about 800 Hz, outside the acoustic spectrum of frequent and rare tones. After the loudness adjustment, subjects were instructed to continuously pay attention to the tones and press the response button immediately after recognizing an odd (high pitch) tone. The paradigm was briefly explained to the subjects prior to the scan.

fMRI acquisition

After appropriate EEG montage (see below), subjects were positioned in the scanner (3 Tesla system, General Electric MR750) with their head being immobilized as far as possible to prevent motion artifacts, and wearing ear protection. Whole brain fMRI time series were acquired using echoplanar imaging (EPI) sequence (time of repetition [TR] 2 000 ms, time of echo [TE] 40 ms, slice orientation according to anterior-commissure/posterior-commissure landmarks, 28 slices, slice thickness 3.5 mm, 0.5 mm gap, in-plane resolution $3.125 \times 3.125\text{mm}^2$) while the acoustic oddball paradigm was applied. Per subject, a total of 307 image volumes was recorded over 10 minutes. Recordings of 8 subjects were suitable for further combined analysis processing.

EEG acquisition

EEG recording was parallel and synchronized to fMRI measurements. An Easy Cap (<http://www.easycap.de/easycap>) with 64 EEG electrodes was used, including an electrocardiogram electrode placed on the left side of participants back close to their spinal cord, referenced against the fronto-central electrode FCz. Data were continuously sampled throughout the experiment at 5 kHz. To allow for optimal artifact correction, EEG recordings and fMRI were synchronized using the scanner's 10 MHz master clock (Mandelkow et al., 2006). Electrode impedance was below 5 k Ω . Additionally, trigger pulses from the MRI system were recorded for subsequent off-line MRI artifact correction.

7.2 Methods

fMRI data preprocessing

All fMRI preprocessing steps were performed on 64-bit-Linux workstations using the SPM software (<http://www.fil.ion.ucl.ac.uk/spm>, version SPM8). First, data were corrected for slice time differences to compensate for different acquisition times due to the interleaved slice acquisition scheme in each volume. Second, images were motion corrected by rigid-body coregistration with the mean image of the uncorrected time series. Third, images were spatially normalized using linear and non-linear transformations to an EPI wholehead template in standard MNI space¹ with default settings of the SPM8 distribution. Intrinsic to the spatial normalization step is an interpolation step that was set to gain voxels sized $4 \times 4 \times 4 \text{ mm}^3$ which corresponded to image volumes of each $40 \times 48 \times 34$ voxels per time point. Last, we calculated two versions of smoothed image trajectories using a symmetric 3D-Gaussian kernel. We generated a very smooth dataset with high signal-to-noise ratio by applying a kernel with $8 \times 8 \times 8 \text{ mm}^3$ FWHM (full width half maximum) and a slightly smoothed dataset with rather low signal-to-noise ratio by applying a kernel with $2 \times 2 \times 2 \text{ mm}^3$ FWHM. For functional analysis, the first 5 images were disregarded due to T1-unequilibration effects.

EEG data preprocessing

Calculation of differential event-related potential trajectories on the subject level

EEG data was corrected for gradient induced and cardioballistic artifacts (Czisch et al., 2009) using Vision Analyzer 1.05 (Brain Products). That is, after MRI artifact correction an independent component analysis (ICA) (Beckmann and Smith, 2005) was performed to clean data from cardioballistic artifacts. Afterwards data was bandpass-filtered with typical settings for evoked potential analysis (low cutoff 0.5 Hz, high cutoff 30 Hz). After baseline correction and DC detrending², EEG time series were segmented into 1 000 ms segments according to odd and even tone onset time points (-200 ms to +800 ms), which after exclusion of obviously corrupted data segments resulted in an average of ~ 490 segments for even tones and ~ 45 segments for odd tones for each electrode (time resolution 250 Hz).

¹The acronym 'MNI' originates from the Montreal Neurological Institute, where this type of brain template was developed.

²A correction method for direct current (DC) drift artifacts.

For further use, segments of each stimulus type are averaged to gain average event-related potential (ERP) trajectories. These are assumed to contain an estimate of the amplitude and morphology of the electrophysiological response. For the odd minus even contrast being of special interest in oddball studies, we calculated a differential ERP by taking the pointwise differences of odd and even ERPs.

Calculation of 3D prior EEG information maps

Based on the derived differential ERP, spatial EEG maps were then calculated for each time point using the sLORETA software (<http://www.uzh.ch/keyinst/loreta.htm>). For later use, EEG source maps had to be reduced to a small number of maps that could be incorporated as prior information for fMRI activation one at a time. Either sLORETA maps of selected time points were considered or maps derived by an aggregation over given sLORETA trajectories. For a subject-specific EEG-prior, EEG source maps were aggregated over time by an ICA. Alternatively, for EEG priors based on the whole subject group, EEG source maps were aggregated over both time and subjects by a multi-session tensor based probabilistic ICA (Tensor-PICA) (Beckmann and Smith, 2005). Both the single-subject ICA and the group Tensor-PICA arrive at spatial component maps that cluster voxels with a common time course. We selected those component maps showing an connections to the differential odd minus even contrast. Selected 3D EEG-prior maps are described in detail in the following section .

Spatial prior information

The following four types of prior information maps were used to enhance fMRI activation detection:

- EEG-based prior information map in form of a group and time aggregated sLORETA map: For this, a spatial Tensor-PICA component (from the group analysis of sLORETA trajectories described in the preceding paragraph) was selected that showed a spatial pattern and time course associated with stimulus presentation (see Figure 7.1). This EEG-based map was chosen for generalizability reasons for being based on group data. The component map was regarded to be more reliable than a 3D map based on single-subject data only. From its spatial pattern, it was supposed to add information to fMRI activation detection especially in upper central parts of the brain. Because we do not differentiate between activation and deactivation (with regard to the sign of effects) we take the absolute value of all map values.

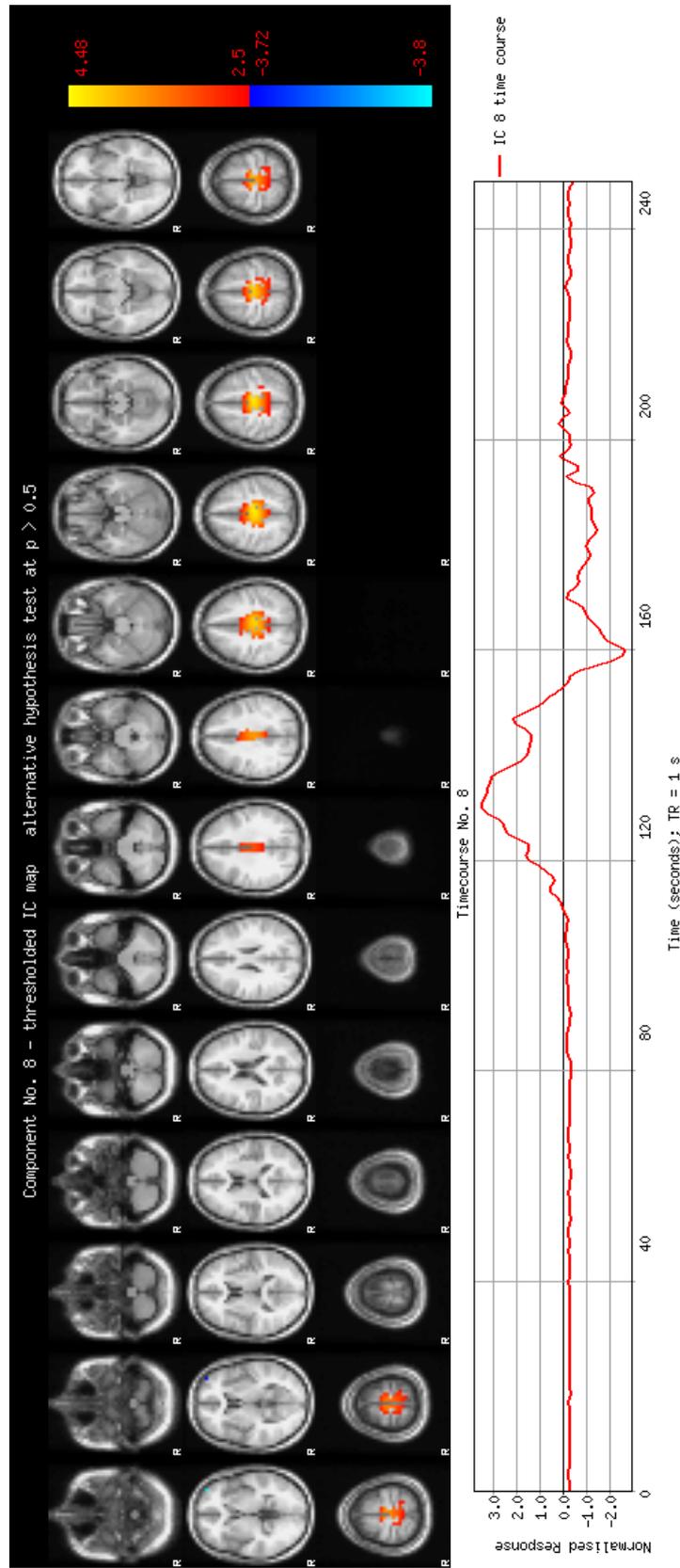


Figure 7.1: Selected result map (spatial component map 8) of a multi-session Tensor-PICA of sLORETA map trajectories of all subjects and its corresponding time course—used as group and time aggregated EEG-based prior information map.

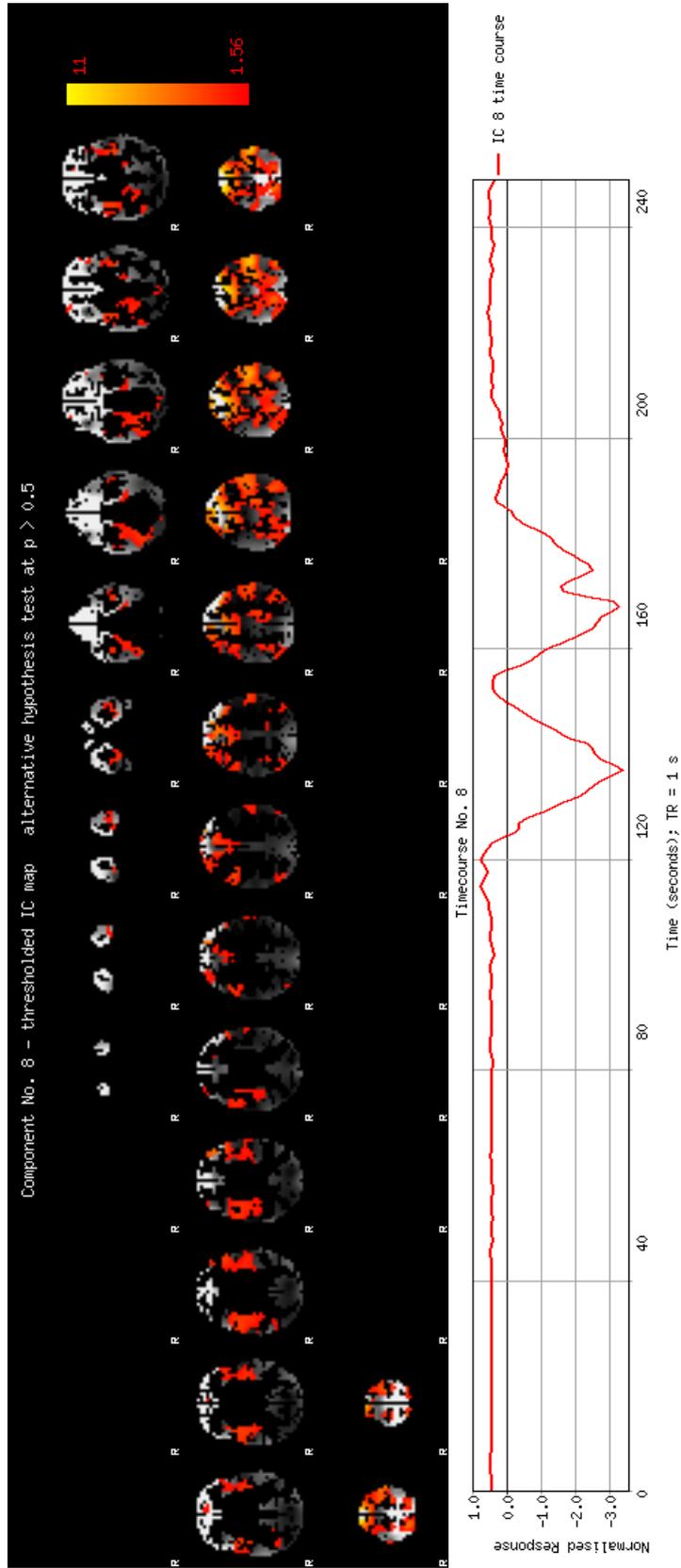


Figure 7.2: Selected result map (spatial component map 8) of a single-session ICA of sLORETA map trajectories of all subjects and its corresponding time course—used as subject-specific, time aggregated EEG-based prior information map

- EEG-based prior information map in form of a subject-specific sLORETA source reconstruction map for a selected time point ($t \approx 480$ ms (PST); sLORETA map no. 172) of the ERP difference profile of odd minus even stimuli: This sLORETA map was selected for showing a strong EEG activation in several brain areas—in particular frontal and in the upper back parts of the brain. In general, this EEG based map contained information closer (i.e. more congruent) to the individual fMRI data than the Tensor-PICA map for being recorded on the same subject. We expected that this prior information map might be easier to handle for proposed algorithms. For numerical stability, we took the logarithm of all map values, which were all positive.
- EEG-based prior information map in form of a subject-specific, time aggregated sLORETA map: For this, an ICA was run on the sLORETA map trajectory from the subject the fMRI data was recorded on. We selected a component map that showed a strong signal in the (neurologically) right part of Brodmann areas 1-4, 6 (containing the primary somatosensory, primary motor and premotor cortex) (see Figure 7.2). This is an area where fMRI activation can also be found. To enlarge the strength of the EEG information, we took the voxelwise square of map values. To adjust for a mismatch between sLORETA analyzing mask and the brain structure found by fMRI, we smoothed the sLORETA map with a Gaussian kernel of $8 \times 8 \times 8$ mm³ FWHM. This should put less emphasize on analyzing mask edges of sLORETA. We considered this EEG-prior to contain congruent information compared to fMRI in an selected part of the brain.
- Redundant prior information map in form of an extended fMRI activation map: This map was based on an activation map from an analysis without EEG component and can, hence, be considered as a redundant prior information map. To increase sensitivity on activation region boundaries the next 28 neighbors (from the surrounding neighborhood cube) were added to the binary activation map. This fMRI based map was taken as a test case for proposed algorithms: The most likely voxels to be additionally found active were supposed to lie on activation region boundaries.

If necessary, the selected prior information maps were resliced to the spatial resolution of the fMRI dataset using the SPM8 coregister procedure. If component maps contained negative values, we took the absolute value of the selected ICA and Tensor-PICA component maps because, on the one hand, strong deviations from zero indicate component membership and, on the other hand, our model predictor relies on the assumption that larger prior information values indicate activation—with the meaning that a specific voxel shows a reaction to a stimulus.

Data analysis

For demonstrating the functioning of proposed single-subject activation detection algorithms, we selected the fMRI session data from one of the recorded subjects (subject ID 8). For (EEG-) enhanced activation detection algorithms, we used the four types of spatial prior information presented in the preceding subsection.

For analysis, we used the proposed algorithms from Chapter 3 and 4. In the following, we use the notation introduced in Chapter 5, p. 90, and denote the proposed algorithms from Chapter 3 with CAR and IGMRF prior as gCAR and iMRF, respectively. In addition, fCAR denotes the fast algorithm for the model with CAR prior proposed in Chapter 4. For all three model types, the exact type of predictor had to be specified: We applied gCAR3, fCAR3 and iMRF3, which rely on predictor type 3 excluding EEG information. For the evaluation of EEG-enhanced fMRI activation detection algorithms, we considered gCAR5, fCAR5 and iMRF5 (predictor type 5 models with an additional spatially-varying EEG component) and gCAR2, fCAR2 and iMRF2 (predictor type 2 models with a global EEG effect restricted to be non-negative).

In case of the CAR prior, the prior mean of the spatial dependency parameter τ^2 (and/or τ_0^2) was set to 0 and its prior variance was chosen to be rather uninformative with value 25. In a testing phase of the gCAR and fCAR algorithms, we have found that convergency of MCMC trajectories is sensitive to both the choice of ξ^2 (ξ_0^2) hyperparameters, i.e. inverse gamma shape and scale parameters of respective priors, and τ^2 (τ_0^2) hyperparameters, i.e. proposal variance ($\sigma_{\tau_0}^2$ resp. σ_{τ}^2) and starting value (τ_0 -start resp. τ -start). Hyperparameters of the CAR algorithms were chosen in such a way that convergency to the equilibrium distribution was acceptable within 1 000 iterations (burnin). Convergency was assessed by visual inspection of trace plots and corresponding diagnostic tools. Suitable hyperparameter values chosen for analysis are listed in Table 7.1.

In case of the IGMRF prior, suitable inverse gamma parameter values had also to be found. Our choices, which lead to an adequately convergency behavior, are listed in Table 7.1 as well.

Note that listed hyperparameters lead to satisfactory results in most runs. However, we had to adapt some of these slightly when incorporating the subject-specific sLORETA map. We do not list these parameters here, but choices can be provided upon request.

The neighborhood underlying both random field priors incorporate the next 6 direct 3D voxel neighbors.

	gCAR	fCAR	iMRF
$\xi_{0,G}^2$ prior:	IG(3, 1)	IG(3, 1)	IG(3, 1)
ξ_G^2 prior:	IG(3, 1)	IG(3, 1)	IG(3, 1)
ξ_0^2 prior:	IG(1802, 16209)	IG(452, 4059)	IG(902, 4054.5)
ξ^2 prior:	IG(452, 4059)	IG(452, 4059)	IG(902, 4054.5)
$\sigma_{\tau_0}^2$	3	0.1	–
σ_{τ}^2	0.25	0.1	–
τ_0 -start	25	0.1	–
τ -start	2	0.1	–

Table 7.1: Hyperparameter settings for gCAR, fCAR and iMRF model runs.

If included in the model (either in gCAR, fCAR or iMRF), global effect parameters had prior mean zero and a variance parameter that followed an inverse gamma distribution with shape parameter 3 and scale parameter 1 (cf. Table 7.1).

For each run, we used 6 000 MCMC iterations including a burnin phase of 1 000 iterations. To remove strong autocorrelations, we thinned out resulting parameter trajectories with a stepping of 5.

We compare our proposed model to the following methods: The Ising model from Smith et al. (2003) and the widely used classical SPM algorithm (Friston et al., 1995) as implemented in SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/software/>).

For all models, we used the canonical HRF with three basis functions to model odd stimulus contribution to the fMRI predictor. Regressors for modeling the even stimulus were not included for introducing collinearity issues. We included the DCT basis set implementing regressors for a highpass filter of 128s and a set of movement parameters and two global signals (containing the average white matter and liquor signal) as confounders. Details about the calculation of these regressors can be found in Chapter 2.

We restricted analysis to all brain voxels lying within a threshold mask defined in analogy to the selection procedure in SPM (cf. Section 4.3). In a second selection step, we removed all small, isolated voxel clusters for these being able to distort analysis results. The fCAR3 model, however, is generally dependent on a cubical mask comprising the used threshold mask. To emphasize this fact, we use a subscript C to distinguish mask choices: fCAR3 $_C$ denotes the (standard) fCAR3 model run with the cubical mask. For better comparability we additionally present fCAR3 $_C$ results on a threshold mask denoted as fCAR3*.

Ising hyperparameters were chosen as follows: Spatial dependency parameter θ was set to a medium level of 0.45, which also had been used in Smith et al. (2003). The external

field map was based on the prior knowledge that about 25% of gray matter is likely to be activated by this experiment. Gray-matter prior information was obtained from SPM segmentation of the first image of the fMRI time series, which provides sufficient anatomical contrast. With this information the external field map could be calculated as in Smith et al. (2003).

7.3 Results

7.3.1 Comparison of different fMRI activation detection algorithms

We first assess the comparative performance of the proposed Bayesian activation detection algorithms (with both IGMRF and CAR priors) without EEG information. Existing methods—so far—provide no means of using additional information in a general form to enhance activation detection. Hence, for fairness we compare only our proposed predictor type 3 models, i.e. gCAR3, fCAR3, iMRF3, with the performance of competing activation detection software.

For being a Bayesian fMRI model with the same activation detection mechanism, the Ising model is directly comparable to the proposed models. Hence, results of Bayesian activation detection schemes comprise the Ising results. To yield an overview of brain wide responses, results from 8 layers between layer 3 and 31 are plotted. Figure 7.3 shows activation difference maps in reference to model gCAR3, and Figure 7.4 shows the underlying activation probability difference maps as well in reference to gCAR3. In Figure 7.5, robustness of activation maps is depicted via threshold dependent activation maps. Robustness of classical SPM results is depicted in significance level dependent activation maps (each controlling the familywise error (FWE) rate) in Figure 7.6. A comparison between activation maps of SPM and gCAR3 is presented in Figure 7.7.

In activation difference maps, we find that all methods identify the most prominent activation regions. Methods differ mostly in the size of found activation clusters. Maps exhibit only minor differences between gCAR3 and iMRF3, though iMRF3 appears to be marginally more sensitive. The Ising model detects much less activation compared to gCAR3: The gCAR3 algorithm possesses the potential to identify larger activation clusters than the Ising model by classifying more voxels on region boundaries as active and to detect some new clusters as well. In activation difference maps of fCAR3_C and

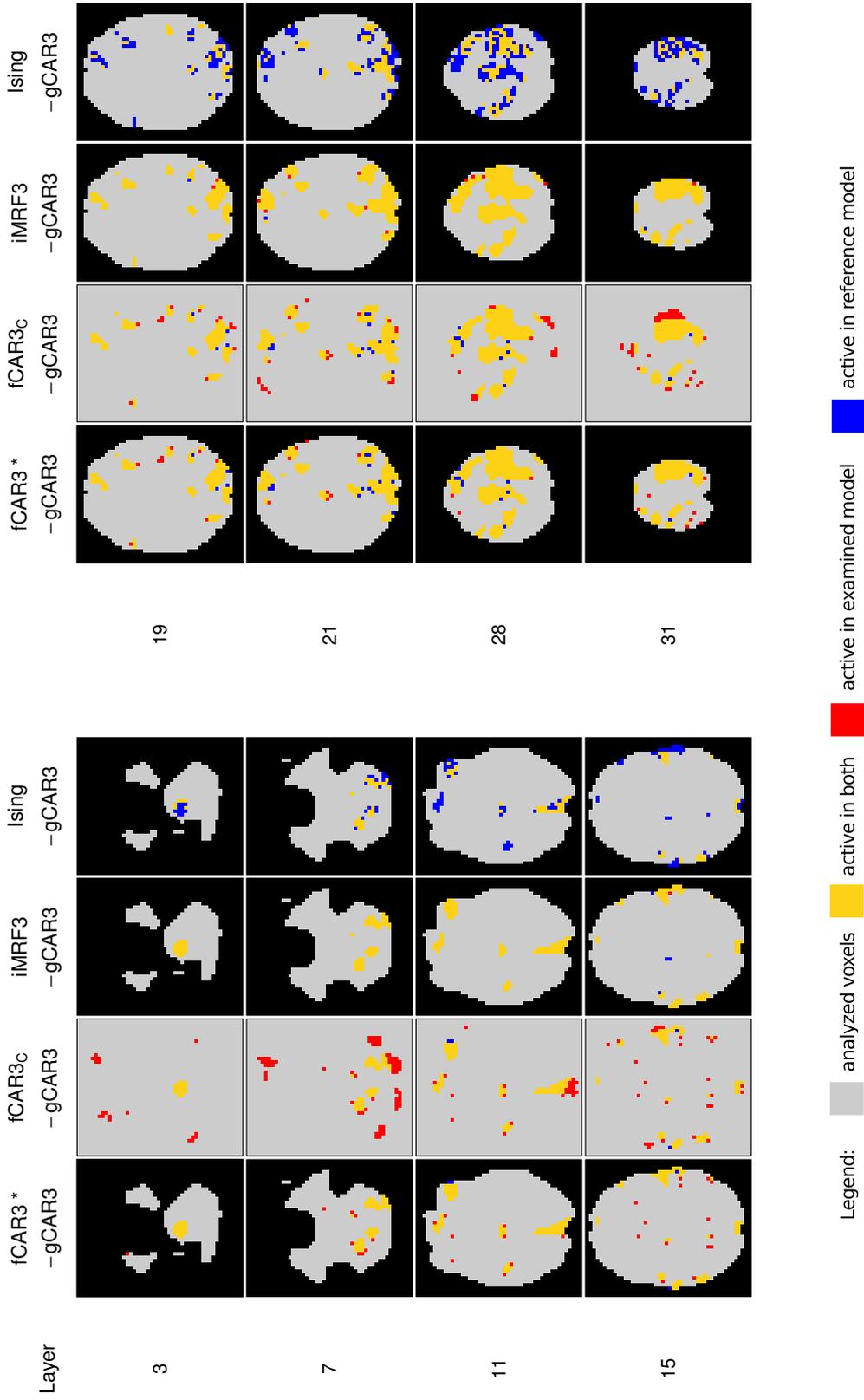


Figure 7.3: Activation difference maps of fCAR3*, fCAR3c, iMRF3 and Ising models in comparison to the activation map of the reference model gCAR3 for selected brain layers.

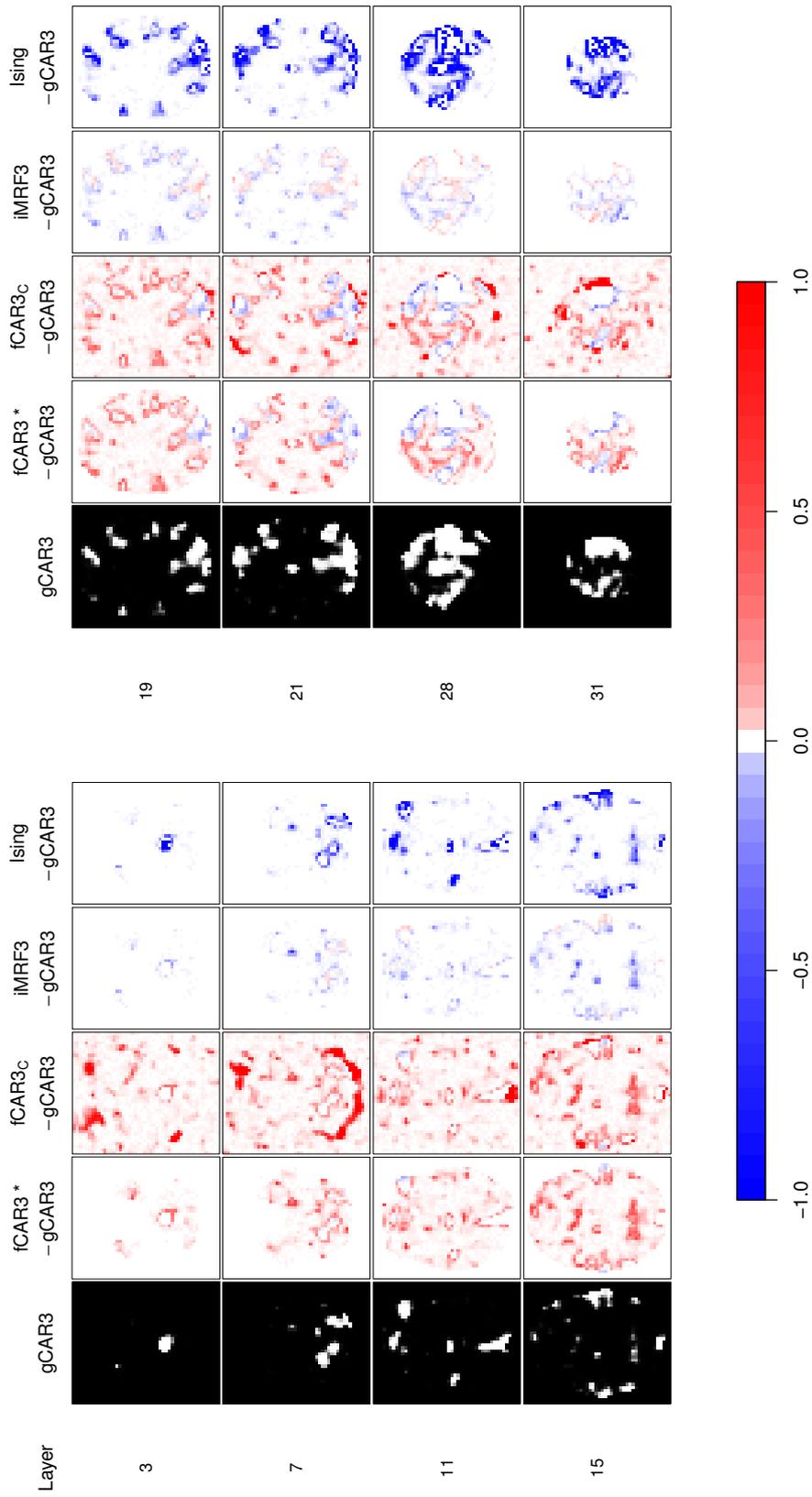


Figure 7.4: Activation probability maps of fCAR3*, fCAR3c, iMRf3 and Ising models in comparison to the activation probability map of gCAR3 for selected brain layers. In the leftmost plot, $\hat{\pi}_{gCAR3}$ is plotted. The following plots depict the differences in π -estimation of the fCAR3*, fCAR3c, iMRf3 and Ising models to $\hat{\pi}_{gCAR3}$. Higher probability values of the gCAR3 model are marked in shades of blue, lower probability values are marked in shades of red.

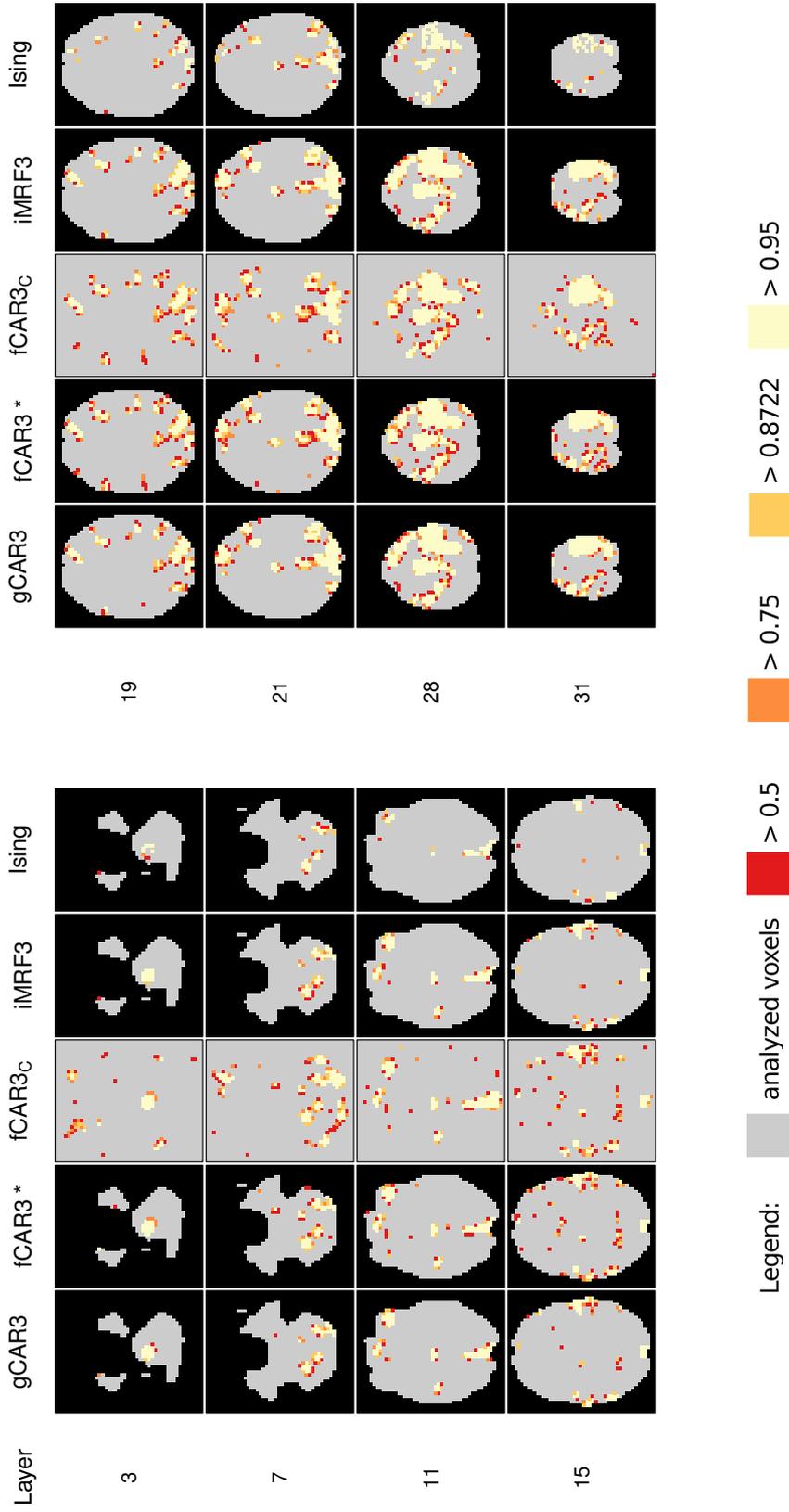


Figure 7.5: Threshold dependent activation maps for the gCAR3, fCAR3*, fCAR3c, iMRF3 and Ising models for selected brain layers. Color codes represent the most extreme threshold the voxelspecific posterior activation estimate exceeds.

	Min.	1% Qu.	25% Qu.	Median	Mean	75% Qu.	99% Qu.	Max
fCAR3*	-0.469	-0.058	0.000	0	0.029	0.016	0.391	0.665
fCAR3 _C	-0.469	-0.019	0.000	0	0.044	0.009	0.991	1.000
iMRF3	-0.654	-0.097	-0.001	0	-0.004	0.000	0.026	0.147
Ising	-1.000	-1.000	-0.007	0	-0.083	0.000	0.001	0.008

Table 7.2: Summary of activation probability difference values of Bayesian activation detection schemes in references to gCAR3.

gCAR3, we notice large activation foci outside the threshold brain mask. These strong differential activations are located in two areas. First, posterior areas in the midline (layer 7) typically represent large venous vessels (e.g. the sagittal sinus) that do not contain neuronal tissue but are strongly influenced by effects of respiration and cardiac pulsation. Correlation with the model in these areas may be explained by some correspondence between these vegetative parameters and the subject's (attentional and motor) response to the odd tone. Second, in anterior areas (layer 3), which correspond to the eye region, correlations were detected that may be explained by paradigm-correlated eye movement or blinking. Both areas tend to be excluded by respective anatomical masks. For better comparability, we additionally look at results of fCAR*, where fCAR3_C results are masked with the threshold mask. In these plots, we notice a number of new, small activation clusters that appear all over the brain. Moreover, changes in large, contiguous activation regions emerge: On the one hand some activation clusters are enlarged by fCAR3, e.g. clusters in layer 19. On the other hand, in layer 31 a group of activated voxels appears on the left posterior area indicating an activation region not found before. However, a local decrease in sensitivity can be observed as well: The larger dependency parameter estimate of the gCAR3 algorithm influences sensitivity positively within activation regions in layers 21 and 28. Increased smoothness leads to more contiguous activation regions in large activation clusters.

A look at the posterior activation probability difference maps allows to gain a deeper insight into the functionality of algorithms. Hardly any differences in probability estimates emerge in centers of most prominent activation centers. The largest differences arise at activation region boundaries hinting at differences in edge preserving resp. smoothing properties or the ability to cope with decreasing signal-to-noise ratio. Differences between iMRF3 and gCAR3 are small and lie mostly in between $[-0.097; 0.026]$. i.e. the 1% and 99% quantiles of the corresponding empirical distribution (cf. Table 7.2). Conform to the results of the binary activation maps, we notice that the Ising model obtains much lower

probability estimates whenever differences to gCAR3 emerge. The fCAR3 model obtains higher posterior activation probabilities in large parts of the brain. Particularly, a number of new clusters with increased posterior probability estimates appear in the fCAR3 difference map leading to isolated small activation foci found in the corresponding activation maps. Apart from this, differences emerge at region boundaries. fCAR3 tends to gain larger activation probabilities on region edges, especially in lower brain layers, speaking for larger cluster sizes. In layers near the top we observe partly the reverse: Within activation regions, probabilities are distinctly decreased compared to smooth gCAR3 results. To check whether changes at boundaries can be traced back to an inefficient separation between head movement and activation effects, we looked at results of a simultaneous F -Test of movement parameter effects, which were included into analysis as confounds (results not shown). The corresponding F -map does not indicate movement effects on region boundaries. Hence, we conclude these differences can be traced back to the fact that the algorithms have different abilities to cope with decreasing signal-to-noise ratios.

The results from the activation threshold maps—plotted for evaluating the robustness of algorithms—go along with the peculiarities from probability difference maps. Predominantly voxels at region boundaries are additionally classified as active with decreasing probability threshold. Thus, the discrimination power of considered Bayesian algorithms between activation and non-activation regions attracts positive attention. Only fCAR3 inclines to find new small clusters.

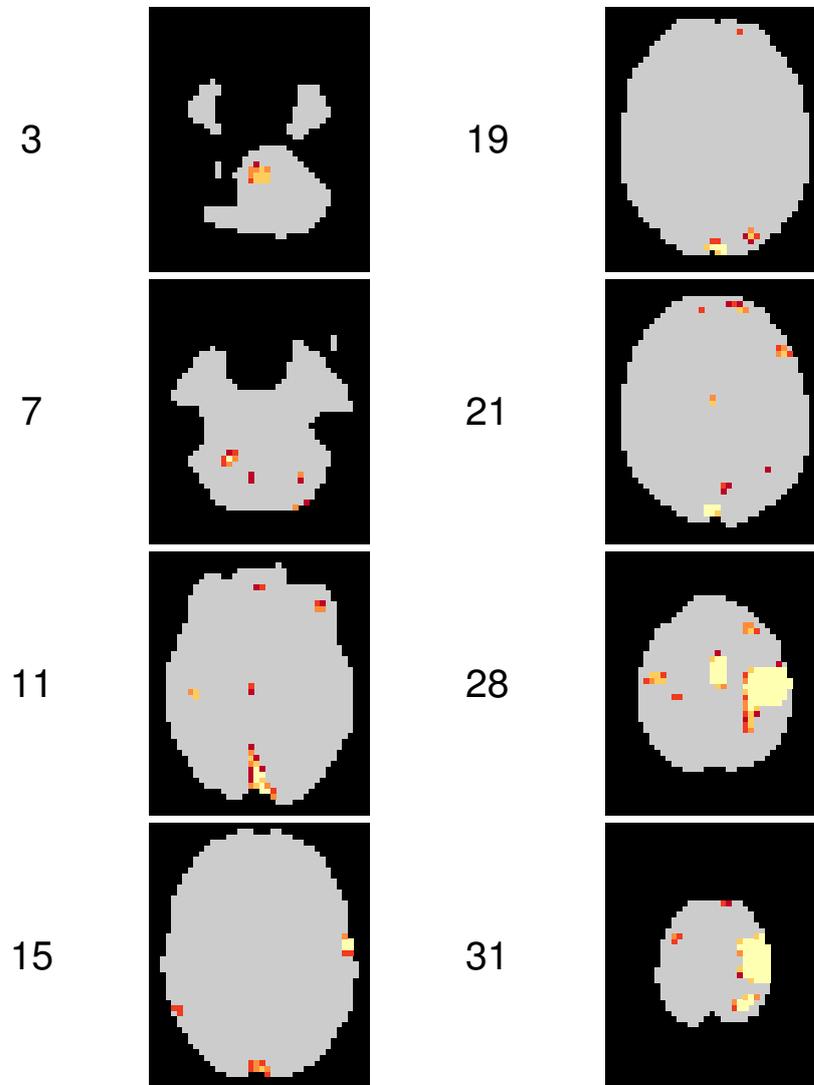
SPM results are not directly comparable to the results of the proposed Bayesian activation schemes (cf. Section 3.4.5 and 6.1.4). Hence, we decided to first evaluate the robustness of resulting SPM activation maps in relation to several choices of significance level—in analogy to probability threshold dependent maps of Bayesian procedures. From significance level dependent activation maps we see that SPM possess a good discrimination power between activated and non-activated sites as well. Varying the significance level leads mostly to marginal changes in cluster sizes. However, there are also several clusters that cannot be detected until a less conservative threshold is applied. Having an idea of robustness of algorithms, we compare next the SPM activation map of the least conservative considered significance level of 0.2 with the gCAR3 activation map obtained by applying the most strict considered probability threshold, i.e. 0.95. As we can see from Figure 7.7, the gCAR3 algorithm yields larger activation regions throughout the brain and appears to be much more sensitive than SPM.

As we can see from these comparisons, the proposed algorithms seem to be much more sensitive than both Ising and SPM algorithms. This might be traced back to a better ability

of proposed models to cope with decreased signal-to-noise ratios. Differences between proposed algorithms emerge most notably at activation region boundaries. Thereby, iMRF3 and gCAR3 yield very similar results: Posterior probability differences are rather small and the estimated activation structures are almost identical. The presumption that iMRF3 and gCAR3 differ substantially in their edge-preserving behavior (cf. Section 6.1.4) can, hence, not be affirmed in this real-world example. Nevertheless, iMRF3 seems slightly more sensitive in upper brain layers. In spite of this, fCAR3 yields higher posterior probability estimates in areas outside main activation centers and hence, inclines to find an increased number of small activation clusters not found by other algorithms. It is not clear whether these small activation areas that are unattached to the larger clusters can be clearly marked as false positives. Specificity of the fCAR3 algorithm, anyhow, has been very high in simulation studies based on an equivalent activation structure (Section 6.2) speaking against the hypothesis of falsely positive clusters. Smoothness of gCAR3 (iMRF3) seems to be advantageous in large activation regions: Compared to fCAR3, sensitivity of gCAR3 (iMRF3) is slightly increased favoring more contiguous activation sites.

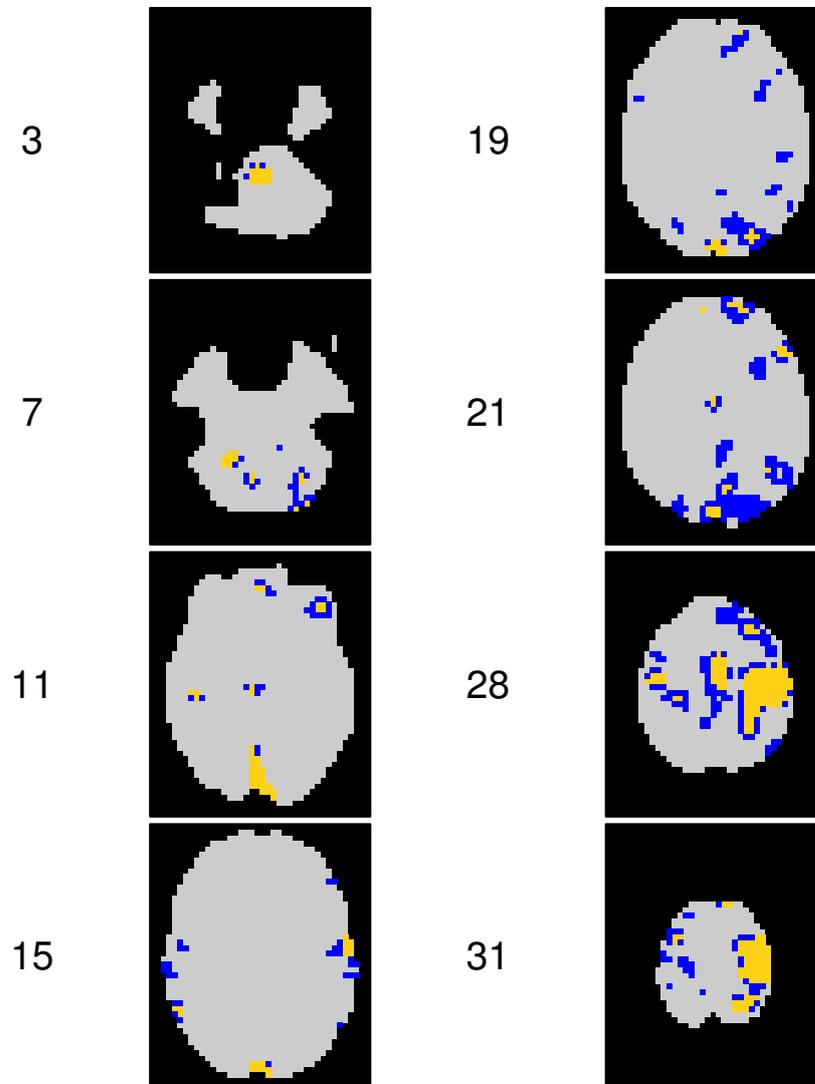
It should be added in this context that fMRI activation detection schemes rely on a very broad definition of activation: A voxel is said to be active if its fMRI signal time series correlates with stimulus presentation. Of course, this definition includes not only neuronal activity but also correlates of it (cf. Section 1.2). Hence, it is open for discussion whether a voxel can be truly marked as false positive if it is not neuronally active in a strict sense, but has an indirect connection to neuronal activation. Anyhow, a careful interpretation of results and possible post-hoc analysis steps are indispensable to identify voxel regions that are neuronally active in a strict sense.

To remove artifact-related activation from analysis results of this study, following aspect can be considered: The masking procedure was liberal in that non-gray matter areas were included. By definition, non-gray matter activation cannot exhibit any neural activation but only correlates of it. Additional masking steps can therefore be applied to remove non-gray matter regions and reveal truly neuronally active voxels. We, however, do not report results on this additional gray-matter mask, because it does not add value to the performance comparison of different algorithms. We, as well, did not use this gray-matter mask as an analysis mask (i.e. a binary mask that indicates voxels to be analyzed) for the following reasons: On the one hand, the gray-matter mask itself is a result of a segmentation routine, which cannot claim to be correct. On the other hand, the analysis mask exhibits many discontinuities making interpretation difficult.



Legend:  analyzed voxels  < 0.2  < 0.1  < 0.05  < 0.01  < 0.001

Figure 7.6: Significance level dependent activation maps for SPM for selected brain layers. Color codes represent activated voxels which survive the critical value corresponding to significance level 0.001, 0.01, 0.05, 0.1 and 0.2.



Legend: analyzed voxels active in both active in examined model active in reference model

Figure 7.7: Activation difference maps between SPM and gCAR3 for selected brain layers. The most conservative activation map from gCAR3 (corresponding to threshold 0.95) is compared to the least conservative activation map from SPM (corresponding to a significance level of 0.2).

7.3.2 Evaluation of EEG-enhanced fMRI detection schemes

To evaluate whether the proposed models are able to increase sensitivity by using EEG resp. external activation information, we compare our EEG-enhanced activation detection algorithms with corresponding uninformed algorithms. We use predictor type 3 models as null models, which do not rely on additional EEG information. The figures used in this section to display results comprise the following contents:

Row	Expression	Description
1	$\mathbf{J} = (J_i)$	EEG resp. prior information map
2	$\hat{\gamma}_{\text{diff}} = (\hat{\gamma}_{i, M_{2/5}} - \hat{\gamma}_{i, M_3})$	Activation difference map (*)
3	$\hat{\pi}_{\text{diff}} = (\hat{p}(\gamma_i = 1 \boldsymbol{\theta}, \mathbf{y}, M_{2/5}) - \hat{p}(\gamma_i = 1 \boldsymbol{\theta}, \mathbf{y}, M_3))$	Posterior activation probability difference map (*)
4	$\hat{\Phi}_{\text{diff}} = (\Phi(\hat{\eta}_{i, M_{2/5}}; 0, 1) - \Phi(\hat{\eta}_{i, M_3}; 0, 1))$	Marginal prior probability difference map (*)
5	$\hat{\eta}_{EEG} = (\hat{\alpha}_i J_i)$	EEG predictor component map

(*) Predictor type 2 or 5 models ($M_{2/5}$) in reference to corresponding predictor 3 models (M_3) without EEG/additional prior information

To be more precise on marginal prior probability difference values (displayed in figure row 4), note the following. The calculation of the estimated marginalized prior activation probability is based on Equation 3.20, p. 64. Here, we insert plugin estimates of the predictor to calculate $\hat{p}(\gamma_i = 1 | \mathbf{U}_{j \neq i}, \boldsymbol{\theta}_{-U}) = \Phi(\hat{\eta}_i; 0, 1)$ for each voxel i . This quantity can be interpreted as the EEG dependent prior probability for activation—although it is no prior probability in a strict sense, but rather an adaptive prior that has been estimated. It contains posterior estimates in form of $\hat{\eta}_i, i = 1, \dots, n$, which in turn contains information about all available data including fMRI information. It can be used to see to which extent EEG activation influences fMRI activation estimation. To visualize the EEG influence on the estimate, we display the difference to the marginal prior probability estimate of the model run without EEG term (predictor type 3 model). Probability values possess the further advantage of being on a common scale of range [0; 1] leading to a range for difference values of [-1; 1].

Group and time aggregated sLORETA map

In this section, we report the results of the EEG-enhanced fMRI activation detection algorithms using the aggregated EEG-based prior information map from the Tensor-PICA

group analysis of sLORETA trajectories described in Section 7.2. As noted above, it is supposed to add information to fMRI activation detection especially in upper central parts of the brain.

In Figures 7.8a, 7.9a and 7.10a the results of the predictor type 5 model runs fCAR5, gCAR5 and iMRF5 with spatially-varying intercept and spatially-varying EEG coefficient are depicted, respectively. We select result maps from brain layers 21, 23, 25 and 27, because these layers comprise the parts of the brain where the EEG information is strong and is supposed to especially add value to the activation detection.

We know from the results of the simulation study presented in Chapter 6 that predictor type 5 algorithms are able to downweight EEG-prior information when it contradicts fMRI activation information (contained in voxelwise likelihood values of the model), whereas it shows a positive EEG effect if activation information is overlapping. An excellent example of the downweighting-mechanism can be seen in the central part of layer 25. On the contrary, in central parts of layer 27, a positive EEG effect, for example, supports fMRI activation. This property of predictor type 5 models can be observed with all three predictor types: gCAR5, fCAR5 and iMRF5. The differences in EEG predictor component maps between models are small indicating that all algorithms recover reliable estimates. Compared to gCAR5, EEG predictor values of fCAR5 contain somewhat sharper edges and more coarse structures, whereas iMRF5 effect maps are slightly smoother.

Marginal prior probability difference maps seem to be more suitable to assess algorithm specific properties. Though all algorithms obtain increased marginal prior probability values in areas with congruent fMRI and EEG information and decrease prior activation probabilities in non-congruent parts, we observe the following differences: The gCAR5 estimates differ from its reference, the gCAR3 estimates, in all parts of the brain regardless of whether EEG values deviate from their zero baseline or not. This seems to be an undesirable property of the algorithm. In the fCAR5 model, sharp separation between congruent and non-congruent brain areas are visible. Wherever positive EEG values and fMRI activation overlap, increased marginal prior probability values emerge. If positive EEG values arise in regions without fMRI activation, a strong decrease in marginal prior probability can be observed. This might imply that the inclusion of EEG-prior information does not provide any additional value: fCAR5 only seems to emphasize activation effects that are already there. In iMRF5 models, we observe changes mainly in parts of the brain where some kind of activation happens. This hints at the stability of the algorithm. Marginal prior activation probability is especially increased on activation region boundaries, e.g. on the central activation focus in layer 27.

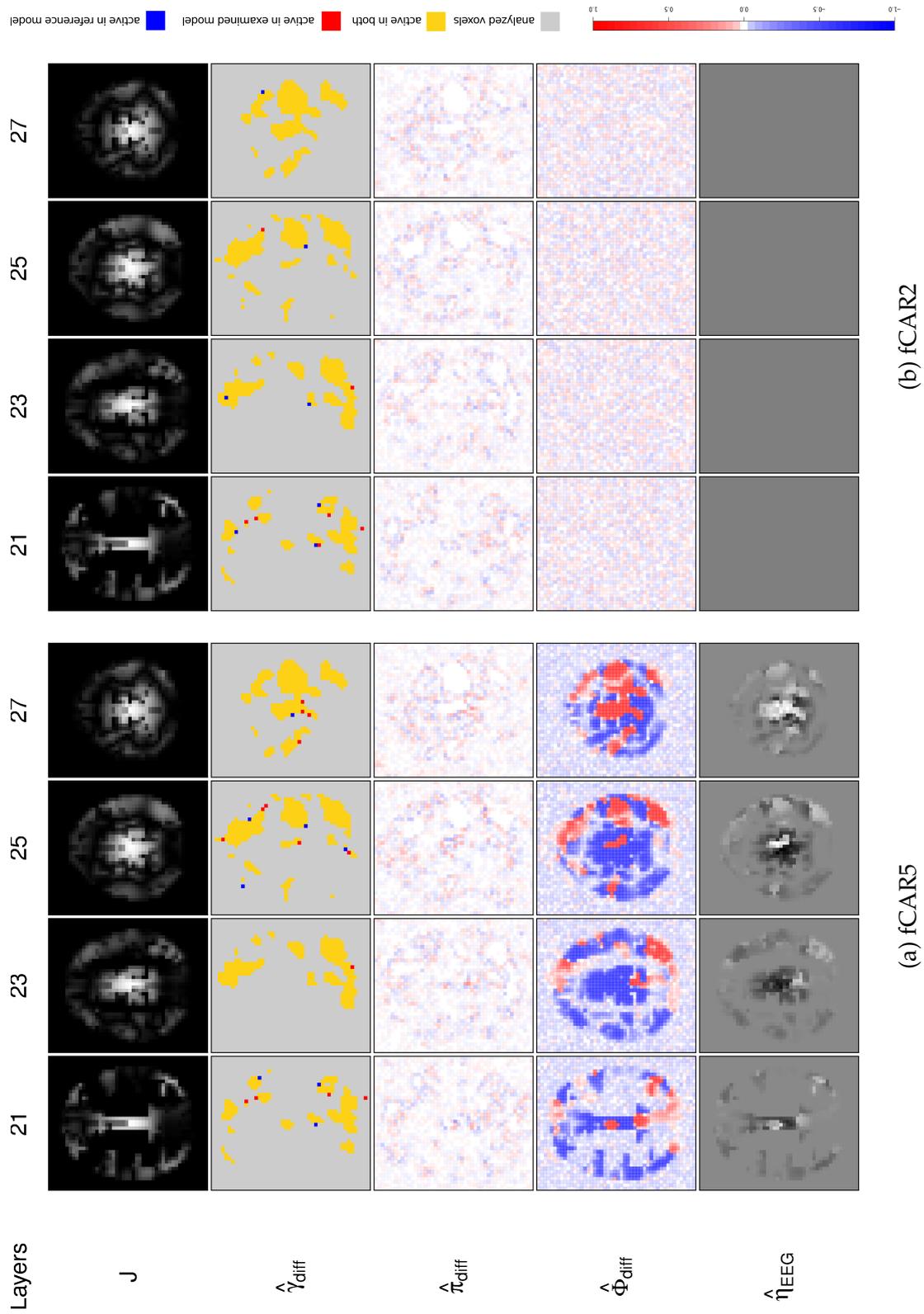


Figure 7.8: Aggregated sLORETA map: Summary of results from the fCAR5 (left) and fCAR2 (right) model in comparison to fCAR3 (selected brain layers: 21, 23, 25, 27).

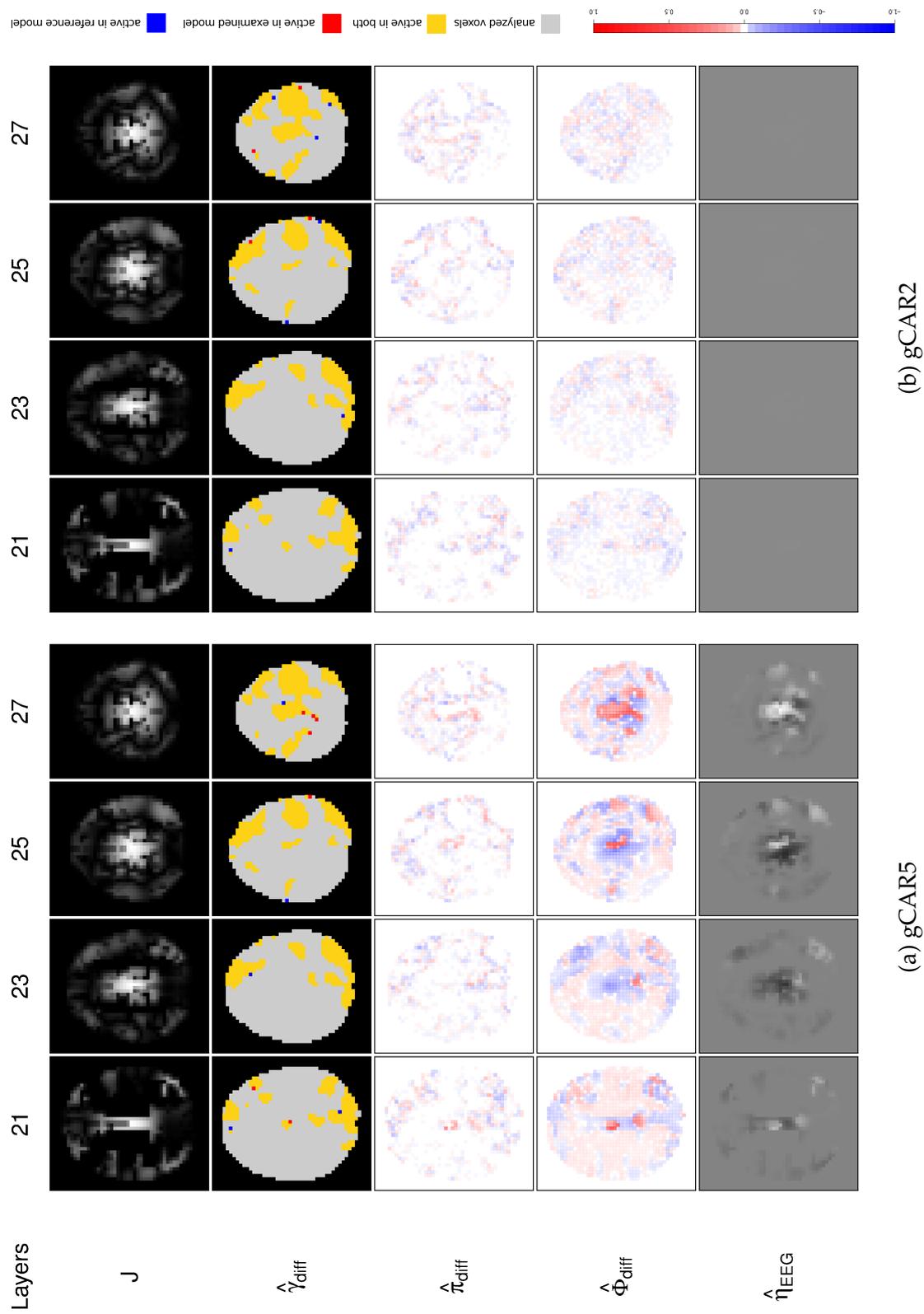


Figure 7.9: Aggregated sLORETA map: Summary of results from the gCAR5 (left) and gCAR2 (right) model in comparison to gCAR3 (selected brain layers: 21, 23, 25, 27).

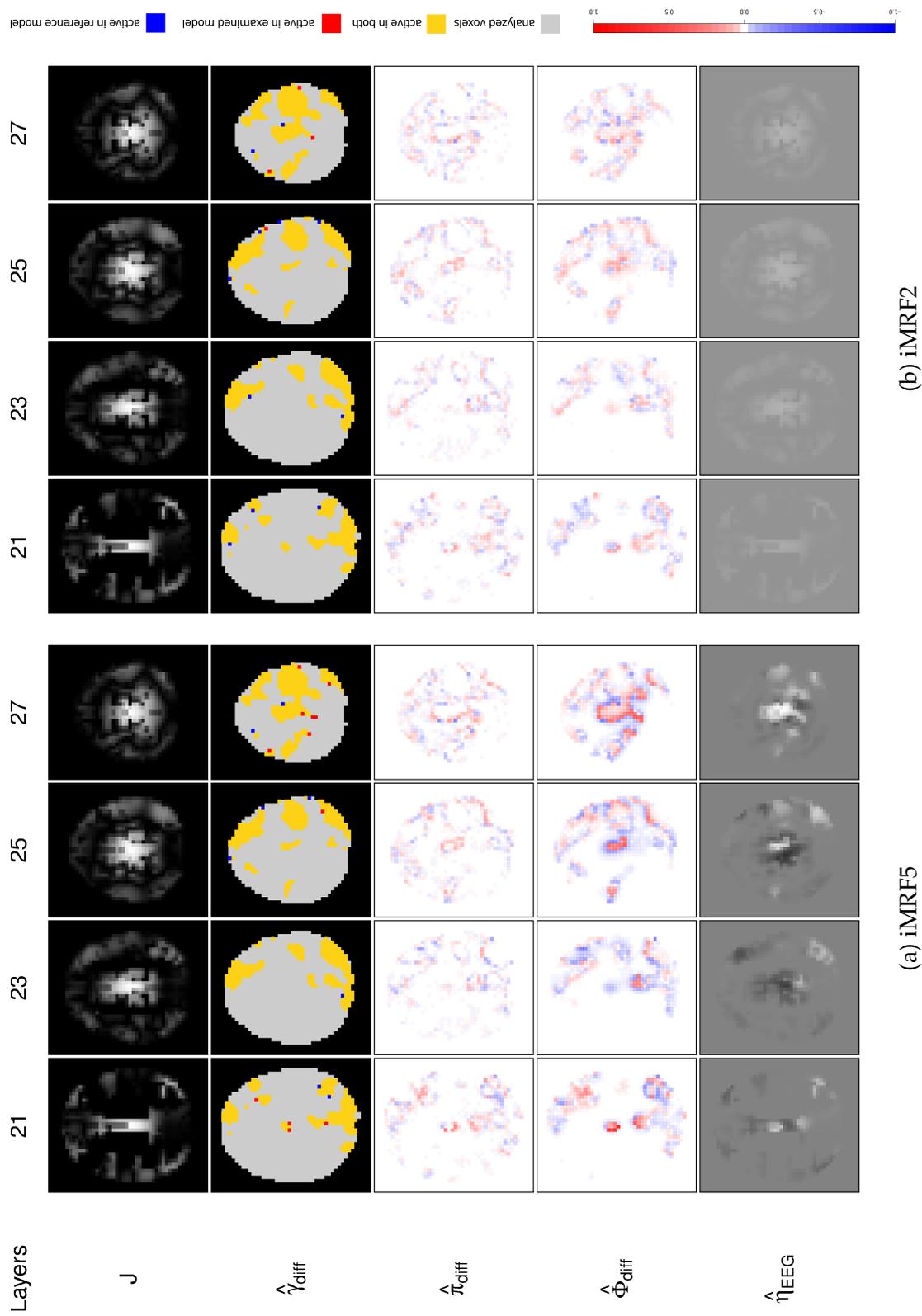


Figure 7.10: Aggregated sLORETA map: Summary of results from the iMRF5 (left) and iMRF2 (right) model in comparison to iMRF3 (selected brain layers: 21, 23, 25, 27).

Marginal prior probability difference maps can be used to evaluate whether latent probit effects make sense. In posterior activation maps, though, we see whether EEG-prior information has an impact on activation estimation at all. The fCAR5 results show differences to its reference, the fCAR3 results, all over the brain: Besides increased activation probability values in central parts of the brain, regional differences appear that cannot be traced back to changes in EEG information. Moreover, the map appears to be very coarse, i.e. changing from positive to negative difference values within the distance of 1 voxel (and back) without obvious cause. This hints at instability issues of the algorithm. In activation maps, this is reflected in voxels becoming non-active without clear reason. Nevertheless, fCAR5 enlarges the activation focus in the upper central part of the brain.

The gCAR5 and iMRF5 algorithms locate largest posterior probability differences on activation region boundaries with supporting EEG information, i.e. in central parts of the brain. These are large enough to slightly increase region sizes in targeted areas (e.g. in the central part in layer 27). A few voxels become non-active in regions where algorithms emphasize sharp edges to non-activation. These regions are visible in marginal prior probability maps where a strong decrease in prior activation probability can be detected on the right margin of the central activation focus in layer 27.

One of the most interesting observations can be made in the central part of layer 21: Both gCAR5 and iMRF5 can find additional activation there and support an existing damped activation foci. We know from the analysis of different time subsets of the fMRI algorithm (cf. Appendix D) that this region is subject to a time-varying activation effect: Activation can be observed in the beginning of the experiment, then the activation focus vanishes. If the whole fMRI trajectory is analyzed, activation probability estimates are averaged over all scans and hence, are reduced. With the iMRF5 and gCAR5 models the prior activation probability can be increased again using external EEG information. However, as a fact supporting the fCAR algorithm, we observe that the fCAR algorithm (in particular already the fCAR3 algorithm) finds this activation foci in any case. Here, EEG information support is dispensable.

Considering the results of all three predictor 5 algorithms, we come to the following conclusions. The iMRF5 and gCAR5 algorithms provide interpretable results and show (slightly) increased sensitivity in targeted parts of the brain—with iMRF5 being the marginally more sensitive one. The fCAR5 algorithm retrieves coarse map estimates. Changes in activation maps can only partly be traced back to EEG information. Nevertheless, fCAR5 is also able to slightly enhance fMRI activation detection by the use of the aggregated sLORETA map (in upper central parts of the brain). Whether the algorithm is susceptible to noise or has

an improved adaptability to coarse activation structures, is not obvious. However, the sensitivity increase with all algorithms is not as large as expected. This might be traced back to the following. The fMRI data at hand contains activation peaks with sharp edges to non-activated areas, whereas the EEG source map is by far smoother. Hence, congruent areas adjoin non-congruent areas. An EEG-enhanced activation detection algorithm detecting such structures is likely to be hardly able to differentiate between actual edges and proximate voxels with a noisy fMRI activation signal. Thus, EEG-enhancement does not take effect. This does not seem to be a problem if congruent and non-congruent areas are well separated—as in the analysis of the artificial dataset in Chapter 6.

In the simulation study presented in Section 6.1.5, predictor type 2 models with a global EEG coefficient (restricted to be positive) showed an equally high performance as the predictor type 5 models. Hence, we also used fCAR2, gCAR2 and iMRF2 for analysis. Results can be found in Figure 7.8b, 7.9b and 7.10b. Note that the color scale of the EEG predictor component maps is chosen with respect to the range of the whole predictor. Hence, if the EEG component is not sufficiently strong, it is not visible on the plotted maps. This happens with both gCAR2 and fCAR2 models. Though the gCAR2 model exhibits some structure (which is visible in marginalized prior probability maps in comparison to gCAR3), a connection to an EEG dependent decrease/increase in activation probability can, however, hardly be noticed. Differences between the fCAR2 and fCAR3 model seem to be completely within random error. The iMRF2 model yields a global EEG effect size that stands out from the intercept component. It is even strong enough to affect posterior activation probabilities. Nevertheless, none of the three models—neither fCAR2, gCAR2 nor iMRF2—seems able to increase sensitivity in this application (substantially). As soon as EEG contradicts fMRI in larger parts of the brain, the global EEG effect is estimated to be so small that it hardly influences activation probability.

In general, the increase in sensitivity by using EEG-enhanced fMRI activation detection algorithms with predictor type 2 and 5 seems limited or, at least, not as promising as expected. We suppose that this might be traced back to incompatible EEG and fMRI maps. The sLORETA analyzing mask, for example, might not fit to the fMRI data structure on the whole provoking meaningless sharp edges in effect maps. Hence, we decided to examine the results of alternative EEG maps.

Subject-specific sLORETA map

EEG-enhanced fMRI activation detection results seem less promising when using an aggregated sLORETA component map. As noted in the preceding section, this might

be due to incompatible information contained in EEG and fMRI maps. Amongst other things, incompatibility may result from the Tensor-PICA specific decomposition of the sLORETA trajectory into distinct spatial components. Voxels with similar time response are grouped into one spatial component. This may lead to an adverse decomposition of activation structures. For example when upper central and lower frontal parts of the brain are activated at a specific time component, but only the frontal part shows activation also at another time, then these two activation structures are broken apart into different component maps. Hence, if we select just one spatial EEG component to combine it with fMRI, missing activation structures introduce non-congruency possibly influencing analysis results negatively. Hence, we decided to also look at results from analyses using a non-aggregated, subject-specific sLORETA map as prior information. The EEG data this sLORETA map is based on was measured simultaneously on the same individual as the fMRI data. We choose the sLORETA source distribution map for a selected time point ($t \approx 480$ ms (PST); sLORETA map no. 172) of the ERP difference profile of odd and even stimuli: This sLORETA map was selected for showing a strong EEG activation in several brain areas—in particular frontal and in the upper back parts of the brain. In general, this EEG based map contained information closer (i.e. more congruent) to the fMRI data than the Tensor-PICA map. We expected that this prior information map might be easier to handle for proposed algorithms. For having a strong gradient on EEG activation boundaries (range $[0; 6.5 * 10^3]$), we decided to logarithmize the map to enhance numerical stability and spatial dependency estimation. Results from predictor type 5 models are shown in Figures 7.11a, 7.12a and 7.13a. Results from predictor type 2 models are shown in Figures 7.11b, 7.12a and 7.13a. Figure contents are described in detail in the beginning of Section 7.3.2. We select result maps from brain layers 9, 11, 27 and 29, because these layers comprise the parts of the brain where the EEG information is strong and is supposed to especially add value to the activation detection.

The predictor type 5 models exhibit mainly the same properties as in the runs with the aggregated sLORETA map as described in the preceding subsection, e.g. EEG is downweighted when it contradicts fMRI information. The fCAR5 model seems to (over-)emphasize EEG effects with sharp edges between congruent and non-congruent activation regions as can be seen in the marginal prior probability difference maps. Resulting posterior probability activation maps seem noisy. In this example, however, these properties seem useful to detect an increased number of active voxels in the frontal part of the brain in layers 9 and 11 and in the center of layer 29. The found activation structure in the frontal part of layers 9 and 11 is very coarse. Activation in the fMRI signal seems

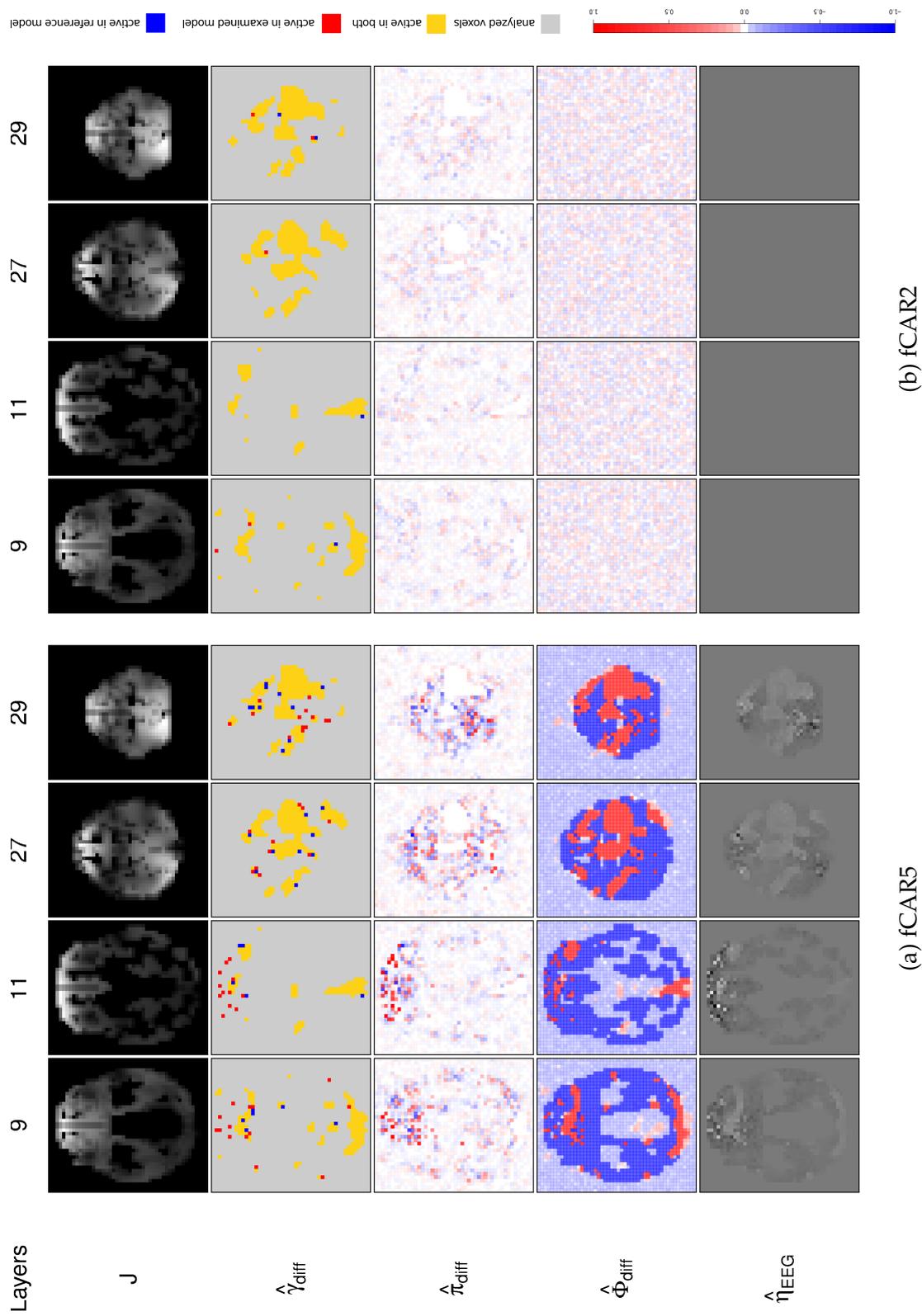


Figure 7.11: Subject-specific logarithmized sLORETA map: Summary of results from the fCAR5 (left) and fCAR2 (right) model in comparison to fCAR3 (selected brain layers: 9, 11, 27, 29).

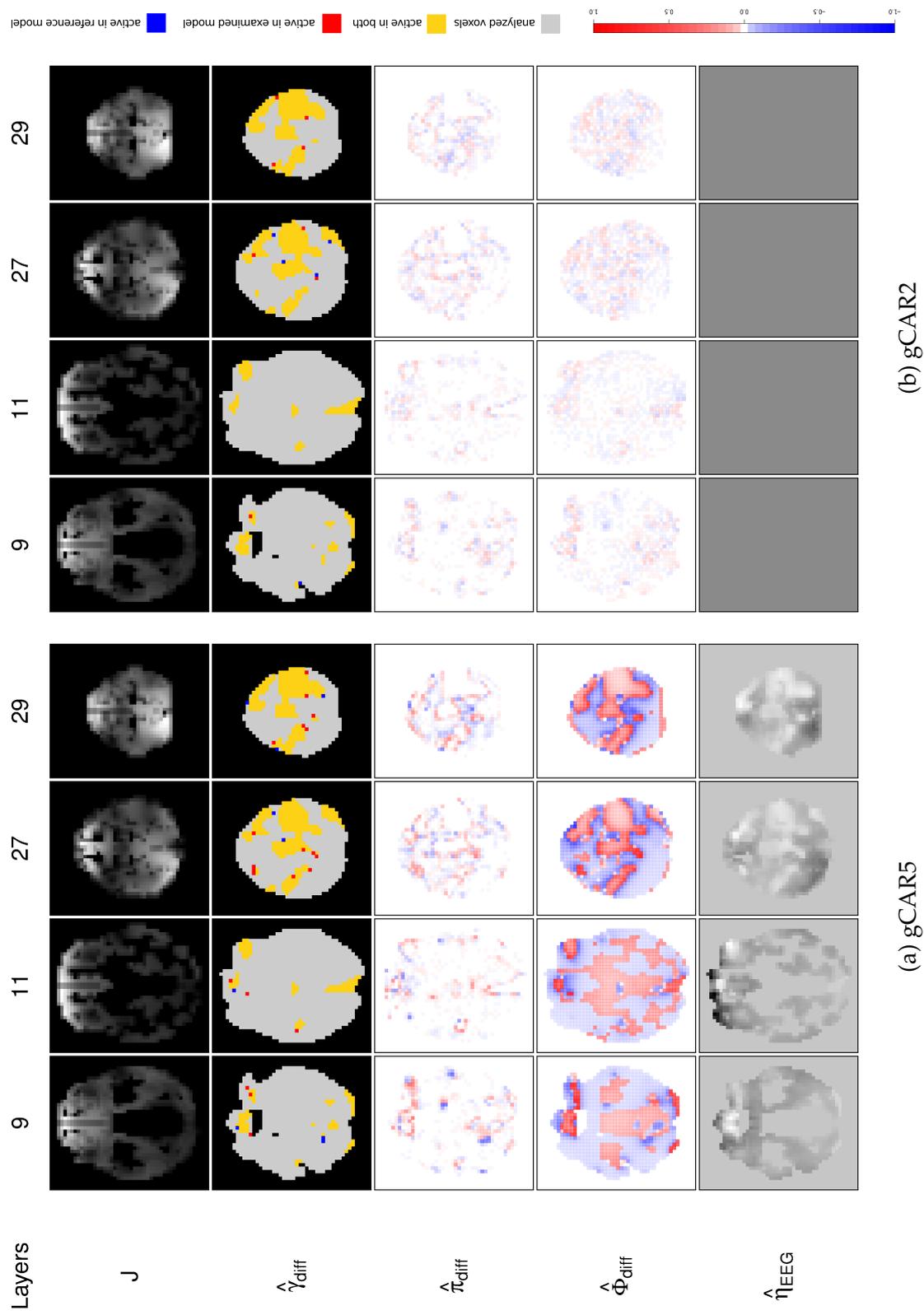
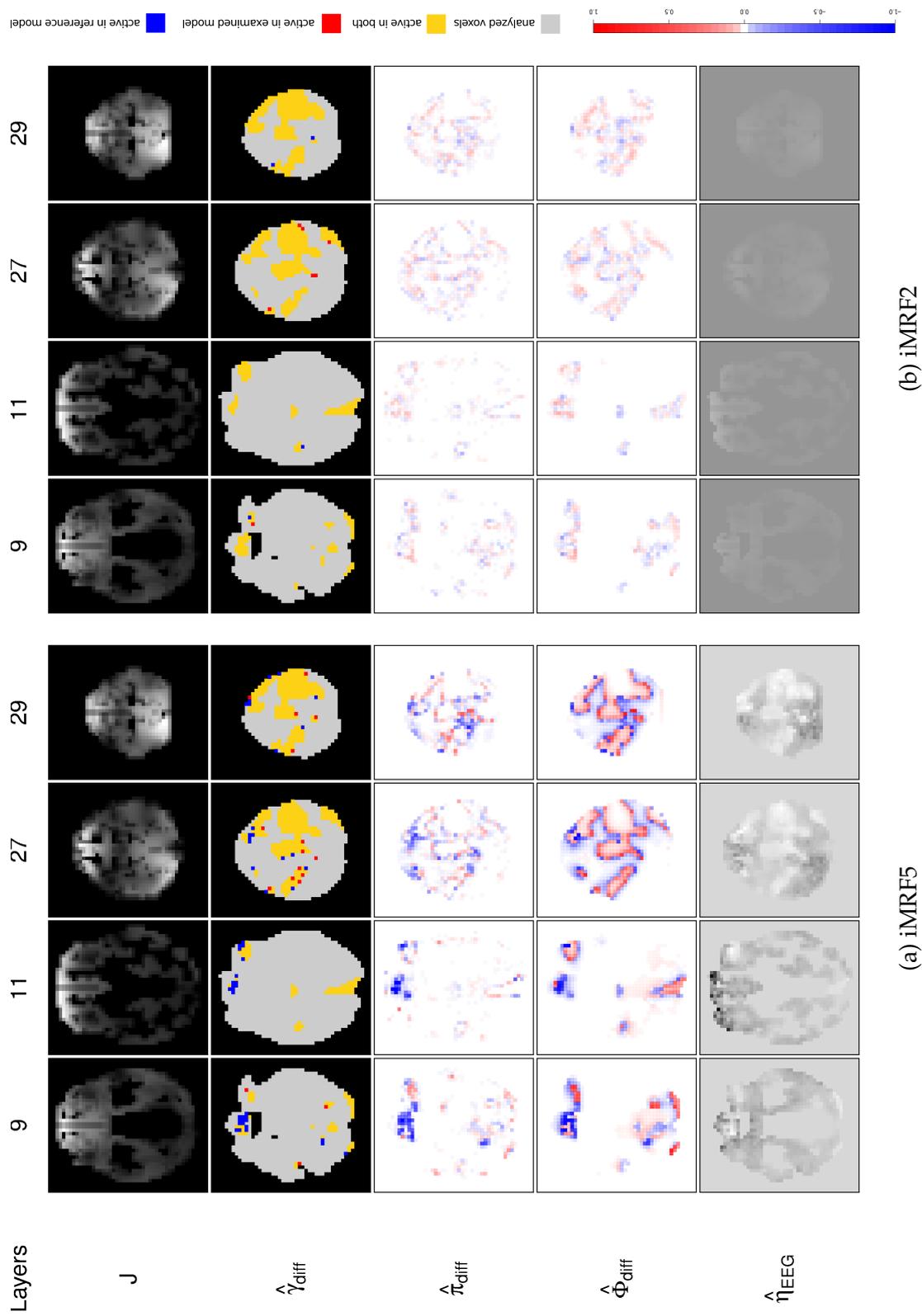


Figure 7.12: Subject-specific logarithmized sLORETA map: Summary of results from the gCAR5 (left) and gCAR2 (right) model in comparison to gCAR3 (selected brain layers: 9, 11, 27, 29).



(b) iMRF2

(a) iMRF5

Figure 7.13: Subject-specific logarithmized sLORETA map: Summary of results from the iMRF5 (left) and iMRF2 (right) model in comparison to iMRF3 (selected brain layers: 9, 11, 27, 29).

to be very weak: Single voxels contain a minimum quantity of activation, whereas direct neighbors do not. Strong EEG information now brings out this coarse activation pattern. However, found activation can be due to unwanted blinking artifacts.

The gCAR5 model obtains smooth EEG effect estimates. In contrary to the fCAR5 pattern, it now seems to smooth out the occasional, isolated activation peaks, e.g. in frontal parts. Nevertheless, gCAR5 utilizes the EEG-prior information to slightly increase sensitivity in the center of layers 27 and 29.

The iMRF5 results, however, point into the opposite direction of fCAR5: It uses the fMRI information on non-activation in the neighborhood to obtain a smooth area of decreased posterior probability estimates and reduces activation structures in the frontal part of layers 9 and 11 substantially. In layer 27, EEG can increase sensitivity, though iMRF5 results also pronounce edges leading to some voxels becoming non-active at activation boundaries.

As we can see from these results, each of the three algorithms comes to different results in the presence of strong EEG activation and weak and noisy fMRI activation signals. fCAR5 finds isolated active voxels, whereas gCAR5 is conservative, i.e. it strengthens existing activation foci in this area (found already by gCAR3), but smooths out isolated voxels with strong EEG activation. The iMRF5, on the contrary, reduces any activation found in this area. In general, we see that all three algorithms cope in different ways with strong EEG activation. However, it is not evident which strategy might generally be the most accurate one. This is likely to vary from case to case. In this example, the iMRF5 behavior is preferable for eliminating erroneous blinking artifacts.

When looking at the predictor 2 results, we come to the same conclusions as with the aggregated sLORETA map described before: The EEG component containing the global EEG effect is not strong enough to be visible on plotted predictor maps for the intercept component dominating the predictor. Posterior probability differences between fCAR2 and fCAR3 as well as between gCAR2 and gCAR3 seem to be mainly within random error. Activation changes in binary activation maps seem to be related to this fluctuations. Again, the iMRF2 model yields a global EEG effect size that stands out from the intercept component. It is even strong enough to find some additional active voxels in layer 27. Nevertheless, none of the three models—neither fCAR2, gCAR2 nor iMRF2—seems able to increase sensitivity in this application (substantially). Again, the subject-specific sLORETA map and the fMRI activation structure seem to be fairly non-congruent provoking these results.

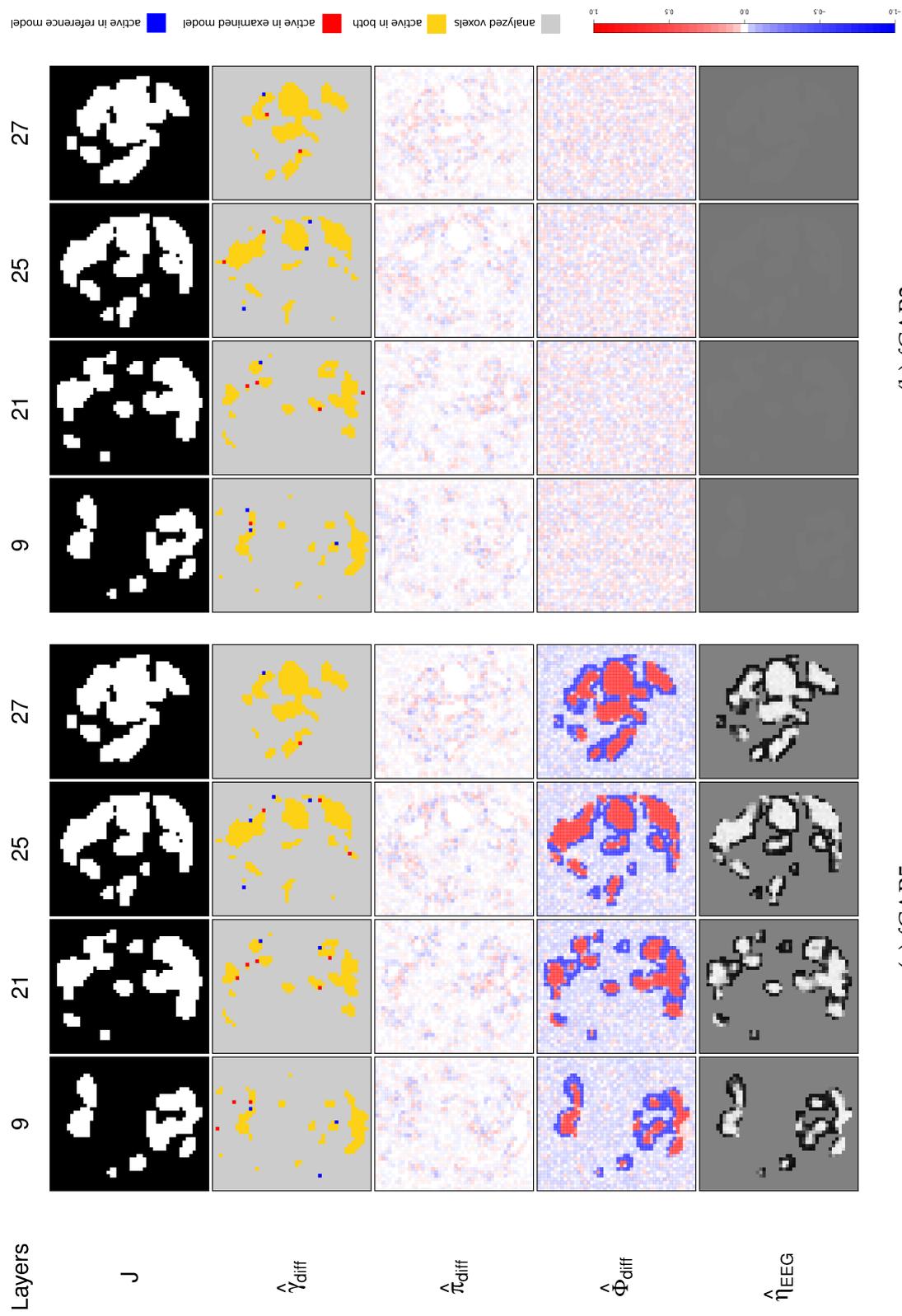
Test case with congruent prior information

As noted before, component maps derived by a Tensor-PICA of sLORETA map trajectories contain non-congruent activation information compared to fMRI in several parts of the brain. From the analysis results presented in the preceding section, we see that the subject-specific sLORETA map does not contain suitable information to enhance sensitivity substantially as well. Therefore, we decided to run a test and use prior information congruent with fMRI activation to see whether sensitivity can be increased in this case. For this, we take a binary fMRI activation map from a run without EEG information (gCAR3) and enlarge activation regions by adding the next 28 nearest neighbors of the cubic neighborhood. The resulting binary map is used as prior information in the enhanced fMRI activation algorithms. We expect this prior information to enlarge activation regions at boundaries. Results from predictor type 5 models are shown in Figures 7.14a, 7.15a and 7.16a. Results from predictor type 2 models are shown in Figures 7.14b, 7.15b and 7.16b. Figure contents are described in detail in the beginning of Section 7.3.2. We select result maps from brain layers 9, 21, 25 and 27.

The EEG predictor component estimated by the fCAR2 algorithm is virtually zero. Hence, the marginal prior probability difference maps exhibit no EEG correlated structure. Structures visible in the posterior probability difference maps seem to be due to fMRI likelihood influence. In activation difference maps, we notice some voxel state changes from active to non-active and vice-versa. Regarding the negligibly small EEG effect, this changes might be traced back to the instability of the fCAR algorithm: Having a negligible dependency between neighboring voxels, the EEG effect component seems to be subsumed into the spatially-varying intercept estimation. Possibly, identification problems arise.

The same potentially instable estimation results can be observed in the brain-background regions of fCAR5. There, both prior and posterior probability maps exhibit a lot of noise. However, compared to fCAR2 a strong EEG effect can be retrieved with sharp edges distinguishing activation regions from their boundaries. At some boundary locations, nevertheless, EEG-enhancement successfully detects activation not found before, but also the reverse happens: some voxels are now classified as non-activated.

The gCAR2 algorithm retrieves a reasonable large global EEG effect. This is what we expected with congruent fMRI and prior information. In marginal prior probability maps, hence, an increase in prior activation probability can be observed on activation boundaries, which is also reflected in posterior probability maps as well as in activation maps. Here, an increase in activated voxels can be observed. Two properties of the algorithm seem to be



(b) fCAR2

(a) fCAR5

Figure 7.14: Congruent prior information map: Summary of results from the fCAR5 (left) and fCAR2 (right) model in comparison to fCAR3 (selected brain layers: 9, 21, 25, 27).

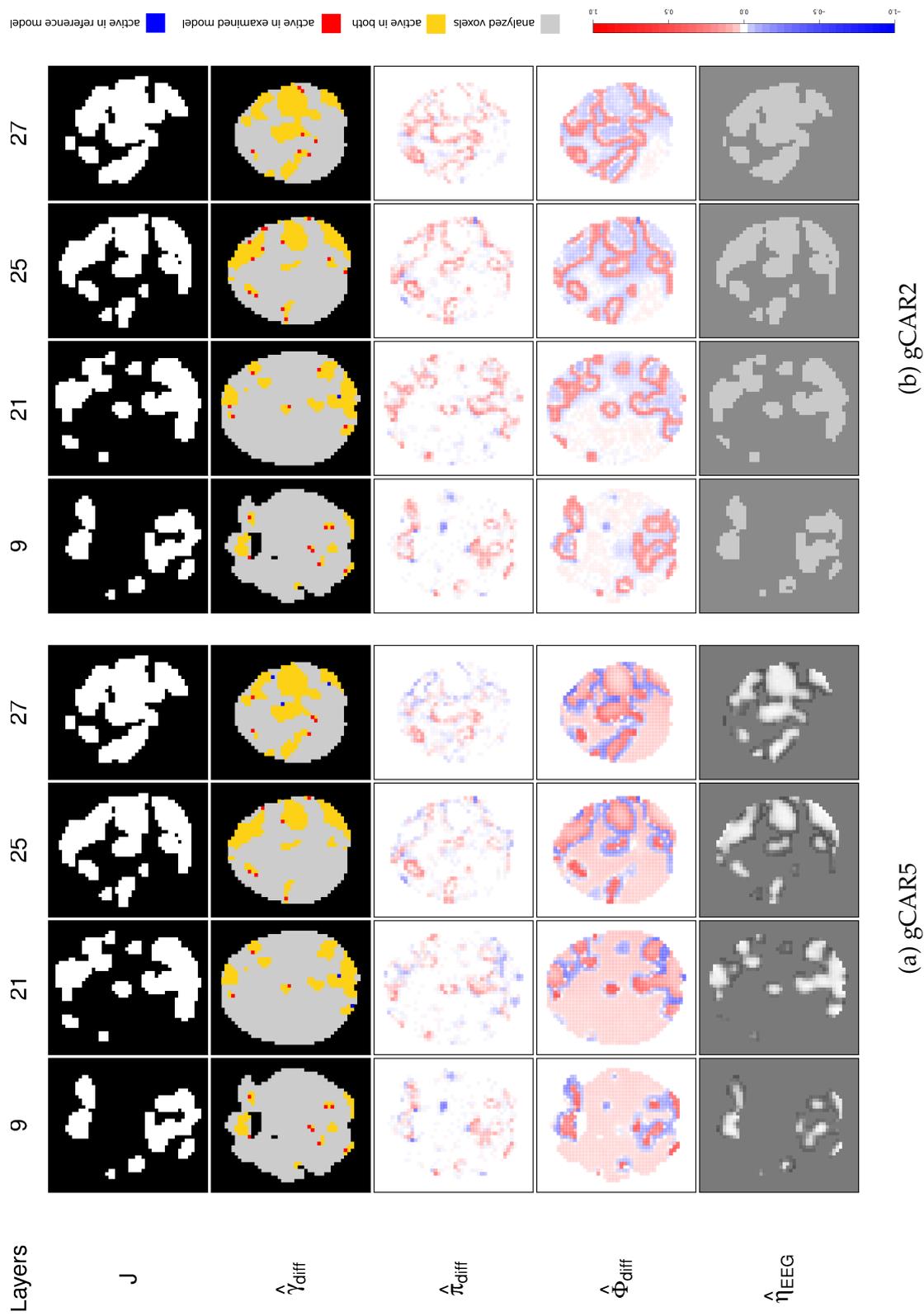


Figure 7.15: Congruent prior information map: Summary of results from the gCAR5 (left) and gCAR2 (right) model in comparison to gCAR3 (selected brain layers: 9, 21, 25, 27).

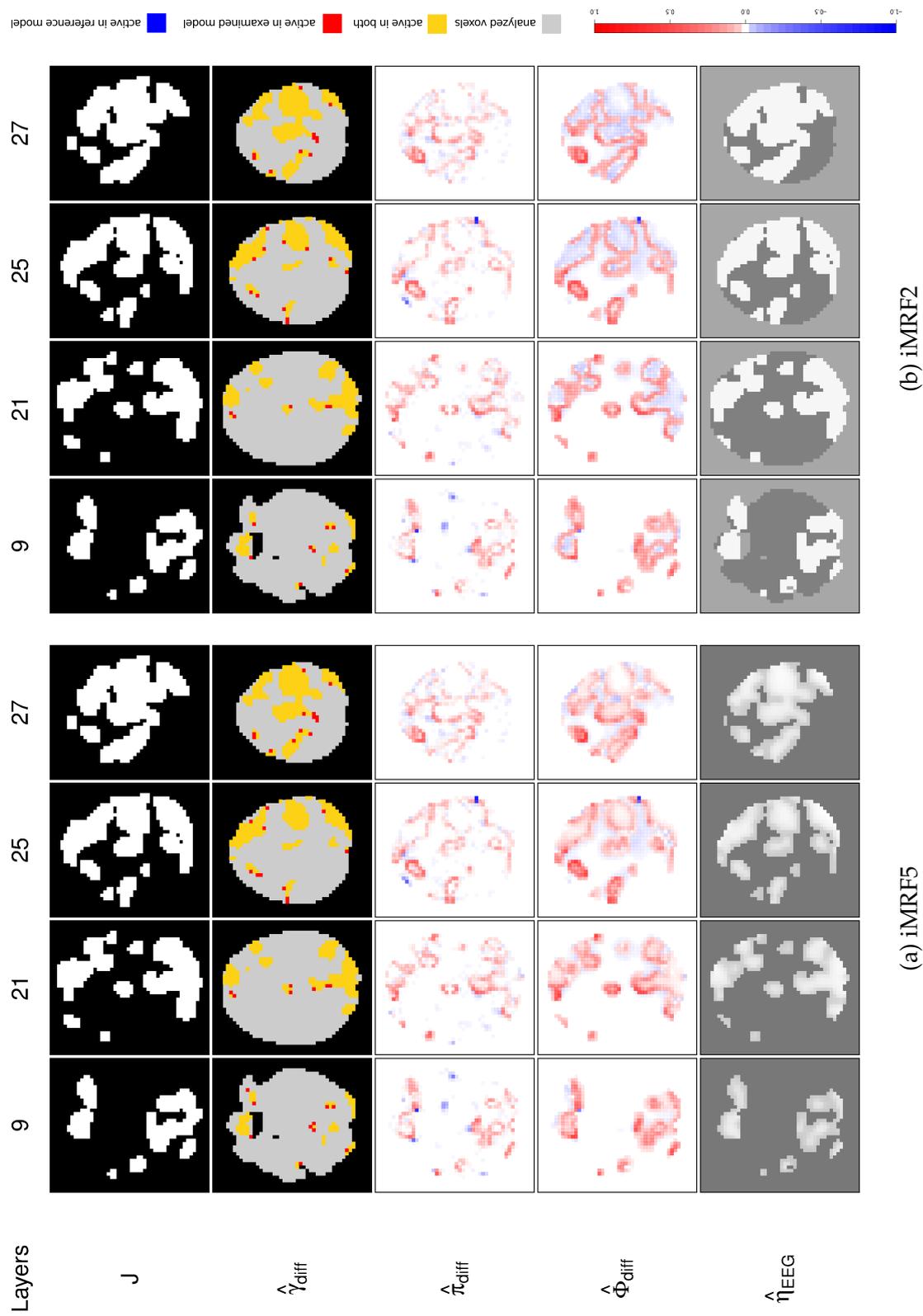


Figure 7.16: Congruent prior information map: Summary of results from the iMRF5 (left) and iMRF2 (right) model in comparison to iMRF3 (selected brain layers: 9, 21, 25, 27).

less desirable: As we can see in marginal probability maps, the intercept component of the probit stage seems, on the one hand, to overcompensate the large global EEG effect and, on the other hand, to provoke a decrease in prior probability in inner- and inter-region spaces. This, however, does not affect posterior probability maps.

The gCAR5 algorithm does not show this overcompensation. On the contrary, marginal prior activation probabilities are increased in the inside of activation regions as well as on several boundaries—which in turn leads to increased posterior probabilities and more activated voxels. Good edge preserving properties can be observed by down-weighting EEG influence on extended region boundaries where fMRI activation is entirely absent. An undesirable property of the model may be the arbitrary increase in prior probability in non-activation regions. Fortunately, this is not reflected in posterior estimation maps.

The iMRF2 results are very similar to gCAR2 results. We observe, however, a weaker overcompensating effect of the probit intercept. An even stronger increase in posterior activation maps on region boundaries leads to the detection of comparatively more activated voxels than with gCAR2.

The iMRF5 model retrieves a strong EEG effect, which leads to increased prior activation probability values on activation regions and corresponding boundaries. Compared to gCAR5, this increase affects posterior estimation maps even more positively leading to increased posterior probability values on region boundaries and additionally activated voxels. Though the EEG component values does not drop to sub-zero values on region edges, prior and posterior probability map structures hint at good edge-preserving properties. In non-activation regions minor changes between the iMRF3 and iMRF5 algorithms can be observed.

Considering the results of all algorithms, we come to the following conclusions: When using congruent prior information, the gCAR and iMRF algorithms with both predictor types yield an increase in sensitivity. These four algorithms have slightly different properties: The predictor type 2 models, gCAR2 and iMRF2, possess the disadvantage to incline to overcompensate strong global EEG effects in some regions. The gCAR5 algorithm seems to overvalue edges and retrieves large-area changes to gCAR3 results in regions where no activation happens. The iMRF5 algorithm seems to be the most promising one of the four algorithms with global updates: It has neither of the mentioned disadvantages, but appears to be one of the most sensitive algorithms. This might be traced back to its advantageous smoothness properties. However, compared to the fCAR algorithm (particularly already fCAR3), the sensitivity of algorithms with global updates is decreased in regions where

underlying activation structures are coarse and signals are weak. This happens particularly in lower layers of the brain scan. As noted in Section 7.3.1, advantages of gCAR and iMRF lie in intensifying large contiguous activation regions—as in upper central brain parts. Despite the high sensitivity of fCAR3, the fCAR2 and fCAR5 algorithms seem hardly be able to increase sensitivity by utilizing external prior information.

Subject-specific, time aggregated sLORETA map

In preceding analyses with highly smoothed fMRI data and two different types of real EEG-prior information, we did not observe a substantial gain in performance of EEG-enhanced activation schemes. The test case with congruent prior activation information showed that the iMRF2/5 and gCAR2/5 algorithms had a slightly increased sensitivity compared to algorithms without EEG-prior. The results of these analyses and the results of simulation studies from Chapter 6 hint at characteristics the combined data must possess for a successful application of EEG-enhanced fMRI activation detection. In Chapter 8 we summarize these findings and conduct an additional study to substantiate these and to come up with a more precise description of suitable data settings. Following derived recommendations, we now consider a new combination of fMRI and EEG data supposed to be suitable for illustrating the usefulness of our approach.

For this analysis, we use a version of the fMRI data that is only slightly smoothed. The spatially aligned, warped fMRI times series is now smoothed with a Gaussian kernel of $2 \times 2 \times 2 \text{ mm}^3$ FWHM (instead of $8 \times 8 \times 8 \text{ mm}^3$). There are several reasons for smoothing fMRI data before analysis (Friston et al., 2008, p. 225), e.g. increasing the signal-to-noise ratio in the data. Applying a gaussian kernel, however, smooths over structural edges and might not preserve image structure. Hence, reducing the kernel size and analyzing less smooth data might generally be of interest for practitioners.

For the EEG-prior, we take an subject-specific, time aggregated sLORETA map from the subject the fMRI data was recorded on. An ICA analysis of the sLORETA trajectory was used to find spatial component maps. We select a spatial component map that shows a strong signal in the (neurologically) right part of Brodmann areas 1-4, 6 (containing the primary somatosensory, primary motor and premotor cortex). This area also contains a strong, but spatially non-smooth, fMRI signal. After reslicing, we smooth the sLORETA map with a Gaussian kernel of $8 \times 8 \times 8 \text{ mm}^3$ (FWHM) to adjust for a mismatch between sLORETA analyzing mask and the brain structure found by fMRI. This should put less emphasize on analyzing mask edges of sLORETA.

In the following, we restrict data analysis to the (neurologically) right part of Brodmann areas 1-4, 6 to circumvent non-congruency issues arising in other parts of the brain.

Results from predictor type 5 models are shown in Figures 7.17a, 7.18a and 7.19a. Results from predictor type 2 models are shown in Figures 7.17b, 7.18b and 7.19b. Figure contents are described in detail in the beginning of Section 7.3.2. We select result maps from brain layers 25-28.

So far, the fCAR algorithm appeared to be less suitable for EEG-enhanced activation detection—although, in Chapter 8, we describe some settings where fCAR2 had a promising performance. Here, neither fCAR2 nor fCAR5 appears to be useful for EEG-enhancement. On the one hand, EEG-prior information is not able to improve sensitivity with fCAR2 and, on the other hand, fCAR5 slightly deteriorates the already found activation structure, which renders the algorithm to be inapplicable.

There is not much difference between the performance of algorithms with global updates. In contrast to the findings in Chapter 8, where the applicability of gCAR5 was questioned, gCAR5 does not perform worse than iMRF5 regarding activation classification. Both algorithms find approximately the same number of additional active voxels, but also some voxels become non-active. Changes in activation classifications, however, do not occur at the same locations. In spite of this, when looking at posterior activation probability estimation maps, iMRF5 seems to perform better. In contrast to gCAR5, it increases activation probabilities in most targeted regions of the observed activation focus—in parts where EEG-prior information is strong and fMRI is deteriorated by noise, e.g. in the upper left part of the activation focus.

For the fMRI and EEG information being fairly congruent, both gCAR2 and iMRF2 show a promising performance by increasing activation probabilities in most targeted parts of the region of interest. They find some additional voxels that can be classified as active.

Summarizing the results of all algorithms, the fCAR2/5 algorithms appear to be inapplicable for EEG-enhanced fMRI activation detection; the algorithms with global updates had a similar performance with iMRF2 being the most sensitive one (see results of layer 27). Nevertheless, none of the algorithms managed to compensate the noisiness of the fMRI data and fill up all holes in the activation structure.

This might be traced back to the following. On the one hand, the fMRI activation structure is not as suitable for EEG-enhanced activation detection as we thought it to be (cf. Figure 7.20a). Most activation gaps occurred at the outside of the activation focus, where

not enough neighboring information can compensate missing sensitivity. On the other hand, EEG information maybe has to be even more congruent and must contain even stronger values in targeted regions. As can be seen from Figure 7.20b, properties of the EEG information map could have been more advantageous. It seems that even more diligence has to be put into the search of finding a suitable combined dataset.

7.4 Summary

New methods have to be applied to real-world examples to test whether they yield the hoped improvements while producing sensible and interpretable results. Having applied our uninformed as well as EEG-enhanced fMRI detection algorithms to data from an auditory oddball experiment, we come to the following conclusions.

The performance of fMRI activation detection algorithms without additional EEG information was impressive: fCAR3, gCAR3 and iMRF3 outperformed existing activation detection algorithms like the Ising algorithm and classical SPM by far. Simulation studies from Chapter 6 suggested that increased sensitivity is not associated with a loss in specificity. Differences between proposed algorithms were not too pronounced. Nevertheless, the fCAR3 algorithm with fast single-site updates was highly sensitive when the underlying activation structure was rather coarse. It could find small isolated activation clusters/voxels for putting less weight on voxel dependencies. The iMRF3 and gCAR3 had better sensitivity in smooth structures. They were capable of strengthen existing activation foci, but smoothed out small activation peaks in brain regions without dominant activation structures. Thereby iMRF3 seemed to have slightly better performance than gCAR3. Concluding, we cannot recommend to use just one of the proposed models. Depending on the brain region of interest and the inherent activation structure either fCAR3 or iMRF3 should be preferred.

Concerning EEG-enhanced activation detection algorithms, we come to the following conclusions. Proceed with caution when using predictor type 2 models with global EEG effect. They did not have any benefit when EEG (or rather external) prior information contradicts fMRI information in large parts of the brain for yielding virtually zero global EEG effect estimates. Nevertheless, when using EEG-prior information that was mostly congruent to fMRI data, gCAR2 and iMRF2 models had a positive impact on sensitivity. For further analyses, one can make use of this result. Congruent EEG-based prior information maps have to be retrieved from the original set of EEG recordings that can add information

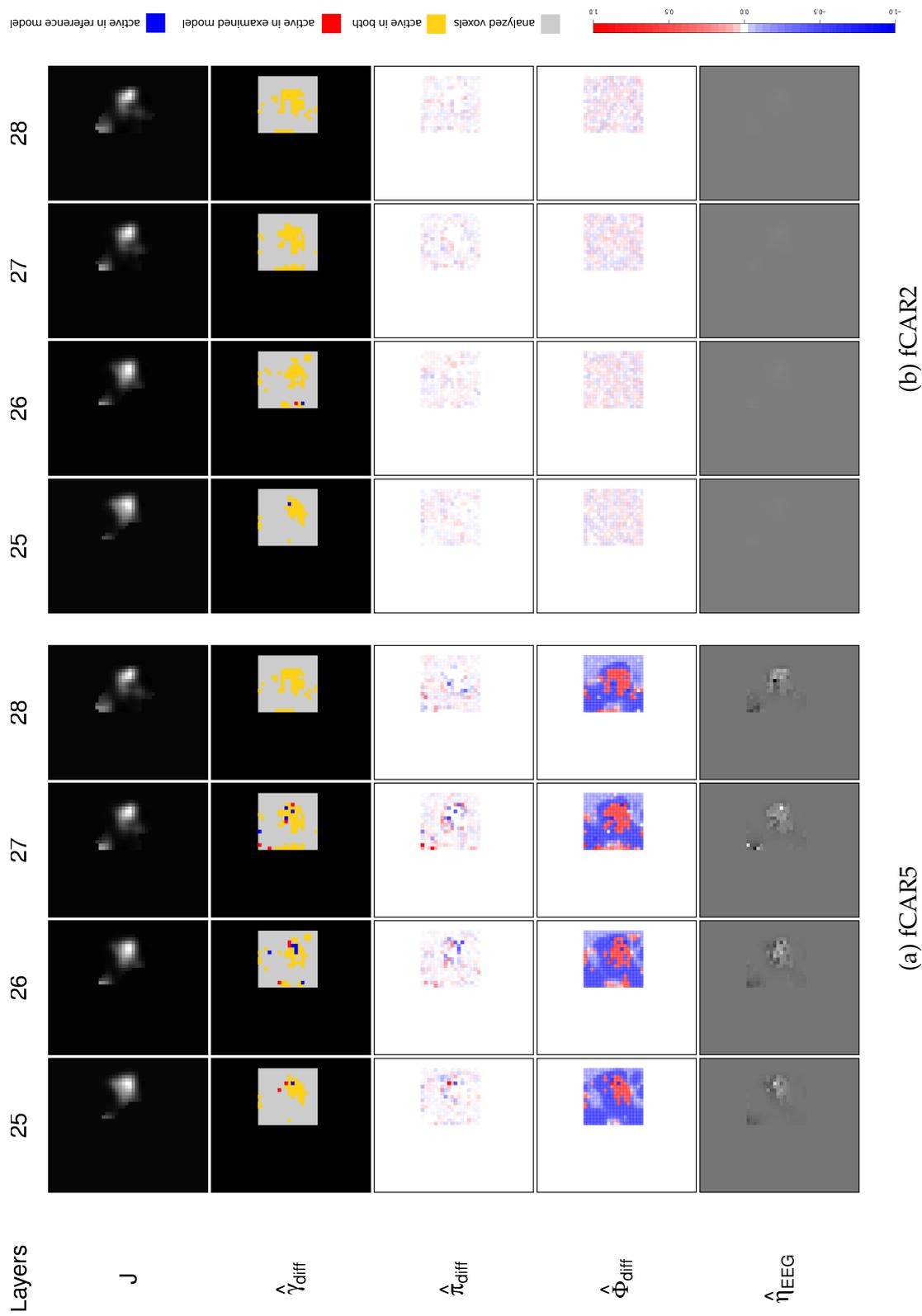


Figure 7.17: Subject-specific, time aggregated sLORETA map: Summary of results from the fCAR5 (left) and fCAR2 (right) model in comparison to fCAR3 (selected brain layers: 25–28).

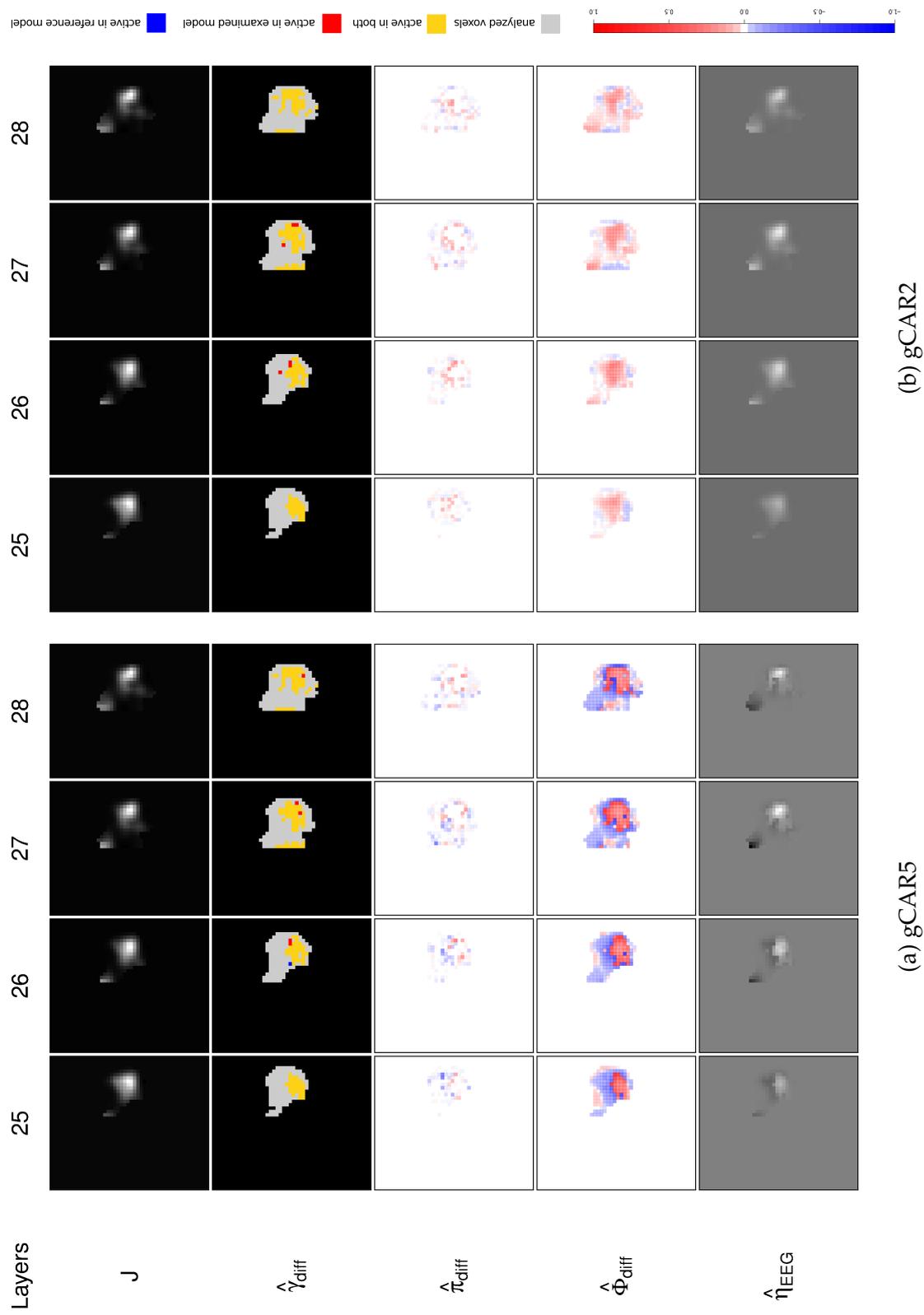


Figure 7.18: Subject-specific, time aggregated sLORETA map: Summary of results from the gCAR5 (left) and gCAR2 (right) model in comparison to gCAR3 (selected brain layers: 25–28).

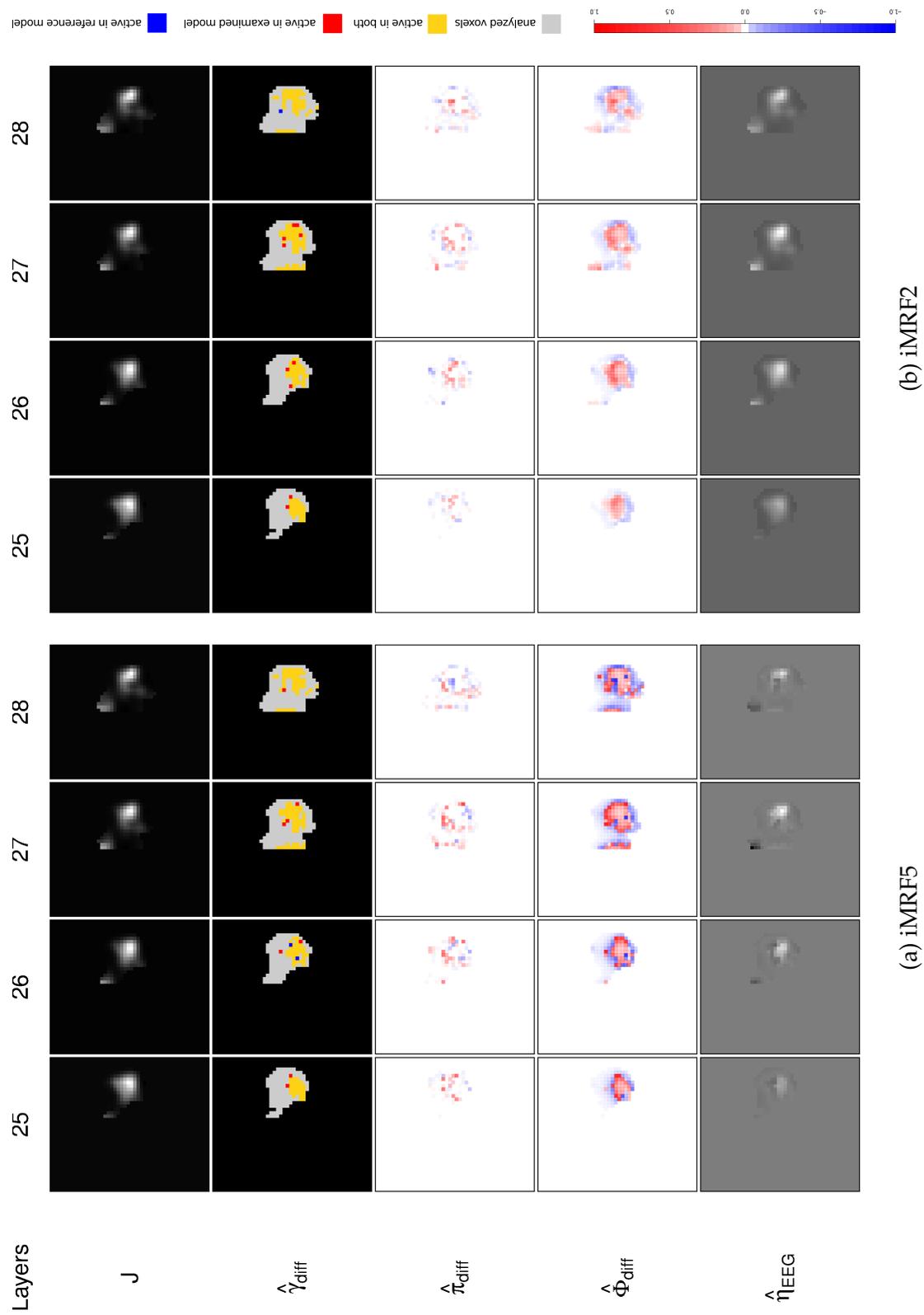
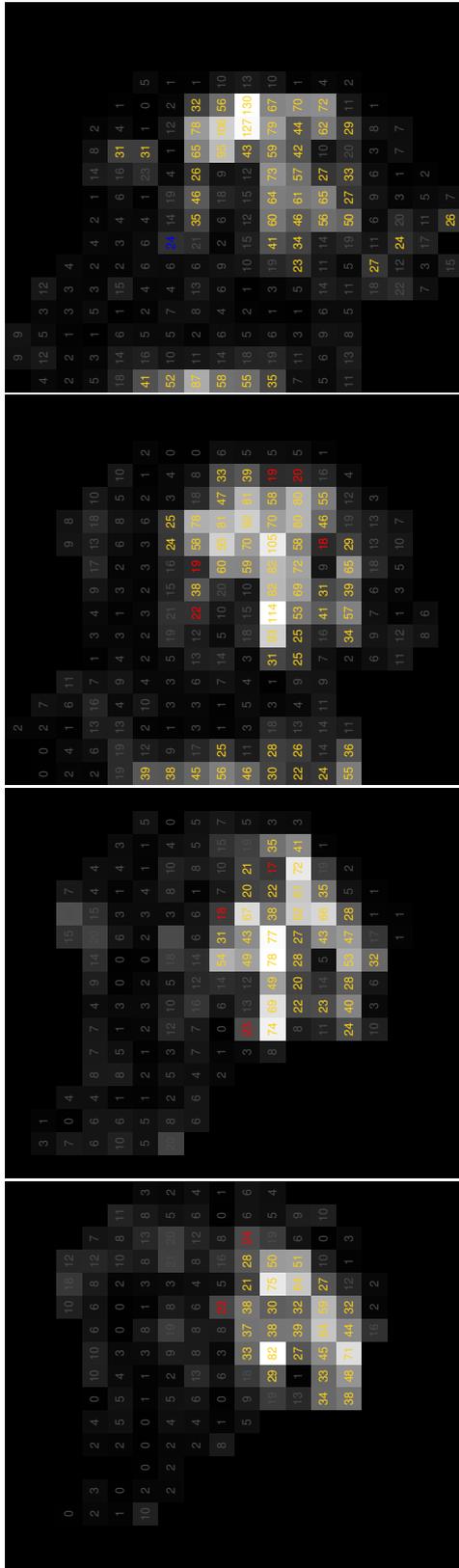
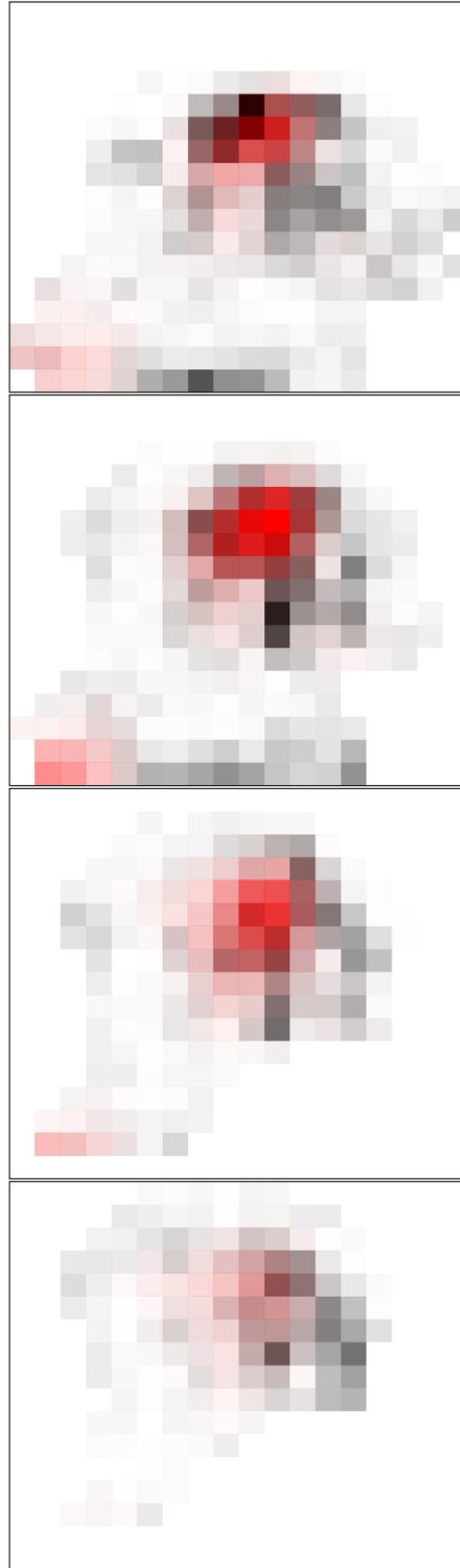


Figure 7.19: Subject-specific, time aggregated sLORETA map: Summary of results from the iMRF5 (left) and iMRF2 (right) model in comparison to iMRF3 (selected brain layers: 25–28).



(a) $\log(\text{LR})$ statistic map of fMRI data smoothed with kernel $2 \times 2 \times 2 \text{ mm}^3$ FWHM.



(b) Overlap of fMRI (gray-scale, in form of $\log(\text{LR})$ values) and EEG (red-scale) information.

Figure 7.20: Informational content of combined fMRI and EEG dataset.

at targeted regions where fMRI shows unexplainable gaps in its activation structure. If no EEG-based prior map can be found that is congruent to fMRI information in whole-brain analysis, analysis must be restricted to corresponding regions of interest.

We see some potential in using predictor type 5 algorithms with both spatially-varying intercept and spatially-varying EEG effect. The fCAR5 model could help to find activation in coarse activation structures by being able to recover some isolated active voxels. Whether these voxels were rather false than true positives has to be decided by brain research experts. The iMRF5 and gCAR5 algorithms used supporting EEG information to strengthen existing pronounced activation structures and smoothed out probably erroneous activation peaks in brain background. Thereby, the iMRF5 algorithm should be preferred, because its effect map estimates appeared to be more stable and it had a slightly better performance. Concerning the size of the sensitivity increase, we observed that the number of additional voxels found to be active was small. There were also some voxels becoming non-activated when EEG-prior information was included. Whether this can be traced back to the fact that the prior is also informative for non-activation remains to be clarified. If there is interest in examining the overlap between EEG and fMRI data, predictor 5 EEG map estimates can give information about this by highlighting congruent and non-congruent areas through positive and negative effect regions, respectively. All algorithms appeared to have sufficient edge preserving properties. It seems that it is disadvantageous when congruent areas adjoin non-congruent areas. That is, an algorithm that is able to retrieve such sharp edges is likely to be hardly able to differentiate between actual boundaries and neighboring voxels with a noisy activation signal. Hence, it seems to be necessary that congruent and non-congruent areas are well separated like in the artificial dataset in Chapter 6. Generally, we saw that EEG-enhanced fMRI activation schemes could increase sensitivity, but they had not lived up to expectations.

8 Proof of concept

EEG-enhanced fMRI activation detection algorithms did not achieve a substantial gain in sensitivity in the application study described in Chapter 7. In Chapter 6, however, they had a promising performance in artificial data settings. There are some reasons that might explain these findings. In the following, we motivate some requirements for successful EEG-enhanced fMRI activation detection and test these by looking at results from a modified oddball dataset. The study presented in this chapter is, thereby, conceptually a proof of concept of (a subset of) proposed algorithms in realistic data settings and not a rigorous test of hypotheses about sufficient conditions for EEG-enhancement.

From the simulation studies described in Section 6.1.5, we know that the global algorithms (iMRF2/5 and gCAR2/5) can increase sensitivity by the use of EEG-prior information at medium fMRI data noise levels. These settings are characterized by a smooth underlying activation structure that has been distorted by noise—though the smooth structure is still recognizable. From the findings of the real data scenario simulation study (cf. Section 6.2), we know that voxels at a log-likelihood ratio ($\log(\text{LR})$) statistic level of about 20 can additionally be found active. Thereby, we use voxelwise $\log(\text{LR})$ statistic values—for voxelwise hypothesis tests for testing whether the stimulus regressor component is 0—as a characterization of activation information contained in the fMRI data. Figure 8.1 depicts the voxelwise $\log(\text{LR})$ statistic values of the R_3 region of the artificial simulation datasets (cf. Section 6.1.1). Note that the gray-scale colors of the voxel areas and annotated numbers indicate the size of the $\log(\text{LR})$ value. The font color of these values is yellow if both the iMRF3 and iMRF5 algorithm classify a voxel as active. In analogy to the color code in activation difference maps presented in preceding chapters, red font color indicates that iMRF5 finds activation not found by iMRF3; blue font color indicates iMRF3 activation not found by iMRF5. In Figure 8.1, we see that above mentioned requirements for a successful application of EEG-enhanced schemes are fulfilled in medium noise levels, which leads to a substantial sensitivity increase in these settings.

When looking at the data structure of the smooth fMRI dataset (smoothed by a gaussian kernel with $8 \times 8 \times 8 \text{ mm}^3$ FWHM) from Chapter 7 in Figure 8.2, we notice that several

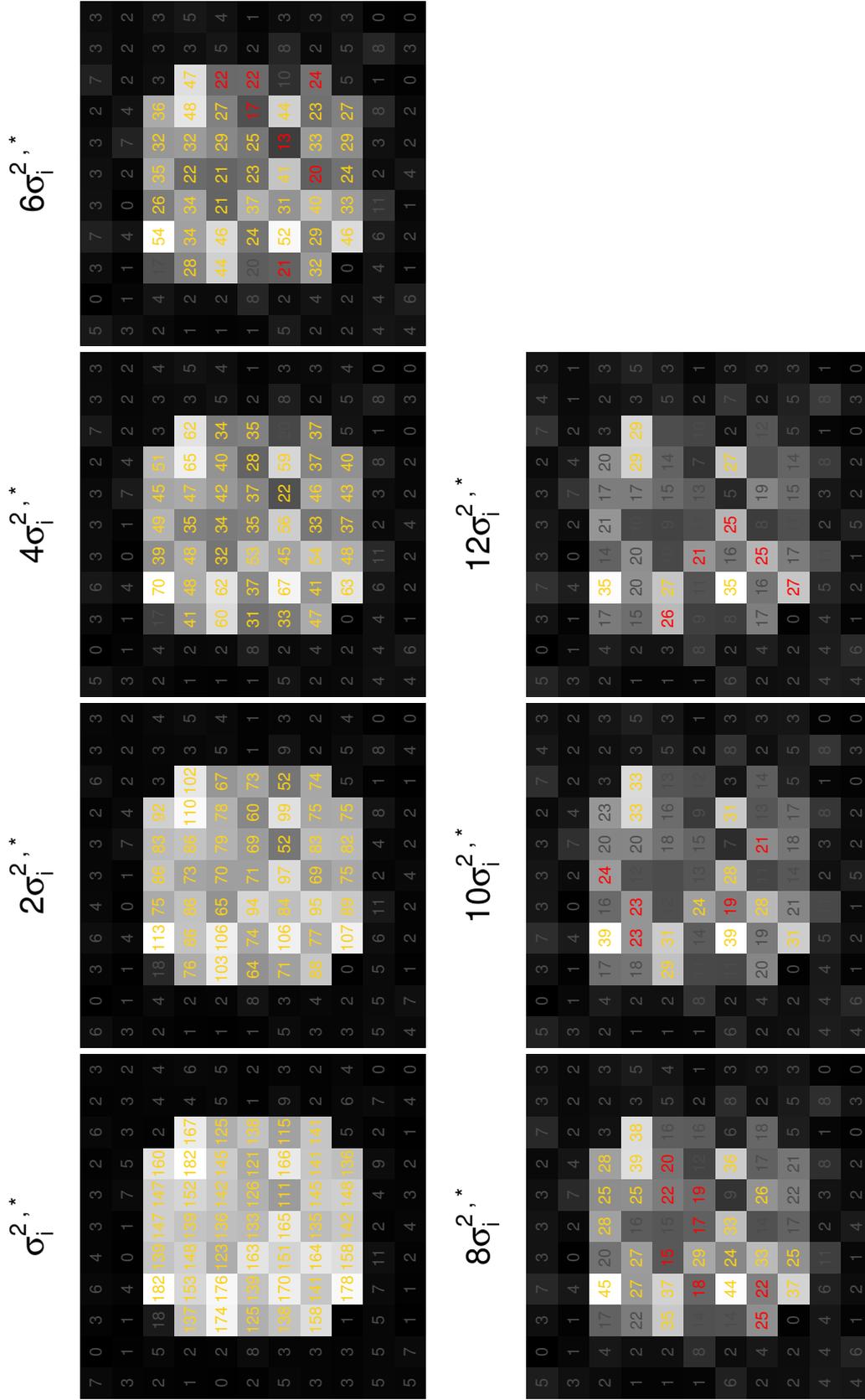


Figure 8.1: log(LR) statistic maps (R_3 -region only) with activation information (iMR5 vs. iMR3) from the simulation discussed in Section 6.1.5.

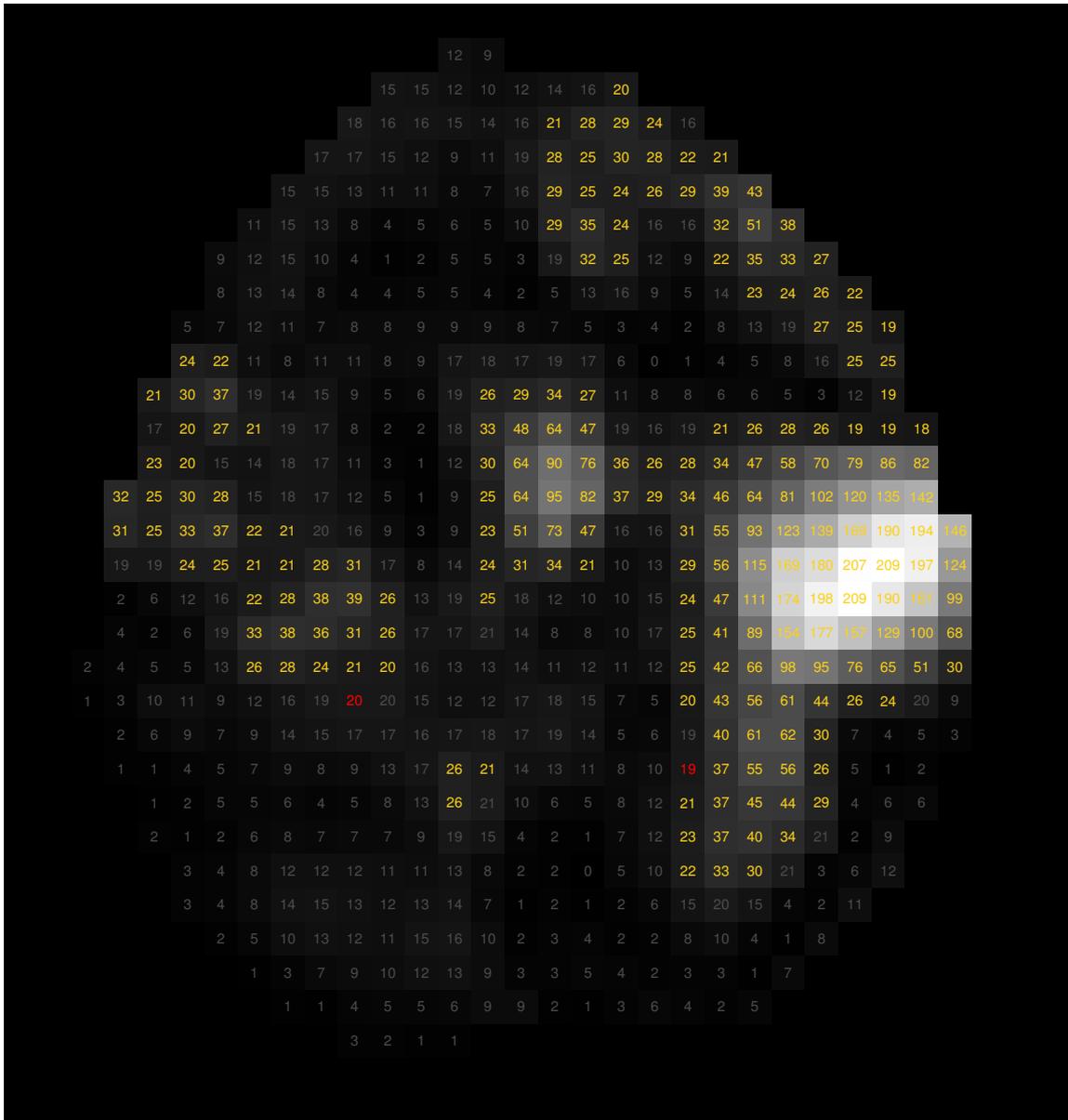


Figure 8.2: log(LR) statistic map (layer 29) with activation information (iMRF5 vs. iMRF3) for the application discussed in Section 7.

properties may be disadvantageous for a performance demonstration of the EEG enhancement: The data exhibits well pronounced, smooth activation peaks. The interesting $\log(\text{LR})$ statistic values of about 20 are only observed at some locations at activation borders—where there is a balance between active and non-active neighbors. This balance may cause that an additional number of active voxels can hardly be found there even if prior EEG information supports activation at borders. This also indicates why performance demonstration in the simulation study based on the real data setting (Section 6.2) was rather disappointing. Though the noise level was increased in some runs, the smooth peak structure of the data was preserved for affecting mainly voxels on region borders. Lacking suitable neighborhood information, this made the use of EEG-prior information dispensable.

Considering all these findings, we hypothesize that the data should possess the following properties for a successful application of EEG-enhanced fMRI activation detection algorithms that provoke a local increase in sensitivity. For the fMRI activation structure:

- The underlying activation structure contains activity regions of a reasonably large size (not just small isolated peaks).
- Some of these large-scale activation regions are distorted by noise leading to a perforated activation pattern.
- Activation gaps comprise voxels with a $\log(\text{LR})$ statistic value of about 20.

In the analyses of preceding chapters, we have seen that it is also necessary that additional conditions for the prior EEG information structure have to be satisfied. That is:

- Prior EEG information must indicate activation at targeted locations.
- For predictor type 2, models the prior EEG information must be fairly congruent to fMRI activation information overall the analyzed brain region.
- For predictor type 5 models, regions with congruent activation information should not adjoin regions with non-congruent activation information.

Additionally, we suppose:

- Prior EEG information must be sufficiently strong.

To evaluate whether realistic datasets that fulfill (most of these) conditions are suitable for a successful application of EEG-enhanced schemes, we modify the application dataset from Chapter 7 in such a way that it possesses above described properties. To decrease

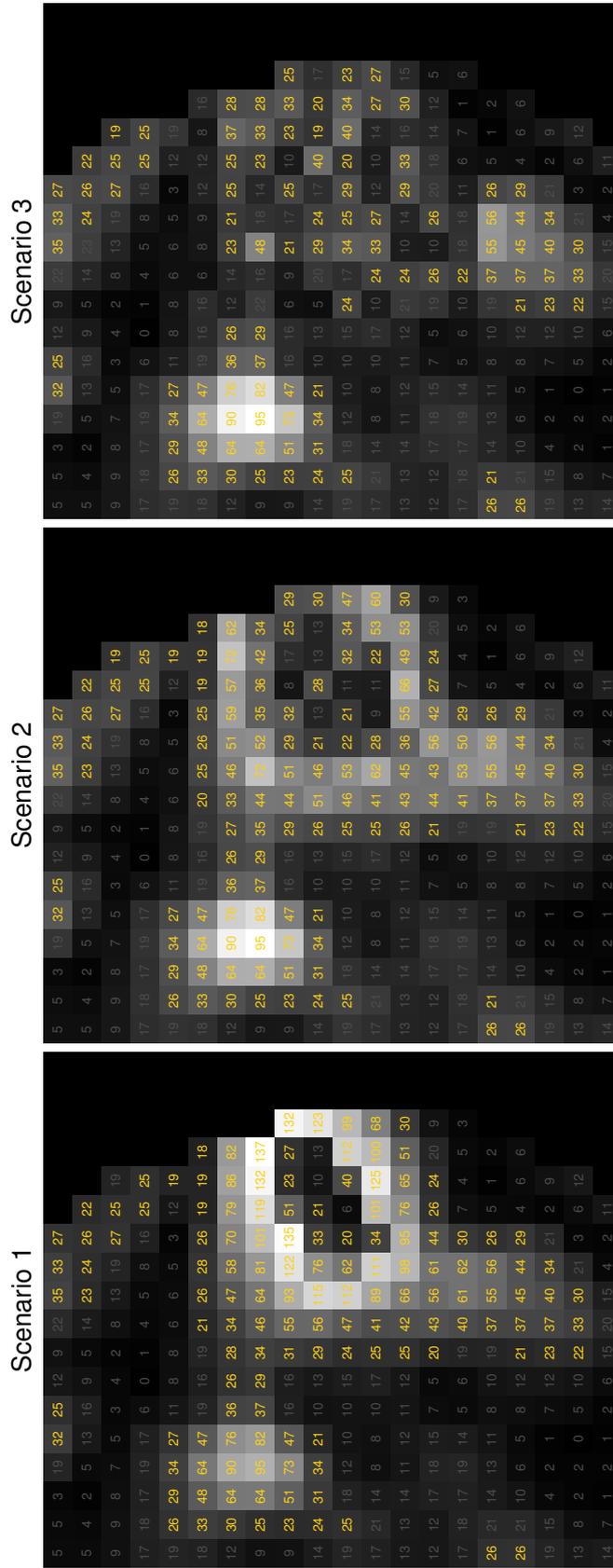


Figure 8.3: log(LR) statistic maps for the three examined noise scenarios with activation information (iMRFS).

computation time, we take a subset of the data ($x \in [15, 38]$, $y \in [14, 33]$ and $z \in [22, 33]$) incorporating the (neurologically) right part of Brodmann areas 1-4, 6 (containing the primary somatosensory, primary motor and premotor cortex), which showed a strong fMRI signal in prior analyses. Then, we use three different strategies to add noise to the observed activation structure (cf. Figure 8.3). Scenario 1 contains a hole within the activation structure with surrounding strongly activated voxels. Scenario 2 contains a slightly larger hole, but surrounding voxels obtain much smaller log-likelihood ratio values than in scenario 1. In scenario 3, the whole activation region is distorted by noise. As artificial prior EEG information, we use two congruent information maps in form of the binary extended fMRI activation map as described in Section 7.2 (denoted as binary prior) and the log(LR) statistic map (denoted as log(LR)-statistic prior) from the original dataset without additional noise. As a form of local prior information, we additionally use the log(LR) statistic map that is set to zero outside a rectangular region of 8×9 voxels covering the targeted activation focus across all slices (denoted as cut log(LR)-statistic prior). This prior introduces non-congruency in large parts of the analyzed brain, whereupon non-congruent regions adjoin the targeted activation focus with congruent information at some borders.

Figure 8.4 depicts the results of all three data scenarios with binary prior. Some voxels at the borders of the activation region change their state (mainly from non-active to active), but no changes occur in the targeted region inside the activation focus. This supports the hypothesis that if the EEG-prior information is not strong enough (like with this binary prior), EEG-enhanced activation schemes cannot make use of it.

In Figure 8.5, the results with log(LR)-statistic prior are displayed. We notice that each algorithm has a similar performance across all data scenarios. Both the iMRF2 and iMRF5 algorithm perform very well compensating the lost activation in any case. The fCAR2 and gCAR2 algorithm also find most activation in targeted regions, but these are not as sensitive as the iMRF algorithms at region borders. This indicates that the CAR prior possesses better edge-preserving properties, which might interfere with the aim to generally increase sensitivity. In contrast to this, the iMRF prior inclines to smooth across these edges (compare Chapter 6.1.3). Despite the good performance of corresponding predictor type 2 models, fCAR5 and gCAR5 degrade the activation structure by additionally deactivating formerly active voxels. Hence, fCAR5 and gCAR5 appear to be not applicable. They might suffer from identifiability issues.

In Figure 8.6, the results with the cut log(LR)-statistic prior are shown. Whereas all predictor type 2 models (fCAR2, gCAR2, iMRF2) perform well in data scenarios 1 and

2, they cannot compensate the loss in data scenario 3—indicating that if the activation structure is too noisy, non-congruent EEG information cannot compensate the loss. This might be traced back to the fact that in the non-congruent setting at hand, there is not enough information in the data to achieve a sufficiently large global EEG effect estimate. Similar to the findings with congruent $\log(\text{LR})$ -statistic prior, the fCAR5 and gCAR5 deteriorate the activation structure in each scenario. This also happens with the iMRF5 algorithm in data scenario 3.

In data scenario 1 and 2, the iMRF5 algorithm indeed finds all activation in targeted regions with the cut $\log(\text{LR})$ -statistic, but all other activation is lost in regions where prior values are zero. In this case, it seems that either iMRF5 uses zero EEG measurements as information on non-activation or that some kind of overcompensation happens. When looking at the estimates (not shown here), we notice a strongly negative global intercept estimate that might be appropriate in the inside of the activation focus—where the congruent EEG effect field is operative. However, outside the region the spatially-varying effect fields are not able to adapt to local activation structures and to compensate this negative estimate. This behavior might hint at identifiability issues that affect the iMRF5 algorithm in reasonably non-congruent settings.

Considering all results, we come to the following conclusions. The three data scenarios were suitable to prove that selected EEG-enhanced activation detection algorithms are useful under certain conditions. Hence, above listed data properties seem to constitute a basis for a successful application of proposed algorithms. At last, we can give more precise advice on favorable conditions for a local increase in sensitivity. For the fMRI data structure still holds:

- The underlying activation structure contains activity regions of a reasonably large size (not just small isolated peaks).
- Some of these large-scale activation regions are distorted by noise leading to a perforated activation pattern. It is advantageous when the smooth structure is preserved to some degree.
- Activation gaps comprise voxels with a $\log(\text{LR})$ statistic value of about 20.

For the EEG-prior information structure:

- Prior EEG information must indicate activation at targeted locations.
- Prior EEG information is preferably on a continuous scale indicating activation with high peak values, i.e. prior EEG information is sufficiently strong.

- Especially in higher noise settings, prior EEG information must be fairly congruent to fMRI activation information overall the analyzed brain region. This is a prerequisite for the use of all predictor types.
- For predictor type 5 models, particularly, congruent regions should not adjoin non-congruent regions.

If these properties can be observed in applicational data, predictor type 2 models can help to identify activation covered by noise. We suggest to proceed with caution with the predictor type 5 models. If one of these should be used, e.g. if it seems indispensable to adapt to local brain response, iMRF5 might be a suitable choice.

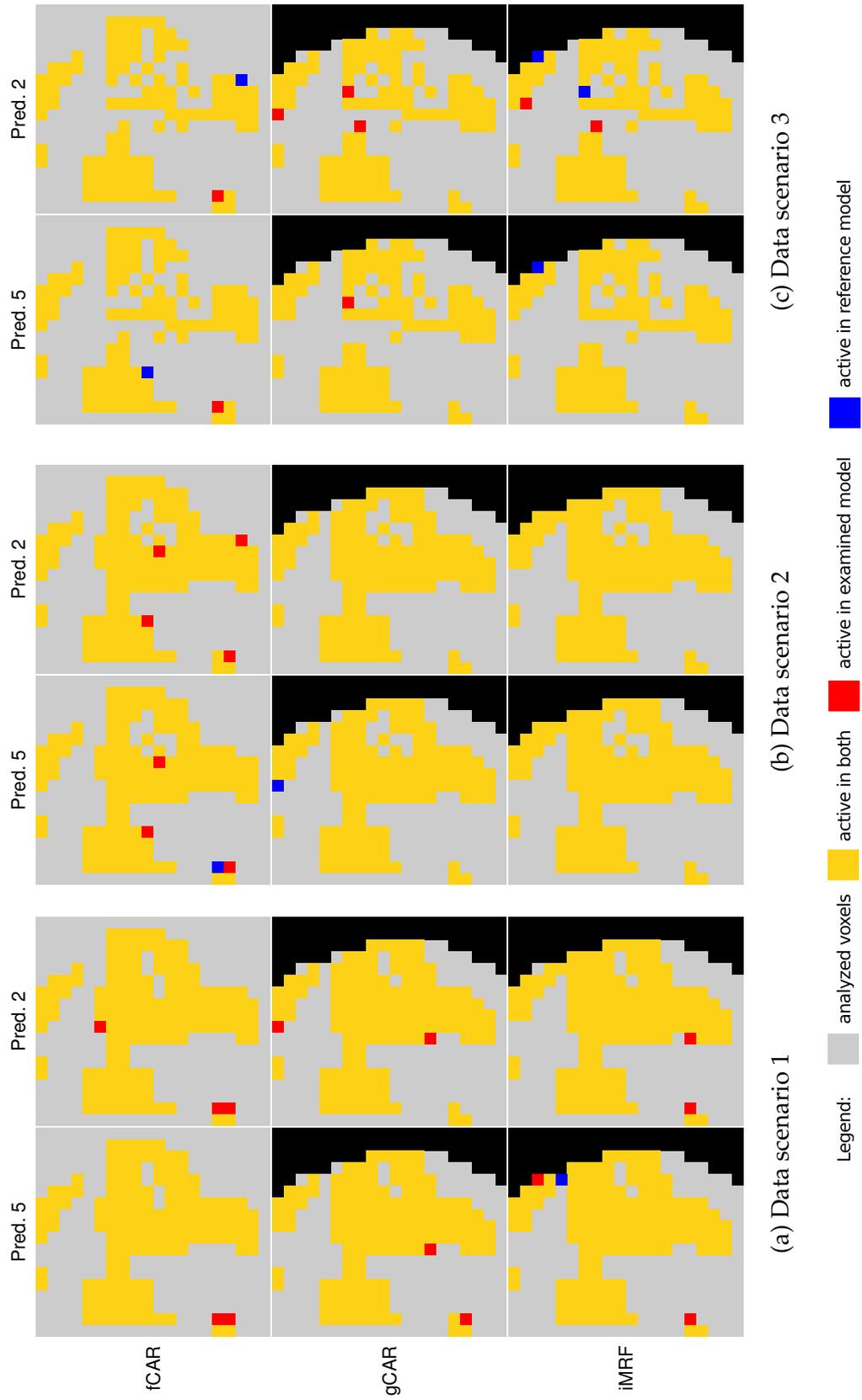


Figure 8.4: Binary prior: Activation difference maps in reference to corresponding predictor 3 algorithms.

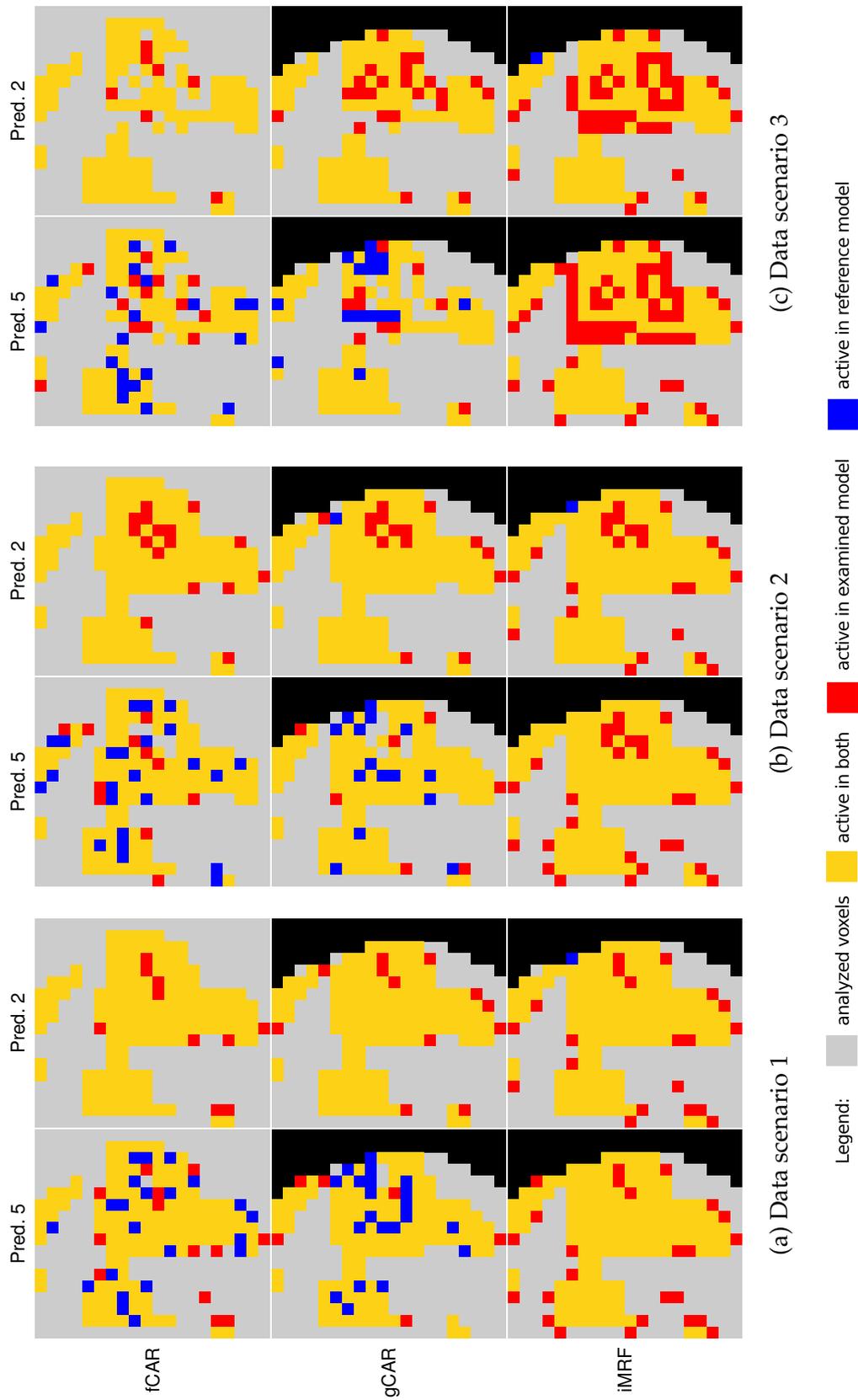


Figure 8.5: $\log(\text{LR})$ -statistic prior: Activation difference maps in reference to corresponding predictor 3 algorithms.

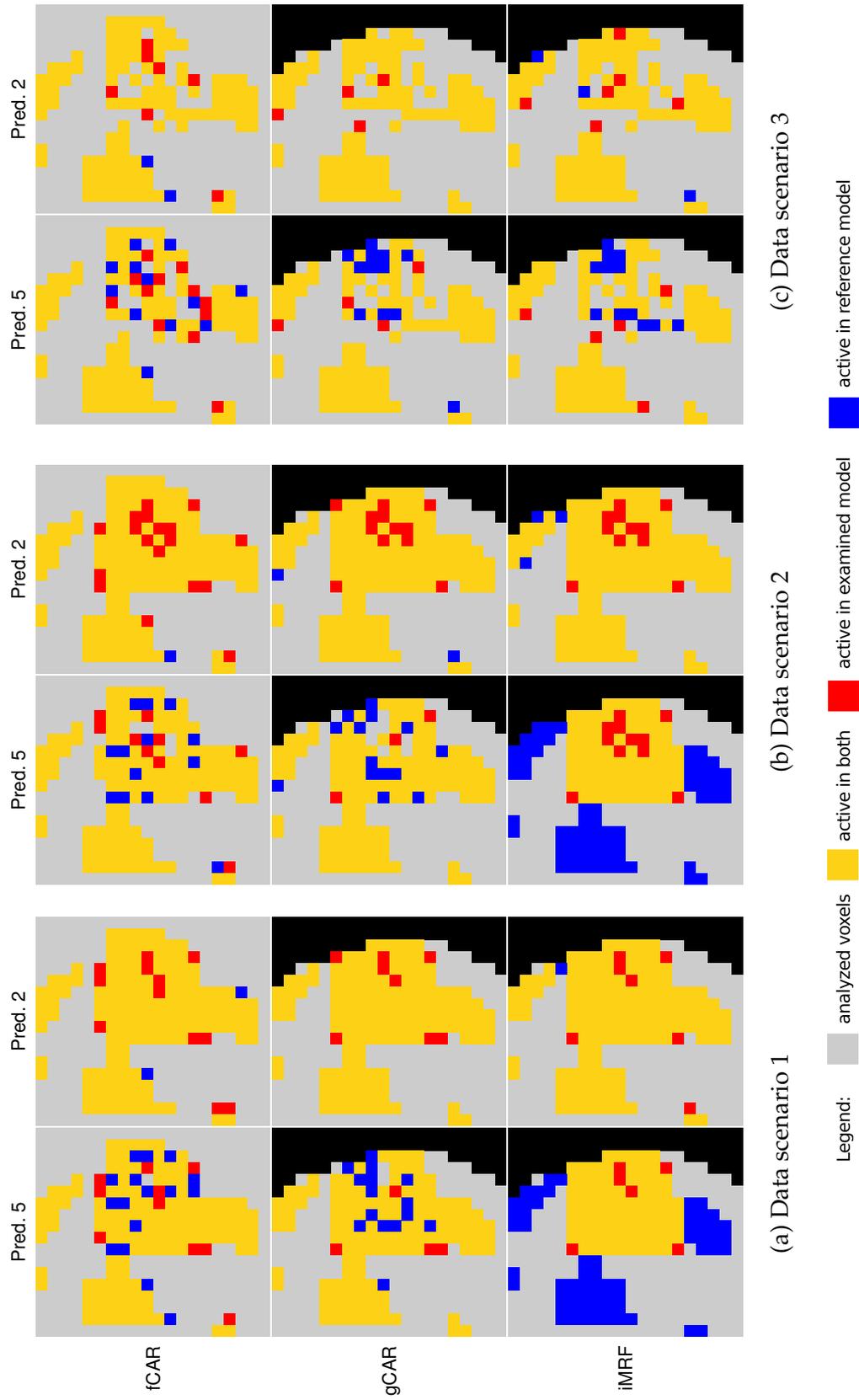


Figure 8.6: Cut $\log(\text{LR})$ -statistic prior: Activation difference maps in reference to corresponding predictor 3 algorithms.

9 Conclusion and Perspective

Multimodal neuroimaging approaches shed light on different dynamical and structural aspects of neuronal activity. Hence, data-fusion strategies seem indispensable for capturing its complex nature. In recent years, the combination of fMRI and EEG data has been a matter of particular interest for neuroscientists. After simultaneous recording techniques were firmly established, new statistical approaches have been developed to fuse information contained in both data types. Combining fMRI with EEG is a field of ongoing research.

In this work, we proposed a novel strategy for the fMRI-EEG data fusion—unequal to any fusion approach found in the literature. We developed an EEG-enhanced fMRI activation detection algorithm that is related to asymmetrical EEG-to-fMRI approaches (Rosa et al., 2010). Current EEG-to-fMRI approaches incorporate non-spatial EEG information in form of predictor variables in mass-voxelwise general linear models to localize EEG generators. Unlike these, we extended the Bayesian fMRI activation detection approach from Smith et al. (2003) to incorporate spatial EEG information in form of prior activation information to enhance sensitivity. More precisely, a high-dimensional Bayesian variable selection approach was used to relate spatial EEG-prior information to voxelwise selection probabilities of a global stimulus regressor component. These probabilities were then used as indicators for activation, i.e. as activation probabilities. Continuous or binary EEG information was thereby connected to activation probabilities by using a latent probit regression stage. The probit predictor was chosen to consist of at least one spatially-varying effect (intercept or EEG component) to adapt to local brain response. Spatially-varying effects were regularized by either an intrinsic Gaussian Markov random field (IGMRF) prior or a Gaussian conditional autoregressive (CAR) prior to ensure identifiability of voxelwise effects and to impose a dependency structure on neighboring voxels. In this thesis, inference was based on a Markov chain Monte Carlo (MCMC) approach relying on global or single-site updates of effect maps. The algorithms with global updates were denoted as gCAR and iMRF, respectively. Additionally, a faster algorithm for the CAR prior choice was proposed based on single-site updates circumventing the computationally

intensive, high-dimensional, sparse Cholesky decompositions. This algorithm was denoted as fCAR.

Theoretical considerations led to the choice of five different predictor types (depending on whether intercept or EEG component was included—and if so—as global or spatially-varying effect). The resulting algorithms were denoted as gCAR1–gCAR5, fCAR1–fCAR5 and iMRF1–iMRF5. This set of algorithms included uninformed fMRI activation detection schemes without EEG-prior information, which rely on a spatially-varying intercept only (gCAR3, fCAR3 and iMRF3). A major part of this thesis was concerned with the building process of a model suitable for EEG-enhanced fMRI activation detection. An extensive simulation study on an artificial dataset was conducted to base model building decisions upon the results of data arising from a controlled simulation design. Algorithms were assessed in terms of interpretability, possible adverse side effects and convergency properties. Afterwards, the performance of selected models was examined in both simulated settings and a real-world example. On the one hand, fMRI detection schemes (without EEG) were compared to existing models, i.e. Bayesian activation detection regularized by the Ising prior (Smith et al., 2003) and SPM (<http://www.fil.ion.ucl.ac.uk/spm/software/>, version SPM8). On the other hand, selected EEG-enhanced models were compared to their corresponding reference (without EEG-prior) to evaluate their use. Data from the real-world example was recorded during an auditory oddball experiment. Corresponding EEG-prior information was available in form of either an individual, an individual time-aggregated or a group- and time-aggregated source distribution map. Time series of source distribution maps were calculated by the sLORETA software (<http://www.uzh.ch/keyinst/loreta.htm>). Finally, we assessed the performance of EEG-enhanced activation schemes in an additional study with slightly modified fMRI data and artificial congruent EEG data from the oddball study to identify data constellations that are most suitable for EEG-enhanced fMRI activation detection.

Considering all results, we arrived at the following conclusions: The performance of fMRI activation detection algorithms without additional EEG information turned out to be impressive. Both the algorithms with global updates (gCAR3 and iMRF3) and the fast, single-site CAR algorithm (fCAR3) outperformed existing activation detection algorithms like the Ising algorithm and classical SPM by far. Thereby, findings did not indicate that increased sensitivity goes at the expense of specificity. Minor differences between proposed models exist that led to slightly different recommendations on when to use one of them. The fCAR3 algorithm performed best when the underlying activation structure

is rather coarse or signals are damped by a rather low signal-to-noise ratio. The iMRF3 and gCAR3 algorithms achieved higher sensitivity values when activation effects were relatively smooth. In contrast to fCAR3, iMRF3 and gCAR3 generally yielded estimates with higher smoothness and were, thus, able to better accommodate the spatial structure. When using these, the iMRF3 algorithm should be preferred for having a slightly better sensitivity than gCAR3.

Arriving at recommendations for EEG-enhanced fMRI activation detection was not as clear-cut as with the algorithms without additional prior information. Considering all results, we observed that the iMRF algorithms consistently had the most promising performance, because they appeared to recover the most reliable estimates by retaining a high sensitivity level. Performance of the iMRF algorithms was always superior to the gCAR algorithms, which even suffered from instability issues like implausible EEG effect map estimates. The fCAR algorithms seemed unsuitable for EEG-enhanced fMRI activation detection in many analysis runs for hardly being able to increase sensitivity compared to fCAR3 by using EEG-prior information. However, if the data structure was quite coarse, the fCAR3 algorithm possessed the best performance of all algorithms. This might indicate that fCAR algorithms have superior edge-preserving properties. However, we suppose that strong edge-preserving properties might interfere with the ability to enhance sensitivity by the use of EEG prior information. The combination of sufficiently strong EEG information and the incorporation of neighborhood information (regulated by the degree of estimated smoothness) appears to be a prerequisite for a local sensitivity increase.

Two types of predictors emerged as suitable choices: On the one hand, predictors consisting of both a spatially-varying intercept and a spatially-varying EEG effect, i.e. predictor type 5 models, and, on the other hand, predictors consisting of a spatially-varying intercept and a non-negative global EEG effect, i.e. predictor type 2 models.

Based on all findings presented in this thesis, we suppose that predictor types and according predictor components possess the following basic properties. For both predictor types, the spatially-varying intercept captures the brain-wide activity profile contained in the fMRI data.

In case of predictor type 5, the spatially-varying EEG coefficient regulates the EEG influence on activation detection and adapts it to local conditions. That is, it increases EEG influence in brain parts where EEG and fMRI are in accordance with each other and downweights EEG influence where it contradicts the fMRI signal. Hence, the algorithm is quite robust against misspecifications in EEG-prior choice. Nevertheless, a requirement for increasing

sensitivity in regions of special interest seems that these are not too close to regions with non-overlapping fMRI and EEG information. If these areas are distant enough, the downweighting mechanism does not interfere with the positive EEG influence in other regions. Thereby, the corresponding minimal distance depends on the degree of smoothness in the data and is likely to span only a small number of voxels. To assess congruency, EEG effect map estimates can give information about the overlap of fMRI and EEG data and, hence, give hints on possible problems.

In case of predictor type 2, the global EEG effect possesses the ability to increase prior activation probabilities proportionally to the observed EEG measurements as soon as a reasonable large, positive EEG effect is retrieved. If necessary, the spatially-varying intercept can have a region-specific compensating effect to improve the goodness-of-fit. The size of the EEG effect depends on the degree of congruency of fMRI and EEG data. If EEG information does not match fMRI information in large parts of the analyzed brain, a virtually zero global EEG effect estimate is recovered. Hence, this algorithm is equally robust against misspecifications, but, in contrast to predictor type 5 models, it loses its ability to increase sensitivity in remaining congruent areas.

Comparing the performance of both predictor types, we generally expect that predictor type 5 models possess a better performance than predictor type 2 models when non-congruency is present in the data. The iMRF2 algorithm retrieves a global EEG effect by averaging over all analyzed brain parts, so that the effect may be decreased (even to virtually zero estimates) if too much non-congruency is in the data. In contrast to this, iMRF5 is able to adapt the EEG effect to local conditions. As long as regions are sufficiently separated, effects in non-congruent areas do not substantially influence effects in congruent areas, which can retain their high sensitivity level. Generally, we observed that both predictor types are at least equally sensitive as its corresponding mere fMRI activation detection algorithms in most analyses.

Despite our expectations, EEG-enhanced fMRI activation detection schemes only revealed their benefit in data settings with certain properties. From these findings, we suppose that they are useful in regions where fMRI data possess activation structures of reasonable large size (not just isolated peaks) with unnatural gaps in the structure (e.g. due to noise). Within this region, non-active voxels containing a damped activation signal can then be classified as active if prior information—in form of sufficiently strong EEG measurements—indicates activation at these locations. Thereby the prior information must be fairly congruent to fMRI activation information overall the analyzed brain.

In our real-world example, the data structure deviated from these guidelines. Therefore, only a relatively small number of voxels switched their status from being non-active to active. Hence, we saw indeed that our EEG enhancement possesses some potential to increase sensitivity, but in the application at hand, it had not lived up to expectations. In controlled settings, like in simulation studies and analyses of modified real data, we observed, however, a substantial performance increase with data that possessed above properties. In the conducted artificial simulation study, iMRF5 could increase sensitivity in the order of 8.8%.

For further applications of the EEG-enhanced algorithms, we suggest to proceed as follows. Having applied an analysis run without EEG-prior (e.g. by use of fCAR3 or iMRF3), regions of interest have to be identified where neuronal activation appears to be damped by noise in the fMRI signal. Looking at the log-likelihood ratio statistic map for evaluating the significance of the stimulus regressor in the fMRI regression stage may be helpful for this. Then, an EEG (or any other non-redundant) prior information map has to be selected that supports the hypothesis of activation exactly at these locations. For EEG-enhanced activation detection, we suggest to use either the iMRF2 algorithm with non-negatively restricted global EEG effect or the iMRF5 algorithm. The user may start with an iMRF5 analysis. If sensitivity could not be substantially increased, the EEG effect map estimate may give hints on problems in the data. For example, positive effect areas being located directly besides negative effect areas indicate that non-congruent and congruent areas are too close. If the degree of congruency in the combined data is high, the iMRF2 algorithm can be applied. To ensure a high congruency level, analysis may be restricted to the region of interest to focus on those parts of the brain, where carefully selected EEG information is mostly congruent to fMRI information or rather supplements it. In this case, iMRF2 has the potential to increase sensitivity at targeted locations.

For researcher that want to use either the uninformed or EEG-enhanced fMRI activation scheme, we have implemented a user-friendly software library making all discussed Bayesian algorithms available. High diligence was put into arriving at a user-friendly, but computationally efficient software solution. The software is freely available as R package `Rfmri EEG`. Alternatively, C++ source code is provided for a binary `Cfmri EEGMain` program. Both kinds of software packages can be obtained from the author.

Summarizing, the approaches proposed in this thesis were based on a completely new idea on how to combine fMRI and EEG data. The use of a high-dimensional hierarchical model was indispensable for this, though no experience with this very specialized modeling approach existed at all. Hence, large parts of our work consisted of arriving at suitable

statistical model building decisions and solving problems arising in high-dimensional data settings like convergence issues and computational feasibility. Generally, we assert that the proposed models possess several good properties and can be considered as a good starting point for further developments.

Several minor modelling decisions may be questioned, like, e.g. the use of independent error terms in voxelwise fMRI regressions or the use of a g-prior for fMRI regression effects. It might be worth a try to implement alternatives to these choices, but we do not expect that the performance of proposed algorithms is strongly influenced. In particular, the exact form of fMRI regression model enters the latent EEG stage very indirectly—only in form of transformed voxelwise likelihood ratio statistics. Nevertheless, we are not entirely sure how these changes can affect the algorithms, because their derivation has to be adapted to alternative modelling choices. Marginalization steps to fMRI regression parameters may not be feasible anymore.

The most promising extension to our modeling approach is the incorporation of brain partitioning information (see, for example, the description in Appendix B of Daunizeau et al., 2007). Estimation of random fields and corresponding variance and spatial dependency parameters should be broken down into parts and be accomplished within (more) homogeneous brain areas. This is likely to resolve several problems. On the one hand, the speed of models with global updates can be increased substantially and convergence issues are likely to be resolved when reducing the high-dimensional dataset into smaller subgroups. On the other hand, the dependency resp. variance parameter estimation can be varied over partitions so that the model allows for different smoothness levels over the brain. The procedure would then incorporate a locally adaptive variance parameter estimation scheme. This implicates that we finally might arrive at a single model incorporating advantageous properties of the fast CAR algorithm (with the tendency to lower dependency level estimates and resulting better performance in coarse structures) and of algorithms with global updates (with the tendency to higher dependency level estimates and better performance in smooth structures). For EEG-enhanced fMRI activation approaches in form of predictor type 2 models, it seems likely that non-congruency issues can be resolved by this as well. If the constant (i.e. formerly global) EEG effect is estimated within partitions, it is no longer averaged over inhomogeneous areas, which has led to implausibly small effect sizes.

Another extension to EEG-enhanced fMRI detection schemes is the incorporation of a further tuning phase. After one of our algorithms has arrived at estimations for the latent probit effect estimates (after a predetermined number of MCMC iterations), it seems

conceivable to set effect maps constant and increase the corresponding (marginal) prior probability level by a selected multiplication factor. Then, an additional number of MCMC iterations, in which only the binary activation field is updated, can lead to a more sensitive solution—though taking account of the general activity structure. Whether this idea is generally applicable, remains to be tested.

The existing approaches can further be modified by changing the fitting algorithm. Though the full Bayesian MCMC approach finds the exact posterior distribution of model parameters, it is rather slow. Another popular approach in neuroimaging is a variational Bayes approach to Bayesian inference (see e.g. Titterton, 2004). It approximates the joint posterior distribution of all unknown parameters with a simpler distribution usually positing further independence assumptions than those implied by the original generative model. By this, the computational demand can be decreased substantially. However, a comparison between full and approximate Bayesian approaches has to be conducted to check whether the speed is achieved at the expense of accuracy.

A further point of criticism may be the restricted scope of application of our algorithms. In this thesis, we focussed on modeling activation provoked by just one stimulus type. Nevertheless, it is conceivable that other stimulus types, which are presented within the same experimental procedure, are either included as confounders or as a further additive stimulus component that is subject to Bayesian activation detection. However, so far, no strategy has been developed for testing general contrasts between incorporated stimulus types. Generally, our methodology is not applicable to stimulus-free experimental procedures like resting state observations. For being inevitably connected to a regressor modeling some kind of stimulus that provokes neuronal activation, an extension to stimulus-free settings is not conceivable.

As we have seen in an analysis of different fMRI time subsets, data may contain time-varying activation profiles. Hence, the assumption of time-constant activation effects—a standard assumption in most fMRI analyses—may be strongly violated. If checks (e.g. the analysis of different fMRI subsets) indicate deviations from this assumption, time-varying activation effects should be incorporated in the fMRI activation detection model. An example for a time-dependent resp. vigilance dependent fMRI stimulus effect at the fMRI regression stage can be found in Bothmann (2012). It seems also conceivable to incorporate time effects into proposed models. The latent probit predictor can, for example, be extended to include a time trend. The model, however, then would have to handle a time-varying latent binary activation field. It remains to be tested if the resulting model for spatio-temporal activation detection can handle this challenging situation.

At last, we question whether the specific application used to test proposed algorithms was suitable to show their strengths. We have seen that the analyst has to exercise reasonable care when selecting both the exact form of fMRI data and EEG-prior information. That is, there must be room for improvements in fMRI activation detection and, at the same time, EEG-prior information must both be strong (exactly where it is needed to improve the result) and mostly congruent with the analyzed fMRI data. In the auditory oddball study presented in Chapter 7, selecting a suitable combined dataset appeared to be problematic. On the one hand, the oddball paradigm produces well-defined activation peaks with reasonable size in smoothed fMRI data. There were few sites, with suitable medium signal-to-noise level ratio, showing great promise for EEG-enhanced activation detection. Hence, there might be other applications providing data structures better suited for EEG-enhanced activation detection. On the other hand, it was a challenging task to find matching 3D prior EEG information maps that were mostly congruent to fMRI data. There are several issues possibly introducing non-congruency. To list a few examples, first, both data types contain different informational content about neuronal activity; second, 3D EEG maps are prone to pre-processing errors introduced by source construction techniques like sLORETA (e.g. by localization errors and non-matching analyzing masks) and by other aggregation techniques like ICA (e.g. by splitting spatial information into independent components); third, if used, group-level EEG information does not necessarily match the subject-specific fMRI dataset. To adopt a rather optimistic attitude towards EEG-enhanced fMRI detection, we conclude that with other pre-processing options more suitable 3D EEG maps might be derived. Thereby, finding a suitable EEG-prior map is only important for gaining a performance increase. The fact, that EEG-prior information cannot be used to arbitrarily increase activation in any brain region hints at the robustness of our algorithms.

In the fascinating field of neuroscience, a vast variety of experimental strategies is available to investigate human brain functioning. Hence, there is no doubt that other applications exist that can benefit from EEG-enhanced activation detection. For these, we recommend a hypothesis-driven approach to data analysis. If algorithms are applied to carefully selected regions of interest, where a suitable combined dataset is at hand, these can bring out otherwise damped fMRI activation.

Appendix A

Derivation of further components of the EEG-enhanced fMRI model

A.1 Marginalized conditional distributions of binary indicators

A marginalized full conditional distribution for γ_i is based on the following factorization

$$\begin{aligned} p(\gamma_i | \gamma_{j \neq i}, \boldsymbol{\theta}, \mathbf{y}) &\propto p(\gamma_i, \gamma_{j \neq i}, \boldsymbol{\theta} | \mathbf{y}) = \int \int p(\gamma_i, \gamma_{j \neq i}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2 | \mathbf{y}) \partial(\boldsymbol{\beta}, \boldsymbol{\sigma}^2) \\ &= \int \int \prod_{i=1}^N p(\mathbf{y}_i | \gamma_i, \boldsymbol{\beta}_i(\gamma_i), \sigma_i^2) p(\boldsymbol{\beta}_i(\gamma_i) | \gamma_i, \sigma_i^2, \mathbf{y}_i) p(\sigma_i^2) \partial(\boldsymbol{\beta}, \boldsymbol{\sigma}^2) \\ &\quad \times p(\gamma_i | \gamma_{j \neq i}, \boldsymbol{\theta}) p(\gamma_{j \neq i} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) \\ &\propto \int \int p(\mathbf{y}_i | \gamma_i, \boldsymbol{\beta}_i(\gamma_i), \sigma_i^2) p(\boldsymbol{\beta}_i(\gamma_i) | \gamma_i, \sigma_i^2, \mathbf{y}_i) p(\sigma_i^2) \partial(\boldsymbol{\beta}_i, \sigma_i^2) p(\gamma_i | \gamma_{j \neq i}, \boldsymbol{\theta}) \\ &= p(\mathbf{y}_i | \gamma_i) p(\gamma_i | \gamma_{j \neq i}, \boldsymbol{\theta}). \end{aligned}$$

The integrals in $p(\mathbf{y}_i | \gamma_i)$ can be solved by noting a multivariate normal density in $\boldsymbol{\beta}_i$ and an inverse gamma distribution in σ_i^2 . This means in detail,

$$\begin{aligned} p(\mathbf{y}_i | \gamma_i) &\propto \int \int \left(\frac{1}{\sigma_i^2} \right)^{T/2} \exp \left(-\frac{1}{2\sigma_i^2} (\mathbf{y}_i - \mathbf{X}(\gamma_i) \boldsymbol{\beta}_i(\gamma_i))' (\mathbf{y}_i - \mathbf{X}(\gamma_i) \boldsymbol{\beta}_i(\gamma_i)) \right) \\ &\quad \times |T\sigma_i^2 (\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1}|^{-1/2} \exp \left(-\frac{1}{2\sigma_i^2} (\boldsymbol{\beta}_i(\gamma_i) - \hat{\boldsymbol{\beta}}_i(\gamma_i))' \frac{1}{T} \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) (\boldsymbol{\beta}_i(\gamma_i) - \hat{\boldsymbol{\beta}}_i(\gamma_i)) \right) \\ &\quad \times \frac{1}{\sigma_i^2} \partial(\boldsymbol{\beta}_i, \sigma_i^2) \end{aligned}$$

$$\begin{aligned}
&= |T(\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1}|^{-1/2} \int \int \left(\frac{1}{\sigma_i^2} \right)^{(T+q_i+2)/2} \\
&\quad \times \exp \left(-\frac{1}{2\sigma_i^2} [\mathbf{y}'_i \mathbf{y}_i - \mathbf{y}'_i \mathbf{X}(\gamma_i) \boldsymbol{\beta}_i(\gamma_i) - \boldsymbol{\beta}_i(\gamma_i)' \mathbf{X}(\gamma_i)' \mathbf{y}_i + \boldsymbol{\beta}_i(\gamma_i)' \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \boldsymbol{\beta}_i(\gamma_i)] \right) \\
&\quad \times \exp \left(-\frac{1}{2\sigma_i^2} \left[\boldsymbol{\beta}_i(\gamma_i)' \frac{\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i)}{T} \boldsymbol{\beta}_i(\gamma_i) - \hat{\boldsymbol{\beta}}_i(\gamma_i)' \frac{\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i)}{T} \boldsymbol{\beta}_i(\gamma_i) \right. \right. \\
&\quad \quad \left. \left. - \boldsymbol{\beta}_i(\gamma_i)' \frac{\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i)}{T} \hat{\boldsymbol{\beta}}_i(\gamma_i) + \hat{\boldsymbol{\beta}}_i(\gamma_i)' \frac{\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i)}{T} \hat{\boldsymbol{\beta}}_i(\gamma_i) \right] \right) \partial(\boldsymbol{\beta}_i, \sigma_i^2) \\
&= |T(\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1}|^{-1/2} \int \left(\frac{1}{\sigma_i^2} \right)^{(T+q_i+2)/2} \\
&\quad \times \exp \left(\frac{1}{2\sigma_i^2} \left[\mathbf{y}'_i \mathbf{y}_i + \frac{1}{T} \hat{\boldsymbol{\beta}}_i(\gamma_i)' \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \hat{\boldsymbol{\beta}}_i(\gamma_i) \right] \right) \\
&\quad \times \int \exp \left(-\frac{1}{2\sigma_i^2} \left[\boldsymbol{\beta}_i(\gamma_i)' \left(1 + \frac{1}{T} \right) \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \boldsymbol{\beta}_i(\gamma_i) \right. \right. \\
&\quad \quad \left. \left. - \boldsymbol{\beta}_i(\gamma_i)' \left(\mathbf{X}(\gamma_i)' \mathbf{y}_i + \frac{1}{T} \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \hat{\boldsymbol{\beta}}_i(\gamma_i) \right) \right. \right. \\
&\quad \quad \left. \left. - \left(\mathbf{X}(\gamma_i)' \mathbf{y}_i + \frac{1}{T} \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \hat{\boldsymbol{\beta}}_i(\gamma_i) \right)' \boldsymbol{\beta}_i(\gamma_i) \right] \right) \partial(\boldsymbol{\beta}_i, \sigma_i^2),
\end{aligned}$$

where $q_i = \sum_{j=1}^p \gamma_{i,j}$ is the number of nonzero regressors in regression i .

The inner integral can be solved by recognizing a multivariate normal distribution in $\boldsymbol{\beta}_i(\gamma_i)$ with mean $\hat{\boldsymbol{\beta}}_i(\gamma_i)$ and covariance matrix $\sigma_i^2 \left(\left(1 + \frac{1}{T} \right) \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \right)^{-1}$. After completion with the inverse of amended terms, $p(\mathbf{y}_i | \gamma_i)$ can be rewritten as

$$\begin{aligned}
p(\mathbf{y}_i | \gamma_i) &\propto \int \left(\frac{1}{\sigma_i} \right)^{(T+q_i+2)/2} |T(\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1}|^{-1/2} \left| \sigma_i^2 \left(\left(1 + \frac{1}{T} \right) \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \right)^{-1} \right|^{1/2} \\
&\quad \times \exp \left(-\frac{1}{2\sigma_i^2} \left[\mathbf{y}'_i \mathbf{y}_i + \frac{1}{T} \hat{\boldsymbol{\beta}}_i(\gamma_i)' \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \hat{\boldsymbol{\beta}}_i(\gamma_i) \right] \right) \\
&\quad \times \exp \left(\frac{1}{2\sigma_i^2} \left[\hat{\boldsymbol{\beta}}_i(\gamma_i)' \left(1 + \frac{1}{T} \right) \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \hat{\boldsymbol{\beta}}_i(\gamma_i) \right] \right) \partial\sigma_i^2 \\
&= T^{-q_i/2} \left(1 + \frac{1}{T} \right)^{-q_i/2} \\
&\quad \times \int \left(\frac{1}{\sigma_i} \right)^{T/2+1} \exp \left(-\frac{1}{\sigma_i^2} \frac{1}{2} [\mathbf{y}'_i \mathbf{y}_i - \hat{\boldsymbol{\beta}}_i(\gamma_i)' \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \hat{\boldsymbol{\beta}}_i(\gamma_i)] \right) \partial\sigma_i^2.
\end{aligned}$$

The integral is solved by recognizing an inverse gamma distribution with shape parameter $T/2$ and scale parameter $S_i(\gamma_i)/2$, where

$$S_i(\gamma_i) = \mathbf{y}'_i \mathbf{y}_i - \hat{\boldsymbol{\beta}}_i(\gamma_i)' \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \hat{\boldsymbol{\beta}}_i(\gamma_i)$$

$$= \mathbf{y}'_i \mathbf{y}_i - \mathbf{y}'_i \mathbf{X}(\gamma_i) (\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1} \mathbf{X}(\gamma_i)' \mathbf{y}_i$$

is the sum of squares in regression i corresponding to subset γ_i . Thus,

$$p(\mathbf{y}_i | \gamma_i) \propto S_i(\gamma_i)^{-T/2} (1 + T)^{-q_i/2}.$$

A.2 Conditional posterior means of fMRI stage parameters

In this section, the conditional posterior means of the regression coefficients β_i and variance parameters σ_i^2 are derived.

The marginalized full conditional of β_i is calculated by

$$\begin{aligned} p(\beta_i | \gamma, \theta, \mathbf{y}) &\propto \int p(\beta, \sigma^2, \gamma, \theta | \mathbf{y}) \partial \sigma^2 \\ &\propto \int p(\mathbf{y}_i | \gamma_i, \beta_i(\gamma_i), \sigma_i^2) p(\beta_i(\gamma_i) | \gamma_i, \sigma_i^2, \mathbf{y}_i) p(\sigma_i^2) \partial \sigma_i^2 \\ &\propto p(\beta_i | \gamma_i, \mathbf{y}_i). \end{aligned}$$

This means that the full conditional distribution of β_i marginalized with respect to σ_i^2 is only dependent on γ_i (and \mathbf{y}_i). With this,

$$\begin{aligned} p(\beta_i | \gamma_i, \mathbf{y}_i) &\propto \int \left(\frac{1}{\sigma_i^2} \right)^{\frac{T}{2}} \exp \left(-\frac{1}{\sigma_i^2} \frac{1}{2} (\mathbf{y}_i - \mathbf{X}(\gamma_i) \beta_i(\gamma_i))' (\mathbf{y}_i - \mathbf{X}(\gamma_i) \beta_i(\gamma_i)) \right) \\ &\quad \times |T \sigma_i^2 (\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1}|^{-1/2} \\ &\quad \times \exp \left(-\frac{1}{\sigma_i^2} \frac{1}{2} (\beta_i(\gamma_i) - \hat{\beta}_i(\gamma_i))' \frac{1}{T} \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) (\beta_i(\gamma_i) - \hat{\beta}_i(\gamma_i)) \right) \\ &\quad \times \frac{1}{\sigma_i^2} \partial \sigma_i^2 \\ &\propto \int \left(\frac{1}{\sigma_i^2} \right)^{\frac{T+q_i}{2}+1} \exp \left(-\frac{1}{\sigma_i^2} \frac{1}{2} b \right) \partial \sigma_i^2, \end{aligned} \tag{A.1}$$

where $q_i = \sum_{j=1}^p \gamma_{i,j}$ is the number of nonzero regressors in regression i and

$$\begin{aligned} b &= \mathbf{y}'_i \mathbf{y}_i - \beta_i(\gamma_i)' \mathbf{X}(\gamma_i)' \mathbf{y}_i - \mathbf{y}'_i \mathbf{X}(\gamma_i) \beta_i(\gamma_i) + \beta_i(\gamma_i)' \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \beta_i(\gamma_i) \\ &\quad + \beta_i(\gamma_i)' \frac{1}{T} \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \beta_i(\gamma_i) - \beta_i(\gamma_i)' \frac{1}{T} \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \hat{\beta}_i(\gamma_i) \end{aligned}$$

$$\begin{aligned}
& -\hat{\beta}_i(\gamma_i)' \frac{1}{T} \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \beta_i(\gamma_i) + \hat{\beta}_i(\gamma_i)' \frac{1}{T} \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \hat{\beta}_i(\gamma_i) \\
& = S_i(\gamma_i) + (\beta_i(\gamma_i) - \hat{\beta}_i(\gamma_i))' \left(1 + \frac{1}{T}\right) \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) (\beta_i(\gamma_i) - \hat{\beta}_i(\gamma_i)) \quad (\text{A.2}) \\
& = 1 + (\beta_i(\gamma_i) - \hat{\beta}_i(\gamma_i))' \left(\frac{T+1}{TS_i(\gamma_i)}\right) \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) (\beta_i(\gamma_i) - \hat{\beta}_i(\gamma_i))
\end{aligned}$$

where $S_i(\gamma_i) = \mathbf{y}_i' \mathbf{y}_i - \hat{\beta}_i(\gamma_i) \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \hat{\beta}_i(\gamma_i) = \mathbf{y}_i' \mathbf{y}_i - \mathbf{y}_i' \mathbf{X}(\gamma_i) (\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1} \mathbf{X}(\gamma_i)' \mathbf{y}_i$, as before.

The form of conditional posterior in (A.1) can be completed to an inverse gamma distribution with shape parameter $\frac{T+q_i}{2}$ and scale parameter $\frac{1}{2}b$. After multiplying inverse terms and using a proportionality statement once more,

$$p(\beta_i | \gamma_i, \mathbf{y}_i) \propto b^{\frac{T+q_i}{2}} \Gamma\left(\frac{T+q_i}{2}\right),$$

which is a multivariate T-distribution with mean $\hat{\beta}_i(\gamma_i)$.

The marginalized conditional mean of σ_i^2 is derived by similar arguments as for β_i . First, we note that the full conditional for σ_i^2 is only dependent on β_i and γ_i to yield

$$\begin{aligned}
p(\sigma_i^2 | \gamma, \boldsymbol{\theta}, \mathbf{y}) & \propto \int p(\beta, \sigma^2, \gamma, \boldsymbol{\theta} | \mathbf{y}) \partial \beta \\
& \propto \int p(\mathbf{y}_i | \gamma_i, \beta_i(\gamma_i), \sigma_i^2) p(\beta_i(\gamma_i) | \gamma_i, \sigma_i^2, \mathbf{y}_i) p(\sigma_i^2) \partial \beta_i \\
& \propto p(\sigma_i^2 | \gamma_i, \mathbf{y}_i).
\end{aligned}$$

Like for β_i , it can be shown that

$$\begin{aligned}
p(\sigma_i^2 | \gamma_i, \mathbf{y}_i) & \propto \left(\frac{1}{\sigma_i^2}\right)^{\frac{T+q_i}{2}+1} |T(\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1}|^{-1/2} \int \exp\left(-\frac{1}{2\sigma_i^2} b\right) \partial \beta_i \\
& = \left(\frac{1}{\sigma_i^2}\right)^{\frac{T+q_i}{2}+1} |T(\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1}|^{-1/2} \exp\left(-\frac{1}{2\sigma_i^2} S_i(\gamma_i)\right) \\
& \quad \times \int \exp\left(-\frac{1}{2\sigma_i^2} (\beta_i(\gamma_i) - \hat{\beta}_i(\gamma_i))' \left(1 + \frac{1}{T}\right) \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) (\beta_i(\gamma_i) - \hat{\beta}_i(\gamma_i))\right) \partial \beta_i
\end{aligned}$$

where b is as in (A.2). The parameter β_i can be integrated out by noting a multivariate normal distribution with mean $\hat{\beta}_i(\gamma_i)$ and covariance matrix $\sigma_i^2 \left(\left(1 + \frac{1}{T}\right) \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i)\right)^{-1}$.

Hence,

$$\begin{aligned}
p(\sigma_i^2 | \gamma_i, \mathbf{y}_i) &\propto \left(\frac{1}{\sigma_i^2}\right)^{\frac{T+q_i}{2}+1} |T(\mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i))^{-1}|^{-1/2} \exp\left(-\frac{1}{2\sigma_i^2} S_i(\gamma_i)\right) \\
&\quad \times \left| \sigma_i^2 \left(\left(1 + \frac{1}{T}\right) \mathbf{X}(\gamma_i)' \mathbf{X}(\gamma_i) \right)^{-1} \right|^{1/2} \\
&\propto \left(\frac{1}{\sigma_i^2}\right)^{\frac{T}{2}+1} \exp\left(\frac{1}{\sigma_i^2} \frac{1}{2} S_i(\gamma_i)\right).
\end{aligned}$$

This can be complete to an inverse gamma distribution with shape parameter $T/2$ and scale parameter $S_i(\gamma_i)/2$. Therefore, $E(\sigma_i^2 | \gamma_i, \mathbf{y}_i) = S_i(\gamma_i)/(T - 2)$.

A.3 Marginal prior probability for activation

In the following, we calculate the marginalized prior probability of \mathbf{U} given variance parameters of probit effect priors: $p(\mathbf{U} | \xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2)$. For this, we have to integrate out effect parameters of the probit stage. For vectors of spatially-varying coefficients, we use a zero vector as prior mean. To show that $p(\gamma_i = 1) = p(\mathbf{U}_i > 0) = 0.5$, we have to set prior means of global parameters to 0 as well.

$$\begin{aligned}
p(\mathbf{U} | \xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2) &= \int p(\mathbf{U} | \boldsymbol{\theta}) p(\alpha_{0,G} | \xi_{0,G}^2) p(\alpha_G | \xi_G^2) p(\boldsymbol{\alpha}_0 | \xi_0^2) p(\boldsymbol{\alpha} | \xi^2) \partial(\alpha_{0,G}, \alpha_G, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}) \\
&= \left(\frac{1}{2\pi}\right)^{N/2} \int \exp\left(-\frac{1}{2}(\mathbf{U} - \alpha_{0,G}\mathbf{1} - \alpha_G\mathbf{j} - \boldsymbol{\alpha}_0\mathbf{I} - \boldsymbol{\alpha}\mathbf{J})'(\mathbf{U} - \alpha_{0,G}\mathbf{1} - \alpha_G\mathbf{j} - \boldsymbol{\alpha}_0\mathbf{I} - \boldsymbol{\alpha}\mathbf{J})\right) \\
&\quad \cdot \left(\frac{1}{2\pi}\right)^{1/2} \left(\frac{1}{\xi_{0,G}^2}\right)^{1/2} \exp\left(-\frac{1}{2\xi_{0,G}^2} \alpha_{0,G}^2\right) \\
&\quad \cdot \left(\frac{1}{2\pi}\right)^{1/2} \left(\frac{1}{\xi_G^2}\right)^{1/2} \exp\left(-\frac{1}{2\xi_G^2} \alpha_G^2\right) \\
&\quad \cdot \left(\frac{1}{2\pi}\right)^{(N-k)/2} \left(\frac{1}{\xi_0^2}\right)^{(N-k)/2} (|\mathbf{Q}|^*)^{1/2} \exp\left(-\frac{1}{2\xi_0^2} \boldsymbol{\alpha}'_0 \mathbf{Q} \boldsymbol{\alpha}_0\right) \\
&\quad \cdot \left(\frac{1}{2\pi}\right)^{(N-k)/2} \left(\frac{1}{\xi^2}\right)^{(N-k)/2} (|\mathbf{Q}|^*)^{1/2} \exp\left(-\frac{1}{2\xi^2} \boldsymbol{\alpha}' \mathbf{Q} \boldsymbol{\alpha}\right) \partial(\alpha_{0,G}, \alpha_G, \boldsymbol{\alpha}_0, \boldsymbol{\alpha})
\end{aligned}$$

where \mathbf{Q} is the precision matrix of spatial GMRFs and k its possible loss in rank. Let $\tilde{\boldsymbol{\alpha}} = (\alpha_{0,G}, \alpha_G, \boldsymbol{\alpha}_0, \boldsymbol{\alpha})$ be the concatenated vector of probit effects and $\tilde{\mathbf{X}} = (\mathbf{1}, \mathbf{j}, \mathbf{I}, \mathbf{J})$ the

according design matrix, then we can rewrite the integral as

$$\begin{aligned}
& p(\mathbf{U}|\xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2) \\
&= C_{tmp} \cdot \int \exp\left(-\frac{1}{2}(\mathbf{U} - \tilde{\mathbf{X}}\tilde{\boldsymbol{\alpha}})'(\mathbf{U} - \tilde{\mathbf{X}}\tilde{\boldsymbol{\alpha}})\right) \\
&\quad \cdot \exp\left(-\frac{1}{2}\left(\frac{\alpha_{0,G}^2}{\xi_{0,G}^2} + \frac{\alpha_0^2}{\xi_G^2} + \boldsymbol{\alpha}'_0 \frac{\mathbf{Q}}{\xi_0^2} \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}' \frac{\mathbf{Q}}{\xi^2} \boldsymbol{\alpha}\right)\right) \partial(\alpha_{0,G}, \alpha_G, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}) \\
&= C_{tmp} \cdot \int \exp\left(-\frac{1}{2}(\mathbf{U} - \tilde{\mathbf{X}}\tilde{\boldsymbol{\alpha}})'(\mathbf{U} - \tilde{\mathbf{X}}\tilde{\boldsymbol{\alpha}})\right) \exp\left(-\frac{1}{2}\tilde{\boldsymbol{\alpha}}'\tilde{\mathbf{Q}}\tilde{\boldsymbol{\alpha}}\right) \partial(\alpha_{0,G}, \alpha_G, \boldsymbol{\alpha}_0, \boldsymbol{\alpha})
\end{aligned}$$

where $\tilde{\mathbf{Q}} = \text{diag}(1/\xi_{0,G}^2, 1/\xi_0^2, \mathbf{Q}/\xi_0^2, \mathbf{Q}/\xi^2)$ and C_{tmp} is a factor not depending on \mathbf{U} and $\tilde{\boldsymbol{\alpha}}$. Expanding the exponent and amending with missing terms, a multivariate normal density in $\tilde{\boldsymbol{\alpha}}$ can be recognized:

$$\begin{aligned}
p(\mathbf{U}|\xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2) &= C(\xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2) \int \left(\frac{1}{2\pi}\right)^{\frac{2(N-k+1)}{2}} \left|(\mathbf{X}'\mathbf{X} + \tilde{\mathbf{Q}})\right|^{1/2} \\
&\quad \cdot \exp\left(-\frac{1}{2}\left[\tilde{\boldsymbol{\alpha}}'(\mathbf{X}'\mathbf{X} + \tilde{\mathbf{Q}})\tilde{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}}'\mathbf{X}\mathbf{U} - \mathbf{U}'\mathbf{X}\tilde{\boldsymbol{\alpha}} + \mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X} + \tilde{\mathbf{Q}})^{-1}\mathbf{X}'\mathbf{U}\right]\right) \partial(\alpha_{0,G}, \alpha_G, \boldsymbol{\alpha}_0, \boldsymbol{\alpha}) \\
&\quad \cdot \exp\left(\frac{1}{2}\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X} + \tilde{\mathbf{Q}})^{-1}\mathbf{X}'\mathbf{U}\right) \exp\left(-\frac{1}{2}\mathbf{U}'\mathbf{U}\right),
\end{aligned}$$

where $C(\xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2)$ depends on all remaining constant terms and variance parameters. Hence, the integral has value 1 and the marginalized density of \mathbf{U} given variance parameters is

$$p(\mathbf{U}|\xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2) = C(\xi_{0,G}^2, \xi_G^2, \xi_0^2, \xi^2) \exp\left(-\frac{1}{2}\mathbf{U}'\left(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X} + \tilde{\mathbf{Q}})^{-1}\mathbf{X}'\right)\mathbf{U}\right).$$

Appendix B

Robustness of variance parameter estimation

Convergency diagnostics of Bayesian activation detection schemes with a spatially-varying effect in a latent probit stage revealed difficulties in variance-covariance parameter estimation. Convergency of variance parameters ξ^2 and ξ_0^2 are strongly dependent on prior settings, i.e. inverse gamma shape and scale parameters. Convergency of spatial dependency parameters τ^2 and τ_0^2 in the models with CAR prior are dependent on the corresponding starting value and proposal variance of the Metropolis step. Although a tedious search for optimal hyperparameters has been conducted, some variance parameter trajectories still show mixing problems like lacking stationarity and strong autocorrelations (see Figure 5.3 for an exemplary gCAR5 trace plot). Convergency to an equilibrium distribution cannot be assessed for sure. Being nuisance parameters of the lowest hierarchical stage of our models, we decided to conduct a small robustness study to evaluate the influence these convergency issues have on remaining parameter estimates. Foremost, we are interested in influences on activation maps. Evaluation is based on the artificial dataset from Chapter 6. We focus on models with predictor type 5, i.e. the probit predictor with both spatially-varying intercept and spatially-varying EEG effect and use as a starting point the settings and results from Section 6.1.3.

B.1 Robustness of gCAR5 results

The robustness study for the gCAR5 model is conducted as follows. We take the variance-covariance parameter estimates from gCAR5 runs in Section 6.1.3. As estimates $\hat{\xi}^2, \hat{\tau}^2, \hat{\xi}_0^2$

and $\hat{\tau}_0^2$, we take the mean of thinned parameter trajectories (with stepping equal to 5) after a burnin of 1 000 iterations of a total of 6 000 MCMC iterations. These estimates are assigned to reference model M_5 of this robustness study. In M_5 the analysis is conducted again with variances and spatial dependency parameters set constant to these values. The remaining runs evaluate the influence of different variance parameters on a small grid of values. For this we take a τ^2 -set of $\{0.50, \hat{\tau}^2 = 2.19, 4.00\}$ and a τ_0^2 -set of $\{28, \hat{\tau}_0^2 = 33.13, 38.00\}$ and form all 2-way combinations of these. Note that these τ^2 resp. τ_0^2 values are rather far apart from each other, i.e. the margins of deviation of estimated posteriori distributions from the original model run are smaller to some extent. It is advantageous if robustness can be established for these rather extreme choices.

Variance parameters ξ^2 and ξ_0^2 are set in such a way that the ratios ξ^2/τ^2 and ξ_0^2/τ_0^2 are equal to the ratios of estimates $\hat{\xi}^2/\hat{\tau}^2$ and $\hat{\xi}_0^2/\hat{\tau}_0^2$. This restriction is not as limiting as it may seem at the first sight: The relative weight of spatial dependence in relation to the total variance is still varied. However, by this, the number of analysis runs can be reduced to 9. The constant ratio between variance-covariance components can be motivated by the following observation: Useful hyperparameter values leading to acceptable variance parameter trajectories in the gCAR5 model were not as difficult to be found as for spatial dependency parameters. Hence, we put less weight on the evaluation of ξ^2 (ξ_0^2). A summary of all parameter constellations can be found in Table B.1.

	M_1	M_2	M_3
ξ_0^2	26.77	31.67	36.32
τ_0^2	28.00	33.13	38.00
ξ^2	1.15	1.15	1.15
τ^2	0.50	0.50	0.50
	M_4	M_5	M_6
ξ_0^2	26.77	31.67	36.32
τ_0^2	28.00	33.13	38.00
ξ^2	5.02	5.02	5.02
τ^2	2.19	2.19	2.19
	M_7	M_8	M_9
ξ_0^2	26.77	31.67	36.32
τ_0^2	28.00	33.13	38.00
ξ^2	9.19	9.19	9.19
τ^2	4.00	4.00	4.00

Table B.1: Design for robustness evaluation of gCAR5 variance parameter estimation.

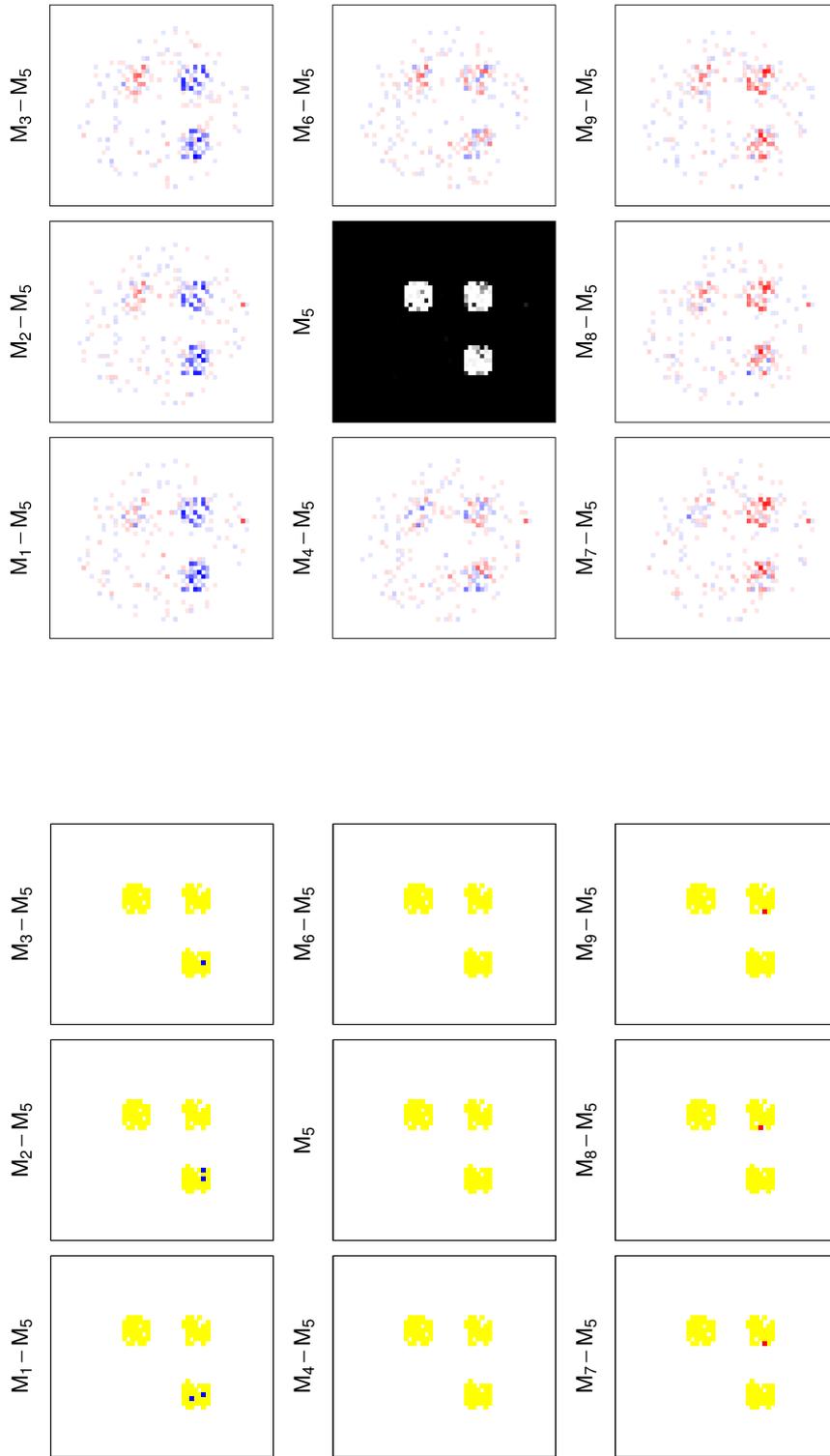
In the following result tables and figures, we look at differences in activation probability estimation and activation classification. We note that varying τ^2 leads to stronger variations in results than varying τ_0^2 —which might be due to our choice of values where τ^2 varies more strongly relative to its size.

In Table B.2 the sensitivity values of M_1 to M_9 are listed. We note that with increasing τ^2 sensitivity increases slightly. Differences, however, are small and values are in the range of 0.812 and 0.847. This can also be inferred from the activation difference plots in Figure B.1a, which are exemplary shown for the medium layer 3. Only a small increase in voxels being classified as active can be observed with increasing τ^2 .

M_1	M_2	M_3
0.812	0.812	0.820
M_4	M_5	M_6
0.828	0.839	0.835
M_7	M_8	M_9
0.839	0.847	0.843

Table B.2: Sensitivity estimates of gCAR5 runs.

In Figure B.1b differences in posterior activation probabilities of respective models are plotted compared to M_5 . The runs with $\tau^2 = 0.50$, i.e. M_1 , M_2 and M_3 , yield smaller activation probabilities in activation regions than M_5 . The runs with $\tau^2 = 4.00$, i.e. M_7 , M_8 and M_9 , yield larger activation probabilities in activation regions than M_5 . The remaining two models, M_4 and M_6 , with τ^2 being equal to M_5 have activation probability values similar to M_5 —though larger variations within activation regions can be observed. As can be inferred from Table B.3, variations in posterior activation probability values are well centered around 0 and are maximal in the order of magnitude of ± 0.15 . We assume that changes in posterior activation probability values must be larger to affect binary activation maps.



(a) Activation maps of models M_i calculated by the gCAR5 algorithm in comparison to the activation map of the reference model M_5 . The yellow color denotes voxels that are active in M_i but not in M_5 . The blue color denotes voxels that are active in M_5 but not in M_i . Red color denotes voxels that are active in M_i but not in M_5 . The maps contain the results of the central brain layer, i.e. layer 3.

(b) Activation probability estimates in the gCAR5 models. In the middle plot, $\hat{\pi}_{M_5}$ is plotted. The surrounding plots depict the differences in π -estimation of the remaining models to $\hat{\pi}_{M_5}$. Higher probability values of the M_5 model are marked in shades of blue, lower probability values are marked in shades of red. The maps contain the results of the central brain layer, i.e. layer 3.

Figure B.1: Robustness of gCAR5

	Min.	Mean	Max.
$M_1 - M_5$	-0.1490	-0.0004	0.0530
$M_2 - M_5$	-0.1430	-0.0004	0.0410
$M_3 - M_5$	-0.1380	-0.0004	0.0390
$M_4 - M_5$	-0.0330	-0.0000	0.0450
$M_6 - M_5$	-0.0430	0.0001	0.0530
$M_7 - M_5$	-0.0410	0.0002	0.0930
$M_8 - M_5$	-0.0370	0.0002	0.0840
$M_9 - M_5$	-0.0240	0.0002	0.0850

Table B.3: Summary of difference values between posterior activation probability values of models M_i and reference model M_5 in the gCAR5 model.

B.2 Robustness of fCAR5 results

The evaluation runs of the fCAR5 algorithm are constructed in an analogous fashion as for the gCAR5 algorithm. We take the variance-covariance parameter estimates of fCAR5 from Section 6.1.3 and set corresponding parameters constant to these values in M_5 . Then, we evaluate robustness by varying τ^2 and τ_0^2 -values on a 2D grid. Variances ξ^2 and ξ_0^2 are again only varied by keeping ξ^2/τ^2 and ξ_0^2/τ_0^2 constant. See Table B.4 for a summary of conducted runs.

	M_1	M_2	M_3
ξ_0^2	2.14	6.77	17.14
τ_0^2	0.50	1.58	4.00
ξ^2	1.40	1.40	1.40
τ^2	0.01	0.01	0.01
	M_4	M_5	M_6
ξ_0^2	2.14	6.77	17.14
τ_0^2	0.50	1.58	4.00
ξ^2	4.32	4.32	4.32
τ^2	0.03	0.03	0.03
	M_7	M_8	M_9
ξ_0^2	2.14	6.77	17.14
τ_0^2	0.50	1.58	4.00
ξ^2	14.02	14.02	14.02
τ^2	0.10	0.10	0.10

Table B.4: Design for robustness evaluation of fCAR5 variance parameter estimation.

In Table B.5, the sensitivity values of M_1 to M_9 are listed. Only slight changes in sensitivity evolve with no specific trend. This can also be inferred from Figure B.2a.

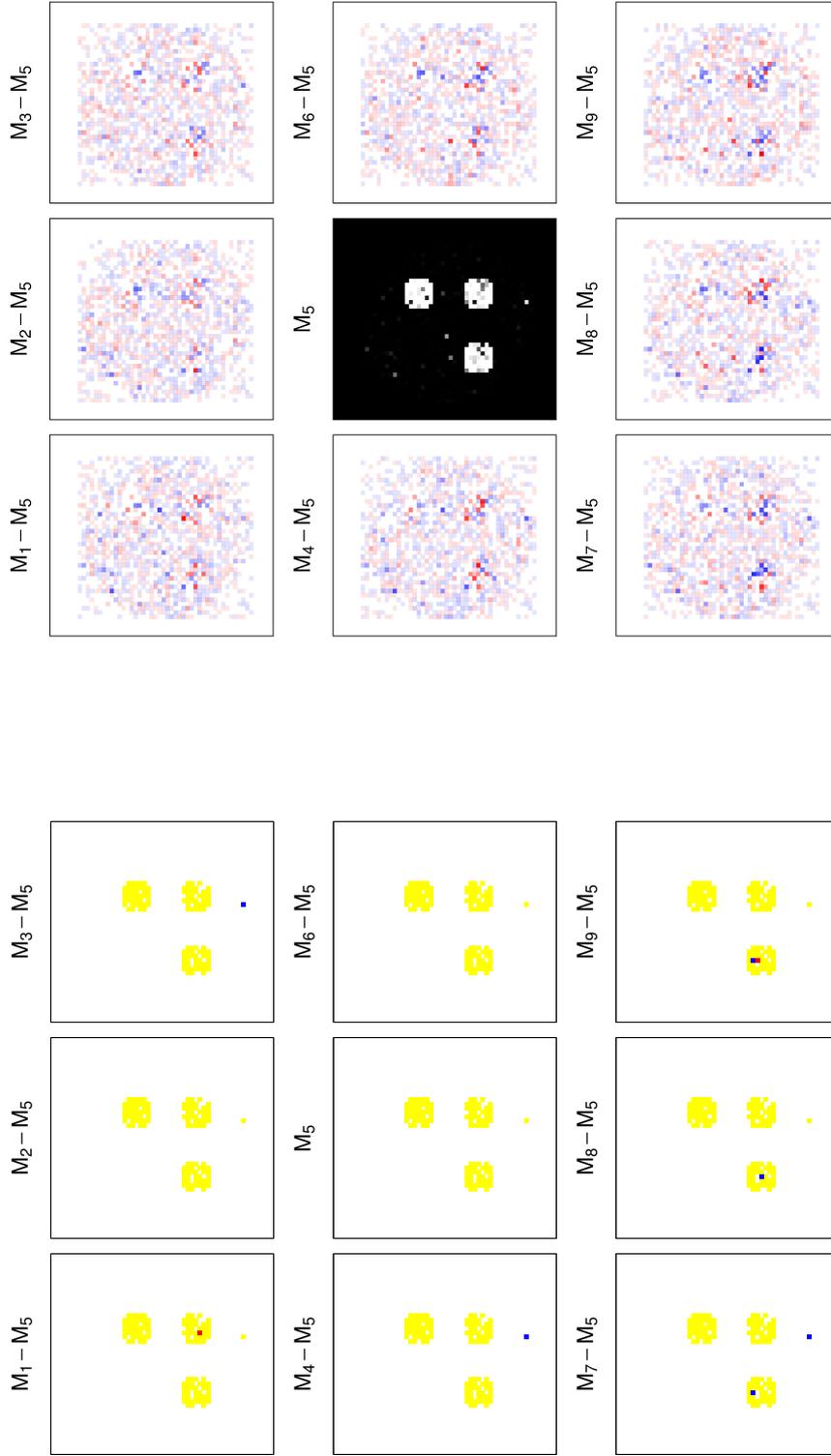
M_1	M_2	M_3
0.862	0.854	0.858
M_4	M_5	M_6
0.854	0.854	0.854
M_7	M_8	M_9
0.851	0.851	0.854

Table B.5: Sensitivity estimates of fCAR5 runs.

In Figure B.2b, differences in posterior activation probabilities of respective models are plotted compared to M_5 . In all models changes in activation probability occur in both directions well scattered over the brain. Outside the brain (which is included by mask in the fCAR5 model) minor changes can be observed, whereas in activation regions larger variations evolve. As can be inferred from Table B.6, variations in posterior activation probability values are well centered around 0 and are maximal in the order of magnitude of ± 0.1 .

	Min.	Mean	Max.
$M_1 - M_5$	-0.0620	-0.0003	0.0860
$M_2 - M_5$	-0.0530	-0.0000	0.0780
$M_3 - M_5$	-0.0750	0.0001	0.0590
$M_4 - M_5$	-0.0740	-0.0003	0.0880
$M_6 - M_5$	-0.0440	0.0002	0.0710
$M_7 - M_5$	-0.0650	-0.0003	0.0680
$M_8 - M_5$	-0.0780	-0.0000	0.0720
$M_9 - M_5$	-0.0610	0.0001	0.1000

Table B.6: Summary of difference values between posterior activation probability values of models M_i and reference model M_5 in the fCAR5 model.



(a) Activation maps of models M_i calculated by the fCAR5 algorithm in comparison to the activation map of the reference model M_5 . The yellow color denotes voxels that are active in M_i but not in M_5 . The blue color denotes voxels that are active in M_5 but not in M_i . Red color denotes voxels that are active in M_i but not in M_5 . The maps contain the results of the central brain layer, i.e. layer 3.

(b) Activation probability estimates in the fCAR5 models. In the middle plot, $\hat{\pi}_{M_5}$ is plotted. The surrounding plots depict the differences in π -estimation of the remaining models to $\hat{\pi}_{M_5}$. Higher probability values of the M_5 model are marked in shades of blue, lower probability values are marked in shades of red. The maps contain the results of the central brain layer, i.e. layer 3.

Figure B.2: Robustness of fCAR5

B.3 Robustness of iMRF5 results

The IGMRF prior depends on only one variance-covariance parameter (per spatially-varying effect) and hence, models both variance and spatial dependency simultaneously. Parameter ξ^2 is the variance-covariance parameter of the spatially-varying EEG effect and ξ_0^2 of the spatially-varying intercept. Like for the gCAR5 and fCAR5 models, we take the variance-covariance parameter estimates $\hat{\xi}^2$ and $\hat{\xi}_0^2$ of the iMRF5 model from Section 6.1.3 and set corresponding parameters constant to these values in M_5 . Then we evaluate robustness by varying ξ^2 and ξ_0^2 -values on a 2D grid (cf. Table B.7 for a summary of parameter constellations).

	M_1	M_2	M_3
ξ_0^2	2.00	5.10	8.00
ξ^2	2.00	2.00	2.00
	M_4	M_5	M_6
ξ_0^2	2.00	5.10	8.00
ξ^2	4.48	4.48	4.48
	M_7	M_8	M_9
ξ_0^2	2.00	5.10	8.00
ξ^2	8.00	8.00	8.00

Table B.7: Design for robustness evaluation of iMRF5 variance parameter estimation.

In Table B.8, the sensitivity values of M_1 to M_9 are listed. We note that with increasing ξ_0^2 sensitivity increases slightly. Differences, however, are very small and values are in the range of 0.881 and 0.912. This can also be inferred from the activation difference plots in Figure B.3a, which are exemplarily shown for the medium layer 3. Only a small increase in voxels being classified as active can be observed with increasing τ_0^2 .

M_1	M_2	M_3
0.889	0.912	0.912
M_4	M_5	M_6
0.893	0.908	0.908
M_7	M_8	M_9
0.881	0.897	0.908

Table B.8: Sensitivity estimates of iMRF5 runs.

	Min.	Mean	Max.
$M_1 - M_5$	-0.1670	-0.0003	0.0290
$M_2 - M_5$	-0.0410	0.0000	0.0570
$M_3 - M_5$	-0.0320	0.0001	0.0550
$M_4 - M_5$	-0.1550	-0.0003	0.0210
$M_6 - M_5$	-0.0490	0.0001	0.0670
$M_7 - M_5$	-0.1860	-0.0004	0.0510
$M_8 - M_5$	-0.0820	-0.0000	0.0310
$M_9 - M_5$	-0.0490	0.0001	0.0680

Table B.9: Summary of difference values between posterior activation probability values of models M_i and reference model M_5 in the iMRF5 model.

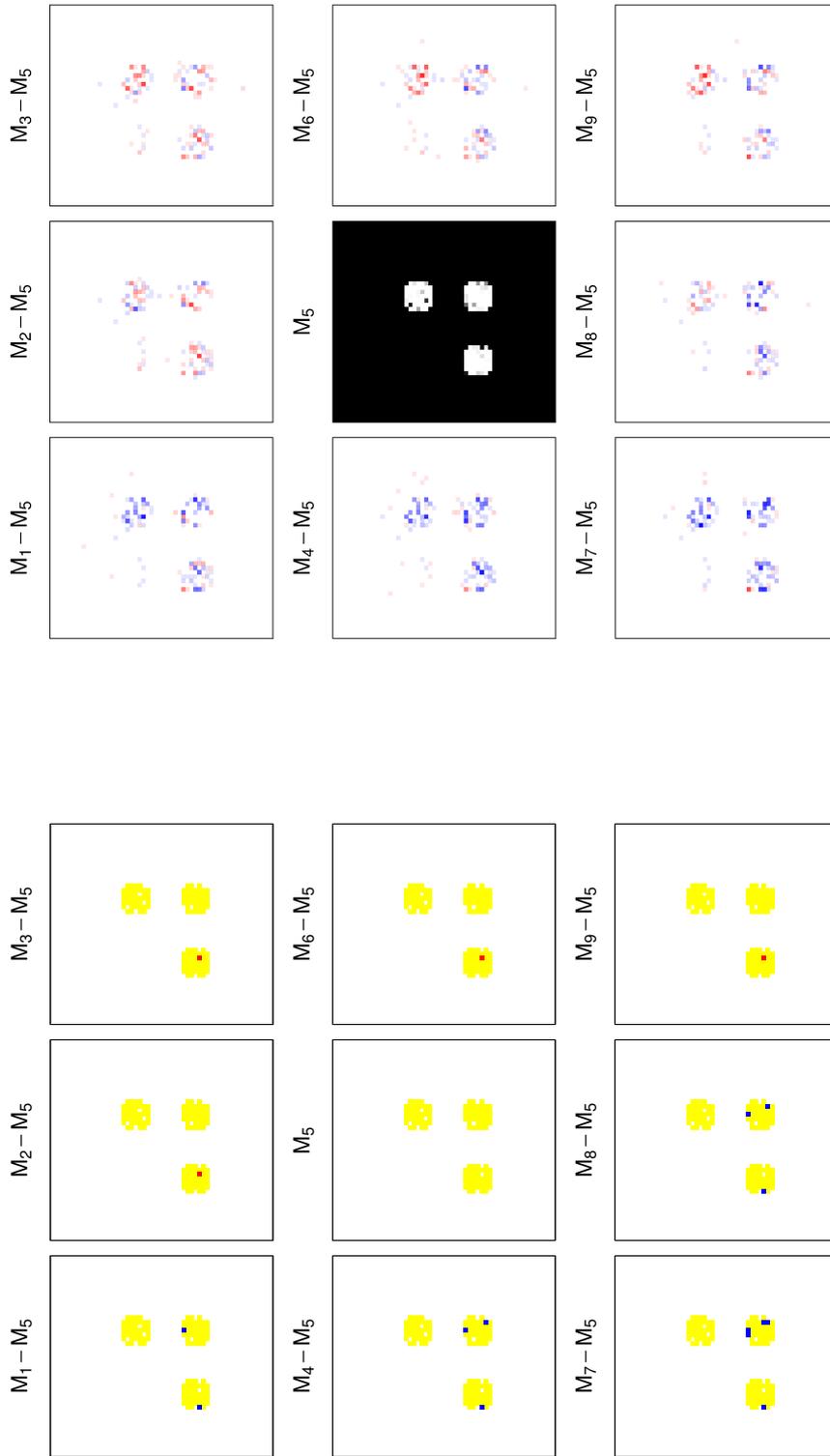
In Figure B.3b, differences in posterior activation probabilities of respective models are plotted compared to M_5 . Conform to activation classifications, we note that increasing τ_0^2 values lead to increasing activation probability values. Variations, however, are mainly restricted to activation regions. Like in the models based on the CAR prior, variations in posterior activation probability values are well centered around 0 with maximal values in the order of magnitude of ± 0.17 (cf. Table B.9).

B.4 Summary

In this chapter, we reported results from a robustness study evaluating the influence of variance and spatial dependency parameter values on estimates of interest. Ultimately, we are interested in binary activation maps and underlying posterior activation probability maps. Despite the rather extreme choices of variance parameters, only minor changes in activation classification occur for all model types (gCAR5, fCAR5, iMRF5). Hence, sensitivity of each model type remains on a similar level for all runs. Specificity values were not shown for staying on the same high level (i.e. most of the time equal to 1) for all runs—which might be due to the simulation design.

The fCAR5 model (which works with single-site updates) appeared to be the most robust concerning the variation in sensitivity values. Variations in posterior activation probability maps happen all over the brain and are not restricted to activation regions.

In both gCAR5 and iMRF5 models, which have in common that spatially-varying effects are updated globally, variations in posterior activation probability maps happen mainly



(a) Activation maps of models M_i calculated by the iMRF5 algorithm in comparison to the activation map of the reference model M_5 . The yellow color denotes voxels that are classified as active in both maps. The blue color denotes voxels that are active in M_i but not in M_5 . Red color denotes voxels that are active in M_5 but not in M_i . The maps contain the results of the central brain layer, i.e. layer 3.

(b) Activation probability estimates in the iMRF5 models. In the middle plot, $\hat{\pi}_{M_5}$ is plotted. The surrounding plots depict the differences in π -estimation of the remaining models to $\hat{\pi}_{M_5}$. Higher probability values of the M_5 model are marked in shades of blue, lower probability values are marked in shades of red. The maps contain the results of the central brain layer, i.e. layer 3.

Figure B.3: Robustness of iMRF5

in activation regions. With increasing spatial dependency parameter more voxels are classified as active. For the gCAR5 model this is due mainly to the increase in τ^2 . For the iMRF5 model this is due particularly to the increase in ξ_0^2 .

We conclude that results are fairly robust against choices of variance and spatial dependency parameters and thus, against mixing issues within the corresponding Bayesian estimation process. Hence, the results presented in Chapter 6 can be considered to be reliable.

Appendix C

Implementation: Arguments for a call to the `fmrieeeg` algorithm

A data independent implementation of the `fmrieeeg` algorithm collection is inevitably connected with a large number of arguments needed for configuring different model runs. For starting the algorithm, up to 58 arguments have to be specified—depending on the kind of model to be fitted. These arguments have to be passed to the algorithm using a convenient configuration text-file. In Chapter 5, we demonstrated an easy way to control the setting of arguments using the utilities of the `Rfmrieeeg` R-package, but also a carefully documented configuration file template is available that can be modified to ones needs.

In the following, we give an overview of arguments that can be controlled by the user. We differentiate between arguments being independent of the exact choice of Bayesian activation detection model. These are listed as common parameters in Table C.1. Arguments for the latent Ising model (Smith et al., 2003), which serves as a reference to our models, are listed in Table C.2. Arguments for the latent probit models, which are proposed in this thesis in Chapters 3 and 4, are listed in Table C.3 combined with the reference to corresponding algorithm types using each.

With the categorical argument `modelType`, the model is selected: '1' selects a latent probit model with CAR prior, '2' selects a latent probit model with an IGMRF prior (iMRF algorithm), '3' selects an Ising model (which is the default). If the CAR model is selected (`modelType=2`), the boolean probit parameter `speedup` (cf. Table C.3) switches between the CAR model with global updates (gCAR) and the fast CAR algorithm (fCAR): If `speedup=FALSE`, gCAR is selected and otherwise fCAR. For `modelType` equal to 1 and 2, the five-level `linPred` argument contains the type of predictor to be fitted (cf. Table 3.1).

Name	Default	Description
<code>setConfig.pathResults(...)</code>		
<i>General parameters</i>		
<code>pathResults</code>	*	Path to folder where a results folder can be generated.
<code>setConfig.fmri(...)</code>		
<i>Arguments related to the scanning procedure</i>		
<code>t0</code>	0	Exact time of first (slice-time corrected) scan in seconds since recording onset.
<code>tr</code>	2	Repetition time of scans in seconds.
<i>Information about the stimulus presentation</i>		
<code>nStimTypes</code>	1	Number of stimulus types to be model. (Note that Bayesian activation detection—so far—can only be applied to one stimulus type specified in <code>stimActDect</code> . The remainder types are modeled as confounders.)
<code>stimActDect</code>	1	Stimulus type for which Bayesian activation detection should be applied to. (This is a number within range 1 - <code>nStimTypes</code> .)
<code>nPerStimType</code>	*	Integer array with length <code>nStimTypes</code> : Each element contains the number of presented stimuli of one of the <code>nStimTypes</code> stimulus types.
<code>pathStimTimes</code>	*	Path to text-file with an one-dimensional array of stimulus presentation times of all stimuli. (Note that presentation times of one stimulus type must be grouped together. Stimulus groups are ordered according to the same sequence as in <code>nPerStimType</code> .)
<i>Arguments related to the fMRI data</i>		
<code>pathNifti</code>	*	Path to fMRI data, which is available as a 4D nifti file (i.e. in form of one .nii file).
<code>maskType</code>	threshold	Type of mask for selecting voxels for analysis. Choose between <code>threshold</code> and <code>implicit</code> ; or specify a path to an explicit binary mask image (file-format '.nii').
<code>grandMeanScaling</code>	FALSE	Indicator if grand mean scaling is requested.
<i>Arguments related to the fMRI design matrix: confounders and high pass filter</i>		
<code>confounder</code>	FALSE	Indicator if confounder variables should be modeled.
<code>pathConf</code>	*	If <code>confounder==TRUE</code> : path to text-file with confounder matrix.
<code>freqHPF</code>	128	High pass filter frequency. Upon this DCT regressors are calculated regressing out frequencies below this value.

Continued on next page

Table C.1 – continued from previous page

Name	Default	Description
<code>setConfig.stimCanonicalBF(...)</code> or <code>setConfig.stimGammaBF(...)</code>		
<i>Arguments related to the HRF modeling at the fMRI regression stage</i>		
<code>typeStimBF</code>	<code>canonical</code>	Type of HRF basis function (BF). Choose between type <code>canonical</code> or <code>gamma</code> .
<code>derivativesStimBF</code>	<code>2</code>	Number of derivatives used for generating basis functions. For <code>canonical</code> : 0 (just canonical), 1 (plus 'time' derivative), 2 (plus 'dispersion' derivative). For <code>gamma</code> : 0 (no derivative) or 1 (plus 'time' derivative);
<i>Arguments for the gamma type of HRF BF:</i>		
<code>shapeVec</code>	<code>4, 8, 16</code>	Vector of shape parameters for gamma basis functions.
<code>noShapeParams</code>	<code>3</code>	Length of <code>shapeVec</code> .
<code>scaleVec</code>	<code>1</code>	Vector of scale parameters for gamma basis functions.
<code>noScaleParams</code>	<code>1</code>	Length of <code>scaleVec</code> .
<code>setConfig.MCMC(...)</code>		
<i>Arguments for the MCMC procedure</i>		
<code>chainSize</code>	<code>6000</code>	Total number of MCMC iterations to be conducted.
<code>burnin</code>	<code>1000</code>	Number of iterations after which convergency of Markov chains to their equilibrium distribution is assumed.
<code>step</code>	<code>5</code>	Thinning parameter.
<code>diagnostics</code>	<code>TRUE</code>	Indicator if diagnostics should be written out: <code>TRUE</code> if diagnostics should be saved, <code>FALSE</code> otherwise.
<code>randomSample</code>	<code>TRUE</code>	Indicator if the sample of voxels for which the diagnostics should be saved should be randomly chosen (then set <code>randomSample = TRUE</code>) or indices of voxels are specified in a vector (then set <code>randomSample = FALSE</code> and specify an array of integers for argument <code>subSample</code>).
<code>noVoxSample</code>	<code>20</code>	This is either the number of randomly selected voxels (if <code>randomSample==TRUE</code>) or the length of <code>subSample</code> (if <code>randomSample==FALSE</code> and an array of indices is specified in <code>subSample</code>).
<code>subSample</code>	<code>0</code>	If <code>randomSample == FALSE</code> , these are the 1D indices (1-based) of selected voxels for which diagnostics are saved. Note that <code>noVoxSample</code> has to be specified nevertheless (for reading in the vector).

Table C.1: Common arguments for the `fmriieg` C++ algorithm. Mark '*' denotes arguments that are mandatory to be specified by the user and have no default values. When using the `Config` class functions within the R-package `Rfmriieg`, the arguments listed underneath the R setter function names are set by the corresponding function.

Name	Default	Description
<code>piPrior</code>	0.25	Global prior probability for activation within gray-matter.
<code>gmMap</code>	1	Optional gray-matter probability map, which is available as a 3D nifti file (i.e. a .nii file). If not specified, all voxels in the analyzing mask have gray-matter probability 1.
<code>theta</code>	0.45	Spatial dependency parameter for Ising field.

Table C.2: Model specific arguments for an Ising model run provided by the `fmrieeg` C++ algorithm.

Argument	Comment	Default	Predictor type				
			1	2	3	4	5
speedup	gCAR (FALSE) vs. fCAR (TRUE)	FALSE	CAR	CAR	CAR	CAR	CAR
pathEEG	Path to EEG data (3D nii-file)	*	iMRF, CAR	iMRF, CAR	CAR	iMRF, CAR	iMRF, CAR
<i>Prior parameter values for estimating variances τ_0^2, τ^2</i>							
tauStart	τ^2 starting value	1	CAR			CAR	CAR
muTau	μ_p	0	CAR			CAR	CAR
sigma2ProposalTau	$\sigma_{\tau^2}^2$	1	CAR			CAR	CAR
tau0Start	τ_0^2 starting value	1		CAR	CAR		CAR
muTau0	$\mu_{0,p}$	0		CAR	CAR		CAR
sigma2ProposalTau0	$\sigma_{\tau_0}^2$	1		CAR	CAR		CAR
sigma2Tau	$\sigma_{0,p}^2$	25	CAR	CAR	CAR	CAR	CAR
<i>Prior parameter values for estimating global effect parameters $\alpha_G, \alpha_{0,G}$</i>							
muAlphaGlobal	$\mu_{0,G}$	0	iMRF	iMRF, CAR		iMRF	iMRF
sigma2ProposalAlphaG	$\xi_{prop,G}^2$	1	iMRF, CAR	iMRF, CAR			
muAlpha0Global	μ_G	0	iMRF, CAR	iMRF		iMRF	iMRF
<i>Prior parameter values for estimating variances $\xi_G^2, \xi_0^2, \xi_0, \xi_0^2, \xi_0^2, \xi_0^2$</i>							
aGlobalIGpriori	a_G	3	iMRF	iMRF, CAR		iMRF	iMRF
bGlobalIGpriori	b_G	1	iMRF	iMRF, CAR		iMRF	iMRF
a0GlobalIGpriori	$a_{0,G}$	3	iMRF, CAR	iMRF		iMRF	iMRF
b0GlobalIGpriori	$b_{0,G}$	1	iMRF, CAR	iMRF		iMRF	iMRF
a0IGpriori	a_0	204.5		iMRF, CAR	iMRF, CAR		iMRF, CAR
b0IGpriori	b_0	915.75		iMRF, CAR	iMRF, CAR		iMRF, CAR
aIGpriori	a	204.5	iMRF, CAR			iMRF, CAR	iMRF, CAR
bIGpriori	b	915.75	iMRF, CAR			iMRF, CAR	iMRF, CAR
<i>If requested: set variance and spatial dependency parameters constant</i>							
xi20Const	boolean	FALSE		iMRF, CAR	iMRF, CAR		iMRF, CAR
xi2Const	boolean	FALSE	iMRF, CAR			iMRF, CAR	iMRF, CAR
tau0Const	boolean	FALSE	CAR	CAR		CAR	CAR
tauConst	boolean	FALSE	CAR			CAR	CAR
xi20Val	constant ξ_0^2 value	5		iMRF, CAR	iMRF, CAR		iMRF, CAR
xi2Val	constant ξ_0^2 value	5	iMRF, CAR			iMRF, CAR	iMRF, CAR
tau0Val	constant τ_0^2 value	5		CAR		CAR	CAR
tauVal	constant τ^2 value	5	CAR			CAR	CAR

Table C.3: Model specific arguments for `fMRIIEEG`. The 2nd column contains the according parameter name used in Chapters 3 and 4 or a short key word or explanation. Model types can be inferred by a combination of predictor type and prior choice ('CAR' and 'iMRF') in the last 5 columns. Mark '*' denotes arguments with no default value.

Appendix D

Supplementary results: Results of shortened fMRI data trajectories

In the simulation study presented in Section 6.1.5, an increased performance of EEG-enhanced fMRI activation detection algorithms was observed at medium signal-to-noise ratio level. The signal-to-noise level, however, seems to be different in our application: An fMRI trajectory of $T = 303$ images may contain sufficient information in the fMRI likelihood to suppress prior EEG information. We hypothesize that the EEG-enhanced fMRI activation detection algorithms may show their strengths if fMRI trajectories contain less measurements. If sensitivity is substantially increased, this finding may be useful to reduce observation time of subjects in experiments. To test this hypothesis, we rerun some analyses from Chapter 7 with shortened fMRI data trajectories with fMRI subsets of images 1-200 (referred to as case 'A') and images 101-300 (referred to as case 'B').

We first compare activation detection results on fMRI subsets to the analysis of all 303 fMRI scans with predictor 3 algorithms, which do not use additional EEG information. Pairwise difference maps of each algorithm and subset can be found in Figure D.1. The difference between the fCAR3, gCAR3 and iMRF3 algorithms are small and in line with findings described in Section 7.3.1. However, the discrepancy between cases 'A' and 'B' are striking: A strong time-varying activation effect emerges overall brain regions. This might be traced back to the specific experimental design and/or characteristics of the observed test person. Activation changes occur in both directions with the majority of voxels (subject to change) changing from being non-active to active. Case 'B' contains more activation information compared to both case 'A' and the entire fMRI dataset. The latter can be explained by the fact that estimates derived from the entire dataset are an average of activated and non-activated contents of the time series. Hence, case 'B' contains more

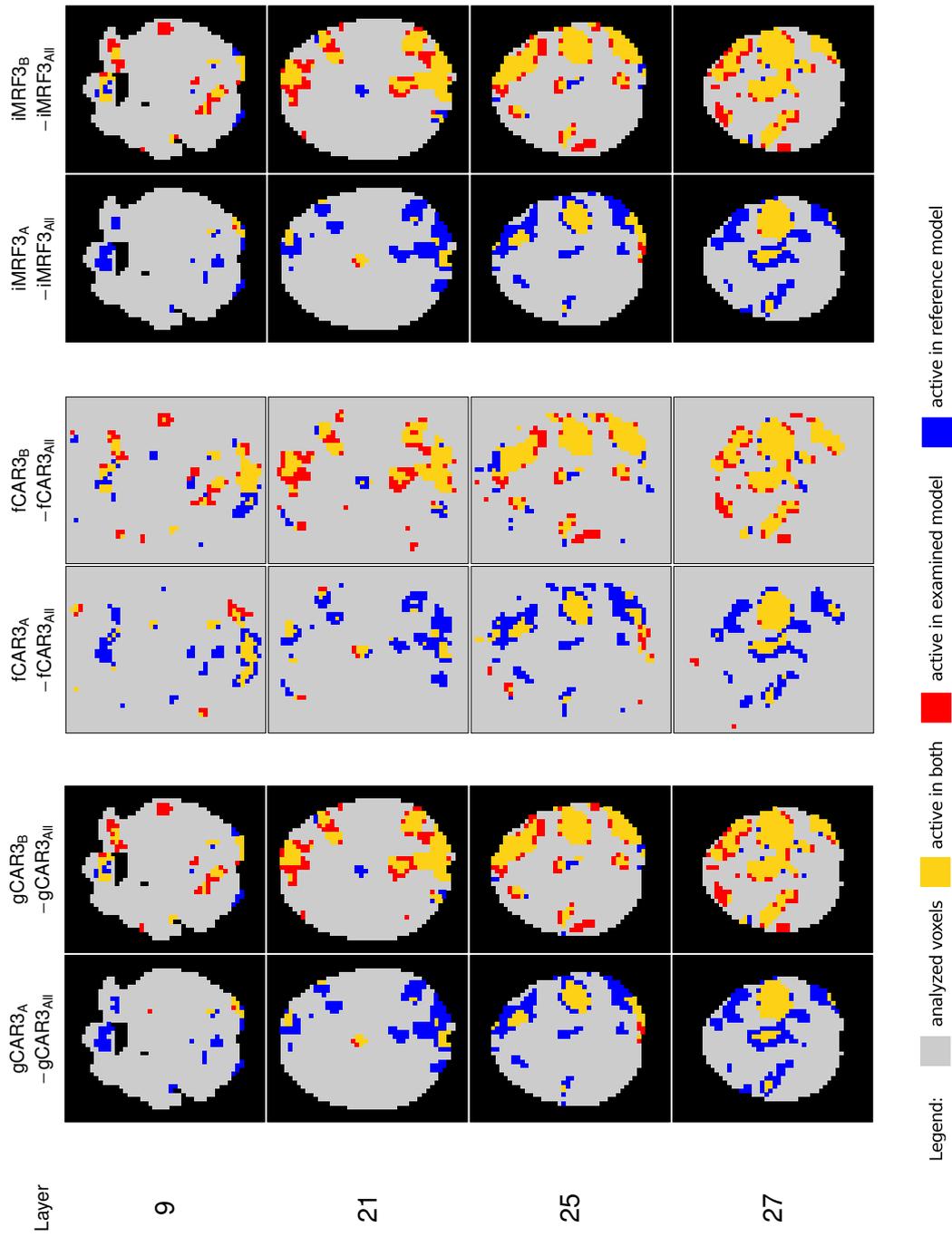


Figure D.1: Activation difference maps of gCAR3 (left), fCAR3 (middle) and iMRF3 (right) runs with fMRI scans 1-200 (A) and 101-300 (B) in comparison to the activation map of an analysis with all scans, 1-303 (All).

activation for only averaging over activated time sections. In the central part of layer 21, we observe an activation change to the other direction: A clear activation focus vanishes over time.

To test whether external EEG information can increase the sensitivity of runs on fMRI subsets, we use the extended fMRI activation map from the test case described in Section 7.3.2. With this map, we do not expect any complications with non-congruent informational content of EEG and fMRI and can directly assess whether beneficial external information can be used to increase sensitivity. The results of the runs with the reduced fMRI dataset 'A' (scans 1-200) are presented in Figures D.2, D.3 and D.4. The results of the runs with the reduced fMRI dataset 'B' (scans 101-300) are presented in Figures D.5, D.6 and D.7. A detailed figure description can be found in Section 7.3.2.

As noted above, subsetting the fMRI dataset to scans 1-200 leads to an enormous loss in sensitivity: The number of activation foci as well as the size of remaining activation foci is strongly decreased. Neither of the six EEG-enhanced algorithms can be used to compensate this loss. Though EEG information identifies more active voxels at few further locations, e.g. gCAR5 enlarges the central activation foci in layer 27, the algorithms are not capable of finding the large activation structures identified with the complete $T = 303$ fMRI series. For the fCAR2 and gCAR2 models not recovering an influential EEG effect in case 'A', it seems that external prior information and fMRI activation is fairly non-congruent. That is, much of the total activation information is either not contained in scans 1-200 or it is not strong enough to effect results. At large, fCAR algorithms are more sensitive than the gCAR and iMRF models in most parts of the brain, which might be traced back to relatively small activation region sizes of this data example. Here smoothness properties, i.e. generally an advantage of the algorithms with global updates, does not have an effect. Subsetting the fMRI dataset to scans 101-300 leads to the large activation foci found overall the brain as described above. Now all predictor type 2 models recover non-negligible global EEG effects leading to similar activation changes than the predictor type 5 models. Though sensitivity is high without using external information, some additional voxels can be found to be active on region boundaries, only a few become non-active. The gain in sensitivity is, however, limited. There is one region, the central part of layer 21, where activation algorithms exhibit an unequal ability to detect activation: Whereas gCAR2/5 and iMRF2/5 are only able to increase local posterior activation probabilities to some extent, fCAR2/5 probability estimates are large enough, so that central voxels are classified as active (supporting and intensifying the results of fCAR3). These are the voxel found to become non-active over the time of the experiment. Hence, compared to case 'A',

corresponding activation information is far less pronounced in scans 101-300. Comparing fCAR results to the algorithms with global updates, we notice again (cf. Section 7.3.1) that fCAR is better when the underlying activation structure is coarse (inter activation region spaces), but is outperformed when intensifying large activation structures (activation region boundaries).

Generally, when examining fMRI subsets of the auditory oddball experiment at hand, we come to the following conclusions. The observation of activation changes over time imply that the assumption of time-constant activation effects—a standard assumption in most fMRI analyses—may be strongly violated. If checks (e.g. the analysis of different fMRI subsets) indicate deviations from this assumption, time-varying activation effects should be incorporated in the fMRI activation detection model. The interested reader is referred to Bothmann (2012) for an excellent description of a time-dependent resp. vigilance dependent fMRI activation effect.

With both fMRI subsets, we found that the benefit of using additional EEG information is limited—though some increase in sensitivity can be observed. Essentially, this might be traced back to the following reason: If the fMRI trajectory contains no or insufficient activation information, the additional prior information hardly manages to increase sensitivity substantially. The interval of suitable signal-to-noise ratios—in which EEG enhancement is possible—seems to be too narrow to be of substantial practical use. Hence, only a few additional activated voxels can be found by EEG. Nevertheless, the fCAR2/5 performance in case ‘B’ appears promising for supporting the central activation foci in layer 21, which become non-active over time.

For a more detailed discussion of reasons why EEG-enhancement might fail in this subset analysis, see Chapter 8.

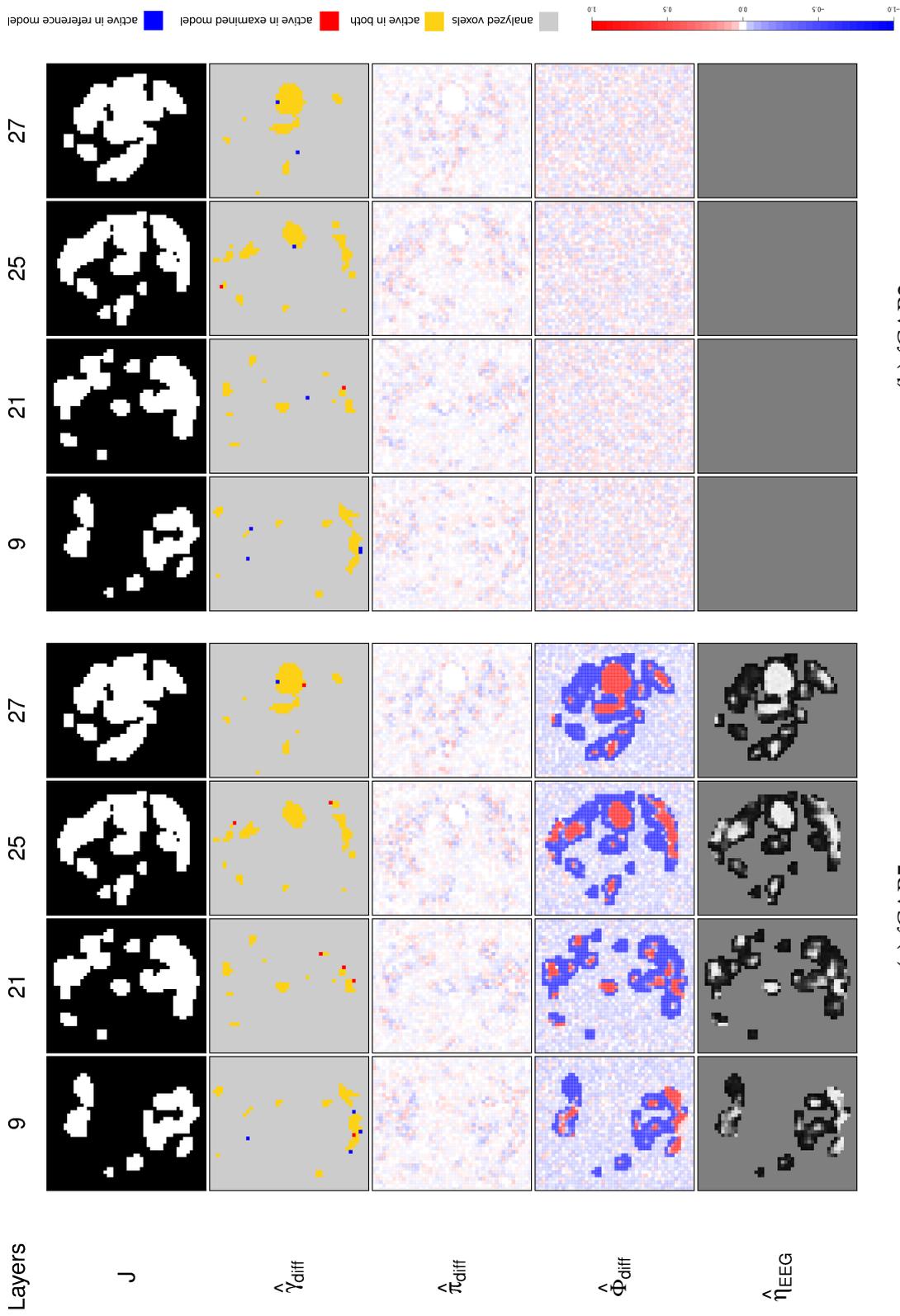


Figure D.2: Runs with fMRI scans 1-200 using the congruent prior information map: Summary of results from the fCAR5 (left) and fCAR2 (right) model in comparison to fCAR3 (selected brain layers: 9, 21, 25, 27).

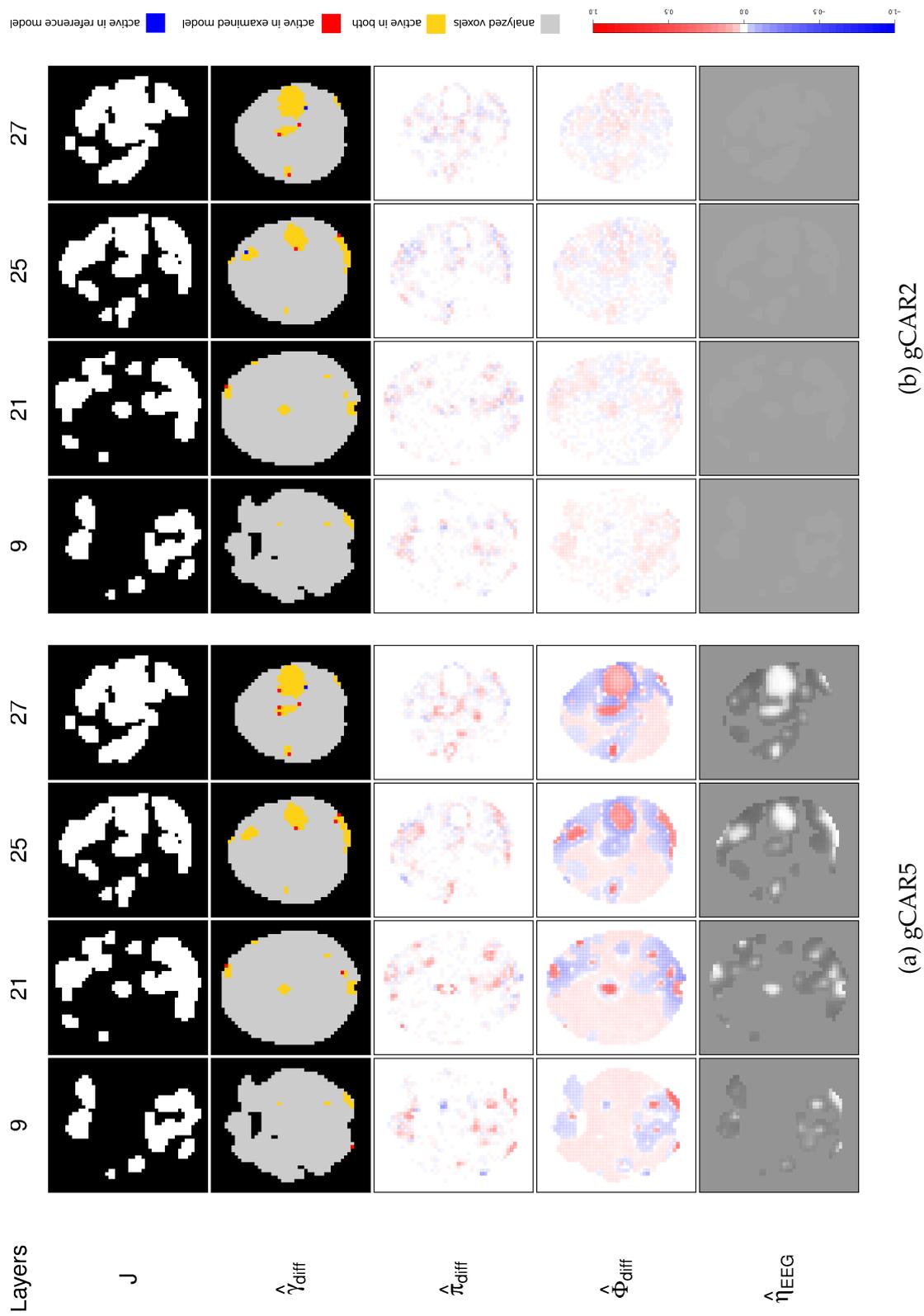


Figure D.3: Runs with fMRI scans 1-200 using the congruent prior information map: Summary of results from the gCAR5 (left) and gCAR2 (right) model in comparison to gCAR3 (selected brain layers: 9, 21, 25, 27).

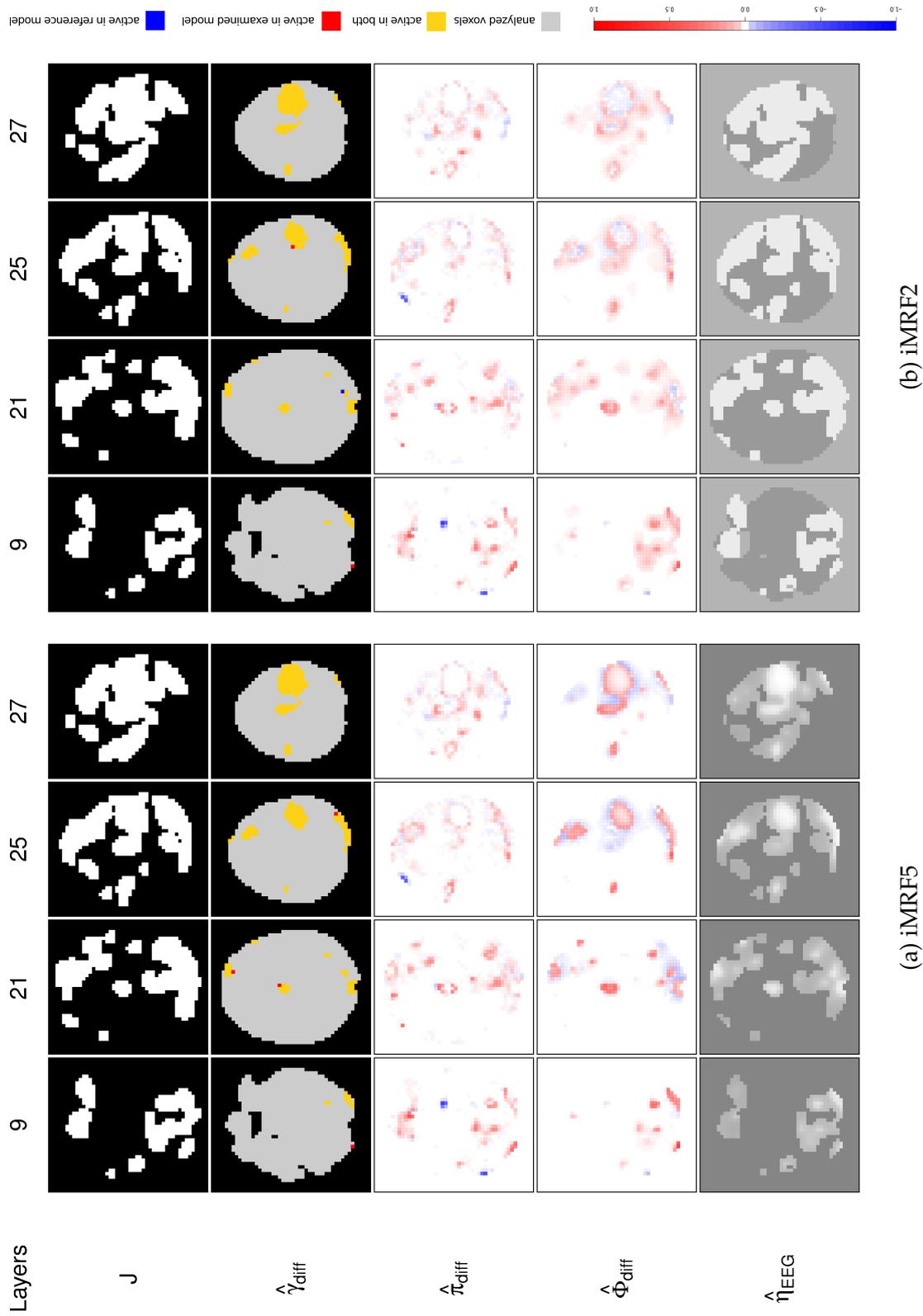


Figure D.4: Runs with fMRI scans 1-200 using the congruent prior information map: Summary of results from the iMRF5 (left) and iMRF2 (right) model in comparison to iMRF3 (selected brain layers: 9, 21, 25, 27).

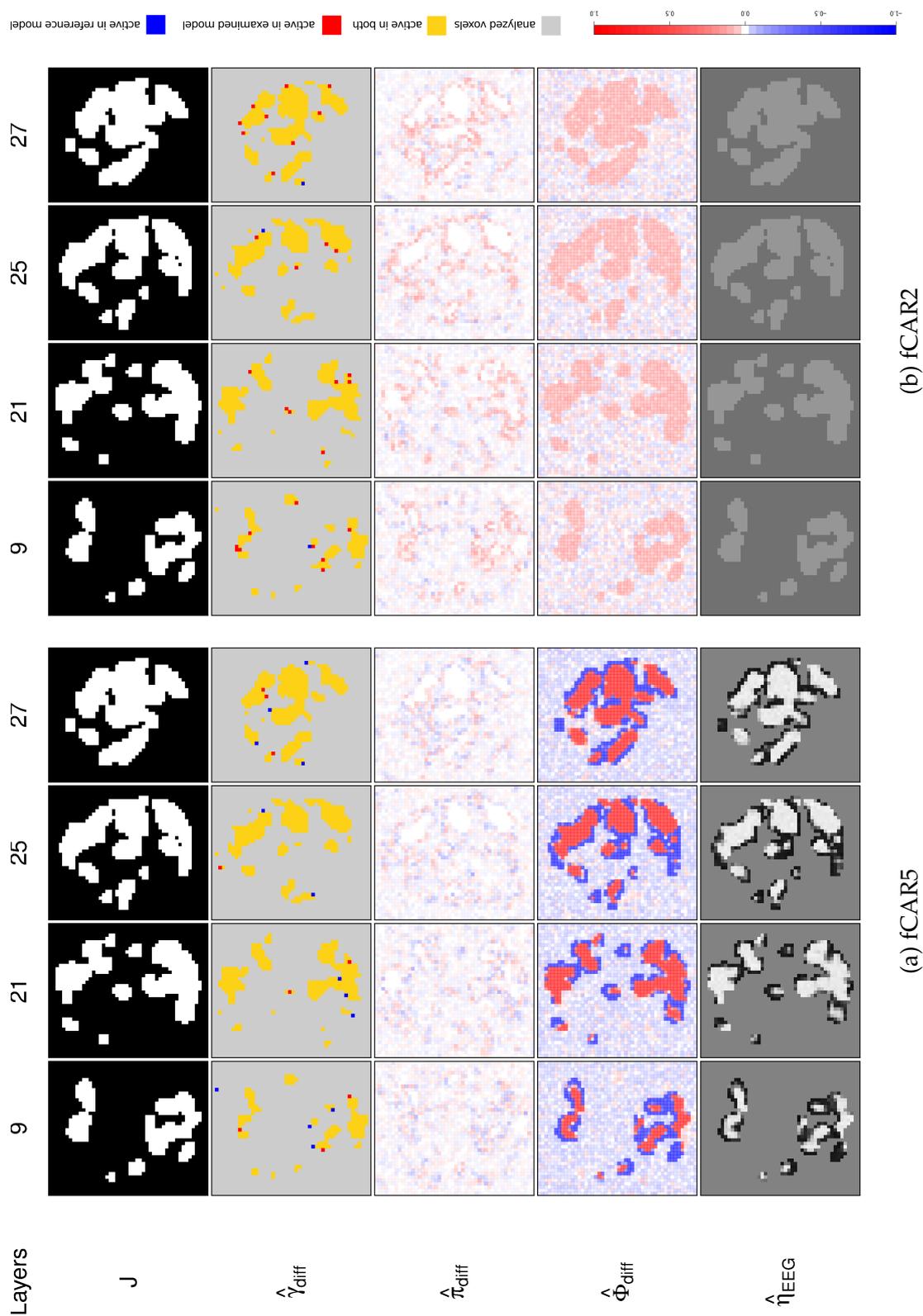


Figure D.5: Runs with fMRI scans 101-300 using the congruent prior information map: Summary of results from the fCAR5 (left) and fCAR2 (right) model in comparison to fCAR3 (selected brain layers: 9, 21, 25, 27).

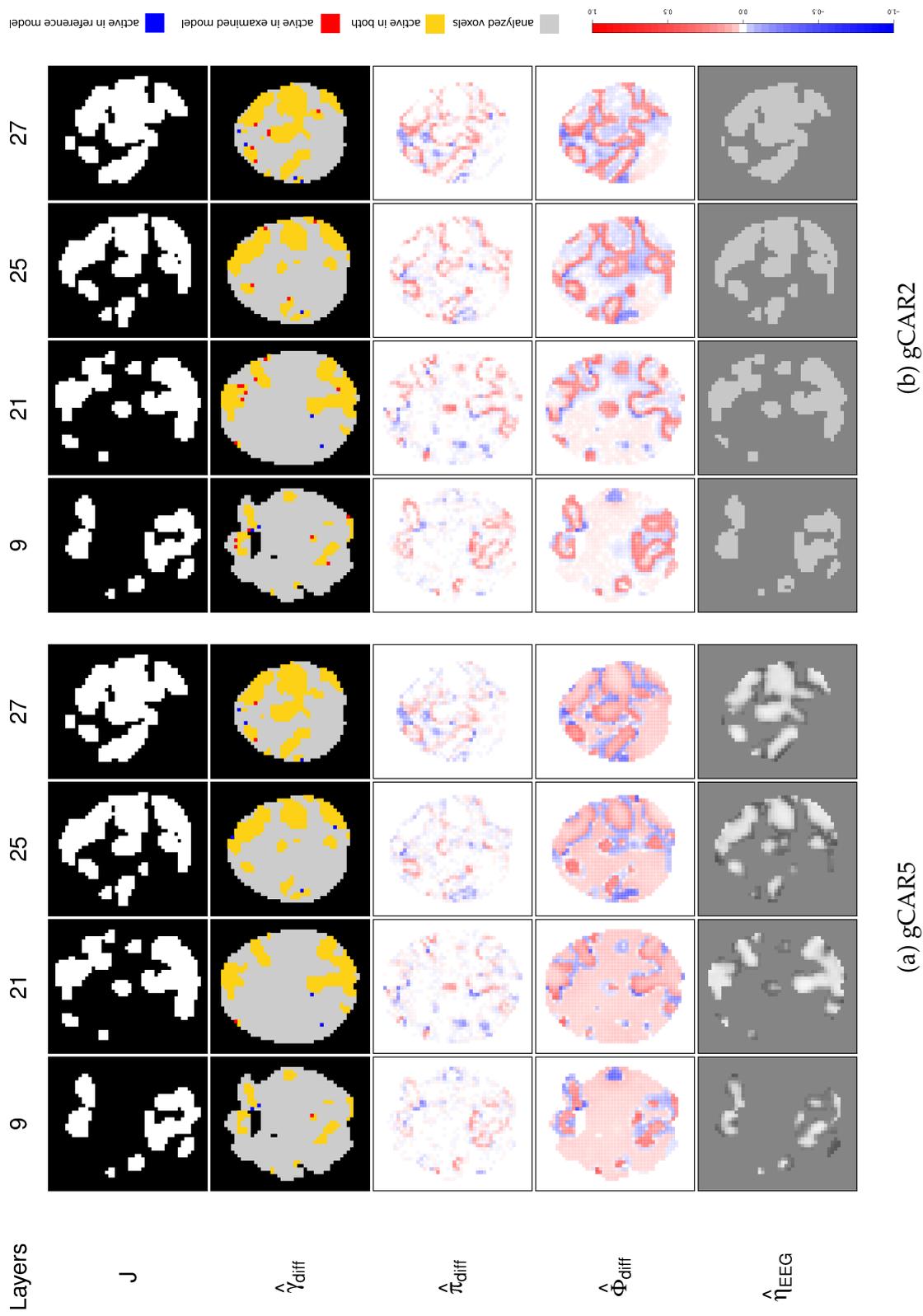


Figure D.6: Runs with fMRI scans 101-300 using the congruent prior information map: Summary of results from the gCAR5 (left) and gCAR2 (right) model in comparison to gCAR3 (selected brain layers: 9, 21, 25, 27).

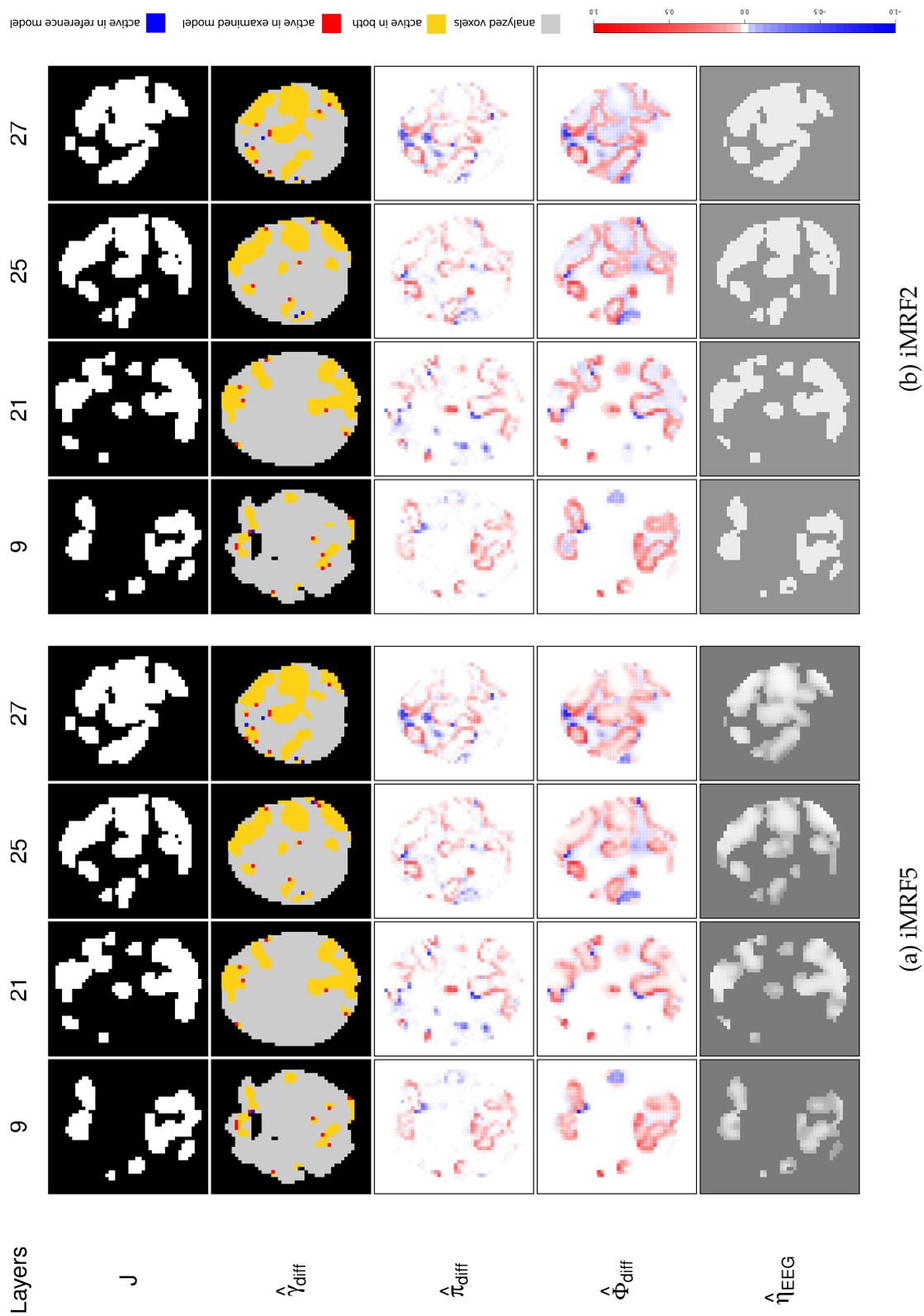


Figure D.7: Runs with fMRI scans 101-300 using the congruent prior information map: Summary of results from the iMRF5 (left) and iMRF2 (right) model in comparison to iMRF3 (selected brain layers: 9, 21, 25, 27).

Bibliography

- Ahlfors, S. P., Simpson, G. V., 2004. Geometrical interpretation of fMRI-guided MEG/EEG inverse estimates. *NeuroImage* 22 (1), 323–332.
- Ahmed, N., Natarajan, T., Rao, K. R., 1974. Discrete cosine transform. *IEEE Transactions on Computers* C-23 (1), 90–93.
- Albert, J. H., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association* 88 (422), 669–679.
- Amestoy, P. R., Davis, T. A., Duff, I. S., 1996. An approximate minimum degree ordering algorithm. *SIAM Journal on Matrix Analysis and Applications* 17 (4), 886–905.
- Amestoy, P. R., Enseeiht-Irit, Davis, T. A., Duff, I. S., 2004. Algorithm 837: AMD, An approximate minimum degree ordering algorithm. *ACM Transactions on Mathematical Software* 30 (3), 381–388.
- Andersen, E., Bai, Z., Bischof, C., Blackford, L. S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., Sorensen, D., 1999. *LAPACK users' guide*, 3rd Edition. Society for Industrial and Applied Mathematics, Philadelphia, PA.
- Arbenz, P., Hetmaniuk, U. L., Lehoucq, R. B., Tuminaro, R. S., 2005. A comparison of eigensolvers for large-scale 3D modal analysis using AMG-preconditioned iterative methods. *International Journal for Numerical Methods in Engineering* 64 (2), 204–236.
- Bagshaw, A. P., Warbrick, T., 2007. Single trial variability of EEG and fMRI responses to visual stimuli. *NeuroImage* 38 (2), 280–292.
- Bai, Z., Demmel, J., Dongarra, J., Ruhe, A., van der Vorst, H. (Eds.), 2000. *Templates for the solution of algebraic eigenvalue problems: A practical guide*. SIAM, Philadelphia.
- Bai, Z., Fahey, M., Golub, G., 1996. Some large-scale matrix computation problems. *Journal of Computational and Applied Mathematics* 74 (1-2), 71–89.

- Bandettini, P. A., 2012. 20 years of fMRI [Special issue]. *NeuroImage* 62 (2), 575–1324.
- Bandettini, P. A., Jesmanowicz, A., Wong, E. C., Hyde, J. S., 1993. Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine* 30 (2), 161–173.
- Barbieri, M. M., Berger, J. O., 2004. Optimal predictive model selection. *The Annals of Statistics* 32 (3), 870–897.
- Beckmann, C. F., Smith, S. M., 2005. Tensorial extensions of independent component analysis for multisubject fMRI analysis. *NeuroImage* 25 (1), 294–311.
- Berger, J. O., Pericchi, L. R., 2001. Objective bayesian methods for model selection: Introduction and comparison. *IMS Lecture Notes-Monograph Series: Model Selection* 38, 135–207.
- Blackford, L. S., Demmel, J., Dongarra, J., Duff, I. S., Hammarling, S., Henry, G., Kaufman, L., Lumsdaine, A., Petitet, A., Pozo, R., Remington, K., Whaley, R. C., 2002. An updated set of basic linear algebra subprograms (BLAS). *ACM Transactions on Mathematical Software* 28 (2), 135–151.
- Bothmann, L., 2012. Statistische Modellierung von EEG-abhängigen Stimuluseffekten in der fMRT-Analyse. Diploma thesis, Ludwig-Maximilians University, Munich.
- Brezger, A., Fahrmeir, L., Hennerfeind, A., 2007. Adaptive Gaussian Markov random fields with applications in human brain mapping. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 56 (3), 327–345.
- Czisch, M., Wehrle, R., Stiegler, A., Peters, H., Andrade, K., Holsboer, F., Sämann, P. G., 2009. Acoustic oddball during NREM sleep: a combined EEG/fMRI study. *PLoS ONE* 4 (8), e6749.
- Dale, A. M., Sereno, M., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience* 5 (2), 162–176.
- Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., Beckmann, C. F., 2006. Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences of the United States of America* 103 (37), 13848–13853.

- Daunizeau, J., Grova, C., Marrelec, G., Mattout, J., Jbabdi, S., Pélégrini-Issac, M., Lina, J.-M., Benali, H., 2007. Symmetrical event-related EEG/fMRI information fusion in a variational Bayesian framework. *NeuroImage* 36 (1), 69–87.
- Daunizeau, J., Laufs, H., Friston, K. J., 2010. EEG-fMRI information fusion: Biophysics and data analysis. In: Mulert, C., Lemieux, L. (Eds.), *EEG - fMRI: Physiological Basis, Technique, and Applications*. Springer-Verlag, Berlin Heidelberg, Ch. 25, pp. 511–526.
- Daunizeau, J., Mattout, J., Clonda, D., Goulard, B., Benali, H., Lina, J.-M., 2006. Bayesian spatio-temporal approach for EEG source reconstruction: Conciliating ECD and distributed models. *IEEE Transactions on Biomedical Engineering* 53 (3), 503–516.
- Davis, T. A., 2005. Algorithm 849: A concise sparse Cholesky factorization package. *ACM Transactions on Mathematical Software* 31 (4), 587–591.
- Davis, T. A., Gilbert, J. R., Larimore, S. I., Ng, E. G., 2004. Algorithm 836: COLAMD, a column approximate minimum degree ordering algorithm. *ACM Transactions on Mathematical Software* 30 (3), 377–380.
- Davis, T. A., Hager, W. W., 1999. Modifying a sparse Cholesky factorization. *SIAM Journal on Matrix Analysis and Applications* 20 (3), 606–627.
- Debener, S., Ullsperger, M., Siegel, M., Engel, A. K., 2006. Single-trial EEG-fMRI reveals the dynamics of cognitive function. *Trends in Cognitive Sciences* 10 (12), 558–563.
- Eichele, T., Specht, K., Moosmann, M., Jongsma, M. L. A., Quiroga, R. Q., Nordby, H., Hugdahl, K., 2005. Assessing the spatiotemporal evolution of neuronal activation with single-trial event-related potentials and functional MRI. *Proceedings of the National Academy of Sciences of the United States of America* 102 (49), 17798–17803.
- Fahrmeir, L., Hamerle, A., Tutz, G., 1996. *Multivariate statistische Verfahren*, 2nd Edition. de Gruyter, Berlin; New York.
- Fahrmeir, L., Kneib, T., 2009. Propriety of posteriors in structured additive regression models: Theory and empirical evidence. *Journal of Statistical Planning and Inference* 139 (3), 843–859.
- Frackowiak, R. S., Friston, K. J., Frith, C. D., Dolan, R. J., Price, C. J., Zeki, S., Ashburner, J. T., Penny, W. D. (Eds.), 2004. *Human brain function*. Elsevier Academic Press, London.

- Friston, K. J., 2005. A theory of cortical responses. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 360 (1456), 815–836.
- Friston, K. J., Ashburner, J. T., Kiebel, S. J., Nichols, T. E., Penny, W. D., 2008. *Statistical parametric mapping - the analysis of functional brain images*. Academic Press.
- Friston, K. J., Fletcher, P., Josephs, O., Holmes, A. P., Rugg, M. D., Turner, R., 1998a. Event-related fMRI: Characterizing differential responses. *NeuroImage* 7 (1), 30–40.
- Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P., Williams, S., Frackowiak, R. S. J., Turner, R., 1995. Analysis of fMRI time-series revisited. *NeuroImage* 2 (1), 45–53.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-B., 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.
- Friston, K. J., Josephs, O., Rees, G., Turner, R., 1998b. Nonlinear event-related responses in fMRI. *Magnetic Resonance in Medicine* 39, 41–52.
- Friston, K. J., Penny, W. D., Phillips, C., Kiebel, S. J., Hinton, G., Ashburner, J. T., 2002. Classical and Bayesian inference in neuroimaging: Theory. *NeuroImage* 16 (2), 465–483.
- Fuchs, M., Kastner, J., Wagner, M., Hawes, S., Ebersole, J. S., 2002. A standardized boundary element method volume conductor model. *Clinical Neurophysiology* 113 (5), 702–712.
- Gelfand, A. E., Smith, A. F. M., 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85 (410), 398–409.
- Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B., 2004. *Bayesian data analysis*, 2nd Edition. Chapman & Hall/CRC, Boca Raton.
- Genovese, C. R., 2000. A Bayesian time-course model for functional magnetic resonance imaging data. *Journal of the American Statistical Association* 95 (451), 691–703.
- George, A., 1973. Nested dissection of a regular finite element mesh. *SIAM Journal on Numerical Analysis* 10 (2), 345–363.
- George, A., Liu, J. W. H., 1981. *Computer solution of large sparse positive definite systems*. Prentice-Hall, Englewood Cliffs, New Jersey.
- George, E. I., McCulloch, R. E., 1993. Variable selection via Gibbs sampling. *Journal of the American Statistical Association* 88 (423), 881–889.

- George, E. I., McCulloch, R. E., 1997. Approaches for Bayesian variable selection. *Statistica Sinica* 7 (2), 339–373.
- Gilks, W. R., Richardson, S., Spiegelhalter, D. J. (Eds.), 1996. *Markov chain Monte Carlo in practice*. Chapman & Hall/CRC, London.
- Glover, G. H., Li, T. Q., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance in Medicine* 44 (1), 162–167.
- Gössl, C., 2001. *Bayesian models in functional magnetic resonance imaging: Approaches for human brain mapping*. Ph.D. thesis, Ludwig-Maximilians-Universität, München.
- Gössl, C., Auer, D. P., Fahrmeir, L., 2000. Dynamic models in fMRI. *Magnetic Resonance in Medicine* 43 (1), 72–81.
- Gössl, C., Auer, D. P., Fahrmeir, L., 2001a. Bayesian spatiotemporal inference in functional magnetic resonance imaging. *Biometrics* 57 (2), 554–562.
- Gössl, C., Fahrmeir, L., Auer, D. P., 2001b. Bayesian modeling of the hemodynamic response function in BOLD fMRI. *NeuroImage* 14 (1), 140–148.
- Gotman, J., Kobayashi, E., Bagshaw, A. P., Bénar, C.-G., Dubeau, F., 2006. Combining EEG and fMRI: A multimodal tool for epilepsy research. *Journal of Magnetic Resonance Imaging* 23 (6), 906–920.
- Goto, K., Van De Geijn, R., 2008. High-performance implementation of the level-3 BLAS. *ACM Transactions on Mathematical Software* 35 (1), 1–14.
- Granert, O., 2012. *Rniftlib - R Interface to NIFTICLIB*. R package version 0.0-31.
- Groves, A. R., Chappell, M. A., Woolrich, M. W., 2009. Combined spatial and non-spatial prior for inference on MRI time-series. *NeuroImage* 45 (3), 795–809.
- Hämäläinen, M. S., Ilmoniemi, R. J., 1994. Interpreting magnetic fields of the brain: Minimum norm estimates. *Medical and Biological Engineering and Computing* 32 (1), 35–42.
- Henson, R., Rugg, M. D., Friston, K. J., 2001. The choice of basis functions in event-related fMRI. *NeuroImage* 13 (6), 149.
- Huettel, S. A., Song, A. W., McCarthy, G., 2004. *Functional magnetic resonance imaging*. Sinauer Associates, Sunderland, Massachusetts, USA.

- Hurn, M. A., Husby, O. K., Rue, H., 2003. A Tutorial on Image Analysis. In: Møller, J. (Ed.), *Spatial Statistics and Computational Methods*. Springer, pp. 87–117.
- Ipsen, I. C. F., Lee, D. J., 2003. Determinant approximations. Tech. Rep., North Carolina State University.
- Jezzard, P., Matthews, P. M., Smith, S. M., 2001. *Functional MRI - an introduction to methods*. Oxford University Press, New York.
- Josephs, O., Turner, R., Friston, K. J., 1997. Event-related fMRI. *Human Brain Mapping* 5, 243–248.
- Karypis, G., Kumar, V., 1998. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing* 20 (1), 359–392.
- Kecher, C., 2009. *UML 2 - Das umfassende Handbuch*. Galileo Computing.
- Kiebel, S. J., Daunizeau, J., Phillips, C., Friston, K. J., 2008. Variational Bayesian inversion of the equivalent current dipole model in EEG/MEG. *NeuroImage* 39 (2), 728–741.
- Kiehl, K. A., Stevens, M. C., Laurens, K. R., Pearlson, G., Calhoun, V. D., Liddle, P. F., 2005. An adaptive reflexive processing model of neurocognitive function: Supporting evidence from a large scale (n = 100) fMRI study of an auditory oddball task. *NeuroImage* 25 (3), 899–915.
- Kohn, R., Smith, M., Chan, D., 2001. Nonparametric regression using linear combinations of basis functions. *Statistics and Computing* 11, 313–322.
- Lang, S., Brezger, A., 2004. Bayesian P-Splines. *Journal of Computational and Graphical Statistics* 13 (1), 183–212.
- Laufs, H., Daunizeau, J., Carmichael, D. W., Kleinschmidt, A., 2008. Recent advances in recording electrophysiological data simultaneously with magnetic resonance imaging. *NeuroImage* 40 (2), 515–528.
- Laufs, H., Kleinschmidt, A., Beyerle, A., Eger, E., Salek-Haddadi, A., Preibisch, C., Krakow, K., 2003. EEG-correlated fMRI of human alpha activity. *NeuroImage* 19 (4), 1463–1476.
- Lazar, N. A., 2008. *The statistical analysis of functional MRI data*. Springer.
- Lehn, M., Stippler, A., Urban, K., 2005. FLENS: Flexible library for efficient numerical solutions. In: *Proceedings of Equadiff 11*. pp. 467–473.

- Liston, A. D., Lund, T. E., Salek-Haddadi, A., Hamandi, K., Friston, K. J., Lemieux, L., 2006. Modelling cardiac signal as a confound in EEG-fMRI and its application in focal epilepsy studies. *NeuroImage* 30 (3), 827–834.
- Liu, A. K., Belliveau, J. W., Dale, A. M., 1998. Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. *Proceedings of the National Academy of Sciences of the United States of America* 95 (15), 8945–8950.
- Luessi, M., Babacan, S. D., Molina, R., Booth, J. R., Katsaggelos, A. K., 2011. Bayesian symmetrical EEG/fMRI fusion with spatially adaptive priors. *NeuroImage* 55 (1), 113–132.
- Mandelkow, H., Halder, P., Boesiger, P., Brandeis, D., 2006. Synchronization facilitates removal of MRI artefacts from concurrent EEG recordings and increases usable bandwidth. *NeuroImage* 32 (3), 1120–1126.
- Matsumoto, M., Nishimura, T., 1998. Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation* 8 (1), 3–30.
- Mattout, J., Phillips, C., Penny, W. D., Rugg, M. D., Friston, K. J., 2006. MEG source localization under multiple constraints: An extended Bayesian framework. *NeuroImage* 30 (3), 753–767.
- Menon, V., Crottaz-Herbette, S., 2005. Combined EEG and fMRI studies of human brain function. *International Review of Neurobiology* 66, 291–321.
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. *Clinical Neurophysiology* 115 (10), 2195–2222.
- Mitchell, T. J., Beauchamp, J. J., 1988. Bayesian variable selection in linear regression. *Journal of the American Statistical Association* 83 (404), 1023–1032.
- Mohar, B., 2004. Graph Laplacians. In: Beineke, L. W., Wilson, R. J. (Eds.), *Topics in algebraic graph theory*. Cambridge University Press, Cambridge, U.K., pp. 113–136.
- Mosher, J. C., Lewis, P. S., Leahy, R. M., 1992. Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Transactions on Biomedical Engineering* 39 (6), 541–557.

- Mulert, C., Lemieux, L. (Eds.), 2010. EEG - fMRI: Physiological Basis, Technique, and Applications. Springer-Verlag, Berlin Heidelberg.
- Padhani, A. R., Ah-See, M.-L., Makris, A., 2005. MRI in the detection and management of breast cancer. *Expert Review of Anticancer Therapy* 5 (2), 239–252.
- Parker, G. J. M., Padhani, A. R., 2003. T1-w DCE-MRT: T1-weighted Dynamic Contrast-enhanced MRI. In: Tofts, P. (Ed.), *Quantitative MRI of the Brain*. Wiley, Chichester, England, pp. 341–364.
- Pascual-Marqui, R. D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Methods and Findings in Experimental and Clinical Pharmacology* 24 Suppl D, 5–12.
- Pascual-Marqui, R. D., Michel, C. M., Lehmann, D., 1994. Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology* 18 (1), 49–65.
- Penny, W. D., Trujillo-Barreto, N. J., Friston, K. J., 2005. Bayesian fMRI time series analysis with spatial priors. *NeuroImage* 24 (2), 350–62.
- Pettitt, A. N., Weir, I. S., Hart, A. G., 2002. A conditional autoregressive Gaussian process for irregularly spaced multivariate data with application to modelling large sets of binary data. *Statistics and Computing* 12, 353–367.
- Phillips, C., Rugg, M. D., Friston, K. J., 2002. Anatomically informed basis functions for EEG source localization: Combining functional and anatomical constraints. *NeuroImage* 16 (3), 678–695.
- Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News* 6 (1), 7–11.
- R Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Raftery, A. E., 1996a. Approximate Bayes factors and accounting for model uncertainty in generalised linear models. *Biometrika* 83 (2), 251–266.
- Raftery, A. E., 1996b. Hypothesis testing and model selection. In: Gilks, W., Richardson, S., Spiegelhalter, D. (Eds.), *Markov chain Monte Carlo in practice*. Chapman & Hall/CRC, pp. 163–187.

- Raftery, A. E., Madigan, D., Hoeting, J. A., 1997. Bayesian model averaging for linear regression models. *Journal of the American Statistical Association* 92 (437), 179–191.
- Reusken, A., 2001. Approximation of the determinant of large sparse symmetric positive definite matrices. *SIAM Journal on Matrix Analysis and Applications* 23 (3), 799–818.
- Ritter, P., Villringer, A., 2006. Simultaneous EEG-fMRI. *Neuroscience and Biobehavioral Reviews* 30 (6), 823–838.
- Rosa, M. J., Daunizeau, J., Friston, K. J., 2010. EEG-fMRI integration: A critical review of biophysical modeling and data analysis approaches. *Journal of Integrative Neuroscience* 9 (4), 453–476.
- Rosen, B. R., Buckner, R. L., Dale, A. M., 1998. Event-related functional MRI: Past, present, and future. *Proceedings of the National Academy of Sciences of the United States of America* 95 (3), 773–780.
- Rue, H., 2001. Fast sampling of Gaussian Markov random fields. *Journal of the Royal Statistical Society. Series B (Methodological)* 63 (2), 325–338.
- Rue, H., Held, L., 2005. *Gaussian Markov random fields: Theory and applications*. Chapman & Hall, London.
- Salek-Haddadi, A., Friston, K. J., Lemieux, L., Fish, D. R., 2003. Studying spontaneous EEG activity with fMRI. *Brain Research Reviews* 43, 110–133.
- Sarkar, D., 2008. *Lattice: Multivariate data visualization with R*. Springer, New York.
- Schabus, M., Dang-Vu, T. T., Albouy, G., Balteau, E., Boly, M., Carrier, J., Darsaud, A., Degueldre, C., Desseilles, M., Gais, S., Phillips, C., Rauchs, G., Schnakers, C., Sterpenich, V., Vandewalle, G., Luxen, A., Maquet, P., 2007. Hemodynamic cerebral correlates of sleep spindles during human non-rapid eye movement sleep. *Proceedings of the National Academy of Sciences of the United States of America* 104 (32), 13164–13169.
- Schwarz, G., 1978. Estimating the dimension of a model. *The Annals of Statistics* 6 (2), 461–464.
- Simpson, D. P., Turner, I. W., Pettitt, A. N., 2008. Fast sampling from a Gaussian Markov random field using Krylov subspace approaches. Tech. Rep., Queensland University of Technology.

- Smith, D., Smith, M., 2006. Estimation of binary Markov random fields using Markov chain Monte Carlo. *Journal of Computational and Graphical Statistics* 15 (1), 207–227.
- Smith, M., Fahrmeir, L., 2007. Spatial Bayesian variable selection with application to functional magnetic resonance imaging. *Journal of the American Statistical Association* 102 (478), 417–431.
- Smith, M., Kohn, R., 1996. Nonparametric regression using Bayesian variable selection. *Journal of Econometrics* 75 (2), 317–343.
- Smith, M., Pütz, B., Auer, D. P., Fahrmeir, L., 2003. Assessing brain activity through spatial bayesian variable selection. *NeuroImage* 20 (2), 802–815.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., Matthews, P. M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23 Suppl 1, S208–S219.
- Stone, M., 1979. Comments on model selection criteria of Akaike and Schwarz. *Journal of the Royal Statistical Society. Series B (Methodological)* 41 (2), 276–278.
- Strang, G., 1999. The discrete cosine transform. *SIAM Review* 41 (1), 135–147.
- Titterton, D. M., 2004. Bayesian methods for neural networks and related models. *Statistical Science* 19 (1), 128–139.
- Volpe, U., Mucci, A., Bucci, P., Merlotti, E., Galderisi, S., Maj, M., 2007. The cortical generators of P3a and P3b: a LORETA study. *Brain Research Bulletin* 73, 220–230.
- Wagner, M., Fuchs, M., 2001. Integration of functional MRI, structural MRI, EEG and MEG. *International Journal of Bioelectromagnetism* 3 (1), 1–15.
- Weir, I. S., Pettitt, A. N., 2000. Binary probability maps using a hidden conditional autoregressive Gaussian process with an application to Finnish common toad data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 49 (4), 473–484.
- Whaley, R. C., Petitet, A., 2005. Minimizing development and maintenance costs in supporting persistently optimized BLAS. *Software: Practice and Experience* 35 (2), 101–121.

- Woolrich, M. W., Behrens, T. E. J., Smith, S. M., 2004a. Constrained linear basis sets for HRF modelling using variational Bayes. *NeuroImage* 21 (4), 1748–1761.
- Woolrich, M. W., Jenkinson, M., Brady, J. M., Smith, S. M., 2004b. Fully Bayesian spatio-temporal modeling of fMRI data. *IEEE Transactions on Medical Imaging* 23 (2), 213–231.
- Worsley, K. J., Friston, K. J., 1995. Analysis of fMRI time-series revisited - again. *NeuroImage* 2 (3), 173–181.
- Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., Evans, A. C., 1996. A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping* 4 (1), 58–73.
- Zellner, A., 1986. On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In: Zellner, A., Goel, P. (Eds.), *Bayesian inference and decision techniques: Essays in honor of Bruno de Finetti*. North Holland, Amsterdam, pp. 233–243.
- Zschocke, S., Hansen, H.-C., 2012. *Klinische Elektroenzephalographie*. Springer, Berlin.

Danksagung

Zu allererst gilt mein Dank meinem Doktorvater Herrn Prof. Dr. Ludwig Fahrmeir. Aufgrund seines Vertrauens in mich und die zugrundeliegende Idee, gab er mir die Möglichkeit am Institut für Statistik der Ludwig-Maximilians-Universität München zu promovieren und mich damit ganz auf die statistische Methodenentwicklung zu konzentrieren. Sein Vertrauen war darüber hinaus mit einem hohen Maß an Freiheit bei der Ausgestaltung der Arbeit verbunden. Ohne seine vielfältige Unterstützung durch immerwährendes Interesse, Anregungen und Zuversicht wäre diese Arbeit nicht zustande gekommen.

Mein Dank gilt auch Herrn Dr. Philipp Sämann und Herrn Dr. Michael Czisch, die eine nicht wegzudenkende Stütze dieses Kooperationsprojektes darstellten. Sie waren nicht nur die Ideengeber für die zugrundeliegende Fragestellung dieser Arbeit, sondern stellten zusammen mit Sara Alessia Kiem und den Hilfskräften Manuela Hölzswimmer und Ludwig Bothmann mit beträchtlichem Arbeitsaufwand die Daten bereit. Zahlreiche Diskussionen halfen mir die praktischen Aspekte der Statistik nicht aus den Augen zu verlieren und gewährten mit einem wertvollen Einblick in die interdisziplinäre Forschung an der Schnittstelle von Medizin, Physik, Psychologie, Biologie und Statistik.

Des Weiteren bedanke ich mich herzlich bei Prof. Dr. Thomas Kneib und PD Dr. Christian Heumann für die wohlwollende Begleitung meines Promotionsverfahrens. Im Rahmen des Promotionsprogramms stand mir PD Dr. Christian Heumann auch als Zweitbetreuer mit Rat und Tat bei der Bewältigung rechnerischer Probleme zur Seite—ebenso wie Dr. Manuel Eugster.

Die vorliegende Arbeit entstand innerhalb eines Kooperationsprojektes zwischen der Arbeitsgruppe Neuroimaging (Max Planck Institut für Psychiatrie, München) und dem Institut für Statistik (Ludwig-Maximilians-Universität München), finanziert durch die Deutsche Forschungsgemeinschaft im Rahmen des Projektes *Biostatistische Modellierung und Analyse kombinierter fMRT- und EEG-Messungen*.

Ich möchte diesen Rahmen nutzen, um mich bei den Menschen zu bedanken, die einen nicht unbeträchtlichen Anteil zum Gelingen dieser Promotion beigetragen haben. In Freundschaft und Dankbarkeit für Ratschläge, Hilfeleistungen, Zuspruch, Anregungen und Ablenkungen im Laufe meines Promotionsprojekts bin ich Christiane Fuchs, Mareike Kohlmann, Ingrid Kreuzmair, Susanne Meßler, Julia Sommer, Sonja und Boris Stuck und—in besonderer Weise—Viola Obermeier verbunden.

Nicht zuletzt danke ich meiner Familie für ihre unentwegte, gleichermaßen selbstlose wie tatkräftige Unterstützung während meiner Promotion. Meinem Ehemann Franz bin ich zu tiefstem Dank verpflichtet. Sein stetiger Rückhalt ist das Fundament dieser Arbeit.

München, im Dezember 2012

Stefanie Kalus

Eidesstattliche Versicherung

(Gemäß Promotionsordnung vom 12.07.11, § 8, Abs. 2 Pkt. 5)

Hiermit erkläre ich an Eidesstatt, dass die vorliegende Dissertation von mir selbstständig, ohne unerlaubte Beihilfe angefertigt ist.

München, den 17. September 2012

Stefanie Kalus