EFFICIENT FULLY BAYESIAN APPROACHES TO BRAIN ACTIVITY MAPPING WITH COMPLEX-VALUED fMRI DATA: ANALYSIS OF REAL AND IMAGINARY COMPONENTS IN A CARTESIAN MODEL AND EXTENSION TO MAGNITUDE AND PHASE IN A POLAR MODEL

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Mathematical Statistics

---

by
Zhengxin Wang
May 2024

Accepted by:
Dr. D. Andrew Brown, Committee Chair
Dr. Xinyi Li, Committee Co-chair
Dr. Daniel B. Rowe
Dr. Xiaoqian Sun

Abstract

Functional magnetic resonance imaging (fMRI) plays a crucial role in neuroimaging, enabling the exploration of brain activity through complex-valued signals. Traditional fMRI analyses have largely focused on magnitude information, often overlooking the potential insights offered by phase data, and therefore, lead to underutilization of available data and flawed statistical assumptions. This dissertation proposes two efficient, fully Bayesian approaches for the analysis of complex-valued functional magnetic resonance imaging (cv-fMRI) time series.

Chapter 2 introduces the model, referred to as CV-sSGLMM, using the real and imaginary components of cv-fMRI data and sparse spatial generalized linear mixed model prior. This model extends the Cartesian model proposed by Lee et al. (2007) through the incorporation of Gaussian Markov random fields (GMRFs) and autoregressive models, enabling the capture of both spatial and temporal correlations in cv-fMRI data. Notably, CV-sSGLMM utilizes brain parcellation and parallel computation techniques to achieve reduction in computational time comparing with the current state-of-the-art, without sacrificing predictive accuracy.

Chapter 3 presents the model characterizing magnitude and phase of the cv-fMRI data, referred to as CV-M&P, which builds upon the polar model proposed by Rowe (2005a). This model provides a nuanced mapping of brain activity by mapping magnitude and phase activations individually in response to a stimulus. In doing so, it addresses a significant gap in the current literature—specifically, the lack of models that efficiently incorporate phase information through Bayesian methods.

Collectively, both models outperform existing approaches in some key predictive metrics and deepen our understanding of the inherent complexities of cv-fMRI signals. These advancements are promising for enhancing our understanding of healthy brain functions and may offer valuable insights for early diagnosis of neurological disorders.

Dedication

This work is dedicated to those living with brain diseases. Through advancements like fMRI, we strive for earlier diagnosis and more effective treatment strategies. May progress in this field bring hope and improved care for all affected.

Acknowledgments

First and foremost, I extend my deepest gratitude to Dr. Andrew Brown. His foundational teachings in statistics have paved the way for my academic progress. His guidance and consistent support have been instrumental in navigating the early stages of my research. I appreciate the autonomy he gave me to explore and learn. He never pushed me to do anything, but instead encouraged me to dive into different areas, even when they weren't directly related to our research.

My co-advisor, Dr. Xinyi Li, merits special mention for her intellectual contributions. Her expansive ideas and meticulous attention to refining writings have substantially enhanced the quality and rigor of my work.

I owe a significant portion of my academic achievements to Dr. Daniel Rowe. His unparalleled expertise in the domain of complex-valued fMRI signals and brain map detection has been crucial. He patiently equipped me with essential background knowledge, guided me in model improvement, meticulously completed derivations, checked results, and rectified errors. His hands-on guidance to my work has been invaluable. Without Dr. Rowe's precious support, this journey would not have been possible.

I would also like to thank Dr. Xiaoqian Sun. As a committee member, his insightful feedback during my comprehensive oral presentation played a pivotal role in shaping my research.

Lastly, I express my heartfelt gratitude to my wife. She has been my anchor, providing incomparable mental and emotional support throughout this journey. My favorite moment was to cook for her while waiting for the running of code. That's an expectation of both life and academy.

To all these remarkable individuals, I extend my most sincere thanks. Their wisdom, patience, and support have been the pillars of my Ph.D. journey.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is a non-invasive imaging technique that has revolutionized the field of medical diagnostics and research, particularly in the realm of neuroimaging. Developed in the late 20th century, MRI employs magnetic fields and radiofrequency pulses to create high-resolution, cross-sectional images of internal structures. It has become indispensable in a variety of applications ranging from diagnosis of pathological conditions to the investigation of complex physiological processes in the human brain.

The core technology of MRI relies on the phenomenon of nuclear magnetic resonance (NMR). In an MRI study, a gradient magnetic field is applied, primarily affecting the hydrogen protons in the brain due to their abundance. This magnetic field serves dual functions. Firstly, it aligns the spin of protons, and when a coil is placed adjacent to the brain, the spinning protons generate a changing magnetic field that induces a current in the coil, as per Faraday's law of electromagnetic induction. This induced electromotive force (EMF) serves as the initial signal captured by the MRI system. Subsequent signal processing techniques, such as demodulation, are employed to transform this raw EMF into a more interpretable form. Secondly, the gradient magnetic field spatially encodes the region of interest. The recorded time-domain signals are transformed into k-space representations through Fourier transformation methods. The signal in k-space is a representation of the proton spin densities (PSDs) across spatial domains, thus enabling image reconstruction through inverse Fourier transformations. Upon conversion back to the spatial domain, images are rendered based on contrasts in PSDs (Brown et al., 2014).

Functional magnetic resonance imaging (fMRI) is a subtype of MRI specifically optimized for higher temporal resolution and functional analyses. While conventional MRI captures static anatomical details, fMRI extends the utility by enabling the examination of metabolic functions over time.

1.2 Blood-Oxygenation-Level-Dependent Signal and Brain Activation

The blood-oxygenation-level-dependent (BOLD) signal serves as a proxy for neuronal activity, enabling fMRI to monitor dynamic changes in brain activity over time indirectly but effectively (Bandettini et al., 1992). In task-based fMRI experiments, subjects are exposed to specific intermittent stimuli, such as visual cues or motor tasks. Neuronal activation in targeted brain regions increases in response to these stimuli, elevating local oxygen consumption. This metabolic alteration triggers changes in the BOLD signal within the affected regions. Furthermore, the BOLD fluctuations impact local magnetic susceptibility, thereby affecting the resulting fMRI signal (Lindquist, 2008).

The existing literature has demonstrated that the expected BOLD response in a brain region undergoing activation can be accurately modeled by the convolution of a "boxcar" binary stimulus function with either a gamma or double-gamma hemodynamic response function (HRF) (Boynton et al., 1996; Lindquist et al., 2009). This "boxcar" stimulus comprises recurring on-off periods, as illustrated in Figure 1.1 in the upper section. Upon convolution with the double-gamma HRF, a modeled expected BOLD response is generated, which is represented in the lower section of Figure 1.1. Armed with this expected BOLD response, one can then identify regions within the brain where the temporal fluctuations in the fMRI signal significantly align with the modeled BOLD response. Such regions are thereby classified as being "active" in response to the stimulus.

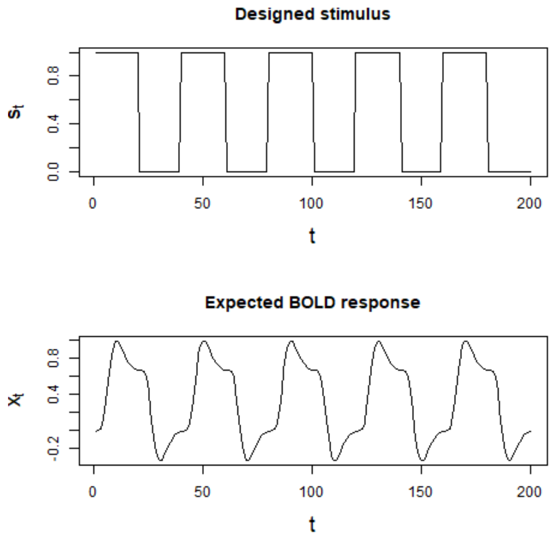


Figure 1.1: Designed stimulus and expected BOLD response.

1.3 Magnitude-Only and Complex-Valued fMRI Signal Analyses

MRI inherently generates complex-valued signals characterized by real and imaginary components. This complex structure arises from the forward and inverse Fourier transformations executed in the data collection process, which are affected by phase imperfections (Brown et al., 2014). Traditionally, fMRI studies that aim to map brain activity have predominantly focused on analyzing only the magnitude of these MR signals (Friston et al., 1994; Lindquist, 2008). The phase components are frequently disregarded during the preprocessing steps.

The magnitude-only analytical framework has its limitations. The first major limitation is that the omission of phase data results in the underutilization of valuable information that could be pertinent to understanding neurophysiological mechanisms. The second limitation, particularly relevant in studies that employ statistical modeling, concerns the statistical assumptions made during the identification of active voxels (volumetric pixels in the imaging data). In such analyses, a voxel is identified as "active" if the magnitude of its complex-valued fMRI (cv-fMRI) signal significantly varies with the expected BOLD response to a given stimulus. This practice assumes that the error terms in the models are normally distributed. However, while the original real and imaginary components may follow a normal distribution, the magnitude actually adheres to a Ricean distribution (Rice, 1944; Gudbjartsson and Patz, 1995). The Ricean distribution approximates a normal distribution only when the signal-to-noise ratio (SNR) is sufficiently high. Given that high SNRs are not universally guaranteed in fMRI studies, this statistical assumption becomes less reliable and exacerbates the limitation, consequently diminishing the power and reliability of the analysis.

In contrast, emerging research utilizing cv-fMRI data offers a more nuanced and comprehensive approach (Rowe and Logan, 2004, 2005; Rowe, 2005a,b; Lee et al., 2007; Rowe et al., 2007; Rowe, 2009; Adrian et al., 2018; Yu et al., 2018, 2023). By incorporating both the real and imaginary components of the MR signals, cv-fMRI studies pave the way for the development of more robust and statistically powerful models. These models are better to handle variations in SNR and can fully exploit the available data, thereby offering potentially deeper and more accurate insights into task-related neuronal activity.

1.4 Research Objectives

Existing approaches to fMRI data analysis exhibit limitations that either hinder the interpretability of findings or sacrifice the richness of the available data. For instance, non-model-based techniques like support vector machine (SVM; Wang et al., 2007), independent component analysis (ICA; Calhoun et al., 2002; Yu et al., 2015), and linear discriminant analysis (LDA; Mandelkow et al., 2016) fall short in capturing the inherent spatiotemporal correlations in fMRI time series. Moreover, although these methods offer some level of prediction accuracy, their lack of interpretability renders them suboptimal for comprehensive scientific inquiry.

Conversely, model-based strategies, particularly those employing a fully Bayesian framework, have the advantage of capturing spatiotemporal correlations naturally occurring in the fMRI data. However, most of these approaches have been applied solely to magnitude-only fMRI data (Zhang et al., 2015; Musgrove et al., 2016), thereby discarding half of information from complex-valued signals and introducing problematic statistical assumptions, as detailed in Section 1.3.

Emerging research in cv-fMRI signal analysis does show promise. These approaches attempt to address the limitations of magnitude-only studies by incorporating both the real and imaginary components of the MR signals. Despite these advances, many of these studies rely on hypothesis testing and lack the flexibility to formulate comprehensive spatiotemporal models (Rowe and Logan, 2004, 2005; Rowe, 2005a,b; Lee et al., 2007; Rowe et al., 2007; Rowe, 2009; Adrian et al., 2018), thereby leaving room for methodological improvement.

Given this landscape, we propose two fully Bayesian approaches for brain activity mapping using single-subject cv-fMRI time series. These models aim to achieve a confluence of important objectives:

- To effectively utilize both the real and imaginary components of the fMRI data, thereby overcoming the limitations associated with magnitude-only analyses.
- To capture the inherent spatiotemporal correlations within the fMRI data, thus enhancing the predictive accuracy of the models.
- To deliver computational efficiency, thereby making the models feasible for application in both academic and clinical settings.

Chapter 2 builds upon the Cartesian model proposed by Lee et al. (2007), with a focus on modeling both the real and imaginary components of cv-fMRI signals. Chapter 3 extends the polar model proposed by Rowe (2005a), aiming to characterize both the magnitude and phase components of cv-fMRI signals. These models are designed to not only accurately predict active voxels but also to quantify the intensity of their activation. Successful completion of these objectives will advance our understanding of healthy brain function and hold promise for the non-invasive and early diagnosis of neurological disorders, such as Alzheimer's disease (Machulda et al., 2003).

Chapter 2

Efficient Fully Bayesian Approaches to Brain Activity Mapping with Complex-Valued fMRI Data: Analysis of Real and Imaginary Components in a Cartesian Model

2.1 Introduction

In this chapter, we propose a fully Bayesian approach based on Lee et al. (2007)'s Cartesian model for brain activity mapping using single-subject cv-fMRI time series. Specifically, we aim to determine which voxels' fMRI signal, in terms of either real or imaginary components, change significantly in response to a particular task, as well as the amount of the change. Our proposed approach uses autoregressive models for the temporal correlations and Gaussian Markov random fields (GMRFs; Rue and Held, 2005) to capture spatial associations in the cv-fMRI data. Moreover, we employ image partitioning and parallel computation to facilitate computationally efficient Markov chain Monte Carlo (MCMC; Gelfand and Smith, 1990) algorithms.

The remainder of the chapter is organized as follows. Section 2.2 details our proposed model, outlines the priors and posteriors, and explains our strategy for brain partitioning. We demonstrate estimation and inference in Section 2.3, where we use simulated datasets to test the performance of our model in terms of the determination of brain activity maps. Section 2.4 shows the results of implementing our proposed approach on cv-fMRI data obtained from a real finger-tapping experiment. Lastly, Section 2.5 summarizes our findings, highlights our contributions, and outlines potential work for future research in this domain.

2.2 Model

In this section, we present our model for brain activity mapping with cv-fMRI data, including an equivalent real-valued representation. We also describe the brain parcellation strategy for parallel computation. We derive the posterior distribution of the parameters of interest, as well as an MCMC algorithm for accessing it.

2.2.1 Model Formulation

The fMRI signals, both real- and complex-valued, are known to exhibit temporal correlations. This can be captured by autoregressive (AR) error structure. Thus, our complex-valued model is based on that proposed by Lee et al. (2007), with some modifications. For the v^{th} voxel, $v = 1, ..., V$, the measured signal is modeled as

$$\mathbf{y}^v = \mathbf{x}\beta^v + \mathbf{r}^v\rho^v + \boldsymbol{\varepsilon}^v, \tag{2.1}$$

where all terms are complex-valued except $\mathbf{x}$. The term $\mathbf{y}^v \in \mathbb{C}^T$ is the vector of signals at voxel v collected at evenly-spaced time points, where T is the total observed time points, and $\mathbf{x} \in \mathbb{R}^T$ is the vector of the expected BOLD response associated with a particular stimulus, with $\beta^v \in \mathbb{C}$ the associated regression coefficient. We assume that low-frequency trends in $\mathbf{y}^v$ have been removed by preprocessing, and that both $\mathbf{y}^v$ and $\mathbf{x}$ are centered. The term $\mathbf{r}^v \in \mathbb{C}^T$ is the vector of lag-1 prediction errors for the assumed AR(1) model, with $\rho^v \in \mathbb{C}$ the scalar autoregression coefficient. The AR(1) model has been shown to often be sufficient for capturing temporal dynamics in fMRI data Cox (1996). We suppose that the error term $\boldsymbol{\varepsilon}^v$ follows the standard complex normal distribution, that is, $\boldsymbol{\varepsilon}^v \sim \mathcal{CN}_T(\boldsymbol{\mu}^v = \mathbf{0}, \boldsymbol{\lambda}^v = 2\sigma_v^2\mathbf{I}, \mathbf{C}^v = \mathbf{0})$, where $\mathcal{CN}_T$ denotes a complex normal distribution of dimension T with mean $\boldsymbol{\mu}^v$, complex-valued, Hermitian and non-negative definite covariance matrix $\boldsymbol{\lambda}^v$, and complex-valued symmetric relation matrix $\mathbf{C}^v$. In Appendix A.1, we provide details similar to those presented by Rowe (2009) that demonstrate the equivalence between the model of Lee et al. (2007) and the cv-fMRI model proposed by Rowe and Logan (2004) with constant phase.

Picinbono (1996) and Yu et al. (2018) provide an equivalent real-valued representation of model (2.1) as

$$\underbrace{\begin{pmatrix} \mathbf{y}_{Re}^v \\ \mathbf{y}_{Im}^v \end{pmatrix}}_{\mathbf{y}_r^v} = \underbrace{\begin{pmatrix} \mathbf{x} & \mathbf{0} \\ \mathbf{0} & \mathbf{x} \end{pmatrix}}_{\mathbf{X}_r} \underbrace{\begin{pmatrix} \beta_{Re}^v \\ \beta_{Im}^v \end{pmatrix}}_{\boldsymbol{\beta}_r^v} + \underbrace{\begin{pmatrix} r_{Re}^v & -r_{Im}^v \\ r_{Im}^v & r_{Re}^v \end{pmatrix}}_{\mathbf{R}_r^v} \underbrace{\begin{pmatrix} \rho_{Re}^v \\ \rho_{Im}^v \end{pmatrix}}_{\boldsymbol{\rho}_r^v} + \underbrace{\begin{pmatrix} \boldsymbol{\varepsilon}_{Re}^v \\ \boldsymbol{\varepsilon}_{Im}^v \end{pmatrix}}_{\boldsymbol{\varepsilon}_r^v}, \tag{2.2}$$

where all terms are real-valued. Using the symbols in the underbraces, this is more concisely written as

$$\mathbf{y}_r^v = \mathbf{X}_r\boldsymbol{\beta}_r^v + \mathbf{R}_r^v\boldsymbol{\rho}_r^v + \boldsymbol{\varepsilon}_r^v, \qquad \boldsymbol{\varepsilon}_r^v \sim \mathcal{N}_{2T}(\mathbf{0}, \boldsymbol{\Sigma}^v),$$

where

$$\boldsymbol{\Sigma}^v = \begin{pmatrix} \boldsymbol{\Sigma}_{Re,Re}^v & \boldsymbol{\Sigma}_{Re,Im}^v \\ \boldsymbol{\Sigma}_{Im,Re}^v & \boldsymbol{\Sigma}_{Im,Im}^v \end{pmatrix},$$

and

$$\begin{aligned} \boldsymbol{\Sigma}_{Re,Re}^v &= \tfrac{1}{2}Re(\boldsymbol{\lambda}^v + \mathbf{C}^v) = \sigma_v^2\mathbf{I}_T, \quad & \boldsymbol{\Sigma}_{Re,Im}^v &= \tfrac{1}{2}Im(-\boldsymbol{\lambda}^v + \mathbf{C}^v) = \mathbf{0}_T, \\ \boldsymbol{\Sigma}_{Im,Re}^v &= \tfrac{1}{2}Im(\boldsymbol{\lambda}^v + \mathbf{C}^v) = \mathbf{0}_T, \quad & \boldsymbol{\Sigma}_{Im,Im}^v &= \tfrac{1}{2}Re(\boldsymbol{\lambda}^v - \mathbf{C}^v) = \sigma_v^2\mathbf{I}_T. \end{aligned}$$

Observe that our assumption on the covariance structure here simply means that $\boldsymbol{\Sigma}^v = \sigma_v^2\mathbf{I}_{2T}$. We assign the voxel-specific variances σ_v^2 and autoregression coefficient $\boldsymbol{\rho}_r^v$ Jeffreys prior and uniform prior, respectively. That is, $p(\sigma_v^2) = 1/\sigma_v^2$ and $p(\boldsymbol{\rho}_r^v) = 1$, for $v = 1, \ldots, V$.

2.2.2 Brain Parcellation and Spatial Priors

In addition to temporal dependence, fMRI signals also exhibit spatial associations. These spatial dependencies can originate from several sources, including the inherent noise of the data

(Krüger and Glover, 2001), unmodeled neuronal activation (Bianciardi et al., 2009), and preprocessing steps, such as spatial normalization (Friston et al., 1995), image reconstruction (Rowe et al., 2009), and spatial smoothing (Mikl et al., 2008). Hence voxels, as artificial partitions of the human brain, often exhibit behavior similar to that of their neighbors. These spatial dependencies can be modeled by imposing spatial structure in the prior on β^v or the hyperparameters in such priors.

Brain Parcellation Musgrove et al. (2016) propose a brain parcellation technique that seeks to identify active voxels within each parcel/partition, and subsequently combines these results to generate a comprehensive whole-brain activity map. The authors partition their brain images into initial parcels of size approximately 500 voxels each. If a parcel is found to be too large or too small, it is broken down into voxels and these voxels are merged into adjacent parcels while ensuring the merged parcels contain less than 1000 voxels each. Alternatively, the partitioning strategy could be based on anatomical atlases such as Brodmann areas (Amunts et al., 2000; Tzourio-Mazoyer et al., 2002), or based on equal geometric size in the image rather than equal numbers of contained voxels. Musgrove et al. (2016) remark that this method of partitioning induces negligible edge effects, that is, the classification of voxels on the borders of parcels is not strongly affected.

In our study, we partition the two- or three-dimensional fMRI image into G parcels of approximately equal geometric size. We then process each parcel independently using the same model and method, facilitating parallel computation and hence computational efficiency. We find that our parcellation strategy incurs minimal edge effects, echoing the observations of Musgrove et al. (2016). We discuss the optimal number of parcels and corresponding number of voxels in each parcel in Section 2.3.

Prior Distribution of β^v For parcel g, $g = 1, \ldots, G$, containing V_g voxels, a voxel v ($v = 1, \ldots, V_g$) is classified as an active voxel under the stimulus if its regression coefficient of slope $\beta^v = \beta^v_{Re} + i\beta^v_{Im} \neq 0$, where i is the imaginary unit. As this is a variable selection problem, we use a spike-and-slab prior (Mitchell and Beauchamp, 1988; Yu et al., 2018):

$$\beta^v \mid \lambda_v \sim \lambda_v \mathcal{CN}_1(0, 2\tau_g^2, 0) + (1 - \lambda_v)\mathcal{I}_0, \tag{2.3}$$

where $\mathcal{I}_0$ denotes the point mass at 0. The binary indicator $\lambda_v \in \{0, 1\}$ reflects the status of a voxel. Specifically, $\lambda_v = 1$ indicates that voxel v is responding to the task, while $\lambda_v = 0$ otherwise. We take

$\tau_g^2 \in \mathbb{R}$ to be constant across all voxels within each parcel. Yu et al. (2018) shows that a real-valued representation of (2.3) is given by:

$$\boldsymbol{\beta}_r^v = \begin{pmatrix} \beta_{Re}^v \\ \beta_{Im}^v \end{pmatrix} \mid \lambda_v \sim \mathcal{N}_2(\mathbf{0},\ \lambda_v \tau_g^2 \mathbf{I}).$$

The parcel specific variances τ_g^2 are assigned a Jeffreys prior, $p(\tau_g^2) = 1/\tau_g^2,\ g = 1, \ldots, G$.

Spatial Prior on λ_v To further reduce computational effort and to capture pertinent spatial structure with a low-dimensional representation, we employ the sparse spatial generalized linear mixed model (sSGLMM) prior, as developed by Hughes and Haran (2013) and Musgrove et al. (2016), which is in turn an extension of the the prior proposed by Reich et al. (2006). Such priors use GMRFs and reduce the dimension by examining the spectra of the associated Markov graphs. For voxel v $(v = 1, ..., V_g)$ within parcel g $(g = 1, ..., G)$, we suppose that

$$\begin{aligned} \lambda_v \mid \eta_v &\overset{iid}{\sim} \mathcal{B}ern\left\{\boldsymbol{\Phi}(\psi + \eta_v)\right\}, \\ \eta_v \mid \boldsymbol{\delta}_g &\sim \mathcal{N}_1\left(\mathbf{m}_v' \boldsymbol{\delta}_g, 1\right), \\ \boldsymbol{\delta}_g \mid \kappa_g &\sim \mathcal{N}_q\left\{\mathbf{0}, (\kappa_g \mathbf{M}_g{}' \mathbf{Q}_g \mathbf{M}_g)^{-1}\right\}, \\ \kappa_g &\sim \mathcal{G}amma\left(a_\kappa, b_\kappa\right), \end{aligned} \tag{2.4}$$

where $\boldsymbol{\Phi}(\cdot)$ denotes the CDF of standard normal distribution and $\psi \in \mathbb{R}$ is a fixed tuning parameter. The terms $\mathbf{m}_v'$, $\mathbf{M}_g$, and $\mathbf{Q}_g$ are derived from the adjacency matrix $\mathbf{A}_g$ of parcel g. The adjacency matrix $\mathbf{A}_g \in \{0,1\}^{V_g \times V_g}$ is such that $\mathbf{A}_{g,uv} = 1$ if voxels u and v are neighbors in the image, and 0 otherwise, where "neighbor" is defined by the user. Typically, voxels that share an edge or a corner are taken to be neighbors. The matrix $\mathbf{M}_g \in \mathbb{R}^{V_g \times q}$ contains the first q principal eigenvectors of $\mathbf{A}_g$, typically with $q \ll V_g$. The term $\mathbf{m}_v'$ is a $1 \times q$ row vector of "synthetic spatial predictors" (Hughes and Haran, 2013) corresponding to the v^{th} row of $\mathbf{M}_g$. The matrix $\mathbf{Q}_g = \text{diag}(\mathbf{A}_g \mathbf{1}_{V_g}) - \mathbf{A}_g$ is the graph Laplacian. The term $\boldsymbol{\delta}_g$ is a $q \times 1$ vector of spatial random effects, and κ_g is the spatial smoothing parameter.

The design of the prior distribution for binary indicator λ_v aims to capture both spatial dependencies and the sparsity of active voxels. This reflects the hypothesis that a voxel is more likely to be active/inactive if their neighboring voxels are also active/inactive (Friston et al., 1994;

Smith and Fahrmeir, 2007). Furthermore, in the context of simple tasks, only a small percentage of voxels across the entire brain are expected to be active (Rao et al., 1996; Epstein and Kanwisher, 1998). Thus the sSGLMM prior is well-suited to the work and compatible with the parcellation approach. Hughes and Haran (2013) remark that $\mathbf{M}_g$ is capable of capturing smooth patterns of spatial variation at various scales.

The parameters ψ, q, a_κ, and b_κ are fixed *a priori* and determined based on several factors. In our simulation studies, we examine various values of ψ to identify the one providing the highest prediction accuracy. For real human datasets, the initial value of ψ is set to $\mathbf{\Phi}^{-1}(0.02) = -2.05$ for all voxels, following the suggestion of Musgrove et al. (2016). This value can be further adjusted based on the proportion of active voxels detected in previous experiments. We set $q = 5$ (when V_g is approximately 200) per Hughes and Haran (2013), indicating that such a reduction is often feasible. We find there is no detectable difference using larger q. The shape and scale parameters of the gamma distribution, $a_\kappa = \frac{1}{2}$ and $b_\kappa = 2000$ respectively, are selected to yield a large mean for κ_g ($a_\kappa b_\kappa$=1000). This choice serves to reduce the chances of creating misleading spatial structures in the posterior distribution, mitigating the risk of identifying spurious brain activity patterns that could be attributed to noise or other confounding factors.

2.2.3 MCMC Algorithm and Posterior Distributions

We use Gibbs sampling to obtain the joint and marginal posterior distributions of parameters of interest. The necessary full conditional distributions and derivations are outlined in Appendix A.2. The fixed-width approach proposed by Flegal et al. (2008) is used to diagnose convergence. Specifically, we consider the algorithm to have converged if the Monte Carlo standard error (MCSE) of any λ_v is less than 0.05. In our numerical studies that follow, we run 10^3 iterations. We take the means of the sampled parameters (after discarding burn-in iterations) as the point estimates. Active voxels are determined by $\widehat{\lambda}_v > 0.8722$ (Smith and Fahrmeir, 2007), and $\widehat{\beta}^v_{Re}$ and $\widehat{\beta}^v_{Im}$ are used to construct the estimated strength maps, indicating how strongly they react with the simulus, computed as $\sqrt{(\widehat{\beta}^v_{Re})^2 + (\widehat{\beta}^v_{Im})^2}$.

2.3 Simulation Studies

In this section, we simulate three types of two-dimensional complex-valued time series of fMRI signals: data with *iid* noise, data with noise following AR(1) temporal dependence, and a more realistic simulated *iid* dataset imitating the human brain. We evaluate three models based on their performance in both classification and estimation fidelity. The models under consideration include:

- The model of Musgrove et al. (2016), which uses a sSGLMM prior for magnitude-only data and incorporates brain parcellation (denoted as MO-sSGLMM).

- The model of Yu et al. (2018) for cv-fMRI, which does not incorporate a spatial prior or brain parcellation (denoted as CV-nonSpatial). In this model, the prior for λ_v in model (2.4) is taken to be $\lambda_v \mid \eta_v \overset{iid}{\sim} \mathcal{B}ern(\eta_v), \eta_v \sim \mathcal{B}eta(1,1)$.

- Our proposed model, which uses an sSGLMM prior for complex-valued data and incorporates brain parcellation (denoted as CV-sSGLMM).

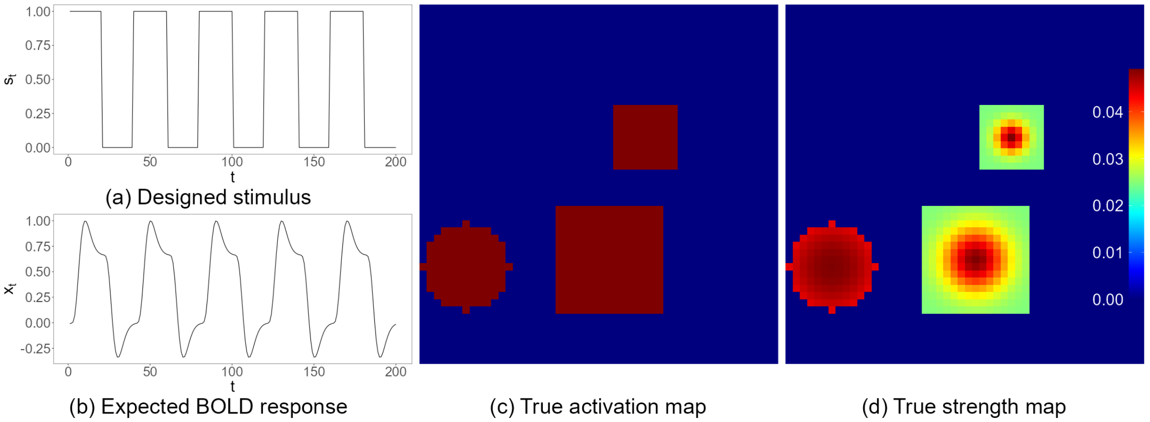
All three models are fully Bayesian, suitable for autoregressive noise, and leverage Gibbs sampling to approximate their respective posterior distributions. Both MO-sSGLMM and CV-sSGLMM use the best combination of parcel number G and tuning parameter ψ in terms of the prediction accuracy ($G = 9$ and $\psi = \Phi^{-1}(0.47)$ for both), and determine the active voxels by thresholding at $\widehat{\lambda}_v > 0.8722$. The CV-nonSpatial model uses a threshold of 0.5, as suggested by Yu et al. (2018).

Following the model comparisons, we concentrate on our proposed CV-sSGLMM model to examine the impacts of the tuning parameter ψ, the number of parcels G, and the length of time series T. Additional results for marginal posterior distributions, time series, and phase are provided in Appendix A.3.

All of the results are generated by running the code on a custom-built desktop computer with an Intel Core i9-9980XE CPU (3.00GHz, 3001 Mhz, 18 cores, 36 logical processors), NVIDIA GeForce RTX 2080 Ti GPU, 64 GB RAM, and operating on Windows 10 Pro.

2.3.1 Simulated Datasets with IID Noise and AR(1) Noise

We discuss how we generate the true maps and simulate fMRI signals here, followed by the results.



(a) Designed stimulus

(b) Expected BOLD response

(c) True activation map

(d) True strength map

Figure 2.1: (a) Designed stimulus; (b) Expected BOLD response; (c) True activation map; (d) True strength map.

Designed Stimulus, Expected BOLD Response, and True Activation/Strength map We use the same pattern of stimulus as simulated by Yu et al. (2018). The designed stimulus is a binary signal $\mathbf{s}$ consisting of five epochs, each with a duration of 40 time points, resulting in a total of $T = 200$ time points. Within each epoch, the stimulus is turned on and off for an equal duration of 20 time points. The expected BOLD response, denoted as $\mathbf{x}$, is generated by convolving the stimulus signal with a double-gamma HRF. Both the designed stimulus and expected BOLD response, depicted in Figures 2.1a and 2.1b, are shared for all simulated datasets.

To simulate 100 replicates on a 50×50 panel, we use the `specifyregion` function in the `neuRosim` library (Welvaert et al., 2011) in `R` (R Core Team, 2023). Each map features three non-overlapping active regions with varying characteristics such as centers, shapes, radii, and decay rates as shown in Table 2.1. The central voxel of an active region has a strength of one, indicating it reacts strongest with the stimulus, while the strengths of the surrounding active voxels decrease based on their distance to the center and the decay rate ϱ. These strengths are further scaled by a multiplier of 0.04909 (which determines the contrast-to-noise ratio via Eq. (2.5)), yielding a range of 0 to 0.04909. Examples of the true activation map and true strength map are shown in Figures 2.1c and 2.1d.

Simulating fMRI Signals with Non-AR Noise and AR(1) Noise We simulate 100 datasets with *iid* noise using the expected BOLD response and each true strength map for CV-nonSpatial

Table 2.1: Characteristics of true maps.

Map size	Number of active regions	Radius	Shape	Decay rate (ϱ)
50×50	3	2 to 6	sphere or cube	0 to 0.3

and CV-sSGLMM. We then extract the moduli to use with MO-sSGLMM. The cv-fMRI signal of voxel v at time t is simulated by:

$$\begin{aligned} y^v_{t,Re} &= (\beta_0 + \beta^v_1 x_t)\cos(\theta) + \varepsilon^v_{t,Re}, \quad \varepsilon^v_{t,Re} \sim \mathcal{N}(0, \sigma^2), \\ y^v_{t,Im} &= (\beta_0 + \beta^v_1 x_t)\sin(\theta) + \varepsilon^v_{t,Im}, \quad \varepsilon^v_{t,Im} \sim \mathcal{N}(0, \sigma^2), \end{aligned} \tag{2.5}$$

where x_t represents the expected BOLD response from Figure 2.1b at time t, and β^v_1 refers to the true strength of voxel v taken from Figure 2.1d. The phase, θ, is set to be the constant $\pi/4$, and σ is set to the constant 0.04909. As a result, the maximum contrast-to-noise ratio (CNR) is $\max \beta^v_1/\sigma = 1$. We determine the intercept β_0 based on the signal-to-noise ratio (SNR) such that SNR $= \beta_0/\sigma = 10$, leading to $\beta_0 = 0.4909$.

Next, we generate 100 datasets with AR(1) noise in a similar manner as Eq. (2.5). The difference lies in the simulation of error terms, which is done so that

$$\begin{pmatrix} \varepsilon^v_{t,Re} \\ \varepsilon^v_{t,Im} \end{pmatrix} = \begin{pmatrix} 0.2 & -0.9 \\ 0.9 & 0.2 \end{pmatrix} \begin{pmatrix} \varepsilon^v_{t-1,Re} \\ \varepsilon^v_{t-1,Im} \end{pmatrix} + \begin{pmatrix} \xi^v_{Re} \\ \xi^v_{Im} \end{pmatrix}, \quad \begin{pmatrix} \xi^v_{Re} \\ \xi^v_{Im} \end{pmatrix} \sim \mathcal{N}_2\left(\mathbf{0}, \sigma^2 \mathbf{I}\right).$$

This is a real-valued equivalent of the complex AR(1) error model:

$$\varepsilon^v_t = (0.2 + 0.9i)\varepsilon^v_{t-1} + \xi_v, \quad \xi_v \sim \mathcal{CN}_1(0, 2\sigma^2, 0). \tag{2.6}$$

Results Results from our simulations are displayed in Figure 2.2, which depicts the estimated maps for a single dataset. The yellow grid lines correspond to the partitions in cases of brain parcellation. The performance across the three models reveals a consistent trend. All models perform well for the *iid* case, while MO-sSGLMM fails to detect any activity in the presence of the AR(1) noise. This is because the complex-valued AR structure in equation (2.6) cannot be recovered after extracting the moduli of the data. Further quantitative results, such as the receiver operating characteristic area under curve (ROC-AUC), true vs estimated strength regression slope, true vs estimated strength concordance correlation coefficient (CCC), and true vs estimated strength pairwise mean square error (X-Y pairwise MSE), are illustrated in Figure 2.3. These offer a comprehensive performance evaluation in terms of classification and estimation. Figure 2.3 shows similar comparative performance as can be gleaned from Figure 2.2. All procedures do well in the presence of *iid* noise, whereas both complex-valued models considerably outperform the magnitude-only model when the errors are correlated. In each case, we can observe slightly better MSE, CCC, and estimation fidelity (Figure 2.3(b), (c), (d), (f), (g), (h)), but these are small when compared to the outperformance of the complex-valued models versus magnitude only.

Table 2.2 summarizes the average metrics across 100 *iid* noise and 100 AR(1) noise replicated datasets. In the *iid* case, the F1-score, slope, CCC, and X-Y MSE clearly favor MO-sSGLMM, followed by our CV-sSGLMM, and CV-nonSpatial ranks last. This demonstrates the proficiency of MO-sSGLMM on datasets where the necessity to capture complex-valued noise dependence is not crucial. The ROC-AUC score of MO-sSGLMM is comparable to that of CV-nonSpatial, and slightly surpasses that of our proposed CV-sSGLMM.

In the analysis of AR(1) datasets, our proposed CV-sSGLMM shows a clear advantage over the two competitors. Due to MO-sSGLMM's limitations already shown, we focus our comparison here between CV-nonSpatial and CV-sSGLMM. The CV-sSGLMM outperforms CV-nonSpatial across multiple metrics, such as F1-score, slope, CCC, and X-Y MSE. The superior performance of the CV-sSGLMM in terms of both classification and estimation can be attributed to the inclusion of the sSGLMM prior. In addition to our results, the value of using spatial priors to enhance the model's performance on correlated datasets has been demonstrated by Yu et al. (2023). Perhaps the most notable and favorable performance of our proposed model is in the vastly computational efficiency due to the brain parcellation and parallel computation, 5.39 seconds with CV-sSGLMM versus 42.2 seconds for the CV-nonSpatial. In other words, we obtain results as good or better than current state-of-the-art, but are able to do so 87% faster.

Effects of Experimental and Parameter Settings on CV-sSGLMM The performance of our CV-sSGLMM is determined in part by three choices: the tuning parameter ψ, the parcel number G, and the time length T. Here we assess their influence using the AR(1) data exclusively. For a single dataset, estimated activation maps generated from varying these settings are depicted in Figure 2.4, with their corresponding estimated strength maps displayed in Figure 2.5. A summary of average

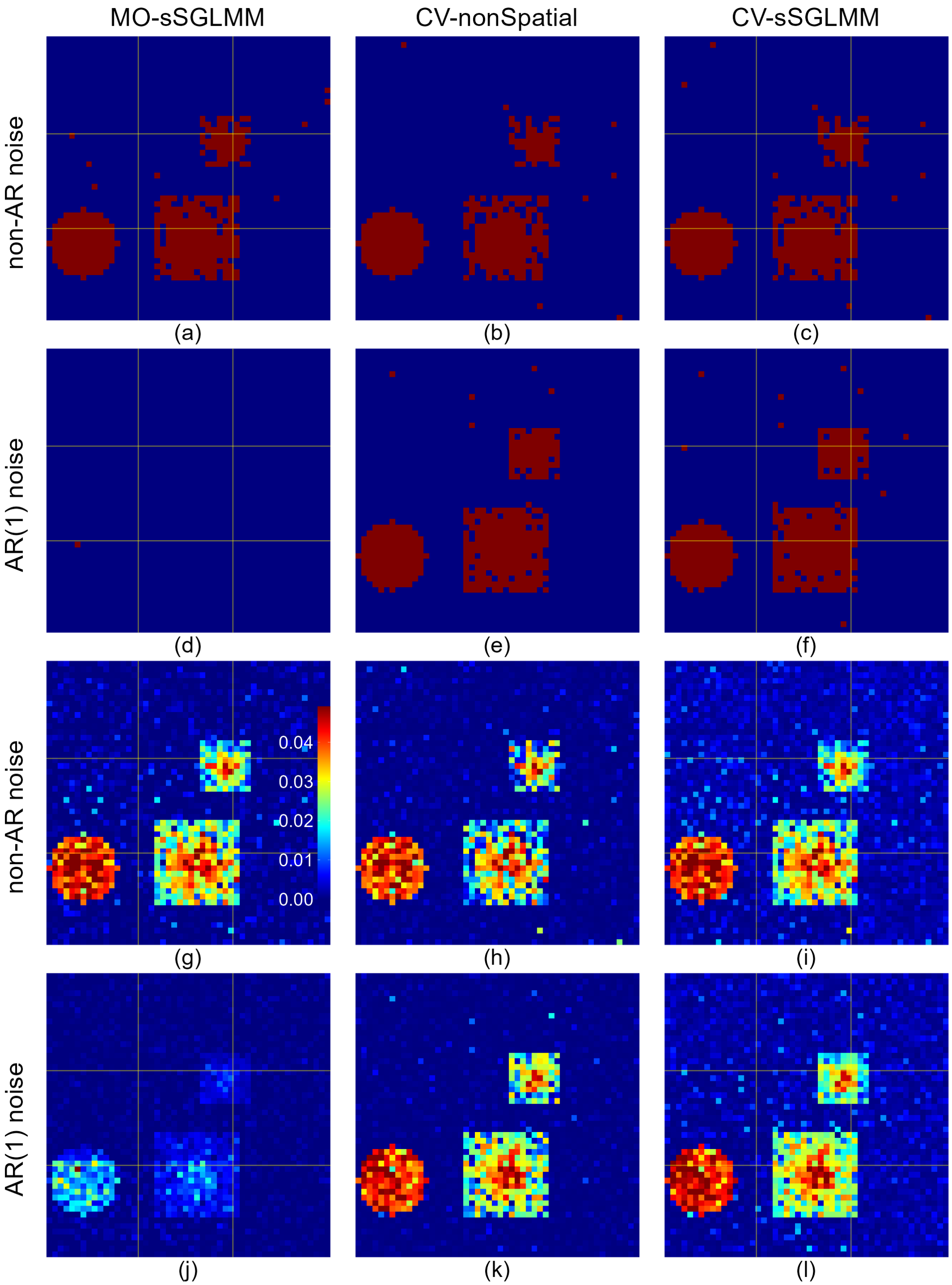


Figure 2.2: (a)-(c) are estimated activation maps for a non-AR dataset as produced by the MO-sSGLMM, CV-nonSpatial, and CV-sSGLMM models, respectively. (d)-(f) are estimated activation maps for an AR(1) dataset, as derived from the same models. (g)-(l) are the corresponding estimated strength maps.

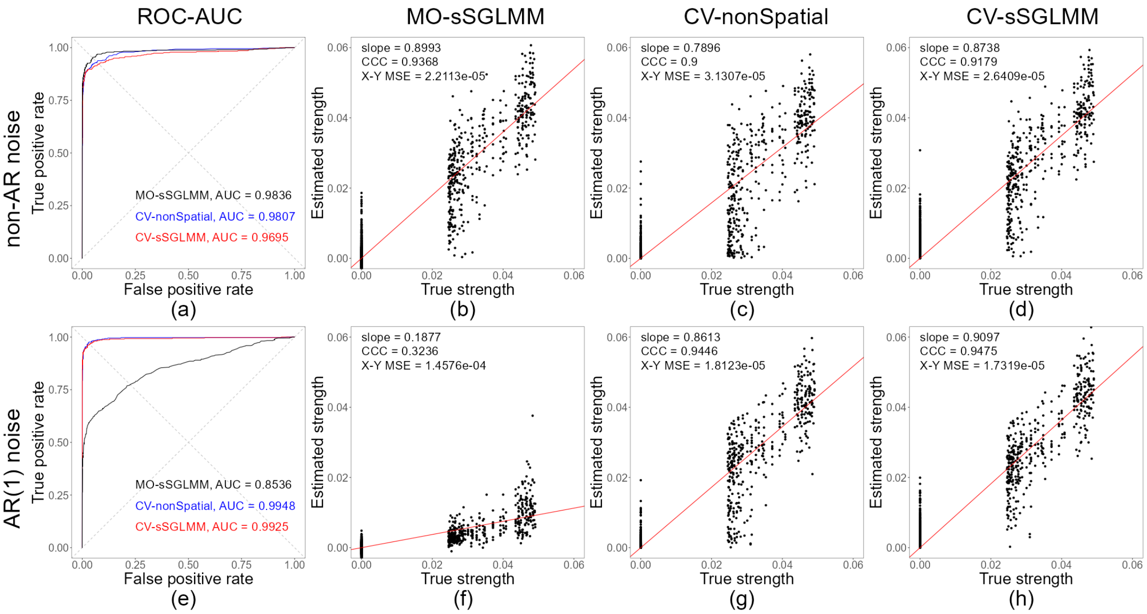


Figure 2.3: (a)-(d) are the ROC curves and plots comparing true versus estimated strength for a non-AR dataset. (e)-(h) are analogous plots for an AR(1) dataset.

Table 2.2: Summary of average metrics across 100 non-AR and 100 AR(1) datasets produced by the MO-sSGLMM, CV-nonSpatial, and CV-sSGLMM models.

AR type	Mode	Accuracy	Precision	Recall	F1 Score	AUC	Slope	CCC	X-Y MSE	Time (s)
non-AR	MO-sSGLMM	**0.9693**	0.9440	**0.8160**	**0.8741**	**0.9774**	**0.8586**	**0.9008**	**2.06e-5**	**2.4**
	CV-nonSpatial	0.9540	**0.9632**	0.6687	0.7853	0.9751	0.6771	0.8222	3.04e-5	41.9
	CV-sSGLMM	0.9622	0.9277	0.7742	0.8424	0.9625	0.8186	0.8627	2.54e-5	5.51
AR(1)	CV-nonSpatial	0.9765	**0.9733**	0.8407	0.9012	**0.9927**	0.8040	0.9096	1.69e-5	42.2
	CV-sSGLMM	**0.9797**	0.9381	**0.9039**	**0.9201**	0.9879	**0.8816**	**0.9145**	**1.60e-5**	**5.39**

metrics over 100 replicated datasets is shown in Table 2.3.

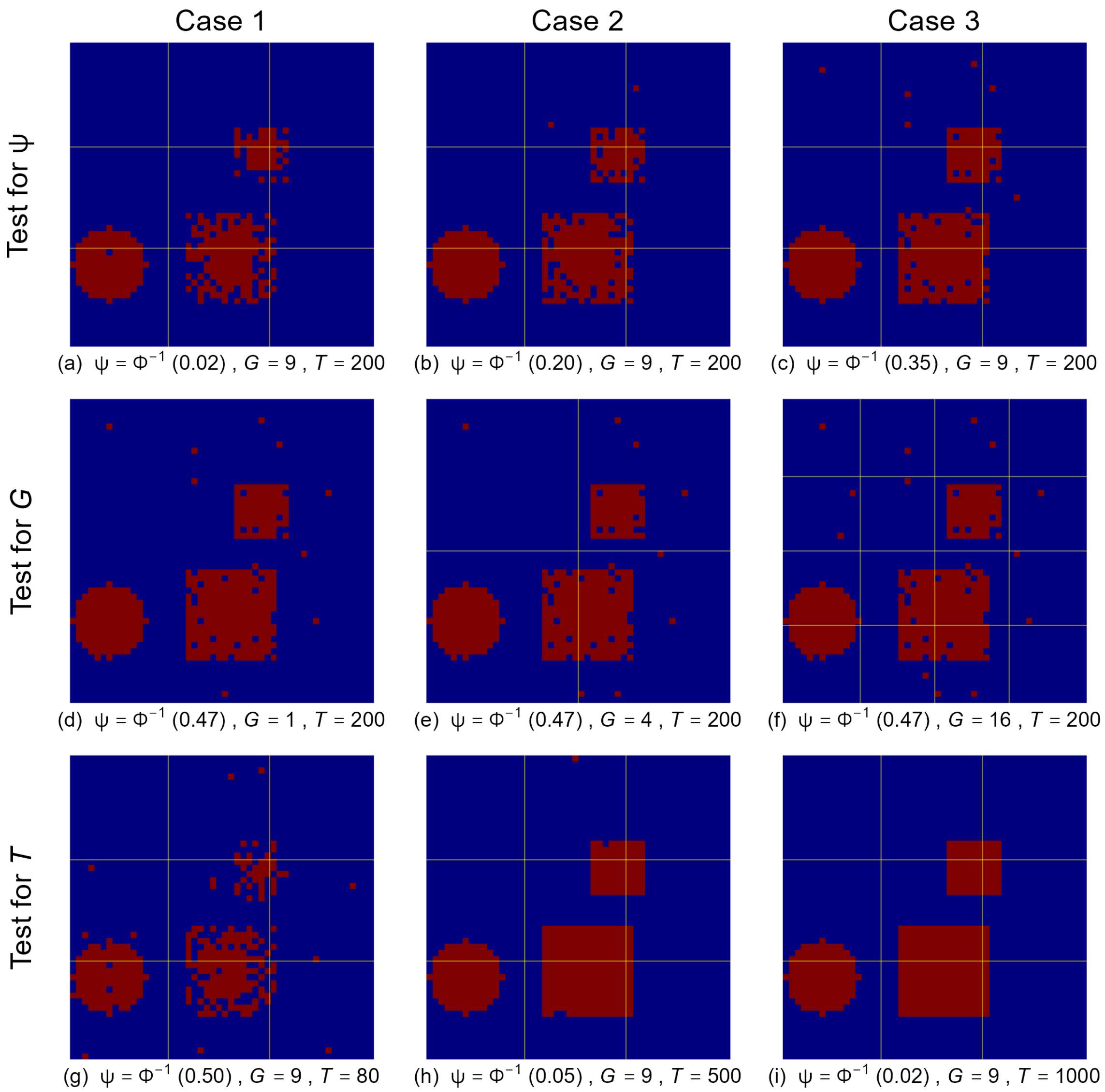
Figure 2.4(a)-(c) illustrates the results using ψ values of $\Phi^{-1}(0.02)$, $\Phi^{-1}(0.20)$, $\Phi^{-1}(0.35)$, respectively, which govern the *a priori* likelihood of a voxel being determined active. Along with Figure 2.2(f) using $\psi = \Phi^{-1}(0.47)$, we can observe a trade-off in selecting ψ: larger values lead to an increase in active voxels and false positives, whereas smaller values result in fewer active voxels and increased false negatives, all of which are as expected. In a simulated scenario, the optimal ψ can be determined by maximizing metrics like prediction accuracy or F1-score. In practical applications, ψ can be tuned to achieve a target percentage of active voxels based on prior experiments, cross-validation, WAIC (Watanabe, 2010), etc.

The effects of varying $G = 1, 4, 16$ are exhibited in Figure 2.4(d)-(f), respectively. Along with Figure 2.2(f) using $G = 9$, we observe negligible edge effects, that is, voxel classifications at parcel borders remain unaffected. Some metrics, such as F1-score, slope, CCC, and X-Y MSE, even exhibit slight improvements through $G = 1, 4, 9$. Moreover, the computation time drops significantly as G increases, as expected. These results coincide with the findings of Musgrove et al. (2016). However, with $G = 16$, performance starts decreasing compared to that of using $G = 9$ due to insufficient number of voxels within each parcel. The choice of G and corresponding parcel size V_g can be guided by prior experience or domain-specific knowledge of, e.g., anatomical regions.

Figure 2.4(g)-(i) depicts the impact of varying the time length $T = 80, 500, 1000$, respectively. The length of each epoch remains the same as 40 time points so that the number of epochs will change correspondingly. Along with Figure 2.2(f) using $T = 200$, we observe improvements in both classification and estimation as T increases. In this case, an accuracy of 100% is achieved when $T = 1000$, and its estimated strength map almost perfectly reproduces the truth. It is worth noting that we adopt a relatively low $\psi = \Phi^{-1}(0.02)$ for $T = 1000$, suggesting a stringent selection of active voxels. Thus, when an ample number of repeated epochs are available for the stimulus, the signal is strong enough to let us select most of the positive voxels while avoiding false positives. This suggests that choosing a low ψ can enhance discriminative capability.

2.3.2 Realistic Simulation

Here we simulate a dataset similar to that has done by Yu et al. (2018) in which we mimic the environmental conditions of a human brain. The data contain *iid* noise. The dataset comprises seven slices, each of size 96×96 voxels, with signals generated across $T = 490$ time points. The

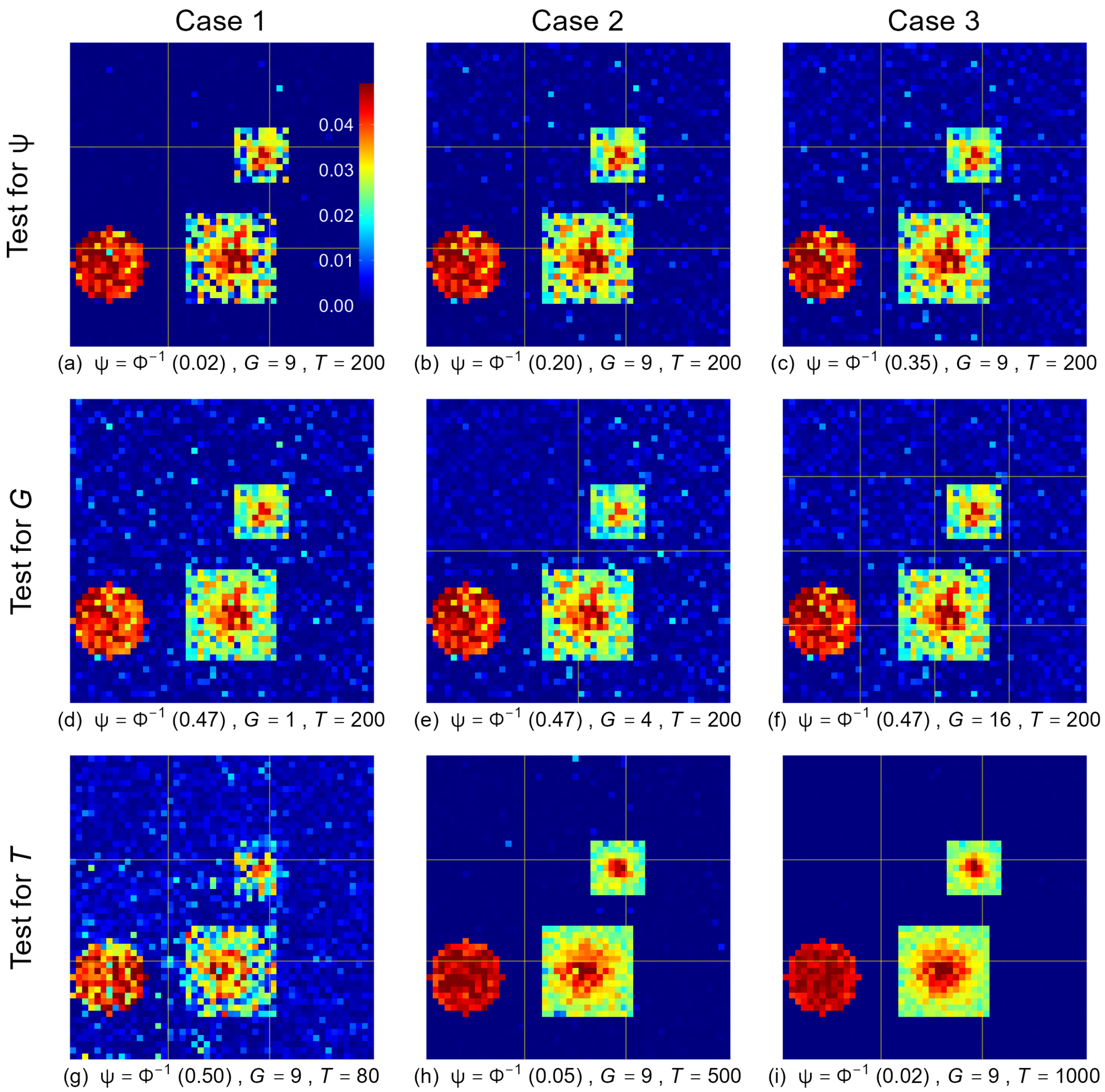


(a) $\psi = \Phi^{-1}(0.02)$, $G = 9$, $T = 200$ (b) $\psi = \Phi^{-1}(0.20)$, $G = 9$, $T = 200$ (c) $\psi = \Phi^{-1}(0.35)$, $G = 9$, $T = 200$

(d) $\psi = \Phi^{-1}(0.47)$, $G = 1$, $T = 200$ (e) $\psi = \Phi^{-1}(0.47)$, $G = 4$, $T = 200$ (f) $\psi = \Phi^{-1}(0.47)$, $G = 16$, $T = 200$

(g) $\psi = \Phi^{-1}(0.50)$, $G = 9$, $T = 80$ (h) $\psi = \Phi^{-1}(0.05)$, $G = 9$, $T = 500$ (i) $\psi = \Phi^{-1}(0.02)$, $G = 9$, $T = 1000$

Figure 2.4: (a)-(c) are estimated activation maps for an AR(1) dataset as produced by the CV-sSGLMM model using various tuning parameters ψ's. (d)-(f) are estimated activation maps using various parcel numbers G's. (g)-(i) are estimated activation maps derived from datasets with various time lengths T's.



(a) $\psi = \Phi^{-1}(0.02)$, $G = 9$, $T = 200$ (b) $\psi = \Phi^{-1}(0.20)$, $G = 9$, $T = 200$ (c) $\psi = \Phi^{-1}(0.35)$, $G = 9$, $T = 200$

(d) $\psi = \Phi^{-1}(0.47)$, $G = 1$, $T = 200$ (e) $\psi = \Phi^{-1}(0.47)$, $G = 4$, $T = 200$ (f) $\psi = \Phi^{-1}(0.47)$, $G = 16$, $T = 200$

(g) $\psi = \Phi^{-1}(0.50)$, $G = 9$, $T = 80$ (h) $\psi = \Phi^{-1}(0.05)$, $G = 9$, $T = 500$ (i) $\psi = \Phi^{-1}(0.02)$, $G = 9$, $T = 1000$

Figure 2.5: (a)-(c) are estimated strength maps for an AR(1) dataset as produced by the CV-sSGLMM model using various tuning parameters ψ's. (d)-(f) are estimated strength maps using various parcel numbers G's. (g)-(i) are estimated strength maps derived from datasets with various time lengths T's.

Parameter	Accuracy	Precision	Recall	F1 Score	AUC	Slope	CCC	X-Y MSE	Time (s)
$\psi = \Phi^{-1}(0.02)$	0.9486	**0.9985**	0.6179	0.7585	0.9908	0.8180	0.8924	2.16e-5	5.39
$\psi = \Phi^{-1}(0.20)$	0.9728	0.9823	0.8123	0.8880	**0.9878**	**0.8894**	**0.9316**	**1.38e-5**	5.66
$\psi = \Phi^{-1}(0.35)$	**0.9783**	0.9628	**0.8706**	**0.9136**	0.9876	0.8893	0.9251	1.47e-5	5.71
$G = 1$	0.9784	0.9220	0.9096	0.9151	**0.9929**	0.8129	0.8818	2.09e-5	63.37
$G = 4$	**0.9796**	**0.9381**	**0.9352**	0.9064	0.9908	0.8464	0.9010	1.83e-5	12.06
$G = 16$	0.9787	0.9306	0.9045	**0.9167**	0.9874	**0.8944**	**0.9142**	**1.68e-5**	**3.74**
$T = 80$	0.9325	0.9070	0.5449	0.6765	0.8952	0.6961	0.7537	4.22e-5	**3.62**
$T = 500$	0.9986	0.9982	0.9919	0.9951	0.9999	0.9749	0.9881	2.59e-5	11.98
$T = 1000$	**0.9999**	**0.9997**	**1**	**0.9999**	**1**	**0.9889**	**0.9950**	**0.11e-5**	21.17

Table 2.3: Summary of average metrics across 100 AR(1) datasets produced by the CV-sSGLMM model using different parameters.

brain's active regions are two 5 × 5 × 5 cubes formed by two 5 × 5 squares within each of slice 2-6. In contrast to the data produced by Eq. (2.5), which exhibits a constant phase, this dataset has a dynamic phase. The cv-fMRI signal for voxel v at time t is thus simulated as

$$\begin{aligned} y_{t,Re}^v &= (\beta_0 + \beta_1^v x_t)\cos(\theta_0 + \theta_1^v x_t) + \varepsilon_{t,Re}^v, \quad \varepsilon_{t,Re}^v \sim \mathcal{N}(0, \sigma^2), \\ y_{t,Im}^v &= (\beta_0 + \beta_1^v x_t)\sin(\theta_0 + \theta_1^v x_t) + \varepsilon_{t,Im}^v, \quad \varepsilon_{t,Im}^v \sim \mathcal{N}(0, \sigma^2). \end{aligned}$$

The slice with the greatest maximum magnitude and phase CNR is slice 4 (Eq. (2.7)):

$$\begin{aligned} \text{CNR}_{\text{Mag}} &= (\max \beta_1^v)/\sigma = 0.5/1, \\ \text{CNR}_{\text{Ph}} &= (\max \theta_1^v)/\text{SNR}_{\text{Mag}} = (\pi/120)/25. \end{aligned} \tag{2.7}$$

Activation then decreases from slice 4 to slices 3 and 5 and is weakest in slices 2 and 6. Slices 1 and 7 exhibit no activation. It is important to note that, with dynamic phase, the model from Lee et al. (2007) is not equivalent to that from Rowe (2005a) as indicated in Rowe (2009). This discrepancy suggests the proposed model is under model misspecification in this scenario. However, as both β_{Re}^v and β_{Im}^v in model (2.2) include magnitude and phase information, and given that prior studies (Yu et al., 2018, 2023) have used the Lee et al. (2007)-based model to process this dataset, we deem it worthwhile to test our model on these data. We set $G = 49$ and a threshold of 0.8722 for both MO-sSGLMM and CV-sSGLMM, with ψ set to $\Phi^{-1}(0.50)$ and $\Phi^{-1}(0.11)$, respectively. For CV-nonSpatial, the threshold is set to 0.5, again following the advice of Yu et al. (2018). Activation maps are presented in Figure 2.6. We indeed observe that our model tends to overestimate the strength. Since the strengths are overestimated, we scale the estimated strength to the range of true strength in the corresponding slice. True and (scaled) estimated strength maps are displayed in Figure 2.7.

Further numerical results, displayed in Table 2.4, show a pattern of the CV-sSGLMM model outperforming both the MO-sSGLMM and CV-nonSpatial models across different slices in terms of detecting true positives (TP). It should be noted, however, that the MO-sSGLMM model achieves a 100% precision (no false positives, FP) for most slices, albeit at the cost of a low recall rate (high false negatives, FN), indicating that the model is more conservative in identifying activated voxels. For the CV-nonSpatial model, although it exhibits good precision across the slices, the recall rates remain lower, specifically in the slices with weaker activation strengths (slices 2 and 6). This performance pattern suggests that the model struggles to detect activations in areas with low CNR, highlighting a limitation when dealing with real-world fMRI datasets that often feature low CNR. In comparison, the CV-sSGLMM model consistently detects a higher number of true positives across all slices, demonstrating a stronger detection power even in slices with weak activations (slices 2 and 6). This underscores the benefit of incorporating spatial information, which enhances the model's capacity to detect weaker activations in the presence of complex noise conditions. The model also maintains a 100% precision across all slices, suggesting that the inclusion of spatial information does not lead to an increase in false positives. As anticipated, both the MO-sSGLMM and CV-sSGLMM models, which employ brain parcellation, demonstrate superior computational efficiency, even when the parallel computation is gated by a 16-core CPU. This advantage becomes even more pronounced when handling larger datasets.

2.4 Analysis of Human CV-fMRI Data

In this study, we consider the fMRI dataset that is analyzed by Yu et al. (2018), which is acquired during a unilateral finger-tapping experiment on a 3.0-T General Electric Signa LX MRI scanner. The experimental paradigm involves 16.33 epochs of alternating 15s on and 15s off periods, leading to $T = 490$ time points. The data are sourced from seven slices, each of size 96×96. For the MO-sSGLMM and CV-sSGLMM models, we set the parcel number to $G = 25$ and again use a threshold of 0.8722 on the inclusion probabilities. The tuning parameter ψ is set to $\Phi^{-1}(0.02)$ and $\Phi^{-1}(0.1)$, respectively. For CV-nonSpatial, the threshold is set to 0.5 as before. The consequent activation and strength maps generated from these analyses are depicted in Figure 2.8 and Figure 2.9. With computation times closely paralleling those in Section 2.3.2 due

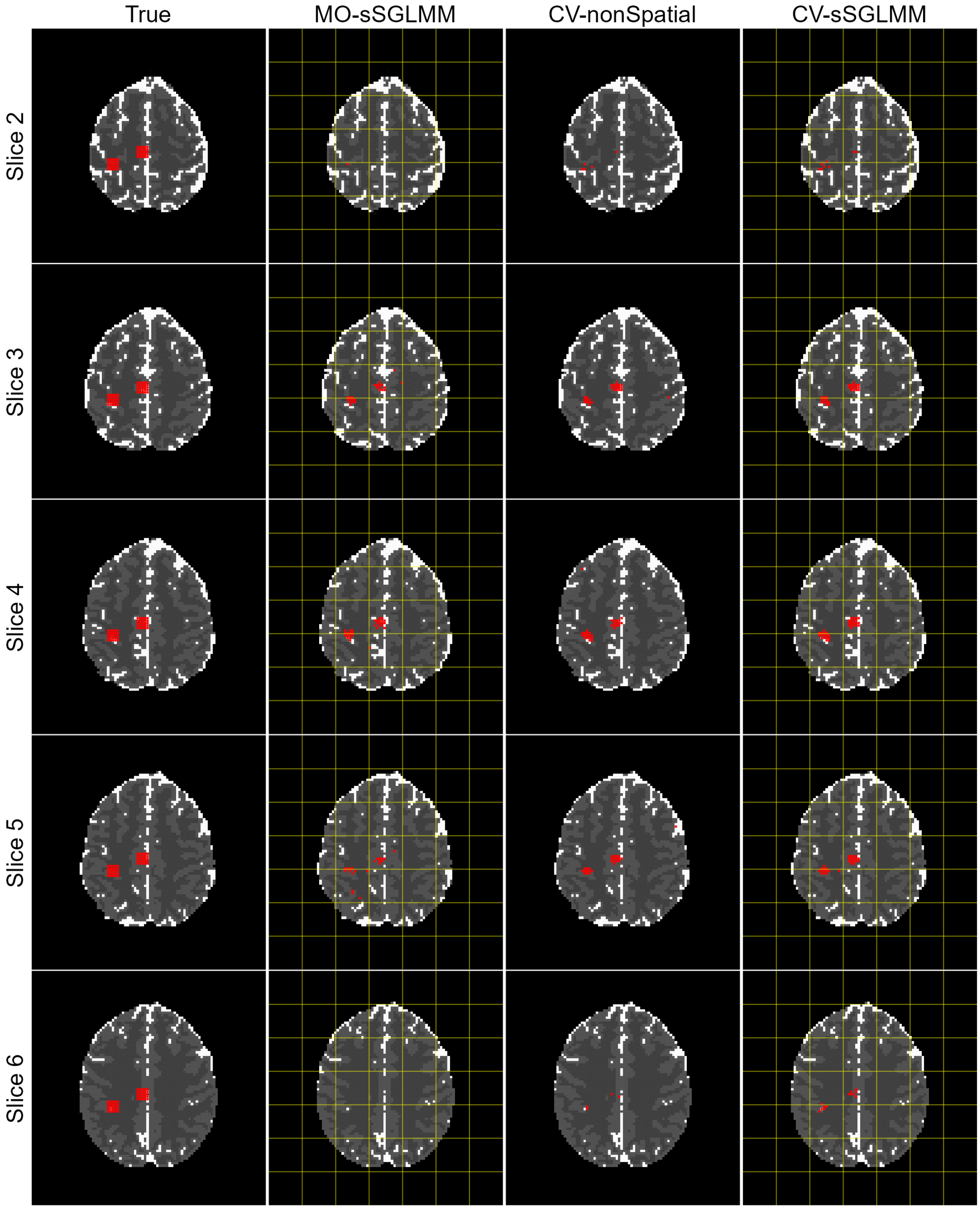


Figure 2.6: True and estimated activation maps for a realistic simulation as produced by the MO-sSGLMM, CV-nonSpatial, and CV-sSGLMM models.

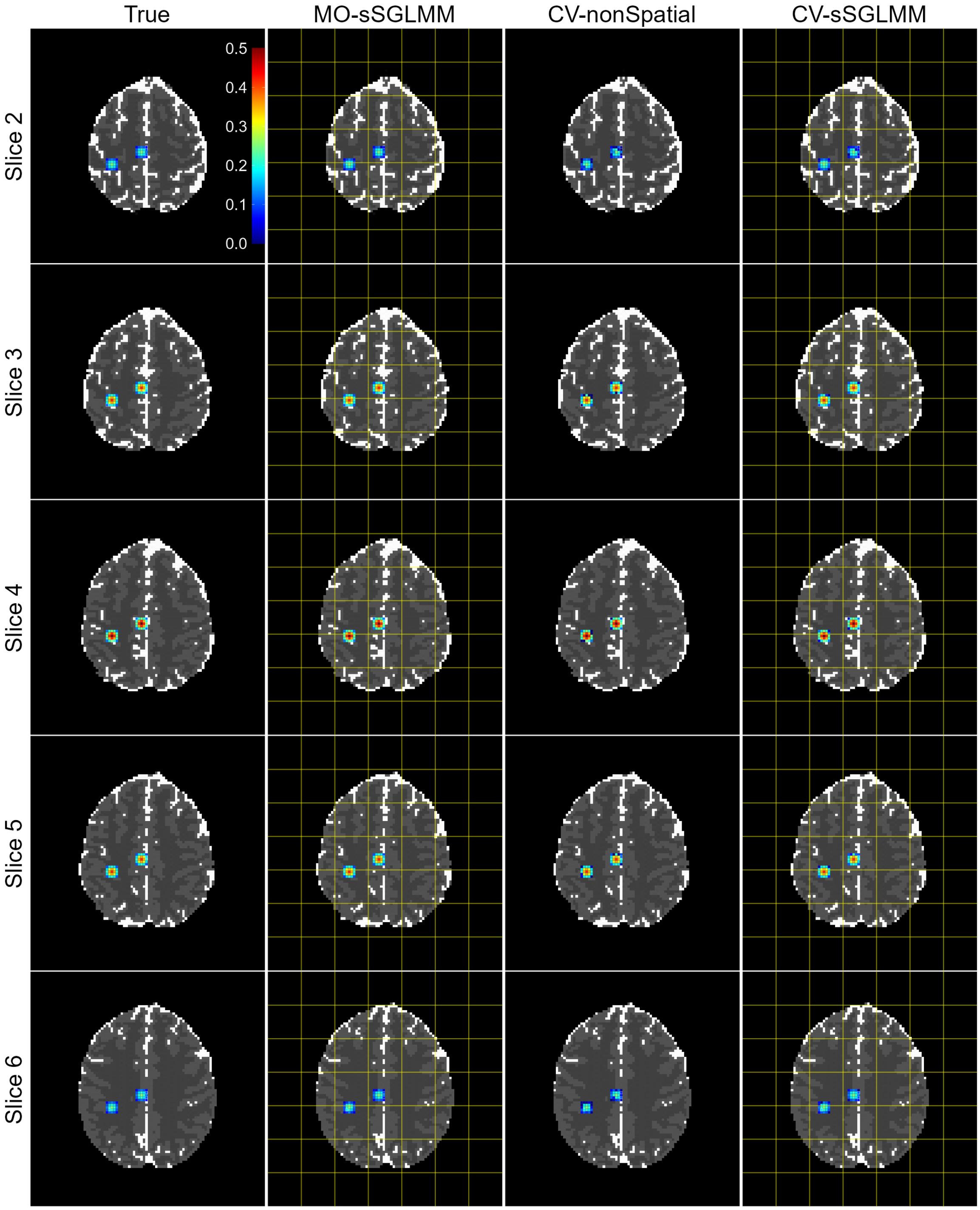


Figure 2.7: True and (scaled) estimated strength maps for a realistic simulation as produced by the MO-sSGLMM, CV-nonSpatial, and CV-sSGLMM models.

Slice	Model	TP	FP	FN	TN	Precision	Recall	Time (s)
2	MO-sSGLMM	1	0	49	9166	1	0.02	**11.59**
	CV-nonSpatial	5	0	45	9166	1	0.1	311.55
	CV-sSGLMM	**8**	0	**42**	9166	1	**0.16**	27.86
3	MO-sSGLMM	22	2	28	9164	0.9166	0.44	same as Slice 2
	CV-nonSpatial	25	1	25	9165	0.9615	0.50	
	CV-sSGLMM	**27**	**0**	**23**	**9166**	**1**	**0.54**	
4	MO-sSGLMM	30	1	20	9165	0.9677	0.60	
	CV-nonSpatial	30	1	20	9165	0.9677	060	
	CV-sSGLMM	**35**	**0**	**15**	**9166**	**1**	**0.70**	
5	MO-sSGLMM	16	5	34	9161	0.7619	0.32	
	CV-nonSpatial	25	**1**	25	**9165**	0.9615	0.50	
	CV-sSGLMM	**28**	**1**	**22**	**9165**	**0.9655**	**0.56**	
6	MO-sSGLMM	0	0	50	9166	NA	0	
	CV-nonSpatial	4	0	46	9166	1	0.08	
	CV-sSGLMM	**13**	0	**37**	9166	**1**	**0.26**	

Table 2.4: Metrics of slices (50 positives and 9166 negatives on each slice) produced by the MO-sSGLMM, CV-nonSpatial, and CV-sSGLMM models.

to comparable dataset sizes, all three models show the same patterns of activation maps. Our CV-sSGLMM consistently demonstrates superior prediction power, particularly evident in the weakly active areas observed in slices 1 and 7, maintaining its consistent performance as discussed in Section 2.3.2. The active regions identified through our CV-sSGLMM method align with those reported in Yu et al. (2018), reinforcing the validity of our results and the efficacy of our proposed approach. More importantly, the active regions correspond to areas of the brain that are known to typically be engaged in finger-tapping tasks, affirming the biological relevance of our findings.

2.5 Conclusion

In this study, we propose an innovative fully Bayesian approach to brain activity mapping using complex-valued fMRI data. The proposed model, which incorporates both the real and imaginary components of the fMRI data, provides a holistic perspective on brain activity mapping, overcoming the limitations of the conventional magnitude-only analysis methods. This model showcases the potential to detect task-related activation with higher accuracy. The adoption of an autoregressive error structure, together with spatial priors, allows us to capture both temporal and spatial correlations in brain activity. Moreover, the employment of brain parcellation and parallel compu-

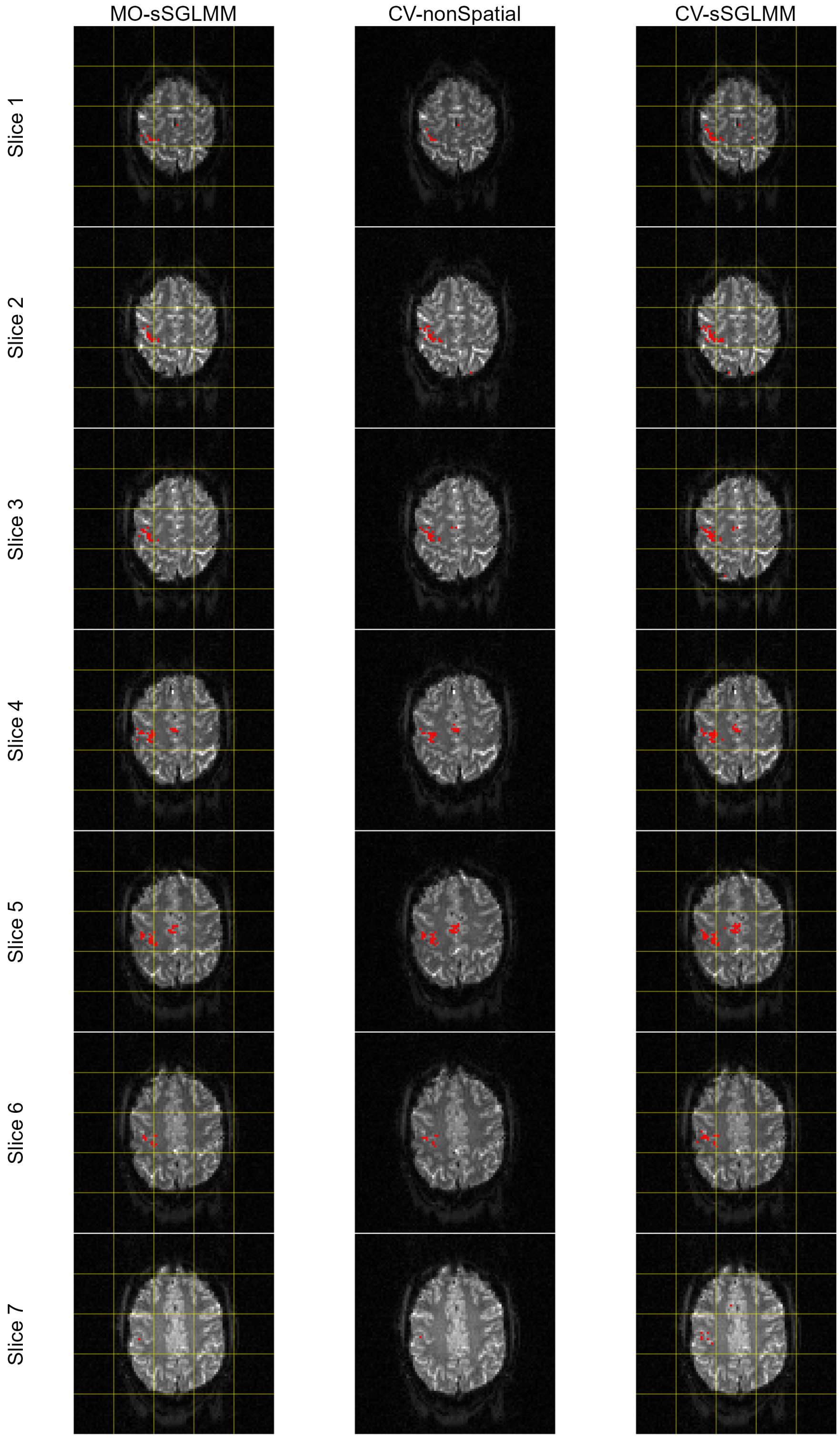


Figure 2.8: Estimated activation maps for a real human brain dataset as produced by the MO-sSGLMM, CV-nonSpatial, and CV-sSGLMM models.

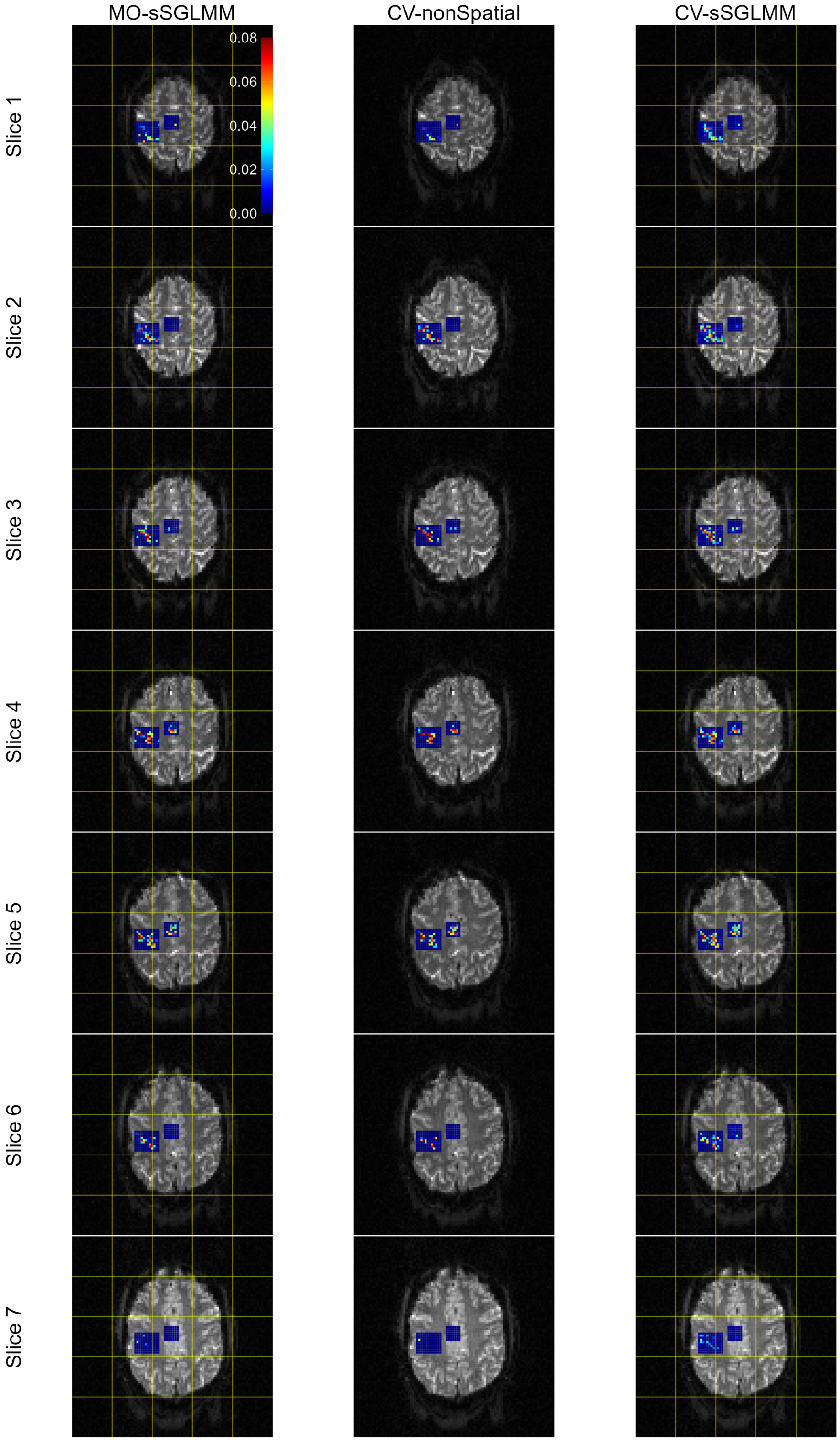


Figure 2.9: Estimated strength maps for a real human brain dataset as produced by the MO-sSGLMM, CV-nonSpatial, and CV-sSGLMM models.

tation significantly enhances the model's computational efficiency. Analyses of both simulated and real fMRI data underscores the benefits of our approach, particularly when temporally-correlated, complex-valued noise is present.

There are still areas for exploration. For instance, while we achieve significant results by assuming the phases are constant, we believe that future Bayesian studies based on the dynamic phase model of Rowe (2005a) should be proposed to account for potential phase variations during brain activity (Petridou et al., 2013). Additionally, our current proposal assumes circular data, that is, $\mathbf{C}^v = \mathbf{0}$ for $\boldsymbol{\varepsilon}^v$ in model (2.1), implying that β^v_{Re} and β^v_{Im} are independent. It would be prudent to develop a more generalized non-circular model where $\mathbf{C}^v \neq \mathbf{0}$ to account for the possibility of non-circular data.

Chapter 3

Efficient Fully Bayesian Approaches to Brain Activity Mapping with Complex-Valued fMRI Data: Analysis of Magnitude and Phase in a Polar Model

3.1 Introduction

Lee et al. (2007)'s Cartesian model we employ in Chapter 2 has gained significant attention for modeling complex-valued fMRI data (Yu et al., 2018, 2023; Wang et al., 2023). However, this model has notable limitations that warrant further investigation and refinement. Before delving into the specifics, it is imperative to understand the inherent properties of cv-fMRI signals, as shown in Figure 3.1. These signals may exhibit changes in magnitude, phase, or both in response to a stimulus.

The magnitude changes in cv-fMRI are fundamentally driven by the BOLD effect, which operates through a cascade of hemodynamic responses. Neuronal activity leads to increased metabolic

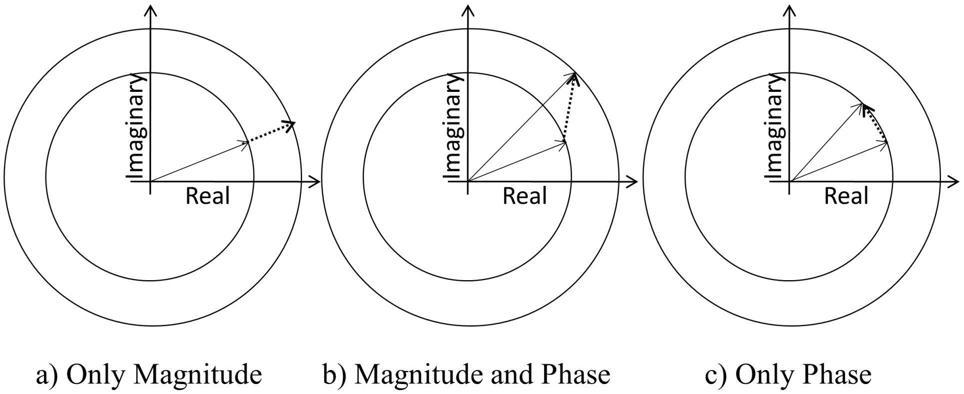


Figure 3.1: Illustration of potential changes in complex-valued fMRI time series (Rowe, 2019).

demand, causing a rise in blood flow to the region. This influx of fresh blood displaces deoxygenated blood, leading to an overall increase in the oxygenation level of the blood in that area. These changes in blood oxygenation cause a change in BOLD signal and magnetic susceptibility, affecting the MR signal magnitude. Thus, the BOLD effect can be considered an indirect measure of neuronal activity, mediated through vascular changes (Boynton et al., 1996; Logothetis, 2008).

On the other hand, phase changes are influenced not only by the BOLD effect but also by the electrical neuronal activities directly. These activities generate moving charges, and therefore create magnetic field which changes the phase of MRI signal. For this reason, the phase changes are able to reveal the aspects of neuronal activity or other phenomena that might be undetected by magnitude-based analyses (Petridou et al., 2006). By accurately modeling these phase changes, researchers can gain insights into the more direct effects of neuronal activity on the MRI signal, potentially leading to more precise and informative interpretations of fMRI data (Feng et al., 2009). This is especially crucial in understanding complex brain functions and improving the accuracy of fMRI in research and clinical applications.

Upon re-examining the CV-sSGLMM model based on Lee's Cartesian model, as discussed in Chapter 2, its contributions to the detection of brain activation are mitigated by two main limitations. First, it fails to properly model both the magnitude and phase components of cv-fMRI signals. As such, it merely identifies active voxels without specifying the type of their activation—whether it is magnitude-related, phase-related, or both. Second, the model only indicates the overall responsive strength of a voxel in reaction to the stimulus, lacking a precise metric for quantification in either the magnitude or phase dimensions.

To overcome these limitations, Rowe (2005a)'s polar model serves as a more fitting alternative. Originating from Rowe and Logan's initial formulation (Rowe and Logan, 2004) with dynamic magnitude and constant phase, the model has undergone several iterations (Rowe and Logan, 2005; Rowe, 2005b) to arrive at its current version to capture both dynamic magnitude and dynamic phase. The primary objective of this chapter is to propose a fully Bayesian approach based on Rowe's polar model for detecting the magnitude and phase activation maps independently. Specifically, we aim to identify which voxels exhibit significant changes in cv-fMRI signals in response to a stimulus, in terms of magnitude, phase, or both, and to quantify how strongly they change in these perspectives. By doing so, it offers a more nuanced and scientific framework for delineating brain activation patterns in task-based fMRI analyses.

We leverage the sparse spatial generalized linear mixed mode (sSGLMM) prior and brain parcellation strategy previously introduced in Chapter 2, aiming to capture spatial correlations and facilitate parallel computation. Section 3.2 elaborates on our proposed model and gives the priors and posteriors involved. Section 3.3 illustrates the model's performance through simulated datasets. Section 3.4 applies our proposed model to the cv-fMRI data obtained from a real finger-tapping experiment, which is the same one as Chapter 2. Finally, Section 3.5 summarizes our research, emphasizes its contributions, and suggests avenues for future exploration in this field.

3.2 Model

In this section, we introduce our model for mapping brain activity using magnitude and phase of cv-fMRI data. We then proceed to derive the posterior distribution for the parameters under consideration and provide an MCMC algorithm to estimate them.

3.2.1 Model Formulation

For a certain voxel v (where $v = 1, \dots, V$) at time t (where $t = 1, \dots, T$), its real and imaginary parts of complex-valued fMRI signal, $y_{v,Rt}$ and $y_{v,It}$, can be modeled as:

$$\begin{pmatrix} y_{v,Rt} \\ y_{v,It} \end{pmatrix} = \begin{pmatrix} \rho_{v,t} \cos \theta_{v,t} \\ \rho_{v,t} \sin \theta_{v,t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{v,Rt} \\ \varepsilon_{v,It} \end{pmatrix}, \quad \begin{pmatrix} \varepsilon_{v,Rt} \\ \varepsilon_{v,It} \end{pmatrix} \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I}_2),$$

where $\rho_{v,t}$ and $\theta_{v,t}$ are temporally varying magnitude and phase given by:

$$\begin{aligned} \rho_{v,t} &= \beta_{v,0} + x_t \beta_{v,1}, \\ \theta_{v,t} &= \gamma_{v,0} + u_t \gamma_{v,1}, \end{aligned}$$

where x_t and u_t are the expected BOLD response and neuronal electromagnetic signal, respectively, at time t. Thus, for all time points:

$$\mathbf{y}_v = \begin{pmatrix} \mathbf{A}_R(\boldsymbol{\gamma}_v) \\ \mathbf{A}_I(\boldsymbol{\gamma}_v) \end{pmatrix} \mathbf{X} \boldsymbol{\beta}_v + \boldsymbol{\varepsilon}_v, \quad \boldsymbol{\varepsilon}_v \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I}_{2T}), \tag{3.1}$$

where $\mathbf{y}_v = [(\mathbf{y}_{v,R})', (\mathbf{y}_{v,I})']' \in \mathbb{R}^{2T}$ stacks real and imaginary components of cv-fMRI signal, and $\mathbf{X} = [\mathbf{1}, \mathbf{x}] \in \mathbb{R}^{T \times 2}$ is the design matrix for the magnitude composed of ones and expected BOLD response $\mathbf{x}$. The matrices $\mathbf{A}_R(\boldsymbol{\gamma}_v), \mathbf{A}_I(\boldsymbol{\gamma}_v) \in \mathbb{R}^{T \times T}$ are diagonal as:

$$\mathbf{A}_R(\boldsymbol{\gamma}_v) = \begin{pmatrix} \cos(\gamma_{v,0} + u_1 \gamma_{v,1}) & 0 & \cdots & 0 \\ 0 & \cos(\gamma_{v,0} + u_2 \gamma_{v,1}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cos(\gamma_{v,0} + u_t \gamma_{v,1}) \end{pmatrix},$$

$$\mathbf{A}_I(\boldsymbol{\gamma}_v) = \begin{pmatrix} \sin(\gamma_{v,0} + u_1 \gamma_{v,1}) & 0 & \cdots & 0 \\ 0 & \sin(\gamma_{v,0} + u_2 \gamma_{v,1}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sin(\gamma_{v,0} + u_t \gamma_{v,1}) \end{pmatrix},$$

with a more compact form:

$$\mathbf{A}_R(\boldsymbol{\gamma}_v) = \text{diag}\left[\cos\left(\mathbf{U}\boldsymbol{\gamma}_v\right)\right], \quad \mathbf{A}_I(\boldsymbol{\gamma}_v) = \text{diag}\left[\sin\left(\mathbf{U}\boldsymbol{\gamma}_v\right)\right],$$

where $\mathbf{U} = [\mathbf{1}, \mathbf{u}] \in \mathbb{R}^{T\times 2}$ is the design matrix for the phase composed of ones and neuronal electromagnetic signal $\mathbf{u}$. Therefore, $\boldsymbol{\beta}_v = [\beta_{v,0}, \beta_{v,1}]' \in \mathbb{R}^2$ and $\boldsymbol{\gamma}_v = [\gamma_{v,0}, \gamma_{v,1}]' \in \mathbb{R}^2$ are magnitude- and phase-related regression coefficients, respectively. The voxel-specific error term $\boldsymbol{\varepsilon}_v$ follows a multivariate normal distribution with the variance-covariance matrix $\sigma_v^2\mathbf{I}_{2T}$, and a Jeffreys prior can be assigned to σ_v^2 as $p(\sigma_v^2) \propto 1/\sigma_v^2$.

3.2.2 Brain Parcellation and Spatial Priors

Brain Parcellation We adopt the same brain parcellation strategy as outlined in Section 2.2.2, wherein the brain is partitioned into G evenly distributed parcels to facilitate parallel computation. This approach has been validated for its minimal edge effects in both the work by Musgrove et al. (2016) and in Chapter 2.

Prior Distributions of $\boldsymbol{\beta}_v$ and $\boldsymbol{\gamma}_v$ For each parcel g (where $g = 1, \ldots, G$) encompassing V_g voxels, we classify a voxel v (where $v = 1, \ldots, V_g$) based on its activity. Specifically, a voxel is classified magnitude-active if $\beta_{v,1} \neq 0$, and phase-active if $\gamma_{v,1} \neq 0$. Adhering to the spike-and-slab prior (Mitchell and Beauchamp, 1988), the model is expressed as follows:

$$\begin{aligned}
\boldsymbol{\beta}_v \mid \lambda_v, \tau_g^2 &\sim \lambda_v \mathcal{N}_2\left(\mathbf{0}, \tau_g^2\mathbf{I}\right) + (1-\lambda_v)\mathcal{N}_2\left(\mathbf{0}, \begin{pmatrix} \tau_g^2 & 0 \\ 0 & 0 \end{pmatrix}\right), \\
\boldsymbol{\gamma}_v \mid \omega_v, \xi_b^2 &\sim \omega_v \mathcal{N}_2\left(\mathbf{0}, \xi_g^2\mathbf{I}\right) + (1-\omega_v)\mathcal{N}_2\left(\mathbf{0}, \begin{pmatrix} \xi_g^2 & 0 \\ 0 & 0 \end{pmatrix}\right).
\end{aligned}$$

In this formulation, $\lambda_v, \omega_v \in \{0, 1\}$ indicate the status of voxel v: $\lambda_v = 1$ for a magnitude-active voxel and $\omega_v = 1$ for a phase-active voxel, with 0 indicating inactivity in respective domains. The parameters τ_g^2 and ξ_g^2 represent parcel-specific variances. These variances are constant for all voxels within a particular parcel but may vary across different parcels, and are assigned a Jeffreys prior, that is, $p(\tau_g^2) \propto 1/\tau_g^2$ and $p(\xi_g^2) \propto 1/\xi_g^2$, for $g = 1, \ldots, G$. The prior distributions can be succinctly represented as:

$$\boldsymbol{\beta}_v \mid \lambda_v, \tau_g^2 \sim \mathcal{N}_2\left(\mathbf{0}, \tau_g^2 \boldsymbol{\Lambda}_v\right), \quad \text{where } \boldsymbol{\Lambda}_v = \begin{pmatrix} 1 & 0 \\ 0 & \lambda_v \end{pmatrix},$$

$$\boldsymbol{\gamma}_v \mid \omega_v, \xi_g^2 \sim \mathcal{N}_2\left(\mathbf{0}, \xi_g^2 \boldsymbol{\Omega}_v\right), \quad \text{where } \boldsymbol{\Omega}_v = \begin{pmatrix} 1 & 0 \\ 0 & \omega_v \end{pmatrix}.$$

Spatial Priors on λ_v and ω_v We employ the sparse spatial generalized linear mixed mode (sSGLMM) prior introduced in Section 2.2.2 to capture spatial structures of fMRI signal and enforce sparsity in the variable selection problem. For voxel v (where $v = 1, ..., V_g$) within parcel g (where $g = 1, ..., G$), from the perspective of the magnitude, we suppose that:

$$\begin{aligned} \lambda_v \mid \eta_{\lambda,v} &\overset{iid}{\sim} \mathcal{B}ern\left\{\boldsymbol{\Phi}(\psi_\lambda + \eta_{\lambda,v})\right\}, \\ \eta_{\lambda,v} \mid \boldsymbol{\delta}_{\lambda,g} &\sim \mathcal{N}_1\left(\mathbf{m}_v' \boldsymbol{\delta}_{\lambda,g}, 1\right), \\ \boldsymbol{\delta}_{\lambda,g} \mid \kappa_{\lambda,g} &\sim \mathcal{N}_q\left\{\mathbf{0}, (\kappa_{\lambda,g} \mathbf{M}_g{}' \mathbf{Q}_g \mathbf{M}_g)^{-1}\right\}, \\ \kappa_{\lambda,g} &\sim \mathcal{G}amma\left(a_\kappa, b_\kappa\right). \end{aligned}$$

From the perspective of the phase:

$$\begin{aligned} \omega_v \mid \eta_{\omega,v} &\overset{iid}{\sim} \mathcal{B}ern\left\{\boldsymbol{\Phi}(\psi_\omega + \eta_{\omega,v})\right\}, \\ \eta_{\omega,v} \mid \boldsymbol{\delta}_{\omega,g} &\sim \mathcal{N}_1\left(\mathbf{m}_v' \boldsymbol{\delta}_{\omega,g}, 1\right), \\ \boldsymbol{\delta}_{\omega,g} \mid \kappa_{\omega,g} &\sim \mathcal{N}_q\left\{\mathbf{0}, (\kappa_{\omega,g} \mathbf{M}_g{}' \mathbf{Q}_g \mathbf{M}_g)^{-1}\right\}, \\ \kappa_{\omega,g} &\sim \mathcal{G}amma\left(a_\kappa, b_\kappa\right). \end{aligned}$$

Two prefixed tuning parameters, ψ_λ and ψ_ω, are employed to regulate the proportion of predicted magnitude-active and phase-active voxels, respectively. Additionally, we set $a_\kappa = 0.5$ and $b_\kappa = 2000$ for both $\kappa_{\lambda,g}$ and $\kappa_{\omega,g}$ in all parcels. Empirical evidence suggests that utilizing the same a_κ and b_κ for both $\kappa_{\lambda,g}$ and $\kappa_{\omega,g}$ in all parcels is acceptable (Musgrove et al., 2016), provided the mean is sufficiently large to minimize the risk of detecting spurious activity due to noise or other confounding factors.

3.2.3 MCMC Algorithm and Posterior Distributions

We employ Gibbs sampling to obtain the joint and marginal conditional distributions of parameters of interest. Only the full conditional posterior distribution of $\boldsymbol{\gamma}_v$ is obtained by Metropolis–Hastings algorithm (Metropolis et al., 1953; Hastings, 1970), the others are in well-known distributions. Detailed derivations of these distributions, along with the full conditional distributions required for sampling, are delineated in Appendix B.1. To assess the convergence of the algorithm, we adopt the fixed-width diagnostic technique suggested by Flegal et al. (2008). Convergence is considered achieved when the Monte Carlo Standard Error (MCSE) for all λ_v and ω_v drops below 0.05, leading us to run 10^3 iterations. After discarding the burn-in phase, the means of the sampled parameters are taken as point estimates. If $\widehat{\lambda}_v > 0.925$, the voxel is magnitude-active; if $\widehat{\omega}_v > 0.925$, it is phase-active. Smith and Fahrmeir (2007) proposed the threshold of 0.8722 regarding the significance level $\alpha = 0.05$. Since our approach is similar to a two-step sequential test, we use Bonferroni correction to make $\alpha = 0.05/2 = 0.025$, leading to the adjustment of threshold from 0.8722 to 0.925.

3.3 Simulation Studies

This section presents two distinct simulation studies. The first study focuses on a single map that comprises three types of active regions: one region is solely magnitude-active, another is solely phase-active, and the third is both magnitude- and phase-active. The second study involves multiple datasets, each containing only one type of activation on their maps. For comparative evaluation, we consider the following models:

- The model proposed by Musgrove et al. (2016), referred to as MO, models magnitude-only data. For a certain voxel v, $v = 1, \ldots, V$, over time T:

  $$\mathbf{y}_{v,M} = \mathbf{X}\boldsymbol{\beta}_{v,M} + \boldsymbol{\varepsilon}_v, \qquad \boldsymbol{\varepsilon}_v \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I}_T)$$

  where $\mathbf{y}_{v,M} \in \mathbb{R}^T$ is the magnitude of complex-valued fMRI signal, and $\mathbf{X} = [\mathbf{1}, \mathbf{x}] \in \mathbb{R}^{T\times 2}$ is the design matrix composed of ones and expected BOLD response $\mathbf{x}$. The vector $\boldsymbol{\beta}_{v,M} = (\beta_{v,M_0}, \beta_{v,M_1})'$ are regression coefficients.

- The model delineated in Chapter 2, based on Lee et al. (2007)'s Cartesian model and referred to as CV-R&I. This model is referred to as CV-sSGLMM in Chapter 2, but we emphasize its usage of real and imaginary components here rather than the sSGLMM prior. It models complex-valued data by modeling the real and imaginary components:

$$\mathbf{y}_v = \begin{pmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{X} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_{v,R} \\ \boldsymbol{\beta}_{v,I} \end{pmatrix} + \boldsymbol{\varepsilon}_v, \quad \boldsymbol{\varepsilon}_v \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I}_{2T}),$$

where $\mathbf{y}_v = [(\mathbf{y}_{v,R})', (\mathbf{y}_{v,I})']' \in \mathbb{R}^{2T}$ is the stack of real and imaginary components of cv-fMRI signal. The vectors $\boldsymbol{\beta}_{v,R} = (\beta_{v,R_0}, \beta_{v,R_1})'$ and $\boldsymbol{\beta}_{v,I} = (\beta_{v,I_0}, \beta_{v,I_1})'$ are regression coefficients regarding real and imaginary components of cv-fMRI signal, respectively.

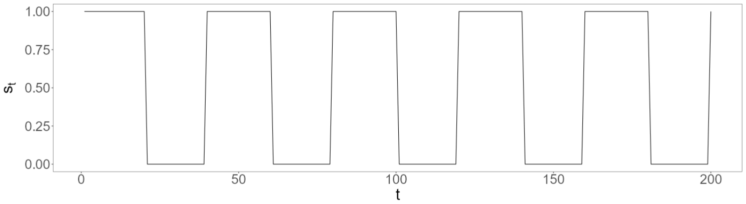
- The model (3.1), referred to as CV-M&P, is based on Rowe (2005a)'s polar model and models complex-valued data while characterizing magnitude and phase.

All three models adhere to a fully Bayesian approach, employ the sSGLMM spatial prior with brain parcellation strategy, and utilize Gibbs sampling to approximate their respective posterior distributions. The number of parcels G is set to 16 for all models. Other tuning parameters, such as $\psi = \Phi^{-1}(0.35)$ for MO, $\psi = \Phi^{-1}(0.30)$ for CV-R&I, and $\psi = \omega = \Phi^{-1}(0.42)$ for CV-M&P, are predetermined to optimize prediction accuracy. The thresholds for identifying active voxels are set at 0.8722 for MO and CV-R&I, as specified in their work, while CV-M&P employs a threshold of 0.925, in accordance with Section 3.2.3.

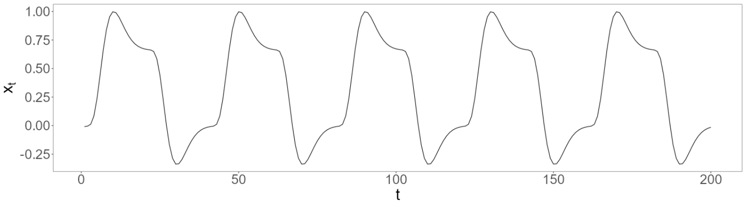
All results are generated by running the code on a custom-built desktop computer with an Intel Core i9-9980XE CPU (3.00GHz, 3001 Mhz, 18 cores, 36 logical processors), NVIDIA GeForce RTX 2080 Ti GPU, 64 GB RAM, and operating on Windows 10 Pro.

3.3.1 Single Simulation

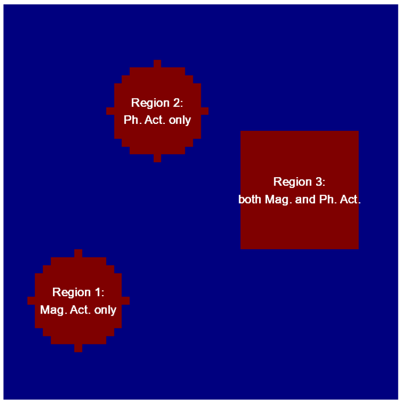
Designed stimulus and expected BOLD response The designed stimulus $\mathbf{s}$ is a binary signal comprised of five repeated epochs, each spanning 40 time points, resulting in a total duration of $T = 200$ time points. Each epoch features the stimulus being alternately active and inactive, with both states persisting for 20 time points. We model the expected BOLD response $\mathbf{x}$ by convolving this stimulus with a double-gamma HRF. Illustrations of both the designed stimulus and the expected BOLD response are provided in Figures 3.2a and 3.2b, respectively, and are consistently used across all our simulation datasets.



(a) Designed stimulus



(b) Expected BOLD response



(c) True activation map

Figure 3.2: (a) Designed stimulus; (b) Expected BOLD response; (c) True activation map.

True activation map and true strength map The true activation map contains three active regions on a 50×50 panel, comprising two circles and one square, each with a radius of five. The exact locations of these regions are depicted in Figure 3.2c. We want to assign distinct types of activation to each region: region 1 exhibits only magnitude activation, region 2 exhibits only phase activation, and region 3 exhibits both magnitude and phase activation, corresponding to the types illustrated in Figures 3.1a, 3.1c, and 3.1b, respectively.

Utilizing the `specifyregion` function in the `neuRosim` library (Welvaert et al., 2011) in `R` (R Core Team, 2023), we initially generate a strength map with decay rates of 0.05, 0.05, and 0.15 for the three regions, respectively. This setup ensures that the central voxel of each active region has a strength of one, diminishing to zero towards the edges at the specified decay rate. For the true magnitude strengths, indicative of voxel response in magnitude to the stimulus, we multiply the strengths in regions 1 and 3 by 0.04909 and nullify the strengths in region 2, as represented in Figure 3.3c. Similarly, for the true phase strengths, reflective of voxel response in phase, we multiply the strengths in regions 2 and 3 by a factor of $\pi/36$ and reduce the strengths in region 1 to zero, as illustrated in Figure 3.3g. This methodology ensures that each region's activation profile is accurately mapped according to its designated stimulus response type.

Simulating fMRI signals We then simulate data according to Eq. (3.2):

$$\begin{aligned} y_{v,Rt} &= (\beta_0 + x_t\beta_{v,1})\cos(\gamma_0 + u_t\gamma_{v,1}) + \varepsilon_{v,Rt}, \quad \varepsilon_{v,Rt} \sim \mathcal{N}(0,\sigma^2), \\ y_{v,It} &= (\beta_0 + x_t\beta_{v,1})\sin(\gamma_0 + u_t\gamma_{v,1}) + \varepsilon_{v,It}, \quad \varepsilon_{v,It} \sim \mathcal{N}(0,\sigma^2), \end{aligned} \tag{3.2}$$

where $\beta_0 = 0.4909$, $\gamma_0 = \pi/4$, and $\sigma = 0.04909$ are set constant for all voxels, and x_t is the expected BOLD response $\mathbf{x}$ from Figure 3.2b at time t. It should be noted that we also use $\mathbf{u} = \mathbf{x}$ as the regressor for phase here when generating the data, but it could be its own neuronal electromagnetic signal $\mathbf{u}$ for the phase in some cases. The signal-to-noise ratio for the magnitude (SNR_{Mag}) is thereby fixed at $\beta_0/\sigma = 10$. The true values of β_1 and γ_1 generated previously in Figures 3.3c and 3.3g are used, yielding the contrast-to-noise ratios for magnitude (CNR_{Mag}) and phase (CNR_{Ph}) as detailed in Eq. (3.3):

$$\begin{aligned} \text{CNR}_{\text{Mag}} &= (\max \beta_{v,1})/\sigma = 0.04909/0.04909 = 1, \\ \text{CNR}_{\text{Ph}} &= (\max \gamma_{v,1})/\text{SNR}_{\text{Mag}} = (\pi/36)/10. \end{aligned} \tag{3.3}$$

Results Figure 3.3 presents both the true and estimated activation maps for magnitude and phase as derived from the CV-M&P model, alongside the corresponding true and estimated parameters β_1 and γ_1. Notably, CV-M&P effectively identifies separate regions that are active in magnitude and phase, and provides proper estimates for the parameters β_1 and γ_1. In the estimated activation maps (Figures 3.3b and 3.3f), the overlap in the predicted active regions corresponds to the square-shaped region 3 in the true map (Figure 3.2c), which is characterized by both magnitude and phase activation. When the predicted region 3 is excluded from these estimated maps, the remaining areas align well with the circular regions 1 and 2 in Figure 3.2c, representing solely magnitude-active and solely phase-active voxels, respectively.

By synthesizing the estimated activation maps for both magnitude and phase (Figures 3.3b and 3.3f), we construct a composite activation map and compare it against results from MO and CV-R&I. Figure 3.4 presents these comparative maps. Performance evaluation reveals that MO fails the competition, primarily due to its inability to detect the phase-only active region 2. Conversely, both CV-R&I and CV-M&P deliver competitive results.

The analysis also extends to comparing the parameter estimations across the three models. As MO and CV-R&I do not explicitly characterize parameters β_1 and γ_1 in their models, we resort

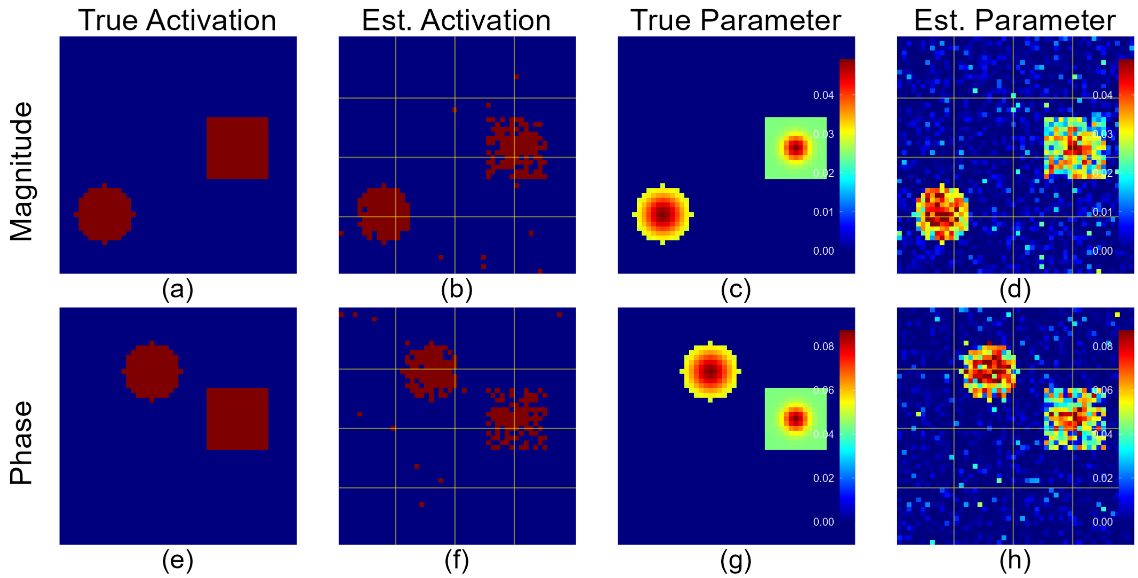


Figure 3.3: (a) and (e) are true magnitude and phase activation maps; (b) and (f) are estimated activation maps as derived from CV-M&P; (c) and (g) are true β_1 and γ_1; (d) and (h) are estimated β_1 and γ_1 as derived from CV-M&P.

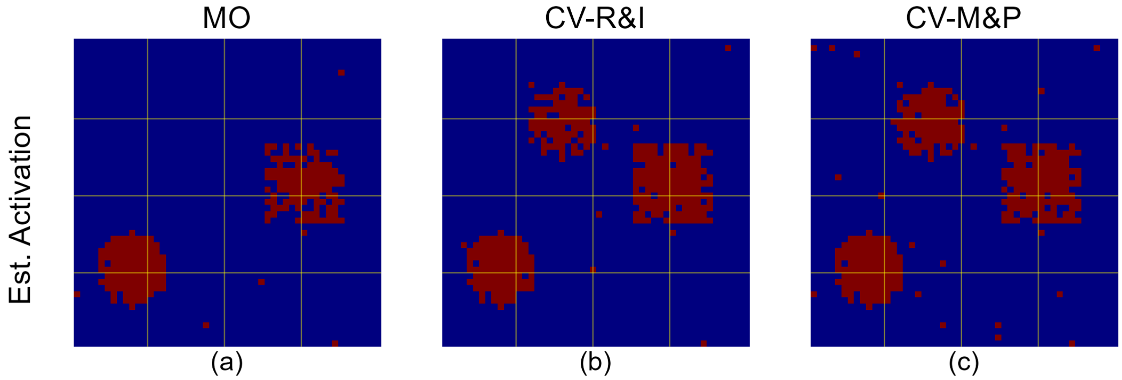
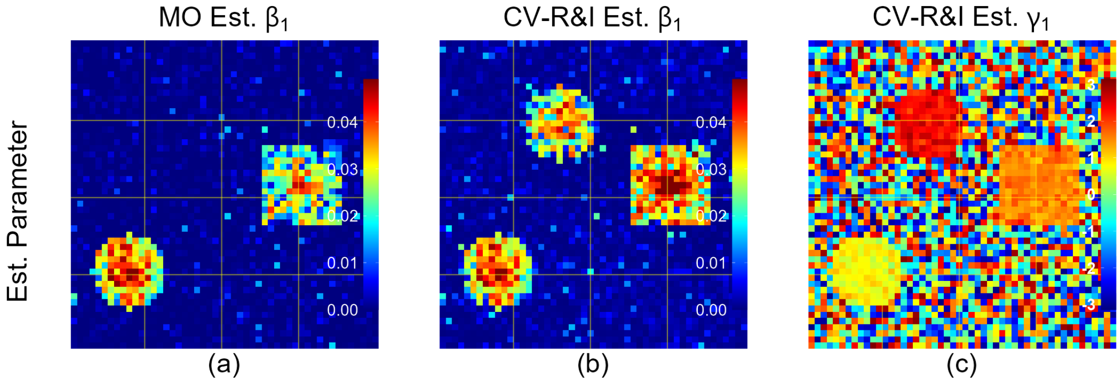


Figure 3.4: (a)-(c) are estimated activation maps as derived from MO, CV-R&I, and CV-M&P, respectively.



(a) (b) (c)

Figure 3.5: (a)-(c) are improperly estimated parameters as derived from MO and CV-R&I.

to indirect methods for their estimation. For MO model, we use the estimated slope of the BOLD signal, $\widehat{\beta}_{v,M_1}$, as an estimate for $\beta_{v,1}$, while for CV-R&I, the square root of the sum of squares of the estimated slopes, $\sqrt{\left(\widehat{\beta}_{v,R_1}\right)^2 + \left(\widehat{\beta}_{v,I_1}\right)^2}$, serves as an estimate for $\beta_{v,1}$. As for γ_1, MO cannot estimate this parameter due to its limitation to magnitude-only data. In contrast, CV-R&I employs $\arctan_4\left(\widehat{\beta}_{v,I_1}/\widehat{\beta}_{v,R_1}\right)$ as an estimate for γ_1. These results are illustrated in Figure 3.5. Upon examination of Figure 3.5a, we observe that while MO's estimated β_1 map appears to closely align with the true β_1 map (Figure 3.3c), it still slightly underestimates values in region 3. Similarly, as seen in Figure 3.5b, CV-R&I not only falsely estimates the non-existent β_1 in the phase-only active region 2, but also tends to overestimate β_1 in region 3. This overestimation of β_1 in region 3 where voxels exhibit both magnitude- and phase-active, is consistent with the findings in Wang et al. (2023). Lastly, Figure 3.5c reveals the CV-R&I's estimated γ_1 map significantly deviates from the true γ_1 map, as showcased in Figure 3.3g.

The numerical evaluation metrics are summarized in Table 3.1, where we employ accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC) to gauge classification performance. We also employ the regression slope between true and estimated parameters to quantify the estimation performance, to expect it be close to one. In terms of classification, CV-M&P outperforms its counterparts in various key metrics, including recall, F1-score, and AUC. While the margin of superiority may not be pronounced, CV-M&P offers two distinct advantages over its counterparts: it allows for the independent prediction of magnitude and phase activation maps, as shown in Figure 3.3b and 3.3f, while the other two approaches cannot, and provides accurate estimation for both β_1 and γ_1, as the slopes (0.9731 and 0.9462) are close to

Table 3.1: Metrics of a single simulated dataset produced by the MO, CV-R&I, and CV-M&P models.

Model	Accuracy	Precision	Recall	F1 Score	AUC	β_1 slope	γ_1 slope	Time (s)
MO	0.9248	0.9726	0.5392	0.6938	0.8910	0.8630	NA	**1.54**
CV-R&I	**0.9688**	**0.9731**	0.8253	0.8931	0.9868	1.0301	33.427	3.93
CV-M&P	0.9680	0.9436	**0.8481**	**0.8933**	**0.9896**	**0.9731**	**0.9462**	17.87

Table 3.2: Characteristics of true maps.

Map size	Number of active regions	Radius	Shape	Decay rate (ϱ)
50×50	3	2 to 6	sphere or cube	0 to 0.3

one. Further evidence from multiple simulation studies, to be discussed in the subsequent section, will reinforce these findings.

3.3.2 Multiple Simulations

Generating Random Maps and Simulating fMRI Signals We generate 100 random true strength maps using the parameters outlined in Table 3.2 and the `specifyregion` function. The true strength maps are then scaled by factors of 0.04909 and $\pi/36$ to obtain 100 true β_1 maps and 100 true γ_1 maps, respectively. Using Eq. (3.2) and the expected BOLD response $\mathbf{x}$ in Figure 3.2b, we generate three datasets from each pair of true β_1 and γ_1 maps with the following assignments:

- β_1 present, γ_1 absent (all active voxels are solely magnitude-active)
- β_1 absent, γ_1 present (all active voxels are solely phase-active)
- β_1 present, γ_1 present (all active voxels are both magnitude- and phase-active)

The values for β_0, γ_0, and σ are held constant as specified in Section 3.3.1, with values 0.4909, $\pi/4$, and 0.04909, respectively, resulting $\text{CNR}_{\text{Mag}} = 1$ and $\text{CNR}_{\text{Ph}} = (\pi/36)/10$. In total, we have 300 datasets for analysis.

Results Table 3.3 presents the performance metrics for each method across these diverse datasets. In terms of classification, MO delivers superior performance in almost all evaluated metrics for datasets featuring exclusively magnitude-active voxels, which is expected given its design specificity

Table 3.3: Summary of average metrics across 100 simulated datasets produced by the MO, CV-R&I, and CV-M&P models. The values in parentheses are min, max, and standard deviation.

Data Type	Measure	MO	CV-R&I	CV-M&P
Mag. -only	Accuracy	**0.9645**(0.9208, 0.9940, 0.0153)	0.9555(0.9032, 0.9916, 0.0188)	0.9598(0.9132, 0.9900, 0.0160)
	Precision	**0.9647**(0.9157, 0.9917, 0.0164)	0.9601(0.9078, 0.9964, 0.0166)	0.9317(0.8515, 0.9805, 0.0224)
	Recall	**0.7629**(0.6052, 0.9680, 0.0707)	0.6966(0.5263, 0.9406, 0.0840)	0.7534(0.6111, 0.9634, 0.0721)
	F1 Score	**0.8502**(0.7366, 0.9716, 0.0437)	0.8046(0.6741, 0.9515, 0.0557)	0.8311(0.7294, 0.9444, 0.0435)
	AUC	0.9760(0.9485, 0.9991, 0.0107)	0.9605(0.9227, 0.9963, 0.0154)	**0.9793**(0.9605, 0.9983, 0.0081)
	β_1 slope	0.8696(0.7927, 0.9466, 0.0327)	0.8337(0.7451, 0.9356, 0.0406)	**0.9771**(0.9337, 1.0170, 0.0190)
	γ_1 slope	NA	NA	NA
Ph. -only	Accuracy	0.8638(0.7576, 0.9428, 0.0418)	0.9390(0.8696, 0.9848, 0.0234)	**0.9459**(0.8868, 0.9832, 0.0201)
	Precision	0.2102(0.0714, 0.6000, 0.1097)	**0.9481**(0.8829, 0.9862, 0.0207)	0.9192(0.8324, 0.9735, 0.0303)
	Recall	0.0059(0.0017, 0.0201, 0.0034)	0.5718(0.4207, 0.9026, 0.0970)	**0.6481**(0.5146, 0.9090, 0.0881)
	F1 Score	0.0114(0.0034, 0.0373, 0.0065)	0.7088(0.5831, 0.9269, 0.0725)	**0.7569**(0.6456, 0.9225, 0.0604)
	AUC	0.5277(0.4898, 0.6001, 0.0214)	0.9326(0.8844, 0.9930, 0.0237)	**0.9544**(0.9216, 0.9952, 0.0150)
	β_1 slope	NA	NA	NA
	γ_1 slope	NA	39.813(30.031, 45.799, 3.3285)	**0.9439**(0.8744, 1.0271, 0.0289)
Both	Accuracy	0.9644(0.9192, 0.9912, 0.0144)	**0.9835**(0.9616, 0.9984, 0.0071)	0.9769(0.9544, 0.9896, 0.0069)
	Precision	0.9643(0.9017, 0.9892, 0.0146)	**0.9798**(0.9354, 0.9967, 0.0106)	0.9134(0.7870, 0.9617, 0.0294)
	Recall	0.7606(0.6358, 0.9662, 0.0651)	0.8949(0.8216, 0.9925, 0.0362)	**0.9073**(0.8457, 0.9927, 0.0329)
	F1 Score	0.8489(0.7703, 0.9592, 0.0398)	**0.9350**(0.8926, 0.9908, 0.0202)	0.9097(0.8299, 0.9539, 0.0208)
	AUC	0.9763(0.9537, 0.9992, 0.0103)	0.9939(0.9846, 0.9999, 0.0035)	**0.9940**(0.9873, 0.9997, 0.0026)
	β_1 slope	0.8710(0.7879, 0.9789, 0.0321)	1.2346(1.1693, 1.3360, 0.0305)	**0.9843**(0.9365, 1.0316, 0.0183)
	γ_1 slope	NA	26.643(19.307, 30.993, 2.4747)	**0.9534**(0.8958, 1.0146, 0.0253)

for magnitude-based activity. However, such an assumption of magnitude-only activity is often unrealistic in real-world applications. When considering datasets comprising solely phase-active voxels, CV-M&P excels in all metrics except precision, thereby establishing its superiority in detecting phase-based activity. For mixed activity involving both magnitude and phase, CV-R&I takes the lead in accuracy, precision, and F1-score metrics, whereas CV-M&P dominates in recall and AUC.

CV-M&P once again stands out with respect to parameter estimation. Specifically, its true vs estimated parameter slopes are close to one when the parameters are present in the simulation, indicating accurate estimations. In contrast, this metric from both MO and CV-R&I deviates from the ideal value of one. It's noteworthy that Wang et al. (2023) elaborates that under conditions where all active voxels are solely magnitude-active, Lee et al. (2007)'s Cartesian model and Rowe (2005a)'s polar model (CV-R&I and CV-M&P) are approximately equivalent. Hence, in such scenario, CV-R&I can properly estimate β_1, although not surpassing the performance of CV-M&P. In other scenarios, both MO and CV-R&I fall short, either failing to estimate γ_1 or inaccurately estimating both β_1 and γ_1.

3.4 Analysis of Human CV-fMRI Data

In this study, we revisit the real data previously analyzed by Yu et al. (2018, 2023) and in our own Section 2.4. This dataset originates from a unilateral finger-tapping experiment conducted using a 3.0-Tesla General Electric Signa LX MRI scanner. The experimental design comprises 16.33 epochs, each consisting of alternating periods of 15s on and 15s off. Consequently, the total number of time points is $T = 490$, excluding the warm-up phase. The acquired dataset has seven slices, each with dimensions 96×96, and our analysis focuses on the initial six slices. For all examined models, we set the number of parcels $G = 25$. Specific tuning parameters are prefixed based on the experience: for the MO model, $\psi = \Phi^{-1}(0.02)$; for CV-R&I, $\psi = \Phi^{-1}(0.1)$; and for CV-M&P, $\psi = \omega = \Phi^{-1}(0.20)$. The thresholds for identifying active voxels are set to values as follows: 0.8722 for both MO and CV-R&I, and 0.925 for CV-M&P, in alignment with Section 3.2.3 and the simulation studies.

In Figure 3.6, we present the results derived from the CV-M&P model. Distinct patterns are observed: the estimated β_0 maps mirror the patterns of magnitude in the background, the estimated γ_0 maps highlight the phase's transition lines across different color zones, and both the estimated β_1 and γ_1 maps reflect patterns consistent with the estimated magnitude and phase activation maps. Such patterns are indicative of the accuracy of our approach in both classification and estimation.

By integrating the magnitude- and phase-activation maps derived by CV-M&P model, we form comprehensive estimated activation maps. They are subsequently compared with activation maps estimated by MO and CV-R&I models, as shown in Figure 3.7, revealing significant alignment. Specifically, the two central and central-left active regions detected by CV-M&P are consistent with the findings reported in Yu et al. (2018, 2023) and Wang et al. (2023). Furthermore, these regions align with known anatomical areas typically activated during finger-tapping tasks. The central region may correspond to the Primary Motor Cortex (M1) or Supplementary Motor Area (SMA), both of which play pivotal roles in voluntary movement and motor planning (Wilder Penfield, 1937; Geyer et al., 1996). Adjacently, the central-left region might represent the Primary Somatosensory Cortex (S1) or the Posterior Parietal Cortex, responsible for tactile sensory information processing and sensory-motor integration, respectively (Culham and Valyear, 2006). Notably, beyond these well-established regions, CV-M&P uncovers additional active regions at the posterior of the brain image. These could be caused by the motion of brain during the data collection.

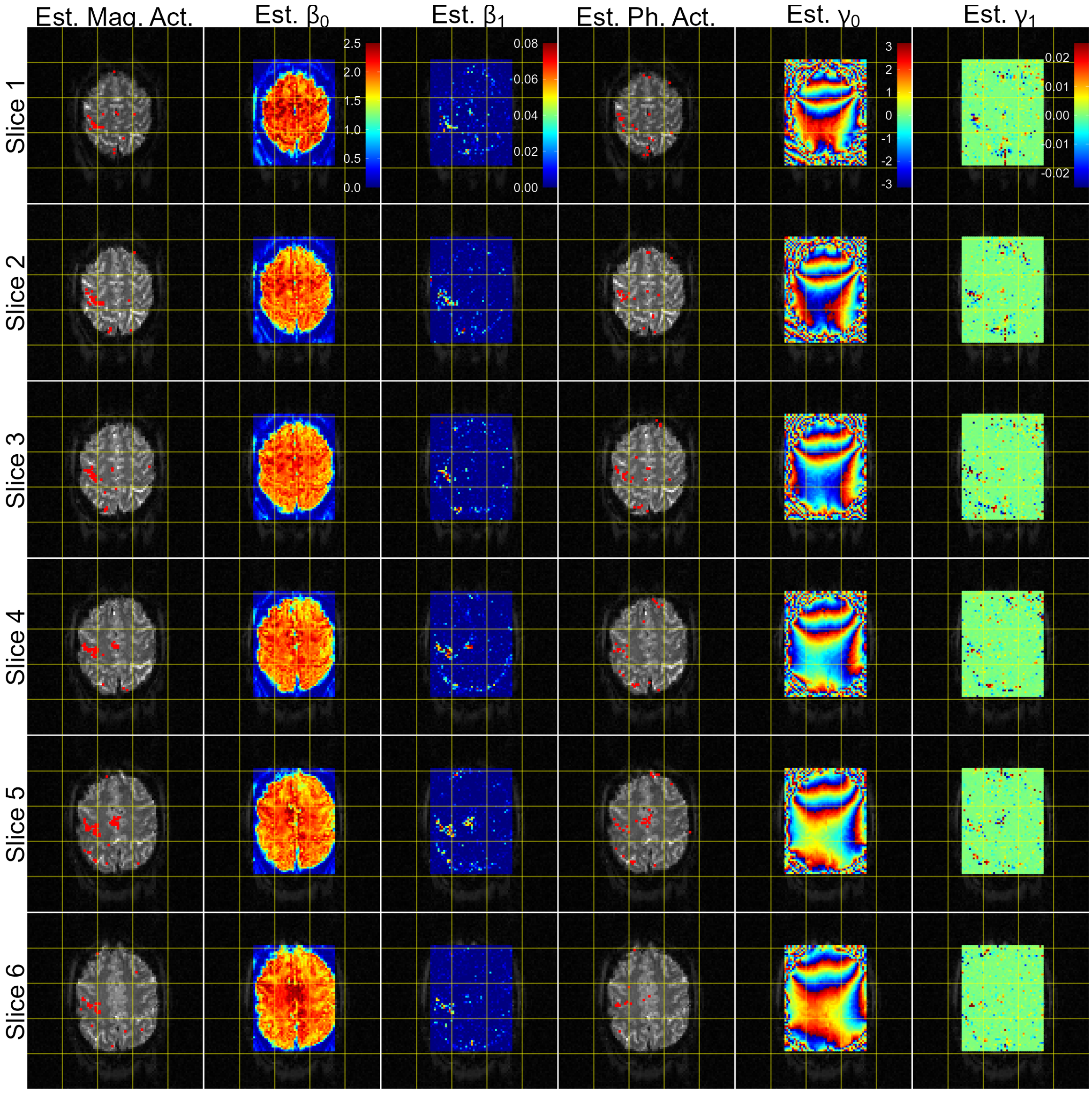


Figure 3.6: Estimated magnitude activation, β_0, β_1, phase activation, γ_1, γ_0 maps for a real human brain dataset as derived by the CV-M&P model.

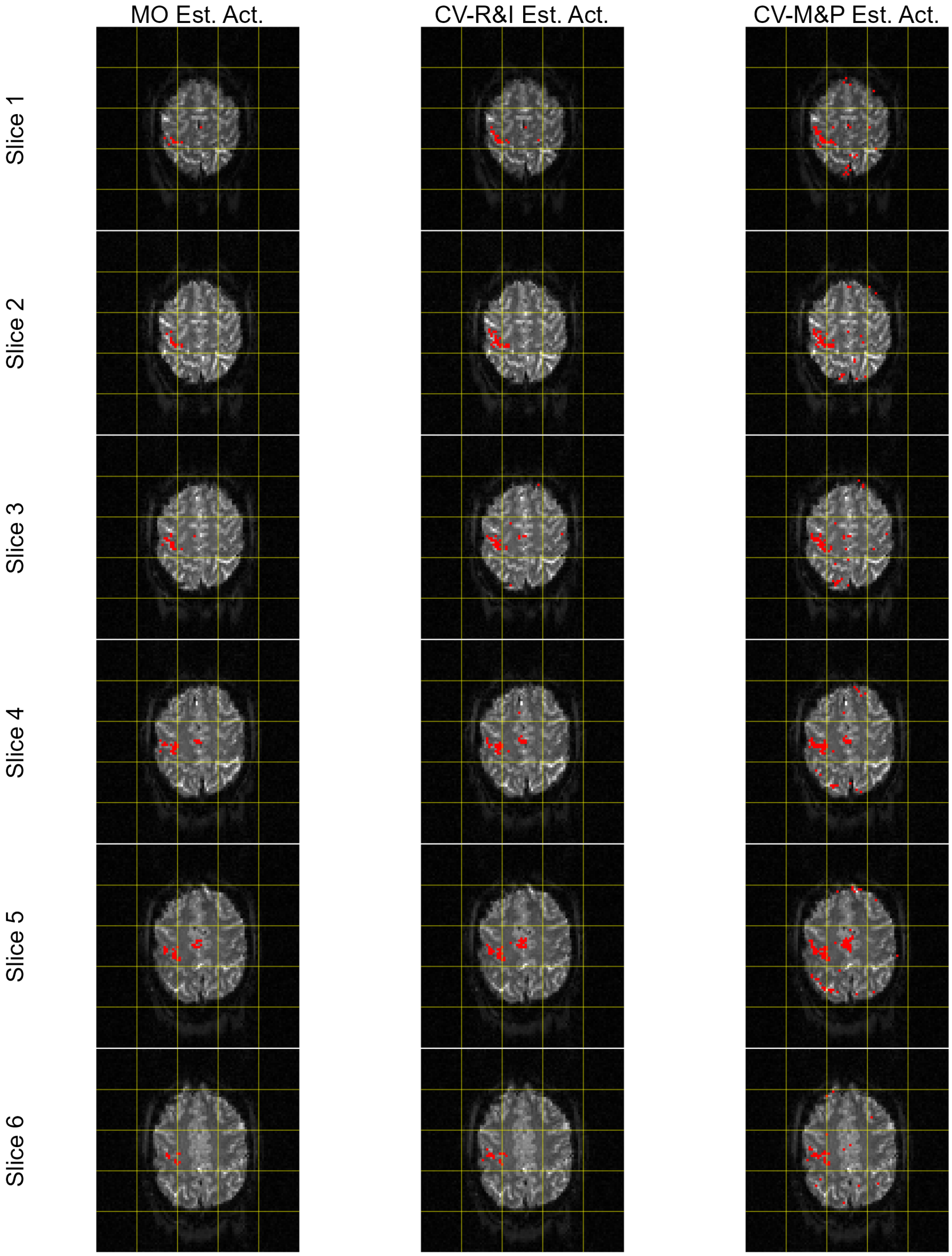


Figure 3.7: Estimated activation maps for a real human brain dataset as derived by the MO, CV-R&I, and CV-M&P models.

3.5 Conclusion

Throughout our investigations on both simulated and real human datasets, the CV-M&P model consistently demonstrates its capability to precisely identify voxels that exhibit significant reactions to stimuli, whether in magnitude, in phase, or in a combination of both. Comparing with the polar model of Rowe (2005a), but using hypothesis testing approaches (Rowe and Logan, 2004, 2005; Rowe, 2005a,b; Rowe et al., 2007; Rowe, 2009; Adrian et al., 2018), our CV-M&P's fully Bayesian framework can capture the spatial correlations of fMRI data, and therefore improve the model flexibly. On the other hand, comparing with other fully Bayesian approaches, but based on the Cartesian model of Lee et al. (2007) (Yu et al., 2023; Wang et al., 2023), our CV-M&P model rectifies the constraints inherent in the Cartesian models, which can detect active voxels but remains ambiguous about the exact type of the activation, as described in Rowe (2009). Moreover, the CV-M&P model excels in providing precise parameter estimates, offering a more nuanced and scientific framework for delineating brain activation patterns in task-based fMRI analyses. As we anticipate future work, there are multiple avenues for advancing this research. These include the exploration of more complex models that account for temporal correlations, models that fit the non-circular data wherein the real and imaginary components of the signal are correlated, and efforts aimed at optimizing computational efficiency.

Chapter 4

Conclusion

4.1 Summary of Research Contributions

The preceding chapters have presented a thorough investigation into brain activity mapping using single-subject task-based complex-valued fMRI time series. Grounded in statistical modeling, the dissertation pioneers two fully Bayesian approaches: CV-sSGLMM model based on the Cartesian model proposed by Lee et al. (2007) and CV-M&P model based on the polar model proposed by Rowe (2005a).

Chapter 2 introduces the CV-sSGLMM model that extends previous work by integrating both spatial and temporal dimensions of complex-valued fMRI data. The CV-sSGLMM model excels in handling cv-fMRI data with *iid* noise as well as AR(1) noise, showcasing an edge over existing models in active voxel detection. Through an array of metrics, such as the accuracy, recall, and F1-score, it establishes that the CV-sSGLMM model significantly outperforms its competitors, particularly in cases involving AR(1) noise. Another noteworthy advancement of CV-sSGLMM is the employment of brain parcellation and parallel computation, which enable a remarkable reduction in computational time by 87% comparing with the current state-of-the-art without compromising the prediction accuracy.

In contrast to the Cartesian model, which has inherent limitations, the poloar model offers a more comprehensive understanding to the nature of cv-fMRI signals by characterizing both magnitude and phase. As delineated in Chapter 3, our CV-M&P model maps the magnitude and phase activations individually and achieves a more precise quantification of activation strength. When compared to existing Bayesian methods in cv-fMRI analyses that primarily rely on a Cartesian model (Yu et al., 2018, 2023; Wang et al., 2023), our approach furnishes a more intricate and complete understanding of brain activation, with special emphasis on phase information. Furthermore, in comparison to cv-fMRI analyses based on the polar model, which predominantly employ hypothesis testing (Rowe and Logan, 2004, 2005; Rowe, 2005a,b; Rowe et al., 2007; Rowe, 2009; Adrian et al., 2018), our Bayesian approach offers a more flexible way of capturing the spatiotemporal correlations present in fMRI data.

4.2 Broader Implications

The methodological advances and findings from this dissertation have broader implications for neuroscience and medical imaging. The approaches can be particularly beneficial for studies focusing on cognitive functions, neurodevelopmental disorders like autism (Dichter, 2012), and neurodegenerative diseases such as Alzheimer's disease (Machulda et al., 2003). The Bayesian models proposed herein have the potential to be integrated into existing software packages, thereby providing researchers and clinicians with enhanced tools for mapping brain activity. The marked reduction in computational time, particularly evident in the CV-sSGLMM model, suggests the feasibility of real-time applications, thus opening doors to potential use in clinical settings.

Additionally, this work contributes to the broader area of brain activation mapping, commonly referred to as encoding. Previous studies have successfully employed encoding techniques in various modalities, including visual (Rombouts et al., 1999), auditory (Jäncke et al., 1998), and linguistic tasks (Jain and Huth, 2018). Such progress sets the stage for advancements in decoding, which involves the reconstruction of tasks—be it images or language—based on brain activation maps or fMRI signals. This holds particular significance for emerging areas like brain-computer interactions (Shen et al., 2019).

4.3 Limitations and Future Directions

While the Bayesian approaches presented in this dissertation offer promising outcomes, several directions for future research merit consideration. First, the existing Bayesian framework could be enhanced to incorporate a more generalized non-circular model. This adaptation would account for potential correlations between the real and imaginary components of cv-fMRI signals, as the current model operates under the assumption of circular data. Second, future efforts could be devoted to refining the computational efficiency of the CV-M&P model, thereby reducing the execution time and making it more practicable for real-world applications. Lastly, future work could explore the extension of the model to a multi-subject Bayesian framework, facilitating more robust generalizations across diverse populations.

4.4 Concluding Remarks

This dissertation has significantly advanced the field of brain activity mapping using fully Bayesian approaches for cv-fMRI data. Through rigorous statistical modeling and validation, it has demonstrated the superiority of the proposed Bayesian frameworks over existing models in various aspects. As this body of work stands at the intersection of statistics, neuroscience, and computer science, it opens up avenues for interdisciplinary research aimed at further unraveling the complexities of the human brain.

Appendices

Appendix A Supplementary Materials for Chapter 2

A.1 Similarities Between Lee's Cartesian Model and Rowe's Polar Model

This appendix is influenced by Rowe (2009), and seeks to demonstrate that, when there's only one stimulus:

- Lee et al. (2007)'s model is approximately equivalent to Rowe (2005a)'s dynamic phase model when the intercept in the magnitude is absent.
- Lee et al. (2007)'s model is fully equivalent to Rowe and Logan (2004)'s constant phase model.

For the first scenario, assuming no intercept in the magnitude, the v^{th} voxel's complex-valued fMRI signal can be simulated using Rowe (2005a)'s dynamic phase model as per equation:

$$\begin{aligned}\mathbf{y}_{Re}^{v} &= D_{Re}^{v}\mathbf{x}\beta^{v},\\ \mathbf{y}_{Im}^{v} &= D_{Im}^{v}\mathbf{x}\beta^{v},\end{aligned}$$

where $\mathbf{y}_{Re}^{v}$ and $\mathbf{y}_{Im}^{v}$ are simulated complex-valued fMRI vectors of length T, and $\mathbf{x}$ is the expected BOLD response of length T with β^{v} as a scalar. The matrices D_{Re}^{v} and D_{Im}^{v} are $T \times T$ and diagonal with $\cos\left(\theta_0 + \theta_1 x_t\right)$ and $\sin\left(\theta_0 + \theta_1 x_t\right)$ as the t^{th} diagonal element, which represent the dynamic phase. By equating this with the means of the Lee et al. (2007)'s model (without intercept), we have:

$$\begin{aligned}\mathbf{x}\beta_{Re}^{v} &= D_{Re}^{v}\mathbf{x}\beta^{v},\\ \mathbf{x}\beta_{Im}^{v} &= D_{Im}^{v}\mathbf{x}\beta^{v},\end{aligned}$$

where β_{Re}^{v} and β_{Im}^{v} are the scalar real and imaginary parts of the regression coefficient, and the maximum likelihood estimators of them are:

$$\begin{aligned}\widehat{\beta}_{Re}^{v} &= \left(\mathbf{x}'\mathbf{x}\right)^{-1}\mathbf{x}'D_{Re}^{v}\mathbf{x}\beta^{v},\\ \widehat{\beta}_{Im}^{v} &= \left(\mathbf{x}'\mathbf{x}\right)^{-1}\mathbf{x}'D_{Im}^{v}\mathbf{x}\beta^{v},\end{aligned}$$

then,

$$
\begin{aligned}
\widehat{\beta}_{Re}^{v,2}+\widehat{\beta}_{Im}^{v,2} &= \left[\left(\mathbf{x}'\mathbf{x}\right)^{-1}\mathbf{x}'D_{Re}^{v}\mathbf{x}\beta^{v}\right]^{2}+\left[\left(\mathbf{x}'\mathbf{x}\right)^{-1}\mathbf{x}'D_{Im}^{v}\mathbf{x}\beta^{v}\right]^{2}\\
&= \beta^{v,2}\left(\mathbf{x}'\mathbf{x}\right)^{-2}\left[\left(\mathbf{x}'D_{Re}^{v}\mathbf{x}\right)^{2}+\left(\mathbf{x}'D_{Im}^{v}\mathbf{x}\right)^{2}\right]\\
&= \beta^{v,2}\left(\mathbf{x}'\mathbf{x}\right)^{-2}\left[\mathbf{x}'D_{Re}^{v}\mathbf{x}\mathbf{x}'D_{Re}^{v}\mathbf{x}+\mathbf{x}'D_{Im}^{v}\mathbf{x}\mathbf{x}'D_{Im}^{v}\mathbf{x}\right]\\
&= \beta^{v,2}\left(\mathbf{x}'\mathbf{x}\right)^{-2}\left[\mathbf{x}'\left(D_{Re}^{v}\mathbf{x}\mathbf{x}'D_{Re}^{v}+D_{Im}^{v}\mathbf{x}\mathbf{x}'D_{Im}^{v}\right)\mathbf{x}\right].
\end{aligned}
$$

Notice that $D_{Re}^{v}\mathbf{x}\mathbf{x}'D_{Re}^{v}$ and $D_{Im}^{v}\mathbf{x}\mathbf{x}'D_{Im}^{v}$ are $T \times T$ symmetric matrices with the following terms as the (i,j)th element, respectively:

$$
\begin{aligned}
&x_i x_j \cos\left(\theta_0+\theta_1 x_i\right)\cos\left(\theta_0+\theta_1 x_j\right),\\
&x_i x_j \sin\left(\theta_0+\theta_1 x_i\right)\sin\left(\theta_0+\theta_1 x_j\right).
\end{aligned}
$$

Using the fact that $\cos(a)\cos(b)+\sin(a)\sin(b)=\cos(a-b)$, we have:

$$
D_{Re}^{v}\mathbf{x}\mathbf{x}'D_{Re}^{v}+D_{Im}^{v}\mathbf{x}\mathbf{x}'D_{Im}^{v} = \mathbf{x}\mathbf{x}'\odot\mathbf{P},
$$

where $\mathbf{P}$ is a $T \times T$ symmetric matrix and $\mathbf{P}_{(i,j)}=\cos\left(\theta_1\left(x_i-x_j\right)\right)$, and $\odot$ denotes the point-wise product. It's important to note that in both simulated and real data, $\mathbf{P}$ closely approximates the all-ones matrix $\mathbf{1}_{T\times T}$. This is because the difference between x_i and x_j is typically small, even when considering the extreme values. After multiplying this small difference with a small θ_1 and then taking the cosine, the result tends to be very close to 1. Thus,

$$
\sqrt{\widehat{\beta}_{Re}^{v,2}+\widehat{\beta}_{Im}^{v,2}} \approx \sqrt{\beta^{v,2}\left(\mathbf{x}'\mathbf{x}\right)^{-2}\left[\mathbf{x}'\left(\mathbf{x}\mathbf{x}'\right)\mathbf{x}\right]}=\beta^{v}.
$$

In this case, Lee et al. (2007)'s model can be considered as approximately equivalent to Rowe (2005a)'s dynamic phase model. For the second scenario, when the phase is constant and the intercept is included in the magnitude, using Rowe and Logan (2004)'s constant phase model to simulate the data, we get:

$$\mathbf{y}_{Re}^{v} = \Lambda_{Re}^{v} \begin{pmatrix} \mathbf{1} & \mathbf{x} \end{pmatrix} \begin{pmatrix} \beta_0^v \\ \beta_1^v \end{pmatrix},$$

$$\mathbf{y}_{Im}^{v} = \Lambda_{Im}^{v} \begin{pmatrix} \mathbf{1} & \mathbf{x} \end{pmatrix} \begin{pmatrix} \beta_0^v \\ \beta_1^v \end{pmatrix},$$

where $\Lambda_{Re}^{v} = \cos(\theta)\,\mathbf{I}_{T\times T}$ and $\Lambda_{Im}^{v} = \sin(\theta)\,\mathbf{I}_{T\times T}$. Upon equating this with the means of the Lee et al. (2007)'s model, we have:

$$\begin{aligned}
\begin{pmatrix} \mathbf{1} & \mathbf{x} \end{pmatrix} \begin{pmatrix} \beta_{Re,0}^v \\ \beta_{Re,1}^v \end{pmatrix} &= \Lambda_{Re}^{v} \begin{pmatrix} \mathbf{1} & \mathbf{x} \end{pmatrix} \begin{pmatrix} \beta_0^v \\ \beta_1^v \end{pmatrix}, \\
\begin{pmatrix} \mathbf{1} & \mathbf{x} \end{pmatrix} \begin{pmatrix} \beta_{Im,0}^v \\ \beta_{Im,1}^v \end{pmatrix} &= \Lambda_{Im}^{v} \begin{pmatrix} \mathbf{1} & \mathbf{x} \end{pmatrix} \begin{pmatrix} \beta_0^v \\ \beta_1^v \end{pmatrix}.
\end{aligned}$$

Since Λ_{Re}^{v} and Λ_{Im}^{v} don't contain $\mathbf{x}$, we can remove the means so that to remove the intercept in the model, which yields:

$$\begin{aligned}
\mathbf{x}_c \beta_{Re,1}^v &= \Lambda_{Re}^{v} \mathbf{x}_c \beta_1^v, \\
\mathbf{x}_c \beta_{Im,1}^v &= \Lambda_{Im}^{v} \mathbf{x}_c \beta_1^v,
\end{aligned}$$

where $\mathbf{x}_c$ is the centered $\mathbf{x}$. This becomes similar to the previous model:

$$\begin{aligned}
\widehat{\beta}_{Re,1}^{v,2} + \widehat{\beta}_{Im,1}^{v,2} &= \beta_1^{v,2} \left(\mathbf{x}_c'\mathbf{x}_c\right)^{-2} \left[\mathbf{x}_c' \left(\mathbf{x}_c\mathbf{x}_c' \odot \mathbf{P}\right)\mathbf{x}_c\right] \\
&= \beta_1^{v,2} \left(\mathbf{x}_c'\mathbf{x}_c\right)^{-2} \left[\mathbf{x}_c' \left(\mathbf{x}_c\mathbf{x}_c'\right)\mathbf{x}_c\right] = \beta_1^{v,2},
\end{aligned}$$

as $\mathbf{P}$ is exactly $\mathbf{1}_{T\times T}$ now. Consequently, Lee et al. (2007)'s model is found to be equivalent to Rowe and Logan (2004)'s constant phase model.

A.2 Full Conditional Posterior Distributions for CV-sSGLMM model

This appendix gives full conditional posterior distributions of $\lambda_v, \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau_g^2, \eta_v, \boldsymbol{\delta}_g, \kappa_g$ for Gibbs sampling. All derivations will omit the subscript of g (parcel index) from the parcel-level parameters τ_g^2, $\boldsymbol{\delta}_g$, and κ_g, since all parcels run the algorithm identically.

Full Conditional Distribution of λ_v

For the voxel v ($v = 1, ..., V$):

$$p(\lambda_v = 1 \mid \mathbf{y}_r^v, \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau^2, \eta_v) = \frac{p(\lambda_v = 1 \mid \eta_v)}{p(\lambda_v = 1 \mid \eta_v) + \frac{L_0}{L_1} \cdot p(\lambda_v = 0 \mid \eta_v)},$$

where

$$L_0 = p(\mathbf{y}_r^v, \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau^2 \mid \lambda_v = 0),$$
$$L_1 = p(\mathbf{y}_r^v, \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau^2 \mid \lambda_v = 1).$$

To determine L_0 and L_1, which are the joint distributions of $\mathbf{y}_r^v, \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau^2$ under the condition of $\lambda_v = 0$ and $\lambda_v = 1$, respectively, we recall the CV-sSGLMM model:

$$\mathbf{y}^v = \mathbf{x}\beta^v + \mathbf{r}^v \rho^v + \boldsymbol{\varepsilon}^v, \qquad \boldsymbol{\varepsilon}^v \sim \mathcal{CN}_T(\mathbf{0}, 2\sigma_v^2 \mathbf{I}, \mathbf{0}).$$

Applying Prais-Winsten transformation (order one backward operator) on $\mathbf{y}^v$ and $\mathbf{x}$, we have:

$$\mathbf{y}^{v*} = \mathbf{y}_{now}^v - \rho^v \mathbf{y}_{lag1}^v,$$
$$\mathbf{x}^{v*} = \mathbf{x}_{now} - \rho^v \mathbf{x}_{lag1},$$

where $\mathbf{y}_{now}^v$ and $\mathbf{y}_{lag1}^v$ are vectors containing the last and the first $T-1$ elements in $\mathbf{y}^v$, respectively. The vectors $\mathbf{x}_{now}$ and $\mathbf{x}_{lag1}$ are from $\mathbf{x}$ by the same rule of truncation. Now it becomes a model without autoregressive errors:

$$\mathbf{y}^{v*} = \mathbf{x}^{v*}\beta^v + \boldsymbol{\varepsilon}^v, \qquad \boldsymbol{\varepsilon}^v \sim \mathcal{CN}_{T-1}(\mathbf{0}, 2\sigma_v^2 \mathbf{I}, \mathbf{0}),$$

with equivalent real-valued representation:

$$\underbrace{\begin{pmatrix} \mathbf{y}_{Re}^{v*} \\ \mathbf{y}_{Im}^{v*} \end{pmatrix}}_{\mathbf{y}_r^{v*}} = \underbrace{\begin{pmatrix} \mathbf{x}_{Re}^{v*} & -\mathbf{x}_{Im}^{v*} \\ \mathbf{x}_{Im}^{v*} & \mathbf{x}_{Re}^{v*} \end{pmatrix}}_{\mathbf{X}_r^{v*}} \underbrace{\begin{pmatrix} \beta_{Re}^v \\ \beta_{Im}^v \end{pmatrix}}_{\boldsymbol{\beta}_r^v} + \underbrace{\begin{pmatrix} \boldsymbol{\varepsilon}_{Re}^v \\ \boldsymbol{\varepsilon}_{Im}^v \end{pmatrix}}_{\boldsymbol{\varepsilon}_r^v}.$$

Using the symbols in underbraces for a more compact form:

$$\mathbf{y}_r^{v*} = \mathbf{X}_r^{v*}\boldsymbol{\beta}_r^v + \boldsymbol{\varepsilon}_r^v, \qquad \boldsymbol{\varepsilon}_r^v \sim \mathcal{N}_{2(T-1)}(\mathbf{0}, \sigma_v^2\mathbf{I}).$$

Therefore, when $\lambda_v = 1$:

$$L_1 = p(\mathbf{y}_r^v, \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau^2) \propto p(\mathbf{y}_r^v \mid \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2)\, p(\boldsymbol{\beta}_r^v \mid \tau^2),$$

where

$$\begin{aligned} p(\mathbf{y}_r^v \mid \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2) &= (2\pi\sigma_v^2)^{-\frac{2(T-1)}{2}} \exp\left\{-\frac{1}{2\sigma_v^2}(\mathbf{y}_r^{v*} - \mathbf{X}_r^{v*}\boldsymbol{\beta}_r^v)'(\mathbf{y}_r^{v*} - \mathbf{X}_r^{v*}\boldsymbol{\beta}_r^v)\right\}, \\ p(\boldsymbol{\beta}_r^v \mid \tau^2) &= (2\pi\tau^2)^{-\frac{2}{2}} \exp\left\{-\frac{1}{2\tau^2}(\boldsymbol{\beta}_r^v)'(\boldsymbol{\beta}_r^v)\right\}. \end{aligned}$$

Similarly, when $\lambda_v = 0$:

$$L_0 = p(\mathbf{y}_r^v, \boldsymbol{\beta}_r^v = \mathbf{0}, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau^2) \propto p(\mathbf{y}_r^v \mid \boldsymbol{\beta}_r^v = \mathbf{0}, \boldsymbol{\rho}_r^v, \sigma_v^2)\, p(\boldsymbol{\beta}_r^v = \mathbf{0} \mid \tau^2),$$

where

$$\begin{aligned} p(\mathbf{y}_r^v \mid \boldsymbol{\beta}_r^v = \mathbf{0}, \boldsymbol{\rho}_r^v, \sigma_v^2) &= (2\pi\sigma_v^2)^{-\frac{2(T-1)}{2}} \exp\left\{-\frac{1}{2\sigma_v^2}(\mathbf{y}_r^{v*})'(\mathbf{y}_r^{v*})\right\}, \\ p(\boldsymbol{\beta}_r^v = \mathbf{0} \mid \tau^2) &= 1. \end{aligned}$$

Integrating $\boldsymbol{\beta}_r^v$ out of L_1 yields:

$$\begin{aligned} L_1^* = \quad & (2\pi\sigma_v^2)^{-\frac{2(T-1)}{2}} \cdot \frac{\sigma_v^2}{\tau^2} \cdot \exp\left\{-\frac{1}{2\sigma_v^2}(\mathbf{y}_r^{v*})'\mathbf{y}_r^{v*}\right\} \left\{\det\left[(\mathbf{X}_r^{v*})'\mathbf{X}_r^{v*} + \frac{\sigma_v^2}{\tau^2}\mathbf{I}\right]\right\}^{-\frac{1}{2}} \\ & \cdot \exp\left\{\frac{1}{2\sigma_v^2}\left[(\mathbf{X}_r^{v*})'\mathbf{y}_r^{v*}\right]' \left[(\mathbf{X}_r^{v*})'\mathbf{X}_r^{v*} + \frac{\sigma_v^2}{\tau^2}\mathbf{I}\right]^{-1} \left[(\mathbf{X}_r^{v*})'\mathbf{y}_r^{v*}\right]\right\}. \end{aligned}$$

Then, the ratio is:

$$\frac{L_0}{L_1^*} = \frac{\tau^2}{\sigma_v^2}\, \frac{\left\{\det\left[(\mathbf{X}_r^{v*})'\mathbf{X}_r^{v*} + \frac{\sigma_v^2}{\tau^2}\mathbf{I}\right]\right\}^{\frac{1}{2}}}{\exp\left\{\frac{1}{2\sigma_v^2}\left[(\mathbf{X}_r^{v*})'\mathbf{y}_r^{v*}\right]' \left[(\mathbf{X}_r^{v*})'\mathbf{X}_r^{v*} + \frac{\sigma_v^2}{\tau^2}\mathbf{I}\right]^{-1} \left[(\mathbf{X}_r^{v*})'\mathbf{y}_r^{v*}\right]\right\}}.$$

Using this ratio and $p(\lambda_v = 1 \mid \eta_v) = \mathbf{\Phi}(\psi + \eta_v)$, the full conditional distribution of λ_v is:

$$\pi(\lambda_v \mid \mathbf{y}_r^v, \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau^2, \eta_v) = \mathcal{B}ern\left(P\right),$$

where

$$P = p(\lambda_v = 1 \mid \mathbf{y}_r^v, \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau^2, \eta_v) = \frac{\mathbf{\Phi}(\psi + \eta_v)}{\mathbf{\Phi}(\psi + \eta_v) + \frac{L_0}{L_1^*} \cdot [1 - \mathbf{\Phi}(\psi + \eta_v)]}.$$

Full Conditional Distribution of $\boldsymbol{\beta}_r^v$

For the voxels with $\lambda_v = 0$, we assign them $\boldsymbol{\beta}_r^v = \mathbf{0}$. For the voxels with $\lambda_v = 1$:

$$\begin{aligned}
\pi(\boldsymbol{\beta}_r^v \mid \mathbf{y}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau^2) &\propto p(\mathbf{y}_r^v \mid \boldsymbol{\beta}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2) p(\boldsymbol{\beta}_r^v \mid \tau^2) \\
&\propto \exp\left\{-\frac{1}{2\sigma_v^2}(\mathbf{y}_r^{v*} - \mathbf{X}_r^{v*}\boldsymbol{\beta}_r^v)'(\mathbf{y}_r^{v*} - \mathbf{X}_r^{v*}\boldsymbol{\beta}_r^v)\right\} \exp\left\{-\frac{1}{2\tau^2}(\boldsymbol{\beta}_r^v)'(\boldsymbol{\beta}_r^v)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left[(\boldsymbol{\beta}_r^v)'\frac{(\mathbf{X}_r^{v*})'\mathbf{X}_r^{v*}}{\sigma_v^2}\boldsymbol{\beta}_r^v - 2(\boldsymbol{\beta}_r^v)'\frac{(\mathbf{X}_r^{v*})'}{\sigma_v^2}\mathbf{y}_r^{v*} + (\boldsymbol{\beta}_r^v)'\frac{1}{\tau^2}(\boldsymbol{\beta}_r^v)\right]\right\} \\
&= \exp\left\{-\frac{1}{2}\left[(\boldsymbol{\beta}_r^v)'\frac{\tau^2(\mathbf{X}_r^{v*})'\mathbf{X}_r^{v*} + \sigma_v^2\mathbf{I}}{\sigma_v^2\tau^2}\boldsymbol{\beta}_r^v - 2(\boldsymbol{\beta}_r^v)'\frac{(\mathbf{X}_r^{v*})'}{\sigma_v^2}\mathbf{y}_r^{v*}\right]\right\},
\end{aligned}$$

which is a kernel of multivariate normal distribution. Thus:

$$\pi(\boldsymbol{\beta}_r^v \mid \mathbf{y}_r^v, \boldsymbol{\rho}_r^v, \sigma_v^2, \tau^2, \lambda_v = 1) = \mathcal{N}_2(\boldsymbol{\mu}_{\boldsymbol{\beta}_r^v}, \mathbf{\Sigma}_{\boldsymbol{\beta}_r^v}),$$

where

$$\begin{aligned}
\boldsymbol{\mu}_{\boldsymbol{\beta}_r^v} &= \left[\frac{\tau^2(\mathbf{X}_r^{v*})'\mathbf{X}_r^{v*} + \sigma_v^2\mathbf{I}}{\sigma_v^2\tau^2}\right]^{-1}\frac{(\mathbf{X}_r^{v*})'}{\sigma_v^2}\mathbf{y}_r^{v*} = \left[(\mathbf{X}_r^{v*})'\mathbf{X}_r^{v*} + \frac{\sigma_v^2}{\tau^2}\mathbf{I}\right]^{-1}(\mathbf{X}_r^{v*})'\mathbf{y}_r^{v*}, \\
\mathbf{\Sigma}_{\boldsymbol{\beta}_r^v} &= \left[\frac{\tau^2(\mathbf{X}_r^{v*})'\mathbf{X}_r^{v*} + \sigma_v^2\mathbf{I}}{\sigma_v^2\tau^2}\right]^{-1} = \sigma_v^2\left[(\mathbf{X}_r^{v*})'\mathbf{X}_r^{v*} + \frac{\sigma_v^2}{\tau^2}\mathbf{I}\right]^{-1}.
\end{aligned}$$

Full Conditional Distribution of $\boldsymbol{\rho}_r^v$

Since $\boldsymbol{\rho}_r^v$ is the autoregression coefficient for AR(1) errors, let:

$$\mathbf{w}^v = \mathbf{y}^v - \mathbf{x}\beta^v$$

be the predicted errors. Let $\mathbf{w}^v_{now}$ and $\mathbf{w}^v_{lag1}$ be the vectors containing the last and the first $T-1$ components in $\mathbf{w}^v$, then:

$$\mathbf{w}^v_{now} = \mathbf{w}^v_{lag1}\rho^v + \boldsymbol{\varepsilon}^v, \qquad \boldsymbol{\varepsilon}^v \sim \mathcal{CN}_{T-1}(\mathbf{0}, 2\sigma^2_v\mathbf{I}, \mathbf{0}),$$

with equivalent real-valued representation:

$$\underbrace{\begin{pmatrix} \mathbf{w}^v_{now,Re} \\ \mathbf{w}^v_{now,Im} \end{pmatrix}}_{\mathbf{w}^v_{now,r}} = \underbrace{\begin{pmatrix} \mathbf{w}^v_{lag1,Re} & -\mathbf{w}^v_{lag1,Im} \\ \mathbf{w}^v_{lag1,Im} & \mathbf{w}^v_{lag1,Re} \end{pmatrix}}_{\mathbf{W}^v_{lag1,r}} \underbrace{\begin{pmatrix} \rho^v_{Re} \\ \rho^v_{Im} \end{pmatrix}}_{\boldsymbol{\rho}^v_r} + \underbrace{\begin{pmatrix} \boldsymbol{\varepsilon}^v_{Re} \\ \boldsymbol{\varepsilon}^v_{Im} \end{pmatrix}}_{\boldsymbol{\varepsilon}^v_r}.$$

Using the symbols in underbraces for a more compact form:

$$\mathbf{w}^v_{now,r} = \mathbf{W}^v_{lag1,r}\boldsymbol{\rho}^v_r + \boldsymbol{\varepsilon}^v, \qquad \boldsymbol{\varepsilon}^v \sim \mathcal{N}_{2(T-1)}(\mathbf{0}, \sigma^2_v\mathbf{I}).$$

Assigning a uniform prior, $p(\boldsymbol{\rho}^v_r) \propto 1$, the full conditional distribution of $\boldsymbol{\rho}^v_r$ is:

$$\pi(\boldsymbol{\rho}^v_r \mid \mathbf{y}^v_r, \cdot) = \mathcal{N}_2(\boldsymbol{\mu}_{\boldsymbol{\rho}^v_r}, \boldsymbol{\Sigma}_{\boldsymbol{\rho}^v_r}),$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\rho}^v_r} &= \left[(\mathbf{W}^v_{lag1,r})'\mathbf{W}^v_{lag1,r}\right]^{-1}(\mathbf{W}^v_{lag1,r})'\mathbf{w}^v_{now,r}, \\ \boldsymbol{\Sigma}_{\boldsymbol{\rho}^v_r} &= \sigma^2_v\left[(\mathbf{W}^v_{lag1,r})'\mathbf{W}^v_{lag1,r}\right]^{-1}. \end{aligned}$$

Full Conditional Distribution of σ^v_r

The full conditional distribution of σ^v_r is also from:

$$\mathbf{w}^v_{now,r} = \mathbf{W}^v_{lag1,r}\boldsymbol{\rho}^v_r + \boldsymbol{\varepsilon}^v, \qquad \boldsymbol{\varepsilon}^v \sim \mathcal{N}_{2(T-1)}(\mathbf{0}, \sigma^2_v\mathbf{I}).$$

Assigning a Jeffreys prior, $p(\sigma^2_v) \propto 1/\sigma^2_v$, we have:

$$\pi(\sigma^2_v \mid \mathbf{y}^v_r, \cdot) = \mathcal{IG}\left(\frac{2(T-1)}{2}, \quad \frac{1}{2}(\mathbf{w}^v_{now,r} - \mathbf{W}^v_{lag1,r}\boldsymbol{\rho}^v_r)'(\mathbf{w}^v_{now,r} - \mathbf{W}^v_{lag1,r}\boldsymbol{\rho}^v_r)\right).$$

Full Conditional Distribution of τ^2

The full conditional distribution of τ^2 should be related to the number of active voxels and could be imposed a Jeffreys prior, $p(\tau^2) \propto 1/\tau^2$. After updating $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_V)'$ and filtering $\boldsymbol{\beta}_r = (\beta_{Re}^1, \cdots, \beta_{Re}^V, \beta_{Im}^1, \cdots, \beta_{Im}^V)'$ by $\boldsymbol{\lambda}$ to make them strictly zeros and non-zeros in each iteration, we have:

$$\pi(\tau^2 \mid \boldsymbol{\beta}_r) = \mathcal{IG}\left(\frac{2\boldsymbol{\lambda}'\boldsymbol{\lambda}}{2}, \frac{1}{2}\boldsymbol{\beta}_r{}'\boldsymbol{\beta}_r\right).$$

Full Conditional Distribution of η_v

Without considering the condition of λ_v, we focus on $\pi(\eta_v \mid \kappa)$ first. Let $\mathbf{Q}_s = \mathbf{M}'\mathbf{Q}\mathbf{M}$ and $\mathbf{Q}_{\kappa s} = \kappa\mathbf{Q}_s = \kappa\mathbf{M}'\mathbf{Q}\mathbf{M}$, then:

$$\begin{aligned}
\pi(\eta_v \mid \kappa) &= \int \pi(\eta_v, \boldsymbol{\delta} \mid \kappa) d\boldsymbol{\delta} \\
&= \int \pi(\eta_v \mid \boldsymbol{\delta})\pi(\boldsymbol{\delta} \mid \kappa) d\boldsymbol{\delta} \\
&= \int \mathcal{N}(\mathbf{m}_v{}'\boldsymbol{\delta}, 1) \times \mathcal{N}\left(\mathbf{0}, \mathbf{Q}_{\kappa s}^{-1}\right) d\boldsymbol{\delta} \\
&\propto \int \exp\left\{-\frac{\eta_v^2 - 2\mathbf{m}_v{}'\boldsymbol{\delta}\eta_v + \boldsymbol{\delta}'\mathbf{m}_v\mathbf{m}_v{}'\boldsymbol{\delta}}{2}\right\} \exp\left\{-\frac{\boldsymbol{\delta}'\mathbf{Q}_{\kappa s}\boldsymbol{\delta}}{2}\right\} d\boldsymbol{\delta} \\
&= \exp\left\{-\frac{\eta_v^2}{2}\right\} \int \exp\left\{-\frac{1}{2}\left[\boldsymbol{\delta}'(\mathbf{Q}_{\kappa s} + \mathbf{m}_v\mathbf{m}_v{}')\boldsymbol{\delta} - 2\mathbf{m}_v{}'\boldsymbol{\delta}\eta_v\right]\right\} d\boldsymbol{\delta} \\
&= \exp\left\{-\frac{\eta_v^2}{2\left[1 - \mathbf{m}_v{}'(\mathbf{Q}_{\kappa s} + \mathbf{m}_v\mathbf{m}_v{}')^{-1}\mathbf{m}_v\right]^{-1}}\right\}.
\end{aligned}$$

Thus, $\eta_v \mid \kappa$ follows normal distribution with mean 0 and variance:

$$\left[1 - \mathbf{m}_v{}'(\mathbf{Q}_{\kappa s} + \mathbf{m}_v\mathbf{m}_v{}')^{-1}\mathbf{m}_v\right]^{-1}.$$

By Woodbury's matrix identity:

$$\left[1 - \mathbf{m}_v{}'(\mathbf{Q}_{\kappa s} + \mathbf{m}_v\mathbf{m}_v{}')^{-1}\mathbf{m}_v\right]^{-1} = 1 + \mathbf{m}_v{}'\mathbf{Q}_{\kappa s}^{-1}\mathbf{m}_v.$$

That is:

$$\pi(\eta_v \mid \kappa) = \mathcal{N}(0, \quad 1 + \mathbf{m}_v{}'\mathbf{Q}_{\kappa s}^{-1}\mathbf{m}_v).$$

If the condition of λ_v is considered, by Albert and Chib (1993):

$$\pi(\eta_v \mid \lambda_v, \boldsymbol{\delta}) = \begin{cases} \mathcal{TN}(\mathbf{m}_v{}'\boldsymbol{\delta},\ 1,\ 0,\ \infty) & \text{if } \lambda_v = 1 \\ \mathcal{TN}(\mathbf{m}_v{}'\boldsymbol{\delta},\ 1,\ -\infty,\ 0) & \text{if } \lambda_v = 0 \end{cases},$$

where $\mathcal{TN}$ denotes the truncated normal distribution. Thus, when $\lambda_v = 1$:

$$\begin{aligned} \pi(\eta_v \mid \lambda_v, \kappa) &= \int \pi(\eta_v, \boldsymbol{\delta} \mid \lambda_v, \kappa) d\boldsymbol{\delta} \\ &= \int \pi(\eta_v \mid \lambda_v, \boldsymbol{\delta}) \pi(\boldsymbol{\delta} \mid \kappa) d\boldsymbol{\delta} \\ &= \int \mathcal{TN}(\mathbf{m}_v{}'\boldsymbol{\delta},\ 1,\ 0,\ \infty) \times \mathcal{N}\left(\mathbf{0}, \mathbf{Q}_{\kappa s}^{-1}\right) d\boldsymbol{\delta} \\ &= \mathcal{TN}\left(0,\ 1 + \mathbf{m}_v{}'\mathbf{Q}_{\kappa s}^{-1}\mathbf{m}_v,\ 0,\ \infty\right). \end{aligned}$$

Similarly, when $\lambda_v = 0$:

$$\pi(\eta_v \mid \lambda_v, \kappa) \ = \mathcal{TN}\left(0,\ 1 + \mathbf{m}_v{}'\mathbf{Q}_{\kappa s}^{-1}\mathbf{m}_v,\ -\infty,\ 0\right).$$

Notice that the variance $1 + \mathbf{m}_v{}'\mathbf{Q}_{\kappa s}^{-1}\mathbf{m}_v = 1 + \mathbf{m}_v{}'(\kappa \mathbf{Q}_s)^{-1}\mathbf{m}_v$. As κ functions as a spatial smoothing parameter, it can be moved out of the parentheses to control the entire variance and play the same role. That is:

$$\pi(\eta_v \mid \lambda_v, \kappa) = \begin{cases} \mathcal{TN}(0,\ \frac{1}{\kappa}\underbrace{\left(1 + \mathbf{m}_v{}'\mathbf{Q}_s^{-1}\mathbf{m}_v\right)}_{\nu_v^2},\ 0,\ \infty) & \text{if } \lambda_v = 1 \\ \mathcal{TN}(0,\ \frac{1}{\kappa}\underbrace{\left(1 + \mathbf{m}_v{}'\mathbf{Q}_s^{-1}\mathbf{m}_v\right)}_{\nu_v^2},\ -\infty,\ 0) & \text{if } \lambda_v = 0 \end{cases}.$$

Since $\mathbf{I} + \mathbf{M}\mathbf{Q}_s^{-1}\mathbf{M}'$ doesn't contain any parameters, it can be pre-calculated, then $\nu_v^2 = 1 + \mathbf{m}_v{}'\mathbf{Q}_s^{-1}\mathbf{m}_v$ is its v^{th} diagonal element. This will accelerate the computation.

Full Conditional Distribution of $\boldsymbol{\delta}$

The full conditional distribution of $\boldsymbol{\delta}$ is:

$$\pi(\boldsymbol{\delta} \mid \boldsymbol{\eta}, \kappa) = \mathcal{N}\left(\left(\mathbf{Q}_{\kappa s} + \mathbf{M}'\mathbf{M}\right)^{-1}\mathbf{M}'\boldsymbol{\eta}, \quad \left(\mathbf{Q}_{\kappa s} + \mathbf{M}'\mathbf{M}\right)^{-1}\right).$$

Similar to how we deal with κ for η_v, this distribution becomes:

$$\pi(\boldsymbol{\delta} \mid \boldsymbol{\eta}, \kappa) = \mathcal{N}\left(\frac{1}{\kappa}\underbrace{(\mathbf{Q}_s + \mathbf{M}'\mathbf{M})^{-1}}_{\widehat{\mathbf{Q}}_s^{-1}}\mathbf{M}'\boldsymbol{\eta}, \quad \frac{1}{\kappa}\underbrace{(\mathbf{Q}_s + \mathbf{M}'\mathbf{M})^{-1}}_{\widehat{\mathbf{Q}}_s^{-1}}\right),$$

where $\widehat{\mathbf{Q}}_s^{-1} = (\mathbf{Q}_s + \mathbf{M}'\mathbf{M})^{-1}$ can be pre-calculated to accelerate the computation.

Full Conditional Distribution of κ

We assume $\eta_1, ..., \eta_V$ are conditionally independent when given κ, thus:

$$\begin{aligned}
\pi(\boldsymbol{\eta} \mid \kappa) &= \prod_{v=1}^{V} \pi(\eta_v \mid \kappa) \\
&= \left[\left(\frac{1}{\kappa}\right)^{-\frac{V}{2}} \prod_{v=1}^{V} \left(1 + \mathbf{m}_v{}'\mathbf{Q}_s^{-1}\mathbf{m}_v\right)^{-\frac{1}{2}}\right] \exp\left\{-\sum_{v=1}^{V} \frac{\eta_v^2}{2 \cdot \frac{1}{\kappa} \cdot (1 + \mathbf{m}_v{}'\mathbf{Q}_s^{-1}\mathbf{m}_v)}\right\} \\
&\propto \kappa^{\frac{V}{2}} \cdot \exp\left\{-\kappa \cdot \frac{1}{2} \cdot \sum_{v=1}^{V} \frac{\eta_v^2}{(1 + \mathbf{m}_v{}'\mathbf{Q}_s^{-1}\mathbf{m}_v)}\right\}.
\end{aligned}$$

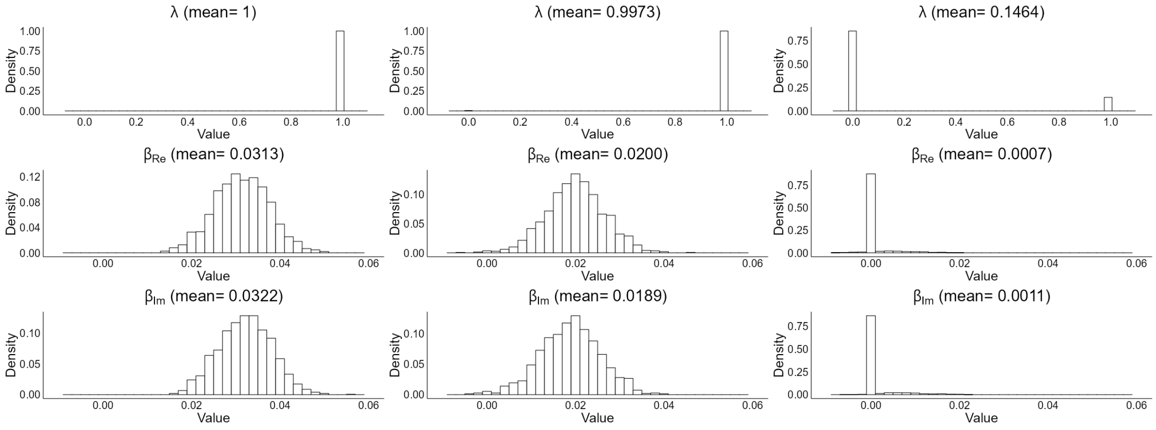
Therefore, the full conditional distribution of κ is:

$$\begin{aligned}
\pi(\kappa \mid \boldsymbol{\eta}) &\propto \pi(\boldsymbol{\eta} \mid \kappa)\pi(\kappa) \\
&\propto \kappa^{\frac{V}{2}} \cdot \exp\left\{-\kappa \cdot \frac{1}{2} \cdot \sum_{v=1}^{V} \frac{\eta_v^2}{(1 + \mathbf{m}_v{}'\mathbf{Q}_s^{-1}\mathbf{m}_v)}\right\} \cdot \kappa^{\frac{1}{2}-1} \cdot \exp\left\{-\frac{\kappa}{2000}\right\} \\
&= \kappa^{\frac{V+1}{2}-1} \exp\left\{-\kappa\left[\frac{1}{2}\left(\sum_{v=1}^{V} \frac{\eta_v^2}{(1 + \mathbf{m}_v{}'\mathbf{Q}_s^{-1}\mathbf{m}_v)}\right) + \frac{1}{2000}\right]\right\}.
\end{aligned}$$

That is:

$$\begin{aligned}
\pi(\kappa \mid \boldsymbol{\eta}) &= \mathcal{G}amma\left(a = \frac{V+1}{2}, \quad b = \left[\frac{1}{2}\left(\sum_{v=1}^{V} \frac{\eta_v^2}{(1 + \mathbf{m}_v{}'\mathbf{Q}_s^{-1}\mathbf{m}_v)}\right) + \frac{1}{2000}\right]^{-1}\right) \\
&= \mathcal{G}amma\left(a = \frac{V+1}{2}, \quad b = \left[\frac{1}{2}\left(\frac{\eta_1^2}{\nu_1^2} + \cdots + \frac{\eta_V^2}{\nu_V^2}\right) + \frac{1}{2000}\right]^{-1}\right),
\end{aligned}$$

where b is the scale, and the details for ν_v^2 are in the full conditional distribution of η_v.

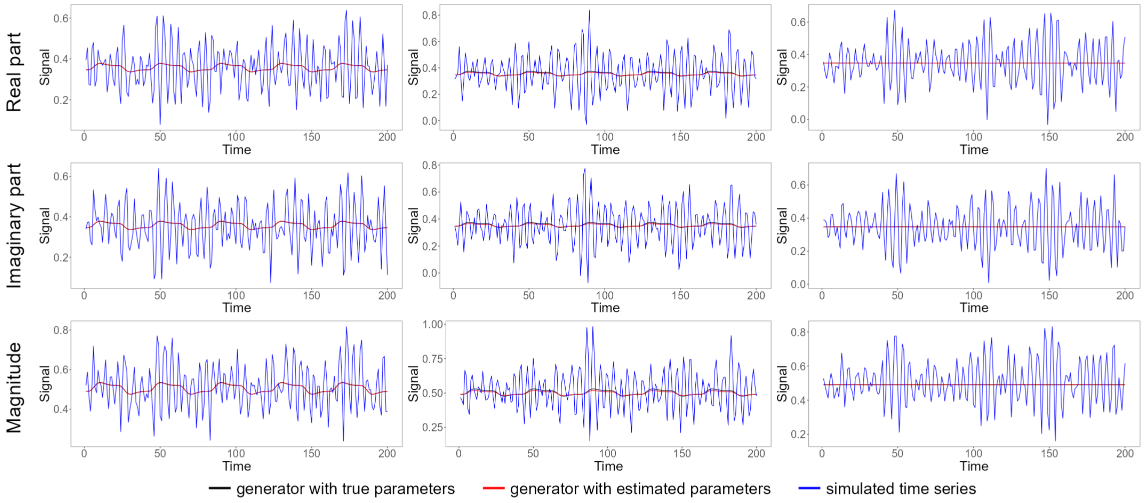


(a) Strong voxel (b) Medium voxel (c) Inactive voxel

Figure 1: (a) is marginal posterior distributions of λ, β_{Re}, and β_{Im} for a voxel exhibiting high strength. (b)-(c) are similar distributions for a medium-strength voxel and an inactive voxel, respectively.

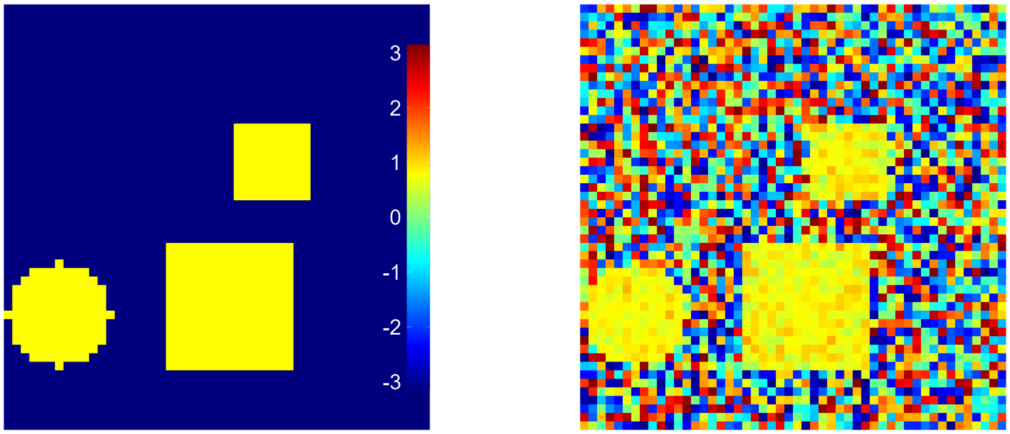
A.3 More Estimations by CV-sSGLMM

The CV-sSGLMM model is applied to estimate the marginal posterior distributions from three distinct types of voxels (strongly active, moderately active, inactive) within an AR(1) dataset, as showcased in Figure 1. The bell-shaped distributions of β_{Re} and β_{Im} corroborate the theoretical derivation and affirm the reliable performance of the MCMC algorithm during the sampling process. The true and estimated time series from these three voxel are presented in Figure 2. The congruence between the generator using true parameters (in black) and that using estimated parameters (in red) is evident. Additionally, both sets of time series aptly capture the pattern of the simulated time series (in blue). This alignment serves as a further testament to the good estimation performance of our CV-sSGLMM model. The phase of voxels is also estimated by the CV-sSGLMM model, and the outcomes are displayed in Figure 3. Figure 3(a) presents the true phase map, which is constant for all voxels as $\theta = \pi/4 \approx 0.79$. Figure 3(b) demonstrates that the CV-sSGLMM model effectively estimated this phase map by $\widehat{\theta}_v = \arctan_4\left(\widehat{\beta}^v_{Im}/\widehat{\beta}^v_{Rm}\right)$.



(a) Strong voxel (b) Medium voxel (c) Inactive voxel

Figure 2: (a) is time series of generator using true parameters, generator using estimated parameters, and simulated time series (generator using true parameters with noise) of a voxel exhibiting high strength. (b)-(c) are similar time series for a medium-strength voxel and an inactive voxel, respectively.



(a) True phase map (b) Estimated phase map

Figure 3: (a) is the true phase map of an AR(1) dataset. (b) is the estimated phase map as produced by the CV-sSGLMM model.

Appendix B Supplementary Materials for Chapter 3

B.1 Full Conditional Posterior Distributions for CV-M&P Model

We need the full conditional posterior distributions of

$$\boldsymbol{\beta}_v, \lambda_v, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2, \tau_g^2, \xi_g^2, \eta_{\lambda,v}, \boldsymbol{\delta}_{\lambda,g}, \kappa_{\lambda,g}, \eta_{\omega,v}, \boldsymbol{\delta}_{\omega,g}, \kappa_{\omega,g}$$

for the Gibbs sampling. All derivations will omit the subscript of g (parcel index) from the parcel-level parameters $\tau_g^2, \xi_g^2, \boldsymbol{\delta}_{\lambda,g}, \kappa_{\lambda,g}, \boldsymbol{\delta}_{\omega,g}, \kappa_{\omega,g}$, since all parcels run the algorithm identically.

Full conditional distributions of λ_v and ω_v

The full conditional distribution of λ_v is

$$\pi(\lambda_v \mid \mathbf{y}_r^v, \boldsymbol{\beta}_v, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2, \tau^2, \xi^2, \eta_{\lambda,v}) = \mathcal{B}ern\left(P_{\lambda_v}\right),$$

where

$$\begin{aligned} P_{\lambda_v} &= p(\lambda_v = 1 \mid \mathbf{y}_r^v, \boldsymbol{\beta}_v, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2, \tau^2, \xi^2, \eta_{\lambda,v}) \\ &= \frac{p(\lambda_v = 1 \mid \eta_{\lambda,v})}{p(\lambda_v = 1 \mid \eta_{\lambda,v}) + \frac{L_0}{L_1} \cdot p(\lambda_v = 0 \mid \eta_{\lambda,v})} \\ &= \frac{\boldsymbol{\Phi}(\psi_\lambda + \eta_{\lambda,v})}{\boldsymbol{\Phi}(\psi_\lambda + \eta_{\lambda,v}) + \frac{L_0}{L_1} \cdot \left[1 - \boldsymbol{\Phi}(\psi_\lambda + \eta_{\lambda,v})\right]}, \end{aligned}$$

and L_0 and L_1 are the joint densities of $\mathbf{y}_v, \boldsymbol{\beta}_v, \lambda_v, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2, \tau^2, \xi^2$ given $\lambda_v = 0$ and $\lambda_v = 1$. Let L be such joint density, that is,

$$L = p(\mathbf{y}_v, \boldsymbol{\beta}_v, \lambda_v, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2, \tau^2, \xi^2) \propto p(\mathbf{y}_v \mid \boldsymbol{\beta}_v, \lambda_v, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2) p(\boldsymbol{\beta}_v \mid \lambda_v, \tau^2).$$

Define:

$$\mathbf{A}_v = \begin{pmatrix} \mathbf{C}_v \\ \mathbf{S}_v \end{pmatrix}, \quad \text{where } \mathbf{C}_v = \operatorname{diag}\left[\cos\left(\mathbf{U}\boldsymbol{\Omega}_v \boldsymbol{\gamma}_v\right)\right], \quad \mathbf{S}_v = \operatorname{diag}\left[\sin\left(\mathbf{U}\boldsymbol{\Omega}_v \boldsymbol{\gamma}_v\right)\right],$$

(Note, $\mathbf{A}_v$ is orthogonal, i.e., $\mathbf{A}_v'\mathbf{A}_v = \mathbf{I}_{2T}$.), then,

$$
\begin{aligned}
&p(\mathbf{y}_v \mid \boldsymbol{\beta}_v, \lambda_v, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2) \\
&= (2\pi\sigma_v^2)^{-\frac{2T}{2}} \exp\left\{-\frac{1}{2\sigma_v^2}\left(\mathbf{y}_v - \mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)'\left(\mathbf{y}_v - \mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)\right\} \\
&= (2\pi\sigma_v^2)^{-\frac{2T}{2}} \exp\left\{-\frac{1}{2\sigma_v^2}\left[\mathbf{y}_v'\mathbf{y}_v - 2\left(\mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)'\mathbf{y}_v + \left(\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)'\mathbf{X}'\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right]\right\},
\end{aligned}
$$

and

$$p(\boldsymbol{\beta}_v \mid \lambda_v, \tau^2) = (2\pi\tau^2)^{-\frac{1+\lambda_v}{2}} \exp\left\{-\frac{1}{2\tau^2}\boldsymbol{\beta}_v'\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right\}.$$

Thus,

$$L \propto (2\pi\tau^2)^{-\frac{1+\lambda_v}{2}} \exp\left\{-\frac{1}{2\sigma_v^2}\left[-2\left(\mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)'\mathbf{y}_v + \left(\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)'\mathbf{X}'\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right] - \frac{1}{2\tau^2}\boldsymbol{\beta}_v'\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right\}.$$

Let $\mathbf{a}_v$ be the flattened version of $\mathbf{A}_v$, that is, $\mathbf{a}_v$ is a $2T \times 1$ vector as

$$\mathbf{a}_v = \mathbf{A}_v\mathbf{1}_T = \begin{pmatrix} \cos\left(\mathbf{U}\boldsymbol{\Omega}_v\boldsymbol{\gamma}_v\right) \\ \sin\left(\mathbf{U}\boldsymbol{\Omega}_v\boldsymbol{\gamma}_v\right) \end{pmatrix}$$

Also, define $\mathbf{x}_{(2)}$ as the second column of $\mathbf{X}$, thus, $\mathbf{x}_{(2)}$ is a $T \times 1$ vector of expected BOLD response; define $\mathbf{x}_{(2)}^* = \begin{pmatrix} \mathbf{x}_{(2)} \\ \mathbf{x}_{(2)} \end{pmatrix}$ as a $2T \times 1$ vector to match the dimension, then,

$$
\begin{aligned}
\frac{L_0}{L_1} &= \frac{L \mid_{\lambda_v=0}}{L \mid_{\lambda_v=1}} \\
&= (2\pi\tau^2)^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_v^2}\left[2\beta_{v,1}\left(\mathbf{x}_{(2)}^* \odot \mathbf{a}_v\right)'\mathbf{y}_v - 2\beta_{v,0}\beta_{v,1}\mathbf{x}_{(2)}'\mathbf{1}_T - \beta_{v,1}^2\mathbf{x}_{(2)}'\mathbf{x}_{(2)}\right] + \frac{1}{2\tau^2}\beta_{v,1}^2\right\}.
\end{aligned}
$$

We flatten $\mathbf{A}_v$ and use Hadamard product $\odot$ here to lessen the computational burden. Similarly, the full conditional distribution of ω_v is

$$\pi(\omega_v \mid \mathbf{y}_r^v, \boldsymbol{\beta}_v, \lambda_v, \boldsymbol{\gamma}_v, \sigma_v^2, \tau^2, \xi^2, \eta_v) = \mathcal{B}ern\left(P_{\omega_v}\right),$$

where

$$P_{\omega_v} = \frac{\boldsymbol{\Phi}(\psi_\omega + \eta_{\omega,v})}{\boldsymbol{\Phi}(\psi_\omega + \eta_{\omega,v}) + \frac{L_0}{L_1}\cdot\left[1 - \boldsymbol{\Phi}(\psi_\omega + \eta_{\omega,v})\right]},$$

where

$$L \propto (2\pi\xi^2)^{-\frac{1+\omega_v}{2}} \exp\left\{\frac{1}{\sigma_v^2}\left(\mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)'\mathbf{y}_v - \frac{1}{2\xi^2}\boldsymbol{\gamma}_v'\boldsymbol{\Omega}_v\boldsymbol{\gamma}_v\right\},$$

and

$$\begin{aligned}\frac{L_0}{L_1} &= \frac{L\mid_{\omega_v=0}}{L\mid_{\omega_v=1}}\\ &= (2\pi\xi^2)^{\frac{1}{2}}\exp\left\{\frac{1}{\sigma_v^2}\left[\left(\mathbf{A}_v\mid_{\omega_v=0} - \mathbf{A}_v\mid_{\omega_v=1}\right)\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right]'\mathbf{y}_v + \frac{1}{2\xi^2}\gamma_{v,1}^2\right\}.\end{aligned}$$

Keep simplifying it, when $\lambda_v = 0$,

$$\frac{L_0}{L_1} = (2\pi\xi^2)^{\frac{1}{2}}\exp\left\{\frac{1}{\sigma_v^2}\beta_{v,0}\left(\mathbf{a}_v\mid_{\omega_v=0} - \mathbf{a}_v\mid_{\omega_v=1}\right)'\mathbf{y}_v + \frac{1}{2\xi^2}\gamma_{v,1}^2\right\}.$$

When $\lambda_v = 1$,

$$\frac{L_0}{L_1} = (2\pi\xi^2)^{\frac{1}{2}}\exp\left\{\frac{1}{\sigma_v^2}\left[\left(\beta_{v,0}\mathbf{1}_{2T} + \beta_{v,1}\mathbf{x}^*_{(2)}\right)\odot\left(\mathbf{a}_v\mid_{\omega_v=0} - \mathbf{a}_v\mid_{\omega_v=1}\right)\right]'\mathbf{y}_v + \frac{1}{2\xi^2}\gamma_{v,1}^2\right\}.$$

Full conditional distribution of $\boldsymbol{\beta}_v$

When $\lambda_v = 1$, the full conditional distribution of $\boldsymbol{\beta}_v$ is

$$\begin{aligned}&\pi(\boldsymbol{\beta}_v \mid \mathbf{y}_v, \lambda_v = 1, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2, \tau^2, \xi^2)\\ &\propto p(\mathbf{y}_v, \boldsymbol{\beta}_v, \lambda_v = 1, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2, \tau^2, \xi^2)\\ &\propto p(\mathbf{y}_v \mid \boldsymbol{\beta}_v, \lambda_v = 1, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2)p(\boldsymbol{\beta}_v \mid \lambda_v = 1, \tau^2)\\ &\propto \exp\left\{-\frac{1}{2\sigma_v^2}\left(\mathbf{y}_v - \mathbf{A}_v\mathbf{X}\boldsymbol{\beta}_v\right)'\left(\mathbf{y}_v - \mathbf{A}_v\mathbf{X}\boldsymbol{\beta}_v\right)\right\}\exp\left\{-\frac{1}{2\tau^2}\boldsymbol{\beta}_v'\boldsymbol{\beta}_v\right\}\\ &\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\beta}_v'\frac{\left(\mathbf{A}_v\mathbf{X}\right)'\left(\mathbf{A}_v\mathbf{X}\right)}{\sigma_v^2}\boldsymbol{\beta}_v - 2\boldsymbol{\beta}_v'\frac{\left(\mathbf{A}_v\mathbf{X}\right)'}{\sigma_v^2}\mathbf{y}_v + \boldsymbol{\beta}_v'\frac{1}{\tau^2}\boldsymbol{\beta}_v\right]\right\}\\ &= \exp\left\{-\frac{1}{2}\left[\boldsymbol{\beta}_v'\frac{\left(\mathbf{A}_v\mathbf{X}\right)'\left(\mathbf{A}_v\mathbf{X}\right) + \frac{\sigma_v^2}{\tau^2}\mathbf{I}}{\sigma_v^2}\boldsymbol{\beta}_v - 2\boldsymbol{\beta}_v'\frac{\left(\mathbf{A}_v\mathbf{X}\right)'}{\sigma_v^2}\mathbf{y}_v\right]\right\}\\ &= \exp\left\{-\frac{1}{2}\left[\boldsymbol{\beta}_v'\frac{\mathbf{X}'\mathbf{X} + \frac{\sigma_v^2}{\tau^2}\mathbf{I}}{\sigma_v^2}\boldsymbol{\beta}_v - 2\boldsymbol{\beta}_v'\frac{\left(\mathbf{A}_v\mathbf{X}\right)'}{\sigma_v^2}\mathbf{y}_v\right]\right\}.\end{aligned}$$

Therefore,

$$\pi(\boldsymbol{\beta}_v \mid \mathbf{y}_v, \lambda_v = 1, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2, \tau^2, \xi^2) = \mathcal{N}_2(\boldsymbol{\mu}_{\boldsymbol{\beta}_v}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}_v}),$$

where

$$\begin{aligned}
\boldsymbol{\mu}_{\boldsymbol{\beta}_v} &= \left(\mathbf{X}'\mathbf{X} + \frac{\sigma_v^2}{\tau^2}\mathbf{I}\right)^{-1} (\mathbf{A}_v\mathbf{X})'\,\mathbf{y}_v, \\
\boldsymbol{\Sigma}_{\boldsymbol{\beta}_v} &= \sigma_v^2 \left(\mathbf{X}'\mathbf{X} + \frac{\sigma_v^2}{\tau^2}\mathbf{I}\right)^{-1},
\end{aligned}$$

where $\mathbf{A}_v\mathbf{X}$ can be calculated as $\left[\mathbf{a}_v, \mathbf{x}^*_{(2)} \odot \mathbf{a}_v\right]$ for faster computation. When $\lambda_v = 0$, it's easy to show:

$$\pi(\boldsymbol{\beta}_{v,0} \mid \mathbf{y}_v, \lambda_v = 0, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2, \tau^2, \xi^2) = \mathcal{N}\left(\frac{(\mathbf{A}_v\mathbf{1}_T)'\,\mathbf{y}_v}{T + \frac{\sigma_v^2}{\tau^2}}, \frac{\sigma_v^2}{T + \frac{\sigma_v^2}{\tau^2}}\right),$$

and $\boldsymbol{\beta}_{v,1} = 0$ with probability 1, where $\mathbf{A}_v\mathbf{1}_T$ is just $\mathbf{a}_v$.

Sampling $\boldsymbol{\gamma}_v$

We apply Metropolis-Hastings algorithm to sample $\boldsymbol{\gamma}_v$. A random walk proposal,

$$\boldsymbol{\gamma}_v^* \mid \boldsymbol{\gamma}_v \sim \mathcal{N}_2\left(\boldsymbol{\Omega}_v\boldsymbol{\gamma}_v, \quad \boldsymbol{\Omega}_v'\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_v}\boldsymbol{\Omega}_v\right),$$

is used, where $\boldsymbol{\gamma}_v^*$ and $\boldsymbol{\gamma}_v$ are proposed parameter and current state, respectively, and $\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_v}$ is a tuning parameter. We use the current indicator of phase status, $\boldsymbol{\Omega}_v$, to secure it proposes $\boldsymbol{\gamma}_v^* = \begin{pmatrix} \gamma_{v,0}^* \neq 0 \\ \gamma_{v,1}^* = 0 \end{pmatrix}$ when the phase is inactive. Let $p_{\boldsymbol{\gamma}_v}(\cdot)$ be the proposal density, then the acceptance ratio is

$$\begin{aligned}
r_{\boldsymbol{\gamma}_v} &= \frac{\pi(\boldsymbol{\gamma}_v^* \mid \mathbf{y}_v, \boldsymbol{\beta}_v, \lambda_v, \omega_v, \sigma_v^2, \tau^2, \xi^2)\,p_{\boldsymbol{\gamma}_v}(\boldsymbol{\gamma}_v \mid \boldsymbol{\gamma}_v^*)}{\pi(\boldsymbol{\gamma}_v \mid \mathbf{y}_v, \boldsymbol{\beta}_v, \lambda_v, \omega_v, \sigma_v^2, \tau^2, \xi^2)\,p_{\boldsymbol{\gamma}_v}(\boldsymbol{\gamma}_v^* \mid \boldsymbol{\gamma}_v)} \\
&= \frac{p(\mathbf{y}_v \mid \boldsymbol{\beta}_v, \lambda_v, \boldsymbol{\gamma}_v^*, \omega_v, \sigma_v^2)\,p(\boldsymbol{\gamma}_v^* \mid \omega_v, \xi^2)}{p(\mathbf{y}_v \mid \boldsymbol{\beta}_v, \lambda_v, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2)\,p(\boldsymbol{\gamma}_v \mid \omega_v, \xi^2)},
\end{aligned}$$

where

$$\begin{aligned}
p(\mathbf{y}_v \mid \boldsymbol{\beta}_v, \lambda_v, \boldsymbol{\gamma}_v, \omega_v, \sigma_v^2) &\propto \exp\left\{-\frac{1}{2\sigma_v^2}\left(\mathbf{y}_v - \mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)'\left(\mathbf{y}_v - \mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)\right\} \\
&\propto \exp\left\{\frac{1}{\sigma_v^2}\left(\mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)'\mathbf{y}_v\right\}, \\
p(\boldsymbol{\gamma}_v \mid \omega_v, \xi^2) &\propto \exp\left\{-\frac{1}{2\xi^2}\boldsymbol{\gamma}_v'\boldsymbol{\Omega}_v\boldsymbol{\gamma}_v\right\}.
\end{aligned}$$

Simplify the ratio, when $\lambda_v = 0$,

$$r_{\boldsymbol{\gamma}_v} = \exp\left\{ \frac{1}{\sigma_v^2}\beta_{v,0}\left(\mathbf{a}_v \mid_{\boldsymbol{\gamma}_v=\boldsymbol{\gamma}_v^*} - \mathbf{a}_v \mid_{\boldsymbol{\gamma}_v=\boldsymbol{\gamma}_v}\right)' \mathbf{y}_v - \frac{1}{2\xi^2}\left(\gamma_{v,0}^{*,2} - \gamma_{v,0}^2\right)\right\}.$$

When $\lambda_v = 1$,

$$r_{\boldsymbol{\gamma}_v} = \exp\left\{ \frac{1}{\sigma_v^2}\left[\left(\beta_{v,0}\mathbf{1}_{2T} + \beta_{v,1}\mathbf{x}_{(2)}^*\right) \odot \left(\mathbf{a}_v \mid_{\boldsymbol{\gamma}_v=\boldsymbol{\gamma}_v^*} - \mathbf{a}_v \mid_{\boldsymbol{\gamma}_v=\boldsymbol{\gamma}_v}\right)\right]' \mathbf{y}_v - \frac{1}{2\xi^2}\left(\boldsymbol{\gamma}_v^{*'}\boldsymbol{\gamma}_v^* - \boldsymbol{\gamma}_v'\boldsymbol{\gamma}_v\right)\right\}.$$

We generate a dummy variable $d_{\boldsymbol{\gamma}_v} \sim \mathcal{U}(0,1)$, and if $d_{\boldsymbol{\gamma}_v} < r_{\boldsymbol{\gamma}_v}$, we update $\boldsymbol{\gamma}_v$ by $\boldsymbol{\gamma}_v^*$, otherwise remain $\boldsymbol{\gamma}_v$.

Full conditional distribution of σ_v^2

Assigning a Jeffreys prior, $p(\sigma_v^2) \propto 1/\sigma_v^2$, we have:

$$\pi(\sigma_v^2 \mid \mathbf{y}_v, \cdot) = \mathcal{IG}\left(\frac{2T}{2}, \quad \frac{1}{2}\left(\mathbf{y}_v - \mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)'\left(\mathbf{y}_v - \mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right)\right).$$

Again, to save computational time, $\mathbf{A}_v\mathbf{X}\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v$ can be calculated as $\beta_{v,0}\mathbf{a}_v$ when $\lambda_v = 0$, or $\left(\beta_{v,0}\mathbf{1}_{2T} + \beta_{v,1}\mathbf{x}_{(2)}^*\right) \odot \mathbf{a}_v$ when $\lambda_v = 1$.

Full conditional distributions of τ^2 and ξ^2

The full conditional distribution of τ^2 should be related to all voxels' β_0's and magnitude-active voxels' β_1's. Assigning a Jeffreys prior, $p(\tau^2) \propto 1/\tau^2$, we have:

$$\pi(\tau^2 \mid \mathbf{y}_v, \cdot) = \mathcal{IG}\left(\frac{1}{2}\sum_{v=1}^{V}\mathbf{1}_2'\boldsymbol{\Lambda}_v\mathbf{1}_2, \quad \frac{1}{2}\sum_{v=1}^{V}\boldsymbol{\beta}_v'\boldsymbol{\Lambda}_v\boldsymbol{\beta}_v\right).$$

Equivalently,

$$\pi(\tau^2 \mid \mathbf{y}_v, \cdot) = \mathcal{IG}\left(\frac{1}{2}\left(V + \sum_{v=1}^{V}\lambda_v\right), \quad \frac{1}{2}\sum_{v=1}^{V}\left[\beta_{v,0}^2 + \left(\lambda_v\beta_{v,1}\right)^2\right]\right).$$

Similarly, the full conditional distribution of ξ^2 should be related to all voxels' γ_0's and phase-active voxels' γ_1's, that is,

$$\pi(\xi^2 \mid \mathbf{y}_v, \cdot) = \mathcal{IG}\left(\frac{1}{2}\sum_{v=1}^{V} \mathbf{1}_2' \boldsymbol{\Omega}_v \mathbf{1}_2, \quad \frac{1}{2}\sum_{v=1}^{V} \boldsymbol{\gamma}_v' \boldsymbol{\Omega}_v \boldsymbol{\gamma}_v\right).$$

Equivalently,

$$\pi(\xi^2 \mid \mathbf{y}_v, \cdot) = \mathcal{IG}\left(\frac{1}{2}\left(V + \sum_{v=1}^{V} \omega_v\right), \quad \frac{1}{2}\sum_{v=1}^{V}\left[\gamma_{v,0}^2 + (\omega_v \gamma_{v,1})^2\right]\right).$$

Full conditional distributions of η_v, $\boldsymbol{\delta}$, and κ

They are the same as in Appendix A.2.

Bibliography

Adrian, D. W., Maitra, R., and Rowe, D. B. (2018). "Complex-valued time series modeling for improved activation detection in fMRI studies." *Annals of Applied Statistics*, 12(3): 1451–1478.

Albert, J. H. and Chib, S. (1993). "Bayesian analysis of binary and polychotomous response data." *Journal of the American Statistical Association*, 88(422): 669–679.

Amunts, K., Malikovic, A., Mohlberg, H., Schormann, T., and Zilles, K. (2000). "Brodmann's areas 17 and 18 brought into stereotaxic space-where and how variable?" *NeuroImage*, 11(1): 66–84.

Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S., and Hyde, J. S. (1992). "Time course EPI of human brain function during task activation." *Magnetic Resonance in Medicine*, 25(2): 390–397.

Bianciardi, M., Fukunaga, M., van Gelderen, P., Horovitz, S. G., de Zwart, J. A., Shmueli, K., and Duyn, J. H. (2009). "Sources of fMRI signal fluctuations in the human brain at rest: a 7T study." *Magnetic Resonance Imaging*, 27(8): 1019–1029.

Boynton, G. M., Engel, S. A., Glover, G. H., and Heeger, D. J. (1996). "Linear systems analysis of functional magnetic resonance imaging in human V1." *Journal of Neuroscience*, 16(13): 4207–4221.

Brown, R. W., Cheng, Y.-C. N., Haacke, E. M., Thompson, M. R., and Venkatesan, R. (2014). *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2nd edition.

Calhoun, V., Adalı, T., Pearlson, G., van Zijl, P., and Pekar, J. (2002). "Independent component analysis of fMRI data in the complex domain." *Magnetic Resonance in Medicine*, 48(1): 180–192.

Cox, R. W. (1996). "AFNI: software for analysis and visualization of functional magnetic resonance neuroimages." *Computers and Biomedical Research*, 29(3): 162–173.

Culham, J. C. and Valyear, K. F. (2006). "Human parietal cortex in action." *Current Opinion in Neurobiology*, 16(2): 205–212.

Dichter, G. S. (2012). "Functional magnetic resonance imaging of autism spectrum disorders." *Dialogues in Clinical Neuroscience*, 14(3): 319–351.

Epstein, R. and Kanwisher, N. (1998). "A cortical representation of the local visual environment." *Nature*, 392(6676): 598–601.

Feng, Z., Caprihan, A., c, K. B. B., and Calhoun, V. D. (2009). "Biophysical modeling of phase changes in BOLD fMRI." *NeuroImage*, 47(2): 540–548.

Flegal, J. M., Haran, M., and Jones, G. L. (2008). "Markov Chain Monte Carlo: can we trust the third significant figure?" *Statistical Science*, 23(2): 250–260.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., and Frackowiak, R. S. J. (1995). "Spatial registration and normalization of images." *Human Brain Mapping*, 3(3): 165–189.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. J. (1994). "Statistical parametric maps in functional imaging: A general linear approach." *Human Brain Mapping*, 2(4): 189–210.

Gelfand, A. E. and Smith, A. F. M. (1990). "Sampling-based approaches to calculating marginal densities." *Journal of the American Statistical Association*, 85(410): 398–409.

Geyer, S., Ledberg, A., Schleicher, A., Kinomura, S., Schormann, T., Bürgel, U., Klingberg, T., Larsson, J., Zilles, K., and Roland, P. E. (1996). "Two different areas within the primary motor cortex of man." *Nature*, 382(6594): 805–807.

Gudbjartsson, H. and Patz, S. (1995). "The Rician distribution of noisy MRI data." *Magnetic Resonance in Medicine*, 34(6): 910–914.

Hastings, W. K. (1970). "Monte Carlo sampling methods using Markov chains and their applications." *Biometrika*, 57(1): 97–109.

Hughes, J. and Haran, M. (2013). "Dimension reduction and alleviation of confounding for spatial generalized linear mixed models." *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 75(1): 139–159.

Jain, S. and Huth, A. G. (2018). "Incorporating context into language encoding models for fMRI." In *32nd Conference on Neural Information Processing Systems (NeurIPS 2018)*. Montréal, Canada.

Jäncke, L., Shah, N. J., Posse, S., Grosse-Ryuken, M., and Müller-Gärtner, H. W. (1998). "Intensity coding of auditory stimuli: an fMRI study." *Neuropsychologia*, 36(9): 875–883.

Krüger, G. and Glover, G. H. (2001). "Physiological noise in oxygenation-sensitive magnetic resonance imaging." *Magnetic Resonance in Medicine*, 46(4): 631–637.

Lee, J., Shahram, M., Schwartzman, A., and Pauly, J. M. (2007). "Complex data analysis in high-resolution SSFP fMRI." *Magnetic Resonance in Medicine*, 57(5): 905–917.

Lindquist, M. A. (2008). "The statistical analysis of fMRI data." *Statistical Science*, 23(4): 439–464.

Lindquist, M. A., Loh, J. M., Atlas, L. Y., and Wager, T. D. (2009). "Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling." *NeuroImage*, 45(1 Suppl): S187–S198.

Logothetis, N. K. (2008). "What we can do and what we cannot do with fMRI." *Nature*, 453(7197): 869–878.

Machulda, M., Ward, H., Borowski, B., Gunter, J., Cha, R., O'Brien, P., Petersen, R., Boeve, B., Knopman, D., Tang–Wai, D., Ivnik, R., Smith, G., Tangalos, E., and C.R. Jack, J. (2003). "Comparison of memory fMRI response among Normal, MCI, and Alzheimer's patients." *Neurology*, 61(4): 500–506.

Mandelkow, H., Zwart, J. A., and H.Duyn, J. (2016). "Linear discriminant analysis achieves high classification accuracy for the BOLD fMRI response to naturalistic movie stimuli." *Frontiers in Human Neuroscience*, 10: 128.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). "Equation of state calculations by fast computing machines." *Journal of Chemical Physics*, 21(6): 1087–1092.

Mikl, M., Mareček, R., Hluštík, P., Pavlicová, M., Drastich, A., Chlebus, P., Brázdil, M., and Krupa, P. (2008). “Effects of spatial smoothing on fMRI group inferences.” *Magnetic Resonance Imaging*, 26(4): 490–503.

Mitchell, T. J. and Beauchamp, J. J. (1988). “Bayesian variable selection in linear regression.” *Journal of the American Statistical Association*, 83(404): 1023–1032.

Musgrove, D. R., Hughes, J., and Eberly, L. E. (2016). “Fast, fully Bayesian spatiotemporal inference for fMRI data.” *Biostatistics*, 17(2): 291–303.

Petridou, N., Italiaander, M., van de Bank, B. L., Siero, J. C. W., Luijten, P. R., and Klomp, D. W. J. (2013). “Pushing the limits of high-resolution functional MRI using a simple high-density multi-element coil design.” *NMR in Biomedicine*, 26(1): 65–73.

Petridou, N., Plenz, D., Silva, A. C., Loew, M., Bodurka, J., and Bandettini, P. A. (2006). “Direct magnetic resonance detection of neuronal electrical activity.” *Proceedings of the National Academy of Sciences*, 103(43): 16015–16020.

Picinbono, B. (1996). “Second-order complex random vectors and normal distributions.” *IEEE Transactions on Signal Processing*, 44(10): 2637–2640.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
URL `https://www.R-project.org/`

Rao, S. M., Bandettini, P. A., Binder, J. R., Bobholz, J. A., Hammeke, T. A., Stein, E. A., and Hyde, J. S. (1996). “Relationship between finger movement rate and functional magnetic resonance signal change in human primary motor cortex.” *Journal of Cerebral Blood Flow and Metabolism*, 16(6): 1250–1254.

Reich, B. J., Hodges, J. S., and Zadnik, V. (2006). “Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models.” *Biometrics*, 62(4): 1197–1206.

Rice, S. O. (1944). “Mathematical analysis of random noise.” *The Bell System Technical Journal*, 23(3): 282–332.

Rombouts, S. A., Scheltens, P., Machielsen, W. C., Barkhof, F., Hoogenraad, F. G., Veltman, D. J., Valk, J., and Witter, M. P. (1999). “Parametric fMRI analysis of visual encoding in the human medial temporal lobe.” *Hippocampus*, 9(6): 637–643.

Rowe, D. B. (2005a). “Modeling both the magnitude and phase of complex-valued fMRI data.” *NeuroImage*, 25(4): 1310–1324.

— (2005b). “Parameter estimation in the magnitude-only and complex-valued fMRI data models.” *NeuroImage*, 25(4): 1124–1132.

— (2009). “Magnitude and phase signal detection in complex-valued fMRI data.” *Magnetic Resonance in Medicine*, 62(5): 1356–1360.

— (2019). *Handbook of Neuroimaging Data Analysis, Chapter 8*. London, United Kingdom: Chapman & Hall.

Rowe, D. B., Hahn, A. D., and Nencka, A. S. (2009). “Functional magnetic resonance imaging brain activation directly from k-space.” *Magnetic Resonance Imaging*, 27(10): 1370–1381.

Rowe, D. B. and Logan, B. R. (2004). “A complex way to compute fMRI activation.” *NeuroImage*, 23(3): 1078–1092.

— (2005). "Complex fMRI analysis with unrestricted phase is equivalent to a magnitude-only model." *NeuroImage*, 24(2): 603–606.

Rowe, D. B., Meller, C. P., and Hoffmann, R. G. (2007). "Characterizing phase-only fMRI data with an angular regression model." *Journal of Neuroscience Methods*, 161(2): 331–341.

Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields*. Boca Raton: Chapman & Hall/CRC.

Shen, G., Dwivedi, K., Majima, K., Horikawa, T., and Kamitani, Y. (2019). "End-to-end deep image reconstruction from human brain activity." *Frontiers in Computational Neuroscience*, 13: 21.

Smith, M. and Fahrmeir, L. (2007). "Spatial Bayesian variable selection with application to functional magnetic resonance imaging." *Journal of the American Statistical Association*, 102(478): 417–431.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and Joliot, M. (2002). "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain." *NeuroImage*, 15(1): 273–289.

Wang, Z., Childress, A. R., Wang, J., , and Detrea, J. A. (2007). "Support vector machine learning-based fMRI data group analysis." *NeuroImage*, 36(4): 1139–1151.

Wang, Z., Rowe, D. B., Li, X., and Brown, D. A. (2023). "Efficient Fully Bayesian Approach to Brain Activity Mapping with Complex-Valued fMRI Data." *arXiv preprint arXiv:2310.18536*. Available at `https://arxiv.org/abs/2310.18536`.

Watanabe, S. (2010). "Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory." *Journal of Machine Learning Research*, 11: 3571–3594.

Welvaert, M., Durnez, J., Moerkerke, B., Berdoolaege, G., and Rosseel, Y. (2011). "neuRosim: an R package for generating fMRI data." *Journal of Statistical Software*, 44(10): 1–18.

Wilder Penfield, E. B. (1937). "Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation." *Brain*, 60(4): 389–443.

Yu, C.-H., Prado, R., Ombao, H., and Rowe, D. B. (2018). "A Bayesian variable selection approach yields improved detection of brain activation from complex-valued fMRI." *Journal of the American Statistical Association*, 113(524): 1395–1410.

— (2023). "Bayesian spatiotemporal modeling on complex-valued fMRI signals via kernel convolutions." *Biometrics*, 79(2): 616–628.

Yu, M.-C., Lin, Q.-H., Kuang, L.-D., Gong, X.-F., Cong, F., and Calhound, V. D. (2015). "ICA of full complex-valued fMRI data using phase information of spatial maps." *Journal of Neuroscience Methods*, 249: 75–91.

Zhang, L., Guindani, M., and Vannucci, M. (2015). "Bayesian Models for fMRI Data Analysis." *Wiley Interdisciplinary Reviews: Computational Statistics*, 7(1): 21–41.