

**A BAYESIAN ADAPTIVE SMOOTHING AND THRESHOLDING
APPROACH FOR ACTIVATION DETECTION IN SINGLE-SUBJECT
fMRI**

by

Juan Esteban Flórez Coronel

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
in
SCIENTIFIC COMPUTING

UNIVERSITY OF PUERTO RICO
MAYAGÜEZ CAMPUS
2024

Approved by:

Alcibiades Bustillo-Zarate, Ph.D.
Member, Graduate Committee

Date

Roberto Rivera-Santiago, Ph.D.
Co-chair, Graduate Committee

Date

Israel Almodóvar-Rivera, Ph.D.
Chairperson, Graduate Committee

Date

FirstName I. LastName, Ph.D.
Representative of Graduate Studies

Date

Omar Colón-Reyes, Ph.D.
Department Chairperson

Date

ABSTRACT

Functional Magnetic Resonance Imaging (fMRI) is a widely used non-invasive medical procedure for studying brain function. Identifying activated regions of the brain is a common challenge in fMRI analysis. Low-signal and small data cases pose significant difficulties for activation detection. These scenarios arise when studying high-level cognitive tasks or single-subject experiments, respectively. In this study, we propose an innovative algorithm, entitled Bayesian Fast Adaptive Smoothing and Thresholding (BFAST), which utilizes smoothing and extreme value theory on probabilistic maps to find threshold values. The algorithm's performance was evaluated on experimental data that simulated a range of signal magnitudes. The results were promising, with an average similarity of 90% with respect to the expected output. Furthermore, the proposed procedure was applied to a study that aimed to identify the cerebral regions responsible for processing beliefs and questions as stimuli. Our findings suggest that the BFAST algorithm holds promise for detecting activated areas in the brain with high accuracy, particularly in cases involving low-signal and small data. Such advancements in fMRI analysis algorithms could lead to more accurate and precise studies of brain function, with significant implications for both clinical and research settings.

RESUMEN

La técnica de Imágenes por Resonancia Magnética Funcional (fMRI) es un procedimiento médico no invasivo ampliamente utilizado para estudiar la función cerebral. Identificar regiones activadas del cerebro es un desafío común en el análisis de fMRI. Los casos de baja señal y datos pequeños plantean desafíos importantes para la detección de activación. Estos escenarios surgen cuando se estudian tareas cognitivas de alto nivel o experimentos con un solo sujeto, respectivamente. En este estudio, proponemos un algoritmo innovador, titulado Umbralizado y Suavizado Adaptativo Bayesiano Rápido (BFAST), que utiliza la teoría de suavizado y valores extremos en mapas probabilísticos para encontrar valores de umbral. El rendimiento del algoritmo se evaluó a partir de datos experimentales que simularon una variedad de magnitudes de señal. Los resultados fueron prometedores, con una similitud promedio del 90% con respecto al resultado esperado. Además, el procedimiento propuesto se aplicó a un estudio que tenía como objetivo identificar las regiones cerebrales responsables de procesar creencias y preguntas como estímulos. Nuestros hallazgos sugieren que el algoritmo BFAST es prometedor para detectar regiones activadas en el cerebro con alta precisión, particularmente en casos que involucran baja señal y datos pequeños. Estos avances en los algoritmos de análisis de fMRI podrían conducir a estudios más exactos y precisos de la función cerebral, con importantes implicaciones tanto para entornos clínicos como de investigación.

ACKNOWLEDGMENTS

I want to express my deepest gratitude to my wonderful wife, Nichool; your patience, understanding, and countless sacrifices made it possible for me to focus on my research. Your faith in my abilities has been my rock. To my family, Edwin, Angy, and Salomé, for their unwavering support, love, and encouragement throughout this academic journey. Your belief in me has been my constant source of motivation. I am also profoundly grateful to my esteemed professor, Dr. Almodóvar, for his mentorship, guidance, and expertise. Your insights and dedication to my academic growth have been invaluable.

This thesis is a culmination of not just my efforts but the collective support and faith of my loved ones. I am thankful beyond words for your presence in my life.

For Sócrates and Toby

Contents

Copyright ©
Juan Esteban Flórez Coronel
2024

Chapter 1

INTRODUCTION

1.1 Motivation and Justification

Functional Magnetic Resonance Imaging (fMRI) is a non-invasive neuroimaging technique that measures brain activity by detecting changes in blood flow [?, ?, ?]. Several fMRI studies explore the brain regions involved in language processing, memory, and decision-making [?, ?, ?]. One of the primary objectives in fMRI is to identify the brain regions that are activated in response to specific stimuli or task [?, ?, ?]. This objective is challenging to approach in low-signal and small-data scenarios. These arise when studying high-level cognitive tasks or single-subject experiments, respectively. This process of identifying activation regions usually involves comparing the Hemodynamic Response Function (HRF) during the presentation of a stimulus or task to the HRF during a resting state or control condition [?, ?, ?].

HRF is a convolution of a discrete variable and some continuous function. This constant function mainly relates to the Blood Oxygenation Level-Dependent (BOLD) response. To compare the BOLD, researchers have used time-series analysis, statistical parametric mapping, multivariate pattern classification, Bayesian modeling, among

other methods [?, ?, ?, ?]. These methods are helpful in different situations, such as analyzing data points over time, processing spatially distributed processes, combining spatial and temporal patterns, and using probabilistic predictions. To some extent, these situations are partially present in some aspects of this research; hence, the methods will be helpful.

Researchers also used methods based in Adaptive Smoothing and Thresholding (AST) for fMRI studies. In their studies, the problems addressed went from finding the extent and shape of the activation region to the identification of the more accurate smoothing technique and procedure [?, ?, ?, ?]. The advantage of the AST method as it has been used is the ability to estimate thresholds between contrast-based maps and reduce noise inherent to the fMRI experiments. This approach yields good results in accurately identifying activated regions in fMRI experiments. However, exploring different methods, such as a Bayesian approach, where probability maps are used instead, can result in more precisely identifying such activated regions.

This study proposes a Bayesian approach using AST methods for the activation detection problem in single-subject fMRI. As opposed to previous research where AST for fMRI is used, in a Bayesian approach, the smoothing procedure will occur in the probability maps, resulting in a more understandable interpretation when finding activated regions. Although results might not be improved compared to a frequentist approach, it is relevant to explore the possible benefits of this kind of technique.

1.2 Objectives

- Perform Bayesian time-series analysis to obtain a posterior probability map of an fMRI image for a single-subject situation.
- Develop an AST method that inputs the probability posterior map and finds the

possible activated voxels.

- Study the proposed algorithm in different simulation frameworks. Study the results in terms of similarity, rate of false positives, and percent of activation.
- Finally, apply the algorithm to a real dataset.

1.3 Chapter Summary

This thesis consists of 7 chapters, which are briefly summarized below:

Chapter 1: Introduction. This chapter introduces the research project and provides an overview of its objectives, significance, and scope. Topics such as fMRI, BOLD, and AST are briefly explained.

Chapter 2: Literature Review. Chapter 2 presents a comprehensive review of the relevant literature on the study. In this chapter, fMRI studies, alongside statistical theory and models, are deeply discussed. The chapter highlights the existing gaps and areas where the current study adds value.

Chapter 3: Methodology. In Chapter 3, the methodology used in our work is described. The chapter details the design and development of the models and algorithms proposed. The methodology is clearly described, ensuring the study's replicability.

Chapter 4: Experimental Simulation. Chapter 4 defines the creation and analysis of simulated data where the ground truth is known. The structure proposed for the simulation enables an evaluation of the accuracy of our methods. By using simulated data, we aim to validate and understand the capabilities of our approach.

Chapter 5: Performance Evaluation Results. In Chapter 5, the proposed algorithm's results applied to experimental simulation data are presented. The chapter summarizes the evaluation of our work in different scenarios of signal magnitudes. The results presented in this chapter ensure that the results from the next chapter are accurate.

Chapter 6: BFAST in a Real Dataset This chapter presents the results of the proposed algorithm in real datasets of fMRI experiments. The objective of this chapter is to show an example of the usage of our work. In the example experiment, the processing of beliefs and questions is taken as stimuli.

Chapter 7: Conclusion and Future Work. In Chapter 7, the study concludes by summarizing the main findings and implications. This chapter also reflects on the study's limitations and identifies areas for improvement. The chapter serves as a closing remark, providing a comprehensive research summary and emphasizing its contributions to the field.

Chapter 2

Literature Review

2.1 Introduction to fMRI

Magnetic Resonance Imaging (MRI) is a powerful medical imaging technique that has revolutionized the field of diagnostic medicine [?]. At its core, MRI relies on the interaction of protons within the human body with strong magnetic fields and radiofrequency pulses [?]. These magnetic fields, often generated by superconducting magnets, align the protons within the body's tissues [?]. Subsequent radiofrequency pulses perturb this alignment, causing the protons to emit radiofrequency signals as they return to their original alignment. By detecting these signals and their variations, MRI scanners create high-resolution anatomical images that provide detailed insights into the body's internal structures [?]. This non-invasive and versatile imaging modality has become indispensable in clinical diagnosis, research, and medical practice, offering a wealth of information for assessing various medical conditions.

As traditional MRI focuses mainly on the generation of static anatomic images of the internal structures of the body [?], fMRI brings a new advantage as it captures the dynamic activities of the body part studied [?, ?, ?]. The critical difference is that

magnetic resonance is based mainly on the interaction of protons with magnetic fields to produce detailed anatomic images in the data acquisition process. At the same time, the fMRI takes advantage of the BOLD contrast to indirectly measure neural activity by detecting changes in the oxygenation level of the blood [?]. This fundamental change of emphasis allows fMRI to study the visualization and mapping of brain regions activated during specific cognitive tasks, making it a very used tool in cognitive neuroscience and neuropsychology [?, ?].

In the domain of fMRI, BOLD contrast is within the most important concepts to be studied [?]. The essence of BOLD contrast relies on the observation that neural activity generates changes in local blood oxygenation levels [?]. As brain regions become more active, they demand increased oxygen and glucose to sustain their functions [?]. In response, blood flow to these regions is expected to be altered to meet the demand. Importantly, hemoglobin, the oxygen-carrying molecule in blood, behaves differently when oxygenated and deoxygenated, affecting its magnetic properties [?, ?, ?]. When oxygenation levels of the blood change, it generates fluctuations in its magnetic properties; this process is all captured by fMRI experiments [?].

As expected, this long data reading process generates a significant amount of noise because of all the factors that are expected to work correctly during the measurements. In addition to that, it is known that high-level cognitive tasks produce low-signal scenarios in fMRI experiments [?]. To quantify the amount of noise concerning the signal studied, researchers use metrics such as the Signal-to-Noise Ratio (SNR) and the Contrast-to-Noise Ratio (CNR) [?]. The SNR quantifies the ratio of the strength of the signal arising from brain activity to the background noise inherent in the imaging process. Higher SNR values indicate a more robust and detectable signal. Similarly, CNR assesses the contrast between activated and non-activated brain regions by comparing the difference in signal intensity between them to the noise level. A higher CNR signifies

a stronger and more discernible activation signal relative to background noise.

2.2 Analysis of fMRI Data - Time Series, Activation Detection and Final Image

Voxels, short for volumetric pixels, are fundamental building blocks in fMRI analysis [?]. They represent 3-Dimensional (3D) units within the image and play a crucial role in discretizing the space studied. Each voxel corresponds to a tiny, well-defined volume in the brain, and within this volume, fMRI data, particularly BOLD signal measurements, are collected over time [?]. These measurements over time can be compiled into a time series and, more specifically, with a linear relation. The General Linear Model (GLM) for time series analysis is a fundamental technique in fMRI data processing because it captures temporal dynamics of neural activity [?, ?].

Detection of neural activity in the Region of Interest (ROI) is a crucial field of study in fMRI research as it enables scientists to identify brain regions that exhibit significant changes in activity in response to specific stimuli or tasks [?]. The identification of the ROI in fMRI is called image masking [?]. The ROI can correspond to anatomically defined brain structures, functionally significant areas, or areas of interest for a particular study [?]. Image masking is employed to improve the precision and efficiency of analyses, as it allows researchers to isolate and concentrate on the neural activity occurring within predefined brain regions [?]. By delineating the ROI, image masking effectively filters out irrelevant data, reducing noise and enhancing the sensitivity of statistical analyses. One method to apply the image masking, as implemented in NiLearn [?], is based on a heuristic proposed by T.Nichols [?]: find the least dense point of the histogram, between a lower cutoff and an upper cutoff of the total image histogram.

Within the area of neural activity, some researchers use frequentist approaches to detect activation in fMRI studies. These approaches can be described as statistical methods that adopt a null hypothesis tested using p-values to determine whether a brain region is significantly activated by a particular stimulus or condition [?, ?]. These methods are widely used in fMRI research [?, ?, ?, ?]. Still, they have been criticized for their limitations, such as their problems addressing hemodynamic variability and the spatio-temporal autocorrelations in fMRI [?].

An essential tool to be discussed that is relevant in generating low-noise activation maps is image smoothing [?]. Image smoothing is a crucial step in activation map analysis because it helps to reduce noise and improve the localization of activated brain regions [?, ?, ?]. By smoothing the probability maps, researchers can more easily identify the brain regions most strongly activated by a particular stimulus or condition [?]. Adaptive smoothing has been a common technique in activation detection in fMRI, and researchers have always complimented this technique with frequentist approaches [?, ?, ?]. These methods yield precise results. However, there is a gap in the literature regarding using adaptive smoothing with Bayesian approaches.

After obtaining the final activation map, researchers must be able to compare methods and test their findings' reliability. Hence, tools like the Jaccard Index (JI) were introduced to the area of fMRI. The JI was initially introduced by Paul Jaccard in 1901 [?]; later, researchers found application in fMRI analysis, as discussed in [?]. The JI measures the similarity between two sets by calculating the intersection over the union of their elements. In fMRI, it assesses the overlap and consistency of brain activation patterns across different subjects, conditions, or studies. A higher JI indicates a more remarkable similarity between activation maps.

2.3 Bayesian Analysis

Bayesian analysis is essential in data analysis and statistical reasoning [?, ?]. It is a probabilistic framework that quantifies uncertainty and makes inferences from data [?]. Unlike traditional frequentist statistics, which treat model parameters as fixed and unknown values, Bayesians treat these parameters as random variables, encapsulating our uncertainty about their values with probability distributions [?]. In the Bayesian analysis, prior beliefs about parameters are combined with observed data through Bayes' Rule to construct the posterior distribution. The prior distribution is the key in Bayesian approaches as it represents the previous knowledge or assumptions about the random variables in question [?, ?].

The selection of a prior distribution is relevant in Bayesian modeling, as it profoundly influences the posterior distribution [?]. When choosing a prior distribution, researchers must balance incorporating relevant domain expertise and ensuring that the prior does not dominate the outcome of the posterior. This requires careful consideration of the prior's shape, scale, and informativeness [?]. Various methods, such as non-informative or weakly informative priors, hierarchical modeling, and empirical Bayes techniques, offer strategies for selecting appropriate priors based on the available information and the specific context of the analysis [?]. A good choice of prior distributions incorporates valuable previous knowledge while preserving the capacity of data to update and refine the result, thus yielding more robust and insightful posterior distributions [?].

2.4 Relevant Distributions

Given the nature of random variables that represent probability values, distributions whose range lies between 0 and 1 are studied. The Truncated Normal (TN) is a relevant probability distribution with applications in modeling extreme values that fall within

a specific range [?]. This distribution is characterized by the constraint that its values lie within a defined interval, effectively truncating the tails of the standard normal distribution.

In [?], the concept of the Domain of Maximal Attraction (DMA) is presented as a fundamental idea of the Extreme Value Theory (EVT). The DMA characterizes the asymptotic behavior of extreme value distributions as it represents a specific class of distributions that exhibit remarkable convergence properties when dealing with extreme values. It is the set of distributions for which the maxima of independent and identically distributed random variables converge to one of the three extreme value distributions: the Gumbel, Fréchet, or Weibull distribution, depending on the characteristics of the underlying distribution [?].

Chapter 3

Methodology

3.1 Time-Series Model

A time series analysis for each voxel of the image will allow temporal fluctuations in BOLD signals to be captured. By tracking changes in BOLD signals over time, the objective is to investigate dynamic patterns of brain activity, allowing for the identification of regions that respond to specific stimuli or tasks. Analyzing time series data at the voxel level provides valuable information into the temporal dynamics of neural processes, enabling a deep understanding of the brain's architecture.

Let \mathbf{y}_i be a vector of the response variable of the i th voxel, and \mathbf{X} be the design matrix of the study containing the expected BOLD and orthogonal drift components to take account of the low-frequency effects during the reading. With $\boldsymbol{\beta}_i$ being the vector of coefficients associated with the stimulus, we will have $\mathbf{y}_i \sim N(\mathbf{X}\boldsymbol{\beta}_i, \Sigma)$. Note that Σ can have a Auto Regressive Moving Average (ARMA) structure. However, if we let $\Sigma = \sigma^2 \mathbf{I}$, the independent model is obtained:

$$\mathbf{y}_i | \boldsymbol{\beta}_i, \sigma, \mathbf{X} \sim N(\mathbf{X}\boldsymbol{\beta}_i, \sigma^2 \mathbf{I}). \quad (3.1)$$

From here, the procedure is presented in [?] explains that to obtain the posterior distribution of the coefficients associated with the stimulus, β_i . We will use a noninformative prior distribution that is uniform on $(\beta_i, \log \sigma)$:

$$\pi(\beta_i, \sigma) \propto \sigma^{-2}. \quad (3.2)$$

Now, let us denote the sampling distribution by $f(\mathbf{y}_i, \sigma | \beta_i)$, then the joint density of \mathbf{y}_i , σ and β_i is given by:

$$f(\mathbf{y}_i, \beta_i, \sigma) = f(\mathbf{y}_i, \sigma | \beta_i) \pi(\beta_i) \quad (3.3)$$

The marginal distribution of \mathbf{y}_i is then given by:

$$m(\mathbf{y}_i) = \iint f(\mathbf{y}_i, \sigma | \beta_i) \pi(\beta_i) d\beta_i d\sigma \quad (3.4)$$

To obtain the posterior distribution of β_i , we calculate the conditional distribution using the Bayes' Rule:

$$\pi(\beta_i | \sigma, \mathbf{y}_i) = \frac{f(\mathbf{y}_i, \beta_i, \sigma)}{m(\mathbf{y}_i)} = \frac{f(\mathbf{y}_i | \sigma, \beta_i) \pi(\beta_i, \sigma)}{\iint f(\mathbf{y}_i, \sigma | \beta_i) \pi(\beta_i) d\beta_i d\sigma} \quad (3.5)$$

Given these results and using the ordinary least squares solution to a linear model $\hat{\beta}_i = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}_i$. The conditional posterior of β_i , given σ is then:

$$\pi(\beta_i | \sigma, \mathbf{y}_i) \sim N\left(\hat{\beta}_i, (\mathbf{X}^T \mathbf{X})^{-1} \sigma^2\right). \quad (3.6)$$

An estimation of σ^2 is then needed. Note that the marginal posterior of σ^2 can be obtained by factoring the joint posterior distribution of β_i and σ^2 as:

$$\pi(\sigma^2 | \mathbf{y}_i) = \frac{\pi(\beta_i, \sigma^2 | \mathbf{y}_i)}{\pi(\beta_i | \sigma^2, \mathbf{y}_i)}. \quad (3.7)$$

This results in:

$$\pi(\sigma^2 | \mathbf{y}_i) \sim Inv - \chi^2(n - k, s^2). \quad (3.8)$$

Where n is the sample size and k is the number of parameters in the data, and:

$$s^2 = \frac{1}{n - k} (\mathbf{y}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_i)^T (\mathbf{y}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_i) \quad (3.9)$$

Finally, for each voxel, i , in the region of interest of our study, calculate the posterior probability that the coefficient associated with the stimulus, t , is not zero, which is roughly estimated using $P(\boldsymbol{\beta}_{i,t} > 0 | \mathbf{y}_i, \mathbf{X})$.

Let $\mathbb{P} = \{P(\boldsymbol{\beta}_{i,t} > 0)\}_{i=[1,v]}$ represent a Posterior Probability Map (PPM), where v is the number of voxels in a fMRI experiment. Our goal now is to calculate a threshold and find activated regions using \mathbb{P} , for which we propose the Bayesian Fast Adaptive Smoothing and Thresholding (BFAST) algorithm.

3.2 BFAST Algorithm

3.2.1 TN Distribution

All the entries of \mathbb{P} are probabilities ranging between 0 and 1. Hence, we will study \mathbb{P} as a TN Distribution in the interval $[0, 1]$, i.e.:

$$\mathbb{P} \sim TN(\mu_{\mathbb{P}}, \sigma_{\mathbb{P}}^2, 0, 1) \quad (3.10)$$

The mean $\mu_{\mathbb{P}}$ and the variance $\sigma_{\mathbb{P}}^2$ can be regarded as a perturbation of the mean $\bar{\mu}$ and variance $\bar{\sigma}^2$ of the parent normal distribution, respectively. Its values can be determined by referencing the normal Probability Density Function (PDF) ϕ and Cumulative

Distribution Function (CDF) Φ . As presented in [?]:

With:

$$\alpha = \frac{-\bar{\mu}}{\bar{\sigma}}; \quad \beta = \frac{1 - \bar{\mu}}{\bar{\sigma}} \quad (3.11)$$

We have:

$$\mu_{\mathbb{P}} = \bar{\mu} - \bar{\sigma} \cdot \frac{\phi(0, 1; \beta) - \phi(0, 1; \alpha)}{\Phi(0, 1; \beta) - \Phi(0, 1; \alpha)} \quad (3.12)$$

And:

$$\sigma_{\mathbb{P}}^2 = \bar{\sigma}^2 \cdot \left(1 - \frac{\beta\phi(0, 1; \beta) - \alpha\phi(0, 1; \alpha)}{\Phi(0, 1; \beta) - \Phi(0, 1; \alpha)} - \left(\frac{\phi(0, 1; \beta) - \phi(0, 1; \alpha)}{\Phi(0, 1; \beta) - \Phi(0, 1; \alpha)} \right)^2 \right) \quad (3.13)$$

3.2.2 EVT

To find the threshold probability value that separates active and inactive voxels, an extreme value distribution for the TN distribution is used [?, ?]. From Theorem 10.5.2 of [?], it is deduced that a TN distribution is in the domain of maximal attraction of a Gumbel distribution (G), See Appendix ??.

Hence, we can say that $\exists a_v > 0$ and b_v and a nondegenerate CDF G such that $TN^v(a_v x + b_v) \rightarrow G(x)$ at all continuity points of G . We can choose:

$$a_v = [v\psi(b_v)]^{-1}; \quad b_v = \Psi^{-1}(1 - 1/v). \quad (3.14)$$

Typically, ψ and Ψ are used as the PDF and CDF of the TN, respectively.

3.2.3 Gaussian Kernel Smoothing

The BFAST algorithm also uses Gaussian Smoothing, a spatial filtering technique commonly used in image and signal processing, to enhance images by reducing noise and preserving essential features [?]. This method applies a Gaussian kernel, characterized by its bell-shaped curve, to each pixel in an image or PPM in this case. The kernel serves as a weighted averaging filter where the central pixel or element has the highest weight while the surrounding pixels or elements contribute with decreasing weights as their distance from the center increases. The mathematical basis of Gaussian Smoothing lies in convolution, where the kernel is convolved with the input data, blurring the image or signal. The smoothing degree depends on the Gaussian kernel's standard deviation, σ_s . A larger standard deviation results in more significant smoothing, and a more minor standard deviation results in less smoothing.

3.2.4 Definition of the JI

A version of the JI is also used in the BFAST algorithm. We define the JI, $J(\mathbf{A}, \mathbf{B})$, as a similarity index between images A and B , and its computed as a quotient:

$$J(\mathbf{A}, \mathbf{B}) = \frac{|\mathbf{A} \cap \mathbf{B}|}{|\mathbf{A} \cup \mathbf{B}|} \quad (3.15)$$

3.2.5 BFAST Pseudocode

The proposed algorithm called BFAST can be described as follows:

1. ***Initial Setup.*** Start with a PPM $\mathbb{P}^{(0)} = \mathbb{P}$. Assume that all voxels are inactive, i.e., $\zeta_i \equiv 0 \forall i$, where ζ_i is 1 when voxel i is activated and 0 otherwise. Set $\zeta_i^{(0)} \equiv \zeta_i$ and $v_0 = v$, where v_k denotes the number of voxels for which $\zeta_i^{(k)} = 0$.
2. ***Iterative Steps,*** For $k = 1, 2, \dots$, iterate as follows:

(a) *Smoothing.* Smooth $\mathbb{P}^{(k-1)}$ using a Gaussian Kernel to obtain $\mathbb{P}^{(k)}$. Let σ_s increase with k .

(b) *Thresholding.* This consists of three steps:

- i. Calculate $\mu_{\mathbb{P}^{(k-1)}}$ and $\sigma_{\mathbb{P}^{(k-1)}}^2$ to estimate $\mathbb{P}^{(k-1)}$ as a TN. Use Equations ?? and ?? with $\bar{\mu}$ and $\bar{\sigma}^2$ being the mean and variance of $\mathbb{P}^{(k-1)}$.
- ii. Calculate a_v and b_v . Use Equations ??, with ψ and Ψ as the PDF and CDF of $TN(\mu_{\mathbb{P}}, \sigma_{\mathbb{P}}^2, 0, 1)$, respectively.
- iii. Calculate the probability threshold, $\eta = a_v \iota_{0.01} + b_v$, with $\iota_{0.01}$ be the upper-tail 0.01-value of the standard Gumbel Distribution.

(c) *Activation:* Set $\zeta_i^{(k)} = 1$ if $\zeta_i^{(k-1)} = 0$ and the value of the i th voxel of $\mathbb{P}^{(k)}$ is greater than η . Finally, calculate $v_k = \sum_{i=1}^v \zeta_i^{(k)}$.

3. ***Termination.***

- (a) Declare no activation and terminate if $\zeta^{(1)} \equiv 0$.
- (b) If $J(\zeta^{(k)}, \zeta^{(k-1)}) \geq J(\zeta^{(k+1)}, \zeta^{(k)})$, the algorithm terminates and the final activation map is $\zeta^{(k)}$.
- (c) The maximum number of iterations is set to $k = 10$.

Chapter 4

Experimental Simulation

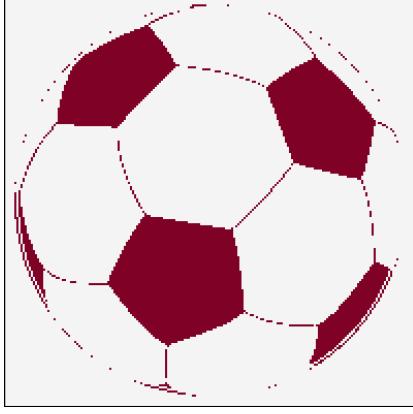
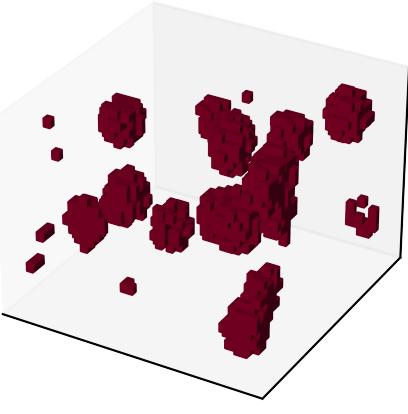
The proposed BFAST algorithm will be studied using an experimental framework mimicking a single-subject fMRI experiment. The datasets from this kind of experiment consist of the measured BOLD signal for each voxel throughout the experiment. Also, it contains a structured log of the events during the experiment, which includes, among other things, stimulus times and duration.

4.1 Generation of the Simulated Framework

4.1.1 Selection of True Maps

The first step to creating a simulated framework is to identify true maps of the activated regions that will be distorted for the simulation. The BFAST algorithm will then be used to reconstruct the activation maps, and the results will be compared to the original true maps to measure the algorithm's accuracy. In these true maps, each voxel will be assigned a value of 0 if it is not activated and 1 if it is activated. To address different spatial structures, 2-Dimensional (2D) and 3D true maps are considered. See Table ?? for more details on the true maps.

Table 4.1: Details of True Maps Considered

Name	2D	3D
Dimensions	200×200	$40 \times 40 \times 25$
Voxels	40000	40000
A%	19.9375	3.9525
Map		
Source	Derived from File in Open Clip Art Library (Public Domain)	See Appendix ??
Description	In both maps, dark voxels are active and light voxels are inactive.	

4.1.2 Creation of Design Matrix

The next step of the framework-building process is the creation of a design matrix of two columns, \mathbf{X} . The second column corresponds to the constant regressor and the first column contains the Glover HRF [?], given specific event descriptions of an fMRI experiment. For the simulated framework, a single type of stimulus will be considered as the event of the fMRI experiment, however, this event will occur at different times. The scan times and durations were arbitrarily selected as follows:

Table 4.2: Event Description of Simulated fMRI Experiment.

Parameter	Value
Number of Scans	100
Time Between Scans	2 seconds
Number of Stimulus	4
Duration of Each Stimulus	10 seconds
Time Between Stimulus	18 - 25 seconds

Resulting then in the HRF shown in Figure ??.

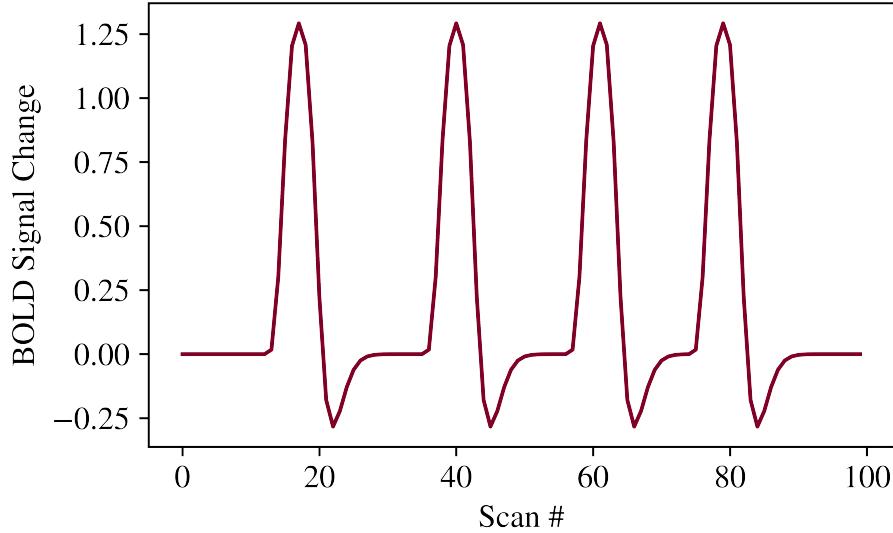


Figure 4.1: Glover HRF Given a Stimulus Described by Table ??.

4.1.3 Computation of BOLD

The final step in building the simulated framework is to compute the BOLD response, \mathbf{y}_i , for each voxel i . This process is divided into two steps: first, the computation of the BOLD response without noise, and then, using an *ARMA* Model to generate noise in the signal [?], i.e.:

$$\mathbf{y}_i = \hat{\mathbf{y}}_i + \boldsymbol{\epsilon}_{\hat{\mathbf{y}}_i} \quad (4.1)$$

To compute the BOLD response without noise ($\hat{\mathbf{y}}_i$) we will set values for the parameter β_i depending on the voxel activation status, ζ_i , on the true map as shown in Table ???. The response is then computed as seen in Equation ??.

$$\hat{\mathbf{y}}_i = \mathbf{X}\boldsymbol{\beta}_i \quad (4.2)$$

Table 4.3: Parameter Selection Based on Activation Status

Activation Status	Parameter Values
$\zeta_i = 0$	$\beta_i = (0, 100)^T$
$\zeta_i = 1$	$\beta_i = (75, 100)^T$

On the other side, the noise ($\epsilon_{\hat{y}_i}$) is a vector of mean $\mu_{ARMA} = 0$ and variance σ_{ARMA}^2 with a baseline structure equivalent to $ARMA_{\epsilon_{\hat{y}_i}}(\{p_1, p_2, \dots\}, \{q_1, q_2, \dots\})$. Note that $P = |\{p_1, p_2, \dots\}|$ and $Q = |\{q_1, q_2, \dots\}|$ are related to the order of the corresponding *ARMA* model. Also, p_a and q_b represent the coefficients of such models. Values of P and Q in the range [0, 3] were chosen to study the model under different noise scenarios. The values of σ_{ARMA} , p_a , and q_b were chosen arbitrarily as parameters, see Table ??.

Table 4.4: Parameter Selection Related to $\epsilon_{\hat{y}_i}$

Order	Values
All	$\sigma_{ARMA} = 25$
$P = 0$	Model without Auto-Regression
$P = 1$	$\{p_1\} = \{0.5\}$
$P = 2$	$\{p_1, p_2\} = \{0.5, 0.3\}$
$P = 3$	$\{p_1, p_2, p_3\} = \{0.5, 0.3, 0.1\}$
$Q = 0$	Model without Moving-Average
$Q = 1$	$\{q_1\} = \{0.5\}$
$Q = 2$	$\{q_1, q_2\} = \{0.5, 0.3\}$
$Q = 3$	$\{q_1, q_2, q_3\} = \{0.5, 0.3, 0.1\}$

Chapter 5

Performance Evaluation Results

5.1 Noise in Simulated Experiments

To quantify the amount of noise in the images after the computation of the BOLD, we use the CNR and SNR [?]. Note that the precision of the calculations depends on the noise present in the images. Therefore, the performance is expected to decrease as the amount of noise increases and the signals from active and inactive voxels are more difficult to differentiate.

For each of the 2 true maps considered and each of the 16 order combinations (P, Q) for the *ARMA* model to generate noise, 50 different BOLD responses were generated, resulting in 1600 simulated fMRI experiments in total. A voxel-wise computation of the CNR and SNR was made for all of them. See Figures ?? for their numerical distributions. Additionally, Figures ?? and ?? present the spatial distributions of the SNR and CNR values in the 2D maps, respectively. Although it is not presented, the spatial distributions of the SNR and CNR values in the 3D maps are expected to observe a similar behavior.

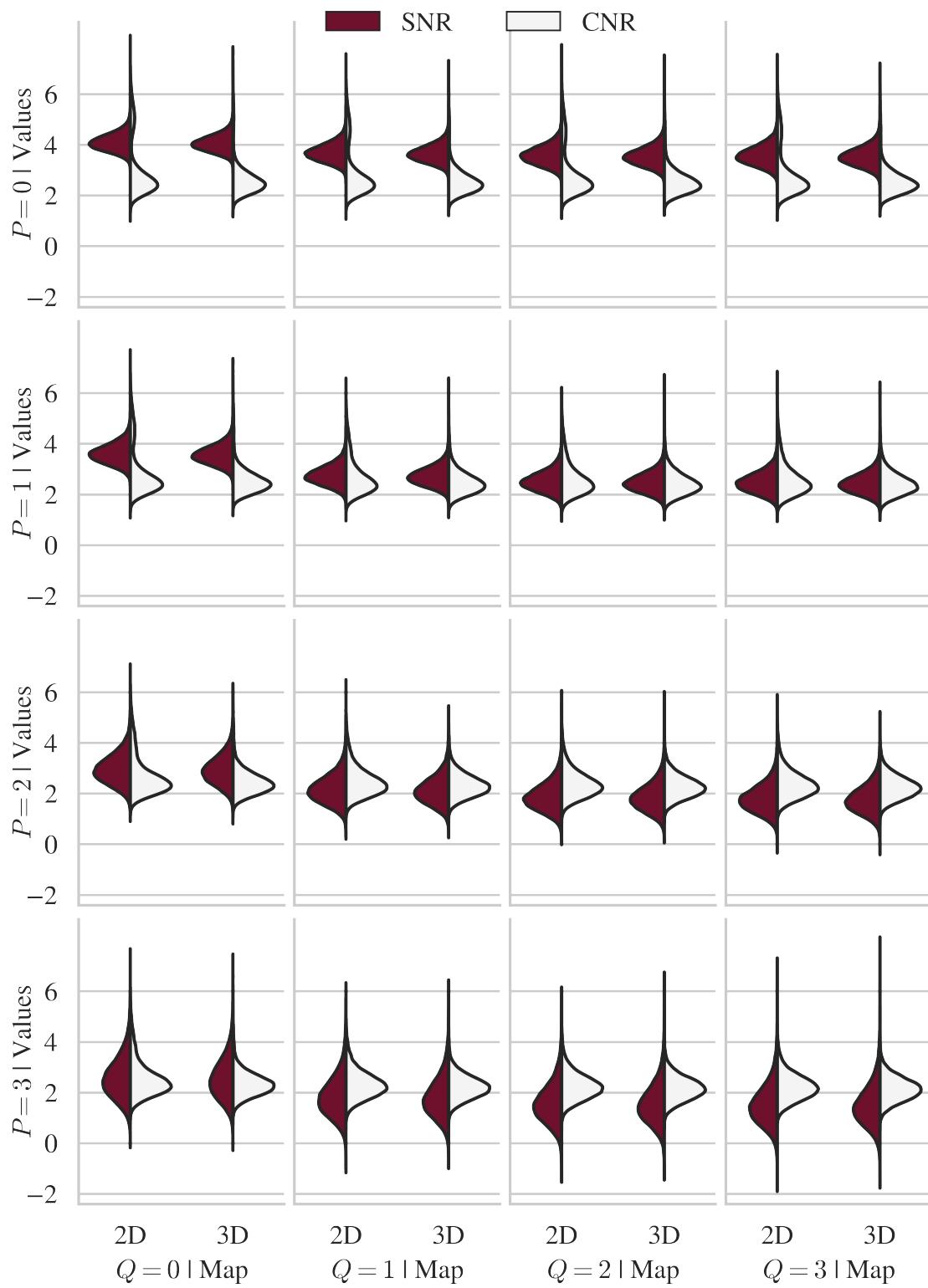


Figure 5.1: Numerical Distribution of the Voxel-Wise SNR and CNR Values

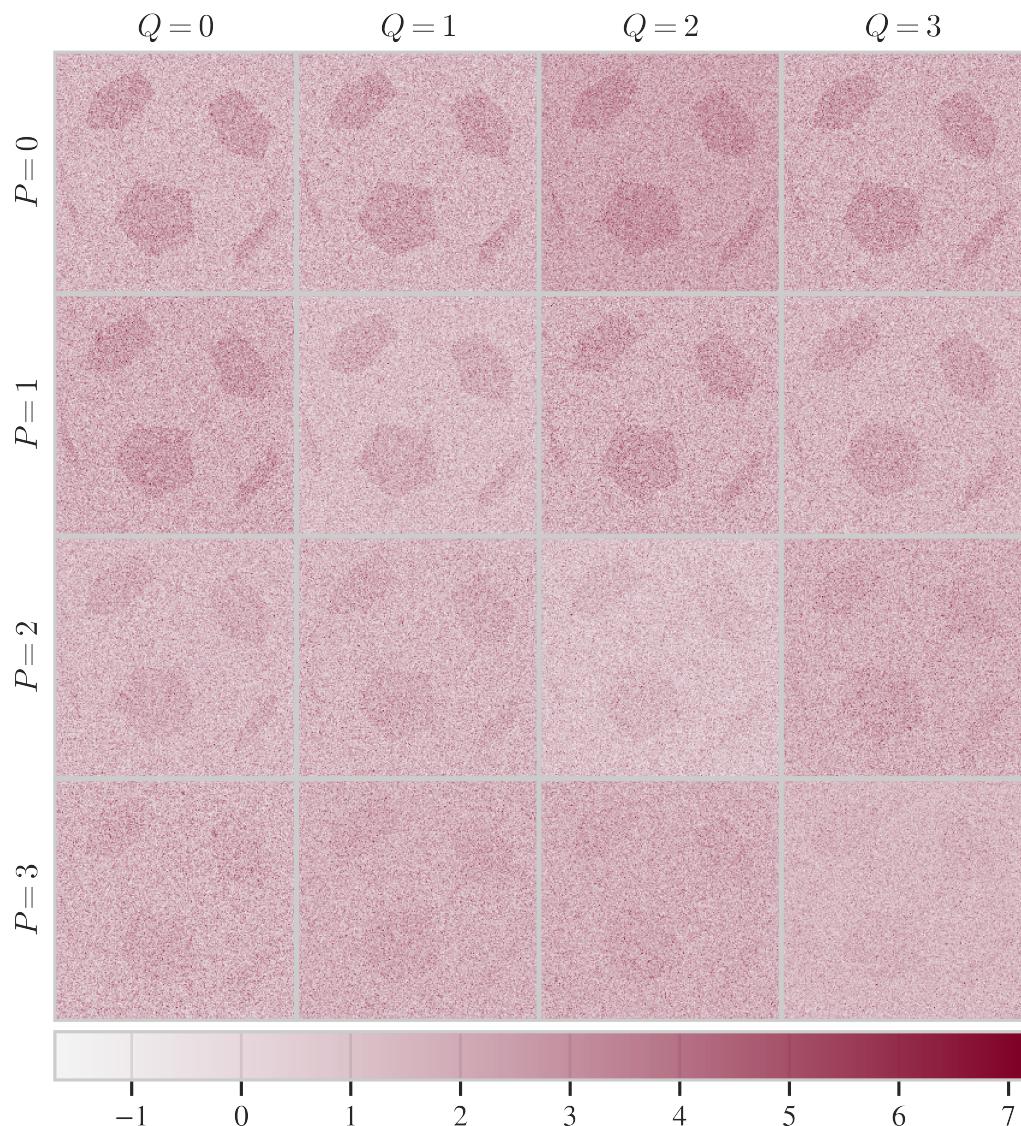


Figure 5.2: Spatial Distribution of the SNR Values in 2D Map

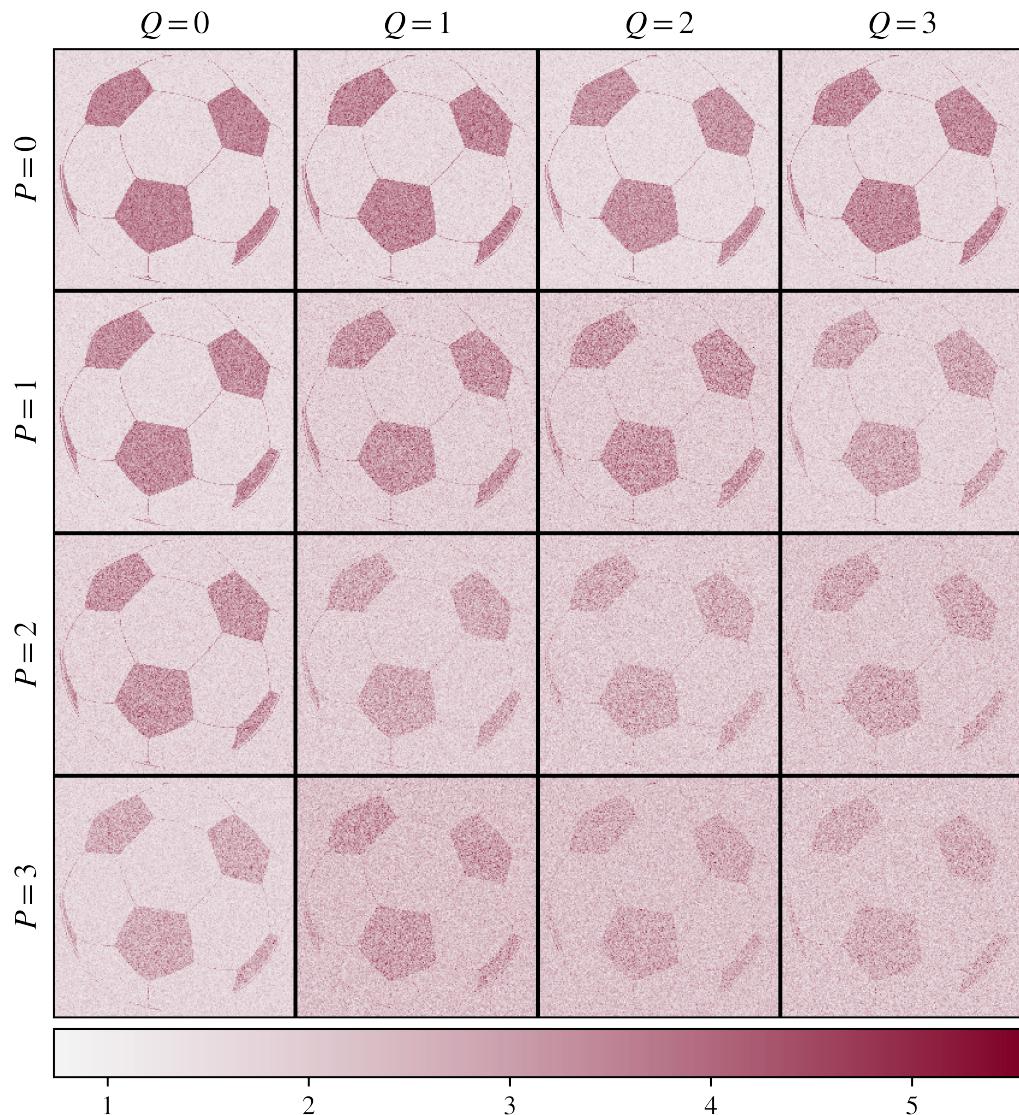


Figure 5.3: Spatial Distribution of the CNR Values in 2D Map

Note from Figure ?? that the SNR and CNR values decrease as the values of P or Q increase, however, we can see that a change in the value of P has a greater impact on the change of the SNR values. Additionally, we can see that the behavior of the SNR values does not change within the two maps. In contrast, the behavior of the CNR values change in the two maps. Note that for the 2D map, the distribution of the CNR for low values of P and Q can be interpreted as bimodal. This is because the number of active voxels in the 2D map is higher than in the 3D map, and active voxels have more contrast than inactive voxels. Finally, note that in both maps, for higher values of P and Q , the SNR appears to be more distributed while the CNR appears to be less distributed. Finally, note from Figures ?? and ?? that the SNR and CNR values in our simulated fMRI experiment range from -2 to 7 and from 0 to 6, respectively. In both the SNR and CNR, the active voxels are more difficult to visually differentiate as the values of P and Q increase.

5.2 Example of the Procedure

5.3 Performance Metrics

The performance of the BFAST algorithm was evaluated by comparing the final activated map with the true activation map using:

- JI: Similarity between the two maps.
- False Positive Rate (FPR): Ratio of the voxels marked as activated that are not really active and the total number of inactive voxels.
- Activation Percentage (A%) Error: Difference in the percentage of active voxels between both maps.

See Tables ?? for the summary of these performance metrics.

Table 5.1: Performance Metrics Summary

P	Q	2D Map				3D Map			
		SNR	CNR	JI	FPR	A% Error	JI	FPR	A% Error
0	0	4.0636	2.8250	0.9256	0.0078	-0.2750	0.7677	0.0046	-0.1450
	1	3.6608	2.7607	0.8925	0.0180	0.5775	0.7249	0.0068	0.0425
	2	3.5568	2.7367	0.8754	0.0227	0.9200	0.6796	0.0086	0.1150
	3	3.5460	2.7310	0.8695	0.0244	1.0575	0.6767	0.0082	0.0425
1	0	3.5770	2.7364	0.8794	0.0222	0.9325	0.6741	0.0092	0.1975
	1	2.7336	2.5914	0.8509	0.0299	1.4600	0.6468	0.0108	0.3125
	2	2.4935	2.5298	0.8488	0.0308	1.5375	0.6410	0.0116	0.4025
	3	2.4497	2.5164	0.8533	0.0292	1.4125	0.6757	0.0099	0.3100
2	0	2.9354	2.5817	0.8685	0.0240	0.9700	0.6979	0.0083	0.1600
	1	2.1446	2.4047	0.8648	0.0267	1.2900	0.6566	0.0108	0.3575
	2	1.8597	2.3326	0.8578	0.0284	1.3950	0.6909	0.0079	0.0550
	3	1.7892	2.3115	0.8581	0.0283	1.3750	0.6859	0.0081	0.0775
3	0	2.5601	2.4581	0.8932	0.0174	0.5075	0.7297	0.0065	0.0125
	1	1.8244	2.2855	0.8882	0.0185	0.5750	0.7194	0.0067	0.0000
	2	1.5739	2.2153	0.8837	0.0200	0.7000	0.7167	0.0074	0.1075
	3	1.5132	2.1919	0.8793	0.0213	0.8025	0.7092	0.0066	-0.0600

Chapter 6

BFAST in a Real Dataset

Chapter 7

Conclusions

APPENDICES

Appendix A

Domain of Maximal Attraction Verification

In this section we are going to verify numerically the Theorem for the Truncated Normal Distribution using Python. For that, we first need to define the generalized normal distribution:

$$\phi(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (\text{A.1})$$

$$\Phi(x; \mu, \sigma^2) = \int_{-\infty}^x \phi(t; \mu, \sigma^2) dt \quad (\text{A.2})$$

In Python:

Now, we define the truncated normal distribution in $(0, 1)$, because we have a distribution of probabilities:

$$\psi(x; \mu, \sigma^2, 0, 1) = \begin{cases} 0 & \text{if } x \leq 0 \\ \frac{\phi(x; \mu, \sigma^2)}{\Phi(1; \mu, \sigma^2) - \Phi(0; \mu, \sigma^2)} & \text{if } 0 < x < 1 \\ 0 & \text{if } x \geq 1 \end{cases} \quad (\text{A.3})$$

$$\Psi(x; \mu, \sigma^2, 0, 1) = \int_0^x \psi(t; \mu, \sigma^2, 0, 1) dt \quad (\text{A.4})$$

In Python:

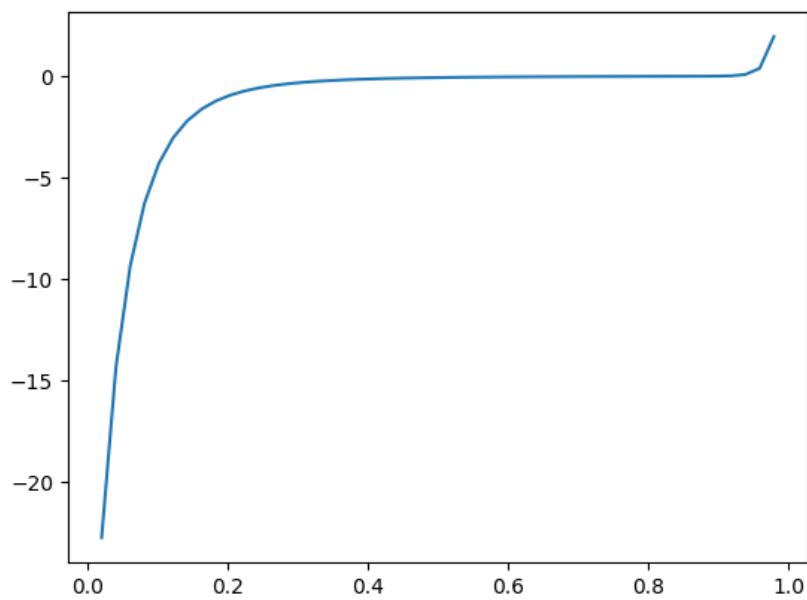
Now, part (c) of the theorem states that sufficient conditions for Ψ to belong to $D(G_3)$, where G_3 is the Gumbel distribution are:

- If $\psi(x) > 0$ and is differentiable for all x in (x_1, ϵ_1) for some x_1 , and

$$\lim_{x \rightarrow \epsilon_1} \frac{d}{dx} \left[\frac{1 - \Psi(x)}{\psi(x)} \right] = 0 \quad (\text{A.5})$$

Numerically, we take arbitrary values for μ and σ , then we substitute values for x within our domain $(0, 1)$ to obtain the following results.

Note that the function inside the limit on Equation ?? tends to 0 inside the $(0, 1)$ interval. Hence, Gumbel distribution can be used as a limiting distribution.



Appendix B

3D True Maps Generation

The procedure

Appendix C

Python Scripts