Summer 2024

New Methods in Electrical Source Imaging Based on EEG and Post-Mortem Pathology Data

Julio Cesar Enciso Alva
University of Texas at Arlington

Recommended Citation

Enciso Alva, Julio Cesar, "New Methods in Electrical Source Imaging Based on EEG and Post-Mortem Pathology Data" (2024). *Mathematics Dissertations*. 162.
https://mavmatrix.uta.edu/math_dissertations/162

New Methods in Electrical Source Imaging Based on EEG and Post-Mortem Pathology Data

by

JULIO CESAR ENCISO-ALVA

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON
August 2024

To my brother E. Ricardo, my mother María Guadalupe, and my late father Nicolás.

Despite the harsh times, we stay afloat and continue. This is the way.

ACKNOWLEDGEMENTS

I want to express my most sincere gratitude to Dr. Jianzhong Su for his encouragement and support when I was unsure of myself, making it possible for me to succeed eventually.

I also want to extend my gratitude to the members of my dissertation committee: Dr. Hristo V. Kojouharov, Dr. Ren-Cang Li, and Dr. Li Wang (in alphabetical order). Thanks for your continuous support and guidance during my time at UTA.

Thanks to Imelda Trejo-Lorenzo for encouraging, helping, and guiding me to pursue a PhD at an outstanding university in the USA.

I want to thank the colleagues I met at UTA for their company and moral support: Miguel A., Zachary V., Dwight W., Saber A., Saul C., Hrishabh K., Kiran M., Talon J., Amanda P., John M., Crystal M., Mondal Z., Oscar A., Madhu G., Zachary C., Jose L., Amairani S., Angela A., Kitham A., among many more. A comprehensive list of the friends who improved my life is an extensive task, and it is beyond the scope of this paragraph. I hope my presence was not a significant burden to y'all.

I want to mention my friends from home who are family in my heart: Omar C., Augusto T., Suri R., Berenice R., Jesus R., Valeria R., Ismael A., Maricela Z, Susana B., among others. Thanks for your friendship.

August 7, 2024

ABSTRACT

New Methods in Electrical Source Imaging Based on EEG and Post-Mortem Pathology Data

Julio Cesar Enciso-Alva, Ph.D.

The University of Texas at Arlington, 2024

Supervising Professor: Jianzhong Su, PhD

A central task for Neuroscience is to determine the location of electrical activity of neural origin inside the brain. Electrical signals can be recorded at a high resolution in time but low resolution in space, thus making it difficult to locate their source unambiguously. Electrical Source Imaging (ESI) is a particular framework for neural electrical source location; it is possible by modeling any additional information we may have about the electrical sources. For instance, minimal-norm estimators assume that the most plausible estimation is that with a lower norm. However, these estimators possess a low resolution in space.

In this work, we construct a novel ESI estimator incorporating binary anatomical data from pathologies observed in the post-mortem. This is done to improve the spatial resolution of the estimator.

This work may be extended to similar types of binary data derived from fMRI, NIRS, and CT, among others.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

One hundred years after its invention, the Electroencephalogram (EEG) is a popular tool for studying the physiological substrate of behavioral phenomena. The study of brain electrophysiology has led to a rather large corpus of electrical signals that are known to be related to normal and pathological states. For example, the interested reader may refer to the book from Niedermeyer [22].

In light of these historical antecedents, an important task for modern Neuroscience has been determining the location of electrical activity of neural origin inside the brain. For instance, we may take a set of EEG recordings related to a specific cognitive task and then ask which brain region is responsible for creating such electrical activity. This task is known as electrical source localization or electrical source reconstruction.

From this vague definition, the task is ill-posed (in the sense of Hadamard) and impossible to achieve without further information/assumptions.

Electrical Source Imaging is a framework for source reconstruction. Its base model reconstructs a group of dipoles whose electric field is equivalent to that produced by neurons. ESI methods are agnostic to the number or extension of electrical sources and require finding parameters from a linear system.

ESI methods have been used extensively in the literature to study pathological conditions. Asadzadeh et al. reported 41 different ESI algorithms between 1970 and June 2019, most of which were used to study clinical conditions.

We explore some ESI methods, their underlying assumptions, and related algorithms. We aim to incorporate a particular set of assumptions into the ESI model and derive an ESI method optimized for those assumptions.

The main focus of this work is the usability of the ESI methods in clinical settings. All the methods described in this work are available for public use in the GitHub repository of the author at https://github.com/EncisoAlva/Region-Priors. These methods are implemented in a format compatible with the Brainstorm toolbox [30], making it possible for non-technical users to incorporate them into a data analysis pipeline without writing code.

1.1 Contents and Organization

Chapter 2 is dedicated to deriving the Forward Problem in ESI from physical principles. At the end of the chapter, it is established that the following equation governs the ESI,

$$\mathbf{Y} = \mathbf{G}\mathbf{S} + \varepsilon$$

with $\mathbf{Y} \in \mathbb{R}^{M \times T}$ encoding the measurements from M point electrodes, $\mathbf{S} \in \mathbb{R}^{N \times T}$ encoding the electrical activity of N dipoles whose electrical field is equivalent to that of the neurons in the brain, $\mathbf{G} \in \mathbb{R}^{M \times N}$ is a mixing operator, and $\varepsilon \in \mathbb{R}^{M \times N}$ encodes some noise. In this context, it is expected that $M \ll N$.

The Forward Problem in ESI consists of computing $\mathbf{G}$ from anatomical data; some additional assumptions are necessary. The derivation presented in chapter 2 is standard; it can be found in the paper by Hallez et al. [12] or the book by Nunez and Srinivasan [24].

The objective of chapter 2 is to present the assumptions required for the model to work. Understanding the derivation of the realistic 4-sphere model, widely used in the literature, should be sufficient to understand the realistic 2-sphere model used in chapter 5.

Chapter 3 presents a basic review of the literature on ESI methods, parameter tuning, and performance evaluation of ESI methods.

This chapter presents some assumptions that have been incorporated into the ESI model, leading to algorithms with improved efficiency. Some of these ideas will be used to construct a novel method reflecting specific assumptions.

The parameter tuning and performance metrics methods will also be used in chapters 4 and 5.

Chapter 4 describes a protocol for creating synthetic data and explores some characteristics induced into the resulting dataset.

Chapter 5 describes a novel ESI method based on anatomical data provided by observed symptoms.

The model is derived from the assumptions, and we propose a way to implement it efficiently. This should prove that the proposed method isn't more computationally expensive than popular fast methods, such as sLORETA.

The proposed method is later used on a real dataset from an induced stroke in an animal model. The characteristics of the real data are the motivation of the model assumptions.

CHAPTER 2

FORWARD MODEL IN ELECTRICAL SOURCE IMAGING

Electroencephalography (EEG) is a type of electrophysiological recording that uses electrodes, whose contact surface is in the order of 20 mm^2, placed on the subject's scalp. The recording circuit is typically completed with a reference electrode at a 'neutral' location, like the contra-lateral earlobe or the nose; another electrode can also be used, resulting in referential channels.

In the century since the invention of the EEG machine, the simultaneous observation of behavioral data and EEG recordings has helped establish strong connections between specific EEG patterns and normal and pathological conditions.

The neural origin of electrical signals measured by EEG can be traced to the post-synaptic potentials in neurons' axons. Although each individual neuron cannot produce an electric potential field large enough to be measured by the EEG electrodes at the scalp, the simultaneous firing of all the neurons in the brain is sufficient to produce the EEG measurements.

It is important to acknowledge that the EEG may also register events whose origin is not neural, referred to as artifacts. These include but are not limited to external electrical fields, eye movements, activity from facial muscles, and movement from the EEG electrodes —if they are not fixated adequately to the scalp or if the subject's movement is excessive. This work does not discuss any methods for dealing with artifacts; the interested reader must refer to specialized manuals, for example, that from Fish and Spehlmann [5].

In the absence of biological or instrumental artifacts, the EEG measurements are produced exclusively by brain activity. The electric potential fields at the EEG scale, often called macro-scale fields, correspond to large groups of neurons that fire simultaneously and have a common physical orientation. It is estimated that groups of neurons of at least 6 mm^2 are required to produce activity measurable at the EEG scale.

2.1 Electrical Source Imaging Framework

Electrical Source Imaging (ESI), or Electrical Source Reconstruction, involves identifying the neural electrical sources responsible for the observed EEG. ESI can be used to identify brain regions related to (and, in some cases, responsible for) the normal or pathological state under study.

In this work, we perform ESI under the distributed dipole paradigm: the current density field produced by neuronal activity is approximated by a finite set of electrical dipoles whose location is known –on a grid that covers the brain region of interest– and whose magnitude is to be determined.

This paradigm can be justified formally as a discretized multi-pole expansion of order 2 (using dipoles) of the current density field; this approximation is sufficient to explain measurements obtained 1 mm or farther from the electrical sources [23].

Within the paradigm of distributed dipoles, ESI's task is to construct a representative set of dipoles that will produce a current density equivalent to that of the real neurons in the brain; the resulting model should operate at the EEG scale. This paradigm is considered to be non-parametric in the sense that it is agnostic to the number and extent of the electric sources. Since an area with no discernible activity will produce a group of dipoles with a magnitude close to zero, this paradigm is considered robust to model misidentification.

When interpreting the results obtained from ESI, one should be aware of the limitations inherited from EEG. For example, the lack of synchronization of time and direction of some active neuron patches may prevent them from producing an electric potential that the EEG can measure. This silent activity can't be measured by ESI methods based on EEG data.

The task of performing ESI within the paradigm of distributed dipoles is an ill-posed inverse problem.

The 'inverse problem' denomination is earned since the reciprocal problem is relatively more straightforward: constructing EEG recordings given the current density. This problem is explored in Section 2.2, following that of Hallez et al. [12].

The 'ill-posed' denomination is earned since, given the data from EEG recordings, the current density is not uniquely defined. In Chapter 3, we review some methods used to solve this problem, while in Chapter 3, we propose a novel method for this purpose.

2.2 Derivation of the Forward Problem

From a Physics framework, we can describe that EEG sensors measure the electric scalar potential field, $V : \mathbb{R}^3 \to \mathbb{R}$, at a finite set of points in space, $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M \in \mathbb{R}^3$. Thus, $V(\mathbf{s}_m)$ represents the measurement from the m-th electrode.

We assume that the electric scalar potential field is caused by a finite number of dipoles with known locations, $\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_N \in \mathbb{R}^3$, and whose momentum, $\mathbf{m}_1, \mathbf{m}_2, \ldots \mathbf{m}_N \in \mathbb{R}$, are to be determined.

We first construct a model for a single sensor and a single dipole with unit moment. The full model will be constructed later by superposing multiple copies of this base model.

Under these circumstances, we consider the electric field, $E : \mathbb{R}^3 \to \mathbb{R}^3$, and the current density field, $K : \mathbb{R}^3 \to \mathbb{R}^3$,

$$E = -\nabla V \tag{2.2.1}$$

$$K = \sigma E = -\sigma \nabla V \tag{2.2.2}$$

with $\sigma : \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ the position-dependent conductivity tensor. In the isotropic case, we assume $\sigma(\mathbf{r}) = \kappa \mathbf{I}_3$ for all $\mathbf{r}$, where $\kappa > 0$ is the constant conductivity and $\mathbf{I}_3 \in \mathbb{R}^{3\times 3}$ is the identity matrix.

The current density field has the property that its divergence vanishes everywhere except near the dipole. In more precise terms, the divergence of the current density field is defined as

$$\nabla \cdot K = \lim_{G(\mathbf{r}) \to \{\mathbf{r}\}} \frac{1}{\operatorname{vol}\left(G\left(\mathbf{r}\right)\right)} \oint_{\partial G(\mathbf{r})} K(\mathbf{r})\, dS \tag{2.2.3}$$

with $G(\mathbf{r}) \subset \mathbb{R}^3$ a set that contains $\mathbf{r}$. The limit must be interpreted in a similar way as Riemann sums, in the sense that the limit is not affected by the selections of $G(\mathbf{r})$.

For the case of one single dipole, $\nabla \cdot K$ can be simplified,

$$\nabla \cdot K(\mathbf{r}) = \delta\left(\mathbf{r} - \mathbf{d}^+\right) - \delta\left(\mathbf{r} - \mathbf{d}^-\right) = \begin{cases} +1, & \text{if } \mathbf{r} = \mathbf{d}^+ \\ -1, & \text{if } \mathbf{r} = \mathbf{d}^- \\ 0, & \text{otherwise.} \end{cases} \tag{2.2.4}$$

where $\mathbf{d}^-, \mathbf{d}^+ \in \mathbb{R}^3$ are the dipole's electric sink and source positions, respectively. For ease of notation, we select $\mathbf{d}^-$ and $\mathbf{d}^+$ so that the dipole position, $\mathbf{d}$, and momentum, $\mathbf{m}$, satisfy the following

$$\mathbf{d} = \frac{1}{2}\left[\mathbf{d}^- + \mathbf{d}^+\right] \tag{2.2.5}$$

$$\mathbf{m} = \left[\mathbf{d}^+ - \mathbf{d}^-\right] \tag{2.2.6}$$

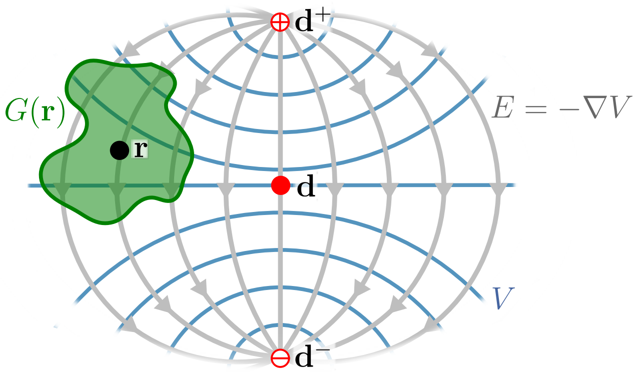

Figure 2.1. Visual representation of a single dipole with electric sink and source located at $\mathbf{d}^-$ and $\mathbf{d}^+$, respectively. The 'position' of the dipole, $\mathbf{d}$, is the midpoint of $\mathbf{d}^-$ and $\mathbf{d}^+$. The electric scalar field, V, is denoted by its isopotential lines, while the electrical field, $E = -\nabla V$, is shown as the lines of the fastest descent. On this diagram is shown an example of a set $G(\mathbf{r}) \subset \mathbb{R}^3$ that contains $\mathbf{r}$; these sets are used in equation (2.2.3) .

Refer to figure 2.1 for a graphical summary of this construction.

With equations (2.2.2) and (2.2.4) at hand, we may determine the electric scalar field, V, in terms of the only dipole as

$$\nabla \cdot (\sigma(\mathbf{r}) \nabla V(\mathbf{r})) = \delta\left(\mathbf{r} - \mathbf{d}^+\right) - \delta\left(\mathbf{r} - \mathbf{d}^-\right) \tag{2.2.7}$$

or, in the isotropic case, as

$$\kappa \Delta V(r) = \delta\left(\mathbf{r} - \mathbf{d}^+\right) - \delta\left(\mathbf{r} - \mathbf{d}^-\right) \tag{2.2.8}$$

It is relevant to recall that the EEG measurement from the only sensor is $V(\mathbf{s})$, with $\mathbf{s}$ the locations of the EEG sensors.

2.2.1 Boundary Conditions

To have a meaningful solution for equation (2.2.7) or (2.2.8) in the context of EEG, we must consider appropriate domain and boundary conditions.

Since the electric conductivity of air is very small compared with that of head tissues, it is straightforward to use the subject's head as a domain with reflective

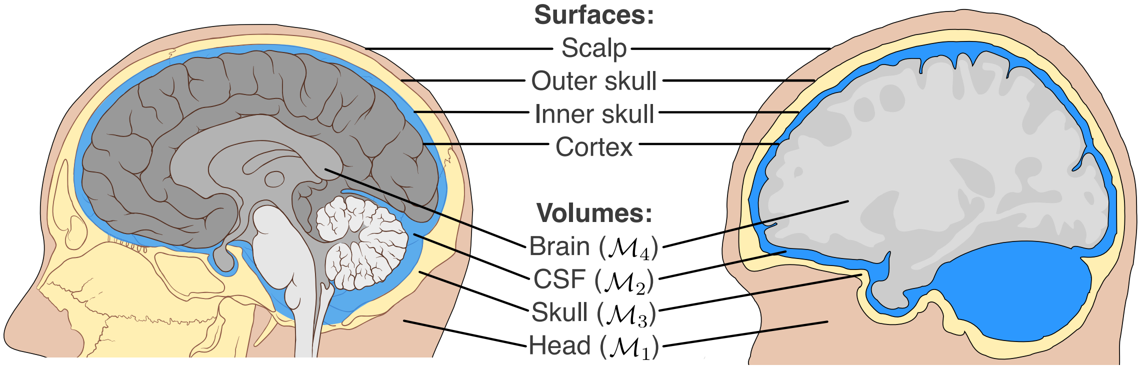

Figure 2.2. Geometric domain defined by relevant tissues in the subject's head. The 4-sphere model considers three media: brain, cerebrospinal fluid (CSF), and the rest of the head. Relevant interfaces between media are displayed: brain cortex, inner and outer skull, and scalp. Notice the anatomical simplifications described in the text. (Left) Realistic illustration of a human head. Composition from two diagrams by Patrick L. Lynch and C. Carl Jaffe [18, 19]. (Right) Cross-section of the surfaces obtained from the ICBM152 template after segmentation.

boundary conditions. Thus, the head is divided into a series of nested volumes, determined by both the biological tissues in place and their respective electric properties.

Under these conditions, the simplest model considers a small number of tissue-based regions as homogeneous isotropic media with constant conductivity; different media may have different conductivity. The 4-sphere model considers only four media: the brain, $\mathcal{M}_4$, cerebrospinal fluid (CSF), $\mathcal{M}_3$, skull, $\mathcal{M}_2$, and the rest of the head, $\mathcal{M}_1$. For computational purposes, we may consider a fifth medium, $\mathcal{M}_0$, representing the outside of the head.

Multiple approximations are used in practice, as illustrated in figure 2.2 where portions of the skull far from the brain are ignored, as well as muscle and connective tissues because those tissues are too far from the EEG sensors to make a significant difference in the model.

Tissue	Conductivity (µS/cm)
White matter	1.43
Gray Matter	3.33
Cerebrospinal Fluid	15.38
Skull Tissue	625.00
Scalp	1.74

Table 2.1. conductivity for different head tissues considered in the Four-sphere model, measured in microsiemens per centimeter. Data from [27]

With the media defined formally, reflective boundary conditions can be written as follows,

$$\sigma(\mathbf{r})\,\nabla V(\mathbf{r})\cdot\mathbf{n}=0,\ \text{ for } \mathbf{r}\in\mathcal{M}_1 \tag{2.2.9}$$

where $\mathbf{n}$ is a vector orthogonal to $\partial\mathcal{M}_1$. At the interfaces between media, the boundary conditions are to have continuous first derivatives; this property can be requested for the whole domain,

$$V\left(\bullet\right),\sigma\left(\bullet\right)V\left(\bullet\right)\in C^1\left(\mathcal{M}_1\cup\mathcal{M}_2\cup\mathcal{M}_3\cup\mathcal{M}_4\right) \tag{2.2.10}$$

The continuity of $\sigma(\mathbf{r})V(\mathbf{r})$ is particularly important on the interfaces between media because the conductivity, σ, is expected to be homogeneous within each media but to change drastically between different media.

To make sense of the magnitudes of the discontinuities of the conductivity, σ, among different media, in table 2.1 are displayed typical values for the four media described before; data is from Ramon et al. [27]. See figure 2.2 for a visual reference of these tissues.

2.2.2 Construction of the leadfield operator

Recall that both equations (2.2.7) and (2.2.8) are constructed for one sensor and one dipole with unit magnitude. We now consider a finite number of EEG sensors at locations $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M \in \mathbb{R}^3$, and dipoles at locations $\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_N \in \mathbb{R}^3$ with momenta $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_N$. For ease of notation, write each momentum as

$$\mathbf{m}_n = \rho_n \, \mathbf{e}_n \tag{2.2.11}$$

with $\rho_n \in \mathbb{R}, \mathbf{e}_n \in \mathbb{R}^3$ such that $\mathbf{e}_n$ has unit norm, i.e. $\|\mathbf{e}_n\|_2 = 1$.

After making the assumptions described in the previous sections, we may conclude that the electric scalar field, V, is given by

$$V(\mathbf{r}) = \sum_{n=1}^{N} \rho_n \tilde{V}_n(\mathbf{r}) \tag{2.2.12}$$

where $\tilde{V}_n$ is the solution to the following equation

$$\nabla \cdot \Big(\sigma(\mathbf{r}) \, \nabla \tilde{V}_n(\mathbf{r})\Big) = \delta\left(\mathbf{r} - \mathbf{d}_n^+\right) - \delta\left(\mathbf{r} - \mathbf{d}_n^-\right) \tag{2.2.13}$$

under the boundary conditions from equation (2.2.9), and with each $\mathbf{d}_n^\pm$ constructed from $\mathbf{d}_n$. For ease of notation, define

$$g(\mathbf{r}, \mathbf{d}_n) = \tilde{V}_n(\mathbf{r}) \tag{2.2.14}$$

and thus, we can write

$$V(\mathbf{r}) = \sum_{n=1}^{N} \rho_n \, g\left(\mathbf{r}, \mathbf{d}_n\right) = \begin{bmatrix} g\left(\mathbf{r}, \mathbf{d}_1\right) & g\left(\mathbf{r}, \mathbf{d}_2\right) & \cdots & g\left(\mathbf{r}, \mathbf{d}_N\right) \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_N \end{bmatrix} \tag{2.2.15}$$

Furthermore, by using the equation for all the sensor locations and stacking the results, we obtain the following vector equation

$$\begin{bmatrix} V\left(\mathbf{s}_1\right) \\ V\left(\mathbf{s}_2\right) \\ \vdots \\ V\left(\mathbf{s}_M\right) \end{bmatrix} = \begin{bmatrix} g\left(\mathbf{s}_1, \mathbf{d}_1\right) & g\left(\mathbf{s}_1, \mathbf{d}_2\right) & \cdots & g\left(\mathbf{s}_1, \mathbf{d}_N\right) \\ g\left(\mathbf{s}_2, \mathbf{d}_1\right) & g\left(\mathbf{s}_2, \mathbf{d}_2\right) & \cdots & g\left(\mathbf{s}_2, \mathbf{d}_N\right) \\ \vdots & \vdots & \ddots & \vdots \\ g\left(\mathbf{s}_M, \mathbf{d}_1\right) & g\left(\mathbf{s}_M, \mathbf{d}_2\right) & \cdots & g\left(\mathbf{s}_M, \mathbf{d}_N\right) \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_N \end{bmatrix} \tag{2.2.16}$$

Equation (2.2.16) can be rewritten as the following generic matrix equation

$$\mathbf{Y} = \mathbf{G}\,\mathbf{S} \tag{2.2.17}$$

where $\mathbf{Y} \in \mathbb{R}^{M\times 1}$ encodes the EEG measurements, $\mathbf{S} \in \mathbb{R}^{N\times 1}$ encodes the dipoles' magnitude, and $\mathbf{G} \in \mathbb{R}^{M\times N}$ is a matrix known as the leadfield matrix, gain matrix, forward operator, among other names; $\mathbf{G}$ encodes the mixture of current density to form the scalar electrical field measured by EEG sensors.

If the dipole's orientation, $\mathbf{e}_n$, is to be determined, we can consider each dipole as the sum of three dipoles parallel to some orthogonal directions. Such as in figure 2.3, we can write

$$\rho_n \mathbf{e}_n = \rho_n^x \mathbf{e}_n^x + \rho_n^y \mathbf{e}_n^y + \rho_n^z \mathbf{e}_n^z \tag{2.2.18}$$

with $\mathbf{e}_n^x, \mathbf{e}_n^y, \mathbf{e}_n^z \in \mathbb{R}^3$ vectors with unit norm located at $\mathbf{d}_n$ and parallel to the respective $x-, y-, z-$axis.

Under these new circumstances, equation (2.2.15) changes to

$$V(\mathbf{r}) = \sum_{n=1}^{N} \left[\rho_n^x\, g\left(\mathbf{r}, \mathbf{e}_n^x\right) + \rho_n^y\, g\left(\mathbf{r}, \mathbf{e}_n^y\right) + \rho_n^y\, g\left(\mathbf{r}, \mathbf{e}_n^z\right)\right] \tag{2.2.19}$$

which can be written as a matrix equation such as (2.2.17).

The overall effect of assuming unknown orientations are (1) the increase of the number of dipole magnitudes to be determined by a factor of 3, as well as the size of the leadfield matrix, and (2) we need to compute the magnitude of these dipoles in order to recover the original interpretation.

The forward model summarized in equation (2.2.17) can be extended trivially to account for T EEG measurements in time by setting $\mathbf{Y} \in \mathbb{R}^{M\times T}$, $\mathbf{S} \in \mathbb{R}^{M\times T}$, and $\mathbf{G}$ unchanged. This is possible due to the quasi-static Maxwell equations. Since the neural electric fields change in time at relatively low frequencies, they don't induce an important change to the magnetic field [24]. The magnetic fields produced by

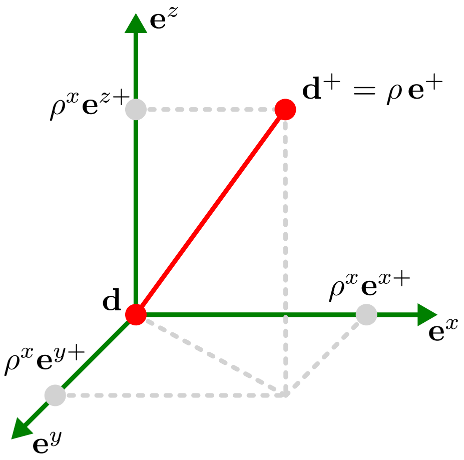

Figure 2.3. Decomposition of a dipole with arbitrary momentum as a sum of dipoles whose momentum is parallel to their respective $x-, y-, z-$axis. By definition, a dipole is made of a sink and a source with equal charges. For visual simplicity, only the source $\mathbf{d}^+$ is shown; the sink $\mathbf{d}^-$ can be constructed similarly.

the electric potentials can be considered constant, and thus, any capacitance can be neglected.

It is important to acknowledge that the changes in the electric and magnetic fields are independent of biological tissue. However, the neural events that generate electrical fields also generate magnetic fields, which can be measured using Magnetoencephalography (MEG).

The quasi-static properties of the system imply that the magnetic and electric fields don’t interact with each other. Any correlation between them comes from their neural generators. A complete discussion of this phenomenon is beyond the scope of this work.

2.3 Practical considerations

2.3.1 Geometry Identification from Data

Recall that the construction of the leadfield matrix, **G**, is based on solving the equation (2.2.7) (or equation (2.2.8) in the isotropic case) subject to the boundary conditions in (2.2.9).

These equations can be solved numerically using Boundary Element Methods (BEM) or Finite Element Methods. A full review of these methods –even in the particular case of ESI– is beyond the scope of this work, and the interested reader may refer to the review by Hallez [12] or that by Vorwerk et al. [32].

These methods are implemented in commercially available software such as OpenMEEG [10], DUNEuro [29], and MNE Python [11].

In this section, the goal is to describe the identification of the tissue media from an operative standpoint. Tissues are often defined from the subject's anatomical data or an appropriately matched template.

The anatomical data often originates as the result of T1 or T2 magnetic resonance imaging (MRI) from the subject. Later, the MRI images are segmented: relevant tissues are identified, and triangulated 3D surfaces are constructed. Relevant algorithms for these tasks are available on commercial software such as CAT12 [7], and Fieldtrip [25].

It is relevant to mention that the Fours Sphere model is not the only option for constructing the leadfield matrix. For instance, the model can be trivially extended to include five media by separating the brain into gray and white matter. Such division requires a robust algorithm to distinguish between those tissues; since the interface is known to be non-smooth, robustness is a strong requirement.

2.3.2 Location of Distributed Dipoles

Over the previous sections, the locations of the distributed dipoles, $\mathbf{d}_n$, were assumed to be known. Two cases were considered: if their orientation, $\mathbf{m}_n$, was either assumed to be known or not.

The goal of ESI is to gain knowledge from the neural sources of electrical activity, including their location. Thus, we shouldn't make assumptions about their location.

In that sense, the most inclusive assumption is that the dipoles should be located 'anywhere' inside the brain. This assumption can be enforced by setting the locations of the distributed dipoles over a grid covering the entire brain volume. In figure 2.4, two of these arrangements are displayed: a cubic grid and an adaptive grid. The adaptive grid, as implemented on the Brainstorm software [30], is constructed following these steps:

1. Use cortex as the initial layer.
2. Downsample to locate dipoles with a prescribed distance between them.
3. Shrink to obtain the next layer.
4. Repeat until no more layers can be constructed.

Adaptive grids offer some advantages since they are conceptually consistent with depth weighting (see section 3.1.2), and they can be constructed so that the outer layer falls within a certain distance from the cortex.

In some scenarios, the assumption of a total lack of information may be replaced by the assumption that the dipoles are located in the cortex and are oriented orthogonally to it. This is justified by the peculiar geometry of the outer layers of the gray matter, which consist of neuron dendrites orthogonal to the cortex surface. In figure 2.4, we may observe an arrangement of distributed dipoles that encodes this assumption.

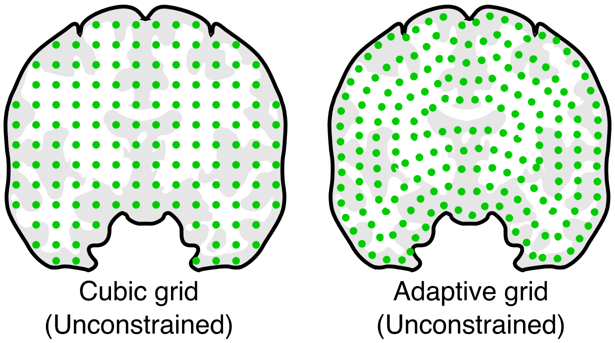

Figure 2.4. Some configurations for distributed dipoles. Constrained dipoles have a known orientation (orthogonal to the cortex surface), while unconstrained dipoles have an unknown orientation..

Whether we may or not assume that the dipoles are located only in the cerebral cortex depends on additional assumptions of the neural processes under study. Many ‘higher’ cognitive processes are known to occur in the cerebral cortex: speech generation and recognition, face recognition, and spatial orientation, among others [28]. Conversely, we may ignore some deeper brain regions if their associated neural functions are not part of the phenomenon under study.

Many software dedicated to ESI refers to the case when dipoles are orthogonal to the cortex surface as ‘constrained’, while the other case is ‘unconstrained’. Other setups are possible, yet their discussion is outside the scope of this work.

CHAPTER 3

REVIEW OF METHODS FOR ELECTRICAL SOURCE RECONSTRUCTION

The forward model discussed in Chapter 2 can be summarized in the following matrix equation

$$\mathbf{Y} = \mathbf{G}\ (\mathbf{S} + \nu) + \varepsilon$$

where $\mathbf{Y} \in \mathbb{R}^{M \times T}$ encodes the EEG measurements over time, $\mathbf{S}$ encodes the magnitudes of a finite set of distributed dipoles with known locations, and $\mathbf{G} \in \mathbb{R}^{M \times N}$ is a mixing operator referred to as the Leadfield matrix.

The terms $\nu \in \mathbb{R}^{N \times T}, \varepsilon \in \mathbb{R}^{M \times T}$ represent additive noise within the true values for the dipoles and sensors, respectively. Different authors have proposed different assumptions for such noise terms, some of which will be discussed later.

In chapter 2, it was discussed that if the orientation is known, then we can use $\mathbf{S} \in \mathbb{R}^{N \times T}$, and otherwise, we should use $\mathbf{S} \in \mathbb{R}^{3N \times T}$ from decomposing each dipole as the sum of dipoles with known orientations, parallel to the x-, y- and z-axis. Since the cases of known and unknown dipole orientations don't significantly change the model formulation, they will be treated as a single model whenever possible.

Within the context of Electrical Source Reconstruction, the computation of the leadfield matrix is referred to as the Forward Problem. This computation implies a large number of steps: identification of tissues from anatomical MRI or similar, construction of triangulated surfaces representing anatomical regions of interest (which may be simplified), setup of several PDEs within the anatomical regions as geometric domain, and the numerical solution of said equations by either BEM or FEM (or a different method).

All these points were discussed in Chapter 2. Thus, after this point, we will consider the leadfield matrix $\mathbf{G}$ to be given for each set of anatomical data and electrode montage.

3.1 Inverse problem formulations

The main aim of this work will be the task of 'recovering' $\mathbf{S}$ once the data $\mathbf{Y}$ and some additional assumptions are considered.

In the absence of biological noise ($\nu = 0$), the most straightforward path of action is to find the $\mathbf{S}$ that minimizes the reconstruction error

$$\hat{\mathbf{S}}_{\text{naive}} = \arg\min_{\mathbf{S}} \|\mathbf{G}\,\mathbf{S} - \mathbf{Y}\|_F^2 \tag{3.1.1}$$

with $\|A\|_F$ the Frobenius norm of $A \in \mathbb{R}^{M\times N}$, defined as

$$\|A\|_F = \left(\sum_{n=1}^{N} \sum_{m=1}^{M} [A(m,n)]^2 \right)^{1/2} \tag{3.1.2}$$

The expression in (3.1.1) represents an ill-posed problem due to (1) not having a unique solution and (2) changing significantly with small changes in the parameters. The first statement is straightforward since we assumed $M \ll N$. The second statement becomes clear if we choose the Least-Square solution,

$$\hat{\mathbf{S}}_{\text{LS}} = \mathbf{G}^T \left(\mathbf{G}\,\mathbf{G}^T\right)^{\dagger} \mathbf{Y} \tag{3.1.3}$$

where $A^\dagger$ is the Moore-Penrose pseudoinverse of $A = \mathbf{G}\,\mathbf{G}^T$.

It can be proven that the eigenvalues of $A^\dagger$ are given by

$$(\sigma_i)^\dagger = \begin{cases} \sigma_i^{-1}, & \text{if } \sigma_i \neq 0 \\ 0, & \text{otherwise.} \end{cases} \tag{3.1.4}$$

where σ_i are the eigenvalues of $A = \mathbf{G}\,\mathbf{G}^T$. Notice that the case $0 \neq \sigma_i \approx 0$ leads to $\hat{\mathbf{S}}_{\text{LS}}$ being 'unstable' as a function of $\mathbf{G}$ and $\mathbf{Y}$.

As is discussed later, one way to alleviate the instability of the Least-Squares estimator is to modify the inverse problem by introducing a regularisation parameter λ, resulting in

$$\hat{\mathbf{S}} = \arg\min_{\mathbf{S}} \|\mathbf{G}\,\mathbf{S} - \mathbf{Y}\| + \lambda\, f\left(\mathbf{S}\right) \tag{3.1.5}$$

with $\lambda > 0$ and $f : \mathbb{R}^{M \times T} \to \mathbb{R}_+$ a penalization function, usually related to the norm of $\mathbf{S}$. It will be shown how this formulation naturally results from certain assumptions, but it can also be designed to impose certain desirable properties on the solution.

3.1.1 Minimal Norm Estimator

For this estimator, we assume that both the sensor and internal noise are Gaussian random variables, independent in time and space, and with constant variance, i.e.

$$\nu \sim \mathcal{N}\left(0, \sigma^2 \mathbf{I}_N \otimes \mathbf{I}_T\right) \tag{3.1.6}$$

$$\varepsilon \sim \mathcal{N}\left(0, \lambda^2 \mathbf{I}_M \otimes \mathbf{I}_T\right) \tag{3.1.7}$$

The Maximum A Posteriori (MAP) estimator of $\mathbf{S}$ is defined as

$$\begin{aligned} \hat{\mathbf{S}}_{\mathrm{MAP}} &= \arg\max_{\mathbf{S}}\ \mathrm{Prob}\left(\mathbf{S} \mid \mathbf{Y}\right) \\ &= \arg\max_{\mathbf{S}}\ \mathrm{Prob}\left(\mathbf{Y} \mid \mathbf{S}\right)\ \mathrm{Prob}\mathbf{S}. \end{aligned} \tag{3.1.8}$$

The above problem can be simplified by (1) taking the logarithm of the expression since log is a monotonic function, and by (2) considering the independence in time so that columns of $\mathbf{S}$ are computed independently, and (3) considering the normal distribution of ν and ε. Thus we have

$$\begin{aligned} \hat{\mathbf{S}}(:,t)_{\mathrm{MAP}} &= \arg\max_{\mathbf{S}}\ \log\left(\mathrm{Prob}\left(\mathbf{Y}(:,t) \mid \mathbf{S}\right)\ \mathrm{Prob}\left(\mathbf{S}\right)\right) \\ &= \arg\max_{\mathbf{S}}\ -\frac{1}{2\sigma^2}\left\|\mathbf{G}\,\mathbf{S} - \mathbf{Y}(:,t)\right\|_2^2 - \frac{1}{2\lambda^2}\left\|\mathbf{S}\right\|_2^2 \end{aligned}$$

$$= \arg\min_{\mathbf{S}} \ \|\mathbf{G}\,\mathbf{S} - \mathbf{Y}(:,t)\|_2^2 + \frac{\sigma^2}{2\lambda^2}\|\mathbf{S}\|_2^2 \tag{3.1.9}$$

Notice how the MAP estimator can be interpreted as a regularized optimization problem in the form of (3.1.5). Since the MAP estimator is obtained by minimizing the norm of $\mathbf{S}$ (in particular, the ℓ_2 norm), it is referred to as the Minimal-Norm Estimator (MNE). This designation is ambiguous since many other estimators depend on minimizing some norm of the sources; thus, 'minimal-norm estimator' refers to a family of estimators. To avoid confusion, this particular estimator will be referred to simply as MNE in this work.

The MNE estimator admits a closed-form solution as follows,

$$\hat{\mathbf{S}}_{\text{MNE}} = \hat{\mathbf{S}}_{\text{MAP}} = \mathbf{G}^T\left(\mathbf{G}\,\mathbf{G}^T + \frac{\sigma^2}{\lambda^2}\mathbf{I}_M\right)^{-1}\mathbf{Y} \tag{3.1.10}$$

3.1.2 Weighted Minimal Norm Estimator

This method incorporates the assumption that smaller coefficients in the lead-field matrix scale the contribution of deep dipoles located far from the scalp. These deeper dipoles must have larger magnitudes to contribute similarly to shallow dipoles in the generation of EEG. The overall effect is that deep dipoles are misrepresented in the MNE estimator and often deleted. This assumption can be encoded by considering the following noise term,

$$\nu \sim \mathcal{N}\left(0, W^{-1} \otimes \mathbf{I}_T\right) \tag{3.1.11}$$

$$W = \operatorname{diag}\left(\|\mathbf{G}(:,1)\|_2, \|\mathbf{G}(:,2)\|_2, \ldots, \|\mathbf{G}(:,N)\|_2\right) \tag{3.1.12}$$

In order to alleviate the effect mentioned above, the depth-weighting matrix W is introduced. The Weighted Minimal Norm Estimator (wMNE) is then defined as

$$\hat{\mathbf{S}}_{\text{wMNE}} = \left(\mathbf{G}^T\mathbf{G} + \lambda\, W^T W\right)^{\dagger}\mathbf{G}^T\mathbf{Y} \tag{3.1.13}$$

$$= \mathbf{G}^T\left(\mathbf{G}\left(W^T W\right)^{-1}\mathbf{G}^T + \lambda\,\mathbf{I}_M\right)^{\dagger}\mathbf{Y} \tag{3.1.14}$$

3.1.3 LORETA

The Low-Resolution Electric Tomography (LORETA) estimator was proposed by Marqui[26]. One key idea from this method is to use a Laplacian operator to impose spatial smoothness over $\mathbf{S}$ and use the same depth-weighting from wMNE.

In its original formulation, the dipoles are assumed to be located on a cubic grid with a distance d between neighboring dipoles.

The LORETA estimator can be expressed as an optimization problem

$$\begin{aligned} \hat{\mathbf{S}}_{\text{sLORETA}} &= \arg\min_{\mathbf{S}} \; \|B\,W\,\mathbf{S}\|_F^2 \\ \text{s.t.}\;\; &\mathbf{G}\,\mathbf{S} = \mathbf{Y} \end{aligned} \tag{3.1.15}$$

with W as in (3.1.12), and B the laplacian operator defined as

$$B(n_i, n_j) = \begin{cases} 6/d^2, & \text{if } n_i = n_j \\ -1/d^2, & \text{if } \left\|\mathbf{r}_{n_i} - \mathbf{r}_{n_j}\right\|_2 = d \\ 0, & \text{otherwise} \end{cases} \tag{3.1.16}$$

where d is the distance between dipoles. On a cubic grid, exactly 6 dipoles satisfy the second condition when the n_i-th dipole is far from the cortex.

The estimator is computed as follows:

$$\hat{\mathbf{S}}_{\text{LORETA}} = \left(W\,B^T\,B\,W\right)^{-1}\mathbf{G}^T\left[\mathbf{G}\left(W\,B^T\,B\,W\right)^{-1}\mathbf{G}^T\right]^{\dagger}\mathbf{Y} \tag{3.1.17}$$

3.1.4 sLORETA

The standardized Low-Resolution Electric Tomography (sLORETA) estimator was proposed by Marqui [26]. Despite their similar name, LORETA and sLORETA are two different methods. The sLORETA estimator considers the following minimization problem

$$\hat{\mathbf{S}}_{\text{sLORETA}} = \arg\min_{\mathbf{S}} \; \|\mathbf{G}\,\mathbf{S} - \mathbf{Y} - c\mathbb{1}\|_F^2 + \lambda\,\|\mathbf{S}\|_F^2 \tag{3.1.18}$$

with $\mathbb{1} = \{1\}^{M \times 1}$ a column vector of ones.

The underlying assumption is that the measurements are re-referenced to the average, i.e., the raw observations $\mathbf{Y}_0 \in \mathbb{R}^{M \times T}$ are pre-processed as follows

$$\mathbf{Y} = (\mathbf{I}_M - H)\,\mathbf{Y}_0 \tag{3.1.19}$$

with $H \in \mathbb{R}^{M \times M}$ an averaging operator constructed as

$$H = \left(\mathbb{1}_M \mathbb{1}_M^T\right) / \left(\mathbb{1}_M^T \mathbb{1}_M\right). \tag{3.1.20}$$

Alternatively, we can construct H element-wise as

$$H(m_1, m_2) = \frac{1}{M} \tag{3.1.21}$$

for $m_i = 1, 2, \ldots, M$ and $i = 1, 2$.

It is relevant to mention that this type of re-reference can be interpreted as constructing a phantom electrode with neutral electric activity and using it to close the recording circuit. This process is expected to eliminate noises that are common to all electrodes.

Assuming that $\mathbf{Y}$ was re-referenced to the average, then the parameter c is always zero, and the estimator is as follows

$$\hat{\mathbf{S}}_{\text{sLORETA}} = \mathbf{G}^T H \left(H\,\mathbf{G}\,\mathbf{G}^T H + \lambda H\right)^{\dagger} \mathbf{Y} \tag{3.1.22}$$

3.1.5 Multivariate Source Prelocalization (MSP)

The MSP estimator proposed by Mattout [20] incorporates the idea of using Single Value Decomposition (SVD) to produce a subspace that captures the most relevant information in $\mathbf{Y}$ and $\mathbf{G}$, then gets the sources more correlated with their projections into that space. The obtained information is later used to construct a weight matrix, so the final estimation is similar to wMNE.

$$\hat{\mathbf{S}}_{\text{MSP}} = \arg\min_{\mathbf{S}} \left\|\mathbf{Y} - \mathbf{G}\mathbf{S}\right\|_F^2 + \lambda \left\|W_{\text{MSP}}\mathbf{S}\right\|_F^2 \tag{3.1.23}$$

The leadfield matrix is column-normalized as in 3.1.12, and the SVD decomposition is performed,

$$\overline{\mathbf{G}} = \mathbf{G}\, W_G^{-1} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T \tag{3.1.24}$$

$$W_G = \operatorname{diag}\left(\|\mathbf{G}(:,1)\|_2, \|\mathbf{G}(:,2)\|_2, \ldots, \|\mathbf{G}(:,N)\|_2\right) \tag{3.1.25}$$

with $\mathbf{U} \in \mathbb{R}^{M\times M}, \mathbf{V} \in \mathbb{R}^{N\times N}$ such that $\mathbf{U}^T\mathbf{U} = \mathbf{I}_M, \mathbf{V}^T\mathbf{V} = \mathbf{I}_N$, and $\boldsymbol{\Lambda} \in \mathbb{R}^{M\times N}$ a matrix with the singular values of $\overline{\mathbf{G}}$ in descending order on its diagonal.

The first s columns of $\mathbf{U}$ are selected based on the Velicer criterion [31],

$$\mathbf{U}_s = [\mathbf{U}(:,1), \mathbf{U}(:,2), \ldots, \mathbf{U}(:,s)] \tag{3.1.26}$$

$\mathbf{Y}$ is normalized on a similar way as $\overline{\mathbf{G}}$, and then it is projected into $\mathbf{U}_s\,\mathbf{U}_s^T$,

$$\overline{\mathbf{Y}}_s = \mathbf{U}_s\mathbf{U}_s^T\overline{\mathbf{Y}} = \mathbf{U}_s\mathbf{U}_s^T W_Y \mathbf{Y} \tag{3.1.27}$$

$$W_Y = \operatorname{diag}\left(\|Y(:,1)\|_2, \|Y(:,2)\|_2, \ldots, \|Y(:,T)\|_2\right) \tag{3.1.28}$$

$$\mathbf{A}_s = \left(\overline{\mathbf{G}}^T\overline{\mathbf{Y}}_s\right)\left(\overline{\mathbf{G}}^T\overline{\mathbf{Y}}_s\right)^T \tag{3.1.29}$$

The authors of MSP propose interpreting $\mathbf{A}_s$ as a correlation matrix of the sources within the resulting space. Following this interpretation, the authors define Activation Probability Map (APM) as

$$D_s(j) = \mathbf{A}_s(j,j) = \left\|\overline{\mathbf{G}}(:,j)^T\overline{\mathbf{Y}}_s\right\|_2^2 \tag{3.1.30}$$

The MSP weigh matrix, W_{MSP}, is constructed as

$$W_{\text{MSP}}(j,j) = 1 - D_s(j) \tag{3.1.31}$$

leaving the MSP estimator as follows,

$$\hat{\mathbf{S}}_{\text{MSP}} = \left[W_{\text{MSP}}\right]^{-2}\mathbf{G}^T\left(\mathbf{G}\left[W_{\text{MSP}}\right]^{-2}\mathbf{G}^T + \lambda\mathbf{I}_m\right)^{-1}\mathbf{Y} \tag{3.1.32}$$

3.1.6 Mixed Norm Estimate (MxNE)

This algorithm, proposed by Gramfort [9], incorporates the assumption that the source patches are sparse in space and smooth in time. This assumption is enforced by solving the following optimization problem.

$$\hat{\mathbf{S}}_{\text{MxNE}} = \arg\min_{\mathbf{S}} \; \|\mathbf{G}\,\mathbf{S} - \mathbf{Y}\|_F^2 + \lambda \; \|\mathbf{S}\|_{2,1} \tag{3.1.33}$$

with $\|\mathbf{S}\|_{2,1}$ the $\ell_{2,1}$ norm of $\mathbf{S}$. In general, the bi-level $\ell_{p,q}$ norm is defined as

$$\|A\|_{p,q} = \left(\sum_{m=1}^{M} \left(\sum_{n=1}^{N} |A(m,n)|^p \right)^{q/p} \right)^{1/q} \tag{3.1.34}$$

The MxNE estimator does not admit a closed-form solution but can be computed using the Alternating Direction Method of Multipliers (ADMM) or similar algorithms. A thorough description of ADMM is beyond the scope of this chapter, and the interested reader should refer to the book by Boyd [2].

3.2 Methods for parameter selection

From all the estimation methods described in the previous section, all of them converge in a general form

$$\hat{\mathbf{S}}_\Theta = F\left(\mathbf{Y}, \mathbf{G}; \Theta\right) \tag{3.2.1}$$

with Θ a set of parameters that must be selected beforehand. One particular case, referred to as linear estimator, occurs when the estimator can be written as

$$\hat{\mathbf{S}}_\Theta = \mathbf{K}_\Theta \, \mathbf{Y} \tag{3.2.2}$$

with $\mathbf{K}_\Theta \in \mathbb{R}^{N\times N}$ referred to as the Wiener kernel, or inversion kernel, which is approximately a right inverse of $\mathbf{G}$. The collection of liner methods includes wMNE, FOCUSS, LORETA and their derivates, and MSP, among others.

On the other hand, methods like MCE, MxNE, SiSSY, etc, are not linear and can't be assigned to a Wiener kernel.

This section reviews some heuristics for selecting the parameter set Θ. The wMNE estimator is a reference model since it depends on a single parameter. However, the heuristics can be generalized to select multiple parameters by finding the minima/maxima of functions of multiple variables. Such a generalization is not made explicit.

3.2.1 L-curve

This heuristic incorporates the idea that the parameter λ balances the factors $\|\mathbf{S}\|_F^2$ and $\|\mathbf{G}\,\mathbf{S}-\mathbf{Y}\|_F^2$, trying to minimize both of them evenly [13]. To do that, define the quantities

$$\rho(\lambda) = \left\|\mathbf{G}\,\hat{\mathbf{S}}_\lambda - \mathbf{Y}\right\|_F^2 \tag{3.2.3}$$

$$\eta(\lambda) = \left\|\hat{\mathbf{S}}_\lambda\right\|_F^2 \tag{3.2.4}$$

The L-curve denominations come from the graph of the parametric curve $(\rho(\lambda), \eta(\lambda))$, which resembles the letter 'L' in some cases. This graph is typically displayed in log-log scale, i.e., the parametric curve $(\log(\rho(\lambda)), \log(\eta(\lambda)))$.

According to this criterion, the optimal value for λ is the one on the point of maximal curvature of the parametric curve, often referred to as the 'elbow' of the curve. Look at the figure 3.1 for visual reference.

It is important to mention that the selection of parameters using the L-curve criterion depends on visual perception, and it may be affected by the precision with which the curve is constructed. Without prior knowledge, η and ρ shall be computed for several potential values of λ. Exploring the obtained parametric curve visually allows us to identify better the range where the optimal values may be found.

3.2.2 Generalized Cross-Validation

This heuristic is based on the principle of leave-one-out cross-validation and some approximations that make it computationally efficient. Although this method was proposed by Golub in 1979 [8], it remains popular for methods that admit an inversion kernel.

Following the leave-one-out principle, we consider $\mathbf{Y}^{[m]} \in \mathbb{R}(M-1) \times T$ and $\mathbf{G}^{[m]} \in \mathbb{R}^{(M-1)\times N}$ with the m-th sensor missing,

$$\mathbf{Y}^{[m]} = \begin{bmatrix} \mathbf{Y}(1,:) \\ \vdots \\ \mathbf{Y}(m-1,:) \\ \mathbf{Y}(m+1,:) \\ \vdots \\ \mathbf{Y}(M,:) \end{bmatrix}, \quad \mathbf{G}^{[m]} = \begin{bmatrix} \mathbf{G}(1,:) \\ \vdots \\ \mathbf{G}(m-1,:) \\ \mathbf{G}(m+1,:) \\ \vdots \\ \mathbf{G}(M,:) \end{bmatrix} \tag{3.2.5}$$

Based on this incomplete information, we construct a restricted Wiener kernel, $K_\lambda^{[m]} \in \mathbb{R}^{N\times(M-1)}$, and a restricted estimation, $\mathbf{S}_\lambda^{[m]} \in \mathbb{R}^{N\times T}$,

$$K_\lambda^{[m]} = \left(\mathbf{G}^{[m]}\right)^T \left(\mathbf{G}^{[m]} \left(\mathbf{G}^{[m]}\right)^T + \lambda \mathbf{I}_M\right)^{-1} \tag{3.2.6}$$

$$\mathbf{S}_\lambda^{[m]} = K_\lambda^{[m]} \mathbf{Y}_\lambda^{[m]} \tag{3.2.7}$$

The cross-validation is incorporated by estimating the measurements from the m-th with information from $\mathbf{S}_\lambda^{[m]}$ and then comparing it with the actual measurements. By performing this procedure over all the sensors and averaging, we obtain the following quantity

$$V(\lambda) = \frac{1}{M} \sum_{m=1}^{M} \left\| \left[\mathbf{G}\mathbf{S}_\lambda^{[m]}\right](m,:) - \mathbf{Y}(m,:) \right\|_F^2 \tag{3.2.8}$$

This quantity is approximated by the following

$$\mathrm{GCV}(\lambda) = \frac{M \, \|(\mathbf{G}\, K_\lambda - \mathbf{I}_M)\, \mathbf{Y}\|_F^2}{\mathrm{trace}^2\left(\mathbf{G}\, K_\lambda - \mathbf{I}_M\right)} \tag{3.2.9}$$

Although the definition of GCV is based on the leave-one-out cross-validation illustrated in equation (3.2.8), it is the expression on (3.2.9) the operative definition of GCV. The latter expression represents a balance between K_λ as a pseudo-inverse of $\mathbf{G}$ in general vs only for the current sensor data.

According to this criterion, the optimal value for λ is the one that minimizes the function in (3.2.9). In other words,

$$\lambda^\star_{\text{GCV}} = \arg\min_\lambda \ \text{GCV}(\lambda) \tag{3.2.10}$$

It is important to emphasize that the approximation for the GCV error in equation (3.2.9) can only be computed for linear methods; for instance, estimators based on the ℓ_1 norm or neural networks can't be tuned using GCV.

3.2.3 Composite Residual and Smoothing Operator (CRESO)

This heuristic incorporates the idea that the residual, $\|\mathbf{G}\,\mathbf{S}-\mathbf{Y}\|_F^2$, and the norm of the solution, $\mathbf{S}_F^2$, should be approximately equal [14]. To achieve that, the residual and norm are written as the following function,

$$\begin{aligned} C\left(\lambda\right) &= \frac{d}{d\lambda}\left[-\left\|\mathbf{G}\mathbf{S}_\lambda-\mathbf{Y}\right\|_2^2+\lambda\left\|\mathbf{S}_\lambda\right\|_2^2\right] \\ &= \frac{d}{d\lambda}\left[-\left\|\left(\mathbf{G}K_\lambda+\mathbf{I}_M\right)\mathbf{Y}\right\|_2^2-\lambda\left\|K_\lambda\mathbf{Y}\right\|_2^2\right] \end{aligned} \tag{3.2.11}$$

with $K_\lambda\mathbf{Y}$ the inversion kernel, understood explicitly as a function of λ. This function is called the Composite Residual and Smoothing Operator (CRESO).

According to this criterion, the optimal value for λ is the smallest solution to the equation $C(\lambda)=0$, or in other words,

$$\lambda^\star_{\text{CRESO}} = \min_\lambda\left\{\lambda\in\mathbb{R}_+ \mid C(\lambda)=0\right\}. \tag{3.2.12}$$

3.2.4 U-curve

This heuristic incorporates the idea that both the largest among the residual and the solution's norm should be small; this is achieved by minimizing their reciprocals [16]. In particular, the following function is constructed

$$U(\lambda) = \|\mathbf{G}\mathbf{S}_\lambda - \mathbf{Y}\|_2^{-2} + \|\mathbf{S}_\lambda\|_2^{-2} \tag{3.2.13}$$

Similar to the L-curve heuristic, the U-curve heuristic gets its denomination from the typical shape of the plot of $U(\lambda)$.

According to this criterion, the optimal value for λ is the global minimum of $U(\lambda)$. In other words,

$$\lambda^\star_{\text{U-curve}} = \arg\min_\lambda \; U(\lambda) \tag{3.2.14}$$

3.3 Performance metrics

3.3.1 Definitions of a Source Patch Location

In many early works, like the original paper for sLORETA [26], synthetic data for ESI methods used 'point sources' in the form of

$$\mathbf{S}_{\text{ref}}(n) = \begin{cases} 1, & \text{if } n = n^* \\ 0, & \text{otherwise} \end{cases} \tag{3.3.1}$$

with n^* the index of one dipole. Point sources are easy to evaluate in the forward model: we only need to select a column of $\mathbf{G}$ instead of performing a multiplication.

The location of a point source indexed by n^* is $\mathbf{r}_{n^*}$, a very straightforward definition. On the other hand, there is no single unified definition for more complicated source distributions, including multiple source patches.

In this work, the location of a source patch, either the estimation or the reference, is defined as the center of mass with respect to magnitude. This particular definition

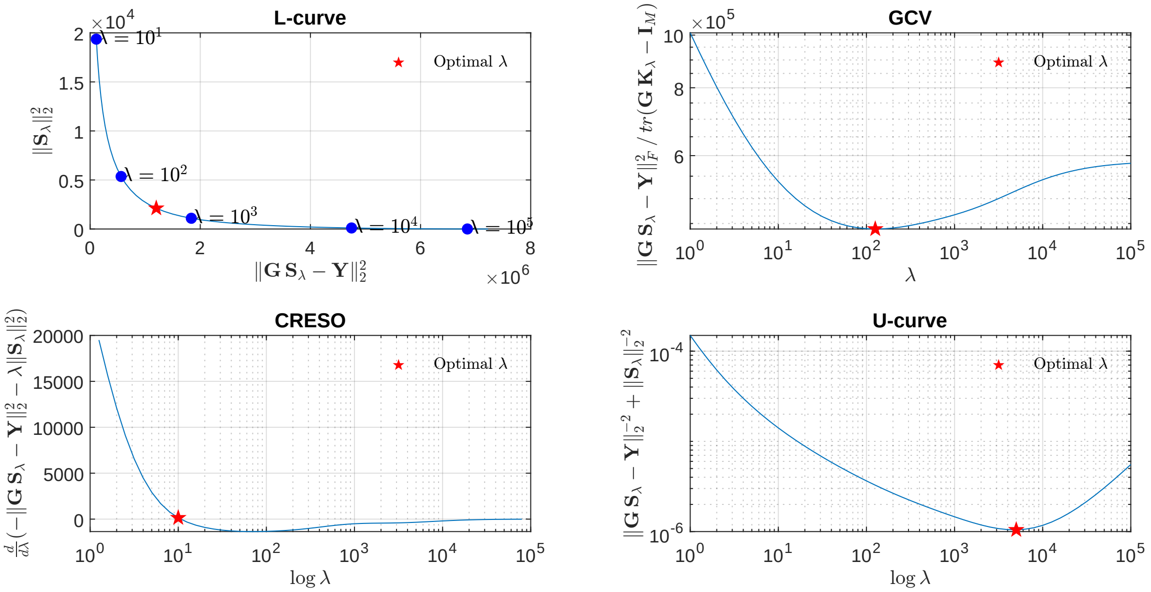

Figure 3.1. Selection of parameters using various criteria: L-curve, Generalized Cross-Validation (GCV), Composite Residual and Smoothing Operator (CRESO), and U-curve. Each criterion's quantity of interest is computed for many candidate values of λ, then the optimal value is decided. Data for these graphs was obtained from a single trial of synthetic data described in Chapter 4. Data $\mathbf{Y}$ is constructed with no-noise, inversion kernel $\mathbf{K}_\lambda$ and the source estimate $\mathbf{S}_\lambda = \mathbf{K}_\lambda \mathbf{Y}$ are computed using sLORETA .

can be traced back to the paper by Molins et al. [21], and the reasoning for its use is discussed in the following subsection.

In order to establish a formal definition, we first must consider the concept of a seed dipole. Informally, the idea of a seed dipole is that the source patch is constructed around it, and thus, the seed dipole is a proto-center for the source patch. This definition only works for sources with known locations; thus, it is unusable for estimated sources.

One key property of a seed dipole, which will be later exploited, is that it must be one of the most active dipoles in the patch, i.e.

$$\left\|\mathbf{S}\left(k_p^*\right)\right\| = \max\left\{\|\mathbf{S}\left(n\right)\| \mid n \in \mathcal{I}_p\right\} \tag{3.3.2}$$

In the case of estimated sources, the role of seed dipoles can be fulfilled by local maxima of the estimated magnitudes.

Let be $k_1^*, k_2^*, \ldots, k_P \in \{1, 2, \ldots, N\}$ the indexes of the seed dipoles. Formally, the p-th source patch, $\mathcal{I}_p$, is defined as the set of dipoles that have significant activity and are closer to the p-th seed dipole, k_p^* than to the other seed dipoles.

$$\mathcal{I}_p = \left\{ n \;\middle|\; \left\|\mathbf{r}_n - \mathbf{r}_{k_p^*}\right\|_2 < \left\|\mathbf{r}_n - \mathbf{r}_{k_j^*}\right\|_2 \text{ for } p \neq j, \; \|\mathbf{S}(n)\|_2 \geq \alpha \left\|\mathbf{S}\left(k_p^*\right)\right\|_2 \right\} \tag{3.3.3}$$

where $\alpha \in [0, 1]$ filters sources whose magnitude is very small; for instance, we may set $\alpha = 10\%$ as a cutoff for significance.

In the case of estimated dipoles, the number of source patches can be adjusted algorithmically by joining or breaking other source patches.

The location of the p-th source patch, $\mathbf{m}_p$, is defined as the center of mass of the dipoles in the source patch with respect to their magnitude,

$$\mathbf{m}_p = \frac{\displaystyle\sum_{n \in \mathcal{I}_p} \|\mathbf{S}(n)\|_2^2 \, \mathbf{r}_n}{\displaystyle\sum_{n \in \mathcal{I}_p} \|\mathbf{S}(n)\|_2^2} \tag{3.3.4}$$

Notice that this definition depends heavily on knowing the true locations of the seed dipoles, if any, so it can't be used with real data.

On the other side, the definition is trivially extended to the location of an estimated source patch,

$$\hat{\mathbf{m}}_p = \frac{\displaystyle\sum_{n \in \mathcal{I}_p} \left\|\hat{\mathbf{S}}(n)\right\|_2^2 \mathbf{r}_n}{\displaystyle\sum_{n \in \mathcal{I}_p} \left\|\hat{\mathbf{S}}(n)\right\|_2^2} \tag{3.3.5}$$

This definition is deficient when an ESI method does not recover all source patches, in which case the performance is overestimated.

3.3.1.1 Other Definitions

Although the definition presented is compatible with other definitions in the case of a single source patch, it is not the only one.

For instance, some authors have used the location of the seed dipole as the location of the source patch

$$\mathbf{m}_p = \mathbf{r}_{k_p^*} \tag{3.3.6}$$

which shouldn't be very different from the center of mass if the source patch is symmetric and centered around the seed. For example, these assumptions are broken for volumetric source patches near the cortex.

Some authors define the location as the local maximum magnitude

$$\mathbf{m}_p = \left\{ \mathbf{r}_k \;\middle|\; k = \arg\max_n \; \{\|\mathbf{S}(n)\| \mid n \in \mathcal{I}_p\} \right\} \tag{3.3.7}$$

which is also very similar if the source patch is symmetrical and decays monotonically far from the seed. This is not the case if $\|\mathbf{S}(n)\|$ is constant in the patch.

The source location based on maximal magnitude can be trivially extended to define the location of an estimated source patch

$$\hat{\mathbf{m}}_p = \left\{ \mathbf{r}_k \;\middle|\; k = \arg\max_n \; \left\{ \left\|\hat{\mathbf{S}}(n)\right\| \;\middle|\; n \in \mathcal{I}_p \right\} \right\} \tag{3.3.8}$$

This definition is not robust to noise if the true source is large in space.

3.3.2 Distance Localization Error (DLE)

Given the definitions of the true and estimated centers for the source patches, the Distance Localization Error is defined as

$$\mathrm{DLE} = \frac{1}{P} \sum_{p=1}^{P} \|\mathbf{m}_p - \hat{\mathbf{m}}_p\|_2 \tag{3.3.9}$$

with P the number of true source patches. This definition does not account for whether the reconstruction has more or fewer source patches than it should.

3.3.3 Spatial Dispersion (SD)

This metric, proposed by Molins et al [21], is defined as

$$\mathrm{SD} = \left(\frac{\sum_{p=1}^{P} \sum_{n \in \mathcal{I}_p} \left\| \hat{\mathbf{S}}(n) \right\|_2^2 \left\| \mathbf{r}_n - \mathbf{m}_p \right\|}{\sum_{n=1}^{N} \left\| \hat{\mathbf{S}}(n)) \right\|_2^2} \right)^{1/2} \tag{3.3.10}$$

This quantity can be understood as a variance in space for the source location. A small spatial dispersion is desirable.

It is possible to replace $\hat{\mathbf{S}}$ for $\mathbf{S}$ in order to compute a 'baseline' spatial dispersion for source patches, which is especially useful if they are large in space.

3.3.4 Relative Mean Square Error (rMSE)

Also referred to as relative Reconstruction Error, this metric represents the estimation error for the dipole magnitudes. It is defined as

$$\mathrm{rMSE} = \frac{\left\| \mathbf{S} - \hat{\mathbf{S}} \right\|_2}{\left\| \mathbf{S} \right\|_2} \tag{3.3.11}$$

3.3.5 Area Under Receiver-Operator Curve (AUROC)

It is possible to interpret the estimation from Electric Source Imaging as a binary classification problem over the dipoles; this is done by considering active/non-active classes. This interpretation is consistent with the usability of ESI data to identify brain regions that are 'active' during specific cognitive or pathological processes.

A natural assumption in this context, already discussed in Chapter 2, is that dipoles with large magnitude are related to brain regions displaying coherent activity.

We consider the target labels as reference, $\mathcal{R} : [0,1] \to \{0,1\}^N$ and the inferred labels as estimations, $\mathcal{E} : [0,1] \to \{0,1\}^N$. The assumption of active dipoles/regions is modeled by constructing the labels as follows,

$$\mathcal{R}_\beta(n) = \begin{cases} 1, & \text{if } \left\|\mathbf{S}(n)\right\|_2 \geq \beta \max_{n'} \left\{\left\|\mathbf{S}(n')\right\|_2\right\} \\ 0, & \text{otherwise} \end{cases} \tag{3.3.12}$$

$$\mathcal{E}_\beta(n) = \begin{cases} 1, & \text{if } \left\|\hat{\mathbf{S}}(n)\right\|_2 \geq \beta \max_{n'} \left\{\left\|\hat{\mathbf{S}}(n')\right\|_2\right\} \\ 0, & \text{otherwise} \end{cases} \tag{3.3.13}$$

The classical interpretation of the Receiver-Operator Curve requires that we affix the target labels, $\mathcal{R}_\beta$, which can be achieved in the context of ESI by selecting a particular value $\beta = \beta_0 \in [0,1]$. See figure 3.2.A for a quick example of creating the target labels using $\beta_0 = 0.5$.

Other authors, like Molins et al. [21], proposed to define $\mathcal{R}_\beta$ dynamically, changing with $\mathcal{E}_\beta$ as β changes. This idea is not used in this work because it is unstable for bad estimations, leading to unintuitive results.

With this notation, we define True Positives, False Positives, True Negatives, and False Negatives as follows

$$\mathrm{TP}(\beta) = \sum_{n=1}^{N} \left[\mathcal{E}_\beta\left(n\right)\right]\left[\mathcal{R}_\beta\left(n\right)\right] \tag{3.3.14}$$

$$\mathrm{FP}(\beta) = \sum_{n=1}^{N} \left[1 - \mathcal{E}_\beta\left(n\right)\right]\left[\mathcal{R}_\beta\left(n\right)\right] \tag{3.3.15}$$

$$\mathrm{TN}(\beta) = \sum_{n=1}^{N} \left[1 - \mathcal{E}_\beta\left(n\right)\right]\left[1 - \mathcal{R}_\beta\left(n\right)\right] \tag{3.3.16}$$

$$\mathrm{FN}(\beta) = \sum_{n=1}^{N} \left[1 - \mathcal{E}_\beta\left(n\right)\right]\left[\mathcal{R}_\beta\left(n\right)\right] \tag{3.3.17}$$

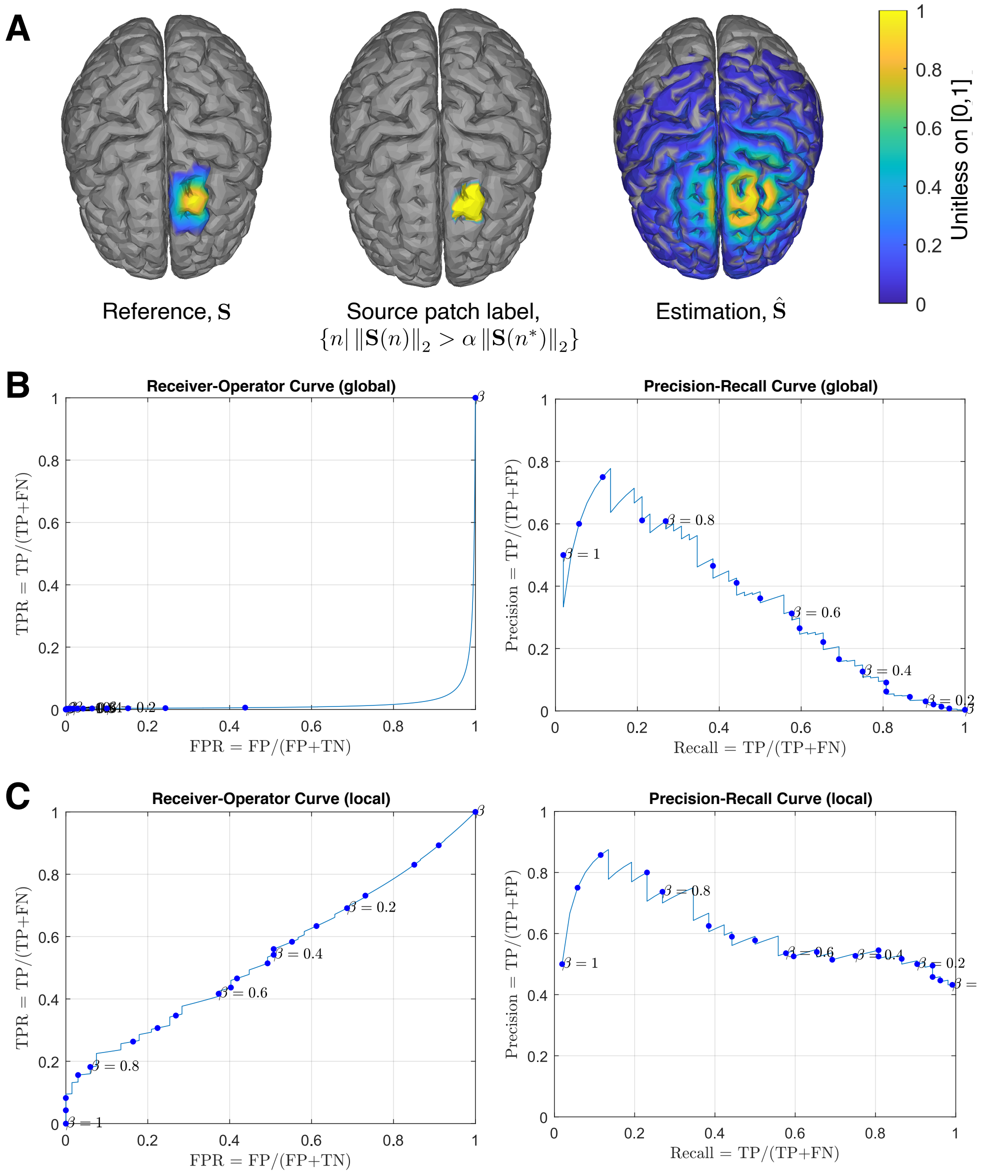

Figure 3.2. A. The true class labels are defined from the true sources, $\mathbf{S}$, while the estimated labels are constructed from the ESI estimation, $\hat{\mathbf{S}}$. Notice the severe class imbalance between yellow and colorless (gray). B. Construction of Receiver Operator Curve (ROC) and Precision-Recall Curve; the area under them is called AUROC and Average Precision, respectively. C. To alleviate the class imbalance, only the dipoles within a certain distance of the patch center are considered; such distance is set to 25.7 mm to consider approximately 50% of dipoles in each category.
Data for these graphs was obtained from a single trial of synthetic data described in Chapter 4. Reference source distribution $\mathbf{S}$ was constructed with no-noise and the source estimate $\hat{\mathbf{S}}$ was computed using sLORETA .

The Receiver-Operator Curve (ROC) itself is the parametric curve defined by $(\mathrm{FPR}\left(\beta\right), \mathrm{TPR}\left(\beta\right))$, where the False Positive Rate (FPR) and True Positive Rate (TPR) are defined as

$$\mathrm{FPR}(\beta) = \frac{\mathrm{FP}\left(\beta\right)}{\mathrm{FP}\left(\beta\right) + \mathrm{TN}\left(\beta\right)} \tag{3.3.18}$$

$$\mathrm{TPR}(\beta) = \frac{\mathrm{TP}\left(\beta\right)}{\mathrm{TP}\left(\beta\right) + \mathrm{FN}\left(\beta\right)} \tag{3.3.19}$$

The interpretation of the Receiver-Operator Curve is that each selection of β balances the ratios of True Positives and False Negatives:

- Large $\beta \implies$ labels will be strict, with very few False Positives but many False Negatives.
- Small $\beta \implies$ labels will be permissive, with very few False Negatives but many False Positives.

In general, extreme values of β will lead to the endpoints (0,0) and (1,1) of the ROC, and a somewhat smooth trajectory will be observed as β increases.

A classifier for which $\mathrm{TPR}(\beta) \approx 1$ and $\mathrm{FPR}(\beta) \approx 0$ for all $beta$'s is preferable whenever possible. Visually, this condition means that the ROC is 'close to the top.'

This concept is formalized by taking the area under the ROC (AUROC): the area inside $[0, 1]^2 \subset \mathbb{R}^2$ below the ROC curve. In practice, the AUROC can be computed numerically using the trapezoid rule as follows

$$\mathrm{AUROC} = \frac{1}{2} \sum_i \left(\mathrm{FPR}\left(\beta_i\right) - \mathrm{FPR}\left(\beta_{i-1}\right)\right) \left(\mathrm{TPR}\left(\beta_i\right) + \mathrm{TPR}\left(\beta_{i-1}\right)\right) \tag{3.3.20}$$

with β_i the finite number of values that β can take.

With this definition at hand, we can rewrite the most desirable scenario where $\mathrm{AUROC} \approx 1$.

3.3.5.1 Local AUROC

AUROC can be biased on the presence of severe class imbalance when a class has many more elements than the other. This phenomenon is represented in figure 3.2: panel A shows how the reference labels created using β_0 only account for a tiny fraction of dipoles available in the model, while panel B shows that the ROC is unusually flat, regardless of the quality of the ESI estimation.

The class imbalance can't be avoided on the ESI since the source patches are globally sparse and compact –or at least we assume that.

One option to alleviate this issue is to consider only a small selection of dipoles for classification. In other words, construct local estimated labels as

$$\mathcal{E}_{\beta}^{(d)}(n) = \begin{cases} 1, & \text{if } \left\|\hat{\mathbf{S}}(n)\right\|_2 \geq \beta \max_{n'} \left\{\left\|\hat{\mathbf{S}}(n')\right\|_2\right\} \text{ and } \|\mathbf{r}_n - \mathbf{m}_p\| < d \\ 0, & \text{otherwise} \end{cases} \tag{3.3.21}$$

with $\mathbf{r}_n$ the location of the n-th dipole, $\mathbf{m}_*$ the location of the center of the closest source patch, and $d > 0$ a parameter controlling how many dipoles are to be considered.

We can define local target labels analogously. To compute local true/false positives/negatives, we must consider only 'local' dipoles. For example,

$$\mathrm{TP}^{(d)}(\beta) = \sum_{\substack{n\in\{1,\ldots,N\} \\ \|\mathbf{r}_n - \mathbf{m}_p\| < d}} \left[\mathcal{E}_{\beta}^{(d)}(n)\right] \left[\mathcal{R}_{\beta}^{(d)}(n)\right] \tag{3.3.22}$$

This construction leads to a local AUROC. In order to satisfy it's purpose, d must be selected so that $\frac{1}{N}\sum_{n=1}^{N} \mathcal{R}_{\beta_0}^{(d)}(n) \approx 0.5$ for a suitable β_0.

3.3.6 Average Precision

A different approach to class imbalance is using a different quantity with properties similar to AUROC. Average Precision, also known as Area Under Precision-

Recall Curve, is based on the parametric curve $(\mathrm{Rec}\,(\beta)\,, \mathrm{Pre}\,(\beta))$, where the Precision (Pre) and Rec (Rec) are defined as

$$\mathrm{Pre}\,(\beta) = \frac{\mathrm{TP}\,(\beta)}{\mathrm{TP}\,(\beta) + \mathrm{FP}\,(\beta)} \tag{3.3.23}$$

$$\mathrm{Rec}\,(\beta) = \frac{\mathrm{TP}\,(\beta)}{\mathrm{TP}\,(\beta) + \mathrm{FP}\,(\beta)} \tag{3.3.24}$$

where the true/false positive/negatives are defined as in equations 3.3.14, 3.3.15, 3.3.16, and 3.3.17.

Similar to AUROC, Average Precision can be computed numerically as follows

$$\text{Avg. Pre.} = \frac{1}{2}\sum_{i} \left(\mathrm{FPR}\,(\beta_i) - \mathrm{FPR}\,(\beta_{i-1})\right)\left(\mathrm{TPR}\,(\beta_i) + \mathrm{TPR}\,(\beta_{i-1})\right) \tag{3.3.25}$$

Average Precision ranges from 0 to 1, with 1 being preferable.

Although the definition of Average Precision is very similar to that of AUROC, it is different in many aspects. For example, the Precision-Recall Curve is somewhat decreasing but displays some sharp zigzags. When the recall is approximately 1, the precision is roughly equal to the ratio of elements between classes.

Similar to AUROC, defining a ‘local’ Average Precision is possible. In figure 3.2, we can observe that the Precision-Recall Curve is quite similar to its global counterpart, or at least when compared against the local and global ROC. Informally, we may observe that Average Precision is more robust to the class imbalance.

CHAPTER 4

NUMERICAL EXPERIMENTS

Following the model equation from the forward model, $\mathbf{Y} = \mathbf{GS} + \varepsilon$, each trial of synthetic data is constructed following these steps:

1. Construct the leadfield matrix $\mathbf{G}$ based on pertinent anatomical data.
2. Select a 'seed' dipole at random to determine the location of the source patch.
3. Construct the source distribution, $\mathbf{S}$, based on the source patch.
4. Construct a noiseless version of the EEG measurements, $\mathbf{Y}_0 = \mathbf{GS}$.
5. Construct some additive noise, ε, in order to achieve a prescribed Signal-to-Noise Ratio ($\mathrm{SNR}_{\mathrm{dB}}$) measured in decibels.
6. Construct the EEG measurements, $\mathbf{Y} = \mathbf{Y}_0 + \varepsilon$.

After each trial of synthetic data is created, ESI is performed to estimate the source distribution by using some of the methods described in Chapter 3. The performance of these methods is then measured using the performance metrics described in the same chapter.

It is important to mention that whenever an ESI method depended on a parameter, its value was determined using either Generalized Cross-Validation (GCV) or U-curve criterion when GCV was not possible.

4.1 Forward Model

For the construction of synthetic data, anatomy data was taken from the ICBM 152 template [6]. Segmentation and surface extraction were omitted since these surfaces are provided with the template. These surfaces are displayed in figure 4.1.

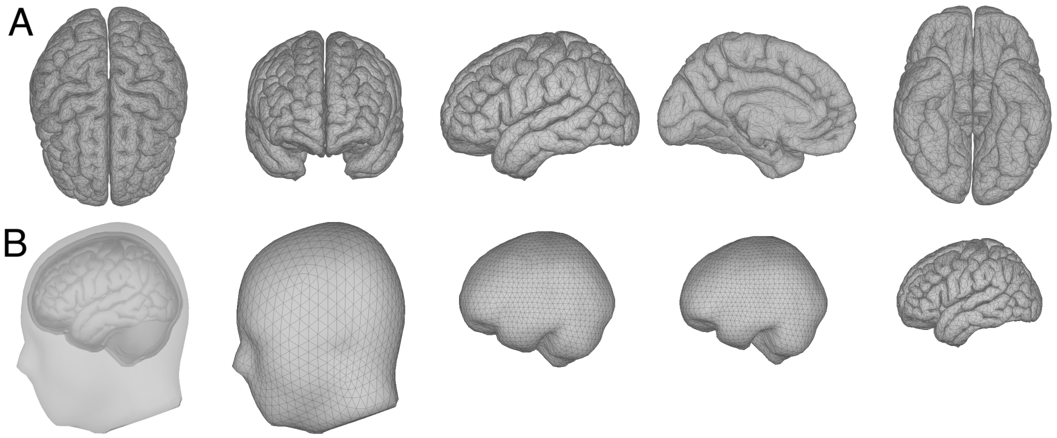

Figure 4.1. A. Cortex surface for the ICBM 152 template, consisting of 15,002 vertices. Views from top, front, left, right (interior), bottom. B. Triangulated surfaces used for the 4-sphere forward model: cortex, inner skull, outer skull, head.

Electrodes were placed according to the 10-10 International System; a total of $M = 90$ electrodes were considered. See either figure 4.2 for visual reference or [15] for a more precise description of the electrode placement protocol. No reference electrode is considered for the synthetic data; the sensor data will be re-referenced to the average.

The positions of the distributed dipoles were set along the brain cortex, and their orientation was orthogonal to the cortex. A total of $N = 15,002$ electrodes were considered.

For illustrative purposes, the duration of the recordings was set to a total of $T = 1$ time stamps.

4.1.1 Source Distribution

The source distribution, $\mathbf{S}$, is constructed by considering one small source patch; the dipoles in the source patch are selected as those being close to a seed dipole,

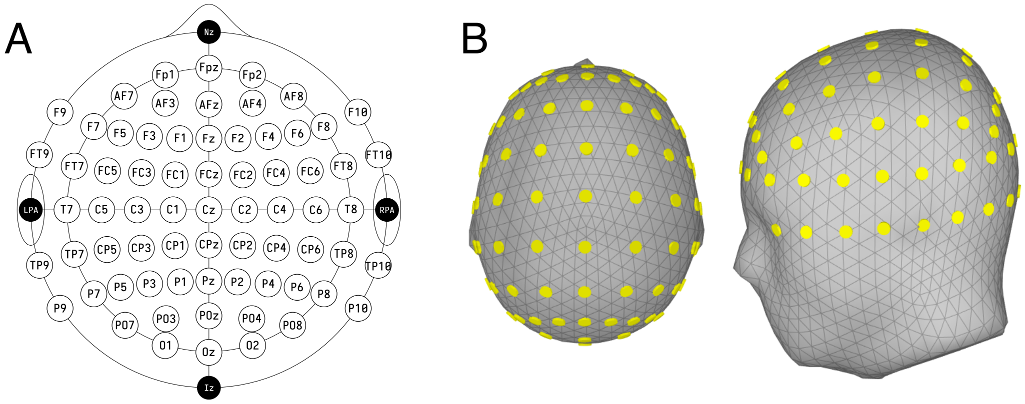

Figure 4.2. Electrode Placement according to the 10-10 International System. A. Schematic with the electrode labels. B. Electrodes at the ICBM152 head.

indexed by n^*. The seed dipole was selected randomly for each trial; a total of 30 trials were simulated.

After selecting the seed dipole, indexed by n^*, $\mathbf{S}$ is constructed as

$$\mathbf{S}(n) = f\left(\left\|\mathbf{r}_n - \mathbf{r}_{n^*}\right\| / \kappa\right) \tag{4.1.1}$$

with $\mathbf{r}_n$ the location of the n-t dipole, $f : \mathbb{R}_+ \to \mathbb{R}$ a decreasing function, and $\kappa > 0$ a scaling parameters.

For illustrative purposes, four different options for f are considered:

$$\text{Square} \qquad f_{\text{sq}}(x) = \begin{cases} 1, & \text{if } x \leq 1 \\ 0, & \text{otherwise.} \end{cases} \tag{4.1.2}$$

$$\text{Gaussian} \qquad f_{\text{Gauss}}(x) = e^{-x^2/\kappa} \tag{4.1.3}$$

$$\text{Exponential} \qquad f_{\text{exp}}(x) = e^{-x/\kappa} \tag{4.1.4}$$

$$\text{Polynomial} \qquad f_{\text{pol}}(x) = \sqrt{\max\left\{1 - x^2, 0\right\}} \tag{4.1.5}$$

Apart from point sources, source patches created using f_{sq} are relatively common; their sharp edges facilitate determining the size of the source patch without ambiguity. Some other papers have used f_{Gauss} to obtain smooth source patches in space.

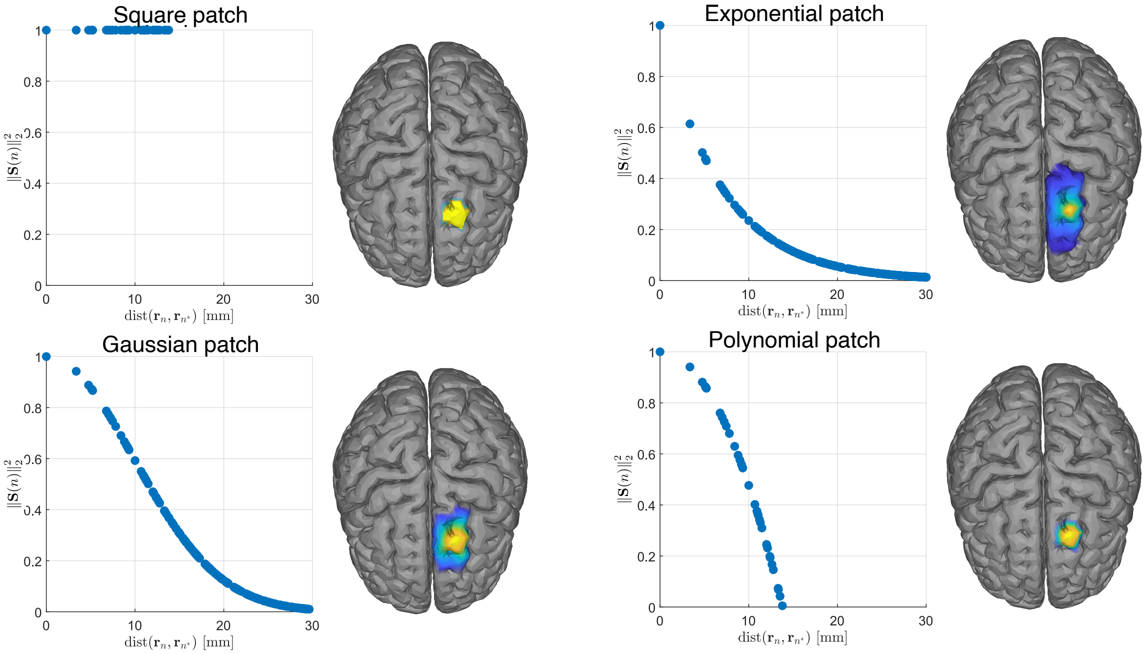

Figure 4.3. Different source patches used to generate synthetic data.

Functions f_{exp} and f_{pol} are introduced here to produce source patches with intermediate properties. The source patches constructed using f_{pol} have a finite support and a smooth peak, while the ones built using f_{exp} present spatial smoothness, a sharp peak, and a 'heavy tail' in space. See figure 4.3 for a visual example of these properties.

For the synthetic data in this chapter, we set $\kappa = 12.6$ mm in order to obtain source patches with a surface area of approximately 5 cm^2.

4.1.2 Noise

For simplicity, only additive white noise is added to the sensors, i.e., we use

$$\varepsilon(:) \sim \mathcal{N}\left(0, \operatorname{diag}\left(\sigma_1, \ldots, \sigma_M\right)^2\right) \tag{4.1.6}$$

with $\sigma_m \geq 0$ for $m = 1, \ldots, M$ the noise variances for each EEG electrode.

Under these conditions the noisy recordings, $\mathbf{Y}$, have the following distribution

$$(\mathbf{Y} \mid \mathbf{S}) = \mathcal{N}\left(\mathbf{G}\,\mathbf{S}, \operatorname{diag}\left(\sigma_1, \ldots, \sigma_M\right)^2\right) \tag{4.1.7}$$

The channel SNR (in decibels) for the m-th sensor is given by

$$\mathrm{SNR}_m = 10\,\log_{10}\left(\frac{\left(\left[\mathbf{G}\,\mathbf{S}\right](m)\right)^2}{\sigma_m^2}\right) \tag{4.1.8}$$

With a prescribed SNR value, the parameters σ_m are selected as follows

$$\sigma_m^2 = 10^{-SNR/10}\left[\mathbf{G}\,\mathbf{S}\right](m) \tag{4.1.9}$$

Based on informal observations, we consider an SNR level of 20 dB a low-noise condition. Notice that the condition of lack of noise is attached to $\sigma_m = 0$ and $\mathrm{SNR} = \infty$.

4.2 Results

Different ESI algorithms, described in Chapter 3, were tested using this data: Weighted Minimal Norm, Multivariate Source Prelocalisation, sLORETA, and SISSY. These algorithms were selected due to their presence in the literature and because of their small computational time.

Notice that only one ℓ_1-norm-based method is accounted for. Although it would be optimal to consider many more methods, the large number of trials reduced the number of methods that can be effectively compared. The priority of this work is the ESI method presented in Chapter 5.

The solutions produced by these methods, referred to as 'solvers' in the figures, are evaluated using the quantities described in Chapter 3. Figure 4.4 displays some of these results.

Before assessing the results, it is worth reviewing the assumption of depth weighting: dipoles closer to the cortex tend to be over-represented on the minimal-

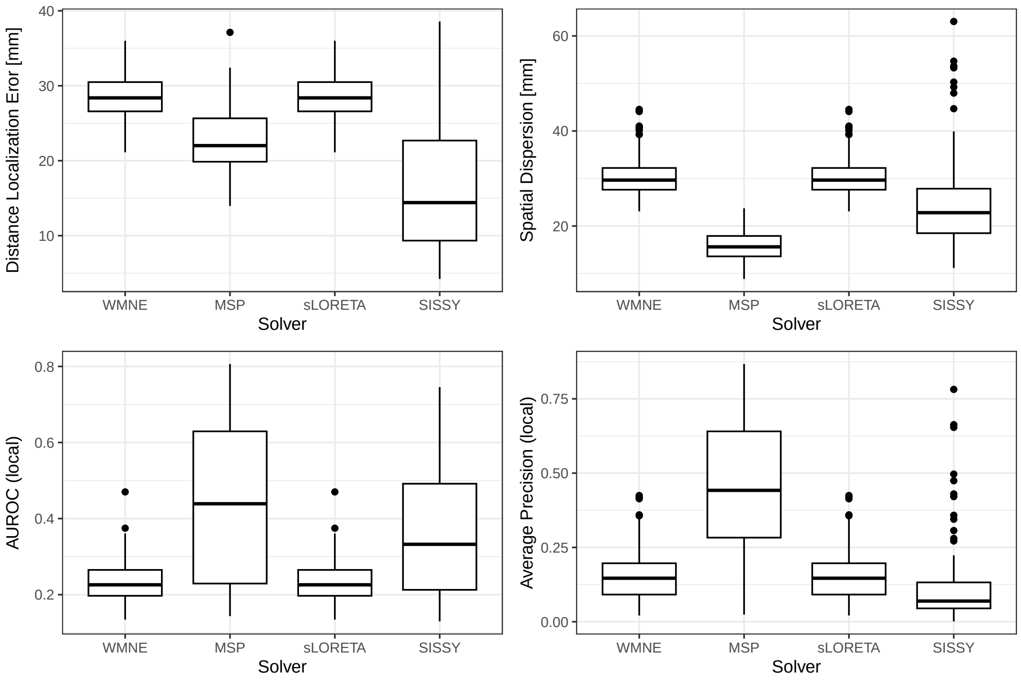

Figure 4.4. Performance metrics for some linear estimators for the synthetic data described in the text: the ‘Human’ dataset. The dataset contains 30 trials per shape of source patch, and the SNR is set to 30 dB.

norm estimates. In previous sections, we described some mechanisms that could alleviate this effect. It is natural to verify whether those mechanisms work or, on the contrary, whether an estimate’s performance depends on the depth of the original source patch.

Table 4.1 shows the Pearson correlation of the performance metrics. We can observe that all the metrics seem to be correlated with each other, but none of them correlate with the depth of the source patch.

4.2.1 Effect of Shape of Source Patch

A large group of factors affects the wide variability observed in the results. For instance, figure 4.5 shows the results for only the SYSSY solver. We can observe that

	SD	AUROC	AP	Depth
DLE	0.403	-0.525	0.036	-0.049
SD		-0.586	-0.464	0.128
AUROC			0.524	-0.114
AP				-0.094

Table 4.1. Pearson correlation between performance metrics with the 'Human' dataset.

some source patch shapes are recovered more accurately than others. For a reference of the shapes of source patches used, see figure 4.3.

This effect was expected since the assumption of minimal norm tends to favor estimations with few high-norm dipoles and many low-norm dipoles. In other words, minimal-norm estimates tend to resemble point sources; if the original source resembled a point source, then the reconstruction was accurate.

Although this phenomenon was observed for all the considered estimators, the effect is exaggerated for SISSY due to its formulation.

4.3 Effect of Noise

Intuitively, we may expect that the performance of the solvers degrade in the presence of stronger noise. In figure 4.6, we may observe that for all the solvers considered, an SNR below 20 dB impairs the performance of the solvers drastically.

Performing the same observations over different shapes of source patches doesn't change this conclusion significantly.

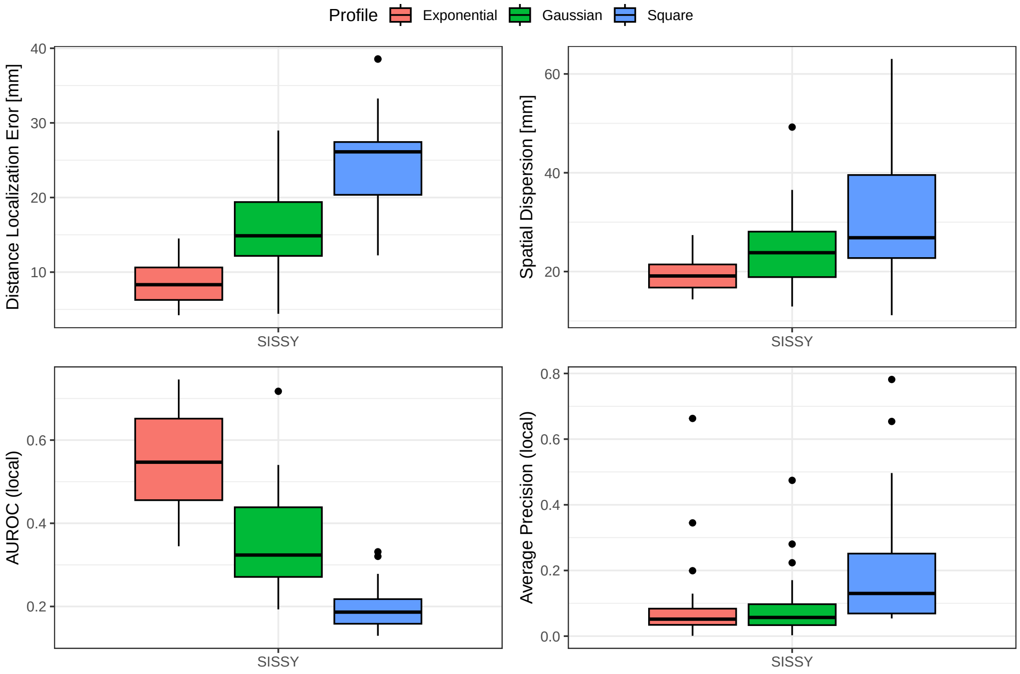

Figure 4.5. Changes in performance metrics for the SISSY estimator with the 'Human' dataset when different shapes of source patches are used.

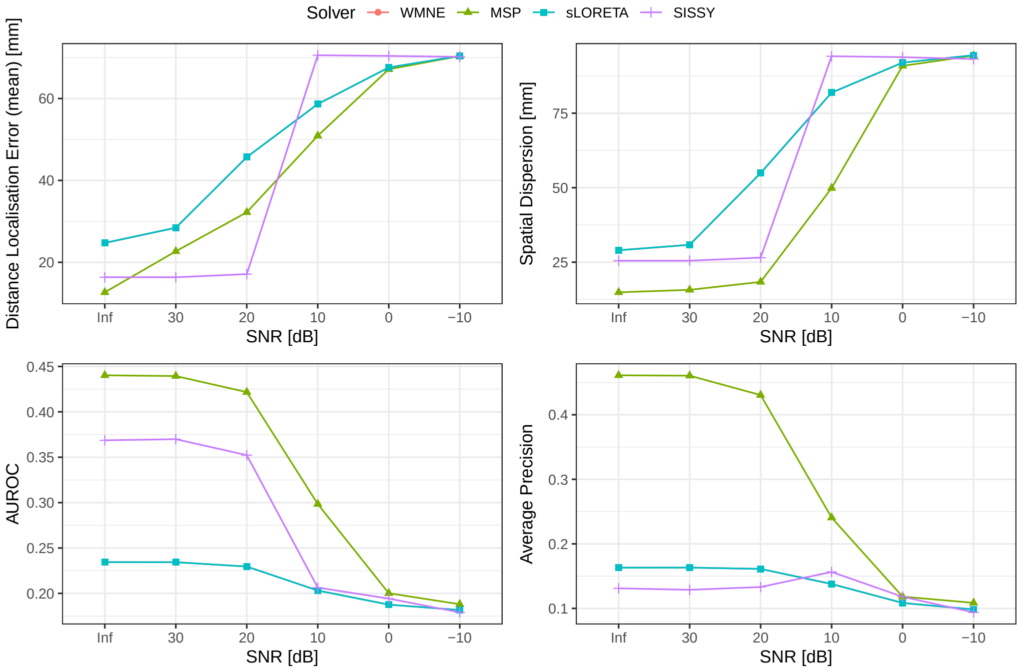

Figure 4.6. Changes in performance metrics for some estimators with the 'Human' dataset as the SNR decreases.

CHAPTER 5

NOVEL METHOD USING ANATOMICAL PRIOR INFORMATION

When considering using more than one data modality for analysis, there are many ways to combine them. The asymmetric multi-modal data analysis paradigm consists of selecting a tool designed for one data modality and incorporating information from all the other data modalities into the analysis. This paradigm fits the purpose of performing the Electronic Source Imaging while incorporating information from additional data modalities.

In particular, we consider the case in which a physical region is known to produce a higher background electrical activity. The location and extent of this region are observed using some imaging techniques independent of EEG.

The motivation for this particular setup comes from a specific experimental setup in which an ischemic stroke is induced, and later, the affected area is determined using histochemical analysis.

5.1 Model Assumptions

The relationship between the recordings from surface electrodes, $\mathbf{Y}$, and the magnitudes of equivalent distributed dipoles located inside the brain, $\mathbf{S}$, is given by the following equation

$$\mathbf{Y} = \mathbf{G}\left(\mathbf{S} + \varepsilon\right), \tag{5.1.1}$$

with $\mathbf{Y} \in \mathbb{R}^{M\times 1}$, $\mathbf{S} \in \mathbb{R}^{3N\times 1}$, $\mathbf{G} \in \mathbb{R}^{3N\times M}$, and $\varepsilon \in \mathbb{R}^{N\times 1}$. This model was described in detail in Chapter 2, including the interpretation and derivation of the leadfield matrix $\mathbf{G}$. Notice that this model considers only internal noise.

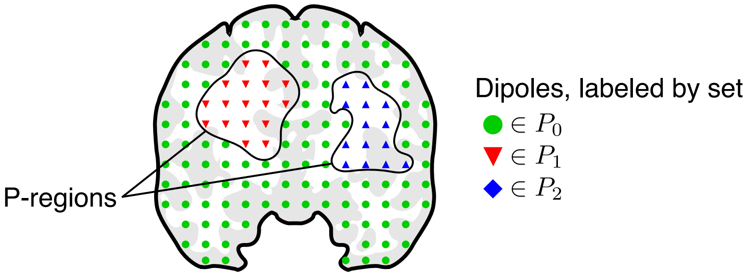

Figure 5.1. Construction of labels based on P-regions, representing regions when a particular symptom, independent of electrophysiological measurements, was observed..

Ideally, the extra information provided by the additional data modality should allow us to identify an anatomical region exhibiting pathological behavior; for simplicity, we refer to this region as a P-region. For generality, we consider the possibility of multiple disjoint P-regions, say K with $1 \leq K < \infty$. P-regions are labeled $P_1, P_2, \ldots, P_K$, with P_0 the dipoles outside all P-regions. For notation, define

$$P_k = \{n \in \{1, 2, \ldots, N\} \mid n\text{-th dipole is in the } k\text{-th region}\} \tag{5.1.2}$$

As a remark, the P-regions are defined independently of the dipoles; each P_k is the set of dipoles inside a specific P-region. If a different set of dipoles is selected for the same subject and pathology data, the P-regions will change but will be constructed in the same anatomical space.

For ease of handling in the model, the P-regions regions are encoded using the matrix $\mathbf{P} \in \{0,1\}^{N \times K}$ defined as

$$\mathbf{P}(n,k) = \begin{cases} 1, & \text{if } n \in P_k \\ 0, & \text{otherwise.} \end{cases} \tag{5.1.3}$$

The relation between the current density, $\mathbf{S}$, and the P-regions is incorporated into the model by adjusting the covariance of $\mathbf{S}$ depending on which P-region it is located. To be specific, consider,

$$\mathrm{Cov}\left(\mathbf{S}\left(n_1\right), \mathbf{S}\left(n_2\right)\right) = \begin{cases} \gamma_k, & \text{if } n_1, n_2 \in P_k, \\ 0, & \text{otherwise} \end{cases} \tag{5.1.4}$$

with $0 < \gamma_0 \ll \gamma_k$ for $1 \leq k \leq K$. This guarantees that P-regions have synchronized activity that is larger in magnitude than the background activity.

The assumption of a perfect synchrony is introduced in order to simplify the model. We can separate the source, $\mathbf{S}$, into a regionally-synchronized activity, $\mathbf{R}$, and a background noise, $\mathbf{B}$, as

$$\mathbf{S} + \varepsilon = \mathbf{B} + \mathbf{PR} = \mathbf{B} + \sum_{k=1}^{K} \mathbf{P}(:, k)\, \mathbf{R}(k) \tag{5.1.5}$$

with $\mathbf{B} \in \mathbb{R}^{N \times 1}, \mathbf{R} \in \mathbb{R}^{K \times 1}$ independent with each other. These assumptions about P-regions can be combined with Gaussian assumptions, leading to

$$\mathbf{B} \sim \mathcal{N}\left(0, \gamma_0 \mathbf{I}_N\right) \tag{5.1.6}$$

$$\mathbf{R} \sim \mathcal{N}\left(0, \left[\mathrm{diag}\left(\gamma_1, \gamma_2, \ldots, \gamma_K\right) - \gamma_0 \mathbf{I}_K\right]\right) \tag{5.1.7}$$

5.2 Proposed Estimator

Given the assumptions described in the previous section, we propose constructing a linear estimator for S similar to those presented in Chapter 3. This choice follows the expectation of short computation times and memory requirements. We expect that the additional information from P-regions will decrease spatial dispersion while not significantly increasing computational speed.

The estimator of $\mathbf{S}$ will be constructed as a Maximum A Posteriori estimator, maximizing the log probability of $\mathbf{S} \mid \mathbf{Y}$, as

$$\begin{aligned}\hat{\mathbf{S}} &= \arg\max_{\mathbf{S}} \quad \log\left(\text{Prob}\left(\mathbf{S} \mid \mathbf{Y}\right)\right) \\ &= \arg\max_{\mathbf{S}} \quad \log\left(\text{Prob}\left(\mathbf{Y} \mid \mathbf{S}\right)\, \text{Prob}\left(\mathbf{S}\right)\right) \qquad (5.2.1)\end{aligned}$$

the structure of P-regions is incorporated by using the following decomposition

$$\mathbf{S} = \mathbf{B} + \mathbf{P}\mathbf{R} = \mathbf{B} + \sum_{k=1}^{K} \mathbf{P}(:,k)\mathbf{R}(k) \qquad (5.2.2)$$

and the model equation $\mathbf{Y} = \mathbf{G}\mathbf{S}$ is incorporated as a constraint, resulting in the following optimization problem

$$\begin{aligned}\left\{\hat{\mathbf{B}}, \hat{\mathbf{R}}\right\} &= \arg\max_{\mathbf{B},\mathbf{R}} \; \frac{\sigma}{2\gamma_0}\left\|\mathbf{B}\right\|^2 + \sum_{k=1}^{K} \frac{\sigma}{2\left(\gamma_0 - \gamma_k\right)} \left\|\mathbf{P}(:,k)\,\mathbf{R}(k)\right\|^2 \\ &\text{s.t.} \quad \mathbf{G}\left(\mathbf{B} + \mathbf{P}\mathbf{R}\right) - \mathbf{Y} = 0 \qquad (5.2.3)\end{aligned}$$

where $\gamma_0, \gamma_1, \ldots, \gamma_k, \sigma \in \mathbb{R}_+$ are parameters related to the assumed distributions of $\mathbf{B}$ and $\mathbf{S}$ given in equations (5.1.6) and (5.1.7). For ease of notation, we define the following equivalent parameters

$$\begin{aligned}\theta_0 &= \frac{1}{\sigma}\gamma_0 \qquad (5.2.4) \\ \theta_k &= \frac{1}{\sigma}\left(\gamma_k - \gamma_0\right), \text{ for } k = 1, 2, \ldots, K. \qquad (5.2.5)\end{aligned}$$

Notice that the parameter σ scales all the other parameters γ_k equally, and thus, it can be ignored when discussing their relative magnitudes only.

The constrained optimization problem in equation (5.2.3) is solved using the method Lagrange multipliers. First, we construct the respective Lagrangian function

$$\mathcal{L}\left(\mathbf{B}, \mathbf{R}, \lambda\right) = \frac{1}{2\theta_0}\left\|\mathbf{B}\right\|^2 + \sum_{k=1}^{K} \frac{1}{2\theta_k}\left\|\mathbf{P}(:,k)\,\mathbf{R}(k)\right\|^2 + \lambda^T\left(\mathbf{G}\left(\mathbf{B} + \mathbf{P}\mathbf{R}\right) - \mathbf{Y}\right) \quad (5.2.6)$$

which can be minimized by solving the normal equations, i.e., the partial derivatives equal to zero. The partial derivatives are given by

$$\frac{\partial}{\partial \mathbf{B}}\mathcal{L}\left(\mathbf{B},\mathbf{R},\lambda\right) = \frac{1}{\theta_0}\mathbf{B} + \mathbf{G}^T\lambda \tag{5.2.7}$$

$$\frac{\partial}{\partial \mathbf{R}(k)}\mathcal{L}\left(\mathbf{B},\mathbf{R},\lambda\right) = \frac{1}{\theta_k}\mathbf{P}(:,k)^T\mathbf{P}(:,k)\,\mathbf{R}(k) + \mathbf{P}(:,k)^T\mathbf{G}^T\lambda \tag{5.2.8}$$

$$\frac{\partial}{\partial \lambda}\mathcal{L}\left(\mathbf{B},\mathbf{R},\lambda\right) = \mathbf{G}\left(\mathbf{B}+\mathbf{P}\mathbf{R}\right) - \mathbf{Y} \tag{5.2.9}$$

Solving the normal equation for (5.2.7) and (5.2.8) lead to the following identities

$$\mathbf{B} = -\theta_0\mathbf{G}^T\lambda \tag{5.2.10}$$

$$\mathbf{R}(k) = -\theta_k\left[\mathbf{P}(:,k)^T\mathbf{P}(:,k)\right]^{-1}\mathbf{P}(:,k)^T\mathbf{G}^T\lambda \tag{5.2.11}$$

Notice that $\mathbf{P}(:,k)^T\mathbf{P}(:,k) \in \mathbb{N}^{1\times 1}$ represents the number of dipoles on the k-th P-region, and thus is invertible. For ease of notation, define

$$|P_k| = \mathbf{P}(:,k)^T\mathbf{P}(:,k) \tag{5.2.12}$$

From equation (5.2.9), we may obtain a closed-form expression for λ,

$$\begin{aligned}\mathbf{Y} &= \mathbf{G}\mathbf{B} + \mathbf{G}\sum_{k=1}^{K}\mathbf{P}(:,k)\,\mathbf{R}(k)\\ &= \mathbf{G}\left(-\theta_0\mathbf{G}^T\lambda\right) + \mathbf{G}\sum_{k=1}^{K}\mathbf{P}(:,k)\,\left(-\theta_k\left|P_k\right|^{-1}\mathbf{P}(:,k)^T\mathbf{G}^T\lambda\right)\\ &= -\mathbf{G}\left(\theta_0\mathbf{I}_N + \sum_{k=1}^{K}\theta_k\left|P_k\right|^{-1}\mathbf{P}(:,k)\,\mathbf{P}(:,k)^T\right)\mathbf{G}^T\lambda \end{aligned} \tag{5.2.13}$$

which leads to the following identity

$$\lambda = -\left[\mathbf{G}\left(\theta_0\mathbf{I}_N + \sum_{k=1}^{K}\theta_k\left|P_k\right|^{-1}\mathbf{P}(:,k)\,\mathbf{P}(:,k)^T\right)\mathbf{G}^T\right]^{-1}\mathbf{Y} \tag{5.2.14}$$

For robustness and to ensure that the resulting matrices are invertible, we propose using a regularized λ_ρ, defined as

$$\lambda_\rho = -W_\rho\mathbf{Y} \tag{5.2.15}$$

$$W_\rho = \left[\mathbf{G}\left(\theta_0 \mathbf{I}_N + \sum_{k=1}^{K} \theta_k \left|P_k\right|^{-1} \mathbf{P}(:,k)\, \mathbf{P}(:,k)^T\right) \mathbf{G}^T + \rho \mathbf{I}_M\right]^{-1} \tag{5.2.16}$$

where $\rho > 0$ is a regularization parameter that must be determined using any of the heuristics defined in Chapter 3.

Using λ_ρ in equations (5.2.10) and (5.2.11) lead to the estimators

$$\hat{\mathbf{B}} = \theta_0 \mathbf{G}^T W_\rho \mathbf{Y} \tag{5.2.17}$$

$$\hat{\mathbf{R}}(k) = \theta_k \left|P_k\right|^{-1} \mathbf{P}(:,k)^T \mathbf{G}^T W_\rho \mathbf{Y} \tag{5.2.18}$$

The estimator for $\mathbf{S}$ is constructed as in equation (5.2.2),

$$\begin{aligned}
\hat{\mathbf{S}} &= \hat{\mathbf{B}} + \sum_{k=1}^{K} \mathbf{P}(:,k) \hat{\mathbf{R}}(k) \\
&= \left(\theta_0 \mathbf{G}^T W_\rho \mathbf{Y}\right) + \sum_{k=1}^{K} \mathbf{P}(:,k) \left(\theta_k \left|P_k\right|^{-1} \mathbf{P}(:,k)^T \mathbf{G}^T W_\rho \mathbf{Y}\right) \\
&= \left(\theta_0 \mathbf{I}_N + \sum_{k=1}^{K} \theta_k \left|P_k\right|^{-1} \mathbf{P}(:,k)\, \mathbf{P}(:,k)^T\right) \mathbf{G}^T W_\rho \mathbf{Y} \qquad (5.2.19)
\end{aligned}$$

Thus, the proposed estimator is given by

$$\hat{\mathbf{S}} = \left(\theta_0 \mathbf{I}_N + \sum_{k=1}^{K} \theta_k \left|P_k\right|^{-1} \mathbf{P}(:,k)\, \mathbf{P}(:,k)^T\right) \mathbf{G}^T W_\rho \mathbf{Y} \tag{5.2.20}$$

$$W_\rho = \left[\mathbf{G}\left(\theta_0 \mathbf{I}_N + \sum_{k=1}^{K} \theta_k \left|P_k\right|^{-1} \mathbf{P}(:,k)\, \mathbf{P}(:,k)^T\right) \mathbf{G}^T + \rho \mathbf{I}_M\right]^{-1} \tag{5.2.21}$$

For ease of notation, the parameter θ_0 will be replaced by rewriting

$$\hat{\mathbf{S}} = \left(\mathbf{I}_N + \sum_{k=1}^{K} \tau_k \left|P_k\right|^{-1} \mathbf{P}(:,k)\, \mathbf{P}(:,k)^T\right) \mathbf{G}^T W_\lambda \mathbf{Y} \tag{5.2.22}$$

$$W_\lambda = \left[\mathbf{G}\left(\mathbf{I}_N + \sum_{k=1}^{K} \tau_k \left|P_k\right|^{-1} \mathbf{P}(:,k)\, \mathbf{P}(:,k)^T\right) \mathbf{G}^T + \lambda \mathbf{I}_M\right]^{-1} \tag{5.2.23}$$

with $\tau_k = \theta_k/\theta_0 > 1$ and $\lambda = \rho/\theta_0$.

5.2.1 Implementation notes

The computation of the estimator described in equations (5.2.22) and (5.2.23) can be accelerated by avoiding explicit matrix multiplications whenever possible. For instance, consider the matrices in the form

$$A_k = |P_k|^{-1}\, \mathbf{P}(:,k)\, \mathbf{P}(:,k)^T \tag{5.2.24}$$

for $1 \le k \le K$. Notice that each $A_k \in \mathbb{R}^{N\times N}$ is an averaging operator with support over the corresponding P-region. In other words, for any vector $V \in \mathbb{R}^N$ we have

$$[A_k V](n) = \begin{cases} \dfrac{1}{|P_k|} \displaystyle\sum_{n' \in P_k} V(n'), & \text{if } n \in P_k, \\ 0, & \text{otherwise.} \end{cases} \tag{5.2.25}$$

With this notation at hand, we can rewrite equation (5.2.22) as

$$\hat{\mathbf{S}} = \left(\mathbf{I}_N + \sum_{k=1}^{K} \tau_k A_k \right) \mathbf{G}^T W_\lambda \mathbf{Y} \tag{5.2.26}$$

Using the matrices A_k, we can construct $\hat{\mathbf{S}}$ row-by-row as follows:

$$\hat{\mathbf{B}}_0 = \mathbf{G}^T W_\lambda \mathbf{Y} \tag{5.2.27}$$

$$\hat{\mathbf{R}}_0(k) = \frac{1}{|P_k|} \sum_{n \in P_k} \hat{\mathbf{B}}_0(n), \text{for } 1 \le k \le K \tag{5.2.28}$$

$$\hat{\mathbf{S}} = \hat{\mathbf{B}}_0 + \sum_{k=1}^{K} \tau_k \mathbf{G}\, \hat{\mathbf{R}}_0(k) \tag{5.2.29}$$

with $\hat{\mathbf{B}}_0 \in \mathbb{R}^{N\times 1}$, $\hat{\mathbf{R}}_0 \in \mathbb{R}^{K\times 1}$. Alternatively, we can construct $\hat{\mathbf{S}}$ row by row as

$$\hat{\mathbf{S}}(n) = \begin{cases} \hat{\mathbf{B}}_0(n) + \tau_k \hat{\mathbf{R}}_0(k), & \text{if } n \in P_k \text{ with } 1 \le k \le K, \\ \hat{\mathbf{B}}_0(n), & \text{if } n \in P_0 \end{cases} \tag{5.2.30}$$

Depending on how large the P-regions are, i.e., how large are each $|P_k|$ with respect to N, it would be more convenient to use either (5.2.29) or (5.2.30). In a practical setting, we informally expect each P-region to contain less than 10% of the total dipoles.

Next, equation (5.2.23) is simplified by considering the properties of the matrices involved. Notice that each matrix A_k is symmetric and idempotent. Furthermore, if the P-regions are non-intersecting, then

$$A_{k_1} A_{k_2} = \begin{cases} A_{k_1}, & \text{if } k_1 = k_2, \\ 0, & \text{otherwise.} \end{cases} \tag{5.2.31}$$

This property can be extended to get powers and decomposition of the averaging operators A_k. In particular, we have the following

$$\left[q_0 \mathbf{I}_N + \sum_{k=1}^{K} q_k A_k \right]^2 = q_0^2 \mathbf{I}_N + \sum_{k=1}^{K} \left(2 q_0 q_k + q_k^2 \right) A_k \tag{5.2.32}$$

which leads to the following identity

$$\left[\mathbf{I}_N + \sum_{k=1}^{K} \left(\sqrt{1+\tau_k} - 1 \right) A_k \right]^2 = \mathbf{I}_N + \sum_{k=1}^{K} \tau_k A_k \tag{5.2.33}$$

For ease of notation, define

$$\mathbf{A}_\Theta = \mathbf{I}_N + \sum_{k=1}^{K} \left(\sqrt{1+\tau_k} - 1 \right) A_k \tag{5.2.34}$$

With this decomposition at hand, we can rewrite W_λ as

$$W_\lambda = \left[\left(\mathbf{G} \mathbf{A}_\Theta \right) \left(\mathbf{G} \mathbf{A}_\Theta \right)^T + \rho \mathbf{I}_M \right]^{-1} \tag{5.2.35}$$

Notice that $\mathbf{A}_\Theta \in \mathbb{R}^{N \times N}$ is a sum of matrices in the form of A_k which, as seen in equation (5.2.25), work as averaging operators. We can conclude that $\mathbf{A}_\Theta$ acts as a weighted average operator over the rows of $\mathbf{G}$. In other words, we can write

$$\left[\mathbf{G} \mathbf{A}_\Theta \right](:, n) = \begin{cases} \mathbf{G}^T(:, n) + \dfrac{\left(\sqrt{1+\tau_k} - 1 \right)}{|P_k|} \displaystyle\sum_{n' \in P_k} \mathbf{G}^T(:, n'), & \text{for } n \in P_k, 1 \leq k \\ \mathbf{G}^T(:, n). & \text{for } n \in P_0 \end{cases} \tag{5.2.36}$$

Clearly, $\left(A_\Theta \mathbf{G}^T\right) \in \mathbb{R}^{N\times M}$ can be computed by updating $\mathbf{G}$ on a similar way as in equation (5.2.29). The similarity can be made more clear by using similar steps.

$$\mathbf{G}_\mathbf{R}(:,k) = \frac{1}{|P_k|}\sum_{n\in P_k} \mathbf{G}^T(:,n), \text{for } 1 \leq k \leq K, \tag{5.2.37}$$

$$[\mathbf{G}\mathbf{A}_\Theta](:,n) = \begin{cases} \mathbf{G}^T(:,n) + \left(\sqrt{1+\tau_k} - 1\right)\mathbf{G}_\mathbf{R}(:,k), & \text{for } 1 \leq k \leq K \\ \mathbf{G}^T(:,n). & \text{for } n \in P_0 \end{cases} \tag{5.2.38}$$

with $\mathbf{G}_\mathbf{R} \in \mathbb{R}^{K\times M}$. Notice that, once $(\mathbf{G}\mathbf{A}_\Theta)$ is computed, it can be subject to SVD decomposition in order to compute W_λ efficiently. This process is similar to all the kernels described in Chapter 2 for linear methods.

We may skip part of the construction and use the eigendecomposition of $(\mathbf{G}\mathbf{A}_\Theta)(\mathbf{G}\mathbf{A}_\Theta)^T$ directly, which is

$$(\mathbf{G}\mathbf{A}_\Theta)(\mathbf{G}\mathbf{A}_\Theta)^T = \mathbf{U}\mathbf{\Sigma}^2\mathbf{U}^T \tag{5.2.39}$$

where $\mathbf{U} \in \mathbb{R}^{M\times M}$ is an orthogonal matrix with eigenvectors of $(\mathbf{G}\mathbf{A}_\Theta)(\mathbf{G}\mathbf{A}_\Theta)^T$ as columns, and $\mathbf{\Sigma} \in \mathbb{R}^{M\times M}$ a diagonal matrix given by

$$\mathbf{\Sigma}(m,m) = \sigma_m \tag{5.2.40}$$

for $m = 1, 2, \ldots, M$ and σ_m an eigenvalue of $(\mathbf{G}\mathbf{A}_\Theta)(\mathbf{G}\mathbf{A}_\Theta)^T$. Without loss of generalization, we may assume

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_M \tag{5.2.41}$$

With these matrices at hand, it is possible to compute W_λ as

$$W_\lambda = \mathbf{U}\left(\mathbf{\Sigma}^2 + \rho\mathbf{I}_M\right)^{-1}\mathbf{U}^T \tag{5.2.42}$$

with the diagonal matrix in the middle given by

$$\left(\mathbf{\Sigma}^2 + \rho\mathbf{I}_M\right)^{-1}(m,m) = \frac{1}{\sigma_m^2 + \rho} \tag{5.2.43}$$

	SD	AUROC	AP	Depth
DLE	0.733	-0.477	-0.054	-0.097
SD		-0.555	-0.393	0.079
AUROC			0.536	-0.103
AP				-0.108

Table 5.1. Pearson correlation between performance metrics for the proposed estimator and the Human dataset.

With this list of considerations, we claim that this new method is not significantly more computationally expensive than the linear methods presented in Chapter 2.

5.2.2 Synthetic Data (Human)

The effectiveness of the proposed ESI method is tested using synthetic data created following the same protocol described in Chapter 4.

In order to use the proposed method, one single P-region is constructed for each trial from the synthetic dataset by following these steps:

1. Recall the position of the seed dipole, $\mathbf{r}_{n^*}$. The source patch has an approximate radius of $\kappa = 12.6$ mm.
2. Select a point at the cortex, $\mathbf{c}$, so that $\|\mathbf{r}_{n^*} - \mathbf{c}\| < \kappa$.
3. Region P_1 consist on a ball with center at $\mathbf{c}$ an radius 2κ.
4. Region P_0 comprises all dipoles not in P_1.

The use of one single P-region is a natural selection since only one connected source patch was considered. Model misspecification is not explored.

The results from these numerical experiments are collected in the figures 5.2, 5.3, and 5.4, as well as in table 5.1. These figures are extensions of figures 4.4, 4.5, and 4.6 with additional data. The findings from these observations may be summarised as follows:

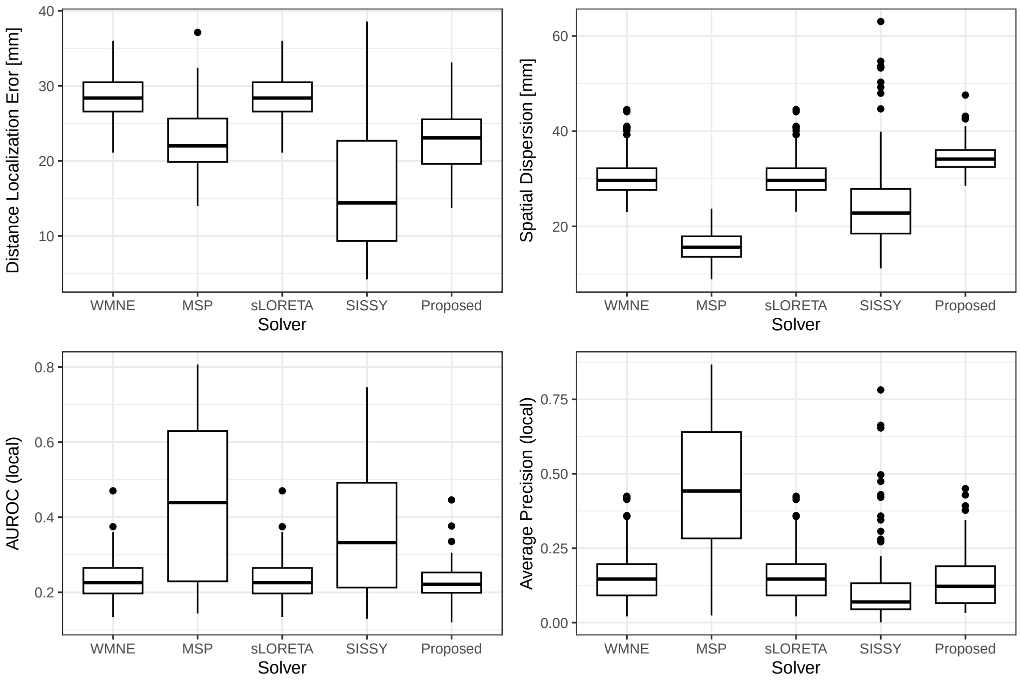

Figure 5.2. Performance metrics for some linear estimators, including the proposed one, for the Human synthetic dataset. The dataset contains 30 trials per shape of source patch, and SNR is set to 30 dB.

- The proposed solver outperforms the linear methods in Distance Localization Error but is outperformed by them in Spatial Dispersion.
- The proposed method is quite sensible to the shape of the source patch.
- The proposed method seems to outperform all methods, including SISSY, in high-noise conditions, as seen by AUROC and Average Precision.
- For the proposed method, performance degradation due to increasing noise is unremarkable –when compared with other linear estimators.

Source patches that resemble point sources, such as the exponential profile, are more easily recovered by all methods; the proposed method is no exception. Average Precision seems to be particularly sensible to the labels created using different shapes of source patches.

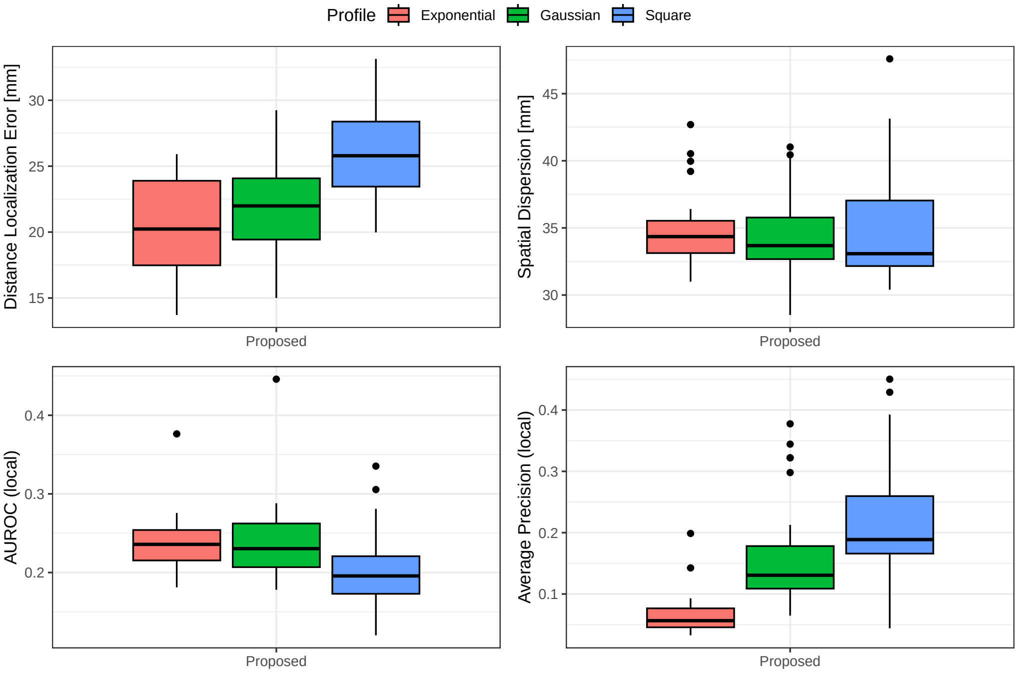

Figure 5.3. Changes in performance metrics for the proposed estimator with the Human dataset when different shapes of source patches are used..

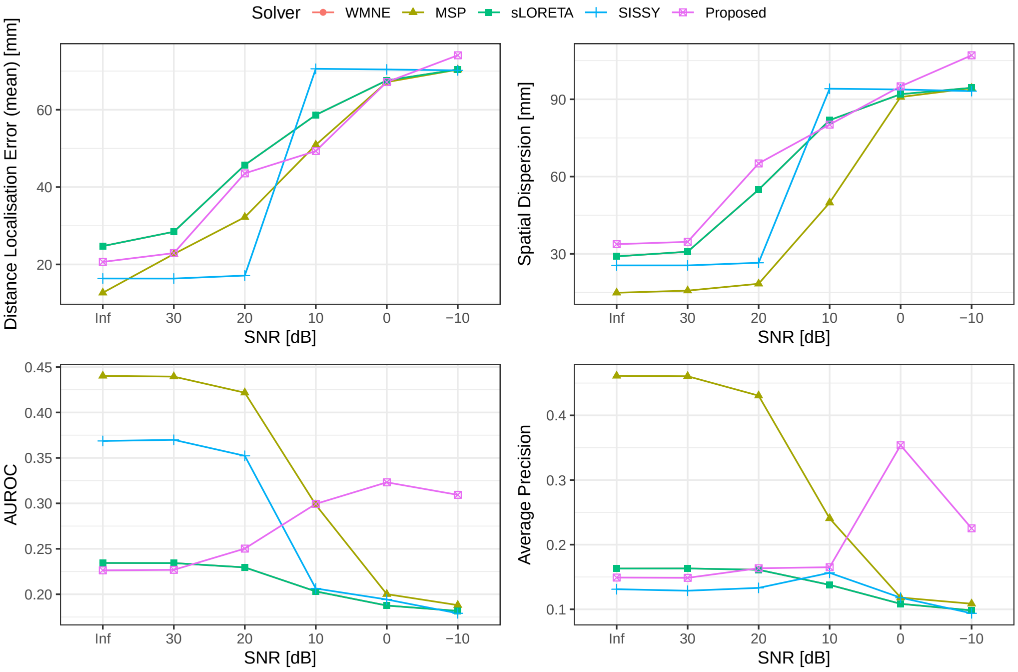

Figure 5.4. Changes in performance metrics for the proposed estimator with the Human dataset as the as the SNR decreases.

The presence of P-regions can explain the proposed method's high performance in the classification metrics: the location of the source is already given, to some extent, independently of noise. In other words, the additional information gives a hot start in creating estimated labels.

The proposed method's advantages seem marginal; they stay within the expected values for linear estimators.

5.2.3 Synthetic Data (Animal Model)

5.2.3.1 Forward Model

We consider a young male from the species *Sus scrofa*: Yucatan minipig for this synthetic dataset. This selection follows from the real data used in the next subsection in preparation for it.

Anatomical data was obtained from an age-matched publicly available MRI template published by Joanne et al. [4].

The skull-stripped MRI was segmented using the CAT12 software [7]. We used the provided skull-tripped MRI to accommodate the anatomical differences concerning human subjects. Later, the cortex surface and its envelope were reconstructed; both surfaces are displayed in figure 5.5.

Following the protocol from the real data, a total of 8 surface electrodes were placed at the cortex; the placement is 10 mm parallel to the Central Line and 10 mm away from the Anterior Edge. See figure 5.6 for a visual guide of the locations of the electrodes.

An adaptive layer of $N = 4,940$ unconstrained dipole sources was constructed over the brain volume. For the forward model, we considered only the surface ECoG electrodes ($M = 8$). Only one single point in time was considered ($T = 1$).

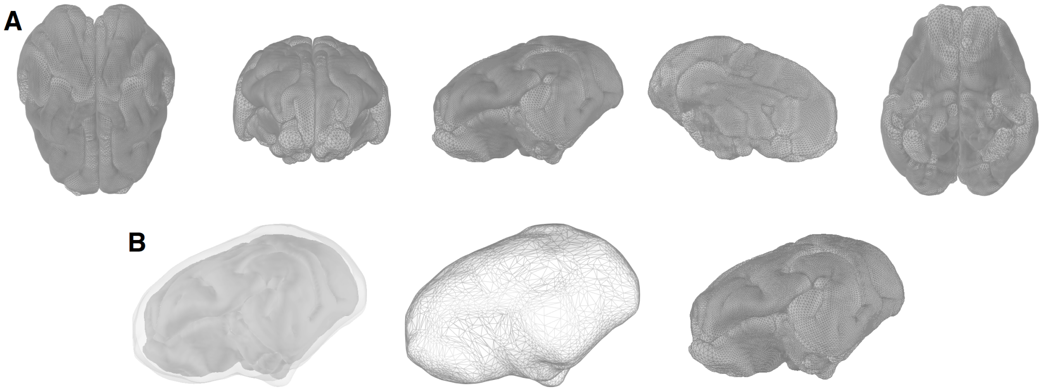

Figure 5.5. A. Cortex surface for the *Sus scrofa* template by Joanne et al. [4], consisting of 8,395 vertices. Views from top, front, left, right, bottom. B. Triangulated surfaces used for the 2-sphere forward model: cortex, cortex envelope.

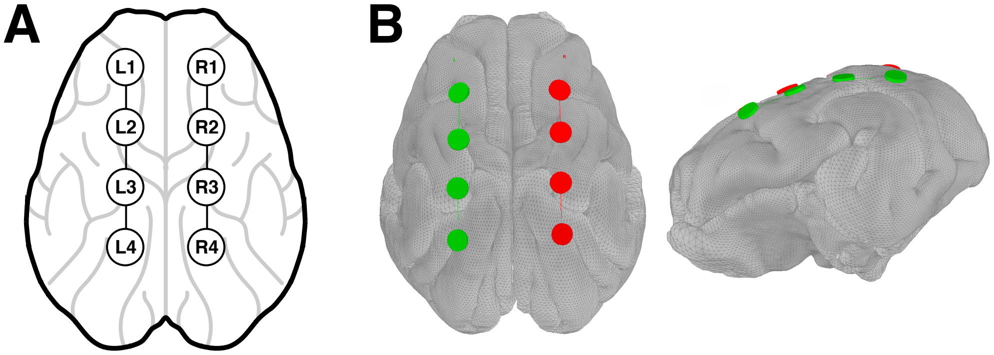

Figure 5.6. A. Approximate location and labels for the surface electrodes on a diagram. B. Electrode locations on the cortex triangulated surface.

For this particular model, a 2-sphere model was constructed with the media being (1) the brain and (2) the space between the cortex and its envelope. This second media is justified because the CSF and connective tissues were not removed from the subject; additionally, it is required computationally due to the reflective conditions at the outer layer. See chapter 2 for more details.

The leadfield matrix was computed using OpenMEEG [10] within the Brainstorm toolbox [30].

	SD	AUROC	AP	Depth
DLE	0.870	0.138	0.105	-0.192
SD		0.028	0.008	-0.144
AUROC			0.957	-0.035
AP				-0.063

Table 5.2. Pearson correlation between performance metrics for the proposed estimator and the Animal Model dataset.

5.2.3.2 Results

As for the previous synthetic dataset, the proposed method is tested. A single P-region is constructed per each trial, following the same protocol with $\kappa = 8.95$ mm. This value of κ was selected so that the resulting source patches have an approximate volume of 3 cm^3, which is similar to the volume of the region with pathological symptoms observed in the real data.

The battery of visualizations from figures 5.7, 5.8, and 5.9 tell us similar information as with the other synthetic dataset: the performance of the proposed method is within the expected, and the estimation is affected by the shape of the source patches.

As for the other datasets, in table 5.2, we have confirmation that the Distance Localization Error and the Spatial Dispersion are correlated, as well as AUROC and Average Precision. None of the variables is significantly correlated with the depth of the source patch.

For this particular dataset, the performance degradation due to increased noise is not very aggressive. This can be explained by the smaller brain volume, which makes the source patches relatively bigger.

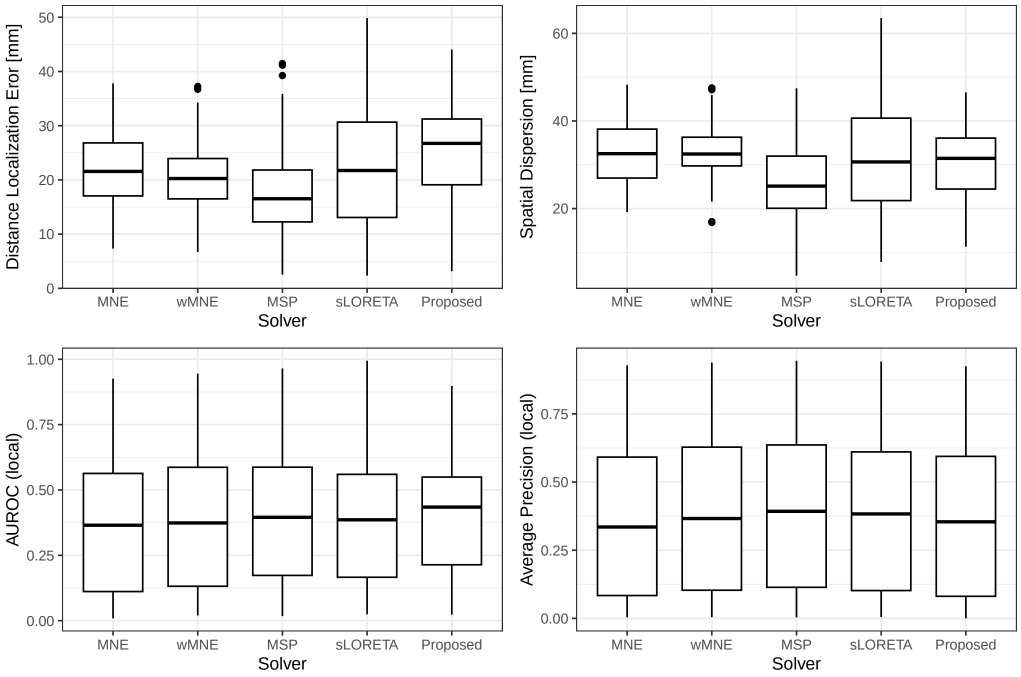

Figure 5.7. Performance metrics for some linear estimators, including the proposed one, for the Animal Model synthetic dataset. The dataset contains 100 trials per shape of source patch, and the SNR is set to 20 dB.

5.2.4 Real Data

The effectiveness of the ESI method is now tested using a dataset from an experiment on acute ischemic stroke in an animal model; this experiment was performed by Dr. J. Pascal et al. at UT Southwestern Medical Center [1, 3].

Acute Ischemic Stroke was induced in the subject by obstruction of the Middle Cerebral Artery. In summary, a stroke is the obstruction of a vein or artery in the brain, which stops the supply of oxygenated blood. The brain regions affected by hypoxia (lack of oxygen at a cellular level) are called the Ischemic Penumbra.

This procedure is referred to as the induced lesion for ease of notation.

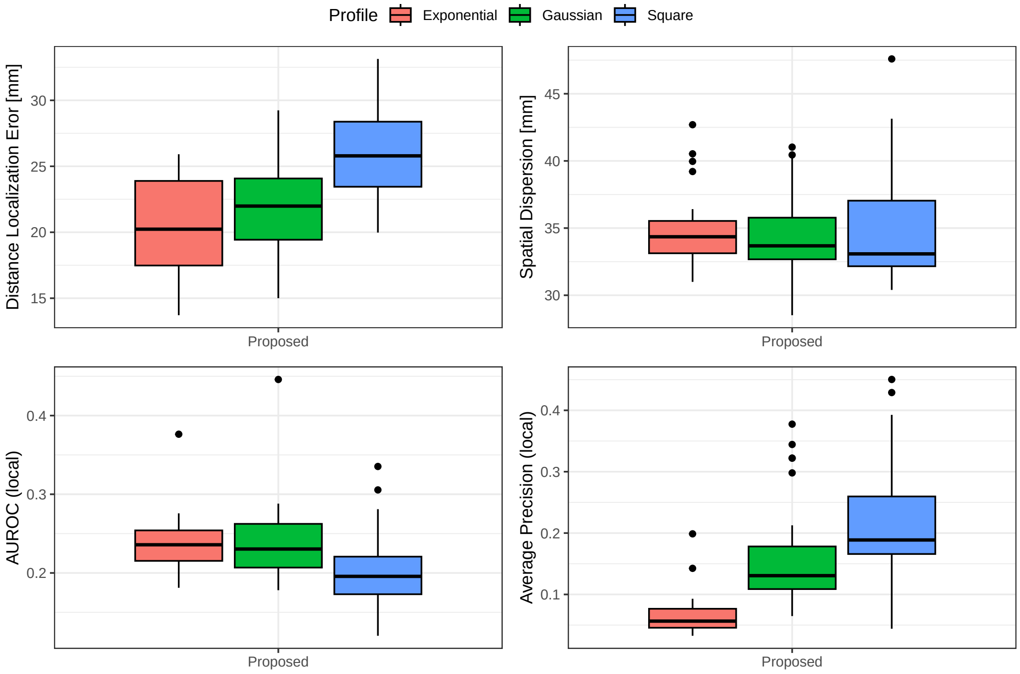

Figure 5.8. Changes in performance metrics for the proposed estimator with the Animal Model dataset when different shapes of source patch are used.

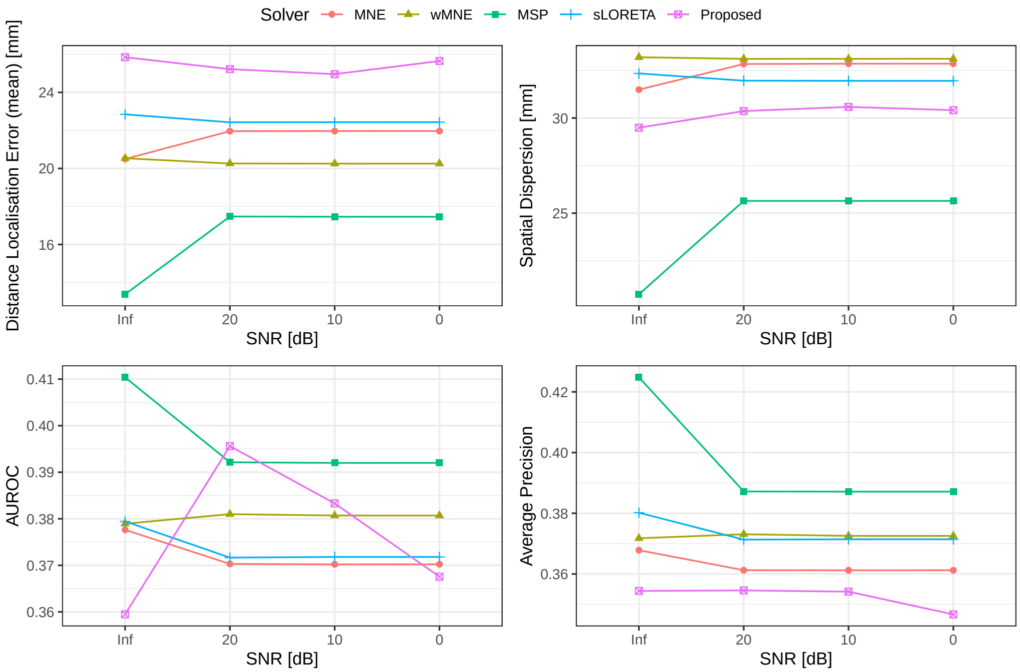

Figure 5.9. Changes in performance metrics for the proposed estimator with the Animal Model dataset as the SNR decreases.

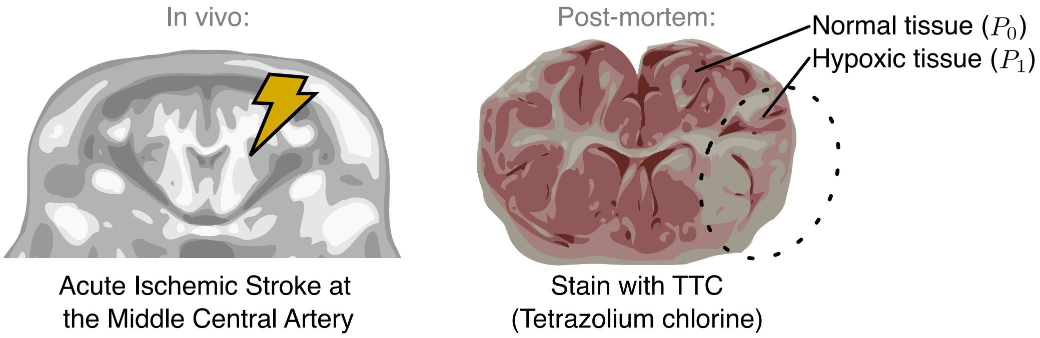

Figure 5.10. P-regions are determined based on observed symptoms from a pathology: the TTC stain colors normal tissue as white and hypoxic tissue as red. Hypoxia is due to an induced acute ischemic stroke in the Middle Cerebral Artery on an animal model (pig) .

5.2.5 Pathology Data

Posterior to the experiment, the subject was sacrificed. During a post-mortem analysis, the subject's brain was extracted and sliced in order to be stained with 2, 3, 5 triphenyltetrazolium (TTC).

The operation of the TTC staining is based on the fact that TTC (white) is degraded to 1,3,5-triphenyl formazan (TPF, red) on the presence of dehydrogenases in metabolically active cells. As a result, tissue colored white was affected by necrosis, and tissue colored red was unaffected [17].

Within the framework of this work, pictures obtained after TTC staining were registered to the template MRI to identify retroactively the Ischemic Penumbra at the time of the ECoG recordings.

With this information, we constructed the P-region P_1 as the Ischemic Penumbra and the rest of the brain as the region P_0. See figure 5.10 for a guide to this process. Although the model can consider multiple P-regions, based on the observed data, only one region is used for this dataset.

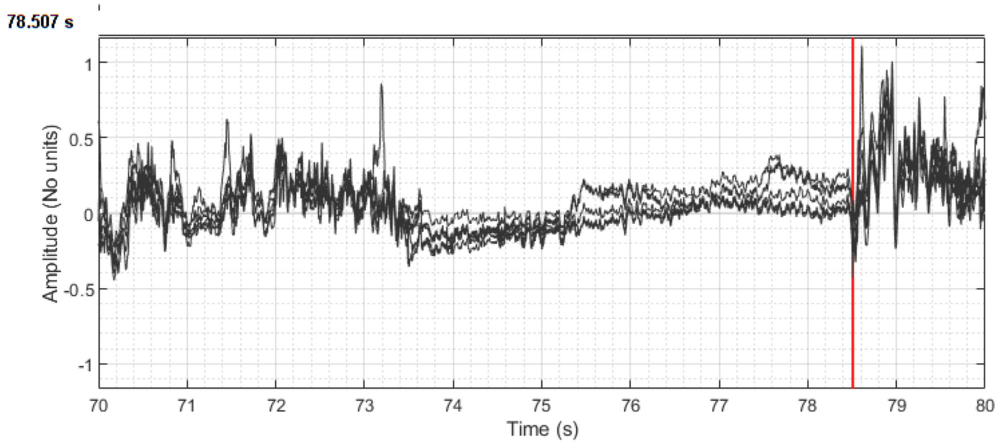

Figure 5.11. Recordings from surface electrodes. The point in time used for figure 5.12 is highlighted. Time is measured after the induced lesion.

5.2.6 Results

After using the TCC data to construct the P-regions, the proposed method was used over recordings from the subject after the induced lesion. Other ESI methods (wMNE, MSP, and sLORETA) were also tested on the same dataset for comparison.

Here, we show an image from the reconstruction at $t = 78.507$ s after the lesion was performed. As shown in figure 5.11, a small peak can observed at this point in time. However, this specific point in time is not of particular interest; it is merely presented as an example to compare the behaviors of different ESI methods.

In figure 5.12, we can observe the cortical activations: the estimated magnitudes of the distributed dipoles superimposed with the subject's MRI. A few coronal slices were used to illustrate the 3D extent of the estimations.

Notice that wMNE and MSP suffer from depth weighting as their estimations are concentrated near the cortex, especially near the sensors. These estimations are also relatively compact in space, which is a positive quality.

On the other hand, sLORETA is spread throughout the whole volume, making it difficult to judge whether it is sensible to the depth of the sources.

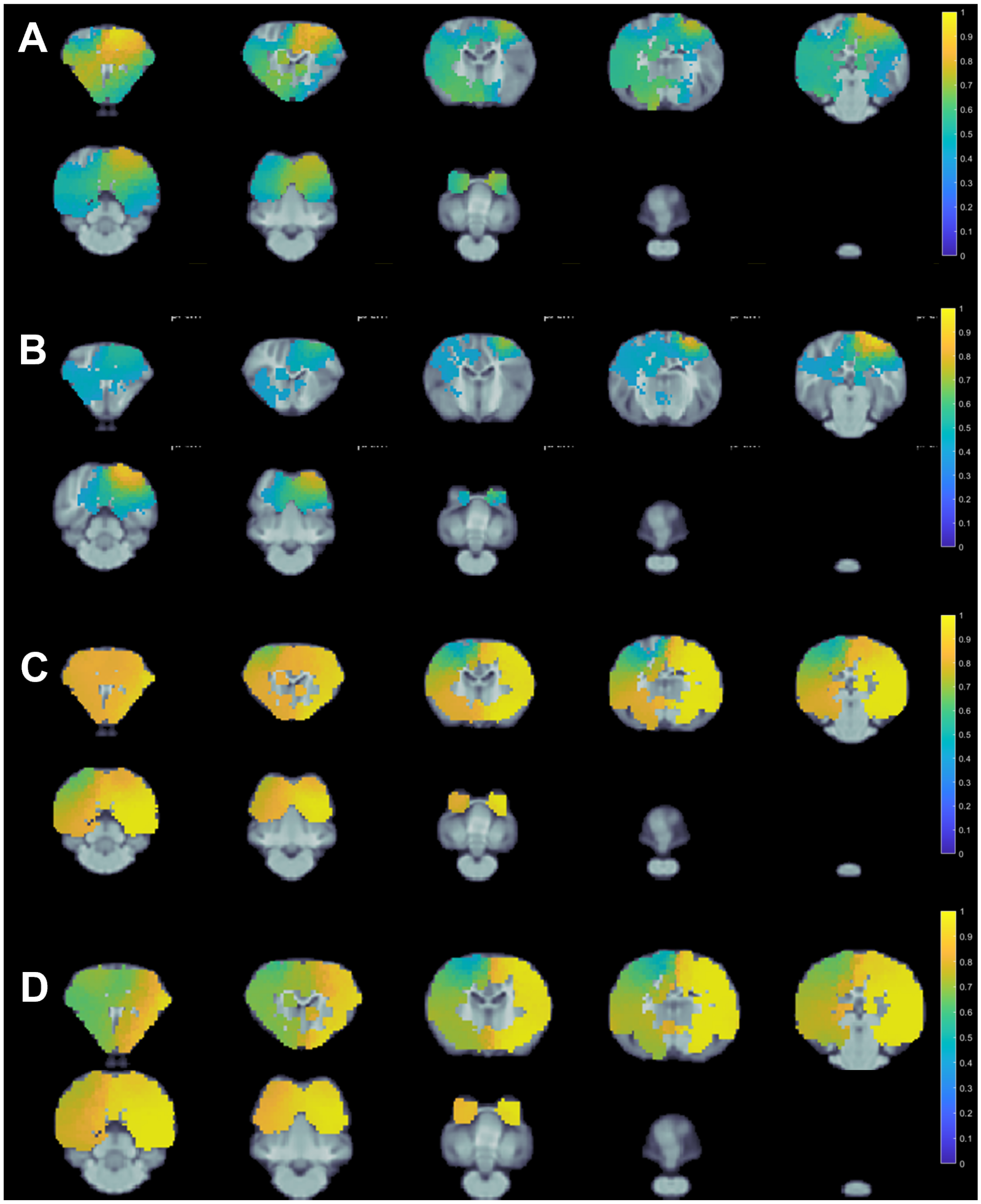

Figure 5.12. Coronal slices of the subject's brain displaying the magnitude of distributed dipoles; maximum magnitude was scaled to 1 for comparison. Results were obtained using various ESI methods: wMNE (A), MSP (B), sLORETA (C), and the proposed method (D).

The proposed method behaves much like sLORETA. Although the spread is similar, we can observe a slightly better focus in space, with regions far from the source possessing lower magnitudes. Notice that the high amplitude region is diffusively aligned with the P-region shown in figure 5.10.

The numerical experiments show that source localization is not heavily correlated with depth for one simple case: a single source patch of relatively small size. We observed that all ESI methods under consideration, including the proposed method, reconstruct better sources whose shape is similar to a point source.

The results from real data suggest that the real source patch may be large in size, and its shape is not similar to a point source.

Based on this data, we can infer that wMNE and MSP enforce the minimal-norm assumption by preferring point-like solutions. In contrast, the sLORETA assumption of smoothness in space is more appropriate for large sources but leads to overspreading. The proposed method's additional information makes it prefer sources located near or at P_1, but it also selects sources that extend on the entirety of P_1 and beyond.

For future work, we suggest using this assumption in combination with stronger assumptions of minimal-norm. Maybe the concept of P-regions can be combined with methods based on the ℓ_1 norm.

REFERENCES

[1] Fuat Arikan et al. “Malignant infarction of the middle cerebral artery in a porcine model. A pilot study”. In: *PLOS ONE* 12.2 (2017), e0172637.

[2] Stephen Boyd et al. “Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers”. In: *Foundations and Trends in Machine Learning* 3.1 (2011), pp. 1–122.

[3] Aksharkumar Dobariya et al. “Recording of pig neuronal activity in the comparative context of the awake human brain”. In: *Scientific Reports* 12.1 (2022), p. 15503.

[4] Joanne E Fil et al. “High-resolution magnetic resonance imaging-based atlases for the young and adolescent domesticated pig (Sus scrofa)”. In: *Journal of Neuroscience Methods* 354 (2021), p. 109107.

[5] Bruce Fisch. *Fisch and Spehlmann’s EEG Primer*. 3rd ed. London, England: Elsevier Health Sciences, 1999.

[6] Vladimir Fonov et al. “Unbiased average age-appropriate atlases for pediatric studies”. In: *Neuroimage* 54.1 (2011), pp. 313–327.

[7] Christian Gaser et al. “CAT–A Computational Anatomy Toolbox for the Analysis of Structural MRI Data”. In: *bioRxiv* (2022), pp. 2022–06.

[8] Gene H Golub, Michael Heath, and Grace Wahba. “Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter”. In: *Technometrics* 21.2 (1979), pp. 215–223.

[9] Alexandre Gramfort, Matthieu Kowalski, and Matti Hämäläinen. "Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods". In: *Physics in Medicine & Biology* 57.7 (2012), pp. 1937–1961.

[10] Alexandre Gramfort et al. "OpenMEEG: opensource software for quasistatic bioelectromagnetics". In: *BioMedical Engineering Online* 9 (2010), pp. 1–20.

[11] Alexandre Gramfort et al. "MEG and EEG Data Analysis with MNE-Python". In: *Frontiers in Neuroscience* 7.267 (2013), pp. 1–13.

[12] Hans Hallez et al. "Review on solving the forward problem in EEG source analysis". In: *Journal of NeuroEngineering and Rehabilitation* 4.1 (2007), pp. 1–29.

[13] Per Christian Hansen. "Analysis of Discrete Ill-Posed Problems by Means of the L-curve". In: *SIAM Review* 34.4 (1992), pp. 561–580.

[14] P.R. Johnston and R.M. Gulrajani. "A New Method for Regularization Parameter Determination in the Inverse Problem of Electrocardiography". In: *IEEE Transactions on Biomedical Engineering* 44.1 (1997), pp. 19–39.

[15] Valer Jurcak, Daisuke Tsuzuki, and Ippeita Dan. "10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems". In: *Neuroimage* 34.4 (2007), pp. 1600–1611.

[16] Dorota Krawczyk-Stańdo and Marek Rudnicki. "Regularization parameter selection in discrete ill-posed problems—the use of the U-curve". In: *International Journal of Applied Mathematics and Computer Science* 17.2 (2007), pp. 157–164.

[17] Li Li, Qiong Yu, and Weimin Liang. "Use of 2, 3, 5-triphenyltetrazolium chloride-stained brain tissues for immunofluorescence analyses after focal cerebral ischemia in rats". In: *Pathology-Research and Practice* 214.1 (2018), pp. 174–179.

[18] Patrick J Lynch and C Carl Jaffe. *Brain human sagittal section*. File: `Brain human sagittal section.svg`. 2006. URL: `https://commons.wikimedia.org/wiki/File:Brain_human_sagittal_section.svg`.

[19] Patrick J Lynch and C Carl Jaffe. *Skull human sagittal section*. File: `Skull human sagittal section.svg`. 2006. URL: `https://commons.wikimedia.org/wiki/File:Skull_human_sagittal_section.svg`.

[20] Jérémie Mattout et al. "Multivariate source prelocalization (MSP): Use of functionally informed basis functions for better conditioning the MEG inverse problem". In: *NeuroImage* 26.2 (2005), pp. 356–373.

[21] Anthony Molins et al. "Quantification of the benefit from integrating MEG and EEG data in minimum $\ell 2$-norm estimation". In: *Neuroimage* 42.3 (2008), pp. 1069–1077.

[22] Ernst Niedermeyer. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 2011.

[23] Paul L Nunez, Michael D Nunez, and Ramesh Srinivasan. "Multi-Scale Neural Sources of EEG: Genuine, Equivalent, and Representative. A Tutorial Review". In: *Brain Topography* 32 (2019), pp. 193–214.

[24] Paul L Nunez and Ramesh Srinivasan. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, USA, 2006.

[25] Robert Oostenveld et al. "FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data". In: *Computational Intelligence and Neuroscience* 2011.1 (2011), p. 156869.

[26] Roberto Domingo Pascual-Marqui et al. "Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details". In: *Methods & Findings in Experimental and Clinical Pharmacology* 24.D (2002), pp. 5–12.

[27] Ceon Ramon, Paul H Schimpf, and Jens Haueisen. "Influence of head models on EEG simulations and inverse source localizations". In: *BioMedical Engineering Online* 5 (2006), pp. 1–13.

[28] Kenneth S. Saladin. *Human Anatomy*. 6th ed. McGraw Hill Educaton, 2020.

[29] Sophie Schrader et al. "DUNEuro—A software toolbox for forward modeling in bioelectromagnetism". In: *PloS one* 16.6 (2021), e0252431.

[30] François Tadel et al. "MEG/EEG Group Analysis with Brainstorm". In: *Frontiers in Neuroscience* 13 (2019), p. 76.

[31] Wayne F Velicer et al. "A Criterion Measurement Model for Health Behavior Change". In: *Addictive Behaviors* 21.5 (1996), pp. 555–584.

[32] Johannes Vorwerk et al. "Comparison of Boundary Element and Finite Element Approaches to the EEG Forward Problem". In: *Biomedical Engineering/Biomedizinische Technik* 57.1 (2012), pp. 795–798.

BIOGRAPHICAL STATEMENT

Julio Cesar Enciso-Alva was born in Hidalgo, Mexico, in 1994. He completed his undergraduate degree in Applied Mathematics in 2017 at Universidad Autónoma del Estado de Hidalgo. He came to the University of Texas at Arlington in 2019 to further his education and was awarded a Master of Science in General Mathematics in 2024.

Julio's interests branch from Applied Mathematics, especially Numerical Methods and Computational Neuroscience.