



CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS.

MAESTRÍA EN CIENCIAS CON ESPECIALIDAD EN
MATEMÁTICAS APLICADAS.

ESTUDIO DEL PROBLEMA INVERSO DE
TOMOGRAFÍA DE IMPEDANCIA ELÉCTRICA
USANDO MÉTODOS BAYESIANOS

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN CIENCIAS

CON ESPECIALIDAD EN
MATEMÁTICAS APLICADAS

P R E S E N T A:

PABLO MIGUEL TADEO TOTOSAUS HERRERA

DIRECTOR DE TESIS:

DR. MARCOS AURELIO CAPISTRÁN OCAMPO.

Guanajuato, Gto. México.

Julio 2016

Datos del Jurado

1 Datos del alumno

Pablo Miguel Tadeo Totosaus Herrera

CIMAT

Maestría en ciencias con especialidad en matemáticas aplicadas

2 Datos del asesor de tesis

Dr. Marcos Aurelio Capistrán Ocampo

CIMAT

3 Datos del sinodal 1

Dr. José Andrés Christen Gracia

CIMAT

4 Datos del sinodal 1

Dr. Miguel Ángel Moreles Vázquez

CIMAT



*A Román y Miriam:
Papá, mamá; ésto es por y para ustedes.*

“Busco la verdad porque soy matemático y creo que cualquiera que acepte la verdad está más cerca de Dios que aquellos que construyen su dignidad humana sobre errores sin sentido.”

“La piel del cielo”, Elena Poniatowska

Resumen

La tomografía de impedancia eléctrica (EIT) tiene como propósito formar una imagen de la conductividad eléctrica en el interior de una sección del cuerpo humano a partir de mediciones tomadas por electrodos en su superficie, esto con el fin de que tal imagen pueda ayudar a conocer el estado clínico del cuerpo en estudio.

Desde el punto de vista matemático, la EIT es un problema inverso no lineal y mal planteado, conocido comúnmente como *Problema Inverso de Calderón*. A lo largo de esta tesis se formula el problema de Calderón desde el enfoque Bayesiano. Se usa elemento finito (FEM) para la resolución del problema directo y métodos Markov Chain Monte Carlo (MCMC) para estudiar la distribución posterior. Se presenta un método para calcular la matriz Jacobiana y así poder calcular la dimensión del subespacio informado por la verosimilitud y conocer la dimensión efectiva del problema.

AGRADECIMIENTOS

A mis padres, a mis hermanos y a mis tías, por todo su apoyo y amor incondicional.

A mi tutor, Marcos Capistrán Ocampo, por asesorar mi desempeño académico y por toda la dedicación y apoyo brindados a este trabajo de tesis.

A mis amigos y compañeros del posgrado, por el apoyo brindado en los momentos necesarios.

A CIMAT, por los apoyos económicos brindados como institución, así como a su personal académico y administrativo.

A CONACYT. por el apoyo económico brindado para poder realizar mis estudios.

ÍNDICE GENERAL

Agradecimientos	III
1. Introducción	2
2. Marco teórico	6
2.1. Planteamiento del problema	6
2.2. El modelo	9
2.3. Formulación débil del problema	12
2.4. Problema inverso	15
2.5. Solución numérica del problema directo	16
2.6. Enfoque Bayesiano	18
2.6.1. Estimadores	21
2.7. Métodos Markov Chan Monte Carlo	23
2.8. Planteamiento Bayesiano de la EIT	26
2.8.1. Distribución a priori de σ	28
2.9. Aplicación de la dimensión efectiva y LIS	31
2.9.1. Matriz jacobiana del mapeo directo	33
3. Resultados	39
3.1. Datos de la simulación numérica	39

3.2. Caso discreto	41
3.3. Espacio informado por la verosimilitud	59
4. Discusión y conclusiones	61
A. Inclusión en el centro, dimensión alta	64
Bibliografía	66

CAPÍTULO 1

INTRODUCCIÓN

La Tomografía de Impedancia Eléctrica (EIT por sus siglas en inglés) es una técnica de formación de imágenes usada para obtener el valor de la conductividad eléctrica (real o compleja) de algún cuerpo o material mediante mediciones de voltaje en la superficie del mismo. Estas mediciones son resultado de inyecciones de corriente eléctrica con frecuencias y amplitudes moderadas a través de electrodos localizados en la superficie del cuerpo en estudio. Este procedimiento tiene varias aplicaciones en la medicina (detección oportuna del cáncer de mama o localización de focos epilépticos) y en la industria (localización de anomalías en estructuras o monitoreo de fluidos en tuberías), y tiene la ventaja de que no es invasivo y cuenta con un bajo costo.

Desde el punto de vista matemático, el recobrar la impedancia eléctrica de un cuerpo a partir de mediciones en su frontera es un problema no lineal y severamente mal planteado conocido como *Problema inverso de Calderón* en honor a Alberto Calderón, quien en el año de 1980 planteó por primera vez el problema en su artículo *On an inverse boundary value problem* Alberto [1].

Calderón formuló una versión abstracta del problema en la que el objetivo era recobrar el coeficiente de la divergencia en una ecuación diferencial parcial elíptica dado su respectivo operador Dirichlet-to-Neumann o Neumann-to-Dirichlet.

Específicamente en el caso Dirichlet-to-Neumann, si Ω es un dominio de \mathbb{R}^d y g es un

operador en $H^{1/2}(\partial\Omega)$, entonces dado $u \in H^1(\Omega)$ tal que satisfice

$$\begin{cases} -\nabla \cdot (\sigma \nabla u) = 0 & \text{en } \Omega \\ u = g & \text{en } \partial\Omega, \end{cases}$$

el problema radica en determinar el valor de σ a partir del conocimiento previo del mapeo

$$\Lambda_\sigma : H^{1/2}(\partial\Omega) \rightarrow H^{-1/2}(\partial\Omega)$$

$$g \mapsto \sigma \frac{\partial u}{\partial \nu},$$

donde, físicamente, g corresponde a las mediciones de voltaje en la frontera y $\Lambda_\sigma(g)$ corresponde a la densidad de corriente en la misma. El espacio de Sobolev $H^{1/2}(\partial\Omega)$ está definido como

$$H^{1/2}(\partial\Omega) = \{u \in L^2(\partial\Omega) \mid \exists \hat{u} \in H^1(\Omega) : u = \text{tr}(\hat{u})\},$$

donde $\text{tr} : H^1(\Omega) \rightarrow L^2(\partial\Omega)$ es el operador traza. A lo largo de este trabajo no abordaremos la formulación abstracta del problema ni sus fundamentos en el análisis funcional, por lo cual sugerimos al lector interesado en profundizar en los conceptos del análisis funcional consultar alguna fuente relacionada con esto, como por ejemplo Cheney [2].

En el año de 1992 se reformuló este problema en un modelo más apropiado físicamente para la implementación de la EIT. En esta reformulación se incluían los electrodos en los que se suministraba una cantidad constante de corriente. Este modelo, el cual será planteado con detalle en las siguientes secciones, ha sido el referente en muchos trabajos anteriores a este y será la base de esta tesis. Sin embargo, es usual encontrarse en la literatura con las dificultades que representa recobrar los valores de la conductividad en las zonas del cuerpo que están alejadas de los electrodos, lo que nos sugiere que estas zonas pueden no estar siendo informadas por los datos. Debido a esto, utilizamos la teoría desarrollada en Cui et al. [5] y Spantini et al. [12] para estudiar el subespacio informado por la verosimilitud del problema, y por consiguiente tener una idea sobre su dimensión efectiva.

En este trabajo abordaremos el problema inverso de tomografía de impedancia eléctrica

mediante métodos Bayesianos que, a diferencia de métodos clásicos de regularización, proponen a los parámetros en cuestión como variables aleatorias y plantea una distribución de probabilidad para estas.

Usamos métodos Markov Chain Monte Carlo (MCMC) para obtener muestras de las densidades de probabilidad obtenidas, las cuales usaremos para construir estimadores de algunas funciones de conductividad a partir de datos sintéticos y también para construir el subespacio informado por la verosimilitud.

Debemos mencionar que la tomografía de impedancia eléctrica sigue siendo un tema de estudio activo tanto en el ambiente de diseño experimental como en la elección apropiada de la a priori, por ejemplo, en Kaipio y Somersalo [9] se plantea una manera óptima de escoger los patrones de corriente en los electrodos, mientras que en Hyvonen y Staboulis [7] se plantea mediante el paradigma Bayesiano una manera de escoger la posición de los electrodos en la superficie del cuerpo a estudiar. Por otro lado, en Dunlop y Stuart [6] se plantea el problema inverso de tomografía impedancia eléctrica en espacios de dimensión infinita, de tal manera que introducen una familia de distribuciones a priori que explotan la geometría del problema. Esto extiende los panoramas del trabajo, pues una posible manera de dar continuación a esta tesis es verificar cuanto influye el diseño del experimento en el subespacio informado por la verosimilitud, o intentar calcular este cuando el problema es formulado con el planteamiento de Stuart.

La presente tesis queda organizada de la siguiente forma: en el capítulo 2 se expone el modelo que será utilizado a lo largo de este trabajo, así como su formulación débil y la implementación del método de elemento finito para resolver el problema directo. Seguidamente se explica de manera breve el paradigma Bayesiano, así como los métodos Markov Chain Monte Carlo. Se plantea el problema de tomografía de impedancia eléctrica en el enfoque Bayesiano y se exponen las a prioris que serán utilizadas. Seguidamente se describe en que consiste el subespacio informado por la verosimilitud para un problema inverso y se da un método para calcular la matriz Jacobiana del mapeo directo para el problema de Calderón. En el capítulo 3 se muestran los resultados obtenidos en la resolución del problema inverso, al igual que en el estudio de la dimensión del espacio informado por los datos. Finalmente, en el capítulo 4

se discuten los resultados mostrados en el capítulo 3 y son confrontados con los resultados de la literatura.

CAPÍTULO 2

MARCO TEÓRICO

2.1. Planteamiento del problema

Algunas propiedades físicas, como la conductividad y la permitividad eléctrica, pueden caracterizar un material o tejido mediante su comportamiento bajo la influencia de campos eléctricos externos. La impedancia eléctrica es la medida de oposición de un cuerpo al flujo de corriente y, dado que esta varía dependiendo del material, es posible distinguir ciertos compuestos de otros dependiendo de la magnitud de esta.

Cuando se trata de tejidos biológicos, la impedancia eléctrica puede depender de la composición del tejido en cuestión Jossinet [8] o de otros factores, como el estado clínico de dicho tejido. Por ejemplo, en Zou y Guo [15] se hace referencia de como el cáncer de mama puede afectar la conductividad eléctrica de los tejidos que componen al seno hasta en un orden de magnitud comparado con un tejido sano.

Tejido	Impedancia (Ω cm)
Músculo	760
Corazón	600
Pulmón	950
Hígado	685
Grasa	≥ 1000

Cuadro 2.1: Valor de impedancia eléctrica para diferentes tejidos humanos

En el cuadro 2.1 se mencionan algunos tejidos del cuerpo humano, así como el valor de conductividad que estos poseen medido en Ohms sobre centímetro.

Dada esta información, es natural pensar que el uso de corrientes alternas en un tejido puede ser de utilidad para conocer propiedades del mismo como su composición o estado clínico, de tal manera que seamos capaces de identificar si nos encontramos en presencia de alguna enfermedad o anomalía que altere la estructura de este. Esta es la idea principal detrás de la tomografía de impedancia eléctrica (EIT).

La tomografía de impedancia eléctrica es una técnica de imagen médica en la cual se infiere una imagen de la conductividad o permitividad de alguna parte del cuerpo a partir de mediciones eléctricas en su superficie. Se han propuesto muchas aplicaciones experimentales para la EIT, como por ejemplo el monitoreo de la función pulmonar, la detección oportuna de cáncer de piel y de seno, y localización de focos epilépticos. En la actualidad la tomografía de impedancia eléctrica juega un papel fundamental en la detección del cáncer de mama, pues en muchos países esta es la única técnica usada para la detección oportuna de la enfermedad.

Con el propósito de generar la imagen médica antes mencionada, se coloca cierta cantidad de electrodos en el tejido o sección del cuerpo humano que es de interés para el estudio, luego se aplican en estos valores controlados de corriente. A través de los mismo electrodos se mide la respuesta del organismo a este voltaje y estos datos son los que después servirán para la formación de la imagen médica.

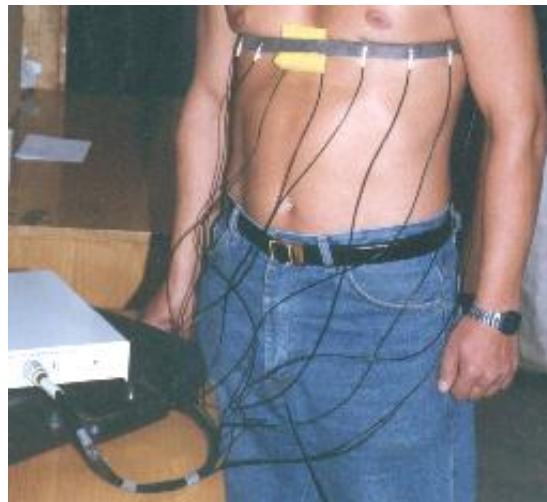


Figura 2.1: Cinturón de electrodos usado comunmente para mediciones en el torax

Notemos que durante todo el proceso no se requiere la aplicación de ninguna sustancia ni se necesita la inserción de algún objeto en el cuerpo estudiado. Esto hace de la tomografía de impedancia eléctrica un estudio no invasivo. Esta observación, combinada con el hecho de que es un estudio de bajo costo, hacen de la tomografía de impedancia eléctrica una alternativa muy conveniente para ciertos estudios específicos.

Matemáticamente hablando, el problema de recobrar la impedancia eléctrica del organismo a partir de las mediciones realizadas en los electrodos es un problema inverso no lineal conocido como *El problema inverso de Calderón*.

Sin embargo, como es común en problemas no lineales y más especialmente en problemas asociados con la realidad, el problema inverso de Calderón es un problema severamente mal planteado, es decir, depende sensiblemente de pequeños cambios realizados en los datos. Dado que nuestra intención es formar una imagen médica, la cual presumiblemente servirá en el diagnóstico médico de alguna persona, el hecho de que el problema dependa sensiblemente a pequeños cambios en los parámetros es crítico. Para la ciencia en general es de vital importancia conocer de que manera se propaga la incertidumbre en un problema y que tan buenas son las predicciones que nosotros somos capaces de realizar antes de poder tomar decisiones. Generalmente nosotros nos basamos en modelos matemáticos que son abstracciones de la realidad y que por consiguiente cuentan con cierto error al no poder capturar completamente la naturaleza del problema, de la misma manera estamos propensos a cometer errores en las mediciones debido a descuidos humanos o por imperfecciones en los instrumentos utilizados para la obtención de estas. Estos errores, unidos a los que se pueden generar de manera computacional, hacen que la sensibilidad del problema a perturbaciones nos pueda llevar a resultados que no coinciden con la realidad, y por consiguiente a tomar decisiones o hacer conclusiones erróneas.

Este último párrafo deja en evidencia la necesidad de desarrollar herramientas matemáticas que no sólo permitan recobrar la impedancia eléctrica a partir de las mediciones en la frontera, sino que también cuantifiquen la incertidumbre presente, es decir, que nos permitan conocer que tan fiables son los resultados obtenidos, así como los alcances y limitaciones de las técnicas desarrolladas.

2.2. El modelo

A lo largo de este trabajo consideraremos al cuerpo en estudio como un dominio conexo y acotado Ω de \mathbb{R}^2 con complemento conexo y frontera suave. Supongamos que los campos electromagnéticos que rodean al cuerpo son de frecuencias bajas y armónicas en el tiempo, es decir, varían de manera periódica o senoidal en la variable temporal.

Hay dos maneras de deducir la ecuación diferencial que gobernará el interior del cuerpo. Primero supongamos que el experimento es estático. Si denotamos por E al campo eléctrico y por B al campo magnético entonces por las leyes de Maxwell $\nabla \times E = 0$, es decir, el campo eléctrico es gradiente y por consiguiente $E = -\nabla u$, donde u es el potencial eléctrico. También por las leyes de Maxwell $\nabla \times H = J$, donde J es la corriente y H es la intensidad magnética. Así que aplicando el operador divergencia en ambos lados de la igualdad anterior y la relación $J = \sigma E$ obtenemos la expresión

$$\nabla \cdot \sigma(\nabla u) = 0 \tag{2.1}$$

que es conocida como la ecuación de conductividad. La función $\sigma = \sigma(x)$ es conocida como la conductividad eléctrica, es el recíproco de la función de impedancia y describe la medida de aceptación de Ω al flujo de la corriente. Desde el punto de vista físico, la ecuación 2.1 nos indica que no hay fuentes ni pozos en el interior de Ω . Otra manera de deducir 2.1 es suponiendo que la permeabilidad magnética del campo es muy pequeña y usando expansión asintótica. Esta formulación es planteada en la referencia Somersalo et al. [11].

Para algunas aplicaciones se considera en la ecuación 2.1 la función de admitividad eléctrica $\gamma(x, \omega) = \sigma(x) + i\omega\epsilon(x)$ donde $\epsilon(x)$ es la permitividad eléctrica, pero para los fines de este trabajo consideraremos únicamente la conductividad σ como una función $\sigma : \Omega \rightarrow \mathbb{R}_+$ tal que satisface las siguientes propiedades de regularidad:

- Existe un número entero $N \geq 1$ y un conjunto de N subconjuntos $\{\Omega_i\}_{i=1}^N$ de Ω dis-

juntos, abiertos y con frontera suave tales que

$$\bar{\Omega} = \bigcup_{j=1}^N \bar{\Omega}_j.$$

- La restricción de σ a cada Ω_j es continua.
- Existen constantes C, c estrictamente positivas y finitas tales que $c < \sigma(x) < C$ para cualquier $x \in \Omega$.

Diremos que la conductividad σ es admisible si satisface los tres puntos anteriores. El conjunto de todas las funciones admisibles en Ω será denotado por $\mathcal{A}(\Omega)$. Los conjuntos Ω_i pueden representar en nuestro contexto órganos o anomalías físicas dentro del cuerpo en estudio. Como se mencionó anteriormente, distintos tejidos toman distintos valores de conductividad que pueden cambiar muy abruptamente dentro del cuerpo, por lo cual tiene sentido considerar a las funciones admisibles como continuas a pedazos.

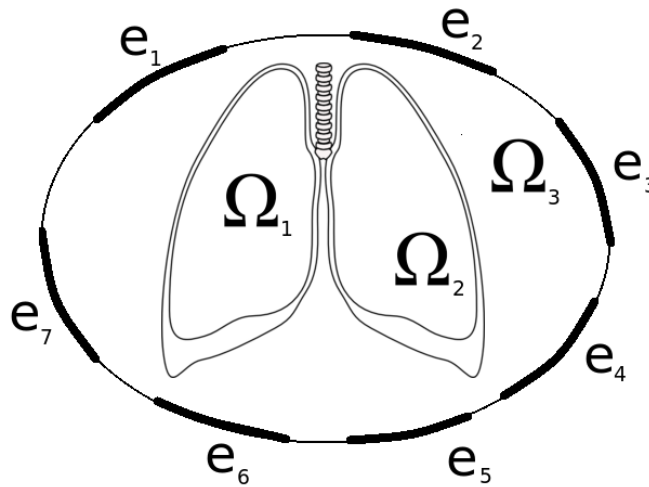


Figura 2.2: Imagen panorámica de la colocación de los electrodos y de los subdominios

Los electrodos que se encargarán de la inyección de corriente y la medición del voltaje resultante en la superficie serán modelados mediante L segmentos e_l en la frontera de Ω tales que son disjuntos por pares, es decir, $e_l \cap e_k = \emptyset$ para $k \neq l$. La corriente eléctrica proporcionada por el electrodo l , así como el voltaje medido serán denotados por I_l y U_l

respectivamente para el electrodo l . En la figura 2.2 se muestra como ejemplo un dominio que posee 7 electrodos diferentes y que está dividido en 3 subdominios, uno por cada pulmón y el complemento de estos como un valor de *background*.

Ahora consideremos a j como la densidad de corriente en cada electrodo y n como el vector normal interior de la superficie. Medimos el flujo de corriente generado en la frontera por el potencial aplicado sobre la misma mediante la aplicación Dirichlet-to-Neumann

$$j = \sigma \cdot \frac{\partial u}{\partial n} \quad \text{en } \partial\Omega, \quad (2.2)$$

así que suponiendo que no hay entrada ni salida de corriente fuera de los electrodos obtenemos que la intensidad de corriente I_l en el electrodo e_l viene dada por

$$I_l = \int_{e_l} \sigma \cdot \frac{\partial u}{\partial n} dS \quad (2.3)$$

mientras que fuera de los electrodos se satisface

$$\sigma \cdot \frac{\partial u}{\partial n} = 0. \quad (2.4)$$

Finalmente, el contacto entre el electrodo y la superficie del cuerpo en estudio genera un efecto electro-químico que se ve reflejado con una caída en el voltaje de la zona, este fenómeno es conocido como *impedancia de contacto efectivo*. Dicho decaimiento es modelado mediante una constante z_l para cada electrodo e_l , por lo que el voltaje en el electrodo está determinado por

$$U_l = u + z_l \sigma \cdot \frac{\partial u}{\partial n}. \quad (2.5)$$

Supondremos que el valor de la impedancia de contacto es conocido y, por simplicidad, real. Después de esta breve discusión presentamos el modelo preliminar para el problema de conductividad en el plano:

$$\begin{aligned}
\nabla \cdot \sigma(\nabla u(x)) &= 0, & x \in \Omega; \\
I_l &= \int_{e_l} \sigma \cdot \frac{\partial u}{\partial n} dS, & 1 \leq l \leq L; \\
\sigma \cdot \frac{\partial u}{\partial n} &= 0, & x \in \partial\Omega, \quad 1 \leq l \leq L; \\
U_l &= u + z_l \sigma \cdot \frac{\partial u}{\partial n}, & x \in e_l, \quad 1 \leq l \leq L;
\end{aligned} \tag{2.6}$$

el cual captura la esencia del problema físico con la precisión de las mediciones tomadas en la referencia Somersalo et al. [11]. Este modelo fue propuesto por primera vez en Cheng et al. [3], y remitimos al lector a dicha referencia para mayores comentarios sobre el mismo. En la siguiente sección establecemos nociones de existencia y unicidad de soluciones para el modelo 2.6, así como la formulación débil del mismo.

2.3. Formulación débil del problema

Supongamos que la conductividad eléctrica σ , el vector de corrientes $(I_l)_{l=1}^L$ y las impedancias de contacto $(z_l)_{l=1}^L$ son fijas. Nuestro interés en este apartado es fijar condiciones necesarias para poder garantizar la existencia de una función $u : \Omega \rightarrow \mathbb{R}$ y un vector $U \in \mathbb{R}^L$ que satisfagan el sistema 2.6 y que además sean únicos. Debido a que el cuerpo humano presenta tejidos con diferentes valores de conductividad, tenemos que suponer que la función σ presenta discontinuidades en el dominio de interés. Esto nos obliga a considerar una solución débil para nuestro problema. Las demostraciones de las proposiciones enunciadas a continuación, así como la notación usada, pueden ser localizadas para el caso complejo en Somersalo et al. [11], en este trabajo únicamente consideraremos el caso real.

Antes que nada definiremos el espacio donde se encontrarán las soluciones de nuestro sistema. Consideremos el espacio $H^1(\Omega)$ definido como

$$H^1(\Omega) = \{u \in L^2(\Omega) \mid D^\alpha u \in L^2(\Omega) \forall |\alpha| \leq 1\}$$

y equipado con la norma

$$\|u\|_{H^1} = \|u\|_{L^2(\Omega)} + \|\nabla u\|_{L^2(\Omega)}.$$

Debido a que queremos nociones de continuidad y suavidad en nuestras soluciones, pediremos que $u \in H^1(\Omega)$, así, dado que la solución viene dada por una dupla (u, U) , definimos nuestro espacio de soluciones como :

$$\mathbb{H} = H^1(\Omega) \oplus \mathbb{R}^L,$$

equipado con la norma usual de la suma directa de espacios. Notar que si $(u, U) \in \mathbb{H}$ es solución de 2.6 entonces también lo es $(u - c, U - c)$ para cualquier $c \in \mathbb{R}$, por lo cual no podemos aspirar a unicidad de soluciones en el espacio \mathbb{H} . Para tener una noción de solución única trabajaremos sobre el espacio

$$\dot{\mathbb{H}} = \mathbb{H}/\mathbb{R},$$

equipado con la norma usual del cociente de espacios.

Denotemos por w' a la dupla (w, W) . El siguiente resultado nos garantiza que podemos debilitar el problema mediante una forma bilineal:

Proposición 2.3.1. Sean $B : \dot{\mathbb{H}} \times \dot{\mathbb{H}} \rightarrow \mathbb{R}$ y $r : \dot{\mathbb{H}} \rightarrow \mathbb{R}$ definidas por

$$B(v', w'; \sigma) = \int_{\Omega} \sigma \nabla v \cdot \nabla w dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (v - V_l)(w - W_l) dS,$$

$$r(w) = \sum_{l=1}^L I_l W_l.$$

Entonces si $v' \in \dot{\mathbb{H}}$ es una solución fuerte de el problema 2.6, también satisface

$$B(v', w'; \sigma) = r(w') \text{ para todo } w' \in \dot{\mathbb{H}}. \quad (2.7)$$

Recíprocamente, en la referencia Somersalo et al. [11], se prueba que si cierto $v' = (v, V) \in \dot{\mathbb{H}}$ satisface 2.7 para cualquier $w' \in \dot{\mathbb{H}}$ entonces es solución de 2.6 en el sentido débil. La siguiente proposición nos afirma que bajo una elección oportuna en el patrón de corrientes aplicado obtenemos unicidad de soluciones en $\dot{\mathbb{H}}$.

Proposición 2.3.2. Supongamos que la conductividad σ , el vector de corrientes $(I_l)_{l=1}^L$ y las impedancias de contacto $(z_l)_{l=1}^L$ satisfacen

1. $\sigma \in L^\infty(\bar{\Omega}; \mathbb{R})$; $\text{ess inf}_{x \in \Omega} \sigma(x) = \sigma_- > 0$;
2. $0 < z_- \leq z_l \leq z_+$, $l = 1, \dots, L$;
3. $\sum_{l=1}^L I_l = 0$.

Bajo estas condiciones, existe un único $[(v, V)] \in \dot{\mathbb{H}}$ tal que satisface 2.7.

La condición (3) de la proposición 2.3.2 es únicamente condicionar las intensidades de corriente de tal manera que se satisfaga ley de conservación de la energía, por lo cual es compatible con la realidad.

Notar que la proposición 2.3.2 concluye que existe una solución única en el espacio $\dot{\mathbb{H}}$, sin embargo nuestra interés es poder garantizar la unicidad en \mathbb{H} . Para esto se escogerá, sin pérdida de generalidad, el representante de la solución dada por la proposición 2.3.2 tal que satisface

$$\sum_{l=1}^L V_l = 0. \quad (2.8)$$

Se escoge este representante debido a que es muy sencillo lograr que los voltajes de corriente satisfagan la ecuación 2.8 mediante una correcta elección de voltaje a tierra. Así, para poder tener existencia y unicidad del modelo 2.6, lo extendemos a

$$\begin{aligned} \nabla \cdot \sigma(\nabla u(x)) &= 0, & x \in \Omega; \\ I_l &= \int_{e_l} \sigma \cdot \frac{\partial u}{\partial n} dS, & 1 \leq l \leq L; \\ \sigma \cdot \frac{\partial u}{\partial n} &= 0, & 1 \leq l \leq L; \\ j &= \sigma \cdot \frac{\partial u}{\partial n}, & \partial\Omega - e_l, 1 \leq l \leq L; \\ \sum_{l=1}^N I_l &= 0 & ; \\ \sum_{l=1}^N U_l &= 0 & . \end{aligned} \quad (2.9)$$

dicho modelo será el utilizado a lo largo de este trabajo.

Ahora supongamos que el patrón de corrientes $(I_l)_{l=1}^L$ y el vector de valores de la impedancia de contacto $(z_l)_{l=1}^L$ son conocidos y satisfacen las hipótesis del teorema 2.3.2, denotaremos por $\mathcal{M}_1 : \mathcal{A}(\Omega) \rightarrow \mathbb{H}$ al mapeo que a cada conductividad σ que cumple ser admisible le asigna la única dupla $(v, V) \in \mathbb{H}$ dada por la proposición 2.3.2 que satisface la condición 2.8, por dicha proposición este mapeo está bien definido. Nosotros nos referiremos a \mathcal{M}_1 como *mapeo hacia adelante*, el cual satisface ser continuo. Referimos al lector a Dunlop y Stuart [6] para la demostración de la continuidad del mapeo hacia adelante y más propiedades sobre el mismo.

2.4. Problema inverso

Hasta el momento se ha presentado el modelo que gobierna la tomografía de impedancia eléctrica, así como el mapeo hacia adelante \mathcal{M}_1 que es aquel que a cada conductividad σ le asigna la dupla (v, V) que satisface 2.9.

A lo largo de este trabajo llamaremos *problema directo* al problema de, dada la conductividad σ , el vector de impedancias de contacto $(z_l)_{l=1}^L$ y el vector de patrones de corriente $(I_l)_{l=1}^L$, obtener el vector (v, V) que satisface 2.9, es decir, el problema directo consiste en evaluar el mapeo hacia adelante. Notar que resolver el problema directo implica resolver una ecuación diferencial parcial. El método utilizado para esto será discutido en la sección 2.5.

Sin embargo, el problema de Calderón consiste más bien en, dados los valores de la impedancia de contacto efectivo, el vector de patrones de corriente $(I_l)_{l=1}^L$ y las mediciones en los electrodos $(V_l)_{l=1}^L$, recuperar la función de conductividad σ . Nos referiremos a este problema como *problema inverso*. Notar que dado que el vector $(V_l)_{l=1}^L$ está formado presumiblemente por datos experimentales, los cuales a su vez están infestados por distintos errores, los métodos de inversión clásicos pueden no ser la mejor opción para la resolución del problema inverso. Por ende, el problema será abordado mediante el enfoque Bayesiano, el cual será discutido brevemente en la sección 2.6.

2.5. Solución numérica del problema directo

El propósito de esta sección es dar un breve resumen sobre la implementación numérica de la solución del sistema 2.9. Una descripción más detallada puede ser encontrada en Zamora [14].

La solución del modelo directo será planteada mediante el método de elemento finito (FEM). Este método es comúnmente utilizado para resolver numéricamente problemas descritos por ecuaciones diferenciales parciales, ecuaciones integrales o ecuaciones variacionales en general. La idea del método es, como su nombre lo dice, dividir el dominio de interés en un número finito de regiones o *elementos* y aproximar la solución de la ecuación diferencial parcial mediante funciones lineales o polinomiales de grado más alto en cada una de estas regiones.

Sea u^h la aproximación de la solución u de 2.9 mencionada anteriormente. Supongamos que u^h está en el subespacio $\mathcal{Q}^h = \text{span}\{\varphi_i | 1 \leq i \leq N_n\}$ de dimensión N_n , entonces es posible escribir a dicha aproximación como una combinación lineal de los elementos de \mathcal{Q}^h de la forma

$$u^h = \sum_{i=1}^{N_n} \alpha_i \varphi_i, \quad (2.10)$$

donde las funciones φ_i son las funciones base relacionadas al método de elemento finito, N_n es el número de nodos y h es el tamaño de la malla.

Recordar que la solución de 2.9 también está conformada por un vector $U \in \mathbb{R}^L$, por lo que también plantearemos un vector aproximado U^h . Para poder garantizar una única solución al problema necesitamos que U satisfaga la condición 2.8, por lo que aproximaremos el vector de voltaje $U^h = (U_1^h, U_2^h, \dots, U_L^h)$ mediante

$$U^h = \sum_{i=1}^{L-1} \beta_i n_i, \quad (2.11)$$

donde los vectores $n_j \in \mathbb{R}^L$ están dados por

$$n_1 = (1, -1, 0, \dots, 0)^T,$$

$$\begin{aligned} n_2 &= (1, 0, -1, \dots, 0)^T, \\ &\vdots, \\ n_{L-1} &= (1, 0, \dots, -1)^T, \end{aligned}$$

de tal modo que por construcción se satisface la condición requerida.

Para aplicar el elemento finito debemos plantear el problema en su forma variacional, la cual está dada por la ecuación

$$B(v', w'; \sigma) = \int_{\Omega} \sigma \nabla v \cdot \nabla w dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (v - V_l)(w - W_l) dS = \sum_{l=1}^L I_l W_l.$$

planteada en la proposición 2.3.1.

Así, sustituyendo 2.10 y 2.11 en la ecuación 2.7, y usando la teoría de elementos finitos, obtenemos la aproximación (u^h, U^h) mediante la resolución del sistema de ecuaciones lineales

$$Ab = f, \tag{2.12}$$

donde $b = (\alpha, \beta) \in \mathbb{R}^{N_n+L-1}$, f es el vector

$$f = \begin{pmatrix} \mathbf{0} \\ \sum_{l=1}^L I_l(n_j)_l \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ C^T I \end{pmatrix}, \tag{2.13}$$

$\mathbf{0}$ denota el vector 0 en \mathbb{R}^{N_n} y $C \in \mathbb{R}^{L \times (L-1)}$ es la matriz definida por

$$C = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 0 & \cdots & & 0 \\ 0 & -1 & 0 & \cdots & \vdots \\ \vdots & & \ddots & & \\ & & & \ddots & \\ 0 & \cdots & & & -1 \end{pmatrix}. \tag{2.14}$$

La matriz $A \in \mathbb{R}^{(N_n+L-1) \times (N_n+L-1)}$ es conocida como la *matriz de ajuste* y está definida

por bloques de la siguiente manera:

$$A = \begin{pmatrix} B & C \\ C^T & G \end{pmatrix}. \quad (2.15)$$

La matriz B es la suma de las matrices $M, W \in \mathbb{R}^{N_n \times N_n}$ conocidas como *matriz de masas* y *matriz de rigidez* respectivamente, cuyas entradas están dadas por

$$W_{i,j} = \int_{\Omega} \sigma \nabla \varphi_i \cdot \nabla \varphi_j dx, \quad M_{i,j} = \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} \varphi_i \varphi_j dS, \quad 1 \leq i, j \leq N_n,$$

por otro lado, la matriz $C \in \mathbb{R}^{N_n \times (L-1)}$ está dada por

$$C_{i,j} = - \left(\frac{1}{z_1} \int_{e_1} \varphi_i dS - \frac{1}{z_{j+1}} \int_{e_{j+1}} \varphi_i dS \right),$$

donde los índices recorren $1 \leq i \leq N_n$, $1 \leq j \leq L-1$, y la matriz $G \in \mathbb{R}^{(L-1) \times (L-1)}$ está determinada por

$$G_{i,j} = \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (n_i)_l (n_j)_l dS = \begin{cases} \frac{|e_1|}{z_1}, & i \neq j \\ \frac{|e_1|}{z_1} + \frac{|e_{j+1}|}{z_{j+1}}, & i = j, \end{cases}$$

donde $1 \leq i$ y $j \leq L-1$.

En Zamora [14] también se pueden hallar algoritmos para la implementación numérica del elemento finito, así como propiedades de la matriz de ajuste que nos serán útiles más adelante.

2.6. Enfoque Bayesiano

En esta sección se hará una breve descripción del enfoque Bayesiano para resolución de problemas inversos en general, para una discusión más detallada referimos al lector a Kaipio y Somersalo [9]. A lo largo de este trabajo supondremos que todas las variables aleatorias en cuestión son absolutamente continuas, por lo que sus distribuciones de probabilidad están

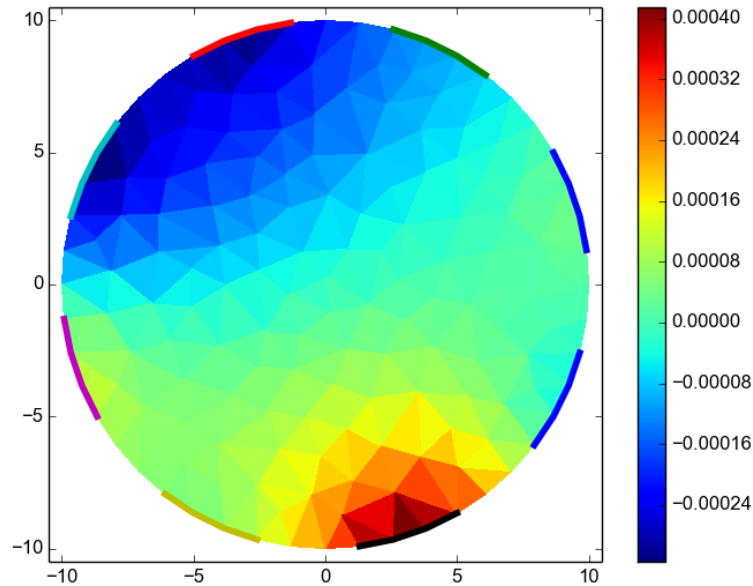


Figura 2.3: Ejemplo de la implementación del elemento finito en tomografía de impedancia eléctrica

dadas por funciones de densidad.

Supongamos que tenemos un conjunto de mediciones $y \in \mathbb{R}^m$ que dependen de un vector de parámetros $x \in \mathbb{R}^n$ mediante cierto modelo de distinta índole (físico, químico, económico, etc.). Notemos que este modelo puede carecer de exactitud debido a ciertos parámetros que nosotros no consideramos en el vector x , o a errores o ruidos en las mediciones y . En el planteamiento clásico de problemas inversos, se escribiría el modelo típicamente como

$$y = f(x, e) \quad (2.16)$$

donde $f : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^m$ es la función de modelado y $e \in \mathbb{R}^k$ es el vector que representa los errores y ruidos en la medición, así como los parámetros que desconocemos del modelo.

En la resolución de problemas inversos estadísticos, todos los parámetros son vistos como variables aleatorias, a partir de este momento, y salvo que se indique lo contrario, las letras mayúsculas serán usadas para representar las variables aleatorias, mientras que las letras minúsculas indicarán las realizaciones de las mismas.

Bajo esta notación, el modelo de nuestro interés será

$$Y = f(X, E). \quad (2.17)$$

Notar que f relaciona a las variables Y, X, E , por lo cual, sus funciones de probabilidad también dependerán las una de las otras.

Las observaciones de la variable de mediciones Y será denotada por y_{obs} .

Supongamos que antes de realizar mediciones contamos con cierto conocimiento de la variable incógnita X . Esta información será concentrada en una densidad de probabilidad $x \mapsto \pi_{pr}(x)$ que llamaremos *densidad a priori*. De igual manera, supongamos que mediante un estudio de la relación entre mediciones e incógnitas, y la información adicional que tengamos del problema, logramos encontrar la probabilidad conjunta de X y Y $\pi(x, y)$, entonces la probabilidad marginal deberá de satisfacer

$$\int_{\mathbb{R}^m} \pi(x, y) dy = \pi_{pr}(x). \quad (2.18)$$

Así que si nosotros tenemos ciertos valores de la incógnita $X = x$, (y $\pi_{pr}(x) \neq 0$) entonces, por definición, la probabilidad condicional de Y dado x vendría dada por

$$\pi(y|x) = \frac{\pi(x, y)}{\pi_{pr}(x)}. \quad (2.19)$$

La probabilidad condicional de Y dado X es llamada la *función de verosimilitud*. Finalmente, si ahora lo que conocemos es $Y = y_{obs}$ y además $\pi(y_{obs}) = \int_{\mathbb{R}^n} \pi(x, y_{obs}) dx \neq 0$ entonces la probabilidad condicional

$$\pi(x|y_{obs}) = \frac{\pi(x, y_{obs})}{\pi(y_{obs})}. \quad (2.20)$$

es conocida como la *distribución posterior* y la denotaremos por π_{post} . Notar que a partir de todo lo comentado anteriormente podemos concluir que, bajo las hipótesis hechas, la posterior esta dada por

$$\pi_{post} = \pi(x|y_{obs}) = \frac{\pi_{pr}(x)\pi(y_{obs}|x)}{\pi(y_{obs})}, \quad (2.21)$$

resultado que es conocido como el *teorema de Bayes para problemas inversos*. Así, el método Bayesiano consiste básicamente en encontrar, dadas las observaciones $Y = y_{obs}$, la distribución posterior.

La densidad marginal $\pi(y)$ es una constante de normalización, y comúnmente no juega un rol muy importante en el problema. Por ende será usual escribir

$$\pi_{post} = \pi(x|y_{obs}) \propto \pi_{pr}(x)\pi(y_{obs}|x).$$

Esta discusión nos proporciona la siguiente estrategia para la resolución del problema inverso:

- Basado en la información que conocemos sobre X , proponer una distribución a priori $\pi_{pr}(x)$.
- Proponer una función de verosimilitud $\pi(y|x)$ que describa adecuadamente la relación entre x y y .
- Una vez obtenidas la a priori y la verosimilitud podemos construir mediante lo expuesto anteriormente la densidad posterior, por lo cual restará el analizar esta función de probabilidad en búsqueda de la información deseada.

2.6.1. Estimadores

Como se mencionó anteriormente, la solución al problema inverso en el enfoque Bayesiano viene dado por la distribución de probabilidad posterior π_{post} . Sin embargo, dicha distribución resulta ser únicamente una abstracción si no contamos con herramientas suficientes para explorarla, es decir, necesitamos ser capaces de construir estimadores con los cuales podamos extraer de la distribución posterior la información que nos interesa de problema, y además tener maneras eficientes de calcularlos.

Lamentablemente, en los problemas reales, es usual encontrarse que los problemas inversos están planteados en dimensiones muy altas, lo cual puede hacer que calcular estimadores (en los cuales usualmente debemos integrar) resulte muy caro computacionalmente hablando.

En la literatura es posible encontrar gran variedad de estimadores usados para el estudio de la posterior, pero en este trabajo usaremos tres. Uno de los estimadores más populares es el llamado *estimador de máxima probabilidad* (o MAP por sus siglas en inglés). Dada una densidad de probabilidad $\pi(x|y)$ de una variable incógnita $X \in \mathbb{R}^n$, el MAP es aquel que satisface

$$x_{MAP} = \arg \max_{x \in \mathbb{R}^n} \pi(x|y).$$

Es decir, el punto que maximiza el valor de la función de densidad.

Notemos que el MAP no existe necesariamente, y en caso de existir no es necesariamente único, por lo que uno tiene que recurrir a métodos de optimización para su búsqueda (por ejemplo, métodos tipo gradiente). Otro estimador muy común es la media condicional (MC) de la incógnita X sujeta a los datos y

$$x_{MC} = E\{x|y\} = \int_{\mathbb{R}^n} x \pi(x|y).$$

Notar que la búsqueda de la media condicional no es un problema de optimización como en el caso del punto más probable, por lo cual uno no se debe preocupar por la suavidad de la densidad de probabilidad. Sin embargo, como mencionamos antes, usualmente las funciones de densidad se encuentran en dimensiones muy altas, por lo cual el cálculo de la media condicional mediante métodos clásicos de integración puede resultar muy arduo numéricamente.

Otro tipo de estimador que puede ser útil en el estudio de la distribución posterior son los *estimadores de dispersión*, como por ejemplo la covarianza condicional

$$cov(x|y) = \int_{\mathbb{R}^n} (x - x_{MC})(x - x_{MC})^T \pi(x|y) dx \in \mathbb{R}^{n \times n},$$

cuya búsqueda es mediante un problema de integración.

2.7. Métodos Markov Chan Monte Carlo

En la sección anterior mencionamos la importancia que tienen los estimadores como la media condicional o la varianza condicional en el estudio de la distribución posterior. Sin embargo, calcular estos estimadores implica resolver un problema de integración, el cual puede ser computacionalmente caro, e incluso irresoluble, por métodos de cuadratura clásicos cuando nos encontramos en dimensiones muy altas.

Una manera de abordar éste problema consiste en evaluar la integral mediante un conjunto *reducido* de puntos muestreados de la misma densidad de probabilidad, tales que estos puntos conforman una muestra significativa de dicha densidad. Algunas de las estrategias para atacar el problema desde esta perspectiva son los métodos *Markov Chain Monte Carlo* (MCMC), los cuales consisten en algoritmos usados para conseguir muestras significantes y poder realizar la integración *Monte Carlo*. A continuación daremos la idea básica detrás de los métodos MCMC.

Consideremos a μ como una medida de probabilidad en \mathbb{R}^n y f una función medible e integrable en \mathbb{R}^n con respecto a la medida μ , es decir, $\mu \in L^1(\mu(dx))$. Supongamos que nuestro problema consiste en aproximar la integral de f con respecto a μ . Si usáramos reglas de cuadratura comunes, deberíamos determinar un conjunto de puntos $\{x_j\}$ en el espacio y un conjunto de pesos $\{w_j\}$ tal que sea posible aproximar la integral de f como

$$\int_{\mathbb{R}^n} f(x)\mu(dx) = \sum_{i=1}^n w_i f(x_i).$$

En el método de integración Monte Carlo, los puntos de soporte $\{x_j\}$ son muestreados aleatoriamente a partir de cierta distribución de probabilidad, y los pesos w_i usados en la regla de cuadratura son determinados por la medida μ . En un caso ideal, estos puntos son muestreados a partir de la misma función de distribución generada por la medida μ . Así, dada una variable aleatoria $X \in \mathbb{R}^n$, nuestro objetivo consiste en diseñar un *generador* de puntos

$\{x_1, \dots, x_N\} \subset \mathbb{R}^n$ distribuidos como μ de tal manera que

$$\int_{\mathbb{R}^n} f \mu(dx) = E\{f(X)\} \approx \frac{1}{N} \sum_{j=1}^N f(x_j).$$

Con el objetivo de encontrar una manera de generar los puntos de soporte mencionados anteriormente, daremos algunos conceptos básicos.

Consideremos $\mathfrak{B} = \mathfrak{B}(\mathbb{R}^n)$ como los conjuntos de Borel sobre \mathbb{R}^n . Un mapeo $P : \mathbb{R}^n \times \mathfrak{B} \rightarrow [0, 1]$ es llamado *kernel de transición de probabilidad*, si

1. Para cada $B \in \mathfrak{B}$, el mapeo $\mathbb{R}^n \rightarrow [0, 1]$ dado por $x \mapsto P(x, B)$ es una función medible.
2. Para cada $x \in \mathbb{R}^n$, el mapeo $\mathfrak{B} \rightarrow [0, 1]$ dado por $B \mapsto P(x, B)$ es una distribución de probabilidad.

Un *proceso estocástico a tiempo discreto* es un conjunto ordenado $\{X_j\}_{j=1}^{\infty}$ de variables aleatorias $X_j \in \mathbb{R}^n$. Una *cadena de Markov a tiempo homogéneo* con kernel de transición P es un proceso estocástico $\{X_j\}_{j=1}^{\infty}$ con las propiedades

$$\mu_{X_{j+1}}(B_{j+1} | x_1, \dots, x_j) = \mu_{X_{j+1}}(B_{j+1} | x_j) = P(x_j, B_{j+1}).$$

Es decir, la probabilidad de que la variable aleatoria X_{j+1} tome valores en el conjunto B_{j+1} sujeto a las observaciones $X_1 = x_1, \dots, X_j = x_j$ es igual a la probabilidad condicionada únicamente a la observación $X_j = x_j$, en otras palabras, la probabilidad de una variable aleatoria depende únicamente de la observación pasada inmediata. También vale la pena recalcar que el kernel de transición no depende del tiempo j en el que estemos evaluando.

Una medida de probabilidad μ se dice que es invariante bajo el kernel de transición P si

$$\mu P = \mu,$$

es decir, la distribución de la variable aleatoria después de un paso del kernel de transición es la misma que antes de dicho paso.

Dada una medida de probabilidad μ , el kernel se dice *irreducible* con respecto a μ (o

simplemente irreducible) si para cada $x \in \mathbb{R}^n$ y $A \in \mathfrak{B}$ tal que $\mu(A) > 0$ existe un entero k tal que $P^{(k)}(x, A) > 0$. Es decir, dado cualquier punto inicial, la cadena de Markov tiene probabilidad positiva de visitar cualquier conjunto con medida mayor que 0.

Un kernel P será llamado *periódico* si para algún entero $m \geq 2$, hay un conjunto de conjuntos disjuntos y no vacíos $\{E_1, \dots, E_m\}$ en \mathbb{R}^n tales que para todo $j = 1, \dots, m$ y cualquier $x \in E_j$, $P(x, E_{j+1(\text{mod } m)}) = 1$. Es decir, si la cadena de Markov llega a tomar algún valor en algún E_j entonces se quedará en alguno de los E_i . En caso contrario decimos que el kernel es *aperiódico*.

La siguiente proposición unifica las definiciones anteriores y es de gran importancia para los métodos Markov Chain Monte Carlo:

Proposición 2.7.1. Sea μ una medida de probabilidad invariante para un kernel de transición P , y supongamos que P es irreducible y aperiódico. Entonces, para cualquier $x \in \mathbb{R}^n$,

$$\lim_{N \rightarrow \infty} P^{(N)}(x, A) = \mu(A) \quad \text{para cualquier } A \in \mathfrak{B},$$

y para cualquier $f \in L^1(\mu(dx))$,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N f(X_j) = \int_{\mathbb{R}^n} f(x) \mu(dx)$$

casi seguramente.

Este resultado nos deja claro el panorama a seguir en el cálculo de los estimadores, éste es construir un kernel aperiódico, invariante e irreducible con el cual generar una cadena de Markov cuya distribución estacionaria sea la distribución posterior y así obtener la muestra significativa que se ha mencionado anteriormente. Así, podremos calcular estimadores a posteriori como nos lo indica la proposición anterior y estudiar dicha distribución. Existe una gran variedad de kernels de transición en la literatura, el propósito de este trabajo no es crear un kernel de transición, por lo cual seremos usuarios del kernel expuesto en Christen et al. [4].

2.8. Planteamiento Bayesiano de la EIT

En esta sección formularemos el problema inverso definido en la sección 2.4 mediante el paradigma Bayesiano expuesto en la sección 2.6.

Antes que nada, notemos que dada una conductividad admisible fija $\sigma \in \mathcal{A}(\Omega)$ y el vector de impedancias de contacto $(z_l)_{l=1}^L$, el mapeo $I \mapsto u'$ tal que a cada vector de patrones de corriente I le asigna el elemento $u' = (u, U)$ tal que satisface 2.9 es lineal, esto es muy de fácil verificar debido a la linealidad de los operadores utilizados en el modelo. Así, podemos afirmar que para cada conductividad $\sigma \in \mathcal{A}(\Omega)$ existe una matriz $R(\sigma) \in \mathbb{R}^{L \times L}$ tal que satisface

$$U = U(\sigma) = R(\sigma)I, \quad (2.22)$$

dicha matriz es conocida como *matriz de resistencia*.

Notar que, con lo que se expuso en la sección 2.5, es posible dar una aproximación numérica de la matriz de resistencia, esto es debido a que por la ecuación 2.11

$$U^h = \sum_{i=1}^{L-1} \beta_i n_j = \mathcal{C}\beta, \quad (2.23)$$

donde \mathcal{C} es como se definió en la ecuación 2.14. Notar que las ecuaciones 2.12 y 2.13 definen a β , así, si definimos a $N(\sigma) \in \mathbb{R}^{(L-1) \times (L-1)}$ como el bloque de la matriz A^{-1} dado por las entradas $A^{-1}(i, j)$, $N_m + 1 \leq i, j \leq N_n + L - 1$ entonces

$$U^h = \mathcal{C}\beta = \mathcal{C}N(\sigma)\mathcal{C}^T I = R^h(\sigma)I, \quad (2.24)$$

donde $R^h(\sigma) = \mathcal{C}N(\sigma)\mathcal{C}^T$. Notar que la notación de la aproximación $R^h(\sigma)$ nos indica que dicha aproximación depende del tamaño de la malla h como en el elemento finito. También vale la pena recalcar de que la matriz de resistencia también depende de las impedancias de contacto, sin embargo suponemos que éstas son fijas, por lo que denotamos $R(\sigma) = R(\sigma, (z_l)_{l=1}^L)$ cuando esto no se preste a confusión.

Recordemos que el problema inverso consiste en recuperar la conductividad σ dadas me-

diciones de voltaje en la frontera. En teoría, nosotros deberíamos ser capaces de inferir más sobre la conductividad si contamos con más datos, por lo cual resulta conveniente aplicar más de un patrón de corriente al dominio con la intención de tener más mediciones de voltaje. Sin embargo, dado que por 2.22 el mapeo $I \mapsto V$ es lineal, únicamente los patrones de corriente linealmente independientes nos proveerán de información que no sea redundante sobre la conductividad en estudio. Debido a que los patrones de corriente son vectores en \mathbb{R}^L , podríamos aplicar hasta L patrones de corriente linealmente independientes, sin embargo, debido a que los patrones deben estar condicionados a la ley de conservación de carga, únicamente podemos aspirar a $L - 1$ patrones linealmente independientes.

Así, supongamos que tenemos una sucesión de J patrones de corriente $(I^{(j)})_{j=1}^J$ donde $1 \leq J \leq L - 1$ y $I^{(j)} = (I_l^{(j)})_{l=1}^L \in \mathbb{R}^L$ para cada j . Supongamos que dichos patrones son aplicados al cuerpo en estudio de manera que recuperamos la sucesión de mediciones ruidosas $V^{(j)} = R(\sigma)I^{(j)}$. Supongamos que el ruido presente en los datos es Gaussiano con media 0 y matriz de covarianza Γ_0 , de manera que podemos escribir

$$y_j = \mathcal{G}_j(\sigma) + \eta_j, \quad \eta_j \sim \mathcal{N}(0, \Gamma_0) \quad (2.25)$$

donde $\mathcal{G}_j(\sigma) = R(\sigma)I^{(j)}$. Concatenando las observaciones, escribimos

$$y = \mathcal{G}(\sigma) + \eta, \quad \eta_j \sim \mathcal{N}(0, \text{diag}(\Gamma_0, \dots, \Gamma_0)). \quad (2.26)$$

Así, buscamos inferir la función σ a partir del vector $y \in \mathbb{R}^{JL}$.

Algunos puntos que se deben de resaltar es que, primeramente, la función σ no se trata de la función de impedancia del cuerpo, sino más bien su función de conductividad. Como se mencionó anteriormente, la función de impedancia es el inverso multiplicativo de la función de conductividad, es decir, si ρ es la impedancia entonces $\rho = \frac{1}{\sigma}$. Así, debido a que la idea original en la tomografía de impedancia eléctrica es precisamente recuperar la función de impedancia ρ , algunos trabajos definen el mapeo de observación como $\mathcal{G}_j(\rho) = R(\rho^{-1})I^{(j)}$, $\rho^{-1} \in \mathcal{A}(\Omega)$ y buscan inferir la función ρ . Debido a que es equivalente recuperar la función de impedancia a recuperar la función de conductividad, nosotros trabajaremos con el mapeo

de observación como se definió en la ecuación 2.25.

Otra observación es que, al utilizar varios patrones de corriente en el problema, se debe resolver muchas veces el problema directo. Debido a que al usar métodos MCMC debemos resolver una gran cantidad de veces el problema directo, utilizar elemento finito desde cero, es decir, calcular las matriz de ajuste y resolver el sistema de ecuaciones 2.12 puede aumentar considerablemente el costo computacional. Por esta razón es importante contar con la aproximación de la matriz de resistencia dada por la ecuación 2.24, pues así es suficiente calcular dicha matriz en una única ocasión y después para calcular cada $V^{(j)}$ sólo necesitamos una multiplicación matriz-vector.

Como se expuso en la sección 2.6, la solución al problema inverso mediante el enfoque Bayesiano está dada por la distribución posterior

$$\begin{aligned}\pi(\sigma|y) &\propto \pi_{pr}(\sigma)\pi(y|\sigma) \\ &\propto \pi_{pr}(\sigma)\pi_{noise}(y - \mathcal{G}(\sigma))\end{aligned}\tag{2.27}$$

Debido a que estamos suponiendo que el ruido se distribuye $\mathcal{N}(0, \text{diag}(\Gamma_0, \dots, \Gamma_0))$, nuestra verosimilitud estará dada por

$$\pi(y|\sigma) \propto \exp\left(-\frac{1}{2}\|y - \mathcal{G}(\sigma)\|_{\Gamma}\right),\tag{2.28}$$

donde $\Gamma = \text{diag}(\Gamma_0, \dots, \Gamma_0)$.

Lo único que nos hace falta para contar con la distribución posterior es la distribución a priori de σ , esta será discutida en la siguiente sección.

2.8.1. Distribución a priori de σ

Nuestro propósito ahora es asignar distribuciones a priori para las funciones de conductividad. Antes que nada, notemos que σ vive en un espacio de funciones, es decir de dimensión infinita, sin embargo, para este trabajo trabajaremos con una discretización de dicha función. Notar que la malla usada en el elemento finito da una partición del dominio Ω en triángulos. Consideremos a \mathcal{T} como el conjunto de triángulos de la malla, diremos que

$\mathcal{P}_m = \{A_1, A_2, \dots, A_m \mid A_i \subset \mathcal{T}, \quad 1 \leq i \leq m\}$ es un *conjunto de pixeles* si

- Los elementos de cada A_i forman un subconjunto conexo de Ω
- $\bigcup_{i=1}^M A_i = \bigcup_{\alpha \in \mathcal{T}} \alpha$.

Los subconjuntos A_i serán conocidos como *pixeles*. Así, dado un conjunto de pixeles \mathcal{P}_m , estaremos interesados en aproximaciones de la forma

$$\sigma(x) = \sum_{A_i \in \mathcal{P}_m} s_i \chi_{A_i}(x) \quad (2.29)$$

donde χ_{A_i} denota la función indicadora del pixel A_i y s_i es un número real. Así, únicamente necesitamos conocer el vector $[s_1, \dots, s_m] \in \mathbb{R}^m$ para determinar en su totalidad la conductividad σ .

Sea $H_{\mathcal{P}_m} \subset \mathcal{A}(\Omega)$ el espacio vectorial generado por las funciones indicadoras de los pixeles de \mathcal{P}_m . Notar que estamos aproximando una función, la cual vive en un espacio de dimensión infinita, mediante un elemento de $H_{\mathcal{P}_m}$ el cual es un espacio vectorial de dimensión finita, lo cual puede ser problemático debido al mal planteamiento del problema. En Dunlop y Stuart [6] se plantea el problema considerando a σ precisamente como un elemento en un espacio infinito-dimensional.

Las a priori que utilizaremos serán de la forma

$$\pi_{pr} \propto \exp(-\alpha L(\sigma)), \quad (2.30)$$

donde el escalar α es nombrado ocasionalmente *parámetro de regularización* y el funcional $L : H_{\mathcal{P}_m} \rightarrow \mathbb{R}$ es conocido como funcional de regularización. Notar que tomando la a priori de esta forma, por 2.27 y 2.28 obtenemos una posterior de la forma

$$\pi(\sigma|y) \propto \exp\left(\frac{1}{2}\|y - \mathcal{G}(\sigma)\|_{\Gamma} - \alpha L(\sigma)\right). \quad (2.31)$$

Lo que resta es determinar el funcional L . En este trabajo utilizaremos los que son planteados en Kaipio et al. [10].

Supongamos que contamos con una función $\sigma_* \in \mathcal{A}(\Omega)$ que sabemos que es una buena aproximación de la verdadera conductividad σ . Así, el primer funcional que utilizaremos para definir la a priori será una aproximación discreta de la norma $\|\sigma - \sigma_*\|_{L^1(\Omega)}$. Como se mencionó anteriormente, la aproximación de la función de conductividad está determinada por el vector $[s_1, \dots, s_m]$ que se definió anteriormente, así, definiremos la a priori L^1 como la que está determinada por el funcional

$$L(\sigma) = L([s_1, \dots, s_m]) = \sum_{i=1}^M |A_i| |s_i - \sigma_*|, \quad (2.32)$$

donde $|A_i|$ denota el área del pixel A_i .

Otra de las a priori que utilizaremos será la llamada *variación total*. Dicha a priori es un caso particular de una clase aun más general de a prioris derivadas de los llamados *Campos aleatorios Markovianos*.

Consideremos una colección de índices indexados

$$\mathcal{N} = \{\mathcal{N}_i | i \leq n\}, \quad \mathcal{N}_i \subset \{1, 2, \dots, n\}. \quad (2.33)$$

Diremos que \mathcal{N} es un *sistema de vecindades* si

- $i \notin \mathcal{N}_i$,
- $i \in \mathcal{N}_j$ si y sólo si $j \in \mathcal{N}_i$.

Así, dado un conjunto de pixeles \mathcal{P}_M definimos un sistema de vecindades sobre $\{1, 2, \dots, m\}$.

Definimos la a priori de variación total como la que está determinada por el funcional

$$L(\sigma) = \sum_{j=1}^m V_j(\sigma), \quad V_j(x) = \frac{1}{2} \sum_{i \in \mathcal{N}_j} l_{ij} |s_i - s_j|, \quad (2.34)$$

donde el vector $[s_1, \dots, s_m]$ es nuevamente el que define a la función σ y l_{ij} es la longitud del borde que tienen en común el pixel A_i y el pixel A_j .

Notar que la a priori de variación total penaliza la oscilaciones en la función de conductividad y además depende de la elección del sistema de vecinos.

La tercera a priori que utilizaremos será una a priori Gaussiana, la que tendrá gran importancia cuando se estudie la dimensión efectiva del problema.

2.9. Aplicación de la dimensión efectiva y LIS

Los parámetros de interés en los problemas inversos son usualmente, como en el caso del problema de Calderón, funciones que dependen del espacio o del tiempo y que por consiguiente forman parte de un espacio de dimensión infinita. En el caso de la tomografía de impedancia eléctrica queremos recuperar una función continua σ la cual es discretizada en un espacio de dimensión finita, por lo que, si queremos una aproximación aceptable, necesitamos una discretización fina de dicha función. Esto puede elevar de gran manera la dimensión del problema inverso.

En la práctica es usual que los problemas inversos estén planteados sobre dimensiones muy altas, lo cual puede afectar la convergencia de los métodos tipo MCMC, y por tanto dificultar la exploración de la distribución posterior.

Como se mencionó en la sección 2.6, la distribución posterior es construida a partir de la distribución a priori de los datos y de la verosimilitud. En un caso ideal, todos los parámetros sobre los cuales estamos haciendo inferencia deberían codificar información de la verosimilitud, o en su defecto la gran mayoría de ellos, sin embargo esto no siempre ocurre, por lo cual es un tema de interés encontrar cuantos son los parámetros que son informados por la verosimilitud (dimensión efectiva) y cuales son estos.

En los artículos Cui et al. [5] y Spantini et al. [12] se plantea un método de reducción de dimensión del problema inverso mediante la identificación del *subespacio informado por la verosimilitud* (LIS, por sus siglas en inglés) cuando la verosimilitud y la a priori son Gaussianas, el cual radica justamente en encontrar el subespacio de parámetros en el cual la verosimilitud tiene un mayor impacto en comparación con la información a priori.

En Cui et al. [5] también se mencionan métodos de factorización de la distribución posterior, lo cual, computacionalmente hablando, agiliza la convergencia de los métodos MCMC, sin embargo para los fines de esta tesis el uso del LIS será únicamente para poder tener una

idea de que tan sensato es recuperar funciones de conductividad que poseen inclusiones muy alejadas de los electrodos. Primeramente describiremos el método planteado para el cálculo del LIS en el caso lineal.

Considerar un mapeo de observación lineal $G(x) = Gx$, con una verosimilitud Gaussiana $\mathcal{N}(\mu_{obs}, \Gamma_{obs})$ y una a priori Gaussiana $\pi_0 = \mathcal{N}(\mu_{pr}, \Gamma_{pr})$. La posterior resultante es también Gaussiana, $\pi(x|y) = \mathcal{N}(\mu_{pos}, \Gamma_{pos})$, donde la media y la matriz de covarianza están escritas en términos de $\mu_{obs}, \Gamma_{obs}, \mu_{pr}$ y Γ_{pr} . Nuestra intención es encontrar un proyector P_r sobre el LIS tal que la aproximación de la posterior original dada por

$$\tilde{\pi}(x|y) = \pi(y|P_r x)\pi_0(x)$$

sea óptima. Dicho proyector, y el sentido en que es óptimo, es determinado en el siguiente resultado de Spantini et al. [12]:

Proposición 2.9.1. Sea $\Gamma_{pr} = LL^T$ una descomposición simétrica de la matriz de covarianza de la a priori y sean (λ_i, v_i) los eigenvalores y eigenvectores del Hessiano preconditionado $L^T H L$ (donde $H = G^T \Gamma_{obs}^{-1} G$, es el Hessiano de la función de ajuste) tales que $\lambda_i \geq \lambda_{i+1}$. Definamos las direcciones $u_i = Lv_i$ y $w_i = L^{-T} v_i$ junto con las matrices $U = [u_1, \dots, u_r]$ y $W_r = [w_1, \dots, w_r]$. Entonces el proyector P_r dado por:

$$P_r = U_r W_r^T,$$

da lugar a la posterior aproximada $\tilde{\pi}(x|y) = \mathcal{N}(\mu_{pos}^{(r)}, \Gamma_{pos}^{(r)})$, la cual es óptima en el sentido de que $\Gamma_{pos}^{(r)}$ minimiza la distancia de Förstner de la matriz de covarianza exacta sobre $\mathcal{M}_r = \{\Gamma_{pr} - KK^T \succ 0 : \text{rank}(K) \leq r\}$ y $\mu_{pos}^{(r)}$ minimiza el riesgo de Bayes sobre las transformaciones lineales de la forma $\mu(y) = Ay$ con $\text{rank}(A) \leq r$.

Referimos al lector interesado al artículo anteriormente citado para más información sobre la distancia de Hellinger y el riesgo de Bayes, y el por que se escogieron estos criterios para determinar la optimalidad del proyector P_r .

Notar que los vectores (u_1, \dots, u_r) generan el rango del proyector óptimo, este rango es al que llamaremos LIS. El teorema 2.9.1 nos da una técnica para calcular el LIS en el caso

lineal, sin embargo el mapeo directo $\mathcal{G} : \sigma \mapsto V$ que a cada conductividad le asigna su respectivo voltaje es no lineal. Para estos casos Cui et al. [5] sugiere considerar la linealización $J(x) \cong \nabla \mathcal{G}$ con el fin de obtener una aproximación local de la sensibilidad de \mathcal{G} . El modelo aproximado $J(x)$ es usado para obtener una aproximación tipo *Gauss-Newton* del Hessiano de la función de ajuste, $H(x) = J(x)^T \Gamma_{obs}^{-1} J(x)$. Una vez que contamos con la aproximación $H(x)$ del Hessiano es posible construir una aproximación local del LIS mediante el método brindado por el teorema 2.9.1, a partir de las l primeras direcciones u_1, \dots, u_l asociadas a los l eigenvalores más grandes tales que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l \geq \tau_{loc}$, para algún valor de truncamiento $\tau_{loc} > 0$. El espacio generado por estas direcciones es conocido como el LIS local en el punto x . Sin embargo nuestro interés radica en conocer el subespacio informado por los datos para el problema global, por lo cual se estima la esperanza de la aproximación del Hessiano preconditionado por la a priori (ppGNH) sobre la posterior, es decir,

$$S = \int_{\mathbf{x}} L^T H(x) L \pi(dx|y). \quad (2.35)$$

En vista de que este es un problema de integración en una dimensión presumiblemente alta, no es posible utilizar métodos de cuadratura convencionales para calcularlo, por lo cual usamos una aproximación Monte Carlo de S dada por

$$\tilde{S} = \frac{1}{n} \sum_{k=1}^n L^T H(x^{(k)}) L,$$

donde $x^{(k)} \sim \pi(x|y)$, $k = 1, \dots, n$ son muestras de la posterior. A partir de este estimador, se calcula el LIS global como en el caso local mediante un valor de truncamiento $\tau_{loc} > 0$.

2.9.1. Matriz jacobiana del mapeo directo

Para calcular el subespacio informado por los datos es necesario contar con la matriz Jacobiana del mapeo directo. En la referencia Kaipio et al. [10] se demuestra que el mapeo directo es Frechet diferenciable y además su derivada satisface ciertas nociones de regularidad. A continuación presentamos una manera de calcular e implementar numéricamente la matriz

Jacobiana del mapeo directo definido por

$$\sigma \mapsto U(\sigma), \quad (2.36)$$

determinando cada una de sus columnas, las cuales están determinadas por la parcial con respecto a cada pixel. Para más información sobre la matriz Jacobiana referimos al lector a Vauhkonen [13].

Notemos que por 2.11,

$$U^h = \sum_{j=1}^{L-1} \beta_j n_j. \quad (2.37)$$

Definamos la matriz $\tilde{\mathcal{C}} \in \mathbb{R}^{L \times N_n + L - 1}$ como $\tilde{\mathcal{C}} = [\mathbf{0}, \mathcal{C}]$. Así, si $b = (\alpha, \beta)$ es la solución del sistema de ecuaciones 2.44 que determina la solución del problema directo, entonces U^h puede ser escrito como

$$U^h = \sum_{j=1}^{L-1} \beta_j n_j = \tilde{\mathcal{C}} b,$$

y por consiguiente, para construir la matriz Jacobiana es suficiente determinar

$$\frac{\partial \tilde{\mathcal{C}} b}{\partial \sigma_l}.$$

Por otro lado, recordemos que $b = A^{-1}f$, donde A es la matriz de precisión definida en secciones anteriores. Teniendo en cuenta de que la matriz $\tilde{\mathcal{C}}$ es constante, entonces para calcular la parcial

$$\frac{\partial \tilde{\mathcal{C}} b}{\partial \sigma_l} = \tilde{\mathcal{C}} \frac{\partial A^{-1}f}{\partial \sigma_l}$$

es suficiente calcular

$$\frac{\partial A^{-1}f}{\partial \sigma_l}.$$

Aplicando la regla de Leibniz a la expresión anterior,

$$\frac{\partial A^{-1}f}{\partial \sigma_l} = -A^{-1} \frac{\partial A}{\partial \sigma_l} A^{-1}f = -A^{-1} \frac{\partial A}{\partial \sigma_l} b,$$

y por consiguiente,

$$\tilde{\mathcal{C}} \frac{\partial b}{\partial \sigma_l} = -\tilde{\mathcal{C}} A^{-1} \frac{\partial A}{\partial \sigma_l} b. \quad (2.38)$$

Recordar que la matriz A es una matriz por bloques definida por

$$\begin{pmatrix} B & C \\ C^T & G \end{pmatrix}, \quad (2.39)$$

donde únicamente la matriz B depende de σ , y está dada por

$$B_{i,j} = \int_{\Omega} \sigma \nabla \varphi_i \cdot \nabla \varphi_j dx + \sum_{l=1}^L \int_{e_l} \varphi_i \varphi_j dS, \quad 1 \leq i, \quad j \leq N_n.$$

Así, la parcial de la matriz A con respecto a σ_l es nuevamente una matriz por bloques que vale 0 en casi todas las entradas, dicha parcial viene dada por

$$\frac{\partial A}{\partial \sigma_l} = \begin{pmatrix} \frac{\partial B}{\partial \sigma_l} & 0 \\ 0 & 0 \end{pmatrix}, \quad (2.40)$$

donde las entradas de $\frac{\partial B}{\partial \sigma_l}$ están definidas como:

$$\frac{\partial B(i,j)}{\partial \sigma_l} = \frac{\partial (\int_{\Omega_l} \sigma \nabla \varphi_i \cdot \nabla \varphi_j dx dy)}{\partial \sigma_l}. \quad (2.41)$$

Notar que el cálculo de (2.38) es equivalente a calcular

$$\tilde{\mathcal{C}} \frac{\partial b}{\partial \sigma_l} = \tilde{\mathcal{C}} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \mathcal{C} a_2, \quad (2.42)$$

donde el vector $a = (a_1, a_2)^T$ está determinado por la solución del sistema de ecuaciones

$$Aa = -\frac{\partial A}{\partial \sigma_l} b.$$

Por (2.40) y (2.44), el sistema de ecuaciones anterior es equivalente a

$$Ba_1 + Ca_2 = -\frac{\partial B}{\partial \sigma_l} \alpha \quad (2.43)$$

$$C^T a_1 + Da_2 = 0. \quad (2.44)$$

Multiplicando (2.43) por la matriz $C^T B^{-1}$, y utilizando la ecuación (2.44) llegamos a que es suficiente resolver el sistema

$$(D - C^T B^{-1} C) a_2 = C^T \frac{\partial B}{\partial \sigma_l} \alpha \quad (2.45)$$

para poder determinar el vector a_2 , que a su vez por (2.42) determina en su totalidad la columna de la matriz Jacobiana que deseamos calcular.

Implementación numérica de la matriz jacobiana.

Notemos que para poder calcular la matriz Jacobiana del mapeo 2.36 es necesario contar con el vector α , el cual es obtenido mediante la resolución del problema directo con elemento finito. Como mencionamos en la sección 2.5, en Zamora [14] se encuentra una manera de crear las submatrices que definen la matriz de ajuste y resolver el problema directo, dado que no es el objetivo de esta tesis profundizar en la solución del problema directo, solo mencionaremos algunas consideraciones planteadas en dicha referencia que nos serán de utilidad para el cálculo de la matriz Jacobiana.

- 1) Considerando una malla formada por elementos triangulares $\{T_k\}$, los nodos, los elementos de la base φ_i y los triángulos T_k pueden ser escogidos y enumerados de manera de que $\varphi_i(x) = 0$ para cualquier $x \in T_k$ siempre que $i \notin \{k_1, k_2, k_3\}$, donde $\{k_1, k_2, k_3\}$ son los vértices del triángulo T_k .
- 2) La integral $\int_{\Omega_l} \sigma \nabla \varphi_i \cdot \nabla \varphi_j dx dy$ es aproximada mediante

$$\int_{\Omega_l} \sigma \nabla \varphi_i \cdot \nabla \varphi_j dx dy \approx \sum_{T \in \mathcal{T}} \int_T \sigma_T \nabla \varphi_i \cdot \nabla \varphi_j dx dy, \quad (2.46)$$

donde \mathcal{T} es el conjunto de todos los triángulos en la malla y σ_T es el valor de σ en el centroide del triángulo T .

- 3) Cada electrodo e_l es representado como la unión de poligonales en la frontera de la malla

$$[N_{r_1}^l, N_{r_2}^l] \cup [N_{r_2}^l, N_{r_3}^l] \cup \dots \cup [N_{r_{m-1}}^l, N_{r_m}^l],$$

por lo que la integral $\int_{e_l} \varphi_i dS$ puede ser aproximada mediante

$$\int_{e_l} \varphi_i dS \approx \sum_{q=1}^{m-1} \int_{[N_{r_q}^l, N_{r_{q+1}}^l]} \varphi_i dS.$$

- 4) Si suponemos que los n_b nodos que forman la frontera de la malla son acomodados de manera que sean los primeros en la enumeración en la lista total de nodos, entonces por la observación 1) y 4), las entradas de la matriz C se anularán para $i > n_b$, esto nos permite escribir a C como una matriz por bloques de la forma

$$C = (C^1, \mathbf{0})^T,$$

con $C^1 \in \mathbb{R}^{n_b \times (L-1)}$.

Notar que por la observación 2), aproximar la parcial 2.41 es una tarea muy sencilla, pues por la linealidad de la derivada dicha parcial viene dada por

$$\frac{\partial B(i, j)}{\partial \sigma_l} = \frac{\partial (\int_{\Omega_l} \sigma \nabla \varphi_i \cdot \nabla \varphi_j dx dy)}{\partial \sigma_l} = \frac{\partial (\sum_{T \in \mathcal{T}} \int_T \sigma_T \nabla \varphi_i \cdot \nabla \varphi_j dx dy)}{\partial \sigma_l} = \int_{T_l} \nabla \varphi_i \cdot \nabla \varphi_j dx dy$$

donde T_l es el l -ésimo triángulo de la malla. Así, por la observación 1), la matriz $\frac{\partial B}{\partial \sigma_l}$ es una matriz rala cuyas únicas entradas distintas de 0 son las $b_{i,j}$ tales que $\{i, j\} \subset \{l_1, l_2, l_3\}$, donde $\{l_1, l_2, l_3\}$ son los elementos que conforman los vértices del triángulo T_l .

Así, notemos que la matriz $\frac{\partial B}{\partial \sigma_l}$ puede escribirse como una matriz por bloques de la forma

$$\frac{\partial B}{\partial \sigma_l} = \begin{pmatrix} \mathbf{0} \\ B^1 \end{pmatrix}$$

donde $\mathbf{0} \in \mathbb{R}^{l \times N_n}$ y $B^1 \in \mathbb{R}^{N_n - l \times N_n}$, a partir de esto y de la observación 4), es fácil ver que si $l > n_b$ entonces el lado derecho del sistema (2.45) es idénticamente 0, y por consiguiente

$$\frac{\partial U^h}{\partial \sigma_l} = \mathbf{0}.$$

En base a lo discutido en esta sección, la manera de calcular la l -ésima columna de la matriz Jacobiana queda resumida en el siguiente cuadro:

Algoritmo 1 Cálculo de la l -ésima columna de la matriz jacobiana

Entrada: Información de la malla, valores de impedancia y voltaje, conductividad σ , número de columna a calcular l .

Salida: l -ésima columna de la matriz jacobiana.

- 1: $L \leftarrow$ Número de electrodos
 - 2: $D \leftarrow 0 \in \mathbb{R}^L$
 - 3: $n_b \leftarrow$ Número de nodos en la frontera
 - 4: **if** $l \leq n_b$ **then**
 - 5: Lista los vértices del triángulo T_l como $\{N_{l_1}, N_{l_2}, N_{l_3}\}$
 - 6: Arma las matrices B, C, D y calcula α mediante el método de elemento finito.
 - 7: $F \leftarrow 0 \in \mathbb{R}^{n_b}$
 - 8: **for** $i \in \{N_{l_1}, N_{l_2}, N_{l_3}\}$ **do**
 - 9: **for** $j \in \{N_{l_1}, N_{l_2}, N_{l_3}\}$ **do**
 - 10: $F[i] \leftarrow F[i] + \alpha[j] \cdot \int_{T_l} \nabla \varphi_i \cdot \nabla \varphi_j dx dy$
 - 11: **end for**
 - 12: **end for**
 - 13: $D \leftarrow \tilde{C}(D - C^T B^{-1} C)^{-1} C^T F$
 - 14: **end if**
 - 15: **return** D
-

La ventaja del algoritmo propuesto, es que todas las matrices a usarse han sido armadas previamente en la resolución del problema directo, de igual forma, el algoritmo propuesto por Zamora [14] implementa la inversa de la matriz $(D - C^T B^{-1} C)$, por lo que el método que proponemos reutiliza los cálculos hechos anteriormente, de igual manera minimiza las multiplicaciones matriz-vector, y por consiguiente calcula la matriz Jacobiana en un tiempo aceptable para su implementación.

CAPÍTULO 3

RESULTADOS

3.1. Datos de la simulación numérica

En esta sección se presentan los resultados obtenidos de la implementación en Python de los métodos discutidos en el capítulo anterior.

Primero mostraremos el caso cuando se desea recobrar la conductividad de un dominio usando una discretización de la función de conductividad, seguido de los resultados obtenidos en el cálculo del subespacio informado por la verosimilitud.

Para las simulaciones numéricas consideramos al dominio Ω como un círculo de radio 1 centrado en el origen. Se colocaron 8 electrodos equiespaciados y de longitud constante de tal manera que ocuparan el 50% de la frontera del dominio. Tomaremos todos los valores de la impedancia de contacto efectivo como $z_l = 0.1$. Por lo expuesto en la sección 2.8, podemos estimular a los electrodos hasta con 7 diferentes patrones de corriente linealmente independientes de tal manera que estos aporten información que no sea redundante a las observaciones, en nuestro caso estimularemos pares adyacentes de electrodos con un valor de 0.1, es decir, nuestros patrones de corriente serán $I = (I^{(j)})_{j=1}^7 \in \mathbb{R}^{8 \times 7}$ donde cada $I^{(j)}$

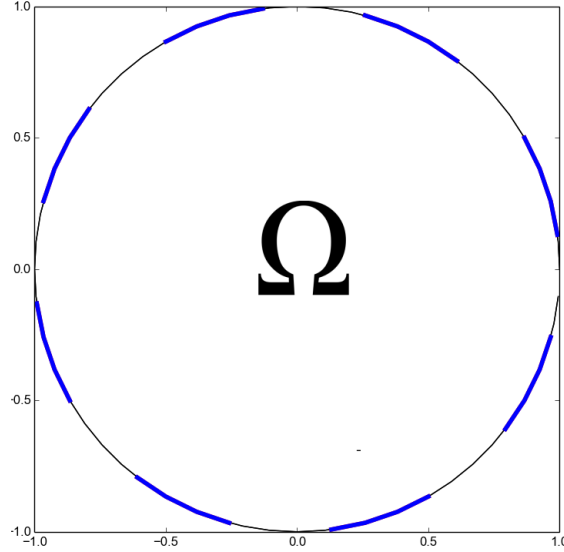


Figura 3.1: Dominio considerado en las simulaciones.

está definido por la j -ésima columna de la matriz

$$I = 0.1 \times \begin{pmatrix} +1 & 0 & \cdots & 0 \\ -1 & +1 & \cdots & 0 \\ 0 & -1 & \ddots & 0 \\ \vdots & \vdots & \ddots & +1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Usando las especificaciones mencionadas anteriormente se generaron datos sintéticos para distintas funciones de conductividad σ_i que serán presentadas más adelante. Para la solución numérica del problema directo se construyó una malla de 484 triángulos usando la paquetería de Python `meshpy`, de tal manera que cada electrodo ocupara 4 nodos de la frontera. Con el propósito de evitar el *crimen inverso*, los datos sintéticos fueron generados usando una malla más fina y agregando un ruido Gaussiano con media 0 y desviación estándar igual al 1% de la media de los voltajes medidos en los electrodos.

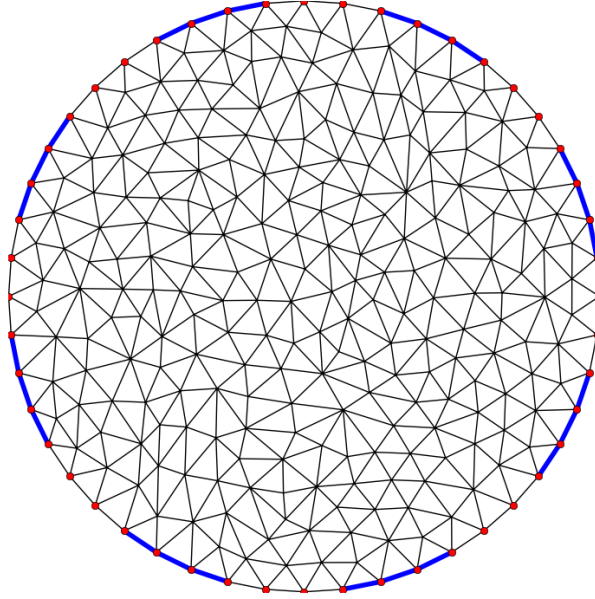


Figura 3.2: Malla utilizada en la solución numérica del problema directo.

3.2. Caso discreto

Como se vio en la sección 2.8.1, una primera aproximación al problema es considerar una discretización de la función de conductividad σ como una combinación lineal de funciones indicadoras en grupos de triángulos de la malla. En esta sección se usa el kernel de transición τ -walk presentado en Christen et al. [4] para muestrear de la distribución posterior resultante. A continuación se presentan algunos casos de estudio para este enfoque del problema.

Caso de estudio 1: Inclusión en una esquina

Supongamos primero que la conductividad de Ω es constante en todo el dominio salvo en una inclusión de área igual a un octavo del área total de Ω situada en una esquina, más específicamente, consideremos que la conductividad del dominio está determinada por la función

$$\sigma_1(x) = \begin{cases} 2 & \|x - (.43, .43)\| < \frac{1}{\sqrt{8}} \\ 1 & \text{otro caso} \end{cases} .$$

Supongamos que se conoce el tamaño, forma y posición de la inclusión, de tal manera

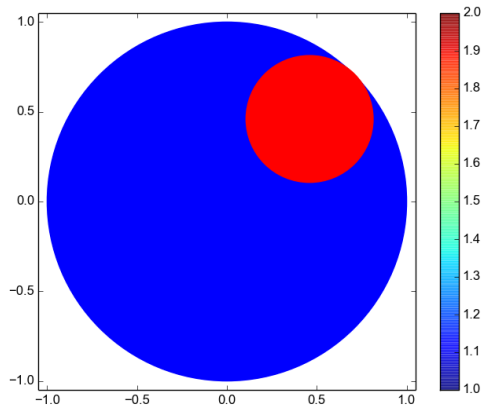


Figura 3.3: Función de conductividad σ_1

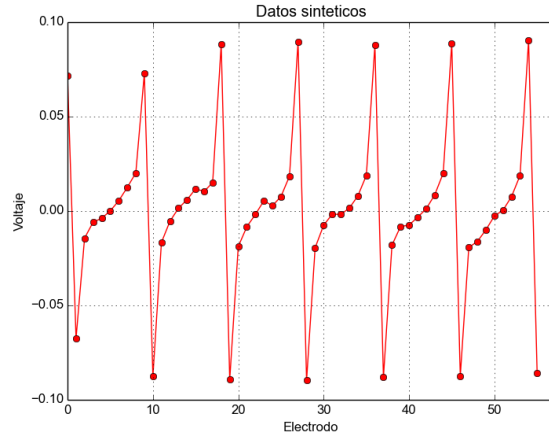


Figura 3.4: Datos sintéticos obtenidos usando la función de conductividad σ_1

que sólo falta determinar el valor de la conductividad dentro y fuera de la misma, así, si A_1 es el conjunto de triángulos que se encuentran dentro de la inclusión y A_2 los que están afuera entonces podemos escribir $\sigma_1 = s_1\chi_{A_1} + s_2\chi_{A_2}$, de tal manera que es suficiente inferir el vector $[s_1, s_2]$ para conocer en su totalidad la función de conductividad. Para este caso se consideró como a priori una normal multivariada con media $[1.5, 1.5]$ y matriz de covarianza igual a $\gamma^2 I$, $\gamma = 2$.

Para calcular el MAP, la media condicional y la covarianza condicional de la distribución posterior mediante lo expuesto en la sección 2.6.1 se construyó una caminata aleatoria de largo 70000, se desecharon las primeras 10000 iteraciones como *burnin*. En las figuras 3.5 y 3.6 se muestran la caminata resultante y los muestreos obtenidos de la distribución posterior respectivamente, en el cuadro 3.2 presentamos los estimadores obtenidos.

	s_1	s_2
Verdad	2	1
MAP	1.9796809	0.99782925
MC	1.98132635	0.99781907
Varianza	1.38281×10^{-3}	1.37297×10^{-5}

Cuadro 3.1: Estimadores para el caso de estudio 1

En el cuadro 3.2 se puede apreciar que tanto el MAP como la media condicional resultan

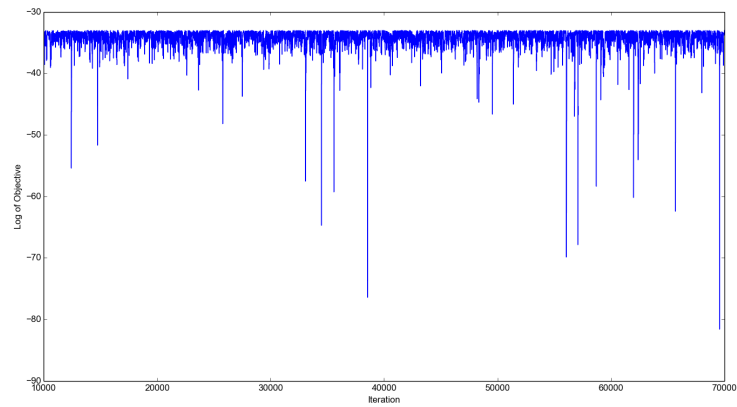


Figura 3.5: Caminata de la posterior asociada al caso de estudio 1

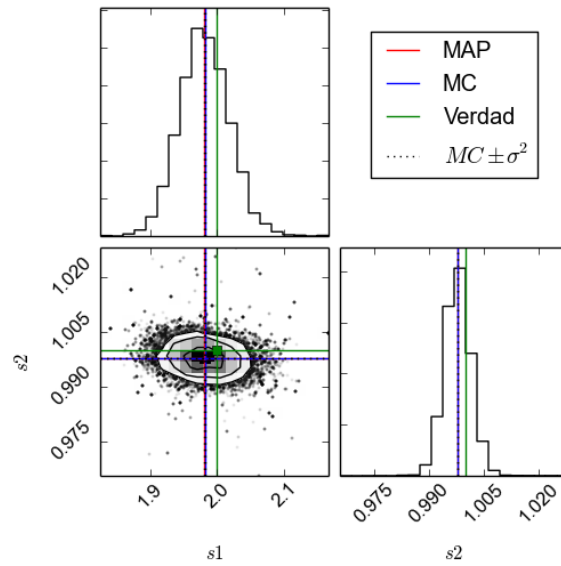


Figura 3.6: Muestreros de la distribución posterior asociada al caso de estudio 1

estimar de una manera aceptable ambos parámetros con una varianza muy pequeña. Para este caso obtuvimos convergencia para la caminata aleatoria en un número relativamente pequeño de pasos, esto debido a que la dimensión en la que planteamos el problema inverso es muy baja (igual a 2). La posterior resultante resulta ser unimodal y la mayor parte de la densidad se encuentra concentrada en un porción muy pequeña del espacio, hecho que refleja el bajo orden de la varianza.

Las imágenes 3.7, 3.8, 3.9 y 3.10 comparan la distribución a priori con la distribución posterior marginal a cada parámetro. Vale la pena recalcar la importancia que tiene la verosimilitud en la posterior, de tal manera que la información proporcionada por la a priori termina siendo prácticamente irrelevante en comparación con la posterior, lo cual refleja que para este caso ambos parámetros son bien informados por los datos.

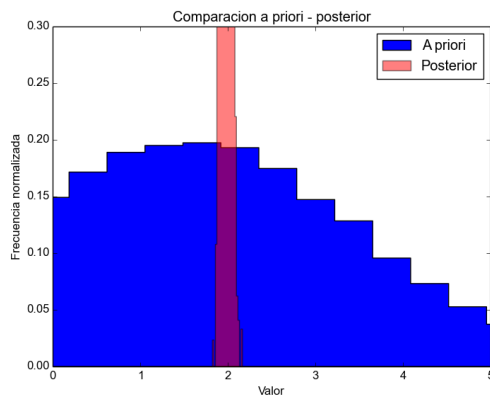


Figura 3.7: Comparación entre a priori y la posterior marginal al parámetro s_1 para el caso de estudio 1 (1)

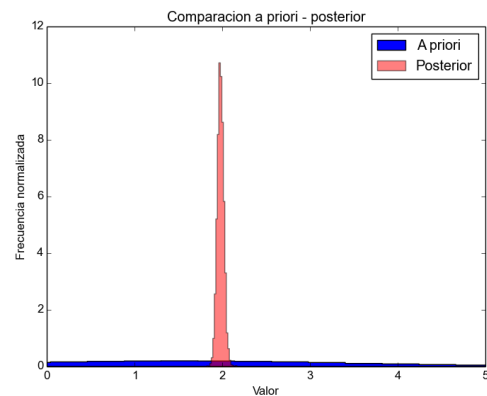


Figura 3.8: Comparación entre a priori y la posterior marginal al parámetro s_1 para el caso de estudio 1 (2)

Caso de estudio 2: Inclusión grande en el centro

Consideremos ahora el caso en que la función de conductividad es constante en Ω salvo en una inclusión grande centrada en el origen, es decir, consideremos la función de conductividad

$$\sigma_2(x) = \begin{cases} 2 & \|x - (0,0)\| < .8 \\ 1 & \text{otro caso} \end{cases} .$$

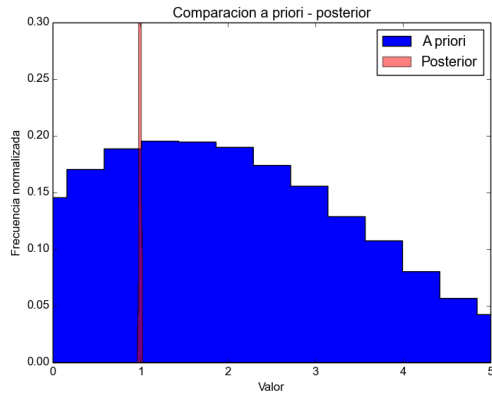


Figura 3.9: Comparación entre a priori y la posterior marginal al parámetro s_2 para el caso de estudio 1 (1)

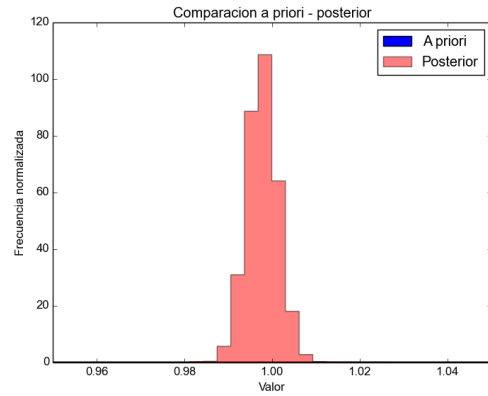


Figura 3.10: Comparación entre a priori y la posterior marginal al parámetro s_2 para el caso de estudio 1 (2)

Como en el caso de estudio anterior, supondremos que el tamaño y ubicación de la inclusión son conocidos. Así, escribiendo $\sigma_2 = s_1\chi_{A_1} + s_2\chi_{A_2}$, con A_1 igual al conjunto de triángulos que se encuentran dentro de la inclusión y A_2 los que están afuera, nuevamente es suficiente inferir el vector $[s_1, s_2]$ para determinar la función de conductividad. Utilizaremos la a priori L_1 con

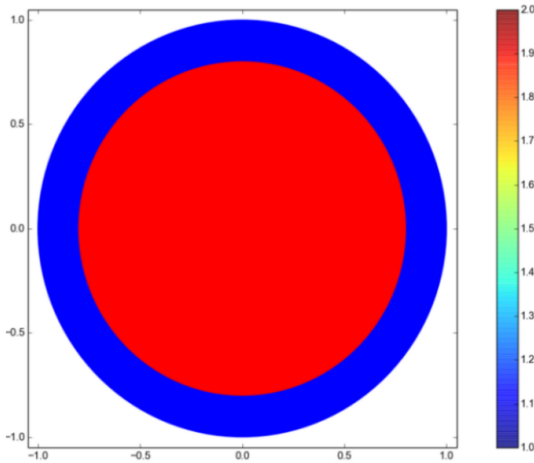


Figura 3.11: Función de conductividad σ_2

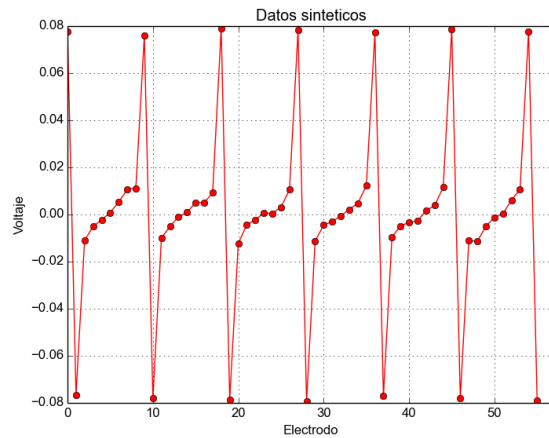


Figura 3.12: Datos sintéticos obtenidos usando la función de conductividad σ_2

un parámetro $\alpha = (1/2)^2$. Nuevamente se generó una caminata aleatoria de largo $N = 70000$ con un burnin de 10000 iteraciones. La caminata resultante, así como los muestreos de la posterior de ambos parámetros, se muestran en las figuras 3.13 y ?? respectivamente. En el cuadro 3.2 se muestran los estimadores resultantes.

	s_1	s_2
Verdad	2	1
MAP	1.99375403	0.99835375
MC	1.9955397	0.99828379
Varianza	1.67124×10^{-3}	8.88723×10^{-5}

Cuadro 3.2: Estimadores para el caso de estudio 2

De las primeras cosas que podemos notar con los resultados obtenidos en este caso es que tanto el MAP como la media condicional resultan estimar mejor al verdadero valor de $[s_1, s_2]$ que en el caso de estudio 1. De igual manera se puede ver en la figura 3.14 que la posterior resultante es un poco más alargada que antes. Esta última observación puede ser interpretada como que la distribución a priori tiene un mayor impacto sobre la posterior que en el caso de estudio pasado, pero como antes, se puede verificar en las imágenes 3.15, 3.16, 3.17 y 3.18 que la verosimilitud aporta mucha más información que la a prior, cosa muy favorable para nuestros fines. Con el fin de explorar que tanto influye la distribución a priori en la posterior resultante se resolvió de nuevo el problema inverso para la distribución σ_2 pero con la a priori Gaussiana usada en el caso 1 y también usando la distribución L_1 pero ahora con la diferencia de que no se fijo el parámetro α , sino que se le impuso una distribución a priori Gamma $Ga(1, 1)$ y se infirió junto con el vector $[s_1, s_2]$. En el cuadro 3.2 se comparan los resultados obtenidos para las distintas a prioris.

	A priori L_1		A priori Gaussiana		A priori L_1 con α variable	
	s_1	s_2	s_1	s_2	s_1	s_2
MAP	1.99375403	0.99835375	1.99341092	0.99836542	1.97674097	1.00693054
MC	1.9955397	0.99828379	1.9976902	0.9979259	1.99445753	0.99878406
Varianza	1.671×10^{-3}	8.887×10^{-5}	1.819×10^{-3}	9.821×10^{-5}	1.908×10^{-3}	9.920×10^{-5}

Cuadro 3.3: Estimadores para el caso σ_2 con diferentes a prioris

En el cuadro se puede apreciar la poca discrepancia que hay para la a priori en la media condicional. Únicamente el MAP varía con respecto a las otras dos a prioris para el caso cuando manejamos el α variable.

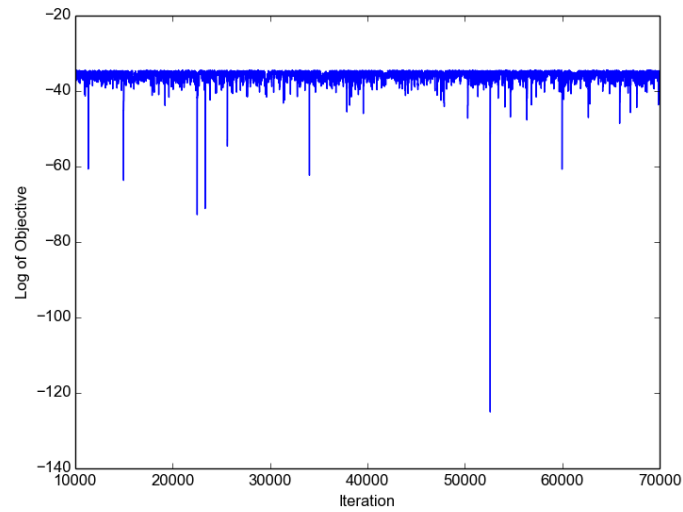


Figura 3.13: Caminata de la posterior asociada al caso de estudio 2

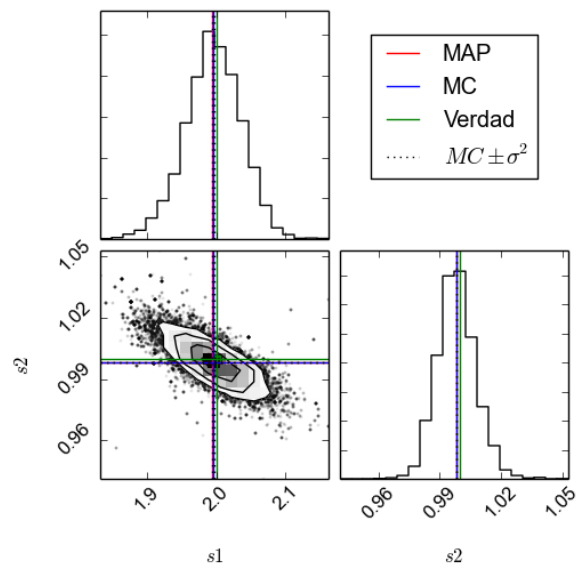


Figura 3.14: Muestras de la distribución posterior asociada al caso de estudio 2

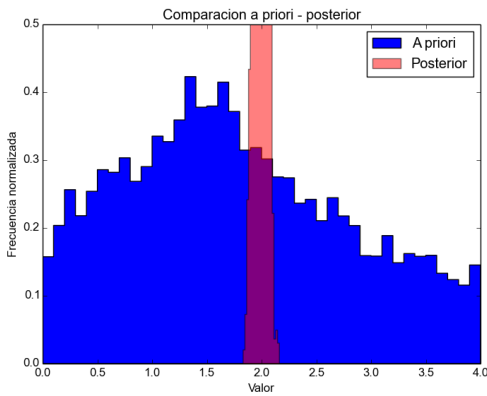


Figura 3.15: Comparación entre a priori y la posterior marginal al parámetro s_1 para el caso de estudio 2 (1)

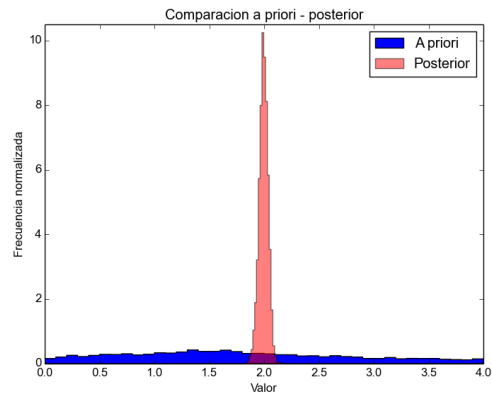


Figura 3.16: Comparación entre a priori y la posterior marginal al parámetro s_1 para el caso de estudio 2 (2)

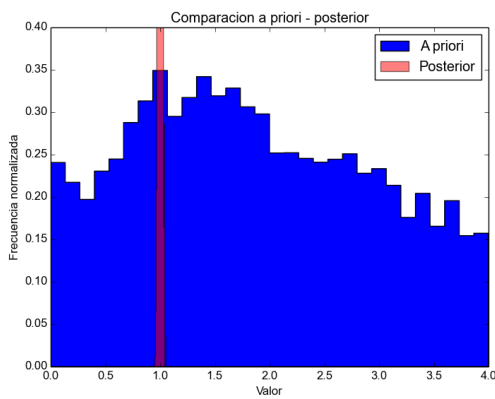


Figura 3.17: Comparación entre a priori y la posterior marginal al parámetro s_1 para el caso de estudio 2 (1)

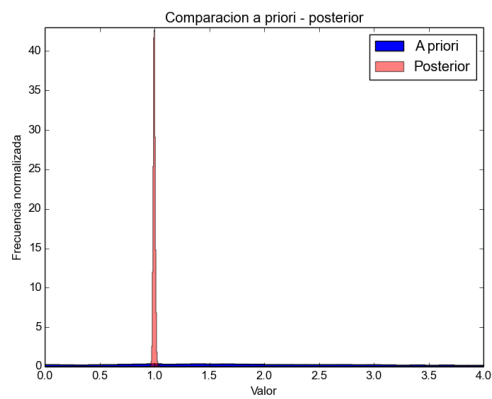


Figura 3.18: Comparación entre a priori y la posterior marginal al parámetro s_2 para el caso de estudio 2 (2)

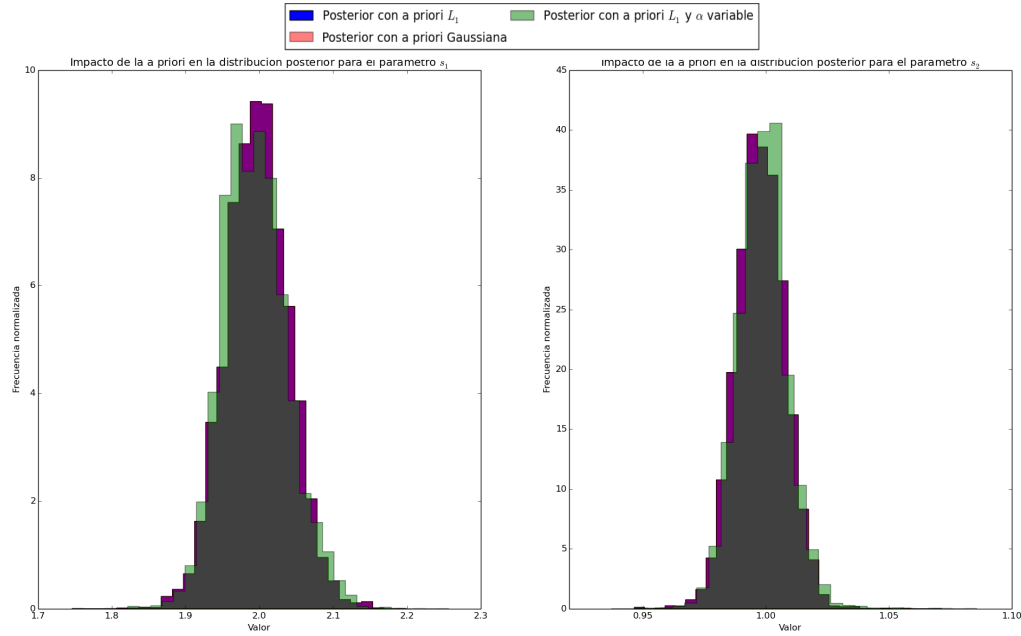


Figura 3.19: Comparación de la posterior a los parámetros con las diferentes a priori

En la figura 3.19 se comparan las distribuciones posteriores marginales para las distintas a priori. Notemos que el caso L_1 y el caso Gaussiano son indistinguibles, cuando tomamos el α variable se tiene que la posterior cambia ligeramente con respecto a las otras dos. Esto refleja nuevamente la poca información con la que contribuye la a priori con respecto a la verosimilitud para este caso de estudio.

En la figura 3.20 se muestran los muestreos de la α , podemos apreciar que la posterior se pega mucho más a 0 que la a priori, favoreciendo valores pequeños de este parámetro.

Caso de estudio 3: Inclusión en una esquina con centro variable

Con el fin de estudiar que pasa conforme aumenta la dimensión del problema, consideraremos el caso en que el centro de la inclusión no es conocido para la conductividad σ_2 .

Sea

$$A(x, y) = \left\{ (a_1, a_2) \in \Omega \mid \|(a_1, a_2) - (x, y)\| < \frac{1}{\sqrt{8}} \right\},$$

es decir, el conjunto de puntos en Ω tales que están en el círculo de radio $\frac{1}{\sqrt{8}}$ centrado en (x, y) . Así, como antes, escribamos a la función de conductividad como $\sigma_2 = s_1\chi_{A(x,y)} + s_2\chi_{\Omega \setminus A(x,y)}$,

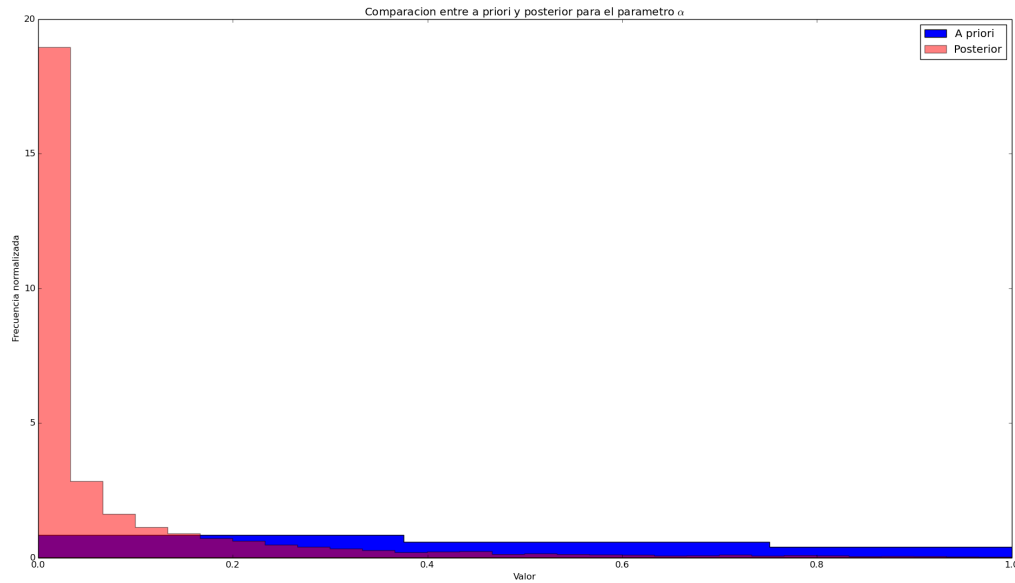


Figura 3.20: Comparación entre a priori y posterior marginal a hiperparámetro α

de manera que buscamos inferir el vector $[x, y, s_1, s_2]$. Como distribución a priori utilizamos normales con media 1.5 y desviación estándar 1 para los valores de conductividad, y para el centro se utilizó una distribución uniforme en el rectángulo $[-1, 1]^2$. Se simuló una cadena de largo $N = 100000$ con un *burnin* de 30000. En el cuadro 3.2 se muestran los resultados obtenidos, en las imágenes 3.21 y 3.22 se muestran respectivamente los muestreos de la posterior marginal, así como una representación gráfica de los centros obtenidos.

	x	y	s_1	s_2
Verdad	.43	.43	2	1
MAP	0.44664601	0.43165962	1.97700355	1.00036626
MC	0.43520775	0.42719285	1.97963443	0.99965322

Cuadro 3.4: Estimadores para el caso de estudio 3

Notar que aunque ahora tenemos un mayor número de parámetros por inferir los estimadores siguen aproximando de una buena manera a los valores reales de la conductividad. La varianza para este caso sigue siendo pequeña. Para este caso no presentamos comparación del impacto de la verosimilitud en la posterior para los parámetros del centro, esto debido a

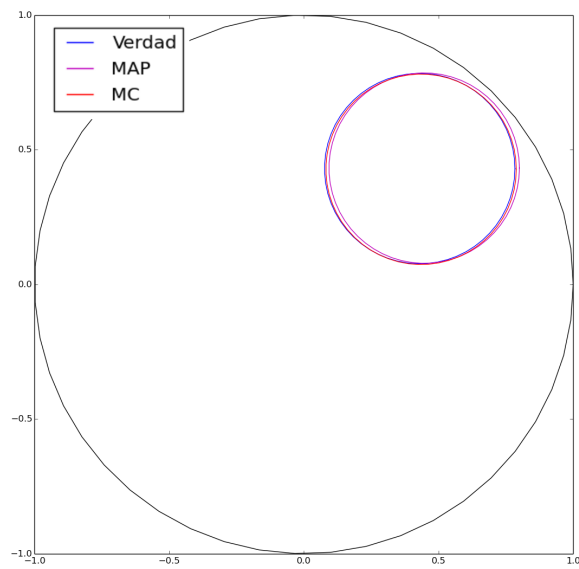


Figura 3.21: Centro de la inclusión para caso 4

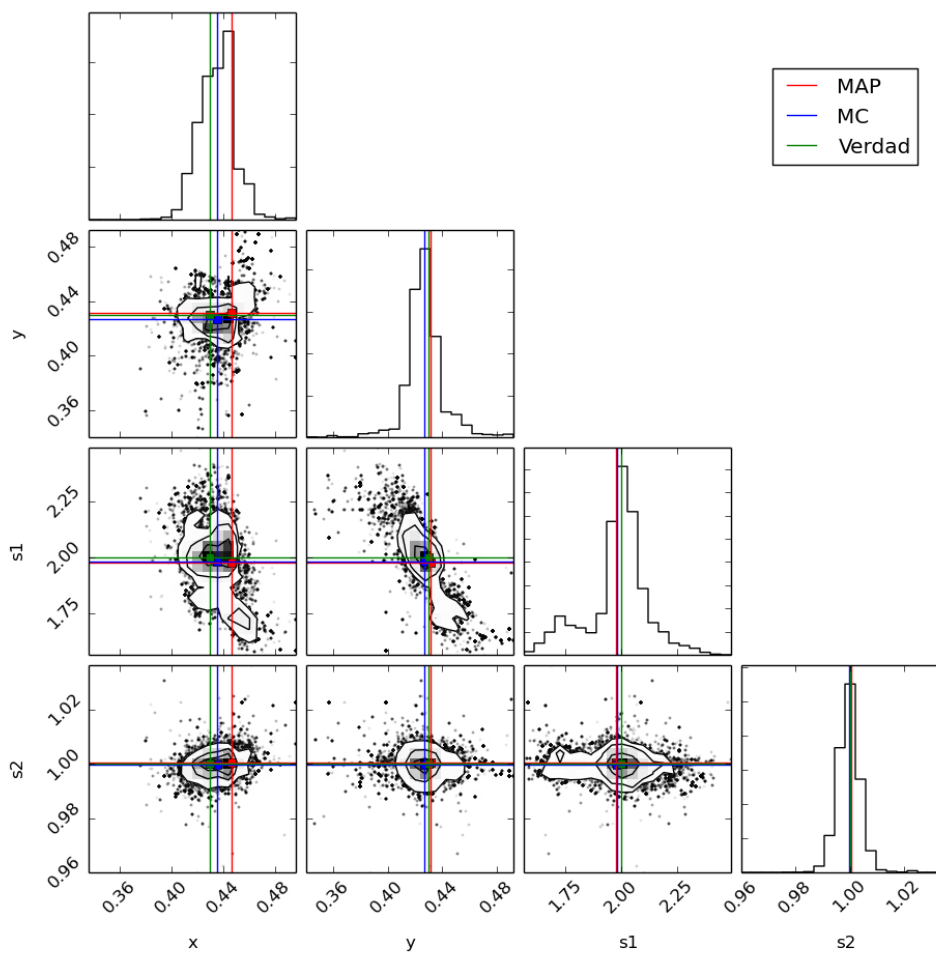


Figura 3.22: Muestras de la distribución posterior asociada al caso 4

que, ya que las a priori fueron uniformes, es claro que estos parámetros si son informados por los datos.

Caso de estudio 4: Inclusión en una esquina con dimensión alta

Tomaremos nuevamente como caso de estudio la conductividad σ_1 planteada anteriormente, pero ahora el problema será planteado en dimensión alta. Esta manera de abordar el problema será particularmente útil cuando estudiemos la dimensión efectiva del mismo.

Con el fin de formar los pixeles, se agruparon los triángulos de la malla utilizada en el elemento finito en grupos de 4 como se muestra en la figura 3.23. En total se formaron 121 pixeles, de manera que nuestra conductividad será inferida mediante el vector $[s_1, s_2, \dots, s_{121}]$, donde los elementos s_l son tales que

$$\sigma_1 = \sum_{l=1}^{121} s_l \chi_{A_l}$$

y A_l son los triángulos que conforman el pixel l . Notar que como la dimensión del problema inverso es ahora 121 se necesita una caminata aleatoria más larga que las anteriores con tal de conseguir convergencia y poder obtener resultados significativos, esto se reflejó en un gran costo computacional a la hora de construir la cadena. Supongamos que estamos en el caso en que se conoce que la conductividad en el cuerpo debe ser constante e igual a 1, esto lo reflejamos utilizando una distribución a priori Gaussiana tal que cada pixel tuviera media 1 y desviación estándar .5, de tal manera que deseamos recuperar casi desde cero la forma y posición de la inclusión en el tejido, así como el valor de la conductividad en el mismo. Se construyó una cadena de largo $N = 1000000$, la cual es mostrada en la figura 3.24.

La media condicional que se obtuvo es presentada de manera gráfica en la figura 3.25. Los pixeles que están fuera de la inclusión se conservaron el el valor dado por la media de la a priori, a excepción de algunos colocados en la esquina inferior izquierda. Para los pixeles que están dentro o cerca de la inclusión se acercaron al valor esperado, sin embargo apreciamos que los pixeles más cercanos a los electrodos obtuvieron un valor más cercano al verdadero que los que están lejos, este hecho será ilustrado nuevamente en el caso de estudio 5. Aunque

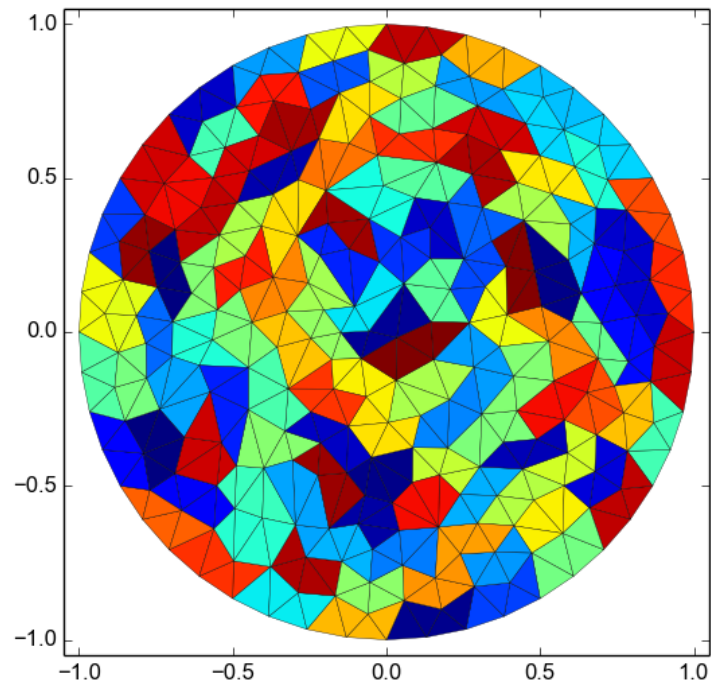


Figura 3.23: Grupos de triángulos para los pixeles

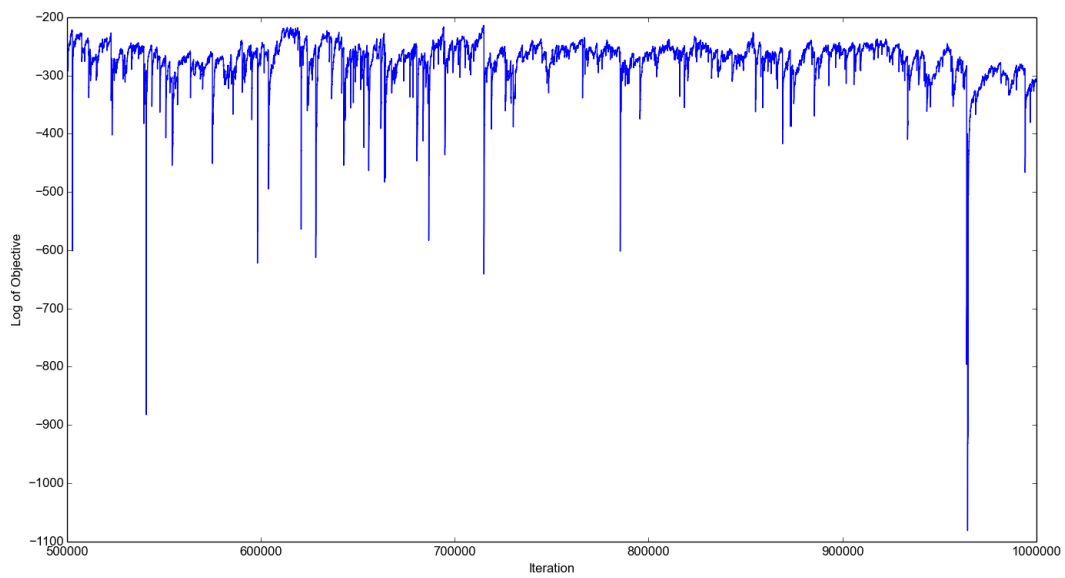


Figura 3.24: Caminata de la posterior asociada al caso de estudio 4

no se logró recuperar completamente la verdadera conductividad, la media condicional deja en evidencia la existencia de una anomalía en el área donde se encuentra la inclusión, aunque también con la incertidumbre de los pixeles localizados en la esquina inferior izquierda, que parecen mostrar otra anomalía que es inexistente.

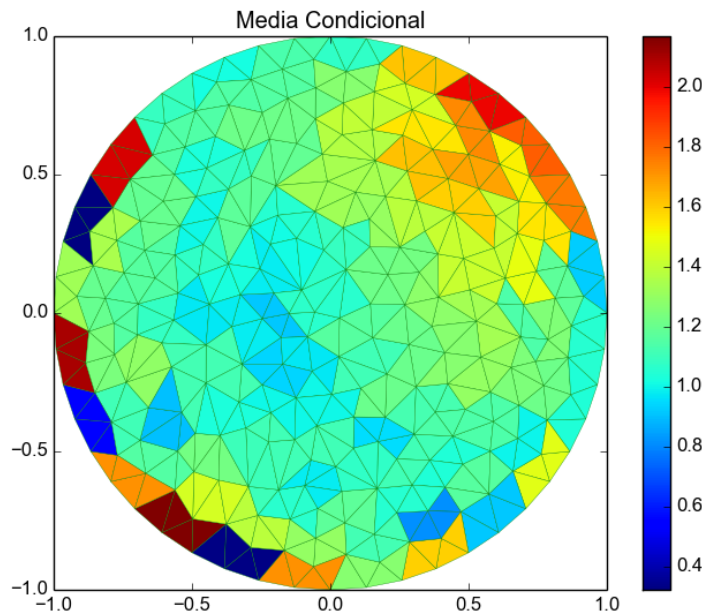


Figura 3.25: Media condicional de la posterior asociada al caso de estudio 4

Para dar una idea de como lucen los muestreos de la posterior marginal a los pixeles se presentan en las figuras 3.27, 3.28 3.29 y 3.30 las frecuencias normalizadas de los pixeles 30, 20, 15 y 27, los cuales son resaltados en la figura 3.26. En dichas figuras también se muestra el impacto que tiene la verosimilitud en la posterior marginal, así como el verdadero valor del pixel contrastado con la aproximación obtenida por la media condicional. Notar que el pixel 30, que es el que está fuera de la inclusión pero más alejado de los electrodos es el que presenta menos discrepancia de la a priori en comparación con la posterior, cosa que no pasaba en los casos de estudio anteriores. En estas imágenes también se aprecia lo mencionado anteriormente sobre los pixeles colocados en la esquina inferior izquierda dan una estimación no tan certera de la conductividad.

Es de resaltarse el hecho de que ya no se presenta un gran contraste entre a priori y

la posterior marginal como era en los casos cuando se planteó el problema en dimensión baja. Esto muestra el hecho planteado en Cui et al. [5], que dice que conforme aumenta la dimensión del problema el impacto de la a priori se vuelve cada vez mayor.

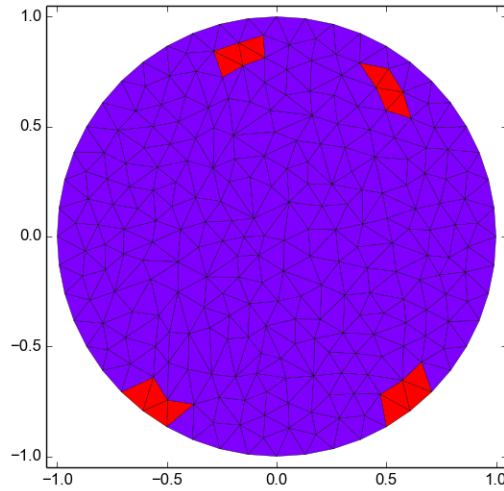


Figura 3.26: Pixeles 30, 20, 15 y 27

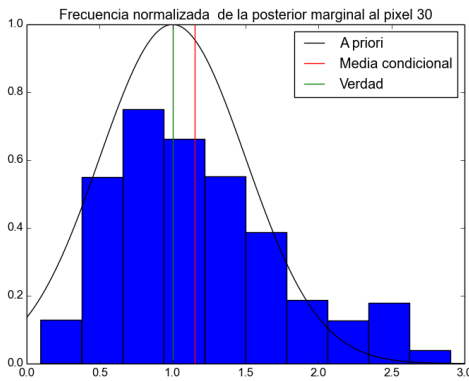


Figura 3.27: Comparación entre a priori y posterior marginal al pixel 30 asociada a la conductividad σ_2

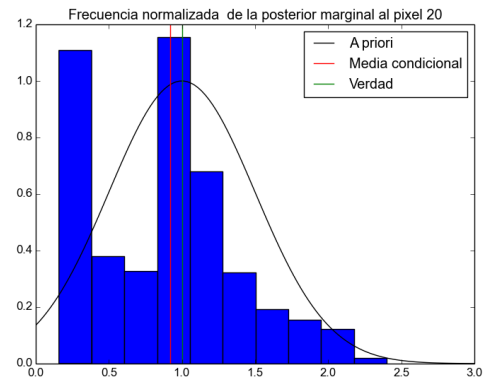


Figura 3.28: Comparación entre a priori y posterior marginal al pixel 20 asociada a la conductividad σ_2

Aprovechando esta manera de plantear el problema probaremos otra función a priori: La variación total. Como se mencionó en la sección 2.8.1, dicha a priori emplea un parámetro α que determina el peso que se le da a la a priori, este fue tomado como un hiperparámetro al que se le asignó una distribución Gamma $Ga(1, 1)$.

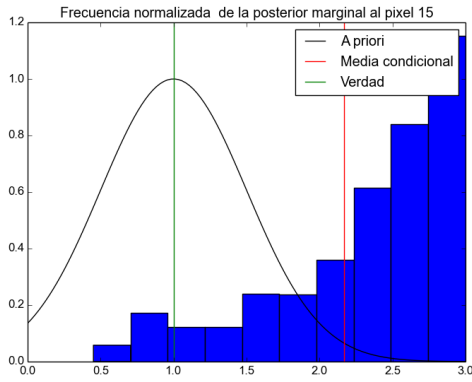


Figura 3.29: Comparación entre a priori y posterior marginal al pixel 15 asociada a la conductividad σ_2

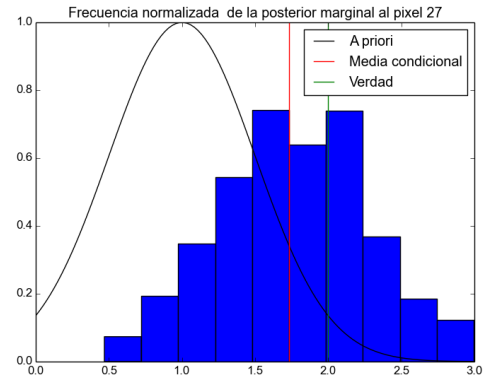


Figura 3.30: Comparación entre a priori y posterior marginal al pixel 27 asociada a la conductividad σ_2

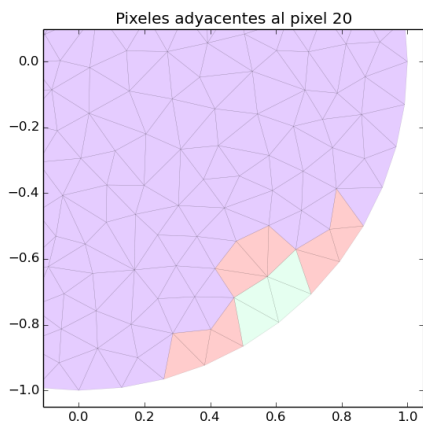


Figura 3.31: Adyacencias del píxel 20

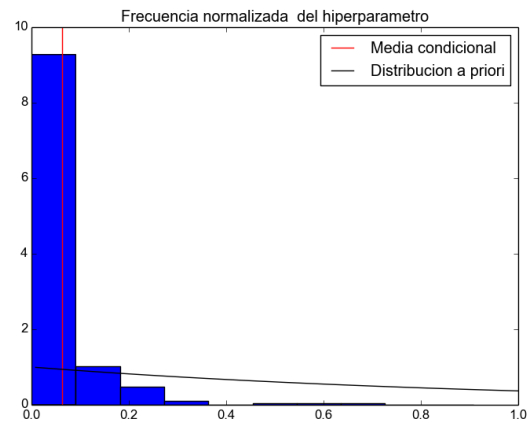


Figura 3.32: Posterior marginal al hiperparámetro

El criterio para determinar el conjunto de vecinos de un pixel fue: dos pixeles son vecinos si y sólo si poseen al menos un triángulo con un vértice en común. Para calcular la media condicional se generó una caminata aleatoria de largo 1100000.

En la figura 3.33 se muestra la media condicional asociada a la distribución posterior. Una de las características de la a priori de variación total es que favorece las funciones *blocky* Kaipio y Somersalo [9], sin embargo, a diferencia de la Gaussiana y la L_1 , no cuenta con un valor de referencia para los valores de la conductividad. Siguiendo la idea de Kaipio et al. [10], se calculó también la media condicional pero ahora fijando los primeros 24 pixeles con un valor igual a 1 como valor de referencia, para este caso sólo fue necesario inferir el vector $[s_{25}, \dots, s_{121}, \alpha]$. El resultado es presentado en la figura 3.34.

Si comparamos los resultados obtenidos para este caso de estudio, los mejores resultados fueron obtenidos utilizando la a priori Gaussiana, esto era de esperarse ya que esta a priori favorecía el verdadero valor de la conductividad fuera de la inclusión, sin embargo, los estimadores obtenidos usando la variación total también son significantes, en especial cuando se tomaron los pixeles que antes habían sido problemáticos en la inferencia fijos como un valor de referencia.

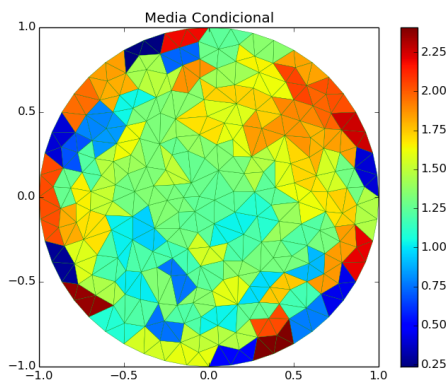


Figura 3.33: Media condicional obtenida con una a priori de variación total

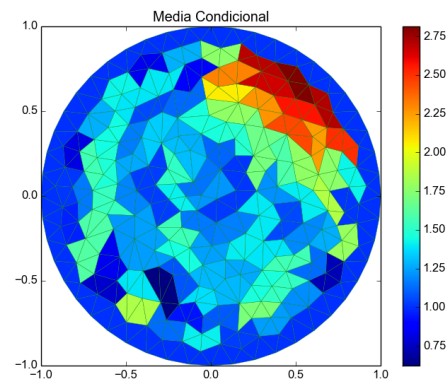


Figura 3.34: Media condicional obtenida con una a priori de variación total usando 24 pixeles de referencia

Caso de estudio 5: Inclusión pequeña en el centro

Por último consideremos de nuevo el caso en que son conocidos el tamaño, forma y posición de la inclusión, por lo que es suficiente determinar los valores de la conductividad dentro y fuera de ella, pero ahora para una inclusión muy pequeña en comparación con el dominio Ω y centrada en el origen. Consideremos la función de conductividad

$$\sigma_3(x) = \begin{cases} 2 & \|x - (0, 0)\| < .3 \\ 1 & \text{otro caso} \end{cases} .$$

El propósito de este ejemplo es ilustrar el hecho de que cuando la inclusión es pequeña, o

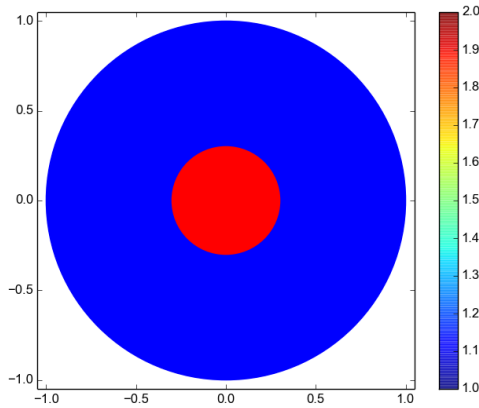


Figura 3.35: Función de conductividad σ_3

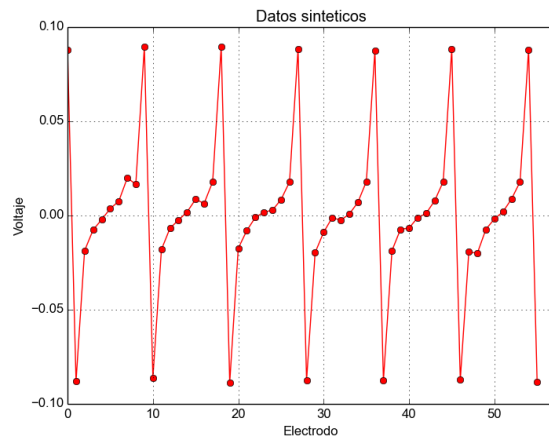


Figura 3.36: Datos sintéticos obtenidos usando la función de conductividad σ_3

está muy alejada de la frontera de Ω se vuelve más complicado recobrar la conductividad. Para ello tomemos el caso más sencillo que se ha considerado, que es, como antes, cuando se plantea $\sigma_3 = s_1\chi_{A_1} + s_2\chi_{A_2}$ donde A_1 y A_2 son los conjuntos que se han definido desde los primeros ejemplos. En el siguiente cuadro se muestran los resultados obtenidos, así como en la figura 3.37.

	s_1	s_2
Verdad	2	1
MAP	1.62593952	1.00345846
MC	1.65051992	1.00336765

Cuadro 3.5: Estimadores para el caso 5

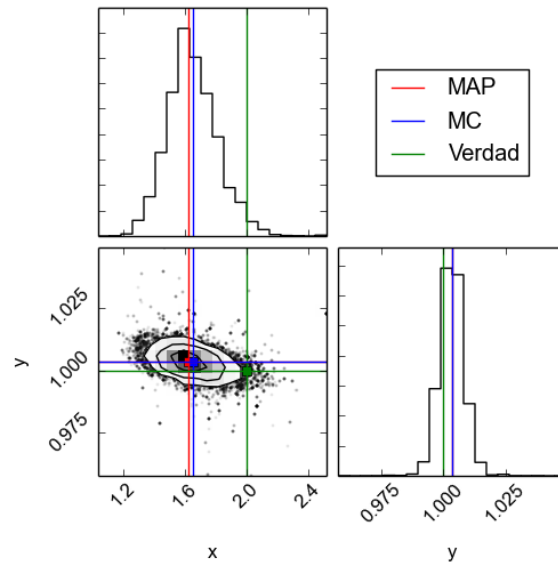


Figura 3.37: Muestreos de la distribución posterior asociada a la conductividad σ_2 para el caso de estudio 5

Como en los casos anteriores, la conductividad fuera de la inclusión fue aproximada de manera aceptable tanto por el MAP como por la media condicional, sin embargo este no fue el caso con el valor de la inclusión dentro de la inclusión. Esto concuerda con lo expuesto en la referencia Zamora [14], donde se argumenta que inclusiones pequeñas o muy en el centro del dominio causan una discrepancia muy pequeña en las mediciones de los voltajes, y por consiguiente hace difícil recuperar el valor verdadero de la conductividad.

3.3. Espacio informado por la verosimilitud

Como se vio en los casos de estudio anteriores, se hace difícil recuperar el valor de la conductividad conforme nos alejamos de la zona donde están colocados los electrodos.

Esto nos lleva a pensar que la verosimilitud no es informativa en los píxeles situados en la parte central del dominio. Con el fin de evidenciar este hecho se calculó el subespacio informado por la verosimilitud (LIS) que fue expuesto en la sección 2.9.

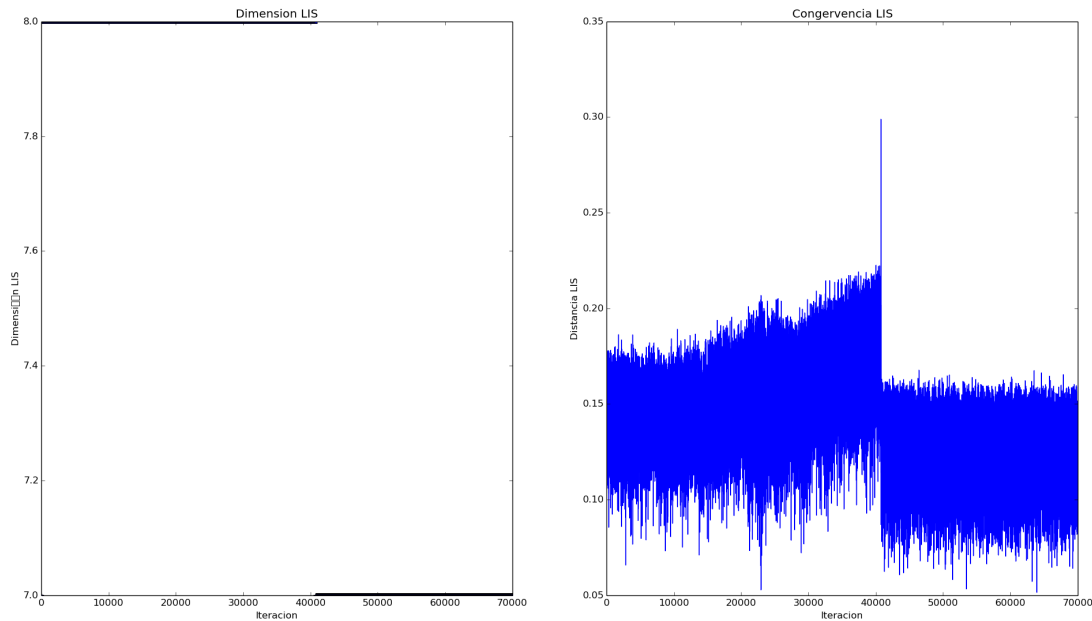


Figura 3.38: Evolución de la dimensión del LIS

Se calculó el LIS con la intención de conocer la dimensión efectiva del problema, para ello, se utilizó la misma caminata aleatoria que se mencionó en la sección 3.2 para la construcción de la aproximación Monte Carlo del Hessiano preconditionado. La caminata aleatoria fue de largo 1000000 y se empleó un burnin de 250000. Para la construcción del espacio informado por la verosimilitud se utilizó un valor de truncamiento igual al 35 % del autovalor más grande del Hessiano preconditionado evaluado en la media condicional calculada en 3.2.

En la figura 3.38 se muestra la evolución de la dimensión del LIS conforme avanza el número de elementos de la cadena utilizados en el cálculo del Hessiano preconditionado, así como el comportamiento del criterio de convergencia empleado. Notar que en realidad la dimensión del LIS no varía mucho conforme varía el número de iteraciones, sin embargo, vale la pena recalcar el hecho de que dicha dimensión resulta ser considerablemente baja con respecto a la dimensión total del espacio de parámetros que es 121. Esto coincide con lo conjeturado anteriormente, que en realidad son muy pocos los pixeles que son verdaderamente informados por los datos con respecto a la a priori. Los ejemplos prácticos presentados sugieren que los pixeles informados están situados cerca de los electrodos.

CAPÍTULO 4

DISCUSIÓN Y CONCLUSIONES

En este trabajo se planteó una estrategia de cuantificación de incertidumbre para el problema inverso de tomografía de impedancia de eléctrica desde el enfoque Bayesiano y se implemento numéricamente en Python usando métodos Markov Chain Monte Carlo para así poder inferir ciertas funciones de conductividad y estudiar la dimensión efectiva del problema.

El problema directo fue resuelto utilizando el esquema de elemento finito planteado en Zamora [14], pero implementando una malla irregular construida con la paquetería `meshpy`.

Se resolvió el problema inverso usando tanto a priori convencionales, como la Gaussiana, así como algunas otras que sacaban provecho de la geometría del problema, como la variación total, primero planteando el problema en una dimensión baja y después en una dimensión alta, para después confrontar los resultados obtenidos.

Se probaron los algoritmos elaborados utilizando diferentes funciones de conductividad. Primeramente se utilizaron funciones que son frecuentemente recurridas en la literatura (vease por ejemplo Kaipio et al. [10]) en las cuales la conductividad es constante en todo el dominio salvo en inclusiones localizadas muy cerca de la frontera del dominio, o en su defecto, las inclusiones son muy grandes en comparación con el tamaño del dominio. Para estos casos se obtuvieron resultados satisfactorios al inferir nuestra función objetivo sin importar la distribución a priori utilizada.

Para contrastar este hecho se utilizaron también funciones de conductividad tales que la inclusión no satisfacía las condiciones mencionadas anteriormente, el resultado fue que el método no fue capaz de recuperar satisfactoriamente la inclusión, esto coincide con lo expuesto en Zamora [14], donde utilizando una conductividad sin inclusiones como valor de referencia para los valores de los voltaje, observaron que las inclusiones pequeñas o alejadas de los electrodos dan origen a una discrepancia muy pequeña para los mismos. También se observó en los resultados obtenidos que conforme aumenta la dimensión del problema el impacto que tiene la a priori en la distribución posterior se vuelve cada vez mayor.

Este último comentario coincide con lo expuesto en los artículos Cui et al. [5] y Spantini et al. [12], los cuales afirman que en los problemas inversos de dimensión alta, el espacio que en realidad es informado por la verosimilitud, y por consiguiente por los datos, es muy pequeño en comparación del espacio de parámetros.

Para corroborar esto se procedió a calcular el subespacio informado por la verosimilitud, por lo cual se tuvo que implementar numéricamente el cálculo de la matriz Jacobiana del mapeo directo, para esto se utilizó un método de diferenciación adjunta. Debido a calcular la matriz Jacobiana implica resolver repetidas veces el problema directo, se propuso un algoritmo eficiente de calcular dicha matriz, intentando minimizar el número de multiplicaciones matriz-vector, esto se logró utilizando las propiedades de las matrices de forma que definen la solución del problema directo por elemento finito expuestas en Zamora [14].

Se calculó el subespacio informado por la verosimilitud del problema con la intención de conocer la dimensión de este. Como era de esperarse, la dimensión del subespacio informado por los datos resultó ser pequeña en comparación de la dimensión del espacio de parámetros. Vale la pena recalcar que esto no implica que los demás parámetros no sean influenciados por los datos, sino más bien significa que la distribución a priori domina a la posterior por encima de la verosimilitud. Así, el problema desde este planteamiento logra identificar anomalías en el dominio de estudio siempre y cuando éstas se encuentren cerca de los pixeles que si son informados por los datos. Los ejemplos presentados sugieren que dichos pixeles son aquellos que se encuentran cerca de los electrodos. Para apoyar esta conjetura se presenta en el apéndice A el caso de una conductividad que tiene una inclusión alejada de los electrodos

planteando el problema en dimensión alta, el resultado es que el método no logra identificar nada de dicha inclusión. Estos últimos comentarios dejan en manifiesto la importancia de escoger una a priori apropiada para el problema inverso de Calderón, como por ejemplo en Kaipio et al. [10], en donde primero calculan una aproximación de la conductividad mediante métodos gradiente y a a partir de esta construyen una a priori Gaussiana que tiene como media dicha aproximación.

APÉNDICE A

INCLUSIÓN EN EL CENTRO, DIMENSIÓN ALTA

A continuación se presentan los resultados obtenidos cuando buscamos inferir la función de conductividad $\sigma_3(x)$ definida en las secciones anteriores cuando el problema es planteado en dimensión alta.

En busca de inferir esta función se propuso nuevamente una a priori Gaussiana con media igual al valor de referencia de la conductividad, y una desviación estandar de .5, notar que esta es la misma a priori utilizada en el caso de estudio 4, para el cual se obtuvieron buenos resultados.

Se simuló de la posterior utilizando los mismos datos del caso de estudio 4. La media condicional obtenida es presentada en la siguiente figura.

Notar que al igual en que en el caso de estudio 5, no es posible recuperar el valor de la conductividad en la inclusión, sin embargo verlo desde esta perspectiva hace mas notorio lo crítico que es tener una inclusión alejada del origen. Esta imagen, junto con las anteriores y el hecho de que la dimensión del subespacio informado por la verosimilitud es muy baja, nos hace conjeturar que los pixeles informados por los datos son aquellos que se encuentran localizados cerca de los electrodos, cosa que era de esperarse.

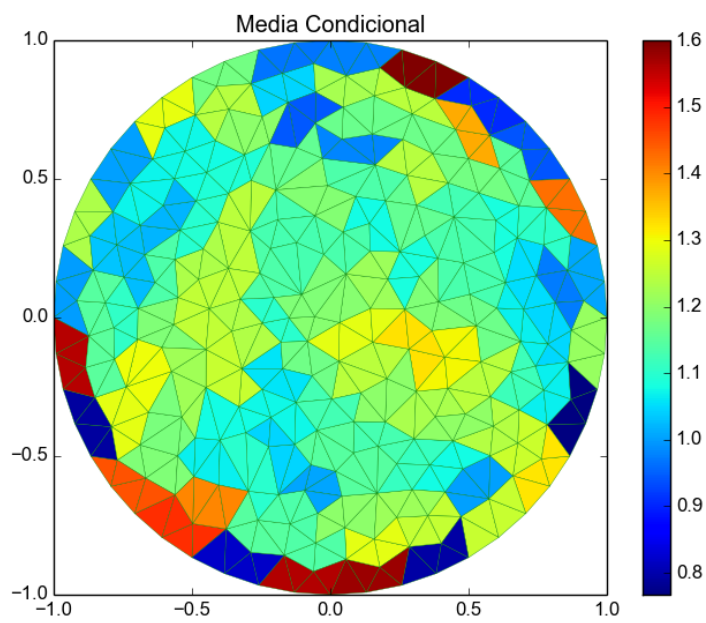


Figura A.1: Media condicional de la posterior para inclusión pequeña y dimensión alta

BIBLIOGRAFÍA

- [1] P Calderón Alberto. On an inverse boundary value problem. *Comp. Appl. Math*, 25(2-3), 2006.
- [2] Ward Cheney. *Analysis for applied mathematics*, tomo 208. Springer Science & Business Media, 2013.
- [3] Kuo-Sheng Cheng, David Isaacson, JC Newell, y David G Gisser. Electrode models for electric current computed tomography. *Biomedical Engineering, IEEE Transactions on*, 36(9):918–924, 1989.
- [4] J Andrés Christen, Colin Fox, et al. A general purpose sampling algorithm for continuous distributions (the t-walk). *Bayesian Analysis*, 5(2):263–281, 2010.
- [5] Tiangang Cui, James Martin, Youssef M Marzouk, Antti Solonen, y Alessio Spantini. Likelihood-informed dimension reduction for nonlinear inverse problems. *Inverse Problems*, 30(11):114015, 2014.
- [6] Matthew M Dunlop y Andrew M Stuart. The bayesian formulation of eit: Analysis and algorithms. *arXiv preprint arXiv:1508.04106*, 2015.
- [7] Seppanen Hyvonen y Stratos Staboulis. Optimizing electrode positions in electrical impedance tomography. *arXiv*, 2014.

-
- [8] Jacques Jossinet. The impedivity of freshly excised human breast tissue. *Physiological Measurement*, 19(1):61, 1998. URL <http://stacks.iop.org/0967-3334/19/i=1/a=006>.
- [9] Jari Kaipio y Erkki Somersalo. *Statistical and computational inverse problems*, tomo 160. Springer Science & Business Media, 2006.
- [10] Jari P Kaipio, Ville Kolehmainen, Erkki Somersalo, y Marko Vauhkonen. Statistical inversion and monte carlo sampling methods in electrical impedance tomography. *Inverse problems*, 16(5):1487, 2000.
- [11] Erkki Somersalo, Margaret Cheney, y David Isaacson. Existence and uniqueness for electrode models for electric current computed tomography. *SIAM Journal on Applied Mathematics*, 52(4):1023–1040, 1992.
- [12] Alessio Spantini, Antti Solonen, Tiangang Cui, James Martin, Luis Tenorio, y Youssef Marzouk. Optimal low-rank approximations of bayesian linear inverse problems. *SIAM Journal on Scientific Computing*, 37(6):A2451–A2487, 2015.
- [13] Päivi J Vauhkonen. *Image reconstruction in three-dimensional electrical impedance tomography*. Kuopion yliopisto, 2004.
- [14] C. Zamora. *Hacia una estrategia de cuantificación de incertidumbre para el problema de tomografía de impedancia eléctrica*. Proyecto Fin de Carrera, Centro de Investigación en Matemáticas, A.C., 2014.
- [15] Y Zou y Z Guo. A review of electrical impedance techniques for breast cancer detection. *Medical engineering & physics*, 25(2):79–90, 2003.