

Universidad Complutense de Madrid

Facultad de Ciencias Matemáticas



TESIS DOCTORAL

Métodos matemáticos y computacionales en elastografía.

Memoria para optar al grado de doctor presentada por

Carolina Abugattas Chacoff

Director/a/es:

Ana María Carpio Rodríguez

Universidad Complutense de Madrid

Facultad de Ciencias Matemáticas

Programa de Doctorado en Investigación Matemática



TESIS DOCTORAL

Métodos matemáticos y computacionales en elastografía.

Memoria para optar al grado de doctor presentada por

Carolina Abugattas Chacoff

Director/a/es:

Ana María Carpio Rodríguez

Agradecimientos

En primer lugar, deseo expresar mi profunda gratitud a mi directora de tesis, Ana Carpio, por aceptar guiarme como su estudiante de doctorado. Le agradezco su apoyo constante, las oportunidades brindadas, su paciencia infinita y, en particular, la guía estructurada y ordenada que ha sido fundamental durante estos años. Extiendo también mi agradecimiento a su equipo de trabajo, cuyo trabajo previo me permitió adquirir conocimientos esenciales para desarrollar este proyecto.

Me gustaría expresar mi agradecimiento a la profesora Elena Cebrián por la ayuda y los conocimientos que me ha brindado, especialmente durante la primera parte de esta tesis, cuando aún tenía dificultades con la implementación del problema inverso en los códigos de Matlab. Asimismo, agradezco su valiosa ayuda en la parte relacionada con las series de Fourier, en la comparación de resultados en las simulaciones numéricas, lo cual fue fundamental para depurar errores en la programación.

También agradezco al profesor Gerardo Oleaga, por su disposición para ayudar, por las reuniones realizadas para la comparación de resultados en las simulaciones numéricas, en las cuales permitía encontrar más de algún error en la programación.

También me gustaría agradecer a Pedro González, el coordinador del programa de doctorado investigación matemáticas, por responder siempre mis preguntas y ayudarme a resolver problemas administrativos. Agradezco al personal de la Facultad de Matemáticas de la Universidad Complutense por siempre querer ayudar en todo momento, lo que me permitió llevar a cabo mis estudios de doctorado.

Me gustaría agradecer al profesor Ernesto Castillo de la Universidad de Santiago de Chile, por introducirme, años atrás, sobre el método de elementos finitos y cálculo numérico, y me brindó la oportunidad de realizar una estancia formativa en su laboratorio. A los miembros de su equipo, gracias por su amabilidad y cálida recepción.

Fuera de la academia, me gustaría agradecer a mi familia, en especial a mi madre Ana María y mi tía Mercedes, que gracias al esfuerzo de ellas y por el apoyo incondicional que me han dado siempre, soy quien soy y he podido superar diferentes barreras. También me gustaría agradecer a mis primos Daniela y Jaime por su cariño y apoyo.

Me gustaría agradecer a Pedro A., por el cariño, paciencia y apoyo que siempre me has brindado a lo largo de todos estos años que hemos estado juntos, en especial durante estos últimos años; gracias por tus consejos y conversaciones de todo tipo que hemos tenido.

Agradezco a compañeros de doctorado que he conocido durante estos años dentro del programa de doctorado y fuera de éste. Agradezco también a mis amigos en Chile, que pese a la distancia, siempre me apoyan y dan ánimos: gracias por los consejos dados, conversaciones que han ayudado a desconectar y a recargar energías.

Este trabajo de tesis doctoral se ha llevado a cabo gracias al grupo de investigación de la profesora Ana Carpio Rodríguez, al Ministerio de Ciencia, Innovación y Universidades de España (Proyecto PID2020-112796RB-C21, AEI) y a la Agencia Nacional de Investigación y Desarrollo /Subdirección de Capital Humano/Doctorado Beca Chile 2022, folio 72220131.

Índice general

Agradecimientos	VII
Abstract	VII
Resumen	IX
Índice de figuras	XI
Índice de tablas	XV
1. Introducción y contexto	1
1.1. Estructura de la tesis	6
1.2. Objetivos de la tesis	8
2. Estudio teórico y numérico del problema directo	11
2.1. Formulación del problema directo	12
2.1.1. Modelo matemático directo planteado en un dominio infinito	13
2.1.2. Operador de observación	14
2.2. Estudio teórico del problema directo truncado	15
2.2.1. Formulación débil del problema directo truncado	15
2.2.2. Resultados de existencia y regularidad mediante el método de Galerkin	16
2.3. Resolución numérica del problema directo truncado: discretización en espacio	20
2.4. Resolución numérica del problema directo truncado: discretización en tiempo	21
2.4.1. Discretización en tiempo de la formulación débil	22
2.4.2. Discretización total de la formulación débil	23
2.5. Condiciones de contorno no reflectantes	24
3. Estudio teórico y numérico del problema inverso en dimensión finita	27
3.1. Problema inverso determinista en dimensión finita	28
3.2. Problema inverso bayesiano en dimensión finita	30
3.2.1. Aproximación de máximo a posteriori	31

3.2.2. Pseudocódigo para cálculo de valor máximo a posteriori	32
3.3. Cuantificación de incertidumbre mediante la aproximación de Laplace	33
3.3.1. Convergencia de la aproximación de Laplace	34
3.3.2. Definición de la aproximación de Laplace de la probabilidad posterior	35
3.4. Cuantificación de incertidumbre mediante métodos de cadenas de Markov Monte Carlo	37
3.4.1. Convergencia del Método Monte Carlo de Cadenas de Markov	37
3.4.2. El Teorema Ergódico (Ley fuerte de los grandes números)	39
3.4.3. Algoritmo de Metropolis-Hastings	39
3.4.3.1. Pseudocódigo Metrópolis-Hastings	40
3.4.4. Muestreador de conjunto invariante afín	41
3.4.4.1. Pseudocódigo de muestreo con invariancia afín	42
4. Estrategias de inversión del campo de ondas basadas en objetos	45
4.1. Operadores de observación aproximados	45
4.1.1. Geometría	46
4.1.2. Discretización espacial	48
4.1.3. Comparación de operadores de observación	49
4.2. Problema bayesiano inverso aproximado	52
4.2.1. Construcción anomalía a priori	53
4.3. Cálculo del Máximo a posteriori (MAP) por optimización	54
4.3.1. Ángulo inicial $\omega_0 = 0$	54
4.3.2. Ángulo inicial $\omega_0 = \frac{\pi}{8}$	56
4.4. Resultados del muestreo	57
4.4.1. Ángulo inicial $\omega_0 = 0$	57
4.4.2. Ángulo inicial $\omega_0 = \frac{\pi}{8}$	60
4.5. Caso aplicado a elastografía médica	63
4.5.1. Generación de datos sintéticos y operador de observación	64
4.5.2. Problema bayesiano inverso aproximado	64
4.5.3. Resultados del muestreo	65
5. Problema bayesiano inverso basado en series de Fourier	69
5.1. Geometría de domo salino	70
5.2. Resolución del problema directo y operadores de observación aproximados	72
5.2.1. Discretización espacial	73
5.2.2. Comparación de operadores de observación	75
5.3. Parametrización mediante series de Fourier: caso bidimensional	77
5.3.1. Aproximación de campos de velocidad y densidad mediante series de Fourier .	79
5.3.2. Comparación de operadores de observación aproximados	83

5.4. Formulación bayesiana inversa aproximada y resultados de muestreo	86
5.4.1. Resultados de muestreo AIES	89
6. Estrategias de inversión basadas en campos aleatorios	91
6.1. Breve recordatorio sobre variables y campos aleatorios	91
6.1.1. Teorema de Karhunen-Loève	92
6.1.2. Núcleo Matérn	93
6.1.3. Muestreo de campos aleatorios para covarianzas arbitrarias	94
6.2. Problema inverso bayesiano en términos de campos aleatorios	95
6.3. Construcción de campos aleatorios de velocidad y densidad asociados a un mallado .	96
6.4. Representación de la información a priori	98
6.4.1. Formulación en términos de los coeficientes \mathbf{m}	98
6.4.2. Formulación en términos de los valores de los campos \mathbf{k}	99
6.5. Definición del operador de observación y la verosimilitud	100
6.6. Definición de la probabilidad a posteriori	100
6.7. Tratamiento computacional del problema bayesiano inverso	101
6.7.1. Método Crank-Nicolson preconditionado como método Monte Carlo basado en cadenas de Markov	102
6.7.2. Muestreador por combinación funcional (FES)	103
6.8. Pseudocódigo de Problema inverso aproximado basado en KL incorporando pCN y FES	104
6.9. Aproximación de campos de velocidad y densidad mediante series Karhunen-Loève .	109
6.10. Comparación de operadores de observación	112
6.11. Resultados de muestreo FES	114
7. Conclusión	117
8. Apéndice	121
8.1. Adimensionalización del problema directo	121
8.1.1. Variables adimensionales	121
8.1.2. Coeficientes adimensionalizados	121
8.1.3. Problema directo adimensional	122
8.2. Serie de Fourier - extensión a 2D	122
8.2.1. Paso 1: desarrollo en x (fijando y)	122
8.2.2. Paso 2: desarrollo de $A_n(y)$ y $B_n(y)$ en y	122
8.2.3. Paso 3: sustitución y serie final 2D	123
8.3. Coeficientes 2D	123
Bibliografía	125

Abstract

Inverse problems aim to reconstruct anomalies present in physical media from measured data. In this work, we consider inverse problems in geophysical set-ups, where the forward problem is a time dependent wave equation which models the propagation of elastic waves, set in a bounded computational domain with non-reflective boundary conditions. We generate observations from the data obtained by receivers positioned at the upper boundary of the computational domain, which define the observation operators. By adding Gaussian noise, we generate synthetic data, which allow us to create bayesian inverse problems to test our methods.

Using Bayes' theorem, we define the posterior probability of finding an anomaly in the domain. To achieve this, we construct prior and conditional (or likelihood) probabilities as multivariate Gaussian distributions. The prior distribution incorporates available knowledge while the likelihood compares the true data with data observed for proposed anomalies by solving the forward problem, weighted by the noise level. Its logarithm defines a cost functional regularized by a Tikhonov type term based on the prior information.

We consider first low-dimensional inverse problems, where the goal is to reconstruct an anomaly with an elliptical shape defined by a few geometric and material parameters. We devise a Fletcher-Levenberg-Marquardt type algorithm to determine the parameters of interest ν . This algorithm starts with prior values ν_0 and updates them iteratively to find the unknown parameters ν that maximize the posterior probability (or equivalently, minimize the regularized cost function) allowing the reconstruction of the anomaly. Once the configuration of maximum probability is determined, we quantify uncertainty by means of a Laplace approximation.

High-dimensional inverse problems, such as the reconstruction of density and velocity fields in a geophysical context, require different techniques. We introduce two approaches. The first one considers parameterizations based on truncated Fourier expansions and seeks a finite number of coefficients considered as random variables. We formulate the inverse bayesian problem and quantify uncertainty by Markov Chain Monte Carlo (MCMC) sampling, more precisely, by means of an affine invariant sampler. The second one considers the unknown fields as random fields and resorts to truncated Karhunen-Loève expansions associated to the prior covariance. We study the bayesian inverse problem by means of a functional ensemble sampler adapted to infinite dimensional problems. We test the proposed methods in the characterization of oil and gas traps near salt domes.

Resumen

Los problemas inversos tienen como objetivo reconstruir anomalías presentes en medios físicos a partir de datos medidos. En este trabajo, consideramos problemas inversos en configuraciones geofísicas, donde el problema directo es una ecuación de onda dependiente del tiempo que modela la propagación de ondas elásticas, establecida en un dominio computacional acotado con condiciones de contorno no reflectantes. Generamos observaciones a partir de los datos obtenidos por receptores ubicados en el límite superior del dominio computacional, que definen los operadores de observación. Al agregar ruido gaussiano, generamos datos sintéticos, que nos permiten crear problemas inversos bayesianos para probar nuestros métodos.

Utilizando el teorema de Bayes, definimos la probabilidad posterior de encontrar una anomalía en el dominio. Para lograr esto, construimos probabilidades a priori y probabilidades condicionales (o de verosimilitud) como distribuciones gaussianas multivariadas. La distribución a priori incorpora el conocimiento disponible mientras que la verosimilitud compara los datos verdaderos con los datos observados para las anomalías propuestas al resolver el problema directo, ponderados por el nivel de ruido. Su logaritmo define un funcional de coste regularizado con un término tipo Tikhonov basado en la información a priori.

Consideramos primero problemas inversos de baja dimensión, en los que el objetivo es reconstruir una anomalía con una forma elíptica definida por unos pocos parámetros geométricos y materiales. Ideamos un algoritmo tipo Fletcher-Levenberg-Marquardt para determinar los parámetros de interés ν . Este algoritmo comienza con valores a priori ν_0 y los actualiza iterativamente para encontrar los parámetros desconocidos ν que maximizan la probabilidad posterior (o equivalentemente, minimizan la función de costo regularizada) permitiendo la reconstrucción de la anomalía. Una vez determinada la configuración de máxima probabilidad, cuantificamos la incertidumbre mediante una aproximación de Laplace.

Los problemas inversos de alta dimensión, como la reconstrucción de campos de densidad y velocidad en un contexto geofísico, requieren técnicas diferentes. Introducimos dos enfoques. El primero considera parametrizaciones basadas en expansiones de Fourier truncadas y busca un número finito de coeficientes considerados como variables aleatorias. Formulamos el problema bayesiano inverso y cuantificamos la incertidumbre mediante muestreo de Markov Chain Monte Carlo (MCMC), más precisamente, mediante un muestreador invariante afín. El segundo considera los campos desconocidos como campos aleatorios y recurre a expansiones truncadas de Karhunen-Loève asociadas a la covarianza a priori. Estudiamos el problema bayesiano inverso mediante un muestreador de conjuntos funcional adaptado a problemas en dimensión infinita. Testamos los métodos propuestos en la caracterización de trampas de petróleo y gas cercanas a domos salinos.

Índice de figuras

2.1. Esquema ilustrativo de la configuración considerada en el problema directo de transmisión de ondas.	11
2.2. Condiciones de contorno de dominio computacional.	24
4.1. Capas de caso 1.	47
4.2. Geometría correspondiente caso 1.	47
4.3. Características de mallado uniforme.	49
4.4. Mallado uniforme utilizado para el objeto elíptico.	49
4.5. Características de mallado estratificado.	50
4.6. Mallado estratificado usado para el objeto elíptico.	50
4.7. Solución en receptores con respecto al tiempo	50
4.8. Diferencia entre mallados.	51
4.9. Mallado estratificado en GID.	51
4.10. Datos obtenidos obtenido por los 51 receptores usando mallado uniforme.....	52
4.11. Datos obtenidos obtenido por los 51 receptores usando mallado estratificado.....	52
4.12. Comparación de elipses: inicial (blanca), verdadera (roja) y parametrizada con ν_{MAP} (negra).	54
4.13. Función de costo por iteración.	55
4.14. Configuración final de la elipse.	56
4.15. Función de costo por iteración.	57
4.16. Posición centro de masa de la elipse para ambos mallados.	58
4.17. Histogramas longitud semi-ejes mayor de elipse.	58
4.18. Histogramas longitud semi-eje menor de elipse.	59
4.19. Histogramas de ángulo de rotación de la elipse.	59
4.20. Histogramas de densidad en la elipse.	59
4.21. Histogramas de velocidad de propagación de la onda en la elipse.	60
4.22. Posición centro de masa de la elipse para ambos mallados.	60
4.23. Histogramas longitud semi-ejes mayor de elipse.	61

4.24. Histogramas longitud semi-eje menor de elipse.	61
4.25. Histogramas de ángulo de rotación de la elipse.	61
4.26. Histogramas de densidad en la elipse.	62
4.27. Histogramas de velocidad de propagación de la onda en la elipse.	62
4.28. Datos obtenidos obtenido en los 51 receptores usando mallado uniforme.	65
4.29. Comparación de anomalía verdadera (contorno negro), con anomalía inicial (color blanco) y anomalía MAP (color rojo).	65
4.30. Histogramas de centro de masa y de semi-eje mayor de la anomalía.	66
4.31. Histogramas de semi-eje menor de la anomalía y de ángulo de rotación de la anomalía.	66
4.32. Velocidad y módulo de elasticidad de la anomalía.	67
5.1. Geometría de domo salino de fondo con yacimientos laterales utilizada en las simulaciones.	70
5.2. Mapa de densidad y velocidad de domo salino incorporando objetos al costado.	71
5.3. Mapa módulo de elasticidad $\chi(\mathbf{x})$	71
5.4. Mallado uniforme utilizado para geometría estratificada de fondo con domo salino.	74
5.5. Mallado estratificado incorporando la estructura de domo salino.	75
5.6. Solución en medida en los receptores en función del tiempo.	76
5.7. Diferencia entre mallados.	76
5.8. Datos medidos por los 13 receptores usando mallado uniforme.	77
5.9. Datos medidos por los 13 receptores usando mallado adaptado al domo salino.	77
5.10. Diagrama de flujo con los principales pasos para la adición de los campos en las zonas de petróleo y gas.	81
5.11. Campo de velocidad $v_{p,A}(x, y)$ aproximado con series de Fourier.	81
5.12. Campo de velocidad $v_{p,A}(x, y)$ aproximado con series de Fourier.	82
5.13. Campo de densidad $\rho_A(x, y)$ aproximado con series de Fourier.	82
5.14. Campo de densidad $\rho_A(\mathbf{x}, \mathbf{y})$ aproximado con series de Fourier.	82
5.15. Campos de fondo de densidad $\rho_F(x, y)$ y de velocidad $v_{p,F}(x, y)$, incorporando las expansiones $\rho_A(x, y)$ y $v_{p,A}(x, y)$	83
5.16. Valores observados en los receptores en función del tiempo para coeficientes que incorporan la contribución de los desarrollos de Fourier de $\rho_A(x, y)$ y $v_{p,A}(x, y)$ definidos por coeficientes ν	84
5.17. Diferencia entre los valores medidos en los receptores para los dos mallados utilizados en la simulación.	85
5.18. Datos obtenidos obtenido por los 13 receptores usando mallado uniforme e incorporando los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$	85
5.19. Datos obtenidos obtenido por los 13 receptores usando mallado adaptado e incorporando el domo salino y los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$	86
5.20. Diagrama de flujo sobre el procedimiento para la definición de la probabilidad condicional y su uso en el método inverso bayesiano.	86

5.21. Campo de densidad con probabilidad máxima obtenido en 1,500 iteraciones, incluyendo $\rho_A(x, y)$	89
6.1. Campos aleatorios de media cero generados con covarianzas Matérn.	98
6.2. Campos aproximados con series Karhunen-Loève truncadas usando $M = 100$ términos.	109
6.3. Campos aproximados con series Karhunen-Loève truncadas usando $M = 200$ términos.	110
6.4. Campos aproximados con series de Karhunen-Loève truncadas usando $M = 400$ términos.	110
6.5. Campos aproximados con series Karhunen-Loève truncadas usando $M = 600$ términos.	110
6.6. Energía acumulada para $\rho_A(x, y)$ y $v_{p,A}(x, y)$	111
6.7. Campo de densidad y velocidad incorporando desarrollos de Karhunen-Loève con 400 modos.	112
6.8. Campo de densidad y velocidad incorporando la expansión Karhunen-Loève con 400 modos.	112
6.9. Diferencia entre los valores medidos en los receptores para los dos mallados utilizados en la simulación.	113
6.10. Datos obtenidos obtenido por los 13 receptores usando mallado uniforme e incorporando los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$	114
6.11. Datos obtenidos obtenido por los 13 receptores usando mallado adaptado e incorporando los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$	114
6.12. Campos totales $\rho(\boldsymbol{\nu})$ y $v_p(\boldsymbol{\nu})$ después de 60 iteraciones.	115

Índice de tablas

4.1. Valores verdaderos de ν para la elipse.	47
4.2. Propiedades físicas adimensionales para cada una de las capas.	47
4.3. Valores de parámetro $\nu_{\text{verdadero}}$ y ν_0 para el primer ensayo.	53
4.4. Valores de parámetro $\nu_{\text{verdadero}}$ y ν_0 para el segundo ensayo.	54
4.5. Valores obtenidos para ν_{MAP} en la optimización.	55
4.6. Valores obtenidos para ν_{MAP} en la optimización.	56
4.7. Valores verdaderos de ν de la anomalía.	63
4.8. Valores verdaderos de ν del fondo.	63
4.9. Valores verdaderos de ν de la anomalía.	65
5.1. Propiedades físicas adimensionales para cada capa del fondo del domo salino.	70
5.2. Propiedades físicas adimensionales de los objetos del costado del domo salino.	71
5.3. Características del mallado uniforme.	73
5.4. Características de mallado estratificado incorporando la estructura domo salino.	75
6.1. Comparación de valores de densidad y velocidad reales en la zona de las trampas de petróleo y gas, con los valores de probabilidad máxima a posteriori.	115

Capítulo 1

Introducción y contexto

En campos como la medicina y la geofísica, la identificación de anomalías en medios heterogéneos es fundamental. En medicina, los métodos de inversión bayesiana de onda completa aplicados a la elastografía por cizallamiento permiten inferir propiedades mecánicas de los tejidos, como el módulo de elasticidad, facilitando la distinción entre tejido sano y patológico y la detección de anomalías. La aplicación de estos métodos en la detección de anomalías ya ha sido estudiada previamente; en el trabajo de [24], por ejemplo, se representan anomalías, como tumores o fibromas, mediante propiedades distintivas en términos de constantes elásticas. Este enfoque permite caracterizar las anomalías mediante pocos parámetros, lo cual reduce significativamente el costo computacional. Este tipo de enfoque ha sido utilizado también en holografía, por ejemplo, en el trabajo [26], donde se utiliza un método de inferencia bayesiana para identificar objetos y un análisis de sensibilidad topológica para la construcción de una distribución previa. En este estudio se obtiene información estadística sobre los objetos como la ubicación espacial, tamaño, orientación y propiedades físicas del material de los objetos. En geofísica, la identificación de anomalías subterráneas es esencial para la exploración y extracción de recursos naturales. En trabajos como [91] y [77] se presentan imágenes obtenidas computacionalmente mediante inversión de onda completa. Mientras que [91] resuelve problemas en alta dimensionalidad utilizando discretizaciones espaciales uniformes y métodos espectrales, en [77] se aplican metodologías inversas bayesianas en problemas geofísicos, pero principalmente en espacios de baja dimensionalidad. El enfoque en estos trabajos resulta ser una estrategia eficaz en problemas donde la precisión espacial y la resolución en regiones complejas son aspectos críticos.

Para la detección de anomalías se utilizan métodos inversos, que se suelen clasificar en deterministas y probabilísticos, como los métodos bayesianos [93]. Estos métodos permiten detectar anomalías y cuantificar sus propiedades a partir de operadores de observación asociados a modelos del proceso y de datos medidos por un dispositivo de imagen o datos sintéticos generados computacionalmente a partir del modelo [16, 24, 26, 98]. Nos centramos en procesos de imagen modelados matemáticamente como procesos de scattering inverso: se lanza un tipo de onda hacia el medio en estudio, que interactúa con él; el resultado de la interacción se mide en una red de receptores dando lugar a los datos. La dinámica de la onda suele venir descrita por un problema de ecuaciones en derivadas parciales, adecuado al tipo de onda. Para ondas armónicas en tiempo, se trata de problemas de contorno elípticos. Para ondas dependientes del tiempo, suelen obtenerse problemas de valores iniciales y de contorno que involucran ecuaciones Maxwell para ondas electromagnéticas, Navier para elasticidad, ondas para acústica, u otros modelos para ondas de otra naturaleza. Estos modelos definen el llamado problema directo. Si se conoce el medio objeto de estudio y las ondas emitidas, la solución del problema directa evaluada en la red de receptores coincidiría con los datos medidos, salvo por el ruido de medida. Para generar el operador de observación y los datos sintéticos, primero es necesario resolver el problema directo. El operador de observación se construye a partir de las mediciones obtenidas por receptores durante cada instante de tiempo. Dado que las mediciones pueden estar sujetas a errores de medición, se añade un componente de ruido Gaussiano (o de otra

naturaleza) al operador de observación; a este conjunto de datos se les llama datos sintéticos y son fundamentales para testar métodos de resolución de problemas de inversión [24].

Para resolver numéricamente un problema directo que involucra coeficientes discontinuos y funciones dependientes del tiempo se emplean aproximaciones de la ecuación en derivadas parciales gobernante del problema, como lo es el método de Galerkin [9]. Este método se basa en la aproximación de la solución reescribiendo primero la EDP en su formulación débil, lo que facilita la integración de las condiciones de frontera y las variaciones espaciales. A continuación, la solución aproximada se representa a través de una combinación lineal de funciones base con coeficientes dependientes del tiempo, los cuales se deben determinar. Este enfoque proporciona una solución aproximada, que debe respetar las condiciones iniciales y de contorno impuestas. El método de Galerkin es la base del método de elementos finitos [57]. El método de elementos finitos es una técnica poderosa para resolver EDPs en dominios computacionales complejos, y se basa en la división del dominio en elementos más pequeños y en la aplicación de funciones de forma local, permitiendo una aproximación precisa de las soluciones de las EDPs [99, 100].

A través del método de elementos finitos, es posible realizar la discretización temporal y espacial de la forma variacional de la EDP, convirtiendo el problema continuo en un sistema de ecuaciones algebraicas que se pueden resolver numéricamente. Esta técnica es crucial para obtener soluciones precisas y estables en problemas de propagación de ondas en medios heterogéneos [100]. El sistema de ecuaciones obtenidas es posible abordarlas a través de esquemas explícitos o implícitos. Un esquema explícito expresa las soluciones futuras de las variables como una función explícita de los actuales, lo que resulta en un avance más rápido, sin embargo, requiere que ese incremento sea suficientemente pequeño para que la solución sea estable. En este caso, el paso de tiempo deberá cumplir con la condición de estabilidad de Courant-Friedrichs-Lewy (CFL) para garantizar la estabilidad numérica del problema directo [35, 100]. En cambio en un esquema implícito, en cada paso se debe resolver un sistema acoplado de ecuaciones que involucra los valores actuales y futuros de la solución, lo cual es costoso de resolver ya que, al tratarse de muchas ecuaciones acopladas, generalmente hay que invertir matrices o iterar para converger a la solución, sin embargo, proporciona mayor estabilidad [100].

En este trabajo consideraremos dos formulaciones matemáticas para los problemas inversos descritos: la formulación determinista en términos de optimización con restricciones de tipo ecuación en derivadas parciales y la formulación bayesiana.

Consideramos reformulaciones del problema inverso determinista en términos de la minimización de una función de costo $J(\boldsymbol{\nu})$, donde $\boldsymbol{\nu}$ es el conjunto de parámetros desconocidos que caracterizan a la anomalía [28]. La función de costo cuantifica la discrepancia entre el operador de observación y los datos para unos parámetros dados. Este método busca encontrar un conjunto único de parámetros que mejor reproduzca las observaciones. Puede ser sensible a la elección de parámetros iniciales $\boldsymbol{\nu}_0$ cuando el funcional de coste no es convexo. Si bien se quiere encontrar un único mínimo de la función de costo $J(\boldsymbol{\nu})$, es posible que pueda presentar mínimos locales adicionales, por lo que se introduce un término de regularización, que por lo general, son regularizaciones de Tikhonov [90] o de tipo variación total para evitar la aparición de mínimos no físicos. La minimización de la función de costo $J(\boldsymbol{\nu})$ implica buscar los parámetros $\boldsymbol{\nu}$ donde el gradiente de $J(\boldsymbol{\nu})$, es decir, $\nabla_{\boldsymbol{\nu}} J(\boldsymbol{\nu})$, sea igual a cero. Esto conlleva la dificultad de que los métodos de optimización de tipo gradiente se estancan. Considerando la curvatura de la función de costo a través de la matriz Hessiana, es posible obtener información importante sobre la forma del valle donde se encuentra el mínimo de la función de costo, acelerando la convergencia y garantizando que se ha alcanzado su valor mínimo. El valor mínimo de $\boldsymbol{\nu}$ es posible aproximarlos a través de métodos numéricos iterativos como Fletcher-Levenberg-Marquardt [24], que hibridizan iteraciones de tipo gradiente y tipo Newton.

El método inverso bayesiano es un enfoque útil en aplicaciones donde el conocimiento previo es importante y la incertidumbre de los parámetro desconocido, o magnitudes derivadas, requiere ser cuantificado. Por ejemplo, en la identificación de anomalías en medios geológicos, donde la

variabilidad de las propiedades del subsuelo puede ser significativa, el método inverso bayesiano ofrece ventajas sobre los enfoques deterministas tradicionales al proporcionar una representación más completa de la incertidumbre y la variabilidad de los parámetros que caracterizan a una anomalía.

El método inverso bayesiano introduce un enfoque probabilístico en el cual se incorporan las incertidumbres tanto en las observaciones como en el modelo. A diferencia del enfoque determinista en el que se busca un único conjunto de valores para los parámetros $\boldsymbol{\nu}$, la inversión bayesiana estima la distribución posterior de estos parámetros, lo que proporciona una descripción más completa sobre la incertidumbre de los parámetros [17, 46]. Para estimar la distribución posterior de los parámetros de interés, se utiliza el teorema de Bayes, el cual relaciona la densidad de probabilidad a priori de los parámetros $\boldsymbol{\nu}$ con la densidad de probabilidad de los datos observados, lo que resulta en una distribución a posteriori que proporciona una estimación probabilística a posteriori de los parámetros de interés $\boldsymbol{\nu}$ [17, 46, 93]. En el método inverso bayesiano, la probabilidad posterior se modela como el producto de la densidad de probabilidad a priori, y la densidad de probabilidad condicionada a las mediciones [46]. A menudo se emplean distribuciones de tipo Gaussiano, lo que conlleva la obtención de una expresión para la probabilidad a posteriori en forma de la exponencial $e^{-J_{\text{reg}}(\boldsymbol{\nu})}$, donde el argumento de la función exponencial corresponde a la suma del coste determinista sin regularizar y un término que engloba la información a priori con un efecto regularizador $J_{\text{reg}}(\boldsymbol{\nu}) = J(\boldsymbol{\nu}) + R(\boldsymbol{\nu})$ [24, 63]. Uno de los estimadores estadísticos en el método inverso bayesiano es el máximo a posteriori o MAP. La estimación MAP se define como el conjunto de parámetros $\boldsymbol{\nu}$ que maximiza la probabilidad a posteriori. Conceptualmente, esto es equivalente a minimizar una función de costo regularizada $J_{\text{reg}}(\boldsymbol{\nu})$ [17, 24]. Este es un problema de optimización determinista que se puede abordar como hemos mencionado mediante un algoritmo de Fletcher-Levenberg-Marquardt [24, 25], que actualiza iterativamente el parámetro inicial $\boldsymbol{\nu}_0$ considerando tanto la dirección del gradiente como la curvatura local, con la matriz hessiana de la función de costo, lo que acelera la convergencia hacia el mínimo global [1, 24, 25].

A diferencia del método de inversión determinista, el método de inversión bayesiano permite en principio cuantificar la incertidumbre asociada a cada parámetro $\boldsymbol{\nu}_{\text{MAP}}$. La principal dificultad en la inversión bayesiana radica en que la distribución a posteriori generalmente no se puede expresar de forma sencilla debido a la multidimensionalidad de los parámetros desconocidos, por lo que, para aproximarla se requieren métodos numéricos avanzados. Técnicas como el Método Monte Carlo basado en Cadenas de Markov (MCMC), son ampliamente utilizadas para esta tarea [1, 24–26, 77], ya que permiten explorar el espacio de parámetros de manera eficiente y obtener estimaciones precisas de las propiedades del medio. En ocasiones, se recurre a aproximar la probabilidad posterior por una Gaussiana multivariada con media $\boldsymbol{\nu}_{\text{MAP}}$ y covarianza como la inversa de la matriz Hessiana de la función de costo [1, 24, 25, 85, 89, 94], lo que se denomina aproximación de Laplace. Esta aproximación es menos costosa pero puede ser pobre en el caso de distribuciones multivariadas o fuertemente asimétricas.

El método MCMC se utiliza para explorar el espacio de parámetros y obtener estimaciones estadísticas de las características del medio y de las anomalías presente. En trabajos como [1, 24–26, 92, 97] se ha aplicado MCMC con implementaciones de conjunto para la reconstrucción de anomalías y de campos de interés. La técnica MCMC se basa en la construcción de una cadena de Markov cuya distribución estacionaria corresponde a la distribución a posteriori deseada [13, 82]. Para generar esta cadena de Markov, podemos emplear algoritmos como Metropolis-Hastings y muestreadores por combinación. [34]. Entre los muestreadores por combinación se encuentra el muestreador por combinación de invariancia afín (AIES) y el muestreador por combinación funcional (FES); ambos son técnicas avanzadas que mejoran la eficiencia y la convergencia del muestreo [34]. AIES es particularmente útil para problemas de baja dimensión pobremente escalados y con múltiples modos, ya que es invariante bajo transformaciones afines y puede adaptarse a la estructura geométrica del espacio de parámetros, lo que acelera la convergencia [50]. FES, por otro lado, combina elementos de AIES, Metropolis-Hastings y el método Crank-Nicolson preconditionado para proporcionar un

muestreo eficiente y robusto, en altas dimensiones [34]. El uso de MCMC en problemas inversos es importante debido a su capacidad para manejar problemas mal condicionados y para proporcionar una caracterización completa de la incertidumbre en las estimaciones de los parámetros desconocidos. Sin embargo, su implementación computacional puede ser costosa, especialmente en problemas de alta dimensión, lo que requiere el uso de técnicas de aceleración y paralelización para mejorar la eficiencia computacional [82].

En aplicaciones complejas, los problemas inversos pueden involucrar un número elevado de parámetros, lo que aumenta la dificultad del análisis. Esta alta dimensionalidad requiere el uso de métodos avanzados para la estimación y muestro de los parámetros ν . En estos casos, la combinación de técnicas avanzadas como la expansión de Karhunen-Loève y el muestreo de MCMC se vuelve crucial para obtener soluciones precisas y eficientes en problemas de alta dimensión [20, 21]. La expansión de Karhunen-Loève permite la reducción dimensional y la modelización eficiente de campos aleatorios, mientras que MCMC facilita el muestreo de distribuciones a posteriori en espacios de alta dimensión, proporcionando soluciones precisas y robustas para problemas inversos complejos. La expansión Karhunen-Loève es una técnica de descomposición que representa un campo aleatorio mediante una serie infinita de funciones ortogonales asociadas a su covarianza, cada una multiplicada por variables aleatorias [62, 67]. Si bien las expansiones de Karhunen-Loève infinitas proporcionan una descripción completa del campo aleatorio, en la práctica se utilizan versiones truncadas para hacer el problema computacionalmente manejable. Estas versiones truncadas retienen un número finito de términos M , lo que permite aproximar el campo aleatorio con un grado de precisión que depende del número de términos incluidos en la expansión [34, 92].

En este trabajo se estudian las formulaciones del problema directo e inverso para dos casos específicos, que difieren en la geometría del fondo, la anomalía a considerar y los parámetros desconocidos a obtener mediante el método inverso bayesiano. Un caso se aborda en baja dimensión y el otro en alta dimensión. Estos dos problemas son los siguientes:

1. El primer caso de estudio se centra en un medio de fondo estratificado, donde cada capa tendrá propiedades diferentes de densidad y velocidad de propagación de la onda. En el centro del dominio de este medio se incorporará una anomalía de forma elíptica. Esta anomalía también tendrá propiedades específicas de densidad y velocidad de propagación de onda. Primero, se resolverá el problema directo para este caso, obteniendo el operador de observación. Posteriormente, a los datos obtenidos se les añadirá un porcentaje de ruido Gaussiano para generar los datos sintéticos. Dado que la resolución numérica del problema directo se basa en la aplicación del método de elementos finitos de la ecuación de ondas, los resultados obtenidos se validarán realizando dos tipos de discretizaciones espaciales para contrastar la convergencia de los mismos. Con los datos sintéticos obtenidos, se aplicarán métodos de inversión para determinar los parámetros desconocidos ν que caracterizan la forma y naturaleza de la anomalía, los cuales proporcionarán información sobre la anomalía, incluyendo su geometría, posición y propiedades físicas como la densidad y velocidad. En este caso de estudio, se aborda un problema de inversión en el que el parámetro ν es de baja dimensionalidad. Para resolverlo, se estimarán los parámetros que maximizan la probabilidad a posteriori ν_{MAP} optimizando el coste regularizado (dado por el logaritmo de la probabilidad a posteriori cambiado de signo) y se analizará su incertidumbre mediante la aproximación de Laplace.

Este enfoque de inversión basada en objetos se enmarca dentro de una línea de investigación más amplia, véase por ejemplo [24]. En dicho estudio, se aplica un enfoque de inversión bayesiana para determinar las constantes elásticas de anomalías, como tumores o fibromas, inmersas en tejido sano. Este método permite una diferenciación precisa entre tejido sano y no sano mediante la incorporación de información a priori y la cuantificación explícita de la incertidumbre en las mediciones. En nuestro contexto nos centramos más en aplicaciones geofísicas, véase la publicación preliminar [25], con el objeto de determinar la presencia de yacimientos de

determinados materiales en el subsuelo, que suele presentar estructura estratificada en lugar de homogénea.

2. El segundo caso de estudio se centra en una estructura geológica compuesta por un domo salino en un medio estratificado. Estas formaciones tienen gran interés pues suelen venir acompañadas de yacimientos de petróleo y gas adyacentes, que es difícil ubicar por el fuerte apantallamiento de las capas salinas sobre las ondas de sondeo. En nuestra estructura geológica, incorporamos anomalías a los costados del domo salino, las cuales representarán trampas de gas y de petróleo. Cada capa, el domo salino y los yacimientos tendrán propiedades geológicas específicas de densidad y velocidad de propagación de onda. El primer paso será la resolución numérica de la ecuación de ondas utilizando dos tipos de discretizaciones espaciales para evaluar la convergencia de los resultados. En la resolución del problema directo, obtendremos el operador de observación y los datos sintéticos necesarios para la formulación del método inverso bayesiano. A diferencia del primer caso de estudio, ahora los parámetros desconocidos no serán objetos, sino funciones: coeficientes que varían en espacio y que nos permitirán caracterizar los campos de densidad y velocidad de propagación de onda. Por lo tanto, en este segundo caso de estudio, el problema se formulará como un modelo bayesiano en alta dimensión, y el parámetro ν , que caracteriza los campos, tendrá información sobre los coeficientes de series de Fourier o de series Karhunen-Loève de dimensión infinita.

Utilizando series de Fourier, el parámetro ν representará los coeficientes de estas una serie truncada capaz de representar los campos de densidad $\rho_A(x, y)$ y velocidad $v_{p,A}(x, y)$ en la zona de las trampas, los cuales dependerán de dichos coeficientes. Se realizará un estudio de sensibilidad para ver cuántos modos son necesarios para aproximar los campos sin que se aprecie el fenómeno de Gibbs en el mallado considerado. Estas aproximaciones de los campos se introducirán en la formulación numérica del problema directo gobernada por la ecuación de ondas, para definir un nuevo operador de observación y obtener datos sintéticos. Incorporando los datos sintéticos y el operador de observación en la formulación del método inverso bayesiano, se reconstruirán los campos de las anomalías que representan las trampas de gas y petróleo. Como método de muestreo de los campos se utilizará el método MCMC de tipo AIES. Observamos que, a pesar del apantallamiento de la señal causado por el domo salino, se aprecia la presencia de yacimientos a sus costados con una buena definición de la región que ocupan. Las técnicas de Fourier requieren trabajar con un elevado número de parámetros pobremente escalados, y no se puede excluir aberraciones aparejadas al fenómeno de Gibbs al ajustar series sin información previa.

Hemos explorado un segundo tipo de desarrollos en serie de estos campos, específicamente adaptados al trabajo con campos aleatorios, basados en la descomposición espectral asociada a sus covarianzas, que proporciona desarrollos en serie Karhunen-Loève para los campos aleatorios de densidad y de velocidad. Para muestrear la distribución de coeficientes truncada ν , se empleará el muestreador MCMC por combinación funcional FES. Este muestreador alternará entre el uso de AIES para muestrear las coordenadas principales de la expansión de Karhunen-Loève y el método Crank-Nicolson preconditionado (pCN) para muestrear el subespacio complementario en alta dimensión. Desde el punto de vista teórico, este enfoque presenta ventajas teóricas, con estabilidad garantizada a medida que la dimensión del espacio de coeficientes aumenta. Sin embargo, se tiene la dificultad de identificar operadores de covarianza adecuados al caso en estudio para definir las funciones de base de la expansión.

Describimos a continuación de forma más precisa la estructura de la tesis.

1.1. Estructura de la tesis

La memoria se estructura en los siguientes capítulos:

1. Estudio teórico y numérico del problema directo.
2. Estudio teórico y numérico del problema inverso en dimensión finita.
3. Estrategias de inversión del campo de ondas basadas en objetos.
4. Formulaciones bayesianas inversas basadas en series de Fourier.
5. Formulaciones bayesianas inversas basadas en campos aleatorios.
6. Discusión de resultados y conclusión.

En el capítulo titulado **Estudio teórico y numérico del problema directo**, se estudiará la formulación y resolución del problema de valores iniciales y de contorno para la ecuación de ondas en un medio heterogéneo en un dominio computacional truncado. Se formulará el modelo matemático que representa el problema directo a resolver, incluyendo la geometría de fuentes emisoras y dispositivos receptores. A continuación, estudiaremos el operador de observación, que contiene la información sobre las ondas captadas por los receptores, y lo relacionaremos con los datos sintéticos, que incluyen un término de ruido modelado como una distribución Gaussiana aleatoria. Este problema planteado en principio en un medio semi-infinito, se truncará a una región finita [24, 25]. En este contexto, se usará el método de Galerkin para obtener la formulación débil de la ecuación de ondas y se analizará la existencia, unicidad y regularidad de la solución $u(\mathbf{x}, t)$ [25, 26]. Tras ello, estudiaremos la resolución numérica del problema directo en un dominio finito. Nos enfocaremos en las discretizaciones espaciales y temporales de la formulación débil, e introduciremos la condición de estabilidad de Courant-Friedrichs-Lewy (CFL) para seleccionar un paso temporal adecuado que garantice la estabilidad de la solución numérica. Finalmente, se introducirán las condiciones de contorno no reflectantes, que se aplicarán en los contornos laterales e inferiores en la geometría rectangular, para asegurar que las reflexiones de la onda en el borde del dominio no afecten a los resultados, y una condición de Neumann cero en el contorno donde estarán ubicados los emisores y receptores [24, 25].

El capítulo **Estudio teórico y numérico del problema inverso en dimensión finita**, se inicia con una introducción al problema inverso desde una perspectiva determinista en un espacio de dimensión finita. A continuación, se abordará el método de inversión bayesiana en dimensión finita, formulando una expresión para la probabilidad a posteriori de encontrar una anomalía en el dominio computacional utilizado. Como veremos en este capítulo, la expresión para la probabilidad a posteriori dependerá de la probabilidad condicional de observar unos datos dados unos parámetros, que está directamente influenciada por el operador de observación y los datos sintéticos, y de la probabilidad a priori, que depende de la información disponible sobre los parámetros de interés. Ambas probabilidades, condicional y a priori, se definirán a partir de distribuciones Gaussianas multivariadas. Al aplicar el teorema de Bayes para obtener la probabilidad a posteriori, se tiene una distribución que varía como la exponencial de $-J_{\text{reg}}(\boldsymbol{\nu})$.

Presentamos las dos opciones empleadas para caracterizar esta probabilidad. La primera se basa en el cálculo de los parámetros de probabilidad máxima mediante optimización, algoritmos iterativos de gradiente, Newton o híbridos Fletcher-Levenberg-Marquardt. A continuación se aproxima la probabilidad a posteriori linearizando en torno a ese máximo mediante la aproximación de Laplace, dando lugar a una distribución multivariada Gaussiana. La segunda recurre a métodos MCMC para explorar la estructura de la probabilidad a posteriori mediante muestreo. Discutiremos distintas técnicas MCMC y su buen uso.

En el capítulo **Estrategias de inversión del campo de ondas basadas en objetos**, exploraremos las estrategias de inversión bayesiana para la detección de un objeto elíptico incorporado en un dominio de fondo estratificado. Los parámetros que caracterizan el objeto se reducen a los que describen su forma y la naturaleza del material. En este capítulo se detallará la geometría, las propiedades y las discretizaciones espaciales del primer caso de estudio, así como el paso de tiempo adecuado para la discretización temporal, con el fin de evitar inestabilidades numéricas. Se describirán las propiedades físicas de densidad y de velocidad de propagación de onda que tendrá el fondo estratificado, así como la anomalía elíptica, la cual se ubicará en el centro del dominio computacional.

Usando la geometría y propiedades descritas, primero, se abordará la resolución del problema directo. Se describirán en detalle las dos discretizaciones espaciales empleadas para el estudio de la convergencia en mallas: una discretización espacial uniforme, que utiliza un tamaño y distribución uniforme de los elementos, y una discretización adaptada a la presencia de estratos. En este capítulo, las variables del parámetro de interés $\boldsymbol{\nu}$ serán: la posición del centro de masa de la elipse (c_x, c_y) , el semi-eje mayor a , el semi-eje menor b , el ángulo de rotación de la elipse ω , la densidad ρ , y la velocidad de propagación v_p . Estas variables se agruparán en un vector $\boldsymbol{\nu} = (c_x, c_y, a, b, \omega, \rho, v_p)$, el cual se considerará tanto en el problema directo como en el inverso.

Se emplearán ambas discretizaciones para resolver numéricamente el problema directo gobernado por la ecuación de ondas, cuyo resultado será el desplazamiento de la onda en diferentes instantes de tiempo. Este desplazamiento será registrado por una red de receptores posicionados en $y = 0$. A este conjunto de datos se le llamará el operador de observación. Se obtendrán dos operadores de observación: uno correspondiente a la discretización uniforme $\mathbf{f}_{\text{uniforme}}(\boldsymbol{\nu})$ y otro a la discretización estratificada $\mathbf{f}_{\text{estratificada}}(\boldsymbol{\nu})$. Estos operadores serán comparados y se calculará la diferencia porcentual entre ellos.

Al contar con dos conjuntos de operadores de observación, se generarán dos conjuntos de datos sintéticos: $\mathbf{d}_{\text{uniforme}}^{\text{obs}}$ y $\mathbf{d}_{\text{estratificada}}^{\text{obs}}$, que se utilizarán en el problema inverso bayesiano aproximado, cuyo objetivo será reconstruir la anomalía elíptica. Para aplicar el método inverso bayesiano, se definirá un vector con información a priori $\boldsymbol{\nu}_0$, donde se usarán dos configuraciones iniciales diferentes; la diferencia será el ángulo de rotación inicial de la elipse. Posteriormente, se obtendrán los valores de $\boldsymbol{\nu}_{\text{MAP}}$, que maximizan la probabilidad posterior de hallar una anomalía en el dominio computacional. Además, se analizará cómo varía la función de costo $J_{\text{reg}}(\boldsymbol{\nu})$ en cada iteración del método inverso bayesiano hasta alcanzar su valor mínimo y estabilizarse. Finalmente, se cuantificará la incertidumbre de cada una de las variables de $\boldsymbol{\nu}_{\text{MAP}}$ utilizando la aproximación de Laplace para la probabilidad posterior. Como caso de aplicación en elastografía médica, previamente estudiado en el trabajo publicado en [24], se utilizarán los parámetros definidos en dicho estudio. En este contexto, mediante el uso del método inverso bayesiano, se buscará estimar las propiedades de la velocidad de propagación de onda y el módulo de elasticidad en tejido no sano, así como reconstruir la anomalía presente. Posteriormente, se cuantificará la incertidumbre de cada variable $\boldsymbol{\nu}_{\text{MAP}}$ utilizando la aproximación de Laplace.

En el capítulo de **Formulaciones bayesianas inversas basadas en series de Fourier** trabajaremos con una geometría estratificada con una estructura de domo salino, que incluye trampas de petróleo/gas al costado del domo salino. Estudiaremos bajo qué modos podemos obtener una buena aproximación en el mallado elegido sin que se aprecie el fenómeno de Gibbs. Los campos de densidad y velocidad parametrizados con series de Fourier los denominaremos $\rho_A(x, y)$ y $v_{p,A}(x, y)$, respectivamente; por lo que el campo total de densidad y velocidad los expresaremos como $\rho(x, y) = \rho_F(x, y) + \rho_A(x, y)$ y $v_p(x, y) = v_{p,F}(x, y) + v_{p,A}(x, y)$, donde $\rho_F(x, y)$ y $v_{p,F}(x, y)$ corresponderán a los campos del fondo estratificado que incluye el domo salino. Posteriormente incorporaremos las aproximaciones de los campos $\rho(x, y)$ y $v_p(x, y)$ en el problema directo, el cual está gobernado por la ecuación de ondas. Para resolver numéricamente el problema directo, utilizaremos dos discretizaciones espaciales: una discretización uniforme y otra adaptada a la estructura del

domo salino. Como resultado obtendremos operadores de observaciones para cada discretización, los cuales se compararán y se calculará la diferencia porcentual entre los conjuntos de datos obtenidos.

Del problema directo obtendremos los operadores de observación y los datos sintéticos para las dos discretizaciones espaciales utilizadas. En el marco del problema inverso bayesiano, utilizaremos el operador de observación correspondiente a la discretización uniforme y los datos sintéticos de la discretización adaptada. En el contexto del problema inverso bayesiano, lo que buscaremos será reconstruir los campos de densidad y de velocidad en las zonas de las trampas de petróleo y gas, por lo que el parámetro desconocido $\boldsymbol{\nu}$ deberá contener esta información. Como los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$ los hemos parametrizado con series de Fourier truncadas, el problema inverso bayesiano se enfocará en determinar los coeficientes de dichas series que permitan la reconstrucción precisa de los campos en las zonas de interés. La dimensión del parámetro desconocido $\boldsymbol{\nu}$ dependerá del número de modos escogidos para aproximar $\rho_A(x, y)$ y $v_{p,A}(x, y)$ sin que aparezca el fenómeno de Gibbs. Para este propósito, se empleará el método de Monte Carlo por cadenas de Markov, utilizando el esquema AIES para realizar el muestreo de los valores de $\boldsymbol{\nu}$.

En el capítulo **Formulaciones bayesianas inversas basadas en campos aleatorios**, utilizaremos el método inverso bayesiano con el fin de reconstruir los campos de densidad y de velocidad de propagación de onda. La geometría de fondo que consideraremos, al igual que en el capítulo anterior, será la de un fondo estratificado con un domo salino y dos anomalías posicionadas al costado inferiores del domo salino, que representan trampas de petróleo y gas formadas en el subsuelo. Utilizaremos los operadores de observación y los datos sintéticos que obtuvimos en la formulación del problema directo del capítulo anterior.

Comenzaremos con un breve recordatorio sobre las variables y los campos aleatorios, centrándonos en el campo Whittle-Matérn con núcleo Matérn. Mediante su descomposición espectral, obtendremos valores y funciones propias. Usando el teorema de Karhunen-Loève, construiremos campos aleatorios $\mathbf{k}(\mathbf{x}, \alpha)$, que simplificaremos truncándolos en M términos. Esta truncación será clave para generar campos aleatorios que describan la densidad y velocidad de propagación en el fondo estratificado con el domo salino. Incorporaremos esta expansión truncada en la formulación del problema directo, obteniendo un nuevo operador de observación $f(\boldsymbol{\nu})$ y datos sintéticos derivados de este operador; donde en este caso $\boldsymbol{\nu}$ son los coeficientes que permiten definir los campos de densidad y de velocidad. Se comenzará construyendo los campos a priori a partir del campo determinista inicial $\boldsymbol{\nu}_0$, es decir, campos de densidades y de velocidad de la onda del fondo incluyendo el domo salino (sin la incorporación de los campos correspondientes a las trampas de petróleo y gas); más la adición de una componente estocástica, que se modelará con la covarianza Matérn. Con el operador de observación, los datos sintéticos del problema directo y la probabilidad a priori, derivaremos una expresión para la probabilidad posterior y para la función de costo regularizada $J_{\text{reg}}(\boldsymbol{\nu})$. Para reconstruir los campos de densidad y de velocidad, utilizaremos el método de Monte Carlo junto con el muestreador FES. Este método aceptará o rechazará los valores propuestos según la condición de aceptación de FES, alternando entre el muestreador AIES, para las coordenadas principales de la expansión de Karhunen-Loève, y pCN, para el subespacio complementario.

Para finalizar, discutiremos los resultados obtenidos en los capítulos anteriores y se presentarán las conclusiones.

1.2. Objetivos de la tesis

Los principales objetivos de la tesis son los siguientes:

1. **Abordar el problema directo:** desarrollar e implementar algoritmos eficientes de bajo coste para la resolución numérica de problemas de propagación de ondas elásticas en medios con

-
- anomalías y en medios estratificados. Utilizar el método de elementos finitos con condiciones no reflectantes en los bordes artificiales del recinto computacional para obtener soluciones precisas.
2. **Abordar el problema inverso:** optimizar funcionales de coste bajo restricciones impuestas por ecuaciones de ondas, utilizando métodos de inferencia inversa bayesiana y métodos deterministas. Incorporar la cuantificación de la incertidumbre en los modelos para mejorar la precisión de las estimaciones.
 3. **Abordar el problema inverso mediante métodos de cadenas de Markov Monte Carlo:** implementar métodos MCMC para resolver el problema inverso, trabajando con campos iniciales aleatorios que utilizan la covarianza Matérn para modelar las incertidumbres en los parámetros del modelo.
 4. **Reducir el coste computacional:** desarrollar y aplicar técnicas rápidas de optimización, como el método de Fletcher-Levenberg-Marquardt, para resolver problemas inversos regularizados de manera eficiente y reducir el coste computacional asociado.
 5. **Aplicar a contextos geofísicos y biomédicos:** extender las metodologías desarrolladas a aplicaciones prácticas en contextos geofísicos y biomédicos, utilizando geometrías bidimensionales simplificadas para facilitar la aplicación y validación de los modelos en escenarios reales.

Capítulo 2

Estudio teórico y numérico del problema directo

El modelo matemático directo describe la propagación de ondas creadas en los puntos fuente o emisores y su interacción con el medio cuya estructura interna deseamos caracterizar. Asimismo, proporciona una caracterización matemática del operador de observación, y conecta las ondas emitidas con la estructura del medio y los datos que se miden en el proceso de imagen.

En la Figura 2.1 se presenta un esquema ilustrativo de una configuración básica, donde los límites de cada capa están representados por líneas horizontales ubicadas en las posiciones $y = 0, -0.4, -1.1, -1.7, -2.5$ e $y = -3$. Las ondas se emiten desde uno o varios emisores (círculos amarillos en la figura 2.1), propagándose a través del medio e interactuando con los estratos y con la inclusión (elipse de color rojo en la figura 2.1). La onda reflejada se mide mediante una red de receptores (cuadrados azules en la figura 2.1).

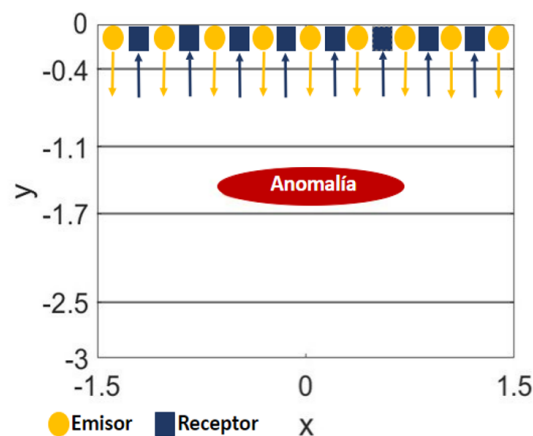


Figura 2.1: Esquema ilustrativo de la configuración considerada en el problema directo de transmisión de ondas.

El problema directo se enfoca a la resolución de la ecuación que gobierna la propagación de estas ondas dependientes del tiempo emitidas desde diferentes fuentes. Estas ondas interactúan con la estructura del medio y la anomalía presente, generando una respuesta en la transmisión de la onda, la cual es posteriormente captada por una red de receptores. La naturaleza de esta respuesta captada está intrínsecamente ligada a las características y propiedades físicas del medio, así como a la presencia de inclusiones de otros materiales en el dominio.

En el contexto del problema directo, la excitación externa así como la estructura, las propiedades elásticas del medio y de las inclusiones, son variables conocidas. El objetivo consiste en formular las ecuaciones que gobiernan el proceso y aplicar métodos numéricos que permitan aproximar computacionalmente el desplazamiento de la onda. Consideraremos situaciones en las que se

desprecia la amortiguación en el medio, lo que implica que los parámetros elásticos, representados por $\chi(\mathbf{x})$, serán reales.

2.1. Formulación del problema directo

Comenzamos el estudio del problema de transmisión de ondas formulando el modelo matemático. Para ello, consideramos la ecuación constitutiva de un material isotrópico con módulo de elasticidad lineal. Esta ecuación relaciona el tensor de tensiones $\sigma_{i,j}$ con el tensor de deformaciones $\epsilon_{i,j}$, y se expresa mediante la ecuación (2.1) [61]

$$\sigma_{i,j} = 2\mu\epsilon_{i,j} + \lambda\delta_{i,j}\epsilon_{k,k}, \quad i, j = 1, 2, 3, \quad (2.1)$$

donde:

- σ_{ij} : componentes del tensor de tensiones, que representan la fuerza por unidad de área en la dirección i sobre una superficie perpendicular a la dirección j .
- ϵ_{ij} : componentes del tensor de deformaciones, definidas como

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right),$$

donde $\mathbf{u} = (u_1, u_2, u_3)$ es el vector de desplazamientos.

- ϵ_{kk} : dilatación volumétrica, correspondiente a la traza del tensor de deformaciones, e implica una suma sobre el índice repetido k , es decir, $\epsilon_{kk} = \sum_{k=1}^3 \epsilon_{kk}$.
- δ_{ij} : delta de Kronecker, definido como

$$\delta_{ij} = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{cases}$$

- μ : módulo de rigidez transversal, calculado como

$$\mu = \frac{E}{2(1 + \nu_{\text{Poisson}})},$$

donde E es el módulo de Young y ν_{Poisson} el coeficiente de Poisson.

- λ : primer parámetro de Lamé, dado por

$$\lambda = \frac{E\nu_{\text{Poisson}}}{(1 + \nu_{\text{Poisson}})(1 - 2\nu_{\text{Poisson}})}.$$

- i, j, k : índices que recorren las dimensiones espaciales 1, 2, 3, correspondientes a las componentes del vector de posición $\mathbf{x} = (x_1, x_2, x_3)$.

En un sólido continuo, isótropo, homogéneo y elástico, la ecuación que rige el movimiento de la onda viene dada por la ecuación de conservación de momento lineal para el desplazamiento $\mathbf{u}(x_1, x_2, x_3, t)$ en cada punto $\mathbf{x} = (x_1, x_2, x_3)$

$$\rho \frac{\partial^2 u_i}{\partial t^2} + \frac{\partial \sigma_{i1}}{\partial x_1} + \frac{\partial \sigma_{i2}}{\partial x_2} + \frac{\partial \sigma_{i3}}{\partial x_3} = \rho f_i, \quad i = 1, 2, 3, \quad (2.2)$$

donde $\mathbf{f} = (f_1, f_2, f_3)$ representa a las fuerzas por unidad de volumen y ρ es la densidad másica del material.

Si se introduce la expresión (2.1) en la ecuación (2.2) y asumimos que λ y μ son constantes, se obtiene la ecuación de Navier

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = \operatorname{div}(\boldsymbol{\sigma}) = \mu \Delta \mathbf{u} + (\lambda + \mu) \nabla(\operatorname{div} \mathbf{u}), \quad (2.3)$$

la cual incorpora los parámetros de Lamé de la ecuación (2.1), véase [2].

Si el desplazamiento se produce sin cambios en el volumen, es decir, si es incompresible, entonces se cumple

$$\nabla \cdot \mathbf{u} = \operatorname{div}(\mathbf{u}) = 0, \quad (2.4)$$

donde el campo de onda \mathbf{u} representa ondas de cizalla (shear waves o S-waves). Insertando la condición (2.4) en la ecuación de Navier (2.3) se obtiene la ecuación de ondas (2.5)

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = \mu \Delta \mathbf{u}. \quad (2.5)$$

En cambio, si el desplazamiento se produce sin cambios en el rotacional entonces se cumple

$$\nabla \times \mathbf{u} = \operatorname{rot}(\mathbf{u}) = 0, \quad (2.6)$$

donde el campo de onda \mathbf{u} representa ondas longitudinales (longitudinal waves o P-waves). Usando $\Delta \mathbf{u} = \nabla \operatorname{div} \mathbf{u} - \operatorname{rot} \operatorname{rot} \mathbf{u}$ en (2.3) y considerando (2.6) se obtiene la ecuación de ondas (2.7)

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + 2\mu) \Delta \mathbf{u}. \quad (2.7)$$

Nos interesan aplicaciones en las que se puede trabajar con ondas escalares de cizalla (elastografía médica) o longitudinales (geofísica). Por tanto en lo sucesivo el campo de desplazamientos será un campo escalar regido por una ecuación de ondas (ecuación (2.5) o ecuación (2.7)).

Al haber cambios bruscos de medio, en general, μ y λ variarán en espacio. Si, por ejemplo, son constantes a trozos, en cada región tendremos una ecuación de ondas de tipo (2.5) o (2.7) en términos de las constantes λ, μ de la región, que se acoplarán a las de la regiones colindantes mediante condiciones de transmisión, es decir, a través de continuidad de los campos de desplazamiento y de las componentes normales de las tensiones. Estos problemas de transmisión se pueden reformular en términos de ecuaciones con divergencias de parámetros dependientes del espacio, como indicamos a continuación.

2.1.1. Modelo matemático directo planteado en un dominio infinito

Consideraremos que la dinámica de la onda, emitida por fuentes dentro del dominio computacional, está regida por una ecuación de onda escalar con fuerzas externas. Tomando como dominio computacional R un semiespacio de \mathbb{R}^n , el problema de la transmisión de ondas queda expresado por el siguiente conjunto de ecuaciones [24],

$$\rho(\mathbf{x}) \frac{\partial^2 u}{\partial t^2} - \operatorname{div}(\chi(\mathbf{x}) \nabla u) = \rho(\mathbf{x}) f(t) g(\mathbf{x}), \quad \mathbf{x} \in R, \quad t > 0, \quad (2.8)$$

$$\frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}, t) = 0, \quad \mathbf{x} \in \partial R, \quad t > 0, \quad (2.9)$$

$$u(\mathbf{x}, 0) = u_t(\mathbf{x}, 0) = 0, \quad \mathbf{x} \in R, \quad (2.10)$$

donde:

- $f(t)$ es la variación temporal de la onda, por ejemplo, $f(t) = f_0(1 - 2\pi^2 f_M^2 t^2)e^{-\pi^2 f_M^2 t^2}$.
- $g(\mathbf{x})$ es perturbación emitida por las fuentes, por ejemplo, $g(\mathbf{x}) = \frac{1}{(\pi\kappa)^{n/2}} \sum_{k=1}^K e^{-\frac{|\mathbf{x}-\mathbf{x}_k|^2}{\kappa}}$, siendo $\kappa > 0$ un parámetro de regularización y n la dimensión del problema. En este caso se considera $n = 2$. El índice k hace referencia a la secuencia de emisores, \mathbf{x}_k .
- $\rho(\mathbf{x})$ es la densidad del medio.
- $\chi(\mathbf{x})$ representa el módulo de elasticidad del medio seleccionado para el estudio ($\lambda(\mathbf{x}) + 2\mu(\mathbf{x})$ o $\mu(\mathbf{x})$). Tomará valores distintos dentro de las inclusiones de otros materiales (o anomalías de un tejido) y en el material matriz (o tejido).

El lado derecho $f(t)g(\mathbf{x})$ de la igualdad de la ecuación (2.8) corresponde a la fuerza externa por unidad de volumen emitida por las fuentes. Cuando $\chi(\mathbf{x})$ es constante por trozos, tenemos ecuaciones de onda con diferentes constantes en el medio R y en las inclusiones $\Omega = \cup_{i=1}^L \Omega_i$, las cuales están acopladas por condiciones de transmisión de onda. En este problema de transmisión de ondas, las condiciones iniciales son desplazamiento y velocidad nulas, es decir, $u(\mathbf{x}, 0) = u_t(\mathbf{x}, 0) = 0$.

En principio, la región R es infinita. Sin embargo, por motivos computacionales, la región se truncará a un recinto finito imponiendo condiciones de contorno adecuadas, como se detalla en la sección 2.2.

2.1.2. Operador de observación

Denotamos por $\boldsymbol{\nu}$ la colección de parámetros de los que depende el modelo, es decir, aquellos parámetros desconocidos que deseamos averiguar a partir de observaciones. En nuestro contexto, serán parámetros que definen la geometría y naturaleza de las anomalías presentes en la región. El operador de observación $f(\boldsymbol{\nu})$ desempeña un papel fundamental al establecer la conexión matemática entre el espacio de los parámetros del modelo $\boldsymbol{\nu}$ y el espacio de las observaciones reales $u_{\boldsymbol{\nu}}$. Su representación matemática es esencial para comprender cómo el modelo matemático se relaciona con las mediciones empíricas.

En nuestro caso, el operador de observación puede tomar varias formas. La más sencilla consiste en observar desplazamientos en los receptores \mathbf{r}_j , $j = 1, \dots, J$, durante una secuencia de tiempos:

$$f(\boldsymbol{\nu}) = (u_{\boldsymbol{\nu}}(r_j, t_m))_{j=1, \dots, J, m=1, \dots, M}, \quad (2.11)$$

siendo $u_{\boldsymbol{\nu}}$ la solución del problema directo con parámetros $\boldsymbol{\nu}$. Los valores de la solución en puntos no siempre están definidos, depende de la regularidad de la solución del problema directo, que estudiaremos más adelante. Con más generalidad se puede considerar trazas de la solución en secuencias de pequeños segmentos o integrales de la solución en torno a los puntos de recepción.

Una vez establecido el operador de observación, se asume que los datos medidos se relacionan con él mediante términos de ruido [20]:

$$\mathbf{d}^{\text{obs}}(\boldsymbol{\nu}) = f(\boldsymbol{\nu}) + \boldsymbol{\eta}, \quad (2.12)$$

donde $\boldsymbol{\eta}$ corresponde al ruido añadido como una distribución Gaussiana aleatoria de media cero y covarianza \mathbf{C}_n . A menudo se considera $\mathbf{C}_n = \sigma^2 \mathbf{I}$, siendo σ la desviación estándar máxima del ruido, e \mathbf{I} una matriz identidad de tamaño $P \times P$ siendo P el número de parámetros desconocidos. Es decir, el ruido añadido es posible escribirlo como $\boldsymbol{\eta} = \mathcal{N}(0, \mathbf{C}_n)$. Con la adición de esta componente estocástica se capturan las imperfecciones y la incertidumbre en el proceso de medición.

2.2. Estudio teórico del problema directo truncado

Como se discutió en la sección 2.1.1, el problema directo se formula inicialmente en un dominio semi-infinito $R \subset \mathbb{R}^2$, cuya frontera superior contiene emisores en las posiciones x_j y receptores en las posiciones r_k . Dado un tiempo $\tau > 0$, y considerando la velocidad finita de propagación de las ondas, es posible reducir el análisis a una región acotada mediante el recorte del dominio R a un subdominio rectangular $R_\tau \subset \mathbb{R}^2$. Esta región truncada R_τ se define como:

$$R_\tau = \{(x, y) \in \mathbb{R}^2 \mid x \in [x_{\min}, x_{\max}], y \in [y_{\min}, y_{\max}]\}$$

donde los valores de $x_{\min} = -r_\tau$, $x_{\max} = r_\tau$ y $y_{\min} = -r_\tau$ se eligen suficientemente grandes para asegurar que las ondas emitidas no alcanzan los bordes en el intervalo considerado. La pared $y_{\max} = 0$ es parte de la frontera original que contiene emisores y receptores y conserva la condición de contorno. En estas condiciones, el problema de propagación de ondas planteado en el dominio original R puede sustituirse por un problema equivalente en R_τ , con condiciones de frontera homogéneas de Dirichlet o Neumann impuestas en los bordes inferiores y laterales durante $t \in [0, \tau]$ (véase sección 2.5 y [1]). Esta reformulación permite una implementación computacional más eficiente. Por lo tanto, el problema directo asociado a las ecuaciones (2.8)-(2.10) puede escribirse de manera equivalente en R_τ , como se presenta a continuación,

$$\rho(\mathbf{x}) \frac{\partial^2 u}{\partial t^2} - \operatorname{div}(\chi(\mathbf{x}) \nabla u) = \rho(\mathbf{x}) f(t) g(\mathbf{x}), \quad \mathbf{x} \in R_\tau, \quad t > 0, \quad (2.13)$$

$$\frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}, t) = 0, \quad \mathbf{x} \in \partial R_\tau, \quad t > 0, \quad (2.14)$$

$$u(\mathbf{x}, 0) = u_t(\mathbf{x}, 0) = 0, \quad \mathbf{x} \in R_\tau. \quad (2.15)$$

La existencia y regularidad de soluciones para el problema descrito en las ecuaciones (2.13)-(2.15) se puede abordar mediante métodos de Galerkin, como detallamos en la siguiente sección. El método de Galerkin busca soluciones aproximadas de una formulación débil del problema directo reduciendo la dimensión del espacio de soluciones mediante bases adecuadas con error controlado.

En la práctica, puede ser conveniente utilizar regiones aún más pequeñas para reducir el coste computacional. En este caso, imponemos en las nuevas paredes artificiales condiciones no reflectantes, como comentaremos más adelante [1].

2.2.1. Formulación débil del problema directo truncado

Para derivar la formulación débil del problema de ondas, se debe multiplicar la ecuación (2.13) en su forma fuerte por una función de prueba $\phi \in H^1(R_\tau)$. El espacio de Sobolev $H^1(R_\tau)$ está formado por funciones de cuadrado integrable en R_τ cuyas derivadas parciales en el sentido de las distribuciones son también de cuadrado integrable en R_τ [80]. A continuación, esta expresión se integra sobre el dominio del problema directo truncado R_τ y se realiza una integración por partes para reemplazar las derivadas de segundo orden por derivadas de primer orden. De esta forma se obtiene a una formulación débil que puede admitir soluciones poco regulares [80]. La nueva formulación proporciona una base sólida para la discretización espacial y temporal del problema mediante métodos de elementos finitos, facilitando así la solución numérica del sistema de ecuaciones resultante [99]. Multiplicando la ecuación por la función test e integrando tenemos

$$\begin{aligned} \int_{R_\tau} \rho(\mathbf{x}) \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} \phi(\mathbf{x}) d\mathbf{x} + \int_{R_\tau} \chi(\mathbf{x}) \nabla u(\mathbf{x}, t) \nabla \phi(\mathbf{x}) d\mathbf{x} - \int_{\partial R_\tau} \chi(\mathbf{x}) \nabla u(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \\ = f(t) \int_{R_\tau} \rho(\mathbf{x}) g(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}, \quad t > 0, \end{aligned} \quad (2.16)$$

supuesto que $u(\mathbf{x}, t)$ sea suficientemente regular (al menos H^1) en espacio. Usamos la condición de contorno (2.14) para llegar a

$$\int_{R_\tau} \rho(\mathbf{x}) \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} \phi(\mathbf{x}) d\mathbf{x} + \int_{R_\tau} \chi(\mathbf{x}) \nabla u(\mathbf{x}, t) \nabla \phi(\mathbf{x}) d\mathbf{x} = f(t) \int_{R_\tau} \rho(\mathbf{x}) g(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}, \quad t > 0. \quad (2.17)$$

La formulación débil del problema truncado (2.13)-(2.15) es entonces encontrar

$$u(\mathbf{x}, t) \in H_\tau = \{u \in C([0, \tau]; H^1(R_\tau)), u \in C^1([0, \tau]; L^2(R_\tau))\}$$

la cual satisface las condiciones iniciales $u(\mathbf{x}, 0) = u_t(\mathbf{x}, 0) = 0$ en el subdominio R_τ y tal que

$$\int_{R_\tau} \rho(\mathbf{x}) u_{tt}(\mathbf{x}, t) \phi(\mathbf{x}) d\mathbf{x} + \int_{R_\tau} \chi(\mathbf{x}) \nabla u(\mathbf{x}, t) \nabla \phi(\mathbf{x}) d\mathbf{x} = f(t) \int_{R_\tau} \rho(\mathbf{x}) g(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}, \quad (2.18)$$

para todas las funciones de prueba $\phi \in H^1(R_\tau)$. Denotamos por H^m al espacio estándar de Sobolev y por $C^m([0, \tau]; B)$ el espacio estándar de las funciones continuas cuyas derivadas temporales son continuas hasta la m -ésima derivada con valores en espacios de Banach [65].

Si la dimensión es $n = 2$, se puede definir un operador de observación que consista en medir desplazamientos de la onda en diferentes puntos y tiempos si

$$u(\mathbf{x}, t) \in C([0, \tau], H^2(\overline{\Omega})), \quad u_t(\mathbf{x}, t) \in C([0, \tau], H^1(R_\tau)), \quad u_{tt}(\mathbf{x}, t) \in C([0, \tau], L^2(R_\tau)),$$

donde $\overline{\Omega}$ denota el cierre de un subdominio $\Omega \subset R_\tau$ que representa la región donde están localizados los receptores.

Habremos de estudiar bajo qué condiciones es ésto posible. En realidad, basta $u(\mathbf{x}, t) \in C([0, \tau], H^2(E))$ en un entorno E de la parte de ∂R que contiene los receptores.

2.2.2. Resultados de existencia y regularidad mediante el método de Galerkin

El método de Galerkin, utilizado en diversos campos de la física y la ingeniería, se fundamenta en el uso de bases numerables para los espacios de funciones en los que se buscan soluciones. A partir de ellas se pueden construir jerarquías de espacios finito-dimensionales en los cuales es posible construir soluciones aproximadas. Esta técnica es particularmente útil en problemas de ecuaciones en derivadas parciales, donde la complejidad de la geometría o las condiciones de contorno pueden hacer que la solución exacta sea difícil o imposible de obtener de forma analítica.

Al seleccionar funciones de base para el espacio de prueba $H^1(R_\tau)$, el método de Galerkin genera aproximaciones en dimensión finita al resolver sistemas de ecuaciones diferenciales que permiten determinar los coeficientes desconocidos. Estos coeficientes definen la aproximación de la solución en el espacio de funciones de base, y conducen a soluciones aproximadas de precisión creciente al problema original.

La característica distintiva del método de Galerkin es su capacidad para garantizar la convergencia de la solución aproximada hacia la solución exacta a medida que se incrementa la dimensionalidad del espacio de funciones de base. Este aspecto es esencial para asegurar la precisión y la fiabilidad de las aproximaciones numéricas obtenidas mediante este método [100].

Siguiendo esta línea, se tiene el siguiente resultado de existencia, unicidad y regularidad de la solución $u(\mathbf{x}, t)$ obtenida a través del procedimiento de Galerkin con bases espectrales [80].

Teorema 1: *Existencia, unicidad y regularidad de la solución $u(\mathbf{x}, t)$.*

Sea el tiempo $\tau > 0$, y el subdominio computacional $R_\tau = (-r_\tau, r_\tau) \times (0, -r_\tau)$ y $R'_\tau = [-r_\tau/2, r_\tau/2] \times [0, -r_\tau/2]$, $r_\tau > 0$. Denotamos $H_\tau = \{u \in C([0, \tau]; H^1(R_\tau)), u \in C^1([0, \tau]; L^2(R_\tau))\}$. Se asume lo siguiente:

- $\rho(\mathbf{x}), \chi(\mathbf{x}) \in L^\infty(R_\tau)$ con $\rho(\mathbf{x}) > k_1 > 0$, $\chi(\mathbf{x}) > k_2 > 0$,
- $f(t) \in C^2([0, \tau])$ y $g(\mathbf{x}) \in C^2(\overline{R_\tau}) \cap C_c(R'_\tau)$ tal que $\frac{\partial g}{\partial \mathbf{n}} = 0$ en $y = 0$.

Entonces, el problema directo (2.18) tiene una solución única $u(\mathbf{x}, t) \in H_\tau$ tal que $u(\mathbf{x}, 0) = 0$ y $u_t(\mathbf{x}, 0) = 0$. Además, $u_t \in H_\tau$.

Demostración: Definimos una base de $H^1(R_\tau)$ de la forma siguiente. Se considera el operador $T : L^2(R_\tau) \rightarrow L^2(R_\tau)$ definido por $Tv = u$, donde $u(\mathbf{x}) \in H^1(R_\tau)$ es la solución débil única de la ecuación $-\text{div}[\chi(\mathbf{x})\nabla u] = \rho(\mathbf{x})v$ con condiciones de contorno Neumann nulas. Procediendo como en el Teorema IX.31 de [14], este operador T es autoadjunto y compacto. Además satisface $(Tv, v) \geq 0$ para toda función $v \in L^2$. Por lo tanto, T admite una secuencia de autovalores no negativos $\mu_1 = 0 < \mu_2 < \dots < \mu_i \rightarrow 0$, con sus autofunciones asociadas ψ_1 constante and $\psi_i \in H^1(R_\tau)$ para $i > 1$. Si ϕ_i es una autofunción del operador T , con autovalor asociado $\mu_i > 0$, entonces, para cualquier función $w \in H^1(R_\tau)$ se cumple la siguiente igualdad:

$$\int_{R_\tau} \chi(\mathbf{x})\nabla\psi_i\nabla w \, d\mathbf{x} = \frac{1}{\mu_i} \int_{R_\tau} \rho(\mathbf{x})\psi_i w \, d\mathbf{x}, \quad \forall w \in H^1(R_\tau). \quad (2.19)$$

Una vez normalizadas, las autofunciones ψ_i asociadas a diferentes autovalores μ_i satisfacen la condición de ortonormalidad $\int_{R_\tau} \rho(\mathbf{x})\psi_i(\mathbf{x})\psi_j(\mathbf{x}) \, d\mathbf{x} = \delta_{ij}$, siendo δ_{ij} la delta de Kronecker. Si $i = j$ entonces $\int_{R_\tau} \rho\psi_i^2 \, d\mathbf{x} = 1$, y si $i \neq j$ entonces $\int_{R_\tau} \rho(\mathbf{x})\psi_i(\mathbf{x})\psi_j(\mathbf{x}) \, d\mathbf{x} = 0$.

Estas autofunciones son una base de $L^2(R_\tau)$. Por tanto, podemos buscar una solución $u(\mathbf{x}, t)$ del problema de ondas como una combinación de autofunciones $\psi_i(\mathbf{x})$. Consideramos el espacio V_N formado por las primeras N autofunciones. Buscamos una solución $u_N(\mathbf{x}, t) = \sum_{i=1}^N c_i(t)\psi_i(\mathbf{x}) \in C^2([0, \tau], V_N)$ de

$$\frac{d^2}{dt^2} \int_{R_\tau} \rho u_N(t) w \, d\mathbf{x} + \int_{R_\tau} \chi \nabla u_N(t) \nabla w \, d\mathbf{x} = f(t) \int_{R_\tau} \rho g w \, d\mathbf{x}, \quad \forall w \in V_N,$$

para $w \in V_N$, con $u_N(0) = u_{N,t}(0) = 0$. Tomando $w = \psi_j$ y sustituyendo $u(\mathbf{x}, t) = \sum_{i=1}^\infty c_i(t)\psi_i(\mathbf{x})$ en la ecuación variacional, obtenemos:

$$\begin{aligned} & \int_{R_\tau} \rho(\mathbf{x}) \frac{\partial^2}{\partial t^2} \left(\sum_{i=1}^N c_i(t)\phi_i(\mathbf{x}) \right) \phi_j(\mathbf{x}) \, d\mathbf{x} + \int_{R_\tau} \chi(\mathbf{x}) \nabla \left(\sum_{i=1}^N c_i(t)\phi_i(\mathbf{x}) \right) \cdot \nabla \phi_j(\mathbf{x}) \, d\mathbf{x} \\ & = f(t) \int_{R_\tau} \rho(\mathbf{x}) g(\mathbf{x}) \phi_j(\mathbf{x}) \, d\mathbf{x}, \end{aligned}$$

para cada $j = 1, \dots, N$. Utilizando la ortonormalidad de las autofunciones y la relación con el operador autoadjunto T , llegamos a:

$$c_j''(t) + \gamma_j c_j(t) = f(t) \int_{R_\tau} \rho(\mathbf{x}) g(\mathbf{x}) \phi_j(\mathbf{x}) \, d\mathbf{x} = f(t) g_j,$$

donde $\gamma_j = \frac{1}{\mu_j}$ y $g_j = \int_{R_\tau} \rho(\mathbf{x}) g(\mathbf{x}) \phi_j(\mathbf{x}) \, d\mathbf{x}$, para $j > 1$. La solución es:

$$c_j(t) = \frac{1}{\sqrt{\gamma_j}} \int_0^t \sin(\sqrt{\gamma_j}(t-s)) f(s) g_j \, ds, \quad \gamma_j > 0.$$

Para el caso especial cuando $\gamma_1 = 0$, la solución es:

$$c_1(t) = \int_0^t \int_0^s f(\sigma) g_1 d\sigma ds.$$

Este proceso define una sucesión de Cauchy u_N en H_τ porque

$$\begin{aligned} & \int_{R_\tau} [\chi |\nabla(u_p(t) - u_m(t))|^2 + \rho |u'_p(t) - u'_m(t)|^2] d\mathbf{x} \\ &= \sum_{i=m+1}^p [\gamma_i |c_i(t)|^2 + |c'_i(t)|^2] \leq 2\tau \int_0^\tau |f(s)|^2 ds \sum_{i=m+1}^p |g_i|^2 \end{aligned}$$

tiende a cero según $p, m \rightarrow \infty$, $p \geq m > 1$. Como H_τ es completo, la secuencia $u(\mathbf{x}, t) \rightarrow u$ en H_τ conforme $N \rightarrow \infty$. Se tiene además $u(0) = u_t(0) = 0$.

Multiplicando la ecuación satisfecha por u_N por $\eta(t) \in C_c^\infty(0, \tau)$ y tomando $w = \psi_i$, obtenemos

$$\begin{aligned} \int_0^\tau \eta''(t) \int_{R_\tau} \rho u_N(t) \psi_i d\mathbf{x} dt + \int_0^\tau \eta(t) \int_{R_\tau} \chi \nabla u_N(t) \nabla \psi_i d\mathbf{x} dt \\ = \int_0^\tau \eta(t) f(t) dt \int_{R_\tau} \rho g \psi_i d\mathbf{x}. \end{aligned}$$

En el límite $N \rightarrow \infty$, u satisface esta ecuación para todo ψ_i , es decir, por densidad para todo $w \in H^1(R_\tau)$, y para todo $\eta(t)$. Por tanto, $u = \sum_{i=1}^\infty c_i(t) \psi_i(\mathbf{x})$ es una solución de nuestro problema. En particular,

$$\rho(\mathbf{x}) u_{tt} = \operatorname{div}(\chi(\mathbf{x}) \nabla u) + \rho(\mathbf{x}) f(t) g(\mathbf{x})$$

en el sentido de las distribuciones en $R_\tau \times (0, \tau)$.

Para obtener $u_t \in H_\tau$, derivamos la ecuación con respecto a t . Vemos que u_t es solución de un problema similar reemplazando f por f' [65], con $\rho(\mathbf{x}) u_{tt}(\mathbf{x}, 0) = f(0) \rho(\mathbf{x}) g(\mathbf{x}) \in L^2(R_\tau)$. Cuando $f(0) = 0$, la argumentación anterior muestra que $u_t \in H_\tau$. En otro caso, hemos de modificar esa argumentación con la fórmula $c_i(t) = \frac{1}{\sqrt{\gamma_i}} f(0) g_i \sin(\sqrt{\gamma_i} t) + \frac{1}{\sqrt{\gamma_i}} \int_0^t \sin(\sqrt{\gamma_i}(t-s)) f'(s) g_i ds$, $\gamma_i > 0$, para probar que tenemos una sucesión de Cauchy [80].

Para demostrar la unicidad, asumimos que tenemos dos soluciones $u_1(\mathbf{x}, t)$, $u_2(\mathbf{x}, t)$ en H_τ . La diferencia entre las dos soluciones $u(\mathbf{x}, t) = u_1(\mathbf{x}, t) - u_2(\mathbf{x}, t)$ satisface el problema de contorno con condiciones iniciales y segundo miembro nulos. Multiplicando la ecuación diferencial por $u_t(\mathbf{x}, t)$ e integrando se obtiene [29]

$$\begin{aligned} & \int_{R_\tau} \rho(\mathbf{x}) |u_t(\mathbf{x}, t)|^2 d\mathbf{x} + \int_{R_\tau} \chi(\mathbf{x}) |\nabla u(\mathbf{x}, t)|^2 d\mathbf{x} \\ &= \int_{R_\tau} \rho(\mathbf{x}) |u_t(\mathbf{x}, 0)|^2 d\mathbf{x} + \int_{R_\tau} \chi(\mathbf{x}) |\nabla u(\mathbf{x}, 0)|^2 d\mathbf{x} = 0. \end{aligned}$$

Esto implica que la solución $u(\mathbf{x}, t)$ es constante en espacio y en tiempo y que las soluciones $u_1(\mathbf{x}, t) = u_2(\mathbf{x}, t)$ son únicas. Por lo tanto, la única solución está dada por una expansión en términos de autofunciones, es decir, $u(\mathbf{x}, t) = \sum_{i=1}^\infty c_i(t) \psi_i(\mathbf{x})$, obtenida como límite de las aproximaciones u_N .

Se puede argumentar regularidad H^2 en un entorno del borde donde están localizados los receptores para $t \leq \tau$ como en [80].

Corolario 1: *Condición de Transmisión en las Interfaces.*

Bajo las hipótesis del Teorema 1, consideremos un subdominio computacional R_τ compuesto por L regiones disjuntas Ω_τ^ℓ , $\ell = 1, \dots, L$, donde las propiedades físicas $\rho(\mathbf{x})$ y $\chi(\mathbf{x})$ son constantes a trozos dentro de cada región. La solución del problema variacional (2.16) satisface las siguientes condiciones:

1. **Regularidad de la Solución:** Para cada región Ω_τ^ℓ y para todo tiempo en el intervalo $[0, \tau]$, la solución $u(\mathbf{x}, t)$ pertenece a $H^2(\Omega_\tau^\ell)$.

Cuando $n = 2$, la solución $u(\mathbf{x}, t)$ pertenece a $C^0(\overline{\Omega_\tau^\ell})$ para cada región Ω_τ^ℓ .

2. **Condición de Transmisión en las Interfaces:** En cada interfaz Γ que separa dos regiones adyacentes Ω_τ^1 y Ω_τ^2 , se cumplen las siguientes condiciones:

$$\begin{aligned} u^+ &= u^-, \\ \chi(\mathbf{x})^- \nabla u^- \cdot \mathbf{n} &= \chi(\mathbf{x})^+ \nabla u^+ \cdot \mathbf{n}, \end{aligned} \quad \text{en } L^2(\Gamma), \quad (2.20)$$

donde u^+ y u^- son los valores de la solución $u(\mathbf{x}, t)$ obtenidos desde cada lado de la interfaz, siguiendo la dirección del vector normal unitario \mathbf{n} .

Demostración: En cada región Ω_τ^ℓ , la ecuación diferencial que describe el comportamiento de la solución $u(\mathbf{x}, t)$ se puede reescribir como:

$$\chi(\mathbf{x})^\ell \Delta u(\mathbf{x}, t) = \rho(\mathbf{x})^\ell u_{tt}(\mathbf{x}, t) - f(t) \rho(\mathbf{x})^\ell g(\mathbf{x}), \quad t > 0,$$

donde $\chi(\mathbf{x})^\ell$ y $\rho(\mathbf{x})^\ell$ son constantes que representan propiedades físicas en la región Ω_τ^ℓ . La regularidad en H^2 se deriva de la teoría de regularidad elíptica [53] y de los resultados de las inyecciones de Sobolev [14]. Estos teoremas establecen propiedades de regularidad para las soluciones de ciertas ecuaciones diferenciales parciales, como la ecuación de Laplace. Aunque existe esta regularidad local en H^2 , las soluciones pueden presentar discontinuidades o singularidades que afectan su regularidad global en $H^2(R_\tau)$.

Con las inyecciones de Sobolev [14], sabemos que en dimensión 2, el espacio $H^2(\Omega^\ell)$ está contenido en el espacio de funciones continuas en el cierre del dominio $\overline{\Omega^\ell}$. En nuestro contexto, las soluciones se definen en la interfaz superior del dominio, que denotamos como Ω^ℓ . Por lo tanto, podemos considerar el operador dado por la ecuación (2.11) en este contexto.

Para obtener la condición de transmisión, consideremos dos regiones adyacentes Ω^1 y Ω^2 con una interfaz común Γ . Dada una función $u(\mathbf{x}, t) \in H^1(\Omega^1 \cup \Omega^2)$, y denotando por u^+ y u^- los valores en el límite de interfaz tomados desde Ω^1 y Ω^2 respectivamente, se tiene $u^+ = u^-$ en la interfaz Γ en el sentido de las trazas. Además, para cualquier función de prueba $\phi \in H_0^1(\Omega^1 \cup \Omega^2)$, se cumple la siguiente identidad,

$$\begin{aligned} \sum_{\ell=1}^2 \int_{\Omega^\ell} \chi^\ell \nabla u \cdot \nabla \phi \, dx &= - \sum_{\ell=1}^2 \int_{\Omega^\ell} \operatorname{div}(\chi^\ell \nabla u) \phi \, dx + \int_{\Gamma} [\chi^- \nabla u^- - \chi^+ \nabla u^+] \cdot \mathbf{n} \phi \, dS_x \\ &= \int_{\Omega^1 \cup \Omega^2} \chi \nabla u \cdot \nabla \phi \, dx - \int_{\Omega^1 \cup \Omega^2} \operatorname{div}(\chi \nabla u) \phi \, d\mathbf{x}. \end{aligned}$$

Esta igualdad se deriva de la aplicación de la versión débil de la fórmula de Green [80] a las ecuaciones diferenciales que describen el comportamiento de $u(\mathbf{x}, t)$ en los dominios Ω^1 y Ω^2 , teniendo en cuenta las condiciones de continuidad en la interfaz Γ . Por lo tanto, las relaciones de transmisión se mantienen en cualquier interfaz de discontinuidad Γ .

2.3. Resolución numérica del problema directo truncado: discretización en espacio

El método de elementos finitos ha surgido como una herramienta poderosa para la resolución numérica de una amplia variedad de problemas en diversas áreas científicas e ingenieriles [78]. Su versatilidad radica en la capacidad de modelar geometrías complejas y resolver ecuaciones diferenciales parciales con precisión y eficiencia. Si bien la demostración de existencia que hemos dado es constructiva y usa una base, esta base depende de los coeficientes de la ecuación. Si cambiamos los coeficientes, hemos de cambiar la base. Por otra parte, es una base global, el cálculo de integrales que involucran las funciones de base es costoso. Los métodos de elementos finitos proporcionan bases que no dependen de los coeficientes del problema, sólo del dominio espacial. Por otra parte, son bases locales, nulas excepto en el entorno de un nodo, lo que facilita el cálculo de integrales que involucran funciones de base.

El núcleo del método de elementos finitos reside en el concepto de discretización espacial, un proceso esencial que convierte un dominio continuo en una representación discreta. Esta estrategia es fundamental para abordar problemas definidos en geometrías complejas y para aproximar soluciones de ecuaciones diferenciales parciales, permitiendo su tratamiento numérico.

La discretización espacial implica la división del dominio computacional en elementos más simples y manejables. Cada elemento se caracteriza por nodos, puntos de interés donde aproximaremos la solución del problema. Al emplear funciones de forma que describen la variación del campo físico entre nodos adyacentes, podemos representar la solución global del problema combinando las soluciones locales de cada elemento.

La elección del tipo y la forma de los elementos, así como la distribución de los nodos, depende de la naturaleza del problema y de las características particulares de la geometría. Los elementos triangulares y cuadriláteros son ejemplos comunes utilizados en problemas bidimensionales, mientras que los tetraedros y hexaedros se emplean en contextos tridimensionales [9].

La discretización espacial confiere al método de elementos finitos la capacidad de aproximar soluciones precisas. Además, la adaptabilidad de la malla y la posibilidad de refinarla en áreas de interés permiten obtener resultados detallados sin sacrificar la eficiencia computacional.

En esta tesis, las geometrías estudiadas son bidimensionales y el dominio a mallar es rectangular. Por lo tanto, utilizaremos mallas rectangulares divididas en elementos triangulares, que cubren completamente el dominio computacional.

Para establecer la discretización espacial, definimos un paso espacial δx y δy , que corresponden a la longitud de los triángulos en las direcciones horizontal y vertical, respectivamente. Construimos un espacio de elementos finitos $V^B \subset H^1(R)$ con funciones de base ϕ_1, \dots, ϕ_B , donde B es el número total de nodos (internos y en la frontera). Cada función u^B se expresa en esta base como [37],

$$u^B(\mathbf{x}, t) = \sum_{b=1}^B a_b(t) \phi_b(\mathbf{x}). \quad (2.21)$$

Esta es una solución de la formulación débil aproximada:

$$\begin{aligned} \frac{\partial^2}{\partial t^2} \int_{R_\tau} \rho u^B(\mathbf{x}, t) w \, d\mathbf{x} + \int_{R_\tau} \chi \nabla u^B(\mathbf{x}, t) \nabla w \, d\mathbf{x} &= f(t) \int_{R_\tau} \rho g w \, d\mathbf{x}, \\ u^B(\mathbf{x}, 0) &= 0, \quad u_t^B(\mathbf{x}, 0) = 0, \end{aligned} \quad (2.22)$$

para todas las funciones w en el espacio V^B , que es el espacio generado por las funciones base ϕ_1, \dots, ϕ_B . Los coeficientes $a_i(t)$ se calculan resolviendo un sistema de ecuaciones diferenciales.

Teorema 2: *Existencia de aproximante.*

Para $f \in C^2([0, \tau])$, existe una única solución $u(\mathbf{x}, t) \in C^2([0, \tau]; V^B)$ que satisface la ecuación diferencial (2.18). Además, si $f \in C^k((0, \tau])$, entonces $u(\mathbf{x}, t) \in C^{k+2}([0, \tau]; V^B)$ para $k \geq 1$.

Demostración. Consideremos la función de prueba $w = \phi_j$, con $j = 1, \dots, B$ y supongamos que la solución tiene la forma $u(\mathbf{x}, t) = \sum_{i=1}^B a_i(t) \phi_i(\mathbf{x})$. Entonces, la ecuación (2.18) es equivalente a un sistema de ecuaciones diferenciales ordinarias, es decir,

$$\sum_{i=1}^B a_i''(t) \int_{R_\tau} \rho \phi_i \phi_j d\mathbf{x} + \sum_{i=1}^B a_i(t) \int_{R_\tau} \chi(\mathbf{x}) \nabla \phi_i \nabla \phi_j d\mathbf{x} = f(t) \int_{R_\tau} \rho(\mathbf{x}) g \phi_j d\mathbf{x}, \quad (2.23)$$

$$a_i(0) = a_i'(0) = 0, \quad i = 1, \dots, B.$$

Escribiendo la ecuación (2.23) en su forma vectorial, obtenemos el sistema:

$$\mathbf{B}\mathbf{a}''(t) + \mathbf{C}\mathbf{a}(t) = f(t)\mathbf{h},$$

donde las matrices y vectores están dados por las siguientes expresiones,

$$B_{j,i} = \int_{R_\tau} \rho(\mathbf{x}) \phi_j \phi_i d\mathbf{x}, \quad C_{j,i} = \int_{R_\tau} \chi(\mathbf{x}) \nabla \phi_j \nabla \phi_i d\mathbf{x}, \quad h_j = \int_{R_\tau} \rho(\mathbf{x}) g \phi_j d\mathbf{x}, \quad (2.24)$$

para $j, i = 1, \dots, B$.

Introduciendo el vector de velocidades $\mathbf{v}(t) = \mathbf{a}'(t)$, podemos escribir el sistema de ecuaciones diferenciales ordinarias de segundo orden como un sistema de primer orden:

$$\mathbf{B}\mathbf{v}'(t) + \mathbf{C}\mathbf{a}(t) = f(t)\mathbf{h}, \quad \mathbf{a}'(t) = \mathbf{v}(t),$$

junto con las condiciones iniciales:

$$\mathbf{a}(0) = \mathbf{0}, \quad \mathbf{v}(0) = \mathbf{0}.$$

Dado que la matriz \mathbf{B} es simétrica y positiva definida, podemos garantizar la existencia y unicidad de la solución para el sistema de ecuaciones diferenciales ordinarias lineales en $[0, \tau]$. La regularidad de la solución $\mathbf{a}(t)$ y $\mathbf{v}(t)$ depende de la regularidad de $f(t)$. Si $f(t) \in C^k((0, \tau])$, entonces $\mathbf{a}(t) \in C^{k+2}([0, \tau])$.

Por lo tanto, existe una única solución $u(\mathbf{x}, t) \in C^2([0, \tau]; V^B)$ que satisface la ecuación diferencial (2.18), y su regularidad aumenta con la regularidad de $f(t)$.

2.4. Resolución numérica del problema directo truncado: discretización en tiempo

En el marco del método de elementos finitos, la discretización temporal desempeña un papel crucial en la resolución numérica del problema directo truncado de la ecuación (2.18). Esta técnica

de discretización temporal nos permite transformar las ecuaciones diferenciales parciales continuas en un sistema algebraico discreto, lo que facilita su implementación en computadoras y nos permite obtener soluciones aproximadas a problemas complejos en ingeniería y ciencias aplicadas. En esta sección, nos centraremos en la discretización temporal del problema directo truncado. Dado que nuestro fin es resolver grandes cantidades de problemas de ondas, elegiremos aproximaciones básicas de bajo coste, evaluando su impacto en la estabilidad, precisión y eficiencia computacional del método de elementos finitos. Al comprender la importancia y los efectos de la discretización temporal en nuestros cálculos numéricos, podremos avanzar hacia la resolución efectiva de problemas reales [37].

2.4.1. Discretización en tiempo de la formulación débil

En la expresión (2.16) aparece una segunda derivada de $u(\mathbf{x}, t)$ con respecto al tiempo. Formalmente, buscamos una aproximación usando serie de Taylor en torno al parámetro t con paso δt , como se observa en las ecuaciones (2.25) y (2.26),

$$u(\mathbf{x}, t + \delta t) = u(\mathbf{x}, t) + \delta t u_t(\mathbf{x}, t) + \frac{\delta t^2}{2} u_{tt}(\mathbf{x}, t) + \frac{\delta t^3}{6} u_{ttt}(\mathbf{x}, t) + \frac{\delta t^4}{24} u_{tttt}(\mathbf{x}, t) + O(\delta t^5), \quad (2.25)$$

$$u(\mathbf{x}, t - \delta t) = u(\mathbf{x}, t) - \delta t u_t(\mathbf{x}, t) + \frac{\delta t^2}{2} u_{tt}(\mathbf{x}, t) - \frac{\delta t^3}{6} u_{ttt}(\mathbf{x}, t) + \frac{\delta t^4}{24} u_{tttt}(\mathbf{x}, t) + O(\delta t^5). \quad (2.26)$$

Combinando los desarrollos de las ecuaciones (2.25) y en (2.26) se obtiene la aproximación u_{tt} de segundo orden,

$$u_{tt}(\mathbf{x}, t) \simeq \frac{1}{\delta t^2} [u(\mathbf{x}, t + \delta t) - 2u(\mathbf{x}, t) + u(\mathbf{x}, t - \delta t)].$$

Introducimos una discretización temporal $t_n = n\delta t$, $n = 0, \dots, N$ en el intervalo de tiempo $[0, T]$ con paso δt . Denotando $u^n = u(\mathbf{x}, t_n) = u(\mathbf{x}, n\delta t)$ se re-escribe la aproximación de la segunda derivada de $u(\mathbf{x}, t)$ con respecto al tiempo como una recurrencia, como se observa en la aproximación (2.27)

$$u_{tt}^n \simeq \frac{u^{n+1} - 2u^n + u^{n-1}}{\delta t^2}. \quad (2.27)$$

El problema aproximado consiste en encontrar una secuencia de funciones u^{n+1} , $n = 1, \dots, N$, que satisfaga la discretización temporal de la expresión (2.16), es decir

$$\begin{aligned} \int_{R_\tau} \rho(\mathbf{x}) \frac{u^{n+1}(\mathbf{x}) - 2u^n(\mathbf{x}) + u^{n-1}(\mathbf{x})}{\delta t^2} \phi_i(\mathbf{x}) \, d\mathbf{x} + \int_{R_\tau} \chi(\mathbf{x}) \nabla u^n(\mathbf{x}) \cdot \nabla \phi_i(\mathbf{x}) \, d\mathbf{x} \\ - \int_{\partial R_\tau} \chi(\mathbf{x}) \frac{\partial u^n(\mathbf{x})}{\partial \mathbf{n}} \phi_i(\mathbf{x}) \, dS_{\mathbf{x}} - f(n\delta t) \int_{R_\tau} \rho(\mathbf{x}) g(\mathbf{x}) \phi_i(\mathbf{x}) \, d\mathbf{x} = 0. \end{aligned} \quad (2.28)$$

Conocidos $u^n(\mathbf{x})$ y $u^{n-1}(\mathbf{x})$, se puede aproximar $u^{n+1}(\mathbf{x})$ por el método de elementos finitos. Por tanto, dados u^0 y u^1 , podemos aproximar toda la recurrencia de niveles.

Si hacemos un desarrollo en serie de Taylor de $u(\mathbf{x}, \delta t)$ con respecto al tiempo y aplicamos las condiciones iniciales nulas se obtiene la expresión

$$u(\mathbf{x}, \delta t) = u(\mathbf{x}, 0) + \delta t u_t(\mathbf{x}, 0) + O(\delta t^2). \quad (2.29)$$

Por tanto, las condiciones iniciales para el esquema (2.28) son

$$u^0 = 0, \quad u^1 = u^0 + \delta t u_t^0 = 0. \quad (2.30)$$

Se puede recurrir a discretizaciones temporales de orden más alto, elegimos esta sencilla discretización por motivos de coste computacional. Los problemas de optimización y de muestreo que consideraremos en secciones posteriores requieren la resolución de un elevado número de problemas directos.

2.4.2. Discretización total de la formulación débil

Combinando la discretización en tiempo mediante diferencias finitas de la sección 2.4.1 y la discretización en espacio mediante elementos finitos de la sección 2.3, obtenemos la discretización total del problema, que conduce a la siguiente formulación.

Dado un espacio de elementos finitos contenido en $H^1(R_\tau)$ con base $\{\phi_1, \dots, \phi_B\}$ y una discretización temporal $t_n = n\delta t$, $n = 0, \dots, N$ del intervalo $[0, T]$, buscamos una aproximación a la solución de (2.16) de la forma $\mathbf{u}^n = \sum_{b=1}^B a_b^n \phi_b$, $n = 0, \dots, N$, donde los coeficientes se determinan a partir de la siguiente recurrencia,

$$\begin{aligned} \sum_{b=1}^B a_b^{n+1} \int_{R_\tau} \rho \phi_b \phi_j \, d\mathbf{x} &= - \sum_{b=1}^B (-2a_b^n + a_b^{n-1}) \int_{R_\tau} \rho \phi_b \phi_j \, d\mathbf{x} \\ &\quad - \delta t^2 \sum_{b=1}^B a_b^n \int_{R_\tau} \chi \nabla \phi_b \cdot \nabla \phi_j \, d\mathbf{x} + \delta t^2 f(n \delta t) \int_{R_\tau} \rho g \phi_j \, d\mathbf{x}, \end{aligned} \quad (2.31)$$

empezando con $a_b^0 = a_b^1 = 0$, $b = 1, \dots, B$. A partir de la ecuación (2.31), se definen dos matrices y un vector

$$\mathbf{M} = \left(\int_{R_\tau} \rho \phi_b \phi_j \, d\mathbf{x} \right)_{b,j}, \quad \mathbf{A} = \left(\int_{R_\tau} \chi \nabla \phi_b \cdot \nabla \phi_j \, d\mathbf{x} \right)_{b,j}, \quad \mathbf{g} = \left(\int_{R_\tau} \rho g \phi_j \, d\mathbf{x} \right)_j. \quad (2.32)$$

Por lo tanto, se trata de resolver de forma numérica los sistemas,

$$\mathbf{M} \mathbf{a}^{n+1} = \mathbf{M} (2\mathbf{a}^n - \mathbf{a}^{n-1}) - \delta t^2 \mathbf{A} \mathbf{a}^n + \delta t^2 f(n \delta t) \mathbf{g}, \quad (2.33)$$

donde \mathbf{a}^n es el vector columna de coeficientes que representa la aproximación numérica de la solución en el paso temporal n . La recurrencia se calcula a partir de los vectores \mathbf{a}^n y \mathbf{a}^{n-1} en cada paso temporal n , comenzando con $\mathbf{a}^0 = \mathbf{a}^1 = \mathbf{0}$. El problema (2.33) es la discretización total del problema de ondas truncado con condición de Neumann nula. Resolviendo (2.33), la solución del desplazamiento de la onda se calcula como una combinación lineal de las funciones de base de elementos finitos. Consideraremos espacios de elementos finitos P^1 de modo que las soluciones son continuas, están definidas en todo punto y podemos evaluar operadores de observación numéricos puntuales.

Para garantizar la estabilidad del proceso a medida que aumenta el tiempo de resolución, utilizamos un paso temporal δt que cumpla con la condición de Courant-Friedrichs-Lewy [35] para el sistema (2.33),

$$\delta t < \frac{\min(\delta x, \delta y)}{2} \min \left(\frac{1}{v_{p1}}, \dots, \frac{1}{v_{p\ell}}, \dots, \frac{1}{v_{pL}} \right), \quad (2.34)$$

donde $\min(\delta x, \delta y)$ es el mínimo paso de la malla espacial y $v_{p\ell}$, $\ell = 1, \dots, L$, son las velocidades de propagación de la onda en las distintas regiones conocidas que componen el medio en estudio, es decir, $v_{p\ell} = \sqrt{\chi_\ell / \rho_\ell}$, $\ell = 1, \dots, L$.

2.5. Condiciones de contorno no reflectantes

En las secciones anteriores hemos considerado el problema directo truncado a una región rectangular finita R_τ con condiciones de contorno de tipo Neumann nulas en todas sus paredes. Este problema es equivalente al problema original en un intervalo de tiempo $[0, \tau]$ con τ tal que las ondas emitidas por las fuentes no han alcanzado las paredes artificiales de R_τ . Si se desea que τ sea suficientemente grande (según la profundidad del terreno que se desea inspeccionar), nos vemos forzados a considerar recintos R_τ de dimensiones suficientemente grandes para evitar reflexiones en sus paredes artificiales, lo cual supone un alto coste computacional. Es posible considerar recintos finitos de menor tamaño imponiendo en las paredes artificiales condiciones de contorno no reflectantes, como se muestra en la expresión (2.35), que hacen que el problema planteado en un rectángulo sea equivalente al problema planteado en el semiespacio

$$\frac{\partial u}{\partial \mathbf{n}} = \frac{-1}{v_p} u_t. \quad (2.35)$$

La figura 2.2 ilustra la nueva geometría. La pared superior Σ donde se ubican emisores y receptores conserva las condiciones de Neumann nulas, mientras que en las demás $\partial R_\tau \setminus \Sigma$ recurrimos a condiciones artificiales de tipo no reflectante, véase [41].

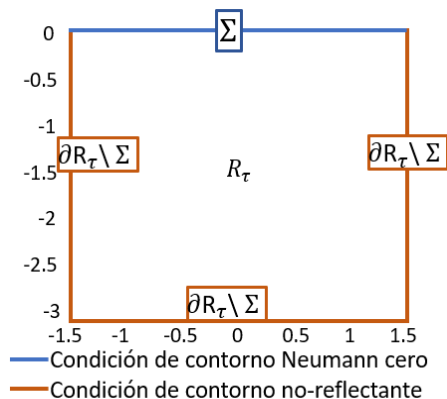


Figura 2.2: Condiciones de contorno de dominio computacional.

Este hecho introduce un nuevo término en la formulación débil del problema truncado. Revisando los cálculos de la sección 2.2.1 obtenemos para $\phi \in H^1(R_\tau)$

$$\begin{aligned} \int_{R_\tau} \rho(\mathbf{x}) \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} \phi(\mathbf{x}) d\mathbf{x} + \int_{R_\tau} \chi(\mathbf{x}) \nabla u(\mathbf{x}, t) \nabla \phi(\mathbf{x}) d\mathbf{x} - \int_{\partial R_\tau} \chi(\mathbf{x}) \nabla u(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \\ = f(t) \int_{R_\tau} \rho(\mathbf{x}) g(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}, \quad t > 0, \end{aligned} \quad (2.36)$$

e imponiendo la nueva condición de contorno (2.35) en la ecuación (2.16) obtenemos

$$\begin{aligned} \int_{R_\tau} \rho(\mathbf{x}) \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} \phi(\mathbf{x}) d\mathbf{x} + \int_{R_\tau} \chi(\mathbf{x}) \nabla u(\mathbf{x}, t) \nabla \phi(\mathbf{x}) d\mathbf{x} - \int_{\partial R_\tau} \chi(\mathbf{x}) \frac{\partial u(\mathbf{x}, t)}{\partial \mathbf{n}} \phi(\mathbf{x}) dS_{\mathbf{x}} \\ = f(t) \int_{R_\tau} \rho(\mathbf{x}) g(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}, \quad t > 0. \end{aligned} \quad (2.37)$$

Discretizamos en tiempo las segundas derivadas como se ha hecho en la sección 2.4 y las derivadas de orden uno mediante una fórmula regresiva $u_t(t) = \frac{u(t) - u(t-\delta t)}{\delta t}$. En la malla $t_n = n\delta t$, $n = 0, \dots, N$, tenemos que la secuencia $u^n(\mathbf{x}) = u(\mathbf{x}, t_n)$ satisface

$$\begin{aligned} & \int_{R_\tau} \rho \frac{u^{n+1} - 2u^n + u^{n-1}}{\delta t^2} \phi(\mathbf{x}) d\mathbf{x} + \int_{R_\tau} \chi \nabla u^n \cdot \nabla \phi(\mathbf{x}) d\mathbf{x} \\ & + \int_{\partial R_\tau \setminus \Sigma} \frac{\chi}{v_p} \frac{u^n - u^{n-1}}{\delta t} \phi(\mathbf{x}) dS_{\mathbf{x}} - f(n\delta t) \int_{R_\tau} \rho g \phi(\mathbf{x}) d\mathbf{x} = 0, \end{aligned} \quad (2.38)$$

partiendo de $u^0 = u^1 = 0$.

Empleando la misma base de elementos finitos $\{\phi_1, \dots, \phi_B\}$ en espacio que se introdujo en la sección 2.4.2, la discretización total del problema conduce a buscar la solución de la forma $\mathbf{u}^n = \sum_{b=1}^B a_b^n \phi_b$, $n = 0, \dots, N$, donde se obtienen los coeficientes a través de la resolución del sistema matricial,

$$\begin{aligned} \sum_{b=1}^B a_b^{n+1} \int_{R_\tau} \rho \phi_b \phi_j d\mathbf{x} = & - \sum_{b=1}^B (-2a_b^n + a_b^{n-1}) \int_{R_\tau} \rho \phi_b \phi_j d\mathbf{x} - \delta t^2 \sum_{b=1}^B a_b^n \int_{R_\tau} \chi(\mathbf{x}) \nabla \phi_b \cdot \nabla \phi_j d\mathbf{x} \\ & - \delta t \sum_{b=1}^B (a_b^n - a_b^{n-1}) \int_{\partial R_\tau} \frac{\chi(\mathbf{x})}{v_p} \phi_b \phi_j dS_{\mathbf{x}} + \delta t^2 f(n\delta t) \int_{R_\tau} \rho g \phi_j d\mathbf{x}, \end{aligned} \quad (2.39)$$

empezando con $a_b^0 = a_b^1 = 0$, $b = 1, \dots, B$. Podemos reescribirlo en forma compacta definiendo las matrices:

$$\mathbf{M} = \left(\int_{R_\tau} \rho \phi_b \phi_j d\mathbf{x} \right)_{b,j}, \quad \mathbf{A} = \left(\int_{R_\tau} \chi(\mathbf{x}) \nabla \phi_b \cdot \nabla \phi_j d\mathbf{x} \right)_{b,j}, \quad \mathbf{B} = \left(\int_{\partial R_\tau} \frac{\chi(\mathbf{x})}{v_p} \phi_b \phi_j dS_{\mathbf{x}} \right)_{b,j} \quad (2.40)$$

y el vector $\mathbf{g} = \left(\int_{R_\tau} \rho g \phi_j d\mathbf{x} \right)_j$. La recurrencia final es

$$\mathbf{M}\mathbf{a}^{n+1} = \mathbf{M}(2\mathbf{a}^n - \mathbf{a}^{n-1}) - \delta t^2 \mathbf{A}\mathbf{a}^n - \delta t \mathbf{B}(\mathbf{a}^n - \mathbf{a}^{n-1}) + \delta t^2 f(n\delta t)\mathbf{g}, \quad (2.41)$$

resultando en un sistema que se puede resolver de forma iterativa, con condiciones iniciales $\mathbf{a}^0 = \mathbf{a}^1 = 0$.

Capítulo 3

Estudio teórico y numérico del problema inverso en dimensión finita

La resolución de problemas inversos, en el contexto de la detección de anomalías, se ha consolidado como una herramienta indispensable en una variedad de disciplinas científicas y tecnológicas. Este enfoque aborda el desafío de inferir propiedades desconocidas de una anomalía a partir de observaciones externas, estableciendo así un puente entre el mundo medible y el dominio subyacente del fenómeno estudiado [60].

El problema inverso encuentra su aplicación en campos tan diversos como la geofísica, la medicina y la ingeniería, entre otros [26]. A medida que las tecnologías de observación y medición avanzan, surge la necesidad crítica de desarrollar métodos efectivos para interpretar y comprender la información recopilada. En este contexto, la formulación inversa se presenta como un marco teórico y computacional poderoso y versátil.

La formulación inversa se basa en la idea fundamental de modelar matemáticamente el comportamiento del sistema a través de ecuaciones directas, que relacionan las variables internas desconocidas con las observaciones externas. Este modelado permite invertir el proceso y estimar los parámetros desconocidos a partir de las observaciones. Sin embargo, esta tarea se complica por la presencia de incertidumbre, ruido y la no linealidad inherente a muchos problemas reales [36].

El desarrollo de métodos para identificar anomalías presentes en un medio requiere la elección de una descripción matemática de las mismas [22]. Una posibilidad consiste en recurrir a inversión basada en objetos caracterizados por un bajo número de parámetros que representan sus propiedades materiales, su geometría y posición. Este punto de vista conduce a formulaciones paramétricas en dimensión finita. Alternativamente, se puede optar por caracterizar las variaciones espaciales de los campos que representan las propiedades materiales del medio en el que están inmersos. Esto conduce a problemas en dimensión infinita, que eventualmente se pueden aproximar por problemas paramétricos en alta dimensión [34]. En este capítulo, nos centramos en formulaciones paramétricas en dimensión finita.

Tradicionalmente, se han enfocado los problemas inversos desde un punto de vista determinista. Dados datos medidos, se buscan parámetros o campos para los cuales el operador de observación asociado a un problema directo de propagación de ondas genere valores próximos a esos datos. Este tipo de problemas está severamente mal puesto. Para datos arbitrarios pueden carecer de solución o presentar más de una, lo que podría dar lugar a soluciones que no dependen continuamente de los datos [32]. Este tipo de problemas se suelen regularizar mediante formulaciones en las que se busca optimizar un funcional de costo que incluye términos que comparan los datos reales y los observados para anomalías propuestas, así como regularizaciones de tipo variación total [40], Tikhonov [90] o iteraciones de Gauss-Newton [55]. Este último método puede en cierto grado proporcionar soluciones adaptadas con un cierto nivel de ruido [55].

Actualmente, se está poniendo de manifiesto la necesidad de obtener soluciones con incertidumbre cuantificada en función del nivel de ruido estimado en los datos medidos. Un enfoque popular

lo proporciona el marco de inferencia estadística [46]. La formulación inversa bayesiana permite incorporar información previa y cuantificar la incertidumbre de cada parámetro de la anomalía de manera explícita.

Como se ha estudiado en el capítulo anterior, el problema directo busca la solución del modelo matemático que gobierna la dinámica de las ondas que se emiten a través de una red de fuentes en el medio y, tras interactuar con éste y la anomalía presente, se mide la respuesta de la onda a través de una red de receptores. En nuestro caso específico, se busca la solución de la ecuación de ondas $u(\mathbf{x}, t)$ dadas las características del medio y los parámetros $\boldsymbol{\nu}$ de la anomalía presente en el dominio. Nos referimos a $\boldsymbol{\nu}$ como el conjunto de parámetros que caracterizan una anomalía en el medio, es decir, una región donde las propiedades físicas — como la densidad, el módulo de elasticidad y la geometría— difieren respecto al material de fondo. En el problema inverso, lo que buscamos es recuperar dichos parámetros $\boldsymbol{\nu}$, dados los datos observados $\mathbf{f}(\boldsymbol{\nu})$ y los datos sintéticos con ruido gaussiano añadido \mathbf{d}^{obs} .

Para abordar el problema inverso, se debe disponer de una caracterización matemática del operador de observación $\mathbf{f}(\boldsymbol{\nu})$ a través de un modelo, el cual desempeña un papel fundamental al establecer la conexión matemática entre los parámetros de la anomalía $\boldsymbol{\nu}$ y las observaciones reales \mathbf{d}^{obs} . En ausencia de ruido, los datos medidos satisfacerían $\mathbf{d}^{\text{obs}}(\boldsymbol{\nu}_{\text{verdadero}}) = \mathbf{f}(\boldsymbol{\nu}_{\text{verdadero}})$ para anomalías definidas por parámetros $\boldsymbol{\nu}_{\text{verdadero}}$. En presencia de ruido, tenemos la relación matemática:

$$\mathbf{d}^{\text{obs}}(\boldsymbol{\nu}_{\text{verdadero}}) = \mathbf{f}(\boldsymbol{\nu}_{\text{verdadero}}) + \boldsymbol{\eta}, \quad (3.1)$$

donde $\boldsymbol{\eta}$ corresponde al ruido añadido como una distribución normal aleatoria de media cero y covarianza $\mathbf{C} = \sigma^2 \mathbf{I}$, siendo σ la desviación estándar máxima del ruido y \mathbf{I} una matriz identidad $P \times P$, donde P es el número de parámetros desconocidos. Es decir, el ruido añadido es posible escribirlo como $\boldsymbol{\eta} = \mathcal{N}(0, \mathbf{C})$. Con la adición de esta componente estocástica se capturan las imperfecciones y la incertidumbre en el proceso de medición.

Tras estos preliminares, formulamos matemáticamente el problema inverso determinista y bayesiano en dimensión finita.

3.1. Problema inverso determinista en dimensión finita

Consideramos anomalías caracterizadas por una colección finita de parámetros $\boldsymbol{\nu} \in \mathbb{R}^P$, donde $\boldsymbol{\nu}$ contiene propiedades de la anomalía como su posición central (x, y) , parámetros geométricos, propiedades materiales como la densidad representada por $\rho(\mathbf{x})$ y la velocidad de propagación de la onda representada por $v_p(\mathbf{x})$. El problema inverso consiste en encontrar los parámetros $\boldsymbol{\nu}$ que, dados las mediciones del operador observador \mathbf{d}^{obs} , hagan que $\mathbf{f}(\boldsymbol{\nu}_{\text{verdadero}}) = \mathbf{d}^{\text{obs}}$.

La formulación matemática del problema inverso es crítica para el éxito del proceso de optimización, especialmente en el contexto del método inverso determinista. Aquí, la función de costo $J(\boldsymbol{\nu})$ se convierte en una herramienta fundamental, ya que mide la discrepancia entre el valor del operador de observación correspondiente $\mathbf{f}(\boldsymbol{\nu})$ y los datos observados \mathbf{d}^{obs} , es decir

$$J(\boldsymbol{\nu}) := \frac{1}{2} \|\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}}\|^2. \quad (3.2)$$

Se trata entonces de encontrar $\boldsymbol{\nu}_{\text{min}}$ que minimiza la función de costo $J(\boldsymbol{\nu})$, es decir

$$\boldsymbol{\nu}_{\text{min}} = \operatorname{argmin} J(\boldsymbol{\nu}) = \operatorname{argmin} \frac{1}{2} \|\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}}\|_W^2, \quad (3.3)$$

donde $\|\cdot\|_W$ denota la norma Euclídea ponderada por la matriz de pesos W , y está definida como $\|\mathbf{x}\|_W^2 = \mathbf{x}^T W \mathbf{x}$, donde T indica la trasposición. En muchos casos, W corresponde a la inversa de la matriz de covarianza del ruido en los datos observados [1, 24].

El problema directo constituye una restricción en el problema de optimización asociado al problema inverso. Si los datos \mathbf{d}^{obs} no tuvieran ruido Gaussiano añadido, la solución exacta sería un mínimo global $J(\boldsymbol{\nu}_{\text{verdadero}}) = 0$. En caso de plantear el problema en dimensión infinita, las variables a optimizar serían funciones definidas en espacios de dimensión infinita, en lugar de parámetros [28], como se está planteando en este capítulo.

Un mínimo $\boldsymbol{\nu}_{\text{min}}$ de la ecuación (3.3) se puede aproximar a través de métodos iterativos como el método iterativo de tipo Fletcher-Levenberg-Marquardt [69]. Se comienza con valores iniciales $\boldsymbol{\nu}^0$ del parámetro, los cuales se actualizan en cada iteración a través de la ecuación (3.4)

$$\boldsymbol{\nu}^{n+1} = \boldsymbol{\nu}^n + \boldsymbol{\xi}^n, \quad (3.4)$$

donde a cada uno de los valores obtenidos $\boldsymbol{\nu}^n$ se añade la corrección $\boldsymbol{\xi}^n$ solución de la ecuación (3.5)

$$\left(\mathbf{H}(\boldsymbol{\nu}^n) + \omega_n \text{diag}(\mathbf{H}(\boldsymbol{\nu}^n)) \right) \boldsymbol{\xi}^n = -\mathbf{g}(\boldsymbol{\nu}^n), \quad (3.5)$$

Cada término de la ecuación (3.5) se detalla a continuación:

- $\omega_n > 0$ es un pequeño parámetro que se ajusta en cada iteración según el funcional crezca (se aumenta) o decrezca (se disminuye).
- $\mathbf{g}(\boldsymbol{\nu}^n)$ es el gradiente de la función de costo con respecto al vector $\boldsymbol{\nu}$: $\mathbf{g}(\boldsymbol{\nu}) = \left(\frac{\partial J(\boldsymbol{\nu})}{\partial \nu_l} \right)_{l=1, \dots, P}$.
- $\mathbf{H}(\boldsymbol{\nu}^n)$ es la matriz Hessiana o la segunda derivada de la función de costo con respecto al vector $\boldsymbol{\nu}$: $\mathbf{H}(\boldsymbol{\nu}) = \left(\frac{\partial^2 J(\boldsymbol{\nu})}{\partial \nu_l \partial \nu_m} \right)_{l, m=1, \dots, P}$.

Las iteraciones del algoritmo se detendrán cuando la norma relativa de la función de costo $J(\boldsymbol{\nu})$ sea más pequeña que el valor de la tolerancia configurada.

La idea tras este tipo de algoritmos se puede fundamentar de la siguiente forma: si se considera una situación en la que el descenso de gradiente $g(\boldsymbol{\nu}^n)$ se acerca a una especie de valle o mínimo, entonces habrá una región donde los valores del gradiente son pequeños. En este escenario, el método del gradiente podría oscilar sobre o bajo el punto donde se produce el valle por las pendientes, o girar muy lentamente en dirección al valle. Sin embargo, la información adicional sobre la curvatura proporcionada por la Hessiana obligaría al descenso del gradiente a orientarse rápidamente hacia el valle, lo que resultaría en una convergencia más rápida hacia el mínimo de la función de costo.

Las funciones de costo asociadas a problemas inversos imponen una restricción que se asemeja a un problema de valores iniciales y de contorno para las ecuaciones de onda, las cuales dependen de los parámetros $\boldsymbol{\nu}$. Estas soluciones son fundamentales para el cálculo de la función de costo.

Para funciones de costo $J(\boldsymbol{\nu})$ definidas en términos de operadores de observación \mathbf{d}^{obs} asociados a las mediciones obtenidas a través del problema directo $\mathbf{f}(\boldsymbol{\nu})$, el gradiente de la función de costo involucra las primeras derivadas de $\mathbf{f}(\boldsymbol{\nu})$ con respecto a $\boldsymbol{\nu}$, y la Hessiana involucra las segundas derivadas de $\mathbf{f}(\boldsymbol{\nu})$ con respecto a $\boldsymbol{\nu}$. Es decir, si desarrollamos la primera derivada de la función de costo con respecto a $\boldsymbol{\nu}$, se obtiene

$$\frac{\partial J(\boldsymbol{\nu}^n)}{\partial \nu_l^n} \approx \frac{1}{\sigma_{\text{ruido}}^2} \sum_{j=1}^J (f(\boldsymbol{\nu}^n) - \mathbf{d}^{\text{ruido}})_j \frac{\partial f(\boldsymbol{\nu}^n)_j}{\partial \nu_l^n}, \quad (3.6)$$

donde ν_l^n es el parámetro l de la anomalía para la iteración n , σ_{ruido}^2 representa la varianza del ruido de los datos observados y $\frac{\partial f(\boldsymbol{\nu}^n)_j}{\partial \nu_l^n}$ es una aproximación por diferencias finitas de primer orden.

Dado que únicamente nos interesa este proceso iterativo para proponer direcciones de descenso y no nos interesan las derivadas de $\mathbf{f}(\boldsymbol{\nu})$ en sí, se puede recurrir a aproximar las derivadas de $\mathbf{f}(\boldsymbol{\nu})$ mediante diferencias finitas

$$\frac{\partial \mathbf{f}(\boldsymbol{\nu}^n)}{\partial \nu_l^n} \approx \frac{\mathbf{f}(\boldsymbol{\nu}^n + \epsilon \delta_l) - \mathbf{f}(\boldsymbol{\nu}^n)}{\epsilon}, \quad (3.7)$$

donde ϵ es un número pequeño que se puede ajustar en cada caso para tener una mejor aproximación de cada derivada parcial y δ_l es el vector director unitario en la componente ν_l^n , que vale uno en l y cero en el resto, véase [24] para detalles.

La Hessiana la aproximaremos de la siguiente forma:

$$\frac{\partial^2 J(\boldsymbol{\nu}^n)}{\partial \nu_l^n \partial \nu_m^n} \approx \frac{1}{\sigma_{\text{ruido}}^2} \sum_{j=1}^J \frac{\partial f(\boldsymbol{\nu}^n)_j}{\partial \nu_m^n} \frac{\partial f(\boldsymbol{\nu}^n)_j}{\partial \nu_l^n}, \quad (3.8)$$

con $m, l = 1, \dots, P$, donde P es el número total de parámetros desconocidos de la anomalía. Esta aproximación de tipo Gauss-Newton desprecia derivadas de segundo orden del operador de observación.

En general, la función de costo $J(\boldsymbol{\nu})$ puede presentar mínimos locales adicionales, por lo que es preciso regularizarlo introduciendo el término $R(\boldsymbol{\nu})$, es decir

$$J_{\text{reg}}(\boldsymbol{\nu}) := J(\boldsymbol{\nu}) + R(\boldsymbol{\nu}), \quad (3.9)$$

donde $R(\boldsymbol{\nu})$ es un término con información adicional que fuerza la selección de un mínimo global. Elecciones típicas se basan en términos de Tikhonov o variación total [90]. Es posible modificar estos planteamientos para considerar niveles de ruido en los datos mediante estrategias de iteraciones regularizantes tipo Gauss-Newton y el uso de criterios de parada basados en discrepancia [55]. Otra opción que veremos más adelante son las regularizaciones bayesianas [24].

De la ecuación (3.9), el primer término cuantifica el ajuste de los datos observados con respecto a los datos medidos y el segundo término $R(\boldsymbol{\nu})$ representa la regularización. Las regularizaciones de Tikhonov son ampliamente empleadas [90] y consideran $R(\boldsymbol{\nu}) = \frac{\lambda}{2} \|\boldsymbol{\nu}\|^2$ para alguna norma $\|\cdot\|$ convenientemente elegida y un factor $\lambda > 0$, el cual es el parámetro de regularización. Como veremos en la sección siguiente, formulaciones bayesianas motivan la consideración de funcionales de costo con términos regularizantes de la forma $R(\boldsymbol{\nu}) = \frac{\lambda}{2} (\boldsymbol{\nu} - \boldsymbol{\nu}_0)^t \boldsymbol{\Gamma}_{\text{prior}}^{-1} (\boldsymbol{\nu} - \boldsymbol{\nu}_0)$, donde $\boldsymbol{\nu}_0$ tiene la interpretación de información inicial o a priori sobre los parámetros buscados y $\boldsymbol{\Gamma}_{\text{prior}}$ es una matriz simétrica definida positiva que representa la covarianza, y engloba información sobre la desviación típica esperada sobre cada parámetro buscado de $\boldsymbol{\nu}_0$. La formulación bayesiana del problema inverso nos permite cuantificar la incertidumbre en los óptimos resultantes.

3.2. Problema inverso bayesiano en dimensión finita

En la formulación bayesiana del problema inverso, se especifica un modelo de probabilidad que incorpora conocimiento previo sobre los parámetros desconocidos $\boldsymbol{\nu}_0$. La inferencia bayesiana combina la información de los datos a través de una función de verosimilitud con la distribución a priori $p_{\text{prior}}(\boldsymbol{\nu}_0)$. Esto implica que el conocimiento del parámetro investigado se actualiza con los datos mediante la aplicación del teorema de Bayes, que utiliza toda la información disponible de $\boldsymbol{\nu}_0$ [17].

En el contexto bayesiano, el problema inverso se formula mediante: 1) la información a priori $p_{\text{prior}}(\boldsymbol{\nu}_0)$, que incluye toda la información inicial sobre el parámetro en estudio $\boldsymbol{\nu}$, 2) la probabilidad condicionada al operador observable $p(\mathbf{d}^{\text{obs}}|\boldsymbol{\nu})$, que describe la probabilidad de que $\boldsymbol{\nu}$ resulte en la medición \mathbf{d}^{obs} , y 3) la probabilidad a posteriori $p(\boldsymbol{\nu}|\mathbf{d}^{\text{obs}})$ dada por la fórmula de Bayes [93]:

$$p_{\text{posterior}}(\boldsymbol{\nu}|\mathbf{d}^{\text{obs}}) = \frac{p(\mathbf{d}^{\text{obs}}|\boldsymbol{\nu}) p_{\text{prior}}(\boldsymbol{\nu})}{p(\mathbf{d}^{\text{obs}})}, \quad (3.10)$$

donde $p(\mathbf{d}^{\text{obs}})$ actúa como un factor de normalización y no depende del parámetro desconocido $\boldsymbol{\nu}$. La solución del problema inverso bayesiano es la probabilidad a posteriori $p(\boldsymbol{\nu}|\mathbf{d}^{\text{obs}})$ y el objetivo es caracterizar su estructura. La probabilidad a posteriori $p(\boldsymbol{\nu}|\mathbf{d}^{\text{obs}})$ puede tener uno o varios máximos locales y máximos globales, dependiendo de la complejidad del problema y la distribución de los datos observados. La estructura global de la probabilidad a posteriori está determinada por la combinación de la probabilidad a priori y la información proporcionada por los datos observados. Puede haber regiones de alta probabilidad que corresponden a valores plausibles de los parámetros, así como regiones de baja probabilidad que son menos probables dadas las observaciones y la información a priori [46].

3.2.1. Aproximación de máximo a posteriori

Para caracterizar la probabilidad a posteriori, se calcula la aproximación de máximo a posteriori (MAP), que corresponde al valor de los parámetros $\boldsymbol{\nu}_{\text{MAP}}$ donde la probabilidad a posteriori es máxima.

En la práctica, las distribuciones a priori y de verosimilitud suelen ser modeladas con distribuciones tipo Gaussiano, como se detalla a continuación:

- **Probabilidad a priori:**

$$p_{\text{prior}}(\boldsymbol{\nu}) = \frac{1}{\sqrt{(2\pi)^{P/2} \det(\boldsymbol{\Gamma}_{\text{prior}})}} \exp\left(-\frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)^T \boldsymbol{\Gamma}_{\text{prior}}^{-1} (\boldsymbol{\nu} - \boldsymbol{\nu}_0)\right), \quad (3.11)$$

donde $\boldsymbol{\nu}_0$ es la media que representa la información a priori sobre los parámetros y $\boldsymbol{\Gamma}_{\text{prior}}$ es la matriz de covarianza a priori.

- **Probabilidad de verosimilitud:**

$$p(\mathbf{d}^{\text{obs}}|\boldsymbol{\nu}) = \frac{1}{\sqrt{(2\pi)^{P/2} \det(\boldsymbol{\Gamma}_{\text{ruido}})}} \exp\left(-\frac{1}{2}(\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}})^T \boldsymbol{\Gamma}_{\text{ruido}}^{-1} (\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}})\right), \quad (3.12)$$

donde $\mathbf{f}(\boldsymbol{\nu})$ es el operador de observación, \mathbf{d}^{obs} son los datos observados y $\boldsymbol{\Gamma}_{\text{ruido}}$ es la matriz de covarianza del ruido en los datos observados.

Sustituyendo las expresiones de $p_{\text{prior}}(\boldsymbol{\nu})$ y $p(\mathbf{d}^{\text{obs}}|\boldsymbol{\nu})$ en la ecuación (3.10), obtenemos la distribución de la probabilidad a posteriori

$$p_{\text{posterior}}(\boldsymbol{\nu}|\mathbf{d}^{\text{obs}}) \propto \exp\left(-\frac{1}{2}\left(\|\boldsymbol{\nu} - \boldsymbol{\nu}_0\|_{\boldsymbol{\Gamma}_{\text{prior}}}^2 + \|\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}}\|_{\boldsymbol{\Gamma}_{\text{ruido}}}^2\right)\right). \quad (3.13)$$

Esta distribución a posteriori puede aproximarse por $e^{-J_{\text{reg}}(\boldsymbol{\nu})}$, en términos de la función de costo

$$J_{\text{reg}}(\boldsymbol{\nu}) = \frac{1}{2}\|\boldsymbol{\nu} - \boldsymbol{\nu}_0\|_{\boldsymbol{\Gamma}_{\text{prior}}}^2 + \frac{1}{2}\|\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}}\|_{\boldsymbol{\Gamma}_{\text{ruido}}}^2, \quad (3.14)$$

donde $\|\mathbf{v}\|_{\mathbf{\Gamma}} = \mathbf{v}^T \mathbf{\Gamma}^{-1} \mathbf{v}$ y representa la norma cuadrática asociada a la matriz de covarianza $\mathbf{\Gamma}$. Optimizando esta función de costo utilizando técnicas descritas en la sección 3.1, se pueden encontrar mínimos locales y globales de $\boldsymbol{\nu}_{\text{MAP}}$, iniciando desde diferentes valores iniciales. Si la información a priori dada por $\boldsymbol{\nu}_0$ y $\mathbf{\Gamma}_{\text{prior}}$ es adecuada, se espera que la función de costo $J(\boldsymbol{\nu})$ sea convexa y tenga un único mínimo global, gracias al efecto regularizador del término $\|\boldsymbol{\nu} - \boldsymbol{\nu}_0\|_{\mathbf{\Gamma}_{\text{prior}}}$.

La ecuación (3.14) es equivalente a la ecuación (3.9) en el marco bayesiano, siendo el término regularizador $R(\boldsymbol{\nu}) = \frac{1}{2} \|\boldsymbol{\nu} - \boldsymbol{\nu}_0\|_{\mathbf{\Gamma}_{\text{prior}}}^2$ y $J(\boldsymbol{\nu}) = \frac{1}{2} \|\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}}\|_{\mathbf{\Gamma}_{\text{ruido}}}^2$.

Para minimizar la función de costo $J_{\text{reg}}(\boldsymbol{\nu})$ se utiliza un algoritmo Fletcher-Levenberg-Marquardt, donde los valores de $\boldsymbol{\nu}$ comienzan de un valor inicial y en cada iteración n se le añade un término $\boldsymbol{\xi}^n$, es decir

$$\boldsymbol{\nu}^{n+1} = \boldsymbol{\nu}^n + \boldsymbol{\xi}_{\text{reg}}^n, \quad (3.15)$$

siendo $\boldsymbol{\xi}_{\text{reg}}^n$ la solución de la ecuación

$$\left(\mathbf{H}_{\text{reg}}(\boldsymbol{\nu}^n) + \omega_n \text{diag}(\mathbf{H}_{\text{reg}}(\boldsymbol{\nu}^n)) \right) \boldsymbol{\xi}_{\text{reg}}^n = -\mathbf{g}_{\text{reg}}(\boldsymbol{\nu}^n), \quad (3.16)$$

siendo $\boldsymbol{\nu}^n$ los parámetro de la anomalía para la iteración n . La ecuación (3.16) es equivalente a la ecuación (3.5) del método determinista, con la diferencia que la función de costo utilizada para el cálculo es $J_{\text{reg}}(\boldsymbol{\nu})$ con el término regularizador añadido. En la ecuación (3.16) tenemos:

- El gradiente $\mathbf{g}_{\text{reg}}(\boldsymbol{\nu}^n)$ de la función de costo $J_{\text{reg}}(\boldsymbol{\nu}^n)$:

$$\mathbf{g}_{\text{reg}}(\boldsymbol{\nu}^n) = \left(\frac{\partial J_{\text{reg}}(\boldsymbol{\nu}^n)}{\partial \nu_p^n} \right)_{p=1}^P \approx \left(\frac{1}{\sigma_{\text{ruido}}^2} \sum_{j=1}^J (f(\boldsymbol{\nu}^n) - \mathbf{d}^{\text{obs}})_j \frac{\partial f(\boldsymbol{\nu}^n)_j}{\partial \nu_p^n} + \left[\mathbf{\Gamma}_{\text{prior}}^{-1} (\boldsymbol{\nu}^n - \boldsymbol{\nu}_0) \right]_p \right)_{p=1}^P \quad (3.17)$$

donde, al igual que en el problema inverso determinista tenemos $\mathbf{f}(\boldsymbol{\nu}^n)$ y \mathbf{d}^{obs} como las mediciones obtenidas del problema directo y los datos medidos respectivamente. La derivada $\frac{\partial f(\boldsymbol{\nu}^n)}{\partial \nu_p^n}$ se puede calcular usando la ecuación (3.7), la cual es una aproximación de primer orden mediante diferencias finitas.

- La matriz hessiana $\mathbf{H}_{\text{reg}}(\boldsymbol{\nu}^n)$ de la función costo $J_{\text{reg}}(\boldsymbol{\nu}^n)$:

$$\mathbf{H}_{\text{reg}}(\boldsymbol{\nu}^n) = \left(\frac{\partial^2 J_{\text{reg}}(\boldsymbol{\nu}^n)}{\partial \nu_p^n \partial \nu_m^n} \right)_{p,m=1}^P \approx \left(\frac{1}{\sigma_{\text{ruido}}^2} \sum_{j=1}^J \frac{\partial f(\boldsymbol{\nu}^n)_j}{\partial \nu_p^n} \frac{\partial f(\boldsymbol{\nu}^n)_j}{\partial \nu_m^n} + \left[\mathbf{\Gamma}_{\text{prior}}^{-1} \right]_{p,m} \right)_{p,m=1}^P, \quad (3.18)$$

con $m, p = 1, \dots, P$, siendo P el número total de parámetros desconocidos de la anomalía. $\mathbf{H}_{\text{reg}}(\boldsymbol{\nu}^n)$ es una aproximación de tipo Gauss-Newton, donde se desprecian los términos de segundo orden.

3.2.2. Pseudocódigo para cálculo de valor máximo a posteriori

A continuación se muestran los pasos principales del código, tomados de [24].

1. Se define el número máximo de iteraciones externas del algoritmo, **ITMAX**, y un número máximo de iteraciones internas, **ITMAX2**.
2. Se inicializan los valores del parámetro $\boldsymbol{\nu}_0$, la covarianza $\mathbf{\Gamma}_{\text{prior}}$, así como el parámetro de iteración ω_0 .
3. Se introducen el operador de observación del problema directo $\mathbf{f}(\boldsymbol{\nu})$ y los datos observados \mathbf{d}^{obs} .

4. Se fija el paso temporal δt y el tiempo de simulación total T . Además, se fijan los pasos espaciales δx y δy .

5. Se resuelve el problema de optimización:

a) Mientras que $n \leq \text{ITMAX}$:

1) Calculamos $J_{\text{reg}}(\boldsymbol{\nu}^n)$ de acuerdo a la ecuación (3.14).

2) Mientras que $l \leq \text{ITMAX2}$:

a' Partimos de $\boldsymbol{\nu}^n$ para $l = 1$.

b' Aproximamos la matriz Hessiana con la ecuación (3.18), y el gradiente de la función costo con la ecuación (3.17).

c' Calculamos $\boldsymbol{\xi}^l$ de acuerdo a la ecuación (3.16) y $\boldsymbol{\nu}^{l+1}$ de acuerdo a la ecuación (3.15).

d' Calculamos $J_{\text{reg}}(\boldsymbol{\nu}^{l+1})$ de acuerdo a (3.14).

e' Calculamos:

$$\text{dif} = \frac{|J_{\text{reg}}(\boldsymbol{\nu}^l) - J_{\text{reg}}(\boldsymbol{\nu}^{l+1})|}{|J_0(\boldsymbol{\nu}^n)|}.$$

Si $\text{dif} < 10^{-6}$, entonces hemos acabado esta iteración interna y ponemos $\boldsymbol{\nu}^{n+1} = \boldsymbol{\nu}^l$.

f' Si $J_{\text{reg}}(\boldsymbol{\nu}^{n+1}) < J_{\text{reg}}(\boldsymbol{\nu}^n)$, entonces $\omega_{n+1} = \omega_n/2$ y volvemos al paso a). En caso contrario, si $J_{\text{reg}}(\boldsymbol{\nu}^{n+1}) \geq J_{\text{reg}}(\boldsymbol{\nu}^n)$, entonces $\omega_n = 2\omega_n$, y volvemos al paso a'.

Ocasionalmente, puede ser necesario introducir un parámetro auxiliar $\lambda > 0$ que multiplique a $\boldsymbol{\Gamma}_{\text{prior}}^{-1}$ en J_{reg} para balancear los dos términos iniciales del coste. Partiendo de λ_0 , este parámetro se actualiza a un valor λ_n en cada iteración incrementándose de forma progresiva hasta alcanzar el valor 1, sin que el proceso se pare antes de que esto ocurra, véase [24, 26] para distintas elecciones.

Esta estrategia gradual, contribuye a mejorar la estabilidad del algoritmo y a evitar oscilaciones o convergencia prematura. El objetivo es controlar de forma adaptativa la influencia del término de regularización: inicialmente se otorga mayor flexibilidad para ajustarse a los datos, mientras que, con el avance de las iteraciones, se refuerza progresivamente la contribución de la información a priori. Esta estrategia incrementa la robustez del algoritmo, permitiendo una transición gradual desde una solución guiada por los datos hacia una más estabilizada, mitigando así riesgos de sobreajuste o estancamiento prematuro.

3.3. Cuantificación de incertidumbre mediante la aproximación de Laplace

Una vez que se obtiene la aproximación a posteriori $\boldsymbol{\nu}_{\text{MAP}}$, es posible aproximar la probabilidad a posteriori $p_{\text{posterior}}(\boldsymbol{\nu})$ a través del método de Laplace. Este método linealiza en torno a $\boldsymbol{\nu}_{\text{MAP}}$ y aproxima la probabilidad a posteriori por una distribución normal multivariada de la forma,

$$\mathcal{N}(\boldsymbol{\nu}_{\text{MAP}}, \boldsymbol{\Gamma}_{\text{posterior}}), \quad \boldsymbol{\Gamma}_{\text{posterior}} = (\mathbf{F}^t \boldsymbol{\Gamma}_{\text{ruido}}^{-1} \mathbf{F} + \boldsymbol{\Gamma}_{\text{prior}}^{-1})^{-1}, \quad (3.19)$$

donde \mathbf{F} es la matriz jacobiana de las observaciones realizadas $\mathbf{f}(\boldsymbol{\nu})$ con respecto a los parámetros $\boldsymbol{\nu}$, evaluada en $\boldsymbol{\nu}_{\text{MAP}}$, y representa en dimensión finita la derivada de Fréchet del operador de observación [94].

Una vez calculada la matriz de covarianza $\boldsymbol{\Gamma}_{\text{posterior}}$, es posible muestrear la aproximación de la distribución a posteriori $\mathcal{N}(\boldsymbol{\nu}_{\text{MAP}}, \boldsymbol{\Gamma}_{\text{posterior}})$ de acuerdo a la relación que se muestra en la ecuación (3.20) a continuación:

$$\boldsymbol{\nu} = \boldsymbol{\nu}_{\text{MAP}} + \boldsymbol{\Gamma}_{\text{posterior}}^{1/2} \mathbf{n}, \quad (3.20)$$

siendo \mathbf{n} un vector de valores aleatorios, independientes y distribuidos de forma idéntica pertenecientes a una distribución $\mathcal{N}(0, 1)$ [94].

Cuando la distribución posterior de un parámetro desconocido, representado por $\boldsymbol{\nu}$, se concentra alrededor de un valor específico, indica que contamos con una gran cantidad de datos y que nuestra incertidumbre sobre el parámetro que intentamos inferir es menor. Esto es beneficioso, ya que nos brinda una mayor confianza en las estimaciones de los parámetros $\boldsymbol{\nu}_{\text{MAP}}$ [46].

Sin embargo, esta alta concentración de la distribución puede plantear desafíos numéricos. A medida que la distribución del parámetro se vuelve más concentrada, los métodos estándar de integración pueden volverse menos efectivos, especialmente cuando el tamaño del conjunto de datos es muy grande [82]. Por lo tanto, es crucial comprender cómo se comportan los métodos de muestreo en situaciones de alta concentración de muestras.

Los métodos numéricos a menudo se basan en la medida a priori, ya que es más simple y permite un muestreo directo. Sin embargo, en conjuntos de datos grandes, la mayoría de los puntos de la muestra pueden estar en regiones donde su contribución a la distribución posterior es mínima. Esto puede llevar a pasar por alto áreas donde la probabilidad posterior es más alta o donde se minimiza la función de costo. Para abordar este problema, es útil utilizar métodos numéricos que se integren con respecto a otra medida de referencia centrada alrededor de los puntos de máximo interés y que se comporte de manera similar a las medidas posteriores a medida que el tamaño de los datos crece. Se consideran métodos numéricos basados en una aproximación Gaussiana de la aproximación de Laplace [51].

Cuando la distribución de probabilidad posterior se concentra más estrechamente alrededor de ciertos valores de $\boldsymbol{\nu}$, el método de Laplace ofrece una manera precisa y explícita de calcular estas integrales. Esto nos permite comprender mejor su comportamiento a medida que el tamaño de los datos aumenta [89].

3.3.1. Convergencia de la aproximación de Laplace

Consideremos el método de Laplace para aproximar la integral \mathbf{I} como una integral asintótica sobre un dominio D en \mathbb{R}^d , dada por [94]:

$$I := \int_D e^{-nJ(\boldsymbol{\nu})} d\boldsymbol{\nu}, \quad \boldsymbol{\nu} \in \mathbb{R}^d, \quad (3.21)$$

donde n es el número de datos observados. Suponemos que la expresión (3.21) cumple las siguientes condiciones [85]:

1. La integral I converge para todo $n \in \mathbb{N}$.
2. Existe un $\boldsymbol{\nu}_{\text{MAP}}$ en el interior del dominio D tal que, para cada $r > 0$, se cumple:

$$\delta_r := \inf_{\boldsymbol{\nu} \in B_r(\boldsymbol{\nu}_{\text{MAP}})} (J(\boldsymbol{\nu}) - J(\boldsymbol{\nu}_{\text{MAP}})) > 0, \quad (3.22)$$

donde $B_r(\boldsymbol{\nu}_{\text{MAP}})$ es una esfera de radio r centrada en $\boldsymbol{\nu}_{\text{MAP}}$.

3. En un vecindario de $\boldsymbol{\nu}_{\text{MAP}}$, la función de costo $J(\boldsymbol{\nu})$ es dos veces continuamente diferenciable, y la matriz Hessiana $\mathbf{H} := \nabla^2 J(\boldsymbol{\nu}_{\text{MAP}})$ es definida positiva.

Si aproximamos $J(\boldsymbol{\nu})$ usando expansión de Taylor de segundo orden alrededor del punto $\boldsymbol{\nu}_{\text{MAP}}$ nos queda

$$J(\boldsymbol{\nu}) \approx J(\boldsymbol{\nu}_{\text{MAP}}) + \frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}})^T \nabla^2 J(\boldsymbol{\nu}_{\text{MAP}})(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}}), \quad (3.23)$$

donde el término lineal desaparece ya que $\boldsymbol{\nu}_{\text{MAP}}$ es un punto crítico y $\nabla^2 J(\boldsymbol{\nu}_{\text{MAP}})$ es la matriz hessiana evaluada en $\boldsymbol{\nu}_{\text{MAP}}$.

Reemplazando la expansión de la ecuación (3.23) en la integral (3.21) obtenemos

$$I_{\text{asintótica}} \approx \int_D e^{-nJ(\boldsymbol{\nu})} d\boldsymbol{\nu} \approx \int_D e^{-n(J(\boldsymbol{\nu}_{\text{MAP}}) + \frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}})^T \nabla^2 J(\boldsymbol{\nu}_{\text{MAP}})(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}}))} d\boldsymbol{\nu}. \quad (3.24)$$

Para calcular la integral de la ecuación (3.24) realizamos un cambio de variable $\boldsymbol{\nu}' = \sqrt{n}(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}})$ lo cual implica $d\boldsymbol{\nu}' = \sqrt{n}d\boldsymbol{\nu}$. Definimos la matriz hessiana $\mathbf{H} = \nabla^2 J(\boldsymbol{\nu}_{\text{MAP}})$, por lo que la integral 3.24 se puede escribir de la siguiente forma:

$$I_{\text{asintótica}} \approx e^{-nJ(\boldsymbol{\nu}_{\text{MAP}})} \int_{D'} e^{-\frac{1}{2}\boldsymbol{\nu}'^T \nabla^2 J(\boldsymbol{\nu}_{\text{MAP}})\boldsymbol{\nu}'} \frac{d\boldsymbol{\nu}'}{\sqrt{n}} \approx e^{-nJ(\boldsymbol{\nu}_{\text{MAP}})} \frac{(2\pi)^{d/2}}{n^{d/2} \sqrt{|\mathbf{H}|}}, \quad (3.25)$$

donde $D' = (\boldsymbol{\nu}' : \boldsymbol{\nu} \in D)$, $|\mathbf{H}|$ es el determinante de la Hessiana de la función de costo evaluada en $\boldsymbol{\nu}_{\text{MAP}}$ y d es la dimensión de $\boldsymbol{\nu}$.

Según [89], a medida que el número de observaciones $n \rightarrow \infty$, el cociente entre la integral exacta I y su aproximación asintótica (ecuación (3.25)) tiende a 1, es decir

$$\lim_{n \rightarrow \infty} \frac{I}{e^{-nJ(\boldsymbol{\nu}_{\text{MAP}})} \frac{(2\pi)^{d/2}}{n^{d/2} \sqrt{|\mathbf{H}|}}} = 1. \quad (3.26)$$

Esto implica que cuando n es grande, la diferencia entre la integral exacta I y su aproximación asintótica dada por la expresión (3.25), disminuye y está acotada por $O(n^{-d/2})$

$$\left| \int_{\mathbb{R}^d} e^{-nJ(\boldsymbol{\nu})} d\boldsymbol{\nu} - e^{-nJ(\boldsymbol{\nu}_{\text{MAP}})} \int_{\mathbb{R}^d} e^{-\frac{1}{2}\|\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}}\|_{\mathbf{H}}^2} d\boldsymbol{\nu} \right| \in O(n^{-d/2}), \quad (3.27)$$

donde $J(\boldsymbol{\nu}_{\text{MAP}})$ es la función de costo evaluada en el punto de máximo a posteriori $\boldsymbol{\nu}_{\text{MAP}}$, y \mathbf{H} es la Hessiana de J evaluada en $\boldsymbol{\nu}_{\text{MAP}}$.

Es decir, se establece que, bajo ciertas condiciones sobre la función de costo $J(\boldsymbol{\nu})$, la aproximación de Laplace converge a la distribución posterior a medida que el número de observaciones n tiende a infinito. Esto implica que la aproximación de Laplace es una herramienta eficaz para aproximar la distribución posterior en problemas de inferencia bayesiana cuando se cumplen las condiciones adecuadas [94].

Además, implica que el comportamiento asintótico de $e^{-nJ(\boldsymbol{\nu})}$ es el mismo que el de la integral de la densidad de probabilidad Gaussiana con media $\boldsymbol{\nu}_{\text{MAP}}$ y covarianza $(n\mathbf{H})^{-1}$.

3.3.2. Definición de la aproximación de Laplace de la probabilidad posterior

Como se mencionó anteriormente, los valores iniciales o a priori de los parámetros desconocidos, $\boldsymbol{\nu}_0$, son continuos con respecto a la medida de Lebesgue y tienen una densidad de probabilidad a priori $p_{\text{priori}} : \mathbb{R}^d \rightarrow [0, \infty)$, definida como [10]:

$$S_0 := \{\boldsymbol{\nu} \in \mathbb{R}^d : p_{\text{priori}}(\boldsymbol{\nu}) > 0\} = \text{supp } p_{\text{priori}}, \quad (3.28)$$

es decir, S_0 es el conjunto de todos los posibles valores de $\boldsymbol{\nu}$ en \mathbb{R}^d donde la densidad de probabilidad a priori $p_{\text{priori}}(\boldsymbol{\nu})$ es mayor que cero.

Para definir la aproximación de Laplace de la probabilidad posterior $p_{\text{posterior}}(\boldsymbol{\nu})$, consideramos la *Suposición 1* y la *Observación 1* que se explican a continuación.

Suposición 1: Se cumple que la función de verosimilitud $p(\mathbf{d}^{\text{obs}}|\boldsymbol{\nu})$ y la probabilidad a priori $p_{\text{priori}}(\boldsymbol{\nu}_0) \in C^2(S_0, \mathbb{R})$, es decir, las funciones $p(\boldsymbol{\nu}|\mathbf{d}^{\text{obs}})$ y $p_{\text{priori}}(\boldsymbol{\nu}_0) : S_0 \rightarrow \mathbb{R}$ son continuamente diferenciables dos veces. Además, la función de costo $\mathbf{J}(\boldsymbol{\nu})$ tiene un mínimo único en el punto $\boldsymbol{\nu}_{\text{MAP}} \in S_0$, satisfaciendo [85]:

$$J(\boldsymbol{\nu}_{\text{MAP}}) = 0, \quad \nabla J(\boldsymbol{\nu}_{\text{MAP}}) = 0, \quad \nabla^2 J(\boldsymbol{\nu}_{\text{MAP}}) > 0. \quad (3.29)$$

Observación 1: Asumiendo que $\min_{\boldsymbol{\nu} \in S_0} J(\boldsymbol{\nu}_{\text{MAP}}) = 0$ es una solución particular y en general no restrictiva, si $\min_{\boldsymbol{\nu} \in S_0} J(\boldsymbol{\nu}) = c > -\infty$, entonces se puede establecer lo siguiente [85]

$$\begin{aligned} \log(\hat{p}(\mathbf{d}^{\text{obs}}|\boldsymbol{\nu})) &:= \log(p(\mathbf{d}^{\text{obs}}|\boldsymbol{\nu})) - c, \\ \hat{J}(\boldsymbol{\nu}) &:= \log(\hat{p}(\mathbf{d}^{\text{obs}}|\boldsymbol{\nu})) + \log(p_{\text{priori}}(\boldsymbol{\nu}_0)). \end{aligned}$$

Así obtenemos la probabilidad posterior $\hat{p}(\boldsymbol{\nu}|\mathbf{d}^{\text{obs}})$:

$$\hat{p}(\boldsymbol{\nu}|\mathbf{d}^{\text{obs}}) = e^{-n\hat{J}(\boldsymbol{\nu})}, \quad (3.30)$$

donde al evaluar $\hat{J}(\boldsymbol{\nu})$ en $\boldsymbol{\nu}_{\text{MAP}}$ se tiene

$$\min_{\boldsymbol{\nu}_{\text{MAP}} \in S_0} \hat{J}(\boldsymbol{\nu}_{\text{MAP}}) = \min_{\boldsymbol{\nu}_{\text{MAP}}} \log(\hat{p}(\mathbf{d}^{\text{obs}}|\boldsymbol{\nu})) + \log(p_{\text{priori}}(\boldsymbol{\nu}_0)) = 0. \quad (3.31)$$

Expandiendo la función de costo $\hat{J}(\boldsymbol{\nu})$ en $\boldsymbol{\nu}_{\text{MAP}}$ mediante la serie de Taylor, obtenemos

$$\hat{J}(\boldsymbol{\nu}) = J(\boldsymbol{\nu}_{\text{MAP}}) + \nabla J(\boldsymbol{\nu}_{\text{MAP}})(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}}) + \frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}})^T \nabla^2 J(\boldsymbol{\nu}_{\text{MAP}})(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}}). \quad (3.32)$$

Dadas las condiciones de la *Suposición 1*, es posible aproximar la función de costo de la ecuación (3.32) como

$$\hat{J}(\boldsymbol{\nu}) \approx \frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}})^T \nabla^2 J(\boldsymbol{\nu}_{\text{MAP}})(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}}), \quad (3.33)$$

por lo tanto, la probabilidad posterior puede escribirse como

$$\hat{p}(\boldsymbol{\nu}|\mathbf{d}^{\text{obs}}) = e^{-\frac{n}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}})^T \nabla^2 J(\boldsymbol{\nu}_{\text{MAP}})(\boldsymbol{\nu} - \boldsymbol{\nu}_{\text{MAP}})}, \quad (3.34)$$

la cual, comparándola con una Gaussiana multivariada, tiene una media $\boldsymbol{\nu}_{\text{MAP}}$ y una covarianza $\mathbf{C} := (\nabla^2 J(\boldsymbol{\nu}_{\text{MAP}}))^{-1} = \mathbf{H}(\boldsymbol{\nu}_{\text{MAP}})^{-1}$.

Comparando la aproximación de Laplace $I(\boldsymbol{\nu})$ de la ecuación (3.21) con la forma de la probabilidad posterior $\hat{p}(\boldsymbol{\nu}|\mathbf{d}^{\text{obs}})$ de la ecuación anterior, es evidente que la aproximación de Laplace busca encontrar una distribución Gaussiana multivariada $\mathcal{N}(\boldsymbol{\nu}_{\text{MAP}}, n^{-1}\mathbf{C})$ que se ajuste a la estructura de la probabilidad posterior cuando n es grande.

Por lo tanto, la aproximación de Laplace de la probabilidad posterior se define como la distribución Gaussiana multivariada centrada en el máximo a posteriori $\boldsymbol{\nu}_{\text{MAP}}$ y con covarianza proporcional a n^{-1} de la inversa de la Hessiana de la función de costo evaluada en $\boldsymbol{\nu}_{\text{MAP}}$.

3.4. Cuantificación de incertidumbre mediante métodos de cadenas de Markov Monte Carlo

Inicialmente, los métodos de Cadenas de Markov Monte Carlo (MCMC) surgieron como una combinación de cadenas de Markov y técnicas de Monte Carlo para calcular integrales complejas [70]. Una cadena de Markov es un modelo secuencial que se mueve de un estado a otro de manera probabilística, siendo el próximo estado condicionado al estado anterior [71]. Bajo ciertas condiciones, si ejecutamos una cadena de Markov lo suficiente, los estados que toma la cadena se muestrean en una distribución de probabilidad [49]. Como hemos estudiado anteriormente en el método inverso bayesiano, podemos utilizar las muestras para investigar el comportamiento de una probabilidad posterior [15]. Además, estas muestras pueden ser utilizadas para evaluar integrales que definen cantidades de interés para los problemas en estudio [66].

Una cadena de Markov es una serie de eventos en la que la probabilidad de que ocurra algo depende únicamente de lo que ocurrió justo antes [12, 49]. En términos estadísticos, una secuencia de elementos aleatorios es una cadena de Markov si y solo si $p(\nu^{n+1}|\nu^n, \dots, \nu^2, \nu^1) = p(\nu^{n+1}|\nu^n)$, donde n es un índice que representa la posición en la secuencia de estados ν en una cadena de Markov. Estas cadenas son fundamentales en la modelización de procesos estocásticos y se utilizan en una amplia gama de aplicaciones, incluyendo el modelado del clima, el análisis de secuencias de ADN, la ingeniería de telecomunicaciones y la optimización de algoritmos, entre otros [12]. La teoría de las cadenas de Markov proporciona herramientas poderosas para analizar la evolución de sistemas dinámicos en los que las transiciones entre estados son probabilísticas y pueden depender únicamente del estado actual.

Una cadena de Markov está definida por tres elementos:

- Un *espacio de estados* S , el conjunto de todas las posibles realizaciones de los términos de la cadena.
- Un *operador de transición* $p(\nu^{n+1}|\nu^n)$ que define la probabilidad de transición del estado ν^n al estado ν^{n+1} .
- Una *distribución inicial* $p_{\text{prior}}(\nu)$ que define la probabilidad de estar en cualquiera de los estados posibles al comienzo ($n = 0$).

Para iniciar una cadena de Markov $\nu^0 \rightarrow \nu^1 \rightarrow \nu^2 \dots \rightarrow \nu^n \rightarrow \dots$, se muestrea un estado inicial ν^0 de acuerdo a $p_{\text{prior}}(\nu)$, y luego se transita de ν^n a ν^{n+1} de acuerdo al operador de transición elegido.

3.4.1. Convergencia del Método Monte Carlo de Cadenas de Markov

Si una cadena de Markov con un espacio de estados no contable es irreducible, recurrente de Harris, aperiódica y tiene una distribución estacionaria con densidad p_s , entonces las distribuciones marginales $p(\nu^n)$ proporcionadas por la cadena convergen a p_s . Es decir, para una matriz de transición estocástica aperiódica con distribución invariante p_s , se cumple que [13, 54],

$$\lim_{n \rightarrow \infty} p(\nu^n) = p_s \quad \text{para todos los estados } \nu \in S.$$

donde:

- $p(\nu^n)$ es la probabilidad de estar en el estado ν en la iteración n , que converge a la distribución estacionaria p_s a medida que n tiende a infinito.

■ ν^n es el estado en la n -ésima iteración de la cadena de Markov.

Para demostrar este resultado, asumimos que $X = (\nu_X^0, \nu_X^1, \dots)$ es una cadena de Markov con núcleo de transición $p(\nu^{n+1} | \nu^n)$ que comienza en un estado fijo ν_X^0 . Además, consideramos la cadena de Markov $Z = (\nu_Z^0, \nu_Z^1, \dots)$ con el mismo núcleo de transición $p(\nu^{n+1} | \nu^n)$ pero con una distribución inicial p_s , es decir, $p(\nu_Z^0 = s) = p_s$ para todo $s \in S$ [82]. Suponemos que X y Z son cadenas independientes entre sí.

Se define T como el primer índice n tal que los estados de las cadenas X y Z son iguales por primera vez [82], es decir

$$T = \min(n | \nu_X^n = \nu_Z^n). \quad (3.35)$$

Ahora, podemos calcular la probabilidad de que las cadenas X y Z no se acoplen en los primeros nN pasos, es decir, la probabilidad de que $T > nN$ es

$$p(T > nN) = p(\nu_X^i \neq \nu_Z^i \text{ para todo } i \leq nN).$$

Dado que en cada bloque de N pasos existe al menos una probabilidad $\epsilon > 0$ de que X y Z se acoplen en T , por lo que podemos acotar la probabilidad de que las cadenas no se acoplen,

$$p(T > nN) \leq (1 - \epsilon)^n,$$

esta expresión significa que en cada bloque de N pasos hay una probabilidad de al menos $1 - \epsilon$ de que las cadenas no sean iguales, si este proceso se repite n veces entonces la probabilidad de que las cadenas no se acoplen es de $(1 - \epsilon)^n$, es decir, la probabilidad de que T sea mayor que nN decae exponencialmente a medida que n aumenta, esto implica que:

$$\lim_{n \rightarrow \infty} (1 - \epsilon)^n = 0. \quad (3.36)$$

De la ecuación (3.36) se implica que la probabilidad de que X y Z se acoplen eventualmente es 1, debido al hecho de que la probabilidad de no acoplarse se reduzca exponencialmente.

Definimos una nueva cadena Y que sigue a X hasta el tiempo T (cuando X y Z se encuentran) y luego sigue a Z desde ese tiempo en adelante:

$$\nu_Y^n = \begin{cases} \nu_X^n & \text{si } n \leq T, \\ \nu_Z^n & \text{si } n > T. \end{cases}$$

Dado que las cadenas X y Z se acoplan en el tiempo T , sus trayectorias coinciden a partir de ese momento. Como la cadena Z tiene una distribución inicial estacionaria p_s , la cadena Y que sigue la dinámica de Z a partir de T , se puede concluir que Y convergerá a la distribución estacionaria p_s porque para $n > T$ la cadena Y tendrá el mismo comportamiento que Z , es decir

$$p(\nu_Y^n = s | n > T) = p_s, \quad (3.37)$$

para todo $s \in S$.

Esto quiere decir que a medida que avanzan los pasos en la cadena Z , esta sigue manteniendo la distribución p_s para cualquier paso n

$$p(\nu_Z^n) = p_s. \quad (3.38)$$

Como la cadena Y sigue el mismo comportamiento que la cadena Z para $n > T$, y Z está en una distribución estacionaria p_s , se concluye que la cadena Y también converge a p_s

$$p(\nu_Y^n = s) = p(\nu_Z^n = s) = p_s. \quad (3.39)$$

Por lo tanto, la distribución de probabilidades de la cadena Y converge a p_s para todos los estados $\nu_Y \in S$, conforme $n \rightarrow \infty$.

3.4.2. El Teorema Ergódico (Ley fuerte de los grandes números)

Consideremos una cadena de Markov con un espacio de estados continuo y no numerable X es irreducible, recurrente de Harris y tiene una distribución estacionaria con densidad p_s , entonces [74, 82],

$$\frac{1}{N} \sum_{n=1}^N f(\boldsymbol{\nu}^n) \rightarrow \int_{x \in X} f(x) p_s(x) dx, \quad N \rightarrow \infty,$$

para cualquier función $f : X \rightarrow \mathbb{R}$ tal que el valor esperado $\int_{x \in X} f(x) p_s(x) dx$ exista y sea finito [82]. N representa el número de iteraciones de la cadena de Markov y a medida que N tiende a infinito, el promedio de las evaluaciones de la función f en los estados de la cadena de Markov convergerá al valor esperado de f bajo la distribución estacionaria p_s .

Una cadena de Markov se dice que es *irreducible* si y solo si hay una probabilidad positiva de que, para cualquier estado inicial $\boldsymbol{\nu}$, la cadena alcance cualquier conjunto A de medida positiva en un tiempo finito. Una cadena de Markov es *recurrente de Harris* si y solo si es irreducible y la probabilidad de que la cadena regrese infinitas veces a cualquier conjunto A de medida positiva (después de haber comenzado desde un punto perteneciente a A) es igual a 1. Una cadena de Markov es *aperiódica* si no queda atrapada en un ciclo de regreso entre el mismo conjunto de estados a intervalos regulares [71].

3.4.3. Algoritmo de Metropolis-Hastings

Dada una densidad de probabilidad objetivo p_s , definida en un espacio de estados S , y computable hasta una constante multiplicativa, $p_s(\boldsymbol{\nu}) \propto \tilde{p}_s(\boldsymbol{\nu})$, el algoritmo de Metropolis-Hastings propone una forma genérica de construir una cadena de Markov en S que es ergódica y estacionaria con respecto a p_s [54]. Esto significa que, si $\boldsymbol{\nu}^n$ es una muestra generada a partir de la distribución $p_s(\boldsymbol{\nu})$, entonces la siguiente muestra $\boldsymbol{\nu}^{n+1}$ también seguirá la distribución $p_s(\boldsymbol{\nu})$, y por lo tanto, las distribuciones marginales de la cadena convergerán a la distribución p_s [70].

Inherente a este muestreador es que cada estado propuesto pasa por un paso de aceptación-rechazo que decide si se elige el estado propuesto o se mantiene la muestra anterior para ser la siguiente muestra. Este paso por sí solo asegura la invarianza requerida y, por lo tanto, otorga al usuario una gran libertad para explorar el espacio de estados [54].

Consideramos el núcleo de transición $p(\boldsymbol{\nu}^{n+1} | \boldsymbol{\nu}^n)$, que describe la probabilidad de transitar desde el estado $\boldsymbol{\nu}^n$ al estado $\boldsymbol{\nu}^{n+1}$ en la cadena de Markov. En un paso n , dada la muestra actual $\boldsymbol{\nu}^n$, se propone un nuevo estado $\boldsymbol{\nu}_{\text{prop}}$ de acuerdo con el núcleo de propuesta $p(\boldsymbol{\nu}_{\text{prop}} | \boldsymbol{\nu}^n)$. La probabilidad de aceptar esta propuesta se calculará de la siguiente manera [54]:

$$\alpha = \min \left\{ 1, \frac{p_s(\boldsymbol{\nu}_{\text{prop}}) p(\boldsymbol{\nu}^n | \boldsymbol{\nu}_{\text{prop}})}{p_s(\boldsymbol{\nu}^n) p(\boldsymbol{\nu}_{\text{prop}} | \boldsymbol{\nu}^n)} \right\}.$$

La nueva muestra será $\boldsymbol{\nu}^{n+1} = \boldsymbol{\nu}_{\text{prop}}$, si la propuesta es aceptada, o bien $\boldsymbol{\nu}^{n+1} = \boldsymbol{\nu}^n$ en caso contrario.

La invarianza detallada de un núcleo de transición $p(\boldsymbol{\nu}^{n+1} | \boldsymbol{\nu}^n)$ se puede demostrar mediante la siguiente condición: se preserva el equilibrio respecto a la distribución estacionaria $p_s(\boldsymbol{\nu})$ si, para todo $\boldsymbol{\nu}^n, \boldsymbol{\nu}^{n+1} \in S$, se cumple que [82],

$$p_s(\boldsymbol{\nu}^n) p(\boldsymbol{\nu}^{n+1} | \boldsymbol{\nu}^n) = p_s(\boldsymbol{\nu}^{n+1}) p(\boldsymbol{\nu}^n | \boldsymbol{\nu}^{n+1}).$$

Esto significa que la probabilidad de transitar de $\boldsymbol{\nu}^n$ a $\boldsymbol{\nu}^{n+1}$, ponderada por la distribución estacionaria p_s , es igual a la probabilidad de transitar de $\boldsymbol{\nu}^{n+1}$ a $\boldsymbol{\nu}^n$, también ponderada por p_s . La

invarianza detallada implica que la distribución estacionaria p_s se conserva a lo largo del tiempo, asegurando la estabilidad y consistencia del proceso.

Sin embargo, la implicación opuesta, es decir, que la invarianza detallada conduce al equilibrio, no siempre se cumple en general, lo que subraya la importancia de verificar esta propiedad en cada caso específico de una cadena de Markov [88].

Una cadena de Markov que preserve el equilibrio detallado con respecto a una distribución p_s se dice que es reversible en equilibrio o simplemente reversible [82]. Esto significa que, si la cadena de Markov está en equilibrio con respecto a p_s , entonces la probabilidad de transición de un estado a otro es simétrica. En otras palabras, la probabilidad de transición entre dos estados cualesquiera es la misma, sin importar la dirección de la transición, es decir

$$p(\boldsymbol{\nu}^{n+1} | \boldsymbol{\nu}^n) = p(\boldsymbol{\nu}^n | \boldsymbol{\nu}^{n+1}).$$

Esta propiedad asegura que la cadena de Markov es reversible con respecto a la distribución p_s , y es fundamental para garantizar que la cadena converja a la distribución deseada.

El algoritmo de Metropolis-Hastings describe una cadena de Markov reversible. Esto es trivial para transiciones de $\boldsymbol{\nu}^{n+1}$ a $\boldsymbol{\nu}^n$ con $\boldsymbol{\nu}^n = \boldsymbol{\nu}^{n+1}$. Para $\boldsymbol{\nu}^n \neq \boldsymbol{\nu}^{n+1}$, consideramos que

$$\begin{aligned} p_s(\boldsymbol{\nu}^{n+1})p(\boldsymbol{\nu}^{n+1}|\boldsymbol{\nu}^n)\alpha(\boldsymbol{\nu}^{n+1}|\boldsymbol{\nu}^n) &= \text{mín} \{p_s(\boldsymbol{\nu}^{n+1})p(\boldsymbol{\nu}^{n+1}|\boldsymbol{\nu}^n), p_s(\boldsymbol{\nu}^n)p(\boldsymbol{\nu}^n|\boldsymbol{\nu}^{n+1})\} \\ &= p_s(\boldsymbol{\nu}^n)p(\boldsymbol{\nu}^n|\boldsymbol{\nu}^{n+1})\alpha(\boldsymbol{\nu}^n|\boldsymbol{\nu}^{n+1}), \end{aligned} \quad (3.40)$$

donde $\alpha(\boldsymbol{\nu}^{n+1}|\boldsymbol{\nu}^n)$ es la probabilidad de aceptación de la transición de $\boldsymbol{\nu}^n$ a $\boldsymbol{\nu}^{n+1}$ y $\alpha(\boldsymbol{\nu}^n|\boldsymbol{\nu}^{n+1})$ es la probabilidad de aceptación de la transición inversa. La aperiodicidad de la cadena de Markov resultante debe cumplirse conjuntamente con la probabilidad de aceptación y la propuesta. La ecuación de la probabilidad de aceptación muestra que una transición no trivial de $\boldsymbol{\nu}^n$ a $\boldsymbol{\nu}^{n+1}$ solo puede tener lugar si la transición inversa correspondiente también es factible, es decir, si $p_s(\boldsymbol{\nu}^n)p(\boldsymbol{\nu}^n|\boldsymbol{\nu}^{n+1}) > 0$. En el caso más extremo, esto podría significar que aplicamos una propuesta irreducible y aperiódica, pero la cadena de Markov resultante de Metropolis-Hastings nunca se alejará del estado inicial y, por lo tanto, no será irreducible [71].

Para garantizar que la cadena sea ergódica, es crucial elegir un núcleo de propuesta $p(\boldsymbol{\nu}_{\text{prop}}|\boldsymbol{\nu})$ que permita explorar todo el espacio de estados. Esto a menudo implica utilizar un núcleo que sea irreducible y aperiódico, asegurando que cualquier estado pueda eventualmente ser alcanzado desde cualquier otro estado [15].

3.4.3.1. Pseudocódigo Metrópolis-Hastings

El algoritmo de Metropolis-Hastings para muestrear p_s se puede resumir en los pasos descritos en el esquema adjunto [24, 34].

La convergencia deseada se refiere a cuando la cadena de Markov ha alcanzado una distribución estacionaria cercana a la distribución objetivo p_s . La convergencia puede ser monitoreada mediante varios métodos, tales como:

- Criterio de Convergencia de Variación Total: Monitorear si la variación total entre la distribución muestral y la distribución objetivo es suficientemente pequeña.
- Diagnóstico de Convergencia de Gelman-Rubin: Usar herramientas estadísticas como el diagnóstico de Gelman-Rubin para evaluar la convergencia [47].
- Visualización de Muestras: Evaluar la convergencia mediante el análisis gráfico de la cadena de Markov, verificando si las muestras parecen estabilizarse alrededor de la distribución objetivo.

Algorithm 1 Algoritmo de Metropolis-Hastings

- 1: **Inicializar** ν_0 en algún estado en S
- 2: **for** $n = 1$ **to** n_{max} **do**
- 3: **Proponer** un nuevo estado ν_{prop} a partir del estado actual ν^n usando la distribución de propuesta $p(\nu_{prop} | \nu^n)$
- 4: **Calcular** la probabilidad de aceptación α :

$$\alpha = \min \left\{ 1, \frac{p_s(\nu_{prop})p(\nu^n | \nu_{prop})}{p_s(\nu^n)p(\nu_{prop} | \nu^n)} \right\}$$

- 5: **Aceptar** el nuevo estado ν_{prop} con probabilidad α :

$$\nu^{n+1} = \begin{cases} \nu_{prop} & \text{con probabilidad } \alpha \\ \nu^n & \text{con probabilidad } 1 - \alpha \end{cases}$$

- 6: **end for**
 - 7: **Repetir** hasta alcanzar el número de iteraciones n_{max} o hasta que se cumpla un criterio de convergencia deseado
-

3.4.4. Muestreador de conjunto invariante afín

El muestreador de conjunto invariante afín (AIES) es un método de muestreo basado en cadenas de Markov que opera en un espacio de estados de dimensión finita. Para muestrear una densidad de probabilidad $p(\nu)$ en un espacio de estados \mathbb{R}^P , donde P es la dimensión del espacio, AIES genera un conjunto de caminantes $\{\nu_1, \dots, \nu_W\}$ en \mathbb{R}^P , donde cada ν_w representa una muestra en este espacio de estados, con $w = 1, \dots, W$. La propiedad de invariancia afín asegura que el algoritmo AIES mantiene su validez para la densidad de producto $p(\nu_1) \cdot \dots \cdot p(\nu_W)$ bajo cualquier transformación afín del espacio de estados, es decir, el producto de densidades $\prod_{n=1}^W p(\nu_n)$ permanece constante bajo transformaciones de la forma $\nu \mapsto A\nu + b$ [34], con A como matriz $P \times P$ y b como vector en \mathbb{R}^P .

En el proceso de muestreo, AIES sigue una variante del algoritmo de Metropolis-Hastings que mezcla cadenas. En cada iteración, se propone un cambio en la posición de un caminante para que se desplace hacia o se aleje de otro caminante, y esta propuesta se acepta o se rechaza según una regla de aceptación basada en el criterio de Metropolis-Hastings. La invariancia afín es fundamental en AIES, ya que asegura que tanto las propuestas de los cambios como las probabilidades de aceptación preservan la distribución objetivo $p(\nu)$ durante el muestreo.

En detalle, para realizar un paso de propuesta en AIES, se elige aleatoriamente un caminante ν_i del conjunto de caminantes $\{\nu_1, \dots, \nu_W\}$ y un segundo caminante ν_j , donde $i \neq j$ y ambos pertenecen al conjunto de W caminantes. Luego, se propone mover ν_i a una nueva posición ν_{prop} según la expresión [34]

$$\nu_{prop} = \nu_i + (1 - Z)(\nu_j - \nu_i),$$

donde Z es una variable aleatoria en el intervalo $[1/a, a]$ con una densidad de probabilidad $g(Z) \propto Z^{-1/2}$ [50]. En aplicaciones prácticas, típicamente se elige $a = 2$, pero a puede ser un parámetro ajustable que controla el tamaño del paso [34]. El principal costo computacional de AIES es la evaluación de la probabilidad de aceptación en el paso de Metropolis-Hastings, que se calcula como:

$$\alpha = \min \left\{ 1, Z^{P-1} \frac{p(\nu_{prop})}{p(\nu_i)} \right\}$$

donde ν_{prop} es la nueva posición propuesta para el caminante ν_i , y $p(\nu_i)$ es la densidad de probabilidad que se muestrea. Este cálculo implica evaluar la densidad de probabilidad $p(\nu)$ en la nueva posición propuesta y en la posición actual.

AIES es un método eficiente para muestrear distribuciones complejas en espacios de moderada dimensión, destacándose por su capacidad para manejar distribuciones mal escaladas y multimodales mediante un conjunto de caminantes y transformaciones afines.

Para dimensiones bajas a moderadas P , AIES proporciona algoritmos de muestreo rápidos y efectivos que permiten la implementación paralela y el manejo eficiente de distribuciones con características difíciles [44]. Sin embargo, a medida que la dimensión P aumenta, el requisito de $W > 2P$ puede resultar en una menor eficiencia del algoritmo.

Para una implementación práctica del muestreador de conjunto invariante afín (AIES), es común descartar una fracción inicial de B muestras generadas para eliminar el periodo de equilibrado. Esta práctica asegura que las muestras descartadas no influyan en el análisis final, permitiendo que las muestras restantes representen adecuadamente la distribución posterior. Se pueden utilizar criterios de convergencia, como el criterio de Gelman-Rubin [47], para evaluar si el muestreo ha sido suficiente o si es necesario generar más muestras para reflejar correctamente la estructura de la distribución posterior.

Una vez que se dispone de suficientes muestras, se pueden emplear herramientas de visualización para analizar la incertidumbre en la predicción de los parámetros de interés. Estas herramientas incluyen histogramas, medias y probabilidades marginales. En algunas ocasiones, la distribución posterior puede ser multimodal, lo que significa que la aproximación de Laplace puede no capturar completamente toda la estructura de la distribución posterior, limitándose a describir el comportamiento cerca de ν_{MAP} [26].

3.4.4.1. Pseudocódigo de muestreo con invariancia afín

Consideramos un algoritmo de muestreo basado en cadenas de Markov diseñado para manejar distribuciones con escalas diferentes en distintas direcciones y distribuciones multimodales [44]. A continuación, se presenta el pseudocódigo del AIES paso a paso para muestrear una distribución a posteriori:

Algorithm 2 Algoritmo de Muestreo AIES

1: **Inicialización**2: Establecer $i = 0$ y el espacio de muestra $X = \mathbb{R}^P$ 3: Establecer el parámetro de escala $a \sim 2$ y definir el número de caminantes W 4: Extraer ν_1^0, \dots, ν_W^0 del espacio de muestra X con una distribución inicial $p_{\text{prior}}(\nu)$

5:

6: **Iteración desde $i = 1$ hasta $i = I_{\text{max}}$** 7: **for $w = 1$ to W do**8: Elegir un índice j al azar del conjunto complementario $\{j \mid j \neq w\}$ 9: Extraer una variable aleatoria Z de la densidad $g(Z) \propto Z^{-1/2}$ para $Z \in [a^{-1}, a]$; de lo contrario, asignar $Z = 0$ 10: Proponer un nuevo estado para el estado ν_w :

$$\nu_{\text{prop},w} = \nu_j^i + (1 - Z)(\nu_w^i - \nu_j^i)$$

11: Calcular la probabilidad de aceptación α :

$$\alpha = \min \left\{ 1, Z^{P-1} \frac{p_{\text{posterior}}(\nu_{\text{prop},w}^i)}{p_{\text{posterior}}(\nu_w^i)} \right\}$$

12: Extraer un número aleatorio u de una distribución uniforme en $U(0, 1)$ 13: Aceptar el nuevo estado con probabilidad α :

$$\nu_w^{i+1} = \begin{cases} \nu_{\text{prop},w}^i & \text{si } u < \alpha \\ \nu_w^i & \text{si } u \geq \alpha \end{cases}$$

14: **end for**

15:

16: **Salida**17: Obtener las muestras de la probabilidad a muestrear $p_{\text{posterior}}$:

$$\nu_1^0, \nu_2^0, \dots, \nu_W^0, \dots, \nu_1^{I_{\text{max}}}, \nu_2^{I_{\text{max}}}, \dots, \nu_W^{I_{\text{max}}}$$

Capítulo 4

Estrategias de inversión del campo de ondas basadas en objetos

En este capítulo, exploraremos las estrategias de inversión bayesianas en el campo de ondas elásticas para la detección de un objeto elíptico incorporado en un dominio estratificado. La inversión del campo de ondas es un área de estudio fundamental en diversas disciplinas, desde la exploración sísmica hasta la reconstrucción de imágenes médicas [24].

Para la inversión del campo de ondas, comenzaremos examinando los operadores de observación aproximados (véase sección 4.1), que desempeñan un papel crucial en la adquisición de datos en entornos donde las mediciones precisas son difíciles de obtener. Con los operadores de observación aproximados obtenidos, nos sumergiremos en el núcleo del problema: el enfoque bayesiano inverso aproximado (véase sección 4.2), que nos permite inferir información sobre un sistema desconocido a partir de observaciones limitadas, datos con ruido incorporado y campos de densidad y velocidad incompletos [25].

Posteriormente, exploraremos el cálculo del Máximo a Posteriori (MAP) mediante optimización (véase sección 4.3), una técnica esencial para estimar los parámetros desconocidos del modelo, los cuales permitirán reconstruir los objetos o anomalías presentes en el dominio verdadero. Finalmente, presentaremos los resultados obtenidos a través de muestras de los parámetros desconocidos, utilizando una aproximación de Laplace como una gaussiana multivariada. La distribución de las muestras se podrá visualizar a través de histogramas (véase sección 4.4), brindando una visión más completa de la efectividad y la aplicabilidad práctica de la estrategia de inversión bayesiana.

4.1. Operadores de observación aproximados

Para la inversión de campos de ondas es necesario definir el operador de observación aproximado $\mathbf{f}(\boldsymbol{\nu})$ que se relaciona con los datos medidos \mathbf{d}^{obs} . Considerando los parámetros de las verdaderas inclusiones $\boldsymbol{\nu}_{\text{exacto}}$, al resolver el problema directo con ellos obtenemos $\mathbf{f}(\boldsymbol{\nu}_{\text{exacto}}) = \mathbf{d}_{\text{exacto}}$. Los datos realmente observados están corrompidos por ruido procedente de diversas causas:

$$\mathbf{d}^{\text{obs}} = \mathbf{f}(\boldsymbol{\nu}) + \boldsymbol{\eta}, \quad (4.1)$$

donde $\mathbf{f}(\boldsymbol{\nu})$ son los valores obtenidos por los receptores en los tiempos de muestreo al resolver el problema directo. El término de ruido $\boldsymbol{\eta}$ tiene componentes de la forma ϵz , con $\epsilon = \frac{\alpha}{100} \sqrt{\frac{\sum_{j=1}^J \sum_{m=1}^M |d_{j,\text{exacto}}^m|^2}{JM}}$, donde J es el número de receptores que en este caso es igual a $J = 51$, M es el número de mediciones temporales por cada uno de los receptores, en este caso se considera $M = 211$, α es el nivel de porcentual de ruido que se escogió; en este caso se escogió 5%, z corresponden a números aleatorios generados con una distribución normal $\mathcal{N}(0, 1)$. Dados datos verdaderos $\mathbf{d}_{\text{exacto}}$ obtenidos calculando $\mathbf{f}_{\text{exacto}}(\boldsymbol{\nu})$, la relación (4.1) nos permite generar datos sintéticos \mathbf{d}^{obs} para ensayos numéricos.

Como se ha mencionado anteriormente, el parámetro ν contiene información del objeto la que corresponde a las características geométricas, posición y propiedades físicas de la anomalía. En esta sección estudiaremos la resolución del problema directo para la obtención de $\mathbf{f}(\nu)$. El problema directo corresponde a la propagación del desplazamiento de ondas elásticas $u(\mathbf{x}, t)$ ¹, las cuales son emitidas por una red de emisores \mathbf{x}_k y que son medidas a través de la red de receptores \mathbf{r}_j . Para las simulaciones de este capítulo, las ondas elásticas son emitidas dentro un dominio truncado y estratificado en cinco capas de diferentes materiales, véase Figura 4.2. En la capa central del dominio se posicionará un objeto o anomalía, que en este capítulo se parametrizará como una elipse. El parámetro ν , que contendrá información de la anomalía, tendrá la siguiente información: la posición de la elipse (centro c_x, c_y), geometría (longitud de semieje mayor a y semieje menor b), ángulo de rotación (ω) y propiedades físicas (densidad ρ y velocidad de propagación de la onda v_p). En este caso ν corresponderá a un vector de dimensión 7, el cual tendrá información de la posición, geometría y parámetros físicos de la elipse, es decir, $\nu = (c_x, c_y, a, b, \omega, \rho, v_p)$.

Para la resolución del problema directo y la obtención del operador de observación $\mathbf{f}(\nu)$, es necesario describir la geometría utilizada como dominio junto con sus propiedades físicas. Para la resolución numérica de la formulación débil de la ecuación de ondas es necesario discretizarla espacialmente y temporalmente, como se explicó en el capítulo sobre estudio teórico y numérico del problema directo. En cuanto a la discretización espacial, se realizarán dos: una será de un espaciado uniforme en todo el dominio y la otra será una discretización que dependerá del tamaño y geometría de cada capa, por lo que el espaciado y la distribución de los elementos no será uniforme. Como resultado de las dos discretizaciones espaciales se tendrán dos mallados numéricos diferentes y por lo tanto, dos conjuntos de mediciones $\mathbf{f}(\nu)$ del problema directo, obteniendo posteriormente dos operadores de observación aproximados \mathbf{d}^{obs} , correspondientes a cada uno de los mallados numéricos. Compararemos los resultados obtenidos para cada una de las discretizaciones.

4.1.1. Geometría

La geometría que se utilizará en este capítulo será una geometría truncada y corresponderá a la región $[-3, 0] \times [-1'5, 1'5]$. Se dividirá en 5 capas y en la capa central se incorporará la anomalía de forma elíptica, como se observa en la figura 4.2. Como se observa en la figura 4.2, los límites superiores e inferiores de cada capa están marcados en el eje vertical de la figura, en total el dominio computacional se dividirá en 5 regiones y cada capa tendrá los límites que aparecen en la tabla de la figura 4.1.

La anomalía elíptica se incorporará a través de una parametrización, que incorpora el ángulo θ que corresponde al ángulo de sistemas polares, esto es

$$\begin{aligned}
 X(\theta) &= a \cos(\theta) \\
 Y(\theta) &= b \sin(\theta) \\
 c_x(\theta) &= c_x + X(\theta) \cos(\omega) - Y(\theta) \sin(\omega) \\
 c_y(\theta) &= c_y + X(\theta) \sin(\omega) + Y(\theta) \cos(\omega)
 \end{aligned} \tag{4.2}$$

El conjunto de ecuaciones (4.2) son las ecuaciones de parametrización de la elipse que dependen de: el semieje mayor de la elipse a , semieje menor b de la elipse, coordenadas del centro c_x y c_y , ángulo de rotación ω y del ángulo de sistemas polares θ .

Como se ha explicado, el operador de observación depende de los parámetros ν de la anomalía, el cual corresponderá a un vector de dimensión 7 y que tendrá información de la posición, geometría y

¹ Donde $\mathbf{x} = (x, y)$ representa la posición espacial en dos dimensiones.

Capa	Límites
1	$[-0'4, 0]$
2	$[-1'1, -0'4]$
3	$[-1'7, -1'1]$
4	$[-2'5, 1'7]$
5	$[-3'0, -2'5]$

Figura 4.1: Capas de caso 1.

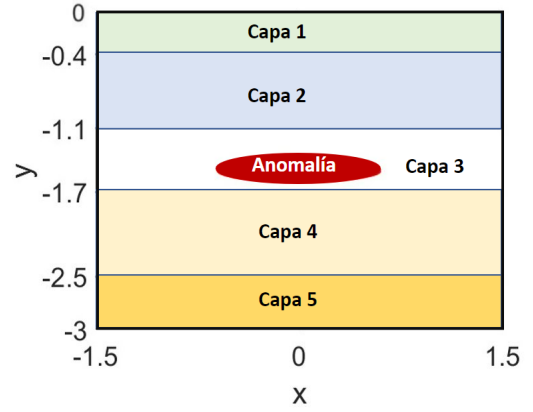


Figura 4.2: Geometría correspondiente caso 1.

parámetros físicos de la elipse, es decir

$$\boldsymbol{\nu} = (c_x, c_y, a, b, \omega, \rho, v_p), \quad (4.3)$$

donde ρ es la densidad y v_p es la velocidad de propagación de la onda. Los valores numéricos verdaderos de cada parámetro de $\boldsymbol{\nu}$ se observan en el cuadro 4.1.

Parámetro	Valor verdadero
c_x	0'0
c_y	-1'5
a	0'6
b	0'1
ω	0'0
ρ	2'1
v_p	4'4

Cuadro 4.1: Valores verdaderos de $\boldsymbol{\nu}$ para la elipse.

Para la resolución numérica del problema de la propagación de onda se requieren las propiedades físicas del dominio truncado, es decir, las propiedades correspondiente a la densidad y a la velocidad de propagación de onda. Como el dominio truncado se divide en 5 capas de diferentes materiales se debe especificar los valores de la densidad $\rho(\mathbf{x})$ y de la velocidad $v_p(\mathbf{x})$ correspondiente a cada una de las capas, como se puede observar en el cuadro 4.2. Además de la densidad y de la velocidad de propagación de onda, se ha calculado el módulo de elasticidad por capa usando la relación $\chi(\mathbf{x}) = \rho(\mathbf{x})v_p^2(\mathbf{x})$.

Capa	Densidad $\rho(\mathbf{x})$	Velocidad de onda $v_p(\mathbf{x})$	Módulo de elasticidad $\chi(\mathbf{x})$
Capa 1	2'00	1'5	4'5000
Capa 2	2'50	2'5	15'6250
Capa 3	2'49	2'8	19'5216
Capa 4	2'49	3'3	27'1161
Capa 5	2'60	3'1	40'6560

Cuadro 4.2: Propiedades físicas adimensionales para cada una de las capas.

Las ondas elásticas se emitirán desde un conjunto de emisores \mathbf{x}_k , con $k = 0, \dots, 50$, los cuales se posicionarán en el borde superior del recinto computacional, es decir en $y = 0$. El barrido horizontal de emisores irá desde la coordenada $(-1'0, 0)$ hasta $(1'0, 0)$ cada $0'04$, en total, para este caso serán $K = 50$.

Para obtener el conjunto de mediciones $\mathbf{f}(\boldsymbol{\nu})$, el cual corresponderá al desplazamiento de la onda en diferentes instantes de posición y de tiempo, se posicionarán receptores \mathbf{r}_j , con $j = 0, \dots, 51$, de forma intercalada a la emisores, éstos irán desde la posición $(-1'02, 0)$ hasta $(1'02, 0)$ cada $0'04$, en total, para este caso, serán $J = 51$ receptores.

Por otra parte, para la resolución del método directo debemos conocer las características de las fuentes, por ejemplo en la variación temporal de la onda, $f(t) = f_0(1 - 2\pi^2 f_M^2 t^2) e^{-\pi^2 f_M^2 t^2}$, la amplitud inicial $f_0 = 100$, la frecuencia media de la onda $f_M = 2$ y la discretización temporal utilizada $t = N\delta t$, con $\delta t = 10^{-3}$. La dependencia espacial de la onda viene dada por $g(\mathbf{x}) = \frac{1}{(\pi\kappa)^2} \sum_{k=1}^K e^{-|\mathbf{x}-\mathbf{x}_k|^2}$, la cual depende de la posición de las fuentes \mathbf{x}_k y del parámetro de regularización $\kappa = 0'04$.

4.1.2. Discretización espacial

Para la resolución numérica de la formulación débil del problema directo, como se explicó anteriormente, se requiere discretizar espacialmente el dominio truncado sobre el que se trabajará computacionalmente. En este caso, donde el dominio se estratifica en 5 capas y se incorporara un objeto parametrizado elípticamente se estudiarán dos casos de discretizaciones espaciales: discretización uniforme y discretización estratificada, para estudiar si existe una similitud de resultados independiente de la discretización espacial usada. A continuación, se explicará cada una de las discretizaciones con sus características.

Discretización espacial uniforme

El primer tipo de discretización espacial que estudiamos es la discretización espacial uniforme. Este tipo de discretización espacial corresponde a un mallado numérico con elementos de igual tamaño y distribución en todo el dominio computacional truncado, independiente de la capa o de la anomalía, es decir, un mallado numérico con elementos de igual tamaño

Para discretizar el dominio de forma numérica usando el software Matlab se usará la función *delaunayTriangulation*, la cual generará el mallado uniforme en todo el dominio truncado, creando nodos y elementos correspondientes a la formulación débil del problema.

El tamaño de los elementos se configurará para que tanto horizontal como vertical, sea igual a $\delta_x = \delta_y = 0'04$. El tipo de elemento en esta discretización es tetraédrica, ya que se trabajará de forma bidimensional. El mallado uniforme utilizado en esta geometría es el que se muestra en la figura 4.4. El número total de elementos en este caso en el dominio computacional es de 11,250 y la geometría de los elementos es de orden lineal.

Discretización espacial estratificada

La segunda discretización espacial estudiada es la discretización estratificada. En este tipo de discretización se mallará cada una de las capas de forma independiente, por lo que el tamaño y distribución de los elementos dependerá del subdominio de cada capa. Para generar el mallado para

Mallado uniforme	Características
Nodos	$2 \times 5,776$
Elementos	$3 \times 11,250$
Orden geométrico	Lineal

Figura 4.3: Características de mallado uniforme.

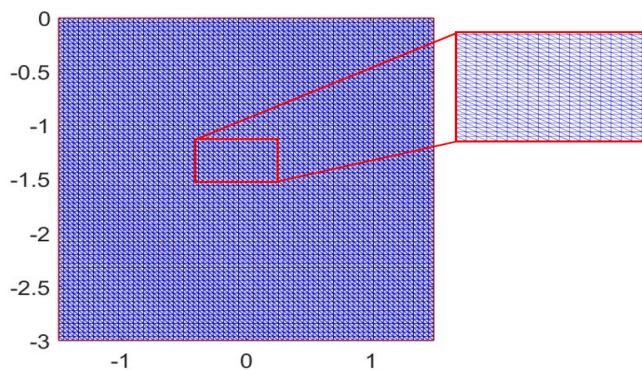


Figura 4.4: Mallado uniforme utilizado para el objeto elíptico.

cada una de las capas se usará la herramienta *GenerateMesh* y *pdetool* de Matlab, la cual generará la distribución de los elementos para cada una de las capas y asegurará la continuidad entre éstos en las interfaces de cada capa, de forma que los subdominios adyacentes sean continuos.

Usando la herramienta *pdetool* se dibujan las 5 capas del dominio con los límites mencionados en 4.2. Como se mencionó, cada uno de los subdominios se mallará de forma independiente usando la función *generateMesh*, por lo que habrán 5 mallados correspondientes a cada uno de los subdominios, sin embargo, la herramienta *pdetool* permitirá la continuidad entre los elementos que se encuentran en los límites de cada uno de los subdominios adyacentes.

Usando la función *generateMesh* se indica el tamaño máximo y mínimo de los elementos, además del orden geométrico de los elementos del mallado. En este caso el tamaño máximo de los elementos será igual a 0'04, el tamaño mínimo igual a 0'02 y el orden geométrico es lineal. El número total de elementos es de 13,008.

En la figura 4.6 se muestra el mallado estratificado para la geometría estratificada por capas, donde R_1, R_2, R_3, R_4 y R_5 representan los 5 subdominios.

4.1.3. Comparación de operadores de observación

Como resultado de la resolución de la formulación débil del problema directo se obtuvieron mapas de los desplazamientos de la onda para cada instante de tiempo. En este caso se aplicarán las discretizaciones espaciales explicadas en la sección 4.1.2. En cuanto a la discretización temporal se aplicará la condición de estabilidad de Courant-Friedrichs-Lewy explicada en el capítulo sobre estudio teórico y numérico del problema directo. En este caso, para asegurar la estabilidad de los resultados imponemos $\delta t < 10^{-3}$. Como paso de tiempo total se ha iterado 2,500 veces, obteniendo un tiempo total de 2'5 segundos.

Mallado estratificado	Características
Nodos	$2 \times 6,656$
Elementos	$3 \times 13,008$
Orden geométrico	Lineal

Figura 4.5: Características de mallado estratificado.

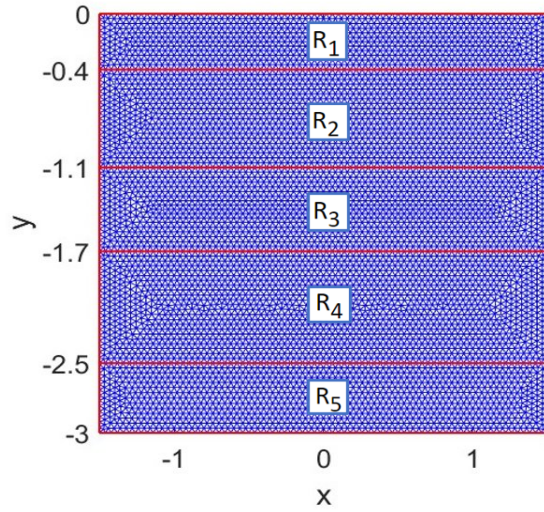


Figura 4.6: Mallado estratificado usado para el objeto elíptico.

En las figuras 4.7a y 4.7b se observan el desplazamiento de la onda para la discretización espacial uniforme y estratificada, respectivamente. Las amplitudes de la onda observadas en los mapas de colores corresponden al operador $\mathbf{f}(\boldsymbol{\nu})$ en el problema directo. Como se observan en ambas figuras las amplitudes son similares entre las dos discretizaciones espaciales.

Aún cuando se puede observar que las mediciones $\mathbf{f}(\boldsymbol{\nu})$ del desplazamiento de onda son similares entre ambos mallados, se cuantificará la diferencia entre ambas mediciones. Para ésto se ha calculado la diferencia porcentual entre $\mathbf{f}_{\text{uniforme}}(\boldsymbol{\nu})$ y $\mathbf{f}_{\text{estratificado}}(\boldsymbol{\nu})$, obteniendo un valor máximo de 0'47% debido a variaciones en la velocidad de propagación al cruzar capas. La diferencia de la solución entre los mallados se puede visualizar a través del mapa de la figura 4.8.

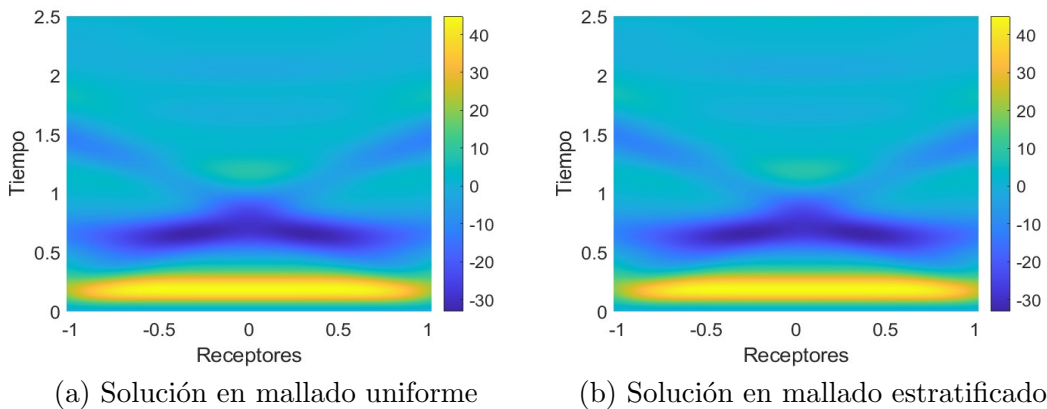


Figura 4.7: Solución en receptores con respecto al tiempo

Debido a las asimetrías que se producen en los contorno de los subdominios con la discretización estratificada de MATLAB (ver figura 4.6), se realizó otra discretización estratificada usando el software GID como se puede observar en la figura 4.9a. Se calculó que la diferencia porcentual máxima entre la discretización uniforme de MATLAB y la discretización estratificada de GID para un tiempo de simulación $t = 2'5$ segundos igual a 0'48%. En la figura 4.9b se observa la diferencia entre estos dos mallados.

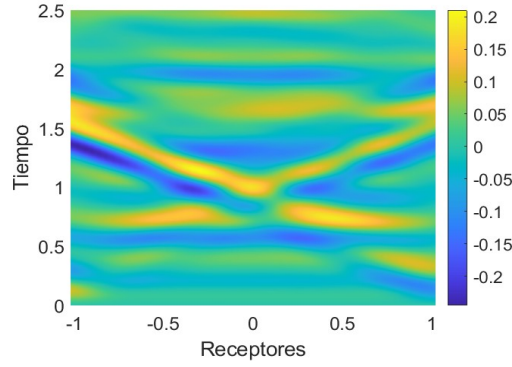
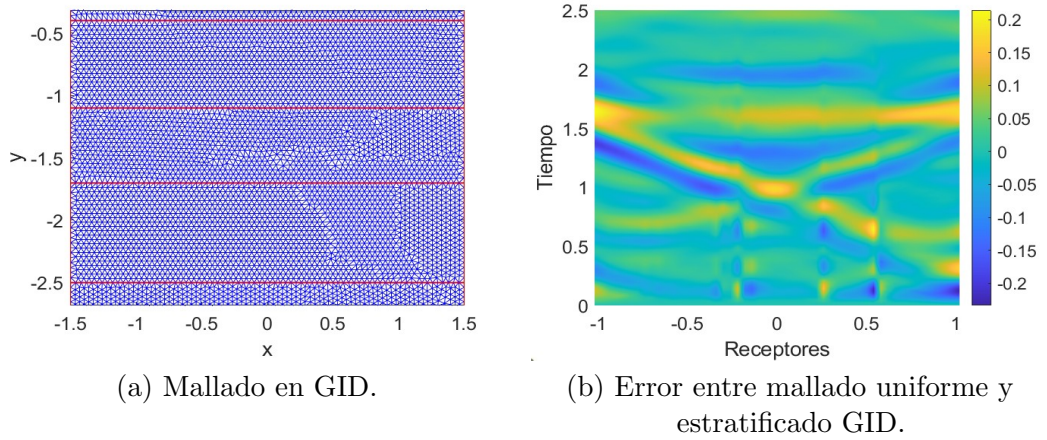


Figura 4.8: Diferencia entre mallados.



(a) Mallado en GID.

(b) Error entre mallado uniforme y estratificado GID.

Figura 4.9: Mallado estratificado en GID.

Como se puede observar de las figuras 4.8 y en 4.9b, la diferencia en los resultados entre los mallados estratificados (MATLAB y GID) y mallado uniforme (MATLAB) se mantiene en un rango similar (de 0'47% y de 0'48%, respectivamente), aunque con una distribución espacial distinta.

Esta pequeña diferencia en precisión contrasta con la ventaja computacional del mallado uniforme, que resolvió el problema directo en 2 minutos, mientras el tiempo de cálculo utilizando el mallado estratificado fue de 8 minutos. Esta reducción de tiempo es particularmente relevante para el problema inverso, donde se requieren múltiples evaluaciones del problema directo.

Como se mencionó anteriormente, para generar datos sintéticos \mathbf{d}^{obs} usamos los valores de las soluciones evaluados en los 51 receptores: $f_{\text{uniforme}}(\boldsymbol{\nu})$ y $f_{\text{estratificado}}(\boldsymbol{\nu})$. Como se explicó anteriormente en la ecuación (4.1), a estos conjuntos de mediciones se les añadirá ruido Gaussiano para su uso posterior como datos sintéticos en las pruebas de inversión

$$\mathbf{d}^{\text{obs}} = \mathbf{f}(\boldsymbol{\nu}) + \boldsymbol{\eta}, \quad (4.4)$$

según detalles ya dados en la sección 4.1.

De la resolución numérica del problema débil es posible graficar $\mathbf{f}(\boldsymbol{\nu})$ con respecto al tiempo, para cada uno de los receptores \mathbf{r}_j . Añadiendo ruido Gaussiano a $\mathbf{f}(\boldsymbol{\nu})$ se obtienen datos sintéticos en los 51 receptores para ensayos numéricos de resolución del problema inverso. En las figuras 4.10b y 4.11b se observan la diferencia de amplitud y forma entre $\mathbf{f}(\boldsymbol{\nu})$ y \mathbf{d}^{obs} , para cada una de las discretizaciones espaciales, para el tiempo total de simulación de $t = 2'5$ segundos.

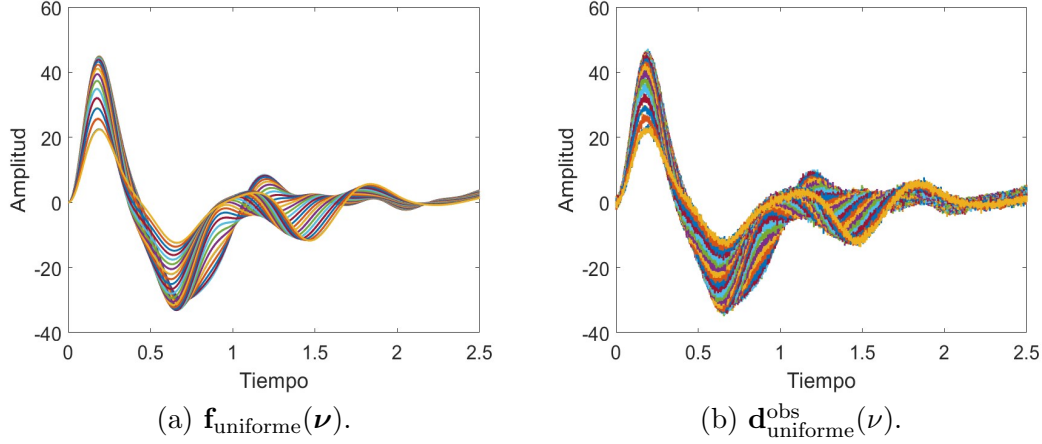


Figura 4.10: Datos obtenidos obtenido por los 51 receptores usando mallado uniforme.

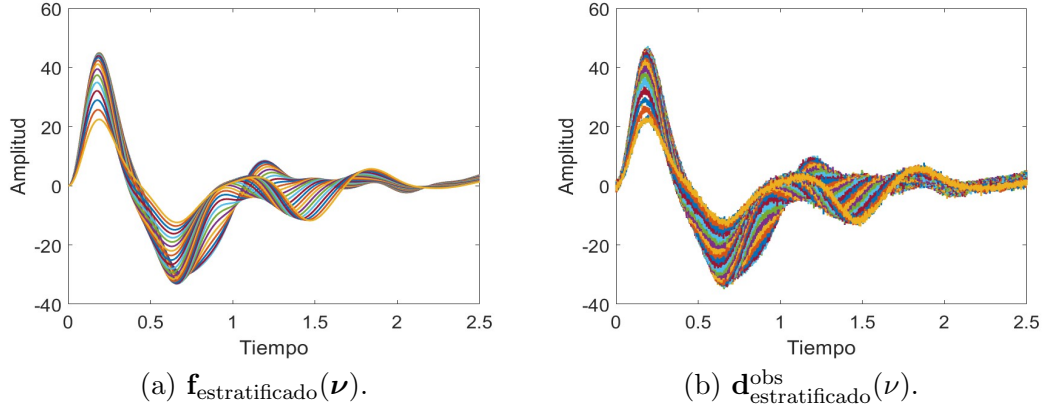


Figura 4.11: Datos obtenidos obtenido por los 51 receptores usando mallado estratificado.

4.2. Problema bayesiano inverso aproximado

Una vez definidos los operadores de observación del problema directo $\mathbf{f}_{\text{uniforme}}(\boldsymbol{\nu})$ y $\mathbf{f}_{\text{estratificado}}(\boldsymbol{\nu})$, y datos sintéticos $\mathbf{d}_{\text{uniforme}}^{\text{obs}}(\boldsymbol{\nu})$ y $\mathbf{d}_{\text{estratificado}}^{\text{obs}}(\boldsymbol{\nu})$, es posible aplicar el método inverso bayesiano para determinar los parámetros de $\boldsymbol{\nu}$ que caracterizan a la elipse y que maximizan la probabilidad posterior.

Con las mediciones $\mathbf{f}(\boldsymbol{\nu})$ y el operador de observación \mathbf{d}^{obs} se construye la primera parte de la función de costo que corresponde a la distancia Euclídea entre $\mathbf{f}(\boldsymbol{\nu})$ y \mathbf{d}^{obs} , que corresponde a un problema inverso determinista, mientras que la segunda parte de la función de costo en el problema bayesiano inverso asume un papel regularizador, la cual se construye con valores los iniciales de la anomalía $\boldsymbol{\nu}_0$, los valores actualizados en cada iteración $\boldsymbol{\nu}$ y la covarianza inicial $\boldsymbol{\Gamma}_{\text{prior}}$:

$$J_{\text{reg}}(\boldsymbol{\nu}) = \frac{1}{2\sigma_{\text{ruido}}^2} \sum_{j=1}^J \sum_{m=1}^M \left| \mathbf{f}_{m,j}(\boldsymbol{\nu}) - \mathbf{d}_{m,j}^{\text{obs}} \right|^2 + \frac{\lambda}{2} (\boldsymbol{\nu} - \boldsymbol{\nu}_0)^T \boldsymbol{\Gamma}_{\text{prior}}^{-1} (\boldsymbol{\nu} - \boldsymbol{\nu}_0), \quad (4.5)$$

donde $\lambda = 1$ en el coste objetivo J , mientras que durante la optimización es un parámetro regularizador $\lambda > 0$ que se introduce para balancear términos inicialmente y va evolucionando hasta llegar a $\lambda = 1$. $\boldsymbol{\Gamma}_{\text{prior}}$ es una matriz diagonal que contiene la varianza de cada parámetro $\boldsymbol{\nu}_0$.

4.2.1. Construcción anomalía a priori

En la formulación inversa bayesiana se genera una densidad de probabilidad a priori $p_{\text{prior}}(\boldsymbol{\nu})$, por lo que se requiere un conocimiento previo sobre los parámetros desconocidos $\boldsymbol{\nu}_0$. Dichos valores a priori de $\boldsymbol{\nu}_0$ pueden influir significativamente los resultados de la inferencia bayesiana inversa.

El parámetro inicial $\boldsymbol{\nu}_0$ contiene información de la elipse inicial: semiejes mayor y menor, posición del centro de la elipse, ángulo de rotación de la elipse, densidad y velocidad de propagación de onda. El parámetro $\boldsymbol{\nu}_0$ se forma con los siguientes valores iniciales: $\boldsymbol{\nu}_0 = [c_x^0, c_y^0, a^0, b^0, \omega^0, \rho^0, v_p^0]$.

Como se estudió anteriormente, la probabilidad a priori se calcula de la forma

$$p_{\text{prior}}(\boldsymbol{\nu}) = \exp\left(-\frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)\boldsymbol{\Gamma}_{\text{prior}}^{-1}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)\right), \quad (4.6)$$

donde $\boldsymbol{\nu}_0$ es la información a priori sobre los valores medios de los parámetros $\boldsymbol{\nu}$ y $\boldsymbol{\Gamma}_{\text{prior}}$ es una matriz diagonal que contiene la varianza de cada variable del parámetro $\boldsymbol{\nu}_0$. En este caso $\boldsymbol{\Gamma}_{\text{prior}}$ tendrá los siguiente valores [25],

$$\boldsymbol{\Gamma}_{\text{prior}} = \text{diag}(1, 1, 0'5, 0'5, 0'1, \sigma_\rho^2, \sigma_{v_p}^2),$$

donde las desviaciones estándar para la densidad y para la velocidad de propagación de la onda se calcularán como la mitad de la diferencia entre los valores máximos y mínimos en toda la secuencia de capas, es decir, $\sigma_\rho = \frac{\rho_{\text{max}} - \rho_{\text{min}}}{2}$, $\sigma_{v_p} = \frac{v_{p,\text{max}} - v_{p,\text{min}}}{2}$.

A continuación, se estudiarán dos casos en los que se añadirá un ruido porcentual de $\alpha = 5\%$ para generar los datos sintéticos \mathbf{d}^{obs} (mallado uniforme y estratificado) correspondiente a una anomalía. Se usará información a priori ligeramente distinta. En un caso tendrá un ángulo de rotación $\omega_0 = 0^\circ$ y en el segundo caso tendrá un ángulo de rotación $\omega_0 = \frac{\pi}{8}$. En ambos casos tendremos como objetivo encontrar los puntos $\boldsymbol{\nu}_{\text{MAP}}$ que maximizan la función de costo $J(\boldsymbol{\nu})$ con regularización bayesiana.

Para la construcción de la primera aproximación a priori se utilizan valores $\boldsymbol{\nu}_0$ que se pueden observar en la tabla 4.3. En la segunda columna de la tabla se observan los valores de $\boldsymbol{\nu}_{\text{verdadero}}$, es decir, los valores reales del parámetro $\boldsymbol{\nu}$ de la elipse.

Parámetro	$\boldsymbol{\nu}_{\text{verdadero}}$	$\boldsymbol{\nu}_0$
Centro c_x	0'0	0'5
Centro c_y	-1'5	-1'4
Semi-eje mayor a	0'6	0'3
Semi-eje menor b	0'1	0'2
Ángulo de rotación ω	0'0	0'0
Densidad ρ	2'1	2'316
Velocidad v_p	4'4	2'9

Cuadro 4.3: Valores de parámetro $\boldsymbol{\nu}_{\text{verdadero}}$ y $\boldsymbol{\nu}_0$ para el primer ensayo.

Las variables del parámetro inicial $\boldsymbol{\nu}_0$ correspondientes al segundo caso de estudio son los que observan en la tabla 4.4, donde el ángulo de rotación de la elipse es $\frac{\pi}{8}$.

Parámetro	$\nu_{\text{verdadero}}$	ν_0
Centro c_x	0'0	0'5
Centro c_y	-1'5	-1'4
Semi-eje mayor a	0'6	0'3
Semi-eje menor b	0'1	0'2
Ángulo de rotación ω	0'0	$\frac{\pi}{8}$
Densidad ρ	2'1	2'316
Velocidad v_p	4'4	2'9

Cuadro 4.4: Valores de parámetro $\nu_{\text{verdadero}}$ y ν_0 para el segundo ensayo.

4.3. Cálculo del Máximo a posteriori (MAP) por optimización

En esta sección recogemos los resultados obtenidos para ambos ensayos.

4.3.1. Ángulo inicial $\omega_0 = 0$

Se calcula los parámetros ν que maximizan la función de costo usando los operadores $\mathbf{f}_{\text{uniforme}}(\nu)$ y $\mathbf{f}_{\text{estratificado}}(\nu)$, con los datos sintéticos $\mathbf{d}_{\text{uniforme}}^{\text{obs}}$ y $\mathbf{d}_{\text{estratificado}}^{\text{obs}}$ obtenidos, respectivamente, al añadirle ruido Gaussiano de un $\alpha = 5\%$, con el fin de comparar si el mallado influye en los resultados obtenidos a través del método inverso bayesiano.

Para el primer caso de estudio usando un parámetro a priori ν_0 con un ángulo de rotación inicial de la elipse igual a $\omega = 0^\circ$ se obtuvieron los gráficos comparativos entre elipse inicial (color blanca), elipse verdadera (color roja) y elipse final parametrizada con los valores ν_{MAP} (elipse contorno negro), como se puede ver en las figuras 4.12a y 4.12b para las discretizaciones espaciales uniforme y estratificada, respectivamente.

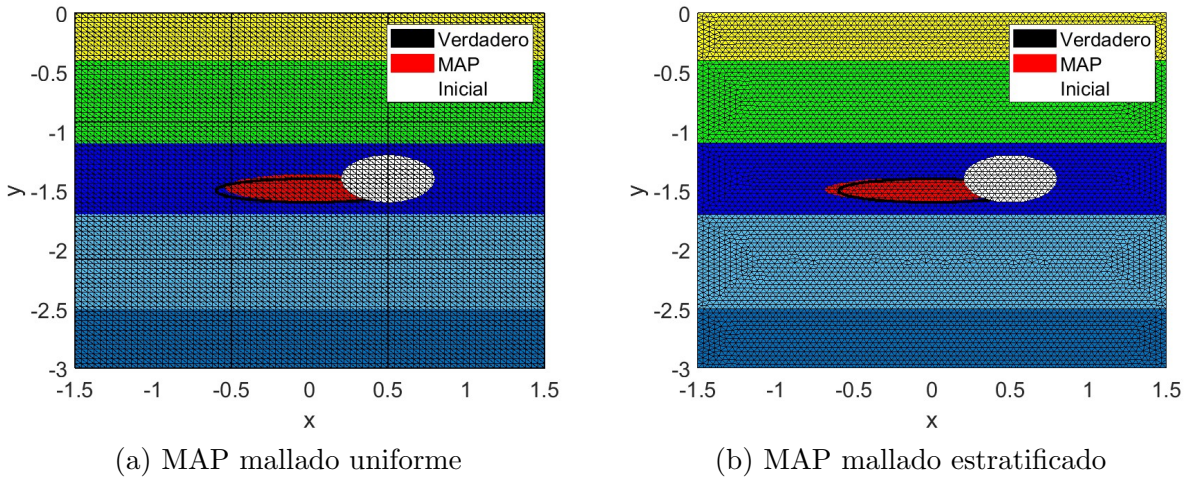


Figura 4.12: Comparación de elipses: inicial (blanca), verdadera (roja) y parametrizada con ν_{MAP} (negra).

Después de 7 iteraciones para las dos discretizaciones, espacial uniforme y estratificada, se obtuvieron los valores de ν_{MAP} o de máximo a posteriori para la elipse, como se observa en el cuadro

4.5, donde los valores de ν_{MAP} obtenidos entre ambas discretizaciones son bastantes similares entre sí.

Parámetro	ν_{MAP} mallado uniforme	ν_{MAP} mallado estratificado
Centro c_x	0'0045	-0'0936
Centro c_y	-1'4766	-1'4920
Semi-eje mayor a	0'5495	0'5978
Semi-eje menor b	0'1160	0'1008
Ángulo de rotación ω	0'0249	0'0297
Densidad ρ	2'6636	2'6158
Velocidad v_p	4'655	4'4189

Cuadro 4.5: Valores obtenidos para ν_{MAP} en la optimización.

En el problema inverso bayesiano, el algoritmo utilizado para hallar los valores ν_{MAP} se detendrá cuando la función de costo $J(\nu)$ haya alcanzado su valor mínimo y se haya estabilizado, es por eso, que para poder comprobar que efectivamente se haya encontrado el mínimo de $J(\nu)$ se graficó el valor de $J(\nu)$ con respecto al número de iteraciones para cada una de las discretizaciones, como se puede observar en las figuras 4.13, donde 4.13a corresponde al gráfico de la función de costo por iteración usando la discretización espacial uniforme, y 4.13b corresponde al gráfico de la función de costo por iteración usando la discretización estratificada.

En la figura 4.13a se observa que el valor de la función de costo disminuye hasta estabilizarse durante la iteración 6 hasta finalizar en la iteración 7. Para el mallado estratificado, la función de costo disminuye de forma casi lineal en cada iteración hasta estabilizarse en la iteración 6 y finalizar en la iteración número 7.

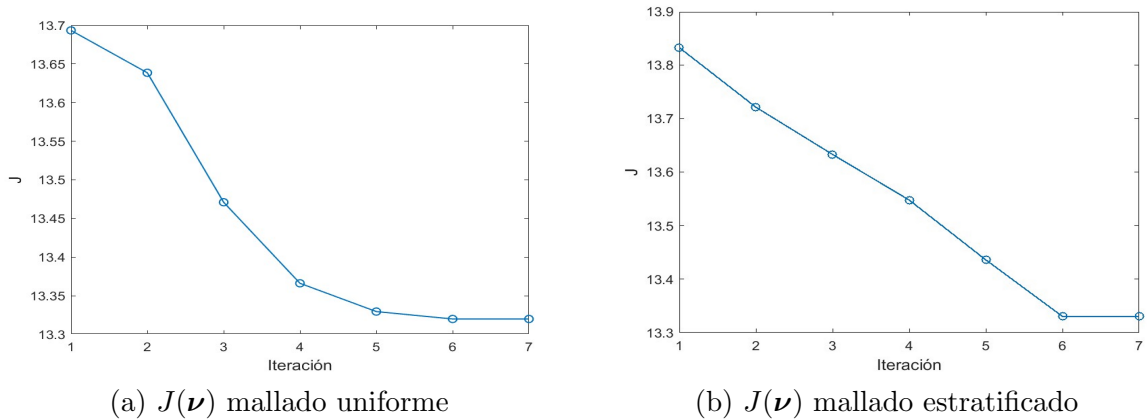


Figura 4.13: Función de costo por iteración.

En este primer caso de estudio, el tiempo de simulación para converger a la estimación MAP de cada variable fue de 13 minutos usando el mallado uniforme, frente a 72 minutos con mallado estratificado adaptado a la geometría de fondo. Esto representa una reducción del 82% en el tiempo computacional al emplear el mallado uniforme, gracias a su estructura de nodos ordenada que optimiza las operaciones matriciales. Si bien el mallado adaptado captura con mayor fidelidad geométrica las capas estratificadas, la disposición irregular de sus nodos introduce un mayor costo computacional sin mejorar significativamente la precisión de los resultados.

4.3.2. Ángulo inicial $\omega_0 = \frac{\pi}{8}$

Para el segundo caso de parametrización inicial ν_0 donde el ángulo de rotación de la elipse inicial es de $\frac{\pi}{8}$ se obtuvo la comparación entre elipse inicial (color blanco), elipse verdadera (color rojo) y elipse máximo a posteriori (contorno color negro), como se observan en las figuras 4.14a y en 4.14b, que corresponden a gráficos comparativos usando la discretización espacial uniforme y discretización espacial estratificada, respectivamente.

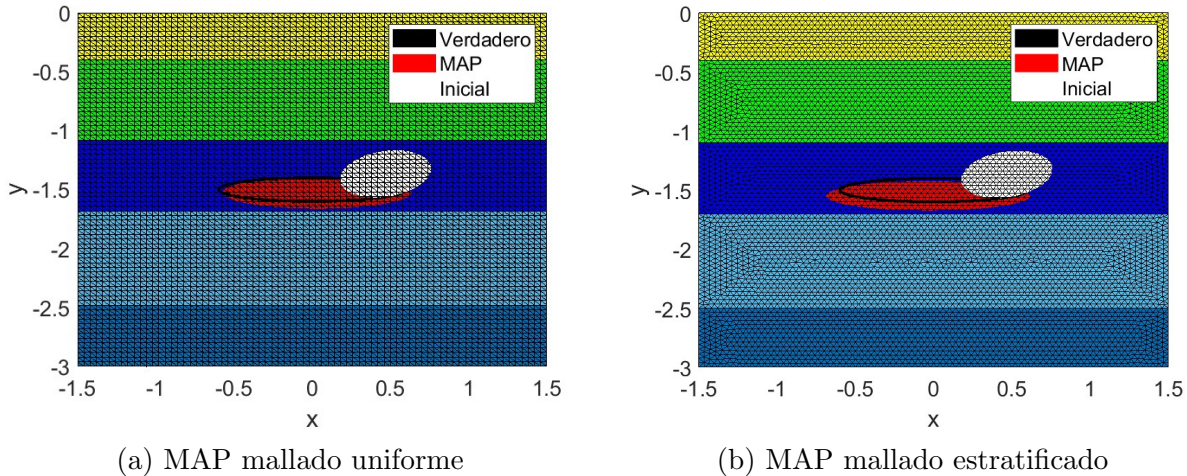


Figura 4.14: Configuración final de la elipse.

Los valores obtenidos del parámetro ν_{MAP} para cada uno de las discretizaciones espaciales se puede observar en el cuadro 4.6.

Parámetro	ν_{MAP} mallado uniforme	ν_{MAP} mallado estratificado
Centro x c_x	0'03069	-0'0349
Centro y c_y	-1'5353	-1'5483
Semi-eje mayor a	0'6079	0'6557
Semi-eje menor b	0'1380	0'1261
Ángulo de rotación ω	-0'0208	-0'0580
Densidad ρ	2'5627	2'6180
Velocidad v_p	3'6211	3'7972

Cuadro 4.6: Valores obtenidos para ν_{MAP} en la optimización.

Como se explicó anteriormente, para hallar los valores de ν_{MAP} la función de coste debe encontrar su valor mínimo, es por esto que para comprobar que así fuera se graficaron los valores de la función de coste $J(\nu_{\text{MAP}})$ por iteración. De las figuras 4.15a y 4.15b es posible observar que usando la discretización espacial uniforme, al problema bayesiano inverso le toma 7 iteraciones hasta encontrar los valores máximo a posteriori, ν_{MAP} , en cambio para la discretización espacial estratificada al problema bayesiano inverso le toma 6 iteraciones para hallar ν_{MAP} . Además, es posible observar de estos gráficos que la función de coste en ambos casos encuentra su valor mínimo.

Para este segundo caso de estudio donde el ángulo de con rotación inicial es de $\frac{\pi}{8}$, el tiempo de cálculo para converger a la estimación MAP de cada parámetro fue de 14 minutos usando mallado uniforme, frente a 107 minutos con mallado estratificado adaptado a la geometría de fondo. Esto

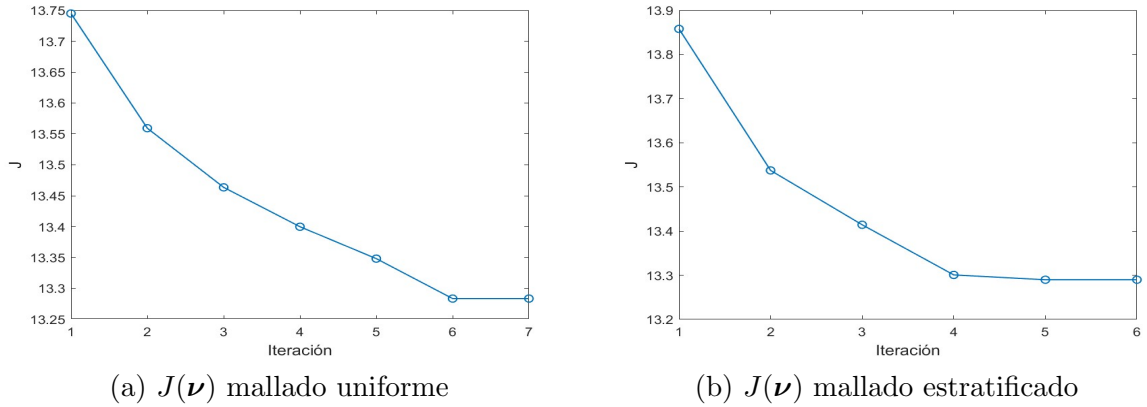


Figura 4.15: Función de costo por iteración.

representa una reducción del 87% en el tiempo computacional al emplear el mallado uniforme. Cabe destacar que, al igual que en el caso anterior, ambas discretizaciones alcanzaron valores estimados similares y cercanos a los reales.

4.4. Resultados del muestreo

Para cuantificar la incertidumbre de cada uno de los parámetros de ν , se aproxima la probabilidad a posteriori como una aproximación de Laplace, por lo que $p_{\text{posterior}} \approx \mathcal{N}(\nu_{\text{MAP}}, \mathbf{H}^{-1})$, según se estudió en el capítulo 3. Para visualizar la distribución de la muestra de cada parámetro de ν se usarán histogramas.

Usando la aproximación de Laplace, se cuantifica la incertidumbre de cada una de las variables del parámetro ν_{MAP} obtenidos anteriormente mediante el método inverso bayesiano. Para cada uno de los parámetros de ν_{MAP} , se emplea la aproximación de Laplace para modelar la distribución posterior como una Gaussiana multivariada. Este proceso se realiza utilizando 10,000 muestras, las cuales se pueden visualizar a través de histogramas que muestran la distribución de los parámetros de ν_{MAP} para ambos mallados numéricos utilizados.

De los histogramas se puede observar, que los valores promedios de cada una de las variables obtenidas de ν_{MAP} son bastantes razonables considerando que están muy cercanos a los valores de $\nu_{\text{verdadero}}$.

A continuación se observan los histogramas para los dos tipos de ensayos.

4.4.1. Ángulo inicial $\omega_0 = 0$

Los resultados de la distribución de las variables de ν_0 cuando el ángulo inicial de rotación de la elipse es 0° se pueden observar en las siguientes figuras. En 4.16 se grafica la posición del centro de masa de la elipse (c_x, c_y). Se comparan los valores de centro de masa del mallado uniforme con el mallado estratificado, donde el símbolo circular de color rojo representa el valor inicial del centro de masa y el diamante de color verde es el valor MAP que se obtuvo como resultado del método inverso bayesiano, además se graficaron el valor verdadero o real del centro de masa (asterisco de color negro). Es posible observar que el centro de masa MAP usando el mallado uniforme es más

cercano al valor centro de masa verdadero, ya que con el mallado estratificado el valor horizontal del centro de masa c_x no es tan cercano al valor real. Al utilizar un mallado uniforme como observamos en la figura 4.16, el valor MAP del centro de masa está muy próximo al valor real. En cambio, en el mallado estratificado, aunque también se alcanza una buena estimación, se produce un pequeño desplazamiento en la componente horizontal c_x . Esto podría deberse a una interacción más compleja entre la geometría de la anomalía y la distribución no uniforme de los elementos en dicho mallado, lo cual introduce cierta asimetría en la sensibilidad espacial.

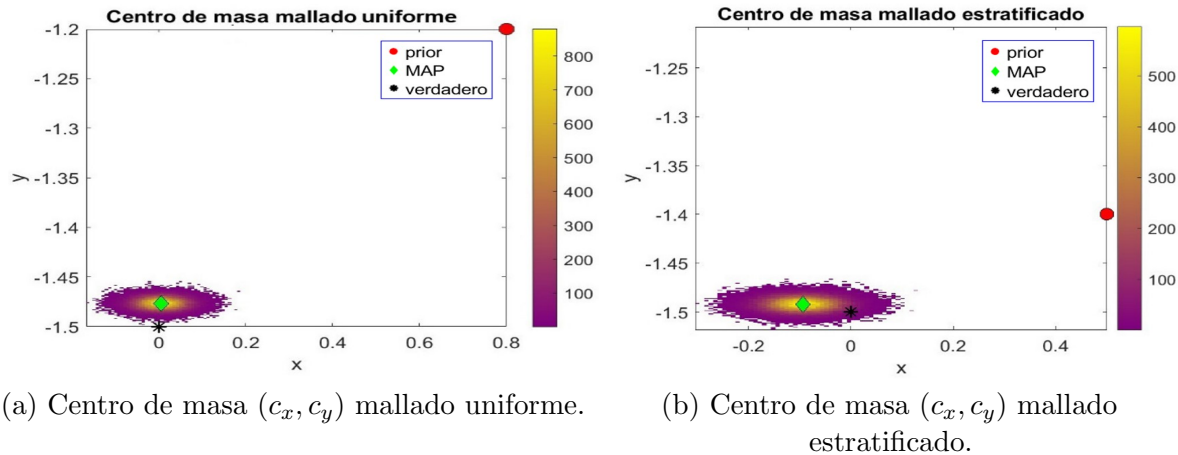


Figura 4.16: Posición centro de masa de la elipse para ambos mallados.

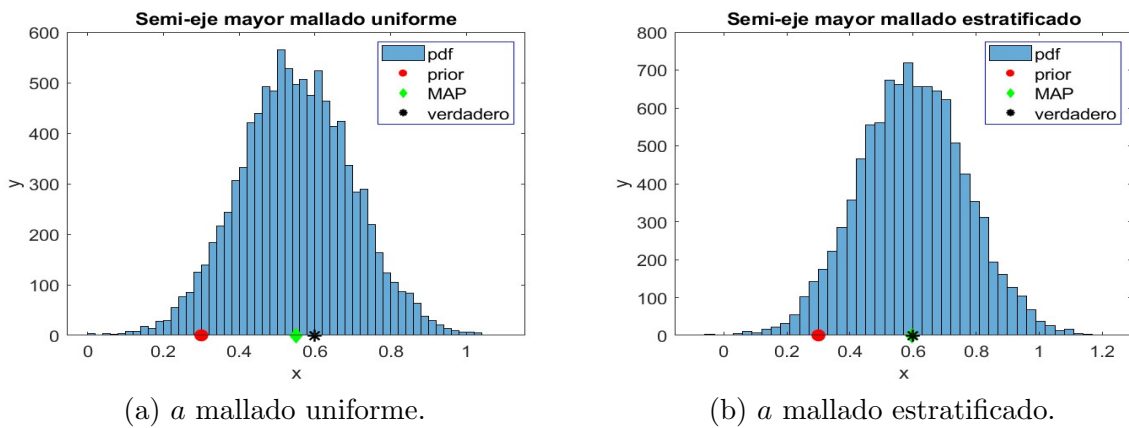
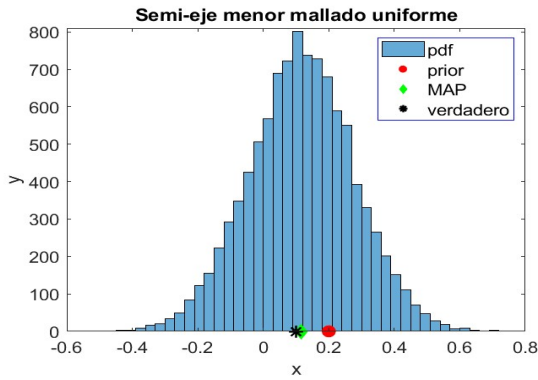


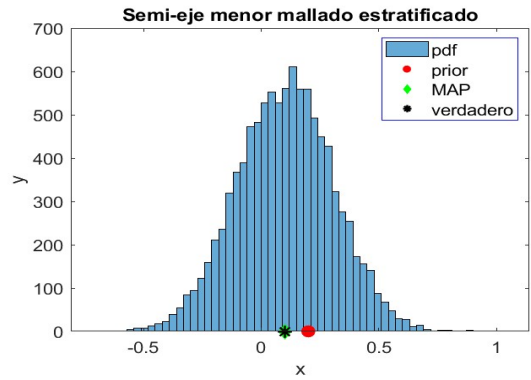
Figura 4.17: Histogramas longitud semi-ejes mayor de elipse.

En la figura 4.17 se puede observar una distribución Gaussiana del semieje mayor de la elipse, donde el valor MAP está representado con el diamante de color verde y el valor verdadero con un asterisco de color negro. Es posible observar que usando un mallado estratificado el valor a_{MAP} es más cercano al valor del semieje mayor verdadero, que si se utiliza el mallado uniforme, ya que observamos una ligera mejora en su estimación, lo que sugiere que una mayor refinación espacial localizada, permite capturar mejor la extensión del cuerpo en la dirección del semieje mayor.

En el histograma de la figura 4.18 se puede observar la distribución del semieje menor de la elipse, para las dos discretizaciones espaciales, donde en ambos casos el valor de MAP de b se encuentra muy cercano al valor verdadero del semieje menor.

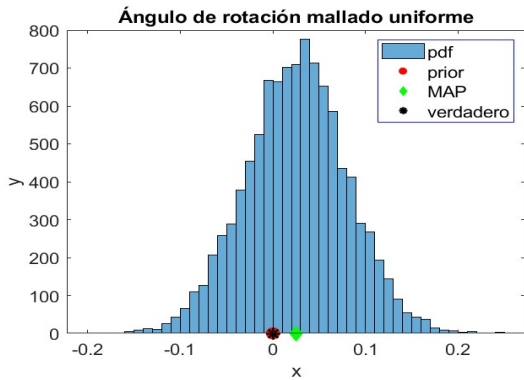


(a) b mallado uniforme.

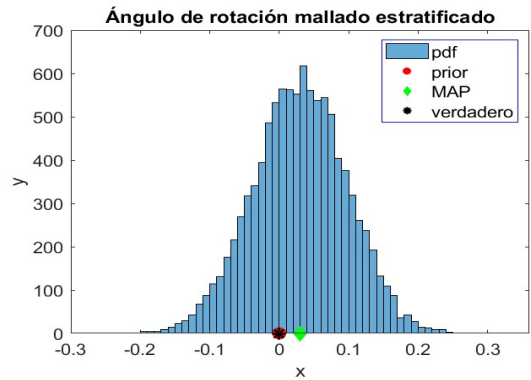


(b) b mallado estratificado.

Figura 4.18: Histogramas longitud semi-eje menor de elipse.

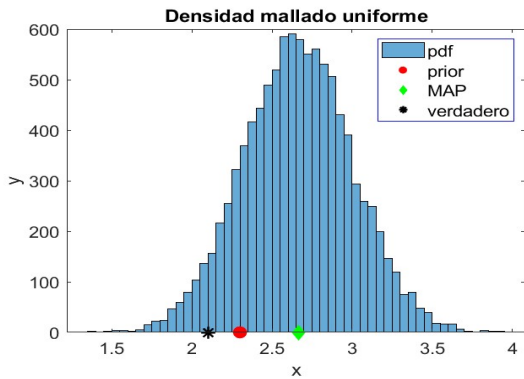


(a) ω mallado uniforme.

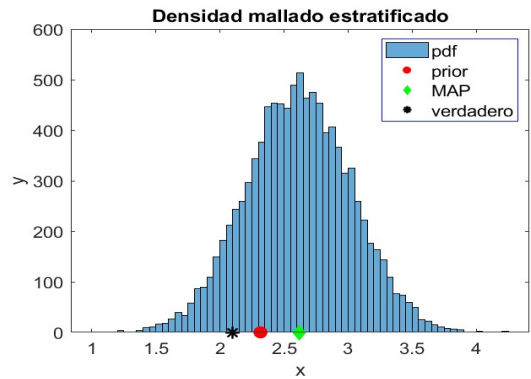


(b) ω mallado estratificado.

Figura 4.19: Histogramas de ángulo de rotación de la elipse.



(a) ρ mallado uniforme.



(b) ρ mallado estratificado.

Figura 4.20: Histogramas de densidad en la elipse.

En el histograma de la figura 4.19 se observa la distribución del ángulo de rotación ω_{MAP} , donde es posible ver que en ambos casos, es decir, con las dos discretizaciones utilizadas, el valor del ángulo es bastante cercano al valor real de $\omega = 0^\circ$.

En los histogramas de las figuras 4.20 y 4.21 se puede observar la distribución de las variables físicas correspondientes a la densidad ρ y a la velocidad de propagación v_p en la elipse MAP. Es

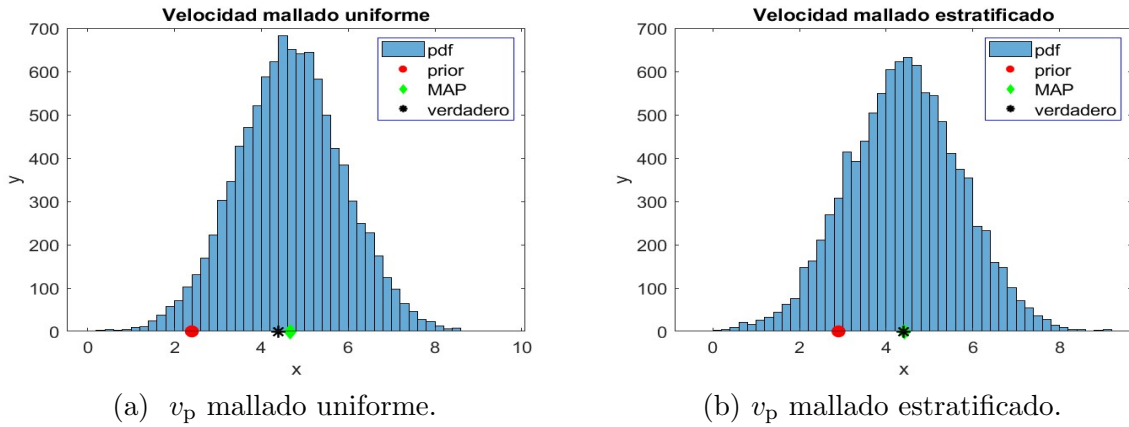


Figura 4.21: Histogramas de velocidad de propagación de la onda en la elipse.

posible observar que si bien los valores de la densidad están dentro del mismo orden de magnitud de los valores verdaderos, estos están un poco alejados del valor real. Sin embargo, no ocurre lo mismo para la velocidad de propagación donde el valor que se obtuvo a través de la optimización es bastante cercano al valor real, en especial usando el mallado estratificado, donde se puede observar que el valor es igual al valor real (difiere a partir del segundo decimal).

4.4.2. Ángulo inicial $\omega_0 = \frac{\pi}{8}$

Para el segundo caso de elipse inicial, es decir, con ángulo de rotación igual a $\omega = \frac{\pi}{8}$ también se cuantificó la incertidumbre a través de la aproximación de Laplace para cada una de las variables de ν_{MAP} . La figura 4.22 grafica la posición del centro de masa (c_x, c_y) , donde el centro de masa MAP se representa con un diamante de color verde y la elipse verdadera como un asterisco de color negro. Además, en el mismo gráfico se ha representado con un círculo de color rojo la posición del centro de masa inicial (prior).

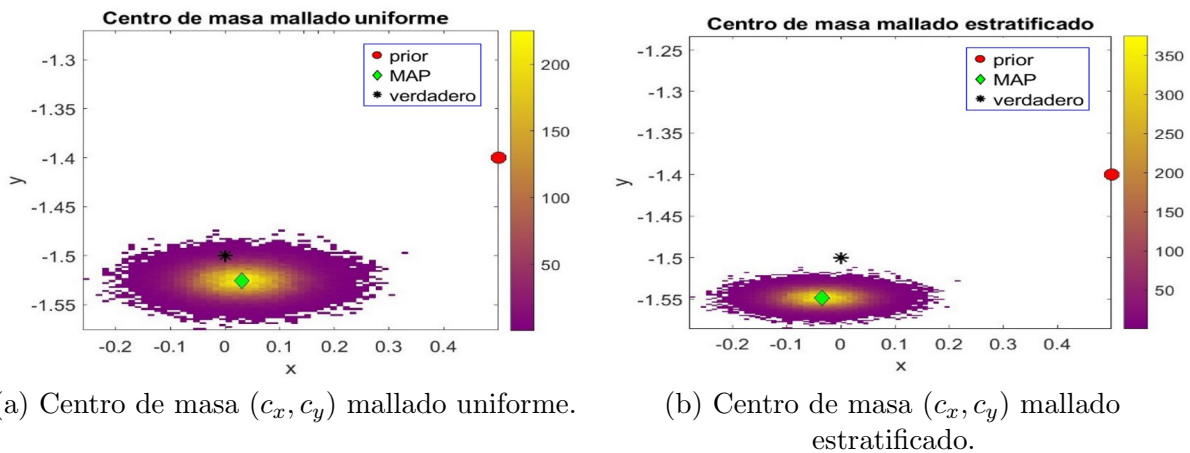
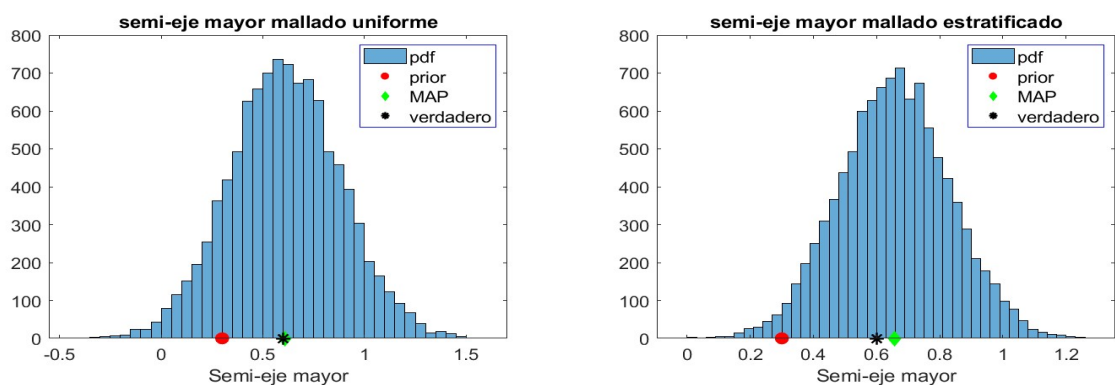


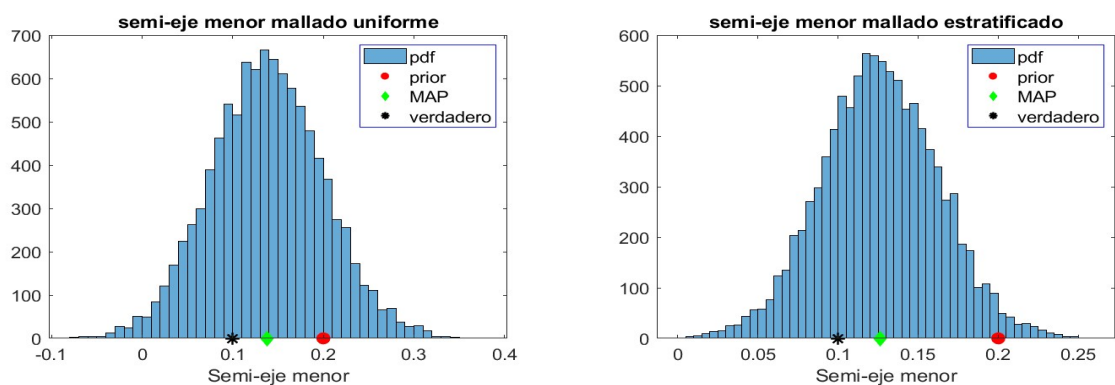
Figura 4.22: Posición centro de masa de la elipse para ambos mallados.



(a) a mallado uniforme.

(b) a mallado estratificado.

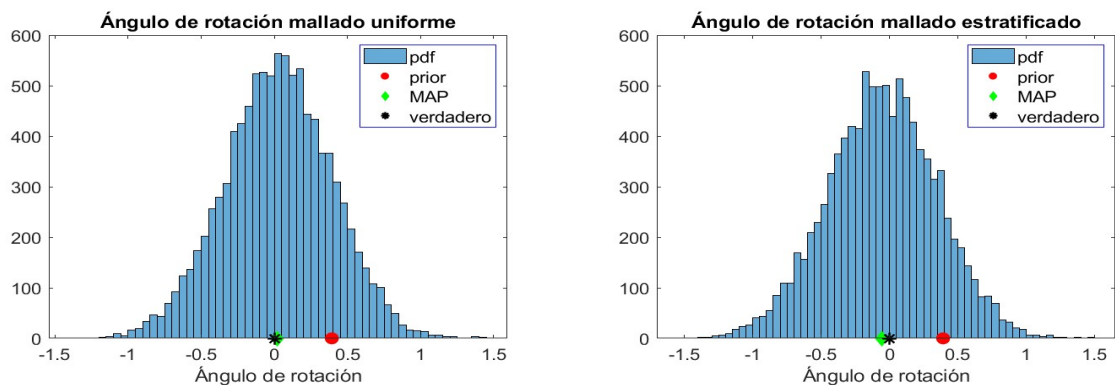
Figura 4.23: Histogramas longitud semi-ejes mayor de elipse.



(a) b mallado uniforme.

(b) b mallado estratificado.

Figura 4.24: Histogramas longitud semi-eje menor de elipse.



(a) ω mallado uniforme.

(b) ω mallado estratificado.

Figura 4.25: Histogramas de ángulo de rotación de la elipse.

En la figura (4.23) se puede observar una distribución Gaussiana del semieje mayor de la elipse a , donde el valor MAP está representado con el diamante de color verde y el valor verdadero con un asterisco de color negro. Es posible observar que usando un mallado uniforme el valor a_{MAP} es más cercano al valor del semieje mayor verdadero, que si se utiliza el mallado estratificado.

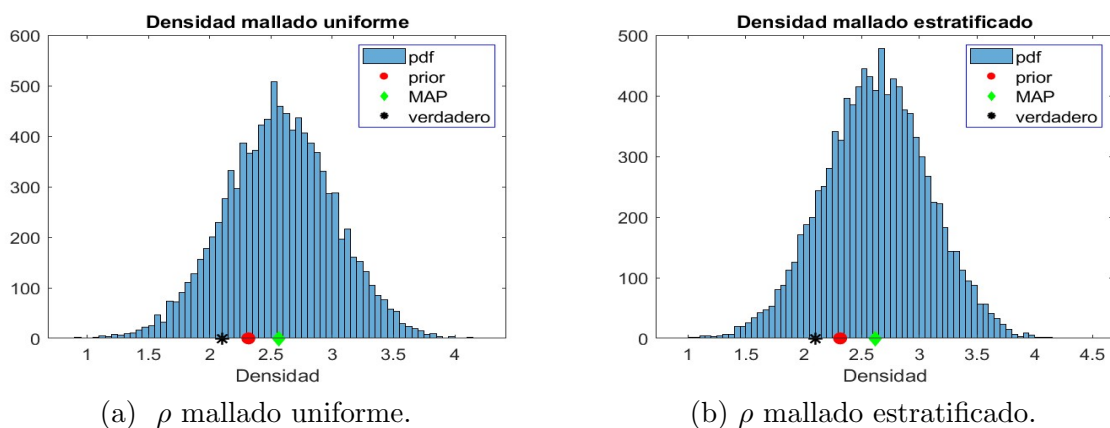


Figura 4.26: Histogramas de densidad en la elipse.

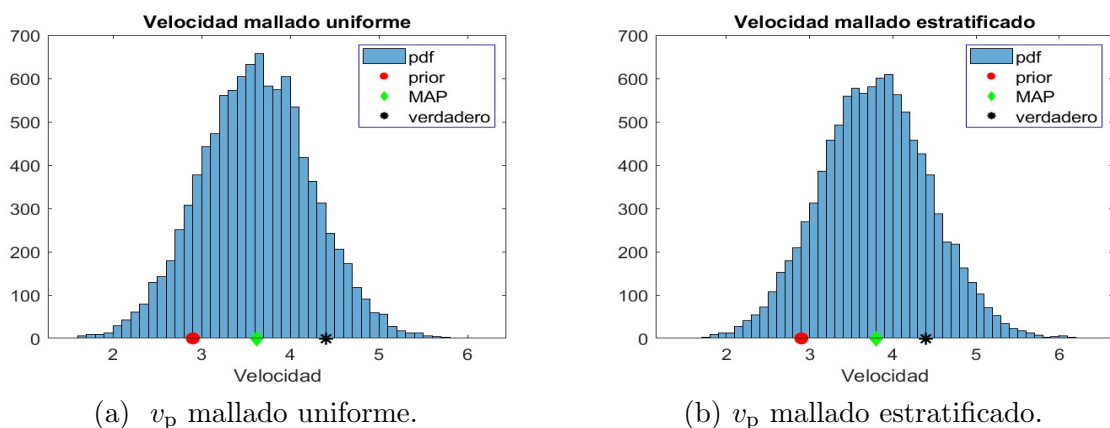


Figura 4.27: Histogramas de velocidad de propagación de la onda en la elipse.

En el histograma de la figura 4.24 se puede observar la distribución del semieje menor de la elipse b , para las dos discretizaciones espaciales. La rotación de la elipse afecta más fuertemente al semieje mayor a que al menor b . Sin embargo, ambos semiejes muestran estabilidad (baja dispersión en los histogramas). Pese a esto, se observa que el semieje menor b presenta un mayor sesgo.

En el histograma de la figura 4.25 se observa la distribución del ángulo de rotación ω_{MAP} , donde es posible ver que en ambos casos el valor MAP es cercano al valor real del ángulo de rotación ($\omega = 0$), sin embargo, si se comparan los histogramas es posible observar que el valor MAP se solapa al valor real del ángulo de rotación cuando se usa un mallado uniforme. Esto se puede interpretar como una ventaja del mallado uniforme cuando se trata de capturar rotaciones no nulas, ya que el mallado estratificado puede introducir anisotropías que afectan la detección precisa de orientaciones oblicuas.

En los histogramas de las figuras 4.26 y 4.27 se puede observar la distribución de las variables físicas de la elipse MAP, como la densidad ρ y la velocidad de propagación v_p . Es posible observar que si bien los valores están dentro del mismo orden de magnitud de los valores verdaderos, estos no son exactamente los mismos, tanto como para la densidad y para la velocidad de propagación, usando las dos discretizaciones espaciales. Para la variable física de densidad, el valor MAP obtenido da sobre el valor verdadero, en cambio, para la variable física de velocidad, el valor MAP obtenido entrega un valor más bajo al valor real, usando ambos mallados.

4.5. Caso aplicado a elastografía médica

Uno de los principales objetivos de la imagen médica es identificar anomalías o estructuras internas presentes en los tejidos mediante métodos no invasivos, tales como radiografía, resonancia magnética y ultrasonido. Estas técnicas abordan problemas inversos, ya que se intenta obtener las propiedades de los tejidos a partir de señales observadas. Generalmente, se envían ondas hacia el tejido y se registran las ondas resultantes para analizar la estructura del medio. Estas técnicas son consideradas no invasivas con el tejido en estudio ya que utilizan fuentes o emisores como la luz, sonido u ondas elásticas. La elastografía médica es una técnica que usa ondas elásticas como fuente y que permite mapear las propiedades elásticas, como el módulo de elasticidad, de los tejidos blandos, ayudando a diferenciar entre tejido sano y enfermo [24].

La aplicación de la elastografía médica para la detección de tejido no sano utilizando métodos inversos ya ha sido abordada en [24], donde se ha tomado densidades similares en todos los tejidos al ser medios celulares. La velocidad es entonces proporcional a la raíz del módulo elástico de cizalla. Incluimos aquí algún resultado por completitud. En esta subsección, se presentará un caso aplicado de elastografía médica con parámetros tomados de [24] donde se buscan propiedades de velocidad de propagación de onda y de módulo de elasticidad de tejido no sano. Primero se resolverá el problema directo, utilizando la misma geometría y usando sólo la discretización uniforme, ya que el fondo no será estratificado, sino que tendrá un sólo un material. La anomalía, al igual que en los casos anteriores, será una elipse, sin embargo, lo que cambiará serán las propiedades del módulo de elasticidad, los cuales representarán a un tejido no sano.

Parámetro	$\nu_{\text{verdadero}}$
Centro x c_x	0
Centro y c_y	-1'5
Semi-eje mayor a	0'6
Semi-eje menor b	0'1
Ángulo de rotación ω	0'0
Módulo de elasticidad χ	16
Velocidad v_p	4

Cuadro 4.7: Valores verdaderos de ν de la anomalía.

El fondo utilizado será el tejido sano, el cual tendrá propiedades de módulo de elasticidad y de velocidad de propagación que se muestran en la tabla 4.8.

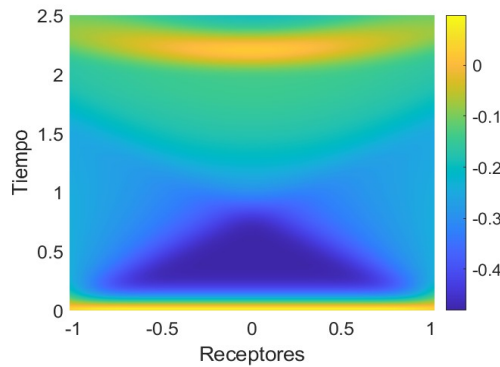
Parámetro	$\nu_{\text{verdadero}}$
Módulo de elasticidad fondo	1'6
Velocidad fondo	1'3

Cuadro 4.8: Valores verdaderos de ν del fondo.

La frecuencia de emisión de la fuente en este caso aplicado será igual a $f_M = 50Hz$.

4.5.1. Generación de datos sintéticos y operador de observación

Como resultado de la resolución de la formulación débil del problema directo se obtuvieron mapas de los desplazamientos de la onda para cada instante de tiempo. En este caso se aplicó la discretización uniforme. En cuanto a la discretización temporal se aplica la condición de estabilidad de Courant-Friedrichs-Lewy, donde en este caso para asegurar la estabilidad de los resultados imponemos $\delta t < 10^{-3}$. Como paso de tiempo total se ha querido iterar 2,500 veces, obteniendo un tiempo total de 2'5 segundos. En la figura se observa el desplazamiento de la onda para la discretización espacial uniforme. Las amplitudes de la onda observadas en los mapas de colores corresponden a los datos medidos $f(\boldsymbol{\nu})$ en el problema directo.



(a) Solución en receptores para problema directo.

Como se mencionó anteriormente, para obtener el operador de observación aproximado \mathbf{d}^{obs} se deben utilizar los datos medidos por los 51 receptores los que se llamarán $\mathbf{f}(\boldsymbol{\nu})$. Como se explicó anteriormente en la ecuación (4.1), a este conjunto de medición se le añadirá ruido Gaussiano para su uso posterior como datos sintéticos en las pruebas de inversión bayesiana. El operador de observación es posible calcularlo usando la expresión (4.1). En este caso se le añadirá un $\alpha = 5\%$ de ruido a $\mathbf{f}(\boldsymbol{\nu})$.

De la resolución numérica del problema directo es posible graficar $\mathbf{f}(\boldsymbol{\nu})$ con respecto al tiempo, para cada uno de los receptores, así como los datos sintéticos generados para una anomalía conocida. En las figuras 4.28a y 4.28b se observan la diferencia de amplitud y forma entre los datos medidos $\mathbf{f}(\boldsymbol{\nu})$ y \mathbf{d}^{obs} , para el tiempo total de simulación de $t = 2'5$ segundos.

4.5.2. Problema bayesiano inverso aproximado

Después de 2 iteraciones utilizando una discretización uniforme se obtuvieron los valores del máximo a posteriori para la elipse que representaría un tejido no sano. A continuación, en la tabla 4.9 se muestran los valores MAP obtenidos para las propiedades de la elipse.

Para comparar las anomalías se realizó el gráfico en conjunto con la elipse inicial (color blanco de la figura 4.29), la elipse verdadera con las propiedades de la tabla 4.7 (elipse de contorno negro en la figura 4.29) y la elipse que se obtuvo con el método bayesiano inverso después de 2 iteraciones (elipse color rojo en la figura (4.29)).

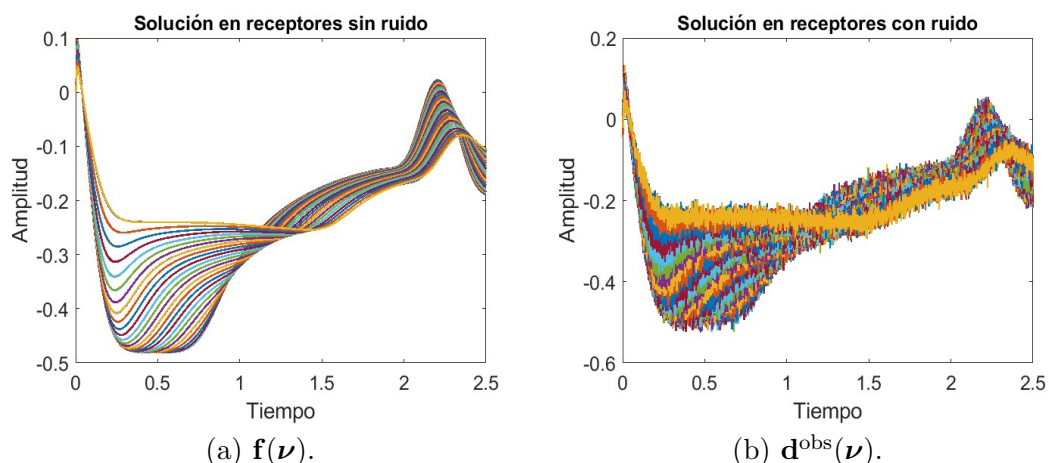


Figura 4.28: Datos obtenidos en los 51 receptores usando mallado uniforme.

Parámetro	$\boldsymbol{\nu}_{\text{MAP}}$ mallado uniforme
Centro x c_x	0'0757
Centro y c_y	-1'4765
Semi-eje mayor a	0'6038
Semi-eje menor b	0'1074
Ángulo de rotación ω	0'0051
Módulo de elasticidad χ	13'8983
Velocidad v_p	4'4001

Cuadro 4.9: Valores verdaderos de $\boldsymbol{\nu}$ de la anomalía.

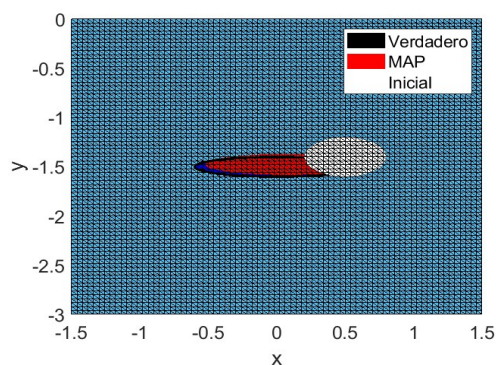


Figura 4.29: Comparación de anomalía verdadera (contorno negro), con anomalía inicial (color blanco) y anomalía MAP (color rojo).

4.5.3. Resultados del muestreo

Para cuantificar la incertidumbre de cada uno de los parámetros de $\boldsymbol{\nu}$, se aproximó la probabilidad posterior utilizando la aproximación de Laplace. Se generaron 10,000 muestras para el muestreo. Para visualizar la distribución de cada variable, se utilizaron histogramas. En las figuras 4.30, se presentan los histogramas correspondientes al centro de masa de la anomalía y al semi-eje mayor de ésta. Del histograma del centro de masa, es posible observar que la posición (c_x, c_y) MAP (diamante color verde) está bastante cerca del valor real (asterisco color negro). En cuanto al histograma del

semi-eje mayor, se puede ver que el valor MAP se solapa con el valor real del semi-eje mayor, lo que indica una buena aproximación.

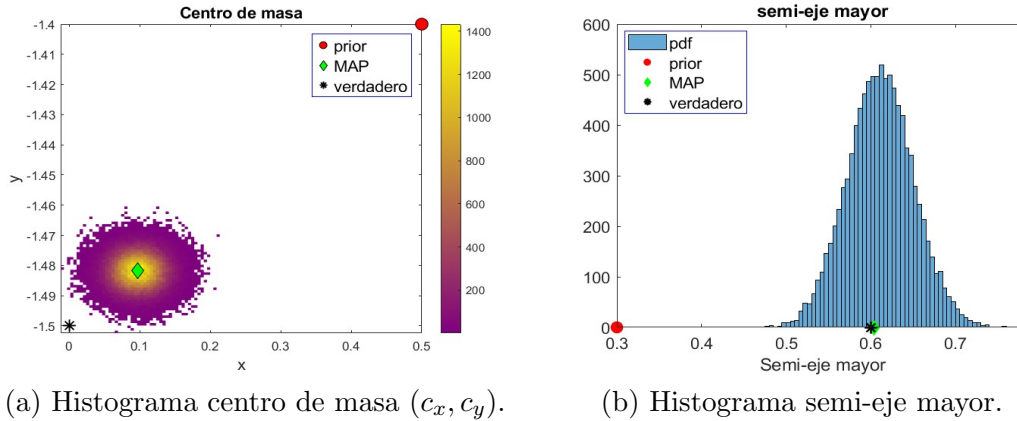


Figura 4.30: Histogramas de centro de masa y de semi-eje mayor de la anomalía.

En el histograma 4.31a se muestra la distribución del semieje mayor de la elipse obtenida a través del método inverso bayesiano, máximo a posteriori (MAP, diamante color verde), comparándola con el valor inicial (círculo rojo) y el valor verdadero (asterisco negro). Es posible observar que el valor MAP se solapa con el valor verdadero del semieje mayor de la elipse. En la figura 4.31b se muestra la distribución del ángulo de rotación de la anomalía, donde se observa que el valor MAP está cercano al valor real del ángulo.

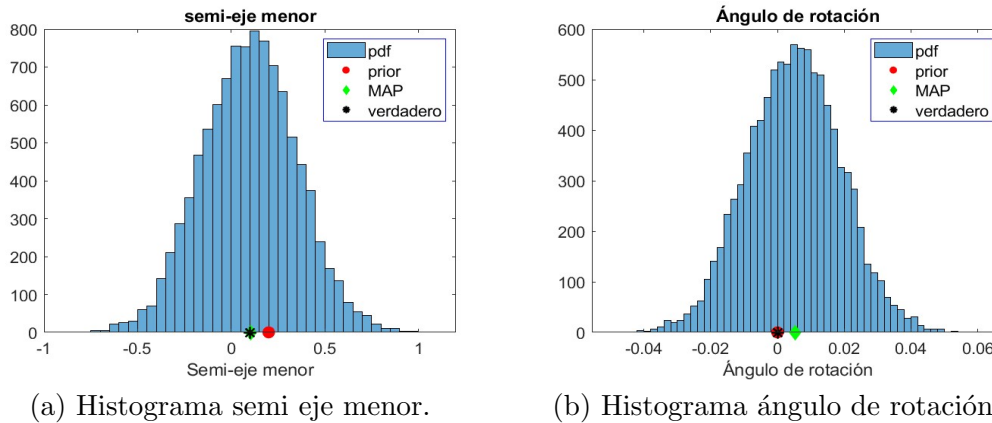
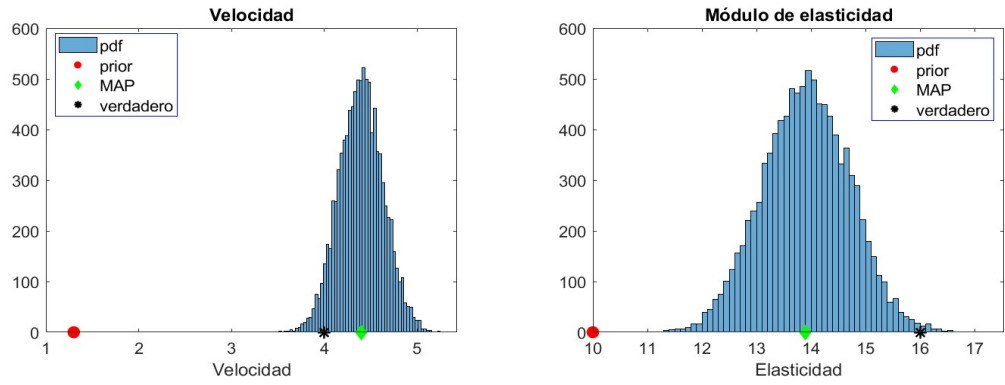


Figura 4.31: Histogramas de semi-eje menor de la anomalía y de ángulo de rotación de la anomalía.

Por último, obtuvimos los histogramas correspondientes a los parámetros físicos: el módulo de elasticidad y la velocidad de propagación de la onda en la anomalía. En el histograma de la figura 4.32a, se observa la distribución de las muestras correspondientes a la velocidad en la anomalía. El valor de velocidad MAP (diamante verde) es cercano al valor de la velocidad real (asterisco negro), lo que indica que la estimación es precisa. Además, la forma del histograma sugiere la dispersión de las muestras alrededor del valor MAP, proporcionando información sobre la incertidumbre de la estimación.

En la figura 4.32b, se muestra la distribución del módulo de elasticidad en la anomalía. Se puede observar que el valor MAP obtenido es del mismo orden que el valor real, y que ambos se encuentran

cercanos dentro del histograma. Esto sugiere una buena concordancia entre la estimación y el valor real, y la forma del histograma revela la variabilidad de las muestras y la precisión de la estimación. No obstante, se aprecia una ligera dispersión en torno a la estimación MAP, lo cual sugiere una mayor incertidumbre asociada a estos parámetros en comparación con los parámetros geométricos.



(a) Histograma de velocidad.

(b) Histograma de módulo de elasticidad.

Figura 4.32: Velocidad y módulo de elasticidad de la anomalía.

Capítulo 5

Problema bayesiano inverso basado en series de Fourier

En las secciones anteriores hemos considerado el problema inverso en geometrías sencillas, en las que se busca inclusiones cuya forma y naturaleza están caracterizadas por un pequeño número de parámetros, insertadas en un medio con estructura conocida. En la práctica, el número de parámetros necesarios para caracterizar una inclusión puede incrementarse, llegando incluso a requerir caracterizaciones infinito-dimensionales de los campos de velocidad y densidad del medio para definirla.

En esta sección vamos a considerar un problema de relevancia práctica: la caracterización de posibles yacimientos de gas y/o petróleo en el entorno de un domo salino. Los domos salinos son configuraciones geológicas en cuyo entorno es frecuente encontrar este tipo de yacimientos, pero su detección con técnicas de imagen estándar se ve dificultada por la fuerte disminución de señal que causan las capas de sal. Vamos a suponer que disponemos de una estimación de la estructura del medio en torno al domo salino (es decir de los campos de densidad y velocidad) que es poco precisa en la zona del domo y no refleja la posible yacimientos de gas/petróleo. A partir de los datos medidos disponibles, se trata de corregir esta ‘información a priori’ mejorando la caracterización de los campos de densidad y velocidad en torno al domo para determinar si refleja la presencia de yacimientos de esos materiales.

En las secciones anteriores, los parámetros que caracterizaban las inclusiones eran parámetros que definían la geometría de una elipse así como valores constantes de la densidad y velocidad en ella. En esta sección, la estructura del medio se caracteriza a partir de las variaciones espaciales de los campos de densidad y velocidad. En principio, estos campos son infinito dimensionales. En esta sección, representaremos los campos de interés mediante series de Fourier y los caracterizaremos por un número finito (moderado) de coeficientes. Tenemos de nuevo un problema bayesiano inverso definido en términos de un número finito de parámetros.

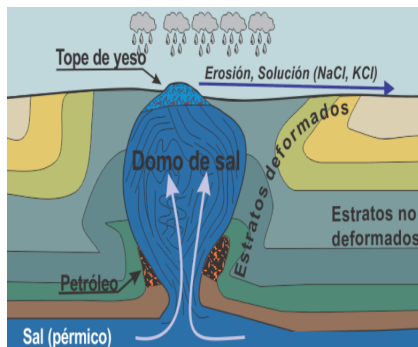
La sección está estructurada como sigue. En primer lugar introducimos la geometría de domo salino y abordamos la resolución del problema directo en ella, así como la definición de un operador de observación aproximado y la generación de datos sintéticos. A continuación, abordamos la caracterización de variaciones locales en los campos de densidad y velocidad mediante series de Fourier truncadas. Tras ello, formulamos el problema bayesiano inverso. Finalmente, el estudio de la probabilidad a posteriori mediante técnicas Monte Carlo conduce a configuraciones de máxima probabilidad que reflejan la presencia de gas/petróleo en los bordes inferiores del domo salino considerado, cuantificando la incertidumbre en esta predicción.

5.1. Geometría de domo salino

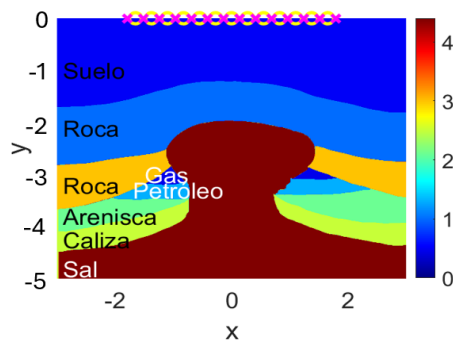
La geometría utilizada en esta sección consiste en un fondo estratificado con una estructura de domo salino o diapiro. Se trata de una formación geológica en forma de cúpula creada por movimientos ascendentes del depósito salino debido a su menor densidad respecto a otros materiales y minerales [3]. El estudio de los domos salinos es crucial, debido a la capacidad de esta estructura para crear trampas de petróleo y gas, lo que los convierte en una característica geológica de gran importancia en la exploración petrolera (ver figura 5.1a).

El dominio computacional truncado lo definimos en el recinto $[-5, 0] \times [-3, 3]$. En este recinto, el fondo estratificado está compuesto por 6 capas, incluyendo un domo salino, que se genera con una parametrización específica. Tanto el domo salino como cada una de las capas de fondo están caracterizados por sus propiedades de densidad $\rho(\mathbf{x})$ y de velocidad de propagación de onda $v_p(\mathbf{x})$. El módulo de elasticidad $\chi(\mathbf{x})$ también varía en cada una de ellas, debido a la relación $\chi(\mathbf{x}) = \rho(\mathbf{x})v_p^2(\mathbf{x})$ [24]. La geometría utilizada en este estudio se muestra en la figura 5.1b. Los valores correspondientes a la densidad, velocidad de propagación de la onda y módulo de elasticidad del fondo, incluyendo a la estructura de domo salino, se muestran en la tabla 5.1 [23].

Como anomalías, en este caso, se incorporarán dos zonas localizadas a los costados del domo salino, que representan trampas de gas y de petróleo. Su forma se construye a través de parametrizaciones específicas. Los valores de los parámetros de densidad, velocidad de propagación de onda y de módulo de elasticidad de los objetos se presentan en la tabla 5.2 [23].



(a) Imagen de domo salino [3].



(b) Domo salino con yacimientos laterales.

Figura 5.1: Geometría de domo salino de fondo con yacimientos laterales utilizada en las simulaciones.

Los valores de densidad $\rho(\mathbf{x})$ y de velocidad de onda $v_p(\mathbf{x})$ de cada estrato se muestran en el cuadro 5.1, una vez adimensionalizados. La unidades son: longitud km , tiempo s , velocidad km/s , densidad kg/1t .

Capa	Densidad $\rho(\mathbf{x})$	Velocidad de onda $v_p(\mathbf{x})$	Módulo de elasticidad $\chi(\mathbf{x})$
Suelo	1'1	0'5	0'27
Capa de roca	2'50	3 - 2 - 1	22'5 - 1' - 2'5
Arena (Reservorio)	2'0	2'1	8'82
Caliza (fuente)	2'7	2'5	16'87
Capa de sal	2'10	4'4	40'65

Cuadro 5.1: Propiedades físicas adimensionales para cada capa del fondo del domo salino.

Al igual que en la sección anterior, las anomalías que representan en este caso yacimientos tendrán propiedades físicas específicas, como podemos ver en el cuadro 5.2.

Objeto	Densidad $\rho(\mathbf{x})$	velocidad de onda $v_p(\mathbf{x})$	módulo de elasticidad $\chi(\mathbf{x})$
Petróleo	0'90	1'3	1'52
Gas	0'7	0'45	0'14

Cuadro 5.2: Propiedades físicas adimensionales de los objetos del costado del domo salino.

Los mapas de la densidad y velocidad correspondientes a la geometría de domo salino de la figura 5.1b) se visualizan en las figuras 5.2a y en 5.2b. Para obtener el mapa del módulo de elasticidad por estratos se usa la relación $\chi(\mathbf{x}) = \lambda(\mathbf{x}) + 2\mu(\mathbf{x}) = \rho(\mathbf{x})v_p^2(\mathbf{x})$. El mapa de módulo de elasticidad lo podemos observar en la figura 5.3.

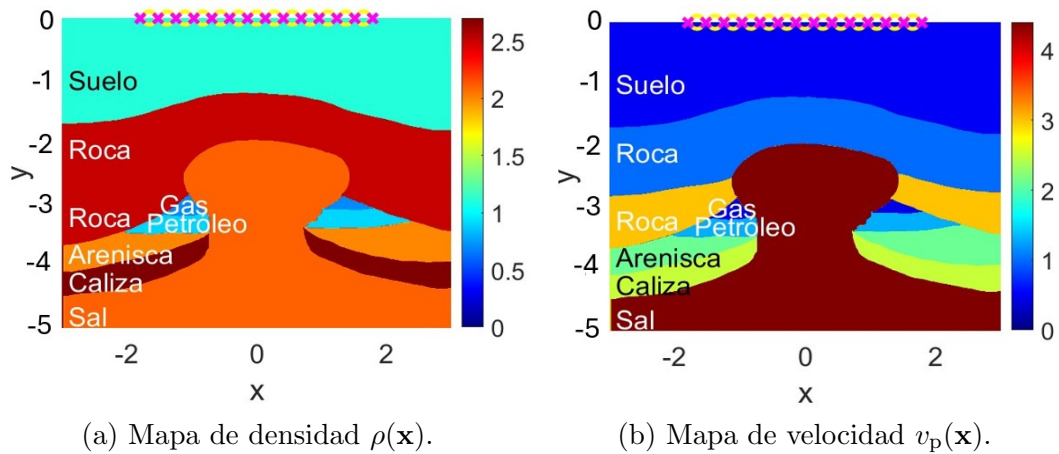


Figura 5.2: Mapa de densidad y velocidad de domo salino incorporando objetos al costado.

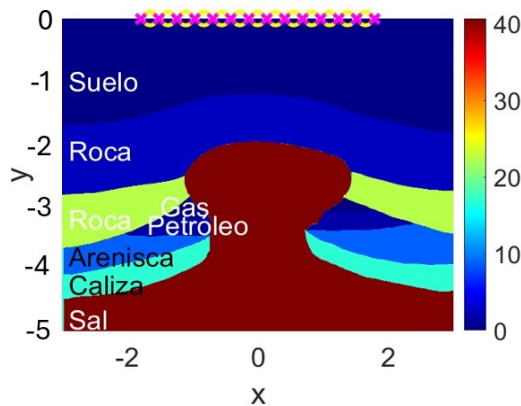


Figura 5.3: Mapa módulo de elasticidad $\chi(\mathbf{x})$.

En los estudios geofísicos del terreno que permiten identificar la estructura bajo la superficie se emiten ondas elásticas desde fuentes situadas a una cierta distancia (según la resolución esperada) y se recoge la señal recibida en la superficie tras la interacción con el medio, en una red de receptores.

En nuestras simulaciones, las ondas elásticas se emiten desde un conjunto de fuentes \mathbf{x}_k equiespaciadas, con $k = 0, \dots, 14$, que se posicionan en la superficie (la interfaz entre aire y suelo), es decir en la posición vertical $y = 0$. El barrido horizontal de emisores varía desde desde la coordenada $(-1'76, 0)$ hasta $(1'76, 0)$. En total, utilizamos $K = 14$ emisores.

Para definir el operador de observación correspondiente al desplazamiento de la onda en diferentes posiciones e instantes de tiempo, colocamos receptores \mathbf{r}_j equiespaciados, con $j = 0, \dots, 13$, de forma intercalada con los emisores. Estos receptores se ubican desde la posición $(-1'68, 0)$ hasta $(1'68, 0)$. En total, utilizamos $J = 13$ receptores. En la vida real, se miden estas señales con aparatos adecuados, definiendo de esta forma el conjunto de datos reales medidos. En nuestras simulaciones utilizamos datos sintéticos generados resolviendo el problema directo asociados perturbados con ruido de magnitud controlada.

Para la resolución del problema directo se deben conocer las características de las fuentes. En el estudio presentado aquí, la variación temporal de la onda se describe mediante la expresión $f(t) = f_0(1 - 2\pi^2 f_M^2 t^2) e^{-\pi^2 f_M^2 t^2}$, donde la amplitud inicial es de $f_0 = 100$ y la frecuencia media de la onda de $f_M = 2$. La discretización temporal utilizada es $t = N\delta t$, con $\delta t = 1 \times 10^{-3}$, mientras que la dependencia espacial de la onda viene dada por $g(\mathbf{x}) = \frac{1}{(\pi\kappa)^2} \sum_{k=1}^K e^{-|\mathbf{x}-\mathbf{x}_k|^2}$, que depende de la posición de las fuentes \mathbf{x}_k y del parámetro de regularización $\kappa = 0'4$.

5.2. Resolución del problema directo y operadores de observación aproximados

Como hemos comentado, para formular un problema inverso bayesiano es necesario ser capaz de evaluar un operador inverso aproximado $\mathbf{f}(\boldsymbol{\nu})$ y disponer de datos \mathbf{d}^{obs} correspondientes a la configuración en estudio, reales o sintéticos. En esta sección nos centraremos en resolver el problema directo gobernado por la ecuación de ondas, utilizando una geometría estratificada con un domo salino y anomalías, que en este caso serán trampas de petróleo y gas. Cada capa de un material distinto se caracterizará a través de su densidad y velocidad de propagación de onda. Desarrollando a nivel global los campos de densidad y velocidad mediante series de Fourier y truncando estas series a una colección finita de términos, los coeficientes $\boldsymbol{\nu}$ permiten caracterizar los campos a determinar mediante una colección finita de parámetros. Resolviendo numéricamente el problema directo para campos de densidad y velocidad definidos mediante parámetros $\boldsymbol{\nu}$, obtendremos el operador de observación aproximado $\mathbf{f}(\boldsymbol{\nu})$. Sean $\boldsymbol{\nu}_{\text{exacto}}$ los parámetros que definen una serie de Fourier truncada que aproxima los campos de densidad y velocidad reales con un error prefijado. Los datos sintéticos \mathbf{d}^{obs} se obtienen resolviendo el problema directo para los parámetros $\boldsymbol{\nu}_{\text{exacto}}$ utilizando la relación (5.1):

$$\mathbf{d}^{\text{obs}} = \mathbf{f}(\boldsymbol{\nu}_{\text{exacto}}) + \boldsymbol{\eta} = \mathbf{d}_{\text{exacto}} + \boldsymbol{\eta}, \quad (5.1)$$

donde el término de ruido $\boldsymbol{\eta}$ tiene componentes de la forma ϵz , con $\epsilon = \frac{\alpha}{100} \sqrt{\frac{\sum_{j=1}^J \sum_{m=1}^M |d_{j,\text{exacto}}^m|^2}{JM}}$, M es el número de mediciones temporales por receptor, siendo, en este caso, $M = 10,000$, α es el nivel porcentual de ruido seleccionado (5%), y z representa un número aleatorio generado con una distribución normal $\mathcal{N}(0, 1)$.

Para la resolución numérica del problema directo de la ecuación de ondas, se realizará una discretización en espacio y en tiempo. El espacio se discretizará de dos maneras. La primera será una discretización uniforme en una malla equiespaciada y la segunda se adaptará a la geometría de fondo estratificado con la estructura de domo salino. Por lo tanto, se construirán dos mallados numéricos diferentes sobre los que definir elementos finitos, lo que resultará al resolver numéricamente en dos conjuntos de mediciones $\mathbf{f}_{\text{uniforme}}(\boldsymbol{\nu})$ y $\mathbf{f}_{\text{adaptado}}(\boldsymbol{\nu})$ correspondientes a cada una de las mallas, obteniendo así dos conjuntos de datos sintéticos \mathbf{d}^{obs} al añadir ruido. En cuanto a la discretización temporal, se considerará la condición Courant-Friedrichs-Lewy para definir el paso de tiempo máximo

que asegure la estabilidad numérica de la solución con un esquema explícito como se ha detallado en capítulos anteriores.

5.2.1. *Discretización espacial*

Para la resolución numérica de la formulación débil del problema directo, como explicamos anteriormente, discretizaremos espacialmente el dominio truncado sobre el que se trabajará computacionalmente. En este caso, el dominio se estratifica y se incorporará la estructura de domo salino como geometría de fondo; además, incluiremos dos anomalías a los costados del domo salino que representarán las trampas de gas y de petróleo. Consideraremos dos tipos de discretizaciones espaciales, con el objetivo de investigar si los resultados dependen de la discretización espacial utilizada. A continuación, se describen cada una de las discretizaciones con sus características.

Discretización espacial uniforme

El primer tipo de discretización espacial que estudiamos es la discretización espacial uniforme. Corresponde a un mallado numérico con elementos de igual tamaño y distribución en todo el dominio computacional truncado, independiente de la geometría de fondo estratificado y de las anomalías.

Para discretizar el dominio de forma numérica utilizando el software MATLAB usaremos la función *delaunayTriangulation*, que genera un mallado uniforme en todo el dominio rectangular truncado. Esto nos proporciona nodos y elementos que incorporamos en la formulación débil del problema directo.

El tamaño de los elementos se configurará para que, tanto horizontal como verticalmente, sean iguales a $\delta_x = \delta_y = 0'04$. El tipo de elemento en esta discretización es triangular, ya que trabajaremos de forma bidimensional. El mallado uniforme utilizado en esta geometría es el que se muestra en la figura 5.4. El número total de elementos en el dominio computacional es de 19,026 y los elementos elegidos son de orden lineal.

Mallado uniforme	Características
Nodos	$2 \times 19,026$
Elementos	$3 \times 37,500$
Orden geométrico	Lineal

Cuadro 5.3: Características del mallado uniforme.

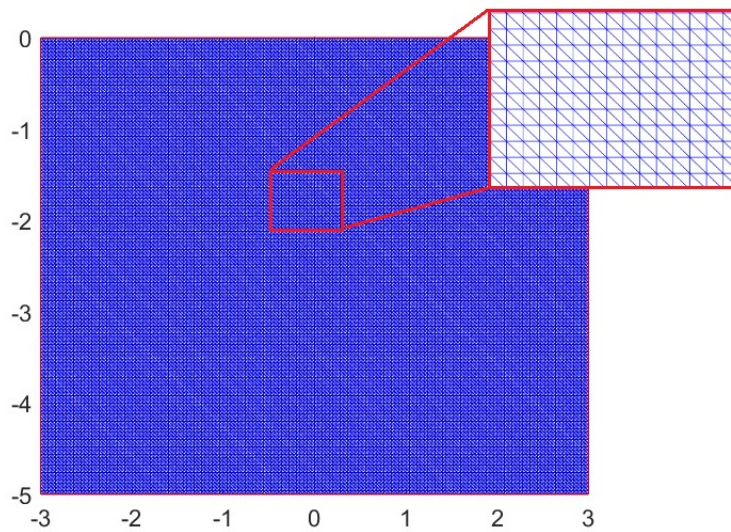


Figura 5.4: Mallado uniforme utilizado para geometría estratificada de fondo con domo salino.

Discretización espacial adaptada a la estructura de domo salino

La segunda discretización espacial estudiada es aquella que se basa en la estructura del domo salino. En este tipo de discretización se malla cada una de las capas de fondo forma independiente, al igual que la estructura de domo salino, por lo que el tamaño y distribución de los elementos dependerá de cada uno de los subdominios de la geometría. Se utiliza un mallado de tipo descomposición de dominio, con nodos comunes en las interfaces y triángulos contenidos en cada subdominio.

Para generar el mallado de cada subdominio se utilizará el software de libre acceso GID, donde primero se deberá dibujar los puntos que delimitan cada subdominio del fondo y del domo salino, para luego unirlos y, a través de diferentes herramientas de GID, dibujar el contorno de cada uno. Luego, con la herramienta *mesh* es posible mallar los subdominios dibujados de forma independiente. A través de los puntos dibujados, que tendrán en común cada una de las interfaces adyacentes, se asegurará la continuidad entre cada subdominio, de forma tal que el dominio y el mallado sean continuos.

Al momento de mallar cada uno de los subdominios se indicará el tamaño máximo y mínimo de los elementos, además del orden geométrico de los elementos del mallado; los elementos se distribuirán de forma aleatoria. En este caso el tamaño máximo de los elementos será igual a 0'04, el tamaño mínimo igual a 0'02 y el orden geométrico es lineal. El número total de elementos será de 6,774. En este caso, el software GID se utilizó como un pre-procesador para dibujar la geometría y luego mallarla; tanto la geometría como los puntos del mallado se pueden extraer para luego importarlos en MATLAB y así incorporar la geometría en la resolución del problema directo e inverso.

En la figura 5.5 se muestra el mallado adaptado a la geometría, pudiéndose visualizar las delimitaciones de cada uno de los subdominios y la estructura de domo salino.

Para la resolución del problema directo usaremos ambas discretizaciones espaciales, con el fin de comparar los operadores de observación y la convergencia de resultados.

Mallado estratificado	Características
Nodos	$2 \times 6,774$
Elementos	$3 \times 13,274$
Orden geométrico	Lineal

Cuadro 5.4: Características de mallado estratificado incorporando la estructura domo salino.

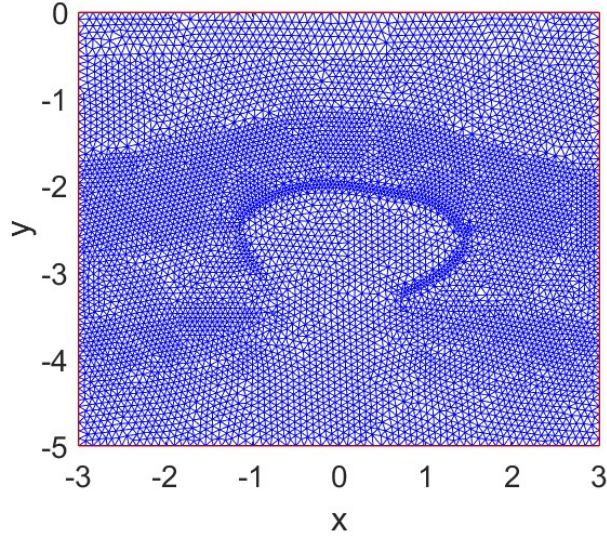


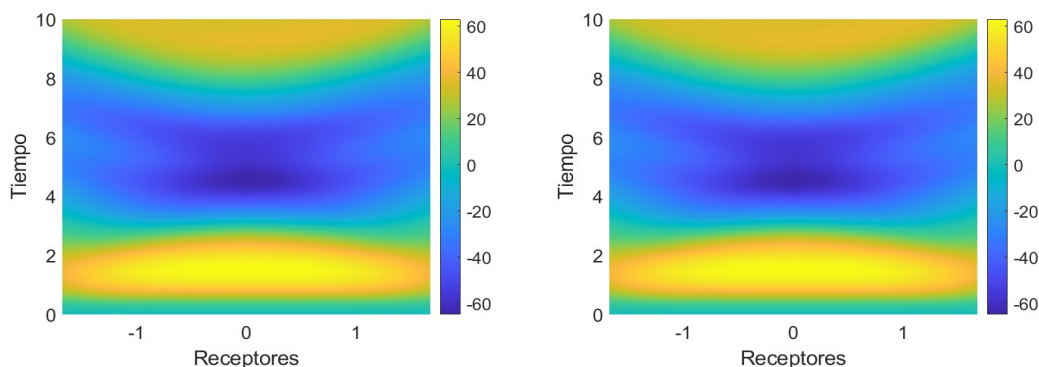
Figura 5.5: Mallado estratificado incorporando la estructura de domo salino.

5.2.2. Comparación de operadores de observación

Como resultado de la resolución del problema directo, obtenemos mapas de los desplazamientos de la onda durante en cada instante de tiempo. Para la resolución, hemos aplicado las discretizaciones espaciales explicadas en la sección 5.2.1. En cuanto a la discretización temporal, se aplicó la condición de estabilidad de Courant-Friedrichs-Lewy, explicada en el capítulo 2, donde para asegurar la estabilidad de los resultados, el paso de tiempo δt debe cumplir con la condición $\delta t < 1 \times 10^{-3}$. Se calcularon 10,000 pasos de tiempo, por lo que el tiempo total en la simulación será de 10 segundos.

En las figuras 5.6a y 5.6b visualizamos el desplazamiento de la onda para la discretización espacial uniforme y para la discretización espacial estratificada incorporando la estructura de domo salino, respectivamente. Las amplitudes de la onda observadas en los mapas de colores corresponden a los datos medidos por cada uno de los receptores, es decir, corresponden al operador de observación aproximado $\mathbf{f}(\boldsymbol{\nu})$ del problema directo, durante el tiempo de medición de 10 segundos. Como observamos en ambas figuras, las amplitudes son similares entre las dos discretizaciones espaciales.

Para cuantificar la diferencia en las mediciones obtenidas usando las dos discretizaciones espaciales, calculamos la diferencia porcentual entre ambos conjuntos de datos, es decir, entre $\mathbf{f}_{\text{uniforme}}(\boldsymbol{\nu})$ y $\mathbf{f}_{\text{adaptado}}(\boldsymbol{\nu})$ para el tiempo de medición $t = 10$ segundos, como visualizamos en la figura 5.7. La diferencia porcentual relativa calculada entre ambos conjuntos de mediciones es de 2'93%. Para calcular la diferencia relativa, primero se obtuvo la diferencia absoluta entre ambos conjuntos de datos. Posteriormente, esta diferencia se dividió por el promedio de los valores absolutos de ambos conjuntos y se multiplica por 100 para así obtener la diferencia porcentual relativa.



(a) Solución en mallado uniforme. (b) Solución en mallado adaptado.

Figura 5.6: Solución en medida en los receptores en función del tiempo.

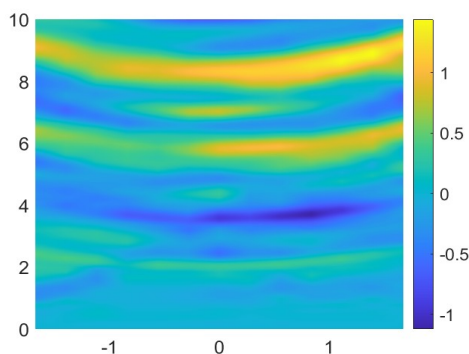


Figura 5.7: Diferencia entre mallados.

Debido a la similaridad en los resultados de los operadores de observación aproximado y a la baja diferencia porcentual entre ambos (2'93%), se concluye que al utilizar diferentes discretizaciones espaciales no observamos grandes diferencias entre los datos observados, y que la formación de asimetrías observadas en el mallado estratificado de GID (ver figura 5.5) no afecta de forma relevante para nuestros propósitos, siempre y cuando el paso temporal y espacial sean suficientemente finos, y similares en ambos. La variación se puede considerar ruido numérico con una cierta estructura. Este hecho es positivo desde el punto de vista de la resolución de problemas inversos bayesianos, en los que se trata de identificar anomalías a partir de colecciones de datos medidos distorsionados con ruido (ya sean reales, con ruido aparejado al dispositivo de medida, o sintéticos, con ruido añadido).

La comparación entre los operadores de observación definidos en distintos mallados adquiere mayor relevancia al considerar su desempeño computacional. Utilizando el mallado uniforme se resolvió el problema directo en 10 minutos, frente a los 35 minutos del adaptado (reducción del 71%). Esta ventaja, combinada con su escalabilidad en operaciones matriciales de gran escala, lo convierte en la elección óptima para problemas de alta dimensión, donde la eficiencia es crítica.

Como mencionamos anteriormente, para obtener los datos sintéticos \mathbf{d}^{obs} se utilizarán los valores de la solución numérica en los receptores \mathbf{r}_j , $j = 1, \dots, J$: $\mathbf{f}_{\text{uniforme}}(\boldsymbol{\nu})$ y $\mathbf{f}_{\text{adaptado}}(\boldsymbol{\nu})$. A estos conjuntos de mediciones se les añadirá ruido Gaussiano para su uso posterior como datos sintéticos en las pruebas de inversión. El conjunto de datos sintéticos es posible obtenerlo utilizando la expresión (5.1). El término de ruido $\boldsymbol{\eta}$ tiene componentes ϵz , donde $z \sim N(0, 1)$ y ϵ está dado por $\epsilon = \frac{\alpha}{100} \sqrt{\frac{\sum_{j=1}^J \sum_{m=1}^M |d_{j,\text{exacto}}^m|^2}{JM}} = 34'7508$, siendo $M = 10,000$ el número de mediciones por receptor y $\alpha = 5\%$.

Usando la solución numérica del problema directo graficamos $\mathbf{f}(\boldsymbol{\nu})$ en función del tiempo, para cada uno de los receptores \mathbf{r}_j . Añadiendo ruido Gaussiano a los datos $\mathbf{f}(\boldsymbol{\nu})$ con la expresión (5.1), obtenemos los datos sintéticos correspondientes a cada uno de los receptores. En las figuras 5.8 y 5.9 visualizamos la diferencia de amplitud y forma entre el operador de observación $\mathbf{f}(\boldsymbol{\nu})$ y los datos sintéticos \mathbf{d}^{obs} , para cada una de las discretizaciones espaciales, para un tiempo total de simulación de $t = 10$ segundos.

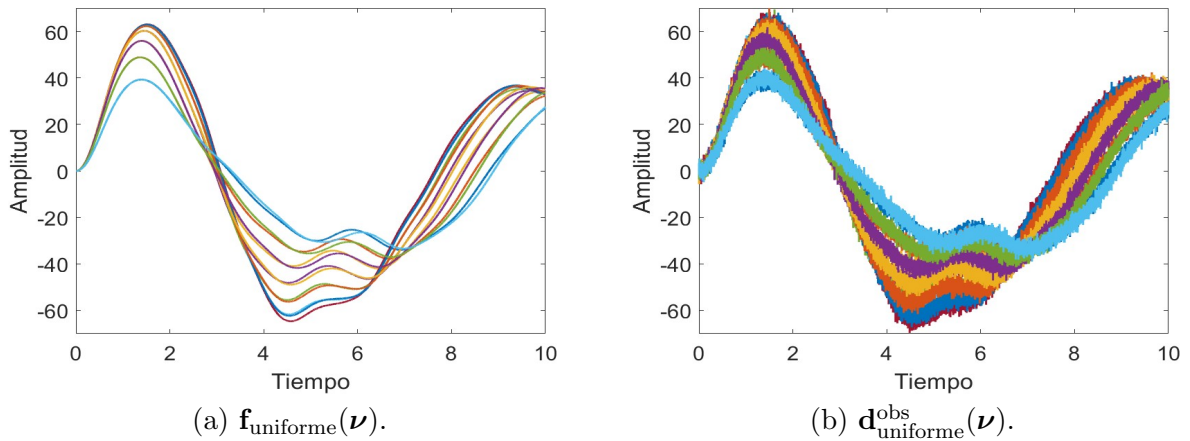


Figura 5.8: Datos medidos por los 13 receptores usando mallado uniforme.

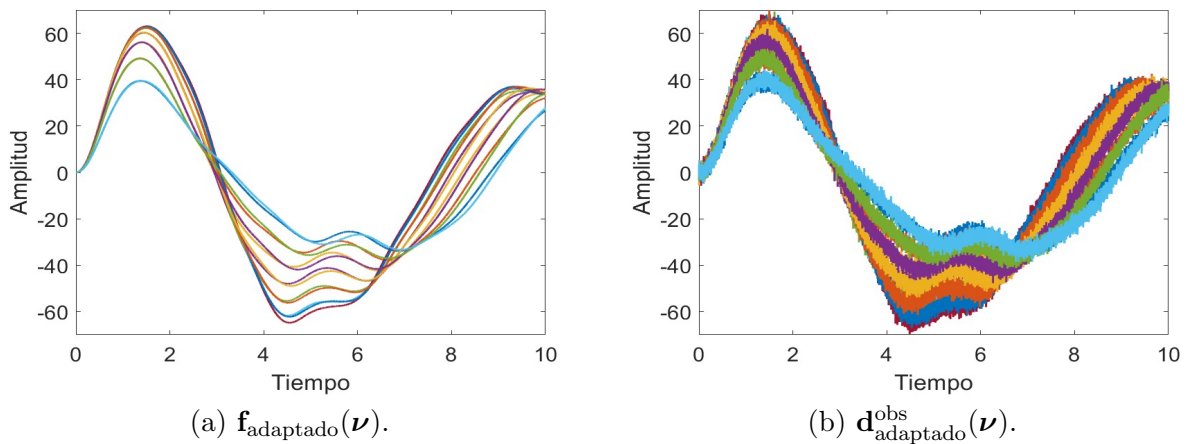


Figura 5.9: Datos medidos por los 13 receptores usando mallado adaptado al domo salino.

5.3. Parametrización mediante series de Fourier: caso bidimensional

En esta sección abordamos la tarea de parametrizar los campos de velocidad y densidad. Las series de Fourier, u otro tipo de representaciones en términos de bases locales empleadas en tratamiento de imágenes, pueden ser útiles para capturar variaciones en regiones específicas. Esta técnica puede ofrecer una representación más precisa de las variaciones locales, especialmente en áreas con

características geológicas complejas, como lo son las trampas de petróleo y gas formadas al costado del domo salino.

Para aplicar la expansión en series de Fourier, es importante preparar adecuadamente los campos espaciales que representan las anomalías del subsuelo. En este caso, consideramos dos campos de interés en las anomalías: el campo de densidad $\rho_A(x, y)$ y el campo de la velocidad de propagación $v_{p,A}(x, y)$ que representan la perturbación de los campos subyacentes en las zonas donde se localizan los yacimientos de gas y petróleo. Los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$ se definen mediante funciones de salto que toman un valor igual a cero en todo el dominio, excepto en la región donde se encuentran las anomalías, es decir, gas y petróleo están presentes (ver figura 5.2). En las zonas específicas donde se encuentran las trampas de gas o petróleo, los valores de densidad y de velocidad se asignan de acuerdo con las propiedades físicas mostradas en la tabla 5.2.

El siguiente paso es representar estas funciones mediante una serie de Fourier. La expansión en series de Fourier es una herramienta poderosa para descomponer funciones en una suma de funciones seno y coseno, facilitando la aproximación de funciones periódicas y no periódicas en un dominio espacial. En particular, para el campo bidimensional de la densidad y de la velocidad de propagación definidos en un dominio rectangular, el desarrollo en serie de Fourier se expresa como [6, 43]:

$$\begin{aligned}
\rho_A(x, y) = & \frac{A_{00}}{4} + \sum_{n=1}^N \frac{A_{n0}}{2} \cos\left(\frac{2n\pi x}{L_x}\right) + \sum_{m=1}^M \frac{A_{0m}}{2} \cos\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \frac{B_{n0}}{2} \sin\left(\frac{2n\pi x}{L_x}\right) + \sum_{m=1}^M \frac{B_{0m}}{2} \sin\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M A_{nm} \cos\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M B_{nm} \sin\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M C_{nm} \cos\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M D_{nm} \sin\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right), \tag{5.2}
\end{aligned}$$

donde $A_{00}, A_{n0}, A_{0m}, B_{n0}, B_{0m}, A_{nm}, B_{nm}, C_{nm}$ y D_{nm} son los coeficientes del desarrollo de Fourier que determinan el campo de densidad $\rho_A(x, y)$ en las zonas de las trampas de gas y petróleo. En este caso de estudio $L_x = 6$ y $L_y = 5$.

Para el campo de velocidad $v_{p,A}(x, y)$, el desarrollo de Fourier viene dado por la expresión (5.3):

$$\begin{aligned}
v_{p,A}(x, y) = & \frac{E_{00}}{4} + \sum_{n=1}^N \frac{E_{n0}}{2} \cos\left(\frac{2n\pi x}{L_x}\right) + \sum_{m=1}^M \frac{E_{0m}}{2} \cos\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \frac{F_{n0}}{2} \sin\left(\frac{2n\pi x}{L_x}\right) + \sum_{m=1}^M \frac{F_{0m}}{2} \sin\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M E_{nm} \cos\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M F_{nm} \sin\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M G_{nm} \cos\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M H_{nm} \sin\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right), \tag{5.3}
\end{aligned}$$

donde E_{00} , E_{n0} , E_{0m} , F_{n0} , F_{0m} , E_{nm} , F_{nm} , G_{nm} y H_{nm} son los coeficientes del desarrollo de Fourier que determinan el campo de velocidad $v_{p,A}(x, y)$ en las zonas de las trampas de gas y petróleo.

En las expresiones (5.2) y (5.3) N y M corresponden al número de coeficientes que se sumarán en la serie para las componentes x e y .

Para una aproximación precisa de los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$ es esencial determinar el número adecuado de modos N y M en la expansión de Fourier [6]. Comenzaremos con una cantidad pequeña de modos y aumentamos gradualmente para observar la mejora en la precisión de la aproximación. La elección del número de modos implica un balance entre la fidelidad de la aproximación y la complejidad computacional. Un número insuficiente de modos puede resultar en una aproximación deficiente, mientras que un número excesivo puede aumentar innecesariamente el costo computacional [6].

5.3.1. Aproximación de campos de velocidad y densidad mediante series de Fourier

En este caso, expresaremos el campo de densidad y de velocidad total como la suma de los campos de fondo estratificado con la incorporación del domo salino, más el campo correspondiente en la zona de las anomalías o trampas, tanto de petróleo como de gas, es decir

$$\begin{aligned}
\rho(x, y) &= \rho_F(x, y) + \rho_A(x, y), \\
v_p(x, y) &= v_{p,F}(x, y) + v_{p,A}(x, y), \tag{5.4}
\end{aligned}$$

donde $\rho_F(x, y)$ es el campo de densidad del fondo estratificado incorporando el domo salino, $\rho_A(x, y)$ es la densidad correspondiente en las trampas de gas o petróleo, $v_{p,F}(x, y)$ es el campo de velocidad del fondo estratificado incorporando el domo salino, y $v_{p,A}(x, y)$ es la velocidad correspondiente en las trampas de gas o petróleo.

Incorporando los desarrollos de Fourier de las expresiones (5.2) y (5.3), correspondientes al campo de velocidad y de densidad en las zonas de las trampas, en las ecuaciones (5.4), obtenemos las aproximaciones para el campo de densidad $\rho(x, y)$ y para el campo de velocidad $v_p(x, y)$. La aproximación para $\rho(x, y)$ es,

$$\begin{aligned}
\rho(x, y) \approx \rho_F(x, y) + & \left[\frac{A_{00}}{4} + \sum_{n=1}^N \frac{A_{n0}}{2} \cos\left(\frac{2n\pi x}{L_x}\right) + \sum_{m=1}^M \frac{A_{0m}}{2} \cos\left(\frac{2m\pi y}{L_y}\right) \right. \\
& + \sum_{n=1}^N \frac{B_{n0}}{2} \sin\left(\frac{2n\pi x}{L_x}\right) + \sum_{m=1}^M \frac{B_{0m}}{2} \sin\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M A_{nm} \cos\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M B_{nm} \sin\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M C_{nm} \cos\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) \\
& \left. + \sum_{n=1}^N \sum_{m=1}^M D_{nm} \sin\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right) \right] \tag{5.5}
\end{aligned}$$

donde la densidad del fondo estratificado incorporando la estructura de domo salino es $\rho_F(x, y)$, que también admite desarrollo de Fourier, sin embargo el enfoque principal aquí es la variación específica $\rho_A(x, y)$ en la zona de gas y petróleo.

La aproximación del campo de velocidad $v_p(x, y)$ incorporando el desarrollo de Fourier para el campo de velocidad en las zonas de las trampas de petróleo y gas, se puede escribir de la siguiente forma,

$$\begin{aligned}
v_p(x, y) \approx v_{p,F}(x, y) & \left[\frac{E_{00}}{4} + \sum_{n=1}^N \frac{E_{n0}}{2} \cos\left(\frac{2n\pi x}{L_x}\right) + \sum_{m=1}^M \frac{E_{0m}}{2} \cos\left(\frac{2m\pi y}{L_y}\right) \right. \\
& + \sum_{n=1}^N \frac{F_{n0}}{2} \sin\left(\frac{2n\pi x}{L_x}\right) + \sum_{m=1}^M \frac{F_{0m}}{2} \sin\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M E_{nm} \cos\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M F_{nm} \sin\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) \\
& + \sum_{n=1}^N \sum_{m=1}^M G_{nm} \cos\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) \\
& \left. + \sum_{n=1}^N \sum_{m=1}^M H_{nm} \sin\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right) \right] \tag{5.6}
\end{aligned}$$

Al igual que para la densidad $\rho_F(x, y)$, la velocidad del fondo estratificado con la incorporación del domo salino, $v_{p,F}(x, y)$, es posible aproximarla por series de Fourier. Sin embargo, el enfoque principal es aproximar el campo $v_{p,A}(x, y)$ en la zona de las trampas de petróleo y gas.

En la figura 5.10 se muestran los pasos principales para incorporar las aproximaciones de los campos de densidad y velocidad en las trampas de petróleo y gas, con el objetivo de evitar discontinuidades en estos campos dentro de la zona de las trampas. Para determinar cuántos modos M y N son necesarios para aproximar los campos de velocidad $v_{p,A}(x, y)$ y de densidad $\rho_A(x, y)$ en la zona de las trampas, realizamos una prueba de sensibilidad, con número de modos crecientes $M = N = 15, 25, 40, 100$.

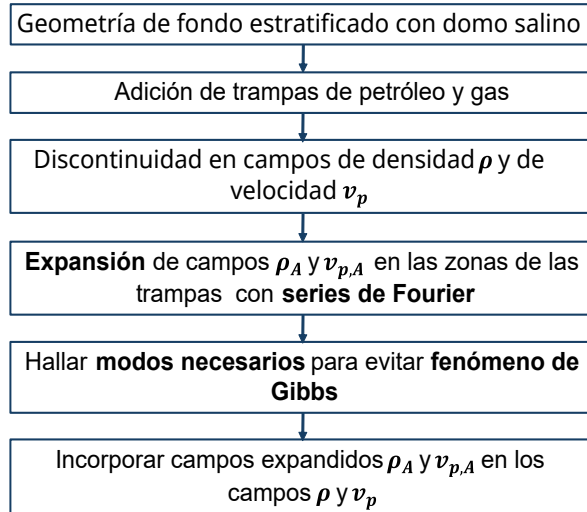


Figura 5.10: Diagrama de flujo con los principales pasos para la adición de los campos en las zonas de petróleo y gas.

A continuación, estudiamos los campos de velocidad $v_{p,A}(x, y)$ correspondientes a las trampas, usando diferente número de modos. Según se visualiza en las figuras 5.11a y 5.11b, con números de modos bajos la aproximación del campo de velocidad en las zonas de las trampas no es adecuada, ya que, para modo $M = N = 15$ modos no se distingue bien la separación entre la trampa de petróleo y la de gas, sin embargo para $M = N = 25$ modos comienzan a distinguirse los dos materiales.

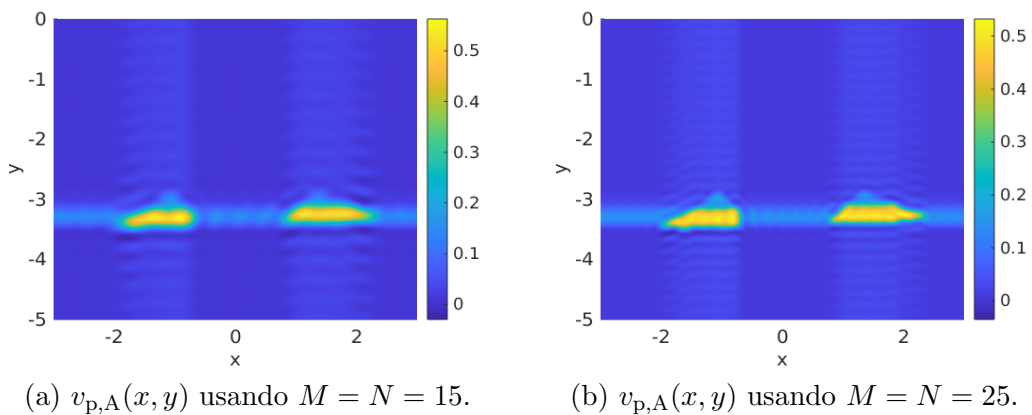
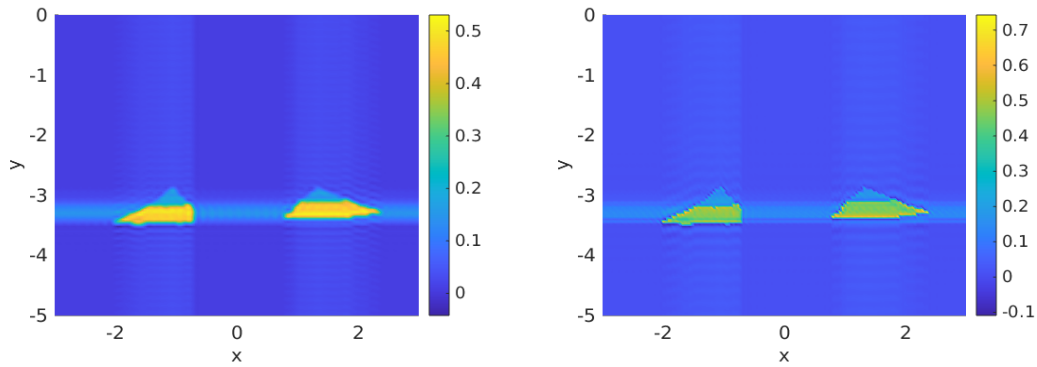


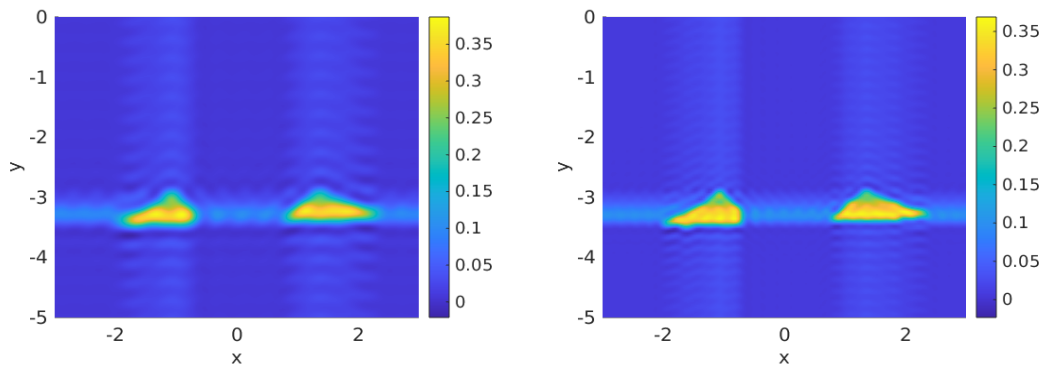
Figura 5.11: Campo de velocidad $v_{p,A}(x, y)$ aproximado con series de Fourier.

En las figuras 5.12a y 5.12b, visualizamos mejores aproximaciones del campo de velocidad, ya que es posible distinguir los dos materiales, podemos observar la separación de los campos para el petróleo y gas dentro de cada trampa. Sin embargo, para $M = N = 40$ y $M = N = 100$ se produce un pequeño rizado debido al fenómeno de Gibbs. Este fenómeno, que es común en las series de Fourier, genera oscilaciones indeseadas cerca de discontinuidades o cambios abruptos. Aunque con $M = N = 40$ y $M = N = 100$ modos se aproximan bien los campos de velocidad para las zonas de petróleo y gas, se amplifica el fenómeno de Gibbs, por lo que se ha optado por reducir el número de modos a $M = N = 25$ (campo de figura 5.11b). Esta reducción de modos atenúa las oscilaciones provocadas por el fenómeno de Gibbs en la malla considerada, permitiendo una aproximación más precisa del comportamiento del campo de velocidad en las regiones de las trampas de petróleo y gas.



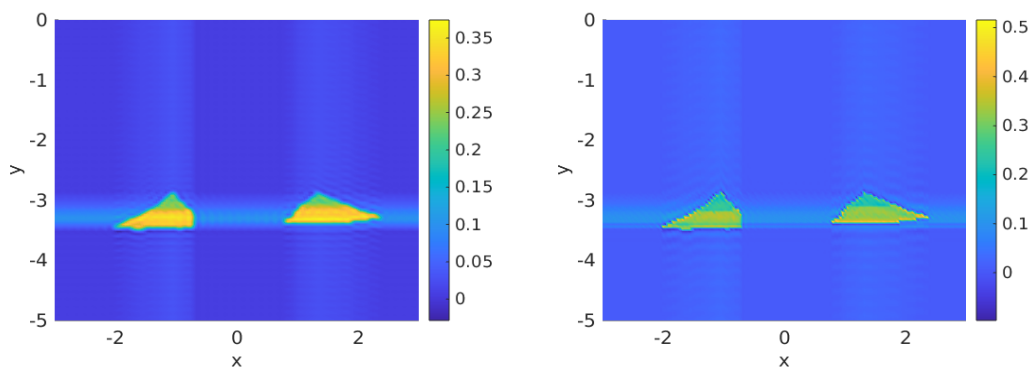
(a) $v_{p,A}(x, y)$ usando $M = N = 40$. (b) $v_{p,A}(x, y)$ usando $M = N = 100$.

Figura 5.12: Campo de velocidad $v_{p,A}(x, y)$ aproximado con series de Fourier.



(a) $\rho_A(x, y)$ usando $M = N = 15$. (b) $\rho_A(x, y)$ usando $M = N = 25$.

Figura 5.13: Campo de densidad $\rho_A(x, y)$ aproximado con series de Fourier.



(a) $\rho_A(x, y)$ usando $M = N = 40$. (b) $\rho_A(x, y)$ usando $M = N = 100$.

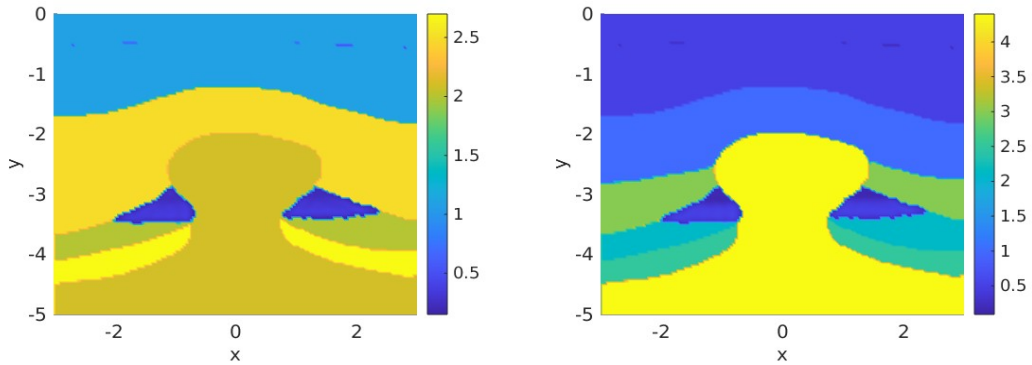
Figura 5.14: Campo de densidad $\rho_A(\mathbf{x}, y)$ aproximado con series de Fourier.

Por otra parte, es conveniente no elegir un número de modos alto con el fin de reducir simplificar la convergencia y coste de las simulaciones a realizar.

Para determinar el número de modos que se utilizarán en la aproximación del campo de densidad en las zonas de las trampas $\rho_A(x, y)$, al igual que en el campo de la velocidad, también se realizaron pruebas con diferentes números de modos en las expansiones de Fourier: $M = N = 15$, $M = N = 25$,

$M = N = 40$, y $M = N = 100$. Los resultados mostraron que con un número menor de modos, como $M = N = 15$, las series de Fourier no logran aproximar adecuadamente el comportamiento del campo de densidad en las trampas de petróleo y gas, como observamos en la figura 5.13a. Por otro lado, con números de modos más elevados, como $M = N = 25$, $M = N = 40$ y $M = N = 100$ ofrecen una buena aproximación general, sin embargo para $M = N = 40$ y $M = N = 100$ se produce un rizado, nuevamente debido al fenómeno de Gibbs. Dado que este efecto es menor en la malla elegida con $M = N = 25$ modos, se ha decidido utilizar el desarrollo del campo de densidad $\rho_A(x, y)$ con este número de modos, garantizando así una representación más suave y precisa.

Una vez seleccionado un número de modos adecuado para aproximar los campos en la zona de las trampas $\rho_A(x, y)$ y $v_{p,A}(x, y)$, estos desarrollos se introducen en las expresiones de las ecuaciones (5.4), que se incorporarán en la formulación del problema directo. En las figuras 5.15a y 5.15b, se muestran los campos de fondo estratificado considerando la estructura de domo salino ρ_F y $v_{p,F}$, con los desarrollos ρ_A y $v_{p,A}$, correspondientes al campo densidad y de velocidad, respectivamente.



(a) Campo $\rho_F(x, y)$ incorporando $\rho_A(x, y)$. (b) Campo $v_{p,F}(x, y)$ incorporando $v_{p,A}(x, y)$.

Figura 5.15: Campos de fondo de densidad $\rho_F(x, y)$ y de velocidad $v_{p,F}(x, y)$, incorporando las expansiones $\rho_A(x, y)$ y $v_{p,A}(x, y)$.

5.3.2. Comparación de operadores de observación aproximados

Para la inversión de campos de ondas, como hemos explicado anteriormente, es necesario ser capaz de evaluar el operador de observación aproximado $\mathbf{f}(\boldsymbol{\nu})$, que corresponde al desplazamiento de la onda captada por el conjunto de los receptores para campos de densidad y velocidad caracterizados por un vector de parámetros $\boldsymbol{\nu}$ que representa los coeficientes de los desarrollos de Fourier. En $\boldsymbol{\nu}$ se incluyen en primer lugar los coeficientes de la densidad y en segundo lugar los coeficientes de la velocidad, ordenados en orden creciente de modos. Este operador se utiliza para evaluar costes en formulaciones deterministas por un lado y para evaluar probabilidades condicionadas en formulaciones bayesianas por otro. En nuestro estudio lo utilizamos además para generar los datos sintéticos \mathbf{d}^{obs} a través de la ecuación (5.1).

La geometría en estudio se describe en la sección 5.1. Las discretizaciones espaciales utilizadas serán la discretización mediante elementos finitos en malla uniforme descrita en 5.2.1 y la discretización espacial mediante elementos finitos en malla adaptada a los estratos y al domo salino descrita en 5.2.1. Para asegurar estabilidad en la solución del problema directo se aplicó la condición de estabilidad de Courant-Friedrichs-Lewy, explicada en el capítulo 2, donde para asegurar la estabilidad

se debe cumplir que $\delta t \leq 1 \times 10^{-3}$, por lo que utilizaremos este paso de tiempo y realizaremos 10,000 iteraciones, obteniendo un tiempo total de simulación de 10 segundos.

Al igual que en la sección 5.1, las ondas elásticas serán emitidas por un conjunto de emisores \mathbf{x}_k , con $k = 0, \dots, K$, los cuales se posicionarán en la interfaz entre aire y suelo, es decir, en $y = 0$. El barrido horizontal de emisores varía desde la coordenada $(-1'76, 0)$ hasta $(1'76, 0)$ con un espaciado de $0'04$. En total, se utilizarán $K = 14$ emisores. Los receptores \mathbf{r}_j , con $j = 0, \dots, J$, al igual que en 5.1, se posicionarán de forma intercalada con los emisores. Estos receptores se ubican desde la posición $(-1'68, 0)$ hasta $(1'68, 0)$, con un espaciado de $0'04$. En total, se utilizarán $J = 13$ receptores. Las características de las fuentes se describen en la sección 5.1.

Se resuelve el problema directo de la propagación de onda, incorporando los desarrollos de los campos de densidad (5.5) y velocidad (5.6). Como resultado, obtenemos el desplazamiento de la onda para cada instante de tiempo durante los 10 segundos de simulación, y por lo tanto, obtenemos el operador de observación aproximado $\mathbf{f}(\boldsymbol{\nu})$ para cada una de las discretizaciones espaciales. En las figuras 5.16a y 5.16b observamos el desplazamiento de la onda para la discretización espacial uniforme y para la discretización espacial adaptada a la estructura de domo salino, respectivamente, durante los 10 segundos para cada receptor. Al comparar, observamos que ambos campos de desplazamiento de onda son bastante similares entre sí.

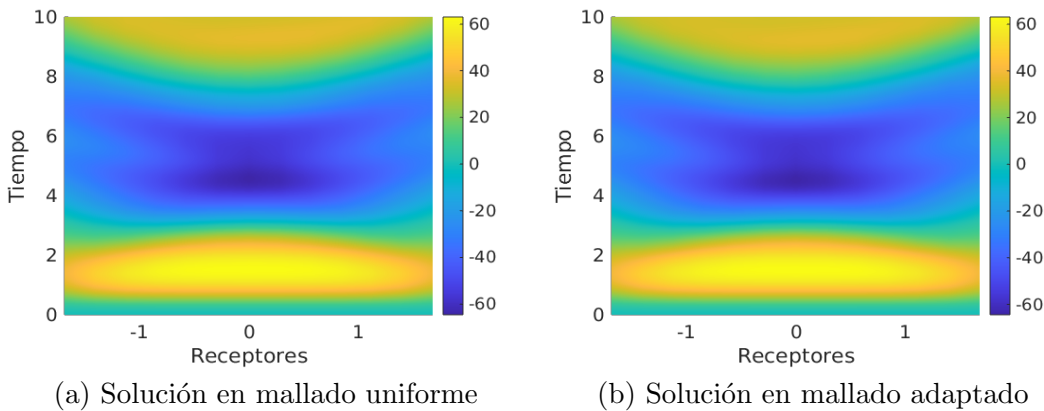


Figura 5.16: Valores observados en los receptores en función del tiempo para coeficientes que incorporan la contribución de los desarrollos de Fourier de $\rho_A(x, y)$ y $v_{p,A}(x, y)$ definidos por coeficientes $\boldsymbol{\nu}$.

Para cuantificar la diferencia en las mediciones obtenidas usando las dos discretizaciones, calculamos la diferencia porcentual relativa entre ambos conjuntos de datos, es decir, entre $f_{\text{uniforme}}(\boldsymbol{\nu})$ y $f_{\text{adaptado}}(\boldsymbol{\nu})$. La diferencia porcentual relativa promedio es de $3'08\%$.

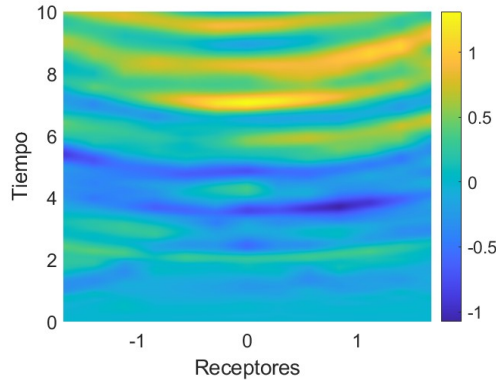


Figura 5.17: Diferencia entre los valores medidos en los receptores para los dos mallados utilizados en la simulación.

Como hemos mencionado anteriormente, a las observaciones numéricas $f_{\text{uniforme}}(\boldsymbol{\nu})$ y $f_{\text{adaptado}}(\boldsymbol{\nu})$, se les añade ruido gaussiano utilizando la expresión de la ecuación (5.1), donde utilizaremos $\epsilon = \frac{\alpha}{100} \sqrt{\frac{\sum_{j=1}^J \sum_{m=1}^M |d_{j,\text{exacto}}^m|^2}{JM}}$, con $\sqrt{\frac{\sum_{j=1}^J \sum_{m=1}^M |d_{j,\text{exacto}}^m|^2}{JM}} = 34'6537$ y un error porcentual de $\alpha = 5\%$. A continuación, visualizamos los gráficos de amplitud de la onda en cada instante de tiempo, durante los 10 segundos de simulación, en cada uno de los receptores \mathbf{r}_j , con $j = 0, \dots, J$, para ambos operadores de observación $\mathbf{f}(\boldsymbol{\nu})$ y para los datos sintéticos \mathbf{d}^{obs} generados con ambos. En la figura 5.18a, observamos la amplitud de la onda correspondiente al operador de observación, $\mathbf{f}_{\text{uniforme}}(\boldsymbol{\nu})$. La figura 5.18b grafica la amplitud de los datos sintéticos $\mathbf{d}_{\text{uniforme}}^{\text{obs}}$ obtenidos usando la discretización uniforme. En la figura 5.19a, se ha graficado el operador de observación con la discretización adaptada al domo salino, mientras que en la figura 5.19b se representa la amplitud de los datos sintéticos generados usando esta misma discretización espacial. En ambos casos, tanto para la discretización uniforme como para la adaptada, se ha utilizado un nivel porcentual de ruido igual al 5% para la obtención de los datos sintéticos.

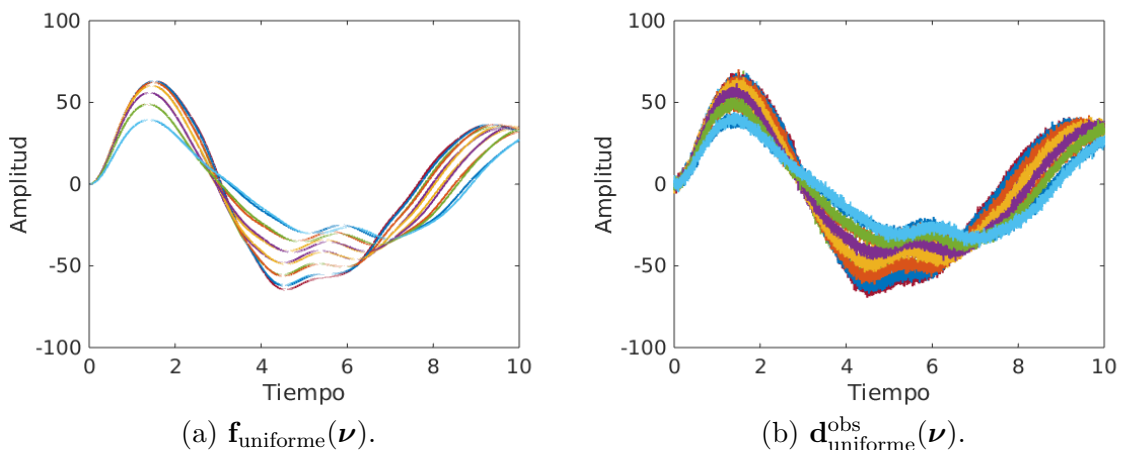


Figura 5.18: Datos obtenidos por los 13 receptores usando mallado uniforme e incorporando los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$.

En la figura 5.20 visualizamos un diagrama con la forma en la que vamos a usar las resoluciones del problema directos y los mallados considerados en el estudio del problema bayesiano inverso. Utilizaremos como datos los generados con el mallado adaptado al domo salido con los campos de

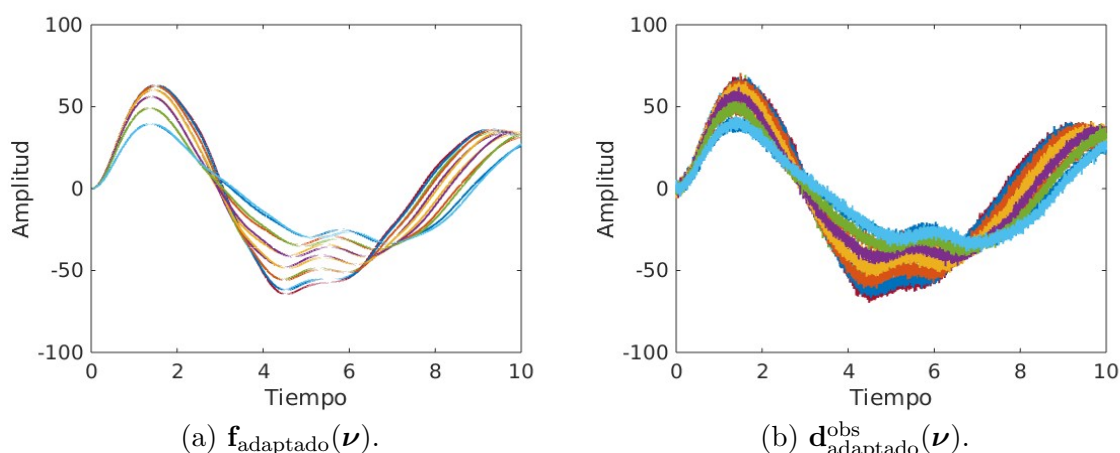


Figura 5.19: Datos obtenidos obtenido por los 13 receptores usando mallado adaptado e incorporando el domo salino y los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$.

densidad y velocidad que incorporan el desarrollo de Fourier en la zona de los yacimientos de gas y petróleo. Al adaptar el mallado a los estratos, se espera mejor reproducción de la propagación de ondas en el medio. Sin embargo, por motivos de reducción de coste computacional, la verosimilitud o probabilidad condicional emplea el operador de observación asociado al mallado uniforme.

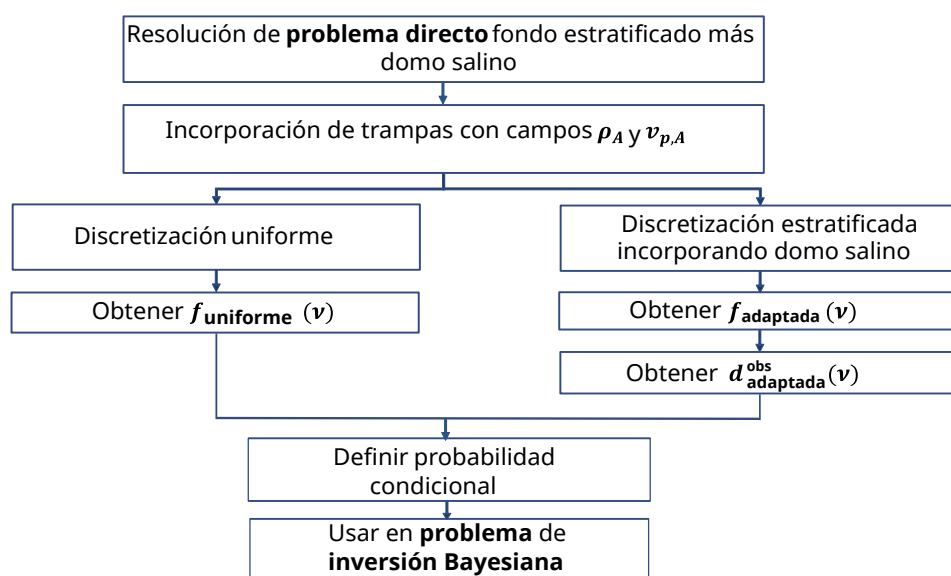


Figura 5.20: Diagrama de flujo sobre el procedimiento para la definición de la probabilidad condicional y su uso en el método inverso bayesiano.

5.4. Formulación bayesiana inversa aproximada y resultados de muestreo

Como estudiamos anteriormente, el problema inverso bayesiano consiste en estimar los parámetros de interés a partir de información a priori con incertidumbre cuantificada, utilizando el operador de

observación y los datos sintéticos que se obtienen a través de la resolución del problema directo (resultados obtenidos en la sección 5.3.2). En este caso, los parámetros de interés $\boldsymbol{\nu}$ corresponden a los coeficientes que generan campos de densidad y de velocidad en la zona de las trampas de petróleo y gas, es decir, $\rho_A(x, y)$ y de $v_{p,A}(x, y)$, aproximados utilizando series de Fourier (ecuaciones (5.2) y 5.3)). Definimos el vector de parámetros $\boldsymbol{\nu}$ como

$$\boldsymbol{\nu} = \begin{bmatrix} A_{00}, & E_{00}, \\ A_{1,0}, \dots, A_{N,0}, & E_{1,0}, \dots, E_{N,0}, \\ B_{1,0}, \dots, B_{N,0}, & F_{1,0}, \dots, F_{N,0}, \\ A_{0,1}, \dots, A_{0,M}, & E_{0,1}, \dots, E_{0,M}, \\ B_{0,1}, \dots, B_{0,M}, & F_{0,1}, \dots, F_{0,M}, \\ A_{1,1}, \dots, A_{N,M}, & E_{1,1}, \dots, E_{N,M}, \\ B_{1,1}, \dots, B_{N,M}, & F_{1,1}, \dots, F_{N,M}, \\ C_{1,1}, \dots, C_{N,M}, & G_{1,1}, \dots, G_{N,M}, \\ D_{1,1}, \dots, D_{N,M}, & H_{1,1}, \dots, H_{N,M} \end{bmatrix}, \quad (5.7)$$

en total, los parámetros de $\boldsymbol{\nu}$ son:

- 2 constantes (A_{00}, E_{00}).
- $4 \times (M + N)$ términos unidimensionales: los coeficientes $A_{N,0}, B_{N,0}, A_{0,M}, B_{0,M}$ correspondientes al campo $\rho_A(\mathbf{x}, \mathbf{y})$, y los coeficientes $E_{N,0}, F_{N,0}, E_{0,M}, F_{0,M}$ correspondientes al campo $v_{p,A}(x, y)$.
- $8 \times M \times N$ términos bidimensionales: los coeficientes $A_{N,M}, B_{N,M}, C_{N,M}, D_{N,M}$ correspondientes al campo de densidad $\rho_A(\mathbf{x}, \mathbf{y})$, y los coeficientes $E_{N,M}, F_{N,M}, G_{N,M}, H_{N,M}$ correspondientes al campo de velocidad $v_{p,A}(x, y)$, con $N = M = 1, \dots, 25$ modos.

En este caso, el enfoque del método inverso bayesiano, se basa en el muestreo de los campos $\rho_A(\boldsymbol{\nu})$ y $v_{p,A}(\boldsymbol{\nu})$, por lo que analizamos el comportamiento de la distribución de la probabilidad posterior con respecto a los parámetros de interés $\boldsymbol{\nu}$. Utilizamos el teorema de Bayes, para caracterizar la distribución a posteriori de los parámetros (y por tanto de los campos $\rho_A(\boldsymbol{\nu})$ y $v_{p,A}(\boldsymbol{\nu})$) [24] conocidos los datos

$$p_{\text{posterior}}(\boldsymbol{\nu} \mid \mathbf{d}^{\text{obs}}) = \frac{p(\mathbf{d}^{\text{obs}} \mid \boldsymbol{\nu}) p_{\text{prior}}(\boldsymbol{\nu})}{p(\mathbf{d}^{\text{obs}})}. \quad (5.8)$$

En la expresión (5.8) tenemos,

- $p(\mathbf{d}^{\text{obs}} \mid \boldsymbol{\nu})$ es a probabilidad condicionada del operador observable (o verosimilitud), que modela cómo se generan los datos observacionales a partir de los campos verdaderos de la densidad y de la velocidad,
- $p_{\text{prior}}(\boldsymbol{\nu})$ es la distribución de la probabilidad a priori, que incluye información inicial sobre los parámetros de interés, en este caso, incluye la información inicial sobre los campos de densidad y de velocidad en la zona de las trampas de petróleo y de gas,
- $p(\mathbf{d}^{\text{obs}})$ es un factor de normalización.

Al incorporar los parámetros de interés (5.7) en la probabilidad condicional $p(\mathbf{d}^{\text{obs}} \mid \boldsymbol{\nu})$ obtenemos la siguiente expresión:

$$p(\mathbf{d}^{\text{obs}} \mid \boldsymbol{\nu}) = \frac{1}{\sqrt{(2\pi)^{P/2} \det(\boldsymbol{\Gamma}_{\text{ruido}})}} \exp\left(-\frac{1}{2} \left(\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}}\right)^\top \boldsymbol{\Gamma}_{\text{ruido}}^{-1} \left(\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}}\right)\right), \quad (5.9)$$

donde

- $\mathbf{f}(\boldsymbol{\nu})$ es el operador de observación del modelo directo con el mallado uniforme,
- \mathbf{d}^{obs} son los datos sintéticos obtenidos según se indica en la subsección 5.3.2 con el mallado adaptado,
- $\boldsymbol{\Gamma}_{\text{ruido}}$ es la matriz de covarianza del ruido en los datos observados. En nuestro caso, será una matriz con diagonal constante σ_{ruido}^2 .

Para la probabilidad a priori, se incorpora el parámetro inicial $\boldsymbol{\nu}_0$, que representa la información disponible sobre los coeficientes de los desarrollos de Fourier truncados a N y M términos de las perturbaciones $\rho_A(x, y)$ y $v_{p,A}(x, y)$ de los campos de densidad y velocidad de referencia en el entorno del domo salido, donde se espera que pueda haber trampas de gas y/o petróleo. Tomaremos $\boldsymbol{\nu}_0$ nulo, lo que implica que los campos de densidad y velocidad en estas regiones se considerarán con los valores de fondo especificados en la tabla 5.1. Por lo tanto, los campos iniciales de densidad y de velocidad en la zona de las trampas de petróleo y gas se definen como,

$$\begin{aligned}\rho_{A_0}(x, y) &= 0, \\ v_{p,A_0}(x, y) &= 0.\end{aligned}\tag{5.10}$$

Suponemos que la probabilidad a priori de los parámetros $\boldsymbol{\nu}$ (los coeficientes de los desarrollos de $\rho_A(x, y)$ y $v_{p,A}(x, y)$) en la probabilidad a priori $p_{\text{prior}}(\boldsymbol{\nu})$ viene dada por una distribución normal de media $\boldsymbol{\nu}_0$

$$p_{\text{prior}}(\boldsymbol{\nu}) = \frac{1}{\sqrt{(2\pi)^{P/2} \det(\boldsymbol{\Gamma}_{\text{prior}})}} \exp\left(-\frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)^\top \boldsymbol{\Gamma}_{\text{prior}}^{-1}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)\right),\tag{5.11}$$

siendo $\boldsymbol{\Gamma}_{\text{prior}}$ es matriz de covarianza a priori de los campos de interés y P la longitud de $\boldsymbol{\nu}$. Para simplificar, tomamos una matriz diagonal definida a partir de las varianzas de la densidad $\sigma_\rho^2 = 0'01$ y de la velocidad $\sigma_v^2 = 0'18$. En nuestro caso, estos valores los calculamos a partir de las desviaciones observadas en la tabla 5.2.

Incorporando las expresiones (5.9) y (5.11) en la probabilidad a posteriori del teorema de Bayes, obtenemos

$$\begin{aligned}p_{\text{posterior}}(\boldsymbol{\nu} \mid \mathbf{d}^{\text{obs}}) &\propto \frac{1}{(2\pi)^P \sqrt{\det(\boldsymbol{\Gamma}_{\text{ruido}})}} \frac{1}{\sqrt{\det(\boldsymbol{\Gamma}_{\text{prior}})}} \\ &\exp\left(-\frac{1}{2}\left[\left(\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}}\right)^\top \boldsymbol{\Gamma}_{\text{ruido}}^{-1}\left(\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}}\right) \right. \\ &\quad \left. + \left(\boldsymbol{\nu} - \boldsymbol{\nu}_0\right)^\top \boldsymbol{\Gamma}_{\text{prior}}^{-1}\left(\boldsymbol{\nu} - \boldsymbol{\nu}_0\right)\right]\right).\end{aligned}\tag{5.12}$$

La probabilidad posterior dada por la expresión (5.12) puede expresarse en términos de un función de costo, en la forma $\exp\{-J_{\text{reg}}(\boldsymbol{\nu})\}$, donde

$$J_{\text{reg}}(\boldsymbol{\nu} \mid \mathbf{d}^{\text{obs}}) = \frac{1}{2} \left\| \mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}} \right\|_{\boldsymbol{\Gamma}_{\text{ruido}}}^2 + \frac{1}{2} \left\| \boldsymbol{\nu} - \boldsymbol{\nu}_0 \right\|_{\boldsymbol{\Gamma}_{\text{prior}}}^2,\tag{5.13}$$

siendo $\|\boldsymbol{\nu}\|_{\boldsymbol{\Gamma}} = \boldsymbol{\nu}^\top \boldsymbol{\Gamma}^{-1} \boldsymbol{\nu}$ la norma cuadrática asociada a la matriz de covarianza $\boldsymbol{\Gamma}$.

Tanto el cálculo de la verosimilitud en la probabilidad a posteriori como el cálculo del coste, conllevan la evaluación del operador de observación para cada propuesta de parámetros $\boldsymbol{\nu}$. Para ello, es preciso resolver el problema directo para campos de densidad y velocidad que corrigen los campos de fondo con los desarrollos de Fourier truncados con coeficientes dados por $\boldsymbol{\nu}$.

La distribución a posteriori no normalizada se puede muestrear mediante métodos de cadenas de Markov Monte Carlo (MCMC) utilizando el esquema AIES presentado en capítulos anteriores.

Si la información a priori proporcionada por $\boldsymbol{\nu}_0$ y Γ_{prior} es adecuada, se espera que la función de costo $J_{\text{reg}}(\boldsymbol{\nu})$ sea convexa, lo que podría indicar la presencia de un único máximo global, gracias al efecto regularizador del término $\|\boldsymbol{\nu} - \boldsymbol{\nu}_0\|_{\Gamma_{\text{prior}}}$. No obstante, dado que el enfoque se basa en el muestreo, la exploración de la probabilidad a posteriori no está enfocada a estudiar probabilidades con un único máximo, sino a converger a la distribución de probabilidad posterior de los parámetros aunque sea multimodal.

5.4.1. Resultados de muestreo AIES

Para obtener un muestreo de los campos de densidad y de velocidad en la zona de las trampas, hemos utilizado un método de cadenas de Markov Monte Carlo de tipo AIES (explicado en el capítulo 3). Como comentamos anteriormente, el parámetro inicial $\boldsymbol{\nu}_0$ es nulo, es decir, $\rho_{A_0}(x, y) = 0$ y $v_{p,A_0}(x, y) = 0$.

La figura 5.21 ha sido elaborada a partir de 1,650 cadenas y 1,500 iteraciones. Utilizamos las muestras generadas para visualizar la estructura de los campos de densidad y velocidad en las zonas de las trampas de petróleo y gas. Estas muestras representan campos distribuidos según la distribución a posteriori, proporcionando una representación indirecta de la variabilidad e incertidumbre de los parámetros de interés de $\rho_A(x, y)$ y $v_{p,A}(x, y)$. La figura 5.21 representa el campo de densidad de fondo ρ_F corregido con el campo $\rho_A(\boldsymbol{\nu})$ correspondiente a la muestra de probabilidad máxima $\boldsymbol{\nu}_{\text{MAP}}$. El análogo para la velocidad refleja fluctuaciones menos apreciables.

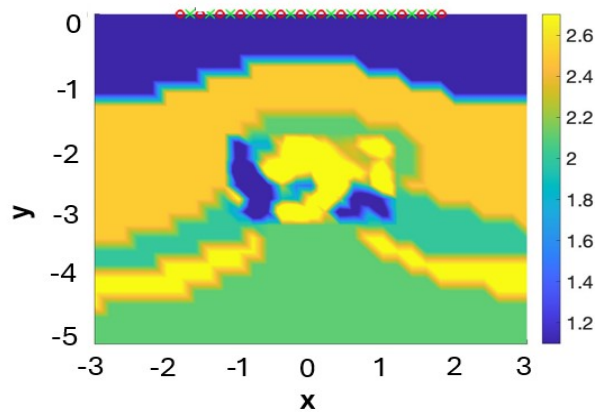


Figura 5.21: Campo de densidad con probabilidad máxima obtenido en 1,500 iteraciones, incluyendo $\rho_A(x, y)$.

En la figura 5.21, se puede observar que, después de 1,500 iteraciones, se logra identificar la presencia de otros materiales de baja densidad en la zona de los yacimientos reales de gas y petróleo.

Cabe destacar que la información a priori introducida es limitada, pues hemos usado el valor cero para todos los coeficientes de Fourier. A pesar de esta simplificación, el método logra identificar la presencia de trampas con distintos materiales.

No obstante, las representaciones de Fourier presentan desafíos para caracterizar con precisión el contenido de las trampas, particularmente en la selección de matrices de covarianza adecuadas. En este capítulo, la matriz de covarianza se ha definido mediante un enfoque genérico, sin incorporar conocimiento específico previo sobre la estructura espacial de los campos.

Capítulo 6

Estrategias de inversión basadas en campos aleatorios

En el capítulo anterior hemos utilizado series de Fourier para parametrizar funciones de densidad y velocidad que deseamos precisar a partir de datos indirectos disponibles. Es posible recurrir a otros tipos de caracterizaciones, que serán más o menos adecuados según el caso. En este capítulo formulamos un marco basado en el uso de desarrollos Karhunen-Loève. Técnicamente, este punto de vista aporta una nueva faceta, al permitir considerar los campos a determinar como campos aleatorios. En todos los capítulos anteriores hemos reducido el problema al estudio de un conjunto finito de parámetros que en el marco bayesiano se traducen en un número finito de variables aleatorias. Los capítulos 2-4 abordan una drástica reducción de dimensionalidad al caracterizar los campos buscados mediante funciones constantes a trozos, definidas por unos pocos parámetros que controlan la geometría y naturaleza de las regiones donde se presentan variaciones abruptas. El capítulo anterior representa un paso intermedio, puesto que los campos en estudio vienen descritos por funciones infinito dimensionales, pero se caracterizan por una colección finita de coeficientes. Si bien en esta sección también truncaremos los desarrollos de Karhunen-Loève por simplicidad computacional, desde el punto de vista teórico, es un paso en la dirección del estudio de formulaciones aleatorias infinito dimensionales, que se alcanzarían en el límite de número infinito de modos. Los desarrollos de Karhunen-Loève están asociados a covarianzas que se esperan sean adecuadas para los campos aleatorios en estudio. Una dificultad en la aplicación práctica de estos desarrollos es la determinación de operadores de covarianza adecuados, ya que pobres elecciones pueden conducir a pobres resultados. Otra dificultad es la construcción práctica del conjunto completo de funciones de base del desarrollo para operadores de covarianzas generales.

Comenzaremos este capítulo recordando algunos conceptos sobre campos aleatorios y la teoría de Karhunen-Loève. Tras ello, explicaremos cómo representar los campos de densidad y de velocidad en este marco en la geometría de domo salino considerada en el capítulo anterior. Revisaremos la formulación del problema bayesiano de domo salino en este contexto y discutiremos las técnicas de muestreo adecuadas, concluyendo con detalles sobre la implementación.

6.1. Breve recordatorio sobre variables y campos aleatorios

Cuando los parámetros de interés varían de forma aleatoria en el espacio, es necesario modelar esta incertidumbre utilizando campos aleatorios. Estos permiten describir cómo los parámetros pueden cambiar de manera incierta en cada punto del espacio, añadiendo una dimensión estocástica al problema [20, 22]. Un campo aleatorio puede ser entendido como un conjunto de variables aleatorias asociadas a cada punto del espacio en un dominio continuo. Esto implica que en cada ubicación espacial se asigna una variable aleatoria que describe el valor incierto del parámetro en ese punto.

Una variable aleatoria X se define como una función medible $X : A \rightarrow E$ desde un espacio muestral A (conjunto de posibles resultados de un experimento) hasta el espacio medible E . El espacio de medición o de probabilidad (A, Σ, μ) está formado por: un conjunto A (el espacio de muestras), una σ -álgebra Σ de eventos medibles en A (cerrado por unión, intersección y complementación) y una medida $\mu : \Sigma \rightarrow \mathbb{R} \cup \{\pm\infty\}$ (satisface $\mu(\emptyset) = 0$ y la medida de la unión de conjuntos disjuntos es la suma de las medidas) [52, 56].

Dados $\Omega \subset \mathbb{R}^N$ y un espacio de probabilidad (A, Σ, μ) , un campo aleatorio es una función $k : \Omega \times A \rightarrow \mathbb{R}$, tal que la función $k(\mathbf{x}, \cdot) : A \rightarrow \mathbb{R}$ es una variable aleatoria para cada $\mathbf{x} \in \Omega$, mientras que cada función $k(\cdot, \alpha) : \Omega \rightarrow \mathbb{R}$, $\alpha \in A$ es una realización. De forma intuitiva, podemos decir que un campo aleatorio es un conjunto de variables aleatorias que representan valores inciertos de los parámetros de interés en cada punto espacial [56].

Un campo aleatorio es de segundo orden, si $k(\mathbf{x}, \cdot)$ es cuadrado integrable para todo $\mathbf{x} \in \Omega$. Su promedio lo podemos definir como [52],

$$\bar{k}(\mathbf{x}) = \mathbb{E}[k(\mathbf{x}, \cdot)] = \int_A k(\mathbf{x}, \alpha) d\mu(\alpha).$$

y su covarianza [52],

$$C(\mathbf{x}, \mathbf{y}) = \mathbb{E}[(k(\mathbf{x}, \cdot) - \bar{k}(\mathbf{x}))(k(\mathbf{y}, \cdot) - \bar{k}(\mathbf{y}))],$$

para todo $\mathbf{x}, \mathbf{y} \in \Omega$.

Un campo aleatorio es Gaussiano si, dados D puntos $\mathbf{x}_1, \dots, \mathbf{x}_D \in \Omega \subset \mathbb{R}^N$, se tiene que $[k(\mathbf{x}_1, \cdot), \dots, k(\mathbf{x}_D, \cdot)]$ es una variable aleatoria Gaussiana multivariada. Los campos aleatorios Gaussianos están determinados por su promedio $\bar{k}(\mathbf{x})$ y covarianza $C(\mathbf{x}, \mathbf{y})$.

Si un campo Gaussiano es estacionario entonces su promedio es constante en todo el espacio, es decir, $\bar{k}(\mathbf{x}) = \bar{k}$ y su covarianza se puede expresar como $C(\mathbf{x}, \mathbf{y}) = C(\mathbf{x} - \mathbf{y})$. Si además el campo Gaussiano es isotrópico entonces la covarianza depende de la distancia de los puntos, $C(\mathbf{x}, \mathbf{y}) = C\|\mathbf{x} - \mathbf{y}\|$ [52, 56].

6.1.1. Teorema de Karhunen-Loève

El teorema de Karhunen-Loève (KL) proporciona una representación espectral de campos aleatorios que facilita su análisis y simulación. Este teorema establece que cualquier campo aleatorio estacionario con una función de covarianza adecuada puede ser representado a través de una expansión infinita en términos de sus valores propios y funciones propias.

Dado un campo aleatorio k con una función de covarianza $C(\mathbf{x}, \mathbf{y})$, existe una descomposición espectral en términos de valores propios no negativos $\{\lambda_i\}_{i \in \mathbb{N}}$, ordenados de forma decreciente $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ y funciones propias $\phi_1, \phi_2, \dots, \phi_i, \dots$ en $L^2(\Omega)$ tales que [95],

$$\int_{\Omega} C(\mathbf{x}, \mathbf{y}) \phi_i(\mathbf{x}, \mathbf{y}) d\mathbf{x} = \lambda_i \phi_i(\mathbf{x}, \mathbf{y}), \quad i \in \mathbb{N},$$

donde $\lambda_i \geq 0$ y ϕ_i son ortogonales, $i \in \mathbb{N}$. La colección $\{\phi_i\}_{i \in \mathbb{N}}$ forma una base en $L^2(\Omega)$. Además, el campo aleatorio $k(\mathbf{x}, \alpha)$ admite la expansión de Karhunen-Loève [5],

$$k(\mathbf{x}, \alpha) = \bar{k}(\mathbf{x}) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \phi_i(\mathbf{x}) Z_i(\alpha), \quad \mathbf{x} \in \Omega,$$

donde \bar{k} es la media del campo aleatorio Gaussiano y Z_i son variables aleatorias independientes y no correlacionadas (i.i.d) [62,67]. Esta expansión permite representar el campo aleatorio de manera eficiente en términos de funciones ortogonales ϕ_i y valores propios λ_i^2 .

La base ortonormal $\{\phi_i\}_{i \in \mathbb{N}}$ es fundamental para la representación del campo aleatorio en el espacio $L^2(\Omega)$. Para obtener λ_i y ϕ_i , se resuelve el problema de valores propios asociados con el operador de covarianza definido por la función de covarianza del campo aleatorio. Para el núcleo Matérn, este operador está relacionado con el operador diferencial fraccionario [64] dado por la ecuación (6.3). En general, los desarrollos KL proporcionan bases de convergencia rápida para la generación de muestras de distribuciones Gaussianas $\mathcal{N}(0, C)$ [34].

La expansión espectral del campo aleatorio se realiza utilizando la serie de Karhunen-Loève (KL), truncada a M términos para aproximar realizaciones del campo [22],

$$k(\mathbf{x}, \alpha) \approx k^{(m)}(\mathbf{x}, \mathbf{Z}) = \bar{k}(\mathbf{x}) + \sum_{i=1}^M \sqrt{\lambda_i} \phi_i(\mathbf{x}) Z_i. \quad (6.1)$$

Esta expansión permite aproximar el campo aleatorio en términos de una serie truncada de funciones propias ponderadas por variables aleatorias independientes. Sólo se retienen un número finito de términos en la expansión, seleccionados según la magnitud decreciente de sus valores propios. Los valores propios más pequeños contribuyen menos a la variabilidad total del campo, por lo que truncar la expansión en estos términos permite una reconstrucción precisa del campo aleatorio con un número manejable de términos [22], y un error estimable.

El truncamiento de las desarrollos KL es crucial para la eficiencia computacional, ya que reduce la dimensión del problema sin sacrificar significativamente la precisión. Es una técnica estándar de muestreo de campos aleatorios cuando se conocen los autovalores y las autofunciones, éste es el caso de los núcleos de tipo Matérn, por ejemplo.

6.1.2. Núcleo Matérn

El núcleo de Whittle-Matérn es una función covarianza $C_\eta(\mathbf{x}, \mathbf{y})$, que se usa para modelar la correlación espacial de los valores del campo en función de la distancia entre dos puntos. Se define como [5],

$$C_\eta(\mathbf{x}, \mathbf{y}) = C_\eta(\|\mathbf{x} - \mathbf{y}\|) = \sigma^2 \frac{2^{1-\eta}}{\Gamma(\eta)} \left(\frac{\sqrt{2\eta} \|\mathbf{x} - \mathbf{y}\|}{\ell} \right)^\eta K_\eta \left(\frac{\sqrt{2\eta} \|\mathbf{x} - \mathbf{y}\|}{\ell} \right), \quad (6.2)$$

donde σ^2 es la varianza, η es el parámetro de regularidad, ℓ es el parámetro longitud de correlación, y K_η es la función de Bessel modificada de segunda clase.

El campo aleatorio asociado con la covarianza Matérn se puede muestrear a través de la siguiente ecuación en derivadas parciales estocástica en \mathbb{R}^N [4, 5],

$$\frac{1}{\sqrt{\beta \ell^N}} (I - \ell^2 \Delta)^{(\eta+N/2)/2} u(\mathbf{x}, \mathbf{y}) = W(\mathbf{x}, \mathbf{y}),$$

donde $W(\mathbf{x}, \mathbf{y})$ es un proceso de ruido blanco Gaussiano con varianza unitaria y $\beta = \sigma^2 \frac{(4\pi)^{N/2} \Gamma(\eta+N/2)}{\Gamma(\eta)}$ es una constante que depende de los parámetros del núcleo Matérn. Para generar muestras del campo aleatorio, se requieren condiciones de frontera en $\Omega \subset \mathbb{R}^N$, típicamente Dirichlet, Neumann

² Como el problema que se resolverá es en dos dimensiones, las expansiones KL dependerán de variables \mathbf{x} con dos componentes.

o periódicas. Se puede resolver en un dominio más grande que contenga Ω para reducir el efecto de estas condiciones de frontera en las correlaciones cerca de la frontera [20]. El operador de covarianza correspondiente a la función de covarianza de la ecuación (6.2) está dado por [5, 79],

$$\beta \ell^N (I - \ell^2 \Delta)^{-\eta - \frac{N}{2}},$$

donde Δ es el operador Laplaciano, I es el operador identidad y N es la dimensión espacial. Este operador de la covarianza actúa sobre las funciones $u(\mathbf{x}, y)$ del campo aleatorio, y tiene la forma de un operador diferencial regularizado, que atenúa las fluctuaciones bruscas (asociada a Δ) según la escala de longitud ℓ [64]. Sin embargo, para evitar complicaciones relacionadas con el factor de escala β y la forma fraccionaria de los operadores, es posible utilizar un operador alternativo de la forma,

$$\tau^2 I - \Delta,$$

donde $\tau = \frac{1}{\ell}$ es el inverso de la escala de longitud. La elección de este operador es útil en aplicaciones prácticas, ya que evita problemas de estabilidad en las simulaciones numéricas al trabajar con potencias fraccionarias del Laplaciano.

Para resolver el sistema en un dominio acotado $\Omega \subset \mathbb{R}^N$ con condiciones de frontera homogéneas (por ejemplo, Dirichlet o Neumann), es útil expresar el operador en términos de sus valores propios y funciones propias. Si μ_i y ϕ_i , $i \in \mathbb{N}$, son los valores propios y funciones propias del operador $\tau^2 I - \Delta$, entonces el operador potencia fraccionaria $(\tau^2 I - \Delta)^s$ actúa sobre una función $u(\mathbf{x}, y)$ como [5],

$$(\tau^2 I - \Delta)^s u = \sum_{i=1}^{\infty} \mu_i^s u_i \phi_i, \quad u_i = \int_{\Omega} u(\mathbf{x}, y) \phi_i(\mathbf{x}, y) d\mathbf{x}, \quad (6.3)$$

donde u_i son los coeficientes de la expansión de $u(\mathbf{x}, y)$ en la base de funciones propias $\phi_i(\mathbf{x}, y)$. Aquí, μ_i son los valores propios del operador $\tau^2 I - \Delta$, y $s = \frac{\eta + N/2}{2}$ corresponde a la potencia fraccionaria que aparece en la ecuación diferencial asociada con la covarianza Matérn.

Los valores propios λ_i del operador fraccionario $(\tau^2 I - \Delta)^s$ están dados por,

$$\lambda_i = \mu_i^{-s},$$

donde μ_i son los valores propios originales del operador $\tau^2 I - \Delta$. Las funciones propias ϕ_i permanecen iguales, ya que son las soluciones del problema de valores propios asociados a $\tau^2 I - \Delta$. De esta forma, podemos generar realizaciones de campos con esta covarianza mediante el desarrollo (6.1).

Sin embargo, no se suele disponer de caracterizaciones análogas de autovalores y autofunciones para operadores de covarianza no Matérn, por lo que es preciso recurrir a otras estrategias de muestreo de campos aleatorios.

6.1.3. Muestreo de campos aleatorios para covarianzas arbitrarias

Consideremos un campo aleatorio Gaussiano $k(\mathbf{x}, \alpha)$ con covarianza c . En primer lugar, numeramos los puntos del mallado de la región computacional (u otra colección de puntos en esa región) como $\mathbf{x}_1, \dots, \mathbf{x}_D \in \Omega \subset \mathbb{R}^N$. Tras ello, consideramos la variable aleatoria Gaussiana multivariada [22]

$$[k(\mathbf{x}_1, \dots), \dots, k(\mathbf{x}_D, \dots)]$$

que tiene media $\bar{\mathbf{k}} = [\bar{k}(\mathbf{x}_1), \dots, \bar{k}(\mathbf{x}_D)]$ y covarianza. $\mathbf{C} = (c(\mathbf{x}_i, \mathbf{x}_j))_{i,j}$.

A continuación calculamos la descomposición espectral $\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^t$, donde $\mathbf{\Lambda}$ es una matriz diagonal cuya diagonal está formada por los autovalores $\lambda_1, \dots, \lambda_D$ y \mathbf{U} es una matriz ortogonal

cuyas columnas son los autovectores normalizados. Podemos generar realizaciones aproximadas del campo k utilizando $\mathbf{k} \sim \mathcal{N}(\bar{\mathbf{k}}, \mathbf{C}) = \mathcal{N}(\bar{\mathbf{k}}, \mathbf{U}\mathbf{\Lambda}\mathbf{U}^t)$, es decir

$$\mathbf{k} = \bar{\mathbf{k}} + \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{r} = \bar{\mathbf{k}} + \sum_{i=1}^D \sqrt{\lambda_i} U_i \mathbf{r}_i, \quad (6.4)$$

con $\mathbf{r} = [r_1, \dots, r_D] \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Para D grande, podemos truncar a q términos con error $\sum_{i=q+1}^D \lambda_i$.

Si cambiamos los puntos $\mathbf{x}_1, \dots, \mathbf{x}_D \in \Omega \subset \mathbb{R}^N$, es preciso recalculer todo, a diferencia de lo que ocurre con los desarrollos de Karhunen-Loève. Caso de que los autovalores y autofunciones para el desarrollo KL infinito dimensional sean conocidos, el empleo de la relación (6.1) requiere evaluar las autofunciones en todo el mallado. Según el número de autofunciones que se preserve y la dimensión, la estrategia expuesta aquí puede ser más eficiente computacionalmente.

6.2. Problema inverso bayesiano en términos de campos aleatorios

La cuantificación de la incertidumbre en problemas de las EDP con parámetros parcialmente conocidos se realiza modelando estas cantidades desconocidas como variables aleatorias con funciones de distribución de probabilidad adecuadas. Al tratar con campos hemos de recurrir a campos aleatorios. Los métodos espectrales son comúnmente utilizados para resolver EDPs que involucran campos aleatorios debido a su alta precisión y eficiencia [7, 96]. En ellos, los campos aleatorios se representan en términos de desarrollos de tipo KL o aproximaciones discretas. Como hemos visto, los desarrollos KL expresan los campos aleatorios como una serie infinita de funciones ortogonales ponderadas por variables aleatorias [34]. En principio, el problema inverso a plantear para determinar tales campos sería infinito dimensional. El objetivo sería caracterizar una probabilidad a posteriori

$$\pi(du) \propto \exp(\phi(u))\pi_0(du) \quad (6.5)$$

donde u es una función de cuadrado integrable en un dominio de \mathbb{R}^L , $\phi(u)$ es el logaritmo de funcional de verosimilitud y $\pi_0 = \mathcal{N}(0, C)$ es una distribución a priori Gaussiana [34]. Bajo una distribución π_0 de este tipo, los campos aleatorios se descomponen en la forma $\sum_{i=1}^{\infty} \sqrt{\lambda_i} \phi_i Z_i$, donde Z_i son variables Gaussianas de media cero y desviación unidad. Bajo la distribución posterior π , Z_i pasan a estar distribuidas según esta distribución que ha de ser aproximada mediante muestreo. En implementaciones prácticas, este problema infinito dimensional se trunca a formulaciones finito dimensionales que han de tener un comportamiento consistente según el número de términos preservado tiende a infinito.

En esta sección, formularemos el problema inverso bayesiano finito dimensional, en términos de aproximaciones de los campos aleatorios que definen los campos de densidad y de velocidad de propagación de onda, con el objetivo de identificar realizaciones que maximicen la probabilidad a posteriori y de estimar la incertidumbre en torno a ellos. Para abordar el problema inverso bayesiano en este contexto [97] conviene tener en mente los siguientes puntos

1. **Objetivo del problema inverso:** el objetivo es inferir los campos de velocidad $v_p(\mathbf{x})$ y densidad $\rho(\mathbf{x})$ de toda o parte de la región en exploración, con incertidumbre cuantificada.
2. **Generación de datos:** a falta de datos reales con errores y geometrías suficientemente caracterizados, utilizamos datos sintéticos. Como se ha explicado en el capítulo anterior, resolvemos el problema directo usando campos deterministas de velocidad y de densidad conocidos en una geometría prefijada y con unas fuentes emisoras determinadas, evaluamos la solución numérica en una red de receptores dada y perturbamos los valores obtenidos con ruido de magnitud estipulada. De esa forma obtenemos los datos sintéticos \mathbf{d}^{obs} . Mantendremos en los ensayos numéricos los mismos datos sintéticos empleados en el capítulo anterior.

-
3. **Generación de campos aleatorios:** se generan muestras de los campos aleatorios de densidad y velocidad que aparecen como coeficientes en el modelo directo usando el método descrito en la sección 6.3.
 4. **Resolución del problema directo y operador de observación:** en este contexto, las magnitudes de interés en el método inverso bayesiano, que son el campo de velocidad de propagación y el campo de densidad, serán modelados como campos aleatorios $v_p(\mathbf{x}, \alpha)$ y $\rho(\mathbf{x}, \alpha)$. Mediante el procedimiento descrito en la sección 6.3, aproximamos estos campos en dimensión finita, en términos de una colección finita de variables aleatorias. Para cada realización de estos campos, podemos resolver el problema directo y construir el operador de observación $\mathbf{f}(v_p(\mathbf{x}, \alpha), \rho(\mathbf{x}, \alpha))$ evaluando la solución en los receptores.
 5. **Problema bayesiano inverso:** por el teorema de Bayes, expresamos la probabilidad a posteriori de tener unos campos de velocidad y densidad generados por una colección determinada de variables aleatorias dados los datos medidos a partir de la verosimilitud y la probabilidad a priori. La solución del problema bayesiano inverso es precisamente esta probabilidad.
 6. **Muestreo de campos:** Para caracterizar la probabilidad a posteriori se recurre a generar muestras de los campos distribuidas según ella. Es preciso recurrir a métodos de cadenas de Markov Monte Carlo específicos de problemas de alta dimensión o infinito dimensionales. Recurriremos al muestreador funcional de conjuntos (FES) especialmente diseñado para este tipo de problemas.

Detallamos a continuación los puntos principales.

6.3. Construcción de campos aleatorios de velocidad y densidad asociados a un mallado

Como hemos comentado, en este capítulo, consideraremos la densidad y velocidad como campos aleatorios. Dada la complejidad/elevado coste del cálculo y de autovalores y autovectores para el problema espectral de Karhunen-Loève, una alternativa eficaz desde el punto de vista práctico consiste en recurrir a descomposiciones ortogonales asociadas a matrices de covarianza definidas sobre mallados.

Fijado un mallado, por ejemplo el mallado a utilizar para la resolución del problema directo, tendremos la matriz \mathbf{C}_{η_ρ} para el campo de densidad, y para el campo de velocidad de propagación, tenemos $\mathbf{C}_{\eta_{v_p}}$. Estas matrices se construyen en función de la distancia euclidiana entre los puntos del mallado bidimensional, $r_{m,k}$ con $m, k = 1, \dots, N$. En las pruebas de inversión, el mallado bidimensional que utilizaremos será equiespaciado, por cuestiones de coste computacional. Para la obtención de las matrices de covarianza de cada campo utilizaremos la expresión (6.2). La elección de los parámetros ℓ , η y σ depende de la información que tengamos sobre los campos. Este es un punto muy difícil, y es preciso ser consciente de que una elección inadecuada impone en todo el proceso el trabajo con campos de velocidad y densidad poco adecuados [24].

Para ilustrar el método, y por simplicidad de evaluación de la matriz de correlación, consideramos un factor de suavizado $\eta = 0.5$ y una longitud de correlación $l = 1$. Dado que el factor de suavizado es $\eta = 0.5$, si reemplazamos por este valor en la ecuación (6.2), la expresión se reduce a una exponencial, por lo que la covarianza Matern \mathbf{C}_{η_ρ} y $\mathbf{C}_{\eta_{v_p}}$ se calculará utilizando las expresiones (6.6) y (6.7),

$$C_{\eta_\rho}(m, k) = \sigma_\rho^2 \exp\left(\frac{-r_{m,k}}{l}\right), \quad (6.6)$$

$$C_{\eta_{v_p}}(m, k) = \sigma_{v_p}^2 \exp\left(\frac{-r_{m,k}}{l}\right), \quad (6.7)$$

donde σ_ρ^2 y $\sigma_{v_p}^2$ representan a las varianzas de los campos de fondo ρ y v_p , respectivamente, $r_{m,k}$ es la distancia euclidiana entre los puntos m y k en el mallado bidimensional (x_{i_m}, y_{j_m}) y (x_{i_k}, y_{j_k}) , es decir, $r_{m,k} = \sqrt{(x_{i_m} - x_{i_k})^2 + (y_{j_m} - y_{j_k})^2}$.

En el caso modelo de domo salino, las varianzas del campo de la densidad y de la velocidad que aparecen en las ecuaciones (6.6) y (6.7) se calculan usando los datos de la densidad $\rho(\mathbf{x})$ y de la velocidad de progación $v_p(\mathbf{x})$ correspondiente al fondo estratificado incorporando el domo salino, y sin la incorporación de las anomalías de gas y de petróleo (ver las propiedades en la tabla 5.1),

$$\sigma_\rho^2 = \frac{1}{D} \sum_{n=1}^D (\rho_n - \bar{\rho})^2, \quad \sigma_{v_p}^2 = \frac{1}{D} \sum_{n=1}^D (v_{p,n} - \bar{v}_p)^2,$$

donde D es el número de puntos en el mallado bidimensional uniforme, $\bar{\rho}$ y \bar{v}_p corresponden al valor promedio de la densidad $\rho(\mathbf{x})$ y de la velocidad de propagación $v_p(\mathbf{x})$. Las varianzas calculadas son $\sigma_\rho^2 = 0'3401$ y $\sigma_{v_p}^2 = 2'6239$.

A continuación, construimos la matriz de covarianza Matérn total para ambos parámetros de interés, $\mathbf{C}_{\eta,2D}$, donde se concatenan las matrices \mathbf{C}_{η_ρ} y $\mathbf{C}_{\eta_{v_p}}$ en bloques diagonales, es decir

$$\mathbf{C}_{\eta,2D} = \begin{bmatrix} \mathbf{C}_{\eta_\rho} & 0 \\ 0 & \mathbf{C}_{\eta_{v_p}} \end{bmatrix}. \quad (6.8)$$

De las matrices de covarianza \mathbf{C}_{η_ρ} y $\mathbf{C}_{\eta_{v_p}}$ se obtienen los valores propios $\lambda_{\rho,i}$ y $\lambda_{v_p,i}$, y los autovectores ortonormales $\phi_{\rho,i}$ y $\phi_{v_p,i}$ correspondientes al campo de densidad y de velocidad, respectivamente. Los valores propios se ordenan de mayor a menor y se pueden seleccionar los primeros M valores propios y funciones propias de modo que

$$\mathbf{k}_\rho \approx \sum_{i=1}^M \sqrt{\lambda_{\rho,i}} \phi_{\rho,i} Z_i, \quad (6.9)$$

$$\mathbf{k}_{v_p} \approx \sum_{i=1}^M \sqrt{\lambda_{v_p,i}} \phi_{v_p,i} Z_i. \quad (6.10)$$

proporcionen una buena aproximación porque el resto de términos es de pequeño tamaño. Según se observa en [34], los autovalores de estas covarianzas decaen rápido. Aquí, \mathbf{k}_ρ y \mathbf{k}_{v_p} son vectores que representan campos aleatorios de media cero con covarianza dada evaluados en los nodos ordenados del mallado y Z_i corresponde a muestras de una distribución Gaussiana estándar, que son independientemente e idénticamente distribuidas (i.i.d.). En las figuras 6.1a y 6.1b se muestran campos obtenidos usando (6.9)-(6.10) con $M = 10$ términos.

Para construir campos aleatorios que representen la densidad y velocidad incorporando nuestra información a priori sobre los valores en los distintos estratos consideramos esta información como media y le sumamos los campos aleatorios de media cero generados. En el mallado bidimensional uniforme, evaluamos la densidad $\rho(\mathbf{x})$ y la velocidad $v_p(\mathbf{x})$ del medio estratificado sin incluir las anomalías del costado del domo salino (trampas de petróleo y gas), es decir, sólo considerando el fondo estratificado junto con la estructura de domo salino (ver tabla 5.1). A estos campos deterministas se les añadirá la componente estocástica generada a través de la covarianza Matérn $\mathbf{C}_{\eta,2D}$ según (6.9) y (6.10) para obtener $\mathbf{k}_{0\rho}$ y \mathbf{k}_{0v_p} .

Los campos resultantes pueden alcanzar valores negativos, cuestión que se puede evitar en la práctica de diversas formas. Una opción consiste en recurrir a distribuciones de tipo log-normal, en las que $\rho = \exp(\tilde{\rho})$ y $v_p = \exp(\tilde{v}_p)$, y se trabaja con los desarrollos KL de $\tilde{\rho}$ y \tilde{v}_p . Otra opción al implementar el problema inverso consiste en truncar la probabilidad a cero si aparecen valores negativos.

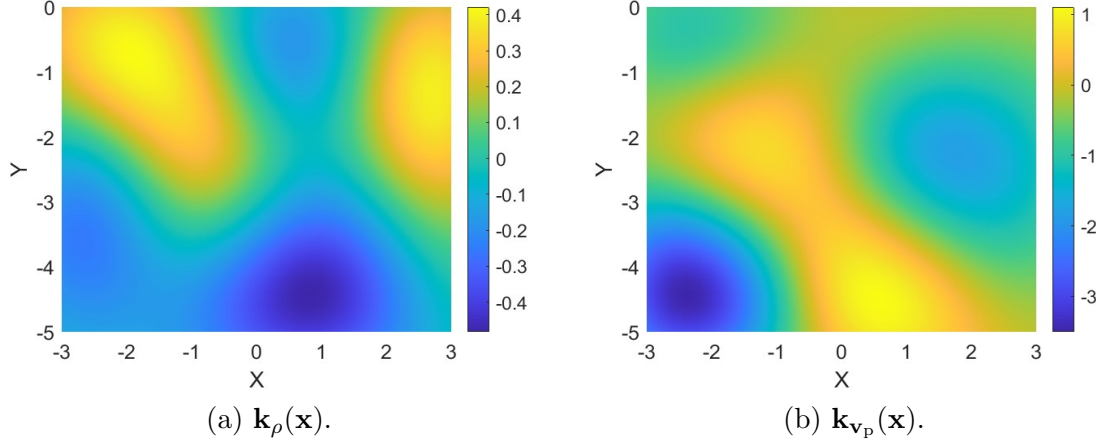


Figura 6.1: Campos aleatorios de media cero generados con covarianzas Matérn.

6.4. Representación de la información a priori

Sean $\bar{\mathbf{k}}_\rho$ y $\bar{\mathbf{k}}_{v_p}$ la densidad y la velocidad de propagación del medio estratificado de referencia (geometría de domo salino sin yacimientos a los costados) en nuestro caso modelo (geometría de domo salino con yacimientos) evaluadas en los puntos de un mallado uniforme ordenados $\mathbf{x}_1, \dots, \mathbf{x}_D$. Como hemos señalado en la sección 6.3, los campos

$$\mathbf{k}_\rho = \bar{\mathbf{k}}_\rho + \sum_{i=1}^D \sqrt{\lambda_{\rho,i}} \phi_{\rho,i} Z_i, \quad Z_i \text{ i.i.d.}, \quad (6.11)$$

$$\mathbf{k}_{v_p} = \bar{\mathbf{k}}_{v_p} + \sum_{i=1}^D \sqrt{\lambda_{v_p,i}} \phi_{v_p,i} W_i, \quad W_i \text{ i.i.d.}, \quad (6.12)$$

proporcionan una realización aproximada de campos aleatorios de densidad y velocidad que incluye información disponible sobre el medio de referencia, así como restricciones sobre su variabilidad y regularidad a través de las covarianzas Matérn empleadas para el cálculo de los autovalores λ_i y las autofunciones ϕ_i . Sin embargo, pueden tomar valores negativos. Podemos optar por dos estrategias distintas para imponer la positividad, así como por formulaciones en términos de los vectores \mathbf{k} o \mathbf{m} .

6.4.1. Formulación en términos de los coeficientes \mathbf{m}

Dada una aproximación a priori $\mathbf{k}_{\rho,0}$ y $\mathbf{k}_{v_p,0}$, definimos $\mathbf{m}_0 = [\mathbf{m}_{\rho,0}, \mathbf{m}_{v_p,0}]$ ensamblando los coeficientes ordenados de los desarrollos primero de $\mathbf{k}_{\rho,0}$ y a continuación de $\mathbf{k}_{v_p,0}$. En total, se tienen $2D$ elementos para un mallado con D puntos. Como se ha hecho en las secciones anteriores, podemos definir la probabilidad a priori como una distribución normal truncada con media $[\mathbf{m}_{\rho,0}, \mathbf{m}_{v_p,0}]$

$$p_{\text{prior}}(\mathbf{m}) = \frac{1}{\sqrt{(2\pi)^D \det(\mathbf{\Gamma}_{\text{prior}})}} \exp\left(-\frac{1}{2}(\mathbf{m} - \mathbf{m}_0)^\top \mathbf{\Gamma}_{\text{prior}}^{-1} (\mathbf{m} - \mathbf{m}_0)\right), \quad (6.13)$$

cuando los campos \mathbf{k} dados por los coeficientes \mathbf{m} no toman valores negativos, y $p_{\text{prior}}(\mathbf{m}) = 0$ en otro caso, siendo $\mathbf{\Gamma}_{\text{prior}}$ la matriz de covarianza a priori. Por la definición de los vectores \mathbf{m} en relación a los desarrollos (6.11)-(6.12), tenemos que $\mathbf{\Gamma}_{\text{prior}} = \mathcal{N}(0, \mathbf{I})$. En el caso modelo en estudio, tomaremos $\mathbf{m}_0 = 0$.

Dadas muestras $\mathbf{m} = [\mathbf{m}_\rho, \mathbf{m}_{v_p}]$, los campos $\mathbf{k}_\rho(\mathbf{m})$ y $\mathbf{k}_{v_p}(\mathbf{m})$ a introducir en el problema directo asociado vendrían dados por

$$\mathbf{k}_\rho(\mathbf{m}) = \bar{\mathbf{k}}_\rho + \sum_{i=1}^D \sqrt{\lambda_{\rho,i}} \phi_{\rho,i} \mathbf{m}_{\rho,i}, \quad (6.14)$$

$$\mathbf{k}_{v_p}(\mathbf{m}) = \bar{\mathbf{k}}_{v_p} + \sum_{i=1}^D \sqrt{\lambda_{v_p,i}} \phi_{v_p,i} \mathbf{m}_{v_p,i}. \quad (6.15)$$

Si se desea trabajar con campos forzosamente positivos, se puede escribir

$$\begin{aligned} \tilde{\mathbf{k}}_{\rho,0} &= \ln(\bar{\mathbf{k}}_\rho) + \sum_{i=1}^D \sqrt{\tilde{\lambda}_{\rho,i}} \tilde{\phi}_{\rho,i} \tilde{Z}_i, \quad \tilde{Z}_i \text{ i.i.d.}, \\ \tilde{\mathbf{k}}_{v_p,0} &= \ln(\bar{\mathbf{k}}_{v_p}) + \sum_{i=1}^D \sqrt{\tilde{\lambda}_{v_p,i}} \tilde{\phi}_{v_p,i} \tilde{W}_i, \quad \tilde{W}_i \text{ i.i.d.}, \end{aligned}$$

asumiendo que son los logaritmos de la densidad y de la velocidad quienes siguen distribuciones Matérn con desviaciones $\tilde{\sigma}_\rho = \ln(\sigma_\rho)$ y $\tilde{\sigma}_{v_p} = \ln(\sigma_{v_p})$, con autovalores $\tilde{\lambda}_{\rho,i}$, $\tilde{\lambda}_{v_p,i}$, y autovectores $\tilde{\phi}_{\rho,i}$, $\tilde{\phi}_{v_p,i}$, respectivamente. Los autovectores resultan ser los mismos que en el caso anterior. Los autovalores están relacionados por $\tilde{\lambda}_{\rho,i} = \lambda_{\rho,i} \frac{\tilde{\sigma}_\rho^2}{\sigma_\rho^2}$ y $\tilde{\lambda}_{v_p,i} = \lambda_{v_p,i} \frac{\tilde{\sigma}_{v_p}^2}{\sigma_{v_p}^2}$. En este caso, construimos $\mathbf{m}_0 = [\tilde{\mathbf{m}}_{\rho,0}, \tilde{\mathbf{m}}_{v_p,0}]$ ensamblando los coeficientes ordenados de los desarrollos primero de $\tilde{\mathbf{k}}_{\rho,0}$ y a continuación de $\tilde{\mathbf{k}}_{v_p,0}$.

La probabilidad a priori se define otra vez como (6.18), pero sin truncamiento ya que los campos a introducir en el problema directo serán positivos. Dadas muestras $\mathbf{m} = [\tilde{\mathbf{m}}_\rho, \tilde{\mathbf{m}}_{v_p}]$ se generan los campos de densidad y velocidad a través de las fórmulas:

$$\tilde{\mathbf{k}}_\rho = \ln(\bar{\mathbf{k}}_\rho) + \frac{\tilde{\sigma}_\rho}{\sigma_\rho} \sum_{i=1}^D \sqrt{\lambda_{\rho,i}} \phi_{\rho,i} \tilde{\mathbf{m}}_{\rho,i}, \quad \mathbf{k}_\rho(\mathbf{m}) = \exp(\tilde{\mathbf{k}}_\rho), \quad (6.16)$$

$$\tilde{\mathbf{k}}_{v_p} = \ln(\bar{\mathbf{k}}_{v_p}) + \frac{\tilde{\sigma}_{v_p}}{\sigma_{v_p}} \sum_{i=1}^D \sqrt{\lambda_{v_p,i}} \phi_{v_p,i} \tilde{\mathbf{m}}_{v_p,i}, \quad \mathbf{k}_{v_p}(\mathbf{m}) = \exp(\tilde{\mathbf{k}}_{v_p}), \quad (6.17)$$

de modo que $\mathbf{k}_\rho(\mathbf{m})$ y $\mathbf{k}_{v_p}(\mathbf{m})$ serían los valores a introducir en la resolución del problema directo asociado.

6.4.2. Formulación en términos de los valores de los campos \mathbf{k}

Alternativamente, se puede trabajar directamente con correcciones $\hat{\mathbf{k}}$ a los campos de referencia $\bar{\mathbf{k}}$. Para $\hat{\mathbf{k}}_0 = 0$ y $\mathbf{I}_{\text{prior}} = \mathbf{C}_{\eta,2D}$, la probabilidad a priori queda definida como

$$p_{\text{prior}}(\hat{\mathbf{k}}) = \frac{1}{\sqrt{(2\pi)^D \det(\mathbf{C}_{\eta,2D})}} \exp\left(-\frac{1}{2} \hat{\mathbf{k}}^\top \mathbf{C}_{\eta,2D}^{-1} \hat{\mathbf{k}}\right), \quad (6.18)$$

cuando los campos $\bar{\mathbf{k}} + \hat{\mathbf{k}}$ no toman valores negativos y $p_{\text{prior}}(\hat{\mathbf{k}}) = 0$ en otro caso. Dadas muestras $\hat{\mathbf{k}} = [\hat{\mathbf{k}}_\rho, \hat{\mathbf{k}}_{v_p}]$, los campos a introducir en el problema directo asociado vendrían dados por $\mathbf{k}_\rho = \bar{\mathbf{k}}_\rho + \hat{\mathbf{k}}_\rho$ y $\mathbf{k}_{v_p} = \bar{\mathbf{k}}_{v_p} + \hat{\mathbf{k}}_{v_p}$.

Si se desea trabajar con campos forzosamente positivos, se puede trabajar con distribuciones log-normal de modo que

$$p_{\text{prior}}(\tilde{\mathbf{k}}) = \frac{1}{\sqrt{(2\pi)^D \det(\tilde{\mathbf{C}}_{\eta,2D})}} \exp\left(-\frac{1}{2}\tilde{\mathbf{k}}^\top \tilde{\mathbf{C}}_{\eta,2D}^{-1} \tilde{\mathbf{k}}\right), \quad (6.19)$$

asumiendo que son los logaritmos de la densidad y de la velocidad quienes siguen distribuciones Matérn $\tilde{\mathbf{C}}_{\eta,2D}$ con desviaciones $\tilde{\sigma}_\rho = \ln(\sigma_\rho)$ y $\tilde{\sigma}_{v_p} = \ln(\sigma_{v_p})$. Dadas muestras $\tilde{\mathbf{k}} = [\tilde{\mathbf{k}}_\rho, \tilde{\mathbf{k}}_{v_p}]$, los campos a introducir en el problema directo asociado vendrían dados por $\mathbf{k}_\rho = \bar{\mathbf{k}}_\rho \exp(\tilde{\mathbf{k}}_\rho)$ y $\mathbf{k}_{v_p} = \bar{\mathbf{k}}_{v_p} \exp(\tilde{\mathbf{k}}_{v_p})$.

6.5. Definición del operador de observación y la verosimilitud

Dadas realizaciones \mathbf{k}_ρ y \mathbf{k}_{v_p} de los campos aleatorios de densidad y velocidad, se resuelve el problema directo según se ha descrito en el capítulo 5 utilizando el mallado uniforme equiespaciado y se evalúa la solución numérica en los receptores elegidos. En la sección 6.4 hemos discutido distintas parametrizaciones $\boldsymbol{\nu}$ posibles y la forma de obtener los campos $\mathbf{k} = \mathbf{k}(\boldsymbol{\nu})$ a partir de ellas. Resolviendo el problema directo para esos campos, obtenemos el operador de observación numérico $\mathbf{f}(\boldsymbol{\nu})$, asociado a cada realización. Se mantiene la geometría, los esquemas de discretización y la elección de parámetros del capítulo anterior en los ensayos numéricos de este capítulo. Preservamos igualmente los datos sintéticos \mathbf{d}^{obs} . Utilizaremos los datos sintéticos generados para la geometría exacta de domo salino resolviendo el problema directo en un mallado adaptado a esa geometría y añadiendo ruido, según se detalla en la sección 5.2.1 del capítulo 5.

Dados los datos y elegida una parametrización de la densidad y velocidad, se define la verosimilitud a través de la expresión

$$p(\mathbf{d}^{\text{obs}} | \boldsymbol{\nu}) = \frac{1}{\sqrt{(2\pi)^{P/2} \det(\boldsymbol{\Gamma}_{\text{ruido}})}} \exp\left(-\frac{1}{2}(\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}})^\top \boldsymbol{\Gamma}_{\text{ruido}}^{-1} (\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}})\right), \quad (6.20)$$

donde:

- $\mathbf{f}(\boldsymbol{\nu})$ es el operador de observación numérico definido,
- $\boldsymbol{\nu}$ son los coeficientes que permiten definir los valores de los campos de densidad $\mathbf{k}_\rho(\boldsymbol{\nu})$ y velocidad $\mathbf{k}_{v_p}(\boldsymbol{\nu})$ en los puntos ordenados del mallado a través de desarrollos de tipo KL,
- \mathbf{d}^{obs} son los datos sintéticos, un vector de longitud P ,
- $\boldsymbol{\Gamma}_{\text{ruido}}$ es la matriz de covarianza del ruido en las observaciones, ya introducida en el capítulo anterior como una matriz diagonal con diagonal constante σ_{ruido}^2 .

6.6. Definición de la probabilidad a posteriori

Elegida una parametrización caracterizada por coeficientes $\boldsymbol{\nu}$, el teorema de Bayes en nuestro contexto proporciona la relación,

$$p_{\text{posterior}}(\boldsymbol{\nu} | \mathbf{d}^{\text{obs}}) = \frac{p(\mathbf{d}^{\text{obs}} | \boldsymbol{\nu}) p_{\text{prior}}(\boldsymbol{\nu})}{p(\mathbf{d}^{\text{obs}})}, \quad (6.21)$$

donde:

- $p(\mathbf{d}^{\text{obs}} | \boldsymbol{\nu})$ es la verosimilitud o probabilidad condicionada al operador observable, que modela cómo se generan los datos observacionales a partir de campos dados,

- $p_{\text{prior}}(\boldsymbol{\nu})$ es la distribución a priori, que incluye información inicial sobre los parámetros en estudio, en este caso, incluye información inicial sobre los campos de la densidad y de la velocidad de propagación,
- $p(\mathbf{d}^{\text{obs}})$ es un factor de normalización y no depende de los parámetros desconocidos.

En la expresión de la probabilidad a posteriori $p_{\text{posterior}}(\boldsymbol{\nu} \mid \mathbf{d}^{\text{obs}})$ se combinan las expresiones verosimilitud y la probabilidad a priori definidas en la sección 6.4, obteniendo la expresión usual

$$p_{\text{posterior}}(\boldsymbol{\nu} \mid \mathbf{d}^{\text{obs}}) \propto \exp\left(-\frac{1}{2}\left[(\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}})^\top \boldsymbol{\Gamma}_{\text{ruido}}^{-1} (\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}}) + (\boldsymbol{\nu} - \boldsymbol{\nu}_0)^\top \boldsymbol{\Gamma}_{\text{prior}}^{-1} (\boldsymbol{\nu} - \boldsymbol{\nu}_0)\right]\right). \quad (6.22)$$

La probabilidad a posteriori de la expresión (6.22) puede aproximarse por $\exp(-J_{\text{reg}}(\boldsymbol{\nu}))$, en términos de una función de costo regularizada, es decir

$$J_{\text{reg}}(\boldsymbol{\nu}) = \frac{1}{2} \left\| \mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}^{\text{obs}} \right\|_{\boldsymbol{\Gamma}_{\text{ruido}}}^2 + \frac{1}{2} \left\| \boldsymbol{\nu} - \boldsymbol{\nu}_0 \right\|_{\boldsymbol{\Gamma}_{\text{prior}}}^2, \quad (6.23)$$

donde $\|\boldsymbol{\nu}\|_{\boldsymbol{\Gamma}} = \boldsymbol{\nu}^\top \boldsymbol{\Gamma}^{-1} \boldsymbol{\nu}$ representa la norma cuadrática asociada a la matriz de covarianza $\boldsymbol{\Gamma}$. En este caso, $\boldsymbol{\Gamma}_{\text{prior}}$ representa una matriz covarianza Matérn análoga a la estudiada en la sección 6.1.2. Si la información a priori proporcionada por $\boldsymbol{\nu}_0$ y $\boldsymbol{\Gamma}_{\text{prior}}$ es adecuada, se espera que la función de costo $J_{\text{reg}}(\boldsymbol{\nu})$ sea convexa, lo que podría indicar la presencia de un único mínimo global, gracias al efecto regularizador del término $\|\boldsymbol{\nu} - \boldsymbol{\nu}_0\|_{\boldsymbol{\Gamma}_{\text{prior}}}$, y en consecuencia, un único máximo de la probabilidad. A menudo esta información es escasa y la probabilidad a posteriori puede tener una estructura compleja.

En nuestro caso modelo, hemos considerado en la sección 6.4, cuatro formas de parametrizar ligeramente distintas que comparten el hecho de que $\boldsymbol{\nu}_0 = 0$. La información sobre los campos de densidad y velocidad de referencia se introduce al definir los campos que entran en el operador de observación a partir de la corrección dada por $\boldsymbol{\nu}$. En dos de ellas $\boldsymbol{\Gamma}_{\text{prior}} \sim \mathcal{N}(0, \mathbf{I})$ y en las otras dos se tienen covarianzas de Matérn por bloques con distinta varianza, calibrada al conocimiento de los campos reales. Todas las construcciones están gobernadas por la elección $\eta = 1/2$ (por simplicidad de evaluación de los núcleos) y $\ell = 1$, que puede necesitar ajuste para recoger variaciones en la escala que nos interesa.

Para explorar la estructura de la probabilidad a posteriori, se empleará el método de muestreo funcional de conjuntos (FES) junto con el método de cadenas de Markov Monte Carlo (MCMC). En este proceso, se generarán muestras que permiten explorar el espacio de parámetros de manera efectiva, utilizando la información contenida en la verosimilitud y en la distribución a priori.

6.7. Tratamiento computacional del problema bayesiano inverso

Para caracterizar la probabilidad a posteriori utilizamos método de Monte Carlo basados en cadenas de Markov (MCMC). Este enfoque permite explorar la distribución a posteriori de manera eficiente y aproximar los parámetros que definen los campos de densidad ρ y de velocidad de propagación v_p en la base espectral definida por la covarianza elegida, cuantificando la incertidumbre.

En el contexto de los métodos de Monte Carlo mediante cadenas de Markov discutidos en la sección *Cuantificación de incertidumbre mediante métodos de cadenas de Markov Monte Carlo* del capítulo 3, uno de los desafíos es manejar la aceptación de propuestas en el espacio de parámetros de alta dimensionalidad [15, 45]. En ese contexto, las propuestas generadas bajo el método Crank-Nicholson preconditionado (pCN) están diseñadas para mejorar la aceptación y explorar el espacio de parámetros de manera más efectiva en altas dimensiones.

El método pCN resuelve el problema de la convergencia a cero de las probabilidades de aceptación en MCMC cuando se trabaja con espacios de parámetros en alta dimensión. Al proponer movimientos específicos que se basan en la distribución a priori Gaussiana, el método pCN contribuye a que la probabilidad de aceptación se mantenga en niveles razonables. Esta propiedad de estabilidad asegura que, incluso al aumentar la dimensionalidad del espacio de parámetros, la probabilidad promedio de aceptación se mantenga alejada de cero [88].

El método de muestreo por combinación funcional (FES) mejora aún más el proceso de muestreo al combinar pCN con el muestreador de conjuntos de invariancia afín (AIES) [34]. FES utiliza AIES para las componentes de baja frecuencia de la expansión KL y pCN para los componentes de alta frecuencia [11]. Esta combinación permite que FES aproveche las fortalezas de ambos métodos: AIES proporciona una mayor eficiencia en subespacios de baja dimensión, mientras que pCN mantiene la estabilidad en espacios de alta dimensión. Como resultado, FES logra una mezcla más rápida y una mayor eficiencia en el muestreo de distribuciones posteriores complejas en espacios de alta dimensión [34].

A continuación, se detallan el método Crank-Nicolson preconditionado y el muestreador funcional de conjuntos, para luego integrarlos en el algoritmo que explorará la distribución posterior de los campos de densidad y de velocidad de propagación.

6.7.1. *Método Crank-Nicolson preconditionado como método Monte Carlo basado en cadenas de Markov*

El método Crank-Nicolson se utiliza en el contexto de MCMC para mejorar la eficiencia de las cadenas de Markov mediante la generación de propuestas que son aceptadas según una distribución Gaussiana previa. Este enfoque es particularmente útil en problemas de alta dimensión, donde las técnicas tradicionales pueden enfrentar problemas de baja probabilidad de aceptación [75], y explota correlaciones definidas por las matrices de covarianza en altas dimensiones para avanzar.

En el método pCN, la propuesta de movimiento desde una posición actual $\boldsymbol{\nu}$ a una nueva posición $\boldsymbol{\nu}_{\text{prop}}$ se define mediante [81],

$$\boldsymbol{\nu}_{\text{prop}} = \sqrt{1 - \omega^2} \boldsymbol{\nu} + \omega \boldsymbol{\xi}, \quad (6.24)$$

donde $\boldsymbol{\xi} \sim \mathcal{N}(0, \mathbf{C})$ es una muestra aleatoria de la distribución Gaussiana con matriz de covarianza \mathbf{C} , y $\omega \in (0, 1]$ es un parámetro de paso que controla la magnitud de la perturbación aplicada a $\boldsymbol{\nu}$. Cuando $\omega \rightarrow 1$, la propuesta representa una perturbación grande de $\boldsymbol{\nu}$, mientras que cuando ω es pequeño, la propuesta es una pequeña perturbación de la posición actual [38].

Dado que pCN utiliza una matriz de covarianza \mathbf{C} para generar propuestas, esta matriz puede ser directamente influenciada por la estructura de covarianza del campo aleatorio. Es decir, pCN puede aprovechar la estructura de la expansión de Karhunen-Loève para generar propuestas que sean consistentes con la estructura del campo aleatorio modelado. Por ejemplo, en el caso de campos aleatorios con núcleo Matérn, la matriz de covarianza \mathbf{C} de la ecuación (6.24) podría estar relacionada con los valores propios λ_i y las funciones propias ϕ_i obtenidas de la expansión de Karhunen-Loève. En este contexto, las propuestas generadas por pCN pueden ser interpretadas en términos de la estructura espectral del campo aleatorio.

El principal costo computacional de pCN proviene de la evaluación de la probabilidad de aceptación [34],

$$\alpha = \min \left\{ 1, \frac{L(\boldsymbol{\nu}_{\text{prop}})}{L(\boldsymbol{\nu})} \right\}, \quad (6.25)$$

donde $L(\boldsymbol{\nu}_{\text{prop}})$ representa la verosimilitud evaluada en $\boldsymbol{\nu}_{\text{prop}}$ y $L(\boldsymbol{\nu})$ representa la verosimilitud evaluada en $\boldsymbol{\nu}$. La evaluación de cantidades puede ser costosa dependiendo de la complejidad del modelo pues requiere la evaluación del operador de observación [34].

Las ventajas clave del muestreador pCN en el contexto de MCMC son las siguientes:

- **Propiedades de Estabilidad:** una de las características más importantes de pCN es su capacidad para mantener la probabilidad de aceptación de propuestas lejos de cero a medida que aumenta la dimensionalidad del espacio de aproximación [33].
- **Escalabilidad:** pCN es particularmente útil en espacios de alta dimensión. La estructura de la propuesta de movimiento asegura que, incluso en dimensiones altas, el método sigue siendo efectivo para explorar el espacio de parámetros [46].
- **Simplicidad Computacional:** pCN es un método simple de implementar que requiere sólo la generación de propuestas aleatorias y la evaluación de la función de verosimilitud, lo que lo hace relativamente barato en términos computacionales.

Sin embargo, también tiene limitaciones:

- **Convergencia Lenta:** la principal desventaja del método pCN es la convergencia lenta de las estadísticas cuando la distribución posterior está mal escalada o es multimodal. Aunque este método es efectivo para explorar el espacio de parámetros, en algunos casos puede necesitar ajustes finos para mejorar la convergencia [82].

6.7.2. Muestreador por combinación funcional (FES)

Para muestrear de manera eficiente en espacios de funciones, se propone un algoritmo que combina dos técnicas de muestreo: el muestreador por combinación de invariancia afín, o AIES estudiados en 3.4.4, y pCN. Este algoritmo, llamado muestreador por combinación funcional (FES), emplea un enfoque Metrópolis within Gibbs. Lo detallamos a continuación para variables de media nula. El muestreo se realiza en dos etapas:

1. **AIES:** se utiliza para las componentes de baja frecuencia de la expansión KL. Las componentes de baja frecuencia representan las variaciones de gran escala en el campo aleatorio. Se trabaja con W cadenas $\boldsymbol{\nu}_w$, $w = 1, \dots, W$ simultáneamente. En cada paso, se calcula una permutación sin repeticiones de todas ellas. Cada cadena w se actualiza mediante actualización la siguiente expresión,

$$\boldsymbol{\nu}_{\text{prop}_w}^i = \boldsymbol{\nu}_w^i + (1 - Z)\mathbf{P}(\boldsymbol{\nu}_{\text{per}(w)}^i - \boldsymbol{\nu}_w^i), \quad w = 1, \dots, W, \quad (6.26)$$

donde per denota la permutación, $Z \in [1/a, a]$ es un número aleatorio elegido con densidad $g(z) \propto 1/\sqrt{z}$ y a es un parámetro de tamaño de paso que típicamente se establece en $a = 2$ [34, 50]. La matriz $\mathbf{P} = \mathbf{J}\mathbf{J}^T$ donde \mathbf{J} es una matriz cuyas columnas son los autovectores de la covarianza \mathbf{C} correspondientes a los M mayores autovalores. Ponemos $\boldsymbol{\nu}_w^{i+1} = \boldsymbol{\nu}_{\text{prop}_w}^i$ con probabilidad [34],

$$\text{mín} \left\{ 1, Z^{M-1} \frac{p_{\text{posterior}}(\boldsymbol{\nu}_{\text{prop}_w}^i)}{p_{\text{posterior}}(\boldsymbol{\nu}_w^i)} \right\}, \quad (6.27)$$

en otro caso la cadena no cambia $\boldsymbol{\nu}_w^{i+1} = \boldsymbol{\nu}_w^i$.

2. **pCN:** se emplea para las componentes de alta frecuencia de la expansión KL. Las componentes de alta frecuencia capturan las variaciones de pequeña escala en el campo aleatorio. Efectuada

la etapa anterior para todas las cadenas, una nueva propuesta de actualización para cada una se lleva a cabo según la relación

$$\boldsymbol{\nu}_{\text{prop}_w}^{i+1} = \mathbf{P}\boldsymbol{\nu}_w^{i+1} + \mathbf{Q} \left(\sqrt{1 - \omega^2} \boldsymbol{\nu}_w^{i+1} + \omega \boldsymbol{\xi} \right), \quad w = 1, \dots, W, \quad (6.28)$$

donde $\boldsymbol{\xi} \sim \mathcal{N}(0, C)$, $\mathbf{Q} = \mathbf{I} - \mathbf{J}\mathbf{J}^T$. Ponemos $\boldsymbol{\nu}_w^{i+2} = \boldsymbol{\nu}_{\text{prop}_w}^{i+1}$ con probabilidad [34],

$$\text{mín} \left\{ 1, \frac{L(\boldsymbol{\nu}_{\text{prop}_w}^{i+1})}{L(\boldsymbol{\nu}_w^{i+1})} \right\} \quad (6.29)$$

siendo L la verosimilitud. En otro caso la cadena no cambia $\boldsymbol{\nu}_w^{2i+1} = \boldsymbol{\nu}_w^{i+1}$.

El proceso concluye en un número prefijado de etapas máximas I y la colección de valores de \mathbf{m} generadas entre todas las cadenas y etapas constituyen las muestras que se utilizan para explorar el comportamiento de la probabilidad. En el caso en que M sea igual a la dimensión de la matriz de covarianza, el muestreador se reduce a AIES y la segunda etapa se suprime.

Al combinar AIES y pCN de esta manera, el muestreador FES aprovecha las fortalezas de cada método para diferentes partes del espacio de funciones, logrando un muestreo más eficiente y efectivo. FES es un método novedoso que mejora el enfoque estándar de pCN. En el muestreador FES podemos ajustar el parámetro M para lograr un cálculo numérico más rápido [33]. El parámetro M controla cuántas coordenadas KL se incluyen en el muestreo de AIES. Cuando $M = 0$, no se realiza muestreo de AIES, por lo que el algoritmo se reduce a pCN. Sin embargo, si M supera los 20, el rendimiento se puede deteriorar, a menudo sólo es eficiente para subespacios de dimensión 20 o menor [34]. El número preciso de coordenadas KL a incluir en el muestreo de AIES es una decisión de ajuste, con el número óptimo dependiendo del problema de estimación.

En la sección 6.8 se muestran los principales pasos del pseudocódigo del problema inverso bayesiano basado en expansiones KL. En este algoritmo incorporamos el preconditionador pCN y el muestreador AIES, siendo nuestro objetivo es encontrar los parámetros de interés $\boldsymbol{\nu}$ que garanticen la convergencia de la probabilidad posterior $p_{\text{posterior}}(\boldsymbol{\nu})$, utilizando p_{prior} como la probabilidad inicial y $p(\mathbf{d}^{\text{obs}}|\boldsymbol{\nu})$ como la probabilidad de verosimilitud o condicionada al operador de observación y a los datos sintéticos.

6.8. Pseudocódigo de Problema inverso aproximado basado en KL incorporando pCN y FES

En esta sección, aplicamos las metodologías previamente descritas para resolver un problema inverso bayesiano. El objetivo es inferir las propiedades de los parámetros de interés, que en este caso corresponden a los campos de densidad y velocidad de propagación en de onda.

Para ello, primero se necesitará resolver el problema directo de la ecuación de ondas con los campos de densidad y de velocidad deterministas verdaderos, en un mallado adaptado, para obtener los datos sintéticos \mathbf{d}^{obs} , tal como explicamos en el capítulo 5. Los datos sintéticos \mathbf{d}^{obs} se utilizan para calcular la probabilidad condicional o de verosimilitud.

A continuación, construimos el mallado uniforme y ordenamos los nodos (x_q, y_s) de este mallado $q = 1, \dots, Q$, $s = 1, \dots, S$ siguiendo la numeración $d(q, s) = (s - 1)S + q$, para generar la lista \mathbf{x}_d , $d = 1, \dots, D$. El nodo d corresponde a la posición (x_q, y_s) , donde q es el resto de dividir D entre S y $s - 1$ es el cociente. Evaluamos en el mallado uniforme los campos de densidad y de velocidad de la geometría de domos salino sin yacimientos de gas y petróleo a los costados y ordenamos el resultado igual que los nodos, obteniendo los vectores $\bar{\mathbf{k}}_\rho$ y $\bar{\mathbf{k}}_{v_p}$. Como en secciones anteriores, utilizaremos la

notación $\mathbf{k} = [\mathbf{k}_\rho, \mathbf{k}_{v_p}]$ para referirnos al vector de valores de los campos en los nodos ordenados concatenados siguiendo ese orden.

El siguiente paso es calcular las matrices de covarianza \mathbf{C}_{η_ρ} y $\mathbf{C}_{\eta_{v_p}}$ para el mallado uniforme elegido. Tras ello, hallamos sus autovalores y autovectores. Ordenando los autovalores de mayor a menor definimos el orden en el que aparecen autovalores y autovectores en los desarrollos (6.14)-(6.15) de \mathbf{k}_ρ y \mathbf{k}_{v_p} . Si bien en la sección 6.4 mencionamos distintos tipos de parametrización, las parametrizaciones que hemos considerado se basan en la sección 6.4.1, donde dadas las muestras \mathbf{m} calculamos los campos \mathbf{k}_ρ y \mathbf{k}_{v_p} con las expresiones (6.16) y (6.17) respectivamente, para luego introducirlos en el problema directo y obtener los operadores de observación asociados a los campos $\mathbf{k}(\boldsymbol{\nu})$ resultantes.

A los vectores $\boldsymbol{\nu}_w^0$ se les aplica la etapa AIES de FES en primer lugar. Al resultado se le aplica la etapa pCN, y así sucesivamente hasta alcanzar el número de etapas fijadas. La convergencia a la probabilidad muestreada se valida mediante criterios de tipo Gelman-Rubin [47]. Descartadas las muestras obtenidas para las cadenas en B etapas iniciales como periodo de equilibrado, el resto nos proporciona una colección de muestras a partir de las cuales identificar los coeficientes KL, y por tanto, los campos de densidad y velocidad, de máxima probabilidad y la incertidumbre en torno a ellos.

A continuación, detallamos los algoritmos utilizados para estimar los coeficientes de los campos aleatorios de densidad y de velocidad de propagación solución del problema inverso. Los algoritmos presentados a continuación permiten una estimación de las propiedades del medio utilizando un enfoque inverso bayesiano en alta dimensión con el método de muestreo FES.

Algorithm 3 Inicialización: generación de datos sintéticos

- 1: **Paso 1:** Crear un mallado espacial en $[-3, 3] \times [-5, 0]$ adaptado a la geometría real del domo salino con yacimientos.
- 2: **Paso 2:** Crear un mallado temporal con paso $h_x = h_y = 0.2$ en el intervalo $x \in [-3, 3]$ e $y \in [-5, 0]$.
- 3: **Paso 3:** Definir los emisores, las fuentes y los receptores. Las fuentes utilizadas van desde $-1'76$ hasta $1'76$, siendo en total 14 emisores. Los receptores utilizados van desde $-1'68$ hasta $1'68$, considerando un total de 13 receptores.
- 4: **Paso 4:** Evaluar los campos de densidad y velocidad en los puntos del mallado espacial, según los datos reales de esta geometría, y construir las matrices asociadas a la discretización espacial de elementos finitos.
- 5: **Paso 5:** Resolver el problema directo mediante esquema explícito de diferencias finitas en tiempo.
- 6: **Paso 6:** Evaluar la solución numérica para calcular el operador de observación $\mathbf{f}(\boldsymbol{\nu}_{\text{exacto}})$ en los receptores r_j , $j = 1, \dots, J$ y en la red de tiempos t_ℓ , $\ell = 1, \dots, L$.
- 7: **Paso 7:** Obtener los datos sintéticos \mathbf{d}^{obs} usando la expresión,

$$\mathbf{d}_{j,\ell}^{\text{obs}} = \mathbf{f}(\boldsymbol{\nu}_{\text{exacto}})_{j,\ell} + \alpha \frac{\sigma_{\text{ruido}}}{100} r,$$

donde $\alpha = 5$ corresponde al nivel porcentual de ruido de 5%, $\sigma_{\text{ruido}} = \sqrt{\frac{\sum_{j=1}^J \sum_{\ell=0}^L |U(r_j, t_\ell)|^2}{JL}}$, siendo J el número de receptores, que en este caso son 13, y L el número de datos temporales por receptor que en este caso son 10,000, y $r \sim \mathcal{N}(0, 1)$.

Algorithm 4 Inicialización: información de referencia

- Paso 1:** Crear un mallado equiespaciado en $[-3, 3] \times [-5, 0]$ con $hx = hy = 0.2$, generando $D = 806$ puntos \mathbf{x}_d , $d = 1, \dots, D$ ordenados como sigue. El punto (x_q, y_s) , $q = 1, \dots, Q$, $s = 1, \dots, S$ ocupa el puesto $d(q, s) = (s - 1)S + q$. Inversamente, el nodo d corresponde a la posición (x_q, y_s) , donde q es el resto de dividir D entre S y $s - 1$ es el cociente.
- 2: **Paso 2:** Evaluar los campos de densidad y velocidad de la geometría de referencia (domo salino sin yacimientos laterales) en los nodos del mallado uniforme para obtener los vectores $\bar{\mathbf{k}}_\rho$ y $\bar{\mathbf{k}}_v$.
- Paso 3:** Generar las matrices de covarianza Matérn \mathbf{C}_{η_ρ} y $\mathbf{C}_{\eta_{v_p}}$, con (6.6) y (6.7), respectivamente.
- 4: **Paso 4:** Formar la matriz de covarianza conjunta $\mathbf{C}_{\eta, 2D}$ mediante (6.8).
- Paso 5:** Obtener autovalores y autovectores de \mathbf{C}_{η_ρ} y $\mathbf{C}_{\eta_{v_p}}$.
- 6: **Paso 6:** Ordenar autovalores de \mathbf{C}_{η_ρ} y $\mathbf{C}_{\eta_{v_p}}$ de mayor a menor, y reordenar los autovectores siguiendo el mismo orden en cada bloque. Nótese que la ordenación de autovalores y el valor de los autovectores no cambia si usamos $\tilde{\mathbf{C}}_{\eta_\rho}$ y $\tilde{\mathbf{C}}_{\eta_{v_p}}$.
- Paso 7:** Almacenar los primeros M autovectores de tamaño relevante de cada matriz, ordenados, en las columnas de matrices \mathbf{J}_ρ y \mathbf{J}_{v_p} .
-

Algorithm 5 Muestreador FES

- 1: **Paso 1:** Calcular las matrices $P_\rho = \mathbf{J}_\rho \mathbf{J}'_\rho$ y $P_{v_p} = \mathbf{J}_{v_p} \mathbf{J}'_{v_p}$. Crear la matriz bloque diagonal,

$$P = \begin{bmatrix} P_\rho & 0 \\ 0 & P_{v_p} \end{bmatrix}$$

y calcular $\mathbf{Q} = \mathbf{I} - \mathbf{P}$, donde \mathbf{I} es una matriz diagonal identidad.

- 2: **Paso 2:** Fijar la matriz de covarianza \mathbf{C} como $\mathbf{C}_{\eta, 2D}$ o $\tilde{\mathbf{C}}_{\eta, 2D}$, según parametrización elegida.
- 3: **Paso 3:** Seleccionar el número de cadenas W e inicializar $\boldsymbol{\nu}_w$ muestreando $\mathcal{N}(0, \mathbf{C})$ para $w = 1, \dots, W$.
- 4: **Paso 4:** Calcular los campos de densidad y velocidad asociados $\mathbf{k}(\boldsymbol{\nu})_w = \bar{\mathbf{k}} + \boldsymbol{\nu}_w$ o $\mathbf{k}(\boldsymbol{\nu})_w = \bar{\mathbf{k}} \exp(\boldsymbol{\nu}_w)$, según parametrización elegida, para $w = 1, \dots, W$.
- 5: **Paso 5:** Tomar $\boldsymbol{\nu} = 0$ y calcular el logaritmo de la distribución a priori logprior_w :

$$\text{logprior}_w = -\frac{\boldsymbol{\nu}_w^T \mathbf{C}^{-1} \boldsymbol{\nu}_w}{2}.$$

- 6: **Paso 6:** Resolver el problema directo con campos $\mathbf{k}(\boldsymbol{\nu})_w$ para $w = 1, \dots, W$ y evaluar en los receptores en los tiempos de medida para determinar el operador de observación $\mathbf{f}(\boldsymbol{\nu}_w)$.
- 7: **Paso 7:** Calcular el logaritmo de la probabilidad de verosimilitud loglike para $w = 1, \dots, W$:

$$\text{loglike}_w = -\frac{\|\mathbf{f}(\boldsymbol{\nu}_w) - \mathbf{d}^{\text{obs}}\|_2^2}{2\sigma_{\text{ruido}}^2}.$$

- 8: **Paso 8:** Calcular el logaritmo de la probabilidad posterior para $w = 1, \dots, W$

$$\text{logpost}_w = \text{logprior}_w + \text{loglike}_w.$$

- 9: **Paso 9:** $\boldsymbol{\nu}_w$ se llamará $\boldsymbol{\nu}_{\text{prev}_w}$ y se actualizará según el algoritmo FES. Asimismo, logpost_w se llamará logpostprev_w , loglike_w se llamará loglikeprev_w y logprior_w pasará a logpriorprev_w .
-

Algorithm 6 Muestreador FES (cont.)

10: **Paso 10:** Configurar el número de pasos I_{\max} .

11: **for** $i = 1$ **to** I_{\max} **do**

12: **Paso 10.1:** Calcular una permutación **per** sin repetición de $(1, 2, 3, \dots, W)$.

13: **Paso 10.2:** Generar los números aleatorios:

14: $\text{vecz} = ((a - 1) \times \text{rand}(1, W) + 1)^2 / a, a = 2,$

15: $\log\text{vecz} = \log(\text{vecz}).$

16: **Paso 10.3:** Aplicar AIES.

17: **for** $w = 1$ **to** W **do**

18: **Paso 10.3.1:** Proponer una actualización $\boldsymbol{\nu}_{\text{prop},w}$ usando AIES:

$$\boldsymbol{\nu}_{\text{prop},w} = \boldsymbol{\nu}_{\text{prev},w} + (1 - \text{vecz}_w) \mathbf{P}(\boldsymbol{\nu}_{\text{prev},\text{per}(w)} - \boldsymbol{\nu}_{\text{prev},w})$$

19: **Paso 10.3.2:** Calcular el logaritmo negativo de la probabilidad a priori de la propuesta

$$\log\text{prior}_w = -\frac{\boldsymbol{\nu}_w^T \mathbf{C}^{-1} \boldsymbol{\nu}_w}{2}.$$

20: **Paso 10.3.3:** Generar $\log r_w$ como $\log(\mathcal{U}(0, 1))$.

21: **if** $\log r_w < \log\text{priorprop}_w - \log\text{postprev}_w + \log\text{vecz}_w \times (2M - 1)$ **then**

22: **Paso 10.3.3.1:** Calcular los campos de densidad y velocidad asociados $\mathbf{k}(\boldsymbol{\nu}_{\text{prop},w}) = \bar{\mathbf{k}} + \boldsymbol{\nu}_{\text{prop},w}$
o $\mathbf{k}(\boldsymbol{\nu}_{\text{prop},w}) = \bar{\mathbf{k}} \exp(\boldsymbol{\nu}_{\text{prop},w})$, según parametrización elegida. Resolver el problema directo con campos $\mathbf{k}(\boldsymbol{\nu}_{\text{prop},w})$ y evaluar en los receptores en los tiempos de medida para determinar el operador de observación $\mathbf{f}(\boldsymbol{\nu}_{\text{prop},w})$.

23: **Paso 10.3.3.2:** Calcular el logaritmo negativo de la verosimilitud correspondiente

$$\log\text{likeprop}_w = -\frac{\|\mathbf{f}(\boldsymbol{\nu}_{\text{prop},w}) - \mathbf{d}^{\text{obs}}\|_2^2}{2\sigma_{\text{ruido}}^2}.$$

24: **Paso 10.3.3.3:** Calcular el logaritmo negativo de la probabilidad posterior

$$\log\text{postprop}_w = \log\text{priorprop}_w + \log\text{likeprop}_w.$$

25: **Paso 10.3.3.4:** Calcular el logaritmo negativo de la probabilidad de aceptación α

$$\log(\alpha) = \log\text{postprop}_w - \log\text{postprev}_w + \log\text{vecz}_w \times (2M - 1)$$

26: **Paso 10.3.3.5:** Aceptar o no la propuesta

27: **if** $\log r_w < \log(\alpha)$ **then**

28: $\boldsymbol{\nu}_{\text{prev},w} = \boldsymbol{\nu}_{\text{prop},w}$

29: $\log\text{postprev}_w = \log\text{postprop}_w$

30: $\log\text{likeprev}_w = \log\text{likeprop}_w$

31: **end if**

32: **end if**

33: **end for**

Algorithm 7 Muestreador FES (cont.)

34: **Paso 10.4:** Seleccionar $\omega = 0'02$ y aplicar pCN.

35: **for** $w = 1$ **to** W **do**

36: **Paso 10.4.1:** Proponer la actualización

$$\boldsymbol{\nu}_{\text{prop},w} = \mathbf{P} \boldsymbol{\nu}_{\text{prev},w} + \mathbf{Q} \left(\sqrt{1 - \omega^2} \boldsymbol{\nu}_{\text{prev},w} + \omega \boldsymbol{\xi}_w \right),$$

donde $\boldsymbol{\xi}_w \sim \mathcal{N}(0, \mathbf{C})$.

37: **Paso 10.4.2:** Calcular los campos de densidad y velocidad asociados $\mathbf{k}(\boldsymbol{\nu}_{\text{prop},w}) = \bar{\mathbf{k}} + \boldsymbol{\nu}_{\text{prop},w}$ o $\mathbf{k}(\boldsymbol{\nu}_{\text{prop},w}) = \bar{\mathbf{k}} \exp(\boldsymbol{\nu}_{\text{prop},w})$, según parametrización elegida. Resolver el problema directo con campos $\mathbf{k}(\boldsymbol{\nu}_{\text{prop},w})$ y evaluar en los receptores en los tiempos de medida para determinar el operador de observación $\mathbf{f}(\boldsymbol{\nu}_{\text{prop},w})$.

38: **Paso 10.4.3:** Calcular el logaritmo negativo de la verosimilitud correspondiente

$$\text{loglikeprop}_w = - \frac{\|\mathbf{f}(\boldsymbol{\nu}_{\text{prop},w}) - \mathbf{d}^{\text{obs}}\|_2^2}{2\sigma_{\text{ruido}}^2}.$$

39: **Paso 10.4.4:** Calcular el logaritmo de la probabilidad

$$\log(\alpha) = \text{loglikeprop}_w - \text{loglikeprev}_w.$$

40: **Paso 10.4.5:** Aceptar la propuesta o no. Generar logr_w como $\log(\mathcal{U}(0, 1))$.

41: **if** $\text{logr}_w < \log(\alpha)$ **then**

42: Calcular el logaritmo negativo de la probabilidad a priori de la propuesta

$$\text{logpriorprop}_w = - \frac{\boldsymbol{\nu}_{\text{prop},w}^T \mathbf{C}^{-1} \boldsymbol{\nu}_{\text{prop},w}}{2},$$

43: $\boldsymbol{\nu}_{\text{prev},w} = \boldsymbol{\nu}_{\text{prop},w}$,

44: $\text{loglikeprev}_w = \text{loglikeprop}_w$,

45: $\text{logpostprev}_w = \text{loglikeprop}_w + \text{logpriorprop}_w$.

46: **end if**

47: **end for**

48: **end for**

6.9. Aproximación de campos de velocidad y densidad mediante series Karhunen-Loève

Expresaremos el campo de densidad y de velocidad total como la suma de los campos de fondo estratificado con la incorporación del domo salino, más el campo correspondiente a la zona de las anomalías o trampas, tal como se hizo en la sección 5.3.1, es decir

$$\rho(x, y) = \rho_F(x, y) + \rho_A(x, y), \quad (6.30)$$

$$v_p(x, y) = v_{p,F}(x, y) + v_{p,A}(x, y), \quad (6.31)$$

donde $\rho_F(x, y)$ y $v_{p,F}(x, y)$ son los campos de densidad y velocidad del fondo estratificado incorporando el domo salino, respectivamente, $\rho_A(x, y)$ es la densidad correspondiente en las trampas de gas o petróleo y $v_{p,A}(x, y)$ es la velocidad correspondiente en las trampas de gas o petróleo.

Aproximaremos los campos en las zonas de las anomalías usando una serie de Karhunen-Loève, donde $\rho_A(x, y)$ y $v_{p,A}(x, y)$ se definen a través de funciones de salto, es decir, toman valor igual a cero en todo el dominio, excepto en la región donde se encuentran las anomalías de gas y de petróleo. Los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$ representados mediante desarrollos de Karhunen-Loève se expresan como,

$$\rho_A(x, y) \approx \bar{\mathbf{k}}_\rho + \sum_{i=1}^M \sqrt{\lambda_{\rho,i}} \phi_{\rho,i} \mathbf{m}_{\rho,i}, \quad (6.32)$$

$$v_{p,A}(x, y) \approx \bar{\mathbf{k}}_{v_p} + \sum_{i=1}^M \sqrt{\lambda_{v_p,i}} \phi_{v_p,i} \mathbf{m}_{v_p,i}, \quad (6.33)$$

donde \mathbf{m}_ρ y \mathbf{m}_{v_p} son coeficientes de longitud M . Para determinar el número de modos M necesarios para aproximar los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$, realizamos una prueba de sensibilidad usando $M = 100$, $M = 200$, $M = 400$ y $M = 600$ modos. A continuación, en las figuras 6.2, 6.3, 6.4 y 6.5 se observan las aproximaciones de los campos en la zona de las anomalías para los cuatro casos.

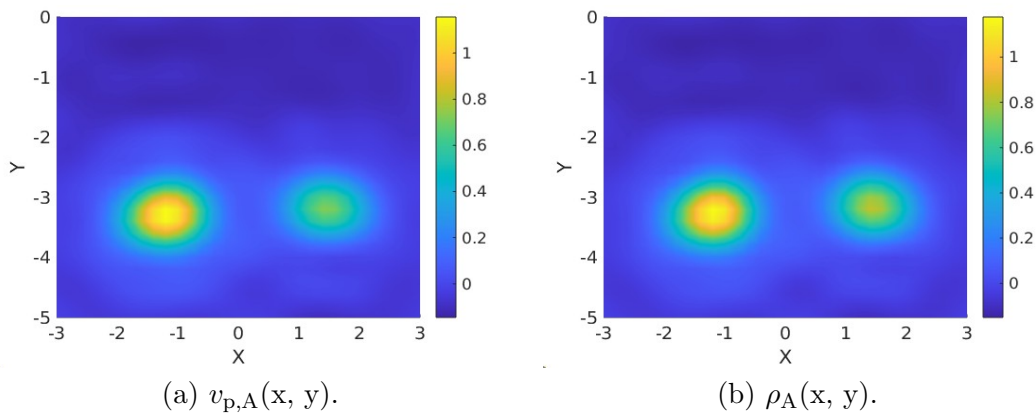


Figura 6.2: Campos aproximados con series Karhunen-Loève truncadas usando $M = 100$ términos.

Como se observa en las figuras 6.2 y 6.3, con un número de modos menor a $M = 400$, las aproximaciones en los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$ aún no muestran una convergencia, sin embargo, a partir $M = 400$ modos la magnitud de los campos se estabiliza, asegurando la convergencia de la aproximación, como se puede observar en las figuras 6.4 y 6.5. Por lo tanto, se ha optado por truncar la serie de Karhunen-Loève a $M = 400$ modos para ambos campos en las zonas de las anomalías.

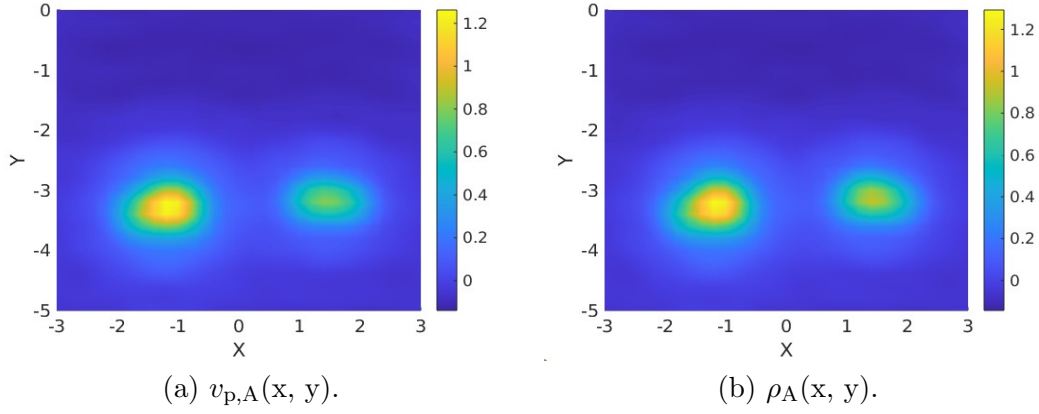


Figura 6.3: Campo aproximados con series Karhunen-Loève truncadas usando $M = 200$ términos.

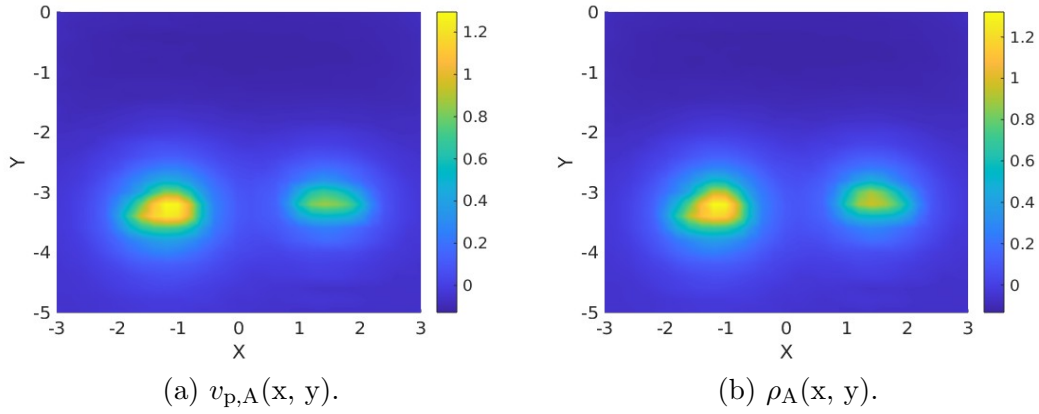


Figura 6.4: Campos aproximados con series de Karhunen-Loève truncadas usando $M = 400$ términos.

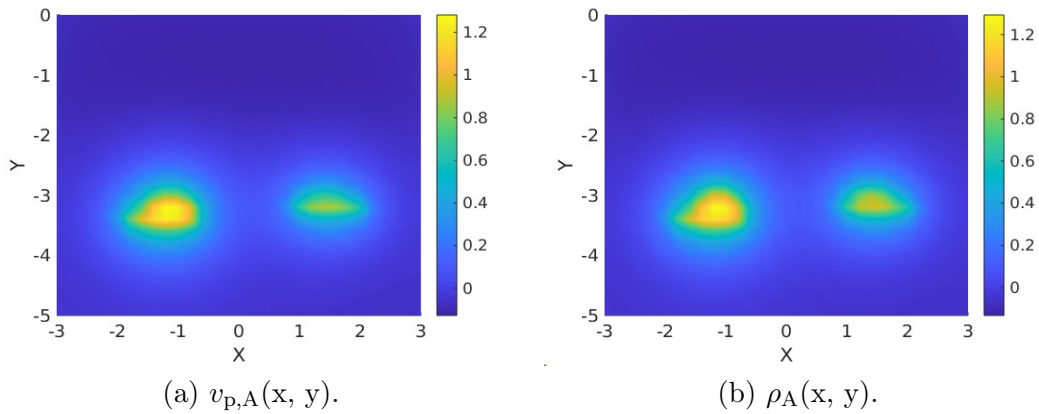


Figura 6.5: Campos aproximados con series Karhunen-Loève truncadas usando $M = 600$ términos.

El número de modos M utilizado para truncar las series en las ecuaciones (6.32) y (6.33) se ha tomado idéntico para las expansiones de densidad y velocidad por dos razones principales:

- Consistencia física: ambos campos comparten la misma geometría de dominio y estructura de correlación espacial, lo que genera espectros cualitativamente similares.
- Simplicidad computacional: utilizar el mismo M simplifica la implementación sin pérdida de precisión, dado que el análisis de sensibilidad mostró que 400 modos son suficientes para ambos campos.

Para determinar M , se realizó,

- Análisis de convergencia visual (figuras 6.2 a 6.5), donde observamos:
 - Para $M < 400$, los campos muestran fluctuaciones sin convergencia clara.
 - Para $M \geq 400$, las aproximaciones se estabilizan.
- Se calculó la energía acumulada para los campos de densidad $\rho_A(x, y)$ y de velocidad $v_{p,A}(x, y)$, donde ambos alcanzan el 95 % de energía acumulada en $M = 400$ modos (ver figura 6.6). El umbral del 95 % garantiza que los modos seleccionados capturen la mayor parte de la información física relevante de los campos ³

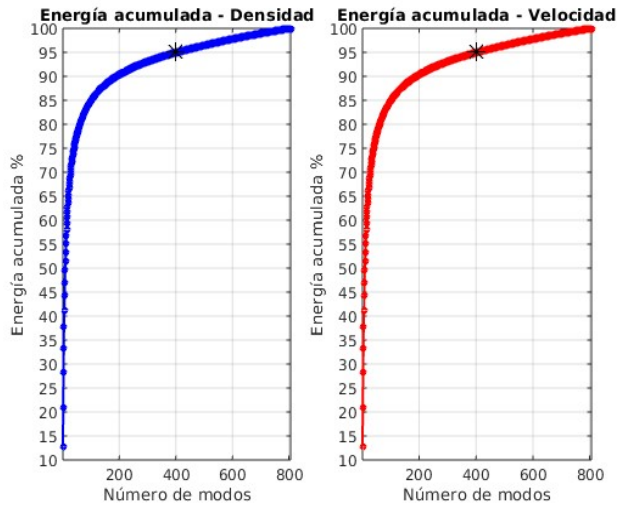
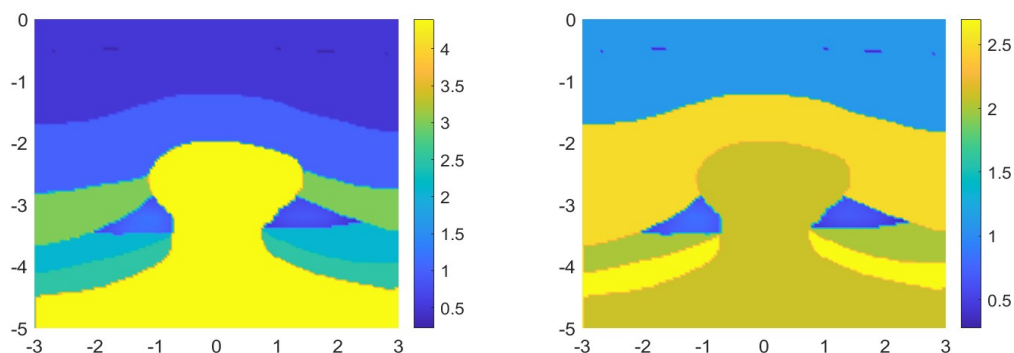


Figura 6.6: Energía acumulada para $\rho_A(x, y)$ y $v_{p,A}(x, y)$.

Esta elección equilibra precisión y eficiencia, ya que reduce el coste computacional en un 50'4 % (de 806 a 400 modos) manteniendo el 95 % de energía acumulada.

Una vez seleccionado el número de modos, los desarrollos se introducirán en las expresiones (6.30) y en (6.31), que a su vez se introducirán en la formulación del problema directo. En las imágenes 6.7a y 6.7b observamos los campos de fondo, donde se incorpora el domo salino y los desarrollos de Karhunen-Loève en las zonas de las trampas de petróleo y de gas.

³ La energía acumulada se calculó como $E(M) = \frac{\sum_{i=1}^M \lambda_i}{\sum_{i=1}^N \lambda_i} \times 100 \%$, donde λ_i son los autovalores ordenados en orden descendente [59].



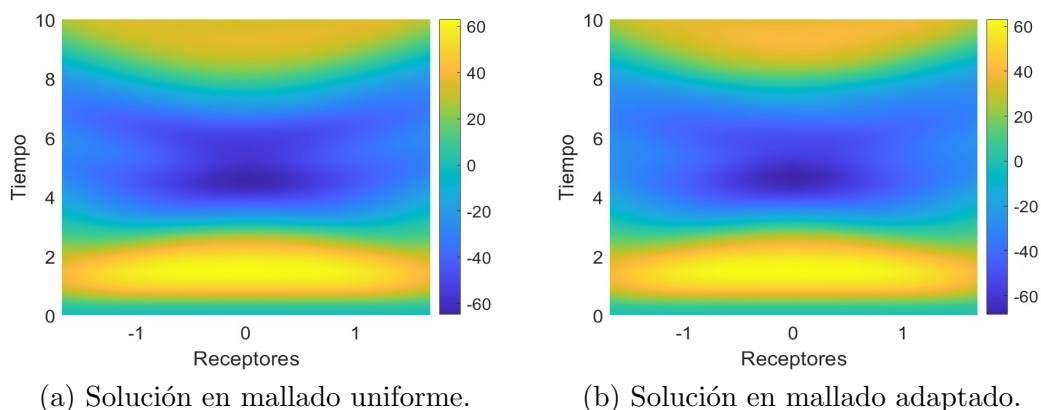
(a) Campo $v_p(x, y) = v_{p,F}(x, y) + v_{p,A}(x, y)$. (b) Campo $\rho(x, y) = \rho_F(x, y) + \rho_A(x, y)$.

Figura 6.7: Campo de densidad y velocidad incorporando desarrollos de Karhunen-Loève con 400 modos.

6.10. Comparación de operadores de observación

Como se mencionó anteriormente, para trabajar en el problema inverso bayesiano necesitamos el operador de observación del problema directo, para luego obtener los datos sintéticos d^{obs} . Como se estudió en la sección 5.2.1, en la resolución numérica del problema directo utilizamos dos discretizaciones espaciales: una discretización uniforme y otra adaptada a la geometría. En esta sección mostramos los resultados obtenidos para el operador de observación utilizando las discretizaciones estudiadas en 5.2.1 e incorporando los campos de densidad y de velocidad en la zona de las trampas de petróleo/gas obtenidos mediante la aproximación Karhunen-Loève con 400 modos.

En la figuras 6.8a y 6.8b observamos el desplazamiento de la onda durante los 10 segundos de simulación en la región donde se encuentran los 13 receptores. Estos resultados corresponden a los operadores de observación $f_{uniforme}(\nu)$ y $f_{adaptado}(\nu)$.



(a) Solución en mallado uniforme.

(b) Solución en mallado adaptado.

Figura 6.8: Campo de densidad y velocidad incorporando las expansión Karhunen-Loève con 400 modos.

Como se observa en las figuras 6.8a y 6.8b los resultados en los operadores de observación son bastantes similares entre sí. La diferencia porcentual entre $f_{uniforme}(\nu)$ y $f_{adaptado}(\nu)$ es de 1'44%, por lo que se concluye que al utilizar diferentes discretizaciones espaciales no observamos grandes

diferencias entre los datos observados. El mapa de la diferencia entre $f_{\text{uniforme}}(\boldsymbol{\nu})$ y $f_{\text{adaptado}}(\boldsymbol{\nu})$ se observa en la figura 6.9.

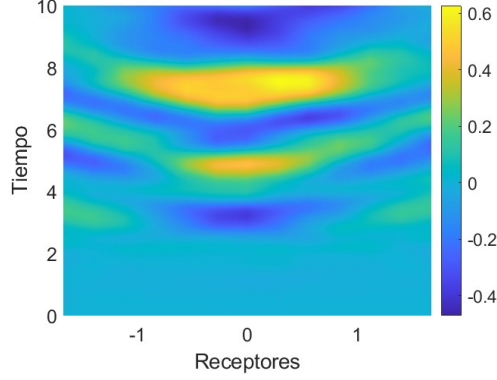


Figura 6.9: Diferencia entre los valores medidos en los receptores para los dos mallados utilizados en la simulación.

Como mencionamos en 5.2.2, para obtener los datos sintéticos \mathbf{d}^{obs} utilizaremos los valores de la solución numérica en los receptores r_j , $j = 1, \dots, J$, con $J = 13$, para $\mathbf{f}_{\text{uniforme}}(\boldsymbol{\nu})$ y $\mathbf{f}_{\text{adaptado}}(\boldsymbol{\nu})$. A estos conjuntos de mediciones se les añadirá ruido Gaussiano para su uso posterior como datos sintéticos en las pruebas de inversión. Los datos sintéticos es posible calcularlos a través de la expresión (6.34), es decir

$$\mathbf{d}^{\text{obs}} = \mathbf{f}(\boldsymbol{\nu}_{\text{exacto}}) + \boldsymbol{\eta} = \mathbf{d}_{\text{exacto}} + \boldsymbol{\eta}, \quad (6.34)$$

donde el ruido añadido está representado como $\boldsymbol{\eta}$, con componentes ϵz , $\epsilon = \frac{\alpha}{100} \sqrt{\frac{\sum_{j=1}^J \sum_{l=0}^L |d_{j,\text{exacto}}^l|^2}{JL}}$, con $\sqrt{\frac{\sum_{j=1}^J \sum_{l=0}^L |d_{j,\text{exacto}}^l|^2}{JL}} = 34'7540$. El nivel porcentual de ruido que añadiremos será de $\alpha = 5\%$ y z serán números aleatorios generados con una distribución normal $\mathcal{N}(0, 1)$.

En las figuras 6.10a y 6.10b se han graficado los datos del operador de observación con mallado uniforme $f_{\text{uniforme}}(\boldsymbol{\nu})$ y los datos sintéticos calculados con (6.34). En las figuras 6.11a y 6.11b se han graficado los datos del operador de observación con mallado adaptado a la geometría $f_{\text{adaptado}}(\boldsymbol{\nu})$ y sus datos sintéticos, respectivamente. En ambos casos se ha añadido un nivel porcentual de 5% de ruido a los datos de los operadores de observación.

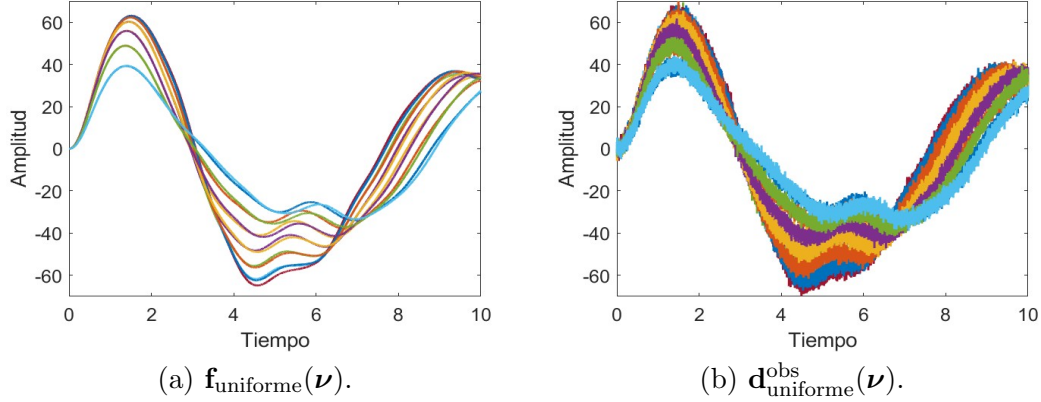


Figura 6.10: Datos obtenidos obtenido por los 13 receptores usando mallado uniforme e incorporando los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$.

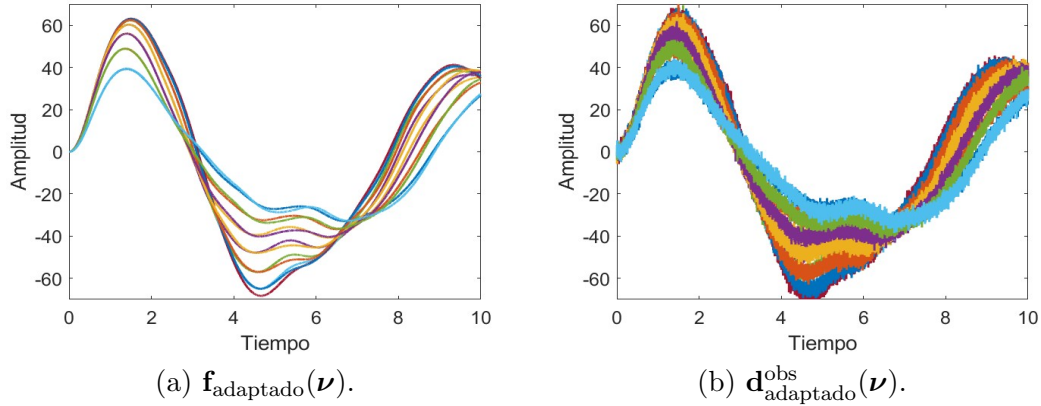


Figura 6.11: Datos obtenidos obtenido por los 13 receptores usando mallado adaptado e incorporando los campos $\rho_A(x, y)$ y $v_{p,A}(x, y)$.

6.11. Resultados de muestreo FES

Para resolver el problema inverso bayesiano en alta dimensión usando el método Monte Carlo cadena de Markov con un esquema FES, correspondiente al algoritmo estudiado en la sección 6.8, utilizamos $W = 1,650$ cadenas, que cumple con la condición de $W > 4M$, donde M son los modos utilizados para truncar la expansión de Karhunen-Loève. En este caso los modos utilizados son $M = 400$, por lo que el número de cadenas utilizadas en la simulación debe ser $W > 1600$.

Los parámetros de interés $\boldsymbol{\nu}$ que queremos estimar en el problema inverso bayesiano son los coeficientes \mathbf{m}_ρ y $\mathbf{m}_{\mathbf{v}_p}$, los cuales aparecen en las ecuaciones (6.14) y (6.15), respectivamente. Estos coeficientes se estiman utilizando el algoritmo FES dentro del esquema MCMC, como se describe en la sección 6.8.

Una vez obtenidos estos coeficientes, se reconstruyen los campos de densidad y de velocidad $\rho_A(x, y)$ y $v_{p,A}(x, y)$ en la zona de anomalías mediante la expansión de Karhunen-Loève truncada a $M = 400$ modos, utilizando las expresiones (6.16) y (6.17). Esto permite reducir la dimensión del espacio de parámetros sin perder información relevante sobre las anomalías en las trampas de petróleo y gas.

Definimos el vector de parámetros $\boldsymbol{\nu}$ como,

$$\boldsymbol{\nu} = [\mathbf{m}_{\rho,1}, \dots, \mathbf{m}_{\rho,j}, \dots, \mathbf{m}_{\rho,M}, \mathbf{m}_{v_p,1}, \dots, \mathbf{m}_{v_p,j}, \dots, \mathbf{m}_{v_p,M}], \quad (6.35)$$

donde el vector de parámetros $\boldsymbol{\nu}$ tiene longitud $2M = 800$, ya que está compuesto por la concatenación de los coeficientes \mathbf{m}_{ρ} y \mathbf{m}_{v_p} , cada uno con $M = 400$ términos. Para la resolución del problema inverso, utilizaremos una discretización espacial uniforme, en lugar de una discretización adaptada, debido al gran costo computacional que supone el uso de una discretización adaptada.

Inicialmente el vector de parámetros $\boldsymbol{\nu}$ de la expresión (6.35) será nulo, lo que implica que los campos ρ_A y $v_{p,A}$ en las zonas de las trampas también lo serán. Los resultados que se observan en las figuras 6.12a y 6.12b, son los campos total $\rho(x, y)$ y $v_p(x, y)$, donde los campos de las trampas, $\rho_A(\boldsymbol{\nu})$ y $v_{p,A}(\boldsymbol{\nu})$, han sido reconstruidos utilizando los coeficientes obtenidos mediante el método FES y MCMC y la expansión de Karhunen-Loève truncada a $M = 400$ modos, tras 60 iteraciones con 1,650 cadenas.

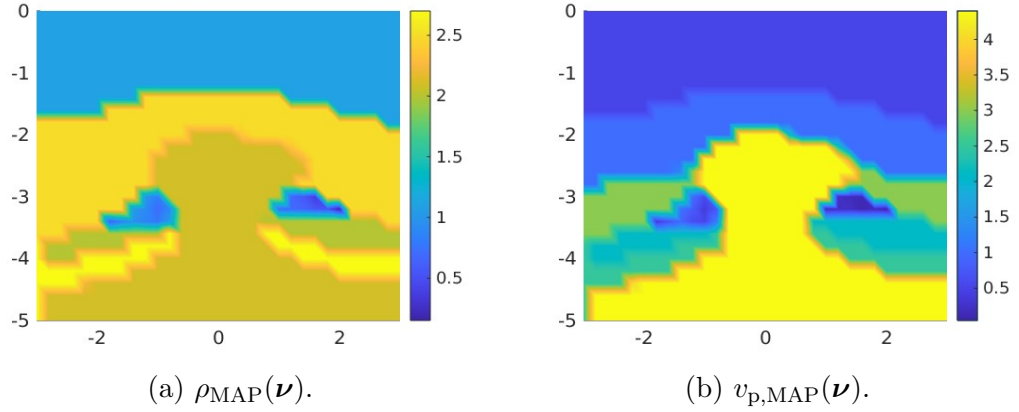


Figura 6.12: Campos totales $\rho(\boldsymbol{\nu})$ y $v_p(\boldsymbol{\nu})$ después de 60 iteraciones.

De la figura 6.12, se puede observar que, después de 60 iteraciones, se logran identificar la presencia de otros materiales diferentes al del fondo, en la zona de los yacimientos. El valor promedio de la densidad en la zona de gas es de 0'65 y, en la de petróleo, de 0'77. Para la velocidad de propagación de la onda, el valor promedio en la zona de gas es de 0'49 y, en la de petróleo, de 0'90.

En la tabla 6.1 comparamos los valores reales de densidad y de velocidad de propagación del petróleo y gas con los valores obtenidos a través del método inverso bayesiano para la muestra de probabilidad más alta $\rho_{A,\text{MAP}}$ y $v_{p,A,\text{MAP}}$.

Objeto	ρ_A	$\rho_{A,\text{MAP}}$	$v_{p,A}$	$v_{p,A,\text{MAP}}$
Petróleo	0'90	0'77	1'30	0'90
Gas	0'70	0'65	0'45	0'49

Cuadro 6.1: Comparación de valores de densidad y velocidad reales en la zona de las trampas de petróleo y gas, con los valores de probabilidad máxima a posteriori.

Comparando los resultados obtenidos con los valores reales de densidad y velocidad presentados en la tabla 6.1, observamos que los valores promedio de densidad $\rho_{A,\text{MAP}}$ obtenidos en la simulación son bastante similares a los valores reales, al igual que el valor promedio MAP de velocidad de propagación de la onda en la zona del gas. En cuanto a la zona de la anomalía correspondiente

al petróleo, el valor promedio $\rho_{A,MAP}$ tiene un diferencia porcentual con respecto al valor real en un 14'43 %, y la diferencia porcentual entre el valor promedio $v_{p,A,MAP}$ con respecto al valor real tiene una diferencia porcentual igual a 30'77 %. En cuanto a la zona de la anomalía correspondiente al gas, la diferencia porcentual entre el valor promedio $\rho_{A,MAP}$ y su valor real es de 7'14 %, y la diferencia porcentual entre el valor promedio $v_{p,A,MAP}$ con respecto a su valor real es de 8'89 %.

Capítulo 7

Conclusión

A lo largo de esta tesis doctoral se ha estudiado, tanto teórica como numéricamente, la formulación inversa bayesiana para la reconstrucción de anomalías y de campos, como el de densidad y de velocidad de propagación de onda, así como la formulación directa del problema de propagación de ondas.

En el capítulo 2 presentamos la formulación del problema directo de propagación de ondas elásticas en un medio heterogéneo, que son generadas por una red de fuentes, y son recibidas por una red de receptores. Como nos interesa la dinámica de la onda, a través de la ecuación de conservación del momento lineal obtenemos la EDP que gobierna la formulación del problema directo, que toma la forma de una ecuación de ondas. Para poder resolverla hemos acotado el dominio artificialmente, para así obtener una solución aproximada que nos permite definir operadores de observación. A partir de la formulación clásica de la ecuación de ondas en un dominio truncado hemos derivado la formulación débil de ésta. Se ha discretizado en espacio usando el método de elementos finitos y en tiempo mediante diferencias finitas explícitas que satisfacen la condición CFL. En la formulación débil hemos incorporado las condiciones de contorno no reflectantes en ciertos bordes del dominio computacional. A través de la resolución de un esquema numérico explícito ha sido posible obtener la solución numérica aproximada, que representa el desplazamiento de la onda con dependencia temporal y espacial. Con este fin hemos desarrollado un algoritmo en el software Matlab, donde primero se incorpora la geometría con el mallado numérico, para luego asignar las propiedades del medio estratificado, y de las anomalías presentes, a cada nodo del mallado. Hemos optado por métodos de orden bajo para reducir el coste computacional, ante la necesidad de resolver grandes cantidades de este tipo de problemas en las formulaciones bayesianas.

Según detallamos en los capítulos 4 y 5, hemos utilizado dos geometrías; la primera es una geometría estratificada con una anomalía de forma elíptica en el centro del dominio computacional; y la otra geometría ha sido un fondo estratificado con una estructura de domo salino al centro del dominio y anomalías presentes que representan trampas de petróleo y gas; esta geometría es una aplicación directa en la geofísica. En ambos casos de estudio, hemos resuelto el problema directo de la propagación de ondas elásticas emitidas por una red de fuentes. Además, en ambos casos de estudio hemos utilizado dos tipos de discretizaciones espaciales. La primera es una discretización uniforme, donde los elementos tienen el mismo tamaño y vienen asociados a una red equiespaciada. La segunda discretización adapta el mallado a los dominios presentes, es decir, se adapta a los estratos de distintos materiales presentes, así como a las inclusiones de otros materiales y a la estructura de domo salino. La resolución del problema directo define los operadores de observación en la red de receptores, que se corresponderían con los datos que se medirían en implementaciones en la vida real del sistema de imagen modelizado. En las pruebas numéricas elegimos configuraciones con geometrías y materiales predefinidos, y definimos datos sintéticos, es decir, generados numéricamente añadiendo un ruido Gaussiano a los valores observados numéricamente. Utilizamos los datos sintéticos como sustitutos de mediciones reales para la configuración elegida.

En el primer caso de estudio donde la anomalía era una elipse, la diferencia porcentual en la comparación de operadores de observación para las dos discretizaciones espaciales cerca de un 0'48 %, mientras que en el segundo caso de estudio donde usamos una geometría más compleja la cual es un fondo estratificado con domos salinos y trampas de petróleo/gas, obtuvimos una diferencia porcentual cerca de 2'93 %. Pese a tener una diferencia porcentual de 2'93 % en el segundo caso de estudio considerando la región computacional entera, las evaluaciones de la solución numérica en los receptores durante el tiempo de computación (véanse figuras 4.8 y 5.7) son bastantes similares entre sí. Dado que los datos que se medirían en una situación real vendrían distorsionados por ruidos de distintos orígenes y que pretendemos desarrollar métodos de inversión estables en presencia de ruido, trabajaremos con ambas. Los mallados adaptados los utilizaremos para generar datos sintéticos y los mallados uniformes para los métodos de inversión.

En el capítulo 3 hemos estudiado de forma teórica y numérica el problema inverso paramétrico en dimensión finita, es decir, cuando se pretende determinar anomalías caracterizadas por un número finito de parámetros. En primer lugar, revisamos la formulación del problema inverso determinista como un problema de optimización con restricciones en forma de problema de valores iniciales y de contorno para ecuaciones en derivadas parciales (el problema directo), así como las estrategias de descenso utilizadas para resolverlo. Tras ello planteamos el problema inverso bayesiano y discutimos dos estrategias de resolución. A partir de la formulación bayesiana obtenemos un funcional de coste regularizado con información a priori. Desarrollamos algoritmos de descenso iterativos de tipo Fletcher-Levenberg-Marquardt para calcular mínimos. Estos métodos funcionan como métodos de gradiente en zonas de gradientes grandes y como métodos de Newton en zonas de gradientes pequeños. Las parametrizaciones óptimas corresponden a valores de probabilidad máxima. Explicamos cómo usarlas para construir aproximaciones de Laplace de la probabilidad a posteriori dada por la fórmula de Bayes, en términos de distribuciones Gaussianas fáciles de muestrear. Esta técnica tiene un alcance limitado cuando hay varios parámetros óptimos, o fuertes asimetrías. Para un estudio más detallado de la probabilidad a posteriori de distintos rangos paramétricos se utilizan técnicas de muestreo Monte Carlo basado en cadenas de Markov, que revisamos brevemente.

En el capítulo 4 estudiamos en detalle un problema inverso en baja dimensión, en el que se busca caracterizar una anomalía definida por un pequeño número de parámetros. Dado un medio estratificado con un objeto elíptico, el vector de parámetros desconocidos ν es un vector que contiene información geométrica y propiedades física de la elipse. Con el método inverso bayesiano buscamos obtener los valores de ν que permitan la reconstrucción de la elipse cuantificando la incertidumbre en estos parámetros. Dada la baja dimensionalidad, hemos procedido a caracterizar los parámetros de máxima probabilidad optimizando un coste con regularización bayesiana. Además, usando la aproximación de Laplace, pudimos cuantificar la incertidumbre en torno a cada variable del parámetro, que visualizamos a través de histogramas. En este caso de estudio, hemos realizado simulaciones combinando datos sintéticos generados con distintos tipos de mallado y operadores de observación calculados igualmente con distintos tipos de mallado (uniforme, estratificado, completamente adaptado a la geometría con inclusión). En todos los casos los valores de máxima probabilidad de los parámetros para los niveles de ruido considerados son cercanos a los parámetros reales de la inclusión, lo que avala el uso de estos métodos de inversión para el desarrollo de técnicas de imagen. Incluimos una extensión del método al tratamiento de problemas de elastografía médica.

En los dos siguientes capítulos subimos la dimensionalidad en una configuración de domo salino con datos sintéticos obtenidos con un mallado adaptado a la geometría. En este caso se trabaja con campos de densidad y velocidad que varían espacialmente, y se caracterizan los yacimientos mediante la variación espacial de estos campos.

En el capítulo 5, representamos estos campos mediante desarrollos de Fourier truncados. El parámetro desconocido ν corresponde a los coeficientes de serie de Fourier que representa las variaciones en la zona de los yacimientos. Caracterizamos los coeficientes de estas series a través del método inverso bayesiano mediante técnicas de Monte Carlo basadas en cadenas de Markov.

Optamos por incluir la información disponible sobre la densidad y velocidad del medio estratificado subyacente en el problema directo, y parametrizar la corrección, con información a priori nula sobre ella. Con esta estrategia, se ha logrado visualizar la presencia de yacimientos en los costados del domo salino. Si bien, los valores de los campos obtenidos son alejados de los valores reales, es importante mencionar que pese a que la información inicial de los parámetros desconocidos fueron nulos y que la presencia de domo salino amortigua la señal recibida, aún así se logra identificar materiales distintos a los de los campos de fondo. La falta de información a priori sobre los yacimientos dificulta la elección adecuada de las matrices de covarianza, lo que a su vez afecta en la obtención de los resultados.

Cabe mencionar que hemos utilizado en la definición de la verosimilitud en la formulación bayesiana un operador de observación definido en mallado uniforme grosero, se ha disminuido el número de nodos en la discretización para disminuir el coste computacional a la mitad comparado con el caso anterior. Si bien, es posible mejorar la discretización espacial, esto conlleva a un coste computacional extra en la resolución del problema inverso, que no asegura tener mejores resultados. En particular, para el caso del domo salino, se ha demostrado que las soluciones de la ecuación de ondas generadas con mallado uniforme y con mallado adaptado al domo salino presentan una diferencia relativa máxima del 2'93 % (sección 5.2.2). Esta diferencia sugiere que ambos operadores de observación son suficientemente similares. Dado que los datos sintéticos se construyen como la suma del operador de observación con una componente de ruido gaussiano, es razonable asumir que estas pequeñas diferencias en los datos sintéticos no tienen un efecto significativo sobre la función de verosimilitud. Por tanto, el uso del operador de observación generado con mallado uniforme se considera una aproximación válida en el contexto del modelo bayesiano desarrollado. Esta elección adquiere mayor relevancia al considerar el coste computacional del método inverso: para cada nueva propuesta del parámetro ν en el método inverso bayesiano, se debe resolver el problema directo para obtener el operador de observación correspondiente. El mallado uniforme reduce el tiempo de resolución del problema directo, lo que conlleva una disminución en el tiempo total del método inverso. Dado que este proceso se repite varias veces en algoritmos MCMC, la optimización del problema directo mediante mallados uniformes resulta crítica para la viabilidad práctica de la inversión bayesiana en alta dimensión, sin comprometer la calidad de los resultados. Sin embargo, lo que sí podría influir en tener una mejor estimación de los campos es la consideración de otro valor medio ν_0 en la definición de la distribución a priori, distinto de cero.

Finalmente, en el capítulo 6, se ha formulado el problema inverso bayesiano en la configuración de domo salino en alta dimensión, en el límite de dimensión infinita, representando los campos de densidad y velocidad como campos aleatorios y aproximándolos en dimensión finita mediante desarrollos de tipo Karhunen-Loève asociados a los operadores de covarianza elegidos para cada campo. Como en el capítulo anterior, la información disponible sobre el medio estratificado de fondo se incluye en campos de densidad y velocidad en el problema directo, se busca la corrección que representa la presencia de posibles yacimientos en torno al domo. Para muestrear la probabilidad a posteriori definida por la fórmula de Bayes, hemos usado un método Monte Carlo basado en cadena de Markov cuyos resultados permanecen estables a medida que la dimensión aumenta, el método FES, que combina una técnica AIES (utilizada en los casos anteriores) eficiente para parámetros mal escalados en baja dimensión con una técnica de tipo Crank-Nicolson preconditionado estándar en problemas infinito dimensionales para el resto de parámetros. Aplicando el método inverso bayesiano buscamos los parámetros desconocidos ν , que representan los coeficientes \mathbf{m}_ρ y \mathbf{m}_{v_p} de la serie de Karhunen-Loève truncada. Para la truncación de la serie, seleccionamos $M = 400$ modos, basado en pruebas de sensibilidad realizadas para los campos de densidad $\rho_A(x, y)$ y de velocidad $v_{p,A}(x, y)$ en la zona de trampas. Las aproximaciones de $\rho_A(x, y)$ y $v_{p,A}(x, y)$ fueron incorporadas en el problema directo para obtener los operadores de observación, los cuales se utilizaron para realizar el muestreo de la probabilidad posterior.

Como resultado de la resolución del problema inverso bayesiano, obtuvimos los coeficientes que parametrizan ρ_A y $v_{p,A}$ y que maximizan la probabilidad a posteriori en la región de las trampas de petróleo y gas. Encontramos que, para el campo de densidad en la zona de petróleo, la diferencia porcentual entre el valor real y el valor promedio MAP es de 14'44 %, mientras que en la zona del gas, esta diferencia es de 7'14 %. En cuanto al campo de velocidad, la diferencia porcentual entre los valores reales y los valores promedio MAP es de 30'77 % en la zona del petróleo y de 8'89 % en la zona del gas. Cabe mencionar que hemos obtenido estos resultados sin información a priori de los campos ρ_A y $v_{p,A}$, es decir, ν_0 nulo, por lo que pese a esta consideración, el método inverso si ha logrado identificar la presencia de las trampas de petróleo/gas.

En los capítulos 5 y 6, las simulaciones se han realizado suponiendo una información a priori nula, es decir, con media $\nu_0 = 0$, para los campos de densidad $\rho_A(x, y)$ y velocidad $v_{p,A}(x, y)$. Esta elección responde principalmente a dos motivos: (i) la ausencia de una caracterización cuantitativa detallada de las anomalías antes de observar los datos, lo cual dificulta la construcción de una media a priori informada, y (ii) el interés en analizar el efecto puro de la información contenida en las observaciones a través de la formulación bayesiana. No obstante, la metodología desarrollada admite de forma natural el uso de $\nu_0 \neq 0$ no nulos. De hecho, incorporar una media a priori $\nu_0 \neq 0$, basada por ejemplo en modelos geológicos previos, podría mejorar significativamente la precisión de los estimadores MAP o de las muestras posteriores. En particular, una estimación a priori mejor informada puede guiar mejor el proceso inverso en zonas con escasa resolución.

Como línea futura, se propone realizar simulaciones con $\nu_0 \neq 0$ no nulo y con matrices de covarianza a priori adecuadas al problema. Esta extensión permitiría evaluar cuantitativamente la mejora en la reconstrucción al incorporar un conocimiento previo sobre la estructura y variabilidad de las anomalías.

El trabajo llevado a cabo en esta tesis muestra que las técnicas de inversión bayesiana tienen un gran potencial en la mejora de técnicas de imagen, al permitir obtener las imágenes más probables y cuantificar la incertidumbre señalando otras configuraciones probables así como en rango de incertidumbre en las predicciones. Comenzando por una configuración simple en baja dimensión, hemos evolucionado hacia el estudio de configuraciones de interés práctico en alta dimensión. Nuestro estudio apunta también la relevancia de disponer de buena información a priori, que puede provenir de otras técnicas de imagen, y de elegir adecuadamente las covarianzas.

Capítulo 8

Apéndice

8.1. Adimensionalización del problema directo

Para adimensionalizar el problema directo, seleccionamos las siguientes escalas características:

- **Longitud:** $L = 1,000$ m
- **Tiempo:** $T = 1$ s
- **Densidad de referencia:** $\rho_0 = 1,000$ kg/m³.

8.1.1. Variables adimensionales

Definimos las variables adimensionales (denotadas con tilde) como [1, 24, 25]:

$$\tilde{x} = \frac{x}{L}, \quad \tilde{t} = \frac{t}{T}, \quad \tilde{u} = \frac{u}{L}, \quad \tilde{R} = \frac{R}{L}, \quad \tilde{\Sigma} = \frac{\Sigma}{L}.$$

8.1.2. Coeficientes adimensionalizados

- **Densidad:** $\tilde{\rho} = \frac{\rho}{\rho_0}$
- **Velocidad de propagación de la onda:** $\tilde{v}_p = \frac{v_p T}{L}$
- **Función de la fuente temporal:**

$$\tilde{f}(\tilde{t}) = \frac{T^2}{L} f_0 (1 - 2\pi^2 f_M^2 T^2 \tilde{t}^2) e^{-\pi^2 f_M^2 T^2 \tilde{t}^2}$$

donde $f_M = 2$ Hz, $f_0 = 0'1 \frac{L}{T^2}$, y por lo tanto:

$$f_M T = 2, \quad \frac{f_0 T^2}{L} = 0'1.$$

- **Distribución espacial de la fuente:**

$$\tilde{g}(\tilde{\mathbf{x}}) = \frac{1}{(\pi\kappa)^{n/2}} \sum_{k=1}^K \exp\left(-\frac{|\tilde{\mathbf{x}} - \tilde{\mathbf{x}}_k|^2}{\kappa}\right), \quad n = 2.$$

8.1.3. Problema directo adimensional

Tras el cambio de variables y omitiendo las tildes por simplicidad, el problema adimensionalizado queda [1, 24, 25]:

$$\rho(\mathbf{x}) \frac{\partial^2 u}{\partial t^2} - \nabla \cdot (\rho(\mathbf{x}) v_p^2(\mathbf{x}) \nabla u) = \rho(\mathbf{x}) f(t) g(\mathbf{x}), \quad \mathbf{x} \in R, t > 0, \quad (8.1a)$$

$$\frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}, t) = 0, \quad \mathbf{x} \in \partial R, t > 0, \quad (8.1b)$$

$$u(\mathbf{x}, 0) = \frac{\partial u}{\partial t}(\mathbf{x}, 0) = 0, \quad \mathbf{x} \in R. \quad (8.1c)$$

8.2. Serie de Fourier - extensión a 2D

Para una función $f(x, y)$ definida en el rectángulo $[x_0, x_1] \times [y_0, y_1]$, generalizamos la serie de Fourier a dos dimensiones [30].

8.2.1. Paso 1: desarrollo en x (fijando y)

Para cada y fijo, expandimos $f(x, y)$ como una serie de Fourier 1D en x (con $L_x = x_1 - x_0$):

$$f(x, y) \sim \frac{A_0(y)}{2} + \sum_{n=1}^{\infty} \left(A_n(y) \cos\left(\frac{2n\pi x}{L_x}\right) + B_n(y) \sin\left(\frac{2n\pi x}{L_x}\right) \right),$$

donde los coeficientes dependen de y :

$$A_n(y) = \frac{2}{L_x} \int_{x_0}^{x_1} f(x, y) \cos\left(\frac{2n\pi x}{L_x}\right) dx, \quad B_n(y) = \frac{2}{L_x} \int_{x_0}^{x_1} f(x, y) \sin\left(\frac{2n\pi x}{L_x}\right) dx.$$

8.2.2. Paso 2: desarrollo de $A_n(y)$ y $B_n(y)$ en y .

Ahora desarrollamos $A_n(y)$ y $B_n(y)$ en series de Fourier para y (con $L_y = y_1 - y_0$):

■ Para $A_n(y)$:

$$A_n(y) \sim \frac{A_{n0}}{2} + \sum_{m=1}^{\infty} \left(A_{nm} \cos\left(\frac{2m\pi y}{L_y}\right) + C_{nm} \sin\left(\frac{2m\pi y}{L_y}\right) \right),$$

donde:

$$A_{nm} = \frac{2}{L_y} \int_{y_0}^{y_1} A_n(y) \cos\left(\frac{2m\pi y}{L_y}\right) dy, \quad C_{nm} = \frac{2}{L_y} \int_{y_0}^{y_1} A_n(y) \sin\left(\frac{2m\pi y}{L_y}\right) dy.$$

■ Para $B_n(y)$:

$$B_n(y) \sim \frac{B_{n0}}{2} + \sum_{m=1}^{\infty} \left(D_{nm} \cos\left(\frac{2m\pi y}{L_y}\right) + B_{nm} \sin\left(\frac{2m\pi y}{L_y}\right) \right),$$

donde:

$$D_{nm} = \frac{2}{L_y} \int_{y_0}^{y_1} B_n(y) \cos\left(\frac{2m\pi y}{L_y}\right) dy, \quad B_{nm} = \frac{2}{L_y} \int_{y_0}^{y_1} B_n(y) \sin\left(\frac{2m\pi y}{L_y}\right) dy.$$

8.2.3. Paso 3: sustitución y serie final 2D

Sustituyendo $A_n(y)$ y $B_n(y)$ en la expansión original de $f(x, y)$, obtenemos la serie de Fourier 2D completa:

$$\begin{aligned} f(x, y) \sim & \frac{A_{00}}{4} + \sum_{n=1}^{\infty} \left(\frac{A_{n0}}{2} \cos\left(\frac{2n\pi x}{L_x}\right) + \frac{B_{n0}}{2} \sin\left(\frac{2n\pi x}{L_x}\right) \right) \\ & + \sum_{m=1}^{\infty} \left(\frac{A_{0m}}{2} \cos\left(\frac{2m\pi y}{L_y}\right) + \frac{B_{0m}}{2} \sin\left(\frac{2m\pi y}{L_y}\right) \right) \\ & + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left(A_{nm} \cos\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right) + B_{nm} \sin\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) \right) \\ & + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left(C_{nm} \cos\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) + D_{nm} \sin\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right) \right). \end{aligned}$$

8.3. Coeficientes 2D

Los coeficientes se calculan mediante integrales dobles:

$$A_{00} = \frac{4}{L_x L_y} \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y) dy dx$$

$$A_{n0} = \frac{4}{L_x L_y} \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y) \cos\left(\frac{2n\pi x}{L_x}\right) dy dx$$

$$A_{0m} = \frac{4}{L_x L_y} \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y) \cos\left(\frac{2m\pi y}{L_y}\right) dy dx$$

$$B_{n0} = \frac{4}{L_x L_y} \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y) \sin\left(\frac{2n\pi x}{L_x}\right) dy dx$$

$$B_{0m} = \frac{4}{L_x L_y} \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y) \sin\left(\frac{2m\pi y}{L_y}\right) dy dx$$

$$A_{nm} = \frac{4}{L_x L_y} \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y) \cos\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right) dy dx$$

$$B_{nm} = \frac{4}{L_x L_y} \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y) \sin\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) dy dx$$

$$C_{nm} = \frac{4}{L_x L_y} \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y) \cos\left(\frac{2n\pi x}{L_x}\right) \sin\left(\frac{2m\pi y}{L_y}\right) dy dx$$

$$D_{nm} = \frac{4}{L_x L_y} \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y) \sin\left(\frac{2n\pi x}{L_x}\right) \cos\left(\frac{2m\pi y}{L_y}\right) dy dx.$$

Bibliografía

- [1] Abugattas C., Carpio A., Cebrián E., Oleaga G., *Quantifying uncertainty in inverse scattering problems set in layered environments*, Applied Mathematics and Computation, vol. 500, 129453, 2025.
- [2] Aghajani A., Haghpanahi M., Nikazad T., *The ultrasound elastography inverse problem and the effective criteria*, Journal of Engineering in Medicine, vol. 227, número 11, pp. 1203-1212, 2013.
- [3] Aminzadeh F., Dasgupta S. N., *Chapter 2 - Fundamentals of Petroleum Geology*, elsevier, vol. 60, número 1, pp. 15-36, 2013, ISBN: 9780444506627.
- [4] Antil H., Pfefferer J., Rogovs S., *Fractional Operators with Inhomogeneous Boundary Conditions: Analysis, Control, and Discretization*, Communications in Mathematical Sciences, vol. 16, número 5, pp. 1395-1426, 2018.
- [5] Antil H., Saibaba A.K., *Efficient algorithms for bayesian inverse problems with Whittle-Matérn priors*, SIAM Journal on Scientific Computing, vol 46(2), 2022.
- [6] Arfken G., *Métodos matemáticos para físicos*, Academic Press, 1981, ISBN: 9681306333.
- [7] Babuska I., Nobile F., Tempone R., *A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data*, SIAM Review, vol. 52, número 2, pp. 317-355, 2010.
- [8] Bamberger A., Glowinski R., Tran Q. H., *A Domain Decomposition Method for the Acoustic Wave Equation with Discontinuous Coefficients and Grid Change*, SIAM Journal on Numerical Analysis, vol. 34, número 2, pp. 603-639, 1997.
- [9] Bathe K.J., *Finite Element Procedures*, Prentice Hall, 1996, ISBN: 0133014584.
- [10] Bernardo J.M., Smith A. F. M., *Bayesian Theory*, John Wiley and Sons, 1994, ISBN: 9780471494645.
- [11] Beskos A., Roberts G., Stuart A., Voss J., *MCMC methods for diffusion bridges*, Stochastics and Dynamics, vol. 118, número 8, pp. 1928-1952, 2008.
- [12] Bishop C. M., *Pattern Recognition and Machine Learning*, Springer-Verlag New York, 2006, ISBN: 9780387310732.
- [13] Brémaud P., *Markov Chains, Gibbs Fields, Monte Carlo, Simulation and Queues*, Springer, 2020, ISBN: 978303045989.
- [14] Brézis H., *Analyse fonctionnelle, Théorie et applications*, Masson, Paris, 1983, ISBN:2225771987.
- [15] Brooks S., Gelman A., Jones G. L., Meng X. L., *Handbook of Markov Chain Monte Carlo*, Chapman and Hall/CRC, 2011, ISBN: 9780429138508.

-
- [16] Bui-Thanh T., Ghattas O., *An Analysis of Infinite Dimensional Bayesian Inverse Shape Acoustic Scattering and Its Numerical Approximation*, Society for Industrial and Applied Mathematics and American Statistical Association (SIAM/ASA), vol. 2, número 1, pp. 203–222, 2014.
- [17] Carlin B.P., Louis T.A., *Bayesian Methods for Data Analysis*, Chapman and Hall/CRC, 2010, ISBN: 9780429191879.
- [18] Carpio A., *Course on Uncertainty Quantification in Scientific Computing - Lecture I: Problems involving random fields*, Conference: XX Spanish-French Jacques Louis Lions School on Numerical Simulations in Physics Engineering: CIMNE, Barcelona, 2023.
- [19] Carpio A., *Course on Uncertainty quantification in scientific computing - Lecture II: Analysis of a posterior probability*, Conference: XX Spanish-French Jacques Louis Lions School on Numerical Simulations in Physics Engineering: CIMNE, Barcelona, 2023.
- [20] Carpio A., *Course on Uncertainty quantification in scientific computing - Lecture III: Problems involving random fields*, Conference: XX Spanish-French Jacques Louis Lions School on Numerical Simulations in Physics Engineering: CIMNE, Barcelona, 2023.
- [21] Carpio A., *Course on Uncertainty quantification in scientific computing - Lecture IV: Reducing dimensionality*, Conference: XX Spanish-French Jacques Louis Lions School on Numerical Simulations in Physics Engineering: CIMNE, Barcelona, 2023.
- [22] Carpio A., *Uncertainty quantification in scientific computing: A short introduction*, Mathematical and Computational Modelling Across the Scales, EHF-JLL 2023, SEMA SIMAI Springer Series, vol 39. Springer, Cham.
- [23] Carpio A., Cebrián E., Abugattas C., *Seeing under salt domes*, preprint 2025.
- [24] Carpio A., Cebrián E., Gutiérrez A., *Object based Bayesian full-waveform inversion for shear elastography*, Inverse Problems 39, 075007, 2023.
- [25] Carpio A., Cebrián E., Oleaga G., Abugattas C., *Object-Based full waveform inversion with quantified uncertainty in stratified media*, Proceedings of The 9th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2024, Lisbon), 2024. URL: https://www.scipedia.com/public/Carpio_et_al_2024a
- [26] Carpio A., Iakunin S., Stadler G., *Bayesian approach to inverse scattering with topological priors*, Inverse Problems 36, 105001, 2020.
- [27] Carpio A., Pierret E., *Uncertainty quantification in covid-19 spread: lockdown effects*, Results in Physics 35, 105375, 2022.
- [28] Carpio A., Rapún M.L., *Hybrid topological derivative and gradient-based methods for electrical impedance tomography*, Inverse Problems 28, 095010, 2012.
- [29] Cazenave T., Haraux A., *An Introduction to Semilinear Evolution Equations, Oxford Lecture Series in Mathematics and Its Applications*, Clarendon Press, vol. 13, 1998, ISBN: 9780198502777.
- [30] Cebrián, E., *Notas sobre series de Fourier*, Universidad de Burgos, 2025.
- [31] Chung E. T., Chan T. F., Tai X. C., *Electrical Impedance Tomography Using Level Set Representation and Total Variational Regularization*, Journal of Computational Physics, vol. 205, número 1, pp. 357-372, 2005.
- [32] Colton D., Kress R., *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer Science and Business Media, 2019, ISBN: 9783030303518.
- [33] Cotter S. L., Roberts G. O., Stuart A. M., White D., *MCMC methods for functions: modifying old algorithms to make them faster*, Statistical Science, vol. 28, número 3, pp. 424-446, 2013.

-
- [34] Coullon J., Webber R. J., *Ensemble sampler for infinite-dimensional inverse problems*, Statistics and Computing, vol. 31, número 28, pp. 28-31, 2021.
- [35] Courant R., Friedrichs K., Lewy H., *Über die partiellen Differenzgleichungen der mathematischen Physik*, *Mathematische Annalen*, vol. 100, pp. 32-74, 1928.
- [36] Dashti M., Stuart A.M., *The Bayesian Approach to Inverse Problems*, Computer Methods in Applied Mechanics and Engineering, vol. 4, pp. 1-107, 2015.
- [37] Dautray R., Lions J.-L., *Mathematical Analysis and Numerical Methods for Science and Technology*, Physics Today, vol. 44, número 12, 1991.
- [38] Doucet A., De Freitas N., Gordon N. J., *Sequential Monte Carlo Methods in Practice*, Springer, 2009, ISBN: 9781475734379.
- [39] Dunlop M.M., *Analysis and Computation for bayesian Inverse Problems*, PhD Thesis, University of Warwick, 2016.
- [40] Engl H.W., Hanke M., Neubauer A., *Regularization of Inverse Problems*, Mathematics and Its Applications, Springer, 1996, ISBN: 9780792361404.
- [41] Engquist B., Majda A., *Radiation Boundary Conditions for Acoustic and Elastic Wave Calculations*, Communications on Pure and Applied Mathematics, vol. 32, número 3, pp.312-358, 1979.
- [42] Esser E., Guasch L., van Leeuwen T., Aravkin A. Y., Herrmann F. J., *Total-Variation Regularization Strategies in Full-Waveform Inversion*, SIAM Journal on Imaging Sciences, vol. 11, número 1, pp. 376-406, 2018.
- [43] Folland G. B., *Fourier Analysis and Its Applications*, American Mathematical Society, 1992, ISBN: 9780821847909.
- [44] Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., *emcee: The MCMC Hammer*, Publications of the Astronomical Society of the Pacific, vol. 125, número 925, pp. 306–312, 2013.
- [45] Gamerman D., Lopes Hedibert F., *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Chapman and Hall/CRC, 2006, ISBN: 9781584885870.
- [46] Gelman A., Carlin J. B., Stern H. S., Dunson D. B., Vehtari A., Rubin D. B., *Bayesian Data Analysis*, Chapman and Hall/CRC, 2015, ISBN: 9780429113079.
- [47] Gelman A., Rubin D.B., *Inference from iterative simulation using multiple sequences*, Statistical Science 7, número 4, pp. 457-472, 1992.
- [48] Ghanem R. G., Spanos P. D., *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, 1991, ISBN: 9781461277958.
- [49] Gilks W. R., Richardson S., Spiegelhalter D.J., *Markov Chain Monte Carlo in Practice*, Chapman and Hall/CRC, 1996, ISBN: 9780429170232.
- [50] Goodman J., Weare J., *Ensemble samplers with affine invariance*, Communications in Applied Mathematics and Computational Science, vol. 5, número 1, pp. 65-80, 2010.
- [51] Green P.J., *Reversible jump Markov chain Monte Carlo computation and bayesian model determination*, Biometrika, vol. 82, número 4, pp. 711-732, 1995.
- [52] Grimmett G. R., Stirzaker D.R., *Probability and Random Processes*, Oxford University Press, 2001, ISBN: 9780198572237
- [53] Grisvard P., *Elliptic Problems in Nonsmooth Domains, Classics in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), 2011, ISBN: 9781611972023.

-
- [54] Hastings W.K., *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*, Biometrika, vol. 57, número 1, p. 97, 1970.
- [55] Hohage T., *Iterative methods in inverse obstacle scattering: regularization theory of linear and nonlinear exponentially ill-posed problems*, PhD thesis, University of Linz, 1999.
- [56] Hristopulos D. T., *Random Fields for Spatial Data Modeling*, Springer, 2009, ISBN:9789402419184.
- [57] Hughes T. J., *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Dover Publications, 2012, ISBN: 0486411818.
- [58] Isaacson E., Keller H. B., *Analysis of Numerical Methods*, SIAM Review, vol. 11, número 1, pp. 100-101, 1969.
- [59] Jolliffe I.T., *Principal Component Analysis*, Springer Series in Statistics, 2002, ISBN: 0387954422.
- [60] Kaipio J.P., Somersalo E., *Statistical and Computational Inverse Problems*, Springer-Verlag New York, 2006, ISBN: 9780387271323.
- [61] Kang F., Zhong-Ci S., *Mathematical Theory of Elastic Structure*, Springer Berlin, 1996, ISBN: 9783662032862.
- [62] Karhunen K., *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, *Annales Academiae Scientiarum Fennicae*, Series A, vol. 37, 1947.
- [63] Kekkonen H., Korolev Y., *Inverse Problems*, Lecture notes, University of Cambridge, 2020.
- [64] Lindgren F., Rue H., Lindström J., *An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach*, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 73, número 4, pp. 423–498, 2011.
- [65] Lions J. L., Magenes E., *Problèmes aux limites non homogènes*, Dunod, Paris, 1968.
- [66] Liu J.S., *Monte Carlo Strategies in Scientific Computing*, Springer Series in Statistics, 2008, ISBN: 9780387763712.
- [67] Loève M., *Probability Theory I*, Springer, 1978, ISBN: 9780387902104.
- [68] Malcolm A., Guzina B., *On the Topological Sensitivity of Transient Acoustic Fields*, *Wave Motion*, vol. 45, número 6, pp. 821-834, 2008.
- [69] Marquardt D.W., *An algorithm for least-squares estimation of nonlinear parameters*, *Journal of the Society for Industrial and Applied Mathematics (SIAM)*, vol. 11, número 2, pp. 431-441, 1963.
- [70] Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H., Teller E., *Equation of State Calculations by Fast Computing Machines*, *Journal of Chemical Physics*, vol. 21, número 6, pp. 1087-1092, 1953.
- [71] Meyn S. P., Tweedie R. L., *Markov Chains and Stochastic Stability*, Cambridge University Press, 2009, ISBN: 9780511626630.
- [72] Minasny B., McBratney A.B., *The Matern function as a general model for soil variograms*, *Geoderma*, vol. 128, número 3, pp. 192-207, 2007.
- [73] Mishra A., *Hologram the future of medicine - From Star Wars to clinical imaging*, *Indian Heart J.* vol. 69, número 4, pp. 566-567, 2017.
- [74] Nadkarni M. G., *Basic Ergodic Theory*, Birkhäuser, Hindustan Book Agency Gurgaon, 1998, ISBN: 9789386279538.

-
- [75] Neal R. M., *Probabilistic Inference and Machine Learning*, Technical Report, University of Toronto, 1993.
- [76] Norris J., *Markov Chains*, Cambridge University Press, 1997, ISBN: 9780511810633.
- [77] Oh S.H., Kwon B.D., *Geostatistical approach to bayesian inversion of geophysical data: Markov chain Monte Carlo method*, Earth Planets Space, vol. 53, pp. 777–791, 2001.
- [78] Rao S. S., *The Finite Element Method in Engineering*, Butterworth-Heinemann, Oxford, 2005, ISBN: 9780750678285.
- [79] Rasmussen C. E., Williams C. K. I., *Gaussian Processes for Machine Learning*, Massachusetts Institute of Technology, 2006, ISBN: 026218253X.
- [80] Raviart P. A., Thomas J. M., *Introduction à l'analyse numérique des équations aux dérivées partielles*, Masson, Paris, 1983.
- [81] Reuschen S., Jobst F., Nowak W., *Discretization-Independent Bayesian Inversion of High-Dimensional Multi-Gaussian Priors Using a Hybrid MCMC*, Water Resources Research, vol. 57, número 8, 2021.
- [82] Robert C. P., Casella G., *Monte Carlo Statistical Methods*, Springer Texts in Statistics, ISBN: 9781475741452, 2013.
- [83] Rue H., Held L., *Gaussian Markov Random Fields: Theory and Applications*, Chapman and Hall/CRC, 2005, ISBN: 9780429208829.
- [84] Sandrin L., Tanter M., Catheline S., Fink M., *Shear modulus imaging with 2-D transient elastography*, IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, vol. 49, número 4, pp. 426-435, 2002.
- [85] Schillings C., Sprungk B., Wacker P., *On the convergence of the Laplace approximation and noise-level-robustness of Laplace-based Monte Carlo methods for Bayesian inverse problems*, Numerische Mathematik, vol. 145, número 4, pp. 915–971, 2020.
- [86] Steinwart I., *Convergence types and rates in generic Karhunen-Loève expansions with applications to sample path properties*, Potential Analysis, vol. 51, pp. 361-395, 2019.
- [87] Tarantola A., *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics (SIAM), 2005, ISBN: 9780898715729.
- [88] Tierney L., *Markov Chains for Exploring Posterior Distributions*, The Annals of Statistics, vol. 22, número 4, pp. 1701 - 1728, 1994.
- [89] Tierney L., Kadane J.B., *Accurate approximations for posterior moments and marginal densities*, Journal of the American Statistical Association, vol. 81, número 393, pp. 82-86, 1986.
- [90] Tikhonov A. N., Goncharsky A. V., Stepanov V. V., Yagola A. G., *Numerical Methods for the Solution of Ill-Posed Problems*, Springer-Science+Business Media, 1977, ISBN: 9789401584807.
- [91] Tromp J., *Seismic Wavefield Imaging of Earth's Interior across Scales*, Nature Reviews Earth and Environment, vol. 1, número 1, pp. 40-53, 2020.
- [92] Uribe F., Papaioannou I., Betz W., Straub D., *Bayesian inference of random fields represented with the Karhunen-Loève expansion*, Computer Methods in Applied Mechanics and Engineering, vol. 358, 2020.
- [93] Van de Schoot R., Depaoli S., King R., Kramer B., Märtens K., Tadesse M., Vannucci M., Gelman A., Veen D., Willemsen J., Yau C., *Bayesian statistics and modeling*, Nature Reviews Methods Primers, vol. 1, número 1, 2021.
- [94] Van der Vaart A.W., *Asymptotic Statistics*, Cambridge University Press, 2000, ISBN:

9780511802256.

- [95] Wang R., *Introduction to Orthogonal Transforms Karhunen-Loève transform and principal*, Cambridge University Press, 2012, ISBN: 9781139015158.
- [96] Xiu D., Hesthaven J. S., *High-Order Collocation Methods for Differential Equations with Random Inputs*, SIAM Journal on Scientific Computing, vol. 27, número 3, pp. 1118-1139, 2005.
- [97] Xu Z., Liao Q., Li J., *Domain-decomposed Bayesian inversion based on local Karhunen-Loève expansions*, Journal of Computational Physics, vol. 504, 2024.
- [98] Zhu H., Li S., Fomel S., Stadler G., Ghattas O., *A Bayesian approach to estimate uncertainty for full-waveform inversion using a priori information from depth migration*, Geophysics, vol. 81, número 5, pp. 307-323, 2016.
- [99] Zienkiewicz O. C., Taylor R. L., *The Finite Element Method: Volume 1, The Basis*, Wiley, 2000, ISBN:9780470395042.
- [100] Zienkiewicz O. C., Taylor R. L., Zhu J. Z., *The Finite Element Method: Its Basis and Fundamentals*, Butterworth-Heinemann, 2013, ISBN: 0750663200.