



CIMAT

Centro de Investigación en Matemáticas, A.C.

CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS

SENSADO COMPRIMIDO EN IMÁGENES DE
RESONANCIA MAGNÉTICA PESADAS POR
DIFUSIÓN DE AGUA

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

**Maestría en ciencias computacionales y matemáticas
industriales**

PRESENTA:

Ricardo Alonso Rios Carrillo

TUTORES:

Dr. Julio C. Estrada Rico

Dr. Alonso Ramírez Manzanares

Guanajuato, Guanajuato, 2017

A mi padre, mi madre y mi tía, por su invaluable apoyo. Los amo.

Reconocimientos

Quisiera agradecer principalmente y sobre todo a mis asesores el Dr. Julio Estrada y el Dr. Alonso Ramírez. Este último año ha sido particularmente complicado para mi, pero gracias a su paciencia, dedicación, consejo y apoyo me han salvado (dicho sin exageración) de posibles caminos muy oscuros. Es gracias a ellos que la elaboración de esta tesis y mi vida académica ha podido salir adelante, ambos se han hecho un modelo a seguir para mi persona, y por ello estoy mas convencido de seguir por el noble camino de la ciencia.

A todos mis compañeros, amigos y profesores que he tenido la fortuna de conocer en CIMAT, todos y cada uno de ellos han hecho invaluable contribuciones a mi persona, lo cual han hecho de CIMAT una de las experiencias mas trascendentales en mi vida. Espero haber podido devolver el gesto con creces.

Quiero agradecer al Dr. Salvador Ruíz Correa (IPICYT), la Dra. Rebeca Romo (UdeG) y el Dr. Alejandro Morales (UdeG), que en primer lugar es gracias a su ejemplo y apoyo por lo cual he dado y encontrado con mi lugar en CIMAT.

Agradezco a toda mi familia, que siempre esta conmigo en los momentos mas difíciles, su apoyo incondicional es invaluable.

Finalmente quiero agradecer a mi *alma mater*, CIMAT, por todo el apoyo económico, académico y personal. Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por su apoyo económico a lo largo de mis estudios de Maestría, debido a ello estudiantes como yo que dependemos de ello para continuar nuestra búsqueda por la ciencia.

Declaración de autenticidad

Por la presente declaro que, el contenido de esta tesis es original y no se ha presentado total o parcialmente para su consideración para cualquier otro título o grado en esta o cualquier otra Universidad. Esta tesis es resultado de mi propio trabajo y no incluye nada que sea el resultado de algún trabajo realizado en colaboración, salvo que se indique específicamente en el texto.

Ricardo Alonso Rios Carrillo. Guanajuato, Guanajuato, 2017

Índice general

Índice de figuras	IX
Índice de tablas	XI
1. Introducción	1
1.1. Objetivo - Motivación	1
1.2. Metodología	2
1.3. Contribuciones	3
1.4. Estructura de la tesis	3
2. El sensado comprimido	5
2.1. El muestreo de señales	5
2.2. Principios del sensado comprimido	6
2.2.1. Raleza	7
2.2.2. Incoherencia	7
2.3. Recuperación de una señal con sensado comprimido	9
2.3.1. Sensado Comprimido Robusto	10
2.3.1.1. Lidiando con el ruido: El problema LASSO	11
2.3.2. Matrices de sensado	12
3. Principios y técnicas de imagenología de resonancia por difusión	13
3.1. Principios físicos de difusión	13
3.2. Imágenes de resonancia pesadas por difusión	14
3.3. Imágenes por tensor de difusión	18
3.3.1. Modelo de tensor multicompartimiento	19
3.4. Imagenología del espacio Q	21
3.5. Imágenes de espectro por difusión	21
3.5.1. Función de distribución-orientación	22
3.6. Ruido en DW-MRI	23
4. Metodología	25
4.1. Diseño de experimento	25
4.1.1. Métricas de error	27

4.2.	Generación sintética de la señal S	28
4.2.1.	Señales S con cruces de fibras	31
4.2.2.	Vectorizando el espacio 3D	31
4.3.	Diseño de la matriz de sensado comprimido A	32
4.3.1.	Construcción de matriz de sensado de Fourier	32
4.3.2.	Construcción de diccionario de representación basado en distri- buciones de difusión	32
4.4.	Muestreo Ω de la señal $S(q)$	36
4.5.	Optimizando parámetros	38
4.5.1.	Valores iniciales	38
4.5.2.	Parámetro Ω	39
4.5.3.	Parámetro λ	41
5.	Resultados y análisis	43
5.1.	Desempeño de métricas bajo variaciones de ruido Riciano	43
5.2.	Resultados de optimización de parámetros	44
5.2.1.	Convergencia parámetro λ	46
5.2.2.	Convergencia de Ω	48
5.3.	Desempeño de recuperación del propagador conjunto promedio	49
5.3.1.	Ejemplos de reconstrucción	50
6.	Conclusiones	53
6.1.	Trabajo a futuro	54
Apéndice A. Algoritmo ADMM		55
Apéndice B. Información adicional sobre el valor q		56
Apéndice C. Métricas entre distribuciones		57
Bibliografía		59

Índice de figuras

3.1. Secuencia de pulsos PGSE	15
3.2. Decaimiento de la señal de DW-MRI en presencia de difusión.	18
3.3. Matrices de tensores de difusión y sus elipsoides correspondientes.	20
3.4. Estimación de la ODF en DSI	23
4.1. El grid en el protocolo de DSI	29
4.2. Diccionario P	33
4.3. Direcciones en la media esfera del diccionario P	36
4.4. Ejemplo de muestreos en 2D	38
5.1. Ajuste lineal visual de métricas de error	45
5.2. Ajuste lineal visual de métricas de error para $SNR \geq 30$	46
5.3. Rango aproximado de valores óptimos λ	47
5.4. Rango aproximado de valores óptimos λ de muestreos con handicap	47
5.5. Ejemplo de reconstrucción 1	51
5.6. Ejemplo de reconstrucción 2	52
5.7. Ejemplo de reconstrucción 3	52

Índice de tablas

4.1. Parámetros del protocolo DSI	28
4.2. Total de muestras m	39
5.1. Desempeño promedio de métricas de error a diferentes SNR	44
5.2. Varianza del desempeño promedio de métricas de error a diferentes SNR	45
5.3. Valores óptimos de λ para un muestreo elite	48
5.4. Desempeño de muestreos Gaussianos	49
5.5. Desempeño de muestreos Gaussianos con handicap	49
5.6. Desempeño de muestreos elite	50

Introducción

Las imágenes de resonancia magnética pesadas por difusión (DW-MRI en sus siglas en inglés: “Diffusion Weighted Magnetic Resonance Imaging”) aprovechan el fenómeno de difusión de las moléculas de agua en tejido biológico, produciendo imágenes médicas que reflejan el mapeo del proceso de difusión en el tejido biológico, permitiendo inferir la micro estructura del tejido en cuestión, tales como células, fibras, membranas etc.

Para algunas aplicaciones de imágenes de resonancia magnética pesadas por difusión (como en tractografía), se requiere obtener el *conjunto propagador promedio* (EAP, por sus siglas en inglés: “Ensemble Average Propagator”) [1] de cada vóxel en el estudio. El EAP describe una distribución de probabilidad de como las moléculas de hidrógeno (en general estamos hablando de agua) se han desplazado dentro del vóxel, lo cual ayuda a inferir la infraestructura dentro del vóxel, tales como fibras axonales etc.

Por otro lado, en los últimos años ha surgido en el campo del procesamiento de señales un esquema de muestreo y recuperación de señales conocido como sensado comprimido (CS por sus siglas en inglés: “compressed sensing” o “compressive sampling”) [2], el cual dicta que, si se cumplen ciertas suposiciones en el esquema de muestreo, uno puede muestrear y recuperar la señal adecuadamente con un número de muestras por debajo de metodologías habituales.

1.1. Objetivo - Motivación

Para aproximar el EAP existen varios métodos (como imágenes por tensor de difusión [3]). Sin embargo, el método más preciso para obtener el EAP son las *imágenes de espectro por difusión* (DSI por sus siglas en inglés: “Diffusion Spectrum Imaging”) [4], el cual obtiene una discretización del EAP aprovechando la relación de Fourier que existe entre el EAP y el conocido como espacio Q [1] (este siendo el espacio donde se mide directamente la señal de DW-MRI). Sin embargo, para que el DSI sea efectivo se requiere una cantidad considerable de muestras del espacio Q, lo cual implica un problema práctico muy grande para este tipo de imágenes: **Mucho tiempo de muestreo.**

El objetivo de esta tesis es aplicar adecuadamente la técnica del sensado comprimido para realizar DSI en imágenes de DW-MRI con un número de muestras significativamente menor al requerido en el DSI. Esto implica mucho menor tiempo de muestreo, y obtener aun así imágenes con suficiente calidad para diferentes aplicaciones.

Otro objetivo consiste en aprovechar el esquema de la teoría del CS para recuperar el EAP al proponer un diccionario que introduce información de distintas distribuciones de hidrógeno, lo cual no se ha visto en nuestro conocimiento en otro trabajo con objetivos similares.

1.2. Metodología

La teoría del sensado comprimido requiere resolver el siguiente problema de optimización convexo [2]:

$$\min_x \|A_\Omega x - y_\Omega\|_{\ell_2} + \lambda \|x\|_{\ell_1} \quad (1.1)$$

donde:

- y : Es la señal a muestrear de tamaño (en el caso discreto) n . y_Ω implica que ya está muestreada, con un tamaño m .
- A : Una matriz de tamaño $n \times l$ compuesta de dos bases, conocida como la matriz de sensado comprimido.
- x : Una representación rala de la señal a recuperar, de tamaño l .
- Ω : Conjunto de índices con los cuales se muestrea a la señal, existirán m índices en tal conjunto.
- λ : Parámetro de regularización, relacionado a la tolerancia al ruido.

La señal y corresponde a señales sintéticas de DW-MRI bajo los parámetros del protocolo estándar de DSI. Para validar la robustez de la reconstrucción de la señal obtenida en la solución, estas señales corresponden a simulaciones de cruces de fibras (donde modelos mas simples como tensores de difusión fallan) donde también agregamos cierto nivel de ruido sintético. Para validar la calidad de las reconstrucciones obtenidas, se exploran algunas métricas de distancia entre distribuciones, las cuales también sirven de apoyo en la optimización de la metodología.

La matriz de sensado comprimido A se compone de una combinación de dos bases tal que $A = \Phi\Psi$, donde:

- Φ : Es la base de sensado, de tamaño $n \times n$. En este caso se trata de una matriz de bases de Fourier debido a que este es el esquema utilizado en imágenes DW-MRI.
- Ψ : Es una base de representación rala de la señal a recuperar, de tamaño $n \times l$. En esta tesis, esta base corresponde a un diccionario, el cual contiene información sobre posibles distribuciones de difusión. Es importante destacar que el EAP (distribución p) se obtiene al realizar la conversión $p = \Psi x$.

Los parámetros λ y Ω se estiman de una manera aproximadamente óptima mediante un sencillo esquema de aprendizaje de máquina supervisado el cual será visto más adelante.

1.3. Contribuciones

Algunos trabajos con objetivos similares a los presentados en esta tesis usan bases wavelets como base de representación sparse Ψ [5] [6], otros usan bases sparse más elaboradas y específicas al modelo [7] [8]. Sin embargo, a pesar de que cumplen en mayor o menor medida con la rareza requerida, estas bases no aportan información extra sobre el EAP que buscamos recuperar.

Esta tesis hace dos principales contribuciones

- Propone un esquema para hacer uso de sensado comprimido en imágenes de resonancia magnética pesadas por difusión, específicamente para imágenes con DSI.
- La base de representación Ψ propuesta en esta tesis aprovecha el modelo para introducir información de distribuciones de difusión de hidrógeno, y al mismo tiempo haciendo que la señal x sea lo más rara posible, como es requerido en el CS.

De esta forma, la señal x no sólo cumple con los requerimientos del CS, sino que además también contiene información de la posible microestructura dentro del vóxel debido a que proviene de un modelo que supone cierto tipo de micro estructura desde un inicio, lo cual, en nuestro conocimiento, no ha sido explorado en trabajos con objetivos similares a los de esta tesis.

1.4. Estructura de la tesis

Este trabajo está dividido en 6 capítulos, siendo el primer capítulo la introducción a la tesis, la cual hace mención de sus objetivos y motivaciones. El segundo y tercer capítulo engloban el marco teórico necesario para el desarrollo de esta tesis, los cuales detallan respectivamente a la teoría del sensado comprimido y de imágenes de resonancia magnética pesadas por difusión. El cuarto capítulo contiene la metodología de trabajo propuesta en esta tesis para cubrir los objetivos. El quinto capítulo presenta los resultados obtenidos mediante la metodología propuesta, así como un análisis y detalles de los mismos. Finalmente el sexto capítulo contiene unas conclusiones breves que resumen la calidad de los resultados obtenidos y el posible trabajo a futuro que surge gracias a los resultados de esta tesis.

El sensado comprimido

En este capítulo se introduce la teoría relacionada con el esquema del sensado comprimido, el cual es un modelo reciente para muestrear señales con un número de muestras menor a los utilizados en métodos habituales, de tal forma que podamos recuperar dicha señal sin pérdida de información. Este capítulo hace una breve introducción del esquema habitual de recuperación de señales y posteriormente los principios necesarios para el sensado comprimido, así como algunas propuestas para resolver el problema que este supone.

2.1. El muestreo de señales

En el campo de procesamiento de señales, para poder muestrear y representar correctamente una señal f , requerimos una cantidad suficiente de n muestras que componen a la señal f . El proceso general de muestreo se puede definir como [2]

$$y_k = \langle f, \phi_k \rangle, k \in [1, 2 \dots n]. \quad (2.1)$$

Esto implica que vamos a correlacionar el objeto que queremos adquirir f con una base de elementos ϕ_k , la cual podemos representar en una matriz de sensado Φ . En el ejemplo más clásico, si estos elementos son funciones delta de Dirac, entonces y es un vector de valores muestreados de f en el dominio del tiempo o espacio. En otro ejemplo, si estas funciones son formas de onda con sinusoidales, entonces y es un vector de coeficientes de Fourier. Este último ejemplo es la modalidad de sensado utilizado en imágenes de resonancia magnética.

Actualmente está demostrado que es posible reconstruir una señal f de manera exacta sobre un dominio $[a, b]$ si obtenemos las suficientes muestras con una frecuencia de muestreo F adecuada [9]. Esto es conocido como el teorema de muestreo o el teorema de Nyquist, y dicta que para que la reconstrucción de una señal muestreada sea exacta, se requiere que la frecuencia de muestreo sea al menos dos veces más grande que la máxima frecuencia, F_{max} , contenida en la señal, es decir que haya n muestras contenidas en el dominio $[a, b]$ con una muestra tomada cada $\frac{2}{F_{max}}$, ver Teorema 1.

Teorema 1 Teorema de Nyquist. *Si la máxima frecuencia contenida en una señal analógica x es F_{max} y la señal es muestreada a una tasa de muestreo $F_s > 2F_{max}$, entonces la señal x puede ser reconstruida de manera exacta con sus valores muestreados mediante una función de interpolación [10].*

La mayoría de la tecnología actual, basada en el procesamiento digital (muestreado) de señales, tiene como base este criterio. Desafortunadamente, en muchas importantes aplicaciones emergentes, cumplir este criterio puede ser muy costoso, o si lo cumple, la cantidad de muestras requeridas puede ser muy elevada [11]. Por eso, esquemas como el Sensado Comprimido, presentado en la siguiente sección, se ha propuesto como una herramienta para este tipo de necesidades, donde en vez de tomar n muestras (el mínimo necesario para representar la señal continua) tomaríamos una cantidad de m muestras mucho menor, y posteriormente recuperaríamos la señal de dimensión n que representa a la señal continua f .

2.2. Principios del sensado comprimido

El sensado comprimido (CS por sus siglas en inglés: “compressed sensing” o “compressive sampling”) [2] es un paradigma en el procesamiento de señales que innova dentro de los esquemas de muestreo-recuperación habituales en la adquisición de datos. La teoría del CS muestra que es posible recuperar señales con menores muestras de los que usan métodos tradicionales, esto es por debajo del establecido por el Teorema de Nyquist.

De hecho, el CS sugiere formas económicas de traducir datos analógicos a datos digitales ya comprimidos [12] como a continuación se explica. Se sabe que ciertas señales tienen cierta estructura, por lo tanto pueden ser comprimidas sin mucha pérdida de información. Por ejemplo, algunos codificadores modernos aprovechan el hecho de que señales tienen representaciones raras, y por lo tanto sólo requieren transmitir estos coeficientes de manera selectiva en vez de toda la señal. Generalmente para esto se requiere muestrear toda la señal, se computan todos los coeficientes y sólo se codifican de vuelta los más significativos, descartando el resto. Este proceso de adquisición masivo de datos desperdicia muchos recursos, ya que si estas señales son comprimibles, posiblemente no sea necesario adquirir todos los datos. El CS muestra que efectivamente no es necesario hacerlo.

Además, existen formas de utilizar optimización numérica para reconstruir las señales completas a partir de esta pequeña cantidad de datos recolectados. Por lo tanto el CS es un protocolo eficiente de adquisición de datos que muestrea a una frecuencia muy baja y utiliza poder computacional para reconstruir la señal de lo que parecía ser un conjunto incompleto de muestras.

El CS se basa principalmente en 3 principios [7]:

- **Raleza:** Expresa la idea de que la información de la señal está contenida en pocos elementos de una base de representación.
- **Incoherencia:** Expresa la idea de que la señal de interés debe estar desparramada en el dominio de adquisición. Esta idea se encuentra relacionada con el principio de incertidumbre de la transformada de Fourier, donde información localizada en el espacio directo, se encuentra esparcida en el espacio de Fourier y viceversa.
- Un método de recuperación basado en una regularización que promueva soluciones ralas, como la norma l_1 , con un protocolo eficiente de muestreo.

Aunque la teoría del CS se puede aplicar en señales continuas [2], en esta tesis vamos a enfocar el estudio de la teoría del sensado comprimido en el caso de señales discretas tales que $f \in \mathbb{R}^n$.

2.2.1. Raleza

Cuando una señal tiene una representación rala (sparse), implica que la señal contiene pocos coeficientes que no tienen valor igual a cero. En la práctica no siempre es posible tener señales ralas con muchos coeficientes que son exactamente igual a cero. Por ello, podemos decir que la raleza se define con la proporción de información relevante de una señal que se concentra en pocos coeficientes, con el resto siendo relativamente muy pequeños. En este trabajo llamaremos a una señal S-rala si contiene S elementos significativamente diferentes de cero en ella.

Matemáticamente hablando, si tenemos un vector $f \in \mathbb{R}^n$ el cual expandimos mediante la base $\Psi = [\psi_1, \psi_2, \dots, \psi_l]$ tal que

$$f_k = \sum_{i=1}^l x_i \psi_{ki}, \quad (2.2)$$

donde sea x_i el coeficiente en la secuencia $x_i = \langle f, \psi_i \rangle$, nos resulta conveniente expresar a la señal de la forma matricial

$$f = \Psi x. \quad (2.3)$$

Cuando una señal tiene una representación rala, uno puede descartar los coeficientes pequeños sin mucha pérdida de información, pues conservar aquellos grandes va a ser suficiente para recuperar a f sin mucha diferencia numérica.

2.2.2. Incoherencia

La coherencia mide la correlación existente más larga entre los elementos de una base de sensado Φ y la base de representación rala Ψ . La incoherencia naturalmente se refiere a la falta de coherencia entre estas dos bases, lo cual expresa la idea de que la información entre ambas bases se encuentre “desparramada”.

2. EL SENSADO COMPRIMIDO

La coherencia entre la base de sensado Φ (compuesta de los vectores fila ϕ_k) y la base de representación Ψ se define como [2] :

$$\mu(\Phi, \Psi) = \sqrt{n} \max_{k=[1, \dots, n], j=[1, \dots, l]} |\langle \phi_k, \psi_j \rangle|, \quad (2.4)$$

esto partiendo de la suposición de que la base de sensado Φ y la base de representación sparse Ψ son ortobases en \mathbb{R}^n . Esta restricción no es esencial pero simplifica el concepto aquí presentado.

De forma equivalente, consideremos a la matriz resultante de multiplicar estas dos bases $A = \Phi\Psi$, tal que $a_{ij} = \langle \phi_i, \psi_j \rangle$, podemos expresar la coherencia como

$$\mu(A) = \sqrt{n} \max_{i=[1, \dots, n], j=[1, \dots, l]} |a_{ij}|. \quad (2.5)$$

También podemos interpretar a $\mu(A)$ como que tan concentrada está la energía de A : tener alta coherencia implica tener muy concentrada la energía de la base en alguna columna de A .

Si entre Φ y Ψ hay elementos correlacionados, la coherencia es grande, si no lo están, entonces tenemos alta incoherencia. De que tan grande o pequeña es la incoherencia, cabe decir que (ver en [2])

$$\mu(\Psi, \Phi) \in [1, \sqrt{n}], \quad (2.6)$$

donde la cota superior es el caso donde un elemento de algún vector columna en A (para nuestro caso con A normalizado) es igual a 1 (esto implica que es el único elemento con un valor diferente de cero para dicha columna). La cota inferior es el caso especial donde los componentes de los vectores columna en A tienen el mismo valor $a_{1j} = a_{2j} = \dots a_{nj}$ para $j \in [1, l]$, es decir, son vectores planos, tal que $\sqrt{\sum_{i=1}^n a_{ij}^2} = 1$, lo cual resulta de $a_{1j} = a_{2j} = \dots a_{nj} = \frac{1}{\sqrt{n}}$ [7].

Para comprender más adecuadamente el concepto de incoherencia, consideremos la matriz A expuesta anteriormente en un escenario general donde queremos recuperar la señal S-rala $x \in \mathbb{R}^l$ de un vector de observaciones $y \in \mathbb{R}^n$, tal que $y = Ax$. Idealmente, debido a que x es S-rala, requerimos que A admita cierto nivel de incoherencia para la correcta recuperación. Si $\mu(A)$ es grande, esto implica que la información significativa para recuperar a x se encuentra en sólo algunas muestras. Por el contrario, si la energía de la base está poco concentrada, entonces la información en x se encuentra distribuida en las observaciones.

Para visualizar este concepto, consideremos el siguiente ejemplo. Supongamos que $x \in \mathbb{R}^l$ es S-rala con $S = 1$. El coeficiente no-cero es el $k \in [1, l]$, esto es

$$\begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1l} \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nl} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ x_k \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1k}x_k \\ \vdots \\ \vdots \\ \vdots \\ a_{nk}x_k \end{bmatrix}. \quad (2.7)$$

Suponga que $a_{ik} = 1$, eso implica que $\mu(A)$ se encuentra en la cota superior de coherencia. Debido a que la norma de esta columna es igual a 1, entonces $a_{kl} = 0$ para cualquier $j \neq i$, por lo tanto

$$\begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ a_{ik}x_k \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ x_k \\ \vdots \\ 0 \end{bmatrix}. \quad (2.8)$$

En este caso, estamos obligados a muestrear a y_i , de otra forma la señal x desaparece.

Ahora supongamos que $\mu(A)$ llega a su cota inferior. Esto implica que A es una matriz con todos sus componentes iguales a $\frac{1}{\sqrt{n}}$, entonces

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \frac{1}{\sqrt{n}} \begin{bmatrix} x_k \\ \vdots \\ x_k \end{bmatrix}. \quad (2.9)$$

En este caso, podemos muestrear en cualquier punto y aún así tener información sobre x_k . La información de x se encuentra distribuida en la observación. El sensado comprimido busca tener pares de bases poco coherentes entre sí, esto se hará más evidente una vez hayamos visto el problema de CS en la Sección 2.3.

2.3. Recuperación de una señal con sensado comprimido

En el CS se busca recuperar la señal $f \in \mathbb{R}^n$ de un vector de observaciones $y \in \mathbb{R}^m$ donde $m < n$, tal que

$$y_\Omega = \Phi_\Omega f, \quad (2.10)$$

donde Ω de cardinalidad m , define un subconjunto de coordenadas indicando en que puntos evaluamos y adquirimos a y_Ω . La matriz Φ_Ω la definimos como resultado de la operación $\Phi_\Omega = M_\Omega \Phi$, donde la matriz auxiliar de muestreo M_Ω (de tamaño $m \times n$), corresponde a los m renglones de una matriz identidad (de tamaño $n \times n$) correspondientes a las coordenadas contenidas en Ω .

Este problema está mal determinado al tener menos ecuaciones que variables conocidas ($m < n$). Sin embargo, en el sensado comprimido, podemos recuperar la señal f , representándola mediante $f = \Psi x$, donde la representación $x \in \mathbb{R}^l$ debe ser muy rala y es la solución del problema de optimización convexa:

$$\begin{aligned} \min_x \quad & \|x\|_{\ell_1} \\ \text{s.t.} \quad & y_\Omega = \Phi_\Omega \Psi x, \end{aligned} \quad (2.11)$$

donde

$$\|x\|_{\ell_1} := \sum_i |x_i|. \quad (2.12)$$

Para fines de claridad de esta tesis, formulamos este problema expresando la matriz de sensado Φ y la matriz de representación rala Ψ en la matriz de sensado comprimido A

$$A = \Phi\Psi, \quad (2.13)$$

tal que $A_\Omega = M_\Omega A$, por lo tanto $y_\Omega = \Phi_\Omega \Psi = A_\Omega x$.

El problema de optimización busca principalmente dos objetivos:

- Tener consistencia con los datos observados: La restricción $y_\Omega = \Phi_\Omega \Psi x$ agrega consistencia a los datos. De esta forma se quiere que la solución se mantenga lo más cercana posible a los datos adquiridos.
- Promover soluciones rala: La minimización con norma ℓ_1 promueve soluciones rala en x (en el CS buscamos que la señal x sea lo más rala posible).

La teoría del CS dicta que al seleccionar m muestras aleatorias, tal que si:

Teorema 2 $m \geq C \cdot \mu^2(\Phi, \Psi) \cdot S \cdot \log n$,

donde S es el número de componentes que no son cero en x y para alguna constante positiva C , la solución a la Ecuación (2.11) es muy probablemente exacta [13]. Del Teorema 2 hacemos énfasis en algunas cosas:

- Es evidente la importancia de la incoherencia: a menor coherencia, se requieren menos muestras.
- Igualmente mientras más rala sea x , se van a requerir menos muestras.
- Uno puede muestrear m de mucho menor tamaño que n sin pérdida de información.
- El esquema (en general) no requiere información previa de los coeficientes que no son cero de x . Con este esquema se asume que todo es desconocido en un principio, al resolver el problema y si la señal es lo suficientemente rala, la reconstrucción ocurre.

2.3.1. Sensado Comprimido Robusto

Para que el CS sea útil en aplicaciones prácticas, se requiere que éste sea capaz de funcionar con señales casi rala y con señales ruidosas. Generalmente las señales en el mundo práctico no son completamente ralas, sino aproximadamente rala. De la misma

forma, en casi cualquier muestreo los datos se obtendrán con cierto nivel de ruido, el cual es inevitable en dispositivos físicos de sensado o debido al modelado de la señal.

Un concepto clave en el estudio de la robustez del sensado comprimido es el conocido como “Restricted Isometry Property” (RIP). Con este concepto, se han establecido diversos resultados [2] sobre la calidad de la reconstrucción cuando la matriz de sensado comprimido A_Ω obedece el RIP, definido como:

Definición 1 Para cada entero $s = 1, 2, \dots$, se define la constante isométrica δ_s de la matriz A_Ω como el número más pequeño tal que

$$(1 - \delta_s)\|x\|_{\ell_2} \leq \|A_\Omega x\|_{\ell_2} \leq (1 + \delta_s)\|x\|_{\ell_2}$$

se mantiene para todos los vectores x S -ralo, esto es aquellos vectores x que tienen al menos s entradas que no son cero.

Se puede decir que A_Ω obedece el RIP de orden s si δ_s no es muy cercano a 1. Cuando esta propiedad se mantiene, los vectores S -ralos no pueden estar en el espacio nulo de A_Ω , lo cual es muy útil, si no, no habría forma de reconstruir estos vectores. Una descripción equivalente al RIP es decir que todos los subconjuntos con s columnas tomadas de A_Ω con casi ortogonales [2].

Particularmente Candés demostró en [2] que para $\delta_s < \sqrt{2} - 1$ en una recuperación sin ruido, x^* (el cual es la solución exacta para la Ecuación (2.11)) obedece que:

Teorema 3 $\|x^* - x\|_{\ell_2} \leq C_0 s^{-1/2} \|x_s - x\|_{\ell_1}$,

con C_0 una constante, y x_s es el vector x pero solo conservando los s valores más altos e igualando el resto de los valores a 0. Entonces, si x tiene exactamente s coeficientes no cero ($x = x_s$), la solución es exacta.

2.3.1.1. Lidiando con el ruido: El problema LASSO

Ahora considere que tenemos una adquisición de datos ruidosa, tal que

$$y_\Omega = A_\Omega x + z, \tag{2.14}$$

donde z es un termino de error desconocido. Replanteamos el problema descrito en la Ecuación (2.11) con una restricción relajada:

$$\begin{aligned} \min_x \quad & \|x\|_{\ell_1} \\ \text{s.t.} \quad & \|A_\Omega x - y_\Omega\|_{\ell_2} \leq \epsilon, \end{aligned} \tag{2.15}$$

donde ϵ es el nivel de ruido en los datos. El problema anterior tiene una forma muy conocida y estudiada en el campo de estadística y matemáticas computacionales, conocida como el problema LASSO [14]. Este problema lo podemos escribir como

$$\min_x \quad \|A_\Omega x - y_\Omega\|_{\ell_2} + \lambda \|x\|_{\ell_1}, \tag{2.16}$$

donde es ahora el parámetro λ el que regula la tolerancia al ruido en el problema (elevar el valor de λ relaja la tolerancia al ruido), además de promocionar lo ralo de la solución x . Los problemas en las Ecuaciones (2.15) y (2.16) son equivalentes para una elección adecuada de λ para valores de ϵ . El problema en la Ecuación (2.16) es a su vez convexo y puede ser resuelto eficientemente. Una forma de hacerlo es con el algoritmo ADMM descrito en el Apéndice A.

Candés establece en [2] que asumiendo $\delta_{2s} < \sqrt{2} - 1$, el error entre x^* y x en la Ecuación (2.16) es

Teorema 4 $\|x^* - x\|_{\ell_2} \leq C_0 s^{-1/2} \|x_s - x\|_{\ell_1} + C_1 \epsilon$.

Donde las constantes C_0 y C_1 suelen ser pequeñas. Aquí se observa que el error de la reconstrucción está limitado a la suma de dos términos. El primero que ocurriera si fueran datos sin ruido, el segundo relacionado al nivel de ruido ϵ .

2.3.2. Matrices de sensado

El RIP establece una condición para definir adecuadamente la matriz de sensado comprimido A_Ω , ya que nos permite evaluar la eficiencia de nuestro esquema de muestreo para obtener una reconstrucción adecuada. Por ello, se busca encontrar matrices de sensado Φ que cumplan con el RIP.

Candés menciona [2] diferentes esquemas para construir matrices aleatorias Φ tal que cumplan el RIP con una probabilidad muy elevada. Uno también puede establecer el RIP para pares de bases Φ, Ψ como han sido descritas a lo largo de este capítulo. Es importante destacar que estas matrices aleatorias tienden a tener alta incoherencia con matrices no aleatorias, lo cual ayuda al modelo de manera general.

Principios y técnicas de imagenología de resonancia por difusión

En el campo de la imagenología médica, una de las técnicas más poderosas y utilizadas en estudios de diagnóstico, entre otras aplicaciones, son las Imágenes de Resonancia Magnética (MRI por sus siglas en inglés). El MRI es una técnica de imagenología tomográfica no invasiva que produce imágenes con características internas de objetos, mediante mediciones de señales de resonancia magnética nuclear (NMR en sus siglas en inglés). Es utilizada principalmente en el campo médico para producir imágenes de alta calidad dentro del cuerpo humano. Físicamente el MRI aprovecha propiedades intrínsecas de las moléculas de hidrógeno del agua en el cuerpo humano y la interacción que estas tienen al absorber y emitir energía de radio frecuencia cuando son expuestas a un campo magnético fuerte externo. Se deja al lector varias referencias para mayor información sobre el MRI si es requerido en [15] [16] [17] .

Las imágenes de resonancia magnética pesadas por difusión hacen un uso específico de secuencias de MRI que, aprovechando el fenómeno de difusión de las moléculas de agua en tejido biológico, produce imágenes de contraste de MRI. Esas imágenes reflejan el mapeo del proceso de difusión en el tejido biológico, permitiendo inferir la micro estructura del tejido en cuestión, tales como presencia de células, fibras, membranas etc.

En este capítulo, se hará una introducción a los principios básicos de DW-MRI, así como una introducción general a algunas técnicas utilizadas en este tipo de imágenes que están relacionadas con la propuesta de investigación de esta tesis.

3.1. Principios físicos de difusión

La difusión es un proceso de transporte existente en la naturaleza, en la cual las moléculas o átomos en regiones muy concentradas se mueven a regiones menos concentradas, esto sin necesidad de energía “externa” para ello. En el campo de la física, las

3. PRINCIPIOS Y TÉCNICAS DE IMAGENOLOGÍA DE RESONANCIA POR DIFUSIÓN

bases de este fenómeno fueron descritas en la “Primera ley de Fick” [18]

$$J = -D\delta C \quad (3.1)$$

donde J es el flujo de las partículas (total de partículas atravesando una unidad de área por segundo), δC es el gradiente de la concentración de las partículas, y D es el coeficiente de difusión. El coeficiente de difusión D es una propiedad intrínseca del medio. El valor del D es determinado por el tamaño de las moléculas de difusión y la temperatura de las microestructuras en el ambiente, y esto ha permitido su estudio como medio de las propiedades biológicas del tejido [19]. Las unidades del D en el S.I. son m^2/s .

El proceso de difusión se puede entender como un modelo de “caminantes aleatorios”, donde en cada cierto intervalo de tiempo una molécula se mueve cierta distancia en una dirección aleatoria [20]. Este esquema de “caminantes aleatorios” tiene sus orígenes en las observaciones realizadas por Robert Brown [21]. Él reportaba que las partículas que observaba en un microscopio (en granos de polen) se movían sin causa aparente (este movimiento aleatorio producido por el choque de moléculas se le conoce actualmente en la literatura como “movimiento Browniano”). Sin embargo, fue Albert Einstein [22] quien estableció un esquema probabilístico que describía el movimiento de un conjunto de moléculas bajo efectos de la difusión, el cual resultaba en una descripción coherente del proceso de difusión que reconciliaba los dos esquemas presentados.

Einstein introdujo el concepto de “distribución de desplazamiento”, que cuantifica la fracción de partículas que van a desplazarse cierta distancia en cierto tiempo. Con este concepto fue posible derivar una relación explícita entre el desplazamiento cuadrático promedio (con movimiento ahora conocido como Browniano) y el coeficiente de difusión D

$$\langle r^2 \rangle = 6D\Delta, \quad (3.2)$$

para el caso tridimensional, donde $\langle r^2 \rangle$ es el desplazamiento cuadrático promedio de las partículas durante el tiempo de difusión Δ , y D el coeficiente de difusión.

3.2. Imágenes de resonancia pesadas por difusión

La DW-MRI observa el desplazamiento de las moléculas sujetas al movimiento Browniano dentro de un material. En el caso de imágenes médicas, estas partículas son agua, la cual está presente en la mayor parte del tejido humano. Estas partículas de agua en el tejido, se encuentran en constante movimiento aleatorio debido a las fluctuaciones térmicas naturales entre ellas [23].

La finalidad específica del DW-MRI es medir la función de la densidad de probabilidad p de los desplazamientos r de partículas sobre un tiempo fijo t . Esta función de probabilidad p también es conocida en la literatura como el *Propagador Conjunto promedio* (EAP, por sus siglas en inglés: *Ensamble Average Propagator*), debido a que

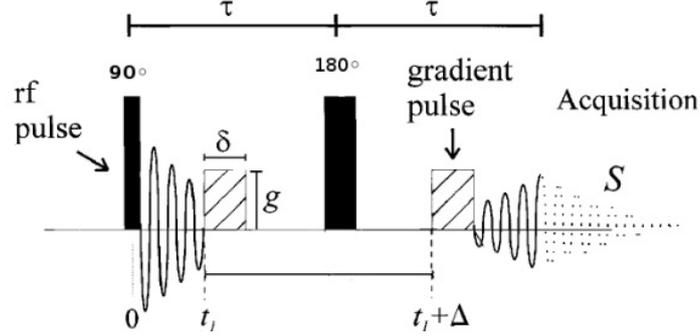


Figura 3.1: Secuencia de pulsos PGSE. La fuerza del gradiente es denotada por g , t es el tiempo inicial donde el primer gradiente es aplicado, el cual tiene una duración δ , mientras que Δ es el tiempo entre el inicio de ambos gradientes. Figura tomada de [24]

es una función de probabilidad de los desplazamientos promedio que ocurren en todo el espacio medido en cuestión. A su vez, las características presentes en el EAP p nos van a dar información de la estructura en el tejido, esto será visto más a detalle en este capítulo.

Las DW-MRI adquieren mediciones que son sensibles al movimiento del núcleo de hidrógeno en las moléculas de agua (los spins) al estar en presencia de un campo magnético lineal no uniforme (un gradiente). Los spins que se mueven de manera aleatoria, sufren cambios en su velocidad de precesión, lo cual resulta en un desfazamiento de estos spins. Esto provoca una reducción en la señal de MRI.

La Figura 3.1 muestra la secuencia *pulsed-gradient spin-echo* (PGSE), la cual es la secuencia más común para DW-MRI y fue propuesta por Stejskal y Tanner [25]. Esta secuencia consiste en parte de lo siguiente [23]:

- La máquina de MRI mantiene un campo magnético H_0 constante y aproximadamente homogéneo. La fuerza de campo magnético H_0 será denominada como B_0 .
- Los spins se alinean con H_0 .
- El pulso de radiofrecuencia de 90° , realizado en el tiempo $t = 0$, pone los spins en el plano transversal perpendicular a H_0 .
 - Los spins precesan sobre H_0 a la frecuencia de Larmor ω_0 . La cual está relacionada con B_0 mediante $\omega_0 = \gamma B_0$, donde γ es la razón giromagnética de los spins. En el caso del agua este valor ronda en $\gamma = 2.675 \times 10^8 \frac{1}{sT}$.
- Desperfectos en la homogeneidad en H_0 causan que la precesión de los spins decaigan gradualmente.

3. PRINCIPIOS Y TÉCNICAS DE IMAGENOLOGÍA DE RESONANCIA POR DIFUSIÓN

- El pulso de 180° centrado en la secuencia invierte la fase de cada spin.
 - Sin la presencia de los pulsos de gradiente, la razón de desfase es la misma antes y después del pulso de 180° tal que los spins vuelven en fase justo al tiempo de adquisición de la señal (también conocido como *Tiempo Echo*).

Los gradientes (que son espacialmente homogéneos) de difusión compensan la fase de cada spin mediante una función lineal de la posición de cada spin. En este esquema, el cambio de fase inducido por el primer pulso de gradiente está dado por

$$\Phi_1 = -q^T x_1, \quad (3.3)$$

donde $x_1 \in \mathbb{R}^3$ es la posición de la partícula durante la aplicación del primer pulso de gradiente. En la expresión anterior, se combinan varios parámetros experimentales en el vector $q \in \mathbb{R}^3$, tal que

$$q = \gamma \delta G, \quad (3.4)$$

donde γ es la constante giromagnética, δ es la duración del pulso de gradiente, y $G \in \mathbb{R}^3$ es el vector con magnitud y dirección de tal gradiente. De manera similar, si la partícula se encuentra en la posición $x_2 \in \mathbb{R}^3$ al momento de la aplicación del segundo gradiente, el cambio de fase está dado por

$$\Phi_2 = -q^T x_2. \quad (3.5)$$

Como se mencionó, el pulso de radio frecuencia de 180° aplicando entre ambos gradientes invierte las fases que pudieron haber sido inducidas por el primer gradiente. El cambio de fase agregado que cada partícula sufre está dada por:

$$\Phi_2 - \Phi_1 = -q^T (x_2 - x_1) = -q^T r, \quad (3.6)$$

donde $r = (x_2 - x_1)$ expresa el desplazamiento entre estas dos posiciones x . Si la partícula en cuestión estuvo estacionaria, tal que $x_2 = x_1$, el cambio de fase se desvanecería. En este caso, la magnitud de la señal no cambiaría. Sin embargo, si las partículas hicieron difusión aleatoriamente sobre el medio, el incremento de fase que ganan durante el primer periodo no cancela necesariamente el decremento de fase adquirido durante el segundo periodo. Esta cancelación incompleta resulta en una dispersión de fase, por lo tanto, la señal resultante es atenuada.

En la literatura, generalmente se utiliza una convención particular para la señal de difusión: $E(q)$, llamada la atenuación de la señal de resonancia magnética. La $E(q)$ se obtiene al dividir la señal de difusión atenuada $s(q)$ entre la señal sin la presencia de ningún gradiente $s_0 = s(0)$, tal que $E(q) = s(q)/s_0$. La señal de atenuación de MR está dada por:

$$E(q) = \int \varrho(x_1) dx_1 \int p(r, \Delta) e^{-iq^T r} dr, \quad (3.7)$$

donde $\varrho(x_1)$ cuantifica la probabilidad de encontrar un spin en la posición x_1 al momento de aplicar el primer pulso de gradiente. Generalmente esta función se toma como una constante de valor unitario (resultado de dicha integral sobre una región con agua). Por lo tanto, generalmente estaremos tratando con la expresión de la forma

$$E(q) = \int p(r, \Delta) e^{-iq^T r} dr, \quad (3.8)$$

(note que es la transformada de Fourier de la función p en la onda q). Donde $p(r, \Delta)$ (el EAP) denota la probabilidad de que una partícula haya tenido un desplazamiento r (donde inicialmente estaba la posición x_1 y terminó en x_2) después de un tiempo Δ , el cual es el tiempo de separación entre los dos gradientes. Además,

$$e^{-iq^T r} = \cos(q^T r) - i \sin(q^T r), \quad (3.9)$$

representa de alguna forma el kernel de Fourier, análogo a la relación frecuencia-tiempo comúnmente utilizada en el procesamiento de señales. La señal de difusión es compleja, y en este caso la parte real (coseno) e imaginaria (seno) representan los dos ejes componentes del vector de magnetización en el plano perpendicular al campo magnético principal. Su integración representa la suma sobre todos los posibles desplazamientos del conjunto.

Si sólo consideramos la parte real de la Ecuación (3.9) en la Ecuación (3.8) tendríamos que

$$E(q) = \int p(r, \Delta) \cos(q^T r) dr. \quad (3.10)$$

Esta última expresión da una idea más clara de como funciona el método de PGSE, donde es más sencillo observar como el cambio de fase debido a los cambios en el campo se cancelan. En ausencia de difusión, los cambios de fase debido a los dos pulsos de gradiente (o en presencia de difusión pero sin gradientes) se cancelan, debido a que $qr = 0$ y dado que $\cos(qr) = 1$, la señal máxima es obtenida. Pero en caso de difusión, como ya mencionamos en la Ecuación (3.6), no necesariamente se cancelan, por lo tanto la señal resultante tendría una atenuación. Un ejemplo de atenuación de la señal sobre tejido se encuentra en la Figura 3.2.

Es importante destacar que la derivación de la Ecuación (3.10) asume que el movimiento de las partículas cuando el gradiente está encendido es despreciable. Esta suposición se justifica en el caso de que δ es muy pequeña en comparación al tiempo de separación entre pulsos Δ , que generalmente es el tiempo de difusión $t = \Delta$. Sin embargo, en la práctica clínica, los valores de δ y Δ son de magnitud similar. Cuando δ no es despreciable, la compensación de fase de los spins también dependen de su trayectoria (mientras el gradiente está encendido) y no sólo de su desplazamiento. Esto complica el modelado a las mediciones relacionadas con el EAP p . Sin embargo, bajo ciertas suposiciones (si el EAP tiene forma Gaussiana, y la forma de los gradientes son rectangulares como en la Figura 3.1) el tiempo no despreciable δ reduce a t en un tiempo efectivo de difusión $t = \Delta - \delta/3$ [23].

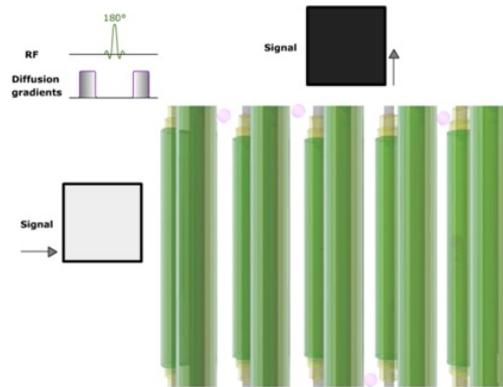


Figura 3.2: Decaimiento de la señal de DW-MRI en presencia de difusión.

En esta figura suponemos un conjunto de fibras orientadas verticalmente. Al aplicar un gradiente en dirección vertical, hay difusión libre y la señal de DW-MRI resultante es muy baja. En cambio, al aplicar un gradiente horizontal la difusión está restringida, lo que resulta en una señal DW-MRI muy elevada. Figura tomada y adaptada de [26]

3.3. Imágenes por tensor de difusión

Si la difusión es libre, el propagador de difusión es Gaussiano y la señal de atenuación es dada por otra Gaussiana, lo cual resulta en una generalización de la Ecuación (3.20) tomando en cuenta la duración de los pulsos, tal que

$$E(q) = e^{-q^T D q (\Delta - \delta/3)}, \quad (3.11)$$

$$E(q) = e^{-bD}, \quad (3.12)$$

donde para pulsos de gradiente rectangulares como en la Figura 3.1, el valor b (en una dimensión) es

$$b = (q^2)(\Delta - \delta/3), \quad (3.13)$$

y es conocido en la literatura como el valor b (b -value). En el Apéndice B se detalla una característica referente al valor que puede tomar q en esta ecuación, relacionado al tipo de estudio en cuestión.

Este formalismo fue utilizado [27] para un caso de difusión libre en un esquema donde se utilizó un “tensor de difusión” D , una matriz de tamaño 3×3 que representa la orientación natural de la difusión anisotrópica con respecto a los 3 ejes de referencia de laboratorio. A partir de esto se desarrolló un esquema más general llamado *Imágenes por tensor de difusión* (DTI, por sus siglas en inglés: *Diffusion Tensor Imaging*) [3]. El DTI calcula el tensor de difusión aparente bajo la suposición de que p es una distribución

Gaussiana trivariada de media cero. Si definimos a

$$p_G(r, D, t) = ((4\pi t)^3 \det(D))^{-1/2} \exp -\frac{r^T D^{-1} r}{4t}, \quad (3.14)$$

entonces

$$p(r) = p_G(r, D, t). \quad (3.15)$$

La substitución de la Ecuación (3.15) en la Ecuación (3.10) da

$$E(q) = e^{-tq^T D q}, \quad (3.16)$$

donde el tensor de difusión D es una matriz simétrica de 3×3 que caracteriza los desplazamientos en 3D

$$\begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix}. \quad (3.17)$$

Los elementos diagonales de la matriz corresponden a la difusividad sobre los tres ejes principales, mientras que los elementos que no están en la diagonal corresponden a la correlación entre desplazamientos entre estos ejes. Teniendo en cuenta la Ecuación (3.2), la matriz del tensor de difusión se puede ver como una matriz de covarianza tridimensional de los desplazamientos sobre cierto intervalo de tiempo.

El tensor de difusión usualmente suele verse en términos de una superficie elipsoide que representa la distancia que una molécula puede desplazarse con la misma probabilidad desde el origen. Los formalismos del tensor de difusión dan un marco de referencia interno, conocido en la literatura como el *eigensystem* [28]. Los principales ejes de este elipsoide están dados por los eigenvectores del tensor, mientras que los tamaños son dados por la distancia de difusión en un intervalo de tiempo. Si el tensor de difusión es sólo la diagonal, estos valores corresponden a sus eigenvalores (denotados como $\lambda_1, \lambda_2, \lambda_3$) y estos corresponden a la difusión sobre cada eje principal del tensor. En la Figura 3.3 se muestran algunos ejemplos de tensores teóricos y su correspondiente elipsoide.

El tensor de difusión es muy utilizado en DW-MRI, dado que para cada voxel en el estudio podemos obtener un tensor que denota la anisotropía de la difusión en el tejido. Sus principales ventajas son su fácil implementación e interpretación, sin embargo, el modelo Gaussiano que asume el DTI no siempre será suficiente para explicar micro estructuras más complejas en el tejido, como un cruce de fibras. Por lo tanto, el DTI tiene muchas limitaciones.

3.3.1. Modelo de tensor multicompartimiento

Una generalización de DTI reemplaza el modelo Gaussiano de p en la Ecuación (3.15) con una mezcla de Gaussianas [23]

$$p(r) = \sum_{i=1}^n a_i p_G(r, D_i, t), \quad (3.18)$$

3. PRINCIPIOS Y TÉCNICAS DE IMAGENOLÓGÍA DE RESONANCIA POR DIFUSIÓN

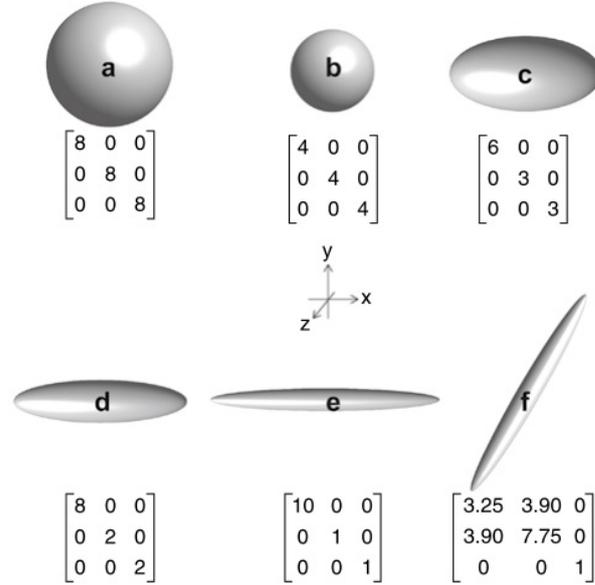


Figura 3.3: Matrices de tensores de difusión y sus elipsoides correspondientes.

Los tensores (a) y (b) son ambos isotrópicos (difusión libre), donde cada uno tiene sus tres valores en la matriz iguales. Sin embargo (a) duplica el tamaño de (b). Los tensores (c), (d) y (e) están alineados con respecto al eje x , por lo tanto sus valores fuera de la diagonal son 0. El tensor en (f) ha sido rotado sobre el eje z en un ángulo de 60° , y esto se refleja en que valores relacionados con x, y fuera de la diagonal sufren cambios. Figura tomada de [28].

donde cada $a_i \in [0, 1]$ y $\sum_{i=1}^n a_i = 1$. Los desplazamientos de partículas en medios que contienen n compartimientos distintos, entre los cuales no ocurre intercambio de partículas, tienen la distribución de la Ecuación (3.18) si la función de densidad de desplazamientos del i -ésimo compartimiento, el cual tiene una fracción de volumen a_i , es la Ecuación (3.15).

Al tomar la transformada de Fourier de la Ecuación (3.18) y al substituir en la Ecuación (3.10) obtenemos

$$E(q) = \sum_{i=1}^n a_i e^{-tq^T D_i q}. \quad (3.19)$$

Este modelo tiene la ventaja de poder obtener estimaciones de p que no necesariamente vienen de una fibra. Sin embargo, tiene varias complicaciones, como por ejemplo que se requieren de algoritmos de optimización no lineal para poder optimizar los parámetros de este modelo.

En esta tesis nos interesa la definición del tensor de difusión para poder simular señales de resonancia con la Ecuación (3.16) y la Ecuación (3.19). En las referencias [28] [23] [29] se encuentra más información si el lector desea conocer más a fondo los detalles sobre el DTI, tales como el cálculo de dicho tensor y sus propiedades.

3.4. Imagenología del espacio Q

En la Ecuación (3.8) se mencionó que existe una relación de Fourier entre la señal de MRI y la densidad de probabilidad $p(r)$. Formalmente la imagenología del espacio Q [1] procura aprovechar esta característica para obtener la probabilidad $p(r)$. En la imagenología del espacio Q, la Ecuación (3.8) generalmente se escribe de la siguiente forma

$$E(q) = \int P(r) e^{-i2\pi q^T r} dr, \quad (3.20)$$

donde ahora

$$q = \frac{\gamma \delta G}{2\pi}, \quad (3.21)$$

y con un Δ que será un tiempo de difusión fijo para todo el experimento. Nota: En la Ecuación (3.20) el vector q no se encuentra en términos de radianes, esto hace que la ecuación ahora tenga un 2π no presente en la Ecuación (3.8). Para mayores detalles referirse al Apéndice B.

La relación fundamental entre la señal $E(q)$ y el propagador de difusión $p(r)$ está dada por la siguiente transformada de Fourier [1]

$$p(r) = \int E(q) e^{-i2\pi q^T r} dq, \quad (3.22)$$

donde el vector de espacio de Fourier (espacio Q) está en coordenadas cartesianas mediante $q = (q_x, q_y, q_z)$, y el vector en el espacio real está dado por $r = (r_x, r_y, r_z)$.

Es importante volver a mencionar que la relación en la Ecuación (3.20) está basada en la suposición de que el tiempo del pulso de gradiente δ es muy pequeño, tal que las moléculas no hagan de manera significativa difusión durante su aplicación. Este criterio es usualmente conocido como la *aproximación de pulso pequeño* [30]. Sin embargo, bajo condiciones ideales y un tiempo de difusión Δ fijo, la señal $E(q)$ depende directa y solamente del vector q , el cual podemos controlar variando la fuerza y orientación de los campos magnéticos de los gradientes G .

3.5. Imágenes de espectro por difusión

En el DSI [4] (por sus siglas en inglés: *Diffusion Spectrum Imaging*) se propone un esquema que procura medir directamente a p sin la necesidad de un modelo previo o

3. PRINCIPIOS Y TÉCNICAS DE IMAGENOLOGÍA DE RESONANCIA POR DIFUSIÓN

suposiciones de p (cosa que se hacia en DTI por ejemplo). Esta técnica se basa en el modelo del espacio Q de la sección pasada. En la práctica, para excluir contaminación de la fase proveniente de las inhomogeneidades y movimiento biológico, el propagador de difusión es reconstruido tomando la transformada de Fourier del módulo de la señal compleja de MRI

$$p(r) = \int |E(q)| e^{-i2\pi q^T r} dq, \quad (3.23)$$

como resultado, los propagadores de difusión p recuperados mediante la Ecuación (3.23) son simétricos ($p(r) = p(-r)$), debido a que la información asimétrica en la señal de MRI se encuentra contenida en la parte imaginaria [30], además de que la transformada de Fourier de una señal simétrica y real también es una señal simétrica y real.

DSI adquiere una medida para cada q en una retícula rectilínea tridimensional de vectores de onda q . De esta configuración de muestras, podemos utilizar la transformada de Fourier rápida (FFT) para invertir la transformada de Fourier y obtener una retícula similar de muestras del EAP p . Por lo tanto, DSI muestrea la señal E en una retícula de q y la FFT obtiene muestras de p en una retícula de desplazamientos r [31]. Los primeros dos paneles de la Figura 3.4 ilustran en un ejemplo simple este procedimiento.

Note que el DSI ignora los efectos del *aproximación de pulso pequeño* discutidos en la sección pasada. Aunque en la práctica (como ya se ha mencionado) no es posible realizar mediciones con pulsos g de duración δ infinitamente pequeños, una aproximación del muestreo de la retícula de E como una función de q , es obtenida mediante mediciones con distintos valores b (Ecuación (3.13)), hasta un máximo valor b establecido, y distintas orientaciones del gradiente.

3.5.1. Función de distribución-orientación

En la práctica, no sólo estamos interesados en tener muestras discretas de p , también estamos interesados en la estructura angular de la forma de p . El siguiente paso en el DSI es simplificar más aún los datos, tomando la suma radial pesada de p , mediante la *Función de distribución-orientación* (ODF, por sus siglas en inglés: *Orientation Distribution Function*) [4], definida para el caso continuo como:

$$ODF(u) = \int p(\rho u) \rho^2 d\rho, \quad (3.24)$$

donde $\|u\| = 1$ y $\rho = \|r\|$. La ODF mide la cantidad de difusión en la dirección del vector unidad u . Para el caso discreto, se resuelve la integral como una sumatoria discreta sobre un rango discreto para ρ . La Figura 3.4 es un ejemplo de la obtención de la ODF para un caso sencillo con DSI en 2D.

La ODF tiene picos en las direcciones donde el EAP p tiene mayor masa. Por lo tanto, en DSI los picos de la ODF muestran estimaciones de las orientaciones de fibras. La ODF puede tener múltiples picos, lo cual permite al DSI obtener orientaciones de varias fibras que cruzan en un mismo vóxel [23].

DSI tiene múltiples ventajas sobre DTI y técnicas de multicompartimiento, dado que puede resolver múltiples orientaciones de fibras, sin asumir ningún modelo o requerir resolver problemas de optimización no lineales. Sin embargo, en la práctica no es muy usado. Esto se debe a que esta técnica requiere una gran cantidad de muestras para obtener suficientes detalles en la reconstrucción del EAP p , lo cual, resulta en tiempos de adquisición muy grandes. El protocolo más común de DSI (propuesto por Wedeen [4]) involucra tomar 515 muestras en el espacio Q , en comparación a sólo los 7 requeridos en el esquema más simple en DTI.

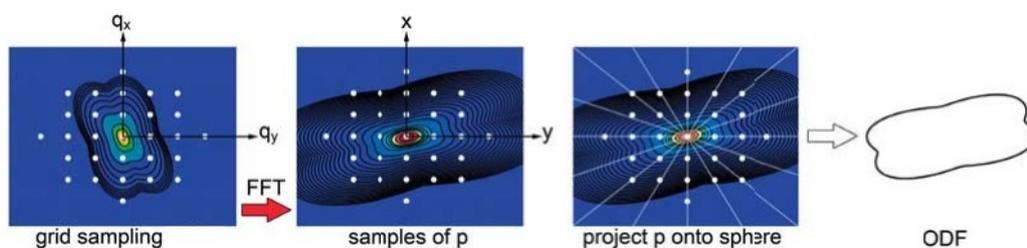


Figura 3.4: Estimación de la ODF en DSI. En este ejemplo en 2D observamos el procedimiento de DSI y estimación de la ODF. El primer panel muestra las mediciones realizadas en función de q . Los puntos blancos muestran los valores de q donde fue muestreada $E(q)$, estos forman una retícula. El segundo panel muestra a p , ese siendo la transformada de Fourier del espacio q , donde los puntos blancos representan los vectores de desplazamiento, en la misma retícula anterior. Para obtener la ODF en el cuarto panel, interpolamos la retícula de muestras de p y lo integramos sobre los ejes radiales desde el origen, como lo describe el tercer panel. Figura tomada de [31]

3.6. Ruido en DW-MRI

Aunque las señales de DW-MRI son complejas, es común utilizar la magnitud de estas señales en la práctica para evitar artefactos relacionados con la fase, o por requisito de ciertas técnicas como en DSI. Generalmente se asume que el ruido en la parte real y compleja de la señal de DW-MRI es Gaussiano con media cero. Sin embargo, calcular la magnitud de una señal compleja implica un mapeo no lineal, por lo tanto, la distribución de su ruido no es Gaussiano. El ruido obtenido en la magnitud de la señal está descrito mediante una distribución Rician [32].

Si la señal sin ruido está denotada en I y la señal medida está denotada en M , la

3. PRINCIPIOS Y TÉCNICAS DE IMAGENOLOGÍA DE RESONANCIA POR DIFUSIÓN

distribución riciana de M se encuentra definida por [33]

$$R(M) = \frac{M}{\sigma^2} e^{-(M^2+I^2)/2\sigma^2} B_0\left(\frac{IM}{\sigma^2}\right), \quad (3.25)$$

donde B_0 es la función Bessel de orden cero modificada del primer tipo, σ denota la desviación estándar de la parte real e imaginaria de la imagen. En altos niveles de la relación señal-ruido (SNR por sus siglas en inglés: *Signal to Noise-Ratio*, definida como la relación existente entre la potencia de la señal y el ruido que contiene), el ruido Riciano tiende a una distribución Gaussiana, mientras que a bajo niveles es aproximadamente de una distribución Rayleigh (de hecho, si no hay señal presente ambas distribuciones serían iguales).

Una forma de añadir ruido Riciano a una señal de DW-MRI, es añadir ruido Gaussiano a la parte real y la parte imaginaria de la señal, y posteriormente obtener la magnitud. Asumiendo que sólo tenemos la parte real (o la magnitud) E , y la parte imaginaria sea 0, añadir ruido Riciano sería:

$$\hat{E} = \sqrt{(E + X)^2 + Y^2}, \quad (3.26)$$

donde X y Y son variables independientes de una normal con cierta varianza σ^2

$$X, Y \sim N(0, \sigma^2), \quad (3.27)$$

donde la desviación estándar de la señal está en función de la señal sin difusión S_0 y el SNR de la misma

$$\sigma = \frac{S_0}{SNR}. \quad (3.28)$$

Metodología

En la Sección 1.1 se mencionó que uno de los principales objetivos de esta tesis es presentar un esquema de trabajo para realizar sensado comprimido en DSI, de tal forma que tomando las menos muestras posibles (y manteniendo la calidad de reconstrucción) reduzcamos de manera efectiva el tiempo de escaneo, siendo este el principal problema del DSI.

En este capítulo se presenta la metodología propuesta para dicho esquema de trabajo. Se muestra el esquema para hacer uso de la técnica de sensado comprimido en DSI. Se detalla bajo qué criterios y parámetros son simuladas las señales de DW-MRI sintéticas utilizadas en esta tesis. Se explica el procedimiento para calcular la matriz de sensado comprimido A , y se presenta una propuesta de metodología para optimizar algunos de los parámetros requeridos en resolver dicho problema de sensado comprimido.

4.1. Diseño de experimento

La teoría del sensado comprimido requiere resolver el siguiente problema de optimización convexo LASSO :

$$\min_x \|A_\Omega x - y_\Omega\|_{\ell_2} + \lambda \|x\|_{\ell_1}, \quad (4.1)$$

donde:

- Ω : Conjunto de índices de cardinalidad m , con los cuales se muestrea a la señal.
- y_Ω : Vector de tamaño m , con las muestras de la señal y (de tamaño n).
- A : Es la matriz de sensado comprimido (Sección 4.3). Esta se compone de dos bases ($A = \Phi\Psi$):
 - Φ : Matriz de sensado de tamaño $n \times n$.
 - Ψ : Matriz de representación rala de tamaño $n \times l$.

4. METODOLOGÍA

La matriz A_Ω de tamaño $m \times l$, denota que sólo se mantienen las filas de A indicadas en el conjunto Ω .

- x : Vector de tamaño l , el cual es una representación rala de la señal a recuperar.
- λ : Parámetro de regularización, relacionado a la tolerancia del ruido y penalización de la representación rala de la señal.

En este capítulo se definen todas las variables y parámetros necesarios para el planteamiento y resolución de este problema, aplicándolo a un esquema de muestreo a utilizar en el protocolo de DSI. Si definimos a:

- c : Un cubo discretizado en n posiciones. Donde cada posición corresponde a distintos valores q en un espacio Q 3D discretizado.
- S^c : Es la señal DW-MRI a muestrear, la cual se encuentra distribuida en el cuboide c . Cada elemento en S^c será denotada mediante $s(i)$. El Sección 4.2 detalla como simularlas.
- Ω : Conjunto de m índices del cuboide c con los cuales se muestrea a la señales S^c . Estrategias de muestreo se discuten en la Sección 4.4.
- S_Ω : Vector de tamaño m de las señales muestradas mediante las coordenadas Ω en S^c .
- M_Ω : Matriz auxiliar de muestreo de tamaño $m \times n$. Corresponde a los m renglones de una matriz identidad $n \times n$ correspondientes a las coordenadas contenidas en Ω .
- F : Matriz de sensado de tamaño $n \times n$. Es una matriz de bases de Fourier, el cual es el esquema utilizado en DW-MRI (Sección 4.3.1).
- P : Matriz de representación de tamaño $n \times l$. Esta matriz es un diccionario basado en distribuciones de probabilidad de difusión (Sección 4.3.2).
- x : Un vector de tamaño l el cual corresponde a los coeficientes correspondientes al diccionario P . Este vector debe ser lo mas ralo posible.

Entonces si $\Phi = M_\Omega F$ y $\Psi = P$, podemos reescribir la Ecuación (4.1) como:

$$\min_x \|M_\Omega F P x - S_\Omega\|_{\ell_2} + \lambda \|x\|_{\ell_1}, \quad (4.2)$$

donde buscamos recuperar el EAP p en c , el cual obtenemos con $p = P x$ y acomodando en c . En esta tesis buscamos evaluar la calidad del p obtenido haciendo variaciones en los parámetros de Ω y λ para diferentes tamaños de muestreo m . Estrategias para la optimización de Ω se encuentran en la Secciones 4.4 y 4.5.2, mientras que estrategias para optimizar λ se encuentran en la Sección 4.5.3. En esta tesis vamos a hacer uso del concepto de la matriz de sensado comprimido A , definida como $A_\Omega = M_\Omega F P$.

En esta tesis se resuelve el problema en la Ecuación (4.2) mediante un solver contenido en el lenguaje de cómputo matemático Matlab, que usa el algoritmo ADMM (Apéndice A), siendo éste (de un conjunto de pruebas realizadas) el más eficiente con respecto a otros solvers que se probaron [34] [35]. El Algoritmo 1 resume el procedimiento para resolver este problema y evaluar la calidad de la reconstrucción, donde definimos a la señal obtenida al realizar el DSI completo en S^c como el EAP p_{gt} ground-truth, la cual nos va a servir para evaluar la calidad de la reconstrucción obtenida.

Algorithm 1 Diseño y prueba de un experimento

Require: señal S^c , λ , Ω , matriz A , diccionario P , SNR

- 1: Obtenemos p_{gt} ground-truth de S^c como sigue: $p_{gt} = FFT(S^c)$.
 - 2: Obtenemos la señal ruidosa \hat{S}^c agregando ruido Riciano a S^c con el SNR deseado.
 - 3: Formamos vector de muestras S_Ω a partir de la señal \hat{S}^c como sigue: Vectorizamos \hat{S}^c para posteriormente $S_\Omega = M_\Omega \hat{S}^c$.
 - 4: Tomamos los renglones correspondientes en A como sigue: $A_\Omega = M_\Omega A$.
 - 5: Optimizamos x en la Ecuación (4.1).
 - 6: Obtenemos vector p solución a partir de x como sigue: $p = Px$.
 - 7: Pasamos p a espacio 3D en cubo c .
 - 8: Obtenemos error entre el p_{gt} y p solución.
 - 9: **return** p
-

Notas:

- El ruido Riciano lo agregamos mediante el método descrito en la Sección 3.6.
- El cubo c es de un tamaño de $16 \times 16 \times 16$. Éste se encuentra en una potencia de 2 para poder abarcar los 515 datos requeridos en DSI y poder usar la transformada rápida de Fourier (FFT).
- Las métricas de error utilizadas se describen en la Sección 4.1.1.

4.1.1. Métricas de error

Para comparar la calidad del EAP reconstruido de la metodología contra el EAP ground-truth, se propone utilizar un conjunto de métricas adecuadas para comparar distribuciones:

- D_{euc} Distancia Euclideana.
- EMD Earths Mover's Distance.

- D_{kl} Divergencia Kullback-Leibler.
- D_{js} Divergencia Jensen-Shannon.

La definición de estas métricas se encuentran en el Apéndice C. Para medir la sensibilidad de estas métricas se propone un simple experimento, comparando múltiples señales vs. esas mismas señales con ruido Riciano, a diferentes niveles de SNR. Lo anterior nos permite explorar la sensibilidad de las métricas.

4.2. Generación sintética de la señal S

Las señales sintéticas $E(q)$ de DW-MRI de esta tesis cumplen con el protocolo propuesto por Wedeen para DSI [4]. Este consiste en tomar $N = 515$ valores codificados mediante su valor q , teniendo en cuenta dentro del espacio Q los puntos en un cubo reticulado dentro de una esfera de 5 unidades reticuladas, el cual llamaremos el grid. Representamos q como:

$$q = aq_x + bq_y + cq_z, \quad (4.3)$$

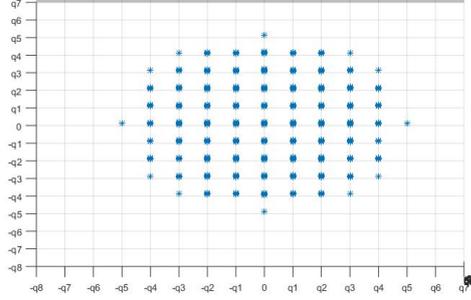
con a, b, c siendo números enteros y $\sqrt{a^2 + b^2 + c^2} \leq 5$. Los valores q_x, q_y, q_z contenidos en el vector q denotan la modulación de fase con respecto a la dirección coordenada. La Figura 4.1 muestra un esquema visual de dicho grid. Los parámetros del protocolo DSI se encuentran en la Tabla 4.1.

Parámetro/Variable	Valor
g_{max}	$40mT/m$
δ	$60ms$
Δ	$66ms$
N	515
b_{max}	$17000s/mm^2$

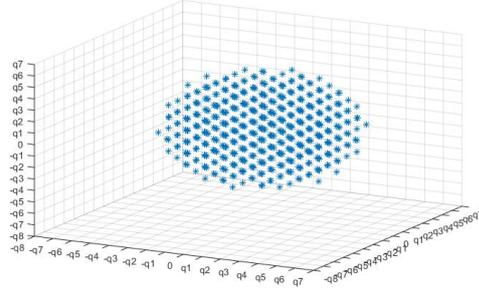
Tabla 4.1: Parámetros del protocolo DSI - Estos son los valores de parámetros utilizados en el diseño y las simulaciones, propuesto para DSI [4].

Estas señales son simuladas mediante un modelo de tensor (Sección 3.3), específicamente mediante la Ecuación:

$$E(q_i) = \frac{s(q_i)}{s_0} = e^{-(\Delta - \delta/3)q_i^T D q_i}, \quad (4.4)$$



(a) El grid en un plano 2D



(b) El grid en 3D

Figura 4.1: El grid en el protocolo de DSI. El panel (a) muestra un plano en 2D del grid del protocolo, el valor máximo del grid se alcanza en q_5 . El panel (b) muestra el grid en 3d, en este se observan los puntos del cubo reticulado dentro de la esfera de 5 unidades

donde s_0 es la imagen sin difusión. En este trabajo suponemos todas nuestras señales han sido normalizadas, lo cual implica que $s_0 = 1$, por lo tanto $E(q) = s(q)$.

En el DSI el valor q irá modulando la dirección y tamaño de paso para cada muestra del protocolo. Esto lo hace modulando mediante la dirección y potencia de gradiente, pues supone un mismo tiempo efectivo de difusión para todo el estudio. Pasa visualizar esto podemos expandir la definición del gradiente g en q como

$$q = \frac{\gamma \delta g G}{2\pi}, \quad (4.5)$$

donde g es la dirección unitaria del gradiente,

$$g = \frac{[a, b, c]}{\|[a, b, c]\|_{\ell_2}}, \quad (4.6)$$

y G es la magnitud del gradiente, que como dicta el protocolo llegará hasta un g_{max} en los puntos esquinados de la retícula.

4. METODOLOGÍA

Por requerimientos de la transformada rápida de Fourier (proceso requerido para obtener el EAP en DSI), este grid de 515 puntos se centra dentro de un cuboide reticulado, c , de tamaño $16 \times 16 \times 16$, centrando el grid en la coordenada $[9, 9, 9]$. La Figura 4.1b es un ejemplo de este procedimiento. Esto resulta en un conjunto de $n = 4096$ valores. Sin embargo, todos aquellas valores fuera del grid tienen valor 0 como lo indica la siguiente regla:

$$\begin{cases} S(q_i), & \sqrt{a^2 + b^2 + c^2} \leq 5 \\ 0, & \text{En otro caso.} \end{cases} \quad (4.7)$$

Los valores de los tensores de difusión D se encuentran en un rango de valores anatómicamente plausibles para fibras comunes en tejido cerebral. En nuestras simulaciones siempre suponemos que existe una dirección de fibra dominante, aunque el modelo de tensores engloba la difusión extra e intraxonal. La forma de los tensores propuestos es

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_{23} & 0 \\ 0 & 0 & \lambda_{23} \end{bmatrix} \quad (4.8)$$

donde

$$\lambda_1 \in \Lambda_1, \quad (4.9)$$

$$\lambda_{23} \in \Lambda_{23}, \quad (4.10)$$

$$\Lambda_1 = [1.5 \times 10^{-9}, 1.9 \times 10^{-9}], \quad (4.11)$$

$$\Lambda_{23} = [.1 \times 10^{-9}, .5 \times 10^{-9}]. \quad (4.12)$$

Para obtener una imagen en una dirección específica requerida, solo es necesario rotar el tensor canónico anterior en la dirección requerida, mediante una matriz de rotación. El Algoritmo 2 describe una función para obtener una señal S con la metodología descrita.

Algorithm 2 Función $\text{getS}(D, Q, t)$

Require: t, D , Colección Q

- 1: **for** i de 1 hasta $\text{lenght}(Q)$ (515 puntos en DSI) **do**
 - 2: $q_i = Q(i)$.
 - 3: $s(q_i) = s_0 e^{-t q_i^T D q_i}$.
 - 4: Colocamos $s(q_i)$ en el grid en su coordenada correspondiente.
 - 5: Centramos el grid en cuboide c de tamaño $16 \times 16 \times 16$ para formar S^c .
 - 6: **return** S^c
-

4.2.1. Señales S con cruces de fibras

Para simular señales S con cruces de fibras, utilizamos un modelo multicompartimiento (Sección 3.3.1). Todos los experimentos en este trabajo se realizan suponiendo un cruce de dos fibras con cierto ángulo \angle de separación, tal que la Ecuación (3.19) sea

$$s(q_i) = a_1 s_1(q_i) + a_2 s_2(q_i) = a_1 e^{-tq_i^T D_1 q_i} + a_2 e^{-tq_i^T D_2 q_i}, \quad (4.13)$$

donde $a_1 \in [.4, .6]$ y $a_2 = 1 - a_1$. Entre $s_1(q_i)$ y $s_2(q_i)$ existe un ángulo \angle de separación en el rango de $\angle \in [60^\circ, 90^\circ]$. Esto lo logramos rotando D_2 en un eje con respecto a D_1 con el ángulo requerido.

En esta tesis se harán una serie de experimentos, donde para simular distintos tamaños y direcciones de posibles microestructuras en vóxeles, simularemos de manera aleatoria señales con los parámetros y métodos anteriormente descritos. El Algoritmo 3 resume la construcción de una señal aleatoria.

Algorithm 3 Función $\text{randomS}(D_i, Q, t)$

Require: t , Colección Q , Λ_1 , Λ_{23} , ang1 , ang2

- 1: Generar D_1 con $\lambda_1 \sim \mathcal{U}(\Lambda_1)$, $\lambda_{23} \sim \mathcal{U}(\Lambda_{23})$
 - 2: Generar D_2 con $\lambda_1 \sim \mathcal{U}(\Lambda_1)$, $\lambda_{23} \sim \mathcal{U}(\Lambda_{23})$
 - 3: Rotar D_1 en dirección aleatoria r en la esfera
 - 4: Rotar D_2 en dirección r
 - 5: Rotar D_2 en sobre un eje con ángulo $\angle \sim \mathcal{U}(\text{ang1}, \text{ang2})$
 - 6: $S_1 = \text{getS}(D_1, Q, t)$
 - 7: $S_2 = \text{getS}(D_2, Q, t)$
 - 8: $\alpha_1 \sim \mathcal{U}(.4, .6)$, $\alpha_2 = 1 - \alpha_1$
 - 9: $S = \alpha_1 S_1 + \alpha_2 S_2$
 - 10: **return** S
-

4.2.2. Vectorizando el espacio 3D

En la Ecuación (4.2) la señal muestreada $S_\Omega \in \mathbb{R}^m$, que contiene muestras de S^c , tiene forma vectorial. Nuestras señales S^c se encuentran en un cubo c 3D de $16 \times 16 \times 16$, por lo tanto requerimos pasar estos espacios a una forma vectorial. Para ello, simplemente se vectoriza el cuboide ordenando por columnas, sólo hay que tener en cuenta la ubicación donde se encuentran las muestras del grid.

4.3. Diseño de la matriz de sensado comprimido A

La matriz de sensado comprimido A de tamaño $n \times l$ se compone de una combinación de dos bases tal que $A = \Phi\Psi$ (Sección 2.2), donde:

- Φ : Es la matriz de sensado de tamaño $n \times n$. En este caso se trata de una matriz de bases de Fourier F , debido a que este es el esquema utilizado en imágenes DW-MRI, por lo tanto $\Phi = F$.
- Ψ : Es una matriz de representación rala de la señal a recuperar de tamaño $n \times l$. En este caso se trata de un diccionario P basado en distribuciones de difusión de hidrógeno. Por lo tanto, en este trabajo $\Psi = P$. Es importante destacar que en este estudio el EAP p se obtiene al realizar la conversión $p = Px$.

En esta sección explicamos ciertos detalles en la construcción de ambas matrices.

4.3.1. Construcción de matriz de sensado de Fourier

En la Sección 2.3.2 se hace mención de un par de estrategias para la construcción de la matriz Φ cumpliendo los requisitos del sensado comprimido (Sección 2.2). Sin embargo, también es importante modelar esta matriz en función del problema a tratar.

En el caso de DW-MRI, la señal $E(q)$ esta en función de la transformada de Fourier que se realiza sobre el vóxel (Sección 3.2). Por lo tanto, en este esquema nuestra matriz Φ se compone de la base de Fourier correspondientes a esta transformación [36]. En esta tesis utilizamos las funciones contenidas dentro del paquete computacional Matlab para el cálculo de las transformaciones rápidas de Fourier.

Dado que la señal S en la Ecuación (4.2) es un vector de un espacio 3D, es necesario armar una matriz de transformación de Fourier correspondiente a una transformación 3D de un espacio 3D vectorizado. Para ello optamos por la estrategia en el Algoritmo 4, donde obtenemos para cada columna de la matriz de sensado Ψ , una base de Fourier en forma vectorial correspondiente a la transformación 3D de una imagen 3D con solo un coeficiente encendido en la posición deseada. La matriz Φ resultante será denotada como matriz F , y tiene un tamaño de 4096×4096 para el caso de nuestro cubo c con $n = 4096$ posiciones.

4.3.2. Construcción de diccionario de representación basado en distribuciones de difusión

Como se vio en la Sección 2.2, para que el sensado comprimido sea efectivo, se requiere que la señal a recuperar f sea lo más rala posible bajo una base de representación Ψ , tal que

$$f = \Psi x, \tag{4.14}$$

Algorithm 4 Matriz de transformación Fourier 3D**Require:** Dimensión de tamaño r en cada arista de cubo c

- 1: **for** i de 1 hasta r **do**
- 2: **for** j de 1 hasta r **do**
- 3: **for** k de 1 hasta r **do**
- 4: Construimos un cubo I donde $I_{ijk} = 1$ y el resto es 0.
- 5: Obtenemos If : la transformada de Fourier 3D de I .
- 6: Vectorizamos If en vf .
- 7: El vector vf corresponde a la columna $(i - 1) + (j - 1) + k$ de la matriz de transformación 3D final de Fourier F
- 8: **return** F

donde x es una señal rala. En la Sección 1.3 se mencionan algunas estrategias en la literatura para el diseño de esta matriz para DW-MRI, sin embargo, estas estrategias desaprovechan el hecho de que bajo una base adecuada es posible extraer información adicional sobre el EAP de la señal rala x , pues sólo se enfocan en mantener la rareza.

En este trabajo la matriz de representación rala Ψ es construida bajo la estrategia de que sea un diccionario P , tal que $\Psi = P$, donde cada átomo de este diccionario corresponde a una distribución de probabilidad de difusión de hidrógeno, un posible EAP. De esta forma, la solución contenida en la señal x se formaría bajo una combinación lineal de cada uno de estos átomos, los cuales fueron construidos suponiendo un modelo de microestructura interna en cada átomo. Por lo tanto, de los coeficientes “encendidos” en la señal x (aquellos que tengan un valor mayor a un treshold propuesto) podemos intuir tamaños, formas y ángulos que forman el EAP reconstruido (Figura 4.2).

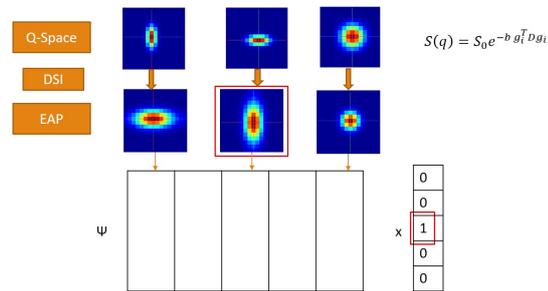


Figura 4.2: Diccionario P . El diccionario P se compone de EAPs provenientes de distintos posibles tamaños y direcciones de fibras, y x contendría los coeficientes de la combinación lineal entre ellos que componen el EAP solución.

4. METODOLOGÍA

Note que cuando elegimos la matriz de Fourier como matriz de sensado, estamos haciendo una conversión del espacio real al espacio q

$$S = Fp, \quad (4.15)$$

donde p se encuentra en el espacio real. Por lo tanto el EAP resultado p , lo podemos obtener al realizar la operación

$$p = Px. \quad (4.16)$$

El EAP resultante de dicha operación estaría descrito por

$$p = \sum_{i=1}^n \alpha_i \text{Atom}(D_i, Q, t), \quad (4.17)$$

donde

$$\alpha_i = \begin{cases} x_i, & x_i > .01 \\ 0, & x_i < .01 \end{cases}. \quad (4.18)$$

En este caso no se garantiza que $\sum_i \alpha_i = 1$, aunque podemos hacer una normalización temporal en x para tener una interpretación mas clara de los porcentajes de los átomos en el diccionario P que componen al EAP resultante .

La función $\text{Atom}(D, Q, t)$ (Algoritmo 5) realiza la propuesta de EAP para cada átomo del diccionario. Esta función realiza tal procedimiento simulando un proceso de DSI, donde ya teniendo un espacio Q armado como en la Sección 4.2, realiza la transformada rápida de Fourier y posteriormente vectoriza este espacio para que sea un átomo del diccionario. Este espacio Q lo simula mediante el procedimiento descrito en la Sección 4.2. Esta función recibe como parámetros el tiempo efectivo de difusión $t = (\Delta - \delta/3)$ (Tabla 4.1), la colección Q que se compone de los vectores q donde se realizan las supuestas muestras para DSI (Sección 4.2), y un tensor de difusión D_j , elemento de la colección $D_{coleccion}$, en la que cada de uno de los elementos D_j en ella es único en sus valores y direcciones. Un resumen de este procedimiento se encuentra en el Algoritmo 6.

Para los eigenvalores en los D_j suponemos solo hay una dirección de difusión anisotrópica dominante $\lambda_{23} = \lambda_2 = \lambda_3$ y $\lambda_1 > \lambda_{23}$. Estas variables toman valores en rangos de tamaños anatómicamente plausibles en tejido cerebral, discretizando en los puntos

$$\lambda_1 \in [1.5, 1.6, 1.7, 1.8, 1.9] \times 10^{-9}, \quad (4.19)$$

$$\lambda_{23} \in [.1, .2, .3, .4, .5] \times 10^{-9}, \quad (4.20)$$

para un total de 25 posibles combinaciones de tamaños. Cada uno de estos tamaños es proyectado en 256 posibles diferentes direcciones equidispaciadas en la media esfera (el EAP es simétrico, sólo basta discretizar media esfera para capturar bien esas direcciones), la Figura 4.3 muestra las direcciones tomadas en la media esfera. Por lo tanto, en

Algorithm 5 Función $\text{Atom}(D, Q, t)$

Require: t, D , Colección Q

- 1: **for** i de 1 hasta $\text{length}(Q)$ (515 puntos en DSI) **do**
 - 2: $s(q_i) = s_0 e^{-tq_i^T D q_i}$.
 - 3: Colocamos $s(q_i)$ en el grid en su coordenada correspondiente.
 - 4: Centramos el grid en cubo c de tamaño $16 \times 16 \times 16$.
 - 5: $p3D = \text{FFT}$ de cuboide.
 - 6: **return** $p3D$
-

Algorithm 6 Construcción de Diccionario P

Require: t , colecciones Q y $D_{\text{coleccion}}$

- 1: **for** j de 1 hasta $\text{length}(D_{\text{coleccion}})$ **do**
 - 2: $D_j = D_{\text{coleccion}}(j)$
 - 3: $p3D = \text{Atom}(D_j, Q, t)$.
 - 4: Vectorizamos $p3D$ en p_j .
 - 5: El vector p_j corresponde a la columna j del Diccionario P .
 - 6: **return** P
-

4. METODOLOGÍA

este caso, el diccionario P se compone de un total de $l = 6400$ átomos, que se componen de la combinaciones de tamaños y direcciones de los D_j contenidos en $D_{coleccion}$ recién mencionados. Note que el diccionario P viene de modelos de difusión sin ruido, por lo tanto el EAP reconstruido bajo esta metodología será una solución sin ruido, lo cual implica que estaremos filtrando ruido de la señal muestreada en la reconstrucción final.

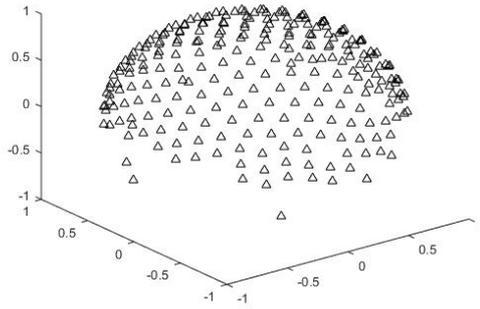


Figura 4.3: Direcciones en la media esfera del diccionario P . Esta figura muestra las 256 direcciones en la media esfera utilizadas para rotar los tensores que conforman el diccionario P . Debido a la simetría del EAP solo requerimos media esfera.

4.4. Muestreo Ω de la señal $S(q)$

En el CS buscamos tomar el menor número de m muestras posibles de una señal de tamaño n , tal que si $m \ll n$, aún podamos reconstruir dicha señal con una calidad práctica suficiente. En este trabajo, dichas muestras corresponden a un subconjunto de mediciones del espacio Q discretizado en el cubo c (de tamaño de $16 \times 16 \times 16$), tomando solamente de aquellas 515 posiciones en el grid (Figura 4.1b). Los índices correspondientes a estas muestras se encuentran contenidas en el conjunto Ω . Dependiendo de la naturaleza de las bases, un muestreo aleatorio uniforme puede ayudar a mantener la incoherencia (Sección 2.2). Sin embargo, al tratarse de un problema más específico (imágenes de DW-MRI), podemos tomar en cuenta algunas consideraciones especiales con el muestreo que van a ayudar a optimizar el resultado:

- Menzel [5] propone distintas estrategias de muestreo para el espacio- Q . De sus resultados, es seguro decir que una estrategia adecuada para el muestreo del espacio- Q es con un muestreo Gaussiano. Este resultado es lógico, ya que el muestreo Gaussiano va a favorecer el centro del espacio Q , donde sabemos *a-priori* que son justamente aquellos valores q que contienen más información sobre la forma general del EAP (de manera análoga, piense en los valores q del centro del espacio

Q como aquellas que caracterizan la menor frecuencia de la señal, por lo tanto son las que conservan más información, mientras que aquellos valores q elevados conservan más información de los detalles de pequeña magnitud de la señal). El muestreo Gaussiano se realiza mediante el Algoritmo 7, con una varianza que asegure abarque el grid.

- En todos los casos de Ω se toma la muestra correspondiente a la medición s_0 , siendo s_0 la medición en el centro del espacio Q ($s_0 = s(q_0)$).
- El espacio Q es por definición simétrico (ver Sección 3.5), por lo tanto, resulta conveniente inyectar la información de esta simetría en la reconstrucción. Esto lo logramos tomando a cada muestra con su par simétrico. Esta información al ser redundante nos permite tomar virtualmente aún menos muestras, al solo repetir la muestra con su punto simétrico en cubo c .

Algorithm 7 Muestreo Ω Gaussiano

Require: m, var

- 1: Se inicia Ω como un conjunto vacío.
 - 2: **while** El número de elementos en Ω sea menor a m **do**
 - 3: $\hat{a} \sim \mathcal{N}(0, var), \hat{b} \sim \mathcal{N}(0, var), \hat{c} \sim \mathcal{N}(0, var)$.
 - 4: $a = round(\hat{a}), b = round(\hat{b}), c = round(\hat{c})$.
 - 5: **if** $\sqrt{a^2 + b^2 + c^2} \leq 5$ **then**
 - 6: Añadimos elemento de coordenadas [a b c] a Ω .
 - 7: **return** Ω
-

En este trabajo se van a explorar algunos muestreos Ω que denominaremos muestreos con **handicap**. Estos muestreos (además de cumplir con las tres características anteriormente mencionadas), aseguran muestrear todo el centro del espacio Q . Este centro corresponde a un cubo subconjunto de c , de tamaño $3 \times 3 \times 3$ ubicado en la coordenada central de c .

La razón de compresión (RC) se define como el entero más próximo posible de la razón entre el tamaño de la señal a recuperar y sus muestras tomadas

$$RC = \frac{N}{m}. \quad (4.21)$$

En esta tesis se manejan razones de compresión que van desde 2 hasta 8, y los posibles valores m para estas razones se encuentran en la Tabla 4.2



(a) Muestreo Gaussiano. (b) Muestreo uniforme.

Figura 4.4: Ejemplo de muestreos en 2D. Esta figura presenta un par de ejemplos de muestras tomadas de un grid en un plano 2D. El muestreo Gaussiano favorece más el centro de este espacio. Nota: El tamaño de este grid 2D no corresponde al utilizado en esta tesis, es sólo un ejemplo ilustrativo. Figuras tomada de [5].

4.5. Optimizando parámetros

Finalmente en esta tesis buscamos explorar optimizar los parámetros del problema LASSO (Ecuación (4.1)) que pueden influir en la calidad de la reconstrucción:

- λ . Este parámetro es una penalización a la no rareza del vector solución x , además está relacionada con la cantidad la tolerancia al ruido del problema.
- Ω Índices de las muestras a tomar. Buscamos saber si existe un conjunto de muestreos que resuelvan de manera más eficiente el problema independientemente de la señal de entrada.

Muy probablemente no existe un muestreo óptimo para todas las posibles estructuras que se puedan simular, menos para datos *in-vivo*, pues habrá muestreos que favorezcan cierto tipo de estructura y dirección que otros. Sin embargo (y sobre todo en razones de compresión elevadas) buscamos explorar la posibilidad de obtener un conjunto de muestreos Ω élites que de manera general obtengan resultados aceptables.

Para optimizar estos parámetros se toma una estrategia de optimización alternada, donde dejamos fijo a Ω y optimizamos para λ , para posteriormente fijar a λ y optimizar para Ω . Esta optimización alternada se realiza mediante un simple esquema de aprendizaje de máquina supervisado, descrito en las siguientes secciones.

4.5.1. Valores iniciales

Para iniciar el método iterativo se realiza el siguiente procedimiento para elegir una λ_0 inicial.

Razón de compresión	m
2	259
3	173
4	129
5	103
6	87
7	75
8	65

Tabla 4.2: Total de muestras m - Total de muestras m adquiridas según su razón de compresión

1. Se propone un conjunto de $nTest$ señales de prueba creadas aleatoriamente (Algoritmo 3).
2. Obtenemos el p_{gt} de cada una de ellas mediante su transformada de Fourier.
3. Agregamos ruido Riciano a estas señales.
4. Cada una de estas señales es muestreada mediante un conjunto Ω , muestreo proveniente de una distribución Gaussiana (Será un muestreo aleatorio distinto para cada señal).
5. El método *lasso* en Matlab acepta resolver el problema LASSO sin que el valor de λ este especificado. Cuando esto sucede, Matlab calcula el valor más grande posible de lambda que da una solución casi nula, a partir de esta, calcula una secuencia de posibles valores lambda, donde el máximo valor en la secuencia es el obtenido en el paso anterior.
6. De este conjunto de posibles valores λ , conservamos aquel donde la solución x propuesta por Matlab es aquella que nos da el menor error con respecto al p_{gt} de la señal S en cuestión.
7. Finalmente promediamos las $nTest$ valores λ obtenidos, para tener una aproximación de un λ .

4.5.2. Parámetro Ω

Es muy poco probable que exista un solo Ω que sea capaz de muestrear de manera óptima (para su posterior reconstrucción del EAP) todas las posibles formas y

4. METODOLOGÍA

estructuras contenidas en señales S , sin embargo, en este trabajo procuramos buscar al menos un conjunto de muestreos Ω élites que funcionen en promedio mejor que todo el conjunto de posibles combinaciones de muestreos aleatorios. En una iteración de optimización en Ω , fijamos el parámetro λ (la primera iteración se realiza con la propuesta de λ_0 inicial, para el resto de las iteraciones con el λ óptimo calculado de la iteración de λ anterior). Para la convergencia de este parámetro, idealmente buscamos no haya muchos cambios en el conjunto de élites Ω entre iteraciones. La metodología se resume en:

1. Se propone un conjunto de $nTest$ señales de prueba creadas aleatoriamente (Algoritmo 3).
2. Obtenemos el p_{gt} de cada una de ellas mediante su transformada de Fourier.
3. Agregamos ruido Riciano a estas señales.
4. Se proponen un conjunto de muestreos Ω a probar.
 - Este conjunto se compone en parte de muestreos creados con una distribución Gaussiana, y el resto de muestreos Gaussianos con handicap que aseguraran muestrear el centro del espacio Q (ver Sección 4.4).
5. En caso de no ser la primera iteración, realizamos “elitismo” al conjunto de muestreo a probar, esto es que agregamos el sub-conjunto de los mejores muestreos obtenidos de la iteración anterior a la población a probar.
6. Para cada muestro Ω en el conjunto propuesto, resolveremos el problema LASSO (Ecuación (4.1)) con cada una de las señales ruidosas propuestas en el conjunto de señales.
 - Esto lo hacemos con las λ obtenidas en el procedimiento de optimización de λ anterior.
 - Para cada solución obtenida en este paso, mediremos el error contra los p_{gt} de las señales sin ruido.
7. Para cada muestro Ω en el conjunto propuesto, obtenemos el error promedio obtenido del las soluciones obtenidas en el problema LASSO con cada una de las señales a reconstruir.
8. Guardamos el sub-conjunto de aquellas Ω que dieron en promedio el menor error obtenido del paso pasado, con fines de realizar “elitismo” en futuras iteraciones.
9. Elegimos de este subconjunto como muestreo óptimo a aquella Ω que obtuvo menor error.
10. Checar convergencia.

4.5.3. Parámetro λ

El parámetro λ es específico para cada problema, pero es posible aproximarnos al rango de soluciones validas para el problema. Para verificar dicho rango se propone la siguiente heurística:

1. Definimos un rango de búsqueda.
2. Resolvemos el problema LASSO con una población aleatoria de señales S con alguna propuesta de λ contenido en el espacio de búsqueda, estos se hacen con un esquema de muestreos aleatorios.
3. Promediamos el error obtenido de las soluciones obtenidas en LASSO con dicha población de señales.
4. Basándonos en estos errores promedios podemos obtener cierta intuición de ese espacio de valores buenos para λ .
5. Repetimos aproximando el espacio de búsqueda, hasta quedar satisfecho con el rango.

Sin embargo, para la optimización de Ω requerimos dejar fijo el valor de λ , para ello, fijamos el parámetro Ω con el mejor muestreo Ω del conjunto de élites de la iteración pasada para el entrenamiento y entrenamos el valor de λ con la siguiente metodología:

1. Se propone un conjunto de $nTest$ señales de prueba creadas aleatoriamente (Algoritmo 3).
2. Obtenemos el p_{gt} de cada una de ellas mediante su transformada de Fourier.
3. Agregamos ruido Riciano a estas señales.
4. Cada una de estas señales es muestreada mediante el muestreo óptimo Ω , resultado de la optimización de Ω anterior.
5. Obtenemos con Matlab un conjunto de posibles valores λ donde existe una solución valida.
6. De este conjunto de posibles valores λ , conservamos aquel donde la solución x propuesta por Matlab es aquella que nos da el menor error con respecto al p_{gt} de la señal S en cuestión.
7. Finalmente promediamos las $nTest$ valores λ obtenidos, para una aproximación de la λ óptima del Ω óptimo en la iteración.
8. Checar convergencia.

Para la convergencia de este parámetro, se busca que los cambios en λ no sean tan fuertes entre iteraciones, pero con un criterio muy suave, por que es probable que el mejor élite sea distinto en cada iteración, entonces se espera haya algo de variación en este parámetro.

Resultados y análisis

En este capítulo presentamos los resultados obtenidos respecto a las metodologías del capítulo anterior. En este capítulo se explica cuales con las métricas de error más adecuadas para el estudio en cuestión. Se resumen los resultados de la propuesta de optimización de los parámetros λ y Ω , y se detallan algunas conclusiones al respecto. Además se presentan una serie de pruebas y resultados para probar la robustez del método, así como una serie de ejemplos de reconstrucciones correctamente obtenidas mediante la metodología propuesta en esta tesis.

5.1. Desempeño de métricas bajo variaciones de ruido Riciano

Para medir la sensibilidad de las métricas de error presentadas en la Sección 4.1.1, se propone un sencillo experimento donde generamos señales aleatorias, les agregamos ruido Riciano en distintos niveles de SNR , extraemos sus EAP y comparamos la diferencia con las métricas presentadas. Los resultados del desempeño promedio de las distintas métricas se encuentra en la Tabla 5.1, mientras que la varianza de estos promedios se encuentra en la Tabla 5.2.

Idealmente nos gustaría que existiera una relación lineal entre el nivel de ruido y nivel de error, esto daría una buena intuición del significado del valor numérico obtenido en la métrica del error con respecto el ruido en la señal. En las Figuras 5.1 y 5.2, hacemos una visualización de los errores obtenidos en diferentes SNR de la Tabla 5.1. En la Figura 5.1, podemos observar que existe cierta estructura del nivel de SNR y las métricas, sin embargo las señales más ruidosas ($SNR < 30$) afectan considerablemente dicha forma. En la Figura 5.2 sólo graficamos para aquellos valores de $SNR \geq 30$, y en esta podemos observar que hay una relación más lineal entre el nivel de ruido y las métricas, siendo la divergencia Kullback-Leibler (Apéndice C) la que se comporta de manera más lineal. También es importante destacar que es la divergencia KL la que tiene mayor varianza en sus resultados, siendo esto interesante pues implica que

SNR	Euc	EMD	Kull-Leib	Jen-Shan
10	0.6824	0.9599	5.6536	0.1907
20	0.4334	0.5280	1.9542	0.0950
30	0.2864	0.3527	1.3171	0.0608
40	0.2393	0.2645	1.0763	0.0453
50	0.1930	0.2171	0.9021	0.0357
60	0.1594	0.1770	0.7718	0.0297
70	0.1413	0.1553	0.6808	0.0251
80	0.1181	0.1311	0.5904	0.0217
90	0.1047	0.1159	0.5138	0.0187
100	0.0981	0.1076	0.4657	0.0167

Tabla 5.1: Desempeño promedio de métricas de error a diferentes SNR

tiene mayor sensibilidad para captar estas diferencias que las diferentes métricas. Por estas razones, además de que es muy fácil de calcular, será **la divergencia Kullback-Leibler la métrica estándar de error** que usaremos para las distintas pruebas en esta tesis.

5.2. Resultados de optimización de parámetros

En la Sección 4.5 se presentó una estrategia para optimizar los parámetros λ y Ω . Esta consiste en alternar la optimización de ambos parámetros, hasta lograr un punto de convergencia. Sin embargo, algunas observaciones en los experimentos surgen con respecto a la convergencia de estos parámetros:

- El valor de λ depende de la naturaleza de las muestras a tomar, esto es, del ruido en la señal, de la forma natural de la señal y de la población muestreada. Por lo tanto va a existir un valor específico de λ óptimo para cada caso distinto. Lo mejor que podemos hacer es procurar aproximarnos al rango donde se encuentran estos valores óptimos.
- De la misma forma, va a existir un muestreo óptimo Ω distinto para cada caso específico, es de esperar ya que ciertos muestreos aleatorios van a favorecer muestrear la forma del EAP más que otros. Esto se hace más evidente con los experimentos realizados, y a pesar de eso se verá que hacer muestreos aleatorias con las consideraciones propuestas en esta tesis es una buena estrategia.

SNR	Euc	EMD	Kull-Leib	Jen-Shan
10	1.23E-02	6.40E-03	2.04E+00	2.43E-04
20	5.30E-03	2.96E-03	1.57E-01	4.51E-05
30	2.14E-03	1.40E-03	8.26E-02	1.90E-05
40	1.48E-03	8.78E-04	2.91E-02	1.87E-05
50	1.22E-03	1.00E-03	1.89E-02	9.88E-06
60	8.58E-04	4.25E-04	1.26E-02	5.85E-06
70	5.20E-04	2.73E-04	9.28E-03	4.97E-06
80	3.81E-04	2.21E-04	8.39E-03	3.39E-06
90	3.90E-04	1.68E-04	5.66E-03	4.26E-06
100	3.74E-04	2.32E-04	6.24E-03	2.60E-06

Tabla 5.2: Varianza del desempeño promedio de métricas de error a diferentes SNR

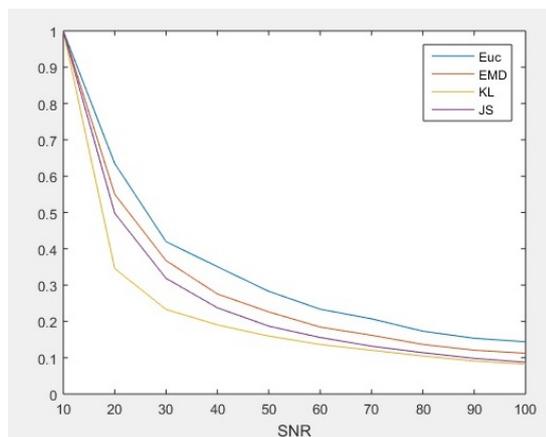


Figura 5.1: Ajuste lineal visual de métricas de error. Esta gráfica se encuentra normalizada con respecto al error máximo obtenido por cada métrica. En esta figura podemos notar cierta forma curva en el desempeño de las métricas, donde la curva sucede debido a desempeño de valores SNR bajos.

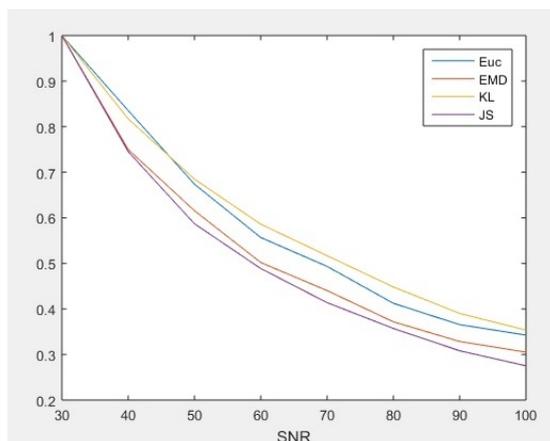


Figura 5.2: Ajuste lineal visual de métricas de error para $SNR \geq 30$. Esta gráfica se encuentra normalizada con respecto al error máximo obtenido por cada métrica para $SNR \geq 30$. En esta figura podemos notar que efectivamente hay una forma de relación lineal entre el error obtenido en las diferentes métricas en $SNR \geq 30$, siendo la divergencia Kullback-Leibler la que mejor se ajusta a una línea.

En las siguientes secciones se elaboran más a detalle estas observaciones.

5.2.1. Convergencia parámetro λ

En algunos resultados previos de la metodología de optimización propuesta en la Sección 4.5.3, se observó que el parámetro λ no daba indicios de converger, pues este es aproximadamente óptimo solamente al muestreo Ω elegido como óptimo en la iteración anterior pero no a toda la población de muestreos a probar. Esto no es sorprendente, pues el valor óptimo para λ es único para cada problema, aunque es posible que nos podamos aproximar al rango que pueda definirse como un espacio de solución para este parámetro.

Para aproximarnos a un rango que contenga posibles valores adecuados para el parámetro λ , seguimos la metodología propuesta en la Sección 4.5.3. A partir de los primeros resultados se determinó que un rango de búsqueda inicial aceptable se encuentra en $[\cdot001, \cdot5]$. En la Figura 5.3 se muestra una gráfica con los resultados correspondientes de dicha prueba, probando con tamaños de pasos en el valor λ de $\cdot001$. De esta figura podemos empezar a notar que un rango más preciso de búsqueda se encuentra en $[\cdot005, \cdot015]$, aunque también podemos observar que para razones de compresión (RC) elevadas el valor λ tiene un comportamiento inestable. Por otro lado, al realizar la misma prueba con muestreos Gaussianos con handicap (que como será visto en la siguiente sección, tienen un mejor desempeño y resultan más adecuados a nuestro problema) este comportamiento es más estable. Los resultados de dicha prueba de muestreos con

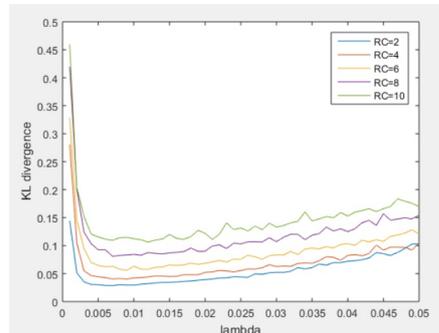


Figura 5.3: Rango aproximado de valores óptimos λ . Se promedia el error obtenido para un conjunto de pruebas en diferentes λ , el mínimo se encuentra aproximadamente en 0.011.

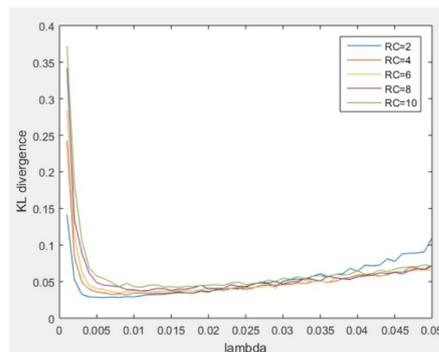


Figura 5.4: Rango aproximado de valores óptimos λ de muestreos con handicap. Se promedia el error obtenido para un conjunto de pruebas en diferentes λ , el mínimo se encuentra aproximadamente en 0.011.

handicap se encuentran en la Figura 5.4. De esta podemos observar que existe mayor estabilidad para RC más elevados, esto gracias a la información del handicap ayuda considerablemente a pruebas con menores muestras, además podemos notar que un rango adecuado para valores λ sigue siendo $[\.005, \.015]$. También podemos observar que a mayor RC mayores valores λ son requeridos, pues note los valores más adecuados de λ se encuentran justo después de la primera gran curva en cada función de la gráfica, la cual sucede cada vez mas próxima a $\.015$ conforme el RC aumenta.

En la Sección 4.5.3 también se estableció una metodología para optimizar el parámetro λ dado un muestreo fijo, proceso que es parte de un método iterativo más general que procura optimizar alternadamente ambos parámetros. En la siguiente sección veremos que es posible obtener un conjunto de muestreos élite que son óptimos para un conjunto de señales de prueba. Para dichos élites es posible establecer un valor aproximadamente óptimo de λ con la metodología propuesta. Un ejemplo de valores λ obtenidos para un

muestreo Ω élite obtenido en diferentes RC se encuentran en la Tabla 5.3. De esta tabla es más claro que a mayor RC se requiere un valor de λ más elevado (aunque de una magnitud similar), esto es lógico pues elevar λ implica que relajamos el problema para tolerar un poco más error, lo cual es de esperar a valores de RC elevados.

RC	λ óptima
2	0.006
4	0.010
6	0.011
8	0.013
10	0.015

Tabla 5.3: Valores óptimos de λ para un muestreo élite

5.2.2. Convergencia de Ω

La metodología propuesta en la Sección 4.5.2 hace una búsqueda en el conjunto de posibles Ω , al estar proponiendo conjuntos de muestreo en cada iteración del método. En cada iteración hacemos “elitismo” de la población, esto es que se conservan aquellos muestreos Ω que dieron mejores resultados en la iteración y los inyecta en la siguiente población de muestreos a probar. Con esto buscamos explorar si existe algún conjunto de posibles muestreos Ω que puedan fungir como muestreos universales.

En la metodología, para una iteración donde se va optimizar una nueva población de muestreos, se proponía un nuevo conjunto de señales aleatorias con que optimizar esta población, tal que en la siguiente iteración, un nuevo conjunto de señales optimice la nueva población (con elitismo) de muestreos. Si los élites después de varias iteraciones empezaran a converger, daríamos por terminado el método. Sin embargo, en base a pruebas iniciales, se observó que los élites entre las iteraciones eran distintos para cada iteración, evitando la convergencia.

Por otro lado, se hicieron una serie de pruebas con las misma metodología pero fijando el conjunto de señales con las cuales entrenar el muestreo para todas las iteraciones del experimento. En estas pruebas se pudo observar que había elitismo de manera efectiva, conservando entre iteraciones élites que provenían inclusive de primeras iteraciones. Esto refuerza la idea de que efectivamente hay muestreos que favorecen cierto tipo de estructura, pero muy difícilmente podemos encontrar un conjunto que funcione mejor o peor en todos los casos, sino que tenemos un conjunto de soluciones optimales que podemos usar de manera indistinta para sentido práctico.

5.3. Desempeño de recuperación del propagador conjunto promedio

En esta sección se presentan resultados de los errores obtenidos promediados al reconstruir el EAP de un conjunto de señales aleatorias ruidosas mediante el esquema de CS presentado. Para esto, bajo un mismo nivel de ruido de $SNR = 30$, generamos un conjunto de 250 señales S aleatorias, las cuales son muestreadas bajo distintas razones de compresión para posteriormente reconstruir su EAP con la metodología propuesta, fijando el parámetro λ según los resultados anteriores. Las métricas de error presentadas son con la distancia Euclideana y la divergencia Kullback-Leibler.

La Tabla 5.4 presenta los resultados de realizar este experimento con muestreos aleatorios Gaussianos, tal que cada muestreo será distinto para cada señal a reconstruir. En la Tabla 5.5 se realiza el experimento con muestreos aleatorios Gaussianos pero esta vez nos aseguramos de muestrear el centro del espacio-Q (handicap). En la Tabla 5.6 se realiza el experimento pero fijando el muestreo con alguno de los muestreos élite obtenido en los resultados anteriores.

RC	Prom. Euc	Var. Euc	Prom. KL	Var. KL
2	0.1253	0.0006	0.0298	0.0001
4	0.1391	0.0011	0.0400	0.0004
6	0.1594	0.0016	0.0639	0.0012
8	0.1629	0.0018	0.0661	0.0013
10	0.2205	0.0035	0.1344	0.0074

Tabla 5.4: Desempeño de muestreos Gaussianos

RC	Prom. Euc	Var. Euc	Prom. KL	Var. KL
2	0.1226	0.0006	0.0279	0.0002
4	0.1263	0.0011	0.0333	0.0003
6	0.1282	0.0012	0.0374	0.0004
8	0.1339	0.0015	0.0395	0.0005
10	0.1423	0.0022	0.0440	0.0007

Tabla 5.5: Desempeño de muestreos Gaussianos con handicap

RC	Prom. Euc	Var. Euc	Prom. KL	Var. KL
2	0.1216	0.0006	0.0271	0.0001
4	0.1192	0.0009	0.0287	0.0002
6	0.1171	0.0010	0.0299	0.0002
8	0.1214	0.0014	0.0326	0.0004
10	0.1254	0.0019	0.0357	0.0006

Tabla 5.6: Desempeño de muestreos élite

En base a los resultados de estas tablas queremos realizar las siguientes observaciones:

- La diferencia de error promedio obtenido entre aquellos muestreos aleatorios y aquellos muestreos aleatorios con cierta ayuda (muestreos élites y muestreos con handicap) son considerables, por lo tanto es buena estrategia optar por alguna de estas alternativas.
- Los muestreos élites son aquellos con los mejores resultados, aunque realizar muestreos aleatorios con handicap en el centro es una buena estrategia que no requiere optimización.
- La metodología filtra el ruido Riciano. Note que los errores promedios presentados son significativamente menores a aquellos obtenidos en la Tabla 5.1, esto es un indicativo de que los resultados tienen menor ruido, el cual sensibilizaba a las métricas.
- El error es mayor conforme el RC aumenta. Esto es de esperar, pues el problema se vuelve más complejo conforme menos muestras.

5.3.1. Ejemplos de reconstrucción

En esta sección presentamos algunos ejemplos de reconstrucciones obtenidas con la metodología propuesta, visualizando el ODF (Sección 3.5.1), el cual condensa el EAP en una figura donde podemos distinguir los principales picos y cruces del EAP, obtenido mediante un software para visualizar ODFs en DSI Studio [37]. Cada figura corresponde a la reconstrucción del EAP mediante la metodología, partiendo de una señal aleatoria generada con base en una microestructura plausible con $SNR = 30$, con su respectivo error KL. Recuerde estos errores comparan el EAP reconstruido con aquel que obtendríamos de la señal completa sin ruido.

Las Figuras 5.5 a 5.7 contienen dichos ejemplos. De estas figuras surgen las siguiente observaciones:

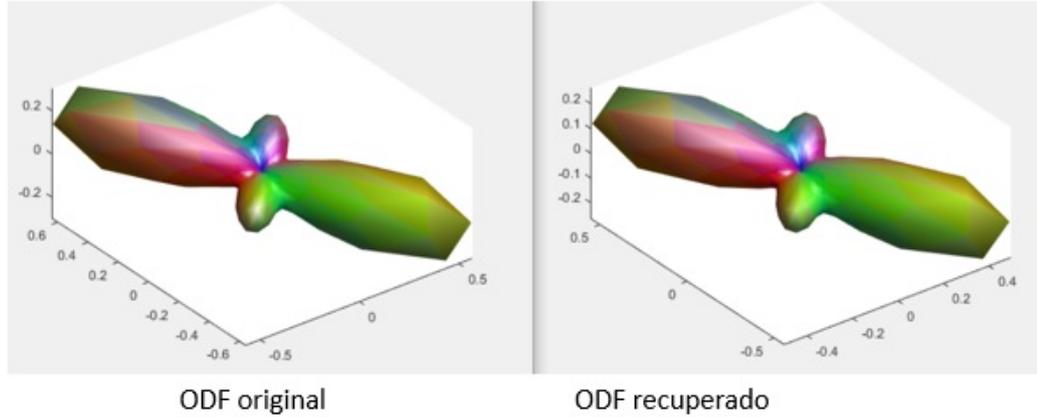


Figura 5.5: Ejemplo de reconstrucción 1. $RC = 2$, $SNR = 30$, $D_{kl} = 0.0193$, $D_{euc} = 0.0657$. Visualización del ODF obtenido al reconstruir el EAP con la metodología presentada.

- Reiteramos en el protocolo más común de DSI se requieren 515 muestras del espacio Q . En estos ejemplo se requieren considerablemente menos muestras (259,129,65), además debido a la simetría del espacio Q , en realidad solo se requiere la mitad de estas muestras (129,65,33) y usar como muestra su punto simétrico, virtualmente teniendo un radio de compresión aún menor en todos los casos.
- Aunque el ODF obtenido no es una reconstrucción perfecta (esto se debe principalmente debido al ruido Riciano), la aproximación obtenida es más que suficiente para su uso práctico.
- Otra ventaja interesante del esquema dicho, es que debido a que el diccionario P propuesto parte de un modelo sin ruido, entonces el resultado final también es sin ruido, haciendo un filtrado de ruido efectivo en la reconstrucción final.
- Además de las reconstrucciones obtenidas, el vector solución x propiamente normalizando puede dar una intuición del porcentaje de composición de microestructura que tienen estos EAP. Esta idea puede explorarse más a fondo con diccionarios que posean información aún mas específica como difusión intra-xonal.

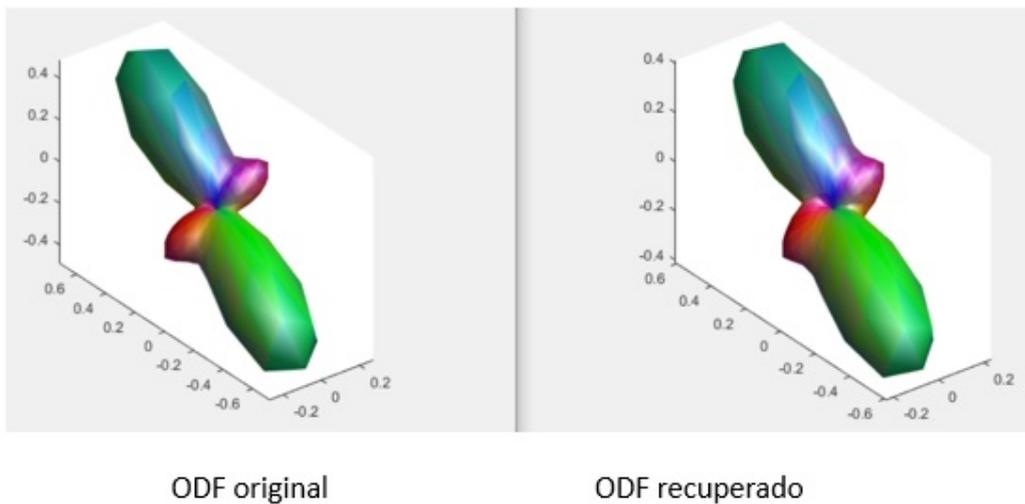


Figura 5.6: Ejemplo de reconstrucción 2. $RC = 4, SNR = 30, D_{kl} = 0.0235, D_{euc} = 0.0899$. Visualización del ODF obtenido al reconstruir el EAP con la metodología presentada.

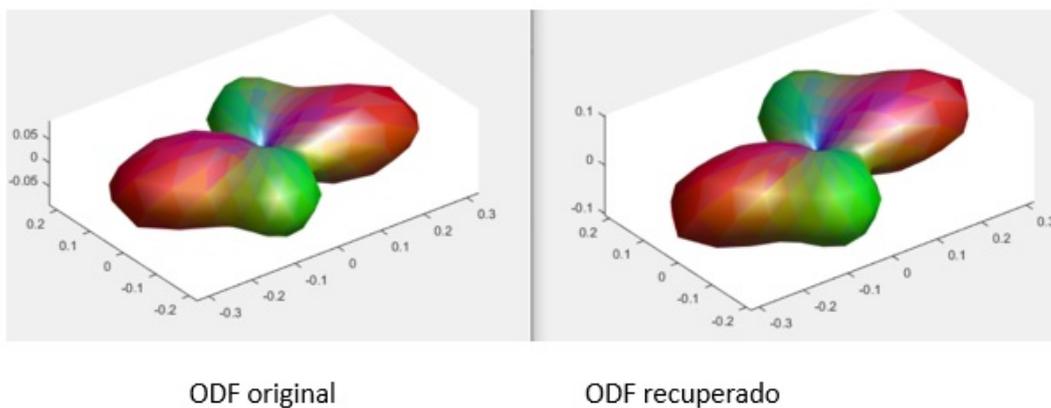


Figura 5.7: Ejemplo de reconstrucción 3. $RC = 8, SNR = 30, D_{kl} = 0.0107, D_{euc} = 0.0766$. Visualización del ODF obtenido al reconstruir el EAP con la metodología presentada.

Conclusiones

En el capítulo anterior se hicieron algunos comentarios como análisis de los resultados obtenidos con la metodología propuesta. De estos queremos enfatizar algunos puntos:

- La divergencia Kullback-Leibler presenta una sensibilidad interesante al ruido Riciano, esto nos permite usarla como base de métrica de error para comparar y optimizar los resultados obtenidos con los esperados.
- Es posible obtener un rango de valores óptimos para el parámetro λ , y aunque no es posible obtener un valor específico debido a la naturaleza del problema, este rango puede servir como base de selección.
- La teoría del CS sugiere que los muestreos sean aleatorios, pero gracias al conocimiento que tenemos del problema (imágenes de DW-MRI) podemos asegurar mejores muestreos al realizarlos con una distribución Gaussiana, y aún mejores muestreos si aseguramos el centro del espacio \mathcal{Q} .
- A sí mismo, es posible obtener un conjunto de muestreos élites que aseguran obtener menor error en promedio con la metodología propuesta.
- El error obtenido en las reconstrucciones del EAP se encuentra en rangos muy aceptables, por lo tanto el diccionario P propuesto en esta tesis es apto y funcional para este esquema.

Podemos dejar como conclusión que la metodología presentada cumple con los objetivos propuestos en esta tesis, al poder realizar de manera efectiva sensado comprimido en DSI, haciendo uso de un diccionario que contiene información de distribuciones de difusión de hidrógeno, lo cual nos ayuda a inferir aún más la micro-estructura interna de los vóxeles del estudio. Además, este esquema abre las puertas para futuras investigaciones aprovechando la forma de construcción de dicho diccionario.

6.1. Trabajo a futuro

El diccionario con distribuciones de difusión de hidrógeno presentado ha funcionado adecuadamente, por lo tanto el vector solución contiene información de las posibles microestructuras contenidas en la señal de difusión, esto abre las puertas para explorar a mayor detalle este tipo de propuestas, las cuales no se han observado (en nuestro conocimiento) en otros trabajos. Una posibilidad para ampliar el esquema mencionado es crear átomos del diccionario que corresponden a “sticks”. Estos sticks buscan explicar la difusión intraxonal dentro de los vóxeles del estudio. Esto es posible mediante un modelo descrito por Neuman para difusión en medios restringidos cilíndricos [38]. Con estos sticks sería posible obtener una estimación de la difusión intraxonal, lo cual sería información muy útil para algunas aplicaciones sobre estimación de propiedades microestructurales y detección de daño axonal.

En este trabajo todas las señales de entrenamiento y prueba son sintéticas (mediante modelos de tensor multicompartimiento con ruido sintético). Es importante llevar este método con pruebas *in-vivo* al campo clínico, para poder validar el esquema con señales reales. De ser posible esta validación, este método podría ser utilizado en aplicaciones clínicas para DSI, con un tiempo considerablemente menor de muestreo que los protocolos habituales de dicha técnica.

Algoritmo ADMM

Una forma de resolver el problema Lasso en ecuación (2.16) es un algoritmo basado en *Alternating Direction Method of Multipliers* (ADMM) [39]. La siguiente nomenclatura se conserva del paper en cuestión. Este método resuelve el problema de la forma

$$\begin{aligned} \min_{x, z} \quad & l(x) + g(z) \\ \text{s.t.} \quad & Ax + Bz = c \end{aligned} \tag{A.1}$$

Con esta notación, el problema de regresión Lasso es

$$\begin{aligned} \min_{x, z} \quad & l(x) + g(z) = \frac{1}{2} \|Ax - b\|_{\ell_2} + \lambda \|z\|_{\ell_1} \\ \text{s.t.} \quad & x - z = c \end{aligned} \tag{A.2}$$

Este es un algoritmo iterativo, donde las actualizaciones hechas en cada iteración son:

$$x^{k+1} = (A^T A + \rho I)^{-1} (A^T b + \rho(z^k - u^k)) \tag{A.3}$$

$$z^{k+1} = S_{\lambda/\rho}(x^{k+1} + \mu^k) \tag{A.4}$$

$$\mu^{k+1} = \mu^k + x^{k+1} - z^{k+1} \tag{A.5}$$

donde A es el conjunto de datos (arreglo alto, nuestra matriz de sensado comprimido), x contiene los coeficientes (o la solución en este caso sparse que buscamos al problema), ρ es el parámetro de penalización (parámetro de Lagrangiano aumentado), b es la respuesta (arreglo grande, en este caso nuestras muestras), y S es el operador de *soft thresholding*

$$S_k(a) = \begin{cases} a - k, & a > k \\ 0, & \|a\| \leq k \\ a + k, & a < -k \end{cases} \tag{A.6}$$

Información adicional sobre el valor q

En la literatura suelen existir dos definiciones del valor de q , naturalmente ambas son equivalentes, sólo hay tener en claro una distinción del tipo de unidades en las que se encuentra, dependiendo del contexto en el que se maneja.

El valor q se define como

$$q = \gamma\delta G, \tag{B.1}$$

donde

- γ Es la proporción giromagnética, comúnmente sus unidades son $\frac{rad}{sT}$
- δ Es el tiempo del pulso de gradiente, sus unidades son s
- G Es la magnitud (también dirección) del gradiente de difusión, sus unidades son $\frac{T}{m}$

Por lo tanto las unidades del valor q son $\frac{rad}{m}$. En la imagenología del espacio Q, este espacio Q es análogo a un espacio de frecuencias en un análisis clásico de Fourier. Podemos ver a los valores q como los componentes frecuenciales de este espacio.

En el *Q-space imaging* el valor q_i se define como

$$\hat{q} = \frac{\gamma\delta G}{2\pi}. \tag{B.2}$$

De esta forma hacemos la conversión de radianes a ciclos tal que el valor \hat{q} tiene sus unidades en $\frac{1}{m}$, análogo a un componente frecuencial de m .

Métricas entre distribuciones

Dado a que los EAP son funciones de densidad de probabilidad discretas, es posible utilizar algunas métricas comunes en el campo de la probabilidad y aprendizaje de máquina para medir similitud entre ellos. En este anexo definimos algunas de las utilizadas en esta tesis.

Distancia euclidiana

La definimos como

$$D_{euc} = \frac{\|P - Q\|_f}{\|P\|_f}, \quad (\text{C.1})$$

donde consideramos a Q el EAP reconstruido a comparar el el EAP de referencia P .

La norma Frobenius para el caso 3D se define como

$$\|X\|_f = \sqrt{\sum_k \sum_j \sum_i (x_{ijk})^2}. \quad (\text{C.2})$$

Earths Mover's Distance (EMD)

Obtener esta distancia implica minimizar el siguiente problema de optimización [40]

$$\begin{aligned} \min_x \quad & \sum_{ij} c_{ij} x_{ij} \\ \text{s.t.} \quad & x \geq 0 \\ & \sum_j x_{ij} = wp_j \\ & \sum_i x_{ij} = wq_i \end{aligned} \quad (\text{C.3})$$

donde:

- C : Matriz de pesos, para este caso simple distancia euclideana entre dos puntos para x .
- w_p, w_q : Valor del pin en el histograma 3D. Deben estar normalizados.

Divergencia Kullback-Leibler

Para el caso discreto se encuentra definida como [41]

$$D_{kl}(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}. \quad (\text{C.4})$$

Notas:

- Esta divergencia no es simétrica $D_{kl}(P||Q) \neq D_{kl}(Q||P)$.
- Solo se encuentra definida si en $Q_i = 0$ implica $P_i = 0$.
- Si $P_i = 0$, esa contribución se toma como 0.

Divergencia Jensen-Shannon

Para el caso discreto se encuentra definida como [42]

$$D_{js}(P||Q) = \frac{1}{2}D_{kl}(P||M) + \frac{1}{2}D_{kl}(Q||M), \quad (\text{C.5})$$

donde

$$M = \frac{1}{2}(P + Q). \quad (\text{C.6})$$

La D_{js} es una versión simétrica y suave de la divergencia Kullback-Leibler.

Bibliografía

- [1] P. T. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford University Press, 1991. [1](#), [21](#)
- [2] E. J. Candes and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 21, 2008. [1](#), [2](#), [5](#), [6](#), [7](#), [8](#), [11](#), [12](#)
- [3] P. J. Basser *et al.*, “MRI diffusion tensor spectroscopy and imaging,” *Biophysical Journal*, vol. 66, pp. 259–267, 1994. [1](#), [18](#)
- [4] V. J. Wedeen *et al.*, “Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging,” *Magnetic Resonance in Medicine*, vol. 54, pp. 1377–1386, 2005. [1](#), [21](#), [22](#), [23](#), [28](#)
- [5] M. I. Menzel *et al.*, “Accelerated diffusion spectrum imaging in the human brain using compressed sensing,” *Magnetic Resonance in Medicine*, vol. 66, pp. 1226–1233, 2011. [3](#), [36](#), [38](#)
- [6] S. L. Merlet, M. Paquette, R. Deriche, and M. Descoteaux, “Compressive Sensing DSI,” in *International Symposium on BIOMEDICAL IMAGING: From Nano to Macro - ISBI HARDI Reconstruction Challenge*, (San Francisco, United States), 2013. [3](#)
- [7] S. L. Merlet and R. Deriche, “Continuous diffusion signal, MRI and MRI estimation via compressive sensing in diffusion MRI,” *Medical Image Analysis*, 2013. [3](#), [6](#), [8](#)
- [8] J. Sun, E. Sakhaee, A. Entezari, and B. C. Vemuri, “Leveraging EAP-sparsity for compressed sensing of MS-HARDI in (k, q)-space,” *Inf Process Med Imaging*, vol. 24, pp. 375–386, 2015. [3](#)
- [9] C. E. Shannon, “Communication in the presence of noise,” *Proc. Institute of Radio Engineers*, vol. 37, pp. 10,21, 1949. [5](#)
- [10] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing*. Prentice-Hall, 1996. [6](#)

BIBLIOGRAFÍA

- [11] Y. C. Eldar and G. Kutyniok, eds., *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012. 6
- [12] E. J. Candes, “Compressive sampling,” *Proceedings of the International Congress of Mathematicians, Madrid, Spain*, 2006. 6
- [13] E. J. Candes and J. Romberg, “Sparsity and incoherence in compressive sampling,” *Inverse Problems*, vol. 23, pp. 969–985, 2007. 10
- [14] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society*, vol. 58, pp. 267–288, 1996. 11
- [15] Z.-P. Liang and P. C. Lauterbur, *Principles of magnetic resonance imaging: A Signal Processing Perspective*. IEEE Press, 2000. 13
- [16] J. P. Hornak, “The basics of MRI.” <https://www.cis.rit.edu/htbooks/mri/index.html>, 2017. 13
- [17] J. G. Webster, ed., *Medical Instrumentation Application and Design*. Jhon Wiley and Sons, 2010. 13
- [18] A. Fick, “Concerns diffusion and concentration gradient,” *Phys Lpz*, 1855. 14
- [19] P. J. Basser and E. Ozarslan, “Introduction to diffusion MRI,” in *Diffusion MRI From quantitative measurement to in-vivo neuroanatomy* (H. Johansen-Berg and T. E. Behrens, eds.), Elsevier, 2009. 14
- [20] J. R. Patiño López, “Design and validation of a robust diffusion-weighted-MRI monte carlo simulator,” Master’s thesis, Centro de Investigación en Matemáticas, 2015. 14
- [21] R. Brown, “A brief account of microscopical observations made in the months of june, july, and august, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies,” *Phil Mag*, 1828. 14
- [22] A. Einstein, “Über die von der molekularkinetischen theorie der wärme gefordete bewegung von in ruhenden flüssigkeiten suspendierten teilchen,” *Ann Physik*, 1905. 14
- [23] D. C. Alexander, “An introduction to computational diffusion MRI: the diffusion tensor and beyond,” in *Visualization and image processing of tensor fields* (J. Weickert and H. Hagen, eds.), Springer, 2006. 14, 15, 17, 19, 21, 22
- [24] W. S. Price, “Pulsed-field gradient nuclear magnetic resonance as a tool for studying translational diffusion: part 1. basic theory,” *Water Research Institute*, 1997. 15

-
- [25] E. O. Stejskal and J. E. Tanner, “Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient,” *J Chem Phys*, vol. 42, pp. 288–292, 1965. [15](#)
- [26] I. SAS, “e-MRI: an interactive online course about MRI physics.” <https://www.imaios.com/en/e-Courses/e-MRI>, 2017. [18](#)
- [27] E. O. Stejskal, “Use of spin echoes in a pulsed magnetic field gradient to study anisotropic, restricted diffusion and flow,” *J Chem Phys*, vol. 43, pp. 3597–3603, 1965. [18](#)
- [28] D. K. Jones, “Gaussian modeling of the diffusion signal,” in *Diffusion MRI From quantitative measurement to in-vivo neuroanatomy* (H. Johansen-Berg and T. E. Behrens, eds.), Elsevier, 2009. [19](#), [20](#), [21](#)
- [29] P. J. Basser and D. K. Jones, “Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review,” *NMR in Biomedicine*, vol. 15, p. 456:467, 2002. [21](#)
- [30] E. J. Canales-Rodríguez, L. Melie-García, and Y. Iturria-Medina, “Mathematical description of q-space in spherical coordinates: Exact q-ball imaging,” *Magnetic Resonance in Medicine*, vol. 61, pp. 1350–1367, 2009. [21](#), [22](#)
- [31] D. C. Alexander and K. K. Seunarine, “Mathematics of crossing fibers,” in *Diffusion MRI theory methods and applications* (D. K. Jones, ed.), Oxford University Press, 2011. [22](#), [23](#)
- [32] H. Gudbjartsson and S. Patz, “The rician distribution of noisy MRI data,” *Magn Reson Med.*, vol. 34, pp. 910–914, 1995. [23](#)
- [33] S. Rice, “Mathematical analysis of random noise,” *Bell System Tech*, vol. 23, 1944. [24](#)
- [34] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1.” <http://cvxr.com/cvx>, Mar. 2014. [27](#)
- [35] J. P. J. Mairal, F. Bach and G. Sapiro, “Online learning for matrix factorization and sparse coding,” *Journal of Machine Learning Research*, vol. 11, pp. 19–60, 2010. [27](#)
- [36] K. Khare, J. C. Hardy, *et al.*, “Accelerated 3D phase-contrast imaging using adaptive compressed sensing with no free parameters,” *Proc. Intl. Soc. Mag. Reson. Med.*, vol. 18, 2010. [32](#)
- [37] F.-C. Yeh, “Documentation of dsi studio: Odf visualization.” <http://dsi-studio.labsolver.org/Manual/odf-visualization>, 2013. [50](#)
- [38] C. H. Neuman, “Spin echo of spins diffusing in a bounded medium,” *Chem Phys*, vol. 60, 1974. [54](#)
-

BIBLIOGRAFÍA

- [39] S. Boyd *et al.*, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends in Machine Learning*, vol. 3, 2010. [55](#)
- [40] Y. Rubner, C. Tomas, and L. J. Guibas, “A metric for distributions with applications to image databases,” *In IEEE International Conference on Computer Vision*, pp. 59–66, 1998. [57](#)
- [41] D. J. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University, 2003. [58](#)
- [42] J. Lin, “Divergence measures based on the shannon entropy,” *IEEE Transactions on information theory*, vol. 37, 1991. [58](#)