

Centro de Investigación en Matemáticas, A.C.

CIMAT

Método eficiente para análisis de DW-MRI
multicapa basado en diccionarios adaptables

Tesis presentada para la obtención de grado de

Maestro en Ciencias

con Especialidad en Computación

y Matemáticas Industriales.

P R E S E N T A:

Odín Fernando Eufrazio Vázquez

Director de tesis:

Dr. Mariano J.J. Rivera Meraz

Guanajuato, Gto, México. Octubre, 2015.

Resumen

La tesis presenta un método de análisis de datos de Resonancia Magnética Pesados por Difusión. El propósito es determinar la estructura intravoxel de la materia blanca en cerebros (humanos) vivos. La estrategia seguida es:

- Usar método basado en diccionarios de funciones de difusión.
- Realizar la adaptación del diccionario dependiendo de la señal para obtener mejor resolución angular y representación de los compartimientos.
- Proponer un estrategia para usar protocolos con múltiples capas (shells).
- Proponer una nueva penalización a la solución para exigir rareza mediante la combinación de una aproximación de la norma L_0 y una estrategia de continuación de parámetros (templado determinista, deterministic annealing).

Agradecimientos

A mi esposa Lulu y la recién llegada Fernanda, por su gran amor. A mi madre Rosa y mi hermano Claudio, por su apoyo incondicional. A mi amigo Edmundo, por su motivación y consejos.

Por supuesto, al Dr. Mariano, por permitirme realizar este trabajo bajo su asesoría. Y claro, a todos los amigos del CIMAT.

Índice general

1. Introducción	2
2. Marco Teórico	5
2.1. Resonancia Magnética Nuclear	5
2.2. Imágenes de Resonancia Magnética Pesadas por Difusión.	9
2.3. Modelo Mono-Tensorial de DW-MRI (DTI)	12
2.4. Modelo Multi-Tensorial de DW-MRI (DBF)	15
2.5. Diccionario Ralos y Adaptables de Difusión (SADD)	18
3. Modelo Propuesto	21
3.1. Factor de Escala para múltiples protocolos	21
3.2. Término de Regularización	24
3.3. Adaptación de Tensores	26
3.4. Algoritmo Propuesto	28
4. Experimentos	30
4.1. Métricas	30
4.2. Datos Sintéticos	32
4.3. ISBI-WMMC 2015	40
4.4. SPARC-dMRI 2014	42
5. Conclusiones y Trabajo Futuro	45
Bibliografía	47

Capítulo 1

Introducción

El título de un trabajo de investigación es lo primero que se lee, pero a veces el título no ayuda a comprender o valorar dicho trabajo de investigación. Cuando leemos una definición algo complicada, una estrategia para comprenderla es identificar palabras claves que resumen aquello que se quiere transmitir. En este trabajo de investigación existen palabras o conceptos que se deben tener claros al leer el título, estas son: difusión, resonancia magnética, análisis y diccionarios adaptables. Por tanto, esta introducción se enfocará a describir brevemente estos conceptos claves.

Difusión es un proceso de transporte que resulta de un movimiento aleatorio por el choque de las moléculas sin involucrar un movimiento de masa Johansen-Berg and Behrens (2013). Esta definición parece algo contradictoria, ¿cómo hay un proceso de transporte sin movimiento? Un ejemplo común de este fenómeno se presenta cuando en un vaso con agua en reposo se suelta una gota de tinta; en un inicio la tinta permanece concentrada, pero al transcurrir el tiempo la tinta se extiende radialmente, y esto no se debe a un movimiento o flujo en el agua, sino al choque de las moléculas de agua. Además, el fenómeno de difusión está influenciado por la estructura o geometría de su medio ambiente, por tanto, se puede usar la difusión para inferir dicha estructura.

En el caso del cerebro, en específico de la materia blanca, se puede aprovechar el fenómeno de difusión para inferir la estructura de las fibras. Pero, ¿cómo medir la difusión en la materia blanca de una forma no invasiva?. La resonancia magnética es una técnica no invasiva que permite medir la difusión de las moléculas de agua en la materia blanca. Así, las mediciones

de difusión de las moléculas de agua son representadas en imágenes, imágenes de resonancia magnética pesadas por difusión.

Por lo tanto, en este contexto, cuando se dice analizar, significa usar las mediciones de difusión y aplicar modelos matemáticos que ayudan a reconstruir la estructura de las fibras en la materia blanca. Basser et al. (1994) presentó un modelo tensorial capaz de ajustar un solo tensor cuyo eje principal está orientado con el manojo de fibras. A partir de este trabajo pionero, numerosos modelos se han desarrollado para mejorar los resultados: menor sensibilidad al ruido, mejor representación de cruce de fibras, etc.

Ramirez-Manzanares et al. (2007) presentó el modelo Diffusion Basis Function (DBF), el cual es un enfoque capaz de identificar cruce de fibras al usar un diccionario discreto de tensores distribuidos uniformemente en todas las direcciones; así el problema consiste en encontrar la mejor combinación lineal que explican la señal observada. En diferentes evaluaciones, Daducci et al. (2014) y Ramirez-Manzanares et al. (2011), se mencionan ventajas de este enfoque basado en diccionario como, un número reducido de mediciones para alcanzar buenas reconstrucciones y tiempos de ejecución muy eficientes. Por otro lado, una de las principales desventajas es la sobre estimación del número de fibras.

Trabajos recientes como los de Aranda et al. (2015) y Loya and Rivera (2015) han atacado estas desventajas del modelo DBF al modificar la base de tensores fijos del diccionario, es decir diccionarios adaptables; esto con el objetivo de lograr una mejor estimación del número de manojos de fibras con direcciones diferentes dentro del voxel.

Este trabajo de tesis se basa en estos trabajos Ramirez-Manzanares et al. (2007), Aranda et al. (2015) y Loya and Rivera (2015) y se aporta una reformulación del problema de optimización para incluir un factor de escala, también se incorpora un término de regularización que aproxima la norma ℓ_0 y así exigir mayor rareza en el ajuste, además se incluye las ideas para adaptar el diccionario y ver como se comporta con las dos primeras aportaciones. El factor de eficiencia, que es característico de estos modelos, fue una consideración importante durante el desarrollo de las ideas que se aportaron.

Por lo anterior, a continuación se define la estructura de este trabajo de tesis. El capítulo 2 inicia con una revisión general de la técnica de resonancia magnética y su modalidad pesada por difusión. Se describirán los modelos de ajuste tensorial sobre los cuales se ha basado

este trabajo. El capítulo 3 describe las propuestas que se realizan en este trabajo. En primer lugar se muestra la reformulación necesaria en el problema de optimización para deducir un factor de escala. A continuación se discute la necesidad de incorporar un nuevo término de regularización para exigir una mayor rareza. Por último, se explica como otros métodos han adaptado el diccionario de tensores y como se pueden incluir en este trabajo. El capítulo 4 muestra los experimentos realizados con datos sintéticos, los resultados obtenidos en el concurso ISBI 2015 y los resultados de la reconstrucción con los datos proporcionados en el SPARC dMRI 2014. Finalmente, el capítulo 5 cierra con las conclusiones de este trabajo así como el posible trabajo futuro.

Capítulo 2

Marco Teórico

2.1. Resonancia Magnética Nuclear

Todas las cosas están hechas de átomos, incluido el cuerpo humano. El átomo está constituido por protones con carga eléctrica positiva, neutrones sin carga eléctrica y electrones con carga eléctrica negativa; además los protones y neutrones forman parte del núcleo del átomo. Estas subpartículas que forman el núcleo presentan un movimiento, y cuando el núcleo está formado por un número impar de subpartículas (ej., más protones que neutrones) el núcleo por si mismo presenta un movimiento neto, ver figura 2.1. El hidrógeno, presente en el cuerpo humano en gran cantidad, con un protón en el núcleo es un ejemplo de este tipo de átomos.

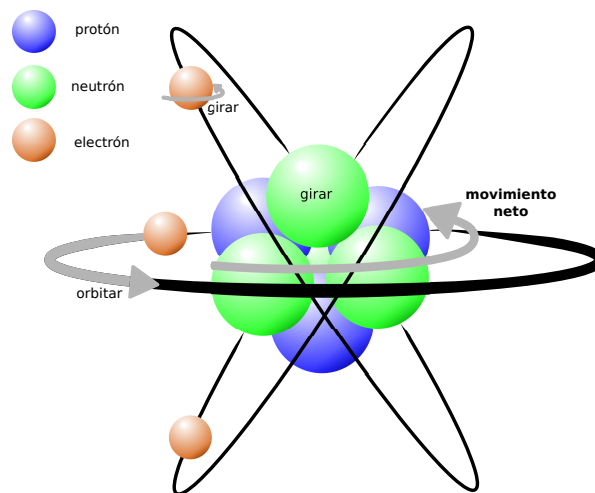


Figura 2.1: Movimientos en el átomo.

La Ley de Inducción Electromagnética de Faraday refiere a tres fuerzas individuales, movimiento, magnetismo y carga eléctrica y establece que si dos de ellas están presentes, la tercera es generada automáticamente. Por tanto, el núcleo de hidrógeno al presentar un movimiento y una carga eléctrica, automáticamente adquiere un momento magnético el cual se puede representar como un vector, figura 2.2.

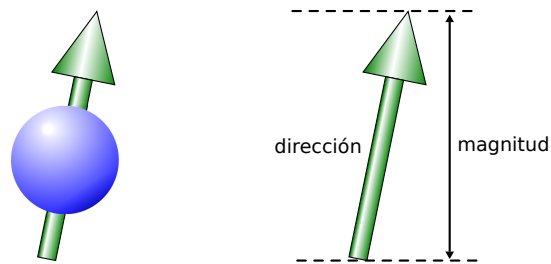


Figura 2.2: Momento Magnético.

En ausencia de un campo magnético externo, los momentos magnéticos de los núcleos de hidrógeno están orientados aleatoriamente. Pero cuando existe un campo magnético externo B_0 , los momentos magnéticos de los núcleos de hidrógeno se tienden a alinear con este campo, algunos en el mismo sentido y otros en sentido contrario al campo externo, figura 2.3.

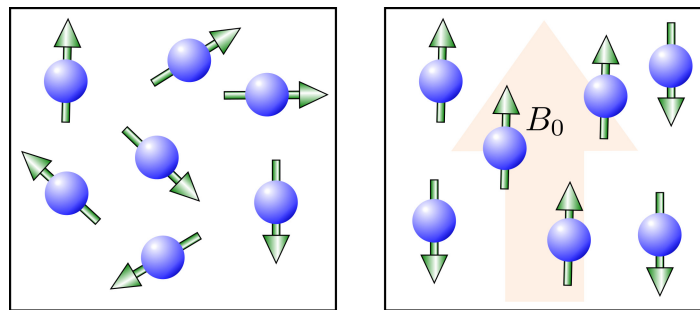


Figura 2.3: Aplicación de un campo magnético externo

Usando la teoría cuántica, los núcleos de hidrógeno se pueden clasificar como núcleos de energía baja, aquellos que alinean su momento magnético con el campo externo, y núcleos con energía alta, aquellos que se alinean en sentido contrario. En equilibrio térmico, siempre habrá más núcleos con energía baja, y estos núcleos cancelarán los momentos magnéticos de los núcleos con energía alta, produciendo así un momento magnético neto, dado por el remanente de núcleos con energía baja. Este momento magnético neto se representa como un Vector de Magnetización Neto (VMN) figura 2.4

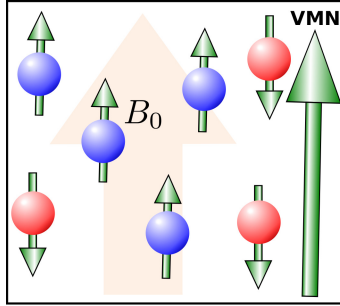


Figura 2.4: Vector de Magnetización Neto (VMN)

De igual forma, con la aplicación del campo magnético externo B_0 , el momento magnético del núcleo de hidrógeno precesa alrededor del campo externo B_0 . A este movimiento circular se le conoce como precesión y la velocidad con la que gira se le conoce como frecuencia de precesión o frecuencia de Larmor, ver figura 2.5.

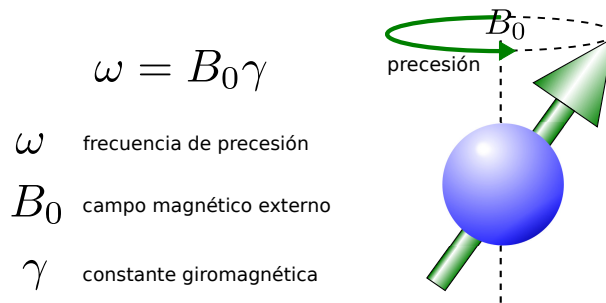


Figura 2.5: Frecuencia de precesión

Cuando un núcleo es expuesto a una perturbación (pulso de radio frecuencia) perpendicular a B_0 y con frecuencia similar a su frecuencia natural, el núcleo gana energía de esta fuerza externa y resuena si la energía es generada a la misma frecuencia de precesión. La aplicación de este pulso de radio frecuencia que genera la resonancia es llamado excitación. Un primer resultado de la excitación es que los momentos magnéticos de los núcleos estarán en fase entre ellos. Además el Vector de Magnetización Neto dependiendo de la magnitud del pulso de RF se irá volcando (ampliara su ángulo de giro o precesión), formando así un ángulo que generalmente es de 90 grados con respecto a B_0 . En la literatura se dice que con este cambio el campo externo se encuentra sobre el plano longitudinal y el Vector de Magnetización Neto sobre el plano transversal, ambos con respecto a B_0 , figura 2.6.

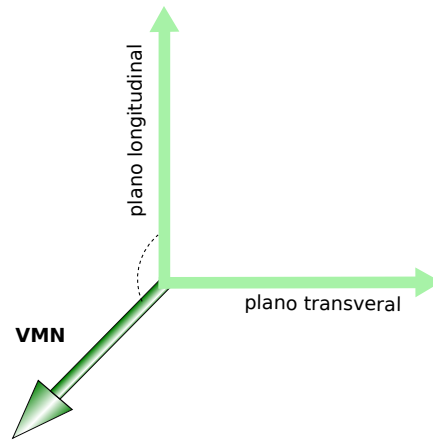


Figura 2.6: Rotación del VMN debido a la excitación

Por la Ley de Faraday, si existe una bobina receptora sobre el plano transversal se generará una señal de resonancia magnética o voltaje cuando el Vector de Magnetización Neto está sobre el plano transversal y corta una bobina receptora, figura 2.7.

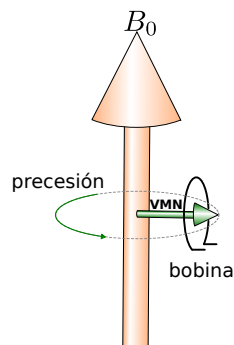


Figura 2.7: Señal de Resonancia Magnética

Lo anterior muestra de una forma muy general la teoría detrás en la obtención de una señal de resonancia magnética. Para ampliar el contenido de esta sección se puede consultar Westbrook et al. (2011). En la siguiente sección se ejemplificará con más detalle un pulso de radio frecuencia usado en resonancia magnética.

2.2. Imágenes de Resonancia Magnética Pesadas por Difusión.

Una vez comprendido el mecanismo para generar una señal por resonancia magnética, el siguiente punto es comprender cómo se puede usar la técnica de resonancia magnética para medir la difusión de las moléculas de agua. Por la sección anterior se sabe que es necesario usar un pulso de radio frecuencia, pero esto no es suficiente para medir la difusión. Stejskal and Tanner (1965) propusieron ciertas innovaciones para hacer esto posible, entre ellas propusieron el *Pulsed Gradient Spin Echo* (PGSE), el cual consiste en pulsos de radio frecuencia y gradientes de corta duración.

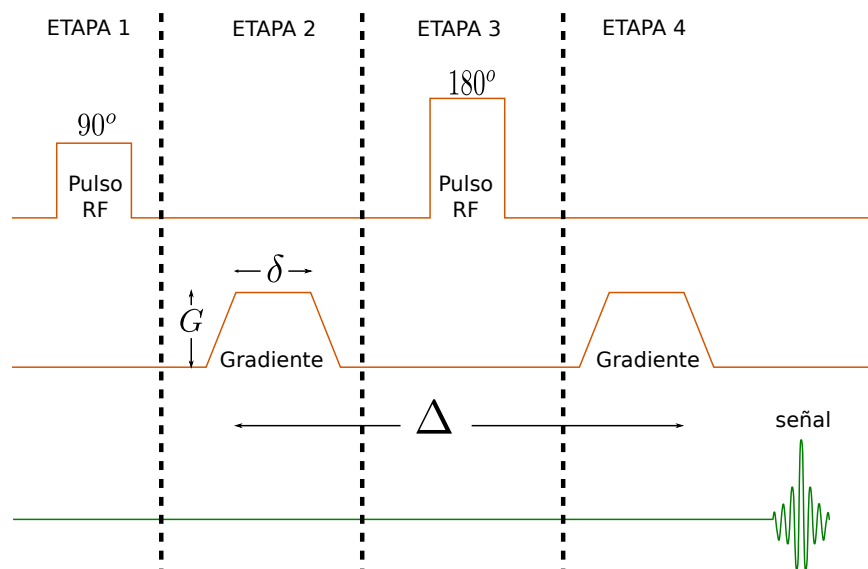


Figura 2.8: PGSE

La figura 2.8 muestra dicha secuencia, donde G es la intensidad del gradiente, δ es la duración del gradiente y Δ es intervalo entre los dos gradientes. La estrategia de este método se puede dividir en 4 etapas, las cuales se ilustran en la figura 2.9. En la etapa 1 todos los núcleos perciben el mismo campo magnético externo B_0 y con el pulso de radio frecuencia los núcleos empiezan a dar señales a la misma frecuencia. En la etapa 2 y con aplicación del primer gradiente los núcleos ahora perciben un diferente campo magnético según su posición y empiezan a resonar a diferente frecuencia. La clave de esto es que ahora cada núcleo tiene una frecuencia diferente dada a su posición, y esto sirve como una etiqueta para identificar en

que posición se encontraban. Al final de este primer gradiente, en la etapa 3 todos los núcleos perciben nuevamente el mismo campo magnético y presentan la misma frecuencia, pero con diferente fase. Finalmente en la etapa 4 se quiere regresar a la condición inicial cuando los núcleos tenían la misma fase, es por eso que se aplica un segundo gradiente pero negativo, y nuevamente cada núcleo percibirá un diferente campo magnético de acuerdo a su posición. Si los núcleos están en la misma posición, podrán regresar a la fase inicial, pero si se encuentran en otra posición no regresarán a la misma fase. Si todos los núcleos logran regresar a la misma fase se obtendrá la misma señal inicial, pero si hay núcleos que no regresaron a la fase inicial habrá una pérdida en la señal. Esta pérdida en la señal será proporcional a la difusión en la dirección de los gradientes aplicados.

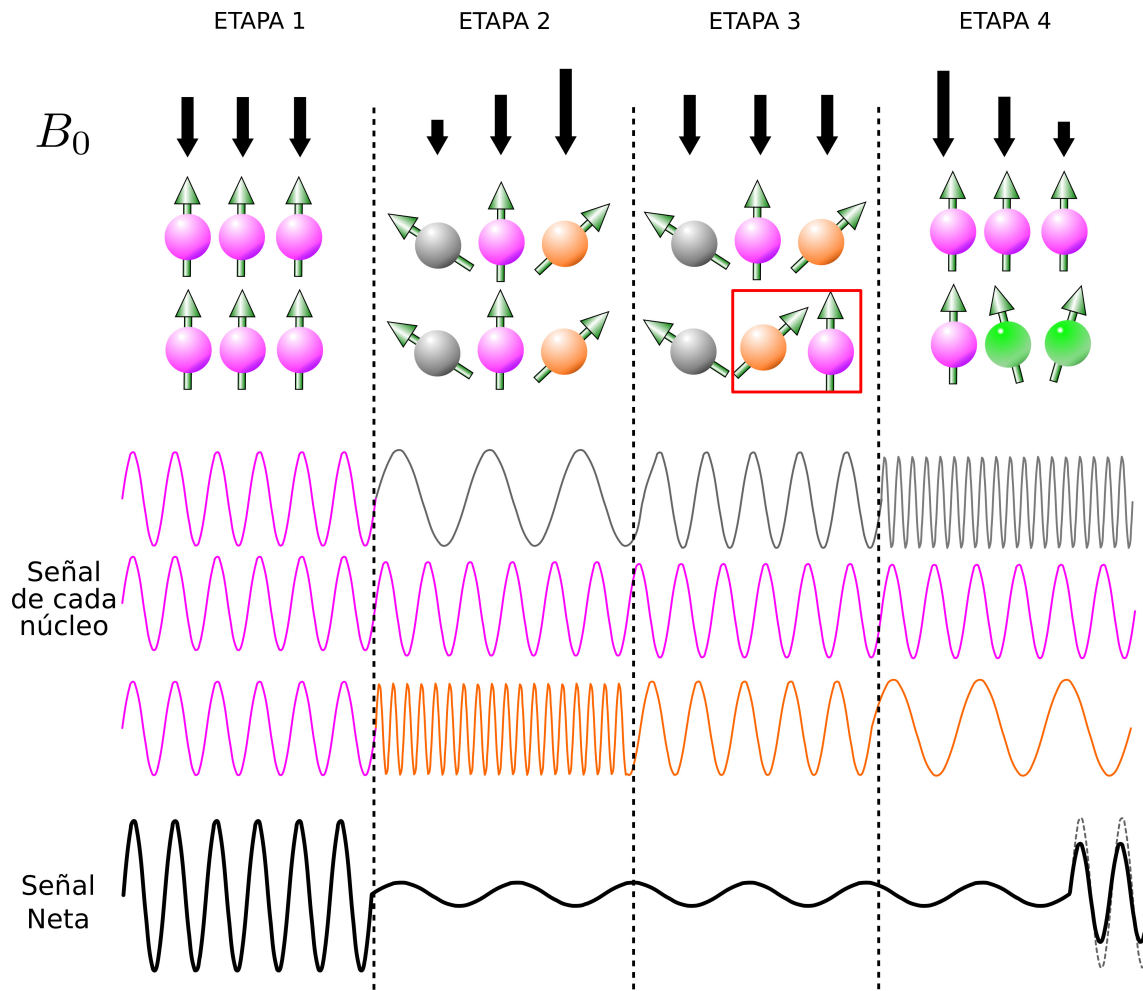


Figura 2.9: Pérdida de señal debido a difusión.

Como se describió en el experimento anterior, para calcular la señal pesada por difusión

es necesario acumular la señal de todas las partículas. La fase de cada partícula según su posición se define de la siguiente forma:

$$\phi(x) = \exp(i\gamma G(t)tx). \quad (2.1)$$

Ahora es necesario determinar la posición de las partículas. Asumiendo difusión libre, la posición de cada partícula queda definida por una Distribución Gaussiana:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right); \quad (2.2)$$

donde el parámetro σ controla que tan lejos se pueden mover en promedio las partículas, y en este caso queda definido por la ecuación de Einstein $\sigma = \sqrt{2Dt}$, donde D representa el coeficiente de difusión. Por tanto, la señal total puede ser calculada al acumular las contribuciones de las partículas en cada posición:

$$\sum_x P(x, \Delta)\phi(x). \quad (2.3)$$

En una situación real (2.3) puede suponerse continua, así que la ecuación se transforma en:

$$\int_x P(x, t)\phi(x)dx = \frac{1}{\sqrt{4\pi Dt}} \int_x \exp\left(-\frac{x^2}{4Dt}\right) \exp(i\gamma G(t)tx)dx. \quad (2.4)$$

Al resolver la ecuación anterior, la pérdida de la señal debido a la difusión queda de la siguiente forma:

$$S = S_0 \exp\left(-\gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right) D\right). \quad (2.5)$$

Para simplificar (2.5), en la literatura es común definir $b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right)$. Por tanto, el decaimiento de la señal debido a la difusión definida por Stejskal and Tanner (1965) queda de la siguiente forma:

$$S = S_0 \exp(-bD). \quad (2.6)$$

2.3. Modelo Mono-Tensorial de DW-MRI (DTI)

Con los trabajos de Stejskal and Tanner (1965) se estableció la relación (2.6), la cual puede ser utilizada para determinar el coeficiente de difusión D . Dicho coeficiente D se puede estimar ajustando un modelo de regresión lineal con una serie de mediciones de DWI. Como resultado importante de estos trabajos es que se identificó que el coeficiente de difusión D depende del ángulo entre el gradiente aplicado y el eje de las fibras, ver figura 2.10; el coeficiente es mayor cuando el gradiente es paralelo a la dirección de las fibras, y es menor cuando es perpendicular. De esta forma la *difusión anisotrópica* puede ser usada para inferir la orientación de las fibras, figura 2.10.

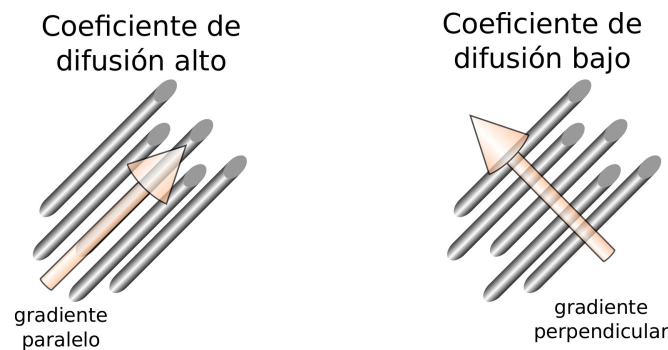


Figura 2.10: Coeficiente de difusión

Por otro lado, una desventaja del coeficiente D es que sólo aporta información del desplazamiento en una sola dirección. Basser et al. (1994) modificó (2.6) para obtener un esquema general y así caracterizar el desplazamiento de las moléculas en el espacio, a esta nueva modalidad la llamó *imágenes de tensor de difusión*, (DTI por sus siglas en inglés). La idea en DTI, ver figura 2.11, es modelar la orientación de las fibras usando un elipsoide; figura 2.11.

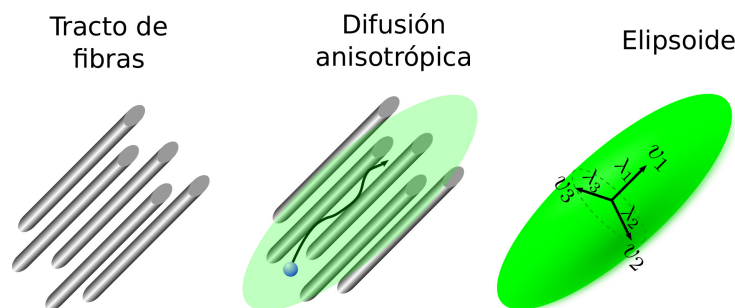


Figura 2.11: Elipsiode de difusión

Como se observa, para construir el elipsoide de difusión son necesarios seis parámetros: los valores $\lambda_1, \lambda_2, \lambda_3$ los cuales definen la forma, y los vectores v_1, v_2, v_3 los cuales definen la orientación. Estos seis parámetros se pueden resumir en un tensor $T_{3 \times 3}$ el cual se denomina tensor de difusión. El tensor de difusión y los seis parámetros se relacionan por un proceso de diagonalización:

$$T = \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{bmatrix} = \lambda_1 v_1 v_1^T + \lambda_2 v_2 v_2^T + \lambda_3 v_3 v_3^T. \quad (2.7)$$

Un punto importante, y la idea central, es que el vector propio v_1 asociado al valor propio más grande λ_1 está alineado con el eje de los tractos. Además, el tensor de difusión T es simétrico y definido positivo, esto implica que son necesarios seis parámetros para definir el elipsoide, sólo la matriz triangular superior de T . Ahora, para determinar el coeficiente de difusión D se proyecta el tensor de difusión T sobre los gradientes aplicados g , $D = g^T T g$. Por tanto (2.6) se transforma en:

$$S = S_0 \exp(-bg^T T g). \quad (2.8)$$

Para caracterizar el tensor T , es posible resolver un problema de ajuste, donde serán necesarias mínimo seis mediciones dados los seis parámetros que se quieren determinar. Para terminar de formular el problema de ajuste, se modificará (2.8):

$$\frac{1}{b} \ln \left[\frac{S_0}{S} \right] = g^T T g \quad (2.9)$$

el siguiente punto es expandir el $g^T T g$:

$$g^T T g = \begin{bmatrix} g_x & g_y & g_z \end{bmatrix} \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{bmatrix} \begin{bmatrix} g_x \\ g_y \\ g_z \end{bmatrix} \quad (2.10)$$

$$g^T T g = g_x^2 T_{xx} + g_y^2 T_{yy} + g_z^2 T_{zz} + 2g_x g_y T_{xy} + 2g_x g_z T_{xz} + 2g_y g_z T_{yz}.$$

Por tanto, para determinar completamente el tensor T son necesarias seis mediciones. Por efectos de ruido en las imágenes, se usa $n > 6$ mediciones para estimar mediante mínimos cuadrados los 6 parámetros.

$$\begin{bmatrix} g_{x1}^2 & g_{y1}^2 & g_{z1}^2 & 2g_{x1}g_{y1} & 2g_{x1}g_{z1} & 2g_{y1}g_{z1} \\ g_{x2}^2 & g_{y2}^2 & g_{z2}^2 & 2g_{x2}g_{y2} & 2g_{x2}g_{z2} & 2g_{y2}g_{z2} \\ & & & \cdot & & \\ & & & \cdot & & \\ & & & \cdot & & \\ g_{xn}^2 & g_{yn}^2 & g_{zn}^2 & 2g_{xn}g_{yn} & 2g_{xn}g_{zn} & 2g_{yn}g_{zn} \end{bmatrix} \begin{bmatrix} T_{xx} \\ T_{yy} \\ T_{zz} \\ T_{xy} \\ T_{xz} \\ T_{yz} \end{bmatrix} = \frac{1}{b} \begin{bmatrix} \ln \left[\frac{S_0}{S_1} \right] \\ \ln \left[\frac{S_0}{S_2} \right] \\ \cdot \\ \cdot \\ \cdot \\ \ln \left[\frac{S_0}{S_n} \right] \end{bmatrix} \quad (2.11)$$

El sistema (2.11) permitirá definir un Tensor, o elipsoide, que mejor se orienta al tracto de fibras definidas en el voxel. Este ajuste será adecuado cuando sólo existe un manojo de fibras en dicho voxel. El ajuste dejará de representar la realidad cuando existe un cruce de fibras, puesto que el modelo formulado por Basser et al. (1994) sólo se obtendrá un tensor promedio, y no dos o más como es deseado. La figura 2.12 ilustra esta desventaja.

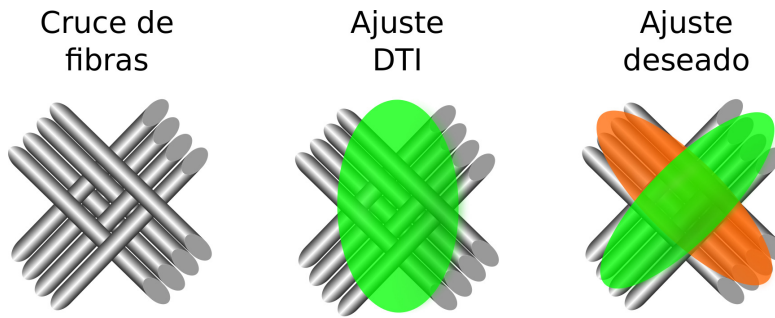


Figura 2.12: Problemas del DTI en cruce de fibras

Esta desventaja del modelo DTI fue un detonador para desarrollar nuevos modelos capaces de identificar cruce de fibras. La importancia que los tensores representen la realidad de los manojos de fibras en cada voxel, es que el ajuste tensorial sirve de entrada para otros procesos, como la tractografía, lo que permite representar en dos o tres dimensiones los tractos neuronales. En la siguiente sección se presentará un modelo multi-tensorial, el cual es la base de las propuestas desarrolladas en esta tesis.

2.4. Modelo Multi-Tensorial de DW-MRI (DBF)

Como ya se observó, una de las desventajas principales del modelo DTI es la incapacidad de describir correctamente el cruce de fibras. Para resolver este problema, Tuch et al. (2002) utilizó un modelo de mezcla de Gaussianas (GMM, por sus siglas en inglés):

$$S_i = \sum_{j=1}^m \alpha_j \exp(-b_i g_i^T T_j g_i) + \eta_i \quad (2.12)$$

donde η_i denota ruido y α_j es la fracción de volumen que aporta el tensor T_j para reconstruir la señal S_i . La no linealidad del problema, así como el incremento en el número de variables en el modelo, han representado un problema para resolverlo eficazmente.

Ramirez-Manzanares et al. (2007) presentó el modelo Funciones Base de Difusión (DBF, por sus siglas en inglés) el cual es una versión discreta del GMM. El modelo DBF busca relajar el problema al construir un conjunto de M tensores bases distribuidas uniformemente en todas las orientaciones posibles; a este conjunto de tensores bases se le conoce como diccionario. La figura 2.12 muestra un ejemplo de diccionario con 129 elementos. Métodos basados en diccionarios pueden consultarse en Ramirez-Manzanares et al. (2007); Jian and Vemuri (2007); Canales-Rodríguez et al. (2010); Ye et al. (2012); Landman et al. (2012); Yap and Shen (2012); Ghosh et al. (2013); Merlet et al. (2013); Cheng et al. (2014); Mani et al. (2015); Aranda et al. (2015); Loya and Rivera (2015).

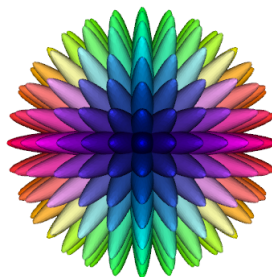


Figura 2.13: Diccionario de tensores bases

Además de fijar las orientaciones de los tensores bases, también se fijan su perfil, es decir, los valores propios del tensor. Los valores propios se pueden establecer, por ejemplo $[\lambda_1, \lambda_2, \lambda_3] = [1 \times 10^{-3}, 2 \times 10^{-4}, 2 \times 10^{-3}] \text{mm}^2/\text{s}$, o pueden ser estimados según los datos analizados. Así, cada tensor T_j del diccionario se define como en la ecuación (2.7).

Con los tensores bases T_j y los gradientes aplicados g_i , es posible construir la matriz ϕ :

$$\phi_{ij} = S_0 \exp(-b_i g_i^T T_j g_i) \quad (2.13)$$

donde cada entrada ϕ_{ij} representa una señal de difusión con gradiente g_i aplicado a la fibra simulada por el tensor T_j . Por lo tanto, DBF modela la señal basada en el siguiente modelo:

$$S_i = \sum_{j=1}^m \alpha_j \phi_{ij} + \eta_i; \quad \alpha_j \geq 0. \quad (2.14)$$

Lo que se busca en (2.14) es determinar la mejor combinación lineal de los elementos del diccionario que mejor reconstruyan la señal, ver figura 2.14. Para determinar dicha combinación lineal se formula un problema de mínimos cuadrados no negativos:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^m \alpha_j \phi_{ij} - S_i \right)^2 \\ \text{s.a.} \quad & \alpha_j \geq 0 \end{aligned} \quad (2.15)$$

donde $i = \{1, 2, \dots, n\}$ es el número de gradientes aplicados y $m = \{1, 2, \dots, m\}$ es el número de elementos del diccionario. Finalmente, el problema de optimización (2.15) conviene formularlo en forma matricial (ver figura 2.14):

$$\begin{aligned} \min_{\alpha} \quad & \|\Phi\alpha - S\|_2^2 \\ \text{s.a.} \quad & \alpha \geq 0. \end{aligned} \quad (2.16)$$

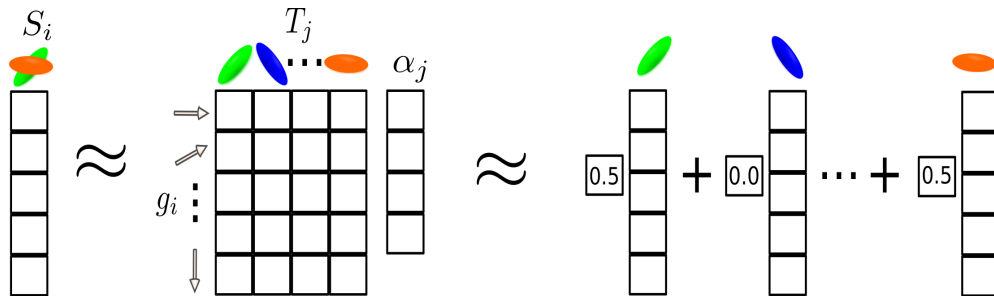


Figura 2.14: Combinación Lineal de señales base de Difusión

El método DBF formuló una versión usando la técnica de *basis pursuit* (BP) como una estrategia para exigir rareza en la combinación lineal de los elementos del diccionario. Otro

punto importante en la solución final es agrupar los tensores cuya orientación es muy cercana, puesto que pueden representar la misma fibra.

En resumen, el método DBF presentó una estrategia capaz de identificar el cruce de fibras. Al contar con un diccionario fijo se logró disminuir el número de grados de libertad, donde las únicas variables son los pesos α de la combinación lineal. El problema de optimización, formulado como un problema de Mínimos Cuadrados No Negativos, permitió usar algoritmos eficientes para resolverlo. En diferentes evaluaciones Daducci et al. (2014) y Ramirez-Manzanares et al. (2011), el método DBF ha sobresalido por obtener errores angulares competitivos, necesitar un número reducido de señales en sus ajustes y un costo computacional muy bajo.

Por otro lado, una de las desventajas del método DBF es la sobrestimación de fibras, es decir, en la solución final hay más elipsoides que manojos de fibras. La Figura 2.15 esquematiza este problema. Si bien, la técnica de *basis pursuit* y el agrupamiento final tratan de aminorar esta sobrestimación, no se ha logrado controlarlo por completo.



Figura 2.15: Problema de sobrestimación de fibras

Como se observa, en algunos casos para ser representado un manajo de fibras requiere más de un tensor; esto no es un error propiamente, es la solución óptima dada la formulación del problema. Hay que recordar que el diccionario de tensores es discreto y fijo, lo cual representa una limitación para lograr un mejor ajuste. Recientemente se han propuesto algunas estrategias para adaptar el diccionario durante las iteraciones del problema; en la siguiente sección se revisarán dichas estrategias.

2.5. Diccionario Ralos y Adaptables de Difusión (SADD)

Recientemente Aranda et al. (2015) propuso el método de Diccionario Ralos y Adaptables de Difusión (SADD, por su siglas en ingles) el cual plantea ciertas modificaciones al método DBF. Si bien el método SADD se centra en adaptar el diccionario de tensores, este método propone un algoritmo de tres etapas denominadas predictora, correctora y generadora:

- En la **etapa predictora** se resuelve el método DBF con un término de regularización ℓ_1

$$\begin{aligned} \min_{\alpha} \quad & \|\Phi\alpha - S\|_2^2 + \mu\|\alpha\|_1 \\ \text{s.a.} \quad & \alpha \geq 0 \end{aligned} \tag{2.17}$$

donde dicho término $\mu\|\alpha\|_1$ exige un solución rala, es decir, una solución con el menor número de tensores posibles. Aquellos átomos activos, $\alpha_j > 0$, serán considerados en la siguiente etapa.

- La **etapa correctora** comienza con actualizar los pesos de los átomos activos, al resolver (2.17) sin el término de regularización. Después se corrigen las orientaciones y perfiles de los tensores del diccionario asociados a los átomos activos. Un punto importante, para rotar los tensores se optimizan los parámetros de unas matrices de rotación.
- Finalmente, en la **etapa generadora** se actualiza el diccionario de tensores bases al incluir las combinaciones (pares y triadas) de los tensores modificados en etapa correctora.

Si bien la metodología SADD inicia usando una base discreta de tensores, gracias a la adaptación de los tensores bases ha logrado un mejor ajuste de la señal. Como resultado de estas mejoras, SADD ha sido capaz de estimar mejor los tensores en cruces de fibras, mejorando así el parámetro de sobrestimación de fibras (n^+). Por otro lado, SADD ha incrementado considerable el costo computacional en comparación al método DBF, esto debido al gran número de problemas de optimización a resolver durante las tres etapas.

Basándose en el método SADD, Loya and Rivera (2015) realizaron ciertas simplificaciones que permitieron reducir considerablemente el costo computacional. Esa versión de SADD fue

llamada LASADD, donde las nuevas siglas enfatizan le eficiencia computacional (Linearly Accelerated). En primer lugar se buscó una representación más sencilla del tensor base:

$$T_j = \chi_1 v_j v_j^T + \chi_2 I \quad (2.18)$$

donde $\chi_1 = \lambda_1 - \lambda_2$ y $\chi_2 = \lambda_2$. El primer término de (2.18) se le denomina *Stick Tensor* (ST) dado que físicamente representa un vector unitario v_j orientado en la dirección principal del manajo de fibras. El segundo término, denominado *Ball Tensor* (BT), representa una esfera con radio χ_2 . Al combinar ambos términos, la dirección del ST y la forma del BT, se obtiene un tensor de difusión base T_j .

El siguiente punto en la simplificación fue diseñar la corrección de los tensores bases utilizando el ST y BT de la nueva formulación (2.18). Para determinar una nueva orientación \tilde{v}_j , se propuso determinar una pequeña perturbación $\delta_j < 1$ tal que:

$$\tilde{v}_j = v_j + \delta_j. \quad (2.19)$$

Para calcular la nueva entrada del diccionario $\tilde{\phi}_{ij}$ se sustituye (2.18) y (2.19):

$$\begin{aligned} \tilde{\phi}_{ij} &= S_0 \exp\left(-b_i g_i^T \tilde{T}_j g_i\right) \\ \tilde{\phi}_{ij} &= S_0 \exp\left(-b_i g_i^T (\chi_1 (v_j + \delta_j)(v_j + \delta_j)^T + \chi_2 I) g_i\right) \\ \tilde{\phi}_{ij} &= S_0 \exp\left(-b_i g_i^T \left(\chi_1 v_j v_j^T + \chi_2 I + 2\chi_1 v_j \delta_j^T + \chi_1 \delta_j \delta_j^T\right) g_i\right) \\ \tilde{\phi}_{ij} &= S_0 \exp\left(-b_i g_i^T (\chi_1 v_j v_j^T + \chi_2 I) g_i\right) \exp\left(-b_i g_i^T (2\chi_1 v_j \delta_j^T) g_i\right) \\ \tilde{\phi}_{ij} &= \phi_{ij} \exp\left(-2b_i \chi_1 v_j^T g_i g_i^T \delta_j\right) \end{aligned} \quad (2.20)$$

en esta reformulación se observa que es posible despreciar $\chi_1 \delta_j \delta_j^T$ puesto que las perturbaciones en las orientaciones permitidas serán pequeñas, por lo tanto este producto tiende a cero.

Además, se utiliza la definición anterior de $\phi_{ij} = S_0 \exp\left(-b_i g_i^T (\chi_1 v_j v_j^T + \chi_2 I) g_i\right)$.

Una última simplificación es remplazar el exponencial hasta el segundo término en la serie de Taylor $e^x = \sum_{n \geq 0} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \dots$, entonces:

$$\tilde{\phi}_{ij} \approx \phi_{ij} \left(1 - 2b_i \chi_1 v_j^T g_i g_i^T \delta_j\right). \quad (2.21)$$

Por lo tanto, para determinar las perturbaciones en las orientaciones δ_j se deberá optimizar la siguiente función de costo:

$$\begin{aligned}
 U(\alpha, x, \delta) &= \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^m \alpha_j \tilde{\phi}_{ij} - S_i \right)^2 \\
 &= \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^m \alpha_j (\phi_{ij}(1 - 2b_i \chi_1 v_j^T g_i g_i^T \delta_j)) - S_i \right)^2 \\
 &= \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^m -\alpha_j b_i \chi_1 v_j^T g_i g_i^T \delta_j - (S_i - \sum_{j=1}^m \alpha_j \phi_{ij}) \right)^2
 \end{aligned} \tag{2.22}$$

y al definir el residual $R_i = S_i - \sum_{j=1}^m \alpha_j \phi_{ij}$ y vector renglon $U_{ij} = -\alpha_j b_i \chi_1 v_j^T g_i g_i^T$ se obtiene:

$$\begin{aligned}
 U(\alpha, x, \delta) &= \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^m U_{ij} \delta_j - R_i \right)^2 \\
 &= \frac{1}{2} \|U\Delta - R\|_2^2.
 \end{aligned} \tag{2.23}$$

Finalmente, una vez determinado las nuevas orientaciones $\tilde{v}_j = v_j + \delta_j$, se debe recordar que v_j debe permanecer como un vector unitario. Al normalizar $\tilde{v}_j = \frac{\tilde{v}_j}{\|\tilde{v}_j\|}$, el parámetro χ_1 absorberá los términos de normalización, así $\chi_1 = \chi_1 \|\tilde{v}_j\|^2$.

Con estas simplificaciones LASADD ha logrado adaptar eficientemente el diccionario de tensores; esto se observa en una reducción considerablemente de los tiempos de cómputo en comparación de la versión SADD de Aranda et al. (2015). También se hace notar que LASADD ha sido generalizado para estimar tres compartimientos: difusión intra-celular, difusión extra-celular y difusión libre.

Capítulo 3

Modelo Propuesto

En este capítulo se describirán las ideas propuestas en esta tesis. En primer lugar, y dado que la señal depende del protocolo de adquisición, se busca reformular el modelo para considerar múltiples protocolos (multi-shell) para ello se incluye un factor de escala relacionado con dicho protocolo. Otro punto importante es la rareza en la solución, por tanto se incluye un término de regularización que aproxima la norma ℓ_0 . Finalmente, se incluye en nuestra formulación la adaptación de los tensores bases propuesta en el método LASADD.

3.1. Factor de Escala para múltiples protocolos

Al igual que en el modelo DBF, se supone que la señal es una combinación lineal de los elementos del diccionario como en (2.14). Por otro lado, cada tensor base del diccionario se forma usando la definición más sencilla del método LASADD (2.18). Al sustituir explícitamente la definición de ϕ_{ij} y T_j en la función de costo (2.15) se obtiene:

$$\begin{aligned} & \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^m \alpha_j \exp(-b_i g_i^T T_j g_i) - \hat{S}_i \right)^2 \\ & \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^m \alpha_j \exp(-b_i g_i^T (\chi_1 v_j v_j^T + \chi_2 I) g_i) - \hat{S}_i \right)^2 \\ & \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^m \alpha_j \exp(-b_i \chi_1 (v_j^T g_i)^2 - b_i \chi_2 g_i^T g_i) - \hat{S}_i \right)^2 \end{aligned} \quad (3.1)$$

se define $\hat{S}_i = \frac{S_i}{S_0}$, como una señal normalizada. Al utilizar la propiedad de la función exponencial, $\exp(x + y) = \exp(x) \exp(y)$, en la última expresión de (3.1) se obtiene:

$$\frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^m \alpha_j \exp(-b_i \chi_1 (v_j^T g_i)^2) \exp(-b_i \chi_2 g_i^T g_i) - \hat{S}_i \right)^2 \quad (3.2)$$

donde la primer exponencial se define como la nueva entrada (i, j) de la matriz Φ

$$\phi_{ij} = \exp(-b_i \chi_1 (v_j^T g_i)^2) \quad (3.3)$$

mientras que la segunda exponencial se considera como un factor de escala, independiente de la dirección del tensor base v_j y relacionado únicamente con los parámetros χ_2 y b_i , es decir la difusión no orientada y el protocolo:

$$\beta_i^0 = \exp(-b_i \chi_2 g_i^T g_i). \quad (3.4)$$

Una forma intuitiva de considerar este factor de escala β se observa en la estructura de la matriz Φ , ver figura 3.1.

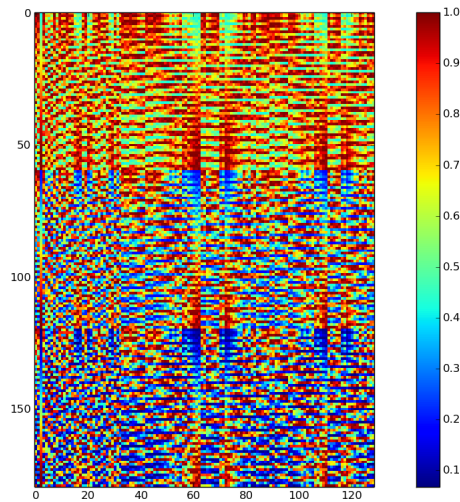


Figura 3.1: Visualización de la matriz Φ . La estructura de tres bloques observados se asocia con tres distintos protocolos de adquisición, ver texto.

La matriz Φ en la figura 3.1 fue generada con un protocolo de 180 gradientes (g_i) y 129 tensores bases (T_j). Los gradientes g_i están distribuidos en tres diferentes valores de $b_i \in \mathcal{B}$; el número posible de valores que puede tomar b es la cardinalidad de \mathcal{B} : esto es, $\#\mathcal{B} = K$.

En este caso en esta caso $\mathcal{B} = [1000, 2000, 3000]$ y $K = 3$. En la matriz Φ se observa tres regiones, las cuales están asociadas a los tres diferentes valores de b_i .

Por lo tanto, el problema de optimización con respecto a β queda definido de la siguiente forma:

$$\min_{\beta} \frac{1}{2} \sum_{i=1}^n \left(\beta_i \sum_{j=1}^m \alpha_j \phi_{ij} - \hat{S}_i \right)^2 + \mu_{\beta} \sum_{i=1}^n (\beta_i^0 - \beta_i)^2 \quad (3.5)$$

donde el término asociado a μ_{β} evitará un sobre-ajuste en los valores de β_i al penalizar la diferencia con sus valores iniciales β_i^0 ; dichos valores iniciales se calculan con (3.4). El problema (3.5) se puede expresar en forma matricial:

$$\min_B \frac{1}{2} \|B\Phi\alpha - \hat{S}\|_2^2 + \mu_{\beta} \|B^0 - B\|_2^2 \quad (3.6)$$

donde B es una matriz diagonal que contiene los coeficientes β_i agrupados en K grupos, donde K representa el número de valores diferentes de b_i . Por tanto, conviene definir los conjuntos $\Omega_k = \{\beta_i | b_i = \mathcal{B}_i\}$, y \mathcal{B}_i denota el i -ésimo elemento de \mathcal{B}

Al resolver (3.6) se obtiene la siguiente fórmula cerrada para cada coeficiente β_i :

$$\beta_i = \frac{[S]_i [\Phi\alpha]_i + \mu_{\beta} [B]_{ii}}{[\Phi\alpha]_i^2 + \mu_{\beta}} \quad (3.7)$$

donde $[.]_i / [.]_{ii}$ representa a la entrada i /ii del vector/matriz correspondiente. Con los diferentes valores de β_i calculados con (3.7), se obtendrá un promedio para cada conjunto Ω_k

$$\beta_{\Omega_k} = \frac{\sum_{i|\beta_i \in \Omega_k} \beta_i}{\#\Omega_k} \quad (3.8)$$

A cada β_i se le asigna el promedio del grupo al que pertenece, $\beta_i \leftarrow \beta_{\Omega_k}, \forall \beta_i \in \Omega_k$. De esta forma, la diagonal de la matriz B tendrá K valores diferentes, lo cual permitirá que B *escale* los renglones de Φ según el número de grupos diferentes de b_i .

Al finalizar el problema, de (3.4) se determina el valor óptimo de χ_2 , como un promedio usando los valores óptimos de β_i

$$\chi_2 = \frac{1}{n} \sum_{i=1}^n \frac{\ln(\beta_i)}{-b_i g_i^T g_i} \quad (3.9)$$

3.2. Término de Regularización

La metodología SADD de Aranda et al. (2015) introdujo un término de regularización basado en la norma ℓ_1 . La raleza que se busca en la solución es una característica heredada del esquema *Compressive Sensing*, es decir, elegir el menor número de átomos de un diccionario para representar de una forma óptima una señal dada. En otras palabras, lo que se busca es reducir la cardinalidad en la solución óptima.

Formular el problema de cardinalidad como un problema de optimización resulta en un problema NP-hard, dada la naturaleza combinatoria del problema. La norma ℓ_1 ha sido usada como una relajación en la reducción de esta cardinalidad, definida como:

$$\|x\|_1 = \sum_i |x_i| \quad (3.10)$$

donde el valor absoluto puede omitirse dado los valores no negativos de las fracciones de volumen α_j . Como se observa en (3.10), sumar los valores de las variables es una forma indirecta de disminuir la cardinalidad en la solución. Por otro lado, la norma ℓ_0 definida como:

$$\|x\|_0 = \#\{i|x_i \neq 0\} \quad (3.11)$$

permite realmente penalizar la cardinalidad en la solución óptima. La figura 3.2 muestra el comportamiento de estas normas:

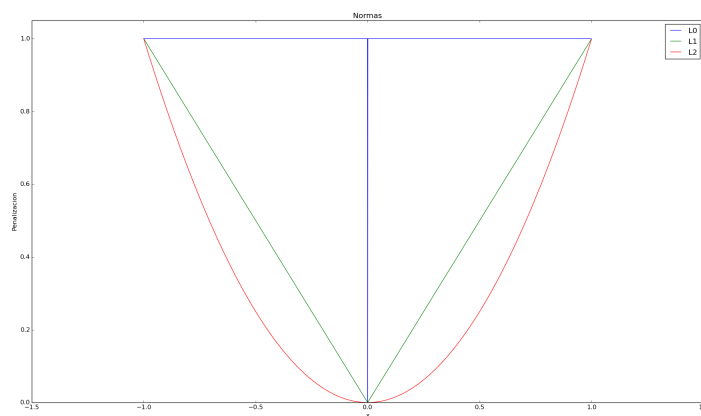


Figura 3.2: Comportamiento de las Normas ℓ_0 , ℓ_1 y ℓ_2

Si bien la norma ℓ_0 penaliza directamente la cardinalidad de la solución óptima, minimizar esta norma también es un problema de los denominados NP-duros. La estrategia que se propone para aproximar la norma ℓ_0 es mediante el siguiente potencial:

$$\rho(\alpha) = \sum_i 1 - \exp(-\mu_\rho |\alpha_i|) \quad (3.12)$$

de la misma forma, el valor absoluto en (3.12) puede omitirse dado los valores no negativos de las fracciones de volumen α_j . El término de regularización $\rho(x)$ tiende a la norma ℓ_0 cuando el valor del coeficiente $\mu_\rho \rightarrow \infty$, ver figura 3.3:

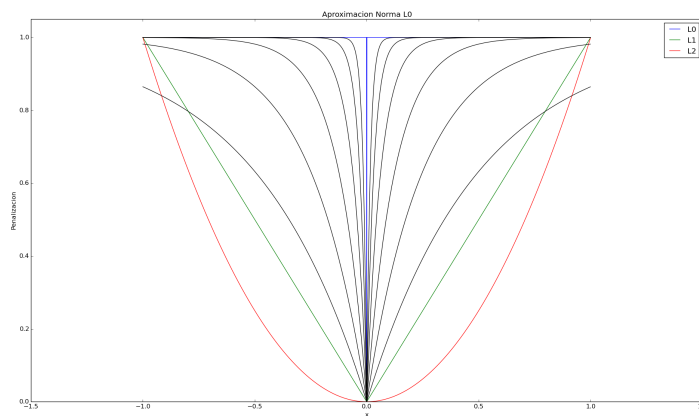


Figura 3.3: Aproximación de la norma ℓ_0

Con este nuevo término de regularización se puede formular el siguiente problema de optimización:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \|B\Phi\alpha - \hat{S}\|_2^2 + \mu_\alpha \rho(\alpha) \\ \text{s.a.} \quad & \alpha \geq 0. \end{aligned} \quad (3.13)$$

donde el término $\rho(\alpha)$ se define como en (3.12), donde la restricción de no negatividad sobre α permite obviar el valor absoluto en (3.12) y facilita la optimización al hacer el potencial continuo. Para resolver (3.13) es posible utilizar una estrategia iterativa como Gauss-Seidel Proyectado (GSP), dado que el problema de optimización (3.13) se transforma en resolver un sistema $Ax = b$, con A simétrica y definida positiva, lo anterior es una condición de convergencia para GSP, Datta (2010). Es importante mencionar que el coeficiente μ_ρ deberá incrementarse paulatinamente durante las iteraciones.

3.3. Adaptación de Tensores

Como se describió anteriormente, los métodos SADD y LASADD implementaron con éxito la adaptación de los tensores bases T_j del diccionario. Esta adaptación de los tensores bases permite que el diccionario deje de ser una discretización fija, para convertirse en una discretización dinámica, adaptándose a las condiciones de cada problema y así lograr un mejor ajuste.

Para incluir esta propiedad dinámica del diccionario nos basamos en la estrategia usada por la versión implementada en el método LASADD, la cual calcula una perturbación δ_j para adaptar las orientaciones. Al incluir esta perturbación en (3.3) se obtiene:

$$\begin{aligned}
\phi_{ij} &= \exp(-b_i \chi_1 (v_j^T g_i)^2) \\
\tilde{\phi}_{ij} &= \exp(-b_i \chi_1 (v_j + \delta_j)^T g_i g_i^T (v_j + \delta_j)) \\
&= \exp\left(-b_i \chi_1 (v_j^T g_i g_i^T v_j + 2v_j^T g_i g_i^T \delta_j + \delta_j^T g_i g_i^T \delta_j)\right) \\
&\approx \exp(-b_i \chi_1 v_j^T g_i g_i^T v_j) \exp(-2b_i \chi_1 v_j^T g_i g_i^T \delta_j) \\
&\approx \phi_{ij} (1 - 2b_i \chi_1 v_j^T g_i g_i^T \delta_j)
\end{aligned} \tag{3.14}$$

de igual forma, el término cuadrático $\delta_j^T g_i g_i^T \delta_j$ se desprecia al permitir sólo perturbaciones pequeñas. Al sustituir la aproximación final de $\tilde{\phi}_{ij}$ (3.14) en la función de costo se obtiene:

$$\begin{aligned}
&\sum_{i \in I} \left[B_i \sum_{j \in J} \alpha_j \tilde{\phi}_{ij} - \hat{S}_i \right]^2 \\
&= \sum_{i \in I} \left[B_i \sum_{j \in J} \alpha_j \phi_{ij} (1 - 2b_i \chi_1 v_j^T g_i g_i^T \delta_j) - \hat{S}_i \right]^2 \\
&= \sum_{i \in I} \left[B_i \sum_{j \in J} \alpha_j \phi_{ij} (1 - 2b_i \chi_1 v_j^T g_i g_i^x \delta_j^x - 2b_i \chi_1 v_j^T g_i g_i^y \delta_j^y - 2b_i \chi_1 v_j^T g_i g_i^z \delta_j^z) - \hat{S}_i \right]^2.
\end{aligned} \tag{3.15}$$

La última expresión de (3.15) se reformula al definir el residual $R_i = \hat{S}_i - B_i \sum_{j \in J} \alpha_j \phi_{ij}$, y $U_{ij}^* = -2b_i \chi_1 v_j^T g_i g_i^* \alpha_j \phi_{ij}$, donde $[.]^*$ se refiere a cada una de las tres dimensiones $[x, y, z]$, obteniéndose:

$$\sum_{i \in I} \left[B_i \sum_{j \in J} (U_{ij}^x \delta_j^x + U_{ij}^y \delta_j^y + U_{ij}^z \delta_j^z) - \hat{R}_i \right]^2. \quad (3.16)$$

La version matricial usando la aproximación (3.16) se define como:

$$\begin{aligned} \min_{\delta} \quad & \|BU\delta - R\|_2^2 \\ \text{s.t.} \quad & \|\delta_j\| \leq \tau \end{aligned} \quad (3.17)$$

donde $U = \begin{bmatrix} U^x & U^y & U^z \end{bmatrix}$, y el vector de perturbaciones $\delta = \begin{bmatrix} \delta^x & \delta^y & \delta^z \end{bmatrix}^T$. Por las supuestos hechos en la formulación, cada vector de perturbación δ_j deberá representar un magnitud pequeña, no mayor un ϵ definido (en este trabajo se utilizó un $\epsilon = 0,13962$, el equivalente de 8 grados a radianes). Para resolver (3.17) también es posible utilizar un estrategia iterativa como Gauss-Seidel Projectado.

Al igual que en el método LASADD, el nuevo vector \tilde{v}_j debe permanecer como un vector unitario. Al normalizar $\tilde{v}_j = \frac{\tilde{v}_j}{\|\tilde{v}_j\|}$, el parámetro χ_1 absorberá el término de normalización, así $\chi_1 = \chi_1 \|\tilde{v}_j\|^2$. De esta forma, el parámetro χ_1 también se actualizará durante las iteraciones del problema.

Para cerrar esta sección, es interesante observar que la última expresión en (3.15) se puede considerar como un solo modelo, que busca optimizar los valores de α_j , β_i y δ_j :

$$\sum_{i \in I} \left[B_i \sum_{j \in J} \alpha_j \phi_{ij} (1 - 2b_i \chi_1 v_j^T g_i g_i^x \delta_j^x - 2b_i \chi_1 v_j^T g_i g_i^y \delta_j^y - 2b_i \chi_1 v_j^T g_i g_i^z \delta_j^z) - \hat{S}_i \right]^2, \quad (3.18)$$

esta formulación representa el caso general; cuando $[\delta_j^x, \delta_j^y, \delta_j^z] = [0, 0, 0]$ se obtiene el modelo más sencillo. La formulación matricial de (3.18) se define como:

$$\begin{aligned} \min_{\delta} \quad & \|B\Phi \odot U\alpha - S\|_2^2 + \mu_\alpha \rho(\alpha) + \mu_\beta \|B^0 - B\|_2^2 \\ \text{s.t.} \quad & \|\delta_j\| \leq \tau \\ & 1^T \alpha = 1 \\ & \alpha \geq 0 \end{aligned} \quad (3.19)$$

donde \odot denota el producto elemento a elemento (llamado también producto de Hadamard)

o de Shur) y la matriz $U = [1|U_x|U_y|U_z]$, la cual representa la composición de las matrices $U_* = -2b_i\chi_1v_j^Tg_i g_i^x \delta_j^*$. En la práctica, esto implicaría un sólo modelo que se resuelve por bloques, y no, varios modelos resueltos en una serie de pasos.

3.4. Algoritmo Propuesto

En seguida presentamos el algoritmo propuesto de versión SADD para múltiples protocolos. Al igual que en el método SADD, el algoritmo propuesto utiliza un estrategia iterativa dividida en 3 etapas:

- **PASO 1.** En esta etapa predictora, se resuelve la versión regularizada (3.13); para iniciar α en el problema de optimización, se usan los valores de la ultima iteración. El objetivo de esta etapa es determinar los elementos del diccionario que contribuyen *significativamente* en la solución, es decir $\alpha_j \geq \epsilon$. Al conjunto de índices cuyos átomos están activos se le denomina conjunto activo $J = \{j|\alpha_j > \epsilon\}$.
- **PASO 2.** Esta etapa consiste en resolver (3.17) para adaptar los tensores bases que pertenecen al conjunto activo. Al resolver este problema de optimización, es posible construir una matriz y vector temporal que contengan sólo los coeficientes relacionados con los átomos activos. Una vez determinadas las perturbaciones δ óptimas, se deberá actualizar las matrices Φ y U .
- **PASO 3.** Finalmente, se actualizarán las fracciones de volumen de los átomos activos α_j al resolver una versión no regularizada de (3.13), al usar $\mu_\alpha = 0$. De igual forma, se actualizarán los factores de escala β_j al resolver (3.5).

Esta estrategia de tres pasos se itera hasta convergencia; esto es cuando la solución no cambie significativamente: $\|\alpha^{k+1} - \alpha^k\|_\infty \leq \epsilon$. Un punto importante es modificar los coeficientes de los términos de regularización; se recomienda incrementarlos paulatinamente, un 5% en cada iteración. Estos pasos descritos se resumen en el **Algoritmo 1**.

Algorithm 1 Algoritmo Propuesto

- 1: Definir $\chi_1 = 0,0009$, $\chi_2 = 0,0001$, $\mu_\alpha = 0,004$, $\mu_\beta = 0,02$, $k = 2,0$
- 2: Inicializar α^0 , B^0 , Φ^0 , U^0 y $J^0 = \{\emptyset\}$ (ver fórmulas para inicializar)
- 3: **for** $k = 0, 1, 2, \dots$

PASO 1. Determinar el conjunto de átomos activos, $J^{k+1} = \{j | \alpha_j > \epsilon\}$.

$$\alpha = \arg \min_{\alpha} \|B^k \Phi^k \alpha - S\|_2^2 + \mu_\alpha \rho(\alpha) \quad \text{s.t. } \alpha \geq 0$$

PASO 2. Adaptar los tensores de los átomos activos α_j , $\forall j \in J^{k+1}$.

$$\delta^{k+1} = \arg \min_{\delta} \|B^k U^k \delta - R^k\|_2^2 \quad \text{s.t. } \|\delta_j\| \leq \tau$$

Actualizar los diccionarios Φ^{k+1} y U^{k+1}

PASO 3. Calcular las fracciones de volumen de los átomos activos α_j , $\forall j \in J^{k+1}$ y factores de escala.

$$\alpha^{k+1} = \arg \min_{\alpha} \|B^k \Phi^{k+1} \alpha - S\|_2^2 \quad \text{s.t. } \mathbf{1}^T \alpha = 1, \alpha \geq 0$$

$$B^{k+1} = \arg \min_B \|B \Phi^{k+1} \alpha - S\|_2^2 + \mu_\beta \|B^0 - B\|_2^2$$

Verificar convergencia

Actualizar parámetros $\epsilon = 1,05\epsilon$ $\mu_\alpha = 1,05\mu_\alpha$ $\mu_\beta = 1,05\mu_\beta$ $\mu_\rho = 1,05\mu_\rho$

end for

Capítulo 4

Experimentos

En el presente capítulo se muestran una serie de experimentos para validar el desempeño de nuestro algoritmo. En primer lugar se describen las métricas utilizadas. A continuación se muestra un primer experimento al comparar nuestro algoritmo contra el modelo DBF. Como segundo experimento se compra el comportamiento de nuestro algoritmo con un protocolo *multi-shell* con mayor número de *shells*. Como tercer experimento se muestran los resultados del ISBI-WMMC 2015, al comparar los datos enviados con nuestro modelo simple (DBF con multi-shell) contra nuestro modelo completo (SAAD con multishell). Finalmente se muestran los resultados de la reconstrucción con los datos proporcionados en el SPARC dMRI 2014.

4.1. Métricas

La métricas más usadas para evaluar el desempeño de los modelos se describen en la literatura como Daducci et al. (2014) y Ramirez-Manzanares et al. (2011). A continuación se describen las métricas implementadas para esta tesis:

1. ***Error de Reconstrucción (ER)*** . Esta métrica representa la eficacia de un método para reconstruir la señal S ; es decir, el mínimo de la función alcanzado en el ajuste:

$$ER = \frac{1}{2} \|S - \hat{S}\|^2 \quad (4.1)$$

donde \hat{S} representa la señal reconstruida.

2. **Tasa de Éxito (SR-Success Rate)**. Esta métrica representa la proporción de vóxeles que estiman correctamente el número de fibras. Para detallar mejor esta métrica, es importante diferenciar si se estimó un menor número de manojos de fibras, **underestimate** (n^-), o un mayor número de fibras, **overestimate** (n^+).
3. **Error Angular** (ϵ_θ) . Esta métrica representa el promedio de los errores en el emparejamiento de las orientaciones entre las fibras verdaderas y las fibras estimadas. El ϵ_θ entre un par de fibras se define como:

$$\epsilon_\theta = \frac{180}{\pi} \arccos(|v_j \cdot \hat{v}_j|) \quad (4.2)$$

donde v_j representa la dirección principal de difusión verdadera (PDD por sus siglas en inglés), mientras que \hat{v}_j es la PDD estimada. Es importante recalcar que se debe probar exhaustivamente todas las combinaciones en el emparejamiento entre los dos conjuntos de fibras, reales y estimadas, y así obtener el menor error posible.

4. **Error de Compartimiento** (ϵ_α) Este métrica representa el promedio de las diferencias absolutas en el emparejamiento entre los pesos o contribuciones de las fibras reales contra las pesos de las fibras estimadas.
5. **Error de Anisotropía Fraccional** (ϵ_{FA}) . Esta métrica representa el promedio de las diferencias absolutas en el emparejamiento entre los perfiles de difusión, reales y estimados. En este caso, el perfil de difusión de un tensor se representa por el escalar *Fractional Anisotropy (FA)*, definido como:

$$FA = \sqrt{\frac{1}{2} \frac{(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_2)^3 + (\lambda_2 - \lambda_3)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (4.3)$$

Para comprender más la razón de un alto ϵ_{FA} , se implementó el promedio de los error relativos en el emparejamiento entre valores propios de los tensores, reales y estimados.

$$\epsilon_{\lambda_i} = \frac{|\lambda_i - \hat{\lambda}_i|}{\lambda_i}; \forall i \in \{1, 2\} \quad (4.4)$$

donde $\hat{\lambda}_i$ representa el valor propio estimado.

4.2. Datos Sintéticos

Para este experimento se utilizaron los datos sintéticos descritos en Daducci et al. (2014). Cada imagen de MR-DWI contiene 910 vóxeles sin estructura espacial, 10 contiene una fibra y 900 dos fibras con diferente ángulo de cruce; la figura 4.1 muestra la distribución de dichos ángulos. Cada esquema de adquisición es multi-shell HARDI, referida como CUSP65. Un esquema consta de 65 imágenes, 5 con $b = 0s/mm^2$ y un par de 30 con dos valores de b . Para cada esquema se generaron 6 versiones con diferentes valores de $SNR = \{10,20,30,40,50,free\}$, donde el término free se refiere a un valor relativamente alto de SNR el cual se considera libre de ruido. En este **primer experimento (Exp01)** se compara el desempeño de nuestro algoritmo contra el método DBF. Se utilizó el esquema CUSP65-2025 el cual contiene los valores de $b = \{2000,2500\}s/mm^2$. Los parámetros para cada algoritmo se describen a continuación:

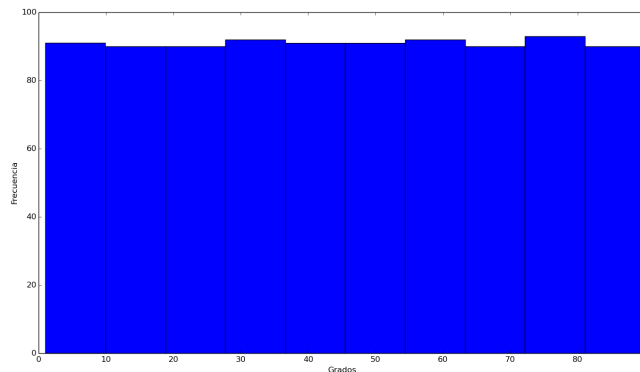


Figura 4.1: Distribución de ángulos de cruce de datos sintéticos.

- **DBF***. Se utilizó una implementación propia en Python-Cython, en la cual se incluyó los pasos predictor y corrector descritos en SADD, así como un término de regularización ℓ_1 , con un peso de $\mu_\alpha = 0,5$. Se utilizó una base de tensores de 129 átomos, distribuidos uniformemente, con un perfil de $\lambda_1 = 0,0015$ y $\lambda_2 = 0,0005$.
- **MS-SADD**. El algoritmo propuesto también se implementó en Python-Cython. Se utilizó una base de tensores de 129 átomos, distribuidos uniformemente, con un perfil de $\lambda_1 = 0,0015$ y $\lambda_2 = 0,0005$. Los pesos para los términos de regularización iniciales se establecieron como $\mu_\alpha = 0,015$ y $\mu_B = 0,0005$.

Algorithm 2 DBF*

-
- 1: Definir $\lambda_1 = 0,001$, $\lambda_2 = 0,0001$, $\epsilon = 0,2$, $\mu_\alpha = 0,5$
 - 2: Inicializar α^0 , $J^0 = \{\emptyset\}$ y Φ
 - 3: **for** $k = 0, 1, 2, \dots$

PASO PREDICTOR. Determinar el conjunto de átomos activos, $J^{k+1} = \{j | \alpha_j > \epsilon \max(\alpha_j)\}$.

$$\alpha = \arg \min_{\alpha} \|\Phi\alpha - S\|_2^2 + \mu_\alpha \|\alpha\|_1 \quad \text{s.t. } \alpha \geq 0$$

PASO CORRECTOR. Corregir el peso de los átomos activos α_j , $\forall j \in J^{k+1}$.

$$\alpha = \arg \min_{\alpha} \|\Phi\alpha - S\|_2^2 \quad \text{s.t. } \alpha \geq 0$$

Verificar convergencia

end for

En la figura 4.2 se observa que ambos algoritmos presentan un comportamiento muy similar.

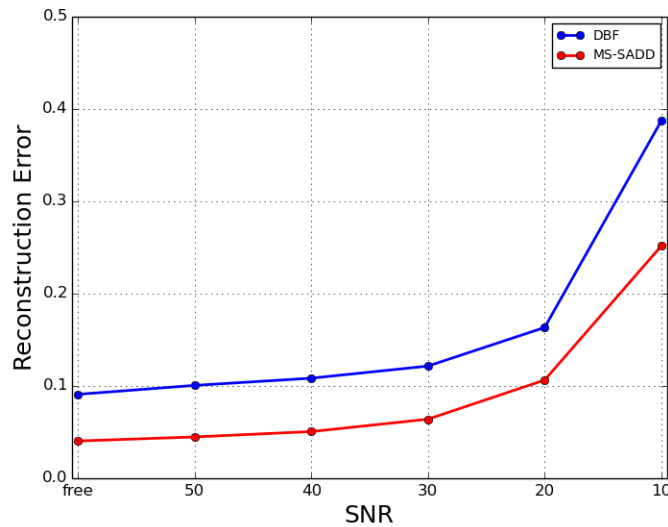


Figura 4.2: Error de Reconstrucción para Exp01.

La figura 4.3 muestra que nuestro algoritmo logra determinar mejor el número de fibras en cada vóxel. El principal problema se observa en una elevada proporción en *underestimation*.

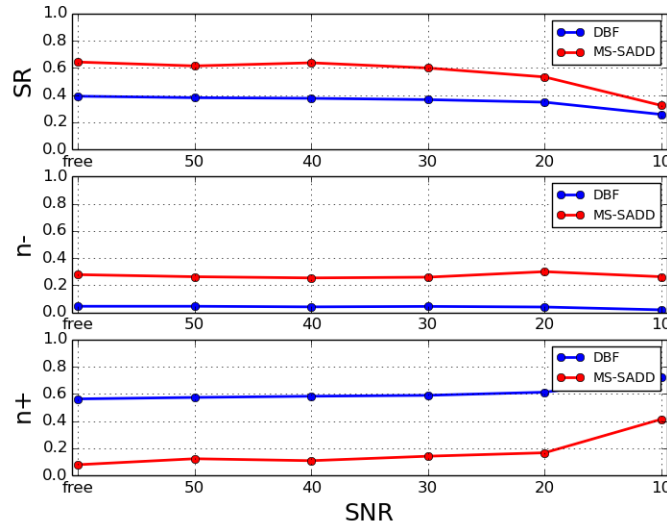


Figura 4.3: Tasa de Éxito para Exp01.

Si se observa la figura 4.1, los cruces entre 0-30 grados representan cerca del 30 % de los datos, y estos cruces son difícil de diferenciar en un algoritmo. Para ahondar más en la métrica anterior, la figura 4.4 muestra los cambios en SR al considerar sólo el caso de $SNR = \{30\}$ y admitir ángulos de cruce cada vez más grandes.

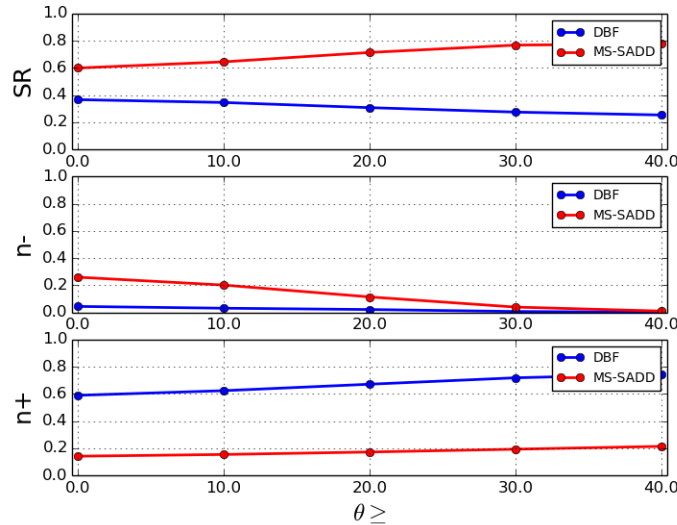


Figura 4.4: Tasa de Éxito para Exp01 para $SNR = \{30\}$ permitiendo ciertos ángulos.

La figura 4.5 muestra que nuestra propuesta obtiene una menor media en error angular. Para el caso en $SNR = 10$, nuestra propuesta presenta un peor comportamiento.

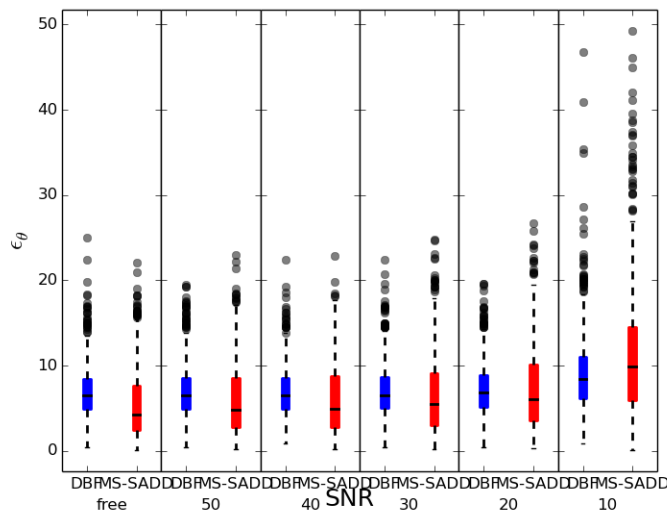


Figura 4.5: Error Angular para Exp01.

Si bien en la figura 4.6 se observa que nuestro algoritmo logra determinar mejor los pesos α , para el caso de $SNR = 10$ el método DBF^+ se comporta mejor. Tal vez cuando la señal es muy ruidosa no es recomendable usar algoritmos que requieran optimizar varios parámetros.

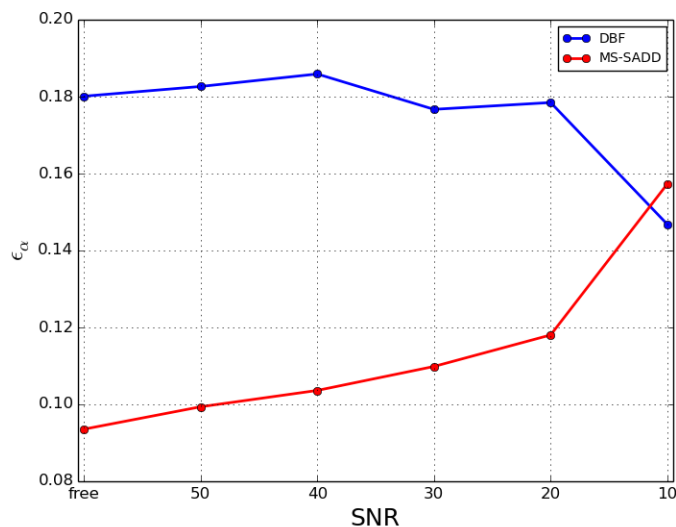


Figura 4.6: Error de Compartimiento para Exp01.

Finalmente en la figura 4.7 se observa que nuestra propuesta obtiene un menor error en la estimación de FA (4.3). Por otro lado, el error ϵ_{λ_2} es más alto en comparación que ϵ_{λ_1} , esto puede deberse a que el parámetro λ_2 se estima a partir de los factores de escala β .

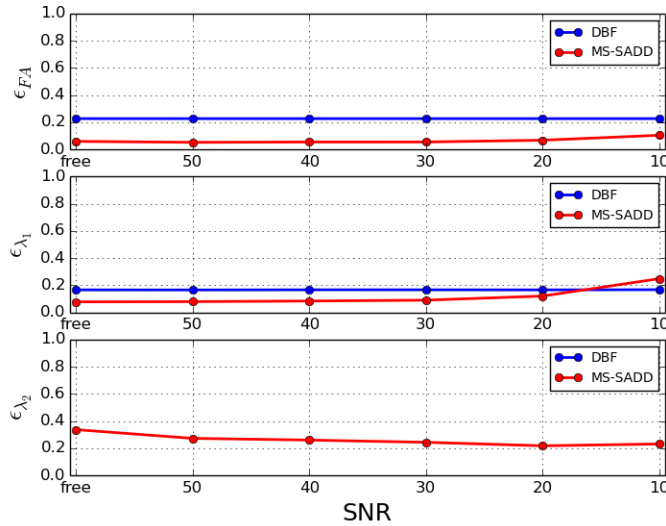


Figura 4.7: Error de Anisotropía Fraccional para Exp01.

En general, se observa que nuestro algoritmo presenta un mejor rendimiento comparado con el método DBF^+ . Uno de los puntos especiales de nuestra propuesta, es la forma de usar los factores de escala β cuando se trabaja con un esquema multi-shell. Por tal motivo, un **segundo experimento (Exp02)** consiste en comparar únicamente el comportamiento de nuestro algoritmo cuando se trabaja con más valores de b .

En este segundo experimento se utilizaron los mismos datos sintéticos CUSP65 y se formaron dos protocolos. El primer esquema consta de 130 imágenes, 10 con $b = 0$ y 120 divididas en $b = \{2000, 2500\}$, 60 imágenes para cada valor de b . El segundo esquema también consta de 130 imágenes, 10 con $b = 0$ y 120 divididas en $b = \{1500, 2000, 2500, 3000\}$, 30 imágenes para cada valor de b . De igual forma, en cada esquema existen 6 versiones con diferentes valores de $SNR = \{10, 20, 30, 40, 50, free\}$. Se utilizan los mismos parámetros descritos en el primer experimento para ejecutar nuestro algoritmo.

En las siguientes gráficas, **MS-1** se refiere a los resultados obtenidos para el esquema con $b = \{2000, 2500\}$, mientras que **MS-2** hace referencia al esquema con $b = \{1500, 2000, 2500, 3000\}$.

Un primer resultado es que nuestro algoritmo logra un mejor ajuste cuando se utilizan más valores de b . Para valores altos de ruido no existe una diferencia significativa en el ajuste de la señal.

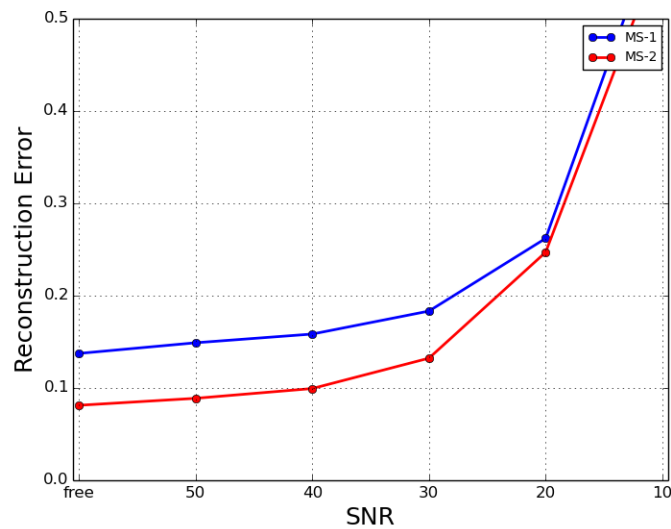


Figura 4.8: Error de Reconstrucción para Exp02.

Si bien se logró una mejor ajuste en la señal usando **MS-2** respecto a **MS-1**, esto no repercutió en una incremento significativo en la métrica SR; solo una ligera ventaja para valores pequeños de ruido.

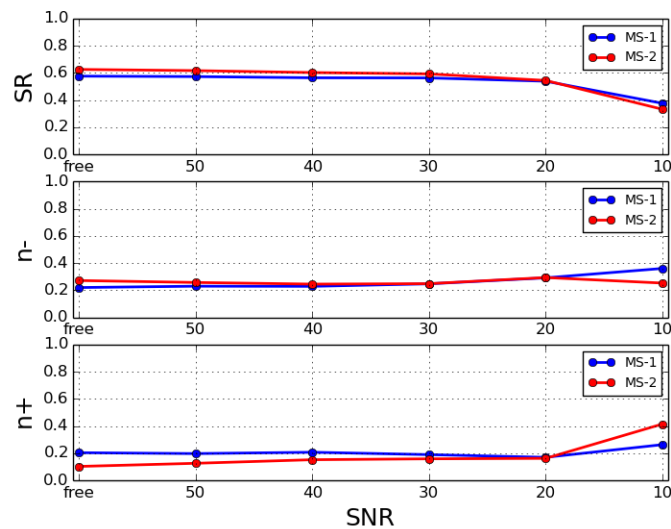


Figura 4.9: Tasa de Éxito para Exp02.

De igual forma, en la gráfica 4.10 se observan los cambios en SR al elegir solo el caso de $SNR = \{30\}$ y considerar ángulos de cruce cada vez más grandes.

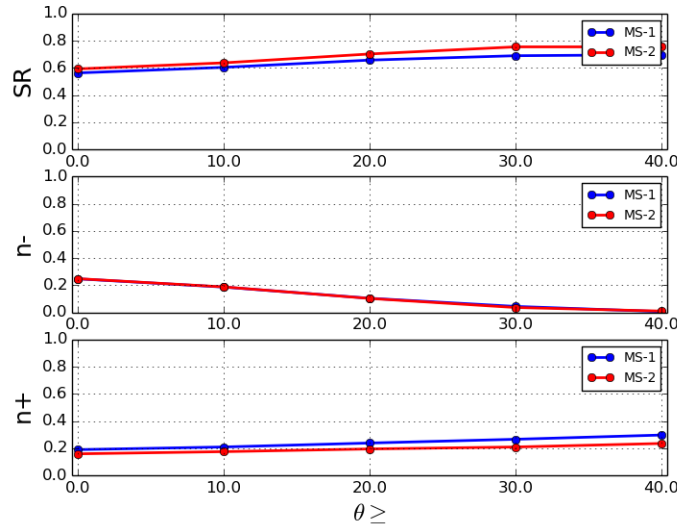


Figura 4.10: Tasa de Éxito para Exp02 para $SNR = \{30\}$ permitiendo ciertos ángulos.

Los errores angulares presentan un comportamiento muy similar, en los primeros tres casos $SNR = free, 50, 40$, se mejoran ligeramente los resultados al usar un esquema multi-shell más variado.

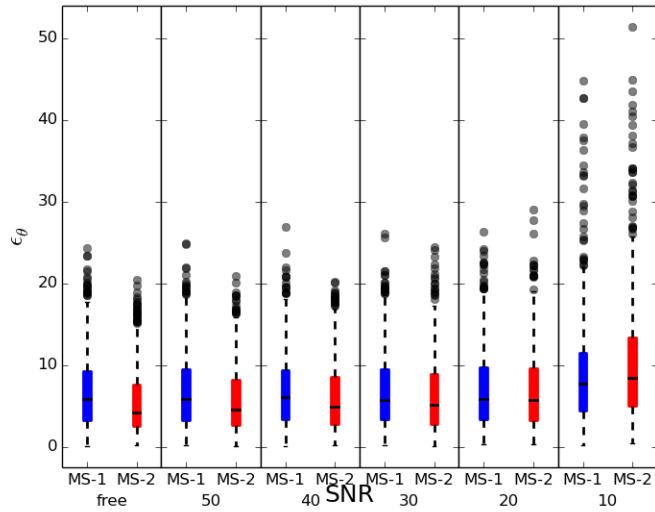


Figura 4.11: Error Angular para Exp02.

En el caso del ϵ_α se observa que mejor rendimiento en el caso un multi-shell más variado. Tal vez, este es el parámetro que explica la mejora en el ajuste de la señal.

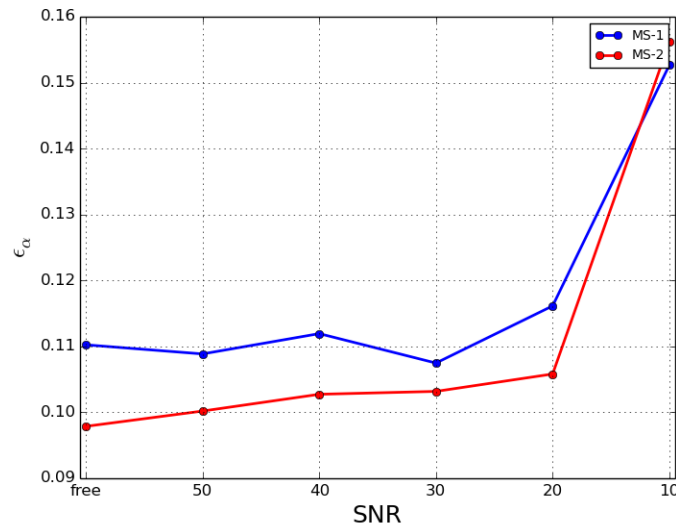


Figura 4.12: Error de Compartimiento para Exp02.

Finalmente, el perfil de los tensores ajustados no tuvo un cambio significativo para el nuevo esquema multi-shell.

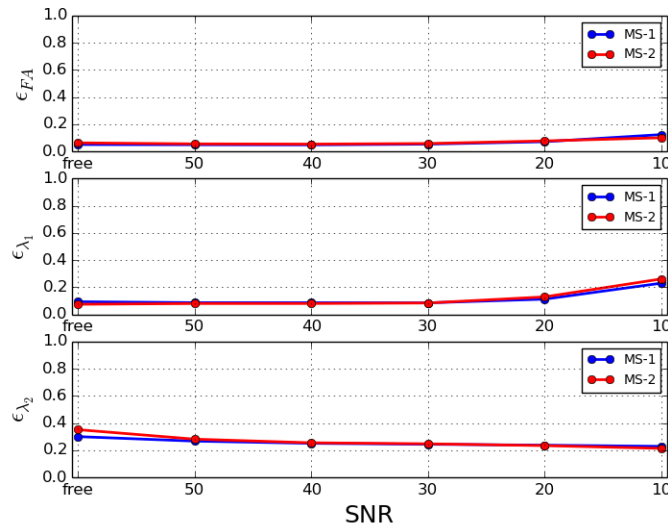


Figura 4.13: Error de Anisotropía Fraccional para Exp02.

Por lo anterior, no se observa una mejora significativa en las métricas, sólo un mejor ajuste en los pesos de los átomos. Lo que sí se puede establecer es que un esquema multi-shell más variado permitirá una mejor reconstrucción de la señal.

4.3. ISBI-WMMC 2015

Como parte de la validación de nuestro modelo, se participó en el *White Matter Modelling Challenge*, el cual forma parte de los ISBI Challenges 2015. Este año se liberaron 3600 mediciones para 12 vóxeles, 6 de la región Genu y 6 del Fornix. El esquema utilizado para obtener las señales incluye un multi-shell con valores de b en el rango de $[50, 42500]$. El desafío consistió en predecir 1200 señales no liberadas; más información sobre este concurso puede consultarse en <http://cmic.cs.ucl.ac.uk/wmmchallenge/>.

Para este concurso, las 1200 predicciones para cada vóxel fueron generadas con una versión incompleta de nuestro modelo, sólo se tenía implementado el factor de escala B , y no incluye la penalización ℓ_0 y la rotación de los átomos del diccionario. Para esa versión, también se utilizó una base de tensores de 129 átomos, distribuidos uniformemente, con un perfil de $\lambda_1 = 0,001$ y $\lambda_2 = 0,00005$. Los pesos para los términos de regularización iniciales se establecieron como $\mu_\alpha = 0,5$ y $\mu_B = 0,02$.

En la figura 4.14 se muestran las predicciones del vóxel 1 de la región Genu.

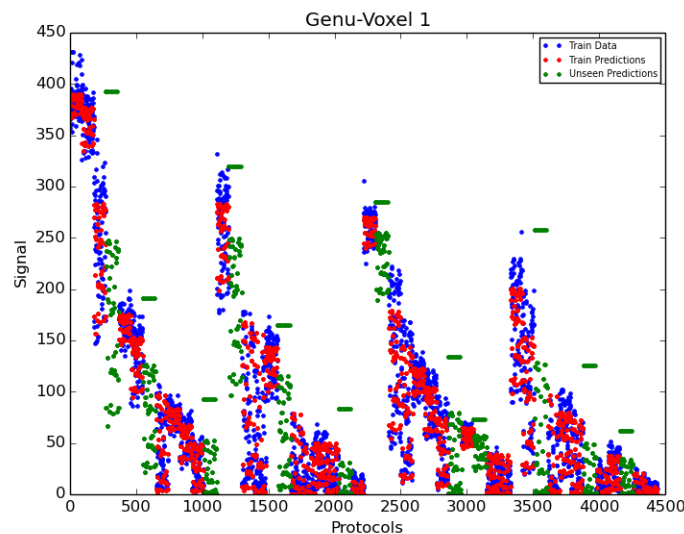


Figura 4.14: Predicciones para el vóxel 1 del Genu con el modelo incompleto.

Ahora con el modelo completo, es interesante comparar los resultados con los datos ya enviados al *challenge*. En este tercer experimento se utilizarán los mismos parámetros que en la versión incompleta. En la figura 4.15 se observan las predicciones con el modelo completo.

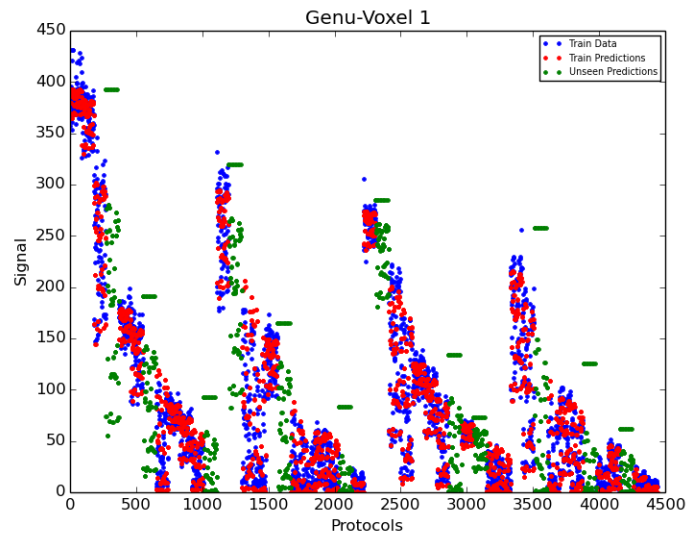


Figura 4.15: Predicciones para el vóxel 1 del Genu con el modelo completo.

Se observa que antes las predicciones estaban compactas; ahora con el modelo final, las predicciones se dispersan un poco, tanto en los datos de entrenamiento como en las predicciones de los datos no liberados. En la figura 4.16 se comparan los errores de reconstrucción para los 6 vóxeles Genu con la version anterior y la nueva.

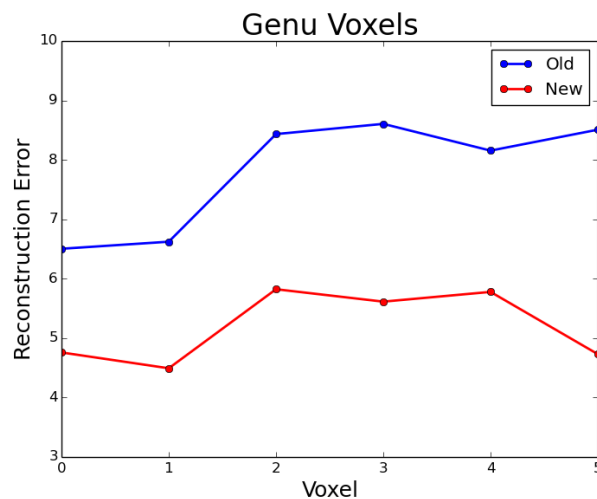


Figura 4.16: Errores de Reconstrucción en los 6 vóxeles Genu.

Con el modelo completo, se reconstruye la señal con mayor precisión, en algunos vóxeles el error de reconstrucción disminuyó a la mitad. Como se vio en el experimento 2, los errores de

reconstrucción disminuyeron en un esquema multi-shell con más shells. Este comportamiento también se observa en estos resultados, al lograr disminuir los errores de reconstrucción.

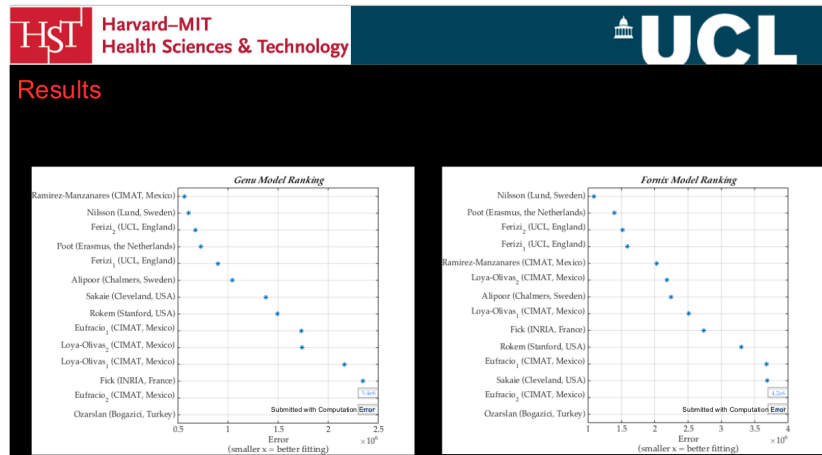


Figura 4.17: Resultados del Challenge 2015

4.4. SPARC-dMRI 2014

En este último experimento se utilizó los datos del *Sparse Reconstruction Challenge for Diffusion MRI 2014*. Este conjunto de datos fue adquirido con valores de $b = \{1000, 2000, 3000\}$ con $K = \{20, 30, 60\}$ direcciones de gradiente por *shell*. En cada esquema se incluyó un $b = 0$, lo que hacen un total de $\{61, 91, 121\}$ imágenes por esquema; más información de este curso puede consultarse en: <http://projects.iq.harvard.edu/sparcdmri>.

En este experimento se trabajó con el conjunto de 121 imágenes. Los parámetros para los algoritmos incluyen una base de tensores de 129 átomos, distribuidos uniformemente; un perfil de los tensores bases con $\lambda_1 = 0,001$ y $\lambda_2 = 0,0001$. Los pesos para los términos de regularización iniciales se establecieron como $\mu_\alpha = 0,004$ y $\mu_B = 0,02$.

En la figura 4.18 se observa el ajuste de tensores por el método DBF. Este método es capaz de identificar vóxeles con cruces aunque muestra problemas de sobre-estimación en ciertos vóxeles.

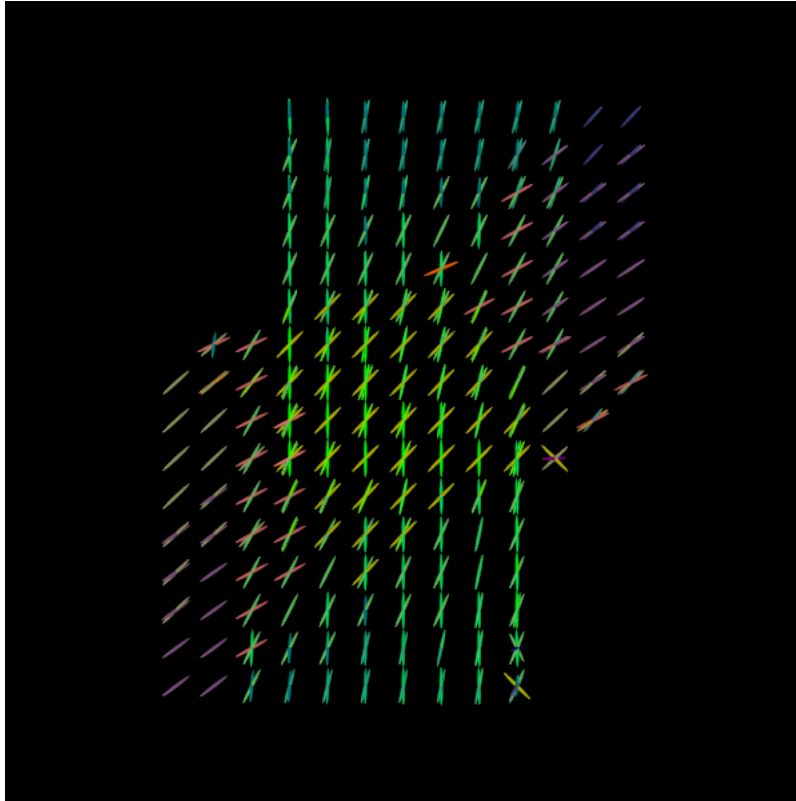


Figura 4.18: Ajuste con el método DBF

En la figura 4.19 se muestra el ajuste por nuestro algoritmo. En este caso, se observan cruces más definidos y en la mayoría de ellos una correcta estimación en el número de fibras.

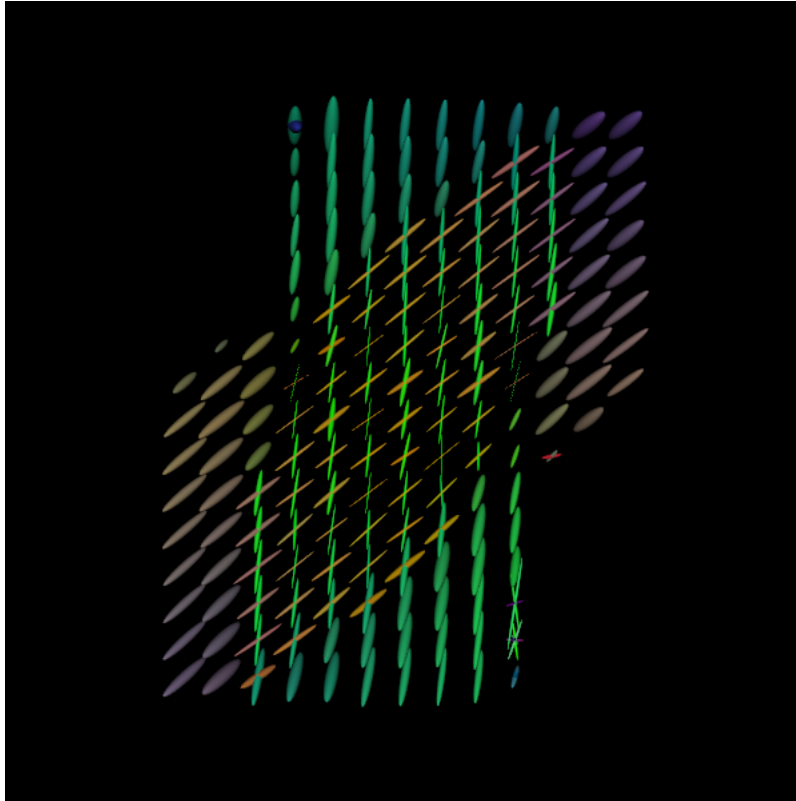


Figura 4.19: Ajuste con nuestra modelo

Con este experimento, más *cualitativo*, se observa que nuestro algoritmo logra un mejor rendimiento en la estimación del número de manojos de fibras para estos datos. Finalmente, la tabla 4.1 resumen los tiempos de ejecución de los métodos para este experimento.

Cuadro 4.1: Tiempos de Ejecución

Método	Vóxeles	Mínimo	Máximo	Mediana	Total
DBF ⁺	168	0.009184	0.01114	0.009229	1.555587
MS-SADD	168	0.028461	1.662814	0.115024	37.67833

Si bien el método DBF⁺ muestra mejores tiempos, se considera que nuestro algoritmo es competitivo en tiempo, dada la complejidad del mismo.

Capítulo 5

Conclusiones y Trabajo Futuro

Como se mostró al inicio de este trabajo, existen modelos en el estado del arte, como DBF y SADD, que recuperan la estructura de las fibras en el cerebro con la información de DW-MRI. Si bien estos modelos logran una gran precisión en sus resultados, el punto de partida de esta tesis mostraba la necesidad de mejorar el rendimiento, ya sea un mejor ajuste en la señal recuperada, mejor estimación del número de fibras o menor costo computacional, etc.

Durante la tesis se implementaron las siguientes ideas claves: la reformulación de la función objetivo para obtener explícitamente un factor de escala, el término de regularización ℓ_0 así como la adaptación de los tensores base. Como se observó en los experimentos, estas ideas permitieron lograr una mejor reconstrucción de la señal y a la vez, una mejor reconstrucción física en la estimación en la estructura de las fibras.

A continuación se presenta los puntos mas importante del trabajo futuro que pueden ayudar a mejorar el trabajo presentado en esta tesis:

- En nuestra propuesta se utilizaron los mismos valores de μ_α , μ_β, λ_1 y λ_2 para los distintos casos de SNR. Por tanto, se podría determinar una estrategia para elegir de una forma óptima los valores de dichos parámetros.
- En los resultados de los diferentes experimentos se observó que el error relativo al estimar λ_2 era mayor en comparación que λ_1 . Así, es importante diseñar otra formulación para obtener λ_2 a partir de los valores óptimos de β .

- Comparar el desempeño del método DBF original contra el método DBF al incluir el término de regularización que aproxima la norma ℓ_0 .
- Medir el desempeño del método MS-SADD al no mover tantos parámetros cuando se trabaja con señales con niveles altos de ruido.
- Al inicio de esta tesis se buscaron ciertas estrategias para disminuir el ruido en la señales; finalmente, estas estrategias no se incorporaron en este trabajo. Resulta importante incorporar en el modelo una estrategia para disminuir dicho ruido.
- Incluir información espacial para incrementar la robustez al ruido.
- Extender el modelo para considerar tres compartimientos: difusión intracelular, extracelular y libre.
- Buscar la incorporación de nuestro algoritmo en software libre, como Dipy.

Bibliografía

Aranda, R., Ramirez-Manzanares, A., and Rivera, M. (2015). Sparse and adaptive diffusion dictionary (sadd) for recovering intra-voxel white matter structure. *accepted in Medical Image Analysis*.

Basser, P., Mattiello, J., and LeBihan, D. (1994). Mr diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1):259–267.

Canales-Rodríguez, E. J., Lin, C.-P., Iturria-Medina, Y., Yeh, C.-H., Cho, K.-H., and Melie-García, L. (2010). Diffusion orientation transform revisited. *NeuroImage*, 49(2):1326–1339.

Cheng, J., Deriche, R., Jiang, T., Shen, D., and Yap, P.-T. (2014). Non-negative spherical deconvolution (nnsd) for estimation of fiber orientation distribution function in single-multi-shell diffusion mri. *NeuroImage*, 101:750–764.

Daducci, A., Canales-Rodriguez, E., Descoteaux, M., Garyfallidis, E., Gur, Y., Lin, Y., Marni, M., Merlet, S., Paquette, M., Ramirez-Manzanares, A., Reisert, M., Reis Rodrigues, P., Seppehrband, F., Caruyer, E., Choupan, J., Deriche, R., Jacob, M., Menegaz, G., Prekowska, V., Rivera, M., Wiaux, Y., and Thiran, J. a. (2014). Quantitative comparison of reconstruction methods for intra-voxel fiber recovery from diffusion mri. *Medical Imaging, IEEE Transactions on*, 33(2):384–399.

Datta, B. N. (2010). *Numerical Linear Algebra and Applications*. s.i.a.m.

Ghosh, A., Megherbi, T., Boumghar, F., and Deriche, R. (2013). Fiber orientation distribution from non-negative sparse recovery. In *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*, pages 254–257.

- Jian, B. and Vemuri, B. (2007). A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted mri. *Medical Imaging, IEEE Transactions on*, 26(11):1464–1471.
- Johansen-Berg, H. and Behrens, T. (2013). *Diffusion MRI: From Quantitative Measurement to In vivo Neuroanatomy*. Elsevier Science.
- Landman, B. A., Bogovic, J. A., Wan, H., ElShahaby, F. E. Z., Bazin, P.-L., and Prince, J. L. (2012). Resolution of crossing fibers with constrained compressed sensing using diffusion tensor mri. *NeuroImage*, 59(3):2175–2186.
- Loya, A. K. and Rivera, M. (2015). Algoritmos para Solución de Modelos No Lineales en DW-MRI. Master’s thesis, Centro de Investigación en Matemáticas, A.C.
- Mani, M., Jacob, M., Guidon, A., Magnotta, V., and Zhong, J. (2015). Acceleration of high angular and spatial resolution diffusion imaging using compressed sensing with multichannel spiral data. *Magnetic Resonance in Medicine*, 73(1):126–138.
- Merlet, S., Caruyer, E., Ghosh, A., and Deriche, R. (2013). A computational diffusion mri and parametric dictionary learning framework for modeling the diffusion signal and its features. *Medical Image Analysis*, 17(7):830–843.
- Ramirez-Manzanares, A., Cook, P., Hall, M., Ashtari, M., and Gee, J. (2011). Resolving axon fiber crossings at clinical b-values: an evaluation study. *Medical Physics*, 38(9):5239–5352.
- Ramirez-Manzanares, A., Rivera, M., Vemuri, B., Carney, P., and Mareci, T. (2007). Diffusion basis functions decomposition for estimating white matter intravoxel fiber geometry. *Medical Imaging, IEEE Transactions on*, 26(8):1091–1102.
- Stejskal, E. and Tanner, J. (1965). Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*, 42(1):288–292.
- Tuch, D. S., Reese, T. G., Wiegell, M., Makris, N., Belliveau, J. W., and Wedeen, V. (2002). Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *Magnetic Resonance in Medicine*, 48(4):577–582.

- Westbrook, C., Kaut-Roth, K., and Talbot, J. (2011). *MRI in Practice*. Wiley-Blackwell.
- Yap, P.-T. and Shen, D. (2012). Spatial transformation of dwi data using non-negative sparse representation. *Medical Imaging, IEEE Transactions on*, 31(11):2035–2049.
- Ye, W., Vemuri, B., and Entezari, A. (2012). An over-complete dictionary based regularized reconstruction of a field of ensemble average propagators. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 940–943.