



Universidad  
Carlos III de Madrid

## **TESIS DOCTORAL**

# **Using state-of-the-art inverse problem techniques to develop reconstruction methods for fluorescence diffuse optical tomography**

**Autor:**

**Judit Chamorro Servent**

**Directores:**

**Manuel Desco**

**Jorge Ripoll**

**DEPARTAMENTO/INSTITUTO**  
**Bioingeniería e Ingeniería Aeroespacial**

**Leganés, julio 2013**



## **TESIS DOCTORAL**

### **Using state-of-the-art inverse problem techniques to develop reconstruction methods for fDOT**

Autor: Judit Chamorro Servent

Directores: Manuel Desco, Jorge Ripoll

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Secretario:

Calificación:

Leganés, de de 2013



*“Most people, if you describe a train of events to them will tell you what the result will be. There are few people, however that if you told them a result, would be able to evolve from their own inner consciousness what the steps were that led to that results. This power is what I mean when I talk of reasoning backward”*

Sherlock Holmes,  
A Study in Scarlet, Part 2, Chapter 7.  
Sir Arthur Conan Doyle (1887)



# Agradecimientos

Una tesis, como todo en esta vida, no tiene un solo autor, sino un conjunto de ellos. Me gustaría agradecer a esa gente que de una manera u otra, directa o indirectamente, han dejado su granito de arena en ella.

En primer lugar me gustaría agradecer a Manuel Desco y Juan José Vaquero por haberme permitido la oportunidad de desarrollar mi tesis en el LIM. Siempre me han dejado investigar en lo que quería y se han preocupado por mantener ese grupo multidisciplinar en el LIM y que todos aprendamos o nos suene, como mínimo, lo que hace un compañero con una formación totalmente distinta a la nuestra y trabajando en otra rama.

Quiero agradecer a Manolo además por haber sido mi director de tesis y haber corregido esta memoria. Manolo sabe como enfocar un paper o una charla y eso como doctorando se agradece. Y a Juanjo que me llevó también durante un tiempo, siempre tuvo la puerta abierta y pese a pertenecer a una rama un tanto distinta a la mía ha hecho el esfuerzo de entender lo que yo intentaba plasmar en esos trabajos llenos de formulotas.

Jorge Ripoll llegó al LIM en el 2012, sin embargo siempre ha seguido mi tesis. Cuando conocí a Jorge, en un descanso de un congreso, ya llevaba casi un año contestando algunas de mis preguntas vía e-mail, pero aun allí, se sentó conmigo en una mesa a revisar los que eran los principios de los capítulos 3 y 4 de esta tesis. En el mismo congreso, un chico que se acercó a mi póster, me preguntó si conocía a Jorge y al hablarle de su disponibilidad me dijo: “¡A qué estudiante de doctorado no le soluciona Jorge sus dudas!”. Como estudiante de doctorado he corroborado muchas veces esa frase. No importa lo ocupado que esté, siempre saca tiempo y es un placer poder disfrutar de sus charlas y aprender de él. Así mismo, me gustaría agradecerle también el haber corregido esta memoria y otros muchos trabajos y papers, su humor y su apoyo

Durante mi transcurso por el LIM he podido conocer mucha gente, por orden cronológico: Alejandro Sisniega, mi compañero de mesa. Una fuente de sabiduría, siempre dispuesto a ayudar. Me llevo muchas cosas aprendidas de Alex y un gran amigo que ha estado siempre ahí. Verónica García, cuando llegué siempre estuvo allí disponible para ayudarme (¿a quién no ha ayudado *Veronique*?). Ella y Alexia Rodríguez que también me mostró siempre su disponibilidad, me ayudaron a editar algunas de las imágenes de los capítulos 3 y 4 con la consola. De ellas me llevo mucho aprendido y dos grandes amistades!. Angelito, uno de los mejores compañeros de despacho, que tiempos los de Pin&Pon y un amigo de los grandes (y a Claudia, un solete, también por cuidármelo). Juanolas, quien, junto a las preguntas que hacía/mos a Jorge, me ayudó en mis principios de la óptica. Me explicó desde hacer un phantom de agar o de resina, adquirir datos y hasta a cambiar los filtros de esa máquina que él mismo creó con sus manitas. De ahí salieron los experimentos del capítulo 3 y 4. Todo esto siempre con un humor inigualable aun en las circunstancias más difíciles como cuando

se nos cayó una máquina encima o intentando crear una tienda oscura para el FMT-CT. Edu, el doctor Lage, siempre disponible también y con su frase de “nadie nace con vaqueros” cuando en aquellos principios le preguntabas la cosa más estúpida de C/C++ o cuando me explicaba como crear la matriz del sistema en PET. Gracias también por tu ayuda toqueteando el código para los experimentos del capítulo 4 de esta tesis. Paula eres un solete muy encantador, gracias por estar siempre y además por las cosillas de reso que me llevo aprendidas. Trajana y Eva, a ellas les tengo que dar mil gracias!!! Cuando uno hace papeleo se da cuenta del enorme trabajo que hay detrás. Gracias “jefa” Tra por tu disponibilidad, sin ti nada de esto funcionaría. Gracias Eva por todos los papeleos y por siempre preocuparte de todos nosotros. Gracias a las dos por vuestra alegría, disponibilidad y por ser tan majetas. Esther, por esos abrazos y sonrisas que nunca falten! has estado para lo bueno y lo malo (Álex tb) y aunque no seas exactamente del LIM para todos creo que lo eres. Sandra y Marina, sin sus manos no se hubieran hecho experimentos como el del ratón del capítulo 3 de esta tesis. Gracias chicas por ser tan encantadoras y cariñosas. Por la siempre disponibilidad de Chemita a mis preguntas de C/C++ o a las muchas dudas resueltas de ordenadores y por sus ánimos siempre. Alvarico, Gus y Josete, el trio del humor. Irina, por ayudarme tanto con el inglés de mi primera charla. El doctor Juan, “Ay pedrín”, le he de agradecer al “doctor Juan” los innumerables ratos dando vueltas a ese código inicial de Split Bregman de Goldstein y a ese paper que tan buenos resultados nos ha brindado. Siempre dispuesto a llenar una pizarra de ecuaciones, discutir un problemilla de mates o recomendar un buen vino o comida. María, no sólo por los buenos momentos en una cenita o tomando algo, sino también por ayudarme a entender toda la bio que he necesitado. Carmen, siempre te recordaré, porque fuiste alguien muy especial que nos enseñó mucho, por no hablar del CASEIB09. Por siempre preocuparte de todos aún cuando peor estabas, por tus ánimos, sonrisas y esa fuerza. Me gustaría estuvieras aquí presente aunque de alguna forma siempre lo estarás. Marco, siempre es un placer desayunar y hablar contigo o ir a una de tus fiestas benéficas, gracias también por los ánimos y a seguir viajando eh. Mónica, por los ánimos y las recomendaciones de sitios qué visitar, espero que ya estés bien cuando leas esto. Marisa y Javi Pascau, seniors desde que llegué pero siempre dispuestos a animar a los juniors. Joost, por tus ánimos y charlas tan chulas de neuro. Santi, quien siempre tenía una palabra de ánimo y se preocupaba por todos, gracias por ser así. Nunca pude traerte la foto del pájaro carpintero pero siempre que la veo me recuerda a ti. Siempre estarás también de alguna forma con nosotros. Yasser, ese cubano sonriente tan atento a los demás. Quique, siempre recordaré las risas del congreso en Cádiz y tu obra de Aladín, siento que mis estancias me llevarán a casi no verte en los últimos tiempos. Javi, quien definió de la forma más graciosa el inverse problem, cuida de Bruno. Lorena, que puede sorprenderte manejando las espadas de esgrima o los bisturís con las ratitas. Elena, esa médica nuclear siempre preocupándose por todos. Las champions, que aunque el orden cronológico os haga llegar el agradecimiento tan tarde habéis dejado una gran huella, por las risas, el cariño, alegría, ánimos y siempre estar ahí. La encantadora Aurora o el saco de la risa que tanta alegría nos deja, el siempre energético solete de Clau, tan majeta ella y una experta en word, y los siempre ánimos de Elia y Merche. Carlos, por acordarte hasta desde la lejanía de nosotros, aquí te espera mucha gente con los brazos abiertos de vuelta. Iván, por los saludos más energéticos del LIM, aunque fue breve, fue un placer compartir inglés contigo. Juanjo peque, gracias por tu cariño y tu humor. Martín, fue un placer y espero q los inmensos destinos te estén tratando bien. Eu, quien hace mapas con las mejores recomendaciones de Granada. Fidel, por crearme esa máquina virtual que tanto me solucionó la vida, poder lanzar un programa allí y otro en mi ordenador, qué lujo. Gracias! Y porque siempre te preocupas por todos. Susana, esa catalana que se está adaptando tan



bien en Madrid. Rigo y Ana, por los paseos a por caféina o un descanso a media tarde y las risas. Natalia, por los ánimos y las sonrisas. Luisa, por ser tan amable y risueña y una crack sobreviviendo a hijos y tesis a la vez. Santi, siempre con su sonrisa alegrando los pasillos. Inés, que aunque hace poquito te conozco estás dejando muy buena huella y ahora seremos vecinas por un tiempo y todo. Y las nuevas adquisiciones del buffer y de abajo que no he tratado tanto pero que no me quiero ir sin mencionar.

Ya fuera del LIM, me gustaría agradecer a dos profesores que directa o indirectamente me llevaron a descubrir cómo aplicar las mates a imagen:

Xavier Bardina, per la seva orientació, per fer-me redescobrir les mates i conduir-me cap a una línia que avui no només és la meva feina, si no també el meu hobby. Gracies per guiar-me a fer aquell erasmus i començar aquell màster que em conduiria fins a la tesis. Gracies també per la teva disponibilitat sempre.

Je voudrais remercier aussi mon coordinateur du Master à Toulouse, Mohamed Masmoudi, pour être toujours disponible, même après le master, et m'apprendre les préliminaires du problème inverse. Aussi pour accepter être part du jury de ma thèse.

Regarding the different collaborations, I would like thank:

Prof. Simon Arridge from UCL, who brought us the Split Bregman paper and his ideas. It was a very good collaboration resulting in three papers. Thank you also for the efforts spent on the pre-evaluation and validation of my thesis. Teresa Coreia also from UCL thank you for your collaboration and encouragement.

Regarding my stays at UBC and DTU:

UBC: Prof. Eldad Haber for allowing me to do a stay in his department and to assist his workshop. During my stay at UBC, he shared with me some of his little free time discussing preconditioning and compressed sensing subjects.

To Sasha, who included me in the seminars from his group and helped me to better understand the SPGL1 code. A Luz Angelica, por su amistad, cariño y ánimos. To Shruti, it was a pleasure to share the lunch times with you and your good mood always. To Manjit for her availability always and her humour.

DTU: I would also like to express my deep appreciation and gratitude to Prof. Per Christian Hansen. His books and papers were to me real manuals to understand inverse problems at the beginning and have made an important mark in this thesis. It was a real pleasure to be allowed to undertake a doctoral visit with his group last year. He always took great care to make me feel very welcome as a visitor. His expertise and advice have been one of the most valuable gains from my visit. Thank you also for your availability and your careful attention to the validation of my thesis and for always being ready to help. Tak!

Thank you also to Prof. Sergios Theodoridis from Athens University, for allowing me to follow his compressed sensing course during my stay at DTU, for addressing all of my questions in person and by e-mail and for giving me such helpful advice. His course, his help and the notes of a chapter of a book he is publishing have been important sources for chapter seven of this thesis.

I would also like to express my gratitude to Jakob Heide Jorgensen from DTU. My best colleague. It was a real pleasure to work with you. I learned a lot from you also. Spanish cheese forever. Tak!

And of course the BDC club at DTU: Jakob again (The Danish BDC member), Yulia (the Russian girl always smiling, energetic and a good friend), Dimitri (BDC would not exist without Dimitri, the coffee cup friendly guy), Roman (the card games and rum guy),

Andrea (my best roommate with music for all occasions), Laura (the home parties lady), Andrea Z (always available for a coffee break), Danielle (the real Italian coffee guy), Nao (Now! The Japanese girl), Dress and Ivana (always available for a party).

No quiero irme sin agradecer a mis amigos y familia, piezas claves para esta tesis. Por orden cronológico empezando por los amigos:

Por estar siempre ahí y demostrar que no importan las distancias ni los años, ¿20 y cuantos ya?, Con vosotros he vivido mis mejores experiencias y no hace falta que os diga cuanto os agradezco el haberos cruzado en mi camino. En especial a Cris R. y el capi Joan, Rosa, Arturo y Montse, Carlos (siempre me animaste a hacer la carrera de mates), los Serrano, Laura, Carlos y Anabel (gracias también por lo otro que ya sabéis y tanto me ha ayudado también en esta tesis ;)), Cris S., Dani, Antonio, Mimi y Marta.

A la Iola y el Pasqui, els meus pseudo-cosins, pels seus ànims sempre i visitar-me vagi on vagi. Sou els millors!

A la Berta per la d'apunts i tardes a la biblio i estar encara avui tant present. Víctor cuida-la. Els dos valeu molt.

Cuando uno tiene que viajar para estudiar o trabajar en lo que le gusta, se agradece encontrar gente que acaban siendo como una *little family*:

Ma petite *famille* en France et qui ont été pendant ma thèse de quelque façon aussi là, vous le savez bien, et bien spécialement à Brice et Céline qui ont été non seulement partie de cette famille mais aussi ont passé avec moi des nuits en blanche en étudiant ( je ne sais pas si ça été ça qui nous a conduit à partager après tellement des expériences ;)).

Meli, un especial también para ti! No importa donde te vayas, sea Londres, África, wherever, siempre tendré una excusa para escaparme a por una de nuestras experiencias. Mil gracias por ser tan solete y por corregir mi inglés de articulillos y demás también.

Gracias también a tres personajes muy importantes en el transcurso de esta tesis, porque sin vosotras nada hubiera sido posible, my *little family* en Madrid, esas San Bernardinas: Clari, Anusca y Albi. Gracias por estar siempre ahí, sois tres soletes inigualables.

Gracias también a la chica más maja de la Roda, Eva, por sus mensajes animadores en la recta final. Y a esos albaceteños por su humor y alegría.

En último lugar, pero no por ello menos importante:

Gracies per fer-ho possible a la meva família: Tiets i cosins sempre animant i per descomptat la meva mare qui tant m'ha animat a fer el que m'agradava, el meu pare, la meva germana i l'Héctor, així com els somriures en la recta final del Pol i el Gerard. També vull deixar un agraïment als meus avis, en especial al meu avi de bcn qui em va ensenyar tant de fotografia i em va despertar tant la curiositat per aprendre, i la meva avia de Sarroca que sempre em va animar tant a seguir lluitant pel que m'agradava. Gracies a tots per tot, i per ser com sou!

Antes de cerrar este apartado, no puedo irme sin agradecer a Vito todo su cariño y el estar ahí. Gracias por estar siempre a mi lado y por siempre animarme a hacer lo que me gusta cueste lo que cueste, incluso, aunque ello suponga estar *lejos* unos mesecillos o seguirme a un país nuevo que descubrir. Para mí eres un grandísimo autor de esta tesis, no sólo por ese ratón para la portada o el video 3D para la presentación, sino por todo lo mucho que has significado en ella.

Y ya por último, agradecer los miembros del tribunal por su disponibilidad.

Judit Chamorro Servent  
Julio 2013

# Contents

<b>Abstract .....</b>	<b>v</b>
<b>Resumen .....</b>	<b>vii</b>
<b>List of Symbols.....</b>	<b>xi</b>
<b>List of Acronyms.....</b>	<b>xiii</b>
<b>Motivation and Objectives .....</b>	<b>1</b>
Structure of the thesis .....	3
<b>Introduction .....</b>	<b>5</b>
1.1. Forward and inverse problems.....	5
1.2. Ill-posed problems and uncertainty of solution .....	6
1.3. Fluorescence diffuse optical tomography (fDOT).....	7
1.3.1. Experimental setup .....	8
1.3.2. Phantoms.....	9
1.3.3. Forward problem.....	9
1.3.4. Simulated data.....	12
1.4. Solving the fDOT inverse problem .....	12
1.4.1. fDOT, an ill-posed problem.....	13
1.4.2. Non-contact fDOT, large datasets .....	13
1.4.3. Reconstruction methods.....	14
<b>Linear <math>l_2</math> regularization reconstruction methods.....</b>	<b>15</b>
2.1. The need for regularization .....	15
2.2. Singular Value Decomposition .....	16
2.2.1. The role of SVD.....	16
2.2.2. SVD for ill-posed problems .....	17
2.3. Truncated singular value decomposition.....	18
2.4. Tikhonov Regularization.....	18
2.5. Dealing with fDOT.....	19
<b>Choosing the regularization parameter for <math>l_2</math> regularization reconstruction methods</b>	
<b>.....</b>	<b>21</b>
3.1. Review of Current Methods .....	22
3.2. Method proposed.....	24
3.2.1. U-curve method .....	24
3.2.1.1. Desirable interval of the regularization parameter.....	24
3.2.1.2. Unicity of U-curve solution.....	26
3.2.2. The role of Discrete Picard's condition.....	26
3.2.3. Feasibility of the U-curve method for fDOT .....	27
3.2.3.1. Phantom experiment.....	27
3.2.3.2. Ex-vivo mouse experiment.....	28
3.2.3.3. Validation of the regularization parameter obtained by the U-curve method.....	28
3.3 Results .....	29
3.3.1 Phantom experiment validation .....	29

3.3.2 Ex-vivo mouse data experiment validation.....	33
3.4. Discussion and Conclusions .....	36
<b>SVA applied to optimizing non-contact fDOT .....</b>	<b>39</b>
4.1. Introduction.....	40
4.2. Methods.....	41
4.3. Results.....	43
4.3.1. Results on density of sources and detectors .....	43
4.3.2. Results on the mesh spatial distribution .....	45
4.4. Discussion and Conclusions .....	47
<b><math>l_2</math>-norm alternatives enforcing sparsity of the solution.....</b>	<b>49</b>
5.1. Sparse solutions .....	50
5.1.1. $l_0$ -norm and $l_1$ -norm.....	50
5.1.2. Total Variation.....	51
5.2. Formulation of the optimization problem.....	52
5.3. Looking for an algorithm .....	53
5.4. A brief review of sparse regularization reconstruction techniques applied to fDOT .....	54
<b>Use of Split Bregman denoising for iterative reconstruction .....</b>	<b>57</b>
6.1. Introduction.....	58
6.2. Methods.....	59
6.2.1. The algebraic reconstruction technique (ART) .....	59
6.2.2. The two-step reconstruction method: ART-SB .....	60
6.2.3. Experimental and simulated data.....	63
6.2.3.1. Experimental phantom data.....	63
6.2.3.2. Simulated data .....	63
6.2.4. Comparison between ART and ART-SB.....	64
6.2.4.1. Simulated data .....	64
6.2.4.2. Experimental phantom data.....	65
6.3. Comparison results of ART versus ART-SB.....	65
6.3.1. Selection of parameters.....	65
6.3.1.1. Selection of relaxation parameter for ART.....	65
6.3.1.2. Selection of weighting and denoising parameters of ART-SB .....	66
6.3.2. Comparison between ART and ART-SB.....	67
6.3.2.1. Simulated data .....	67
6.3.2.2. Experimental phantom data.....	70
6.4. Discussion and Conclusions .....	72
<b>Compressed Sensing in fDOT .....</b>	<b>75</b>
7.1. Mathematical basics.....	76
7.1.1. Sparse synthesis model .....	77
7.1.1.1. The spark of a matrix .....	77
7.1.1.2. The mutual coherence .....	77
7.1.1.3. $l_0$ and $l_1$ -minimizer solutions.....	78
7.1.1.4. The Restricted Isometry Property (RIP).....	78
7.1.2. Co-sparse analysis model.....	79
7.2. CS applied to the fDOT ill-posed problem .....	80
7.2.1. Brief review .....	80
7.2.2. Incoherence of the fDOT forward matrix .....	82
7.3. A novel approach to CS for fDOT: the SB-SVA method.....	84
7.3.1. The Split Bregman (SB) approach to CS.....	84
7.3.2. The SB-SVA method .....	85
7.4. Results.....	86
7.5. Discussion and Conclusions .....	88

<b>Conclusions .....</b>	<b>91</b>
<b>Publications .....</b>	<b>93</b>
Journal papers .....	93
International conference record proceedings.....	94
National conference record proceedings .....	95
<b>References.....</b>	<b>97</b>



# Abstract

An inverse problem is a mathematical framework that is used to obtain info about a physical object or system from observed measurements. It usually appears when we wish to obtain information about internal data from outside measurements and has many applications in science and technology such as medical imaging, geophysical imaging, image deblurring, image inpainting, electromagnetic scattering, acoustics, machine learning, mathematical finance, physics, etc.

The main goal of this PhD thesis was to use state-of-the-art inverse problem techniques to develop modern reconstruction methods for solving the fluorescence diffuse optical tomography (fDOT) problem. fDOT is a molecular imaging technique that enables the quantification of tomographic (3D) bio-distributions of fluorescent tracers in small animals.

One of the main difficulties in fDOT is that the high absorption and scattering properties of biological tissues lead to an ill-posed inverse problem, yielding multiple non-unique and unstable solutions to the reconstruction problem. Thus, the problem requires regularization to achieve a stable solution.

The so called “non-contact fDOT scanners” are based on using CCDs as virtual detectors instead of optic fibers in contact with the sample. These non-contact systems generate huge datasets that lead to computationally demanding inverse problem. Therefore, techniques to minimize the size of the acquired datasets without losing image performance are highly advisable.

The first part of this thesis addresses the optimization of experimental setups to reduce the dataset size, by using  $l_2$ -based regularization techniques. The second part, based on the success of  $l_1$  regularization techniques for denoising and image reconstruction, is

devoted to advanced regularization problem using  $l_1$ -based techniques, and the last part introduces compressed sensing (CS) theory, which enables further reduction of the acquired dataset size.

The main contributions of this thesis are:

1) A feasibility study (the first one for fDOT to our knowledge) of the automatic U-curve method to select the regularization parameter ( $l_2$ -norm). The U-curve method has shown to be an excellent automatic method to deal with large datasets because it reduces the regularization parameter search to a suitable interval.

2) Once we found an automatic method to choose the  $l_2$  regularization parameter for fDOT, singular value analysis (SVA) of fDOT forward matrix was used to maximize the information content in acquired measurements and minimize the computational cost. It was shown for the first time that large meshes can be reduced in the  $z$  direction, without any loss in imaging performance but reducing computational times and memory requirements.

3) Dealing with  $l_1$ -based regularization techniques, we presented a novel iterative algorithm, ART-SB, that combines the advantage of Algebraic reconstruction method (ART) in handling large datasets with Split Bregman (SB) denoising, an approach which has been shown to be optimum for Total Variation (TV) denoising. SB has been implemented in a cost-efficient way to handle large datasets. This makes ART-SB more computationally efficient than previous TV-based reconstruction algorithms and most splitting approaches.

4) Finally, we proposed a novel approach to CS for fDOT, named the SB-SVA iterative method. This approach is based on the analysis-based co-sparse representation model, where an analysis operator multiplies the image transforming it in a sparse one. Taking advantage of the CS-SB algorithm, we restrict the solution reached at each CS-SB iteration to a certain space where the singular values of the forward matrix and the sparsity structure combine in beneficial manner. In this way, SB-SVA forces indirectly the well-conditioning of the forward matrix while designing (learning) the analysis operator and finding the solution. Furthermore, SB-SVA outperforms the CS-SB algorithm in terms of image quality and needs fewer acquisition parameters.

The approaches presented here have been validated with experimental data.



# Resumen

El problema inverso consiste en un conjunto de técnicas matemáticas para obtener información sobre un fenómeno físico a partir de una serie de observaciones, medidas o datos. Dicho problema aparece en muchas aplicaciones científicas y tecnológicas como pueden ser imagen médica, imagen geofísica, acústica, aprendizaje máquina, física, etc.

El principal objetivo de esta tesis doctoral fue utilizar la teoría del problema inverso para desarrollar nuevos métodos de reconstrucción para el problema de tomografía óptica difusiva por fluorescencia (fDOT), también llamada tomografía molecular de fluorescencia (FMT). fDOT es una modalidad de imagen médica que permite obtener de manera no-invasiva la distribución espacial 3D de la concentración de sondas moleculares fluorescentes en animales pequeños in-vivo.

Una de las dificultades principales del problema inverso en fDOT, es que, debido a la alta difusión y absorción de los tejidos biológicos, es un problema fuertemente mal condicionado. Su solución no es única y presenta fuertes inestabilidades, por lo que el problema debe ser regularizado para obtener una solución estable.

Los llamados escáneres fDOT “sin contacto” se basan en utilizar cámaras CCD como detectores virtuales, en vez de fibras ópticas en contacto con la muestras. Estos sistemas, necesitan un volumen de datos muy elevado para obtener una buena calidad de imagen y el coste computacional de hallar la solución llega a ser muy grande. Por esta razón, es importante optimizar el sistema, es decir, maximizar la información contenida en los datos adquiridos a la vez que minimizamos el coste computacional.

La primera parte de esta tesis se centra en optimizar el sistema de adquisición, reduciendo el volumen de datos necesario usando técnicas de regularización basadas en la norma  $l_2$ . La segunda parte se inspira en el gran éxito de las técnicas de regularización

basadas en la norma  $l_1$  para la reconstrucción de imagen, y se centra en regularizar el problema fDOT mediante dichas técnicas. El trabajo finaliza introduciendo la técnica de “compressed sensing” (CS), que permite también reducir el número de datos necesarios sin por ello perder calidad de imagen.

Las contribuciones principales de esta tesis son:

1) Realización de un estudio de viabilidad, por primera vez en fDOT, del método automático U-curva para seleccionar el parámetro de regularización (norma  $l_2$ ). U-curva mostró ser un método óptimo para problemas con un volumen elevado de datos, ya que dicho método ofrece un intervalo donde encontrar el parámetro de regularización.

2) Una vez encontrado el método automático de selección de parámetro de regularización se realizó un estudio de la matriz del sistema de fDOT basado en el análisis de valores singulares (SVA), con la finalidad de maximizar la información contenida en los datos adquiridos y minimizar el coste computacional. Por primera vez se demostró que el uso de un mallado con menor densidad en la dirección perpendicular al plano obtiene mejores resultados que el uso convencional de una distribución isotrópica del mismo.

3) En la segunda parte de esta tesis, usando técnicas de regularización basadas en la norma  $l_1$ , se presenta un nuevo algoritmo iterativo, ART-SB, que combina la capacidad de la técnica de reconstrucción algebraica (ART) para lidiar con problemas con muchos datos con la efectividad del método Split Bregman (SB) para reducir ruido en la imagen mediante su variación total (TV). SB fue implementado de forma eficiente para procesar un elevado volumen de datos, de manera que ART-SB es computacionalmente más eficiente que otros algoritmos de reconstrucción presentados previamente en la literatura, basados en la TV de la imagen y que la mayoría de las técnicas llamadas de “splitting”.

4) Finalmente, proponemos una nueva aproximación iterativa a CS para fDOT, llamada SB-SVA. Esta aproximación se basa en el llamado modelo analítico co-disperso (*co-sparse*), donde un operador analítico multiplica la imagen convirtiéndola en una imagen dispersa. Este método aprovecha el método SB para CS (CS-SB) para restringir la solución alcanzada en cada iteración a un espacio determinado, donde los valores singulares de la matriz del sistema y la dispersión (“sparsity”) de la solución en dicha iteración combinen beneficiosamente; es decir, donde valores singulares muy pequeños no estén asociados a valores distintos de cero de la solución “sparse”. SB-SVA mejora el mal condicionamiento de la matriz del sistema a la vez que diseña el operador apropiado a través del cual la imagen se puede representar de forma dispersa y soluciona el problema de

CS. Además, SB-SVA mostró mejores resultados que CS-SB en cuanto a calidad de imagen, requiriendo menor número de parámetros de adquisición.

Todas las aproximaciones que presentamos en esta tesis fueron validadas con datos experimentales



# List of Symbols

Symbol	
$\Omega$	Domain/Space
$\mu_a$	Absorption term
$\mu_s$	Scattering term
$A$	Forward matrix of a general problem
$x$	Image/Model of a general problem
$x_{true}, x_{exact}$	True image/model or target of a general problem
$b$	Data of a general problem
$b_{exact}$	Data without errors of a general problem
$l_2, \  \cdot \ _2, \  \cdot \ $	2-norm
$W$	Forward matrix of fDOT problem
$f$	Image of fDOT problem/Fluorescent concentration of fDOT problem
$f_{true}$	Target of fDOT problem
$d$	Data of fDOT problem
$M^T$	Transpose of a matrix $M$
$M^{-1}$	Inverse of a matrix $M$
$cond(M)$	Condition number of a matrix $M$
$Rang(M)$	Range of a matrix $M$
$Ker(M)$	Kernel of a matrix $M$
$diag(M)$	Diagonal of a matrix $M$
$I$	Identity matrix

$U$	Matrix of left singular vectors
$S$	Matrix with singular values in its diagonal
$V$	Matrix of right singular vectors
$u_i$	i-th left singular vector
$\sigma_i$	i-th singular value
$\sigma_{\max}, \sigma_0$	Maximum singular value
$\sigma_{\min}$	Minimum singular value
$v_i$	i-th left singular vector
$\tau$	Threshold
$Nd$ or $nd$	Number of detectors
$Ns$ or $ns$	Number of sources
$N$ or $n$	Number of voxels
$\nabla$	Gradient
$l_p, \parallel \parallel_p$	$p$ -norm
$l_0, \parallel \parallel_0$	0-norm
$l_1, \parallel \parallel_1$	1-norm
$   $	Absolute value
$\alpha$	$l_2$ regularization parameter
$f_\alpha$	$l_2$ regularized image/solution of fDOT
$\alpha_u$	$l_2$ regularization parameter obtained with U-curve method
$\alpha_L$	$l_2$ regularization parameter obtained with L-curve method
$x^k, x^{it}$	it-th or k-th iteration of a solution $x$
$\lambda$	Total variation/ $l_1$ regularization parameter
$\nabla_x, \nabla_y, \nabla_z$	Gradient in $x$ /gradient in $y$ /gradient in $z$
$D$	Dictionary
$T$	Analysis operator
$\Lambda$	Co-support
$\mu(M)$	Mutual coherence of a matrix $M$
$G$	Gram matrix

# List of Acronyms

Acronyms	
ART	Algebraic reconstruction technique
CN	Condition number
CS	Compressed Sensing
DOT	Diffuse optical tomography
DPC	Discrete Picard Condition
fDOT	Fluorescence diffuse optical tomography
FMT	Fluorescence molecular tomography
FOV	Field of view
FWHM	Full width half maximum
GN	Gauss-Newton
LS	Least square
NIR	Near infrared
PSF	Point spread function
ROF	Rudin, Osher and Fatemi
SB	Split Bregman
SNR	Signal to noise ratio
SVs	Singular values
SVA	Singular value analysis
SVD	Singular value decomposition
TSVD	Truncated singular value decomposition
TV	Total Variation
VOI	Volume of interest





# Motivation and Objectives

Inverse problems appear in many applications in science and technology, whenever we want to recover “hidden” information about a system from measurements of related variables. Most inverse problems have many particular features that make it advisable a specific study.

Fluorescence diffuse optical tomography (fDOT), also called fluorescence molecular tomography (FMT), is a molecular imaging technique able to retrieve the three dimensional distribution of the concentration of extrinsic fluorophores in small animals, non-invasively and in-vivo. This technique facilitates the detection and quantification of important biological processes, and is being increasingly employed in pre-clinical research and drug development.

Mathematically, fDOT can be modelled as a linear system and its inverse problem focuses on finding the fluorescence concentration at each voxel given the data and the forward matrix.

The high degree of absorption and scattering of light through biological tissues turns fDOT into a highly ill-posed problem that yields multiple non-unique and unstable solutions to the reconstruction problem. Therefore, small perturbations in acquired data may lead to arbitrarily large changes in the reconstructed images. Regularization techniques are required to achieve well-defined and stable approximated solutions.

The so called “non-contact fDOT scanners” are based on using CCDs as virtual detectors instead of optic fibers in contact with the sample. Besides, these non-contact fDOT systems generate huge datasets by using CCDs as a virtual detectors, that leads to

computationally demanding inverse problem. One alternative, explored in this thesis, is to try to minimize the size of the acquired datasets, without losing image quality.

The main goal of this PhD-thesis is to exploit state-of-the-art inverse problem techniques to develop modern reconstruction methods for solving the fDOT problem. With this aim, the specific objectives of this thesis were:

1) To demonstrate the feasibility of the U-curve method to select the best  $l_2$ -based regularization parameter for fDOT. The U-curve has been shown to be an excellent automatic method to deal with large datasets because it allows restricting the search of the regularization parameter to a desirable interval.

2) To study the effect of different settings of the acquisition parameters (distribution of mesh points, density of sources and detectors) of a parallel-plate non-contact fDOT system, in order to achieve the best possible imaging performance, understood as maximizing the information content in acquired measurements while minimizing the computational cost by using the singular value analysis.

3) To develop a novel efficient reconstruction method based on TV regularization and Split Bregman methods to reach solutions with improved resolution compared to classical methods. The method combines and exploits recent advances on  $l_1$ -based regularization techniques for image denoising and reconstruction, and takes advantage of the high sparsity of the fDOT reconstructed images.

4) To apply Compressed Sensing (CS) theory to develop a new algorithm to fDOT. The iterative algorithm developed, named Split Bregman-Singular value Analysis (SB-SVA), is based on the existing Split Bregman algorithm for CS (CS-SB), but it restricts the solution reached at each CS-SB iteration to a certain subspace where the singular values of the forward matrix and the sparsity structure combine in beneficial manner. Indirectly, SB-SVA induces a well-conditioning of the forward matrix, while design/learn the analysis operator and finding the solution.

This thesis was developed in the framework of one of the research lines carried out in the Department of Bioengineering and Aerospace Engineering, University Carlos III of Madrid and the Hospital General Universitario Gregorio Marañón. One important source of motivation for this work was the real need for better and more efficient reconstruction techniques to be applied to the biological experimentation carried out by the group.

The author was awarded with a FPI (fellowship of Spanish Personnel Research Training Program for PhD students) by the Spanish Ministry of Economy and Competitiveness (reference: BES-2009-024463, associated with the project TEC2008-06715-C02-01).

## Structure of the thesis

The first part of this thesis addresses the optimization of experimental setups to reduce the dataset size, by using  $l_2$ -based regularization techniques. The second part, based on the success of  $l_1$  regularization techniques for denoising and image reconstruction, is devoted to advanced regularization using  $l_1$ -based techniques. Finally, compressed sensing (CS) theory is applied to enable a further reduction of the acquired dataset size.

In summary, the thesis is structured as follows:

- Chapter 1 introduces the linear inverse problem and discusses why ill-posedness can be challenging when solving an inverse problem. Finally, it presents the fluorescence diffuse optical tomography (fDOT).

### - *First part:*

- Chapter 2 introduces the theoretical principles of  $l_2$ -based regularization techniques.
- Chapter 3 presents an automatic method that provides a suitable selection of the regularization parameters for  $l_2$ -based techniques.
- In chapter 4 we exploit the singular value decomposition of the forward matrix to study the effect of different fDOT system settings, such as the distribution of mesh points or the density of sources and detectors.

### - *Second part:*

- Chapter 5 introduces the theoretical principles of  $l_1$ -based regularization techniques and a brief review about  $l_1$ -norm regularization techniques used in fDOT.
- In chapter 6 we present a novel iterative algorithm for fDOT, based on Total Variation (TV) techniques, the ART-SB (Algebraic Reconstruction Technique - Split Bregman) method.
- Chapter 7 includes a brief presentation of the CS theory and the sparse synthesis and cospase analysis models, shows the incoherence of the fDOT forward matrix and proposes a novel CS-based approach for fDOT, named the SB-SVA method.

- Finally, chapter 8 summarizes the conclusions and the main contributions of this thesis.
- A list of author's publications is included after chapter 8.

# Chapter 1

## Introduction

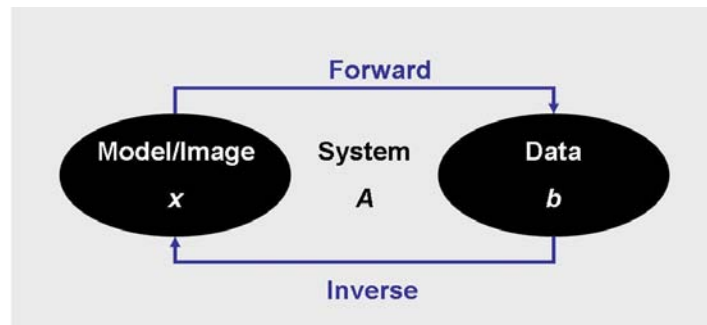
Inverse problems research concentrates on the mathematical theory and practical interpretation of indirect measurements. That is, on estimating certain parameters based on indirect measurements of these parameters. The inverse problems arise in many applications in science and technology such as medical imaging, geophysical imaging, electromagnetic scattering, acoustics, machine learning, physics, astronomy, mathematical financial, etc.

This chapter starts with a brief theoretical review of the *linear* inverse problem. Afterwards, it discusses why ill-posedness can be challenging when solving an inverse problem. Finally, it introduces a molecular medical imaging technique, fluorescence diffuse optical tomography (fDOT), also called fluorescence molecular tomography (FMT) and the ill-posed inverse problem of its reconstruction (the main interest of this thesis).

### 1.1. Forward and inverse problems

In 1976 Keller (Keller 1976) formulated the following very general definition of inverse problems, which is often cited in the literature:

“We call two problems inverses of one another if the formulation of each involves all or part of the solution of the other. Often, for historical reasons, one of the two problems has been studied extensively for some time, while the other is newer and not so well understood. In such cases, the former problem is called the direct/forward problem, while the latter is called the inverse problem”.



**Fig. 1.1.** General inverse problem diagram

To summarize, a forward problem is to deduce consequences (data) of a cause (model/image), while the inverse problem is to find the causes of a known consequence. Mathematically, given a linear problem  $Ax = b$ , the forward problem consists on computing the data (output) vector,  $b$ , given the system/forward matrix,  $A$ , and the model/image (input) vector,  $x$ . On the other hand, the inverse problem consists on computing the model/image vector,  $x$ , or the forward matrix,  $A$ , given the other two quantities, usually when the known data,  $b$ , have errors.

## 1.2. Ill-posed problems and uncertainty of solution

One problem is called ill-posed (or ill-conditioned) in the Hadamard sense (Hadamard 1902) when it does not satisfy at least one of the three conditions required for being well-posed (or well-conditioned)

- 1) Existence of the solution
- 2) Uniqueness of the solution
- 3) Stability of the solution; that is, continuous dependence of the solution on the data.

Note that we can refer to the non-uniqueness of a problem as the ambiguity of the problem.

The need of the third requirement has important consequences for the solution of ill-posed problems modelling physical situations, where the existence of noise or errors in the data may imply the presence of numerical instabilities in the solution.

Inverse problems are very often ill-posed problems (Hadamard 1902).

Without prior knowledge, ill-posed problems have several solutions (not necessary understandable/stable ones). Mathematically, this process of using prior knowledge to well-condition the problem is called regularization. More details about regularization will be provided in chapters 2 and 5 and throughout this thesis.

### 1.3. Fluorescence diffuse optical tomography (fDOT)

Fluorescence diffuse optical tomography (fDOT) is a technique that enables the quantification of tomographic (3D) bio-distributions of fluorescent tracers. fDOT technique is also called fluorescence molecular tomography (FMT) in some studies (Lorenzo 2012; Ntziachristos et al. 2002). Most fluorescent tracers in fDOT emit light in the NIR (Near Infrared, 600nm-900nm) spectrum.

fDOT has several interesting advantages as compared to other molecular imaging techniques (PET, SPECT), such as the use of non-ionizing radiation, or the lower price of the experimental setup.

The first demonstration that the NIR light can be used to monitor the state of cortical tissues non-invasively through the skull was presented by Jöbsis in 1977 (Jobsis 1977).

In 1988, Patterson et al. (Patterson, Chance, and C. 1989) experimentally showed that the propagation of light through biological tissues can be modelled by the diffusion equation.

The investigation on images produced by light propagation through tissues also attracted considerable attention (Boas 1996; Ntziachristos et al. 2000; O'Leary 1996).

The first in-vivo diffuse optical tomography study of the human breast after contrast agent administration was presented by Ntziachristos in 2000 (Ntziachristos et al. 2000).

During the last twenty years, a better understanding of photon propagation through tissues has allowed the researchers to gain a deeper understanding of the correlation between physiological internal changes and optical changes. This, together with the continuous development of NIR fluorescent probes for a wide variety of biological applications, has positioned fDOT as an emerging biomedical imaging technique. fDOT can three-dimensionally resolve markers that can be helpful in many fields: drug discovery (Lorenzo 2012; Ntziachristos, Leroy-Willig, and Tavitian 2007), carcinogenesis, protein

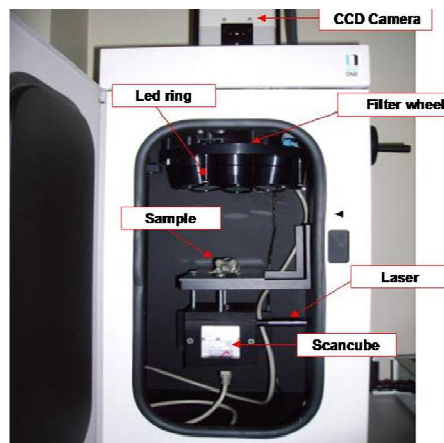
expression, protease activity, and receptor regulation (Garofalakis et al. 2007; Graves et al. 2004; Martin et al. 2008; Ntziachristos et al. 2007; Ntziachristos et al. 2004) .

### 1.3.1. *Experimental setup*

In this thesis all the experimental data were acquired with a non-contact parallel-plate system developed in our laboratory (Aguirre 2012). In the non-contact parallel-plate configuration the sample under study is gently compressed between two parallel anti-reflective transparent plates, thus achieving a slab-like geometry. This slab geometry is often assumed in optical mammography and small animal imaging (Graves et al. 2003; Ntziachristos et al. 2000; Schulz, Ripoll, and Ntziachristos 2004). In this thesis we focused on fDOT problems with slab geometry.

With our setting the excitation laser beam enters the sample perpendicularly through the first plate, and the transmitted light emerging through the opposite plate is recorded with a CCD camera. The light emerging from a constant-intensity laser diode is focused onto the sample at the desired points (source locations) using two mirrors moved by galvanometers, thus making it possible to choose the number and spatial distribution of sources. The laser emitter wavelength is set at  $675 \pm 5$  nm, and the power delivered to the sample is controlled by means of a TTL signal that modulates the laser duty cycle. Typical power values are in the 1-mW range. All the components of the set-up are placed inside a light-shielded box. The acquisition process is controlled by in-house software hosted on a PC workstation.

Figure 1.2 shows our non-contact parallel-plate fDOT experimental setup.



**Fig. 1.2.** fDOT non-contact parallel-plate experimental setup.



Emission fluorescence images are recorded by placing a 10-nm bandwidth filter centered at 720 nm in front of the camera lens, while for transmission excitation images a 10-nm bandwidth filter centered at 675 nm is used. These filters are placed in front of the camera using a motorized wheel (see Fig. 1.2). The acquired images are normalized by their respective laser power.

For each source, a variable number and distribution of detectors can be defined over the CCD sensor field of view (FOV), thus making it possible to retrieve the fluorescent and excitation average intensity at the desired points on the sample surface.

### ***1.3.2. Phantoms***

In order to perform the experimental testing of our algorithms, we studied how to build phantoms that simulate the optical properties of biological media. In (Firbank, and Delpy 1993) the authors described how to prepare an agar-based phantom using India ink and intralipid. The India ink mimics the absorption while the intralipid simulates the scattering. The role of the agar is just to solidify the phantom.

The needed concentration of agar for different concentrations of intralipid can be found in (Firbank, and Delpy 1993).

We also built agar-based phantoms using India ink and  $\text{TiO}_2$ . In these phantoms  $\text{TiO}_2$  simulates the scattering.

However, agar dissipates in less than twenty-four hours, making the life of this kind of phantoms very short and compromising the repetitiveness of the studies. For this reason, we also developed polyester resin-based phantoms with India ink and  $\text{TiO}_2$ . The necessary amounts of the components are described in (Boas 1996).

In this thesis we used both type of phantoms. It should be emphasized, however, that the results are largely independent of the type of phantom used.

### ***1.3.3. Forward problem***

A theoretical model (forward problem) that predicts photon propagation through the diffusive medium is needed before trying to solve fDOT inverse problems.

At NIR wavelengths, scattering of photons is a more significant attenuation component than absorption (Ripoll et al. 2008). In order to model the forward problem, we take into account that, in highly scattering media where light scattering dominates over absorption, light propagation complies with the diffusion equation (Ishimaru 1978)

$$(-\nabla D \nabla + \mu_a) \phi(r) = S(r'), \quad (1.1)$$

where  $D(r, \lambda) = [3(\mu_a(r, \lambda) + \mu_s'(r, \lambda))]^{-1}$  is the diffusion coefficient for a wavelength  $\lambda$  at position  $r$  in a domain  $\Omega$ ,  $\mu_s'(r, \lambda)$  the reduced scattering term,  $\mu_a(r, \lambda)$  the absorption term,  $\phi(r)$  the average intensity, and  $S(r')$  the source term at position  $r'$ .

In fDOT, the excitation intensity  $\phi_{ex}(r, \lambda_{ex})$  at excitation wavelength  $\lambda_{ex}$  and emission intensity  $\phi_{em}(r, \lambda_{em})$  at emission wavelength  $\lambda_{em}$  are given by a pair of diffusion equations (Abascal et al. 2012; Graves et al. 2003; Hyde et al. 2009; Lorenzo 2012; Ntziachristos, and Weissleder 2001).

The excitation intensity is emitted by an external source  $q_0(r_s)$  at a location  $r_s \in \Omega$ , and the emission comes from a fluorescent region characterized by a fluorescence yield  $f(r_{fl})$ , which accounts for its quantum efficiency, its absorption parameter, and its concentration of fluorescence.

Assuming that the presence of the fluorophore does not affect the absorption coefficient and that we are working on the steady-state regime, excitation and emission intensities are given by

$$\begin{aligned} -\nabla D(r, \lambda_{ex}) \nabla \phi_{ex}(r, \lambda_{ex}) + \mu_a(r, \lambda_{ex}) \phi_{ex}(r, \lambda_{ex}) &= q_0(r_s) \\ -\nabla D(r, \lambda_{em}) \nabla \phi_{em}(r, \lambda_{em}) + \mu_a(r, \lambda_{em}) \phi_{em}(r, \lambda_{em}) &= f(r_{fl}) \phi_{ex}(r, \lambda_{ex}) \end{aligned} \quad (1.2)$$

These diffusion equations can be solved using Green's function for a homogeneous medium and canonical geometries (Abascal et al. 2012; Graves et al. 2003; Hyde et al. 2009; Lorenzo 2012; Ntziachristos, and Weissleder 2001).

Generalizing, we define a Green function that solves the heterogeneous problem

$$[-\nabla D(r) \nabla + \mu_a(r)] G(r, r') = \delta(r - r') \quad (1.3)$$

Using this function, the average intensity–solving equations (1.2) become

$$\begin{aligned}\phi_{ex}(r) &= \int dr' G(r, r') q_0(r') \\ \phi_{em}(r) &= \int dr' G(r, r') f(r') \phi_{ex}(r')\end{aligned}\quad (1.4)$$

The *normalized Born approximation* (Born, and Wolf 1999), defined as the quotient between the fluorescence measurement and the excitation measurement for each source-detector pair, is applied to the data, as follows:

$$d_b(r_d) = \frac{\phi_{em}^{meas}(r_d)}{\phi_{ex}^{meas}(r_d)} = \frac{\int dr' G(r_d, r') f(r') \phi_{ex}(r')}{\int dr' G(r_d, r') q(r')} = \frac{\int dr' G(r_d, r') f(r') \int dr'' G(r', r'') q(r'')}{\int dr' G(r_d, r') q(r')} \quad (1.5)$$

To compute the matrix of the linear system,  $W$ , we first differentiate the emission average intensity with respect to  $f$  and discretize the integral as a sum of all finite elements  $\Omega_j$ . The variation of emission average intensity is given by

$$\delta\phi_{em}(r_d) = \sum_j \int_{\Omega_j} dr_j G(r_d, r_j) \phi_{ex}(r_j, r_s) \delta f_j = \sum_j \int_{\Omega_j} dr_j \tilde{\phi}(r_j, r_d) \phi_{ex}(r_j, r_s) \delta f_j \quad (1.6)$$

where  $\phi_{ex}(r_j, r_s)$  is excitation average intensity at  $r_j$  induced by a source at  $r_s$ ,  $\tilde{\phi}(r_j, r_d)$  is the adjoint field at  $r_j$  given by a source  $\tilde{q}_0$  located at the detector position  $r_d$ , taking into account the reciprocity of the Green function. The adjoint field solves the equation

$$[-\nabla D(r) \nabla + \mu_a(r)] \tilde{\phi}(r) = \tilde{q}_0(r_d) \quad (1.7)$$

The element  $ij$  of the forward matrix  $W_{ij}$  relates each measurement  $(d_b)_i$  (where  $i$  denotes each source-detector pair) to the concentration of fluorophore at element  $\Omega_j$  and can be written as

$$W_{ij} = \frac{\partial (d_b)_i}{\partial f_j} = \frac{1}{\phi_{ex}^{meas}(r_d, r_s)} \int_{\Omega_j} dr_j \tilde{\phi}(r_j, r_d) \phi_{ex}(r_j, r_s) \quad (1.8)$$

Combining the  $M \times N$  elements of the matrix  $W$ , the fDOT linear system can be expressed as

$$\begin{pmatrix} d^{nB}(r_{s_1}, r_{d_1}) \\ \vdots \\ d^{nB}(r_{s_M}, r_{d_M}) \end{pmatrix} = \begin{pmatrix} W_{11} & \cdot & \cdot & \cdot & W_{1N} \\ \vdots & & & & \\ \vdots & & & & \\ \vdots & & & & \\ W_{M1} & \cdot & \cdot & \cdot & W_{MN} \end{pmatrix} \begin{pmatrix} f(r_1) \\ \vdots \\ \vdots \\ f(r_N) \end{pmatrix} \quad (1.9)$$

Another way to solve the diffusion equations described by equations (1.2) is using a Garlekin finite element approach. For this latter option we adapted TOAST, a finite elements toolbox for DOT (Schweiger 1994; Schweiger et al. 1995), to fDOT.

#### 1.3.4. Simulated data

We also used the TOAST toolbox (Schweiger 1994; Schweiger et al. 1995) adapted for fDOT (introduced in precedent section) to simulate excitation and fluorescent photon densities and construct the forward matrix. Sources are modelled as isotropic point sources (located at a depth  $1/\mu_s$  below the surface) using Dirichlet boundary conditions. This setting resembles a collimated laser as described in (Schweiger et al. 1995). Measurements were modelled by a Gaussian kernel centered at the detector location and computed as a linear operator  $M$  acting on the photon density at the boundary of the domain. Thus, measured excitation and emission photon densities at the detector position,  $r_d$ , become  $\phi_{ex}^{meas}(r_d) = M\phi_{ex}(r)$  and  $\phi_{em}^{meas}(r_d) = M\phi_{em}(r)$ . Afterwards, similarly to equations (1.5), we calculated the normalized data component  $\left( d_b(r_d) = \frac{\phi_{em}^{meas}(r_d)}{\phi_{ex}^{meas}(r_d)} \right)$  and finally solved the linear system matrix as described in equations (1.6-1.9).

### 1.4. Solving the fDOT inverse problem

The fDOT inverse problem focuses on finding the fluorescence concentration at each voxel given the acquired data and the forward matrix.

### *1.4.1. fDOT, an ill-posed problem*

One of the main issues of fDOT is that the high degree of absorption and scattering of light through the biological tissues leads to a severely ill-posed inverse problem, and reduces the accuracy of the localization of fluorescent targets (Arridge, and Schotland 2009; Dutta et al. 2012; Egger, Freiburger, and Schlottbom 2010; Lorenzo 2012; Markel, and Schotland 2002). Furthermore, fDOT problems involve a large number of unknowns together with a limited set of measurements (Arridge, and Schotland 2009).

The fDOT inverse problem yields multiple non-unique and unstable solutions to the reconstruction problem.

Although the data at our disposal may contain a remarkable amount of information, the ill-posedness of the problem combined with the presence of noise implies that solving the inverse problem is not trivial. Thus, the image reconstruction is highly susceptible to the effects of noise and numerical errors, in which case appropriate priors or penalties are needed to stabilize the reconstruction images.

### *1.4.2. Non-contact fDOT, large datasets*

Initial fDOT systems worked delivering laser light with optical fibers to different points of the surface of the sample under study (sources). The laser was selected at a wavelength appropriate to excite a fluorescent contrast agent. For each source, the outgoing excitation and fluorescent intensity were separately collected, placing fiber-optics detectors at several points around the surface of the sample.

Once the nature of the diffuse light travelling from the sample to a separated detector was better understood and modelled, fDOT setups evolved towards non-contact geometries, using CCD cameras as detectors in a setup described in section 1.3.1. This arrangement retrieves the average light intensity at different points on the surface of the animal (virtual detectors), and also allows us to focus the excitation laser beam directly into different points of the sample. Such systems generate datasets that are orders of magnitude larger than those acquired with fiber-based systems, leading to significant computational challenges for image reconstruction (Arridge, and Schotland 2009).

The size of the forward matrix is  $m \times n$ , where  $m$  is the product of number of detectors ( $N_d$ ) by number of sources ( $N_s$ ), and  $n$  is the number of elements of mesh. For contact geometries, this number ranges between  $10^2 - 10^3$ , whereas for non-contact fDOT setups the size of the forward matrix can easily reach  $10^7 - 10^9$ .

### ***1.4.3. Reconstruction methods***

Different methods have been proposed to solve the fDOT inverse problem. We can classify them into two large categories: linear and non-linear methods. The first group comprises methods such as Tikhonov, Newton method, Landwever and Steepest descent, Krylov methods, and Kaczmark methods. The second group includes methods such as Gauss-Newton, Levenberg-Marquardt, and non-linear Kaczmarz methods, amongst others (Arridge, and Schotland 2009; Egger et al. 2010).

Another way to classify the different reconstruction algorithms is according to the need of a regularization functional. In this case we would have: a) methods on which a regularization functional can be included explicitly in the minimization term, and b) methods that do not include a regularization functional. The first group (a) comprises methods such as direct linear Tikhonov regularization (which might require extensive computational resources in terms of creation and storage of forward matrix), or non-linear gradient-based methods, such as Gauss Newton (which may be unfeasible in terms of Hessian matrix calculus when dealing with large datasets). The second group (b) includes methods such as iterative Kaczmarz methods, or free-matrix methods (Arridge, and Schotland 2009; Egger et al. 2010).

When a regularization functional can be explicitly included in the minimization term, it can be introduced in terms of  $l_2$ -based norms,  $l_1$ -based norms, total variation or combinations of them. The choice of regularization is discretionary and reflects a prior knowledge about the system.

## Chapter 2

# Linear $l_2$ regularization reconstruction methods

One of the difficulties of ill-posed problems is that there is not a single and well-behaved solution. In order to obtain a stable solution which is not too sensitive to the perturbations that approximate the desired solution, the problem must be regularized.

In this chapter we first introduce the need for regularization of a general problem. Afterwards, we present the singular value decomposition (SVD) of a forward matrix to finally introduce some  $l_2$  regularization reconstruction methods such as truncated singular value decomposition (TSVD) or Tikhonov regularization.

This chapter is mainly based on references (Hansen 2010; Vogel 2002).

### 2.1. The need for regularization

Given a linear system  $Ax = b$ , a larger condition number of  $A$ ,  $\text{cond}(A)$ , indicates that the system is sensitive to perturbations of right-hand side,  $b$ . Discrete ill-posed problems are characterized to have coefficient matrices with large condition numbers.

Let's suppose that we have the exact solution from an ill-posed problem  $Ax = b$ , and we name it,  $x_{\text{exact}}$ . Suppose that  $x_{\text{exact}}$  and  $x$  satisfy  $Ax_{\text{exact}} = b_{\text{exact}}$  and  $Ax = b = b_{\text{exact}} + \varepsilon$ , where  $\varepsilon$  designates the perturbation or error. In this case we have

$$\frac{\|x_{exact} - x\|_2}{\|x_{exact}\|_2} \leq \text{cond}(A) \frac{\|\varepsilon\|_2}{\|b_{exact}\|_2}, \text{ i.e., } \frac{\|\Delta x\|_2}{\|x_{exact}\|_2} \leq \text{cond}(A) \frac{\|\Delta b\|_2}{\|b_{exact}\|_2}, \quad (2.1)$$

being  $\Delta x$  and  $\Delta b$  the data ( $x$ ) perturbation and the right-hand side ( $b$ ) perturbation, respectively.

If  $\text{cond}(A)$  is large,  $x$  can be very far from  $x_{exact}$ , even when the perturbation is small ( $\|\varepsilon\|_2 \ll \|b_{exact}\|_2$ ) (Hansen 2010; Vogel 2002). In these cases we need regularization methods to obtain less sensitive and stable solutions (good approximations to  $x_{exact}$ ).

## 2.2. Singular Value Decomposition

Singular Value Decomposition (SVD) is a powerful tool for analyzing discrete inverse problems.

### 2.2.1. The role of SVD

If we have a linear system

$$Ax = b, \quad (2.2)$$

where  $A$  is an invertible matrix, we can express the solution  $x$  in terms of SVD as

$$x = \sum_{i=1}^n \frac{u_i^T b}{\sigma_i} v_i. \quad (2.3)$$

Furthermore, the  $l_2$ -norm of a matrix  $A$  can be expressed in terms of SVs as  $\|A\|_2 = \sigma_{\max}$ , where  $\sigma_{\max}$  denotes the maximum of the SVs of  $A$ . Similarly, we can express the  $l_2$ -norm of inverse matrix of  $A$  ( $A^{-1}$ ), as  $\|A^{-1}\|_2 = \sigma_{\min}^{-1}$ , where  $\sigma_{\min}$  denotes the minimum of the SVs of  $A$ . Thus, it follows that the condition number (CN) of  $A$  is given by the ratio between the largest and the smallest nonzero singular value (Hansen 2010; Vogel 2002),

$$\text{cond}(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_{\max}}{\sigma_{\min}}. \quad (2.4)$$



### 2.2.2. SVD for ill-posed problems

For any linear system  $Ax = b$ , the subspace of  $b$  that can be calculated by  $x$  is called the column space or range of  $A$  ( $Rang(A)$ ). In other words, we can define  $Rang(A)$  as the subspace spanned by the columns of  $A$ ,

$$Rang(A) = span\{a_1, \dots, a_n\} = \{b \in \mathbb{R}^m \mid b = Ax, x \in \mathbb{R}^n\}. \quad (2.5)$$

When dealing with ill-posed problems, we say that  $A$  is a singular matrix, that is, a part of the vector  $x$  projects to zero. This part is called the null space or kernel of  $A$  ( $Ker(A)$ ),

$$Ker(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}. \quad (2.6)$$

If  $Ker(A)$  exists, the solution of the system  $Ax = b$  is not unique, and its solutions are a combination of the general solution with some linear combinations of its null space.

If we decompose the matrix  $A$  by SVD, every singular value close or equal to zero corresponds to a singularity of  $A$ . In this case, the rate of decrease of the SVs,  $\sigma_i$ 's, is an indication of the ill-posedness of the problem. We say that a problem is ill-posed if both the following criteria are satisfied:

- 1) The SVs of  $A$  decay gradually to zero.
- 2) The ratio between the largest and the smallest nonzero singular value (CN) is large.

If condition 2 is fulfilled, by examining equation 2.4 we can see that if there is only one singular value of  $A$  close to zero, the CN of  $A$  becomes very large and the system is ill-conditioned. Furthermore, the null space of  $A$  is given by columns of  $V$  corresponding to the zero SVs and the range of  $A$  is given by columns of  $U$  corresponding to the non-zero SVs (Vogel 2002). The columns of  $V$  corresponding to the zero SVs are vectors having several changes of sign that increase the instability of the solution.

If  $A$  is a singular matrix, when  $\sigma_i = 0$  (or close to zero), we have problems with  $1/\sigma_i$  and extremely large errors in the solution to equation 2.3 arise.

## 2.3. Truncated singular value decomposition

One straightforward regularization technique is the truncated singular value decomposition solution (TSVD) (Hansen 1987), also called selective singular value solution. The basic idea in TSVD is to include only the SVD components corresponding to the largest SVs, i.e., those which make significant contributions to the regularized solution.

To do this, we can set  $\sigma_i^{-1} = 0$  when  $\sigma_i = 0$ ; that is, to reject the null space. Alternatively, we can sum all the components of the solution  $x$  in terms of SVD (equation 2.3) for which the absolute value of the right-hand side of SVD coefficient ( $u_i^T b$ ) is above a certain threshold ( $\tau$ ) (Vogel 2002)

$$x_\tau = \sum_{|u_i^T b| > \tau} \frac{u_i^T b}{\sigma_i} v_i. \quad (2.7)$$

Although TSVD is a very intuitive regularization technique, it requires to calculate the SVD, or at least its  $\tau$  first components. In general, it is best to avoid this calculation for large problems, unless SVD gives us some interesting information, as we will see in chapters 4 and 7.

## 2.4. Tikhonov Regularization

One of the most common regularization methods in the field of inverse problems is the Tikhonov regularization (Golub, and Matt 1997; Golub, and Van Loan 1996). It explicitly incorporates a regularization term, and its solution is given by the following functional minimization problem

$$\min_x \left\{ \|Ax - b\|_2^2 + \alpha \|x\|_2^2 \right\}, \quad (2.8)$$

where the first term  $\|Ax - b\|_2^2$  is the least squares problem, which is equivalent to fitting the predicted data to real data, while the second term  $\|x\|_2^2$  is a penalty term that stabilizes the solution forcing it to have a small  $l_2$ -norm, being  $l_2$ -norm the Euclidian norm

$(\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2})$ . The parameter  $\alpha$ , termed the regularization parameter, achieves a balance between both terms.

Since the approximated solution is governed by high frequencies, the incorporation of the second term,  $\|x\|_2^2$ , has the aim of suppressing most of the large high frequency noise components.

The Tikhonov solution,  $x_{Tik}$ , is obtained by equalling to zero the gradient of equation 2.8,  $\nabla \left( \|Ax - b\|_2^2 + \alpha \|x\|_2^2 \right) = \nabla \left( (Ax - b)^T (Ax - b) + \alpha x^T x \right) = (A^T A)x - A^T b + \alpha x = 0$ .

Then  $x_{Tik} = (A^T A + \alpha^2 I)^{-1} A^T b$  and since  $A = USV^T$  and  $I = VV^T$ , we have

$$x_{Tik} = \left( (USV^T)^T (USV^T) + \alpha^2 VV^T \right)^{-1} V S U^T b = \dots = V \left( S^2 + \alpha^2 VV^T \right)^{-1} S U^T b.$$

In terms of the SVs, the solution can be expressed as

$$x_{Tik} = \sum_{i=1}^{\min(m,n)} \frac{\sigma_i}{\sigma_i^2 + \alpha^2} u_i^T b v_i = \sum_{i=1}^{\min(m,n)} \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \frac{u_i^T b}{\sigma_i} v_i. \quad (2.9)$$

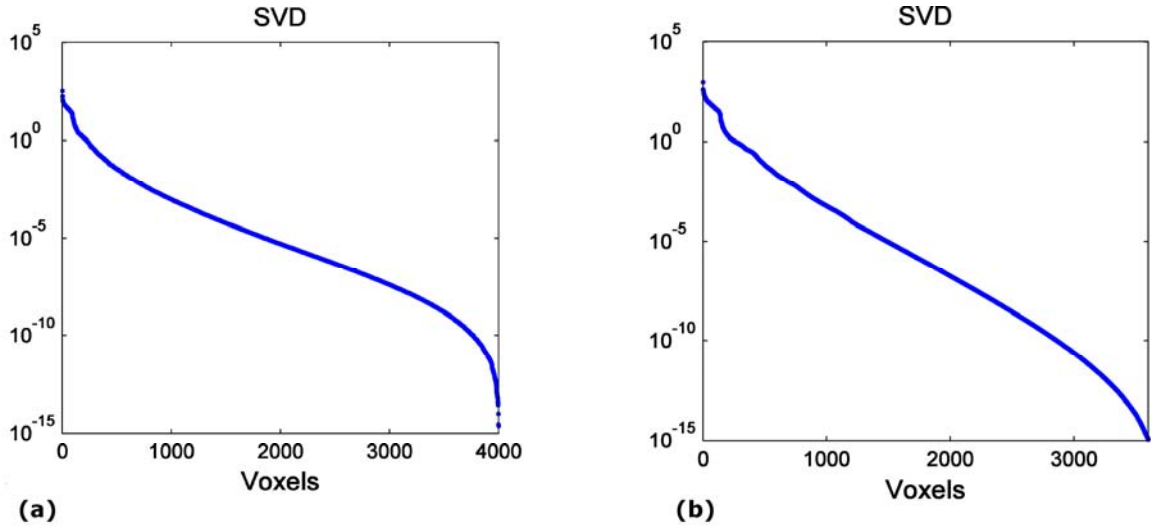
Note that Tikhonov regularization leaves the coefficients of the singular vectors that correspond to the large SVs almost unchanged, while it reduces the weight of the coefficients that correspond to the small SVs (noisy singular vectors).

Thus, it is important to remark that Tikhonov regularization, as opposed to TSVD, takes into account all the available information.

## 2.5. Dealing with fDOT

As mentioned in section 1.4.1, due to the highly absorbing and scattering properties of biological tissues, the fDOT inverse problem is ill-posed (Arridge, and Schotland 2009; Dutta et al. 2012; Egger et al. 2010; Lorenzo 2012).

An indication of the ill-posedness of the problem is the decay-rate of the SVs (section 2.2.2). Figure 2.1 plots representative values of this singular value decay of forward fDOT matrices corresponding to the noise-free simulated data that will be presented later in section 6.2.3.2, and for the ex-vivo mouse data presented in section 3.2.3.2.



**Fig. 2.1.** SVs of forward fDOT matrices. (a) Noise-free simulated data presented in section 6.2.3.2. (b) Ex-vivo mouse data presented in section 3.2.3.2.

We can observe the ill-conditioning nature of fDOT problem, since both examples satisfy the two criteria of ill-posedness presented in section 2.2.2. Note that SVs of forward fDOT matrix for ex-vivo mouse data decay faster than those for the noise-free simulated data. fDOT forward matrices have a very high conditioning numbers (around  $10^{17}$ ) that make image reconstruction highly susceptible to the effects of noise and numerical errors.

The key to obtaining a meaningful solution is to reformulate the fDOT problem in such a way that the new solution is less sensitive to perturbations. This is achieved by adding appropriate priors or penalties to facilitate the stability of the reconstruction. In other words, the problem must be stabilized or regularized.

## Chapter 3

# Choosing the regularization parameter for $l_2$ regularization reconstruction methods

When dealing with ill-posed problems such as fDOT, the choice of the regularization parameter is extremely important for computing a reliable  $l_2$  regularization reconstruction. Although the topic may seem trivial, it still receives high attention by researchers. Several automatic methods for the selection of the regularization parameter have been introduced over the years, but their performance highly depends on the particular inverse problem. Even we can say that there is no perfect choice for the regularization parameter, since each strategy has its advantages and disadvantages.

In this chapter, a U-curve-based algorithm for the selection of the regularization parameter has been applied for the first time to fDOT. Reducing the regularization parameter search to a validity interval increases the computational efficiency for large systems. Using both phantom and ex-vivo mouse data we will show that the U-curve provides a suitable choice of the regularization parameter in terms of Picard's condition, image resolution and image noise. This chapter is organized as follows: First we present a brief review of the different methods for choosing the regularization parameter in the fDOT literature. Then, we introduce the U-curve-based method and the discrete Picard's

condition. Afterwards, we show the feasibility of the U-curve-based method for fDOT experimental data. Finally, we summarize the discussion and present the main conclusions.

### 3.1. Review of Current Methods

In many implementations, the Tikhonov regularization problem (equation 2.8) is solved by manually selecting the regularization parameter,  $\alpha$ . This is done using a sequence of regularization parameters and selecting the value that leads to best results, as judged by the user. Obviously, the procedure is highly subjective and time consuming. To overcome this problem, several automatic methods for selecting regularization parameters have been suggested over the years. We cite some examples classified into two groups:

- Strategies based on the calculation of the variance of the solution, which requires prior knowledge of the noise:

- The Unbiased predictive risk estimator method (UPRE) tries to minimize the unbiased predictive risk. Since the data noise is assumed to be random white noise of known variance, the precision achieved depends on the accuracy of the actual noise estimate.
- The Discrepancy principle method (DP) selects the  $\alpha$  value for which data norm is equal to the data variance.

- Strategies that do not need a-priori information:

- Generalized cross-validation (GCV) is a statistical method whose goal is to find a value  $\alpha$  such that  $Ax_\alpha$  approximates the exact data  $b_{exact}$  as well as possible. The  $\alpha$  provided by GCV criterion has to provide a solution,  $x_\alpha$ , that can fit the data using the smallest possible number of parameters, thereby minimizing the contribution for small singular values (SVs).
- L-curve is a log-log plot of the regularization solution norm  $\|b_\alpha\|_2$  versus the corresponding residual error norm  $\|Ax_\alpha - b\|_2$ , for  $\alpha > 0$ . It is named L-curve after the L-shape of the resulting plot. The key idea behind the L-curve criterion is to choose a value,  $\alpha_L$ , that corresponds to the corner of the L-curve, i.e., a point that

balances  $\|b_\alpha\|_2$  and  $\|Ax_\alpha - b\|_2$ . The corner is the point on the L-curve with maximum curvature (in log-log coordinates).

More details of these methods can be found in a survey paper that describes some of them (Hanke, and Hansen 1993; Hansen, and O'Leary 1993), and in the chapter 7 of the book (Vogel 2002). As Vogel emphasized in his book, some methods may perform better than others depending on the particular inverse problem, and some methods may even fail completely.

Regarding optical tomography (DOT, fDOT, time-resolved (fluorescence) diffuse optical tomography), the reported strategies for the selection of the regularization parameter have been: the manual selection (Graves et al. 2004), upon some acceptable reconstruction variability parameters (non-biased estimators: mean and standard deviation of the reconstruction) (Ducros et al. 2009); choosing from plots of the reconstruction spatial resolution versus estimator variance (Chaudhari et al. 2009); the L-curve method (Corlu 2007; Culver et al. 2003; Serdaroglu, Yazici, and Ntziachristos 2006); a variant of the L-curve method (Xu, Jin, and Bai 2009); and more recently, using a neighbourhood regularization method which yields similar results to L-curve when L-curve works and provides solutions where it fails (Li et al. 2012).

L-curve method has been extensively analyzed (Hansen 1992; Hansen, and O'Leary 1993) and applied to different areas. Recently, it has been reported that L-curve returns good regularization parameter values for electrical impedance tomography of the brain (Abascal et al. 2008) and for diffuse optical tomography (DOT) (Correia et al. 2009). In both studies several methods were compared arriving to the conclusion that in practice, selection methods without a-priori information, such as GCV and L-curve, were more robust and there was no significant difference between GCV and L-curve in terms of accuracy. The only relevant difference is that GCV is more computationally expensive than L-curve for large systems (Busby, and Trujillo 1997).

Regarding DOT, one work (Culver et al. 2003) emphasized that the L-curve analysis yields overly-smooth solution in some cases.

## 3.2. Method proposed

### 3.2.1. U-curve method

The U-curve is a plot of the sum of the inverse of the regularized solution norm  $\|f_\alpha\|_2$  and the corresponding residual error norm  $\|Wf_\alpha - d\|_2$ , for  $\alpha > 0$  on a log-log scale

$$U_{curve}(\alpha) = \frac{1}{\|Wf_\alpha - d\|_2^2} + \frac{1}{\|f_\alpha\|_2^2}. \quad (3.1)$$

The U-curve is a U-shaped function, where the sides of the curve correspond to regularization parameters for which either the solution norm or the residual norm dominates. The optimum value of  $\alpha$ ,  $\alpha_u$ , is the value for which the U-curve has a minimum.

The regularized solution norm ( $\|f_\alpha\|_2$ ) and the corresponding residual error norm ( $\|Wf_\alpha - d\|_2$ ) are calculated numerically.

The U-curve method has been proposed by (Krawczyk-StańDo, and Rudnicki 2007, 2008; Krawczyk-StańDo, Rudnicki, and StańDo 2008) for the selection of the regularization parameter in inverse problems. The U-curve was tested on some numerical examples of the Fredholm integral equation of first kind (Krawczyk-StańDo, and Rudnicki 2007, 2008; Krawczyk-StańDo et al. 2008) and super-resolution problems (Qiangqiang et al. 2010).

#### 3.2.1.1. Desirable interval of the regularization parameter

It can be shown, as in (Krawczyk-StańDo, and Rudnicki 2007), that the function  $U_{curve}(\alpha)$  is strictly decreasing in the interval  $\alpha \in (0, \sigma_r^{2/3})$  and strictly increasing in the interval  $\alpha \in (\sigma_0^{2/3}, \infty)$ , where  $\sigma_0 \geq \sigma_1 \geq \dots \geq \sigma_r > 0$  are the SVs.

For simplicity, we will call  $E(\alpha) = \|Wf_\alpha - d\|_2^2$ , the residual error norm, and  $R(\alpha) = \|f_\alpha\|_2^2$ , the regularized solution norm. Thus,  $U_{curve}(\alpha) = \frac{1}{E(\alpha)} + \frac{1}{R(\alpha)}$  and its first derivative is



$$U_{curve}'(\alpha) = \frac{-E'(\alpha) * R^2(\alpha) - R'(\alpha) E^2(\alpha)}{(E(\alpha) * R(\alpha))^2}.$$

Using the SVD of  $W$  the residual error norm and the regularized solution norm can be expressed respectively as follows

$$E(\alpha) = \sum_{i=1}^r \frac{\alpha^4 d_i^2}{(\sigma_i^2 + \alpha^2)^2} \quad (3.2)$$

$$R(\alpha) = \sum_{i=1}^r \frac{\sigma_i^2 d_i^2}{(\sigma_i^2 + \alpha^2)} \quad (3.3)$$

Regarding the sign of  $U_{curve}'(\alpha)$  is equivalent to considering the factor  $\alpha^3 - \sigma_i^2$ , see demonstration below:

- $E'(\alpha) = -\alpha^2 R'(\alpha)$  by (Hansen, and O'Leary 1993).
- $U_{curve}'(\alpha)$  can be written as  $U_{curve}'(\alpha) = \frac{-R'(\alpha)(E(\alpha) + \alpha R(\alpha))(E(\alpha) - \alpha R(\alpha))}{(E(\alpha) * R(\alpha))^2}$

where  $\frac{-R'(\alpha)(E(\alpha) + \alpha R(\alpha))}{(E(\alpha) * R(\alpha))^2} > 0$ .

Thus, the sign depends only on  $E(\alpha) - \alpha R(\alpha)$ ; i.e., it depends on  $\sum_{i=1}^r \frac{\alpha d_i^2 (\alpha^3 - \sigma_i^2)}{(\alpha^2 + \sigma_i^2)^2}$

(equations 3.2 and 3.3), and  $\sum_{i=1}^r \frac{\alpha d_i^2}{(\alpha^2 + \sigma_i^2)^2} > 0$ .

Therefore,

$$\alpha^3 - \sigma_i^2 > 0 \Leftrightarrow \alpha > \sigma_i^{2/3} \text{ and } \alpha^3 - \sigma_i^2 < 0 \Leftrightarrow \alpha < \sigma_i^{2/3}.$$

In conclusion,

$$U_{curve}'(\alpha) < 0 \quad \text{if} \quad \alpha \in (0, \sigma_r^{2/3}) \quad \text{and} \quad U_{curve}'(\alpha) > 0 \quad \text{if}$$

$\alpha \in (\sigma_r^{2/3}, \infty)$  and the function  $U_{curve}(\alpha)$  is strictly decreasing on the interval  $\alpha \in (0, \sigma_r^{2/3})$

and strictly increasing on the interval  $\alpha \in (\sigma_0^{2/3}, \infty)$ .

Expressing  $U_{curve}(\alpha)$  on SVD terms, it can be proven that the function  $U_{curve}(\alpha)$  accomplishes  $\lim_{\alpha \rightarrow 0} U_{curve}(\alpha) = +\infty$  and  $\lim_{\alpha \rightarrow +\infty} U_{curve}(\alpha) = +\infty$  (Krawczyk-StańDo, and Rudnicki 2007).

From these two results, it can be concluded that the function  $U_{curve}(\alpha)$  has a local minimum in the interval  $\alpha \in (\sigma_r^{2/3}, \sigma_0^{2/3})$ .

The computation of the U-curve for values out of this interval is not necessary. Thus, the use of this interval can greatly increase the computational efficiency in selecting the regularization parameter.

### 3.2.1.2. Unicity of U-curve solution

Krawczyk-Stando (Krawczyk-StańDo, and Rudnicki 2007) pointed out that, if in the SVD there are one or more non-zero values, we can analytically calculate a unique  $\alpha > 0$  for which the U-function will reach a minimum, and this would be the only minimum of the function.

Finally, we can conclude that U-curve provides a desirable interval where the existence and unicity of a minimum can be proven (section 3.2.1.1.). This minimum corresponds to the optimum regularization parameter,  $\alpha_u$ .

### 3.2.2. The role of Discrete Picard's condition

As commented in chapter 2, the rate of decrease of the SVs,  $\sigma_i$ 's, is an indication of the ill-posedness of the problem. The Discrete Picard's condition (DPC) (Hansen 1990, 2010) provides us with an objective assessment of this fact.

The DPC is satisfied if the data space coefficients  $|u_i^T b|$ , on average, decay to zero faster than the respective SVs,  $\sigma_i$ 's.

The representation of  $|u_i^T b|$  and  $\sigma_i$  in the same plot is known as Picard's plot.

To compute a satisfactory solution by means of the Tikhonov regularization, DPC has to be fulfilled (Hansen 1990, 2010), since it determines how well the regularized solution approximates the unknown, exact solution.

In ill-posed problems, there may be a point where the data become dominated by errors and the DPC fails. In these cases, a suitable regularization term should fulfill the DPC.

Thus, the DPC itself can be used as a method for choosing the regularization parameter in Tikhonov regularization. Furthermore, it can be used as a tool to verify that other automatic methods provide a good choice for the Tikhonov regularization parameter.

### *3.2.3. Feasibility of the U-curve method for fDOT*

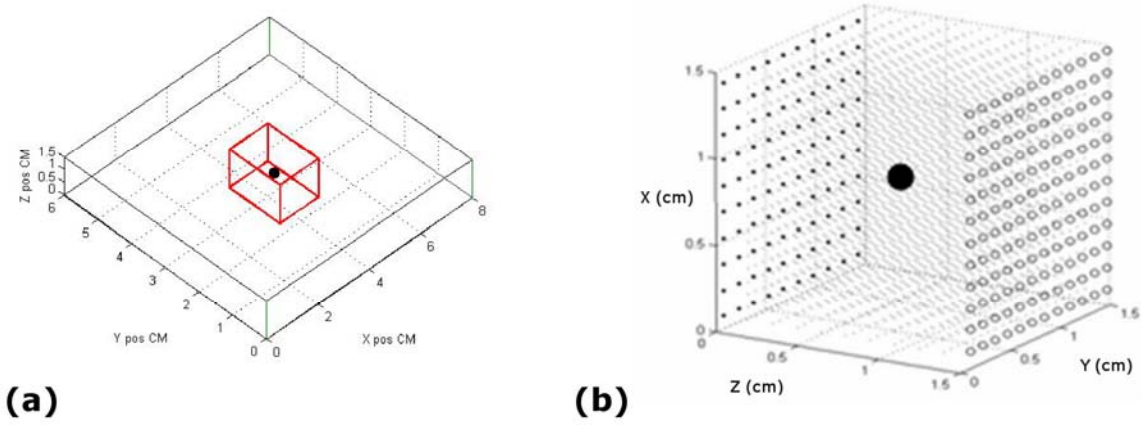
It is not possible to define a regularization parameter optimal for all the imaging applications, since the user may have different noise and resolution requirements according to each case. In this section, we study the feasibility of the U-curve method for fDOT, and evaluate its performance with phantom and real ex-vivo fDOT experiments, acquired with the fDOT experimental setup presented in section 1.3.1.

Furthermore, we validated this method by confirming that Picard's condition is fulfilled and inspecting the noise level of the reconstructed images to ensure that the U-curve method yields a satisfactory regularized solution.

#### *3.2.3.1. Phantom experiment*

We prepared a slab-shaped agar-based phantom (8 x 6 x 1.5 cm) using intralipid and India ink to obtain an absorption coefficient of approximately  $\mu_a = 0.3 \text{ cm}^{-1}$  and a reduced scattering coefficient of  $\mu_s = 10 \text{ cm}^{-1}$  (as described in section 1.3.2). A capillary with its tip filled with 6  $\mu\text{l}$  at 30  $\mu\text{M}$  of Alexa Fluor 700 (Invitrogen, Carlsbad, California, USA) was inserted into the phantom, with the tip positioned at the center of the slab (figure 3.1.a).

We built the fDOT forward matrix (as explained in section 1.3.3) based on 20x20x10 mesh points for a 1.5x1.5x1.5-cm FOV and equally spaced 10x10 sources and 12x12 detectors. The center of the mesh's FOV was aligned with the center of the slab. Figure 3.1 shows the three-dimensional mesh FOV for the reconstruction and how the sources and detectors are located in an area of [1.5x1.5] on the front and back plates (figure 3.1.b).



**Fig. 3.1.** a) Geometrical configuration of the slab (black) and the mesh FOV (red). The capillary tip is represented by the black sphere. b) Detail of the mesh's FOV. Sources are represented by black dots and detectors by black empty circles.

### 3.2.3.2. *Ex-vivo mouse experiment*

An euthanized mouse was imaged with a capillary inserted into the esophagus. The tip of the capillary (<1.5 mm thick) was filled with 6  $\mu$ l at 30  $\mu$ M of Alexa Fluor 700 (Invitrogen, Carlsbad, California, USA).

We constructed the fDOT forward matrix (as described in section 1.3.3) of with 20x20x10 mesh points for a 1.4x1.4x1.5-cm FOV centered on the chest of the mouse. Equally spaced 6x6 sources and 10x10 detectors were selected. The mouse was gently compressed between two transparent anti-reflective plates to a thickness of 1.5 cm approximately, in order to conform its geometry as much as possible to that of a slab. We define the axis x of the FOV along the width of the mouse, the axis y along the length of the mouse and the axis z along the antero-posterior dimension of the mouse.

### 3.2.3.3. *Validation of the regularization parameter obtained by the U-curve method*

The forward matrix was decomposed by SVD, and the images were reconstructed using Tikhonov regularization (section 2.4) for different  $\alpha$  parameters in the  $10^{-1}$  to  $10^{-6}$  range, which included the U-curve-based regularization parameter,  $\alpha_U$ .

We measured resolution and noise of the images for each  $\alpha$  value. To assess image resolution, we followed the procedure described in (Culver et al. 2003), assuming that the FWHM of the point spread function (PSF) of the capillary tip (that can be considered as a single isolated region) is directly related to the resolution performance of the system. The

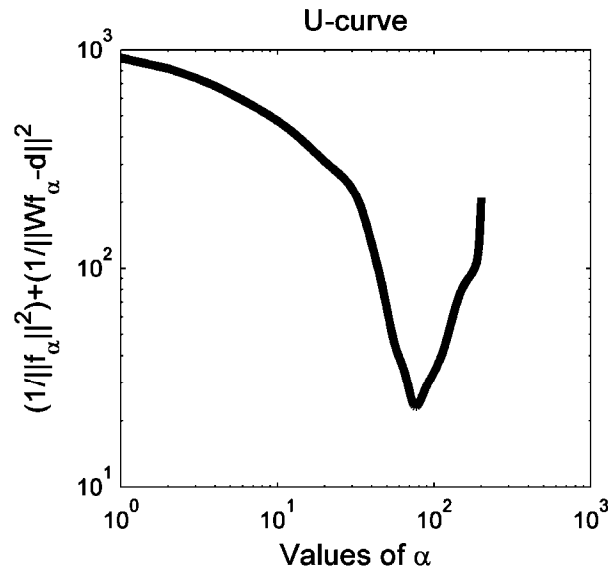
noise present in the images was measured as the standard deviation in a region of the image with no signal.

We have already explained in section 3.2.2, that fulfilment of the DPC is crucial for the existence of a meaningful solution to discrete ill-posed problems. To confirm that Picard's condition was fulfilled, we plotted on the same graph the  $|u_i^T d|$  coefficients, their corresponding SVs, and their quotient, using the Picard routine available in the Matlab Regularization Toolbox (Hansen 2007).

### 3.3 Results

#### 3.3.1 Phantom experiment validation

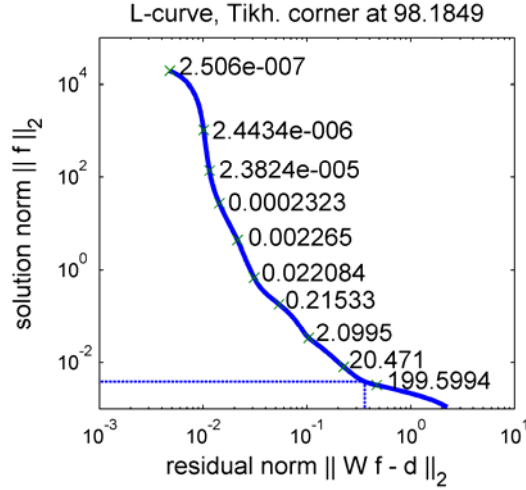
Figure 3.2 shows the U-curve plot on a log-log scale for the phantom experiment.



**Fig. 3.2.** U-curve plots on log-log scale. (Minimum corresponds to  $\alpha_u = 4.38 \times 10^{-2}$ ).

In this case, the U-curve shows a minimum which corresponds to a regularization parameter  $\alpha_u = 4.38 \times 10^{-2}$ .

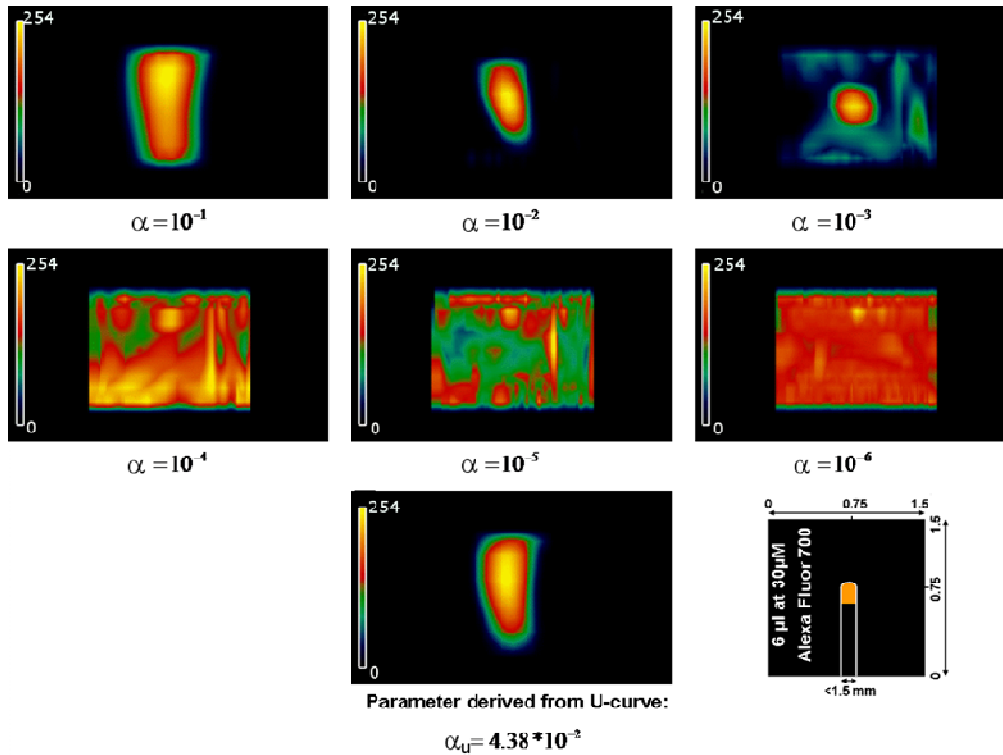
Figure 3.3 shows the L-curve plot on a log-log scale provided using (Hansen 2007):



**Fig. 3.3.** L-curve plot on log-log scale. (Maximum curvature at  $\alpha_L = 5.65 \cdot 10^{-2}$ ).

The L-curve did not exhibit a neat identifiable corner. The failure to find a sharp corner is due to the high ill-posedness of our problem. The value of  $\alpha_L$  was chosen as the point on the L-curve with maximum curvature.

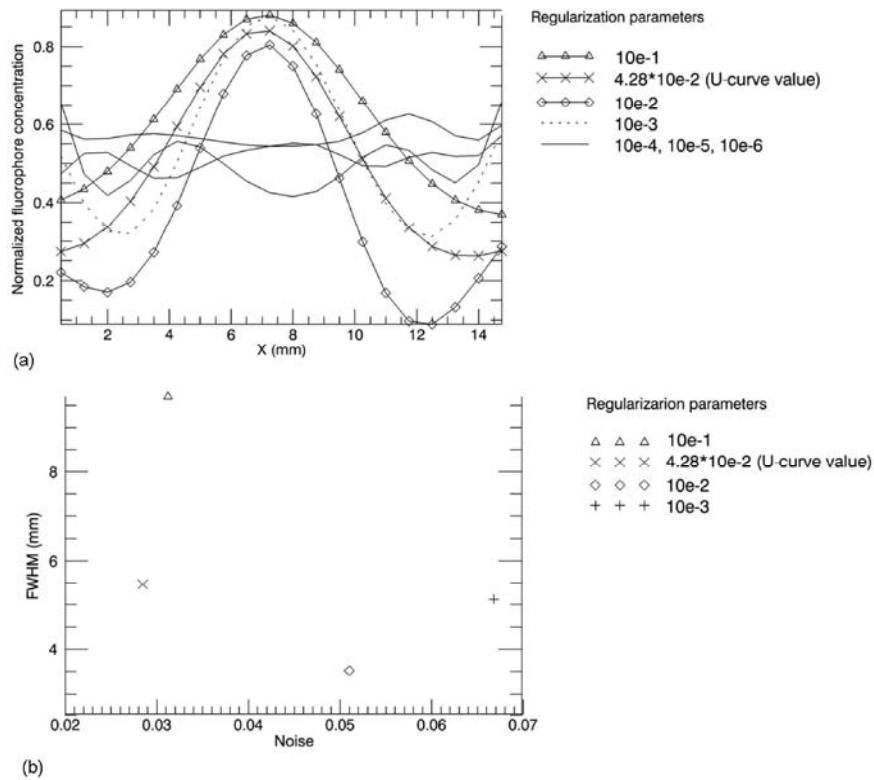
Figure 3.4 shows fDOT reconstructions for  $\alpha$  parameters in the  $10^{-1}$  to  $10^{-6}$  range and with  $\alpha_u = 4.38 \cdot 10^{-2}$ .



**Fig. 3.4.** Coronal view of maximum intensity projection 3D render of the reconstructions obtained for the  $\alpha$  parameter in the  $10^{-1}$  to  $10^{-6}$  range. The result for  $\alpha_u = 4.38 \cdot 10^{-2}$  (obtained from the U-curve) is showed at the bottom center. At the bottom right: drawing indicating the phantom fluorescence concentration.

The resolution versus noise is plotted in figure 3.5. This figure shows profiles taken in the x direction (corresponding to  $z=7.5$  mm and  $y=7.5$  mm) and the FWHM (Full width half maximum) versus noise.

It has been shown in (Culver et al. 2003) that the general trend for resolution is to increase together with image noise while the regularization parameter decreases, and this trend can be seen clearly in figure 3.5.b. For  $\alpha$  values of  $10^{-1}$ ,  $4.38 \cdot 10^{-2}$  ( $\alpha_u$ ), and  $10^{-2}$ , the FWHM of the profiles decreases. For  $\alpha=10^{-3}$  image noise begins to prevail, and for  $\alpha < 10^{-3}$  the reconstruction is noise only, and the object is no longer visible in the reconstructed images. According to these data, we observe a heuristic range of  $\alpha_u$  values that produces reconstructed images with a reasonable amount of noise and resolution, namely,  $10^{-1} \leq \alpha < 10^{-3}$ . This range includes the optimum value obtained by the U-curve method.



**Fig. 3.5.** a) Profiles taken in the x direction, corresponding to the line  $z=y=7.5$  mm, for each regularization parameter. b) FWHM (mm) vs. noise (%) plot.

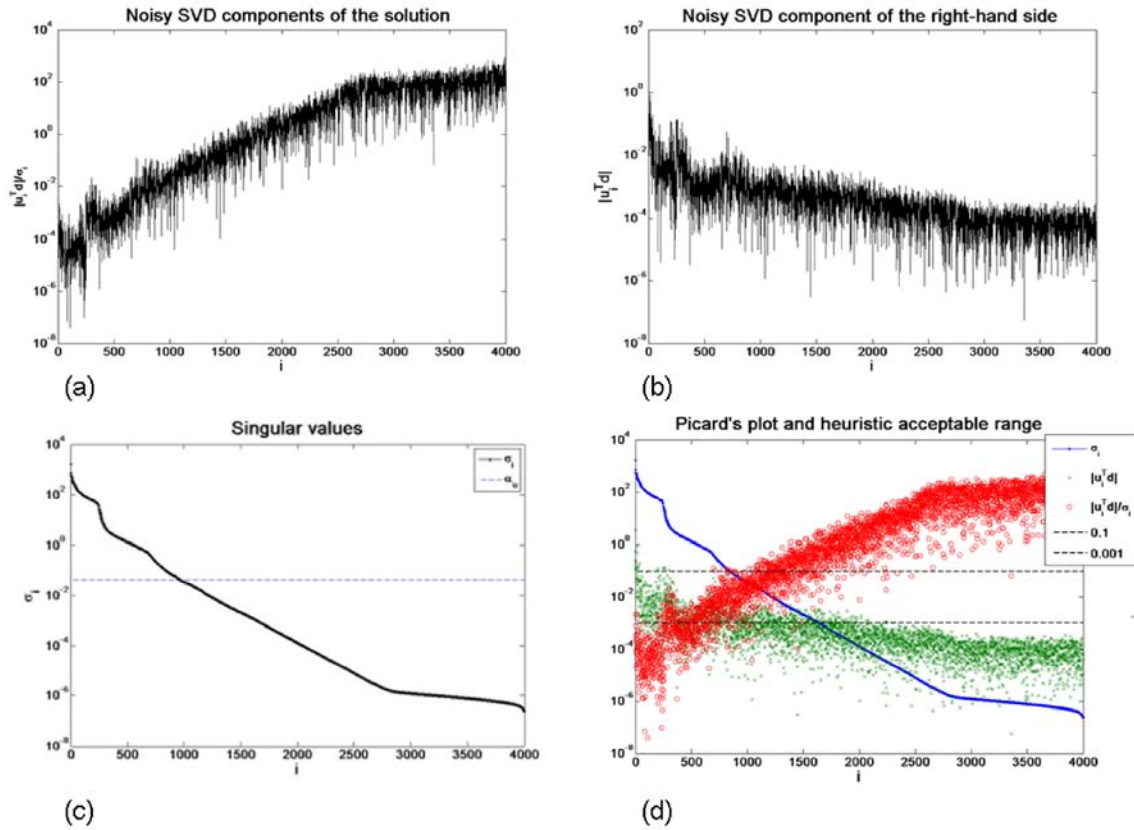
To corroborate that the U-curve regularization parameter leads to a suitable solution we verified that it satisfies the DPC. To this end, in figures 3.6.a and 3.6.b, we plot the noisy SVD components of the solution and the right-hand side of the phantom study. One

interesting aspect is the severe ill-posedness of the problem, indicated by the fact that the SVs decay gradually to zero and the ratio between the largest and the smallest nonzero singular value is large, as explained in section 2.2.2 (figure 3.6.c).

Figure 3.6.d illustrates Picard's plot showing the maximum and minimum values of the heuristically acceptable range plotted as two horizontal dashed lines ( $10^{-1}$  and  $10^{-3}$ ).

Picard's plot makes it possible to compare the SVD coefficients of the right-hand side with the SVs and their quotient. The data vector  $d$  is said to satisfy the DPC if the data space coefficients  $|u_i^T d|$ , on average, decay to zero faster than the SVs  $\sigma_i$  (section 3.2.2)

The blue line represents the decay of the SVs  $\sigma_i$ , the green crosses correspond to  $|u_i^T d|$ , and the red circles represent the quotient  $|u_i^T d|/\sigma_i$ .



**Fig. 3.6.** SVD components and Picard's plot for the phantom experiment. (a) Noisy SVD components of the solution. (b) Noisy SVD components of the right-hand side. (c) Decay of the SVs. The regularization parameter provided by U-curve method ( $\alpha_u = 4.38 \times 10^{-2}$ ) is plotted as a horizontal dashed blue line. (d) Picard's plot with the maximum and minimum parameter of the heuristic acceptable range plotted as two horizontal dashed black lines ( $10^{-1}$  and  $10^{-3}$ ).

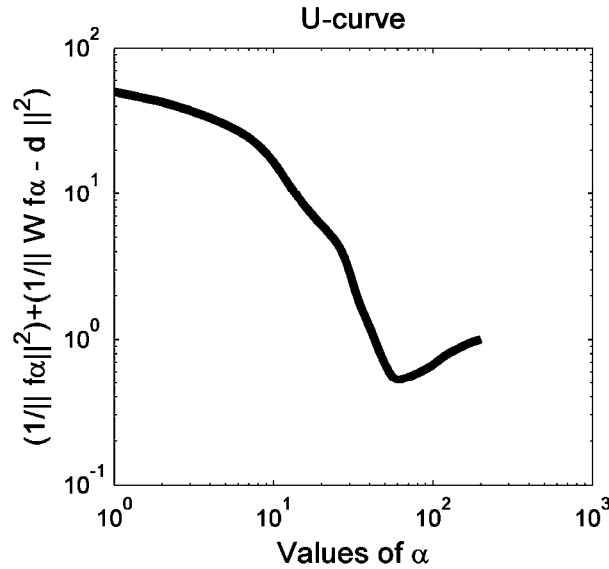


SVs below  $10^{-3}$ , on average, decay to zero faster than those of the respective  $|u_i^T d|$  coefficients.

We can observe that the SVs above the heuristic acceptable range of parameters  $10^{-1}$  and  $10^{-3}$ , and particularly the SVs above the U-curve cut-off ( $4.38 \cdot 10^{-2}$ ), fulfil Picard's condition.

### 3.3.2 Ex-vivo mouse data experiment validation

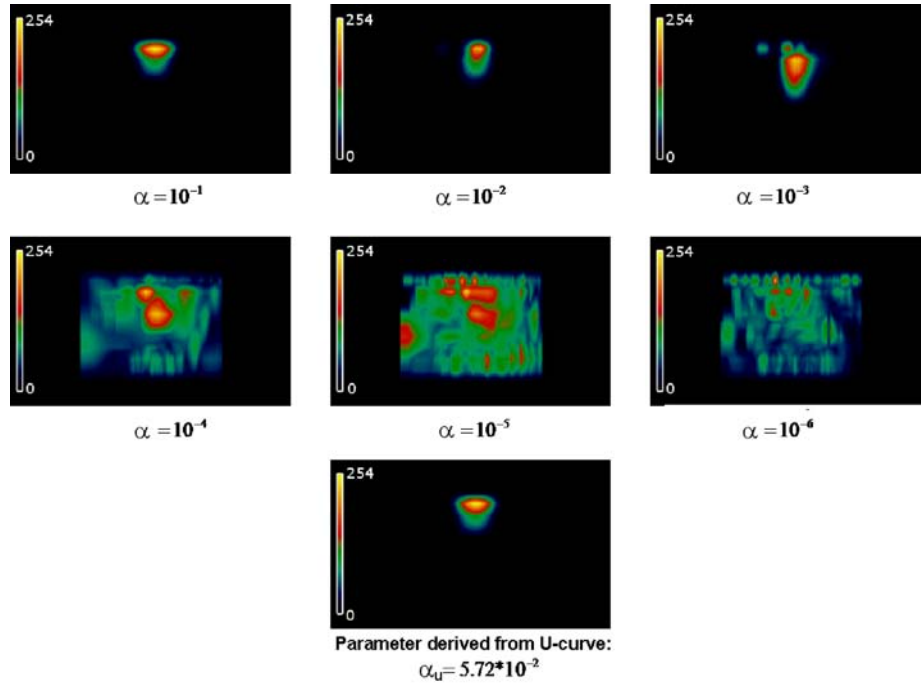
Figure 3.7 shows that the U-curve plotted on a log-log scale has a minimum at the regularization value  $\alpha_u = 5.72 \cdot 10^{-2}$ .



**Fig. 3.7.** U-curve plot in log-log scale. (Minimum corresponds to  $\alpha_u = 5.72 \cdot 10^{-2}$ ).

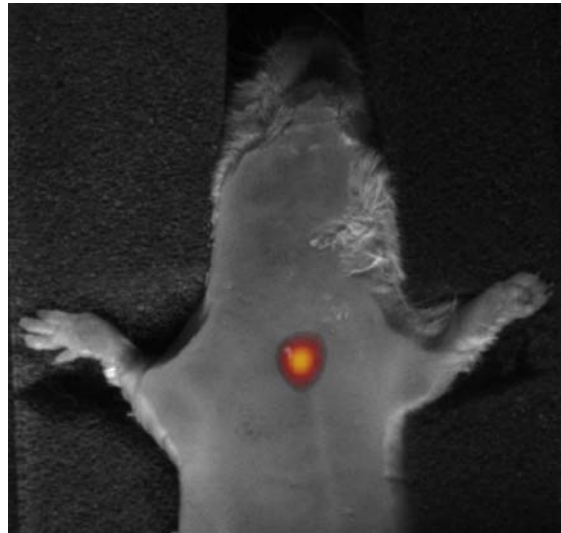
The curve is not really U-shaped, indicating that fewer useful SVs remain when compared to the phantom experiment, because the problem for the ex-vivo mouse data experiment is more ill-posed than the problem for the phantom experiment.

Figure 3.8 shows fDOT reconstructions with  $\alpha$  parameters in the  $10^{-1}$  to  $10^{-6}$  range and with  $\alpha_u = 5.72 \cdot 10^{-2}$ .



**Fig. 3.8.** Coronal maximum intensity projection render of the reconstructions obtained for  $\alpha$  parameters in the  $10^{-1}$  to  $10^{-6}$  range. The result for  $\alpha_u = 5.72 \cdot 10^{-2}$  (obtained from the U-curve) is shown at the bottom center. We used  $6\mu\text{l}$  of Alexa Fluor 700  $30\mu\text{M}$ .

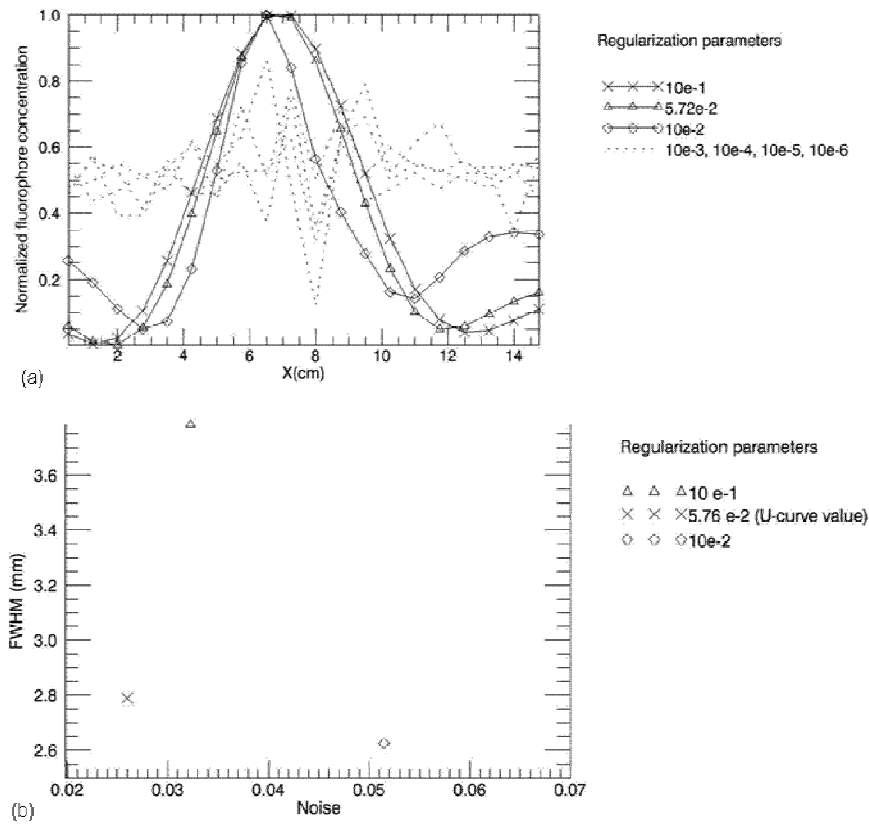
Figure 3.9 shows a coronal view of a 3D render of the Tikhonov reconstruction or the regularization parameter obtained using the U-curve method. The reconstruction is merged with the white light image of the mouse.



**Fig. 3.9.** Coronal of a maximum intensity projection render of the reconstruction for  $\alpha_u = 5.72 \cdot 10^{-2}$  (obtained from U-curve). The white light image is shown for reference image. Dye concentration and volume were  $6\mu\text{l}$  at  $30\mu\text{M}$  Alexa Fluor 700.

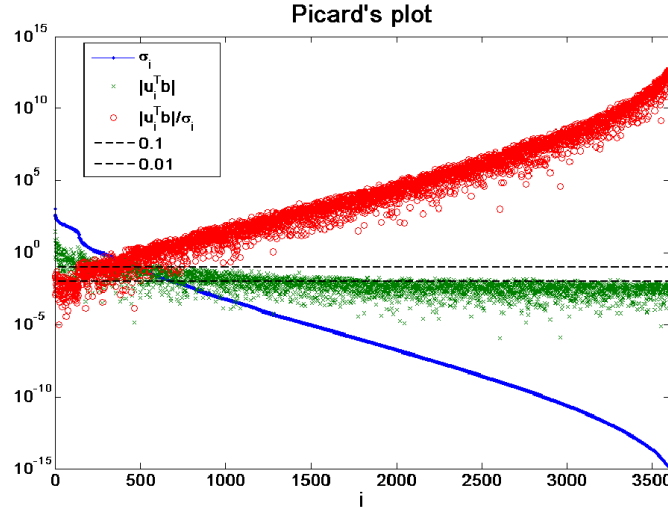
Figure 3.10.a shows profiles taken in the x direction, corresponding to  $z=8$  mm and  $y=0.7$  mm, and the FWHM (resolution) vs. noise plot, which outlines the behaviour of the resolution and image noise of the reconstructions depending on the regularization parameter. Again, we observe a range of  $\alpha$  values that produce reconstructed images with a reasonable compromise between noise and resolution. The U-curve-based value falls within this range, which is  $10^{-1} \leq \alpha \leq 10^{-2}$ .

In this case, we can confirm that the U-curve value leads to the image with minimum noise while retaining the best resolution possible.



**Fig. 3.10.** a) Profiles taken in the x direction, corresponding to the line  $z=8$  mm,  $y=0.7$  mm for each regularization parameter. b) FWHM (mm) vs. noise (%) plot.

Figure 3.11 shows how SVs above the heuristic acceptable range of parameters, particularly the SVs above the U-curve cut-off, fulfill Picard's condition. The blue line is the decay of the SVs,  $\sigma_i$ 's, the green crosses correspond to  $|u_i^T d|$ , the red circles represent the quotient  $|u_i^T d|/\sigma_i$ , and the two horizontal dashed lines represent the heuristic acceptable range ( $10^{-1}$  and  $10^{-2}$ ).



**Fig. 3.11.** Picard's plot indicating the maximum and minimum values of the heuristic acceptable range (horizontal dashed black lines at  $10^{-1}$  and  $10^{-2}$ ).

We can observe that, as in the case of figure 3.6.d, the SVs above the heuristic acceptable range of parameters  $10^{-1}$  and  $10^{-2}$ , particularly the SVs above the U-curve cut-off ( $\alpha_u = 5.72 \cdot 10^{-2}$ ), fulfill Picard's condition. In this case the SVs decay to zero faster than for the phantom experiment (figures 3.6.c -3.6.d). Furthermore, the ratio between the largest and the smallest nonzero singular value (CN) is larger for the ex-vivo mouse data experiment (CN= $10^{17}$ ) than for the phantom experiment (CN= $10^{10}$ ), thus confirming that the problem for the ex-vivo experiment is more ill-posed than the one for the phantom experiment.

### 3.4. Discussion and Conclusions

The U-curve-based method is here used for the first time to select the regularization parameter in Tikhonov regularization reconstruction of fDOT. It has been shown that it provides a suitable selection of the regularization parameters, in terms of Picard's condition, image resolution and image noise. Results are shown both on phantom and mouse data.

Choosing the correct regularization parameter is crucial for the reconstruction of DOT and fDOT data. Singular Value Analysis has been used to optimize experimental setups in optical tomography (Chamorro-Servent et al. 2010; Chamorro et al. 2009; Graves et al. 2004), and chapter 4 of this thesis, and the Tikhonov regularization has recently been used to introduce anatomical a priori information into fDOT reconstructions. Therefore, an

automatic method that enables us to choose the regularization parameter is of paramount interest.

To our knowledge, the L-curve method is the only automatic strategy, not requiring priori knowledge of the noise, that has been successfully applied to fDOT (Corlu 2007; Serdaroglu et al. 2006; Xu et al. 2009).

Recently, it has been found that methods without a-priori information, such as L-curve and GCV were robust in other fields (Abascal et al. 2008; Correia et al. 2009).

It is noteworthy that the L-curve method presents several theoretical limitations and may fail to find a good regularization parameter when the solutions are very smooth (Hanke 1996), and examples of inverse problems where the L-curve does not converge have been found (Vogel 1996).

In diffuse optical tomography applications, the authors of (Culver et al. 2003) emphasized that the L-curve analysis yielded an overly-smooth solution in some cases. Recently, (Jagannath, and Yalavarthy 2012; Li et al. 2012), in agreement with our results, also pointed out that L-curve may fail when there is no optimal turning point in the L-curve plot.

The GCV method, on the other hand, is more computationally expensive than the L-curve for large systems (Busby, and Trujillo 1997).

It can be seen that the L-curve calculated for the phantom experiment did not exhibit a neat corner (figure 3.3). However, the U-curve for the same experiment had a clear minimum (figure 3.2). Furthermore, this minimum was found in the interval given in the section 3.2.1.1 (thus not being necessary to calculate the U-curve for the  $\alpha$  parameters out of this interval).

Besides, figures 3.6.d and 3.11 (Picard's plots) show how the U-curve regularization parameter satisfies Picard's condition and assure a satisfactory regularized solution. From figures 3.4 and 3.8 (reconstruction obtained for different  $\alpha$  parameters) and figures 3.5 and 3.10 (profiles and FWHM versus noise plot), we can choose values for the regularization parameter that are lower than the value at which the reconstructed image started to be noisy. Figures 3.6.d and 3.11 show that, for these lower values, Picard's condition is satisfied, as the U-curve parameter is in this range.

Simple observation of the Picard's plot can reveal a valid regularization parameter; however, the choice is more subjective. An automatic selection of the threshold parameter may be simpler and more objective in most cases. Furthermore, in agreement with (Culver

et al. 2003), we can see clearly in figures 3.5.b and 3.10.b (FWHM versus noise plot) that the general trend is for resolution to increase together with image noise while the regularization parameter decreases.

It is interesting to remark that the resolution achieved for the *ex-vivo* experiment is better than for the phantom experiment (figure 3.10.b versus figure 3.5.b), due to the fact that the tip of the capillary is closer to the surface. As the resolution of DOT systems are depth-dependent, resolution improves when the object is closer to either side of the slab (Pogue et al. 1999).

We have reviewed the different methods used in the literature, focusing on their differences.

Regarding the limitations of this study, we realize that the U-curve criterion may fail in some cases, but in our experience it works well for fDOT experiments in practice. Although only two experiments are presented in this chapter, when we used the U-curve method in other experiments with different aims, as for example the experiments described in chapter 4 or in references (Chamorro-Servent et al. 2010; Chamorro et al. 2009), we always obtained satisfactory reconstructions, both for mice and for phantoms.

We expect the automatic U-curve method for selecting the regularization parameter to yield robust and useful results that can be applied to the reconstruction of fDOT images and studies of image performance by singular value analysis.

## Chapter 4

# SVA applied to optimizing non-contact fDOT

As described in section 1.4.2, non-contact fDOT setups generate very large datasets. For this reason, many works have used singular value analysis (SVA) of the forward matrix to find the optimal experimental parameters for rotating configurations of DOT and fDOT systems, in terms of number, arrangement and size of detectors, mesh density and field of view (FOV), and number of geometrical projections.

To achieve good image quality it is necessary to determine the number of useful singular values (SVs) to retain. The work presented in this chapter exploits the U-curve automatic method, introduced in chapter 3, to appraise the effect of different settings of the acquisition parameters (distribution of mesh points, density of sources and detectors) of a parallel-plate non-contact fDOT (section 1.3.1), in order to achieve the best possible imaging performance. The idea is make use of the minimum number of SVs of the fDOT forward matrix in order to maximize the information content in acquired measurements while minimizing the computational cost.

The results obtained from this study can guide the selection of optimum acquisition parameters for fDOT experiments.

This chapter is organized as follows: First we present a quick introduction on the use of SVA as image performance tool, then, the phantom and real studies carried out, and finally, the results, discussion and conclusions of the work.

## 4.1. Introduction

As explained in section 1.3.3, the collected data can be mathematically arranged as a system of equations  $d = Wf$ , where  $d$  is a vector that contains the measurements corresponding to each source detector pair,  $f$  is the unknown fluorophore concentration at each voxel of the subject, and  $W$  is the forward matrix that represents the contribution of each voxel to the measurement corresponding to each source-detector pair.

The size of the matrix  $W$  is  $mxn$ , where  $m$  is the product of number of detectors ( $N_d$ ) by number of sources ( $N_s$ ) and  $n$  is the number of elements of the mesh. As remarked in section 1.4.2, this size ranges between  $10^2 - 10^3$  for contact geometries, whereas for non-contact fDOT setups it can easily reach  $10^7 - 10^9$ , leading to a computationally demanding inverse problem.

SVA has been shown to be a simple yet powerful tool (Culver et al. 2001; Graves et al. 2004; Graves et al. 2003; Lasser, and Ntziachristos 2007) to assess the usefulness of DOT setups (Culver et al. 2001). Several fDOT systems have been studied with this technique: Graves et al. (Graves et al. 2004; Graves et al. 2003) studied how FOV, number of the detectors and sources, and number of mesh elements affect the system performance, using 2D simulated data corresponding to a parallel plate system. More recently, Lasser et al. (Lasser, and Ntziachristos 2007), using a phantom experiment with a parallel plate setup, studied the effect of varying the number of detectors with fixed number of sources and mesh points.

In this chapter, we use the SVA technique to find the optimal experimental parameters for the fDOT experimental setup presented in section 1.3.1.

We hypothesize that we will achieve better coupling between detected data and reconstructed image if the mesh point distribution is less dense in the direction perpendicular to the plates due to the poor depth information in the acquired data. We also verified that placing the detectors and sources closer than one transport mean free path leads to a slight increase in imaging performance.



## 4.2. Methods

Using the slab-shaped phantom described in section 3.2.3.1, we explored different combinations of density of sources and detectors, and distribution of mesh points. After that, we constructed the forward matrices of fDOT settings as explained in section 1.3.3 and we decomposed them into their SVs.

$$W = USV^T, \quad (4.1)$$

where  $U$  and  $V$  are orthonormal matrices ( $U^{-1} = U^T$ ,  $V^{-1} = V^T$ ) and  $S$  is a diagonal matrix containing the SVs of  $W$ .

Finally, regarding the minimum number of SVs required, we assessed the influence on the imaging performance of the density of sources and detectors, and the influence of the number of the mesh points and their spatial distribution, using the U-curve method presented in chapter 3.

Note that equation (4.1) can be written as:

$$U^T d = SV^T f, \quad (4.2)$$

where the columns of  $U$  can be seen as the detector-space of  $W$  and the columns of  $V$  as the image-space of  $W$ . The SVs of  $W$  couple the image-space and the corresponding detector-space as can be seen in figure 4.1.

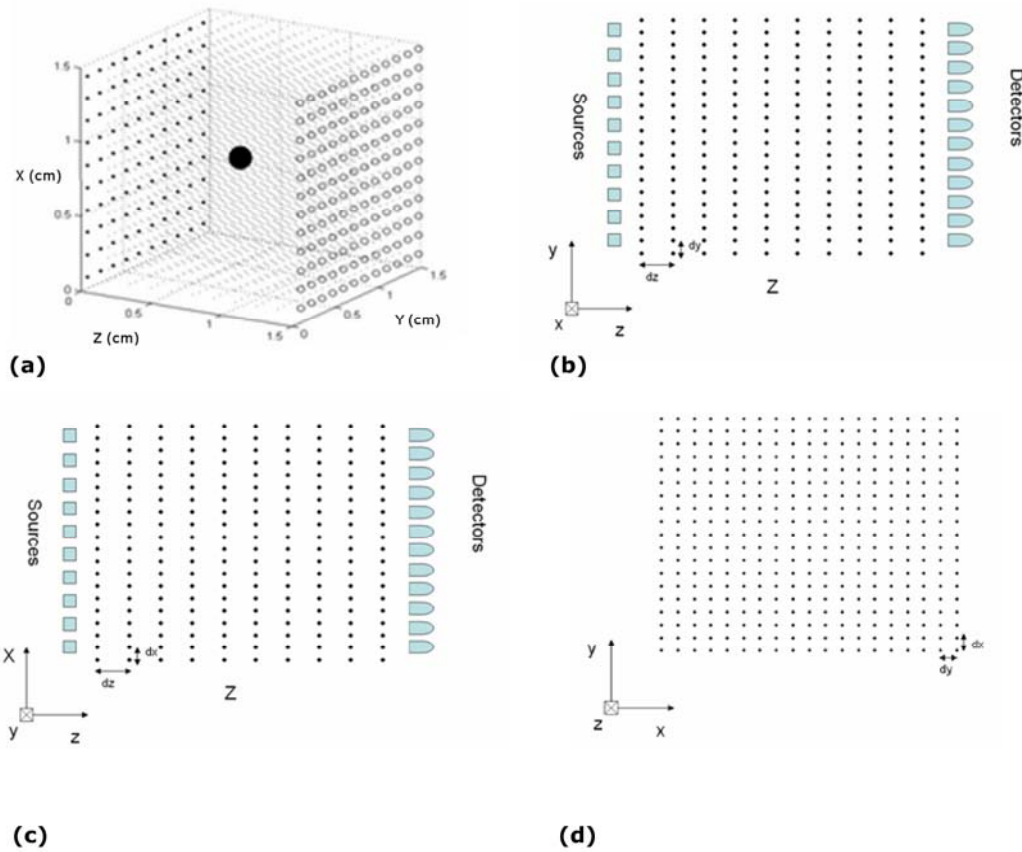
$$\begin{array}{c}
 \begin{pmatrix} d(r_{s1}, r_{d1}) \\ \vdots \\ d(r_{sm}, r_{dm}) \end{pmatrix} = \begin{pmatrix} W_{11} & \dots & W_{1n} \\ \vdots & \ddots & \vdots \\ W_{m1} & \dots & W_{mn} \end{pmatrix} \begin{pmatrix} f(r_1) \\ \vdots \\ f(r_n) \end{pmatrix} \quad n < m \\
 \text{SVD } W = U \Sigma V^T \\
 \begin{pmatrix} d(r_{s1}, r_{d1}) \\ \vdots \\ d(r_{sm}, r_{dm}) \end{pmatrix} = \begin{pmatrix} | & \dots & | \\ u_1 & \dots & u_n \\ | & \dots & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} | & \dots & | \\ v_1 & \dots & v_n \\ | & \dots & | \end{pmatrix}^T \begin{pmatrix} f(r_1) \\ \vdots \\ f(r_n) \end{pmatrix} \\
 \text{Detector Space} \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0 \quad \text{Image Space}
 \end{array}$$

**Fig. 4.1.** : Image and detector-spaces in terms of SVD of matrix  $W$ .

Using SVA, we assessed:

- The influence of the density of sources and detectors on the imaging performance, using a 2x2x1.5 cm volume of interest (VOI), 20x20x10 mesh points, and source and detectors square FOVs of 2x2 cm (Study 1).
- The influence of the number of the mesh points and their spatial distribution, for a mesh VOI of 1.5x1.5x1.5cm, using 12x12 detectors and 10x10 sources equally spaced in their respective FOVs of 1.5x1.5 cm, as depicted in figure 4.2.a (Study 2).

Due to the fact that the mesh VOI is cubic, for each number of mesh points,  $n$ , three anisotropic distributions were constructed:  $dx=dy=2dz$ ,  $2dx=dy=dz$ , and  $dx=2dy=dz$ , where  $dx$ ,  $dy$ , and  $dz$  denote the spacing for the mesh points coordinates  $x$ ,  $y$  and  $z$  respectively (figure 4.2. b, c, d). Afterwards, the anisotropic distribution giving best imaging performance was compared with an isotropic mesh distribution,  $dx=dy=dz$ .

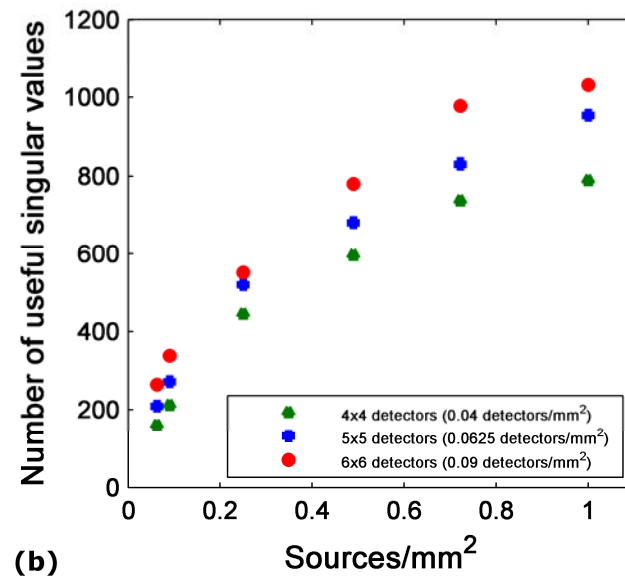
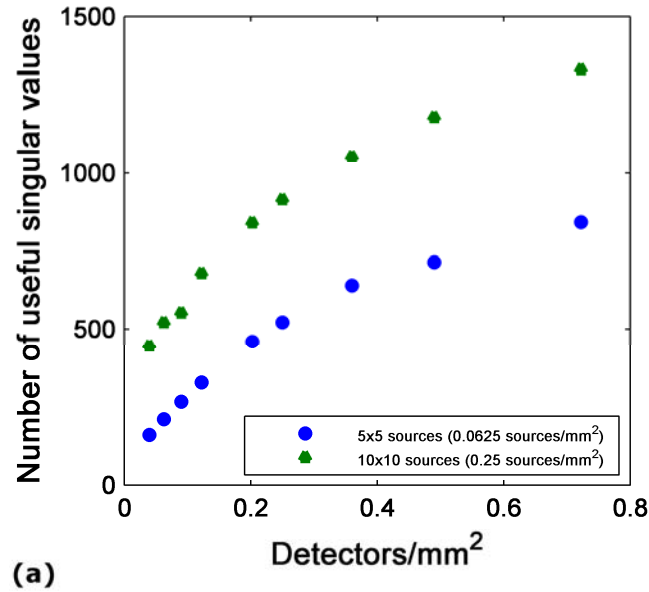


**Fig. 4.2.** (a) Diagram of the mesh, detector and sources FOV perspective view for the second study. The big black sphere represents the target object (the tip of the capillary). Figures (b), (c) and (d) represent the different views of the mesh points arrangement corresponding to the distribution  $dz=2dx=2dy$ , respectively.

## 4.3. Results

### 4.3.1. Results on density of sources and detectors

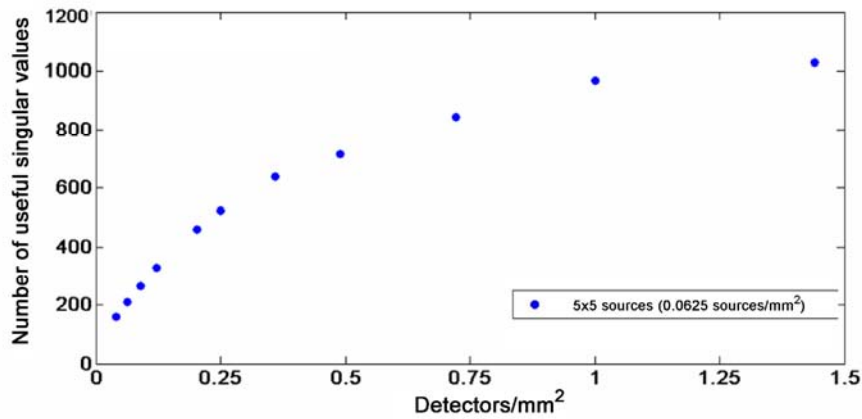
Figures 4.3, 4.4 and 4.5 show the influence on imaging performance of the density of sources and detectors, based on the study of the number of useful SVs.



**Fig. 4.3.** SVA of the influence on the imaging performance: (a) of the density of sources for a fixed number of detectors, (b) of the density of detectors for a fixed numbers of sources.

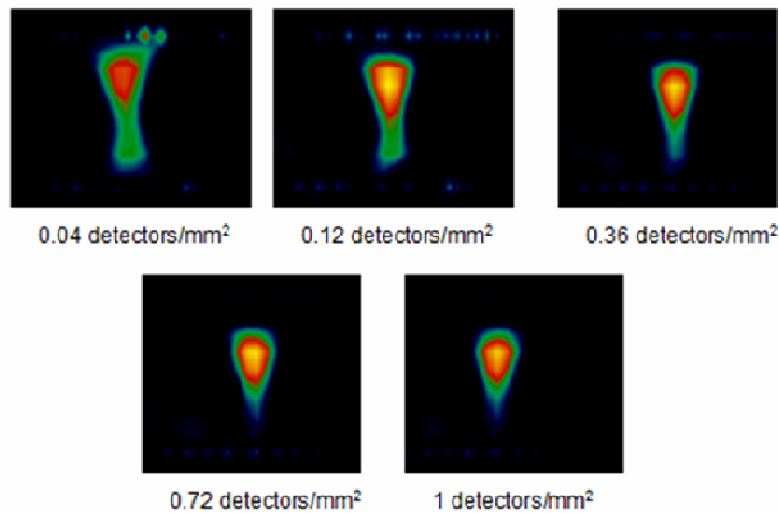
We can see (figure 4.3.a) that for a fixed number of sources, the number of useful SVs as a function of the detectors density quickly increase at the beginning but steadies when detectors get closer than the average mean free path (density of  $1 \text{ mm}^2$ ). Similar results appear (figure 4.3.b) when fixing the number of detectors and varying the source density.

Figure 4.4 shows, for small fixed number of sources, that increasing the density of detectors to distances shorter than  $1 \text{ mm}$  leads to slight improvements in image quality while increasing computational burden.



**Fig. 4.4.** SVA of the influence on the imaging performance of the density of detectors for a fixed numbers of sources.

Figure 4.5 depicts the reconstructions of the phantom presented in section 3.2.3.1, corresponding to the different densities of detectors used during the data acquisition.



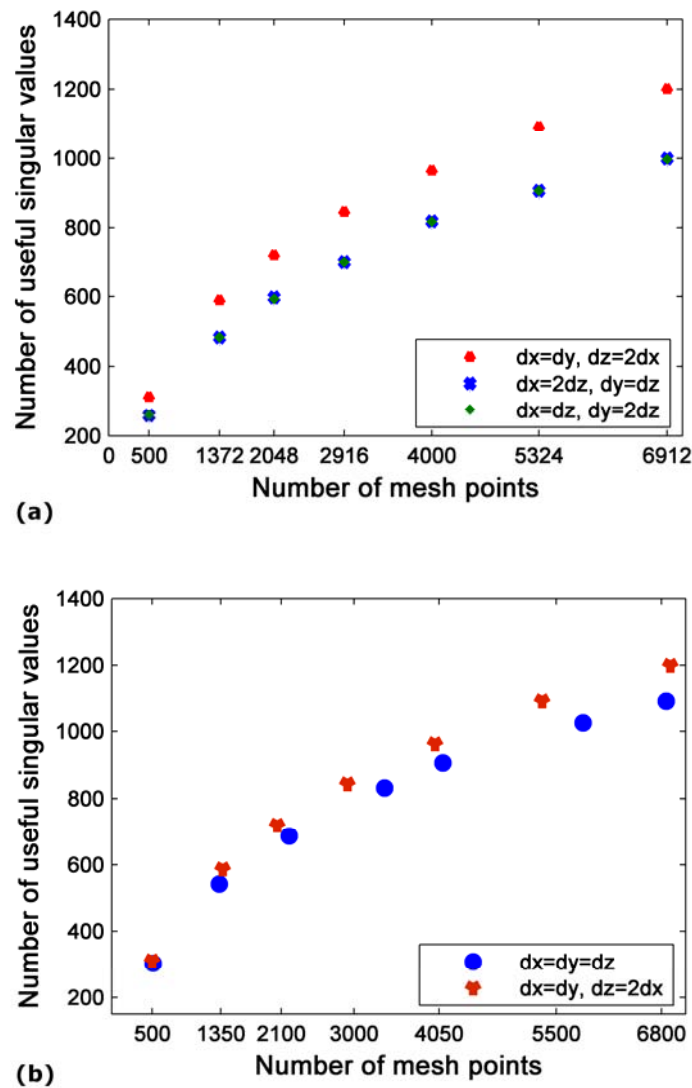
**Fig. 4.5.** Axial views (x direction) of 5 reconstructed images corresponding to different detectors densities for a fixed numbers of sources ( $10 \times 10$ ).

We can observe that the tip of the capillary is better defined when the detectors are at a distance of about 1 mm. This conclusion agrees with the results provided by SVA in figures 4.3 and 4.4, and with those reported in (Graves et al. 2004).

We can conclude that we should not increase detectors/sources density at distances shorter than the average mean free path (1 mm).

#### 4.3.2. Results on the mesh spatial distribution

Regarding the influence of the number of voxels and their spatial distribution,



**Fig. 4.6.** SVA analysis of imaging performance: (a) changing the different voxel anisotropies. (b) isotropic voxel distribution versus anisotropic voxel distribution.

We can see (figure 4.6.a and table 1) that a mesh density  $dx=dy$ ,  $dz=2dx$  gives better results in terms of imaging performance than  $dx=2dz$ ,  $dy=dz$  or  $dx=dz$ ,  $dy=2dz$ . Comparing an anisotropic distribution of voxels ( $dx=dy$ ,  $dz=2dx$ ), against the isotropic distribution ( $dx=dy=dz$ ) (figure 4.6.b and Table 4.1) we can see that the anisotropic distribution achieves better performance than the isotropic one.

The number of useful SVs quickly increase at the beginning with the number of mesh points but steadies for large numbers of elements, similarly to what happened in figures 4.3-4.5.

**Table 4.1.** Study of the number of useful singular values depending on the number of voxels and their distribution

Isotropic mesh density ( $dx=dy=dz$ )			Anisotropic mesh density ( $dx=dy$ , $dz=2dx$ )		
Mesh Elements	Size (mm)	Number of useful SVs	Mesh elements	Size (mm)	Number of useful SVs
8x8x8 (512)	1.87	304	10x10x5 (500)	1.5x1.5x3	312
11x11x11 (1331)	1.36	542	14x14x7 (1372)	1.07x1.07x2.14	584
13x13x13 (2197)	1.15	687	16x16x8 (2048)	0.94x0.94x1.87	721
15x15x15 (3375)	1	832	18x18x9 (2916)	0.83x0.83x1.66	845
16x16x16 (4096)	0.94	907	20x20x10 (4000)	0.75x0.75x1.5	964
18x18x18 (5832)	0.83	1026	22x22x11 (5324)	0.68x0.68x1.36	1092
19x19x19 (6859)	0.79	1090	24x24x12 (6912)	0.63x0.63x1.2	1201

In Table 4.1, note that a similar number of useful SVs was necessary for the isotropic distribution 19x19x19 (1090 SVs) than for the anisotropic distribution 22x22x11 (1092 SVs). However, the first distribution used 6859 voxels while the second one used only 5324 voxels. These numerical results point out that an anisotropic mesh achieves better performance than an isotropic mesh.

Since these results show that an anisotropic distribution of voxels ( $dx=dy$ ,  $dz=2dx$ ) achieves better performance than an isotropic one, we studied to what extent density in  $z$  direction could be reduced without losing information. Table 4.2 shows the result of using different  $z$ -voxel sizes for a similar number of mesh points,  $n$ .

**Table 4.2.** Number of useful singular values for an anisotropic distribution of voxels ( $dx=dy$ ,  $dz=c dx$ ), where  $c$  denotes a integer constant number varying between 1 and 7, that corresponds to varying  $dz$  between  $1mm \sim 4mm$ .

Mesh Elements	Size (mm)	Number of useful SVs	Number total of mesh points ( $n$ )
15x15x15 ( $c=1$ )	1,1,1	832	3375
18x18x9 ( $c=2$ )	0.83x0.83x <b>1.66</b>	845	2916
21x21x7 ( $c=3$ )	0.71x0.71x <b>2.14</b>	913	3087
25x25x5 ( $c=5$ )	0.6x0.6x <b>3</b>	<b>953</b>	3125
28x28x4 ( $c=7$ )	0.53x0.53x <b>3.7</b>	<b>954</b>	3136

We can observe that for a similar number of mesh points ( $n$ ), the number of useful SVs increases with  $dz$  and stabilizes for  $dz$  larger than 3mm.

## 4.4. Discussion and Conclusions

In this chapter we studied the effect of different settings of the acquisition parameters (distribution of mesh points, density of sources and detectors) of a non-contact parallel-plate fDOT.

Analyzing the SVA of our forward matrix for different source and detector configurations, we found that any increase in the number of sources and detectors at distances shorter than the average mean free path leads to slight improvements in image quality but increases computational burden (figures 4.3-4.5).

Besides, the use of a mesh with lower density in the direction perpendicular to the plates (figure 4.6, and table 4.1) achieves better performance than the usual isotropic mesh point distribution used in most fDOT experimental studies. Moreover, these results confirm (table 4.2) that a suitable spacing for the mesh points in the direction perpendicular to the plates is around 3mm, in agreement with the fDOT resolution in  $z$  direction (3 mm).

To our knowledge, it is the first time that this result has been obtained using SVA for a non-contact parallel-plate fDOT.

This finding is especially interesting to optimize the image reconstruction, which is largely a function of the number of points in the reconstruction mesh. Large meshes can be

compressed in the  $z$  direction without any loss in imaging performance, but reducing computational times and memory requirements.

One drawback of this methodology is the high computational cost required by SVD. Besides, we would like to remark that our results correspond to a non-contact parallel-plate fDOT phantom and a mouse experiment using a laser whose emitter wavelength is  $675 \pm 5$  nm; and the forward matrix was modelled using Green functions as explained in section 1.3.3. Although most of the results provided are in broad agreement with other works previously presented in the literature, it would be advisable to confirm the study for different laser wavelengths, sample dimensions or even when the forward matrix is modelled differently.

In conclusion, our findings can serve as a guide for the selection of optimum acquisition parameters for a non-contact parallel-plate fDOT experiment, as the one presented in section 1.3.1.



## Chapter 5

# $l_2$ -norm alternatives enforcing sparsity of the solution

Traditionally, image reconstruction is formulated using a regularized scheme in which  $l_2$ -norms are used to assess both the data misfit and the penalization term. However, there is an increasing interest in the use of  $l_1$ -norm as a regularization term for image reconstruction, derived from a tendency in the inverse problem community to solve many inverse problems with a sparse solution.

Near-infrared (NIR) tracers used in fDOT are designed to be specific and accumulate particular organs or regions. For this reason, most of the fluorophore is usually located within a volume of interest (i.e. a determined number of voxels) of the reconstructed image, making these images highly sparse.

The second part of this thesis and this chapter centre their attention on reconstruction techniques that enforce sparsity, such as  $l_1$  and Total Variation (TV) penalties, as an alternative to regularization based on  $l_2$ -norm.

The chapter is organized as follows: First, we introduce the concepts of sparse solutions,  $l_0$  and  $l_1$ -norm and TV. Afterwards, we introduce the different formulations of the optimization problem and present some algorithms to solve them. Finally, we briefly review the alternatives to  $l_2$ -norm regularization reconstruction techniques used in fDOT.

## 5.1. Sparse solutions

A solution or image  $x$  is sparse if most of its entries are zero. That is, we minimize

$$\|Ax - b\|_2^2 + \alpha \|x\|_1. \quad (5.1)$$

If  $x$  is not sparse, we assume that we can find an orthogonal transformation basis  $T$  where  $x$  is sparse. Then, we can write

$$\|Ax - b\|_2^2 + \alpha \|Tx\|_1. \quad (5.2)$$

### 5.1.1. $l_0$ -norm and $l_1$ -norm

Amongst the most popular norms used in functional analysis are the so-called  $p$ -norms ( $p > 0$ ). If  $x = (x_1, \dots, x_n)$  is a vector, its  $p$ -norm is

$$\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}. \quad (5.3)$$

Thus, the  $l_1$ -norm is the sum of the absolute values of the components of a vector, while the  $l_2$ -norm presented in previous chapters is the Euclidian norm.

Another popular norm is the most fundamental sparsity metric, the  $l_0$ -norm. The  $l_0$ -norm of a vector  $x$  is denoted by  $\|x\|_0$  and defined as the number of non-zero terms of  $x$ . Thus, this norm can be used to measure the sparsity level of a vector. Nevertheless, this norm is not convex. The advantage of convex problems over non-convex problems is that a global optimum can be computed.

Regarding the  $p$ -norm definition (equation 5.3), the smallest value for which convexity is retained is  $p=1$ . The  $l_1$ -norm is a convex relaxation of  $l_0$ -norm and it is often used to enforce sparsity in images. This is the reason why the  $l_1$ -norm has received significant interest in the context of compressed sensing (CS). However, this regularization is neither linear nor differentiable, and is more difficult to implement than the classical  $l_2$ -based regularization.

An interesting equivalence between  $l_1$ -norm and  $l_2$ -norm penalizing techniques was described by Kees van den Doel et al. (van den Doel, Ascher, and Haber 2012). These authors showed the equivalence between minimizing the problem described in equation 5.2 and the TSVD method (see section 2.3) when  $T = V^T$ , being  $V^T$  the transpose of a matrix having as columns the right singular vectors resulting from SVD of the forward matrix.

Remember that the null space of  $W$  ( $\text{Ker}(W)$ ) is given by columns of  $V$  corresponding to the zero SVs (section 2.2.2). Thus, being  $T = V^T$ , through  $V^T x_{true} \neq 0$  we select the non-zero SVs, which is equivalent to the TSVD method. We will return to this theorem in chapter 7, but dealing with other transformations that enable CS.

In terms of Bayesian statistics, the  $l_1$  regularization usually corresponds to setting a Laplacian prior on the regression coefficients when taking a maximum a posteriori hypothesis. Similarly, the  $l_2$  regularization corresponds to Gaussian prior on the regression coefficients when taking a maximum a posteriori hypothesis. Laplace ( $l_1$ ) tends to better tolerate very large or small values of coefficients than Gaussian ( $l_2$ ).

### 5.1.2. Total Variation

Total Variation (TV) has shown large potential in image restoration and image reconstruction since it was introduced by Rudin, Osher and Fatemi (Rudin, Osher, and Fatemi 1992). Since most images have slowly varying pixel intensities (except at the edges), its discrete gradient will be sparse.

The TV of an image  $x$  in a domain  $\Omega$  is defined as the  $l_1$ -norm of the gradient of the image. As mentioned before, by using the  $l_1$ -norm we are promoting the sparsity of the solution. By penalizing its TV we are preserving local smoothness and piecewise constancy (by promoting sparsity in the gradient of the solution).

$$TV(x) = \|\nabla x\|_1 = \int_{\Omega} |\nabla x| dr. \quad (5.4)$$

The non-differentiability of the absolute functional can be overcome by using regularization approximations of TV. Some of these approximations are Huber regularization, TV weight regularization or Perona Malik (PM) regularization (Correia et al. 2011).

Summarizing,  $l_2$ -norm regularization produces in practice the minimum energy solution, suppressing most of the large noise components, while  $l_1$ -norm and TV-norm regularization enforce sparsity. This latter option, furthermore, preserves local smoothness and piecewise constancy.

## 5.2. Formulation of the optimization problem

In this section we use the term  $l_1$ -norm ( $\|x\|_1$ ) to denote  $l_1$ -norm and TV-norm regularization indistinctly.

The optimization problem which addresses the reconstruction problem (finding  $x$  from a linear system  $Ax = b$ ) while enforcing sparsity of solution can be formulated in a number of equivalent ways including:

- Basis pursuit (BP) or linear programming, where the  $l_1$ -norm appears in the cost function and the  $l_2$  data-fitting term appears in the constraint as equality:

$$\min_x \|x\|_1 \text{ such that } b = Ax . \quad (5.5)$$

- Basis pursuit denoising (BPDN) or quadratically constrained linear program, where the  $l_1$ -norm appears in the cost function and the  $l_2$  data-fitting term appears in the constraint as a noise-dependant inequality:

$$\min_x \|x\|_1 \text{ such that } \|b - Ax\|_2^2 \leq \delta . \quad (5.6)$$

- Least absolute shrinkage and selection operator (commonly referred to as LASSO problem or quadratic program), where the  $l_2$  data-fitting term appears in the cost function and the  $l_1$ -norm appears in the constraint:

$$\min_x \|b - Ax\|_2^2 \text{ such that } \|x\|_1 \leq \varepsilon . \quad (5.7)$$

- $L_1$ -penalized least squares ( $LS+L_1$ ) also known by some authors as first-order Tikhonov (Arridge, and Schotland 2009), or  $l_1$ -Tikhonov (Egger et al. 2010), where

the  $l_1$ -norm penalty weighted by a regularization parameter is added to the  $l_2$  data-fitting term to construct the cost function:

$$\min_x \frac{1}{2} \|b - Ax\|_2^2 + \lambda \|x\|_1. \quad (5.8)$$

The use of one or another formulation depends on the prior information available. If we have prior knowledge about the noise behaviour (related to  $\delta$ ) we would use the BPDN formulation, whereas if we have prior information about the sparsity of the solution (related to  $\varepsilon$ ) we would prefer the LASSO formulation. BP works only for undetermined systems while LASSO is more suitable for overdetermined systems.  $LS+L_1$  is the most general formulation but has the difficulty of the appropriate choice of the regularization parameter,  $\lambda$ . In summary, the choice of one or other formulation depends on the nature of problem to be solved.

### 5.3. Looking for an algorithm

As  $l_1$ -norm regularization-based problems are non-linear, the choice of an algorithm for solving this kind of problems may become crucial because computational cost can be excessively large when using classical gradient-based approaches.

Different works have pursued the goal of developing efficient algorithms. In (Theodoridis et al. 2012), the authors present a review on the most commonly used and cited algorithms focusing on schemes that have low computational burden and that scale well to very large problem sizes. With this aim, they categorize the algorithms into three groups: greedy-type algorithms, iterative shrinkage schemes and convex optimization techniques, and stress that drawing definite conclusions about which are the most efficient algorithms is unrealistic.

Regarding 3D reconstruction problems, some examples of algorithms efficient in terms of convergence and stability are: the SPGL1 algorithm (van den Berg, and Friedlander 2007), which solves BP and LASSO problems with  $l_1$ -norm penalty; the TVReg method (Jensen et al. 2012), for  $LS+TV$  problems; the CVX scheme (Grant, Boyd, and Ye 2008), which solves  $LS+L_1$ , BP and LASSO problems; and the Split Bregman algorithm (Goldstein, and Osher 2009), appropriate for  $LS+L_1$  and  $LS+TV$  problems. The

first three algorithms are based on optimization techniques while the last one is based on iterative shrinkage schemes.

Although we have tested the four above cited algorithms obtaining satisfactory solutions, the work presented in this thesis focuses on the Split Bregman (SB) method (Goldstein, and Osher 2009). SB is a simple and efficient algorithm for solving  $l_1$  regularization-based problems, and is particularly effective for problems involving TV regularization. Its efficiency derives from its possibility of splitting the minimization of  $l_1$  and  $l_2$  functionals. Applying the SB method to TV, or Rudin, Osher and Fatemi (ROF) image denoising and compressed sensing the authors in (Goldstein, and Osher 2009) showed that SB was computationally efficient, because the SB formulation leads to a problem that can be solved using Gauss-Seidel and Fourier transform methods.

From now on, we will denote by SB the Split Bregman method applied to TV/ROF denoising and by CS-SB the Split Bregman method applied to CS.

## 5.4. A brief review of sparse regularization reconstruction techniques applied to fDOT

In the context of fDOT reconstruction, several sparsity-enforcing alternatives to  $l_2$ -norm for regularization have been proposed. Arridge and Schotland (Arridge, and Schotland 2009) and Egger et al. (Egger et al. 2010) introduced  $l_1$  and TV-based regularizations in their respective reviews about forward and inverse problems in optical tomography. Baritau et al. (Baritau, Hassler, and Unser 2010) designed a reconstruction algorithm that incorporates a general regularization application of general constraints to fDOT, combined with an efficient matrix-free strategy that enables the algorithm to deal with large reconstruction problems at reduced memory and computational cost. Freiburger et al. (Freiburger, Clason, and Scharfetter 2010) introduced an alternating direction minimization method to solve  $l_1$  regularization; this method splits the reconstruction problem for simulated fDOT data into two subproblems: an  $l_2$  stage, solved using a Gauss-Newton (GN) step, and an  $l_1$  regularization stage, solved by thresholding (or shrinkage). Later, Baritau et al. (Baritau, and Unser 2011) presented an  $l_2$ - $l_1$ -mixed-norm penalization for incorporating a structural prior in fDOT image reconstruction. In the same paper they proposed other sparsity penalizations, such as  $l_1$ -norm and TV-norm, and tested

them with synthetic and experimental data. Correia et al. (Correia et al. 2011) introduced an operator splitting method with nonlinear anisotropic diffusion and edge priors in fDOT reconstructions of simulated, phantom, and *ex-vivo* data. In one paper from our group (Abascal et al. 2011) we reconstructed fDOT focusing on the use of Split Bregman method to solve the optimization problem by imposing a non-negativity constraint; the image was updated using a nonlinear GN step, based on the computation of first and second derivatives of the nonlinear total variation functional. Dutta et al. (Dutta et al. 2012) applied a combination of  $l_1$  and TV penalties to the fDOT inverse problem to simultaneously encourage properties of sparsity and smoothness in the reconstructed images; they concluded that using  $l_1$  or TV regularization, in combination or separately, clearly led to improvements in localizing fluorescent sources in fDOT. Qualitatively, the joint  $l_1$ -TV images showed the most natural appearance in simulation and phantom studies but the quantitative studies did not identify a clear winner. Behrooz et al. (Behrooz et al. 2012) compared  $l_2$ -regularization methods and algebraic reconstruction technique (ART) with TV-reconstruction methods inspired in R.O.F. (see section 5.1.2) and CS-SB; they implemented a preconditioned conjugate gradient method at each iteration of CS-SB. They remarked in their paper that this can lead to slow convergence in some cases and that TV regularization has the potential of offering higher resolution and robustness compared to conventional  $l_2$ -regularization algorithms and ART.





## Chapter 6

# Use of Split Bregman denoising for iterative reconstruction

The algebraic reconstruction technique (ART) is an extensively applied and cost-efficient reconstruction method that yields fast and stable reconstructions for large datasets, as is the case of experimental DOT and fDOT studies. The useful results yielded by more advanced  $l_1$ -regularized techniques for signal recovery and image reconstruction, together with the recent availability of the Split Bregman (SB) procedure, led us to propose a novel iterative algorithm for fDOT image reconstruction in diffusive media, ART-SB. This new algorithm has application in a wide range of areas, in particular in in-vivo imaging. This method alternates a cost-efficient reconstruction step (an ART iteration) with a denoising filtering step based on the minimization of total variation (TV) of the image using the SB method, which can be solved efficiently and quickly. We applied this approach to simulated and experimental fDOT data, and we found that ART-SB provides substantial benefits over conventional ART without increasing the computation time.

The organization of this chapter is as follows. Section 6.1 introduces the chapter. Section 6.2 presents the well-known ART algorithm and describes the proposed ART-SB method. Then, it describes data acquisition and data simulation, and presents the tools used to compare ART with ART-SB. Section 6.3 presents the reconstructions and comparative results of simulated and experimental data. Finally, section 6.4 presents the discussion and conclusions.

## 6.1. Introduction

DOT and fDOT image reconstruction are commonly carried out by means of iterative methods such as the algebraic reconstruction technique (ART) (Arridge, and Schotland 2009; Intes et al. 2002; Zacharopoulos et al. 2009).

However, based on the success of  $l_1$  regularization techniques for denoising and image reconstruction, new iterative image reconstruction procedures have been proposed in the fields of computed tomography and positron emission tomography (Johnston, Johnson, and Badea 2010; Pan et al. 2010; Sawatzky et al. 2008). These procedures alternate an iterative method (such as simultaneous ART or the expectation-maximization algorithm) with a TV-denoising step. To minimize the TV functional, the two first works cited above (Johnston et al. 2010; Pan et al. 2010) used a standard gradient descent method, while the third one applied a dual optimization algorithm (Sawatzky et al. 2008). Consequently, the choice of technique for solving  $l_1$  regularization-based problems may become crucial, as  $l_1$  is non-linear; therefore, the computational burden can increase significantly using classic gradient-based methods.

In our work we focus on the Split-Bregman (SB) method because it is a simple and efficient algorithm for solving  $l_1$  regularization-based problems, and is particularly effective for problems involving TV regularization. Its efficiency derives from the splitting of the minimization of  $l_1$  and  $l_2$  functionals.

SB was recently applied to fluorescence tomography reconstruction (Abascal et al. 2011; Behrooz et al. 2012). From our group, Abascal et al. (Abascal et al. 2011) used SB to solve the optimization problem by imposing a non-negativity constraint. The image was updated using a non-linear Gauss-Newton step (Arridge, and Schotland 2009) based on the computation of first and second derivatives of the non-linear TV functional. Behrooz et al. (Behrooz et al. 2012) compared  $l_2$  regularization methods and ART with TV reconstruction methods. In the last work cited, the authors implemented a preconditioned conjugate gradient method (Arridge, and Schotland 2009) in each iteration of SB that led to slow convergence in some cases. To validate the method and compare reconstructions, they used a non-contact constant-wave trans-illumination fluorescence tomography system and concluded that TV regularization has the potential to offer higher resolution and robustness than conventional  $l_2$  regularization algorithms and ART.

This chapter presents a new approach to the fDOT inverse problem that alternates the Algebraic Reconstruction Technique (ART) with a denoising step based on the minimization of TV, solved by means of the Split Bregman (SB) method. This approach, named ART-SB, has been optimized and thoroughly studied. It has also been validated with real experimental fDOT data, in contrast to most other reports making use of shrinkage algorithms (Douiri et al. 2007; Freiburger et al. 2010), and resulted in substantial benefits over conventional ART without increasing the computation time.

## 6.2. Methods

### 6.2.1. The algebraic reconstruction technique (ART)

The Kaczmarz's method, known in computed tomography as algebraic reconstruction technique (ART) (Kak 1979) is one of the most widespread iterative methods in image reconstruction.

Iterative regularization methods begin with an initial  $f^0$  (often a zero vector) and then iterate different solutions  $f^1, f^2, \dots$  until converging. In some way, the role of the regularization parameter is played by the number of iterations.

ART recalculates each solution by projecting from one to another hyperplane, defined by each row of the system  $Wf = d$ , as:

$$f_i^{it+1} = f_i^{it} + \lambda \frac{d_i - \sum_{n=1}^N w_{in} f_n^{it}}{\sum_{n=1}^N w_{in}^2} w_i, \quad (6.1)$$

where  $f_i^{it}$  is the it-th estimate of the it-th row contribution to the output  $f$ ,  $d_i$  the it-th component of the right-hand side,  $w_i$  the it-th row vector of the system matrix, and  $\lambda$  the relaxation parameter that adjusts the projection step at each iteration.

ART with low values of the relaxation parameter approximates a weighted least square solution that leads to over-smoothed images. On the contrary, high relaxation parameter values lead to high resolution images with noise and artifacts.

It can be shown that ART is equivalent to applying Gauss-Seidel iterations to the problem  $f = W^T y$ ,  $WW^T y = d$  (Hansen 2010).

As proved in (Herman, and Meyer 1993), one advantage of ART when dealing with non-singular matrices is that the residual norm does not increase.

The selection of an appropriate access order to the data (such as randomized access order) has been shown to speed up the iterative algorithm and generate a better output (Hansen 2010; Intes et al. 2003).

It is known that ART exhibits semiconvergence (Hansen 2010). It means that during the first iterations, the iterates  $f^k$  follows the correct trend approximating  $f_{true}$ , but at some stage they start to diverge from  $f_{true}$  to converge to the named “naïve solution”,  $W^{-1}d$ . More details can be found in (Hansen 2010).

### 6.2.2. The two-step reconstruction method: ART-SB

The first step corresponds to the minimization problem

$$f^{it} = \min_{\tilde{f}} \left\| W \tilde{f} - d \right\|_2^2 \quad (6.2)$$

solved by ART (section 6.2).

The second step corresponds to the denoising problem. For every z-projection

$$\tilde{f} = \min_{\hat{f}} TV\left(\hat{f}\right) + \frac{\mu}{2} \left\| \hat{f} - f^{it} \right\|_2^2, \quad (6.3)$$

where  $\mu$  is the weighting parameter for the fidelity term  $\left\| \hat{f} - f^{it} \right\|_2^2$  and  $TV$  is an anisotropic

TV given by

$$TV\left(\hat{f}\right) = \left\| \nabla_x \hat{f} \right\|_1 + \left\| \nabla_y \hat{f} \right\|_1 = \left\| \frac{\partial}{\partial x} \hat{f} \right\|_1 + \left\| \frac{\partial}{\partial y} \hat{f} \right\|_1, \quad (6.4)$$

The solution  $\tilde{f}$  constitutes the estimate for the next ART iteration. Note that  $\tilde{f}^0 = [0, \dots, 0] \in \mathbb{R}^n$  is used as initial guess in the first ART call.

The SB method (Goldstein, and Osher 2009) that solves (equation 6.3) is based on splitting the problem into two subproblems that are easier to solve. To this end, the original unconstrained problem (equation 6.3) is transformed into an equivalent constrained problem

$$\min_{\hat{f}, D_x, D_y} \|D_x\|_1 + \|D_y\|_1 + \frac{\mu}{2} \left\| \hat{f} - f^{it} \right\|_2^2 \text{ such that } D_i = \nabla_i \hat{f}. \quad (6.5)$$

The constraint condition of equation (6.5) is enforced by applying the Bregman iteration (Goldstein, and Osher 2009; Yin et al. 2008)

$$\min_{\hat{f}, D_x, D_y} \|D_x\|_1 + \|D_y\|_1 + \frac{\mu}{2} \left\| \hat{f} - f^{it} \right\|_2^2 + \frac{\beta}{2} \left\| D_x - \nabla_x \hat{f} - b_x^k \right\|_2^2 + \frac{\beta}{2} \left\| D_y - \nabla_y \hat{f} - b_y^k \right\|_2^2, \quad (6.6)$$

where the values of  $b_i^k$  above correspond to the Bregman iteration

$$(b_i^k = b_i^{k-1} + \left( \nabla_i \hat{f}^k - D_i^k \right)) \text{ and } \beta \text{ is the denoising parameter.}$$

The  $l_1$  and  $l_2$  components of this functional can be split and efficiently solved by SB (Goldstein, and Osher 2009), which iteratively minimizes  $\hat{f}$  and  $D_i$  separately:

$$\begin{aligned} \hat{f}^{k+1} &= \min_{\hat{f}} \frac{\mu}{2} \left\| \hat{f} - f^{it} \right\|_2^2 + \frac{\beta}{2} \left\| D_x^k - \nabla_x \hat{f} - b_x^k \right\|_2^2 + \frac{\beta}{2} \left\| D_y^k - \nabla_y \hat{f} - b_y^k \right\|_2^2 \\ D_i^{k+1} &= \min_{D_i} \|D_i\|_1 + \frac{\beta}{2} \left\| D_i - \nabla_i \hat{f}^{k+1} - b_i^k \right\|_2^2. \end{aligned} \quad (6.7)$$

Note that SB decouples  $\hat{f}$  from the  $l_1$  portion of the problem, thus making  $\hat{f}$  differentiable. To optimally solve  $\hat{f}$  in a cost efficient manner, we used the Gauss-Seidel method, as proposed in (Goldstein, and Osher 2009)

$$\begin{aligned}
\hat{f}_{i,j}^{k+1} = & \frac{\beta}{\mu+4\beta} \left( \hat{f}_{i+1,j}^k + \hat{f}_{i-1,j}^k + \hat{f}_{i,j+1}^k + \hat{f}_{i,j-1}^k + D_{x,i-1,j}^k - D_{x,i,j}^k + D_{y,i,j-1}^k - D_{y,i,j}^k - b_{x,i-1,j}^k + b_{x,i,j}^k - b_{y,i,j-1}^k + b_{y,i,j}^k \right) + \\
& + \frac{\mu}{\mu+4\beta} f_{i,j}^{it}
\end{aligned} \tag{6.8}$$

Furthermore, since there is no coupling between elements of  $D$ , we can use shrinkage operators to compute separately the optimal values of  $D_x$  and  $D_y$

$$D_i^{k+1} = \text{shrink} \left( \nabla_i \hat{f}^{k+1} + b_i^k, \frac{1}{\beta} \right) = \frac{\nabla_i \hat{f}^{k+1} + b_i^k}{\left| \nabla_i \hat{f}^{k+1} + b_i^k \right|} * \max \left( \left| \nabla_i \hat{f}^{k+1} + b_i^k \right| - \frac{1}{\beta}, 0 \right). \tag{6.9}$$

Further details about shrinkage operators can be found in (Goldstein, and Osher 2009; Setzer 2009). A pseudocode of the ART-SB algorithm is presented in Table 6.1.

**Table 6.1.** ART-SB algorithm.

---

**ART-SB algorithm**

---

$$\tilde{f}^0 = [0, 0, \dots, 0] \in \mathbb{R}^N$$

while  $\left\| \tilde{f}^{it} - \tilde{f}^{it-1} \right\|_2 > \text{tol\_1}$  (where  $\text{tol\_1}$  is a given tolerance)

**1st step (ART iteration loop):**

$$f^{it} = \min_{\tilde{f}^{it}} \left\| W \tilde{f}^{it} - d \right\|_2^2 \text{ by ART (6.1)}$$

**2nd step (SB for each z-projection):**

for  $\xi = 1 \dots \text{nz}$  (z-projection loop,  $\text{nz}$  is the number of z-slices)

$$f_\xi^0 = f^{it}(x, y, \xi), D_x^0 = D_y^0 = b_x^0 = b_y^0 = 0$$

while  $\left\| \hat{f}_\xi^k - \hat{f}_\xi^{k-1} \right\|_2 > \text{tol\_2}$  (SB loop)

$$\hat{f}_\xi^{k+1} = \min_{\hat{f}_\xi} \frac{\mu}{2} \left\| \hat{f}_\xi - f_\xi^{it} \right\|_2^2 + \frac{\beta}{2} \left\| D_x^k - \nabla_x \hat{f}_\xi - b_x^k \right\|_2^2 + \frac{\beta}{2} \left\| D_y^k - \nabla_y \hat{f}_\xi - b_y^k \right\|_2^2 \text{ by (6.8)}$$

$$D_x^{k+1} = \min_{D_x} \left\| D_x \right\|_1 + \frac{\beta}{2} \left\| D_x - \nabla_x \hat{f}_\xi^{k+1} - b_x^k \right\|_2^2 \text{ by (6.9)}$$

$$D_y^{k+1} = \min_{D_y} \left\| D_y \right\|_1 + \frac{\beta}{2} \left\| D_y - \nabla_y \hat{f}_\xi^{k+1} - b_y^k \right\|_2^2 \text{ by (6.9)}$$


---

---


$$b_x^{k+1} = b_x^k + \left( \nabla_x \hat{f}_\xi^{k+1} - D_x^{k+1} \right)$$

$$b_y^{k+1} = b_y^k + \left( \nabla_y \hat{f}_\xi^{k+1} - D_y^{k+1} \right)$$

*end (end of SB loop)*

*end (end of z-projection loop)*

$$\tilde{f}^{it+1} = \left( \hat{f}_1^{k+1}, \dots, \hat{f}_{nz}^{k+1} \right)$$

*end (end of while)*

---

### 6.2.3. Experimental and simulated data

#### 6.2.3.1. Experimental phantom data

A 10-mm thick slab-shaped phantom was built using a resin base and adding titanium dioxide and India ink to provide a reduced scattering coefficient of  $\mu_s' = 0.8 \text{ mm}^{-1}$  and an absorption coefficient of  $\mu_a = 0.01 \text{ mm}^{-1}$  (as described in section 1.3.2). The phantom had a 5-mm diameter cylindrical hole filled with a fluid that matched the optical properties of the resin (Cubeddu et al. 1997) mixed with Alexa fluor 700 1  $\mu\text{M}$  (Invitrogen, Carlsbad, California, USA). The fDOT fluorescence and excitation data were acquired with the non-contact parallel plate fDOT scanner presented in section 1.3.1, using 9x9 source positions and 9x9 detector positions over a  $12 \times 12 \text{ mm}^2$  surface.

#### 6.2.3.2. Simulated data

Simulated data were calculated for a numerical phantom designed to have the same optical properties than those in the real phantom. For the simulation of the excitation and fluorescent average intensity, and for the construction of the forward matrix we used an in-house version of the TOAST toolbox (Schweiger 1994; Schweiger et al. 1995) introduced in section 1.3.4, adapted for fDOT. Sources and measurements were modelled as explained in section 1.3.4. The number of sources, number of detectors, and the surface covered by them were equal to those used with the experimental data. The phantom was simulated using a fine finite element mesh (145000 nodes). The average intensity for the forward matrix was reconstructed on a coarser finite element mesh (55000 nodes) and mapped into a uniform mesh of  $20 \times 20 \times 10$  voxels.

The simulation was perturbed with different levels of additive Gaussian noise: 1%, 3%, 5% and 10%.

The target,  $f_{true}$ , corresponding to the physical slab geometry phantom with a cylindrical region filled with fluorophore was modelled using the same finite element mesh used for the simulated data and subsequently mapped into a uniform mesh of 20x20x10 voxels.

#### 6.2.4. Comparison between ART and ART-SB

To assess the effect of choosing different ART-SB algorithm parameters, we reconstructed both acquired and simulated data, for a range of relaxation parameters,  $\lambda = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1)$  and a range of weighting parameters,  $\mu = (0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5)$ .

The stop criterion for ART and ART-SB was a change in the relative solution error norm lower than 0.1% from the previous iteration.

##### 6.2.4.1. Simulated data

ART and ART-SB reconstructions were compared in terms of convergence, signal-to-noise ratio (SNR), and image profiles.

Convergence was assessed by visualizing the relative solution error norm against the number of iterations. The relative solution error norm (with respect to the target) was calculated as

$$E_{rel}(f) = \frac{\|f - f_{true}\|_2}{\|f_{true}\|_2}, \quad (6.10)$$

where  $f_{true}$  is the target solution projected onto the reconstruction mesh.

We compared the contrast of ART and ART-SB reconstructions using horizontal profiles drawn at the center of the image. The profiles were normalized by the average of highest voxel values in the corresponding reconstructions within a region of interest (2x2 mm) around fluorescent target.



SNR was calculated as

$$SNR = 20 \log_{10} \frac{\|signal\|_2}{\|noise\|_2}. \quad (6.11)$$

#### 6.2.4.2. Experimental phantom data

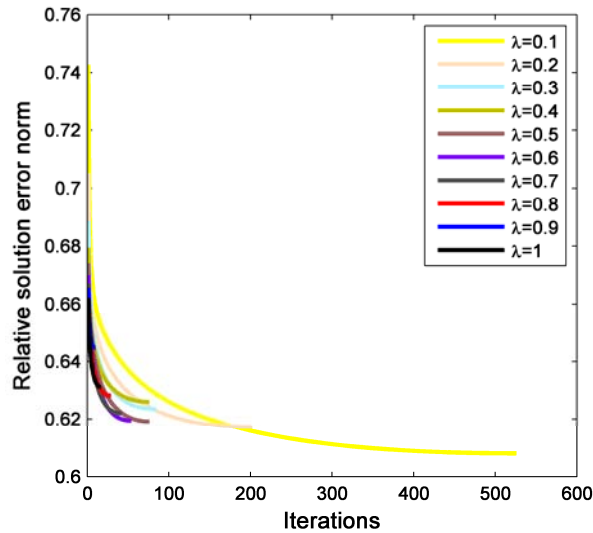
For experimental data only an approximate estimate of the solution target can be estimated. For this reason, ART and ART-SB were only compared in terms of SNR and image profiles, as described above.

### 6.3. Comparison results of ART versus ART-SB

#### 6.3.1. Selection of parameters

##### 6.3.1.1. Selection of relaxation parameter for ART

Figure 6.1 shows the relative solution error norm versus the different  $\lambda$ 's tested for simulated data with 1% Gaussian noise.



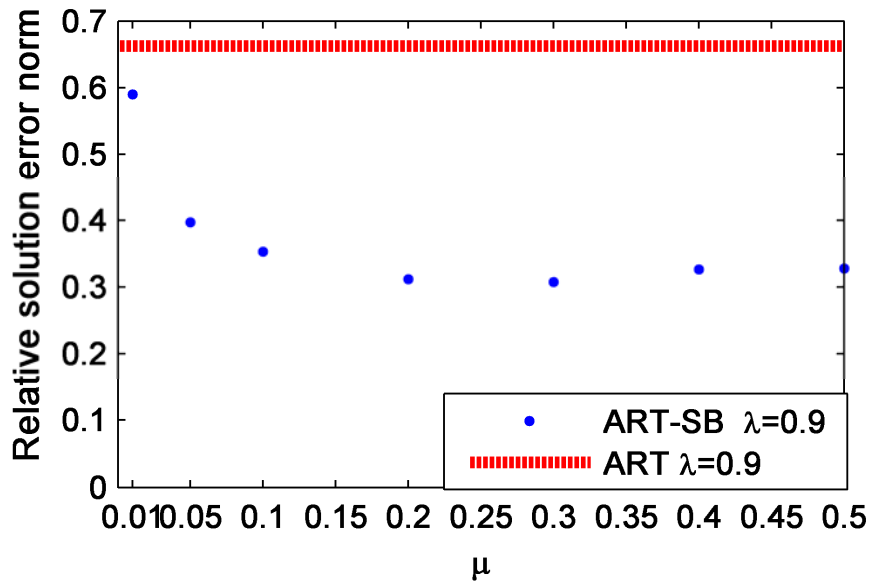
**Fig. 6.1.:** Relative solution error norm of ART reconstruction for every different value of  $\lambda$ . Simulated data with 1% Gaussian noise.

All reconstructions converged approximately to the same relative solution error norm ( $< 0.05\%$  maximum difference) but requiring a different number of iterations, ranging from 22 to 435.

Figure 6.1 confirms that the role of the regularization parameter is now played by the number of iterations and it demonstrates that ART is robust in terms of solution error norm for a wide range of relaxation parameters.

### 6.3.1.2. Selection of weighting and denoising parameters of ART-SB

Once established that, the figure 6.2 shows the minimum solution error norm achieved with ART-SB reconstructions of simulated data with 1% additive normal noise, for different weighting parameters,  $\mu$  ( $\beta = 2\mu$  and  $\lambda = 0.9$ ). The relative solution error norm achieved by ART for the same relaxation parameter  $\lambda = 0.9$  is represented by a horizontal dashed-line.



**Fig. 6.2.** Relative solution error norm of reconstruction of simulated data by ART-SB taking  $\beta = 2\mu$ , and varying the weighting parameters,  $\mu$ , for a relaxation parameter  $\lambda = 0.9$ . The dashed-line indicates the relative solution norm of ART with  $\lambda = 0.9$ . Results correspond to simulated data with 1% additive normal noise.

Figure 6.2 shows a noticeable improvement of relative solution error norm of ART-SB over ART, even at low noise level (1%). Furthermore, figure 6.2 illustrates that there is a value of  $\mu$  above which the relative solution error norm stagnates ( $\mu \geq 2$  in the figure 6.2).

After splitting the problem in equation 6.7,

$$\begin{aligned} \hat{f}^{k+1} &= \min_{\hat{f}} \frac{\mu}{2} \left\| \hat{f} - f^{\text{it}} \right\|_2^2 + \frac{\beta}{2} \left\| D_x^k - \nabla_x \hat{f} - b_x^k \right\|_2^2 + \frac{\beta}{2} \left\| D_y^k - \nabla_y \hat{f} - b_y^k \right\|_2^2 \\ D_i^{k+1} &= \min_{D_i} \left\| D_i \right\|_1 + \frac{\beta}{2} \left\| D_i - \nabla_i \hat{f}^{k+1} - b_i^k \right\|_2^2, \end{aligned}$$

we can see that the choice of  $\beta$  affects to  $D$  and  $f$  subproblems, while the choice of  $\mu$  determines how much the image is regularized ( $f$  subproblem corresponding to equation 6.8). Besides, in the  $D$  subproblem (corresponding to equation 6.9), the solution  $D$  is equal to  $(\nabla f + b)$  after shrinking its vector magnitude by  $1/\beta$  (equation 6.9); this effect is more dramatic when  $\beta$  is small. Thus, once fixed  $\beta = 2\mu$ , lower values of  $\mu$  lead to smoother reconstructions.

### 6.3.2. Comparison between ART and ART-SB

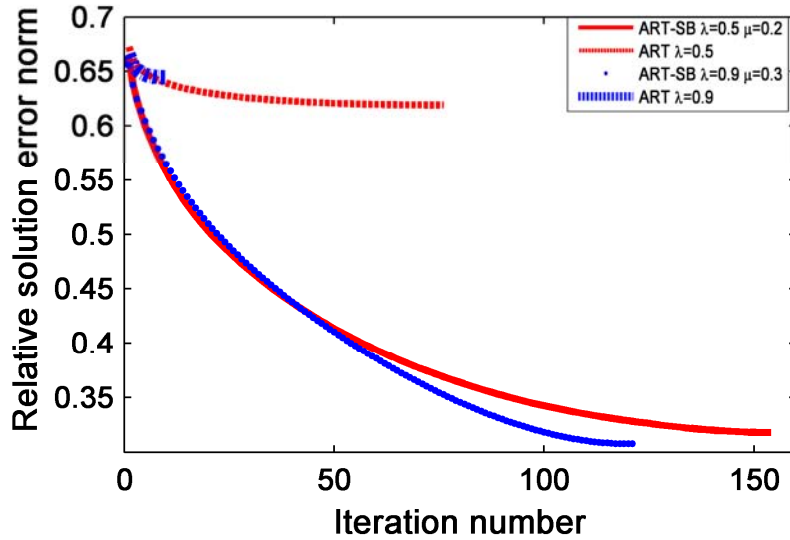
In figure 6.1, we demonstrated that all reconstructions converged approximately to the same relative solution error. However, as commented in section 6.2.1, ART using a low relaxation parameter approximates a weighted least square solution leading to over-smoothed images. On the contrary, high relaxation parameters lead to high-resolution images with noise and artefacts.

In ART-SB method, ART was used to fit the data while SB filtered the noise in the reconstructed image. In view of the above, we compared ART with ART-SB using two relaxation parameter values:  $\lambda = 0.9$  and  $\lambda = 0.5$ .

#### 6.3.2.1. Simulated data

- Convergence:

The faster convergence of ART-SB compared with ART can be observed in a plot of the relative solution error norm versus iteration number, for simulated data (figure 6.3)

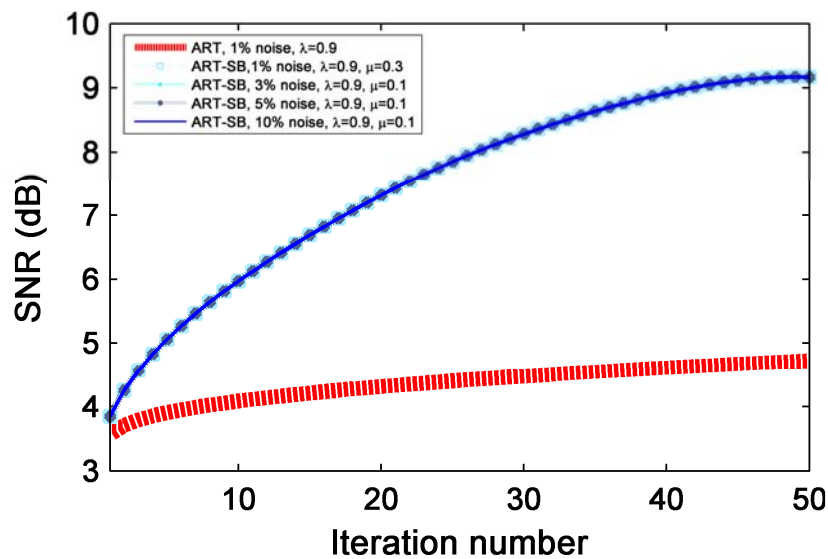


**Fig. 6.3.** Relative solution error norm against iteration number to show the convergence of ART and ART-SB for two different relaxation parameter values (simulated data with 1% additive normal noise).

Note that the mean CPU time for performing the SB denoising for each ART iteration (Intel®-Core™ 2 Quad CPU, 2.40GHz, 4 GB de RAM, Windows Vista) was 0.021 seconds. Therefore, considering that we need about 120-160 iterations for the examples in figure 6.3, less than 4 seconds are necessary for SB denoising.

- SNR:

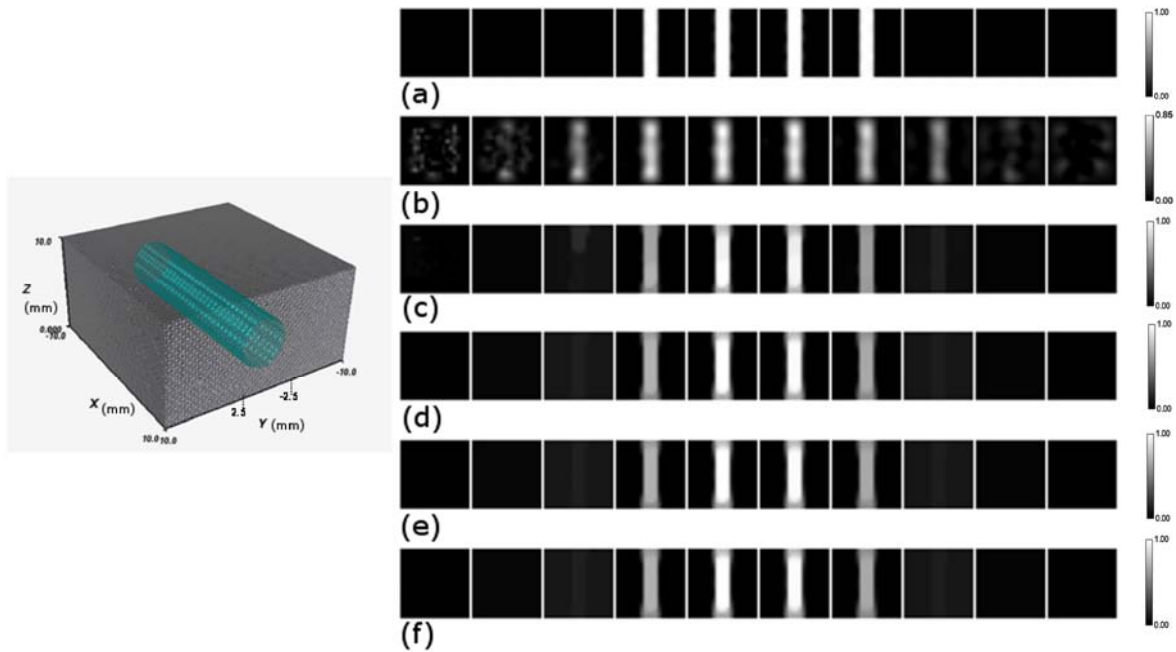
ART-SB led to consistently higher SNR than ART with simulated data (figure 6.4)



**Fig. 6.4.** SNR (dB) plotted against iteration number for ART and ART-SB with relaxation parameter  $\lambda = 0.9$  and denoising parameter  $\beta = 2\mu$ , where  $\mu = 0.1$  (simulated data with different levels of additive Gaussian noise).

- Image reconstruction:

The figure 6.5 shows a comparison of z-slices of ART reconstruction of simulated data with the lowest noise level tested (1%) versus the same z-slices of ART-SB reconstructions with different noise levels (1%, 3% , 5% and 10%).

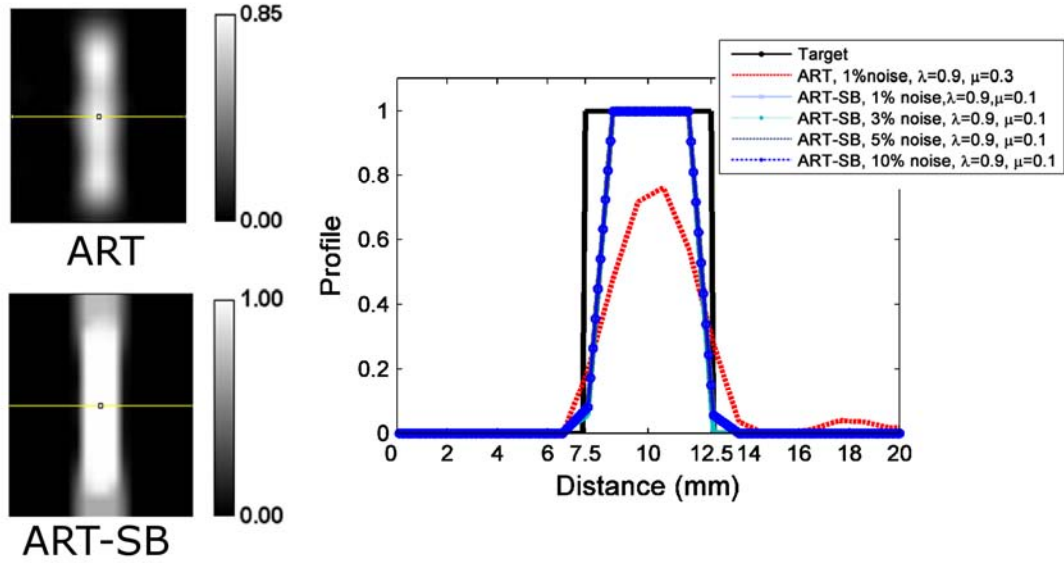


**Fig. 6.5.** Left: Finite element model corresponding to the simulated phantom. Right: 1 mm z-slices (y-x planes) of a) target solution. b) ART reconstruction (1% additive noise). c) ART-SB reconstruction (1% additive noise) with  $\mu = 0.3$ . d) ART-SB reconstruction (3% additive noise) with  $\mu = 0.1$ . e) ART-SB reconstruction (5% additive noise) with  $\mu = 0.1$ . f) ART-SB reconstruction (10% additive noise) with  $\mu = 0.1$ . In all of these cases, the relaxation parameter and the denoising parameter were  $\lambda = 0.9$  and  $\beta = 2\mu$ , respectively.

Figure 6.5 illustrates that even the ART reconstruction for simulated data with the lowest noise level tested (1%) is worse than ART-SB reconstruction for simulated data with higher noise levels.

- Image profiles:

In the Y-profiles drawn over ART and ART-SB reconstructed images (figure 6.6) we can see that the ART-SB profiles are closer to the target than those obtained with ART.

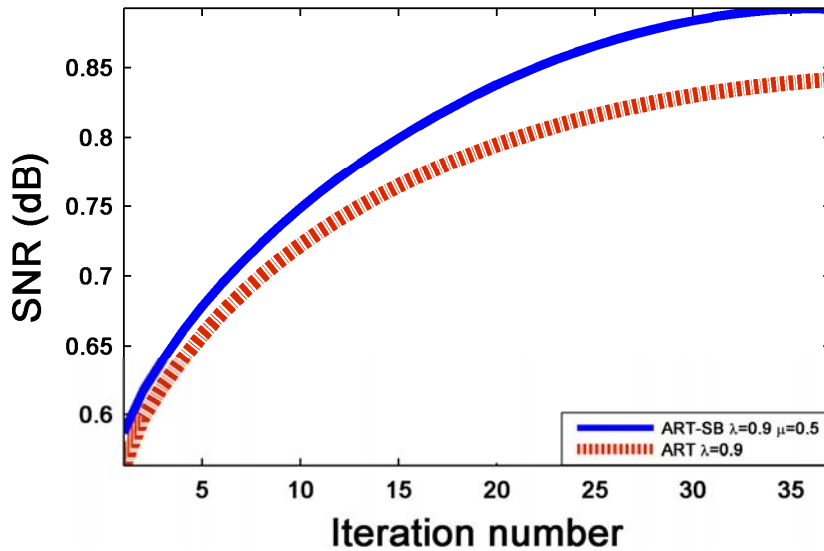


**Figure 6.6.** Y-profiles of central z-slice from ART and ART-SB reconstructions provided in previous figure.

### 6.3.2.2. Experimental phantom data

- SNR:

Similarly to the simulated data case, the plot of SNR against iteration number for ART-SB and ART reconstructions shows a higher SNR for ART-SB.



**Fig. 6.7.** SNR (dB) plotted against iteration number of ART and ART-SB with relaxation parameter  $\lambda = 0.9$  and denoising parameter  $\beta = 2\mu$ , where  $\mu = 0.5$ .

- Image reconstruction:

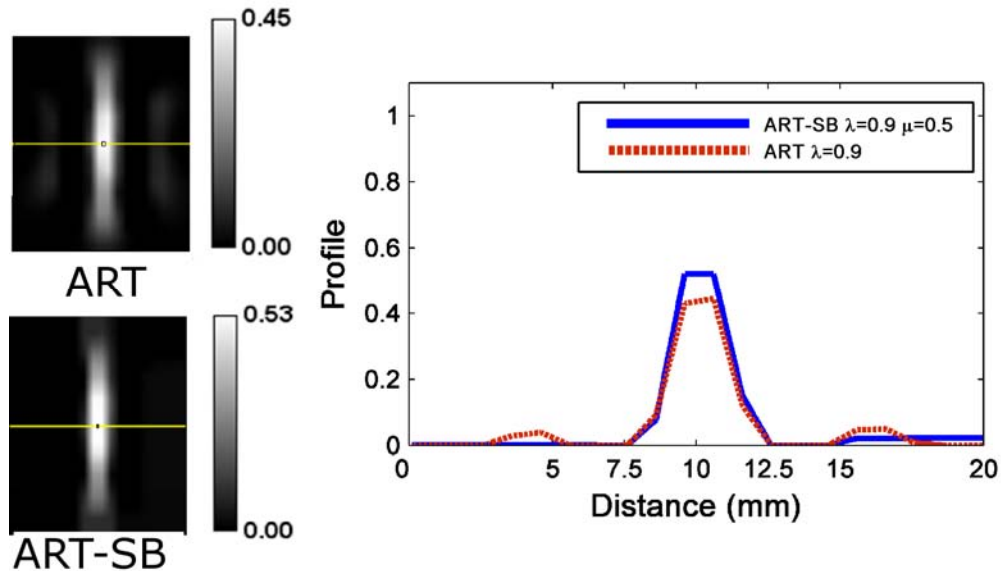
The figure 6.8 shows a comparison of z-slices from ART and ART-SB reconstructions of experimental data. We can observe that ART-SB improved the localization against ART that is accompanied by a loss of resolution along the z-axis.



**Fig. 6.8.** Left: Image of the experimental phantom used. Right: 1 mm z-slices (y-xplanes) of experimental data study. a) ART reconstruction. b) ART-SB reconstruction with denoising parameter  $\beta = 2\mu$ , where  $\mu = 0.5$ . In both ART and ART-SB the relaxation parameter was  $\lambda = 0.9$ .

- Image profiles:

Figure 6.9 shows Y-profiles over the ART and ART-SB reconstructions. The peak-to-valley ratio of ART-SB reconstruction doubled that of ART (ART-SB: 19.326; ART: 9.0427).



**Fig. 6.9.** Y-profiles of ART and ART-SB reconstructions

## 6.4. Discussion and Conclusions

In this chapter we propose a novel iterative algorithm, ART-SB, which alternates a cost-efficient reconstruction method (ART) with a denoising step based on the minimization of TV using SB, which is also solved in a cost-efficient way. Although ART-SB is a state-of-the-art “shrinkage methodology” (Correia et al. 2011; Douiri et al. 2007; Freiburger et al. 2010; Johnston et al. 2010; Pan et al. 2010; Sawatzky et al. 2008) it provides a novel approach to the solution of  $l_1$ -regularized problems, minimizing TV by means of the SB method introduced by (Goldstein, and Osher 2009).

In contrast to (Abascal et al. 2011; Behrooz et al. 2012), we used the SB denoising formulation, which is solved efficiently, without computing first and second derivatives of the TV functional. SB denoising using Gauss-Seidel and shrinkage has a relatively small memory footprint compared with second-order methods that require explicit representations of the Hessian matrix (Goldstein, and Osher 2009). In (Goldstein, and Osher 2009) it was shown that this way of solving SB denoising improves the speed of convergence compared with a gradient descent algorithm or dual formulation of the Rudin Osher Fatemi functional. Thus, ART-SB is a practical method for solving large dataset problems, because ART does not need to hold the system matrix in memory and our implementation of SB denoising does not require an explicit representation of the Hessian matrix.

We compared ART-SB and ART in terms of convergence, SNR, and quality of image profiles for simulated data, and in terms of SNR and image profiles for experimental data. The results indicate that ART-SB enhanced the quality of reconstructions with lower noise and faster convergence than ART. Convergence of ART (figures 6.1 and 6.3) is fast during the first few iterations, after which it stagnates, in agreement with (Hansen 2010).

ART-SB provided significantly improved localization and sharpened edges. ART shows more blurred reconstructions and a loss of resolution along the z-axis (figures 6.5 and 6.8; y-profiles figures 6.6 and 6.9). Furthermore, ART deteriorates with increased noise levels (Hansen 2010), while ART-SB is more robust (see z-slices in figures 6.5 and 6.8 and y-profiles in figures 6.6 and 6.9). These findings are in agreement with the conclusions of other works on two-step reconstruction algorithms for computed tomography and positron emission tomography (Johnston et al. 2010; Pan et al. 2010; Sawatzky et al. 2008) that, as we pointed out in 6.1, combined simulated algebraic



reconstruction technique or expectation maximisation method with TV-denoising solved by gradient based methods requiring high computational burden.

In section 6.3.1.2 we described that lower values of  $\beta = 2\mu$  lead to smoother reconstructions. Since the shrinkage operators (only depending on  $\beta$ ) are evaluated at each iteration, the number of required iterations for convergence depends on the  $\beta$  value (further details can be found in (Goldstein, and Osher 2009)). Thus, lower values of  $\beta = 2\mu$  increase the number of iterations before reaching convergence.

A limitation of our study is that, although we assessed the effect of a range of  $\lambda$  and  $\mu$  parameters, we did not explore other values of  $\beta$  than  $\beta = 2\mu$ , as suggested in (Goldstein, and Osher 2009). A further study about different choices of  $\beta$  is warranted.

Another limitation is that we tested the algorithm with simple phantom geometries only. It remains still unknown whether the improvements are maintained when dealing with real anatomical structures (smoother regions without sharp edges) or not.

One possible future line would be to test whether 3D-SB denoising implementations could lead to even better results.

In conclusion, in this chapter we show that the combination of a cost-efficient linear iterative technique (ART) with a denoising method (anisotropic SB) is well suited for large datasets (such as those involved in DOT and fDOT) and significantly improves the reconstruction of phantom fDOT data.



## Chapter 7

# Compressed Sensing in fDOT

Compressed Sensing (CS) is an increasingly popular technique due to its ability to speed up data acquisition in many modalities. As we pointed out in chapter 5, regularization methods based on the  $l_1$ -norm and total variation (TV) have gained interest as a result of CS theory.

However, most of the CS theory is devoted to undetermined problems and there are few contributions that apply it to ill-conditioned problems. In this chapter we present a CS method for reconstruction of the ill-posed problem of fluorescence diffuse optical tomography (fDOT), based on the analysis co-sparse representation model, where an analysis operator multiplies the image and the outcome is a sparse image. This method combines a Split Bregman algorithm to solve CS (CS-SB) problems with a theorem about the effect of ill-conditioning on  $l_1$  regularization, stating that  $l_1$  regularization problem depends on how the sparsity of the true solution and the singular values (SVs) of the forward matrix relate. Our method, SB-SVA restricts the solution reached at each CS-SB iteration to a determined space where the SVs of the forward matrix and the non-zero values of the solution in each iteration combine in a beneficial manner. In this way, SB-SVA forces indirectly the well-conditioning of the forward matrix while designing (learning) the analysis operator and reconstructing the image. We tested the method with fDOT simulated and experimental data, and found beneficial improvement with respect to the results of standard CS-SB algorithm without the restriction cited above.

This chapter is organized as follows: we first introduce mathematical basics of CS applied to undetermined systems in section 7.1; afterwards, in section 7.2, we present a short review on CS applied to the ill-posed problem of fDOT, highlighting the incoherence of fDOT forward matrix itself. Finally, we present a novel CS approach for fDOT, named Split Bregman-Singular Value Analysis (SB-SVA) method in section 7.3 and we show the results in section 7.4. To close, section 7.5 presents the discussion and conclusions.

## 7.1. Mathematical basics

If we consider an image vector  $x \in \mathbb{R}^n$ , a measurement vector  $b \in \mathbb{R}^m$  and a matrix  $A \in \mathbb{R}^{m \times n}$  that relates  $x$  and  $b$  in a linear way,  $b = Ax$ , being  $m \leq n$ .

The Sparse Synthesis Model

$$\|b - Az\|_2^2 + \lambda \|z\|_1, \quad z = Dx, \quad (7.1)$$

states that this signal could be described as  $x = Dz$ , where  $D \in \mathbb{R}^{n \times p}$  is a redundant dictionary ( $p \geq n$ ), and  $z \in \mathbb{R}^p$  a sparse representation of the signal. The model assumes that any image could be described as a linear combination of few columns from the dictionary  $D$ . As its name suggests, the synthesis model describes a way to synthesize the image. In this case, we expect the number of non-zeros in  $z$  to be smaller than  $p$ . The non-zero entries of  $z$  is named support of  $x$ .

The Co-sparse Analysis Model

$$\|b - Ax\|_2^2 + \lambda \|Tx\|_1, \quad (7.2)$$

states that the analysed vector  $Tx$  is expected to be sparse, where  $T \in \mathbb{R}^{t \times n}$  is a possible redundant analysis operator ( $t \geq n$ ).

In this case, we consider an image  $x$  that minimizes  $\|Tx\|_1$ . Here, we care about the zeros of  $Tx$  (also named co-support of  $x$ ) (Candes et al. 2011; Elad, Milanfar, and Rubinstein 2007; Nam et al. 2013; Theodoridis et al. 2012).

### 7.1.1. Sparse synthesis model

The uniqueness of the synthesis model solution is commonly verified by studying the spark of system matrix (Donoho, and Tanner 2009), the mutual coherence (Mallat, and Zhang 1993) or the Restricted Isometry Property (RIP) condition (Candes, and Romberg 2005). We will now introduce these concepts.

In this section, for the sake of simplicity,  $A \in \mathbb{R}^{m \times n}$  will be termed the sensing matrix with  $m \leq n$  and  $x \in \mathbb{R}^n$  itself sparse (i.e.  $D=I$ ). Note that, if  $x \in \mathbb{R}^n$  is sparsified using a dictionary  $D$ , then  $AD$  is the sensing matrix.

#### 7.1.1.1. The spark of a matrix

The  $\text{spark}(A)$  is defined as the smallest number of linear dependent columns. Note that the  $\text{spark}$  can only be obtained with a combinatorial search of all possible combinations of the columns of the respective matrix. This means that, when the matrix is large, calculating the  $\text{spark}$  can be expensive in terms of computational burden.

If a linear system of equations  $Ax = b$  has a solution that satisfies  $\|x\|_0 < \frac{1}{2} \text{spark}(A)$  then this is the sparsest possible solution. This is, necessarily, the unique solution of  $l_0$ -minimizer (Donoho, and Elad 2003; Gribonval, and Nielsen 2003).

#### 7.1.1.2. The mutual coherence

The mutual coherence of a matrix  $A$  is defined by

$$\mu(A) = \max_{1 \leq i < j \leq n} \frac{|A_i^T A_j|}{\|A_i\| \|A_j\|}, \quad (7.3)$$

where  $A_i$ ,  $i = 1, 2, \dots, n$  denotes the columns of  $A$ .

Normalized mutual coherence is bounded,  $0 \leq \mu \leq 1$ . It provides a measure of the worst similarity between the matrix columns. Note that the sensing matrix  $A$  relates to the unknown vector  $x$  and the measurement vector  $b$ ; i.e.,  $b$  is the result of a combination of the columns of  $A$ , each one weighted by a different component of  $x$ . Then, if the columns of  $A$  are independent as much as possible, the information regarding each component of  $x$  is contributed by a different direction making its recovery easier (Theodoridis et al. 2012). In conclusion, the columns of  $A$  should be as orthogonal as possible.

Mutual coherence is also a powerful tool to find the optimal dictionaries, based on the knowledge that sensing matrices need to have the smallest possible mutual coherence.

A different way to understand mutual coherence is by considering the Gram matrix,  $G = A^T A$ , given that the off-diagonal entries in  $G$  are the inner products that appear in equation 7.3 (Elad 2007).

#### 7.1.1.3. $l_0$ and $l_1$ -minimizer solutions

The concept of mutual coherence contributes to find a correspondence between  $l_0$  and  $l_1$ -minimizer solutions. Given  $Ax = b$ , if a solution exists and satisfies

$$\|x\|_0 < \frac{1}{2} \left( 1 + \frac{1}{\mu(A)} \right), \quad (7.4)$$

then this is the unique solution of both the  $l_0$  and  $l_1$ -minimizers (Donoho, and Elad 2003; Gribonval, and Nielsen 2003).

#### 7.1.1.4. The Restricted Isometry Property (RIP)

The isometry constant  $\delta_k$  of matrix  $A$  for each integer is defined as the smallest number such that

$$(1 - \delta_k) \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k) \|x\|_2^2 \quad (7.5)$$

holds true for all  $k$ -sparse vectors  $x$ .

A matrix obeys RIP of order  $k$  if  $\delta_k$  is not close to one. As we remarked in subsection 7.1.1.2, the columns of  $A$  should be as orthogonal as possible. Note that, if  $\delta_k$  is closed to zero, the subsets of  $k$  columns of  $A$  are orthonormal.

When RIP holds the Euclidian norm of  $x$  is approximately preserved after projecting onto the rows of  $A$ . That is, it preserves Euclidean distances between  $k$ -sparse vectors (Theodoridis et al. 2012).

As in the case of mutual coherence, RIP also relates to the Gram matrix, and more exactly to its condition number (CN). In (Candes, and Romberg 2005) it was stressed that if  $A_r$  denotes the matrix resulting of considering only  $r$  columns of  $A$ , then RIP is equivalent to requiring that the SVs of respective Gram matrix,  $G_r = A_r^T A_r$ , are within the interval  $[1 - \delta_k, 1 + \delta_k]$ . If  $A$  is ill-conditioned, the subspace taking the  $r$  columns of  $A$  will have very likely zero or close to zero SVs. Then, well-conditioning  $A$ , removing the zero or close to zero SVs, is desirable.

### 7.1.2. Co-sparse analysis model

Note that the analysis model in equation 7.2 can also be written as

$$\begin{pmatrix} b \\ 0 \end{pmatrix} = \begin{pmatrix} A \\ T_\Lambda \end{pmatrix} x = Mx \quad (7.6)$$

where  $\Lambda$  denotes the co-support of  $x$ .

To verify the uniqueness of the analysis model solution,  $A \in \mathbb{R}^{m \times n}$  and  $T \in \mathbb{R}^{t \times n}$  must be mutually independent (Nam et al. 2013). Furthermore, it can be shown that when the co-support ( $\Lambda$ ) is known, the null space of  $M$  must to be 0 to be able to uniquely identify  $x$ . That is,

$$\text{Ker}(T_\Lambda) \cap \text{Ker}(A) = \text{Range}(T_\Lambda^T)^\perp \cap \text{Ker}(A) = \{0\} \quad (7.7)$$

Giryes et al. (Giryes et al. 2011) redefined the RIP condition for the analysis model (equation 7.5). The only difference with the RIP condition for the sparse synthesis is that  $x$

is  $k$ -sparse in  $D$  in the synthesis model while  $x$  is  $l$ -cospase with respect to  $T$  in the analysis model.

Besides, the quality of the reconstruction for the different methods based on analysis model highly depends on the right choice of suitable operator, and different works have focused on developing methods to learn the analysis operators (Hawe, Kleinstaubler and Diepold 2013; Nam et al. 2013; Rubinstein, Peleg and Elad 2011; Yaghoobi et al. 2013).

## 7.2. CS applied to the fDOT ill-posed problem

### 7.2.1. *Brief review*

Although most of the CS theory is devoted to undetermined systems (section 7.1), a few authors applied it to highly ill-posed problems.

If the selected submatrices of the sensing matrix are highly ill-posed, there are SVs close (or equal) to zero,  $\delta_k$  becomes arbitrarily close to one and the RIP property is violated. Something similar happens with mutual coherence, which is linked to the CN of submatrices of the sensing matrix. Since RIP and mutual coherence are sufficient but not necessary conditions, their violation does not mean that we cannot obtain the solution.

When dealing with undetermined systems, many studies make use of RIP and mutual coherence to create appropriated sensing matrices (Elad 2007; Theodoridis et al. 2012). However, this cannot be extrapolated when considering an ill-posed problem without first well-conditioning it, that is, eliminating near-zero SVs.

In a recent paper (van den Doel et al. 2012), the authors formulated a theorem stating that the efficiency of an  $l_1$  regularization problem depends on how the sparsity of the true solution and the SVs of the forward matrix relate. This is particularly important when considering ill-posed problems, on which near-zero SVs are involved.

Some authors have applied compressing techniques to fDOT, DOT and fDOT X-ray computed tomography (fDOT-CT). We can divide these works into two categories. The first one includes works based on Fourier or wavelet transformations to sparsify measurements, Born-normalized measurements or the reconstruction images (Ducros et al. 2012; Ripoll 2010; Süzen, Giannoula, and Durduran 2010). The second category groups



works based on preconditioning the forward matrix or submatrices (Cao et al. 2012; Jin et al. 2012; Mohajerani and Ntziachristos 2013) by using the measurements or image redundancy.

Regarding the first group, Ripoll (Ripoll 2010) presented an approach that combines Fourier and real space functions to encode the CCD measurements of DOT. Süzen et al. (Süzen et al. 2010) compared CS using the discrete Fourier transform basis function with SVD reconstruction of DOT simulated data. They concluded that CS improves significantly in terms of contrast, contrast-to-noise ratio, normalized root mean square error and localization error. However, as they pointed out, measurement noise was not considered in their simulations, nor was a study of optimal sparse expansion of the investigated signal. Ducros et al. (Ducros et al. 2012) presented an approach consisting of illuminating the medium with only a few wavelet patterns and compressing the acquired images by means of a wavelet transform, thus reducing acquisition and reconstruction times without sacrificing the reconstruction quality. They investigated the compression ability of different wavelets for the acquired fluorescence images and concluded that Battle-Lemarie functions achieve good compression of fluorescent images with the least degradation, as compared to other bases (Haar, Daubechies, Coiflet, Symlets).

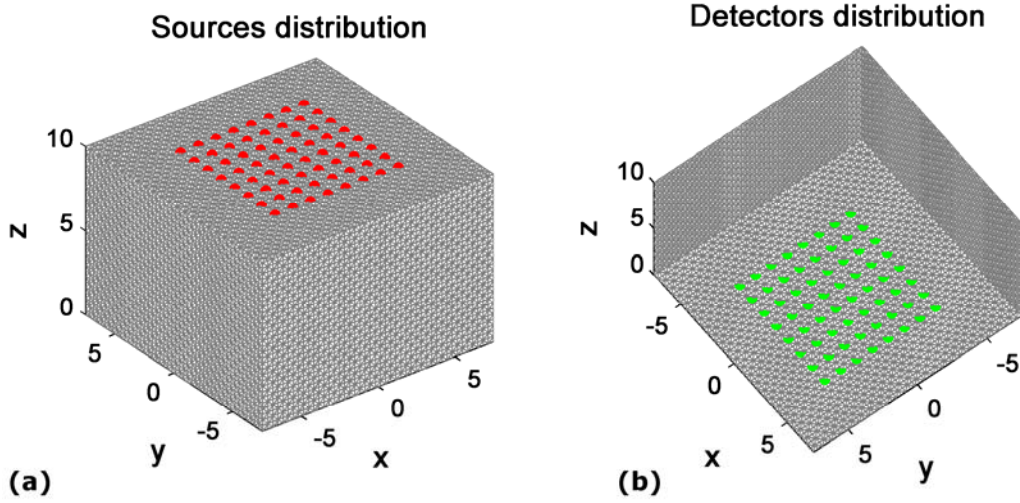
Regarding the second group of techniques, Jin et al. (Jin et al. 2012), focused their research on reducing the coherence of the fDOT forward matrix, based on the fact that sparse signals can be recovered exactly from an undetermined system when the underlying forward matrix is incoherent (Elad 2007). Thus, indirectly, they were well-conditioning the forward matrix. Cao et al. (Cao et al. 2012) considered the correlations of source-detector maps from the same projection and used principal component analysis (PCA) to reduce the dimension of the weight matrix by discarding the less relevant components. Mohajerani and Ntziachristos (Mohajerani and Ntziachristos, 2013) made use of intersource signal dependencies to reduce the size of the fDOT-XCT reconstruction, that is, using PCA they removed the correlation among optical measurements obtained at different sources between adjacent projection angles for 360° rotation geometry of fDOT-XCT.

Nevertheless, the methods presented in this second group are based on the preconditioning (well-conditioning) of the forward matrix.

### 7.2.2. Incoherence of the fDOT forward matrix

Given that most of the fluorophore tracers are designed to label specific tissues, the fluorophore concentration itself is often sparse in the imaging domain. Keeping this in mind, we evaluated the cumulative coherence of the forward matrix of fDOT simulated data, to show that is actually incoherent.

Simulations were done using TOAST (section 1.3.3), keeping constant the mesh volume of interest (VOI) (1.4x1.4x1cm), the detectors FOV (0.8x0.8cm) and the sources FOV (0.8x0.8cm) as shown in figure 7.1.



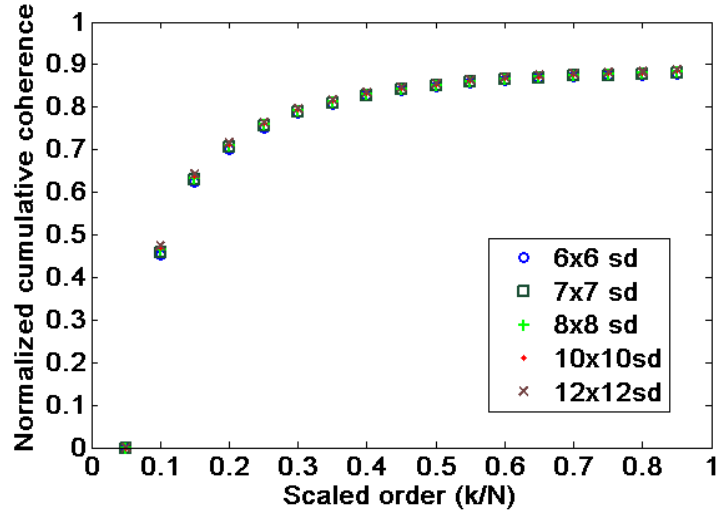
**Fig. 7.1.** Mesh volume of interest. Left: Sources distribution (case 8x8 sources). Right: Detectors distribution (case 8x8 detectors). Note that measurements in the image are given in mm.

We calculated the different forward matrices,  $W$ , for each distribution of sources and detectors as explained in section 1.3.3. After that, we considered the Gram matrix  $G = W^T W$  after normalizing its columns. The off-diagonal entries in  $G$  are the inner products of equation 7.3. The mutual coherence was calculated by taking the off-diagonal entries  $g_{ij}$  with the largest magnitude.

Due to the high computational complexity of this calculation, similarly to (Süzen et al. 2010) the normalized cumulative coherence was calculated for an order of up to  $k_{\max} = 18$ . That is, the number of columns of the forward matrix  $W$  (number of voxels) was chosen to be  $N=3 \times 3 \times 2$ .

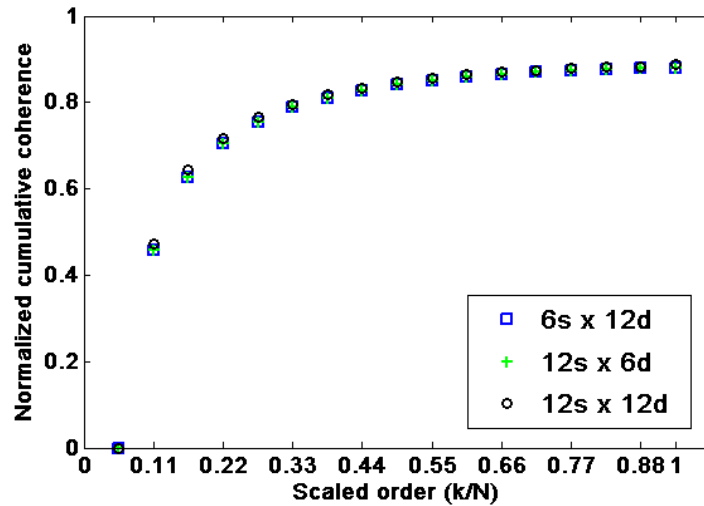
Figure 7.2 shows the normalized cumulative coherence against scaled  $k$ -sparsity order (that is,  $k/N$ , where  $k=2,3,\dots,17$  and  $N=18$ ), using different distribution of source-

detector pairs: from 6x6 sources and 6x6 detectors ( $0.45 \text{ sources/mm}^2$  and  $0.45 \text{ detectors/mm}^2$ ) to 12x12 sources and 12x12 detectors ( $0.1 \text{ sources/mm}^2$  and  $0.1 \text{ detectors/mm}^2$ ).



**Fig. 7.2.** Normalized cumulative coherence of fDOT forward matrix against several scaled orders ( $k/N$ ,  $k=2,3,\dots,17$ ,  $N=18$ ) using different combinations of sources-detector pairs.  $N$  is the number of image voxels.

Figure 7.3 shows the normalized cumulative coherence against scaled  $k$ -sparsity order ( $k/N$ ), using different distributions of source-detectors.



**Fig. 7.3.** Normalized cumulative coherence of fDOT forward matrix for several scaled orders (i.e.  $k/N$ ,  $k=2,\dots,17$ ,  $N=18$ ) using different combinations of source-detectors.  $N$  is the number of image voxels.

Both plots show that, independently of the number of sources/detectors, the increasing trend of the normalized cumulative coherence is always slow, supporting the hypothesis that the fDOT forward matrix is itself incoherent. Thus, finding a basis incoherent with the fDOT forward matrix can be a difficult task.

### 7.3. A novel approach to CS for fDOT: the SB-SVA method

The goal of this section is to present a novel approach, SB-SVA, that solves the analysis-based co-sparse representation model of fDOT reconstruction problem. The novelty of this approach is that it designs (learns) the analysis operator (different to identity matrix) while indirectly forcing the well-conditioning of the forward matrix.

#### 7.3.1. The *Split Bregman (SB) approach to CS*

As we mentioned in chapter 5, it has been shown that the SB method for denoising and CS is computationally efficient (Goldstein and Osher 2009), because the SB formulation leads to a problem that can be solved using Gauss-Seidel and Fourier transforms methods.

In this work, we use SB to solve the co-sparse analysis model problem

$$\min_{f,t} |t| + \frac{\mu}{2} \|Wf - d\|_2^2 + \frac{\lambda}{2} \|t - Tf - b_t\|_2^2 \quad (7.8)$$

where  $W$  is the fDOT forward matrix,  $f$  is a vector representing the concentration of fluorophore at each voxel,  $d$  is a vector containing the acquired measurements,  $T$  represents an analysis operator that provides a sparse representation for  $f$ ,  $Tf \rightarrow t$  and  $b_t$  represents the Bregman iteration that imposes the respective constraint (Goldstein 2009).

In our case  $T$  was chosen as Battle-Lemarie wavelet transform, based on the results of (Ducros et al. 2012). The Battle-Lemarie is a spline wavelet transform; it is symmetric, orthonormal and smooth.

### 7.3.2. The SB-SVA method

The SB-SVA method is based on the combination of an SB algorithm to solve the co-sparse analysis model problem presented in equation (1) with a theorem about the effect of ill-conditioning on  $l_1$  regularization, presented in (van den Doel et al. 2012). This theorem states that in the highly ill-conditioned case and in presence of significant noise, as it is the case of fDOT problems, no small SVs of  $W$  can be tolerated in the set  $Tf_{true} \neq 0$ , being  $f_{true}$  the true solution.

Thus, at each SB iteration we update the co-support in two steps:

- 1) We find the location of the largest entries (in absolute value) of  $Tf^{it} = 0$  and regard them as not belonging in the co-support, similarly to the Greedy Analysis Pursuit algorithm presented in (Nam et al. 2013).

$$L^{it} = \left\{ i : |\alpha_i| \geq th \max_j |\alpha_j| \right\}, \text{ where } \alpha = Tf^{it} \text{ and } th \text{ a certain threshold} \quad (7.9)$$

$$\Lambda^{it} = \Lambda^{it} \setminus L^{it}$$

- 2) Once selected the co-support  $\Lambda^{it}$ , with the goal to well-conditioning our inverse problem, we restrict the solution,  $f^{it}$ , to the subspace where the “sparsity of  $Tf^{it}$  and SVs of  $W$  combine in a beneficial manner”, not allowing small SVs of  $W$  in the set  $Tf^{it} \neq 0$ . Therefore, being  $\Omega$  the image space and  $\Omega_{KerNz}$  the subset of the null-space of  $W$  ( $Ker(W)$ ) corresponding to the complementary set of the co-support of  $f^{it}$  (set  $Tf^{it} \neq 0$  or  $Range(T_{\Lambda^{it}}^T)^\perp$ ), we restricted the transformation  $T$  to the space  $\Omega \setminus \Omega_{KerNz}$  as follows:

$$\begin{cases} T_j = T_j & \text{if } T_j f_j^{it} \in \Omega \setminus \Omega_{KerNz} \\ T_j = 0 & \text{if } T_j f_j^{it} \in \Omega_{KerNz} \end{cases}, \quad (7.10)$$

where the subscript  $j = 1, \dots, n$  indicates voxel indices.

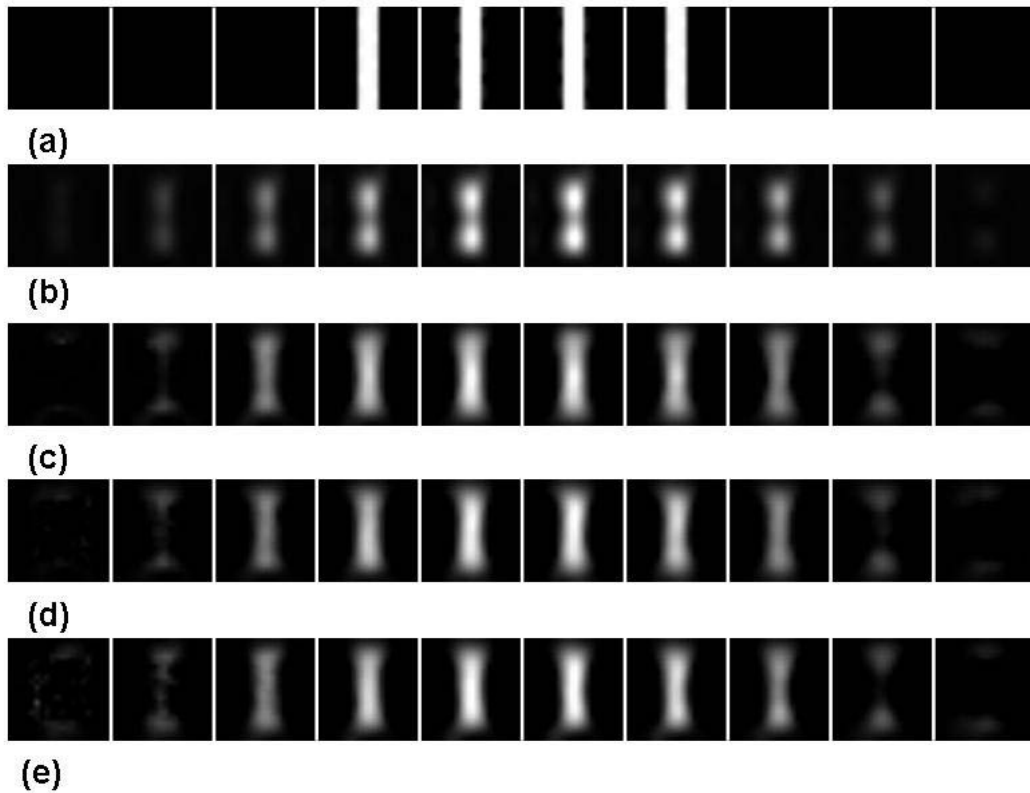
Note that, indirectly, we impose the uniqueness condition introduced in equation (7.7) taking advantage from the fact that in every iteration we know the co-support of  $f^{it}$ .

When  $T$  is a wavelet transformation, like in our case, we scale samples of wavelet transforms following a geometric sequence of ratio two. We have to take this into account when restricting the transformation  $T$  to the space  $\Omega \setminus \Omega_{KerNz}$ . Note that for these cases we are making a slight abuse of notation defining the spaces.

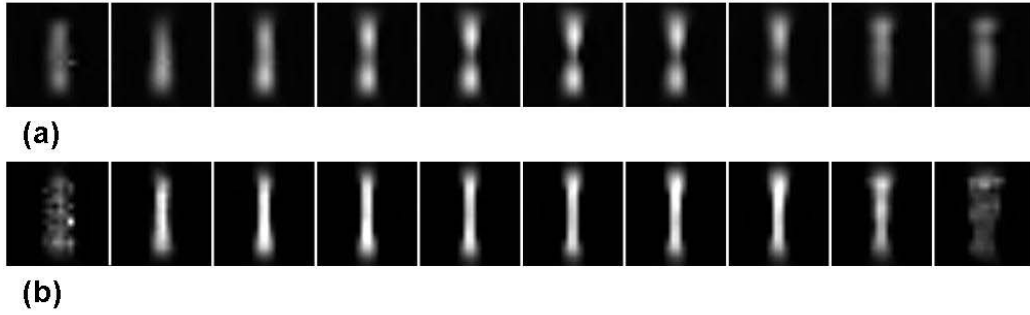
We tested our algorithm using the simulated and experimental fDOT data presented in sections 6.2.3.

## 7.4. Results

Figures 7.3 and 7.4 show z-slices of simulated and experimental data reconstructed with SB and SB-SVA using a Battle-Lemarie wavelet transform.



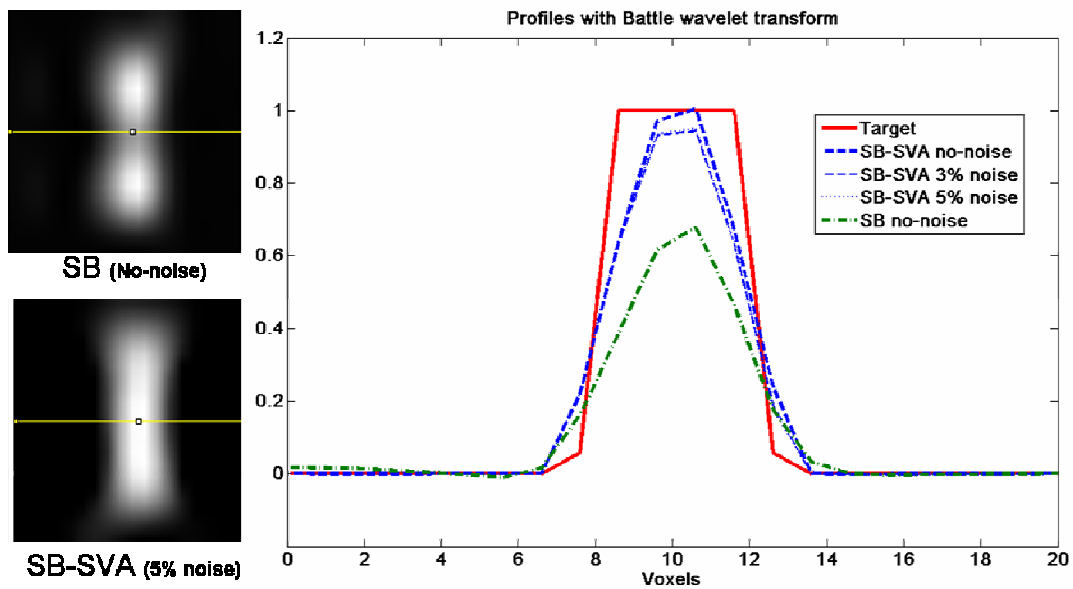
**Fig 7.3.** (a) Target, (b) Reconstruction of no-noise data by CS-SB, (c) Reconstruction of no-noise data by SB-SVA, (d) Reconstruction of 3% noise data by SB-SVA, (e) Reconstruction of 5% noise data by SB-SVA. (f) Reconstruction of 10% noise data by SB-SVA. All of them using Battle-Lemarie wavelet transform as initial transformation and for simulated data.



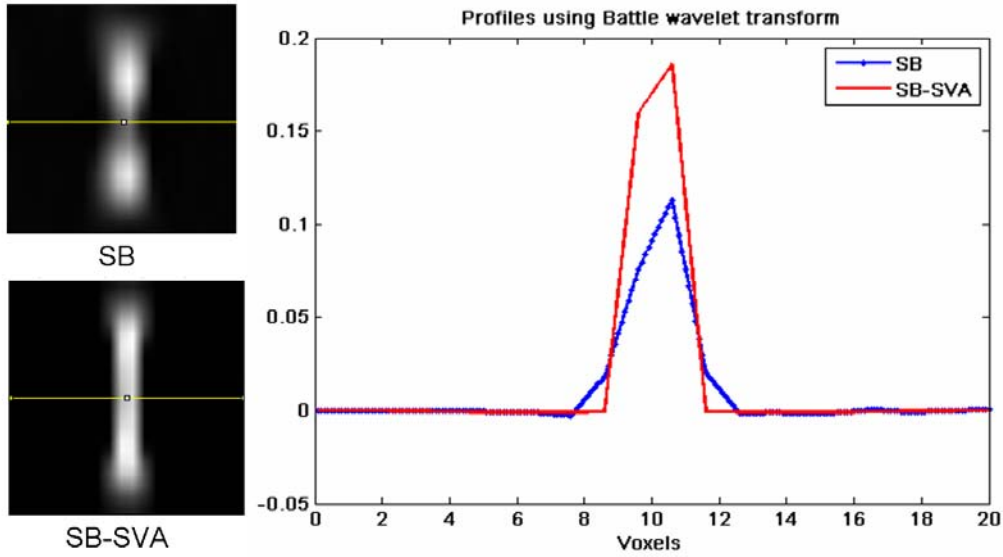
**Fig 7.4.** Reconstructions of experimental data using Battle-Lemarie wavelet transform: (a) by CS-SB, (b) by SB-SVA.

The improvement of SB-SVA versus SB is noticeable for both simulated and experimental phantom data. Furthermore, the solution given by SB-SVA is sparser than the solution given by SB. For example, in the case of no-noise simulated data, the resulting SB-SVA image (given its restriction to  $\Omega \setminus \Omega_{KerNz}$  space) is 11.5% sparser than the SB solution.

In order to compare CS-SB with SB-SVA methods, we also obtained y-profiles. In simulated data, profiles were normalized by the average of highest voxel values in the corresponding reconstructions within a region of interest around fluorescent target.



**Fig. 7.5:** y-profiles of reconstructions of simulated data using Battle-Lemarie wavelet transform. Target (red) and CS-SB (green) for different level of noise (blue).



**Fig. 7.6:** y-profiles of reconstructions of experimental data using Battle-Lemarie wavelet transform. Target (red) and CS-SB (blue).

Profiles of both simulated and experimental phantom data showed a significant improvement provided by SB-SVA versus CS-SB. Furthermore, considering profiles of simulated data, no noticeable differences appear between different noise values.

## 7.5. Discussion and Conclusions

In this chapter, we propose a novel CS reconstruction method, named SB-SVA, for ill-posed fDOT problems that makes use of a co-sparse representation model. The method is based on the CS-SB algorithm and a theorem about the effect of ill-conditioning on  $l_1$  regularization (van den Doel et al. 2012). At each iteration of CS-SB, we restrict the solution to a subspace where the SVs of forward matrix and the sparsity of the iterative solution combine in a beneficial way.

The redundant analysis operator chosen was the Battle-Lemarie wavelet transform. Note that the wavelet coefficients are sparse, however is well-known that low-pass wavelet coefficients of an image (“scaling” coefficients at the lowest resolution scale) are often not sparse. The analysis model does not depend on low-pass coefficients and is therefore not adversely affected when they are not regularized, in contrast to the synthesis model. Thus, the choice of analysis or synthesis model can make substantial differences in the results, in accordance with (Selenick and Figuereido 2009).



Regarding our algorithm, SB-SVA improved reconstruction in terms of image quality and imaging profiles compared with CS-SB (figures 7.3-7.6), even using fewer voxels than SB.

An important fact to point out is that SB-SVA is, indirectly, well-conditioning the sensing matrix while finding the solution due to our restriction to  $\Omega \setminus \Omega_{KerNz}$  space of the solution in each iteration. Note that this restriction is, somehow, eliminating some of the columns of analysis operator  $T$  corresponding to  $Ker(W)$ . Thus, indirectly:

- It detects the desired co-support at each SB iteration. This step is similar to the idea of greedy algorithms for co-sparse analysis models, such as the Greedy analysis pursuit algorithm (GAP) presented in (Nam et al. 2013) that aims to detect the elements outside the set  $\Lambda$  (detecting the desired co-support).
- It reduces the co-support through the iterations, reducing the similarity between the columns of  $T$  and  $W$ , that is, the mutual coherence (section 7.1.1.2).
- It well-conditions the forward matrix (reducing its CN). Besides, the compression of the forward matrix is achieved by maintaining only a few components with large SVs.
- It performs a designing/learning of the analysis operator,  $T$ .

Note that if we change the transformation  $T$  in SB-SVA, the singular value decomposition of the forward matrix does not require recalculation.

SB-SVA is a simple and efficient algorithm since it is based on SB that has been shown its effectiveness for solving  $l_1$ -based regularization problems making it possible to split the minimization of  $l_1$  and  $l_2$  functionals.

Although SB-SVA provides significant improvements in terms of image quality for ill-posed fDOT reconstruction problem, there are some limitations that need further study. The reconstruction along the z-dimension, is not optimal with either SB nor SB-SVA. This poor localization is due to the low resolution of fDOT in the axis perpendicular to the plates (z). Besides, the number of wavelet coefficient to keep in order to select the co-support, was heuristically chosen. An automated and optimized approach would be desirable.

To conclude, dealing with ill-conditioned problems, as fDOT problem, SB-SVA improves SB reconstructions in terms of image quality while it provides simultaneous well-conditioning of the forward matrix and designs the analysis operator.



# Chapter 8

## Conclusions

The main goal of this PhD-thesis was to make use of state-of-the-art inverse problem techniques to develop novel reconstruction methods for solving the fluorescence diffuse optical tomography (fDOT) problem.

The first part of this thesis addressed the optimization of experimental setups to reduce the dataset size, by using  $l_2$ -based regularization techniques. The second part, based on the success of  $l_1$  regularization techniques for denoising and image reconstruction, was devoted to advanced problem regularization using  $l_1$ -based techniques, to finish introducing compressed sensing (CS) theory, which enabled further reduction of the acquired dataset size.

To summarize, the main conclusions of this work are:

1) Regarding  $l_2$ -based regularization techniques, a U-curve-based method was utilized for the first time to select the regularization parameter in  $l_2$  regularization fDOT reconstruction. Since the performance of automatic methods for the selection of this parameter depends on the particular inverse problem, the U-curve method was studied in depth in terms of fulfilment of the Picard's condition, image resolution and image noise. Results showed that the U-curve approach may constitute a good alternative in cases where the well-known L-curve method yields unsatisfactory results. Furthermore, the U-curve method provides an interval for the optimal regularization parameter. This fact increases the computational efficiency of the method in selecting the regularization parameter, and

increases its interest for the study of image performance by singular value analysis (SVA), particularly when dealing with large datasets, as it is the case of fDOT.

2) We proposed a procedure for the selection of optimum acquisition parameters for any specific fDOT experiment, based on the SVA of the fDOT forward matrix for different distributions of the acquisition parameters (mesh points, density of sources and density of detectors). We found that any increase in the number of sources and detectors at distances shorter than the average mean free path leads to slight improvements in image quality while increasing computational burden. Besides, regarding the number of mesh elements and their distribution, we showed for the first time in fDOT that large meshes can be reduced in the  $z$  direction without any loss in imaging performance but decreasing computational time and memory requirements.

3) Regarding  $l_1$ -based regularization techniques, we presented a novel iterative algorithm for image reconstruction in diffusive media, with application to a wide range of areas, particularly in-vivo imaging. The approach, named ART-SB, alternates the ART method with a denoising step based on the minimization of TV, solved by using the Split-Bregman (SB) method. SB has been implemented in a cost-efficient way to handle large datasets. ART-SB provides better results than conventional ART and it is computationally more efficient than previous TV-based reconstruction algorithms and most splitting methodologies. This methodology is particularly well suited for handling large datasets in fDOT biomedical imaging.

4) Regarding the Compressed Sensing (CS) techniques, we proposed a novel approach for fDOT reconstruction, named Split Bregman-Singular Value Analysis (SB-SVA) method, that takes advantage of the existing SB for CS (CS-SB) algorithm, restricting the solution given in each CS-SB iteration to a space where the singular values of forward matrix and the sparsity structure of the solution combine in beneficial manner. Thus, SB-SVA is, indirectly, well-conditioning the forward matrix (reducing its CN) while designing/learning the analysis operator and finding the solution. Dealing with ill-conditioned fDOT reconstruction problem, we demonstrate the existence of improvement as compared to CS-SB algorithm in terms of image performance, and number of voxels required.

# Publications

## Journal papers

- **J Chamorro-Servent**, JF Pérez-Juste Abascal, J Aguirre, S Arridge, T Correia, J Ripoll, M Desco, JJ Vaquero. "*Use of Split Bregman denoising for iterative reconstruction in fluorescence diffuse optical tomography*". J. Biomed. Opt., 18 (7): 076016, 2013.
- JF Pérez-Juste Abascal, J Aguirre, **J Chamorro-Servent**, M Schweiger, S Arridge, J Ripoll, JJ Vaquero, M Desco. "*Influence of absorption and scattering on the quantification of fluorescence diffuse optical tomography using normalized data*". J. Biomed. Opt., 17(3): 036013-1 -- 036013-9, 2012.
- J Pascau, JJ Vaquero, **J Chamorro-Servent**, A Rodríguez-Ruano, M Desco. "*A method for small-animal PET/CT alignment calibration*". Phys. Med. Biol., 57(12): N199-N207, 2012.
- JF Abascal, **J Chamorro-Servent**, J Aguirre, S Arridge, T Correia, J Ripoll, JJ Vaquero, M Desco. "*Fluorescence diffuse optical tomography using the split Bregman method*". Med Phys, 38(11): 6275-6284, 2011.
- **J Chamorro-Servent**, J Aguirre, J Ripoll, JJ Vaquero, M Desco. "*Feasibility of U-curve method to select the regularization parameter for fluorescence diffuse optical tomography in phantom and small animal studies*". Optics Express, 19(12): 11490-11506, 2011.
- T Correia, J Aguirre, A Sisniega, **J Chamorro-Servent**, J Abascal, JJ Vaquero, M Desco, V Kolehmainen, S Arridge. "*Split operator method for fluorescence diffuse optical tomography using anisotropic diffusion regularisation with prior anatomical information*". Biomed Opt Express, 2(9): 2632-2648, 2011.

## International conference record proceedings

- **J Chamorro-Servent**, J F.P-J Abascal, J Ripoll, J J Vaquero, M Desco. *Split Bregman-Singular Value Analysis approach to solve compressed sensing of fluorescence diffuse optical tomography*. Accepted in XIII Mediterranean Conference on Medical and Biological Engineering and Computing, September 2013.
- A Sisniega, J Abascal, M Abella, **J Chamorro-Servent**, M Desco, J J Vaquero. *Iterative dual-energy material decomposition for slow kVp switching: A compressed sensing approach*. Accepted in XIII Mediterranean Conference on Medical and Biological Engineering and Computing, September 2013.
- V García-Vázquez, L Cusso, **J Chamorro-Servent**, I Mirones, J García-Castro, L López, S Peña-Zalbieda, P Montesinos, C Chavarrias, J Pascau, M Desco. *Registration of Small-Animal SPECT/MRI Studies for Tracking Human Mesenchymal Stem Cells*. Accepted in XIII Mediterranean Conference on Medical and Biological Engineering and Computing, September 2013.
- **J Chamorro**, JF Abascal, J Aguirre, S Arridge, T Correia, J Ripoll, M Desco, JJ Vaquero. *"ART with Split Bregman Denoising: a Reconstruction Method for Fluorescence Diffuse Optical Tomography"*. Abstract book of the IEEE Nuclear Science Symposium and Medical Imaging Conference, 268, 2011.
- P Montesinos, JF Pérez-Juste Abascal, **J Chamorro**, C Chavarrias, M Benito, JJ Vaquero, M Desco. *"High-Resolution Dynamic Cardiac MRI on Small Animals using Reconstruction based on Split Bregman Methodology"*. 2011 IEEE Nuclear Science Symposium Conference Record, 3462-3464, 2011.
- **J Chamorro-Servent**, J Aguirre, J Ripoll, JJ Vaquero, M Desco. *"FDOT Reconstruction and Setting Optimization using Singular Value Analysis with Automatic Thresholding"*. 2009 IEEE Nuclear Science Symposium Conference Record, 2827-2829, 2009.
- **J Chamorro-Servent**, J Aguirre, J Ripoll, JJ Vaquero, M Desco. *"Maximizing the information content in acquired measurements of a parallel plate non-contact FDOT while minimizing the computational cost: singular value analysis"*. Abstract book of European Society for Molecular Imaging (ESMI), 161, 2009.
- **J Chamorro-Servent**, J Aguirre, J Ripoll, JJ Vaquero, M Desco. *"An automatic method to select a noise threshold in the singular-value domain for reconstruction of parallel plate*

*non-contact FDOT images*". Abstract book of European Society for Molecular Imaging (ESMI), 162, 2009.

- V García-Vázquez, **J Chamorro-Servent**, A Rodríguez-Ruano, M Benito, J Tejedor Fraile, FJ Carrillo Salinas, L Montoliu, M Desco. *"Mouse eyeball's axial length measurement with MRI"*. Abstract book of European Society for Molecular Imaging (ESMI), 164, 2009.

- A Rodríguez-Ruano, J Pascau, **J Chamorro**, A Sisniega, V García-Vázquez, A Udías, JJ Vaquero, M Desco. *"PET/CT Alignment for Small Animal Scanners based on Capillary Detection"*. 2008 IEEE Nuclear Science Symposium Conference Record, 3832-3835, 2008.

- V García-Vázquez, S Reig, J Janssen, J Pascau, A Rodríguez-Ruano, A Udías, **J Chamorro**, JJ Vaquero, M Desco. *"Use of IBASPM Atlas-based Automatic Segmentation Toolbox in Pathological Brains: Effect of Template Selection"*. 2008 IEEE Nuclear Science Symposium Conference Record, 4270-4272, 2008.

### National conference record proceedings

- **J Chamorro-Servent**, JF P-J Abascal, J Aguirre, J Vaquero, M Desco. *"Reconstrucción de Tomografía Óptica Difusiva por Fluorescencia usando Compressed Sensing Split Bregman"*. Libro de actas del XXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB), 599-601, 2011.

- P Montesinos, JF Pérez-Juste Abascal, **J Chamorro**, C Chavarrías, M Benito, JJ Vaquero, M Desco. *"Uso del método de Split Bregman para la resolución del problema de compressed sensing en imagen de resonancia magnética dinámica cardíaca para pequeño animal"*. Libro de actas del XXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB), 453-456, 2011.

- J Aguirre, **J Chamorro-Servent**, J Abascal, J Ripoll, M Desco, JJ Vaquero. *"Imaging features of an FDOT system with optimized ART parameters"*. Proceedings del XXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB), 2010.

- **J Chamorro-Servent**, J Abascal, J Aguirre, SR Arridge, J Ripoll, JJ Vaquero, M Desco. *"Optimización del diseño experimental FDOT de órganos de animal pequeño a través del análisis de valores singulares"*. Proceedings del XXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB), 2010.

- **J Chamorro-Servent**, J Aguirre, J Ripoll, JJ Vaquero, M Desco. *"Optimización del diseño experimental y reconstrucción FDOT a través del análisis de valores singulares"*. Actas del XXVII Congreso Anual de la Sociedad Española de Ingeniería Biomédica, 173-176, 2009.
- V García-Vázquez, M Benito, **J Chamorro-Servent**, A Rodríguez-Ruano, J Tejedor Fraile, FJ Carrillo Salinas, L Montoliu, M Desco. *"Medida de la longitud axial del globo ocular en ratones utilizando imágenes de resonancia magnética"*. Libro de Resúmenes del CASEIB 2009, 57, 2009.
- **J Chamorro-Servent**, A Rodríguez-Ruano, J Pascau, A Udías, A Sisniega, V García-Vázquez, JJ Vaquero, M Desco. *"Alineamiento de sistemas PET/CT para pequeños animales basado en detección de capilares"*. Libro de Actas del CASEIB 2008, 128-130, 2008.



# References

- Abascal, J.-F. P. J., S. R. Arridge, R. H. Bayford, and D. S. Holder. 2008. "Comparison of methods for optimal choice of the regularization parameter for linear electrical impedance tomography of brain function." *Physiological Measurement* 29(11): 1319.
- Abascal, J. F. P.-J., J. Chamorro-Servent, J. Aguirre, S. Arridge, T. Correia, J. Ripoll, J. J. Vaquero, and M. Desco. 2011. "Fluorescence diffuse optical tomography using the split Bregman method." *Medical Physics* 38(11): 6275-6285.
- Abascal, J. F. P.-J., J. Aguirre, J. Chamorro-Servent, M. Schweiger, S. Arridge, J. Ripoll, J. J. Vaquero, and M. Desco. 2012. "Influence of absorption and scattering on the quantification of fluorescence diffuse optical tomography using normalized data." *Journal of Biomedical Optics* 17(3): 0360131-39.
- Aguirre, J. 2012. "Estudios sobre la tomografía óptica difusiva de fluorescencia." Madrid: University Carlos III.
- Arridge, S. R. and J. C. Schotland. 2009. "Optical tomography: forward and inverse problems." *Inverse problems* 25(12):123010.
- Baritau, J.-C., K. Hassler, and M. Unser. 2010. "An Efficient Numerical Method for General Regularization in Fluorescence Molecular Tomography." *IEEE Trans Med Imaging* 29(4): 1075.
- Baritau, J.-C. and M. Unser. 2011. "A primal-dual reconstruction algorithm for fluorescence and bioluminescence tomography." *Biomedical Imaging: From Nano to Macro, IEEE International Symposium on*. Chicago, Illinois, U.S.A., pp. 960-63.
- Behrooz, A., H.-M. Zhou, A. A. Eftekhari, and A. Adibi. 2012. "Total variation regularization for 3D reconstruction in fluorescence tomography: experimental phantom studies." *Applied Optics* 51(34): 8216-27.
- Boas, D. A. 1996. "Diffuse photon probes of structural and dynamical properties of turbid media: theory and biomedical applications." Pennsylvania: University of Pennsylvania.
- Born, M. and E. Wolf. 1999. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge, United Kingdom: Cambridge University Press.
- Busby, H. and D. Trujillo. 1997. "Optimal regularization of an inverse dynamics problem." *Computers & Structures* 63(2): 243-48.
- Candes, E. J., Y. C. Eldar, D. Needell, and P. Randall. 2011. "Compressed sensing with coherent and redundant dictionaries." *Applied and Computational Harmonic Analysis* 31(1): 59-73.
- Candes, E. J. and J. K. Romberg. 2005. "Signal recovery from random projections." *Electronic Imaging, International Society for Optics and Photonics. SPIE*, San José, California, U.S.A., pp. 76-86.
- Cao, X., X. Wang, B. Zhang, F. Lui, J. Luo, and J. Bai. 2013. "Accelerated image reconstruction in fluorescence molecular tomography using dimension reduction." *Biomedical Optics Express* 4(1): 1-14.
- Corlu, A. 2007. "Three-dimensional in vivo fluorescence diffuse optical tomography of breast cancer in humans." *Optical Society of America* 15(11): 6696-716.
- Correia, T., J. Aguirre, A. Sisniega, J. Chamorro-Servent, J. Abascal, J. J. Vaquero, M. Desco, V. Kolehmainen, and S. Arridge. 2011. "Split operator method for fluorescence diffuse optical tomography using anisotropic diffusion regularisation with prior anatomical information." *Biomedical optics express* 2(9): 2632.

- Correia, T., A. Gibson, M. Schweiger, and J. Hebden. 2009. "Selection of regularization parameter for optical topography." *Journal of Biomedical Optics* 14(3): 034044.
- Cubeddu, R., A. Pifferi, P. Taroni, A. Torricelli, and G. Valentini. 1997. "A solid tissue phantom for photon migration studies." *Phys Med Biol* 42(10): 1971.
- Culver, J., R. Choe, M. Holboke, L. Zubkov, T. Durduran, A. Slemple, V. Ntziachristos, B. Chance, and A. Yodh. 2003. "Three-dimensional diffuse optical tomography in the parallel plane transmission geometry: evaluation of a hybrid frequency domain/continuous wave clinical system for breast imaging." *Medical Physics* 30(2): 235-278.
- Culver, J., V. Ntziachristos, M. Holboke, and A. Yodh. 2001. "Optimization of optode arrangements for diffuse optical tomography: A singular-value analysis." *Optics letters* 26(10): 701-703.
- Chamorro-Servent, J., J. Abascal, J. Aguirre, S. Arridge, J. Ripoll, J. Vaquero, and M. Desco. 2010. "Optimización del diseño experimental FDOT de órganos de animal pequeño a través del análisis de valores singulares." *Proceedings del XXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB)*, Madrid, Spain.
- Chamorro, J., J. Aguirre, J. Ripoll, J. Vaquero, and M. Desco. 2009. "FDOT setting optimization and reconstruction using singular value analysis with automatic thresholding." *Nuclear Science Symposium Conference Record (NSS/MIC)*, IEEE. Orlando, Florida, U.S.A. pp. 2827 - 29.
- Chaudhari, A. J., S. Ahn, R. Levenson, R. D. Badawi, S. R. Cherry, and R. M. Leahy. 2009. "Excitation spectroscopy in multispectral optical fluorescence tomography: methodology, feasibility and computer simulation studies." *Phys Med Biol* 54(15): 4687.
- Donoho, D. and J. Tanner. 2009. "Counting faces of randomly projected polytopes when the projection radically lowers dimension." *Journal of the American Mathematical Society* 22(1): 1-53.
- Donoho, D. L. and M. Elad. 2003. "Optimally sparse representation in general (nonorthogonal) dictionaries via  $\ell_1$  minimization." *Proceedings of the National Academy of Sciences - P.N.A.S., U.S.A.*, 100(5): 2197-202.
- Douiri, A., M. Schweiger, J. Riley, and S. Arridge. 2007. "Anisotropic diffusion regularization methods for diffuse optical tomography using edge prior information." *Measurement Science and Technology* 18(1): 87.
- Ducros, N., L. Hervé, A. Da Silva, J.-M. Dinten, and F. Peyrin. 2009. "A comprehensive study of the use of temporal moments in time-resolved diffuse optical tomography: part I. Theoretical material." *Phys Med Biol* 54(23): 7089.
- Ducros, N., C. d'Andrea, A. Bassi, G. Valentini, and S. Arridge. 2011. "A virtual source pattern method for fluorescence tomography with structured light". *Phys Med Biol* 57(18):3811-32.
- Dutta, J., S. Ahn, C. Li, S. R. Cherry, and R. M. Leahy. 2012. "Joint L1 and total variation regularization for fluorescence molecular tomography." *Phys Med Biol* 57(6): 1459.
- Egger, H., M. Freiberger, and M. Schlottbom. 2010. "On forward and inverse models in fluorescence diffuse optical tomography." *Inverse Problems and Imaging* 4(3): 411-27.
- Elad, M. 2007. "Optimized Projections for Compressed Sensing." *Signal Processing, IEEE Transactions on* 55(12): 5695-702.
- Elad, M., P. Milanfar, and R. Rubinstein. 2007. "Analysis versus synthesis in signal priors." *Inverse problems* 23(3): 947.
- Figueiredo, M. A. and Nowak, R. D. 2001. "Wavelet-based image estimation: an empirical Bayes approach using Jeffrey's noninformative prior." *Image Processing, IEEE Transactions on* 10(9): 1322-1331.
- Firbank, M. and D. Delpy. 1993. "A design for a stable and reproducible phantom for use in near infra-red imaging and spectroscopy." *Phys Med Biol* 38(6): 847.
- Freiberger, M., C. Clason, and H. Scharfetter. 2010. "Total variation regularization for nonlinear fluorescence tomography with an augmented Lagrangian splitting approach." *Applied Optics* 49(19): 3741-7.
- Garofalakis, A., G. Zacharakis, H. Meyer, E. N. Economou, C. Mamalaki, J. Papamatheakis, D. Kioussis, V. Ntziachristos, and J. Ripoll. 2007. "Three-dimensional in vivo imaging of green

- fluorescent protein-expressing T cells in mice with noncontact fluorescence molecular tomography.” *Molecular imaging* 6(2): 96-107.
- Giryes, R., S. Nam, R. Gribonval, and M. E. Davies. 2011. “Iterative cospase projection algorithms for the recovery of cospase vectors.” *The 19th European Signal Processing Conference (EUSIPCO)*. Barcelona, Spain.
  - Goldstein, T. and S. Osher. 2009. “The split Bregman method for L1-regularized problems.” *SIAM Journal on Imaging Sciences* 2(2): 323-43.
  - Golub, G. H. and U. Von Matt. 1997. “Tikhonov Regularization for Large Scale Problems.” *Workshop on Scientific Computing, Springer*. Berlin, Germany. p.p. 3-26.
  - Golub, G. H. and C. F. Van Loan. 1996. *Matrix computations*: Johns Hopkins University Press.
  - Grant, M., S. Boyd, and Y. Ye. 2008. “CVX:Matlab Software for Disciplined Convex Programming.”
  - Graves, E. E., J. P. Culver, J. Ripoll, R. Weissleder, and V. Ntziachristos. 2004. “Singular-value analysis and optimization of experimental parameters in fluorescence molecular tomography.” *JOSA A* 21(2): 231-41.
  - Graves, E. E., J. Ripoll, R. Weissleder, and V. Ntziachristos. 2003. “A submillimeter resolution fluorescence molecular imaging system for small animal imaging.” *Medical Physics* 30(5): 901-12.
  - Gribonval, R. and M. Nielsen. 2003. “Sparse representations in unions of bases.” *Information Theory, IEEE Transactions on* 49(12): 3320-25.
  - Hadamard, J. 1902. “Sur les problèmes aux dérivées partielles et leur signification physique.” *Princeton University Bulletin*. pp. 49-52. Princeton (New Jersey): McGraw-Hill.
  - Hanke, M. 1996. “Limitations of the L-curve method in ill-posed problems.” *BIT Numerical Mathematics* 36(2): 287-301.
  - Hanke, M. and P. C. Hansen. 1993. “Regularization methods for large-scale problems.” *Surv. Math. Ind* 3(4): 253-315.
  - Hansen, P. C. 1987. “The truncated SVD as a method for regularization.” *BIT Numerical Mathematics* 27(4): 534-53.
  - Hansen, P. C. 1990. “The discrete Picard condition for discrete ill-posed problems.” *BIT Numerical Mathematics* 30(4): 658-72.
  - Hansen, P. C. 1992. “Analysis of discrete ill-posed problems by means of the L-curve.” *SIAM Review* 34(4): 561-80.
  - Hansen, P. C. 2007. “Regularization tools version 4.0 for Matlab 7.3.” *Numerical Algorithms* 46(2): 189-94.
  - Hansen, P. C. 2010. *Discrete inverse problems: insight and algorithms*. Philadelphia, USA: Society for Industrial and Applied Mathematics.
  - Hansen, P. C. and D. P. O’Leary. 1993. “The use of the L-curve in the regularization of discrete ill-posed problems.” *SIAM Journal on Scientific Computing* 14(6): 1487-503.
  - Hawe, S., M. Kleinstuber, K. Diepold. 2013. “Analysis Operator Learning and its Application to Image Reconstruction.” *arXiv preprint arXiv:1204.5309*.
  - Herman, G. T. and L. B. Meyer. 1993. “Algebraic reconstruction techniques can be made computationally efficient positron emission tomography application.” *Medical Imaging, IEEE Transactions on* 12(3): 600-09.
  - Hyde, D., R. de Kleine, S. A. MacLaurin, E. Miller, D. H. Brooks, T. Krucker, and V. Ntziachristos. 2009. “Hybrid FMT-CT imaging of amyloid-beta plaques in a murine Alzheimer's disease model.” *Neuroimage* 44(4): 1304-11.
  - Intes, X., V. Ntziachristos, J. P. Culver, A. Yodh, and B. Chance. 2002. “Projection access order in algebraic reconstruction technique for diffuse optical tomography.” *Phys Med Biol* 47(1): N1.
  - Intes, X., J. Ripoll, Y. Chen, S. Nioka, A. G. Yodh, and B. Chance. 2003. “In vivo continuous-wave optical breast imaging enhanced with Indocyanine Green.” *Medical Physics* 30(6): 1039-47.
  - Ishimaru, A. 1978. *Wave propagation and scattering in random media*. New York: Academic press
  - Jagannath, R. P. K. and P. K. Yalavarthy. 2012. “Minimal residual method provides optimal regularization parameter for diffuse optical tomography.” *Journal of Biomedical Optics* 17(10): 106015.

- Jensen, T. L., J. H. Jørgensen, P. C. Hansen, and S. H. Jensen. 2012. "Implementation of an optimal first-order method for strongly convex total variation regularization." *BIT Numerical Mathematics* 52(2): 329-56.
- Jin, A., B. Yazici, A. Ale, and V. Ntziachristos. 2012. "Preconditioning of the fluorescence diffuse optical tomography sensing matrix based on compressive sensing." *Optics letters* 37(20): 4326-28.
- Jobsis, F. F. 1977. "Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters." *Science* 198(4323): 1264-67.
- Johnston, S., G. Johnson, and C. Badea. 2010. "GPU-based iterative reconstruction with total variation minimization for micro-CT." *Proceedings of SPIE, Medical Imaging: Physics of medical Imaging*. San Diego, California, USA. pp. 762238.
- Kak, A. C. 1979. "Computerized tomography with X-ray, emission, and ultrasound sources." *Proceedings of the IEEE* 67(9): 1245-72.
- Keller, H. B. 1976. *Numerical solution of two point boundary value problems*: SIAM.
- Krawczyk-StańDo, D. and M. Rudnicki. 2007. "Regularization parameter selection in discrete ill-posed problems—the use of the U-curve." *International Journal of Applied Mathematics and Computer Science* 17(2): 157-64.
- Krawczyk-StańDo, D. and M. Rudnicki. 2008. "The use of L-curve and U-curve in inverse electromagnetic modelling." *Intelligent Computer Techniques in Applied Electromagnetics, Springer*. - Heidelberg, Berlin, Germany. pp. 73-82.
- Krawczyk-StańDo, D., M. Rudnicki, and J. StańDo. 2008. "Radial neural network learning using U-curve approach." *Polish Journal of Environmental Studies* 17.
- Lasser, T. and V. Ntziachristos. 2007. "Optimization of 360 projection fluorescence molecular tomography." *Med Image Anal* 11(4): 389-99.
- Li, M., X. Cao, F. Liu, B. Zhang, J. Luo, and J. Bai. 2012. "Reconstruction of fluorescence molecular tomography using a neighborhood regularization." *Biomedical Engineering, IEEE Transactions on* 59(7): 1799-803.
- Lorenzo, J. R. 2012. *Principles of Diffuse Light Propagation: Light Propagation in Tissues with Applications in Biology and Medicine*: World Scientific.
- Mallat, S. G. and Z. Zhang. 1993. "Matching pursuits with time-frequency dictionaries." *Signal Processing, IEEE Transactions on* 41(12): 3397-415.
- Markel, V. A. and J. C. Scotthland. 2002. "The inverse problem in optical diffusion tomography. II. Inversion with boundary conditions." *Optical Society of America A* 19(3): 558-66.
- Martin, A., J. Aguirre, A. Sarasa-Renedo, D. Tsoukatou, A. Garofalakis, H. Meyer, C. Mamalaki, J. Ripoll, and A. M. Planas. 2008. "Imaging changes in lymphoid organs in vivo after brain ischemia with three-dimensional fluorescence molecular tomography in transgenic mice expressing green fluorescent protein in T lymphocytes." *Mol Imaging* 7(4): 157-67.
- Mohajerani, P. and V. Ntziachristos. 2013. "Compression of Born ratio for fluorescence molecular tomography/x-ray computed tomography hybrid imaging: methodology and in vivo validation." *Optics letters* 38(13): 2324-6.
- Nam, S., M. E. Davies, M. Elad, and R. Gribonval. 2013. "The cosparsity analysis model and algorithms." *Applied and Computational Harmonic Analysis* 34(1): 30-56.
- Ntziachristos, V., C. Bremer, E. E. Graves, J. Ripoll, and R. Weissleder. 2002. "In vivo tomographic imaging of near-infrared fluorescent probes." *Mol. Imaging* 1(2): 82-88.
- Ntziachristos, V., A. Leroy-Willig, and B. Tavittian. 2007. *Textbook of in vivo Imaging in Vertebrates*: Wiley-Interscience.
- Ntziachristos, V., E. A. Schellenberger, J. Ripoll, D. Yessayan, E. Graves, A. Bogdanov, Jr., L. Josephson, and R. Weissleder. 2004. "Visualization of antitumor treatment by means of fluorescence molecular tomography with an annexin V-Cy5.5 conjugate." *Proceedings of The National Academy Sciences - PNAS. U S A*. 101(33): 12294-9.
- Ntziachristos, V. and R. Weissleder. 2001. "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized Born approximation." *Optics letters* 26(12): 893-5.

- Ntziachristos, V., A. Yodh, M. Schnall, and B. Chance. 2000. "Concurrent MRI and diffuse optical tomography of breast after indocyanine green enhancement." *Proceedings of The National Academy of Sciences - PNAS, U S A* 97 (6): 2767-72.
- O'Leary, M. A. 1996. "Imaging with diffuse photon density waves." University of Pennsylvania.
- Pan, Y., R. Whitaker, A. Cheryauka, and D. Ferguson. 2010. "TV-regularized iterative image reconstruction on a mobile C-ARM CT." *SPIE Medical Imaging International Society for Optics and Photonics*. San Diego, California, USA. pp. 76222L
- Patterson, M. S., B. Chance, and W. C. 1989. "Time resolved reflectance and transmittance for the non-invasive measurement of tissue optical properties." *Applied Optics* 28(12): 2331-36.
- Pogue, B. W., T. O. McBride, J. Prewitt, U. L. Österberg, and K. D. Paulsen. 1999. "Spatially variant regularization improves diffuse optical tomography." *Applied Optics* 38(13): 2950-61.
- Qiangqiang, Y., Z. Liangpei, S. Huanfeng, and L. Pingxiang. 2010. "Adaptive multiple-frame image super-resolution based on U-curve." *Image Processing, IEEE Transactions on* 19(12): 3157-70.
- Ripoll, J., V. Ntziachristos, C. Cannet, A. L. Babin, R. Kneuer, H.-U. Gremlich, and N. Beckmann. 2008. "Investigating pharmacology in vivo using magnetic resonance and optical imaging." *Drugs in R & D* 9(5): 277-306.
- Ripoll, J. 2010. "Hybrid Fourier-real space method for diffuse optical tomography." *Optics letters*, 35(5): 688-690.
- Rubinstein, R., T. Peleg, and M. Elad. 2013. "Analysis K-SVD: A Dictionary-Learning Algorithm for the Analysis Sparse Model." *Signal processing, IEEE Transaction on* 61 (3), 661-677.
- Rudin, L. I., S. Osher, and E. Fatemi. 1992. "Nonlinear total variation based noise removal algorithms." *Physica D: Nonlinear Phenomena* 60(1): 259-68.
- Sawatzky, A., C. Brune, F. Wubbeling, T. Kosters, K. Schafers, and M. Burger. 2008. "Accurate EM-TV algorithm in PET with low SNR." *Nuclear Science Symposium Conference Record.. NSS'08. IEEE*. Dresden, Germany. pp. 5133-37.
- Schulz, R. B., J. Ripoll, and V. Ntziachristos. 2004. "Experimental Fluorescence Tomography of Tissues With Noncontact Measurements." *IEEE Trans Med Imaging* 23(4): 492-500.
- Schweiger, M. 1994. "Application of the finite element method in infrared image reconstruction of scattering media." London, UK: University College London.
- Schweiger, M., S. R. Arridge, M. Hiraoka, and D. T. Delpy. 1995. "The finite element method for the propagation of light in scattering media: boundary and source conditions." *Med Phys* 22(11): 1779-92.
- Serdaroglu, A., B. Yazici, and V. Ntziachristos. 2006. "Fluorescence molecular tomography based on a priori information." *Biomedical Topical Meeting, Optical Society of America*. Fort Lauderdale, Florida. pp. SH46.
- Setzer, S. 2009. "Split Bregman Algorithm, Douglas-Rachford Splitting and Frame Shrinkage." *Proceedings of the Second International Conference on Scale Space and Variational Methods in Computer Vision, Springer, Verlag, Norway*. pp. 464-76.
- Süzen, M., A. Giannoula, and T. Durduran. 2010. "Compressed sensing in diffuse optical tomography." *Optics Express* 18(23): 23676-90.
- Theodoridis, S., Y. Kopsinis, and K. Slavakis. 2012. "Sparsity-Aware Learning and Compressed Sensing: An Overview." *arXiv preprint arXiv:1211.5231*.
- van den Berg, E. and M. P. Friedlander. "SPGL1: A solver for large-scale sparse reconstruction, June 2007." Available in <http://www.cs.ubc.ca/labs/scl/spgl1>.
- van den Doel, K., U. Ascher, and E. Haber. 2012. "The lost honour of l2-based regularization." Submitted.
- Vogel, C. R. 1996. "Non-convergence of the L-curve regularization parameter selection method." *Inverse problems* 12(4): 535.
- Vogel, C. R. 2002. *Computational methods for inverse problems*: SIAM.
- Xu, Z., Y. Jin, and J. Bai. 2009. "Determining the regularization parameter: a hybrid reconstruction technique in fluorescence molecular tomography." *Asia Communications and Photonics, International Society for Optics and Photonics*. Shangai, China. pp. 763407.

- Yaghoobi, M., S. Nam, R. Gribonval, and M. E. Davies. 2011. "Analysis operator learning for overcomplete cospase representations." *European Signal Processing Conference (EUSIPCO'11)*. Barcelona. Spain.
- Yin, W., S. Osher, D. Goldfarb, and J. Darbon. 2008. "Bregman iterative algorithms for  $l_1$ -minimization with applications to compressed sensing." *SIAM Journal on Imaging Sciences* 1(1): 143-68.
- Zacharopoulos, A. D., M. Schweiger, V. Kolehmainen, and S. Arridge. 2009. "3D shape based reconstruction of experimental data in diffuse optical tomography." *Optics Express* 17(21): 18940-56.