

University of Brasilia at Gama – UnB/FGA  
Biomedical Engineering Graduate Program (PPGEB)

**COMPRESSIVE SENSING METHODS  
FOR DIFFUSE OPTICAL TOMOGRAPHY**

**LUCAS SILVA LOPES**

Advisor: CRISTIANO JACQUES MIOSSO, PhD



**UNIVERSITY OF BRASILIA AT GAMA – UNB/FGA  
GRADUATE PROGRAM**

**COMPRESSIVE SENSING METHODS  
FOR DIFFUSE OPTICAL TOMOGRAPHY**

**LUCAS SILVA LOPES**

**ADVISOR: CRISTIANO JACQUES MIOSSO**

**MASTERS DEGREE PROGRAM IN  
BIOMEDICAL ENGINEERING**

**PUBLICATION: 097A/2018**

**BRASILIA/DF, AUGUST 2018**

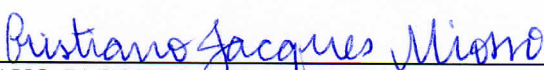
**UNIVERSIDADE DE BRASÍLIA  
FACULDADE DO GAMA  
ENGENHARIA BIOMÉDICA**


**"COMPRESSIVE SENSING METHODS FOR DIFFUSE OPTICAL  
TOMOGRAPHY"**

**LUCAS SILVA LOPES**

**DISSERTAÇÃO DE MESTRADO SUBMETIDA À FACULDADE UNB GAMA DA  
UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA  
A OBTENÇÃO DO TÍTULO DE MESTRE EM ENGENHARIA BIOMÉDICA.**

**APROVADA POR:**

  
\_\_\_\_\_  
**PROF. DR. CRISTIANO JACQUES MIOSSO RODRIGUES MENDES; FGA / UNB  
(ORIENTADOR)**

  
\_\_\_\_\_  
**PROF. DR. RICARDO FREITAS VON BORRIES; UNIVERSITY OF TEXAS AT EL PASO  
(EXAMINADOR EXTERNO)**

  
\_\_\_\_\_  
**PROF. DR. VINICIUS DE CARVALHO RISPOLI; FGA / UNB  
(EXAMINADOR EXTERNO)**

**BRASÍLIA, 14 DE AGOSTO DE 2018**

**FICHA CATALOGRÁFICA**

LOPES, LUCAS SILVA

Compressive Sensing Methods for Diffuse Optical Tomography

[Distrito Federal], 2018.

84p., 210 × 297 mm (FGA/UnB Gama, Mestrado em Engenharia Biomédica, 2018).

Dissertação de Mestrado em Engenharia Biomédica, Faculdade UnB Gama, Programa de Pós-Graduação em Engenharia Biomédica.

1. Diffuse Optical Tomography

2. Compressive Sensing

3. Medical Imaging

4. Near Infrared

I. FGA UnB/UnB.

II. Título (série)

**REFERÊNCIA**

LOPES, LUCAS SILVA (2018). Compressive Sensing Methods for Diffuse Optical Tomography. Dissertação de mestrado em engenharia biomédica, Publicação 097A/2018, Programa de Pós-Graduação, Faculdade UnB Gama, Universidade de Brasília, Brasília, DF, 84p.

**CESSÃO DE DIREITOS**

AUTOR: Lucas Silva Lopes

TÍTULO: Compressive Sensing Methods for Diffuse Optical Tomography

GRAU: Mestre

ANO: 2018

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor.

---

lucaslopes@ieee.org

Brasília, DF – Brasil

## RESUMO ESTENDIDO

### Introdução:

*Diffuse Optical Tomography* (DOT) [5] é uma promissora nova técnica de imagiologia funcional. Seu princípio básico é similar ao da oximetria de pulso, em que luz no infravermelho próximo é utilizada para se estimar a saturação de oxigênio do sangue. Diferentemente da oximetria de pulso, no entanto, a DOT é capaz de reconstruir a saturação de oxigênio em diversos pontos dentro do corpo, informação que pode ser reunida e representada na forma de imagens.

Por ser baseada em luz no infravermelho próximo, a DOT não emite radiação ionizante, ao contrário da tomografia por emissão de pósitrons, e não é perigosa de ser utilizada em pacientes com eletrônicos implantados, ao contrário da ressonância magnética funcional. Além disto, ela é em geral relativamente mais barata e portátil do que estas outras técnicas. Contudo, a DOT ainda não foi aceita como uma técnica de imagem robusta e confiável [1].

Como um dos principais problemas envolvidos com DOT é a qualidade das imagens reconstruídas, esta técnica pode ser beneficiada pelo uso de *compressive sensing* (CS) [9]. Nos casos em que um sistema de aquisição satisfaz determinados critérios, o CS permite reconstruções de sinais esparsos com maior qualidade, usando menos medidas do que técnicas tradicionais. Nos casos em que as imagens não são esparsas no seu domínio original, uma das formas de reconstruí-las usando CS é pelo método da pré-filtragem [7, 8]. Neste método, o processo de reconstrução da imagem é decomposto em dois estágios. No primeiro, versões filtradas da imagem são reconstruídas usando CS. No segundo, a imagem é composta a partir de suas versões filtradas.

Apesar de já ser conhecida uma formulação matemática para o problema de recuperação de imagem em DOT que satisfaz os critérios do CS [2, 3, 6], esta formulação a princípio não permite o uso do método da pré-filtragem. Além disto, ela possui limitações no que diz respeito ao modelo utilizado. Estas características justificam o desenvolvimento de outras formulações matemáticas que não só satisfaçam os critérios de CS, mas que também permitam a utilização do método da pré-filtragem, ou que façam considerações menos restritivas com respeito ao meio imageado.

### Métodos e Resultados:

Além do desenvolvimento de uma nova formulação matemática para o problema de imagem em DOT, que satisfizesse os critérios de CS e permitisse o uso do método da pré-filtragem, também foram desenvolvidos os algoritmos necessários para a reconstrução de imagens, e experimentos numéricos para avaliar estes algoritmos.

A formulação proposta para o problema de DOT melhora a formulação anterior no

sentido de que ela considera os efeitos de borda. Esta formulação é baseada na representação do espectro angular [4], e permite o uso do método da pré-filtragem de uma maneira similar a como ela é utilizada em ressonância magnética.

O método de CS proposto permitiu a reconstrução de imagens esparsas no seu domínio original com relações sinal-erro acima 20 dB, e em um caso particular acima de 50 dB. O método da pré-filtragem, por sua vez, permitiu a reconstrução de imagens não esparsas no seu domínio original com relações sinal-erro acima de 8 dB, e em um caso particular acima de 12 dB.

Além disto, os experimentos numéricos mostraram que as técnicas propostas em geral melhoraram, em termos de relação sinal erro, as reconstruções com relação a regularização de Tikhonov, a técnica mais utilizada em DOT em situações similares.

## Conclusão:

As técnicas propostas assumem que o meio é homogêneo semi-infinito e que ele contém apenas perturbações no coeficiente de absorção. Para usar estas técnicas com êxito na reconstrução de meios mais complexos, pode ser necessário adapta-las para que elas se baseiem em modelos mais apropriados. As pesquisa mais recentes têm utilizado o método dos elementos finitos para resolver o problema direto de DOT, logo idealmente as técnicas propostas deveriam ser adaptadas para se basear em modelos obtidos pelo método dos elementos finitos.

## Referências

- [1] R. J. Cooper. Bioimaging: Watching the brain at work. *Nature Photonics*, 8(6):425–426, 2014.
- [2] B. P. V. Dileep, T. Das, and P. K. Dutta. Greedy algorithms for diffuse optical tomography reconstruction. *Optics Communications*, 410, 2018.
- [3] B. P. V. Dileep, T. Das, and P. K. Dutta. Modified CS-MUSIC for diffuse optical tomography using joint sparsity. *Optik*, 158:1478–1490, 2018.
- [4] T. Durduran, J. P. Culver, M. J. Holboke, X. D. Li, L. Zubkov, B. Chance, D. N. Pattanayak, and A. G. Yodh. Algorithms for 3D localization and imaging using near-field diffraction tomography with diffuse light. *OPTICS EXPRESS*, 4(8):247–262, 1999.
- [5] Y. Hoshi and Y. Yamada. Overview of diffuse optical tomography and its clinical applications. *Journal of Biomedical Optics*, 21(9):091312, 2016.

- [6] O. Lee, J. M. Kim, Y. Bresler, and J. C. Ye. Compressive diffuse optical tomography: Noniterative exact reconstruction using joint sparsity. *IEEE Transactions on Medical Imaging*, 30(5):1129–1142, 2011.
- [7] C. J. Miosso, R. von Borries, and J. H. Pierluissi. Compressive sensing method for improved reconstruction of gradient-sparse magnetic resonance images. In *Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, pages 799–806, 2009.
- [8] C. J. Miosso, R. von Borries, and J. H. Pierluissi. Compressive Sensing With Prior Information: Requirements and Probabilities of Reconstruction in  $l_1$ -Minimization. *IEEE Transactions on Signal Processing*, 61(9):2150–2164, 2013.
- [9] M. Rani, S. B. Dhok, and R. B. Deshmukh. *A Systematic Review of Compressive Sensing: Concepts, Implementations and Applications*, 2018.

## ABSTRACT

Diffuse optical tomography (DOT) is a promising new functional imaging modality. Its basic principle is similar to pulse oximetry, in which the blood oxygen saturation is estimated by measuring the absorption of a thin part of the body with respect to two different wavelengths in the near infrared (NIR) range. However, DOT is capable of recovering images of tissue oxygen saturation. Even though it is expected to extend the reach of functional imaging to several patients with specific needs, DOT is still under development, and has been for the last 30 years. A lot of research is being made in both the areas of DOT hardware and software to overcome the issues related to this technique.

In the general context of medical imaging, several imaging modalities have benefited from the theoretical developments regarding the recovery of sparse signals using compressive sensing (CS). In the cases that the images are sparse in some domain, CS is able to increase the quality of the reconstructed images, while decreasing the number of measurements required, in comparison to the other more traditional reconstruction techniques. In general, medical images are not naturally sparse, so the images are frequently recovered in a transformed domain and then converted back to the original domain. This is commonly carried out by a sparsifying transform, but not any transform can be used, since CS requires that the resulting sensing matrix meet specific requirements. In the context of magnetic resonance imaging (MRI) the prefiltering method has been used as an alternative to an explicit sparsifying transform with particular benefits. One of the advantages of the prefiltering method is that it uses a set of filters that act on the measurements and not on the sensing matrix. As a consequence, there is no restriction on the set of filters that can be used in the prefiltering method.

Although there is already a formulation of the DOT problem proven to satisfy the requirements of CS under certain circumstances, this formulation does not immediately allow the use of the prefiltering method in DOT. Moreover, there are limitations in the assumed model that justify the search for other formulations. Therefore, we proposed the development of another mathematical formulation for the DOT imaging problem that not only satisfies the requirements for CS, but also allows the use of the prefiltering method in DOT. We also proposed the development of the associated algorithms necessary for image reconstruction, and the evaluation of the resulting techniques in simulated settings.

The proposed formulation of the DOT imaging problem improves on the previous formulation in that it accounts for the medium boundaries. This formulation is based on the angular spectrum representation of the diffuse photon density waves, and allows the use of the prefiltering method in DOT in a similar manner to how it is used in MRI.

The proposed CS technique allowed the reconstruction of several sparse images in



their original domains with signal to error ratios (SERs) above 20 dB, and in a specific case above 50 dB. The prefiltering method allowed the reconstruction of several images with SERs above 8 dB, and in a specific case above 12 dB, in cases that the images were not sparse in their original domain. Finally, the numerical experiments also showed that the proposed techniques generally improved the reconstructions over Tikhonov regularization, the technique most commonly used in DOT in similar situations, in terms of SER.

The proposed methods assume that the medium is an homogeneous semi-infinite medium containing only perturbations in absorption coefficient. In order to use the proposed techniques effectively to image more complex mediums, it may be necessary to adapt the techniques to rely on more complex models first. The most recent research in DOT has used the finite elements method (FEM) to solve the DOT forward problem, so ideally the techniques should be adapted to rely on models based on the FEM.

# Contents

<b>Resumo Estendido</b>	<b>v</b>
<b>Abstract</b>	<b>viii</b>
<b>List of Tables</b>	<b>xiii</b>
<b>List of Figures</b>	<b>xiv</b>
<b>Abbreviations</b>	<b>xviii</b>
<b>Notation</b>	<b>xix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 History of DOT . . . . .	5
1.2 Image Reconstruction in DOT . . . . .	6
1.3 Research Objectives . . . . .	8
1.3.1 General Objective . . . . .	8
1.3.2 Specific Objectives . . . . .	9
1.4 Contributions . . . . .	9
1.5 Document Structure . . . . .	11
<b>2 Theoretical Background</b>	<b>13</b>
2.1 Diffuse Optical Tomography . . . . .	13
2.1.1 Imaging perturbations of absorption coefficient in an homogeneous semi-infinite medium using the first Born approximation . . . . .	13
2.1.2 Depth Compensation . . . . .	19

2.1.3	Angular Spectrum Representation of Diffuse Photon Density Waves	20
2.2	Tikhonov Regularization	21
2.3	Compressive Sensing	21
2.3.1	Incoherent Sampling Theorem	22
2.3.2	Restricted Isometry Property	23
2.3.3	Iteratively Reweighted Least Squares	25
2.4	Prefiltering Method	25
<b>3</b>	<b>Methodology</b>	<b>29</b>
3.1	Compressive Sensing with the Angular Spectrum Representation of Diffuse Photon Density Waves	29
3.1.1	Matrix Formulation of the Imaging Problem with Measurements in the Angular Spectrum Representation	29
3.1.2	Depth Compensation	33
3.1.3	Compressive Sensing Guarantees of Accurate Recovery	34
3.1.4	Efficient Method for the Sensing Matrix and its Hermitian	35
3.1.5	CS with a Sparsifying Transform	37
3.2	Prefiltering Method with the Angular Spectrum Representation of Diffuse Photon Density Waves	38
3.3	Numerical Experiments	46
3.3.1	General Procedures	51
<b>4</b>	<b>Results and Discussion</b>	<b>58</b>
4.1	Reconstructions of Perturbations in the Small Volume	58
4.2	Reconstructions of Perturbations in the Large Volume	63
4.3	The Effect of Noisy Measurements	68
4.4	Comparisons with the State-of-the-Art technique	69
<b>5</b>	<b>Conclusion</b>	<b>74</b>
5.1	Limitations	75
5.2	Future research	76



# List of Tables

3.1	Light and medium properties used in the simulations. . . . .	47
3.2	Geometric parameters assumed when reconstructing perturbations in the small volume. . . . .	47
3.3	Geometric parameters assumed when reconstructing perturbations in the large volume. . . . .	47
3.4	Properties of the images of perturbation in absorption coefficient that will be considered in Chapter 4. $d_1$ and $d_2$ refer to the dimensions specified in Figure 3.8 for the rectangular prism and the hexagonal prism perturbations.	50
4.1	Parameters related to the reconstruction of the image I using CS. . . . .	59
4.2	Parameters related to the reconstruction of the image IV using CS. . . . .	59
4.3	Parameters related to the reconstruction of the image IV using the pre-filtering method. . . . .	59
4.4	Parameters related to the reconstruction of the image VI using CS. . . . .	64
4.5	Parameters related to the reconstruction of the image VII using CS. . . . .	64
4.6	Parameters related to the reconstruction of the image VII using the pre-filtering method. . . . .	64

# List of Figures

1.1	A portion of the electromagnetic spectrum containing the wavelengths used in DOT, which partially overlap with the red portion of visible spectrum but are mostly contained in the NIR range. . . . .	2
1.2	Comparison of CT and DOT. In CT, the light travels a straight line from source (circle) to detector (triangle), which are kept at a certain distance from the subject. Several measurements are taken with light impinging at several different angles. In DOT, sources and detectors are placed on the boundary of the body part to be imaged, often in reflection mode. The photons suffer several scattering events before being detected. Measurements from different combinations of sources and detectors are used to reconstruct an image. . . . .	3
1.3	Illustration of DOT in functional brain imaging. (a) Sources (circles) and detectors (triangles) are placed over the cortex region to be imaged, in this case the occipital lobe. (b) Some stimulus cause the activation and inflow of blood to a specific spot of this region, which in turn causes an increase in the nearby absorption coefficients. (c) Some reconstruction technique, usually Tikhonov regularization, is used to recover the location and magnitude of the changes in absorption coefficient. The numerical simulation that led to figures (b) and (c) were based on [37]. . . . .	4
2.1	Illustration of the infinite medium solution when the source is an arbitrary point source $-\frac{v}{D_0}S(\mathbf{r}) = -\frac{v}{D_0}S_0\delta(\mathbf{r} - \mathbf{r}_s)$ . In this case, the solution is a rescaled and phase shifted version of $G(\mathbf{r}, \mathbf{r}_s)$ , which has damped sinusoidal waves as real and imaginary parts. . . . .	15

2.2	Illustration of the first Born approximation of the scattered field when the incident field of Figure 2.1 reaches an arbitrary impulse perturbation $\Delta k^2(\mathbf{r}) = P_0\delta(\mathbf{r} - \mathbf{r}_p)$ . In the first Born approximation, the scattered field can be seen as another homogeneous medium solution, with $-\Delta k^2(\mathbf{r})\Phi^i(\mathbf{r}, \mathbf{r}_s)$ as the source. Therefore, the scattered wave emanating from an impulse perturbation is similar to the incident wave of Figure 2.1, but with different amplitude, phase and point of origin. . . . .	17
2.3	Typical separable two-dimensional analysis filterbank, with four branches. To the right, we represent the sub-bands ideally isolated in each branch. .	26
2.4	Schematic representation of how the spectrums of the filtered versions are used to obtain the image spectrum in the prefiltering method. . . . .	27
2.5	Exemplification of the prefiltering method in the MRI case. The data acquisition process is modeled by a linear system of equations $\mathbf{b} = \mathbf{A}\mathbf{x}$ , where $\mathbf{b}$ are the measurements, $\mathbf{x}$ are the image pixels, and $\mathbf{A}$ are selected rows from the two-dimensional DFT matrix. $\mathcal{F}\{\cdot\}$ represents a filter's two-dimensional DFT, with the filter zero padded to the same size as the image, and $\circ$ represents filtering in frequency domain. . . . .	28
3.1	The proposed sensing matrix can be seen as a selection of rows from a unitary matrix. More specifically, it can be seen as the product of two matrices, where the matrix to the right is unitary, and the matrix to the left has rows selected from a unitary matrix. . . . .	36
3.2	(a) Imaging setup based on [20]. Each square represents 64 detectors and the diamond represents a source. The shaded volume corresponds to the imaged volume. The blue disc represents a perturbation in absorption coefficient. The real and imaginary parts of the image to be reconstructed along the red line crossing the perturbation is shown in (b). . . . .	41
3.3	(a) Proposed imaging setup. Each square represents 64 detectors and each diamond represents a sources. The shaded volume corresponds to the imaged volume. The blue disc represents a perturbation in absorption coefficient. The real and imaginary parts of the image to be reconstructed along the red line crossing the perturbation is shown in (b). . . . .	42
3.4	The domain of the DFTs of the set of filter (2.51), divided in 16 rectangular regions. The regions for which the DFTs of the filters $\mathbf{h}_1\mathbf{h}_1^\top$ , $\mathbf{h}_1\mathbf{h}_0^\top$ , $\mathbf{h}_0\mathbf{h}_1^\top$ and $\mathbf{h}_0\mathbf{h}_0^\top$ are nonzero are indicated by HH, HL, LH, and LL respectively. The filters are assumed to be zero padded to size $J_1 \times J_2$ . . . . .	44

3.5	Divisions of the image spectrum among the filtered versions when: (a) only the LH and HL filters are used; (b) the LH, HL and HH filters are used. The regions labeled with an LH are reconstructed using the LH filtered version, the regions labeled with an HL are reconstructed using the HL filtered version, and so on. . . . .	45
3.6	A representation of the small volume, shaded, and the optode arrangement used to reconstruct it. Each square represents 64 detectors and each diamond represents a source. This optode arrangement has a total of 65536 detectors and 20 sources. The perturbation in absorption coefficient is only illustrative. Its actual shape, size and location will depend on the experiment. . . . .	48
3.7	A representation of the large volume, shaded, and the optode arrangement used to reconstruct it. Each square represents 64 detectors and each diamond represents a source. This optode arrangement has a total of 29584 detectors and 20 sources. The perturbation in absorption coefficient is only illustrative. Its actual shape, size and location will depend on the experiment. . . . .	49
3.8	Generic representation of the perturbations in absorption coefficient that will be considered in Chapter 4. Some perturbations will have the shape of the rectangular prism to the left, and others will have the shape of the hexagonal prism to the right. In the top left corner and the bottom left corner are the frontal and top views of the rectangular prism, respectively, and in the top right corner and the bottom right corner are the frontal and top views of the hexagonal prism, respectively. . . . .	49
3.9	Optode arrangement used with Tikhonov regularization. Each square represents one detector, and each diamond represents one source. This optode arrangement has a total of 4096 detectors and 16 sources. . . . .	57
4.1	(a) Layers of the VOI of the image I. (b) Reconstructed layers of the VOI of the image I using CS. Black indicates a perturbation in squared wavenumber of $10 \text{ cm}^{-2}$ and white a perturbation in squared wavenumber of $-10 \text{ cm}^{-2}$ . . . . .	60
4.2	(a) Layers of the VOI of the image IV. (b) Reconstructed layers of the VOI of the image IV using CS. Black indicates a perturbation in squared wavenumber of $10 \text{ cm}^{-2}$ and white a perturbation in squared wavenumber of $-10 \text{ cm}^{-2}$ . . . . .	61



4.3	(a) Layers of the VOI of the image IV. (b) Reconstructed layers of the VOI of the image IV using the prefiltering method. Black indicates a perturbation in squared wavenumber of $10 \text{ cm}^{-2}$ and white a perturbation in squared wavenumber of $-10 \text{ cm}^{-2}$ . . . . .	62
4.4	(a) Layers of the VOI of the image VI. (b) Reconstructed layers of the VOI of the image VI using CS. Black indicates a perturbation in squared wavenumber of $10 \text{ cm}^{-2}$ and white a perturbation in squared wavenumber of $-10 \text{ cm}^{-2}$ . . . . .	65
4.5	(a) Layers of the VOI of the image VII. (b) Reconstructed layers of the VOI of the image VII using CS. Black indicates a perturbation in squared wavenumber of $10 \text{ cm}^{-2}$ and white a perturbation in squared wavenumber of $-10 \text{ cm}^{-2}$ . . . . .	66
4.6	(a) Layers of the VOI of the image VII. (b) Reconstructed layers of the VOI of the image VII using the prefiltering method. Black indicates a perturbation in squared wavenumber of $10 \text{ cm}^{-2}$ and white a perturbation in squared wavenumber of $-10 \text{ cm}^{-2}$ . . . . .	67
4.7	The effect of different amounts of noise on the reconstruction of the image II using CS. . . . .	69
4.8	The effect of different amounts of noise on the reconstruction of the image III using the prefiltering method. . . . .	69
4.9	SER curves obtained for different variants of the proposed techniques and for Tikhonov regularization when the size of a specific rectangular prism perturbation was varied. . . . .	71
4.10	SER curves obtained for different variants of the proposed techniques and for Tikhonov regularization when the sizes of two specific rectangular prism perturbations were varied. . . . .	72
4.11	SER curves obtained for different variants of the proposed techniques and for Tikhonov regularization when the size of a specific hexagonal prism perturbation was varied. . . . .	73

## ABBREVIATIONS

BOLD	Blood Oxygen Level Dependent
CCD	Charge Coupled Device
CS	Compressive Sensing
CT	Computerized Tomography
CW	Continuous Wave
DFT	Discrete Fourier Transform
DOT	Diffuse Optical Tomography
DWT	Discrete Wavelet Transform
EEG	Electroencephalography
FEM	Finite Elements Method
fMRI	Functional Magnetic Resonance Imaging
HH	High-pass High-pass
HL	High-pass Low-pass
IRLS	Iteratively Reweighted Least Squares
LH	Low-pass High-pass
LL	Low-pass Low-pass
MEG	Magnetoencephalography
MMV	Multiple Measurement Vector
MRI	Magnetic Resonance Imaging
NIR	Near Infrared
NIRS	Near Infrared Spectroscopy
PET	Positron Emission Tomography
RF	Radio Frequency
RIP	Restricted Isometry Property
SMV	Single Measurement Vector
SER	Signal to Error Ratio
US	Ultrasonography
VOI	Volume of Interest

## NOTATION

We use bold uppercase letters for matrices, bold lowercase letters for column vectors, and non-bold letters for scalars. We use  $\sim$  to indicate quantities in spatial frequency domain,  $'$  and  $\hat{\phantom{x}}$  to distinguish a variable from a similar one, and  $*$  to indicate recovered quantities. We use subscripts either for indexing or to further distinguish one variable from another. We use superscripts for exponentiation or also for further distinction. Whenever index notation is used, the counting starts from zero (zero based indexing). We use  $\Delta$  to indicate a quantity that is the difference of other two quantities.

In a system of linear equation, we use one dot above a matrix or vector to indicate that either its rows or columns had been rescaled. If there are any quantities formed by rearranging the elements, or a subset of the elements, of the matrix or vector, we also use a dot above it to indicate that the associated quantity in the system of linear equations had its rows or columns rescaled. Two dots indicate that this process had been repeated in a case that it had already been done before with a different rescaling.

We reserve the characters  $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, X, Y, Z, \mathbf{X}, \mathbf{Y}, \mathbf{X}, \mathbf{Y}, \mathbb{C}, \mathbb{R}$  for sets and the character  $F$  for a distribution. The characters  $x, y$  and  $z$  are the cartesian coordinates, unless otherwise specified. The following examples give more details of the notation and define some operators.

$\mathbb{R}, \mathbb{C}$	The sets of real and complex numbers, respectively.
$\mathbb{R}^{I \times J}, \mathbb{C}^{I \times J}$	The sets of real and complex $I \times J$ matrices, respectively.
$\mathbb{R}^I, \mathbb{C}^I$	The sets of real and complex column vectors of dimension $I$ respectively.
$\Re(z)$	Real part of $z \in \mathbb{C}$ . If $z = a + jb$ , then $\Re(z) = a$ .
$\Im(z)$	Imaginary part of $z \in \mathbb{C}$ . If $z = a + jb$ , then $\Im(z) = b$ .
$\bar{z}$	Complex conjugate of $z \in \mathbb{C}$ . If $z = a + jb$ , then $\bar{z} = a - jb$ .
$ z $	Absolute value of $z \in \mathbb{C}$ . If $z = a + jb$ , then $ z  = \sqrt{a^2 + b^2}$ .
$\arg(z)$	Angle between $z \in \mathbb{C}$ and the real axis, given by the four-quadrant inverse tangent function [39]. If $z = a + jb$ , then
$\arg(z) =$	{
	$\tan^{-1}(b/a)$ if $a > 0$ $\pi/2$ if $a = 0$ and $b > 0$ $-\pi/2$ if $a = 0$ and $b < 0$ $0$ if $a = 0$ and $b = 0$ $\pi + \tan^{-1}(b/a)$ if $a < 0$

where  $\tan^{-1}$  is the inverse of the tangent function defined over the interval  $(-\pi/2, \pi/2)$ .

$\lfloor x \rfloor$	Greatest integer less than or equal to $x \in \mathbb{R}$ .
$x \bmod y$	Remainder after division of $x \in \mathbb{R}$ by $y \in \mathbb{R}$ [51], that is  $x \bmod y = x - y\lfloor x/y \rfloor$ .
$\arg \min_{x \in \mathbb{C}} f(x)$	Value or values of $x \in \mathbb{C}$ that minimize the real valued function $f(x)$ .
$\min_{x \in \mathbb{C}} f(x)$	Minimum value of the real valued function $f(x)$ , when $x \in \mathbb{C}$ .
$\frac{\partial f}{\partial t}$	Partial derivative of $f$ with respect to $t$ .
$\nabla^2 f$	Laplacian of $f$ , that is  $\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$
$\mathbb{E}(a)$	If $a$ is a discrete random variable, then  $\mathbb{E}(a) = \sum_{i=0}^{\infty} a_i p_i,$  where $a_0, a_1, \dots$ are the possible outcomes of the random variable, and $p_i$ is the probability of $a_i$ . If there are only $I$ possible outcomes, then  $\mathbb{E}(a) = \sum_{i=0}^{I-1} a_i p_i.$  If $a$ is a continuous random variable, then  $\mathbb{E}(a) = \int_{a' \in \mathbb{R}} a' f(a') da',$  where $f(a')$ is the probability density function of the random variable. If $a$ is a complex random variable, then $\mathbb{E}(a) = \mathbb{E}(\Re(a)) + j\mathbb{E}(\Im(a))$ .
For $i = 0, \dots, I - 1$ and $j = 0, \dots, J - 1$ :	
$\mathbf{X}_i$	The $i$ 'th matrix from a family of matrices.
$\mathbf{x}_i$	The $i$ 'th vector from a family of vectors.
$x_i$	The $i$ 'th scalar from a family of scalars.
$[\mathbf{x}]_i$ , or $x_i$	The $i$ 'th element of the vector $\mathbf{x}$ .
$[\mathbf{X}]_{i,j}$ , or $x_{i,j}$	The element in the $i$ 'th row and the $j$ 'th column of the matrix $\mathbf{X}$ .
$\mathcal{X}_i$	The $i$ 'th element from the set $\mathcal{X}$ .

$\mathbf{X} \circ \mathbf{Y}$  Hadamard or entrywise product of  $\mathbf{X}$  and  $\mathbf{Y}$ . The matrices  $\mathbf{X}$  and  $\mathbf{Y}$  must be of the same size. If  $\mathbf{X}$  and  $\mathbf{Y}$  are  $I \times J$ , then  $\mathbf{X} \circ \mathbf{Y}$  is also  $I \times J$  and

$$[\mathbf{X} \circ \mathbf{Y}]_{i,j} = [\mathbf{X}]_{i,j} [\mathbf{Y}]_{i,j}.$$

$\mathbf{X}^\top$  Transpose of  $\mathbf{X}$ . If  $\mathbf{X}$  is  $I \times J$  then  $\mathbf{X}^\top$  is  $J \times I$  and

$$[\mathbf{X}^\top]_{j,i} = [\mathbf{X}]_{i,j}.$$

$\mathbf{X}^H$  Conjugate transpose or Hermitian of  $\mathbf{X}$ . If  $\mathbf{X}$  is  $I \times J$  then  $\mathbf{X}^H$  is  $J \times I$  and

$$[\mathbf{X}^H]_{j,i} = \overline{[\mathbf{X}]_{i,j}}.$$

$\mathbf{X} \otimes \mathbf{Y}$  Kronecker product of  $\mathbf{X}$  and  $\mathbf{Y}$ . If  $\mathbf{X}$  is  $I \times J$  then

$$\mathbf{X} \otimes \mathbf{Y} = \begin{bmatrix} [\mathbf{X}]_{0,0} \mathbf{Y} & \cdots & [\mathbf{X}]_{0,J-1} \mathbf{Y} \\ \vdots & \ddots & \vdots \\ [\mathbf{X}]_{I-1,0} \mathbf{Y} & \cdots & [\mathbf{X}]_{I-1,J-1} \mathbf{Y} \end{bmatrix}.$$

$\text{vec}(\mathbf{X})$  If  $\mathbf{x}_0, \dots, \mathbf{x}_{J-1}$  are the columns of  $\mathbf{X}$ , that is  $\mathbf{X} = [\mathbf{x}_0 \ \cdots \ \mathbf{x}_{J-1}]$ , then

$$\text{vec}(\mathbf{X}) = \begin{bmatrix} \mathbf{x}_0 \\ \vdots \\ \mathbf{x}_{J-1} \end{bmatrix}.$$

$\text{diag}(\mathbf{x})$  If  $\mathbf{x}$  is  $I \times 1$ , then  $\text{diag}(\mathbf{x})$  is the  $I \times I$  diagonal matrix with the elements of  $\mathbf{x}$  in the diagonal, that is

$$\text{diag}(\mathbf{x}) = \begin{bmatrix} x_0 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & x_{I-1} \end{bmatrix}.$$

$\|\mathbf{X}\|_F$  Frobenius norm of the matrix  $\mathbf{X}$ . If  $\mathbf{X}$  is  $I \times J$ , then

$$\|\mathbf{X}\|_F = \left( \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} |x_{i,j}|^2 \right)^{1/2}.$$

$\|\mathbf{x}\|_p, p > 0$   $\ell_p$  of  $\mathbf{x}$ . If  $\mathbf{x}$  is  $I \times 1$ , then

$$\|\mathbf{x}\|_p = \left( \sum_{i=0}^{I-1} |x_i|^p \right)^{1/p}.$$

$\text{SER}(\mathbf{x}^*, \mathbf{x})$  SER between a reference vector  $\mathbf{x}$  and a recovered vector  $\mathbf{x}^*$

$$\text{SER}(\mathbf{x}^*, \mathbf{x}) = 10 \log_{10} \left( \frac{\|\mathbf{x}\|_2^2}{\|\mathbf{x} - \mathbf{x}^*\|_2^2} \right).$$

$\text{SER}(x^*, x)$  If  $x[n]$  and  $x^*[n]$ , for  $n = 0, \dots, N - 1$ , are discrete finite length functions, then

$$\text{SER}(x^*, x) = 10 \log_{10} \left( \frac{\sum_{n=0}^{N-1} |x[n]|^2}{\sum_{n=0}^{N-1} |x[n] - x^*[n]|^2} \right).$$

$\Re(\mathbf{X})$  Real part of  $\mathbf{X} \in \mathbb{C}^{I \times J}$ . If  $\mathbf{X} = \mathbf{A} + j\mathbf{B}$ , then  $\Re(\mathbf{X}) = \mathbf{A}$ .

$\Im(\mathbf{X})$  Imaginary part of  $\mathbf{X} \in \mathbb{C}^{I \times J}$ . If  $\mathbf{X} = \mathbf{A} + j\mathbf{B}$ , then  $\Im(\mathbf{X}) = \mathbf{B}$ .

$\overline{\mathbf{X}}$  Complex conjugate of  $\mathbf{X} \in \mathbb{C}^{I \times J}$ . If  $\mathbf{X} = \mathbf{A} + j\mathbf{B}$ , then  $\overline{\mathbf{X}} = \mathbf{A} - j\mathbf{B}$ .

$\mathbb{E}(\mathbf{X})$  If  $\mathbf{X}$  is  $I \times J$ , then

$$\mathbb{E}(\mathbf{X}) = \begin{bmatrix} \mathbb{E}([\mathbf{X}]_{0,0}) & \dots & \mathbb{E}([\mathbf{X}]_{0,J-1}) \\ \vdots & \ddots & \vdots \\ \mathbb{E}([\mathbf{X}]_{I-1,0}) & \dots & \mathbb{E}([\mathbf{X}]_{I-1,J-1}) \end{bmatrix}.$$

$\otimes$  The two-dimensional circular convolution.

$\mathcal{F}, \mathcal{F}^{-1}$  Forward and inverse Fourier transforms.

$\mathcal{W}, \mathcal{W}^{-1}$  Forward and inverse wavelet transforms.

# 1 INTRODUCTION

Medical imaging is of paramount importance in today's clinical practice. It provides clinicians with images of the interior of the human body without the need of surgical intervention. It not only makes possible the noninvasive diagnosis and monitoring of several diseases, but it also improves surgical and treatment plannings. Such techniques are also essential in medical research. They allow the study of the body's normal anatomy and function, and how they break down in unhealthy subjects.

In its early days, medical imaging provided only projection images of the body's interior organs. This was simpler and did not require computer processing. However, as the limitations of this type of imaging became clear, and the development of computers advanced, researches developed cross sectional imaging, or tomographic imaging, of the inside of the human body. This was when techniques such as ultrasound (US), positron emission tomography (PET), computerized tomography (CT) and magnetic resonance imaging (MRI) were developed [8].

Several imaging techniques are available nowadays to facilitate the work of clinicians and medical researches. With respect to their purposes, these types of imaging can be separated into structural and functional imaging. Structural imaging is primarily intended to reconstruct properties associated with the anatomy and morphology of the human body, while functional imaging is primarily intended to reconstruct properties associated with the body's biochemistry and biological activity [1].

Imaging techniques such as US, CT and MRI are classified as structural [1]. The properties reconstructed in each of these techniques that are associated with the body's anatomy and morphology are: localized acoustic impedance mismatches, in the case of US [16]; the concentration of hydrogen along the body, in the case of MRI [55]; and the localized absorption of light in the wavelength range from 0.008 to 0.06 nm, in the case of CT [48]. On the other hand, imaging techniques such as electroencephalography (EEG), magnetoencephalography (MEG), functional MRI (fMRI) and PET are classified as functional [1, 81]. fMRI detects changes in deoxygenated hemoglobin concentration by means of the blood oxygen level dependent (BOLD) signal [55, 66]. PET reconstructs the locations of molecules that were injected inside the body to track their metabolism [69]. EEG and MEG monitor, respectively, the electrical and magnetic activities of the

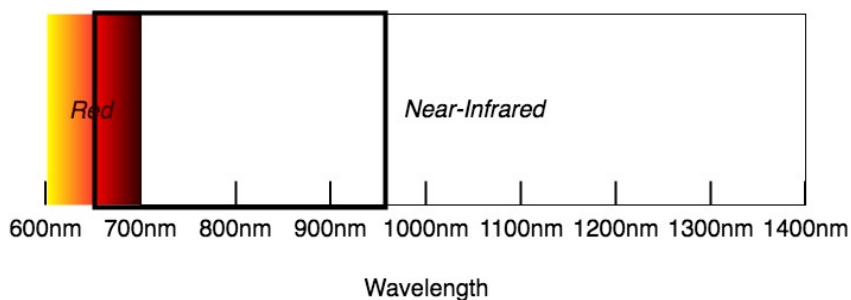
brain [81].

Each structural or functional imaging technique has its advantages, disadvantages and exclusive applications. MRI is considered the imaging modality with the highest spatial resolution, but it requires a very expensive and bulky equipment, and the exams take a considerable amount of time [63, 31, 35]. CT has a lower spatial resolution than MRI, but it also has a lower cost [44]. Another disadvantage of CT is that it uses ionizing radiation, even if in small quantities [31]. fMRI also uses the same costly and bulky equipment as MRI [55]. Both EEG and MEG have a high temporal resolution, but a very low spatial resolution [81]. The main disadvantage of PET is that it uses radioisotopes [69].

Due to the particularities of each of the imaging techniques, all of them tend to be used in the hospitals nowadays. The clinician decides which screening to request based on the risk, convenience, cost and information content of the imaging techniques.

Because of the high demands of healthcare, the research on medical imaging has never really stopped. Great efforts are still made on improving the instruments quality, safety, portability and cost. The approaches adopted by researchers are not limited to further developing the existing imaging techniques, but also include the development of completely new techniques.

One relatively new imaging technique that is still under development is diffuse optical tomography (DOT) [42]. In contrast to CT, this imaging modality uses light from 650nm to 950nm, which is a harmless form of radiation. Figure 1.1 shows a portion of the electromagnetic spectrum that includes these wavelengths. They partially overlap with the red portion of the visible spectrum, but are mostly contained in the near infrared (NIR) range. The reason why these wavelengths are used and not other forms of safe radiation is because light at these wavelengths experience less absorption inside biological tissue [6, 77].

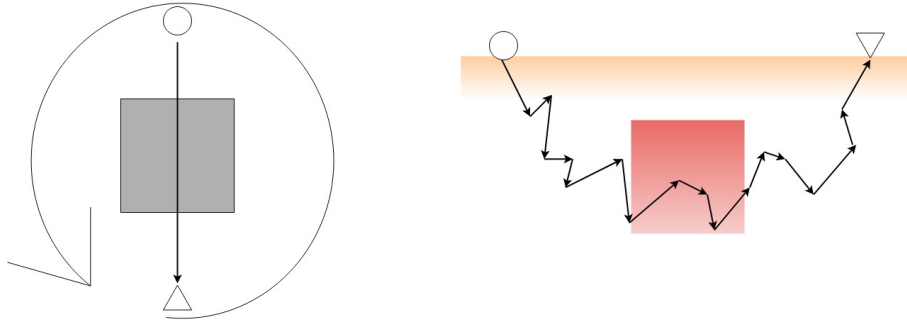


**Figure 1.1.** A portion of the electromagnetic spectrum containing the wavelengths used in DOT, which partially overlap with the red portion of visible spectrum but are mostly contained in the NIR range.

However, the trajectory of these forms of light inside biological tissue is not as the trajectory of the X-rays of CT, which suffer little or no scattering events from source to



detector [42]. These forms of light are intensely scattered, and in some situations can even be detected with sources and detectors in reflection mode. Their absorption is also much more intense than the absorption of X-rays. For these reasons, sources and detectors in DOT are placed on the borders of the body part to be imaged, often in reflection mode to image just a portion close to the surface. Figure 1.2 shows a comparison of CT and DOT with respect to the general setup and light trajectories. Also, Figure 1.3 illustrates the use of DOT to reconstruct the activity of a specific region of a subject’s brain cortex.

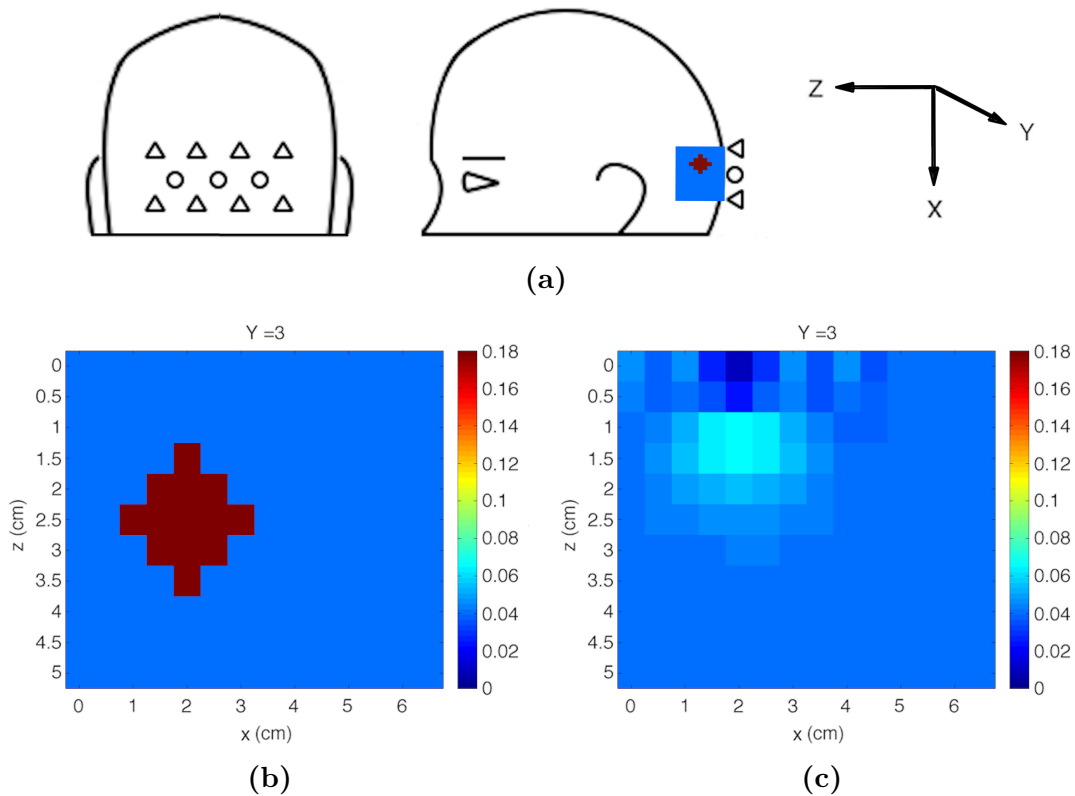


**Figure 1.2.** Comparison of CT and DOT. In CT, the light travels a straight line from source (circle) to detector (triangle), which are kept at a certain distance from the subject. Several measurements are taken with light impinging at several different angles. In DOT, sources and detectors are placed on the boundary of the body part to be imaged, often in reflection mode. The photons suffer several scattering events before being detected. Measurements from different combinations of sources and detectors are used to reconstruct an image.

In contrast to CT, DOT is classified as a functional imaging technique. This is because the absorption of light in the wavelength range from 650nm to 950nm is primarily due to the oxygenated and deoxygenated hemoglobins, which are more related to the body function than to the body anatomy [78]. But DOT can also be used to reconstruct the scattering intensities along the body, since this is also a predominant effect during red and NIR light propagation. The scattering that happens inside biological tissue is caused by cells, cell nuclei, cell organelles and surrounding fluids [29]. This information can also be linked to the body function in some cases [89].

Current areas of DOT applications include: brain imaging, breast cancer imaging, diagnosis of rheumatoid arthritis in fingers, among others [6, 42]. In breast cancer imaging and in functional brain imaging, the focal increases in hemoglobin concentrations caused by angiogenesis in cancer and regional cerebral blood flow, respectively, cause a local increase in absorption which can be detected in DOT [29, 66, 87]. In finger rheumatoid arthritis diagnosis, the presence of rheumatoid arthritis cause the inflammation of the synovial fluid, which in its turn translates into changes in the synovial fluid optical properties that can be detected in DOT [52].

There are many advantages of DOT over other functional imaging tools. It has a



**Figure 1.3.** Illustration of DOT in functional brain imaging. (a) Sources (circles) and detectors (triangles) are placed over the cortex region to be imaged, in this case the occipital lobe. (b) Some stimulus cause the activation and inflow of blood to a specific spot of this region, which in turn causes an increase in the nearby absorption coefficients. (c) Some reconstruction technique, usually Tikhonov regularization, is used to recover the location and magnitude of the changes in absorption coefficient. The numerical simulation that led to figures (b) and (c) were based on [37].

higher temporal resolution than fMRI and PET while having a higher spatial resolution than EEG and MEG, meaning it is midway between the other modalities [11]. If more than one wavelength is used, DOT can recover images of oxygenated hemoglobin and deoxygenated hemoglobin, which means DOT can inform the oxygen saturation along the imaged region, data not provided by the other techniques [89]. Compared to PET, this modality does not use ionizing radiation, and compared to fMRI, it can be used on patients with implanted electronics [32]. DOT is also less susceptible to electrical noise than EEG.

Other important advantages of DOT are that it is generally smaller, less expensive, less restraining, less sensitive to motion and requires little to no maintenance [43]. Because of these reasons, it can be used to study patients difficult to study using the other imaging modalities, such as: newborns, agitated and claustrophobic patients, athletes while exercising, persons in operatory rooms and intensive care units, the elderly, among others [7].

## 1.1 HISTORY OF DOT

The idea of using NIR light to examine the human body is much older than DOT [6, 77]. For example, pulse oximeters are devices very similar to DOT instruments in their working principle that have been used in the hospitals for many decades. Such devices continuously measure the absorption of a thin part of the body with respect to two different wavelengths in the near infrared range, to then estimate the heart rate and the blood oxygen saturation.

However, perhaps the first major milestone towards the development of a tomography using NIR light came with the work of Jobsis [46]. He was the first to show that it was possible to monitor concentrations of oxygenated and deoxygenated hemoglobin in the brain noninvasively, even though in this case there is the presence of the skull [77, 66]. This led to the development of several instruments to measure hemoglobin concentration at a certain region of the brain cortex by illuminating the head from a spot on its surface and then measuring the reflected light at a close spot on the surface. This practice is named near infrared spectroscopy (NIRS) [77].

The next development came by combining several of such measurements at different locations of the head to obtain topographic images of hemoglobin concentration spanning large regions of the cortex [77]. The disadvantage of this simple approach, however, was that it had no depth resolution and was very sensitive to extracerebral compartments of the head [77].

For some time, the development of DOT was delayed by the need of better under-

standing about how light propagates in highly scattering mediums [43]. The radiative transfer equation accurately models the propagation of light in highly scattering mediums, but its application poses high computational load [42]. It was only with the widespread acceptance of the applicability of diffusion theory in this situation that the technology could be further developed [29]. The basic requirements are that the observation point be sufficiently distant from the source, and that the photons undergo several scattering events before being absorbed, which is usually true for light in the near infrared range propagating in biological tissue [45].

Current DOT systems measure, at several different locations, light of a certain wavelength emitted also from several different places. The resulting data include measurements with sensitivity to overlapping regions. In most cases, diffusion theory is then used to recover location dependent absorption and scattering coefficients. In some cases these coefficients are sufficient. In cases that concentrations of oxygenated and deoxygenated hemoglobin must be recovered, it is necessary to repeat the measurements and the reconstruction for a different wavelength. The two maps of absorption coefficient for the two wavelengths are then used to recover maps of concentration of oxygenated and deoxygenated hemoglobin.

There are three types of DOT systems: short pulse systems, radio frequency (RF) systems and continuous wave (CW) systems [6]. Short pulse systems illuminate the tissue in brief periods and, when they detect photons, they also measure the photon's time of arrival [77]. RF systems use amplitude modulated light at megahertz frequencies and measure both the amplitude and the phase of the detected light. CW systems either emit light at a constant amplitude or amplitude modulate at kilohertz frequencies and measure only the amplitude of the detected light [6]. Short pulse systems contain the most amount of information, but are the most complex systems [77]. CW systems are the simplest and cheapest to implement, but it is proven that it cannot recover absorption and scattering simultaneously [2].

## 1.2 IMAGE RECONSTRUCTION IN DOT

The passage of photons through the tissue is usually modeled by the photon diffusion equation. Therefore, reconstruction techniques generally involve the solution of the photon diffusion equation for the fluence rate at the detector locations. This is named the forward problem. Analytical solutions to the forward problem are available, but they are limited to simple mediums in structure and geometry [38]. For this reason, solutions using numerical modeling techniques, like the finite elements method (FEM) are preferred in real life applications. In possession of a way of solving the forward problem, one attempts to solve the inverse problem, that is, given measurements of fluence rate at the boundary

for sources at different locations, find the distributions of absorption and scattering.

The inverse problem also can be solved in several different ways. There are linear model reconstruction techniques and nonlinear model reconstruction techniques. In linear model reconstruction techniques, the measurements are assumed to vary linearly with changes in the medium optical properties. This results in an underdetermined linear system of equations that relates changes in the measurements to changes in optical properties. One can then try to recover images of changes in optical parameters by finding special solutions to the underdetermined linear system of equations, like the least squares solution, in which case the least squares problem must be regularized, since the inverse problem is ill posed.

linear model reconstruction techniques require measurements in two states, one before and other after the changes in optical parameters [26]. They are primarily used in functional brain imaging to detect changes in blood supply to specific parts of the cerebral cortex during different tasks [38]. Increase of blood supply to a specific area means that area is activated [66]. linear model reconstruction techniques can be used with both analytical and numerical forward models.

In nonlinear model reconstruction techniques, the distribution of optical parameters is varied until the error between the experimental and calculated data using the forward model is minimized. This constitutes a nonlinear least squares problem that also requires some regularization technique due to the high ill-posedness of the inverse problem [42].

Even though there is a lot of interest in DOT, there are a lot of difficulties associated with its image quality that limit its current applications. Some of the reasons why the DOT images are of generally low quality are: 1) Due to the high complexity of the DOT problem, and since there are many parameters that are unknown or difficult to measure, the DOT modeling process often involves several approximations, which introduces errors in the reconstruction [6]; 2) The most common reconstruction technique used employs Tikhonov regularization, which has the characteristic of smoothing out the reconstructed image, thus lowering its resolution [38]. In part because of these reasons, and also due to the severe hardware and software requirements imposed for acceptable quality images, the range of applications of DOT has been much more restricted than its actual potential [21].

One solution to the second problem that has been recently adopted is the use of  $\ell_1$  regularization instead of  $\ell_2$  regularization in both the linear and nonlinear DOT least squares problems [57, 33, 82, 71, 41, 80, 67, 4, 18, 49, 79, 85, 15]. The  $\ell_1$  regularization has the characteristic of favoring sparse solutions instead of smooth ones, thus solving the problem of smooth images.

The use of the  $\ell_1$  norm to regularize the DOT least squares problem is closely related

to another reconstruction technique that has been drawing a lot of attention in the past few years: compressive sensing (CS). CS is a technique for recovering signals from much fewer samples than the Nyquist Shannon sampling rate by exploring its sparsity [74]. The conditions however, are that the sensing mechanism cannot be a simple sampling of the signal, and that it must also satisfy other requirements [19].

The current theory of CS applies to underdetermined linear systems of equations. Therefore the area which DOT can mostly benefit from CS are in linear model reconstruction techniques. The research in this area is still underdeveloped, with one publication only dealing with the requirements that the DOT sensing matrix must satisfy to recover sparse signals [53]. Moreover, there is one CS related technique in particular, the prefiltering method, which was developed in the context of Magnetic Resonance Imaging (MRI), that has consistently outperformed some other traditional MRI reconstruction techniques that could also be used in DOT, with some adaptation [61, 62].

The advantage of CS, and its related reconstruction techniques, is its associated theory. The theory of CS predicts the possibility of accurate image restoration under certain circumstances. Such reconstruction guarantees could mean a great improvement in the image quality and reliability of DOT, at least in the context of linear model reconstruction techniques.

## 1.3 RESEARCH OBJECTIVES

### 1.3.1 General Objective

Although there is already a formulation of the DOT problem proven to satisfy the requirements of CS under certain circumstances [53], this formulation does not immediately allow the use of the prefiltering method in DOT. Moreover, there are limitations in the assumed model that justify the search for other formulations. Therefore, we propose the development of another mathematical formulation for the DOT imaging problem that not only satisfies the requirements for CS, but also allows the use of the prefiltering method in DOT, and possibly makes less stringent requirements about the characteristics of the imaged region. Based on this formulation, we also propose the development of the associated algorithms necessary for image reconstruction, and the evaluation of the resulting techniques in simulated settings.

Note that the prefiltering methods rely on specially chosen filters operating on the measurement domain and lead to the reconstruction of several filtered images, based on which a final composition stage generates the desired reconstructed image. It is worth emphasizing that, contrary to the prefiltering approach for MRI, in DOT the measurements are not samples in the frequency domain, so that a new theoretical development

was required in order to compute the measurements associated to filtered versions of the target image. We present this development in the following pages.

In the following, we also define a set of numerical experiments using this mathematical formulation and present the performance results in terms of objective metrics.

### 1.3.2 Specific Objectives

In order to attain the general objective, we must meet the following specific goals:

1. Show that a different formulation of the DOT problem may satisfy the requirements of CS under specific circumstances.
2. Show how to obtain measurements associated with filtered versions of the image in this formulation of the DOT problem.
3. Develop efficient and effective algorithms for the complete prefiltering and CS techniques derived from this formulation.
4. Develop numerical experiments to demonstrate the effectiveness of the proposed algorithms.
5. Evaluate the obtained images in terms of objective metrics.
6. Compare the reconstructions obtained using the proposed techniques with the reconstructions obtained using the state-of-the-art technique, in terms of objective metrics.
7. Analyze and discuss the results based on the theories and principles involved with the techniques.

## 1.4 CONTRIBUTIONS

We found only one other publication in the literature that proposed a formulation of the DOT problem and showed that this formulation could satisfy the requirements for CS under certain circumstances [53]. Their approach has the advantage that the DOT problem is modeled as a multiple measurement vector (MMV) problem, while in ours it is modeled as a single measurement vector (SMV) problem, and the successful recovery rate of the MMV model is generally higher than that of the SMV model [54]. However, there is no reason to assume that our approach cannot be adapted to the MMV model as well.

Both approaches are similar in that they assume constant scattering coefficients throughout the medium and that they reconstruct perturbations in absorption coefficient in an otherwise homogeneous medium. Our approach further assumes that the perturbations are small in magnitude and support due to the first Born approximation, but we could also adopt a similar strategy as in [53] to avoid this approximation.

One of our contributions with respect to this previous publication consists of the acknowledgement of the boundary effects. In the demonstration of their final theorem, they assumed an infinite medium Green’s function. Therefore, even though they considered a planar detection geometry, they did not consider the effects of the boundaries. In our approach, we use a semi-infinite medium Green’s function, which takes into account such effects, and therefore we show that DOT satisfies the requirements of CS in a more realistic setting with respect to the boundaries.

Furthermore, our approach allows the use of the prefiltering method in DOT. The prefiltering method has several advantages over other reconstruction techniques. One of them is that it is more flexible than the use of an explicit sparsifying transformation. When a sparsifying transformation is used in CS, the sensing matrix is modified, and this can have major effects in the eligibility for CS. The prefiltering method on the other hand makes possible the sparsification of the images to be reconstructed without changing the sensing matrix. This means that, no matter the set of filters used, the modified linear system will always satisfy the requirements for CS. When using an explicit sparsifying transform, on the other hand, sometimes a transformation effectively sparsify the image, but the resulting system is not eligible for CS.

Our overall method is also expected to find a solution quicker than the method of reference [53], because the SMV optimization problem is generally simpler to solve than the MMV optimization problem [56, 83]. Our method also includes a depth compensation matrix that is easier to compute than the one used in the references [4, 49, 65, 64]. A depth compensation matrix is a mathematical artifice used in DOT to work around badly scaled sensing matrices. At first, the matrix we propose only makes sense with the DOT formulation we use, but it may be adapted to other formulations. The basic idea is to use the known Green’s function decay with distance to rescale the sensing matrix. This is different from the traditional technique, which uses a numerical reasoning, instead of an analytical reasoning, to rescale the sensing matrix.

Our numerical results also contribute to the DOT literature in that, before the publication of this document, only three articles reported experimental or simulation results of CS in DOT [27, 28, 53], and there was no mention of the prefiltering method in the DOT literature.



## 1.5 DOCUMENT STRUCTURE

The next chapters are organized as follows. Chapter 2 is devoted exclusively to a literature review. Chapter 3 is dedicated to a description of the proposed methods, the numerical experiments and the analysis performed on the results. In Chapter 4 we present and discuss the graphs and objective metrics obtained from the numerical experiments. In Chapter 5 we review the main results and assumptions made in the development of the techniques, and point out topics for future research.

The analytical DOT model on which we base our techniques is described in the Subsection 2.1.1. For didactic reasons we also provide a common matrix formulation derived from this model. This matrix is frequently used together with the depth compensation method described on Subsection 2.1.2.

The aforementioned analytical model relates the medium optical properties with the light intensities in the three-dimensional space. If we convert the  $x$  and  $y$  dimensions to spacial frequency domain, we obtain the angular spectrum representation, which we describe on Subsection 2.1.3. This representation will be useful in the adaptation of the prefiltering method to DOT.

The Section 2.2 contains a brief overview of Tikhonov regularization, the linear model reconstruction technique most commonly used in DOT.

One of the main goals of our research was to find a DOT formulation that satisfied the requirements for CS. In a later section, we will show that a specific DOT formulation satisfies the requirements of the CS theorems reviewed in the Subsections 2.3.1 and 2.3.2. The Subsection 2.3.3 reviews the algorithm that we will use to minimize the  $\ell_1$  norm of the solution, The iteratively reweighted least squares (IRLS). Finally, in order to adapt the prefiltering method to DOT, which is the other main goal of our research, we need to understand how it works. Section 2.4 gives the details of the prefiltering method.

In the Subsection 3.1.1, we derive the matrix formulation that satisfies the requirements for CS and allows the use of the prefiltering method in DOT. To prevent bias towards the surface, we introduce in the Subsections 3.1.2 a depth compensation method specific to this formulation of the DOT problem. In the Subsection 3.1.3 is where we actually show that this formulation satisfies the requirements for CS.

Bearing in mind that any CS reconstruction algorithm will invariably require subsequent applications of the sensing matrix and its Hermitian [76], we provide efficient methods for these two matrices in the Subsection 3.1.4.

In the Subsection 3.1.5, we propose a sparsifying transform based on the Haar discrete wavelet transform (DWT) that maintains the most important properties of the sensing matrix and therefore still allows the use of CS. Later we will compare the reconstructions

of CS using this transform with the reconstructions of the prefiltering method.

In the Section 3.2 we show how measurements associated with filtered versions of the image can be obtained in the proposed formulation of the DOT problem, and how the final image can be composed once the filtered versions are recovered by CS. This process of obtaining measurements associated with filtered versions of the image, recovering the filtered versions using CS, and then composing the the image from its filtered versions is what constitutes the prefiltering method.

In the Section 3.3, we describe the experiments that we made to evaluate the proposed techniques. The results of such experiments are presented on Sections 4.1 to 4.4. In the first two of these sections, we report proof of concept reconstructions of perturbations in two volumes of different sizes. In the third section, we evaluate the proposed techniques in the presence of noise, and in the last section we compare different variants of our proposed techniques to the most common DOT technique.

## 2 THEORETICAL BACKGROUND

### 2.1 DIFFUSE OPTICAL TOMOGRAPHY

#### 2.1.1 Imaging perturbations of absorption coefficient in an homogeneous semi-infinite medium using the first Born approximation

The explanation in this section is based on the references [37, 36, 68, 48]. We begin with the Photon Diffusion Equation, which models the passage of photons through highly scattering mediums:

$$-D(\mathbf{r})\nabla^2\Phi(\mathbf{r},t) + v\mu_a(\mathbf{r})\Phi(\mathbf{r},t) + \frac{\partial\Phi(\mathbf{r},t)}{\partial t} = vS(\mathbf{r},t), \quad (2.1)$$

where

$$D(\mathbf{r}) = \frac{v}{3\mu'_s(\mathbf{r})}. \quad (2.2)$$

In this equation,  $\Phi(\mathbf{r},t)$  is the photon fluence at position  $\mathbf{r} = (x,y,z)$  and time  $t$  in watts per square meter,  $S(\mathbf{r},t)$  is the isotropic source term in watts per cubic meter,  $v$  is the speed of light in the medium in meters per second,  $\mu_a$  is the absorption coefficient in inverse meters,  $\mu'_s$  is the reduced scattering coefficient in inverse meters,  $D$  is the photon diffusion coefficient in square meters per second, and  $v$ ,  $\mu_a$ ,  $\mu'_s$  and  $D$  are all positive [5]. It is common practice to consider only one temporal frequency  $\omega \geq 0$  at a time. The response to a complex sinusoid  $S(\mathbf{r},t) = S(\mathbf{r})e^{-j\omega t}$  is also a complex sinusoid  $\Phi(\mathbf{r},t) = \Phi(\mathbf{r})e^{-j\omega t}$  of the same frequency. Actual sources are real sinusoids, as are their responses, but in this case it is simpler to consider real sinusoids as complex sinusoids projected on the real line. That means, the actual source is  $\Re(S(\mathbf{r})e^{-j\omega t})$ , and the actual response is  $\Re(\Phi(\mathbf{r})e^{-j\omega t})$ . Substituting the expressions for the complex sinusoids in equation (2.1) we get

$$-D(\mathbf{r})(\nabla^2\Phi(\mathbf{r}))e^{-j\omega t} + v\mu_a(\mathbf{r})\Phi(\mathbf{r})e^{-j\omega t} - j\omega\Phi(\mathbf{r})e^{-j\omega t} = vS(\mathbf{r})e^{-j\omega t}. \quad (2.3)$$

Canceling out  $e^{-j\omega t}$  on both sides we get

$$-D(\mathbf{r})\nabla^2\Phi(\mathbf{r}) + v\mu_a(\mathbf{r})\Phi(\mathbf{r}) - j\omega\Phi(\mathbf{r}) = vS(\mathbf{r}). \quad (2.4)$$

For an homogeneous medium,  $D(\mathbf{r}) = D_0$  and  $\mu_a(\mathbf{r}) = \mu_{a0}$  are constants over space. In this case, equation (2.4) reduces to a Helmholtz equation with complex wavenumber

$$\nabla^2 \Phi(\mathbf{r}) + k_0^2 \Phi(\mathbf{r}) = -\frac{v}{D_0} S(\mathbf{r}), \quad (2.5)$$

where

$$k_0^2 = \frac{-v\mu_{a0} + j\omega}{D_0}. \quad (2.6)$$

The solution to this equation, which we will call incident or homogeneous solution  $\Phi^i(\mathbf{r})$ , can be obtained by means of the Green's function as

$$\Phi^i(\mathbf{r}) = \int_V G(\mathbf{r}, \mathbf{r}') \left( -\frac{v}{D_0} S(\mathbf{r}') \right) d\mathbf{r}', \quad (2.7)$$

where the Green's function is the solution of

$$[\nabla^2 + k_0^2]G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'), \quad (2.8)$$

for specific boundary conditions. In equation (2.8),  $\delta$  represents the dirac delta function. In other words, the Green's function contains the information of the solutions of equation (2.5) for every point source  $-\frac{v}{D_0} S(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}')$  with location  $\mathbf{r}' = (x', y', z')$  in the domain of integration  $V$ . Here we use the term source to indicate the right hand side of equation (2.5).

In the case of an infinite medium, i.e. no boundaries, the Green's function can be expressed as [48]:

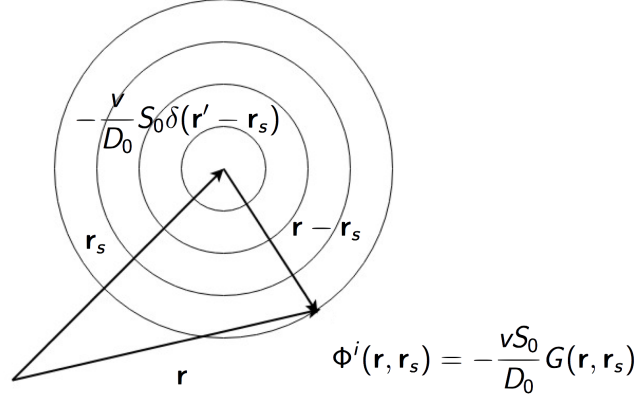
$$G(\mathbf{r}, \mathbf{r}') = -\frac{e^{jk_0\|\mathbf{r}-\mathbf{r}'\|}}{4\pi\|\mathbf{r}-\mathbf{r}'\|} = \frac{e^{-\Im(k_0)\|\mathbf{r}-\mathbf{r}'\|}}{4\pi\|\mathbf{r}-\mathbf{r}'\|} \left( -e^{j\Re(k_0)\|\mathbf{r}-\mathbf{r}'\|} \right) \quad (2.9)$$

where

$$\begin{aligned} k_0 &= \sqrt{\alpha} e^{j(\beta/2)}, \\ \alpha &= |k_0^2|, \\ \beta &= \arg(k_0^2). \end{aligned} \quad (2.10)$$

Notice that  $\Im(k_0) > 0$  and  $\Re(k_0) \geq 0$ . Therefore, from equation (2.9), the real and imaginary parts of the infinite medium Green's function for a particular  $\mathbf{r}'$  are damped sinusoidal waves emanating from  $\mathbf{r}'$ . Figure 2.1 illustrates the solution of the equation (2.5) when the isotropic source term is a punctual isotropic source term,  $S(\mathbf{r}) = S_0 \delta(\mathbf{r} - \mathbf{r}_s) = |S_0| e^{j \arg(S_0)} \delta(\mathbf{r} - \mathbf{r}_s)$ , with magnitude  $|S_0|$ , phase  $\arg(S_0)$  and location  $\mathbf{r}_s = (x_s, y_s, z_s)$ , and the medium is infinite. The solution is basically the same as the  $G(\mathbf{r}, \mathbf{r}_s)$ , apart from a complex number, which basically has the effect of rescaling and shifting the damped

sinusoidal waves that are the real and imaginary parts of  $G(\mathbf{r}, \mathbf{r}_s)$ . In the figure, only the real or the imaginary part is represented.



**Figure 2.1.** Illustration of the infinite medium solution when the source is an arbitrary point source  $-\frac{v}{D_0}S(\mathbf{r}) = -\frac{v}{D_0}S_0\delta(\mathbf{r} - \mathbf{r}_s)$ . In this case, the solution is a rescaled and phase shifted version of  $G(\mathbf{r}, \mathbf{r}_s)$ , which has damped sinusoidal waves as real and imaginary parts.

In the case of a semi-infinite medium, with the  $x - y$  plane as the boundary ( $z = 0$ ) and the top half space as the medium ( $z > 0$ ), the Robin boundary condition

$$\Phi = z_b \frac{\partial \Phi}{\partial z} \quad \text{at} \quad z = 0, \quad (2.11)$$

is usually approximated by the Dirichlet boundary condition

$$\Phi = 0 \quad \text{at} \quad z = -z_b, \quad (2.12)$$

where

$$z_b = \frac{2}{3\mu'_s} \left( \frac{1 + R_{eff}}{1 - R_{eff}} \right) \quad (2.13)$$

is called the extrapolated boundary and  $R_{eff}$  is the effective reflection coefficient from inside the medium [29]. This boundary condition can be obtained by considering an infinite medium and implementing the method of images. In this method every source gets a duplicate with opposite sign placed at its mirror image with respect to the plane of zero field [36]. Doing that to equation (2.8) we get

$$[\nabla^2 + k_0^2]G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') - \delta(\mathbf{r} - \mathbf{r}_i'), \quad (2.14)$$

and the semi-infinite medium Green's function becomes

$$G(\mathbf{r}, \mathbf{r}') = \frac{-1}{4\pi\|\mathbf{r} - \mathbf{r}'\|} e^{jk_0\|\mathbf{r} - \mathbf{r}'\|} + \frac{1}{4\pi\|\mathbf{r} - \mathbf{r}'_i\|} e^{jk_0\|\mathbf{r} - \mathbf{r}'_i\|} \quad (2.15)$$

where  $\mathbf{r}'_i = (x'_i, y'_i, z'_i)$  is the location of the source image, and

$$\begin{aligned} x'_i &= x' \\ y'_i &= y' \\ z'_i &= -z' - 2z_b \end{aligned} \quad (2.16)$$

are the coordinates of the source image.

When there is a perturbation in absorption coefficient in an otherwise homogeneous medium, i.e. when  $\mu_a = \mu_{a0} + \Delta\mu_a(\mathbf{r})$ , equation (2.4) reduces to

$$[\nabla^2 + k_0^2 + \Delta k^2(\mathbf{r})]\Phi(\mathbf{r}) = -\frac{v}{D_0}S(\mathbf{r}), \quad (2.17)$$

where

$$\Delta k^2(\mathbf{r}) = \left( \frac{-v\Delta\mu_a(\mathbf{r})}{D_0} \right). \quad (2.18)$$

We define the heterogeneous or scattered field as  $\Phi^s(\mathbf{r}) = \Phi(\mathbf{r}) - \Phi^i(\mathbf{r})$ . In other words,  $\Phi^s(\mathbf{r})$  is the difference in the fluence rate caused by the presence of the perturbation. Recall that  $\Phi^i(\mathbf{r})$  is the solution of equation (2.5). If we replace  $\Phi(\mathbf{r}) = \Phi^s(\mathbf{r}) + \Phi^i(\mathbf{r})$  in equation (2.17) we get

$$\begin{aligned} [\nabla^2 + k_0^2 + \Delta k^2(\mathbf{r})](\Phi^s(\mathbf{r}) + \Phi^i(\mathbf{r})) &= -\frac{v}{D_0}S(\mathbf{r}), \\ [\nabla^2 + k_0^2](\Phi^s(\mathbf{r}) + \Phi^i(\mathbf{r})) + \Delta k^2(\mathbf{r})(\Phi^s(\mathbf{r}) + \Phi^i(\mathbf{r})) &= -\frac{v}{D_0}S(\mathbf{r}), \\ [\nabla^2 + k_0^2]\Phi^s(\mathbf{r}) - \frac{v}{D_0}S(\mathbf{r}) + \Delta k^2(\mathbf{r})(\Phi^s(\mathbf{r}) + \Phi^i(\mathbf{r})) &= -\frac{v}{D_0}S(\mathbf{r}), \\ [\nabla^2 + k_0^2]\Phi^s(\mathbf{r}) + \Delta k^2(\mathbf{r})(\Phi^s(\mathbf{r}) + \Phi^i(\mathbf{r})) &= 0, \\ [\nabla^2 + k_0^2]\Phi^s(\mathbf{r}) &= -\Delta k^2(\mathbf{r})(\Phi^s(\mathbf{r}) + \Phi^i(\mathbf{r})), \end{aligned} \quad (2.19)$$

which is similar to equation (2.5), except for the right hand side.  $\Phi^s(\mathbf{r})$  can then be obtained using the same Green's function used to obtain  $\Phi^i(\mathbf{r})$  by

$$\Phi^s(\mathbf{r}) = \int_V G(\mathbf{r}, \mathbf{r}') (-\Delta k^2(\mathbf{r}')(\Phi^s(\mathbf{r}') + \Phi^i(\mathbf{r}'))) d\mathbf{r}'. \quad (2.20)$$

In this equation, the relation between  $\Delta k^2(\mathbf{r})$  and  $\Phi^s(\mathbf{r})$  is not linear, but it can be considered approximately linear when  $\Phi^i(\mathbf{r}) \gg \Phi^s(\mathbf{r})$ , which happens when the perturbation  $\Delta k^2(\mathbf{r})$  is not large in magnitude and support. In this case, the scattered field can be

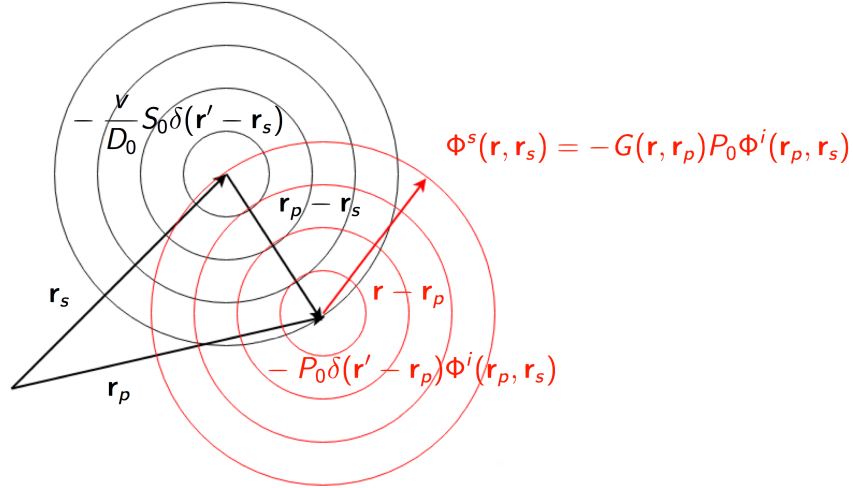
approximated by

$$\Phi^s(\mathbf{r}) = \int_V G(\mathbf{r}, \mathbf{r}') (-\Delta k^2(\mathbf{r}') \Phi^i(\mathbf{r}')) d\mathbf{r}', \quad (2.21)$$

which expresses a linear relationship between  $\Delta k^2(\mathbf{r})$  and  $\Phi^s(\mathbf{r})$ . This is known as the first Born approximation [29]. Notice that the first Born approximation is the solution of the equation

$$[\nabla^2 + k_0^2] \Phi^s(\mathbf{r}) = -\Delta k^2(\mathbf{r}) \Phi^i(\mathbf{r}). \quad (2.22)$$

Therefore, by an analogy with the equation (2.5), the term  $-\Delta k^2(\mathbf{r}) \Phi^i(\mathbf{r})$  can be seen as the source of the approximate scattered field. Figure 2.2 illustrates the scattered field estimated using the first Born approximation when the incident field of Figure 2.1 reaches an arbitrary impulse perturbation  $\Delta k^2(\mathbf{r}) = P_0 \delta(\mathbf{r} - \mathbf{r}_p)$  located at  $\mathbf{r}_p = (x_p, y_p, z_p)$ . The characteristics of the scattered wave are similar to the incident wave, but with different amplitude, phase and point of origin.



**Figure 2.2.** Illustration of the first Born approximation of the scattered field when the incident field of Figure 2.1 reaches an arbitrary impulse perturbation  $\Delta k^2(\mathbf{r}) = P_0 \delta(\mathbf{r} - \mathbf{r}_p)$ . In the first Born approximation, the scattered field can be seen as another homogeneous medium solution, with  $-\Delta k^2(\mathbf{r}) \Phi^i(\mathbf{r}, \mathbf{r}_s)$  as the source. Therefore, the scattered wave emanating from an impulse perturbation is similar to the incident wave of Figure 2.1, but with different amplitude, phase and point of origin.

In diffuse optical tomography, several sources and detectors are placed at a tissue boundary. In RF DOT systems, the sources are amplitude modulated at megahertz frequencies. Usually, only one source is turned on at a time while the amplitude and phase of the diffused light is measured at the detector locations. By the end one gets a series of fluence rate amplitude and phase measurements at the detector locations for several sources that were turned on sequentially. Each source is usually modeled as a

point source one transport mean free path length inside the medium [45], where the transport mean free path length is defined as  $l_{tr} = 1/\mu'_s$ .

In the case that the medium can be approximated by an homogeneous semi-infinite medium with a perturbation in absorption coefficient that is small in magnitude and support, the imaging problem can be considerably simplified. The medium can be approximated by a semi-infinite medium if the boundary is approximately planar where the sources and detectors are placed and if the other boundaries of the tissue are sufficiently far away. From equation (2.7), the incident field for a punctual isotropic source term  $S(\mathbf{r}) = S_0\delta(\mathbf{r} - \mathbf{r}_s)$  can be obtained by

$$\Phi^i(\mathbf{r}, \mathbf{r}_s) = -\frac{v}{D_0} \int_V G(\mathbf{r}, \mathbf{r}') S_0 \delta(\mathbf{r}' - \mathbf{r}_s) d\mathbf{r}' = -\frac{v}{D_0} S_0 G(\mathbf{r}, \mathbf{r}_s), \quad (2.23)$$

where the dependency of  $\Phi^i(\mathbf{r})$  on the location of the source was made explicit. The field measured at a particular detector is a superposition of the incident field coming from the source and the scattered field coming from the perturbations inside the medium. Therefore, if the field measured at  $\mathbf{r}_d$  for a source turned on at  $\mathbf{r}_s$  is  $\Phi(\mathbf{r}_d, \mathbf{r}_s)$ , the scattered field at  $\mathbf{r}_d$  for a source turned on at  $\mathbf{r}_s$  is  $\Phi^s(\mathbf{r}_d, \mathbf{r}_s) = \Phi(\mathbf{r}_d, \mathbf{r}_s) - \Phi^i(\mathbf{r}_d, \mathbf{r}_s)$ . From equation (2.21), the scattered field  $\Phi^s(\mathbf{r}_d, \mathbf{r}_s)$  is related linearly to  $\Delta k^2(\mathbf{r})$  by the equation

$$\Phi^s(\mathbf{r}_d, \mathbf{r}_s) = \int_V G(\mathbf{r}_d, \mathbf{r}') (-\Delta k^2(\mathbf{r}') \Phi^i(\mathbf{r}', \mathbf{r}_s)) d\mathbf{r}'. \quad (2.24)$$

If we assume that the Green's function, fluence rate and perturbation in squared wavenumber are constant over equal sized cubic voxels that divide the medium, this integral can be approximated by the summation

$$\Phi^s(\mathbf{r}_d, \mathbf{r}_s) = \sum_{\mathbf{r}_c \in C} G(\mathbf{r}_d, \mathbf{r}_c) (-\Delta k^2(\mathbf{r}_c) \Phi^i(\mathbf{r}_c, \mathbf{r}_s)) h^3, \quad (2.25)$$

where  $\mathbf{r}_c$  is a voxel center,  $C$  is the set of voxels centers, and  $h$  is the voxel side. Grouping the scattered fields for the different sources and detectors in a vector  $\mathbf{b}$ , stacking the perturbation in squared wavenumber for the different voxels in a vector  $\mathbf{x}$ , using equation (2.25) to relate vector  $\mathbf{x}$  to each element of vector  $\mathbf{b}$ , we get the system of linear equations

$$\begin{aligned} \mathbf{b} &= \mathbf{A}\mathbf{x} \\ \left\{ \begin{array}{l} a_{m,n} = -G(D_t, C_n) \Phi^i(C_n, S_q) h^3 \\ b_m = \Phi^s(D_t, S_q) \\ x_n = \Delta k^2(C_n) \end{array} \right. & \quad (2.26) \\ m &= qT + t \end{aligned}$$



where  $D$  and  $S$  are the sets of detector locations and source locations, respectively, and  $t \in \{0, 1, \dots, T - 1\}$ ,  $q \in \{0, 1, \dots, Q - 1\}$ ,  $n \in \{0, 1, \dots, N - 1\}$  and  $m \in \{0, 1, \dots, M - 1\}$  are the indexes for the set of detector locations, the set source locations, the set of voxels centers  $C$  and the vector of measurements  $\mathbf{b}$ , respectively. The relation  $m = qT + t$  simply means that the measurements were organized such that all detector measurements for every source were placed together. The way the measurements are organized is actually irrelevant, as long as the lines of the matrix  $\mathbf{A}$  are organized accordingly. Similarly, the columns of  $\mathbf{A}$  must be placed the same way as their respective voxels in vector  $\mathbf{x}$ . The system of equations (2.26) provides a means of recovering three-dimensional images of perturbation in squared wavenumber,  $\mathbf{x}$ , from estimates of scattered fluence rate,  $\mathbf{b}$ , obtained from measurements of fluence rate, using linear inversion methods such as truncated SVD and Tikhonov regularization.

### 2.1.2 Depth Compensation

Although already useful, the system of equations (2.26), together with linear inversion methods such as SVD, Tikhonov or  $\ell_1$ -minimization, tend to underestimate the actual depth of perturbations [4, 49, 65, 64]. To appreciate why, assume that the voxels are ordered in vector  $\mathbf{x}$  from smallest to largest  $z$ -coordinate. It does not matter how voxels with same  $z$ -coordinate, here called collectively as a layer, are ordered. Then matrix  $\mathbf{A}$  can be decomposed as

$$\mathbf{A} = \begin{bmatrix} \mathbf{W}_0 & \mathbf{W}_1 & \cdots & \mathbf{W}_{L-1} \end{bmatrix} \quad (2.27)$$

where  $\mathbf{W}_\ell$  is the submatrix of  $\mathbf{A}$  that multiplies the  $\ell$ -th layer of  $\mathbf{x}$ . The order of magnitude of the elements of  $\mathbf{W}_\ell$  decreases as  $\ell$  increases. This is due to the increase in distance between voxels and sources and between voxels and detectors as the  $z$ -coordinate of the layer increases, and this is what causes the recovered images to be concentrated in the first layers. Also the largest singular value of  $\mathbf{W}_\ell$ , here denoted as  $\sigma_\ell$ , decreases as  $\ell$  increases. Therefore, one of the envisioned ways to work around this issue was to multiply  $\mathbf{W}_\ell$  by a power of the largest singular value of layer  $L - 1 - \ell$ , this way the decay in order of magnitude would be counterbalanced [4, 49, 65, 64]. In other words, we should multiply  $\mathbf{A}$  to the right by

$$\begin{bmatrix} \sigma_{L-1}^\gamma \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma_{L-2}^\gamma \mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \sigma_0^\gamma \mathbf{I} \end{bmatrix} \quad (2.28)$$

where  $\gamma$  is an adjustable real number,  $\mathbf{I}$  is the identity matrix of size  $(\frac{N}{L})$ , and  $\mathbf{0}$  is the matrix of zeros of the same size. However, since we modified the matrix  $\mathbf{A}$  but did

not modify the measurements, the images recovered using the modified linear system of equations are in a different domain than the desired domain. In order to convert the recovered images to the desired domain, they must be multiplied to the left by the matrix (2.28).

### 2.1.3 Angular Spectrum Representation of Diffuse Photon Density Waves

Consider equation (2.21) again. Expanding it for the semi-infinite medium we get

$$\Phi^s(x, y, z) = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty G(x, y, z, x', y', z') A(x', y', z') dx' dy' dz', \quad (2.29)$$

where

$$A(x, y, z) = -\Delta k^2(x, y, z) \Phi^i(x, y, z) \quad (2.30)$$

and

$$G(x, y, z, x', y', z') = -\frac{e^{jk_0\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}}}{4\pi\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}} + \frac{e^{jk_0\sqrt{(x-x')^2+(y-y')^2+(z+z'+2z_b)^2}}}{4\pi\sqrt{(x-x')^2+(y-y')^2+(z+z'+2z_b)^2}}. \quad (2.31)$$

From now on, we will represent the Green's function by  $G(x-x', y-y'; z, z')$  instead of  $G(x, y, z, x', y', z')$ . According to equation (2.29), the scattered field at the boundary of the semi-infinite medium is

$$\Phi^s(x, y, 0) = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty G(x-x', y-y'; 0, z') A(x', y', z') dx' dy' dz'. \quad (2.32)$$

If we take the Fourier transform of this equation sequentially in the x and y dimensions and use the convolution theorem we get

$$\begin{aligned} \tilde{\Phi}^s(\omega_x, y, 0) &= \int_0^\infty \int_{-\infty}^\infty \left[ \int_{-\infty}^\infty \left( \int_{-\infty}^\infty G(x-x', y-y'; 0, z') A(x', y', z') dx' \right) e^{-j\omega_x x} dx \right] dy' dz' \\ &= \int_0^\infty \int_{-\infty}^\infty \tilde{G}(\omega_x, y-y'; 0, z') \tilde{A}(\omega_x, y', z') dy' dz', \\ \tilde{\Phi}^s(\omega_x, \omega_y, 0) &= \int_0^\infty \left[ \int_{-\infty}^\infty \left( \int_{-\infty}^\infty \tilde{G}(\omega_x, y-y'; 0, z') \tilde{A}(\omega_x, y', z') dy' \right) e^{-j\omega_y y} dy \right] dz' \\ &= \int_0^\infty \tilde{G}(\omega_x, \omega_y; 0, z') \tilde{A}(\omega_x, \omega_y, z') dz', \end{aligned} \quad (2.33)$$

which is called the angular spectrum representation of the scattered field  $\Phi^s(x, y, 0)$ . Equation (2.33) is the main equation of diffraction tomography with diffuse photon density waves [30], one of the earliest approaches to diffuse optical tomography.

## 2.2 TIKHONOV REGULARIZATION

This section gives a brief overview of the most common method in DOT: Tikhonov regularization [40]. Consider a general linear forward model

$$\mathbf{b}_\eta = \mathbf{A}\mathbf{x} + \mathbf{n}, \quad (2.34)$$

where  $\mathbf{x} \in \mathbb{C}^n$  is a vector of interest,  $\mathbf{b}_\eta \in \mathbb{C}^m$  is a vector of observations contaminated with noise,  $\mathbf{n} \in \mathbb{C}^m$  is the noise vector, and  $\mathbf{A} \in \mathbb{C}^{m \times n}$  is the forward model matrix. When  $m > n$ , the vector  $\mathbf{x}$  is frequently approximated by the least squares solution, which can be described as the minimizer of the optimization problem

$$\min_{\hat{\mathbf{x}} \in \mathbb{C}^n} (\|\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}_\eta\|_2^2). \quad (2.35)$$

However, if the number of equations is not large enough with respect to the number of unknowns, this method results in overfitting [50]. One way to prevent overfitting is to add a term to the objective function to penalize solutions with large  $\ell_2$  norm. This is known as Tikhonov regularization. The Tikhonov regularized least squares solution can be described as the minimizer of the problem

$$\min_{\hat{\mathbf{x}} \in \mathbb{C}^n} (\|\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}_\eta\|_2^2 + \lambda^2 \|\hat{\mathbf{x}}\|_2^2), \quad (2.36)$$

where  $\lambda > 0$  is the Tikhonov regularization parameter. The solution to this problem can be calculated by

$$\mathbf{x}_\lambda = (\mathbf{A}^H \mathbf{A} + \lambda^2 \mathbf{I})^{-1} \mathbf{A}^H \mathbf{b}_\eta. \quad (2.37)$$

The Tikhonov regularization parameter has to be chosen for each specific problem. Some methods for choosing the regularization parameter are the L-curve method and the cross validation method [75][23].

## 2.3 COMPRESSIVE SENSING

Consider the underdetermined linear system of equations

$$\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} \quad (2.38)$$

where  $\mathbf{b} \in \mathbb{C}^m$ ,  $\hat{\mathbf{x}} \in \mathbb{C}^n$ ,  $\mathbf{A} \in \mathbb{C}^{m \times n}$  and  $m \ll n$ . Solving this system for  $\hat{\mathbf{x}}$  is an ill-posed problem since there are infinitely many solutions possible. However, suppose that the desired solution  $\mathbf{x}$  has no more than  $s$  nonzero entries, or, in other words, that  $\mathbf{x}$  is  $s$ -sparse. In some situations it is possible to find  $\mathbf{x}$  by searching for the sparsest solution

possible. What are such situations is one of the main concerns of compressive sensing, a research field that first saw the light of day in the 21<sup>st</sup> century.

The field of compressive sensing is essentially concerned with two questions [34, 24]:

- What matrices  $\mathbf{A}$  allow us to recover the  $s$ -sparse vector  $\mathbf{x}$  by searching for the sparsest solution?
- What algorithms can we use to search for the sparsest solution?

In the next sections we will provide a group of matrices  $\mathbf{A}$ , and one algorithm, that will allow us to recover  $\mathbf{x}$  from  $\mathbf{b} = \mathbf{A}\mathbf{x}$  when  $m \ll n$ . We will use the fact that minimizing the  $\ell_1$  norm, defined as  $\|\hat{\mathbf{x}}\|_1 = \sum_{i=0}^{n-1} |\hat{x}_i|$ , or the  $\ell_p$  quasinorms, defined as  $\|\hat{\mathbf{x}}\|_p = (\sum_{i=0}^{n-1} |\hat{x}_i|^p)^{1/p}$  for  $0 < p < 1$ , tend to coincide with minimizing the  $\ell_0$  function, defined as  $\|\hat{\mathbf{x}}\|_0 = \sum_{\hat{x}_i \neq 0} |\hat{x}_i|^0$ , which is the same as searching for the sparsest solution possible [24].

### 2.3.1 Incoherent Sampling Theorem

Let  $\mathbf{a}$  be a random column vector on  $\mathbb{C}^n$  with probability distribution  $F$ , that is  $\mathbf{a} = [a_0 \ a_1 \ \dots \ a_{n-1}]^\top \sim F$ , and let  $\mathbb{E}[\mathbf{a}\mathbf{a}^H] = \mathbf{I}$ . Define the coherence parameter  $\mu(F)$  as the smallest number such that  $\max_{0 \leq i \leq n-1} |a_i|^2 \leq \mu(F)$  holds either deterministically or stochastically [13]. Let  $\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_{m-1}$  be a sequence of realizations of  $\mathbf{a}$ , which we will call sensing vectors, and let  $\mathbf{A}$  be the matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_0^H \\ \mathbf{a}_1^H \\ \vdots \\ \mathbf{a}_{m-1}^H \end{bmatrix}, \quad (2.39)$$

which we will call sensing matrix. Then the incoherent sampling theorem reads:

**Theorem 2.3.1** [13, 12] *Let  $\mathbf{x}$  be a fixed but otherwise arbitrary  $s$ -sparse vector in  $\mathbb{C}^n$ . Assume that the sensing vectors are isotropic ( $\mathbb{E}[\mathbf{a}\mathbf{a}^H] = \mathbf{I}$ ) and let  $\mathbf{b} = \mathbf{A}\mathbf{x}$  be the data vector. Pick any scalar  $\beta > 0$ . If the number of equations obeys*

$$m \geq C_\beta \cdot \mu(F) \cdot s \cdot \log n \quad (2.40)$$

*then  $\mathbf{x}$  is the unique minimizer to  $\min_{\hat{\mathbf{x}} \in \mathbb{C}^n} \|\hat{\mathbf{x}}\|_1$  s.t.  $\mathbf{b} = \mathbf{A}\hat{\mathbf{x}}$  with probability at least  $1 - 5/n - e^{-\beta}$ . Further,  $C_\beta$  may be chosen as  $C_0(1 + \beta)$  for some numerical constant  $C_0$ .*

One example of situation in which the isotropy condition is satisfied is when the rows

of matrix  $\mathbf{A}$  are selected from the rows of a matrix  $\sqrt{n}\mathbf{U}$ , where  $\mathbf{U}$  is unitary [14]. The isotropy in this situation can be proven by the following argument: suppose all rows are selected such that  $\mathbf{A} = \sqrt{n}\mathbf{U}$ ; in this case  $\mathbf{A}^H\mathbf{A}/n = \mathbb{E}[\mathbf{a}\mathbf{a}^H]$ ; since also  $\mathbf{A}^H\mathbf{A}/n = (\sqrt{n}\mathbf{U})^H(\sqrt{n}\mathbf{U})/n = \mathbf{U}^H\mathbf{U} = \mathbf{I}$ , then  $\mathbb{E}[\mathbf{a}\mathbf{a}^H] = \mathbf{I}$ . The coherence parameter in this case is  $\mu(F) = \max_{0 \leq i, j \leq n-1} |\sqrt{n} u_{i,j}|^2 = n \max_{0 \leq i, j \leq n-1} |u_{i,j}|^2$ . From the theorem it is clear that the smaller  $\mu(F)$  the smaller the number of equations needed. In this case it reaches its minimum  $\mu(F) = 1$  when the rows of  $\mathbf{U}$  are maximally spread, that is when all its elements have absolute value equal  $1/\sqrt{n}$ , and reaches its maximum  $\mu(F) = n$  when the rows of  $\mathbf{U}$  are minimally spread, that is when all but one of its elements have absolute value different from zero and equal to 1. Therefore the theorem requires that the sensing vectors be spread so that the number of equations can be low.

In a particular application,  $\mathbf{U}$  is further decomposed as the product of two other unitary matrices,

$$\Phi = \begin{bmatrix} \phi_0^H \\ \phi_1^H \\ \vdots \\ \phi_{n-1}^H \end{bmatrix} \quad (2.41)$$

and

$$\Psi = \begin{bmatrix} \psi_0 & \psi_1 & \cdots & \psi_{n-1} \end{bmatrix} \quad (2.42)$$

such that  $\mathbf{U} = \Phi\Psi$ , where the set of vectors  $\{\phi_0, \phi_1, \dots, \phi_{n-1}\}$  is called the sensing basis, and the set of vectors  $\{\psi_0, \psi_1, \dots, \psi_{n-1}\}$  is called the sparsifying basis [14]. In this case the coherence parameter can also be calculated as  $\mu(F) = n \max_{i,j} |\langle \phi_i, \psi_j \rangle|^2$ , which is small when there is little similarity between the sensing basis vectors and the sparsifying basis vectors, and is large when there is considerable similarity between them.

### 2.3.2 Restricted Isometry Property

Theorem 2.3.1 gives conditions under which it is possible to recover exactly sparse signals. However, in most applications, the signals we are interested in are only approximately sparse. The essential questions of compressive sensing then becomes: what matrices  $\mathbf{A}$  allow us to *accurately* recover the *approximately*  $s$ -sparse vector  $\mathbf{x}$  by searching for the sparsest solution possible, and what algorithms can we use to search for the sparsest solution.

Even though there is an adaptation of Theorem 2.3.1 for approximately sparse signals, it assumes that the signals are real-valued [13]. In general we are interested in recovering signals that are complex-valued. Another way of addressing the questions of compressive sensing is by means of the restricted isometry property (RIP) [24]. Assume that the columns of  $\mathbf{A}$  had been normalized to have unit  $\ell_2$ -norm. The matrix  $\mathbf{A}$  is said to satisfy the RIP of order  $s$  with parameter  $\delta_s$  if there is some  $\delta_s \in (0, 1)$  such that

$$(1 - \delta_s)\|\hat{\mathbf{x}}\|_2^2 \leq \|\mathbf{A}\hat{\mathbf{x}}\|_2^2 \leq (1 + \delta_s)\|\hat{\mathbf{x}}\|_2^2 \quad (2.43)$$

for every  $s$ -sparse vector  $\hat{\mathbf{x}}$ . The following theorem gives conditions under which accurate recovery of approximately  $s$ -sparse complex-valued signals is possible using  $\ell_1$ -minimization.

**Theorem 2.3.2** [3] *Suppose that  $\mathbf{A}$  satisfies the RIP of order  $2s$  with  $\delta_{2s} \leq 0.3$ . Then the solution  $\mathbf{x}_{\ell_1}$  to  $\min_{\hat{\mathbf{x}} \in \mathbb{C}^n} \|\hat{\mathbf{x}}\|_1$  s.t.  $\mathbf{b} = \mathbf{A}\hat{\mathbf{x}}$  satisfies*

$$\|\mathbf{x}_{\ell_1} - \mathbf{x}\|_2^2 \leq C_0 \frac{\|\mathbf{x} - \mathbf{x}_s\|_1^2}{s} \quad (2.44)$$

where  $\mathbf{x}_s$  is the vector obtained by setting all but the  $s$  largest entries of  $\mathbf{x}$  to zero, and  $C_0 > 0$  is a constant given by

$$C_0 = 4 \left( \frac{1 + \delta_{2s}}{1 - 3\delta_{2s}} \right)^2 \quad (2.45)$$

Similarly to when we discussed the Incoherent Sampling Theorem, matrices constructed by selecting rows of a unitary matrix also satisfy the RIP with high probability, provided that some conditions are met. These conditions are given in Theorem 2.3.3.

**Theorem 2.3.3** [3] *Let  $\mathbf{U}$  be any  $n \times n$  unitary matrix. Choose a subset  $\Omega$  of cardinality  $m \stackrel{\text{def}}{=} |\Omega|$  uniformly at random from the set  $\{0, \dots, n-1\}$ . Further, let  $\mathbf{A}$  be the  $m \times n$  matrix obtained by sampling  $m$  rows of  $\mathbf{U}$  corresponding to the indices in  $\Omega$  and renormalizing the resulting columns so they have unit  $\ell_2$ -norms. For each integer  $n$ ,  $s > 2$ , and for any  $t > 1$  and any  $\delta_s \in (0, 1)$ , let*

$$m \geq (C_3 \cdot \mu_U^2 \cdot t \cdot s \cdot \log n) \cdot \log(t \cdot s \cdot \log n) \cdot \log^2 s \quad (2.46)$$

then the subsampled matrix  $\mathbf{A}$  satisfies the RIP of order  $s$  with parameter  $\delta_s$  with probability exceeding  $1 - 10e^{-C_4 \delta_s^2 t}$ . Here  $\mu_U \stackrel{\text{def}}{=} \sqrt{n} \max_{i,j} |u_{i,j}|$  is termed the coherence of the unitary matrix  $\mathbf{U}$ , and  $C_3, C_4 > 0$  are absolute constants that do not depend on  $m$ ,  $n$ , or  $s$ .

### 2.3.3 Iteratively Reweighted Least Squares

The iteratively reweighted least squares (IRLS) algorithm [17, 60] uses the result that the minimizer of the problem

$$\begin{cases} \min_{\hat{\mathbf{x}} \in \mathbb{C}^n} \hat{\mathbf{x}}^H \mathbf{W} \hat{\mathbf{x}} \\ \text{s.t. } \mathbf{b} = \mathbf{A} \hat{\mathbf{x}} \end{cases}, \quad (2.47)$$

where  $\mathbf{W}$  is a diagonal matrix, is

$$\mathbf{x}_w = \mathbf{W}^{-1} \mathbf{A}^H (\mathbf{A} \mathbf{W}^{-1} \mathbf{A}^H)^{-1} \mathbf{b}. \quad (2.48)$$

Suppose an hypothetical case in which the diagonal elements of  $\mathbf{W}$  could be chosen as  $\mathbf{W}_{ii} = |\hat{x}_i|^{p-2}$ , for  $p > 0$ . Then problem (2.47) would be equivalent to

$$\begin{cases} \min_{\hat{\mathbf{x}} \in \mathbb{C}^n} \|\hat{\mathbf{x}}\|_p^p \\ \text{s.t. } \mathbf{b} = \mathbf{A} \hat{\mathbf{x}} \end{cases}. \quad (2.49)$$

Even though we cannot choose  $\mathbf{W}_{ii} = |\hat{x}_i|^{p-2}$ , we can choose  $\mathbf{W}_{ii} = |\hat{x}_i^0|^{p-2}$ , where  $\hat{\mathbf{x}}^0$  is an approximation to  $\hat{\mathbf{x}}$ . We can then start an iterative process in which at iteration  $k > 1$  we choose  $\mathbf{W}_{ii} = |\hat{x}_i^{k-1}|^{p-2}$ , where  $\hat{\mathbf{x}}^{k-1}$  is the solution found in the previous iteration. This process can be shown to converge. In practice, instead of (2.48), one computes

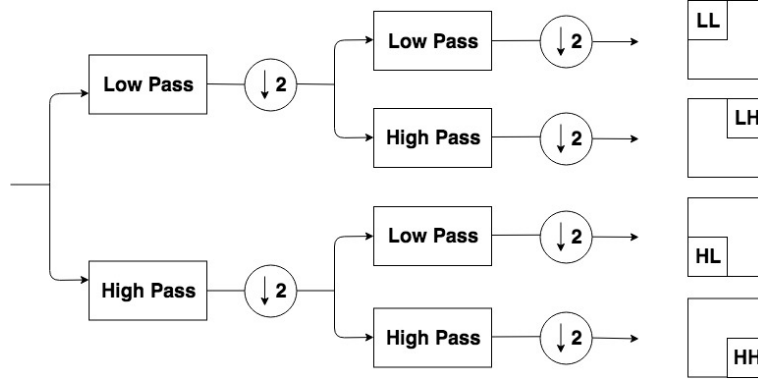
$$\mathbf{x}_w = \mathbf{Q} \mathbf{A}^H (\mathbf{A} \mathbf{Q} \mathbf{A}^H)^{-1} \mathbf{b}. \quad (2.50)$$

where  $\mathbf{Q}_{ii} = |\hat{x}_i^{k-1}|^{2-p}$ . Since  $\hat{\mathbf{x}}^{k-1}$  is expected to have many elements close to zero, which may be a problem for the inversion in (2.50), one starts by adding some relatively large regularization parameter  $\mu$  to every element in the diagonal of  $\mathbf{Q}$ , then reduces  $\mu$  every time the algorithm converges and the regularization parameter becomes less needed [60].

## 2.4 PREFILTERING METHOD

The prefiltering method is an image reconstruction technique built upon the theory of compressive sensing that uses concepts of filterbanks to promote signal recovery [61]. An analysis filterbank is a signal processing device that decomposes a signal into other signals, each carrying information from one frequency sub-band of the original signal [86]. Figure 2.3 shows the basic structure of a typical separable two-dimensional analysis filterbank, used in image processing. This filterbank consists of four branches. In both the first and the second branches, the columns of the input image are low pass filtered, with the rows of the resulting image being low pass filtered in the first branch and high pass

filtered in the second. Conversely, in both the third and fourth branches the columns of the input image are high pass filtered, with the rows of the resulting image being low pass filtered in the third branch and high pass filtered in the fourth. For this reason, these branches are labeled respectively LL, LH, HL, and HH. After filtering one dimension, that dimension is also downsampled by a factor of two, in this filterbank. Figure 2.3 also shows the sub-bands of the original image spectrum ideally isolated in each branch.



**Figure 2.3.** Typical separable two-dimensional analysis filterbank, with four branches. To the right, we represent the sub-bands ideally isolated in each branch.

The low pass filter and the high pass filter of Figure 2.3 are usually chosen so that the decomposition is invertible. A common choice of low pass filter and high pass filter are the Haar analysis filters  $\mathbf{h}_0 = [1 \ 1]^\top$  and  $\mathbf{h}_1 = [1 \ -1]^\top$ . The same decomposition can also be obtained by convolving the input image with each of the filters

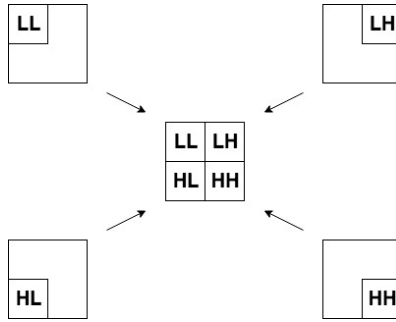
$$\begin{aligned} \mathbf{h}_0 \mathbf{h}_0^\top &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, & \mathbf{h}_0 \mathbf{h}_1^\top &= \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \\ \mathbf{h}_1 \mathbf{h}_0^\top &= \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, & \mathbf{h}_1 \mathbf{h}_1^\top &= \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \end{aligned} \tag{2.51}$$

and then downsampling both the rows and the columns. These filters basically correspond to a moving average, a derivative in the horizontal direction, a derivative in the vertical direction and a derivative in the diagonal direction.

The basic idea of the prefiltering method is to reconstruct filtered versions of the image instead of the image itself, in situations where the filtered versions are sparser than the original image. If a set of filters that covers the whole spectrum of the image is used, such as the set of filters (2.51), then the image can be reconstructed from its filtered versions by a composition stage that basically consists of taking each piece of the image's spectrum from the spectrum of one of its filtered versions. This is illustrated in Figure 2.4. Before any Fourier coefficient from any filtered version is used to compose the image spectrum, it is divided by the matching Fourier coefficient from the corresponding filter to undo any distortions caused by the nonideal magnitude and phase responses of



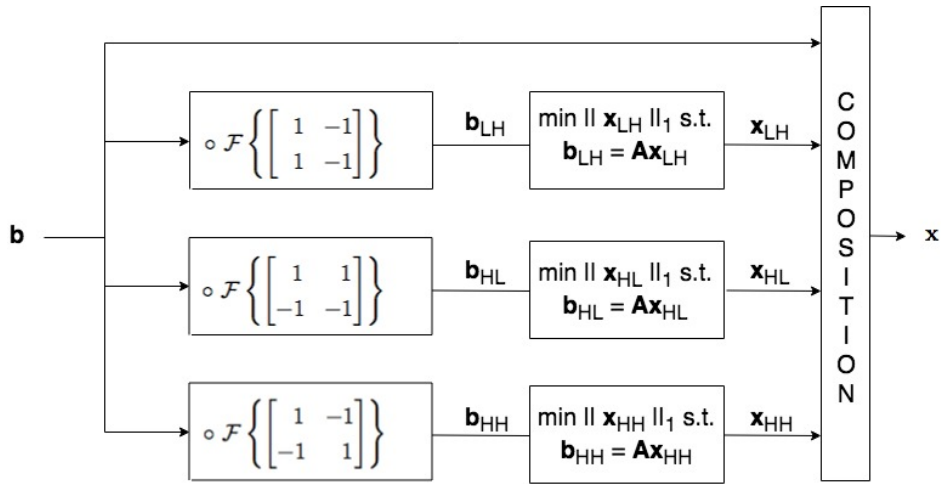
the filters.



**Figure 2.4.** Schematic representation of how the spectrums of the filtered versions are used to obtain the image spectrum in the prefiltering method.

The prefiltering method assumes that one can obtain measurements associated with filtered versions of the image to be reconstructed. It was developed in the context of magnetic resonance imaging (MRI), in which case it is straightforward to compute these measurements, since the data acquired from an MRI scanner correspond to samples in the frequency domain [58]. They are computed multiplying corresponding points of the image spectrum and filter spectrum, since, according to the convolution theorem, product point by point in the frequency domain corresponds to filtering in the space domain [70]. These measurements computation procedure is feasible even in the context of incomplete information as it is performed point by point whereas if it was performed in the space domain the convolution would require complete information.

Since in the MRI case the measurements are already samples in the frequency domain, the Fourier coefficients derived from the filtered versions are used only to fill the gaps in the image spectrum not sampled during data acquisition. Furthermore, since the LL filtered version is not expected to be sparse, it is not reconstructed and its part of the spectrum is obtained from the other filtered versions. It is important to note that the procedure of Figure 2.4 was idealized thinking of ideal filters. True filters are not sensitive only to the parts indicated in Figure 2.4. Therefore, in practice, the composition stage is implemented by choosing for a given part of the spectrum the contribution from the filtered version whose filter has the largest magnitude at that position. This allows us to use only a subset of the filters (2.51). Figure 2.5 represents schematically the prefiltering method in the MRI case, implemented using a subset of the filters (2.51).



**Figure 2.5.** Exemplification of the prefiltering method in the MRI case. The data acquisition process is modeled by a linear system of equations  $\mathbf{b} = \mathbf{A}\mathbf{x}$ , where  $\mathbf{b}$  are the measurements,  $\mathbf{x}$  are the image pixels, and  $\mathbf{A}$  are selected rows from the two-dimensional DFT matrix.  $\mathcal{F}\{\cdot\}$  represents a filter's two-dimensional DFT, with the filter zero padded to the same size as the image, and  $\circ$  represents filtering in frequency domain.

### 3 METHODOLOGY

#### 3.1 COMPRESSIVE SENSING WITH THE ANGULAR SPECTRUM REPRESENTATION OF DIFFUSE PHOTON DENSITY WAVES

##### 3.1.1 Matrix Formulation of the Imaging Problem with Measurements in the Angular Spectrum Representation

In this section, as well as in the next sections, we will use the concepts of circular convolution [72], Kronecker [10, 9, 73, 88] and Hadamard [10, 88] products, Frobenius matrix norm [10, 73] and the vec [10, 73, 88] and diag [73, 88] operations. Notice that we use the diag operation defined in [73, 88] that acts on vectors, and not the one defined in [10, 73, 88] that acts on matrices.

Our objective in this section is to obtain a matrix equivalent for the same relationship expressed in equation (2.33). For this, consider equation (2.32) again, which we reproduce here for convenience

$$\Phi^s(x, y, 0) = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty G(x - x', y - y'; 0, z') A(x', y', z') dx' dy' dz'. \quad (3.1)$$

The Green's function decays rapidly in every direction, and therefore after some distance it can be considered approximately zero. Assume that  $G(x, y; 0, z)$  is zero whenever  $|x| > l_g/2$ ,  $|y| > w_g/2$ , or  $z > h_g$ , which means  $G(x, y; 0, z)$  is nonzero for a rectangular parallelepiped, whose length, width and height are  $l_g$ ,  $w_g$  and  $h_g$  respectively. The region for which  $A(x, y, z)$  is nonzero is basically determined by  $\Delta k^2(x, y, z)$ , which we do not know a priori. However, suppose we can affirm that  $\Delta k^2(x, y, z)$  is zero for  $|x - x_0| > l_A/2$ ,  $|y - y_0| > w_A/2$ , or  $z > h_A$ , or outside a rectangular parallelepiped of length  $l_A$ , width  $w_A$ , height  $h_A$ , and locations in the x and y axes equal to  $x_0$  and  $y_0$ . Call this the volume of interest (VOI). Then,  $\Phi^s(x, y, 0)$  is zero for  $|x - x_0| > l_A/2 + l_g/2$  or  $|y - y_0| > w_A/2 + w_g/2$ , or outside a rectangle of length  $(l_g + l_A)$ , and width  $(w_g + w_A)$ . This is due to the convolutions in the x and y dimensions. Suppose we discretize the medium in equal sized rectangular parallelepiped voxels of length and width  $h$  and height  $\tilde{h}$  and that the Green's function, fluence rate, and perturbation in squared wavenumber are constant over each

of these voxels. Also suppose that  $l_g = J_1^g h$ ,  $w_g = J_2^g h$ ,  $h_g = J_3 \hbar$ ,  $l_A = J_1^A h$ ,  $w_A = J_2^A h$ ,  $h_A = J_3 \hbar$  and that  $J_1^g$  and  $J_2^g$  are odd numbers. Define the discrete finite length functions

$$G[j_1, j_2, j_3] = \begin{cases} G(\mathcal{X}_{j_1}, \mathcal{Y}_{j_2}; 0, \mathcal{Z}_{j_3}), & \text{for } j_1 = 0, \dots, J_1^g - 1, \\ & j_2 = 0, \dots, J_2^g - 1 \text{ and} \\ & j_3 = 0, \dots, J_3 - 1 \\ 0, & \text{otherwise} \end{cases} \quad (3.2)$$

$$\begin{aligned} \mathcal{X}_{j_1} &= j_1 h + \frac{h}{2} - \frac{J_1^g h}{2}, \\ \mathcal{Y}_{j_2} &= j_2 h + \frac{h}{2} - \frac{J_2^g h}{2}, \\ \mathcal{Z}_{j_3} &= j_3 \hbar + \frac{\hbar}{2}, \end{aligned}$$

$$A[j_1, j_2, j_3] = \begin{cases} A(X_{j_1}, Y_{j_2}, Z_{j_3}), & \text{for } j_1 = 0, \dots, J_1^A - 1, \\ & j_2 = 0, \dots, J_2^A - 1 \text{ and} \\ & j_3 = 0, \dots, J_3 - 1 \\ 0, & \text{otherwise} \end{cases} \quad (3.3)$$

$$\begin{aligned} X_{j_1} &= j_1 h + \frac{h}{2} - \frac{J_1^A h}{2} + x_0, \\ Y_{j_2} &= j_2 h + \frac{h}{2} - \frac{J_2^A h}{2} + y_0, \\ Z_{j_3} &= j_3 \hbar + \frac{\hbar}{2}, \end{aligned}$$

where  $j_1 = 0, \dots, (J_1 - 1)$ ,  $j_2 = 0, \dots, (J_2 - 1)$  and  $j_3 = 0, \dots, (J_3 - 1)$ . Many times in discrete domain, it is the circular convolution that is the most appropriate rather than the linear convolution. However equation (3.1) implies a linear convolution between the functions (3.2) and (3.3) as well, in order for the result to be a sampled version of  $\Phi^s(x, y, 0)$ . Fortunately, we can make their circular convolution coincide with their linear convolution by choosing  $J_1 = J_1^g + J_1^A - 1$  and  $J_2 = J_2^g + J_2^A - 1$ . Then

$$\Phi^s[j_1, j_2] = \sum_{j'_3=0}^{J_3-1} \sum_{j'_2=0}^{J_2-1} \sum_{j'_1=0}^{J_1-1} G[(j_1 - j'_1) \bmod J_1, (j_2 - j'_2) \bmod J_2, j'_3] A[j'_1, j'_2, j'_3] h^2 \hbar, \quad (3.4)$$

where

$$\begin{aligned} \Phi^s[j_1, j_2] &= \Phi^s(\mathbf{X}_{j_1}, \mathbf{Y}_{j_2}, 0), \\ \mathbf{X}_{j_1} &= j_1 h + \frac{h}{2} - \frac{J_1 h}{2} + x_0, \\ \mathbf{Y}_{j_2} &= j_2 h + \frac{h}{2} - \frac{J_2 h}{2} + y_0. \end{aligned}$$

Equation (3.4) can be seen as a numerical approximation of equation (3.1) using Riemann sums. The convergence of the Riemann sums to the integrals is not rapid [25], so the grid intervals  $h$  and  $\hbar$  have to be made particularly small in order to achieve low approximation errors. For other more efficient numerical integration methods see [25].

If we take the discrete Fourier transform of equation (3.4) with respect to  $j_1$  and  $j_2$  sequentially and use the circular convolution theorem we get

$$\begin{aligned}
\tilde{\Phi}^s[k_1, j_2] &= \\
&\sum_{j'_3=0}^{J_3-1} \sum_{j'_2=0}^{J_2-1} \left[ \sum_{j_1=0}^{J_1-1} \left( \sum_{j'_1=0}^{J_1-1} G[(j_1 - j'_1) \bmod J_1, (j_2 - j'_2) \bmod J_2, j'_3] A[j'_1, j'_2, j'_3] \right) e^{-\frac{-j_2 2\pi k_1 j_1}{J_1}} \right] h^2 \hbar \\
&= \sum_{j'_3=0}^{J_3-1} \sum_{j'_2=0}^{J_2-1} \tilde{G}[k_1, (j_2 - j'_2) \bmod J_2, j'_3] \tilde{A}[k_1, j'_2, j'_3] h^2 \hbar, \\
\tilde{\Phi}^s[k_1, k_2] &= \sum_{j'_3=0}^{J_3-1} \left[ \sum_{j_2=0}^{J_2-1} \left( \sum_{j'_2=0}^{J_2-1} \tilde{G}[k_1, (j_2 - j'_2) \bmod J_2, j'_3] \tilde{A}[k_1, j'_2, j'_3] \right) e^{-\frac{-j_2 2\pi k_2 j_2}{J_2}} \right] h^2 \hbar \\
&= \sum_{j'_3=0}^{J_3-1} \tilde{G}[k_1, k_2, j'_3] \tilde{A}[k_1, k_2, j'_3] h^2 \hbar,
\end{aligned} \tag{3.5}$$

where  $k_1 = 0, \dots, J_1 - 1$  and  $k_2 = 0, \dots, J_2 - 1$ . From equation (3.5), the normalized two-dimensional discrete Fourier transform  $\tilde{\Phi}^s[k_1, k_2]/\sqrt{J_1 J_2}$  of  $\Phi^s[j_1, j_2]$  can be obtained from

$$\frac{\tilde{\Phi}^s[k_1, k_2]}{\sqrt{J_1 J_2}} = \sum_{j'_3=0}^{J_3-1} \sqrt{J_1 J_2} \frac{\tilde{G}[k_1, k_2, j'_3]}{\sqrt{J_1 J_2}} \frac{\tilde{A}[k_1, k_2, j'_3]}{\sqrt{J_1 J_2}} h^2 \hbar \tag{3.6}$$

where  $\tilde{G}[k_1, k_2, j'_3]/\sqrt{J_1 J_2}$  and  $\tilde{A}[k_1, k_2, j'_3]/\sqrt{J_1 J_2}$  are the normalized two-dimensional discrete Fourier transforms of  $G[j_1, j_2, j_3]$  and  $A[j_1, j_2, j_3]$  with respect to  $j_1$  and  $j_2$ . Define the matrices

$$[\mathbf{A}_{j_3}]_{j_1, j_2} = A[j_1, j_2, j_3], \tag{3.7}$$

$$[\tilde{\mathbf{A}}_{j_3}]_{k_1, k_2} = \frac{\tilde{A}[k_1, k_2, j_3]}{\sqrt{J_1 J_2}}, \tag{3.8}$$

$$[\tilde{\mathbf{G}}_{j_3}]_{k_1, k_2} = \sqrt{J_1 J_2} \frac{\tilde{G}[k_1, k_2, j_3]}{\sqrt{J_1 J_2}} h^2 \hbar, \tag{3.9}$$

$$\tilde{\mathbf{B}}_{k_1, k_2} = \frac{\tilde{\Phi}^s[k_1, k_2]}{\sqrt{J_1 J_2}} \tag{3.10}$$

and the vectorized versions of such matrices as

$$\mathbf{x}_{j_3} = \text{vec}(\mathbf{A}_{j_3}), \tag{3.11}$$

$$\tilde{\mathbf{x}}_{j_3} = \text{vec}(\tilde{\mathbf{A}}_{j_3}), \quad (3.12)$$

$$\tilde{\mathbf{g}}_{j_3} = \text{vec}(\tilde{\mathbf{G}}_{j_3}), \quad (3.13)$$

$$\tilde{\mathbf{b}} = \text{vec}(\tilde{\mathbf{B}}). \quad (3.14)$$

Equation (3.6) can then be written in matrix form as

$$\tilde{\mathbf{B}} = \sum_{j_3=0}^{J_3-1} \tilde{\mathbf{G}}_{j_3} \circ \tilde{\mathbf{A}}_{j_3} \quad (3.15)$$

and in vectorized form as

$$\begin{bmatrix} \tilde{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \text{diag}(\tilde{\mathbf{g}}_0) & \text{diag}(\tilde{\mathbf{g}}_1) & \cdots & \text{diag}(\tilde{\mathbf{g}}_{J_3-1}) \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}_0 \\ \tilde{\mathbf{x}}_1 \\ \vdots \\ \tilde{\mathbf{x}}_{J_3-1} \end{bmatrix} \quad (3.16)$$

where  $\circ$  is the Hadamard product, or the entrywise product. The process of taking the normalized discrete Fourier transform of  $A[j_1, j_2, j_3]$  with respect to  $j_1$ , and then taking the normalized discrete Fourier transform of the result with respect to  $j_2$ , can be expressed layer by layer in matrix form as

$$\tilde{\mathbf{A}}_{j_3} = \mathbf{F}_{J_1} \mathbf{A}_{j_3} \mathbf{F}_{J_2}^\top, \quad (3.17)$$

and in vectorized form as

$$\tilde{\mathbf{x}}_{j_3} = (\mathbf{F}_{J_2} \otimes \mathbf{F}_{J_1}) \mathbf{x}_{j_3} \quad (3.18)$$

where  $\mathbf{F}_{J_1}$  and  $\mathbf{F}_{J_2}$  are the normalized DFT matrices of sizes  $J_1$  and  $J_2$ , respectively,  $\otimes$  is the Kronecker product, and when moving from equation (3.17) to equation (3.18), we used the identity  $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^\top \otimes \mathbf{A}) \text{vec}(\mathbf{B})$ . Then equation (3.16) can be expanded to

$$\begin{bmatrix} \tilde{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \text{diag}(\tilde{\mathbf{g}}_0) & \text{diag}(\tilde{\mathbf{g}}_1) & \cdots & \text{diag}(\tilde{\mathbf{g}}_{J_3-1}) \end{bmatrix} \begin{bmatrix} \mathbf{F}_{2D} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{2D} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{F}_{2D} \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{J_3-1} \end{bmatrix} \quad (3.19)$$

where  $\mathbf{F}_{2D} = (\mathbf{F}_{J_2} \otimes \mathbf{F}_{J_1})$  is the normalized two-dimensional DFT matrix of this case. This equation can be more compactly written as

$$\tilde{\mathbf{b}} = \mathbf{\Gamma} \mathbf{\Theta} \mathbf{x}. \quad (3.20)$$

where  $\mathbf{\Gamma}$  is the matrix to the left,  $\mathbf{\Theta}$  is the matrix to the right and  $\mathbf{x} = [\mathbf{x}_0^\top \mathbf{x}_1^\top \cdots \mathbf{x}_{J_3-1}^\top]^\top$ .

### 3.1.2 Depth Compensation

Reconstruction algorithms based on equation (3.19) also tend to underestimate the actual depth of perturbations, as did those based on equation (2.26). This time, it is the order of magnitude of the elements of  $\text{diag}(\tilde{\mathbf{g}}_{j_3})$  that decreases as  $j_3$  increases. To counteract this issue, we use a similar but different method than Section 2.1.2. Recall that  $G[j_1, j_2, j_3]$  is

$$G(\mathcal{X}_{j_1}, \mathcal{Y}_{j_2}; 0, \mathcal{Z}_{j_3}) = \frac{e^{jk_0 \sqrt{(\mathcal{X}_{j_1})^2 + (\mathcal{Y}_{j_2})^2 + (0 - \mathcal{Z}_{j_3})^2}}}{4\pi \sqrt{(\mathcal{X}_{j_1})^2 + (\mathcal{Y}_{j_2})^2 + (0 - \mathcal{Z}_{j_3})^2}} + \frac{e^{jk_0 \sqrt{(\mathcal{X}_{j_1})^2 + (\mathcal{Y}_{j_2})^2 + (0 - (-\mathcal{Z}_{j_3} - 2z_b))^2}}}{4\pi \sqrt{(\mathcal{X}_{j_1})^2 + (\mathcal{Y}_{j_2})^2 + (0 - (-\mathcal{Z}_{j_3} - 2z_b))^2}}, \quad (3.21)$$

where

$$\begin{aligned} k_0 &= \sqrt{\alpha} e^{j(\beta/2)}, \\ \alpha &= |k_0^2|, \\ \beta &= \arg(k_0^2), \end{aligned} \quad (3.22)$$

and

$$\begin{aligned} k_0^2 &= \frac{-v\mu_{a0} + j\omega}{D_0}, \\ v &> 0, \\ \mu_{a0} &> 0, \\ D_0 &> 0, \\ \omega &\geq 0, \end{aligned} \quad (3.23)$$

whenever it is nonzero. We can write equation (3.21) in terms of the real and imaginary parts of  $k_0$  as

$$G(\mathcal{X}_{j_1}, \mathcal{Y}_{j_2}; 0, \mathcal{Z}_{j_3}) = \frac{e^{(-\Im(k_0) + j\Re(k_0)) \sqrt{(\mathcal{X}_{j_1})^2 + (\mathcal{Y}_{j_2})^2 + (0 - \mathcal{Z}_{j_3})^2}}}{4\pi \sqrt{(\mathcal{X}_{j_1})^2 + (\mathcal{Y}_{j_2})^2 + (0 - \mathcal{Z}_{j_3})^2}} + \frac{e^{(-\Im(k_0) + j\Re(k_0)) \sqrt{(\mathcal{X}_{j_1})^2 + (\mathcal{Y}_{j_2})^2 + (0 - (-\mathcal{Z}_{j_3} - 2z_b))^2}}}{4\pi \sqrt{(\mathcal{X}_{j_1})^2 + (\mathcal{Y}_{j_2})^2 + (0 - (-\mathcal{Z}_{j_3} - 2z_b))^2}}. \quad (3.24)$$

Since  $k_0^2$  belongs to the second quadrant of the complex plane, then  $k_0$  belongs to the first quadrant due to the way it was defined in equation (3.22). More precisely  $\Re(k_0) \geq 0$  and  $\Im(k_0) > 0$ , where  $\Re(k_0) = 0$  when  $\omega = 0$ . Therefore the imaginary part of  $k_0$  is responsible for the exponential decay of the Green's function, while the real part is responsible for the oscillating behavior. If we ignore the oscillations, we can approximate

the decay along the z-axis by

$$\lambda_{j_3} = \left| -\frac{e^{-\Im(k_0)\mathcal{Z}_{j_3}}}{4\pi\mathcal{Z}_{j_3}} + \frac{e^{-\Im(k_0)(\mathcal{Z}_{j_3}+2z_b)}}{4\pi(\mathcal{Z}_{j_3}+2z_b)} \right|. \quad (3.25)$$

Therefore we multiply  $\text{diag}(\tilde{\mathbf{g}}_{j_3})$  not by the largest singular value of  $\text{diag}(\tilde{\mathbf{g}}_{J_3-1-j_3})$ , but by

$$\lambda_{j_3}^{-1} = \left| -\frac{e^{-\Im(k_0)\mathcal{Z}_{j_3}}}{4\pi\mathcal{Z}_{j_3}} + \frac{e^{-\Im(k_0)(\mathcal{Z}_{j_3}+2z_b)}}{4\pi(\mathcal{Z}_{j_3}+2z_b)} \right|^{-1}. \quad (3.26)$$

In other words, we multiply  $\mathbf{\Gamma}$  to the right by

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_0^{-1}\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \lambda_1^{-1}\mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \lambda_{J_3-1}^{-1}\mathbf{I} \end{bmatrix}, \quad (3.27)$$

where  $\mathbf{I}$  is the identity matrix of size  $(J_1 \cdot J_2)$ , and  $\mathbf{0}$  is the matrix of zeros of the same size. When we modify  $\mathbf{\Gamma}$  we are essentially assuming an hypothetical situation in which the sensing matrix and the vector that generated the measurements were others than the actual ones. Notice that if we simultaneously multiply  $\mathbf{\Gamma}$  to the right by  $\mathbf{\Lambda}$  and  $\mathbf{\Theta}\mathbf{x}$  to the left by  $\mathbf{\Lambda}^{-1}$  we do not modify the measurements, since

$$\mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Lambda}^{-1}\mathbf{\Theta}\mathbf{x} = \mathbf{\Gamma}\mathbf{\Theta}\mathbf{x} = \tilde{\mathbf{b}}. \quad (3.28)$$

Also notice that

$$\mathbf{\Lambda}^{-1}\mathbf{\Theta}\mathbf{x} = \mathbf{\Theta}\mathbf{\Lambda}^{-1}\mathbf{x}, \quad (3.29)$$

due to the way that  $\mathbf{\Theta}$  and  $\mathbf{\Lambda}$  were defined. Therefore, we assume an hypothetical sensing mechanism  $\tilde{\mathbf{\Gamma}}\mathbf{\Theta}$  and an hypothetical image  $\tilde{\mathbf{x}}$  such that

$$\tilde{\mathbf{b}} = \tilde{\mathbf{\Gamma}}\mathbf{\Theta}\tilde{\mathbf{x}}, \quad (3.30)$$

where  $\tilde{\mathbf{\Gamma}} = \mathbf{\Gamma}\mathbf{\Lambda}$  and  $\tilde{\mathbf{x}} = \mathbf{\Lambda}^{-1}\mathbf{x}$ . After recovering  $\tilde{\mathbf{x}}^*$  by some reconstruction method, we recover  $\mathbf{x}^*$  by the matrix vector product  $\mathbf{x}^* = \mathbf{\Lambda}\tilde{\mathbf{x}}^*$ , where  $*$  indicates recovered quantities.

### 3.1.3 Compressive Sensing Guarantees of Accurate Recovery

Multiply both sides of equation (3.30) to the left by a diagonal rescaling matrix  $\mathbf{\Pi}$  that normalizes the rows of  $\tilde{\mathbf{\Gamma}}$ . This results in

$$\tilde{\mathbf{b}} = \tilde{\mathbf{\Gamma}}\mathbf{\Theta}\tilde{\mathbf{x}}, \quad (3.31)$$



where  $\check{\mathbf{b}} = \mathbf{\Pi}\tilde{\mathbf{b}}$  and  $\check{\mathbf{\Gamma}} = \mathbf{\Pi}\mathring{\mathbf{\Gamma}}$ . The matrix  $\Theta$  in equation (3.31) is unitary, while the matrix  $\mathring{\mathbf{\Gamma}}$  has orthonormal rows. Also the nonzero elements of the rows of  $\mathring{\mathbf{\Gamma}}$  are evenly spaced, while the nonzero elements of the columns of  $\Theta$  are juxtaposed. Due to this structural difference, we expect the products of the columns of  $\Theta$  with the rows of  $\mathring{\mathbf{\Gamma}}$  to be low. If we manage to complete the set of orthonormal row vectors such that the products of the columns of  $\Theta$  with these rows remain low, we satisfy theorem 2.3.1 for a small  $\mu(F)$  and theorem 2.3.3 for a small  $\mu_U$ .

One way of completing the set of orthonormal row vectors of  $\mathring{\mathbf{\Gamma}}$  and keeping its products with the columns of  $\Theta$  low is by realizing: first that every row of  $\mathring{\mathbf{\Gamma}}$  has  $J_3$  nonzero locations that do not overlap with the  $J_3$  nonzero locations of any other row, and second that, considering only the nonzero elements, each row of  $\mathring{\mathbf{\Gamma}}$  is a  $J_3$ -dimensional vector, and we can find  $J_3 - 1$  vectors in  $\mathbb{C}^{J_3}$  that are orthogonal to it. The nonzero elements of the new rows found this way are also evenly spaced, as were the nonzero elements of the original rows of  $\mathring{\mathbf{\Gamma}}$ , and therefore we expect the products of these new rows with the columns of  $\Theta$  to be as low as before. Figure 3.1 illustrates  $\mathring{\mathbf{\Gamma}}$  as a selection of rows from a unitary matrix found this way.

Theorem 2.3.3 also requires that the columns of the sensing matrix be unit norm. We can satisfy this requirement by pretending that the sensing matrix that generated the measurements was  $\check{\mathbf{\Gamma}}\Theta\Omega$  and the vector that generated the measurements was  $\check{\mathbf{x}} = \Omega^{-1}\hat{\mathbf{x}}$ , where  $\Omega$  is a diagonal rescaling matrix that normalizes the columns of  $\check{\mathbf{\Gamma}}\Theta$ . That is, we assume that the vector of measurements  $\check{\mathbf{b}}$  was obtained by

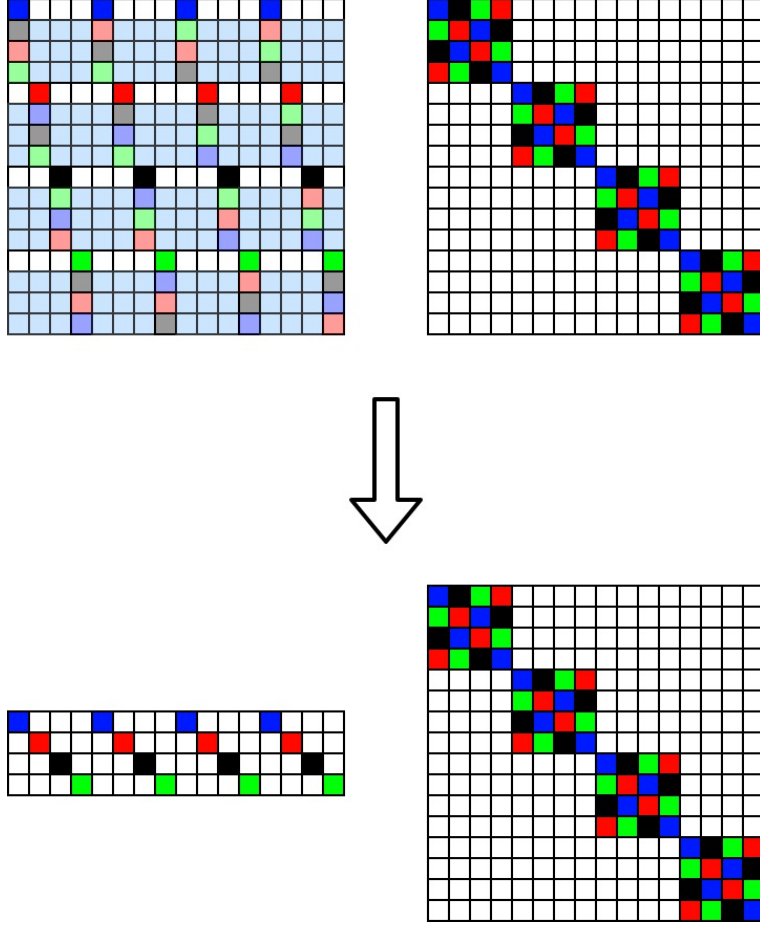
$$\check{\mathbf{b}} = \check{\mathbf{\Gamma}}\Theta\Omega\check{\mathbf{x}}. \quad (3.32)$$

We then try to recover  $\check{\mathbf{x}}$  by solving the optimization problem  $\check{\mathbf{x}}^* = \arg \min_{\hat{\mathbf{x}} \in \mathbb{C}^n} \|\hat{\mathbf{x}}\|_1$  subject to  $\check{\mathbf{b}} = \check{\mathbf{\Gamma}}\Theta\Omega\hat{\mathbf{x}}$ , and  $\hat{\mathbf{x}}$  by calculating the matrix vector product  $\hat{\mathbf{x}}^* = \Omega\check{\mathbf{x}}^*$ , where the superscript \* indicates recovered quantities.

### 3.1.4 Efficient Method for the Sensing Matrix and its Hermitian

The matrix  $\Omega$  that normalizes the columns of  $\check{\mathbf{\Gamma}}\Theta$  can be expressed as

$$\Omega = \begin{bmatrix} \text{diag}(\mathbf{o}_0) & & & & \\ & \text{diag}(\mathbf{o}_1) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \text{diag}(\mathbf{o}_{J_3-1}) \end{bmatrix} \quad (3.33)$$



**Figure 3.1.** The proposed sensing matrix can be seen as a selection of rows from a unitary matrix. More specifically, it can be seen as the product of two matrices, where the matrix to the right is unitary, and the matrix to the left has rows selected from a unitary matrix.

where  $[\mathbf{o}_{j_3}]_i$ , for  $i = 0, \dots, (J_1 \cdot J_2 - 1)$  and  $j_3 = 0, \dots, (J_3 - 1)$ , is the inverted norm of the  $(J_1 J_2 j_3 + i)$ 'th column of  $\ddot{\Gamma} \Theta$ . Define  $\mathbf{O}_{j_3}$  as the  $J_1 \times J_2$  matrix such that  $\text{vec}(\mathbf{O}_{j_3}) = \mathbf{o}_{j_3}$ . Also, let  $\mathbf{v}$  and  $\mathbf{w}$  denote two vectors with the same size as  $\ddot{\mathbf{x}}$ , and  $\mathbf{c}$  and  $\mathbf{d}$  two vectors with the same size as  $\dot{\mathbf{b}}$ . Define  $\mathbf{V}_{j_3}$  and  $\mathbf{W}_{j_3}$  as the  $J_1 \times J_2$  matrices such that  $[\mathbf{v}_0^\top, \mathbf{v}_1^\top, \dots, \mathbf{v}_{J_3-1}^\top]^\top = \mathbf{v}$  and  $[\mathbf{w}_0^\top, \mathbf{w}_1^\top, \dots, \mathbf{w}_{J_3-1}^\top]^\top = \mathbf{w}$  where  $\mathbf{v}_{j_3} = \text{vec}(\mathbf{V}_{j_3})$  and  $\mathbf{w}_{j_3} = \text{vec}(\mathbf{W}_{j_3})$ . Also define  $\mathbf{C}$  and  $\mathbf{D}$  as the  $J_1 \times J_2$  matrices such that  $\text{vec}(\mathbf{C}) = \mathbf{c}$  and  $\text{vec}(\mathbf{D}) = \mathbf{d}$ . With these definitions, the efficient method for the sensing matrix in (3.32) applied to the vector  $\mathbf{v}$  transforms the vectors  $\mathbf{v}_{j_3}$  back into their matrix forms,  $\mathbf{V}_{j_3}$ , then computes

$$\mathbf{C} = \sum_{j_3=0}^{J_3-1} \ddot{\mathbf{G}}_{j_3} \circ \mathcal{F}\{\mathbf{O}_{j_3} \circ \mathbf{V}_{j_3}\}, \quad (3.34)$$

and returns  $\mathbf{C}$  in its vectorized form,  $\mathbf{c}$ . In equation (3.34), the operator  $\mathcal{F}\{\cdot\}$  represents the normalized two-dimensional DFT, and  $\circ$  the entrywise product. The Hermitian of

the sensing matrix in (3.32) is

$$\begin{bmatrix} \text{diag}(\mathbf{o}_0) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \text{diag}(\mathbf{o}_1) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \text{diag}(\mathbf{o}_{J_3-1}) \end{bmatrix} \begin{bmatrix} \mathbf{F}_{2D}^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{2D}^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{F}_{2D}^{-1} \end{bmatrix} \begin{bmatrix} \overline{\text{diag}(\check{\mathbf{g}}_0)} \\ \overline{\text{diag}(\check{\mathbf{g}}_1)} \\ \vdots \\ \overline{\text{diag}(\check{\mathbf{g}}_{J_3-1})} \end{bmatrix} \quad (3.35)$$

where we used the identity  $(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H$ , and the facts that a diagonal matrix is equal to its transpose, and the Hermitian of a unitary matrix is equal to its inverse. The efficient method for the matrix (3.35) applied to the vector  $\mathbf{d}$  transforms the latter into the matrix  $\mathbf{D}$  then computes

$$\mathbf{W}_{j_3} = \mathbf{O}_{j_3} \circ \mathcal{F}^{-1} \left\{ \overline{\check{\mathbf{G}}_{j_3}} \circ \mathbf{D} \right\}, \quad (3.36)$$

for  $j_3 = 0, \dots, (J_3 - 1)$ , converts each matrix  $\mathbf{W}_{j_3}$  into its vectorized form  $\mathbf{w}_{j_3}$ , and returns them stacked into  $\mathbf{w}$ . In the expressions above,  $\overline{(\cdot)}$  indicates complex conjugation, and  $\mathcal{F}^{-1}\{\cdot\}$  the normalized inverse two-dimensional DFT. Note that  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  can be efficiently computed using the fast Fourier transform and the inverse fast Fourier transform.

### 3.1.5 CS with a Sparsifying Transform

Recall that our proposed equation for the DOT forward problem with the measurements in angular spectrum representation, in its pure form, is the equation (3.20):

$$\tilde{\mathbf{b}} = \mathbf{\Gamma} \mathbf{\Theta} \mathbf{x}. \quad (3.37)$$

This equation can also be written for  $\mathbf{x}$  in an arbitrary domain:

$$\tilde{\mathbf{b}} = \mathbf{\Gamma} \mathbf{\Theta} \mathbf{T}^{-1} \mathbf{T} \mathbf{x}. \quad (3.38)$$

Notice that, if the image is not exactly or approximately sparse in its original domain, it may be possible to make it exactly or approximately sparse by choosing an appropriate sparsifying transform  $\mathbf{T}$ . In the next section, we propose an optode arrangement to make the real and imaginary parts of the layers of the image more gradient sparse, so that they are sparsified by three of the Haar filters (2.51). Therefore, consider a sparsifying transform  $\mathbf{T}$  that takes the 1-level 2D Haar discrete wavelet transform (DWT) [84] of each layer of the image.

The 1-level 2D Haar DWT of the layer  $\mathbf{A}_{j_3}$  can be obtained by first transforming

the columns of  $\mathbf{A}_{j_3}$ , which can be expressed by  $\mathbf{W}_{J_1,1}\mathbf{A}_{j_3}$ , where  $\mathbf{W}_{J_1,1}$  is the 1-level 1D Haar DWT matrix of size  $J_1$ , and then transforming the rows of the result, which can be expressed by  $\mathbf{W}_{J_1,1}\mathbf{A}_{j_3}\mathbf{W}_{J_2,1}^\top$ , where  $\mathbf{W}_{J_2,1}$  is the 1-level 1D Haar DWT matrix of size  $J_2$ . In matrix-vector product, this can be expressed by  $(\mathbf{W}_{J_2,1} \otimes \mathbf{W}_{J_1,1})\mathbf{x}_{j_3}$ , where  $\mathbf{x}_{j_3} = \text{vec}(\mathbf{A}_{j_3})$ . Therefore, considering all layers at once:

$$\mathbf{T}\mathbf{x} = \begin{bmatrix} \mathbf{W}_{2D} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{2D} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{W}_{2D} \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{J_3-1} \end{bmatrix} \quad (3.39)$$

where  $\mathbf{W}_{2D} = \mathbf{W}_{J_2,1} \otimes \mathbf{W}_{J_1,1}$  is the 1-level 2D Haar DWT matrix of this case. If we consider only normalized DWTs, then the matrices  $\mathbf{W}_{J_1,1}$ ,  $\mathbf{W}_{J_2,1}$ ,  $\mathbf{W}_{2D}$  and  $\mathbf{T}$  are all orthogonal matrices.

Notice that the matrix  $\Theta\mathbf{T}^{-1}$  has the same block diagonal structure as  $\Theta$ . Also notice that, if we consider only normalized DWTs, then the matrix  $\Theta\mathbf{T}^{-1}$  is also unitary. Therefore, we can make adaptations to the system of equations (3.38), similar to those that we made in the Sections 3.1.2 and 3.1.3 for the system of equations (3.20), to depth compensate the layers of the matrix  $\Gamma$  and to satisfy the requirements of the Theorems 2.3.1 and 2.3.3. Furthermore, the efficient methods for the resulting sensing matrix and its Hermitian are similar to those of Section 3.1.4, but with the inclusion of the efficient methods for the 1-level 2D Haar DWT matrix and its Hermitian, which are based on the fast wavelet transform and the inverse fast wavelet transform.

## 3.2 PREFILTERING METHOD WITH THE ANGULAR SPECTRUM REPRESENTATION OF DIFFUSE PHOTON DENSITY WAVES

In Section 3.1.1, we had already represented the layers of the discrete signal (3.3) by the matrices

$$[\mathbf{A}_{j_3}]_{j_1, j_2} = A[j_1, j_2, j_3]. \quad (3.40)$$

In this section, we will also define the matrices

$$[\mathbf{G}_{j_3}]_{j_1, j_2} = G[j_1, j_2, j_3]h^2\hbar, \quad (3.41)$$

and

$$[\mathbf{B}]_{j_1, j_2} = \Phi^s[j_1, j_2], \quad (3.42)$$

to represent respectively the layers of the discretized Green's function, and the scattered field at the boundary. With these definitions, equation (3.4) can be written as

$$\mathbf{B} = \sum_{j_3=0}^{J_3-1} \mathbf{G}_{j_3} \circledast \mathbf{A}_{j_3}, \quad (3.43)$$

where  $\circledast$  represents the circular convolution. Filtering both sides of equation (3.43) with a purely real discrete two-dimensional finite impulse response filter  $\mathbf{H}$ , zero padded to the same size as the measurements, we obtain

$$\begin{aligned} \mathbf{B} \circledast \mathbf{H} &= \left( \sum_{j_3=0}^{J_3-1} \mathbf{G}_{j_3} \circledast \mathbf{A}_{j_3} \right) \circledast \mathbf{H}, \\ \mathbf{B} \circledast \mathbf{H} &= \sum_{j_3=0}^{J_3-1} [(\mathbf{G}_{j_3} \circledast \mathbf{A}_{j_3}) \circledast \mathbf{H}], \\ \mathbf{B} \circledast \mathbf{H} &= \sum_{j_3=0}^{J_3-1} [\mathbf{G}_{j_3} \circledast (\mathbf{A}_{j_3} \circledast \mathbf{H})], \\ \mathbf{B} \circledast \mathbf{H} &= \sum_{j_3=0}^{J_3-1} \{ \mathbf{G}_{j_3} \circledast [\Re(\mathbf{A}_{j_3}) \circledast \mathbf{H} + j\Im(\mathbf{A}_{j_3}) \circledast \mathbf{H}] \}. \end{aligned} \quad (3.44)$$

That is, filtering the measurements by a purely real filter corresponds to filtering separately the real and imaginary parts of the layers of the image to be reconstructed. The same result could also have been obtained in frequency domain, by filtering both sides of equation (3.15) by the same filter:

$$\begin{aligned} \tilde{\mathbf{B}} \circ \tilde{\mathbf{H}} &= \left( \sum_{j_3=0}^{J_3-1} \tilde{\mathbf{G}}_{j_3} \circ \tilde{\mathbf{A}}_{j_3} \right) \circ \tilde{\mathbf{H}}, \\ \tilde{\mathbf{B}} \circ \tilde{\mathbf{H}} &= \sum_{j_3=0}^{J_3-1} [(\tilde{\mathbf{G}}_{j_3} \circ \tilde{\mathbf{A}}_{j_3}) \circ \tilde{\mathbf{H}}], \\ \tilde{\mathbf{B}} \circ \tilde{\mathbf{H}} &= \sum_{j_3=0}^{J_3-1} [\tilde{\mathbf{G}}_{j_3} \circ (\tilde{\mathbf{A}}_{j_3} \circ \tilde{\mathbf{H}})], \\ \tilde{\mathbf{B}} \circ \tilde{\mathbf{H}} &= \sum_{j_3=0}^{J_3-1} \left\{ \tilde{\mathbf{G}}_{j_3} \circ \left[ \mathcal{F}\{\Re(\mathbf{A}_{j_3})\} \circ \tilde{\mathbf{H}} + j\mathcal{F}\{\Im(\mathbf{A}_{j_3})\} \circ \tilde{\mathbf{H}} \right] \right\}, \end{aligned} \quad (3.45)$$

where  $\tilde{\mathbf{H}}$  is the filter's frequency response, and  $\mathcal{F}\{\cdot\}$  represents the normalized two-dimensional DFT.

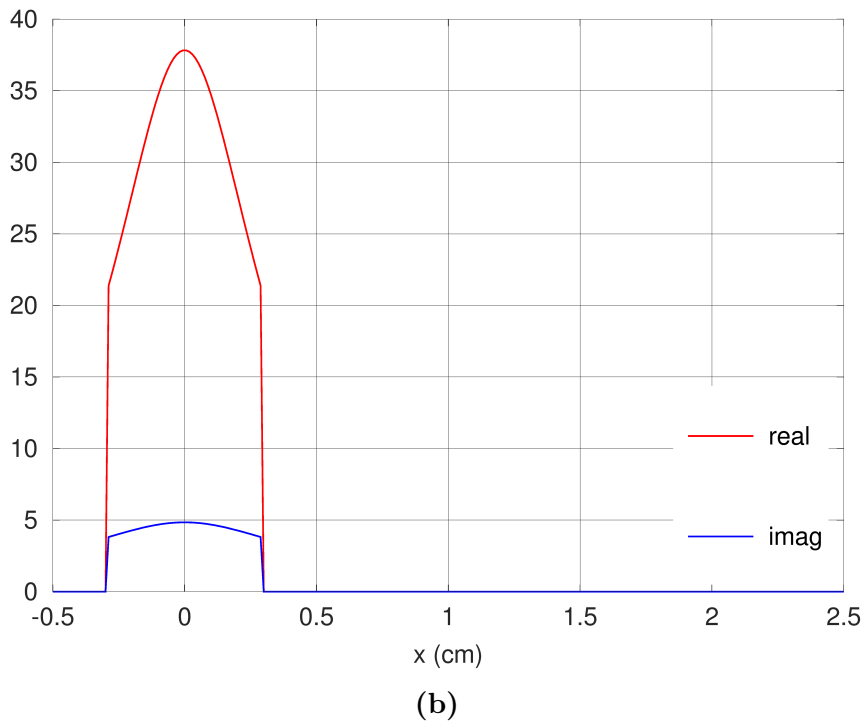
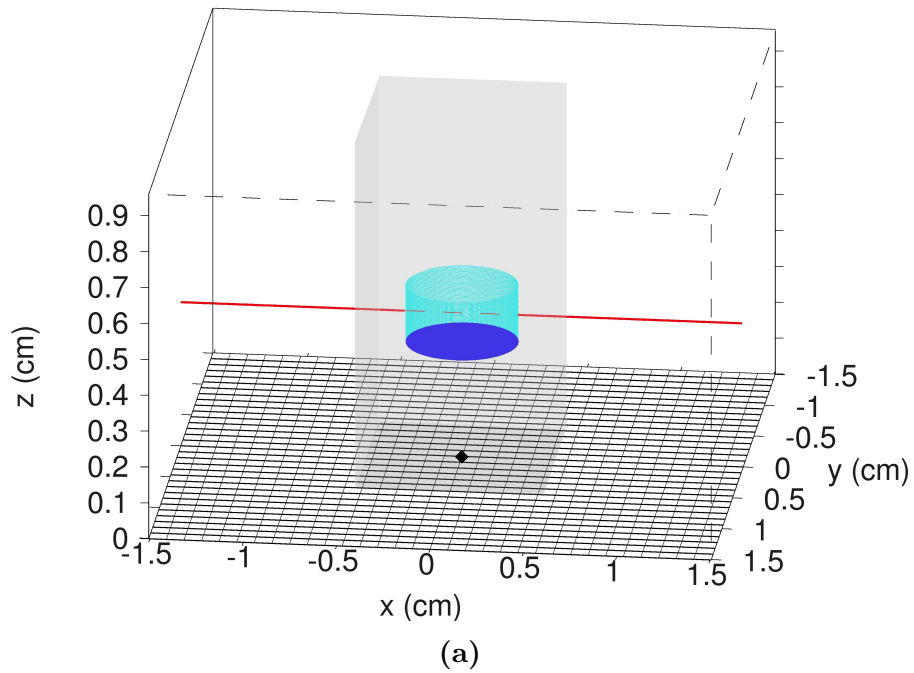
There are some important takeaways from equations (3.44) and (3.45). First of all, even though the measurements are not samples of the image in frequency domain, measurements associated with filtered versions of the image can still be obtained by

filtering the measurements either in space or frequency domain. Second, the signal's filtered versions obtained correspond to variants of the signal filtered in the  $x$  and  $y$  dimensions only. Third, while in the MRI case the prefiltering method based on the set of filters (2.51) assumed that the image was gradient sparse, in this case both the real and imaginary parts of the image have to be gradient sparse for the prefiltering method to work, because this is the only case in which the filtered versions of the image can also be sparse.

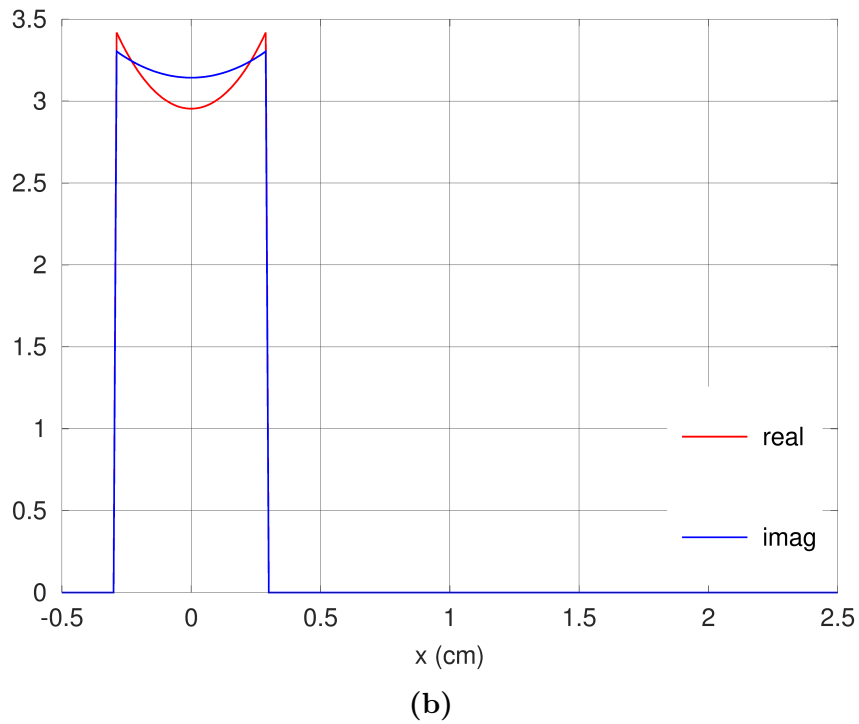
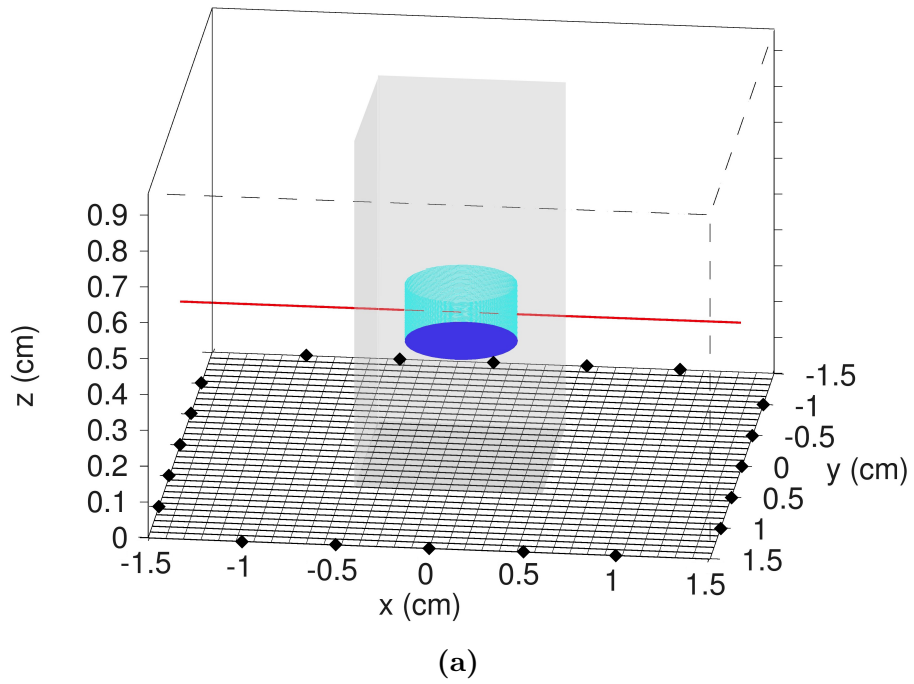
The signal to be reconstructed is the signal (3.3). It corresponds basically to the product point by point of the incident field with the perturbation in squared wavenumber. While the perturbation in squared wavenumber is purely real, the incident field contains overdamped waves as real and imaginary parts [29], meaning they are basically monotonically decreasing functions in every direction from the source. If the perturbation in squared wavenumber is an homogeneous perturbation of a particular shape, then the signal (3.3) is proportional to the incident field inside the region delimited by the perturbation, and zero outside it. This means that at the borders of the perturbation, both its real and imaginary parts will have large gradients and inside the perturbation they will have gradients that depend basically on the decay of the incident field.

Suppose we use an optode arrangement such as the one in [20] to image a volume with a disc perturbation inside it. This setup is represented in Figure 3.2a for a particular choice of volume, perturbation and optode arrangement dimensions. The real and imaginary parts of the signal to be reconstructed, at a line parallel to the  $x$  axis passing through the center of the disc perturbation, is shown in Figure 3.2b. Lines with different orientations or at different layers would show similar profiles. As indicated, the real and imaginary parts still have large gradients inside the perturbation. This is due to the medium being illuminated by a single source, which causes the incident field to be far from uniform at any plane parallel to the surface. One option to make it more uniform, is to turn on several sources in different locations at the same time, for example using the optode arrangement of Figure 3.3a. For this arrangement, the real and imaginary parts are more gradient sparse, as can be inferred from Figure 3.3b.

The advantage of the optode arrangement of Figure 3.3a compared to one that contains sources in the middle of the arrangement is that both the charge coupled device (CCD) camera and the sources can be closer to the surface. For example, for the arrangement of Figure 3.2a to be possible, it is necessary to move the source away from the surface, then rotate it  $10^\circ$  from the surface normal in order to give space for a CCD camera that is also positioned at approximately the same distance from the surface as the source. This way the beam of light, which is assumed collimated, can intersect the surface at the middle of the area imaged by the CCD camera [20].



**Figure 3.2.** (a) Imaging setup based on [20]. Each square represents 64 detectors and the diamond represents a source. The shaded volume corresponds to the imaged volume. The blue disc represents a perturbation in absorption coefficient. The real and imaginary parts of the image to be reconstructed along the red line crossing the perturbation is shown in (b).



**Figure 3.3.** (a) Proposed imaging setup. Each square represents 64 detectors and each diamond represents a sources. The shaded volume corresponds to the imaged volume. The blue disc represents a perturbation in absorption coefficient. The real and imaginary parts of the image to be reconstructed along the red line crossing the perturbation is shown in (b).



The problem of composing the image from its filtered versions can be analyzed layer by layer. In theory the spectrum of each layer can be obtained from its filtered versions as indicated in Figure 2.4. In practice however, the LL filtered version is not reconstructed because it is not expected to be sparse. The lack of a LL filtered version can be circumvented by realizing that the filters are not completely insensitive to the other parts of the spectrum, even though they are particularly sensitive to specific parts. The spectrums of the filters (2.51) are nonzero for the regions indicated in Figure 3.4. Therefore, at these regions, the spectrums of the filtered versions can be divided point by point by the corresponding filter spectrum, and the coefficients found can be used to fill the unknown coefficients of the image spectrum.

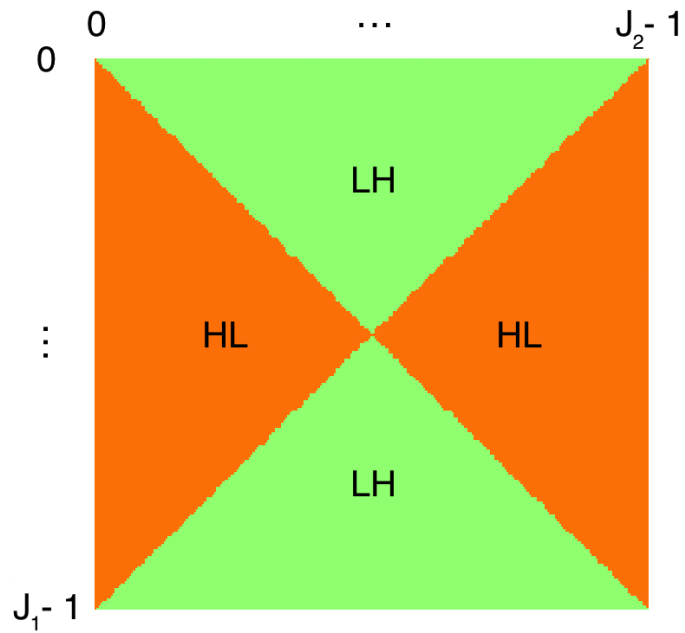
The prefiltering method can be implemented even using only a single filter. However, there are regions from the spectrum in which only specific filters are nonzero, and therefore only them can provide coefficients for that regions. That is the case for example of the zero frequency component, which only the LL filtered version can provide. Therefore, every time a filter is removed from the prefiltering method, a specific part of the spectrum is lost. Images are still recovered by assuming zero the unknown parts of the spectrum, but errors are introduced for every lacking filtered version. In the case that only the LL filtered version is not recovered, the zero frequency component can still be recovered by adjusting the background level to have zero mean, thus reducing the reconstruction error.

The image spectrum is composed from the filtered versions respecting the following rule: the filter with largest magnitude response at a given region defines the filtered version used to reconstruct this region. Therefore, for example, if at a given region the HH filter has a larger magnitude response than the LH and HL filters, then the HH filtered version will be used to reconstruct this region. This results in the divisions of the image spectrum among the filtered versions depicted in the Figures 3.5a and 3.5b, when only the LH and HL filters are used, and when the LH, HL and HH filters are used, respectively.

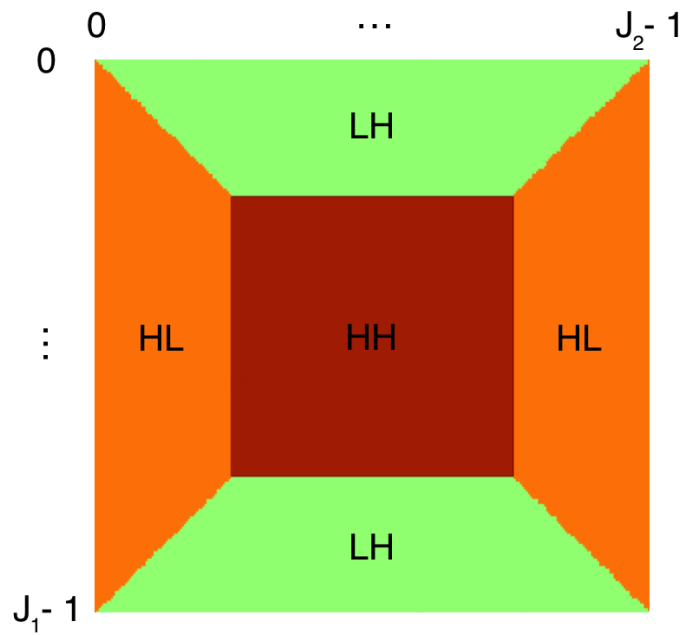
After selecting which filtered version will provide the Fourier coefficients for which part of the spectrum, these coefficients must be divided by the matching coefficients from the corresponding filter. Since the filter spectrum has very small coefficients close to the regions it is zero, the filtered version coefficients should also be close to zero at these regions. If they are just a small amount larger than they should be, then they can explode when we divide by the filter coefficients. To avoid this, we approximate to zero any filtered version coefficient with absolute value below a certain threshold  $\delta$ .

		0	...	$J_2/2$	...	$J_2-1$
0	LL	LL	LH	LH	LH	LL
	LL	LL	LH	LH	LH	LL
.						
.						
	HL	HL	HH	HH	HH	HL
$J_1/2$	HL	HL	HH	HH	HH	HL
	HL	HL	HH	HH	HH	HL
.						
.						
$J_1-1$	LL	LL	LH	LH	LH	LL

**Figure 3.4.** The domain of the DFTs of the set of filter (2.51), divided in 16 rectangular regions. The regions for which the DFTs of the filters  $\mathbf{h}_1\mathbf{h}_1^\top$ ,  $\mathbf{h}_1\mathbf{h}_0^\top$ ,  $\mathbf{h}_0\mathbf{h}_1^\top$  and  $\mathbf{h}_0\mathbf{h}_0^\top$  are nonzero are indicated by HH, HL, LH, and LL respectively. The filters are assumed to be zero padded to size  $J_1 \times J_2$ .



(a)



(b)

**Figure 3.5.** Divisions of the image spectrum among the filtered versions when: (a) only the LH and HL filters are used; (b) the LH, HL and HH filters are used. The regions labeled with an LH are reconstructed using the LH filtered version, the regions labeled with an HL are reconstructed using the HL filtered version, and so on.

### 3.3 NUMERICAL EXPERIMENTS

In Chapter 4, we will present results from simulation experiments designed to evaluate the methods of Sections 3.1 and 3.2. In these experiments, we use the light properties and the medium optical parameters listed in Table 3.1, which were based on the references [37, 36].

We will simulate perturbations in absorption coefficient inside two differently sized volumes of interest (VOIs). The perturbations in the first volume, which we will refer to as the small volume, will be reconstructed assuming the geometric parameters listed in the Table 3.2. On the other hand, the perturbations in the second volume, which we will refer to as the large volume, will be reconstructed assuming the geometric parameters listed in the Table 3.3. In both cases, we will assume that there are five sources equally spaced in each side of the detection array, resulting in a total of twenty sources. The Figures 3.6 and 3.7 illustrate the small and the large volumes and the optode arrangements used to reconstruct them.

The perturbations in absorption coefficient will have the shapes specified in the Figure 3.8. Each image that will be reconstructed will contain one or more of these perturbations. The properties of all the considered images are listed in Table 3.4. Notice that the images reconstructed correspond to the VOI zero padded in the first and second dimensions, because these are the images reconstructed by the algorithms of the Sections 3.1 and 3.2.

In the next chapter, we will first report specific reconstructions of perturbations, either in the small volume or in the large volume. In the graphs that we will report later, we will consider only perturbations in the small volume to reduce the computational load, but in these initial reconstructions, we will also consider perturbations in the large volume.

Our primary aim with these initial reconstructions is to confirm what we predicted in Section 3.1: 1) that specific exactly sparse signals could be perfectly recovered using the proposed technique, and 2) that specific approximately sparse signals could be accurately recovered using the proposed technique. The approximately sparse signals that we will try to reconstruct correspond to the filtered versions of the image, in the context of the prefiltering method. Our secondary aim is to verify the effectiveness of the complete prefiltering method described in the Section 3.2 using all three of the LH, HL and HH filters.

In case of perfect recovery of an exactly sparse signal, the error would be zero, but since the reconstruction algorithms are iterative, it is unlikely that we will get a zero error. However, if we get a signal to error ratio (SER) of a few tens of dBs, and the  $\ell_1$  norm of the reconstructed signal is higher but close to that of the true signal, this

**Table 3.1.** Light and medium properties used in the simulations.

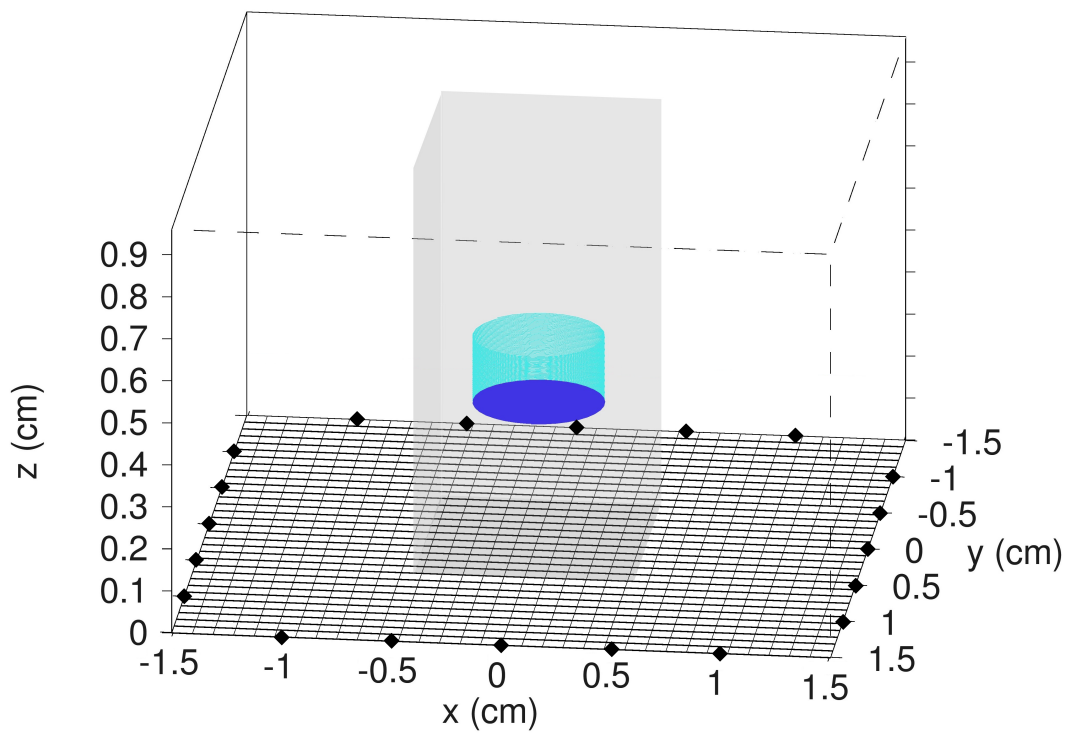
Parameter	Value	Units
$v$	$2.189781 \times 10^{10}$	$cm/s$
$\mu_{a0}$	0.041	$1/cm$
$\mu'_{s0}$	10	$1/cm$
$\omega$	$2\pi(200 \times 10^6)$	$rad/s$
$R_{eff}$	0.4664	-
$\Delta\mu_a$	0.139	$1/cm$
$D_0$	$7.29927 \times 10^8$	$cm^2/s$
$z_b$	0.18321	$cm$
$\ell_{tr}$	0.1	$cm$
$\Delta k^2$	-4.17	$1/cm^2$

**Table 3.2.** Geometric parameters assumed when reconstructing perturbations in the small volume.

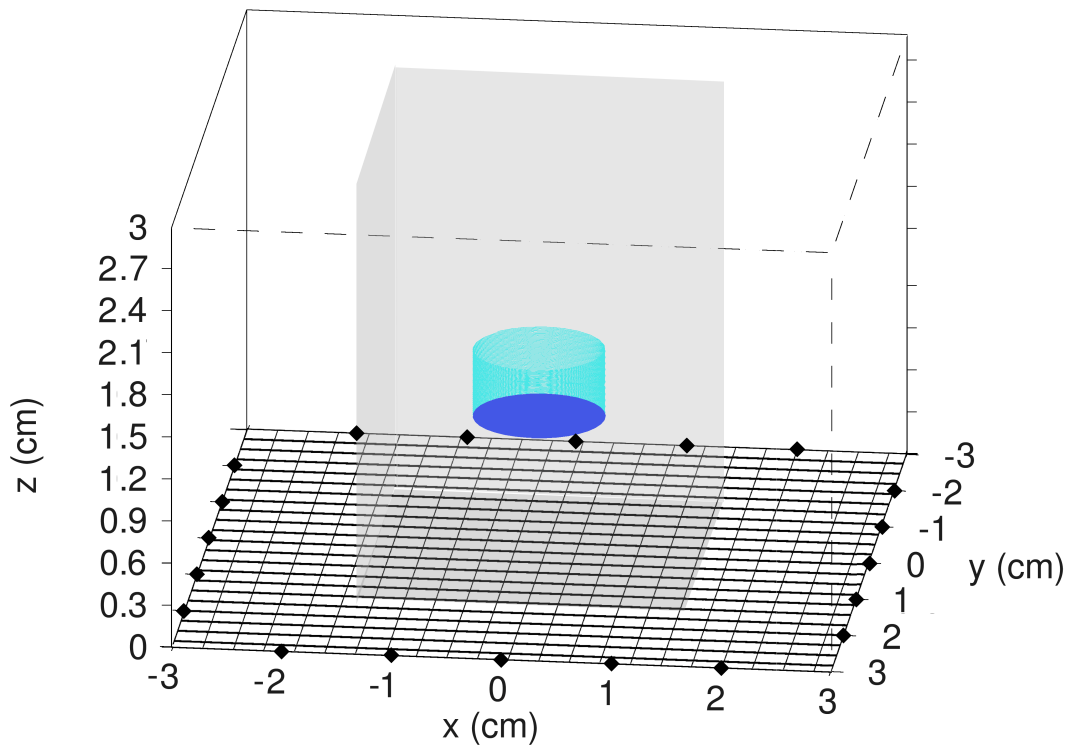
Parameter	Value	Units
$h$	0.012	$cm$
$J_1^A, J_2^A$	86	-
$l_A, w_A$	1.032	$cm$
$J_1^g, J_2^g$	171	-
$l_g, w_g$	2.052	$cm$
$J_1, J_2$	256	-
$(l_g + l_A), (w_g + w_A)$	3.084	$cm$
$\hbar$	0.16	$cm$
$J_3$	6	-
$h_A, h_g$	0.96	$cm$
$x_0, y_0$	0	$cm$

**Table 3.3.** Geometric parameters assumed when reconstructing perturbations in the large volume.

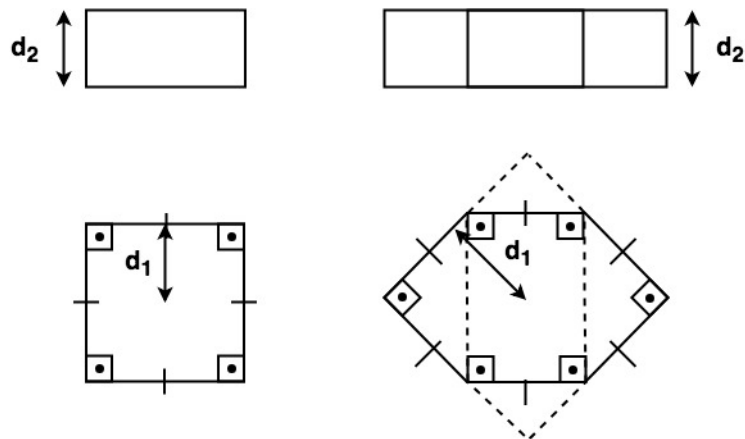
Parameter	Value	Units
$h$	0.035	$cm$
$J_1^A, J_2^A$	86	-
$l_A, w_A$	3.01	$cm$
$J_1^g, J_2^g$	87	-
$l_g, w_g$	3.045	$cm$
$J_1, J_2$	172	-
$(l_g + l_A), (w_g + w_A)$	6.055	$cm$
$\hbar$	0.15	$cm$
$J_3$	20	-
$h_A, h_g$	3	$cm$
$x_0, y_0$	0	$cm$



**Figure 3.6.** A representation of the small volume, shaded, and the optode arrangement used to reconstruct it. Each square represents 64 detectors and each diamond represents a source. This optode arrangement has a total of 65536 detectors and 20 sources. The perturbation in absorption coefficient is only illustrative. Its actual shape, size and location will depend on the experiment.



**Figure 3.7.** A representation of the large volume, shaded, and the optode arrangement used to reconstruct it. Each square represents 64 detectors and each diamond represents a source. This optode arrangement has a total of 29584 detectors and 20 sources. The perturbation in absorption coefficient is only illustrative. Its actual shape, size and location will depend on the experiment.



**Figure 3.8.** Generic representation of the perturbations in absorption coefficient that will be considered in Chapter 4. Some perturbations will have the shape of the rectangular prism to the left, and others will have the shape of the hexagonal prism to the right. In the top left corner and the bottom left corner are the frontal and top views of the rectangular prism, respectively, and in the top right corner and the bottom right corner are the frontal and top views of the hexagonal prism, respectively.

**Table 3.4.** Properties of the images of perturbation in absorption coefficient that will be considered in Chapter 4.  $d_1$  and  $d_2$  refer to the dimensions specified in Figure 3.8 for the rectangular prism and the hexagonal prism perturbations.

image identifier	VOI size	object shape	object $d_1$ (cm)	object $d_2$ (cm)	object centroid $x, y$ (cm) $z$ (cm)		object layers	image sparsity	image size <sup>†</sup>
I	S	rect.	0.1	0.1	0	0.56	1	256	393216
II	S	rect.	0.125	0.2	0	0.48	2	800	393216
III	S	rect.	0.2	0.2	0	0.56	3	3468	393216
IV	S	rect.	0.3	0.3	0	0.56	3	7500	393216
V	S	rect.	0.4	0.4	0	0.56	4	21780	393216
VI	L	rect.	0.3	0.2	0	1.5	2	648	591680
VII	L	rect.	0.6	0.3	0	1.5	4	4624	591680
VIII	S	hex.	0.1	0.1	0	0.56	1	180	393216
IX	S	hex.	0.2	0.2	0	0.56	3	2088	393216
X	S	hex.	0.3	0.3	0	0.56	3	4644	393216
XI	S	hex.	0.4	0.4	0	0.56	4	13688	393216
XII	S	rect.	0.05	0.05	-0.25	0.359	1	128	393216
		rect.	0.05	0.05	+0.25	0.679	1		
XIII	S	rect.	0.1	0.1	-0.25	0.359	1	545	393216
		rect.	0.1	0.1	+0.25	0.679	1		
XIV	S	rect.	0.15	0.15	-0.25	0.359	2	2500	393216
		rect.	0.15	0.15	+0.25	0.679	2		
XV	S	rect.	0.2	0.2	-0.25	0.359	2	4624	393216
		rect.	0.2	0.2	+0.25	0.679	2		

<sup>†</sup> This size corresponds to the VOI zero padded in the first and second dimensions. This is the 3D image actually reconstructed by the algorithms.



indicates that the algorithm is converging to the true solution.

If, however, the  $\ell_1$  norm is lower than that of the true signal, then the algorithm will not converge to the true solution. That is because the IRLS algorithm searches for the solution with minimum  $\ell_1$  norm inside the space of solutions of the system of equations. Therefore if, in any iteration of the IRLS algorithm, we get a signal with  $\ell_1$  norm lower than the true signal, then the true solution is not the one with lowest  $\ell_1$  norm.

The case for approximately sparse signals is particularly different. The theory does not predict exact recovery of approximately sparse signals, but it predicts the recovery of specific approximately sparse signals with a certain accuracy. We can compute the predicted error bounds from the equations (2.44) and (2.45), to check if the error is below predicted value.

In the next chapter, we also report graphs intended to evaluate the proposed techniques under different amounts of noise, and to compare different variants of the proposed techniques with Tikhonov regularization. These different variants correspond to CS with and without the sparsifying transform of Section 3.1.5, and to the prefiltering method using different combinations of the Haar filters.

### 3.3.1 General Procedures

We start our experiments by generating the image of perturbation in absorption coefficient. We start with an image that is zero everywhere, and then define one or more regions where the perturbation in absorption coefficient is  $0.139 \text{ cm}^{-1}$ . The regions that we consider have the shapes illustrated in Figure 3.8.

We calculate the sparsity of the image to be reconstructed by counting the nonzero elements of the image of perturbation in absorption coefficient. Even though the images that we will primarily reconstruct are of perturbation in squared wavenumber multiplied by the negative of the incident field and not of perturbation in absorption coefficient, their sparsities are the same. Strictly speaking the filtered versions of the image to be reconstructed also have approximately the same sparsity as the image itself, but their advantageous property is that they can be well approximated by a few of its nonzero elements, the ones with relatively large absolute values. We estimate how many elements of the filtered versions have relatively large absolute values by filtering the perturbation in absorption coefficient and counting the nonzero elements, even though the filtered versions we will reconstruct are of perturbation in squared wavenumber multiplied by the negative of the incident field and not of perturbation in absorption coefficient, because the locations we expect the gradients to be largest are the same.

We then compute the nonzero values of the discretized Green's function (function (3.2))

by equation (3.21). In order to simulate measurements, we also need to define the values of the signal to be reconstructed (signal (3.3)) inside the VOI, which are

$$A(X_{j_1}, Y_{j_2}, Z_{j_3}) = -\Delta k^2(X_{j_1}, Y_{j_2}, Z_{j_3})\Phi^i(X_{j_1}, Y_{j_2}, Z_{j_3}), \quad (3.46)$$

for  $j_1 = 0, \dots, (J_1^A - 1)$ ,  $j_2 = 0, \dots, (J_2^A - 1)$ , and  $j_3 = 0, \dots, (J_3 - 1)$ . The perturbation in squared wavenumber and incident field inside the VOI are respectively

$$\Delta k^2(X_{j_1}, Y_{j_2}, Z_{j_3}) = \frac{-v\Delta\mu_a(X_{j_1}, Y_{j_2}, Z_{j_3})}{D_0}, \quad (3.47)$$

and

$$\Phi^i(X_{j_1}, Y_{j_2}, Z_{j_3}) = \sum_{q=1}^Q \Phi_q^i(X_{j_1}, Y_{j_2}, Z_{j_3}), \quad (3.48)$$

where the incident field is the result of  $Q$  sources turned on at the same time. We simulate each source as point source with unit amplitude and zero phase moved one transport mean free path length from the boundary to inside the medium. Therefore, the incident field for each source can be computed by

$$\Phi_q^i(X_{j_1}, Y_{j_2}, Z_{j_3}) = \frac{-v}{D_0} \left( - \frac{e^{jk_0\sqrt{(X_{j_1}-\mathbb{X}_q)^2+(Y_{j_2}-\mathbb{Y}_q)^2+(Z_{j_3}-\ell_{tr})^2}}}{4\pi\sqrt{(X_{j_1}-\mathbb{X}_q)^2+(Y_{j_2}-\mathbb{Y}_q)^2+(Z_{j_3}-\ell_{tr})^2}} \right. \\ \left. + \frac{e^{jk_0\sqrt{(X_{j_1}-\mathbb{X}_q)^2+(Y_{j_2}-\mathbb{Y}_q)^2+(Z_{j_3}-(-\ell_{tr}-2z_b))^2}}}{4\pi\sqrt{(X_{j_1}-\mathbb{X}_q)^2+(Y_{j_2}-\mathbb{Y}_q)^2+(Z_{j_3}-(-\ell_{tr}-2z_b))^2}} \right), \quad (3.49)$$

where  $\mathbb{X}_q$  and  $\mathbb{Y}_q$  are the coordinates of source  $q$  in the x-y plane, and  $\ell_{tr}$  is the transport mean free path length.

We then obtain the matrices (3.9) by taking the normalized DFT of the Green's function in the first and second dimensions and multiplying the result by  $h^2\hbar\sqrt{J_1J_2}$ . In order to compensate for the decay in the order of magnitude of  $\tilde{\mathbf{G}}_{j_3}$  as  $j_3$  increases, we redefine these matrices as

$$\dot{\mathbf{G}}_{j_3} = \frac{\tilde{\mathbf{G}}_{j_3}}{\lambda_{j_3}}, \quad (3.50)$$

and simultaneously redefine the layers of the image as

$$\dot{\mathbf{A}}_{j_3} = \mathbf{A}_{j_3}\lambda_{j_3}, \quad (3.51)$$

so the measurements are unchanged. We next normalize the rows of the matrix  $\dot{\mathbf{\Gamma}}$  by making

$$[\ddot{\mathbf{G}}_{j_3}]_{k_1, k_2} = \frac{[\dot{\mathbf{G}}_{j_3}]_{k_1, k_2}}{\sqrt{\sum_{j_3=0}^{J_3-1} ([\dot{\mathbf{G}}_{j_3}]_{k_1, k_2})^2}} \quad (3.52)$$

and the columns of the matrix  $\ddot{\Gamma}\Theta$  by making the diagonal elements of matrix (3.33) equal to

$$[\mathbf{o}_{j_3}]_i = \|\ddot{\mathbf{G}}_{j_3} \circ \mathcal{F}\{\mathbf{E}_i\}\|_{\text{F}}^{-1} \quad (3.53)$$

where  $\mathbf{E}_i$  is the  $J_1 \times J_2$  matrix such that  $\text{vec}(\mathbf{E}_i) = \mathbf{e}_i$ , where  $\mathbf{e}_i$  is the  $i$ 'th column of the identity matrix of size  $(J_1 \cdot J_2)$ . Recall that  $\|\cdot\|_{\text{F}}$  represents the Frobenius matrix norm. We also have to redefine the elements of the image as

$$[\ddot{\mathbf{A}}_{j_3}]_{j_1, j_2} = \frac{[\dot{\mathbf{A}}_{j_3}]_{j_1, j_2}}{[\mathbf{O}_{j_3}]_{j_1, j_2}}, \quad (3.54)$$

where  $\mathbf{O}_{j_3}$  is the  $J_1 \times J_2$  matrix such that  $\text{vec}(\mathbf{O}_{j_3}) = \mathbf{o}_{j_3}$ . Finally we simulate measurements using the efficient method for the sensing matrix

$$\dot{\mathbf{B}} = \sum_{j_3=1}^{J_3-1} \ddot{\mathbf{G}}_{j_3} \circ \mathcal{F}\{\mathbf{O}_{j_3} \circ \ddot{\mathbf{A}}_{j_3}\}, \quad (3.55)$$

where  $\mathcal{F}\{\cdot\}$  and  $\circ$  represent respectively the normalized two-dimensional DFT and the Hadamard, or entrywise, product. To obtain measurements associated with a version of the image modified by a filter  $\mathbf{H}$ , we multiply  $\dot{\mathbf{B}}$  point by point with the filter's frequency response  $\tilde{\mathbf{H}}$ . With the new measurements occupying the place of  $\dot{\mathbf{B}}$ , the steps for the reconstruction of a filtered version are the same as those for the reconstruction of the image directly.

Note that, in our simulations, we generate directly  $\dot{\mathbf{b}}$  or, in other words, rescaled frequency domain measurements of scattered fluence rate. In practice however, we would obtain space domain measurements of scattered fluence rate mixed with incident fluence rate. Therefore, we would have to subtract the estimated incident fluence rate at the detectors from the measurements, take the normalized two-dimensional DFT, and perform the rescaling

$$[\dot{\mathbf{B}}]_{k_1, k_2} = \frac{[\tilde{\mathbf{B}}]_{k_1, k_2}}{\sqrt{\sum_{j_3=0}^{J_3-1} \left([\dot{\mathbf{G}}_{j_3}]_{k_1, k_2}\right)^2}}, \quad (3.56)$$

to obtain the rescaled frequency domain measurements of scattered fluence rate  $\dot{\mathbf{b}}$ .

Next, we attempt to recover the image, or filtered versions of it, using the IRLS algorithm. Let the superscript  $*$  indicate recovered quantities. After recovering the image or one of its filtered versions, we undo the side effects of column normalization by performing the set of operations

$$[\dot{\mathbf{A}}_{j_3}^*]_{j_1, j_2} = [\ddot{\mathbf{A}}_{j_3}^*]_{j_1, j_2} \cdot [\mathbf{O}_{j_3}]_{j_1, j_2}, \quad (3.57)$$

and the side effects of depth compensation by

$$\mathbf{A}_{j_3}^* = \frac{\dot{\mathbf{A}}_{j_3}^*}{\lambda_{j_3}}. \quad (3.58)$$

In case  $\mathbf{A}_{j_3}^*$ ,  $\dot{\mathbf{A}}_{j_3}^*$  and  $\ddot{\mathbf{A}}_{j_3}^*$  refer to a filtered version, we note that it is  $\mathbf{A}_{j_3}^*$  that must be used in the composition stage of the prefiltering method, never  $\ddot{\mathbf{A}}_{j_3}^*$ . Finally, when  $\mathbf{A}_{j_3}^*$  already refers to the unfiltered image, we can estimate the perturbation in squared wavenumber and the perturbations in absorption coefficient inside the VOI by

$$\Delta k^{2*}(X_{j_1}, Y_{j_2}, Z_{j_3}) = \frac{A^*(X_{j_1}, Y_{j_2}, Z_{j_3})}{-\Phi^i(X_{j_1}, Y_{j_2}, Z_{j_3})} \quad (3.59)$$

and

$$\Delta \mu_a^*(X_{j_1}, Y_{j_2}, Z_{j_3}) = \frac{\Delta k^{2*}(X_{j_1}, Y_{j_2}, Z_{j_3})}{(-v/D_0)}, \quad (3.60)$$

for  $j_1 = 0, \dots, (J_1^A - 1)$ ,  $j_2 = 0, \dots, (J_2^A - 1)$ , and  $j_3 = 0, \dots, (J_3 - 1)$ .

In the next chapter, when we calculate the  $\text{SER}(\Delta k^{2*}, \Delta k^2)$ , we consider only the values inside the VOI. When we calculate  $\text{SER}(\ddot{\mathbf{x}}^*, \ddot{\mathbf{x}})$  we consider also the values outside of the VOI.

### 3.3.1.1 CS with a Sparsifying Transform

The steps we follow in the CS reconstructions with the sparsifying transform  $\mathbf{T}$  defined in the Section 3.1.5 are very similar to the steps we follow in the CS reconstructions with no sparsifying transform. The only remarks are: right after calculating the signal (3.3), we transform its layers to the wavelet domain. As a consequence, every instance of  $\mathcal{F}\{\cdot\}$  is replaced by  $\mathcal{F}\{\mathcal{W}^{-1}\{\cdot\}\}$  and every instance of  $\mathcal{F}^{-1}\{\cdot\}$  is replaced by  $\mathcal{W}\{\mathcal{F}^{-1}\{\cdot\}\}$  in the efficient methods for the sensing matrix and its Hermitian, where  $\mathcal{W}\{\cdot\}$  and  $\mathcal{W}^{-1}\{\cdot\}$  represent, respectively, the efficient method for the normalized 1-level 2D Haar DWT, and efficient method for the normalized inverse 1-level 2D Haar DWT. The elements of the matrix (3.33) cannot be computed by (3.53) anymore. Instead, we calculate them by the method that is always applicable, which is by applying the efficient method for the sensing matrix on each column of the appropriate identity matrix, and then taking the norms of the resulting vectors. Finally, after recovering the layers of the signal in (3.58), we convert them back from the wavelet domain to the space domain, before continuing from equation (3.59) onwards.

### 3.3.1.2 Tikhonov Regularization

In some of the experiments, we report reconstructions of Tikhonov regularization as well. The parameters used in these reconstructions are not different than the parameters used in the other reconstructions (Tables 3.1 and 3.2), but some of the specific procedures that we developed for our proposed techniques are not repeated in the Tikhonov reconstructions. We opted to do so both because it gave better results and to comply with the state of the art.

In our proposed techniques, we assume that all sources are turned on, in phase, and at the same time, while the detectors take the measurements. We also assume that the measurements are converted from space to spatial frequency domain after acquisition. In our reconstructions using Tikhonov regularization, on the other hand, we assume that the sources are turned on separately, instead of at the same time, and that the measurements remain in space domain.

Notice that in the first case, the number of measurements is equal to the number of detectors, while in the second case, the number of measurements is equal to the number of detectors times the number of sources. This means that if we use with Tikhonov the same optode arrangement that we use with the other techniques, the number of measurements for this technique will be much greater than for the other techniques. Therefore, in order to make the number of measurements equal for all techniques, we use with Tikhonov the optode arrangement of Figure 3.9. This optode arrangement was obtained from the optode arrangement of Figure 3.6 by removing the center sources of each side, and taking one in every 16 detectors.

Mathematically, the system of equations that we use with Tikhonov regularization is similar to the system of equations (2.26), but with the definitions

$$\begin{cases} a_{m,n} = -G[(j_1 - j'_1) \bmod J_1, (j_2 - j'_2) \bmod J_2, j'_3] \Phi_q^i[j'_1, j'_2, j'_3] h^2 \hbar \\ b_m = \Phi_q^s[j_1, j_2] \\ x_n = \Delta k^2[j'_1, j'_2, j'_3] \end{cases} \quad (3.61)$$

where  $\Phi_q^s[j_1, j_2]$  is the signal (3.4) when only the source  $q$  is turned on,  $\Phi_q^i[j'_1, j'_2, j'_3]$  and  $\Delta k^2[j'_1, j'_2, j'_3]$  is the decomposition of the signal (3.3) when only the source  $q$  is turned on (both are zero whenever  $A[j'_1, j'_2, j'_3]$  is zero),  $G[j_1, j_2, j_3]$  is the signal defined in (3.2),  $n$  is the index of the triple  $(j'_1, j'_2, j'_3)$ ,  $m$  is the index of the triple  $(j_1, j_2, q)$ ,  $q \in \{0, \dots, 15\}$ ,  $j'_1 \in \{0, \dots, 255\}$ ,  $j'_2 \in \{0, \dots, 255\}$ ,  $j'_3 \in \{0, \dots, 5\}$ ,  $j_1 \in \{3, 7, 11, 15, \dots, 255\}$ , and  $j_2 \in \{3, 7, 11, 15, \dots, 255\}$ .

The resulting matrix  $\mathbf{A}$  associated with this system of equations is too large to be stored, so we use an efficient method that can be described by: for source  $q$ , multiply point

by point the input vector with the vectorized incident field; apply the efficient method described in Section 3.1.4 (for  $\mathbf{\Omega}$  an identity matrix); before vectorizing the result, take the normalized inverse two-dimensional DFT; select the values corresponding to where the measurements are taken; vectorize the result; repeat this process for every other source and stack the results in a single vector.

The efficient method that we use for  $\mathbf{A}^H$  can be described by: for source  $q$ , take the subvector of the input vector from  $(4096q)$  to  $(4096q + 4095)$ ; interleave this vector with zeros where measurements were not taken; apply the method described in Section 3.1.4 for the matrix (3.35) (for  $\mathbf{\Omega}$  an identity matrix), but before the first product point by point, take the normalized two-dimensional DFT; multiply the result point by point with the conjugate of the incident field; repeat this process for every source and sum all the resulting vectors.

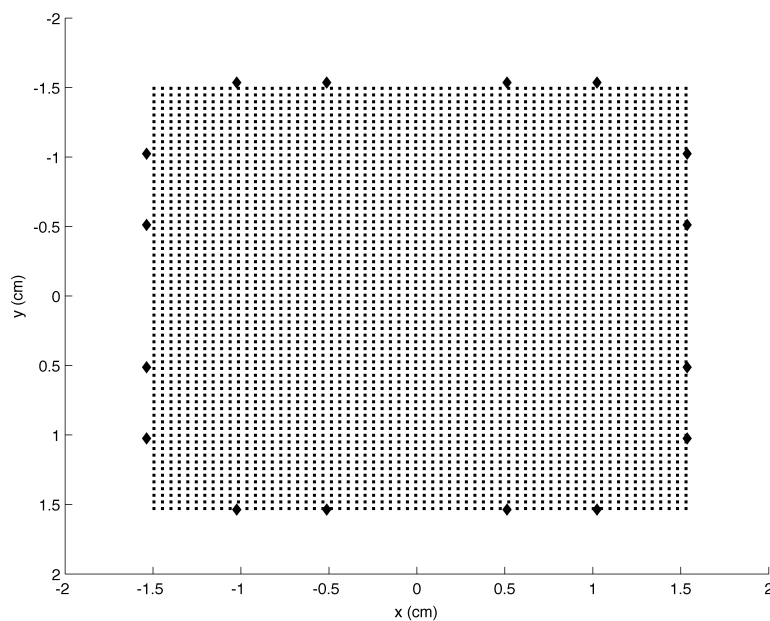
The depth compensation method we use with Tikhonov is again one of the present in the DOT literature, and not the one we use in our proposed techniques. More specifically, we use the depth compensation described in the Section 2.1.2. When the matrix  $\mathbf{A}$  is too large to be stored, we have to use an efficient method to calculate the largest singular value of  $\mathbf{W}_\ell$ , for  $\ell = 0, \dots, L - 1$ . We use the power method [59] to compute the largest eigenvalue of the matrix  $\mathbf{W}_\ell^H \mathbf{W}_\ell$ , and take the square root of the result to obtain the largest singular value of  $\mathbf{W}_\ell$ . In every iteration of the power method, we use efficient methods for  $\mathbf{W}_\ell$  and  $\mathbf{W}_\ell^H$  that we derived from the efficient methods for  $\mathbf{A}$  and  $\mathbf{A}^H$ .

In experiments that we performed, the values of  $\gamma$  that resulted in the highest SERs were about  $\gamma = 0.5$ . Therefore, we use this exponent of the singular values in the depth compensation of the matrix  $\mathbf{A}$ .

Call  $\dot{\mathbf{A}}$  the depth compensated matrix. The efficient method for  $\dot{\mathbf{A}}$  is simply the previously described method for  $\mathbf{A}$ , preceded by the efficient method for the matrix (2.28), which is simply a product point by point. The efficient method for  $\dot{\mathbf{A}}^H$ , on the other hand, is simply the previously described method for  $\mathbf{A}^H$ , but followed by the efficient method for the matrix (2.28).

We used the L-curve method to determine the Tikhonov regularization parameter. For each image that we intended to reconstruct using Tikhonov regularization, we varied the regularization parameter from  $10^{-0}$ ,  $10^{-1}$ ,  $10^{-2}$ ,  $\dots$  until  $10^{-10}$  and visually chose the value closest to the corner of the L-curve. However, the L-curve had no visual corner in the cases that the noise was zero. This may happen when the uncorrelated noise in the measurements does not dominate the highly correlated geometry noise in the forward matrix [47]. Therefore, in these cases we chose the value of  $\lambda$  found by the L-curve method when the noise was close to zero. More specifically, when the standard deviation of the noise in the real and imaginary parts of the measurements was 1.67% the value of the real

or imaginary part. As a consequence, in the reconstructions of the next chapter, we use  $\lambda = 10^{-8}$  in all cases that the noise is zero. We also use  $\lambda = 10^{-8}$  when we reconstruct the images II and III with 1.67% noise and when we reconstruct the image II with 3.33% noise. In all other reconstructions with noise we use  $\lambda = 10^{-7}$ .



**Figure 3.9.** Optode arrangement used with Tikhonov regularization. Each square represents one detector, and each diamond represents one source. This optode arrangement has a total of 4096 detectors and 16 sources.

## 4 RESULTS AND DISCUSSION

### 4.1 RECONSTRUCTIONS OF PERTURBATIONS IN THE SMALL VOLUME

In this section, we report the reconstructions of the images I and IV of Table 3.4 using the methods of Sections 3.1 and 3.2 and the parameters of Tables 3.1 and 3.2. The image I was selected to emphasize the capabilities of CS and the image IV, the capabilities of the prefiltering method. In the reconstructions of this section, we did not use any sparsifying transform and we used all three of the LH, HL and HH filters in the prefiltering method. Also, the regularization parameter of the IRLS algorithm was initialized to 1 and reduced by a factor of 10 after every 30 iterations. The algorithm was interrupted when the regularization parameter reached  $10^{-7}$ .

The reconstruction of the image I using CS resulted in an  $\text{SER}(\tilde{\mathbf{x}}^*, \tilde{\mathbf{x}})$  of 55.2 dB, with  $\|\tilde{\mathbf{x}}^*\|_1 = 14.982$  and  $\|\tilde{\mathbf{x}}\|_1 = 14.976$ . The facts that the SER is a few tens of dBs, the  $\ell_1$  norm is close to that of the true signal, and that the reconstructed image looks like the true image, as can be seen from Figures 4.1a and 4.1b, suggest that the algorithm was converging to the true solution. This indicates that the proposed formulation of the DOT problem does indeed satisfy the requirements for CS. In other words, this indicates the sensing matrix satisfy the RIP of order  $2s$  with  $\delta_{2s} \leq 0.3$  for this level of sparsity. Other parameters related to this reconstruction are given in Table 4.1.

The reconstruction of the image IV using CS resulted in an  $\text{SER}(\tilde{\mathbf{x}}^*, \tilde{\mathbf{x}})$  of -0.8 dB, with  $\|\tilde{\mathbf{x}}^*\|_1 = 359.87$  and  $\|\tilde{\mathbf{x}}\|_1 = 442.27$ . Since the  $\ell_1$  norm is lower than that of the true signal, we can affirm that the algorithm was converging to a solution different than the true solution. This suggests that the level of sparsity of this image is too high for this method and this number of measurements or, in other words, that the sensing matrix does not satisfy the RIP of order  $2s$  with  $\delta_{2s} \leq 0.3$  for this level of sparsity. Parts of the original and recovered images for this reconstruction are given in Figures 4.2a and 4.2b, while other parameters are given in Table 4.2.

The reconstruction of the image IV using the prefiltering method resulted in an  $\text{SER}(\Delta k^{2*}, \Delta k^2)$  of 8.5 dB. this is a great improvement in SER compared to CS, and the reconstructed image looks much more like the true image, as can be seen in Figure 4.3b. other parameters related to this reconstruction are given in the Table 4.3. The filtered



versions were reconstructed with SERs 8.5 dB, 8.5 dB and 12.0 dB. Table 4.3 also includes the squared  $\ell_2$  norm of the error of reconstruction for each filtered version, and error bounds obtained using the equations (2.44) and (2.45), with  $s$  calculated as described in Section 3.3, and the pessimistic estimate  $\delta_{2s} = 0.3$ .

The reconstruction errors satisfy the bounds guaranteed by the RIP of order  $2s$  with parameter  $\delta_{2s} = 0.3$ . In fact, they might satisfy the bounds for a much lower parameter, since the reconstruction errors were much lower than the calculated bounds. Theorem 2.3.3 states that, if the number of measurements is high enough, then the sensing matrix satisfies the RIP of order  $2s$  with parameter  $\delta_{2s} \leq 0.3$  with high probability, and Theorem 2.3.2 states that, if the RIP of order  $2s$  with parameter  $\delta_{2s} \leq 0.3$  is satisfied, then the minimizer recovers the signal with an error bound that depends on  $\delta_{2s}$ . Therefore, since all the reconstruction errors were lower than the calculated error bounds, we have reasons to assume that the sensing matrix does indeed satisfy the RIP of order  $2s$  with parameter  $\delta_{2s} \leq 0.3$ .

**Table 4.1.** Parameters related to the reconstruction of the image I using CS.

SER( $\tilde{\mathbf{x}}^*$ , $\tilde{\mathbf{x}}$ )	SER( $\Delta k^{2*}$ , $\Delta k^2$ )	$\ \tilde{\mathbf{x}}^*\ _1$	$\ \tilde{\mathbf{x}}\ _1$
55.2 dB	55.1 dB	14.982	14.976

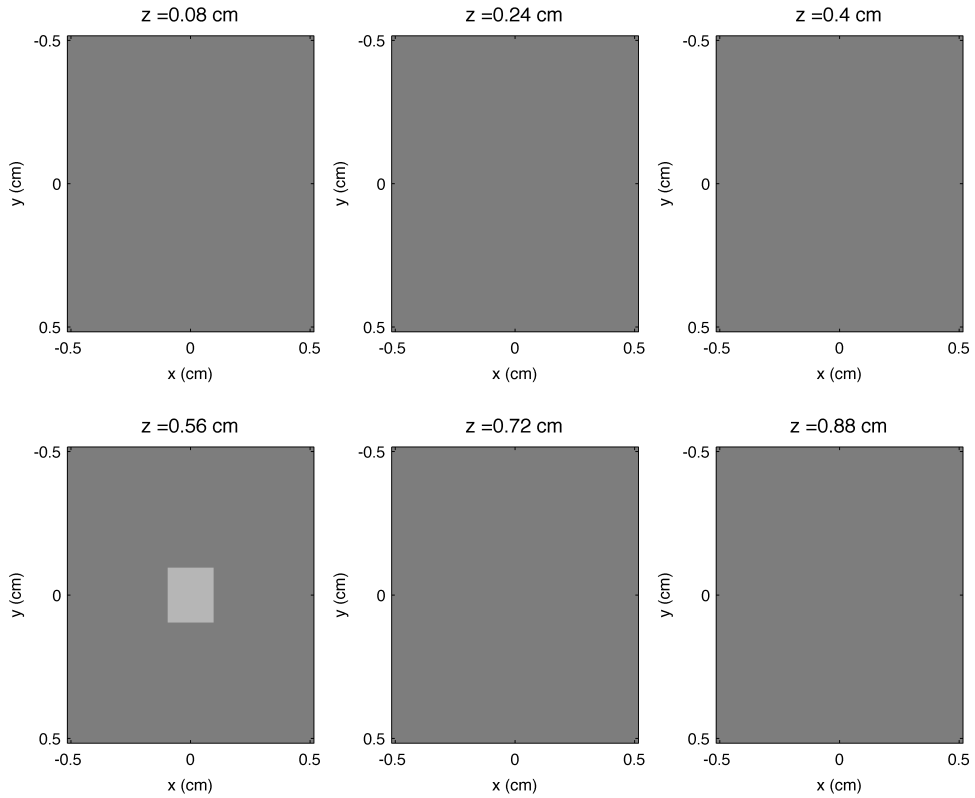
**Table 4.2.** Parameters related to the reconstruction of the image IV using CS.

SER( $\tilde{\mathbf{x}}^*$ , $\tilde{\mathbf{x}}$ )	SER( $\Delta k^{2*}$ , $\Delta k^2$ )	$\ \tilde{\mathbf{x}}^*\ _1$	$\ \tilde{\mathbf{x}}\ _1$
-0.8 dB	-1.0 dB	359.87	442.27

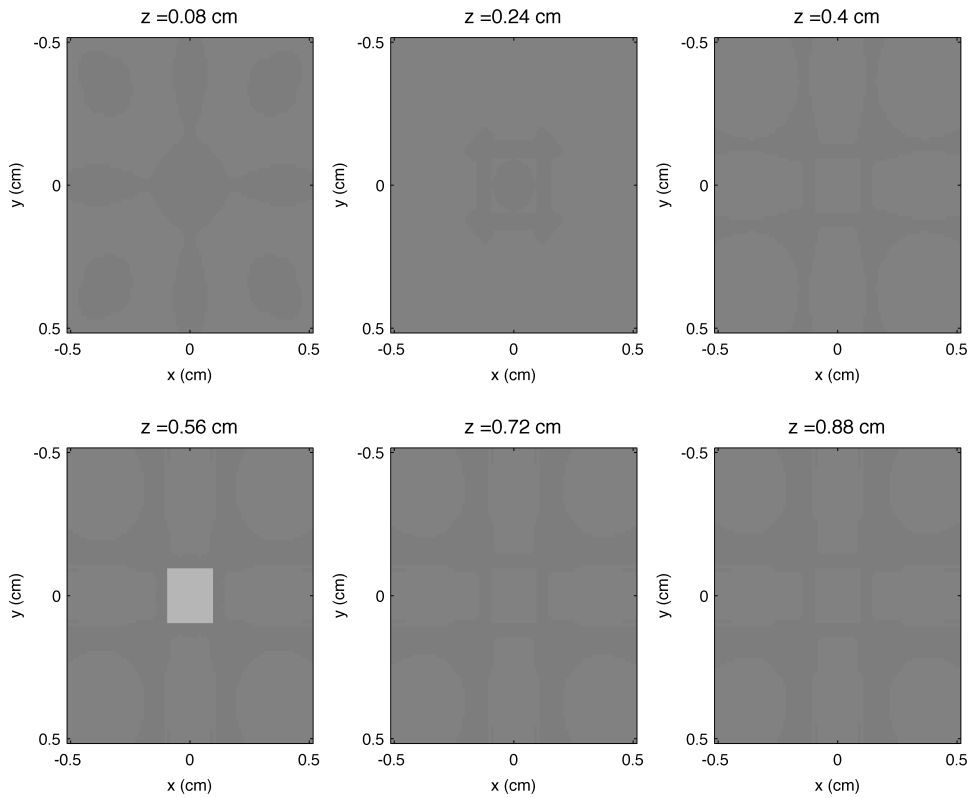
**Table 4.3.** Parameters related to the reconstruction of the image IV using the prefiltering method.

	SER( $\tilde{\mathbf{x}}^*$ , $\tilde{\mathbf{x}}$ )	SER( $\Delta k^{2*}$ , $\Delta k^2$ )	$\ \tilde{\mathbf{x}}^* - \tilde{\mathbf{x}}\ _2^2$	$\frac{C_0 \ \tilde{\mathbf{x}} - \tilde{\mathbf{x}}_s\ _1^2}{s}$	$s$
LH	8.5 dB	8.5 dB <sup>†</sup>	0.2	6.1	306
HL	8.5 dB	8.5 dB <sup>†</sup>	0.2	6.1	306
HH	12.0 dB	8.5 dB <sup>†</sup>	$8.2 \times 10^{-4}$	0.2	12

<sup>†</sup> This value corresponds to the composed image using all the filtered versions.

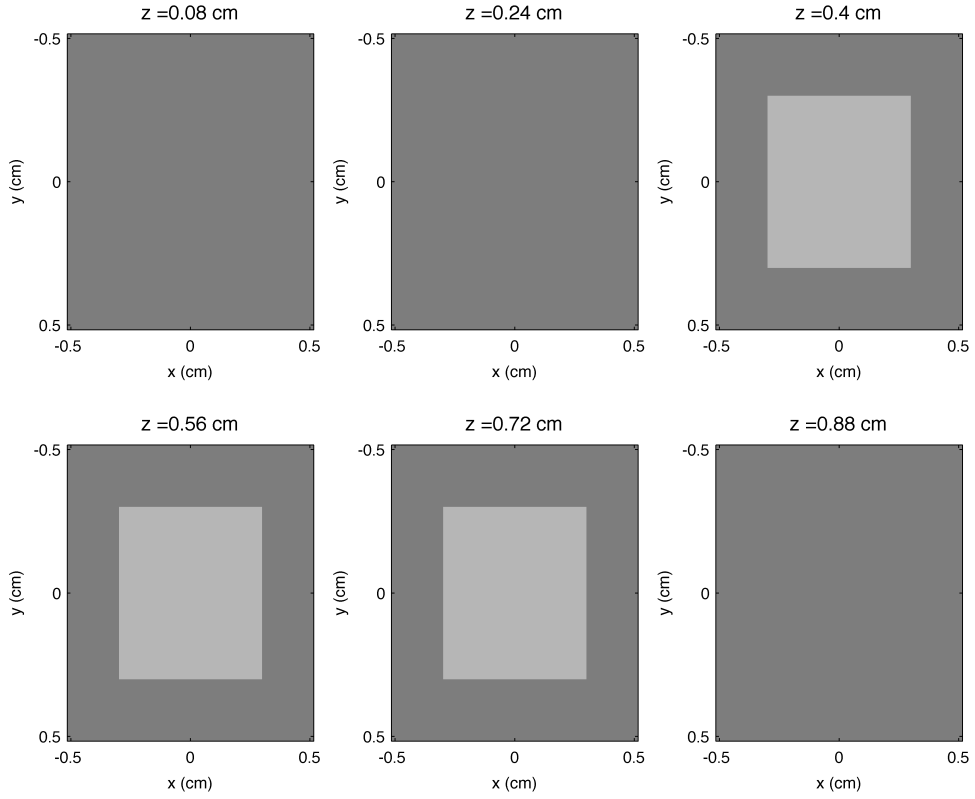


(a)

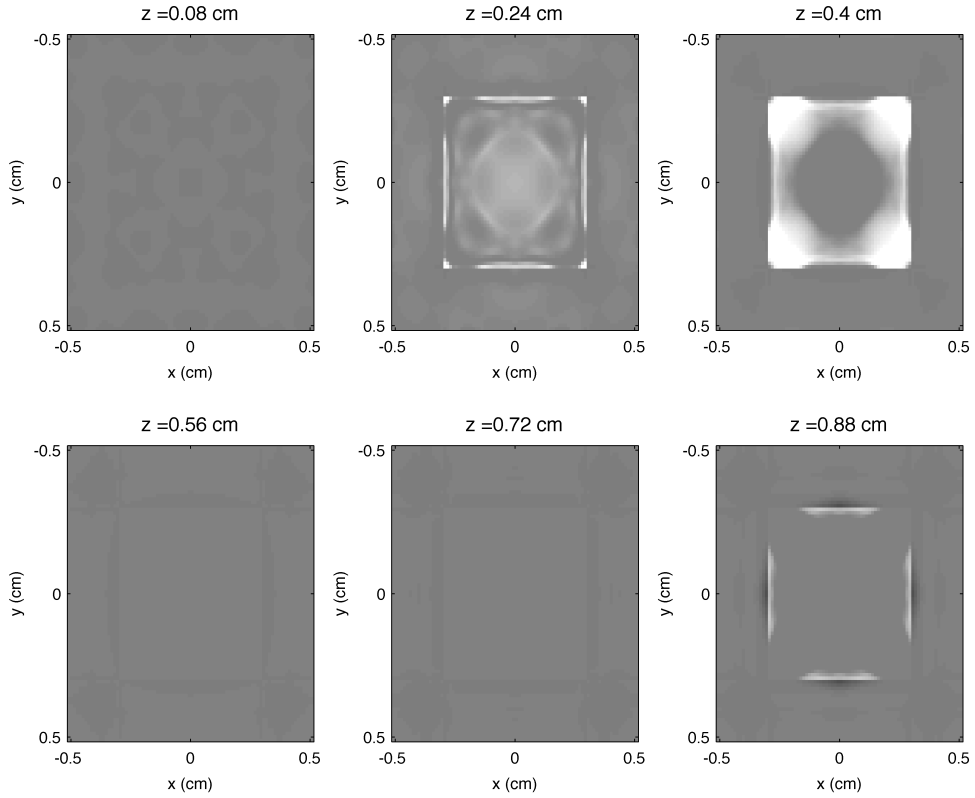


(b)

**Figure 4.1.** (a) Layers of the VOI of the image I. (b) Reconstructed layers of the VOI of the image I using CS. Black indicates a perturbation in squared wavenumber of  $10 \text{ cm}^{-2}$  and white a perturbation in squared wavenumber of  $-10 \text{ cm}^{-2}$ .

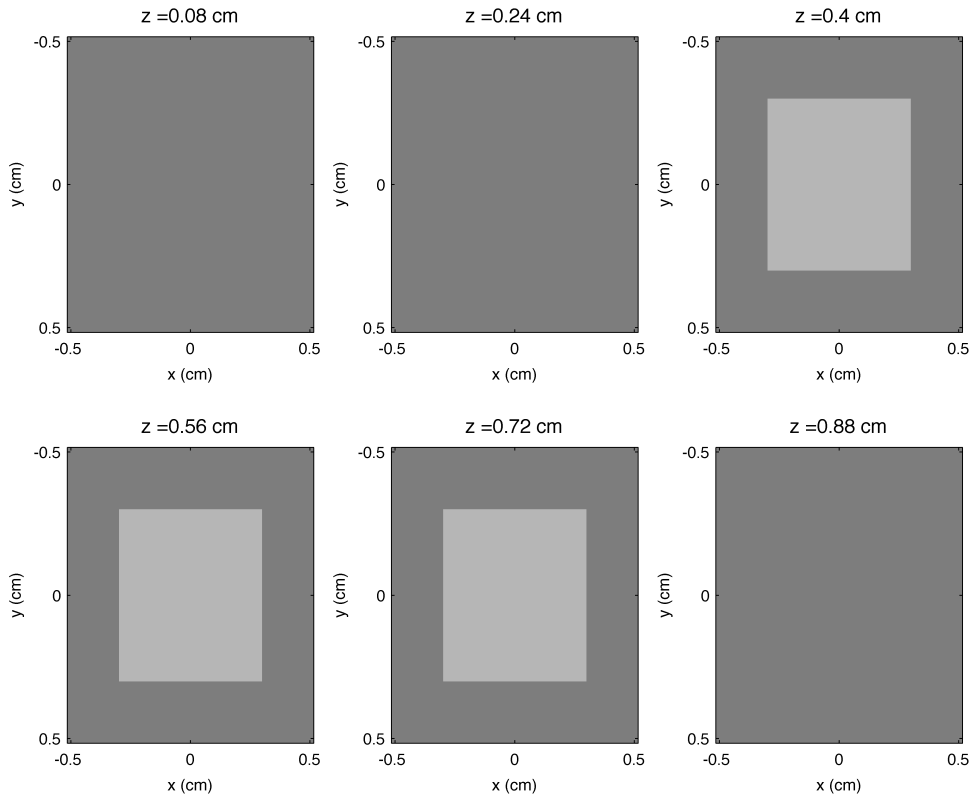


(a)

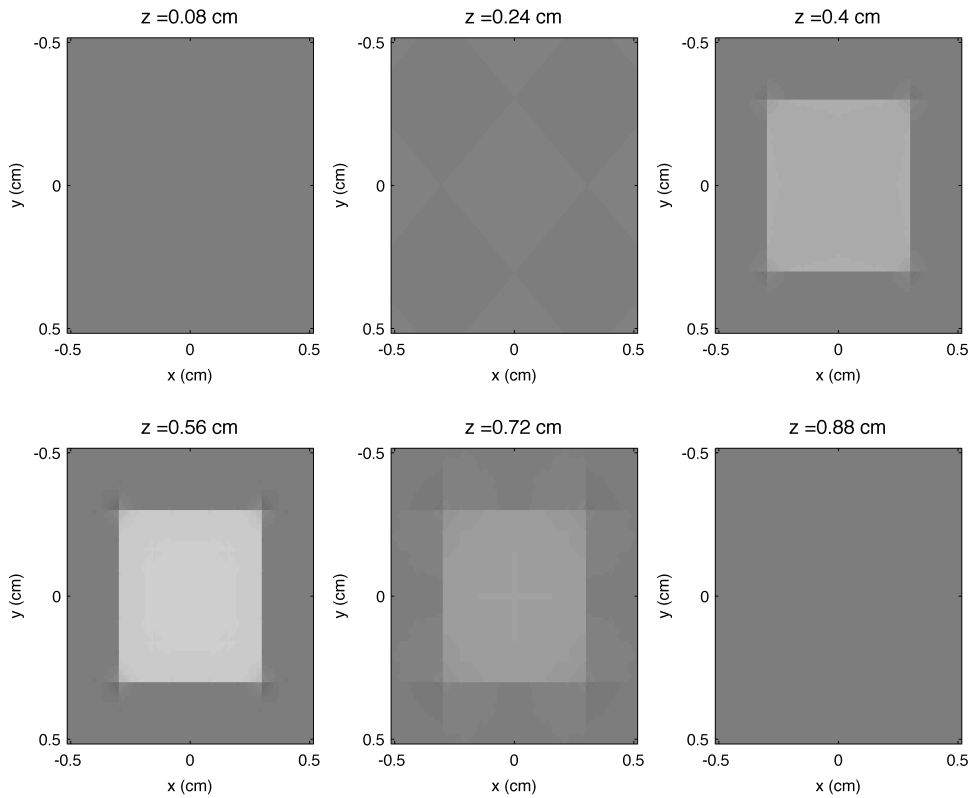


(b)

**Figure 4.2.** (a) Layers of the VOI of the image IV. (b) Reconstructed layers of the VOI of the image IV using CS. Black indicates a perturbation in squared wavenumber of  $10 \text{ cm}^{-2}$  and white a perturbation in squared wavenumber of  $-10 \text{ cm}^{-2}$ .



(a)



(b)

**Figure 4.3.** (a) Layers of the VOI of the image IV. (b) Reconstructed layers of the VOI of the image IV using the prefiltering method. Black indicates a perturbation in squared wavenumber of  $10 \text{ cm}^{-2}$  and white a perturbation in squared wavenumber of  $-10 \text{ cm}^{-2}$ .

## 4.2 RECONSTRUCTIONS OF PERTURBATIONS IN THE LARGE VOLUME

In this section, we report the reconstructions of the images VI and VII of Table 3.4 using the methods of Sections 3.1 and 3.2 and the parameters of Tables 3.1 and 3.3. The images were again selected to emphasize the capabilities of CS and the prefiltering method. In the reconstructions of this section, we did not use any sparsifying transform and we used all three of the LH, HL and HH filters in the prefiltering method. Also, the regularization parameter of the IRLS algorithm was initialized to 1 and reduced by a factor of 10 after every 30 iterations. The algorithm was interrupted when the regularization parameter reached  $10^{-12}$ .

The VOI size in this section was selected to better approximate the conditions in brain imaging. In brain imaging the imaged region must reach at least about 1.5 cm in order for brain cortex activation to be detectable. This minimum distance is based on the facts that the depth of maximum sensitivity is approximately half the source detector separation [81], and commonly used source detector distances that result in clear brain activity are between 3 and 4 cm [89].

The reconstruction of the image VI using CS resulted in an  $\text{SER}(\ddot{\mathbf{x}}^*, \ddot{\mathbf{x}})$  of 20.8 dB. These and other parameters are listed in Table 4.4. This value of SER and the  $\ell_1$  norm of the reconstructed signal suggest that the algorithm was converging to the true solution. We emphasize, however, that to reach this level of SER we had to reduce the minimum regularization parameter of the IRLS algorithm compared to the value we used in the previous section.

Also, we had to reduce the size of the Green's function relative to the size of the VOI in order to achieve higher levels of SER. When reconstructing the small VOI, we assumed that the Green's function had roughly twice the size of the VOI in the  $x$  and  $y$  dimensions. However, when reconstructing the large VOI, we assumed that the Green's function had roughly the same size as the VOI in the  $x$  and  $y$  dimensions. This is because the region where the Green's function was considered nonzero was already large enough. Considering the Green's function nonzero for a region larger than this resulted in a decrease in the reconstruction SER.

The reconstructions of the image VII using CS and using the prefiltering method resulted in the parameters listed in the Tables 4.5 and 4.6, respectively. Again, we selected an image which CS was not capable of recovering. All filtered versions of this image were reconstructed with an SER of 10.5 dB. The final SER after the composition stage and after all other steps was  $\text{SER}(\Delta k^{2*}, \Delta k^2) = 9.3$  dB. The filtered versions reconstruction errors and the bounds predicted by the RIP of order  $2s$  with parameter  $\delta_{2s} = 0.3$  are listed in Table 4.6. Since all the reconstruction errors were below the predicted bounds, this suggests that the RIP of order  $2s$  with parameter  $\delta_{2s} = 0.3$  is satisfied by the sensing

matrix.

The portion of the image VI corresponding to the VOI is depicted in Figures 4.4a. This same portion as reconstructed by CS is shown in Figure 4.4b. The portion of the image VII corresponding to the VOI is depicted in Figure 4.5a or 4.6a. This same portion as reconstructed by CS is shown in Figure 4.5b, and as reconstructed by the prefiltering method, is shown in Figure 4.6b.

**Table 4.4.** Parameters related to the reconstruction of the image VI using CS.

SER( $\ddot{\mathbf{x}}^*$ , $\ddot{\mathbf{x}}$ )	SER( $\Delta k^{2*}$ , $\Delta k^2$ )	$\ \ddot{\mathbf{x}}^*\ _1$	$\ \ddot{\mathbf{x}}\ _1$
20.8 dB	19.7 dB	$6.850 \times 10^{-2}$	$6.847 \times 10^{-2}$

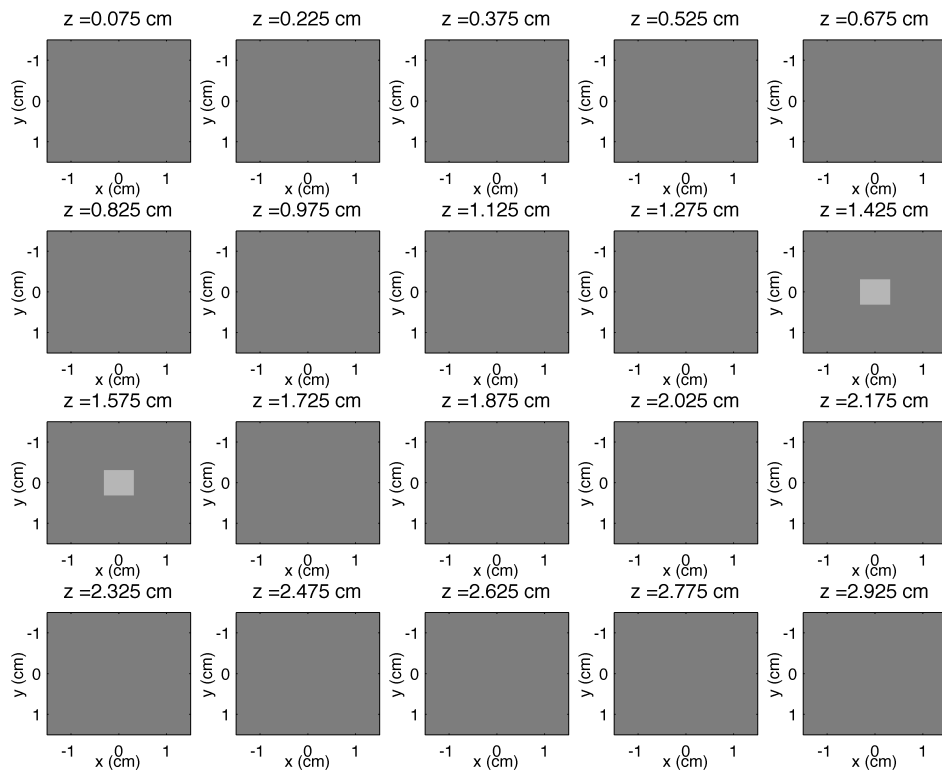
**Table 4.5.** Parameters related to the reconstruction of the image VII using CS.

SER( $\ddot{\mathbf{x}}^*$ , $\ddot{\mathbf{x}}$ )	SER( $\Delta k^{2*}$ , $\Delta k^2$ )	$\ \ddot{\mathbf{x}}^*\ _1$	$\ \ddot{\mathbf{x}}\ _1$
-2.7 dB	-2.1 dB	0.425	0.530

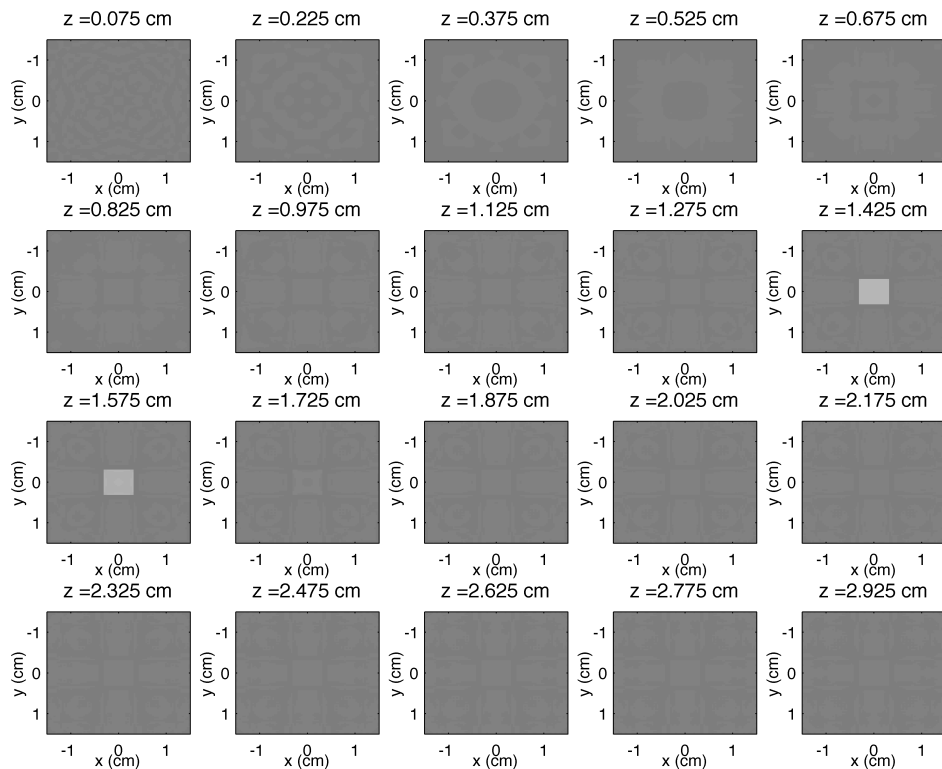
**Table 4.6.** Parameters related to the reconstruction of the image VII using the prefiltering method.

	SER( $\ddot{\mathbf{x}}^*$ , $\ddot{\mathbf{x}}$ )	SER( $\Delta k^{2*}$ , $\Delta k^2$ )	$\ \ddot{\mathbf{x}}^* - \ddot{\mathbf{x}}\ _2^2$	$\frac{C_0 \ \ddot{\mathbf{x}} - \ddot{\mathbf{x}}_s\ _1^2}{s}$	$s$
LH	10.5 dB	9.3 dB <sup>†</sup>	$4.0 \times 10^{-7}$	$9.7 \times 10^{-5}$	280
HL	10.5 dB	9.3 dB <sup>†</sup>	$4.0 \times 10^{-7}$	$9.7 \times 10^{-5}$	280
HH	10.5 dB	9.3 dB <sup>†</sup>	$7.3 \times 10^{-9}$	$5.8 \times 10^{-6}$	16

<sup>†</sup> This value corresponds to the composed image using all the filtered versions.

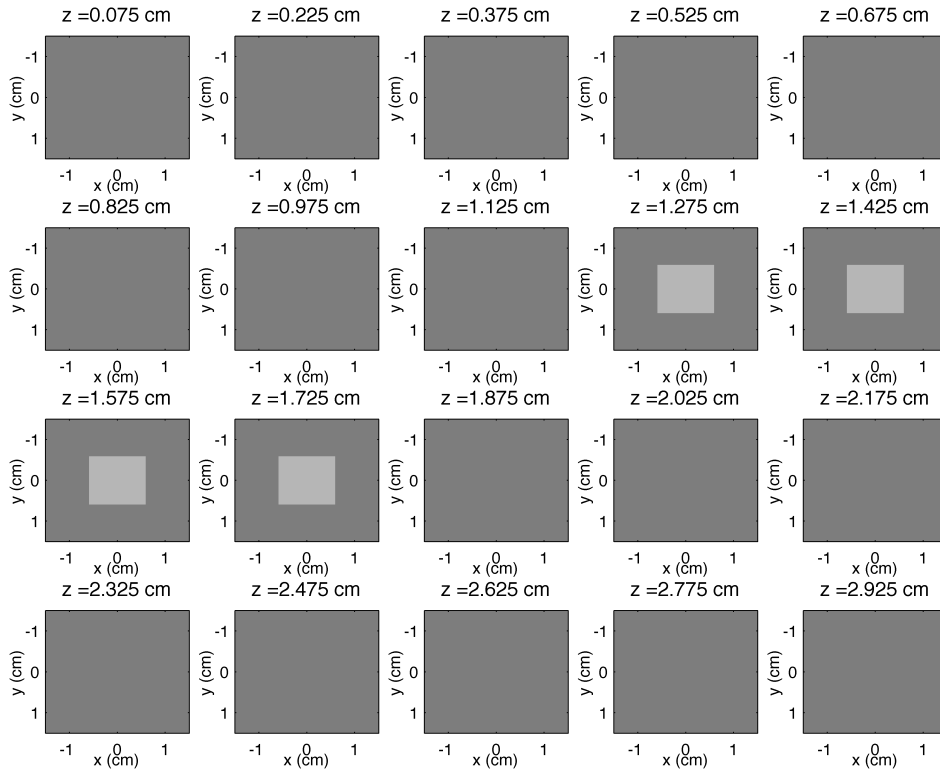


(a)

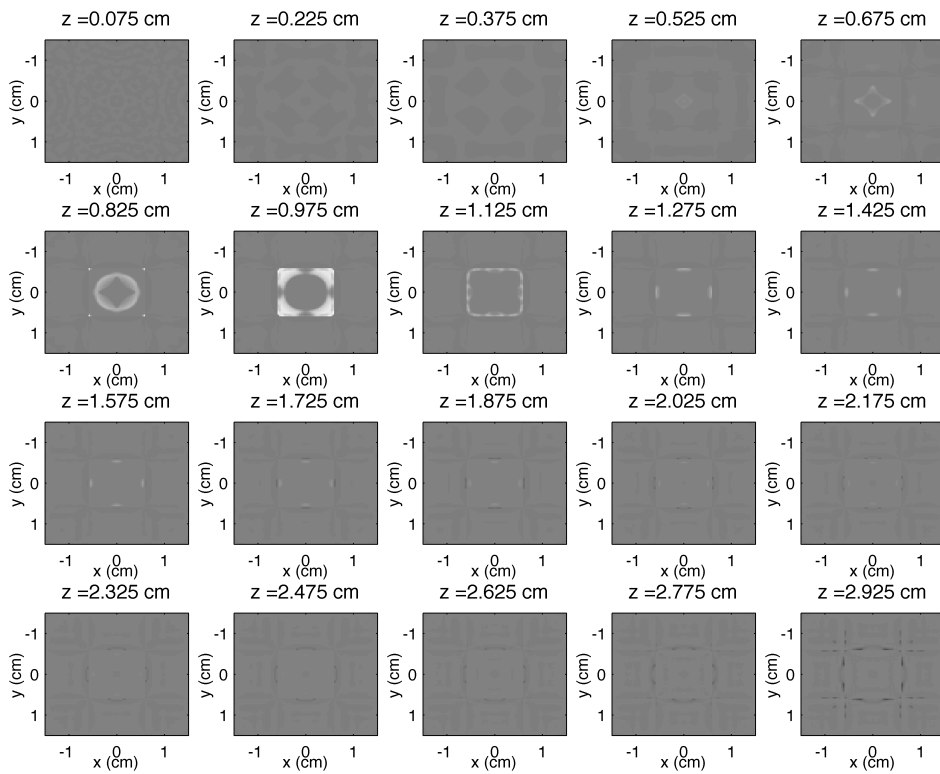


(b)

**Figure 4.4.** (a) Layers of the VOI of the image VI. (b) Reconstructed layers of the VOI of the image VI using CS. Black indicates a perturbation in squared wavenumber of  $10 \text{ cm}^{-2}$  and white a perturbation in squared wavenumber of  $-10 \text{ cm}^{-2}$ .



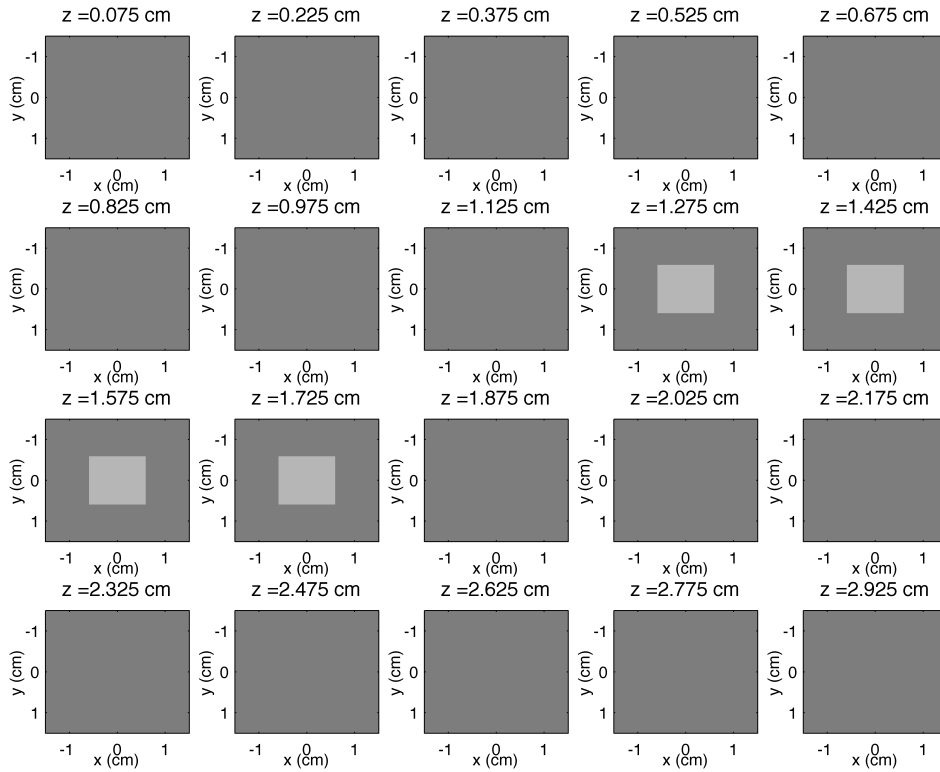
(a)



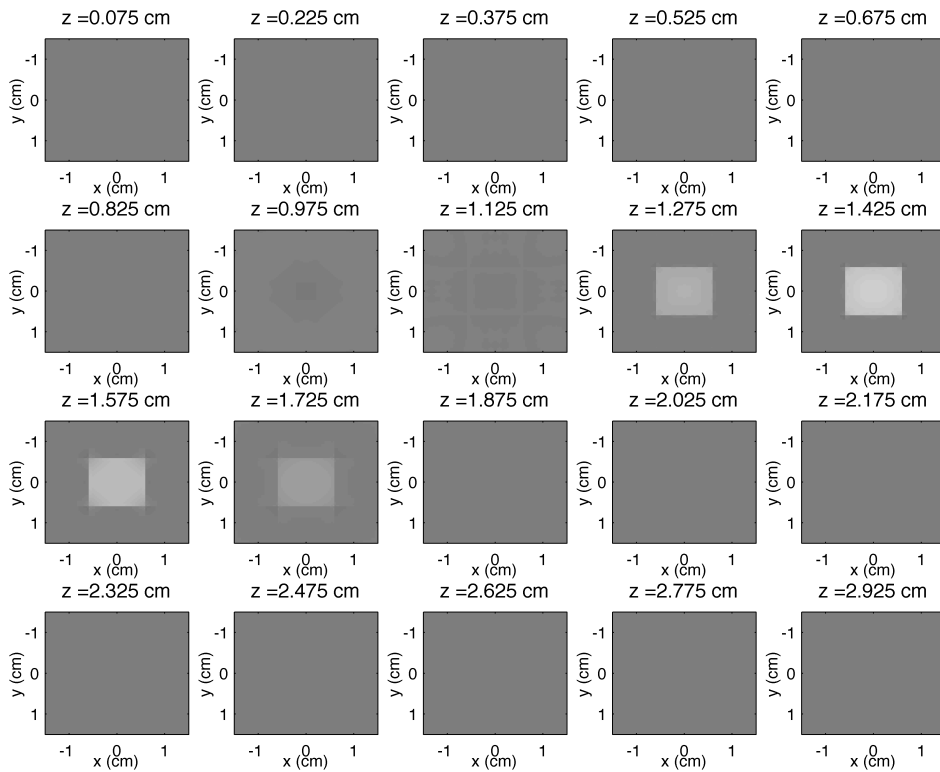
(b)

**Figure 4.5.** (a) Layers of the VOI of the image VII. (b) Reconstructed layers of the VOI of the image VII using CS. Black indicates a perturbation in squared wavenumber of  $10 \text{ cm}^{-2}$  and white a perturbation in squared wavenumber of  $-10 \text{ cm}^{-2}$ .





(a)



(b)

**Figure 4.6.** (a) Layers of the VOI of the image VII. (b) Reconstructed layers of the VOI of the image VII using the prefiltering method. Black indicates a perturbation in squared wavenumber of  $10 \text{ cm}^{-2}$  and white a perturbation in squared wavenumber of  $-10 \text{ cm}^{-2}$ .

### 4.3 THE EFFECT OF NOISY MEASUREMENTS

Since many approximations were made during the modeling process, and since in real life applications the measurements are always contaminated with noise, a relevant question is how different amounts of noise affect the reconstructions of CS and the prefiltering method. To partially answer this question, we selected two images, image II and image III of Table 3.4, and subsequently added noise with more variance to the measurements.

We added noise to each measurement real and imaginary part. The noise in the real part was taken from a zero mean gaussian distribution. Similarly for the noise in the imaginary part. In both cases the standard deviation of the noise varied from 0% to 11.67% of either the real or the imaginary part.

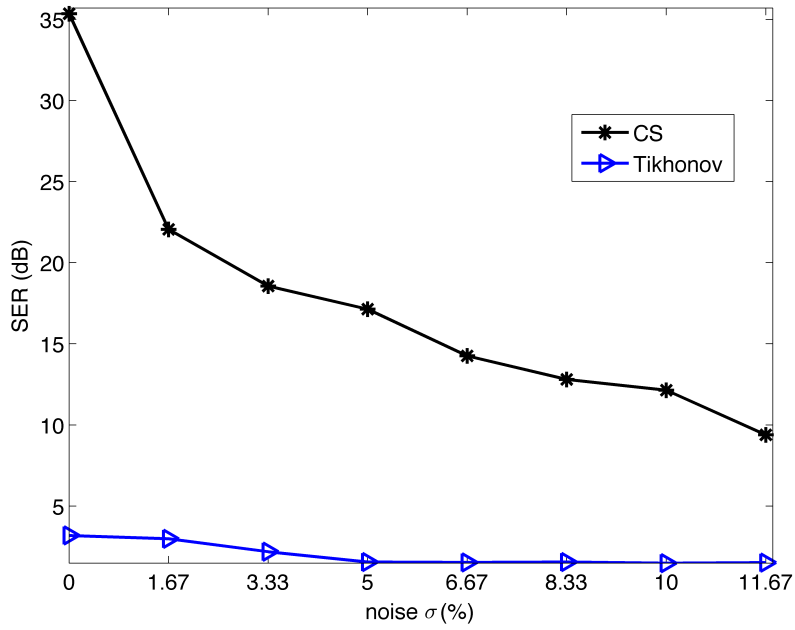
In the reconstructions of this section, we did not use any sparsifying transform. The regularization parameter of the IRLS algorithm was initialized to 1 and reduced by a factor of 10 after every 15 iterations until the regularization parameter reached  $10^{-7}$ .

To assess the robustness to noise of the CS method, we used the image II. Figure 4.7 shows the CS method reconstruction SER when the noise varied from 0% to 11.67%. As a matter of comparison, we also included a curve for Tikhonov regularization. The SERs in this graph are of reconstructed images of perturbation in squared wavenumber with respect to true images of perturbation in squared wavenumber, that is,  $SER(\Delta k^{2*}, \Delta k^2)$ .

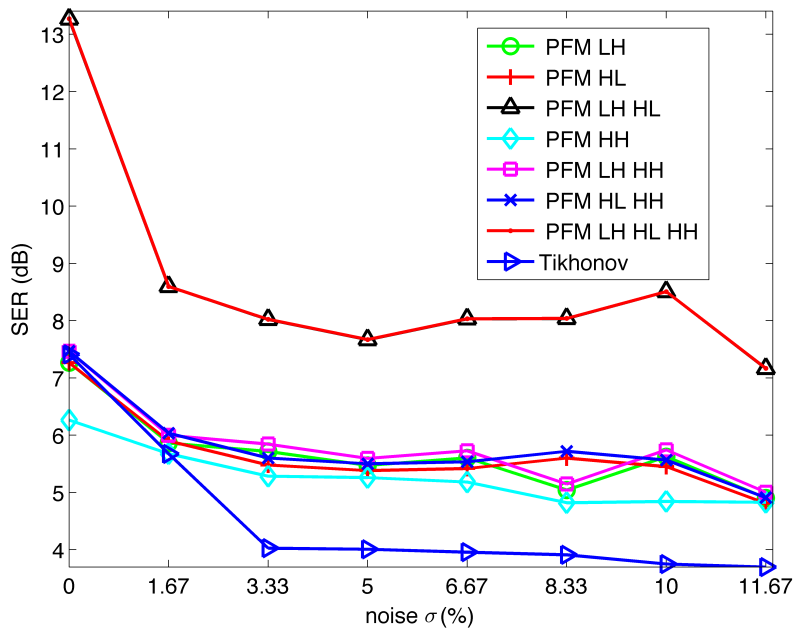
Figure 4.8 is the same as Figure 4.7, but in this case it is the robustness to noise of the prefiltering method that is evaluated, and the image used is image III.

The reconstruction SER of CS initially decayed rapidly from 35 dB to 22 dB, when the noise increased from 0% to 1.67%, but after that it decayed more slowly, reaching the mark of 10 dB when the noise reached the level of 11.67%. Similarly, the reconstruction SER of the prefiltering method using the LH, HL and HH filters initially decayed rapidly from 13 dB to 8 dB, when the noise increased from 0% to 1.67%, but after that it decayed more slowly, reaching the mark of approximately 7 dB when the noise reached the level of 11.67%.

Figure 4.8 also contains curves for the prefiltering method using subsets of the LH, HL and HH filters. Note that, at 0% noise, the reconstruction SERs of some prefiltering methods were approximately the same as the reconstruction SER of Tikhonov regularization, but after the 1.67% noise level, all possible prefiltering methods outperformed Tikhonov regularization.



**Figure 4.7.** The effect of different amounts of noise on the reconstruction of the image II using CS.



**Figure 4.8.** The effect of different amounts of noise on the reconstruction of the image III using the prefiltering method.

#### 4.4 COMPARISONS WITH THE STATE-OF-THE-ART TECHNIQUE

This section is intended to compare different variants of the proposed techniques with Tikhonov regularization. The images reconstructed in this section are the images I, III, IV, V, VIII, IX, X, XI, XII, XIII, XIV and XV. In the reconstructions of this section,

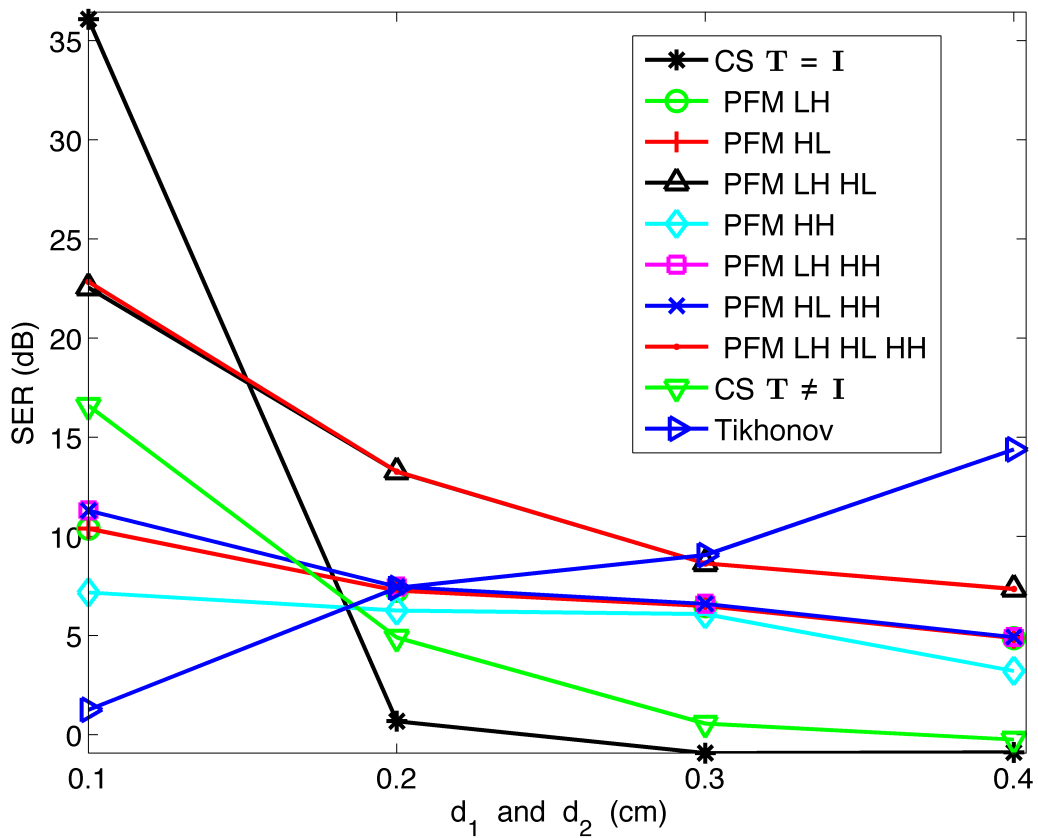
the regularization parameter of the IRLS algorithm was initialized to 1 and reduced by a factor of 10 after every 15 iterations until the regularization parameter reached  $10^{-7}$ .

Figures 4.9, 4.10 and 4.11 show the SER curves obtained for the different methods when we varied the sizes of a rectangular prism perturbation, two rectangular prism perturbations and an hexagonal prism perturbation, respectively. The SERs in these graphs are of reconstructed images of perturbation in squared wavenumber with respect to true images of perturbation in squared wavenumber, that is,  $\text{SER}(\Delta k^{2*}, \Delta k^2)$ .

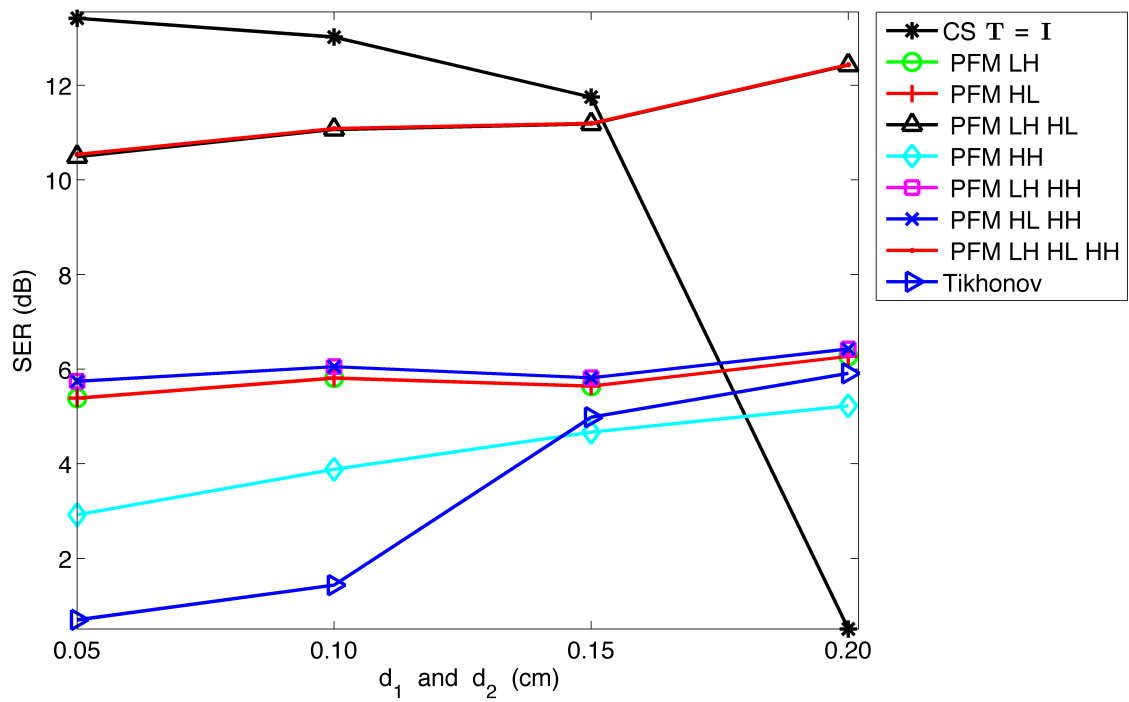
Notice that, in most cases, the SER of Tikhonov regularization was lower than the SER of one of the proposed techniques. The only exceptions being when the dimensions  $d_1$  and  $d_2$  of the rectangular prism perturbation were 0.3 cm or 0.4 cm, and when the dimensions  $d_1$  and  $d_2$  of the hexagonal prism perturbation were 0.4 cm. The SER of the CS method with no sparsifying transform was higher than the SER of Tikhonov regularization only when the dimensions  $d_1$  and  $d_2$  of the perturbations were 0.15 cm or less. This is due to the requirement that the image be sparse in its original domain. However, the highest SER obtained using this proposed method was more than twice the highest SER obtained using Tikhonov regularization.

Figure 4.9 includes an SER curve for the CS method using the sparsifying transform of Section 3.1.5. Compared to the SERs of the prefiltering method using different combinations of filters, the use of this method generally resulted in lower values of SER. If we consider only the prefiltering method using the LH, HL and HH filters, then it always resulted in lower levels of SER. One possible explanation is that the sparsifying transform of Section 3.1.5 uses all of the Haar filters, even the LL filter, which does not sparsify the image. The prefiltering method, on the other hand, uses only the filters that sparsify the image. Notice that, even without the DC component, and sometimes other parts of the image spectrum, the SERs of the prefiltering method using different combinations of filters, were still generally higher than the SERs of the CS method using the Haar DWT.

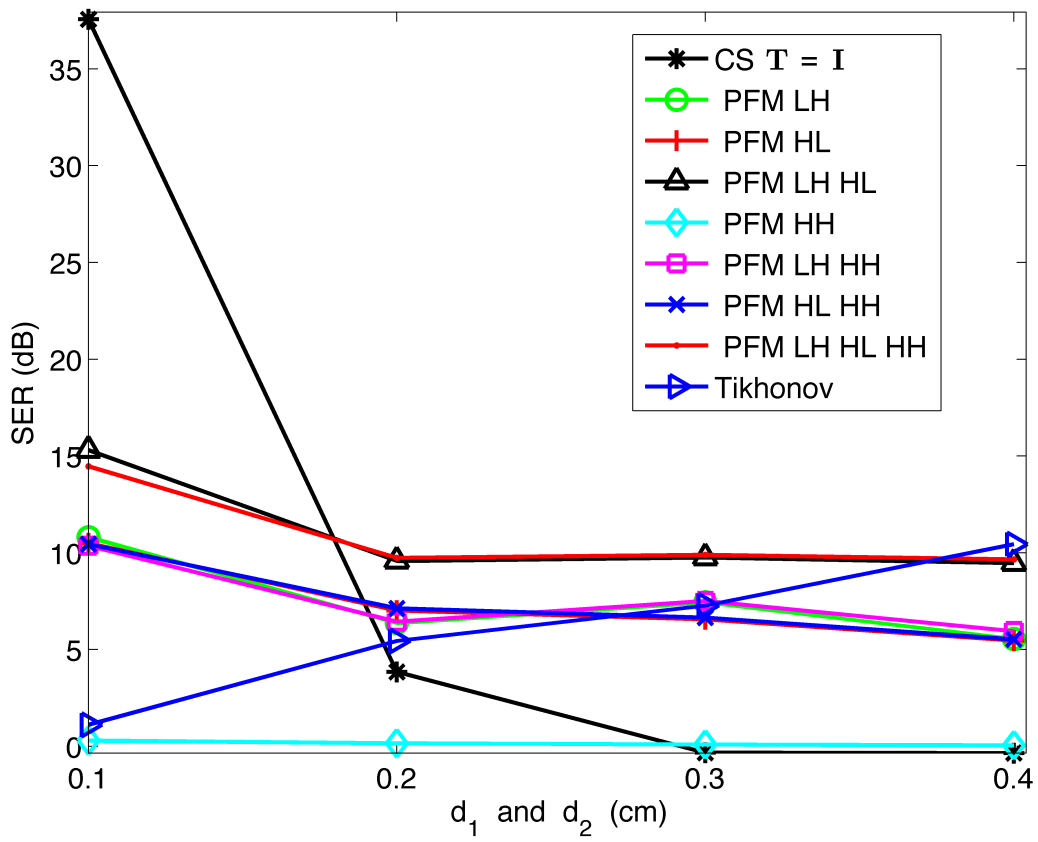
In general, the SERs of the prefiltering method were highest when all three of the LH, HL and HH filters were used. However, the maximum loss in SER caused by the removal of the HH filter was of 0.3 dB (Figure 4.9, for  $d_1$  and  $d_2$  equal to 0.1 cm). Losses of this magnitude were more common for the hexagonal prism perturbation, which has diagonal edges, and therefore more frequency content in the HH part of the spectrum.



**Figure 4.9.** SER curves obtained for different variants of the proposed techniques and for Tikhonov regularization when the size of a specific rectangular prism perturbation was varied.



**Figure 4.10.** SER curves obtained for different variants of the proposed techniques and for Tikhonov regularization when the sizes of two specific rectangular prism perturbations were varied.



**Figure 4.11.** SER curves obtained for different variants of the proposed techniques and for Tikhonov regularization when the size of a specific hexagonal prism perturbation was varied.

## 5 CONCLUSION

In Chapter 3, we showed that a particular formulation of the DOT problem satisfies the requirements for CS, and that this same formulation allows us to use the prefiltering method in DOT. We also designed numerical experiments to verify the usefulness of the proposed techniques. These numerical experiments consisted of a few proof of concept reconstructions, and graphs to compare different variants of the proposed techniques with the state-of-the-art.

In order to obtain more reconstructions in a lesser amount of time, the majority of the volumes reconstructed had dimensions 1 cm by 1 cm by 1 cm. In some experiments, this volume was increased to approximately 3 cm by 3 cm by 3 cm. Using the proposed techniques, we were able to reconstruct images containing the first volume with SERs as high as 55 dB, and images containing the second volume with SERs as high as 20 dB. The prefiltering method was able to increase the SER to about 9 dB in cases that the SER of CS was about -2 dB.

In general, the Thikhonov SER was higher than the other methods SERs for large perturbations, the CS SER was higher than the other methods SERs for small perturbations, and the prefiltering method SER was higher than the other methods SERs for medium sized perturbations. The SER of the prefiltering method using the Haar filters was higher than the SER of CS using the Haar DWT for all sizes of a specific perturbation. Also, the prefiltering method SER was highest when all three of the LH, HL and HH filters were used, but the loss in SER caused by the removal of the HH filter did not exceed 0.3 dB.

The effect of noise on the CS reconstruction was such that the reconstruction SER of a particular signal reduced from 35 dB to 10 dB when the standard deviation of the noise in the real and imaginary parts of the measurements achieved 12%. The effect of this same amount of noise on the prefiltering method reconstruction, on the other hand, was such that the reconstruction SER of another particular signal reduced from 13 dB to 7 dB.

Our numerical experiments confirm that CS and the prefiltering method can be used with our proposed model. In some cases, the reconstruction SER and the  $\ell_1$  norm indi-



cated that the CS algorithm was converging to the true solution. In the cases that it was not, we were able to increase the SER by using the prefiltering method.

## 5.1 LIMITATIONS

In the development of the proposed techniques, we assumed an homogeneous semi-infinite medium with a small perturbation in absorption coefficient, in magnitude and support.

The semi-infinite medium assumption is accurate if the tissue boundary has a small curvature where the sources and detectors are placed and if the other boundaries of the tissue are far away. Then the tissue can be approximated by a semi-infinite medium in a region close to the sources and detectors.

The assumption that the medium is homogeneous could be made if only the outermost layers of the cerebral cortex were imaged, since in this case the imaged volume could be considered approximately homogeneous [66]. In noninvasive studies, however, this assumption cannot be made [66].

The choice to reconstruct only perturbations in absorption coefficient, in its turn, is based on the fact that in many applications of DOT, the scattering coefficient changes negligibly with time while the absorption coefficient does not [66]. This is because the scattering coefficient is associated with the properties of cells, cell nuclei, cell organelles and surrounding fluids [29], while the absorption coefficient is associated with changes in oxygenated and deoxygenated hemoglobin, which in turn are associated with regional cerebral blood flow and angiogenesis in cancer [53], two of the most studied phenomenon using DOT.

Finally, the small perturbation assumption is associated with the Born approximation. The Born approximation assumes that the scattered field is small. This can only be true if the perturbation is small in magnitude and in support. The perturbations of absorption coefficient generally of interest in DOT range from  $0.02 \text{ cm}^{-1}$  to  $0.3 \text{ cm}^{-1}$  [68]. The Born approximation is valid only for perturbations with magnitude of about  $0.1 \text{ cm}^{-1}$  or less [68]. Also, the larger the perturbation is support, the lower the magnitude of the perturbation has to be in order to the Born approximation to be valid.

Our proposed techniques also assume that the changes in absorption coefficient are zero outside the volume where they are unknown, or at least until a distance corresponding to the assumed size of the Green's function. If this is not true, then the measurements will be contaminated by scattered fields coming from other locations, increasing the amount of noise.

In brain imaging, this assumption corresponds to considering that cortical areas

nearby the cortical area of interest are not activated by the task under study. In breast cancer imaging, the breast could be embedded in a box filled with matching material, such as an intralipid with optical properties similar to those of the breast [30], in which case the perturbations in the medium optical parameters outside of the breast would be zero.

## 5.2 FUTURE RESEARCH

One possible way of improving the model is to use the finite elements method (FEM) to derive the Jacobian matrix of the forward model. This would allow to consider much more complex geometries and initial distributions of optical parameters, which could be inferred by subject specific MRI images or a registered 3D atlas head model [22].

However, we showed only that a specific DOT Jacobian matrix satisfies the requirements for CS. This Jacobian matrix relates small changes in absorption coefficient in a semi-infinite homogeneous medium to changes in measurements in angular spectrum representation. In order to use the same techniques with a FEM forward matrix, one would have to show that the resulting forward matrix also satisfies the requirements for CS. Furthermore, the prefiltering method assumes that it is possible to obtain measurements associated with filtered versions of the image. One would also have to demonstrate this to be possible with the FEM forward model in order to use the prefiltering method.

Even though it has not been demonstrated yet that a FEM forward matrix satisfies the requirements for CS and that it is possible to obtain measurements associated with filtered versions of the image, the fact that we were able to show this for other DOT forward matrices, even if simpler ones, makes us hypothesize that it is also possible to show this for more generic matrices.

A natural follow up to our study would be to evaluate the proposed techniques with experimental data. Our numerical simulations could be replicated with a customly designed imaging phantom and a customly designed DOT instrument.

Other possibility would be to use them directly in brain imaging, for example, and compare their reconstructions with other similar, but more established, imaging modalities, such as fMRI. However, it may be necessary to first improve the model assumed by the techniques before using them effectively in human functional imaging.

# References

- [1] P. Aguiar and A. Ruibal. Multimodal imaging in lung cancer : It is time to change. *Journal of Lung Diseases and Treatment*, 1(1), 2015.
- [2] S. R. Arridge and W. R. B. Lionheart. Nonuniqueness in diffusion-based optical tomography. *Optics Letters*, 23(11):882, 1998.
- [3] W. U. Z. Bajwa. *New Information Processing Theory and Methods for Exploiting Sparsity in Wireless Systems*. PhD thesis, University of Wisconsin - Madison, 2009.
- [4] T. Bhowmik, H. Liu, Z. Ye, and S. Orintara. Dimensionality Reduction Based Optimization Algorithm for Sparse 3-D Image Reconstruction in Diffuse Optical Tomography. *Scientific Reports*, 6(February):22242, 2016.
- [5] D. A. Boas. *Diffuse photon probes of structural and dynamical properties of turbid media: Theory and biomedical applications*. PhD thesis, University of Pennsylvania, 1996.
- [6] D. A. Boas, D. H. Brooks, E. L. Miller, C. A. Dimarzio, M. Kilmer, R. J. Gaudette, and Q. Zhang. Imaging the body with diffuse optical tomography. *IEEE Signal Processing Magazine*, 18(6):57–75, 2001.
- [7] D. A. Boas, C. E. Elwell, M. Ferrari, and G. Taga. Twenty years of functional near-infrared spectroscopy: Introduction for the special issue. *NeuroImage*, 85:1–5, 2014.
- [8] W. G. Bradley. History of Medical Imaging. *Source: Proceedings of the American Philosophical Society*, 152(3):349–361, 2008.
- [9] J. W. Brewer. Kronecker Products and Matrix Calculus in System Theory. *IEEE Transactions on Circuits and Systems*, 25(9):772–781, 1978.
- [10] R. Bro. *Multi-way Analysis in the Food Industry*. PhD thesis, UvA, 1998. <https://pdfs.semanticscholar.org/f2bd/c868e33937a52c519bdf13e51a7affbc03c.pdf>.

- [11] M. Caffini, D. Contini, R. Re, L. M. Zucchelli, R. Cubeddu, A. Torricelli, and L. Spinelli. Functional near infrared spectroscopy and diffuse optical tomography in neuroscience. In Vikas Chaudhary, editor, *Advances in Brain Imaging*, chapter 4, pages 51–76. InTech, 2012.
- [12] E. Candes. Mathematics of sparsity (and a few other things). *ICM 2014 Proceedings, to appear*, pages 1–27, 2014.
- [13] E. J. Candès and Y. Plan. A probabilistic and RIPless theory of compressed sensing. *IEEE Transactions on Information Theory*, 57(11):7235–7254, 2011.
- [14] E. J. Candès and J. Romberg. Sparsity and incoherence in compressive sampling. *Inverse Problems*, 23(3):969–985, 2007.
- [15] N. Cao, A. Nehorai, and M. Jacobs. Image reconstruction for diffuse optical tomography using sparsity regularization and expectation-maximization algorithm. *Optics Express*, 15(21):13695, 2007.
- [16] V. Chan and A. Perlas. Basics of ultrasound imaging. In S. N. Narouze, editor, *Atlas of Ultrasound-Guided Procedures in Interventional Pain Management*, pages 13–19. Springer, 2011.
- [17] R. Chartrand and W. Yin. Iteratively reweighted algorithms for compressive sensing. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 3869–3872, 2008.
- [18] C. Chen, F. Tian, H. Liu, and J. Huang. Diffuse optical tomography enhanced by clustered sparsity for functional brain imaging. *IEEE Transactions on Medical Imaging*, 33(12):2323–2331, 2014.
- [19] H. Cheng. The fundamentals of compressed sensing. In *Sparse Representation, Modeling and Learning in Visual Recognition*, pages 21–53. Springer-Verlag London, 2015.
- [20] X. Cheng and D. A. Boas. Diffuse optical reflection tomography with continuous-wave illumination. *Optics Express*, 3(3):118–123, 1998.
- [21] D. Chitnis, R. J. Cooper, L. Dempsey, S. Powell, S. Quaggia, D. Highton, C. Elwell, J. C. Hebden, and N. L. Everdell. Functional imaging of the human brain using a modular, fibre-less, high-density diffuse optical tomography system. *Biomedical Optics Express*, 7(10):4275–4288, 2016.
- [22] R. J. Cooper, M. Caffini, J. Dubb, G. Fang, A. Custo, D. Tsuzuki, B. Fischl, W. Wells, I. Dan, and D. A. Boas. Validating atlas-guided DOT: A comparison of

- diffuse optical tomography informed by atlas and subject-specific anatomies. *NeuroImage*, 62(3):1999–2006, 2012.
- [23] T. Correia, A. Gibson, M. Schweiger, and J. Hebden. Selection of regularization parameter for optical topography. *Journal of Biomedical Optics*, 14(3):034044, 2009.
- [24] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok. Introduction to compressed sensing. In Yonina C. Eldar and Gitta Kutyniok, editors, *Compressed Sensing: Theory and Applications*, pages 1–64. Cambridge University Press, 2012.
- [25] P. J. Davis and P. Rabinowitz. *Methods of numerical integration*, volume 17. Academic Press, 1975.
- [26] H. Dehghani, S. Srinivasan, B. W. Pogue, and A. Gibson. Numerical modelling and image reconstruction in diffuse optical tomography. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 367(1900):3073–93, 2009.
- [27] B. P. V. Dileep, T. Das, and P. K. Dutta. Greedy algorithms for diffuse optical tomography reconstruction. *Optics Communications*, 410, 2018.
- [28] B. P. V. Dileep, T. Das, and P. K. Dutta. Modified CS-MUSIC for diffuse optical tomography using joint sparsity. *Optik*, 158:1478–1490, 2018.
- [29] T. Durduran, R. Choe, W. B. Baker, and A. G. Yodh. Diffuse optics for tissue monitoring and tomography. *Reports on Progress in Physics*, 73(7):076701, 2010.
- [30] T. Durduran, J. P. Culver, M. J. Holboke, X. D. Li, L. Zubkov, B. Chance, D. N. Pattanayak, and A. G. Yodh. Algorithms for 3D localization and imaging using near-field diffraction tomography with diffuse light. *OPTICS EXPRESS*, 4(8):247–262, 1999.
- [31] W. A. Edelstein, M. Mahesh, and J. A. Carrino. MRI: Time is dose - And money and versatility, 2010.
- [32] A. T. Eggebrecht, S. L. Ferradal, A. Robichaux-Viehoever, M. S. Hassanpour, H. Dehghani, A. Z. Snyder, T. Hershey, and J. P. Culver. Mapping distributed brain function and networks with diffuse optical tomography. *Nature Photonics*, 8(6):448–454, 2014.
- [33] A. Farina, M. Betcke, L. Di Sieno, A. Bassi, N. Ducros, A. Pifferi, G. Valentini, S. Arridge, and C. D’Andrea. Multiple-view diffuse optical tomography system based on time-domain compressive measurements. *Optics Letters*, 42(14):2822, 2017.
- [34] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing*. Birkhäuser, 2013.

- [35] D. Ganguly, S. Chakraborty, M. Balitanas, and T. Kim. Medical Imaging: A Review. *springer*, 78(February):504–516, 2010.
- [36] R. J. Gaudette. *Constrained Reconstruction Techniques for Diffuse Optical Tomography*. PhD thesis, Northeastern University, 2000.
- [37] R. J. Gaudette, D. H. Brooks, C. A. DiMarzio, M. E. Kilmer, E. L. Miller, T. Gaudette, and D. A. Boas. A comparison study of linear reconstruction techniques for diffuse optical tomographic imaging of absorption coefficient. *Physics in medicine and biology*, 45(4):1051–1070, 2000.
- [38] A. P. Gibson, J. C. Hebden, and S. R. Arridge. Recent advances in diffuse optical imaging. *Phys. Med. Biol. Phys. Med. Biol*, 50(50):1–43, 2005.
- [39] T. H. Glisson. *Introduction to Circuit Analysis and Design*. Springer, 2011.
- [40] M. Gockenbach. *Linear Inverse Problems and Tikhonov Regularization*. Mathematical Association of America, 2016.
- [41] C. Habermehl, J. Steinbrink, K. Müller, and S. Haufe. Optimizing the regularization for image reconstruction of cerebral diffuse optical tomography. *Journal of Biomedical Optics*, 19(9):096006, 2014.
- [42] Y. Hoshi and Y. Yamada. Overview of diffuse optical tomography and its clinical applications. *Journal of Biomedical Optics*, 21(9):091312, 2016.
- [43] T. J. Huppert, M. A. Franceschini, and D. A. Boas. Noninvasive imaging of cerebral activation with diffuse optical tomography. In R. D. Frostig, editor, *In Vivo Optical Imaging of Brain Function, Second Edition*, chapter 14, pages 393–433. CRC Press, 2009.
- [44] R. Ibrahim, S. Samian, M. Z. Mazli, M. N. Amrizal, and S. M. Aljunid. Cost of Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scan in UKMMC. *BMC Health Services Research*, 12(Suppl 1):P11–P11, nov 2012.
- [45] S. L. Jacques and B. W. Pogue. Tutorial on diffuse light transport. *Journal of Biomedical Optics*, 13(4):041302, 2008.
- [46] F. F. Jobsis. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*, 198(4323):1264–1267, 1977.
- [47] P. R. Johnston and R. M. Gulrajani. Selecting the corner in the L-curve approach to Tikhonov regularization. *IEEE Transactions on Biomedical Engineering*, 47(9):1293–1296, 2000.

- [48] A. C. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. IEEE Press, 1988.
- [49] V. C. Kavuri, Z. Lin, F. Tian, and H. Liu. Sparsity enhanced spatial resolution and depth localization in diffuse optical tomography. *Biomedical Optics Express*, 3(5):943, 2012.
- [50] S. Kim, K. Koh, M. Lustig, S. Boyd, and G. Dimitry. An Interior-Point Method for Large-Scale L1-Regularized Least Squares. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):606–617, 2007.
- [51] D. E. Knuth. *The Art of Computer Programming, Vol. 1: Fundamental Algorithms*, volume 1. Addison-Wesley Professional, 1997.
- [52] J. M. Lasker, C. J. Fong, D. T. Ginat, E. Dwyer, and A. H. Hielscher. Dynamic optical imaging of vascular and metabolic reactivity in rheumatoid joints. *Journal of Biomedical Optics*, 12(5):052001, 2007.
- [53] O. Lee, J. M. Kim, Y. Bresler, and J. C. Ye. Compressive diffuse optical tomography: Noniterative exact reconstruction using joint sparsity. *IEEE Transactions on Medical Imaging*, 30(5):1129–1142, 2011.
- [54] W. Li and N. Wang. Transforming the smv model into mmv model based on the characteristics of wavelet coefficients. *IET Signal Processing*, 11(3):275–284, 2017.
- [55] M. A. Lindquist. The Statistical Analysis of fMRI Data. *Statistical Science*, 23(4):439–464, 2008.
- [56] H. Lu, X. Long, and J. Lv. A fast algorithm for recovery of jointly sparse vectors based on the alternating direction methods. *Journal of Machine Learning Research*, 15:461–469, 2011.
- [57] W. Lu, D. Lighter, and I. B. Styles. L1-norm based nonlinear reconstruction improves quantitative accuracy of spectral diffuse optical tomography. *Biomed. Opt. Express*, 9(4):1423–1444, Apr 2018.
- [58] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed sensing MRI. *Signal Processing Magazine, IEEE*, 25(2):72–82, 2008.
- [59] J. H. Mathews and K. D. Fink. *Numerical Methods Using Matlab*. Prentice-Hall, 2004.
- [60] C. J. Miosso, R. von Borries, M. Argàez, L. Velazquez, C. Quintero, and C. M. Potes. Compressive sensing reconstruction with prior information by iteratively reweighted least-squares. *IEEE Transactions on Signal Processing*, 57(6):2424–2431, 2009.

- [61] C. J. Miosso, R. von Borries, and J. H. Pierluissi. Compressive sensing method for improved reconstruction of gradient-sparse magnetic resonance images. In *Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, pages 799–806, 2009.
- [62] C. J. Miosso, R. von Borries, and J. H. Pierluissi. Compressive Sensing With Prior Information: Requirements and Probabilities of Reconstruction in  $l_1$ -Minimization. *IEEE Transactions on Signal Processing*, 61(9):2150–2164, 2013.
- [63] Nature. A cheaper, quieter MRI machine. *Nature*, 501:138, sep 2013.
- [64] H. Niu, Z. Lin, F. Tian, S. Dhamne, and H. Liu. Comprehensive investigation of three-dimensional diffuse optical tomography with depth compensation algorithm. *Journal of Biomedical Optics*, 15(4):046005, 2010.
- [65] H. Niu, F. Tian, Z. Lin, and H. Liu. Development of a compensation algorithm for accurate depth localization in diffuse optical tomography. *Optics letters*, 35(3):429–431, 2010.
- [66] H. Obrig and A. Villringer. Beyond the visible - Imaging the human brain with light. *Journal of Cerebral Blood Flow and Metabolism*, 23(1):1–18, 2003.
- [67] S. Okawa, Y. Hoshi, and Y. Yamada. Improvement of image quality of time-domain diffuse optical tomography with  $l_p$  sparsity regularization. *Biomedical Optics Express*, 2(12):3334, 2011.
- [68] M. A. O’Leary. *Imaging with Diffuse Photon Density Waves*. PhD thesis, University of Pennsylvania, 1996.
- [69] J. M. Ollinger and J. A. Fessler. Positron-emission tomography. *IEEE Signal Processing Magazine*, 14(1):43–55, 1997.
- [70] A. V. Oppenheim and R. W. Schaffer. *Discrete-Time Signal Processing*. Prentice Hall Press, Upper Saddle River, NJ, USA, 3rd edition, 2009.
- [71] J. Prakash, C. B. Shaw, R. Manjappa, R. Kanhirodan, and P. K. Yalavarthy. Sparse Recovery Methods Hold Promise for Diffuse Optical Tomographic Image Reconstruction. *Selected Topics in Quantum Electronics, IEEE Journal of*, 20(2):74–82, 2014.
- [72] P. Prandoni and M. Vetterli. *Signal Processing for Communications*. Free Online Version, 2008. Available at <http://www.sp4comm.org/download.html>.



- [73] S. Puntanen, G. P. H. Styan, and J. Isotalo. Notation. In *Matrix Tricks for Linear Statistical Models*, pages 427–434. Springer, 2011. Available at <http://www.sis.uta.fi/tilasto/matrixtricks/>.
- [74] M. Rani, S. B. Dhok, and R. B. Deshmukh. *A Systematic Review of Compressive Sensing: Concepts, Implementations and Applications*, 2018.
- [75] A. Ribes and F. Schmitt. Linear inverse problems in imaging. *IEEE Signal Processing Magazine*, 25(4):84–99, 2008.
- [76] Justin K. Romberg. Compressive sensing by random convolution. *SIAM J. Imaging Sciences*, 2:1098–1128, 2009.
- [77] F. Scholkmann, S. Kleiser, A. J. Metz, R. Zimmermann, J. Mata Pavia, U. Wolf, and M. Wolf. A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *NeuroImage*, 85:6–27, 2014.
- [78] M. Schweiger, A. Gibson, and S. R. Arridge. Computational aspects of diffuse optical tomography. *COMPUTING IN SCIENCE & ENGINEERING*, 5:33–41, 2003.
- [79] C. B. Shaw and P. K. Yalavarthy. Effective contrast recovery in rapid dynamic near-infrared diffuse optical tomography using L1-norm-based linear image reconstruction method. *Journal of biomedical optics*, 17(8):086009, 2012.
- [80] C. B. Shaw and P. K. Yalavarthy. Prior image-constrained L1-norm-based reconstruction method for effective usage of structural information in diffuse optical tomography. *Optics letters*, 37(20):4353–5, 2012.
- [81] G. Strangman, D. A. Boas, and J. P. Sutton. Non-invasive neuroimaging using near-infrared light. *Biological Psychiatry*, 52(7):679–693, 2002.
- [82] H. J. Sun, H. W. Hou, C. C. Chou, and W. C. Fang. Diffuse optical tomography image reconstruction based on sparse recovery methods. In *IEEE Biomedical Circuits and Systems Conference: Engineering for Healthy Minds and Able Bodies, BioCAS 2015 - Proceedings*, 2015.
- [83] L. Sun, J. Liu, J. Chen, and J. Ye. Efficient Recovery of Jointly Sparse Vectors. *Advances in Neural Information Processing Systems 22*, pages 1812–1820, 2009.
- [84] D. Sundararajan. *Discrete Wavelet Transform: A Signal Processing Approach*. John Wiley and Sons Singapore, 2015.
- [85] M. Süzen, A. Giannoula, and T. Durduran. Compressed sensing in diffuse optical tomography. *Optics express*, 18(23):23676–23690, 2010.

- [86] P. P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.
- [87] S. van de Ven, S. Elias, A. Wiethoff, M. van der Voort, A. Leproux, T. Nielsen, B. Brendel, L. Bakker, M. van der Mark, W. Mali, and P. Luijten. Diffuse optical tomography of the breast: Initial validation in benign cysts. *Molecular Imaging and Biology*, 11(2):64–70, 2009.
- [88] J. Villares. *Sample Covariance Based Parameter Estimation For Digital Communications*. PhD thesis, Polytechnic University of Catalonia, 2005. Available at <https://spcom.upc.edu/>.
- [89] A. von Luhmann. Design and evaluation of a system for mobile brain activity measurements using functional near-infrared spectroscopy. Master’s thesis, Karlsruhe Institute of Technology, 2014.