

**A novel method to increase depth of imaging in
Optical Coherence Tomography using
ultrasound**

by
Pedro Fernando Pereira Bogado

**A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of**

Master of Science

Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg

Copyright © 2012 by Pedro Fernando Pereira Bogado

Abstract

Optical coherence tomography (OCT) is a biomedical imaging technique with many current applications. A limitation of the technique is its shallow depth of imaging. A major factor limiting imaging depth in OCT is multiple-scattering of light. This thesis proposes an integrated computational imaging approach to improve depth of imaging in OCT. In this approach ultrasound patterns are used to modulate the refractive index of tissue. Simulations of the impact of ultrasound on the refractive index are performed, and the results are shown in this thesis. Simulations of the impact of the modulated refractive index on the propagation of light in tissue are needed. But there is no suitable simulator available. Thus, we implemented a Monte Carlo method to solve integral equations that could be used to perform these simulations. Results for integral equations in 1-D and 2-D are shown.

Acknowledgements

Many thanks to the following persons for their role in the completion of this thesis.

- My adviser Prof. Sherif Sherif whose guidance was invaluable in the process of this research
- My examination committee Prof. Cyrus Shafai and Prof. Stephen Pistorius for the valuable feedback
- Amy Dario for all her help and advice
- Mount-first Ng for all his help with computer related issues
- My parents for all their support
- My lovely wife Liza who was always there to support and encourage me

Contents

Abstract	ii
Acknowledgements	iii
List of Tables	vi
List of Figures	vii
1 Introduction	1
1.1 Organization of this thesis	2
1.2 Contributions of this thesis	3
1.3 Related publications	4
2 Optical Coherence Tomography	5
2.1 Introduction	5
2.2 Time domain OCT	7
2.3 Fourier domain OCT	9
2.3.1 Spectral FD-OCT	9
2.3.2 Swept source FD-OCT	11
2.4 Applications of OCT	12
2.5 Multiple-scattering effect on OCT signal	14
3 Integrated Computational Imaging for OCT	16
3.1 Introduction	16
3.2 OCT imaging in the presence of ultrasound	17
3.3 A novel method: Ultrasound patterns	18
3.4 Delivery of ultrasound and OCT beams	20
3.5 Conclusion	22
4 Relationship between pressure and the refractive index	23
4.1 Introduction	23
4.2 Compression due to acoustic pressure	23
4.3 The relationship between refractive index and density	24
4.3.1 Polarization of materials	25
4.4 The Lorentz field and the Debye Equation	27
4.4.1 Derivation of the Lorentz field	27
4.4.2 Derivation of the Debye Equation	29
4.5 Derivation of the Classius-Mossotti equation	30
4.6 Conclusion	32

5	Results from the simulations of ultrasound patterns	34
5.1	Introduction	34
5.2	COMSOL setup	34
5.3	Ultrasound safety considerations	37
5.4	Results	37
5.5	Conclusions	43
6	Monte Carlo Methods	44
6.1	Introduction	44
6.2	Why use a Monte Carlo approach?	45
6.3	History of Monte Carlo methods	47
6.4	Random number generation	48
6.4.1	Classes of pseudo-random number generators	48
6.5	Monte Carlo integration	50
6.5.1	Assessment of the estimate	51
6.6	Importance Sampling	52
6.7	Sequential Importance sampling	53
6.8	Markov Chain Monte Carlo	54
6.8.1	Evaluating performance of MCMC	56
6.9	Reversible Jump Markov Chain Monte Carlo	58
6.9.1	Evaluating performance of RJMCMC	59
6.10	Summary	59
7	Implementation of a RJMCMC solver for Fredholm integral equations of the second kind	61
7.1	Introduction	61
7.2	Linear integral equations	62
7.3	Monte Carlo methods to solve integral equations	63
7.4	A solution based on sequential importance sampling	64
7.4.1	Limitation of the sequential importance sampling approach	65
7.5	A solution based on RJMCMC	66
7.5.1	Local Solution through RJMCMC	69
7.6	Implementation of a RJMCMC solver for Fredholm integral equations of second kind	70
7.7	Conclusion	75
8	Conclusions and future work	77
8.1	Conclusions	77
8.2	Future work	79
A	MATLAB code for Monte Carlo simulations	81
	References	89
	Subject Index	93

List of Tables

6.1	Comparison of convergence rate of Monte Carlo with different quadrature rules	46
7.1	Results obtained for Eq. (7.18) with 10 runs of 300,000 samples.	71
7.2	Results obtained for Eq. (7.19) with 10 runs of 300,000 samples used for the estimation.	72
7.3	Results obtained for Eq. (7.20) with 10 runs of 300,000 samples used for the estimation.	73
7.4	Results obtained for Eq. (7.21) with 10 runs of 300,000 samples use for the estimation.	74
7.5	Results obtained for Eq. (7.22) averaging 10 runs of 300,000 samples. . . .	74
7.6	Results obtained for equation 7.23 with 10 runs of 4 million samples used for the estimation.	75

List of Figures

2.1	Different imaging modalities with their resolution and penetration depth. OCT fills a gap between ultrasound imaging and confocal microscopy. . . .	6
2.2	TD-OCT scheme with Michelson interferometer.	8
2.3	Class I reflectance represent that obtained from ballistic photons. Class II reflectance is obtained from photons that undergo more than one scattering event. It is clearly seen that after a certain depth Class II reflectance is dominant.	14
3.1	A block diagram of an integrated computational imaging system	16
3.2	A standing wave ultrasound pattern generated by two ultrasound sources in a homogeneous medium	19
3.3	Parallel orientation between ultrasound wave and OCT beam. OCT beam is shown by an arrow, ultrasound wave shaded cone.	20
3.4	Perpendicular orientation between ultrasound wave and OCT beam. OCT beam is shown by an arrow , ultrasound wave shaded cone.	20
3.5	A proposed design for a probe capable of delivering ultrasound and OCT beams to the sample in a coordinated manner.	21
4.1	Spherical region for calculation of the electric field. It's big enough to include a number of particles, yet small enough to consider the external region to be homogeneous. Some fields are also shown.	28
5.1	Screen capture from COMSOL Multiphysics showing the geometry of the model and the material selected for the homogeneous medium.	35
5.2	Screen capture from COMSOL Multiphysics showing the pressure sources applied to the lower boundary of the model.	36
5.3	Refractive index of mediums used in simulations. Homogeneous medium with $n = 1.33$ on top (fig. (a)), and inhomogeneous medium with n varying between 1.3 to 1.5 bottom (fig. (b)).	39
5.4	Fig. (a) shows the effect of an ultrasound plane wave on the refractive index of the homogeneous medium. Fig. (b) shows the effect of an ultrasound plane wave on the refractive index of the inhomogeneous medium.	40
5.5	Effect of an ultrasound standing-wave pattern on the refractive index. Fig. (a) shows the effect on the homogeneous medium, while fig. (b) shows the effect on the inhomogeneous medium.	41

5.6	Pressure applied to the mediums to generate the ultrasound standing-wave patterns shown in fig. 5.5	42
6.1	An example of importance sampling. The samples are concentrated in the area of higher value of the distribution. There are still a few samples in areas of lower value.	52
7.1	The curves of the analytical solution and the estimates for Eq. (7.18) are on top of each other in the graph. Points of evaluation are shown.	71
7.2	Figure shows the curves of the estimate and analytical solution on top of each other.	72

Chapter 1

Introduction

Optical Coherence Tomography (OCT) is an optical imaging technique used mostly in biomedical optics and medicine. OCT provides high resolution cross-sectional imaging of the internal structure of tissues by measuring backscattered light. It is capable of imaging in situ and in real time with axial resolutions between $1 \mu\text{m}$ and $15 \mu\text{m}$ [1]. Its unique combination of resolution and penetration depth makes OCT a very promising and interesting imaging technique.

OCT was originally applied in ophthalmology where it has already established itself as a reliable diagnostic tool. Many more applications of OCT have been developed since its inception. Some of OCT applications include cancer detection in a variety of tissues, optical biopsy, monitoring of coronary diseases, gastrointestinal imaging, caries detection in dentistry, and more. Aside from its applications in biomedical imaging, there are a also number of applications in non-destructive testing of materials. OCT can be used for non-destructive testing of polymers, ceramics, and art objects. However, in the applications outside of ophthalmology a limitation of OCT is its shallow depth of imaging.

The main factor limiting imaging depth in OCT is multiple scattering of light. The OCT model assumes that only single backscattered photons provide useful information about the tissue being imaged. But deeper in tissue the contribution of multiple-scattered photons to the detected signal dominates that of single-scattered ones, reducing the amount

of useful information obtained from these depths. This thesis presents a novel way to reduce the effect of multiple scattering in OCT imaging. By applying ultrasound patterns to modulate tissue's refractive index, multiple-scattering at the detector can be reduced and OCT imaging depth increased. The modulation of the refractive index is considered as a step in an integrated computational imaging approach to OCT. The goal of this thesis is to study the feasibility of an integrated computational imaging OCT system that uses ultrasound patterns to modulate the refractive index of tissue. To do this simulations of the effect of acoustic pressure on the refractive index are performed and the results discussed.

To assess the performance of these patterns in reducing multiple-scattering a simulation of light interaction with tissue in the presence of the patterns is needed. However, there is not a suitable simulation software available to perform such work. It would require a lot of computational resources to simulate the interaction of light and modulated tissue in a volume as large as required in the simulation software available. It would also take a lot of time. Thus we decided to develop our own simulator. We want information on the whole optical field, intensity and phase. So we need to solve the integral equation of scalar scattering. A Monte Carlo method approach is selected to solve the integral equation of scalar scattering. We decided to use the Monte Carlo method for its efficiency in high dimensional problems, and because it is possible to solve the integral equation at specific points of the domain with it. Some 1-D and 2-D integral equations are solved to test the performance of the solver.

1.1 Organization of this thesis

The thesis is organized in the following way. In chapter 2 we discuss Optical Coherence Tomography. We present a description of the main classes of OCT systems, discuss some applications of OCT in various fields, and show how multiple-scattering affects this technique. In chapter 3 we present the novel idea of an *integrated computational OCT system*. In chapter 4 we present a discussion on the relationship of the refractive index and applied

acoustic pressure. Knowledge of this relationship is needed to successfully simulate the effect of an ultrasound beam on the refractive index of tissue. Chapter 5 presents the results of the simulation of the effect of ultrasound patterns on the refractive index of tissue and a discussion of these results. After this chapter we need to perform simulations of the interaction of light and tissue modulated by the ultrasound patterns, however we found that there is not an available simulator that is adequate for our requirements. So we decide to develop our own simulator based on the solution of the integral equation of scattering. The rest of this thesis deal with the development of that simulator. In chapter 6 we introduce some required concepts about Monte Carlo methods, and in chapter 7 these concepts come together for the implementation of a Reversible Jump Markov chain Monte Carlo solver for Fredholm integral equations of the second kind. Finally we present our conclusions in chapter 8.

1.2 Contributions of this thesis

The contributions of this thesis are as follows: We introduce the idea of an integrated computational imaging OCT system. This system uses ultrasound patterns to modulate the refractive index of tissue. We performed simulations of the application of these ultrasound patterns to homogeneous and inhomogeneous tissue. The results of these simulations show that:

1. The ultrasound patterns produce a meaningful change on the refractive index,
2. The change in the refractive index can be obtained within safety levels of application of ultrasound to tissue,
3. Acoustic scattering inside an inhomogeneous medium does not change the ultrasound pattern much, thus it is possible to design the ultrasound pattern considering only homogeneous tissue.

These results indicate that an integrated computational OCT system that uses ultrasound to modulate the refractive index of tissue is viable. In the second part of this thesis we implemented a Reversible Jump Markov chain Monte Carlo (RJCMC) method to solve Fredholm integral equations of the second kind. We tested the implemented RJCMC solver on several 1-D integral equations obtaining very accurate results. We also tested the implemented RJCMC solver on 2-D integral equations, obtaining accurate results and proving that the implemented RJCMC integral equation solver can be applied to problems of higher dimensions. We perform an analysis of error for our implemented RJCMC solver of Fredholm integral equations of the second kind obtaining the relative error of the estimate and its variance.

1.3 Related publications

The research done in this thesis was also of use in the production of the following papers:

1. Pedro F. Pereira and Sherif S. Sherif, "Design of an optimum ultrasound pattern to minimize multiple-scattered light reflected from inhomogeneous tissue", *Proc. SPIE*, vol. 8223, 2012
2. Pedro F. Pereira and Sherif S. Sherif, "Simulation of the interaction of light and tissue in a large volume using a Markov chain Monte Carlo method", *Proc. SPIE Photonics North*, Montreal, 2012
3. S. S. Sherif and P. Pereira, "Methods and systems for use in imaging using interferometry" (PCT patent application filed) 2011.

Chapter 2

Optical Coherence Tomography

2.1 Introduction

Optical Coherence Tomography (OCT) is an optical imaging technique used mostly in biomedical optics and medicine. OCT provides high resolution cross-sectional imaging of the internal structure of tissues by measuring backscattered light. It is capable of imaging in situ and in real time with axial resolutions between $1\ \mu\text{m}$ and $15\ \mu\text{m}$ [1]. Its unique combination of resolution and penetration depth makes OCT a very promising and interesting imaging technique.

OCT has common features with ultrasound and confocal microscopy. It is useful to compare these technologies in order to highlight unique characteristics. Clinical ultrasound has a resolution varying between 0.1 mm and 1 mm depending on the frequency used for imaging. At standard ultrasound frequencies sound waves are only lightly attenuated in tissue and hence it is possible to image a few centimetres in tissue. On the other hand, confocal microscopy is capable of extremely high resolution, of around $1\ \mu\text{m}$. But optical scattering can significantly degrade image signal and contrast, hence it is only possible to image a few hundred micrometres in tissue [1].

OCT stands at the middle ground between ultrasound and confocal microscopy. See Fig. 2.1 It is capable of imaging with axial resolutions between 1 and $15\ \mu\text{m}$. At the same

time, it is possible to penetrate a few millimetres in tissue using OCT. This shallow imaging depth is a great disadvantage of OCT, yet since it is an optical technology, it is possible to use endoscopes and catheters to perform imaging inside the body.

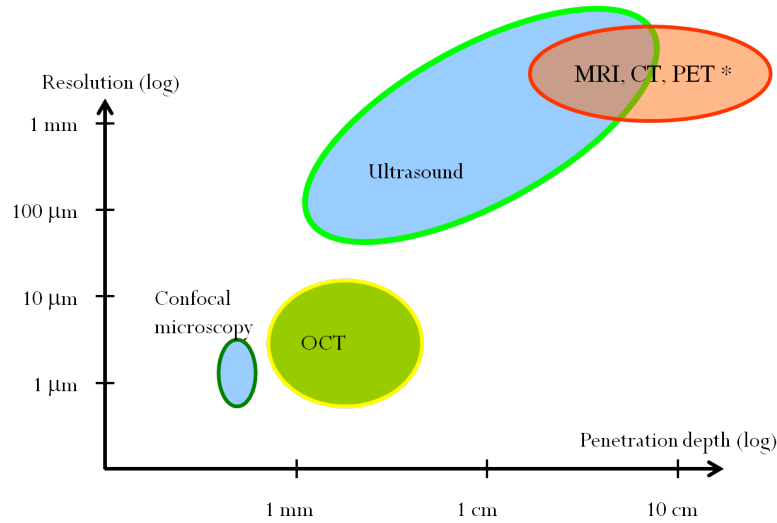


Figure 2.1. Different imaging modalities with their resolution and penetration depth. OCT fills a gap between ultrasound imaging and confocal microscopy.

As mentioned before, many applications of OCT are found in the biomedical imaging field. OCT is an established technique in ophthalmology, where it is easy to image using an optical technique due to the transparency of the eye. A few other applications of OCT include the assessment of coronary diseases, optical biopsy and cancer detection, and applications in dentistry [2–5]. There are also a number of applications being developed outside of the medical imaging field [6].

OCT can be classified as either time-domain OCT (TD-OCT) and Fourier-domain OCT (FD-OCT), depending on how the imaging is done. There are also a number of extensions to OCT that allow the technique to obtain some functional information from the tissues being imaged, e.g.: Doppler OCT, polarization sensitive OCT and combinations of OCT with other techniques like fluorescence imaging and spectroscopic measuring. In the next sections we shall present the more important variants of OCT, namely TD-OCT and FD-OCT.

2.2 Time domain OCT

OCT is based on an optical measurement technique known as low-coherence interferometry [1, 7]. Interferometry techniques perform correlation between light that is backscattered from tissue and light that has traveled a known distance in a reference arm. Interferometry measures the optical field rather than light intensity. A first approach toward tomographic imaging was presented by Fercher at the ICO-15 SAT conference in 1990. Fujimoto and co-workers pioneered OCT using a fibre optic Michelson interferometer in 1991. The first *in vivo* tomograms of human retina were obtained independently by Fercher *et al* and Swanson *et al* in 1993 [7].

A schematic of a Michelson interferometer for a TD-OCT system is shown in Fig. 2.2. Incident light from the source is separated into a reference beam and a signal beam which travel different distances in the two arms of the interferometer. The output is the summation of the reference field and the signal field, and a detector measures the intensity of the output. If we call the reference field $E_r(t)$ and the signal field $E_s(t)$, and if we consider a coherent light source then we can write,

$$I_o \sim |E_r(t)|^2 + |E_s(t)|^2 + 2E_r(t)E_s(t)\cos(2k\Delta L)$$

ΔL is the path length difference between the signal and reference arm in the interferometer. If the reference path length is scanned interference fringes will be generated as a function of time. When using low coherence light these interference fringes will only be generated if the path length of the reference and signal arms is matched within the coherence length of the source.

Operation of a TD-OCT system is based on the occurrence of interference fringes when the optical path length of the signal and reference arms coincide within the round trip coherence length, l_c . The coherence length is dependent on the wavelength of the light source as well as on its spectral width. If we assume a Gaussian spectrum for the light

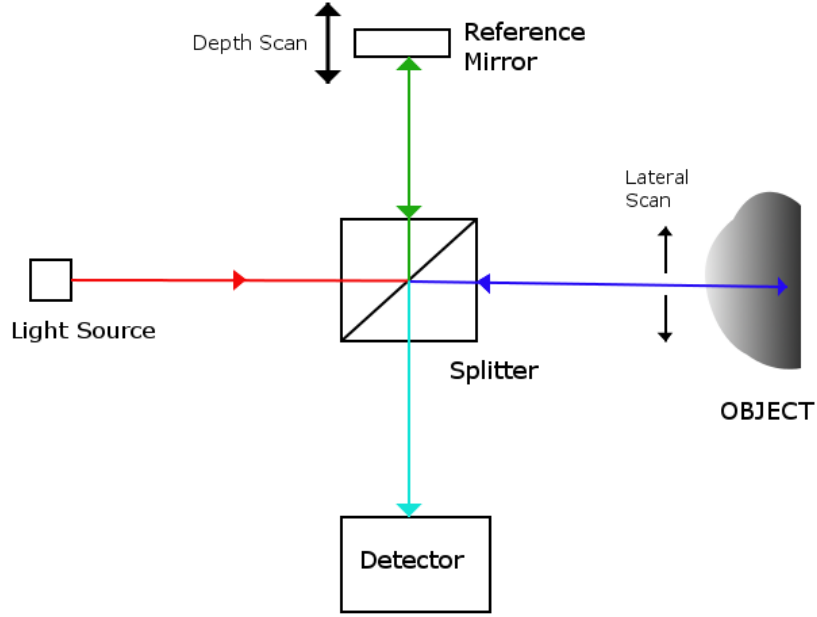


Figure 2.2. TD-OCT scheme with Michelson interferometer.

source then,

$$l_c = \frac{2 \ln 2}{\pi} \frac{\bar{\lambda}^2}{\Delta\lambda} \quad (2.1)$$

where $\bar{\lambda}$ is the mean wavelength of the source and $\Delta\lambda$ is its spectral width. As an example, using a superluminescent light diode (SLD) with $\bar{\lambda} = 820$ and $\Delta\lambda = 20$, both in nanometers (nm), yields a round trip coherence length and thus a depth resolution of $l_c \approx 15 \mu\text{m}$ [7].

Thus, TD-OCT is based on low time-coherence interferometry. It requires a broadband low coherence light source, as the coherence length is related to the bandwidth of the source. Scanning the length of the reference arm allows for the determination of the depth in tissue from which backscattered light is being detected. The depth scan is called, in analogy with ultrasound imaging, an A-scan. To produce a 2-D image a lateral scan is required. TD-OCT is capable of producing 3-D images of volumes by performing a raster scan, as well as its axial depth scan. The sensitivity of TD-OCT ranges between -60 and -90 dB of the incident light intensity, meaning intensities as low as -90 dB with respect with the incident light intensity can be detected [1].

2.3 Fourier domain OCT

FD-OCT differs from TD-OCT in that there is no need to scan the reference path length. FD-OCT measures backscattered light from all depths simultaneously. There are two types of FD-OCT. Spectral domain OCT uses an interferometer with a low-coherence light source and measures the interference fringes by using a spectrometer and a high speed line scan camera. The other approach is called Swept source domain OCT, or by some authors optical frequency domain imaging (OFDI). It uses an interferometer with a narrow-bandwidth, frequency-swept light source and detectors. It measures the interference fringes of the output as a function of time.

2.3.1 Spectral FD-OCT

The first implementation of Spectral domain OCT (SD-OCT) was reported in 1995. Even though the technique presented some advantages over TD-OCT it took some time to be accepted. In SD-OCT, the reference arm is kept stationary, and the depth information is obtained by a Fourier transform of the spectrally resolved interference fringes in the detection arm of a Michelson interferometer [8]. This approach provides an advantage in signal to noise ratio (SNR) when compared to TD-OCT.

Spectral domain OCT uses a very similar scheme as the one shown in Fig. 2.2 for TD-OCT. Unlike TD-OCT there is no need to scan the position of the reference mirror to obtain an A-scan of the sample. Instead in SD-OCT the cross-spectral density in the detection arm of the Michelson interferometer is measured by using a spectrometer. All spectral components of the OCT signal are captured simultaneously by the spectrometer [9].

An A-scan is obtained in SD-OCT by doing an inverse Fourier transform of the detected signal. As in TD-OCT, light from a broad-bandwidth source is split by the Michelson interferometer, part of it goes to the reference arm in which is reflected by a fixed mirror; the other part goes to the tissue to be imaged in which it is reflected back from internal

structures at different depths. The recombined light from both arms of the interferometer is then spectrally separated. The detected interference signal at the spectrometer can be expressed as [8],

$$I(k) = I_r(k) + 2\sqrt{I_r(k)I_s(k)} \sum_n \alpha_n \cos k z_n + I_s(k) \quad (2.2)$$

where $I_r(k)$ and $I_s(k)$ are the wavelength-dependent intensities reflected from the reference and sample respectively and k is the wavenumber. The second term in the right-hand side represents the interference between the reference and sample arms of the interferometer. α_n is the square root of the sample reflectivity at a depth z_n . The depth information is obtained by performing an inverse Fourier transform of Eq. (2.2), which yields [8],

$$|FT^{-1}[I(k)]|^2 = \Gamma^2(z) \otimes \left\{ \delta(0) + \sum_n \alpha_n^2 \delta(z - z_n) + \sum_n \alpha_n^2 \delta(z + z_n) + O[I_s^2/I_r^2] \right\} \quad (2.3)$$

where $\Gamma^2(z)$ representing the envelope of the coherence function. The first term within braces is the autocorrelation signal of the reference arm and has magnitude 1. The second and third term within braces on the right-hand side represent the interference between the reference and sample arms. They provide two mirror images of which only one is kept. The final term on the right-hand side represents autocorrelation noise from the sample arm. I_s and I_r represent the intensity reflected back from the reference and sample arm respectively.

The greatest advantage provided by Spectral-domain OCT over TD-OCT is an considerable increase in imaging speed. Imaging speeds can be 50 to 100 times faster [1]. This increase in imaging speed enables the possibility of higher resolution images, as it is possible to do more axial scans. It can also increase the amount of cross-sectional images acquired to produce 3D volumes. The increase in imaging speed is largely due to an increase in sensitivity when using spectral domain detection. SD-OCT can show an improvement of 21.7 dB in sensitivity when compared to TD-OCT [8].

2.3.2 Swept source FD-OCT

Although SD-OCT provides high-speed imaging and a better SNR than TD-OCT it also has some problems. For example, the use of array detectors may cause problems associated with phase washout by changes in the sample arm length during the pixel integration time. Also it is difficult to implement balanced detection and polarization diversity, which may be needed in some applications, with the scheme. An alternative way of implementing a Fourier domain OCT, hence maintaining the benefits of high-speed imaging and SNR, is optical frequency domain imaging (OFDI) [10] also called swept source OCT. It is based on frequency-domain reflectometry and uses a wavelength tunable laser and single-element detectors.

Again we can refer to Fig. 2.2 as the scheme is very similar. The main differences in OFDI case is the use of a wavelength tunable laser source, and the reference mirror being fixed. Operation can be described in the following way. The output from the light source is split between the reference arm and the sample arm. Interference between the reference and sample arms is detected using a square-law photodetector while the wavelength of the monochromatic source is changed or swept. The path lengths of the reference and sample arm are maintained constant. The axial reflectivity profile is obtained by performing a Fourier transform of the detected signal.

As was the case with SD-OCT, OFDI provides a great increase in imaging speed when compared with TD-OCT. Imaging speeds of 115,000 axial scans per second were obtained in 2005 using novel swept-laser technology. And catheter/endoscope imaging speeds of 100,000 axial scans per second were obtained in animal models [1]. These speeds are 50 to 100 times faster than TD-OCT. These high-imaging speeds enable 3D-OCT of luminal structures.

OFDI presents the advantage over SD-OCT that it does not require the use of line scan cameras. Thus, it can be used with wavelengths of 1000 nm and 1300 nm, where CCD cameras lack sensitivity. These longer wavelengths have more penetration in scattering tissues, hence allowing imaging of these kind of tissues.

2.4 Applications of OCT

As mentioned earlier OCT is mostly applied in biomedical imaging. OCT is an established diagnostics technique in ophthalmology. The earliest clinical studies of OCT were done in this field. OCT has been demonstrated for the detection and monitoring of a variety of macular diseases, The retinal nerve fibre layer thickness, which is a predictor of early glaucoma, can be quantified using the technique. OCT is a promising technique for the diagnosis and monitoring of glaucoma and macular edema associated with diabetes because it provides quantitative information on retinal pathology which is a measure of disease progression. [1, 11].

The fact that OCT is done mostly with fibre optic interferometers allows an easy combination with endoscopes and catheters. The use of endoscopes and catheters with OCT allows imaging inside the body, thus enabling a number of applications. A very interesting application of catheter based OCT is intravascular imaging. OCT is capable of detecting plaque formations in atherosclerotic arteries as well as thrombus [2]. This can be of great help in the treatment of coronary diseases, which are one of the major causes of death in North America. This particular area of application has greatly benefited from FD-OCT techniques. Blood is one of the most strongly scattering tissues in the body. Hence it is necessary to use a saline flush or a balloon occlusion to remove blood from the area to be imaged, this only provide a few seconds to perform imaging. Thanks to FD-OCT techniques high-imaging speeds it is possible to obtain much better images in one flush than it was possible with TD-OCT.

Another application of OCT in the biomedical imaging field is optical biopsy [11, 12]. OCT provides resolution comparable with histology. Though certainly OCT is still not capable of replacing histology in the assessment of cancerous lesions, the capability of performing *in situ*, in real time imaging of tissue morphology opens some interesting possibilities. OCT can be used to assess tissue pathology in regions where excisional biopsy is too dangerous. It also be used to guide conventional biopsy, thus reducing the number

of false results due to samples obtained from regions with no lesions. Finally, due to its real-time imaging capabilities, OCT is a promising option to guide surgical procedures in order to ensure that all of the diseased tissue is removed.

It is worth mentioning that although a great number of OCT applications are found in biomedical imaging, the technique can also be applied to other fields. OCT presents great possibilities for non-destructive testing (NDT) and material sciences [6]. One such application of OCT for NDT is the evaluation of ceramic materials. Advanced ceramic materials are used in many industrial applications. But their performance is highly related to the presence of surface or near-surface damage. One study found 0.5 mm of imaging depth in ceramic materials using a 1,300 nm source and 130 μV of output power [6]. Although very shallow, the penetration depth obtained with OCT is sufficient to evaluate sub-surface cracks generated due to material strain and stress.

One other field of application outside of biomedical imaging can be found in arts objects. One of the first OCT studies related to art featured archaic jade objects. The whitening effect seen in jade objects that have been buried for centuries is similar to the one whitening seen in jade objects that go through a forging method by burning. In this study it was shown that OCT can detect differences in the sub-surface morphology that help authenticate the real archaic jade objects. One of the most studied application of OCT in arts though, is related to paintings. In particular OCT is used to evaluate the thickness of the varnish layer that is applied on top of paintings for protective reasons. In old paintings this varnish cracks and becomes yellow, and have to be removed and re-applied. OCT provides a means to evaluate the removal process and prevent removal of the paint layers, as normally the paint layers present very high absorption [6].

These are just but a few of the applications of OCT. Extensions of the technique, like Doppler OCT, and polarization sensitive OCT enable a number more of applications. The ones presented here are just a sample to show the incredible potential of this imaging technique.

2.5 Multiple-scattering effect on OCT signal

Despite its great potential OCT applications remain limited by the technique's shallow imaging depth. In a scattering medium the imaging depth of OCT is around 3 millimetres. Although improvements in penetration depth can be obtained by imaging using 1,300 nm wavelengths, this is related to absorption in tissue. It has been shown that the dominant process limiting imaging depth in OCT is optical scattering [4, 13].

Multiple scattered light forms a disturbing background in the image. Its effects are a reduction of imaging contrast, a reduction of resolution, and a reduction of penetration depth [7]. Scattering limits imaging depth because it causes attenuation of the light as well as randomization of image information. Simulations done in [13, 14] show the contribution of single-scattered photons and multiply-scattered photons to the OCT signal according to depth in tissue. In Fig. 2.3 obtained from [14] it is visible that, after a certain depth has been reached, the signal contribution from multiple-scattered photons exceeds that of single-scattered photons.

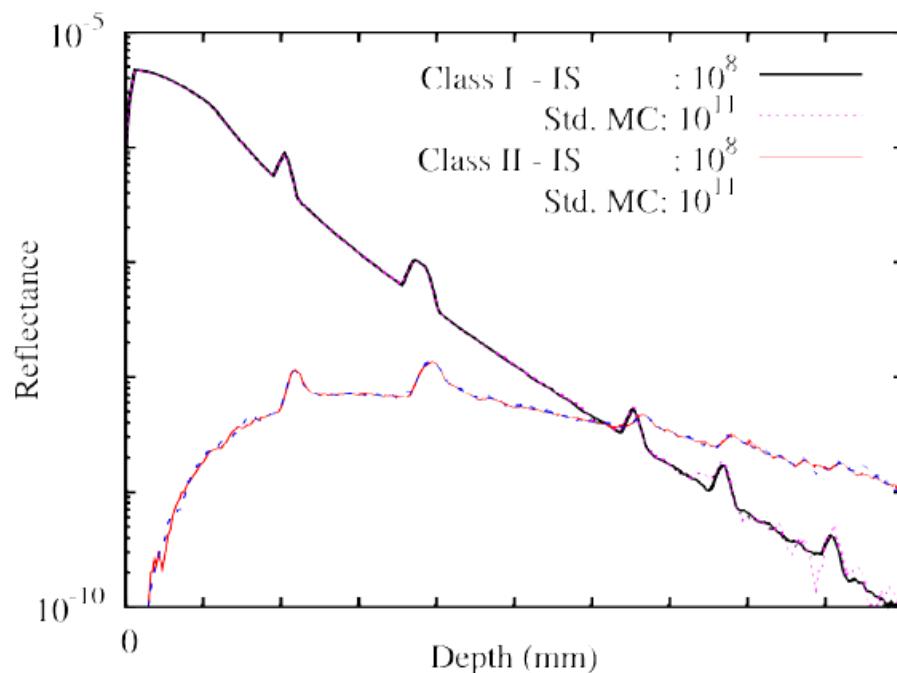


Figure 2.3. Class I reflectance represent that obtained from ballistic photons. Class II reflectance is obtained from photons that undergo more than one scattering event. It is clearly seen that after a certain depth Class II reflectance is dominant.

It can be inferred from Fig. 2.3 that increasing the output power of the light source won't have much effect on improving penetration depth, as the amount of Class II reflectance will also increase. The point at which Class I and Class II reflectance is the same should be noted; in the simulation result shown this occurs at around 0.6 mm in depth. Which indicates that Class II signals increase in contribution very fast.

Since many of OCT applications would greatly benefit from an increased imaging depth, this line of research is very interesting. We shall note that some improvement can be obtained by adjusting the wavelength parameters of the source, or the amount of power of the light source. But the best way to obtain improvement in imaging depth is to tackle the effect of multiply-scattered photons in OCT signal. In chapter 3, we propose an Integrated Computational Imaging method that aims to reduce the effect of multiple-scattering in the OCT signal. This method could help improve OCT imaging depth, thus enabling improvements in current applications of OCT, as well as new applications of the technique.

Chapter 3

Integrated Computational Imaging for OCT

3.1 Introduction

In this chapter we describe an integrated computational imaging system for OCT. An integrated computational imaging system, also known as a hybrid system, combines a modified optical imaging system and a digital post-processing step. A hybrid imaging system is different from a system obtained by cascading a standard imaging system and a digital post-processing step. In a hybrid system, both the optical and digital modules are parts of a single system, and the imaging process is divided between them [15]. A block diagram is shown in Figure 3.1.

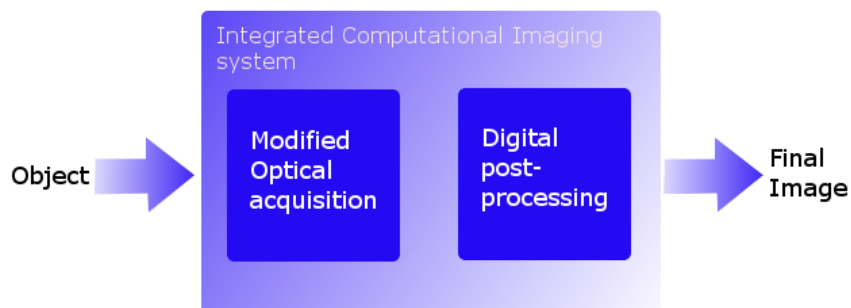


Figure 3.1. A block diagram of an integrated computational imaging system

In the proposed system a continuous-wave (CW) ultrasound is applied to tissue at the same time as OCT imaging is performed. This ultrasound signal changes the density of the tissue to which it is applied, thus changing the refractive index of that tissue. We can call this a modulation of the refractive index, or a modulation of the sample [16]. We propose to manipulate the acoustic signal in such a way as to produce a modulation of the sample that would benefit OCT imaging depth. Let us first present evidence that applying an ultrasound signal simultaneously with the OCT beam is beneficial for the OCT imaging process.

3.2 OCT imaging in the presence of ultrasound

We have stated that the presence of CW ultrasound in tissue at the time of imaging can benefit OCT. There is evidence showing this in [17] and [18]. In these papers the authors show that the presence of an ultrasound beam parallel to the OCT beam reduces multiple scattering as measured by the point spread function (PSF) by at least 20%. The cause for these improvements of the OCT signal is assumed to be the interaction of multiply-scattered photons with the acoustic wave.

An acoustic wave traveling in a homogeneous medium creates a perturbation of the refractive index. The medium becomes inhomogeneous with a time-varying stratified refractive index. Yet due to the great difference in the frequencies of light and sound we can, for the purpose of studying light-sound interaction, regard the refractive index as being static [19]. Considering the refractive index as being static, we obtain in the medium a diffraction grating in the direction of the ultrasound propagation. The simplest form of light sound interaction is the partial reflection of light from the sound created diffraction grating depending on the angle of incidence θ of the light. This interaction is known as *Bragg diffraction*.

Taking these considerations about the interaction of an acoustic wave and an optical one we can now explain the reduction of multiple scattering in [17] and [18] rather simply. Given that in the experiments the ultrasound beam was parallel to the OCT beam, the inci-

dent light and back-reflected photons which underwent just one scattering event had little interaction with the grating. However, photons that underwent more than one scattering event have a higher probability of hitting the grating at an angle cause them to be diffracted to a region away from the photodetector. We note here, that since many OCT systems are based on single-mode fibre optics the detection area is only a few microns in radius [13]. Thus the angle of the diffraction events do not need to be an extremely high one to turn multiple-scattered photons away from the detection region.

3.3 A novel method: Ultrasound patterns

We know from the experimental evidence that the application of an CW ultrasound signal to tissue has a beneficial effect to OCT by increasing imaging depth. And we know, from the same experimental evidence that this benefit is due to a reduction of multiply-scattered photons at the detector. We present here a novel idea to take advantage of this knowledge. That is, instead of considering ultrasound modulation of tissue and OCT imaging as separate events; treat them as parts of an *integrated computational imaging* for OCT.

In a integrated computational imaging system approach modulation of the sample via the application of ultrasound and OCT imaging are considered as steps of a process, and not independent processes themselves. Thus ultrasound application should be designed considering that OCT imaging will be done depending on it. And also, OCT post-processing should take into account that the sample have been modulated using ultrasound. A similar idea was shown in [16], though in this work the authors were more concerned with obtaining new information by post-processing the obtained OCT signal. Our interest is in designing a whole integrated imaging system. But, we shall concern ourselves first with the ultrasound modulation to be applied to the sample.

Previous works in the literature about OCT imaging in the presence of acoustic waves use only one CW ultrasound beam [16–18]. We propose the design and utilization of *ultrasound patterns* to perform the modulation of the sample. This ultrasound patterns are

obtained by the combination of two or more ultrasound beams. In this thesis we design standing-wave ultrasound patterns. These patterns present the interesting characteristic of presenting areas of no-change in the refractive index surrounded by acoustic generated gratings. Thus the standing-wave patterns behave like a waveguide for the optical OCT beam.

A proof of concept image is shown in Figure 3.2. This figure was obtained in MATLAB assuming a medium with a speed of sound of 1500 m/s and a frequency for the acoustic waves of 1 MHz. Two acoustic waves with angles of $\pi/8$ and $-\pi/8$ were used to produce the standing-wave patterns. The obtained pattern shows the areas of no change in the refractive index surrounded by acoustic generated gratings that could be used as waveguides for single-scattered photons.

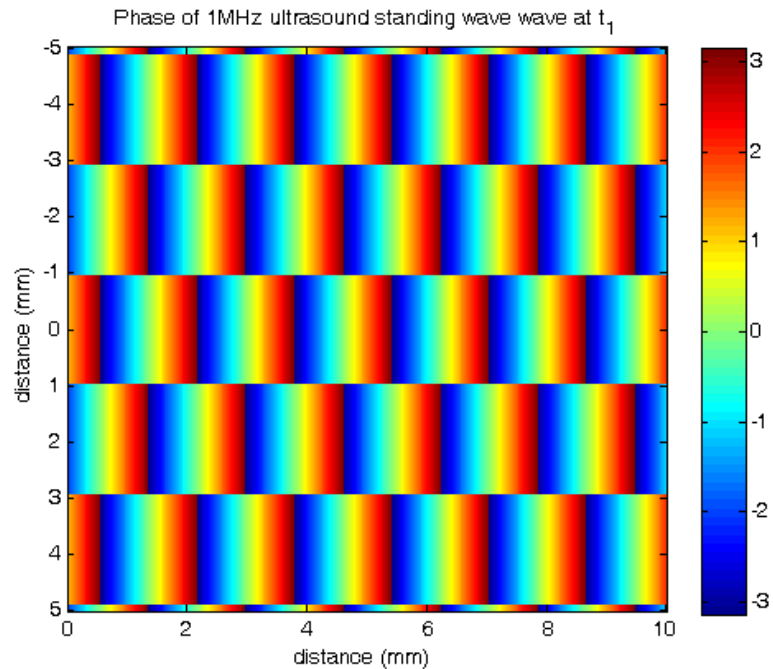


Figure 3.2. A standing wave ultrasound pattern generated by two ultrasound sources in a homogeneous medium

Of course, to be of use the ultrasound patterns should be applied to inhomogeneous tissue. An useful ultrasound pattern will not be much affected by the inhomogeneity of tissue, that is the acoustic scattering resulting from the inhomogeneity will not change the

pattern much. In order to test the feasibility of obtaining patterns like the one shown in Figure 3.2 simulations were performed. These simulations were done using COMSOL Multiphysics software and the results obtained are shown in Chapter 5.

3.4 Delivery of ultrasound and OCT beams

In the literature about performing OCT imaging in the presence of an acoustic wave the orientation of the ultrasound beam is not given much consideration. The acoustic wave can be parallel, perpendicular or anti-parallel to the OCT beam. Although these orientations need only be approximate. For example, [18] mentions an angle of about 10° between the OCT beam and the ultrasound beam, due to alignment difficulties, still considering the beams to be parallel. Fig. 3.3 and Fig. 3.4 show two of these arrangements between the ultrasonic wave and the OCT beam.

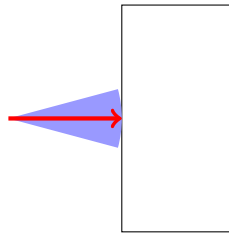


Figure 3.3. Parallel orientation between ultrasound wave and OCT beam. OCT beam is shown by an arrow, ultrasound wave shaded cone.

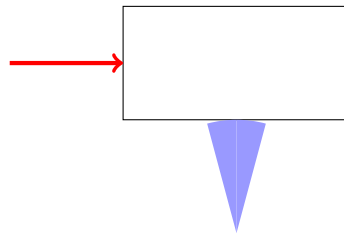
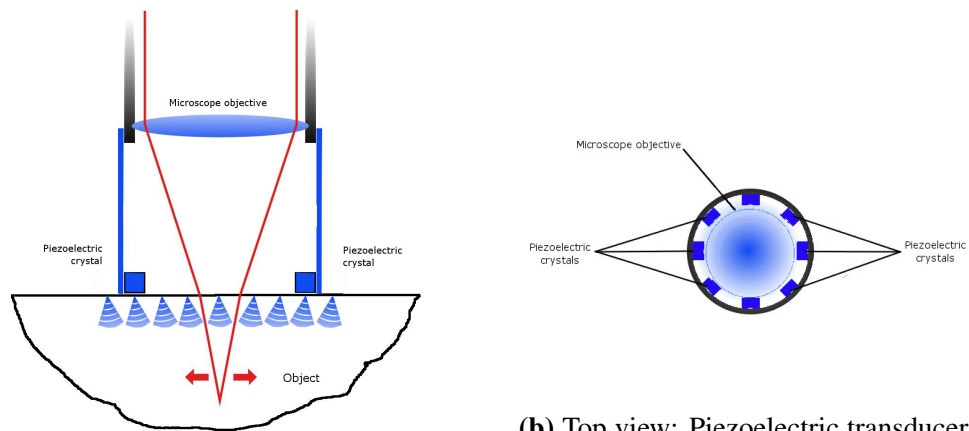


Figure 3.4. Perpendicular orientation between ultrasound wave and OCT beam. OCT beam is shown by an arrow, ultrasound wave shaded cone.

These simple arrangements work well in [17] and [18] because they are not concerned with the exact angle of incidence of the ultrasound wave. That is not the case when trying to obtain a specific ultrasound pattern to modulate the sample. The ultrasound pattern

obtained will depend on the angles between the different ultrasound plane waves used. We may also need to locate the OCT beam in an specific region of the ultrasound generated pattern. Thus we can not simply deliver the ultrasound waves in a separate manner from the OCT imaging beam, we shall require a delivery system that provides ultrasound waves and OCT beam in a coordinated manner. In other words the system will require a specially designed probe, capable of delivering both signals.

In order to design this probe simulations of the interaction of an OCT beam with the designed patterns is required. The goal of these simulations is to determine the optimal ultrasound patterns that minimize multiple-scattering for a particular class of tissue. However there is not an adequate simulation software for this problem. That is, there is no simulator capable of simulating light interaction with tissue in a very big volume, relative to the light wavelength, with reasonable computational resources. Because of this fact, we decided to develop our own simulator. The implementation of an algorithm to perform this simulation is shown in Chapters 6 and 7 In the meantime we show in Fig. 3.5 one proposed design for the probe discussed.



(a) Lateral view: Optical beam focused by microscope objective in red with piezoelectric transducers generating ultrasound in blue.

(b) Top view: Piezoelectric transducers are shown in gray inside a sleeve connected to the microscope objective.

Figure 3.5. A proposed design for a probe capable of delivering ultrasound and OCT beams to the sample in a coordinated manner.

3.5 Conclusion

This chapter introduces the idea of an *integrated computational OCT system*. There is evidence showing that OCT sees improved imaging depth in the presence of ultrasound. Taking this into consideration our proposed integrated computational OCT system uses ultrasound to modulate the refractive index of tissue. Instead of just applying an ultrasound beam as in the literature, we propose that an ultrasound pattern can be designed to minimize the effect of multiple-scattering in OCT imaging. To be useful the ultrasound pattern has to be applied to inhomogeneous tissue. Acoustic scattering caused by the inhomogeneity should not affect the pattern considerably. To test the feasibility of this requirement simulations of the effect of ultrasound pressure on the refractive index were performed.

As a first step to perform the simulations of the effect of ultrasound pressure on the refractive index, we need to establish a relationship between acoustic pressure and the refractive index. In the next chapter, we discuss the relationship between acoustic pressure and the refractive index. The formulae derived in chapter 4 are used to connect the effect of pressure on the density of the medium with a change in the refractive index. Afterwards, in chapter 5 we present the results of our simulations performed in COMSOL Multiphysics software.

To design ultrasound patterns that minimize multiple-scattering in tissue we need to simulate the interaction of light with tissue modulated by these ultrasound patterns. However there is not an adequate simulation software for this problem. So we need to develop our own simulator of light interaction with tissue. We decide to use Monte Carlo methods because of their computational efficiency in high dimensions. Chapter 6 introduces the basic theory of Monte Carlo methods. Chapter 7 discuss the implementation of Reversible Jump Markov chain Monte Carlo solver for Fredholm integral equations of the second kind. In chapter 7 we present the method to solve integral equations with Monte Carlo, we implement the algorithm and MATLAB and show the results obtained in its application to 1-D and 2-D equations.

Chapter 4

Relationship between pressure and the refractive index

4.1 Introduction

In this chapter we will discuss the relationship between pressure and the refractive index. We will first present the effect of acoustic pressure in the density of a medium. Afterwards, we will discuss the polarization of dielectric materials. This serves as a basis to introduce the Clausius-Mossotti equation, which relates the density of the material with its polarizability. Then we will be able to link changes in density due to acoustic pressure to changes in the refractive index of the material. This knowledge is required to simulate the effect of ultrasound patterns on the refractive index of tissue.

4.2 Compression due to acoustic pressure

Let us consider first a fluid under equilibrium. Under equilibrium the equation of state of the fluid is given by [20]

$$k_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right) \Big|_T = -\frac{1}{\rho} \left(\frac{\partial \rho}{\partial P} \right) \Big|_T$$

and

$$\beta = -\frac{1}{V} \left(\frac{\partial V}{\partial T} \right) \Big|_P = -\frac{1}{\rho} \left(\frac{\partial \rho}{\partial T} \right) \Big|_P$$

where k_T is called the *isothermal compressibility* and β is called the *coefficient of thermal expansion of the fluid*. Both k_T and β are functions of P and T .

For sound frequencies below 10^9 cps, in gases, however it is a better approximation to assume adiabatic compression. In this case

$$\left(\frac{\partial \rho}{\partial P} \right) \Big|_S = \rho k_S = \frac{1}{\gamma} \left(\frac{\partial \rho}{\partial P} \right) \Big|_T$$

and

$$\left(\frac{\partial T}{\partial P} \right) \Big|_S = \frac{\gamma - 1}{\gamma} \left(\frac{\partial T}{\partial P} \right) \Big|_P$$

so that a change in density due to the applied pressure is given by

$$\rho + \delta = \rho + \rho k_S p \tag{4.1}$$

and a change in temperature by

$$T + \tau = T + (\gamma - 1) \frac{k_S}{\beta} p \tag{4.2}$$

where γ is the ratio of the specific heat at constant pressure of the gas to the specific heat at constant volume. And k_S is called the adiabatic compressibility.

4.3 The relationship between refractive index and density

We can link the refractive index with the density of the material via the Classius-Mossotti equation. We shall now show the derivation of this relationship.

4.3.1 Polarization of materials

In considering the polarization of molecules we must start by considering the polarization of isolated atoms. In a neutral system, two charges $+e$ and $-e$ a distance x apart have a dipole moment $\mathbf{p} = ex$. The dipole moment of a set of such dipoles is

$$\mathbf{P} = \sum_j ex_j = N\mathbf{p} \quad (4.3)$$

in which N is the density of dipoles.

Polarization of atoms

The simplest case is that of an isolated atom with no dipole moment in the absence of an external field. The application of a homogeneous field induces a dipole moment given by

$$\mathbf{p}_e = 4\pi\epsilon_0\alpha_e\mathbf{E} \quad (4.4)$$

where α_e is the *electronic polarizability* of the atom.

Polarization of molecules

The situation for molecules is a little more complex. Some molecules may have permanent dipoles depending on their charge distributions. Materials with such molecules are called polar materials. When the electronic charge distribution within the molecule is symmetric, it has no permanent dipole. This type of molecule is called nonpolar. For nonpolar materials, the application of an external electric field gives rise to dipole moments described by

$$\mathbf{p}_a = 4\pi\epsilon_0\alpha_a\mathbf{E}_{loc} \quad (4.5)$$

where α_a is the *atomic polarizability* and \mathbf{E}_{loc} is the local field surrounding the molecule. The atomic polarizability is the average effect of all possible polarizations in relation to the applied field. α_a is, in principle, a function of the electric field, but for a wide range of

field strengths it is a constant that characterizes the response of the molecules to an applied field [21]. The expression for the atomic polarization per unit volume can be expressed as

$$\mathbf{P}_a = N\mathbf{p}_a = 4\pi\epsilon_0 N\alpha_a \mathbf{E}_{loc} \quad (4.6)$$

where N is the dipole density.

Permanent dipoles contribution in polar molecules

Permanent dipoles are created in molecules when electron charge distributions are asymmetrical. An example of this is the water molecule where the two hydrogen atoms form an angle with the oxygen atom [22]. The Langevin theory is used to determine the average molecular dipole moment parallel to an applied field as a function of temperature. The potential energy of a molecule with dipole moment p_d in an electric field \mathbf{E} , where ϕ is the angle between p_d and \mathbf{E} is

$$P.E. = p_d \mathbf{E} \cos \phi \quad (4.7)$$

The average dipole moment per unit volume is

$$\mathbf{P}_d = N\mathbf{p}_d \langle \cos \phi \rangle \quad (4.8)$$

where N is the density of molecules. The factor $\langle \cos \phi \rangle$ is given by [22]

$$\langle \cos \phi \rangle = L(a) = \coth a - \frac{1}{a} \quad (4.9)$$

when $a \ll 1$ then $L(a) = a/3$. In our case $a = \mathbf{p}_d \mathbf{E} / k_B T \ll 1$, so

$$L(a) = \frac{1}{3} \frac{\mathbf{p}_d \mathbf{E}}{k_B T} \quad (4.10)$$

plugging this value back into Eq. (4.8)

$$\mathbf{P}_d = N\mathbf{p}_d\langle\cos\phi\rangle = N\mathbf{p}_d \cdot \frac{\mathbf{p}_d\mathbf{E}}{3k_B T} = \frac{N\mathbf{p}_d^2\mathbf{E}}{3k_B T} \quad (4.11)$$

From this the molecular or dipolar orientation polarizability is

$$\alpha_d = \frac{\mathbf{p}_d^2}{3k_B T} \quad (4.12)$$

The total polarizability of a substance must include the electron contribution (Eq. (4.4)) and the atomic polarizability (Eq. (4.6)). Thus the total polarizability is

$$\alpha = \alpha_e + \alpha_a + \alpha_d \quad (4.13)$$

4.4 The Lorentz field and the Debye Equation

4.4.1 Derivation of the Lorentz field

A particular molecule in a liquid is affected by both the applied field and the fields of the dipoles present on other molecules surrounding it. We want to describe now the field at the location of a given particle and to determine its effect on that particle. This field is the *local* field.

The local field is determined at the centre of a small spherical region within the dielectric material. This spherical volume is chosen to be sufficiently small with respect to the dimensions of the dielectric material so that the external volume surrounding the sphere may be regarded as homogeneous and continuous and to have dielectric constant ϵ . However, the sphere must be large with respect to the dimension of the particle (See Figure 4.1).

The internal field is affected by several factors. The first one is the charge density on the plates on either side of the dielectric, given by $E_{vac} = D/\epsilon_0 = q/\epsilon_0$. The effect of the charges induced at the interface adds the depolarizing field $E_D = -P/\epsilon_0$ to the field

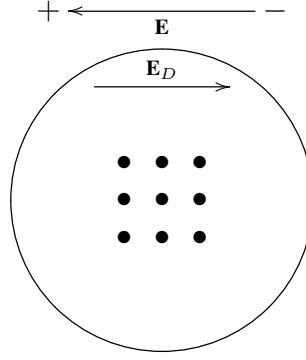


Figure 4.1. Spherical region for calculation of the electric field. It's big enough to include a number of particles, yet small enough to consider the external region to be homogeneous. Some fields are also shown.

strength. Thus, the induced charges on the surface of the region are a factor to be considered. Here, we shall only consider the components parallel to the applied field. An element of surface area on the surface of the spherical region has an area $2\pi r^2 \sin \theta d\theta$. The angle θ determines the surface charge density given by $-(\mathbf{P}/\sqrt{4\pi\epsilon_0}) \cos \theta$. Then $-(\mathbf{P}/\sqrt{4\pi\epsilon_0}) \cos \theta \cdot 2\pi r^2 \sin \theta d\theta$ is the charge on the element of surface area. The field on the spherical surface is

$$\sqrt{4\pi\epsilon_0} d\mathbf{E}_S = \frac{2\pi r^2 \sin \theta d\theta \cdot \mathbf{P} \cos \theta \cdot \cos \theta}{\sqrt{4\pi\epsilon_0} r^2}$$

The charge on the spherical surface is thus found to be

$$\mathbf{E}_S = \frac{\mathbf{P}}{2\epsilon_0} \int_0^\pi \cos^2 \theta \sin \theta d\theta = \frac{1}{3\epsilon_0} \mathbf{P} \quad (4.14)$$

The sum of these three components ($D/\epsilon_0 + \mathbf{E}_D + \mathbf{E}_S$) give the local field as

$$\mathbf{E}_{loc} = \frac{D}{\epsilon_0} - \frac{\mathbf{P}}{\epsilon_0} + \frac{\mathbf{P}}{3\epsilon_0} = D - \frac{2\mathbf{P}}{3\epsilon_0}$$

Using the well known fact that,

$$\mathbf{D} = \epsilon \mathbf{E} = \mathbf{P} + \epsilon_0 \mathbf{E}$$

we can now write,

$$\mathbf{E}_{loc} = \mathbf{E} + \frac{\mathbf{P}}{\epsilon_0} - \frac{\mathbf{P}}{3\epsilon_0} = \mathbf{E} + \frac{1}{3\epsilon_0}\mathbf{P} \quad (4.15)$$

Now we substitute $\mathbf{P} = \mathbf{E}(\epsilon - \epsilon_0)$ into our last equation we obtain

$$\mathbf{E}_{loc} = \mathbf{E} + \frac{\mathbf{E}}{3\epsilon_0}(\epsilon - \epsilon_0) = \mathbf{E}\left[1 + \frac{1}{3\epsilon_0}(\epsilon - \epsilon_0)\right]$$

or after rearranging

$$\mathbf{E}_{loc} = \mathbf{E} \frac{\epsilon + 2\epsilon_0}{3\epsilon_0} \quad (4.16)$$

Eq. (4.16) is called the Lorentz field. It does not include the effects of the dipoles within the internal spherical region. This field depends heavily on the lattice structure of the material. When a material is composed of several different ions, each ion will have a different local field because each has different surroundings. A modification of Eq. (4.15) gives the field for one such ion as

$$\mathbf{E}_{loc} = \mathbf{E} + \beta \frac{\mathbf{P}}{4\pi\epsilon_0} \quad (4.17)$$

where β represents the particular field constant for a given case. Eq. (4.16) is a special case of Eq. (4.17) that holds only in the special case when the net effects of the dipoles within the spherical region are zero.

4.4.2 Derivation of the Debye Equation

Including atomic and dipolar polarizations we can now write,

$$\mathbf{P} = \mathbf{E}(\epsilon - \epsilon_0) = P_a + P_d$$

we can substitute Eq. (4.6) and Eq. (4.12) here to write

$$\mathbf{P} = \mathbf{E}(\epsilon - \epsilon_0) = 4\pi\epsilon_0 N\alpha_a \mathbf{E}_{loc} + N \frac{p_d^2}{3k_B T} \mathbf{E} \quad (4.18)$$

Assuming $\mathbf{E} \approx \mathbf{E}_{loc}$ we get

$$\mathbf{P} = \frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} = \frac{N}{3\epsilon_0} \left(4\pi\epsilon_0\alpha_a + \frac{p_d^2}{3k_B T} \right) \quad (4.19)$$

This is the Debye equation for the polarization of a pure substance. It combines the Langevin theory and the Classius-Mossotti equation. In so doing it accounts for the effects of permanent dipoles and temperature. At this point it is helpful to define a *molar polarization* as

$$A = \frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} \cdot \frac{M}{\rho} \quad (4.20)$$

in which M is the molecular weight and ρ is the density of the substance. Taking Eq. (4.20) into account we can now express Eq. (4.19) as

$$\mathbf{P} = \frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} = \frac{N_A}{3\epsilon_0} \left(4\pi\epsilon_0\alpha_a + \frac{p_d^2}{3k_B T} \right)$$

where $N_A = 6.0221415 \times 10^{23}$ is the Avogadro's number, that express the number of elementary particles in a mol of a substance.

4.5 Derivation of the Classius-Mossotti equation

We decided to work with the well known Lorentz-Lorenz relationship, that is an extension of the Classius-Mossotti equation. The Classius-Mossotti was derived considering nonpolar material, water our material of interest is a polar liquid. But, in the range of optical frequencies the part of the polarization due to permanent dipoles is zero. And because of that a polar liquid behaves as a nonpolar one at these frequencies [23].

We shall now proceed to derive the Classius-Mossotti equation. Assume a local field applied to a substance which main polarization effect is electronic, then

$$\mathbf{P}_e = N4\pi\epsilon_0\alpha_e\mathbf{E}_{loc} = \mathbf{E}(\epsilon - \epsilon_0) \quad (4.21)$$

where N is the density of the dipole. We can obtain an expression for the local field from Eq. (4.17) and Eq. (4.4)

$$\mathbf{E}_{loc} = \mathbf{E} + \beta N \alpha_e \mathbf{E}_{loc}$$

Collecting terms,

$$\mathbf{E}_{loc}(1 - \beta N \alpha_e) = \mathbf{E} \quad (4.22)$$

Substituting Eq. (4.22) is substituted into Eq. (4.21) for \mathbf{E} we get

$$N4\pi\epsilon_0\alpha_e = (1 - \beta N \alpha_e)(\epsilon - \epsilon_0)$$

or

$$\epsilon - \epsilon_0 = \frac{4\pi\epsilon_0 N \alpha_e}{1 - \beta N \alpha_e} \quad (4.23)$$

in this last case if the effects of the dipoles are zero in the spherical region, then $\beta = 1/(3\epsilon_0)$ from Eq. (4.15)

Another useful relationship can be obtained starting from

$$\mathbf{E} = \mathbf{E}_{loc} - \frac{\mathbf{P}}{3\epsilon_0} \quad (4.24)$$

We divide both sides of Eq. (4.21) by \mathbf{E} and Eq. (4.24) is substituted in the denominator to give

$$\chi_e = \frac{P_e}{\mathbf{E}} = \frac{N4\pi\epsilon_0\alpha_e\mathbf{E}_{loc}}{\mathbf{E}} = \frac{N4\pi\epsilon_0\alpha_e\mathbf{E}_{loc}}{\mathbf{E}_{loc} - \frac{\mathbf{P}}{3\epsilon_0}} \quad (4.25)$$

where χ_e is the electric susceptibility. We can obtain another expression for χ_e by considering that $\frac{\epsilon}{\epsilon_0} = \frac{\mathbf{P}}{\epsilon_0\mathbf{E}} + 1$, Equating this to Eq. (4.25) we have

$$\chi_e = \frac{\mathbf{P}}{\mathbf{E}} = (\epsilon - \epsilon_0) = \frac{N4\pi\epsilon_0\alpha_e\mathbf{E}_{loc}}{\mathbf{E}_{loc} - \frac{4\pi\epsilon_0 N \alpha_e \mathbf{E}_{loc}}{3\epsilon_0}} \quad (4.26)$$

where we replaced \mathbf{P} using Eq. (4.6). We can reexpress Eq. (4.26) as

$$\begin{aligned}
 (\epsilon - \epsilon_0) \left(1 - \frac{4\pi}{3} N\alpha_e \right) &= 4\pi\epsilon_0 N\alpha_e \\
 \epsilon - \frac{4\pi\epsilon}{3} N\alpha_e - \epsilon_0 + \frac{4\pi\epsilon_0}{3} N\alpha_e &= 4\pi\epsilon_0 N\alpha_e \\
 \epsilon - \epsilon_0 = \frac{4\pi}{3} N\alpha_e (3\epsilon_0 + \epsilon - \epsilon_0) &= \frac{4\pi}{3} N\alpha_e (\epsilon + 2\epsilon_0)
 \end{aligned}$$

Finally, if we consider $\alpha = \alpha_e + \alpha_a$

$$P = \frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} = \frac{4\pi}{3} N\alpha \quad (4.27)$$

Eq. (4.27) is the Clausius-Mossotti equation for a pure substance. It relates the polarizability and the dielectric constant. It will be noted that this equation does not take orientation effects into account and its not a function of temperature. We can relate the polarizability with the density of the material by remembering our definition of molar polarizability,

$$A = \frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} \cdot \frac{M}{\rho} = \frac{4}{3}\pi N_A \alpha \quad (4.28)$$

where N_A is the Avogadro number. Putting it in function of the density

$$\frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} = \rho \cdot \frac{4\pi N_A \alpha}{3M} = \rho C \quad (4.29)$$

where C is a constant for each given material. The constants C may be calculated from published values of A , the *molar polarizability*.

4.6 Conclusion

In conclusion, the refractive index of a medium is directly related to its density. That relationship can be reduced to a dependence on a single constant C that varies for each material, as shown in Eq. (4.29). This constant can be easily calculated from data about the

material of interest. Thus, it is possible to obtain the variations in the refractive index of tissue due to applied acoustic pressure by first calculating the change in density. And then, converting this change in density into a change of refractive index using Eq. (4.29).

Of course it is also possible to obtain the change in refractive index caused by acoustic pressure using a formula as,

$$n(r, t) = n_0 + \eta \frac{p(r, t)}{\rho c^2} n_0 \quad (4.30)$$

where, ρ is the average density of the medium, c is the average speed of sound in the medium and η is called the elasto-optic coefficient. This formula is found in [24]. The elasto-optic coefficient is much like C in Eq. (4.29), it links variations of the refractive index with the corresponding variation of density. η is also particular to each material. A similar formula, that relates refractive index to the strain caused by acoustic pressure can be found in [19].

Each of these approaches is valid when solving acousto-optic problems. In this particular case, for the simulations shown in the following chapter I chose to use the first approach. Finding first the variations in density due to acoustic pressure and then converting those to variation of the index of refraction.

Chapter 5

Results from the simulations of ultrasound patterns

5.1 Introduction

Simulations of ultrasound in tissue were needed to assess its impact on the refractive index. This chapter presents results of simulations of ultrasound patterns in tissue. These simulations were done using the software COMSOL Multiphysics.

5.2 COMSOL setup

COMSOL Multiphysics is a simulation software capable of working with a number of different physics models. It is based on the finite element method (FEM). COMSOL provides the necessary acoustic simulation features, it also has a very structured work method to setup the simulation. One first defines the model, this includes the geometry and the materials. Then one sets the physics, with the equations that have to be considered. Lastly, one setups the solver, which includes assigning a mesh and setting the parameters for the simulations. We describe here the setup required to perform these simulations with COMSOL Multiphysics.

We first define the geometry, in this simulation it is 2-D rectangular region with dimension of $1 \text{ cm} \times 1 \text{ cm}$. COMSOL allows its user to construct the geometry directly inside the software or to import it from some CAD softwares. Since the region we are defining is simple, we just constructed it inside COMSOL. Water is selected as the material for the simulation of a homogeneous medium. Water is available in COMSOL library of materials, so all the relevant parameters for the simulations in a homogeneous medium are already available in the software. The inhomogeneous medium is obtained by randomly changing the refractive index inside the region of simulation between 1.3 and 1.5. As stated before, this means a change in the density of the medium used for the acoustic simulation. Considering the model given by Eq. (4.30), the speed of sound of the medium is kept constant, and equal to the value of the speed of sound in water. The setup of the homogeneous medium is shown in Figure 5.1.

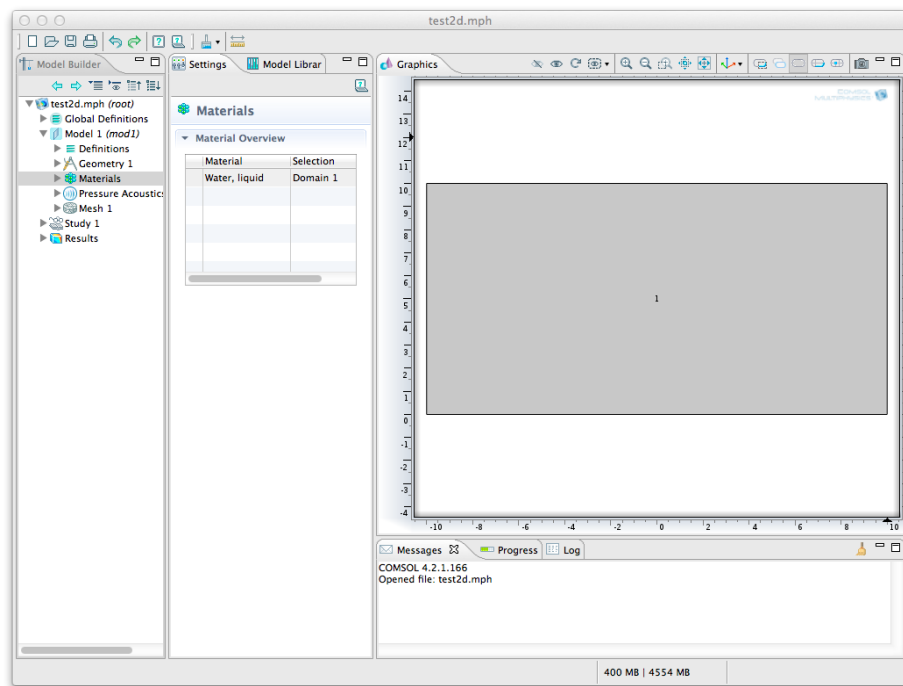


Figure 5.1. Screen capture from COMSOL Multiphysics showing the geometry of the model and the material selected for the homogeneous medium.

After setting the geometry and the material, one needs to set up the physics. In these simulations COMSOL is set to study an acoustic model. The side and top boundaries of

the region of simulation are set as absorbing boundaries. The bottom boundary is set as the pressure source. Pressure sources are set as incident pressure fields on the boundary, with a pressure amplitude of 5×10^5 Pa in the ultrasound standing-wave pattern simulations. Two incident pressure fields are used for the ultrasound standing-wave pattern simulations. For the plane-wave pattern simulations one pressure source of 1 MPa was used. The value of pressure provided by these sources was adjusted in a way as to not surpass the safety limits for ultrasound application in tissue. The setup of the pressure sources is shown in Figure 5.2.

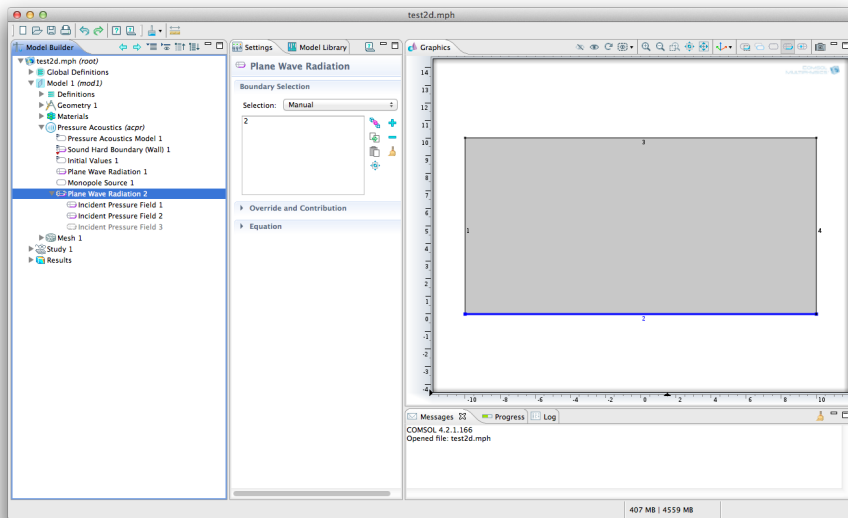


Figure 5.2. Screen capture from COMSOL Multiphysics showing the pressure sources applied to the lower boundary of the model.

Once the physics and all the sources are defined, one has to adjust the size of the mesh to be used. We used the extremely fine setting from the default meshing options in COMSOL. With this setting the constructed mesh had 12480 elements, with a maximum element size of 0.2 mm. Using a coarser mesh produced inaccurate results. Configured like this COMSOL provides results on the applied pressure and the change in density it produces. We use the Lorentz-Lorenz equation defined in Eq. (4.29) to convert this result of change in density to a change in refractive index. The results obtained can be found in section 5.4.

5.3 Ultrasound safety considerations

The design of the hybrid imaging system must consider the effect of ultrasound in tissue. There are two major mechanisms through which an acoustic wave can produce damage in tissue. To prevent any kind of harm, safety limits for the application of ultrasound in tissue are established. These limits include indices for each of the cited mechanisms.

We took particular care to respect the mechanical index. This index is related to the possibility of acoustic cavitation. Acoustic cavitation involves the expansion and then quick implosion of bubbles in liquids or gasses. Equipment marketed in the US can not exceed a mechanical index of 1.9. This means a peak negative pressure of 1.9 MPa at 1 MHz, or 6 MPa at 10 MHz.

There are three thermal indices that are applied under different conditions. One of them (TIS) is applied when only soft tissue is exposed to ultrasound. The other two (TIB and TIC) are used when bone is exposed to ultrasound. Though we payed less attention to these indices we note that the tissue that absorbs ultrasound more strongly is bone, which is not one of the tissues normally imaged with OCT. Also, we note that for the operating conditions of most diagnostic ultrasound equipment damage due to heating is unlikely under most conditions [25].

Our system must adhere to the safety limits for application of ultrasound. At the same time we would like to produce the maximum amount of change in the refractive index, where change in the refractive index is directly related to the amount of pressure applied. At a frequency of 1 MHz, at which the simulations were performed, we took care not to exceed the safety standard limit imposed by the mechanical index. That is, a peak negative pressure of 1.9 MPa [25] was not exceeded.

5.4 Results

We show in this section the results of the performed simulations. We first show the mediums used in the simulations (Fig. 5.3). We note that the refractive index used for the

inhomogeneous medium is randomized between 1.3 and 1.5. This is the range of refractive index commonly found in tissue. We also emphasize the fact that it is unlikely that one shall find this much variation in real tissue, it is more likely that tissue of similar refractive index is grouped together.

Fig. 5.4 shows the effect of an ultrasound plane wave on the refractive index of the mediums. We note that the obtained patterns obtained in both the homogeneous and inhomogeneous mediums are very similar. This suggest that the effect of acoustic scattering on the ultrasound pattern is not significant.

Fig. 5.5 shows the effect on an ultrasound standing-wave pattern on the refractive of the mediums. We can notice that patterns in the inhomogeneous medium are similar to the patterns in the homogeneous medium. Some distortion of the pattern is observed in the inhomogeneous medium, for example comparing Fig. 5.5(a) and Fig. 5.5(b), but most features of the desired ultrasound pattern remain.

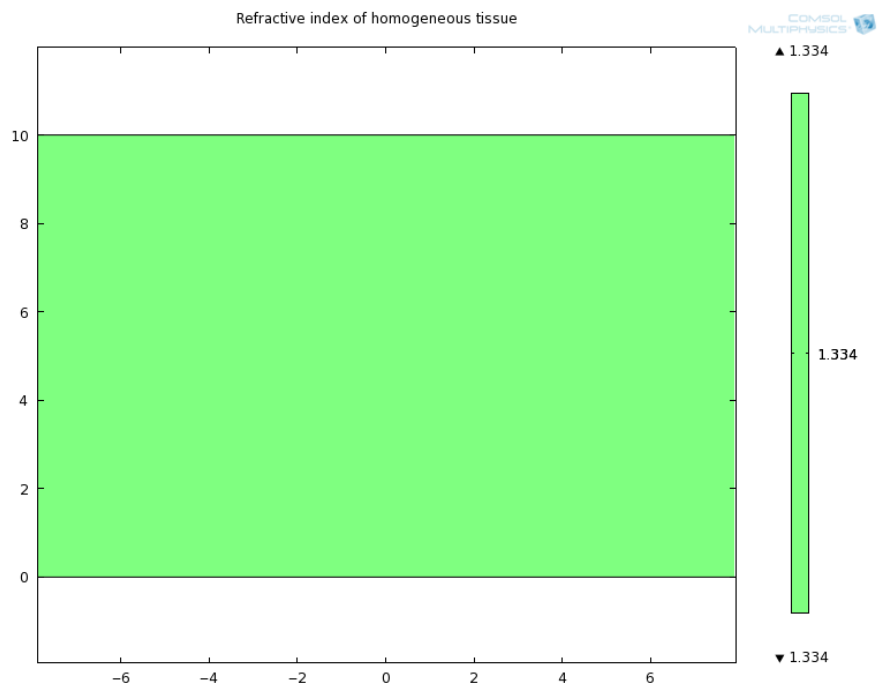
In Fig. 5.6 we show the pressure applied to produce the ultrasound standing-wave patterns. We note that it does not exceed the safety limit of 1.9 MPa at 1 MHz.

We need a measure to assess how similar the produced patterns are. We decided to use the angle in Hilbert space, defined for two vectors \mathbf{x} and \mathbf{y} as

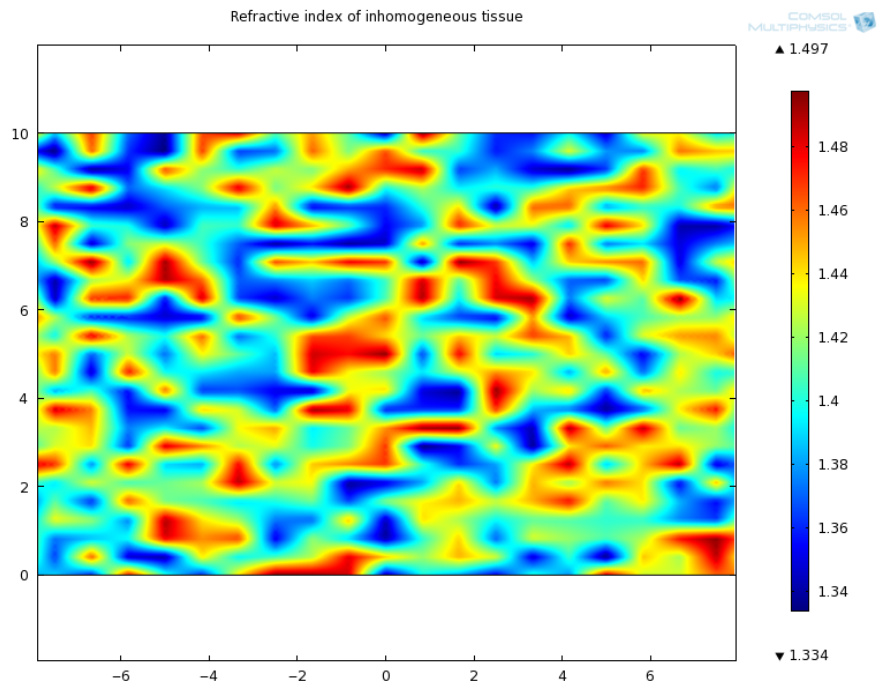
$$\cos \theta = \frac{\mathbf{x}\mathbf{y}}{\|\mathbf{x}\|_2 \|\mathbf{y}\|_2} \quad (5.1)$$

where $0 \leq \theta \leq \pi/2$. With this measure, if $\theta = 0$ both patterns are the same, if $\theta = \pi/2$ the patterns are orthogonal. To perform the comparison we make a vertical cut in the p tterns, and compare the obtained vectors representing A-scans. For the plane-wave patterns of Fig. 5.4 we found that $\theta = 9.8040^\circ$. For the standing-wave patterns of Fig. 5.5 we found that $\theta = 8.6170^\circ$. Thus, we confirm that the patterns obtained for the homogeneous and inhomogeneous mediums are similar.

From looking at Figs. 5.4 and 5.5 we notice that it is possible to produce significant change in the refractive index without exceeding safety limits. The change obtained is

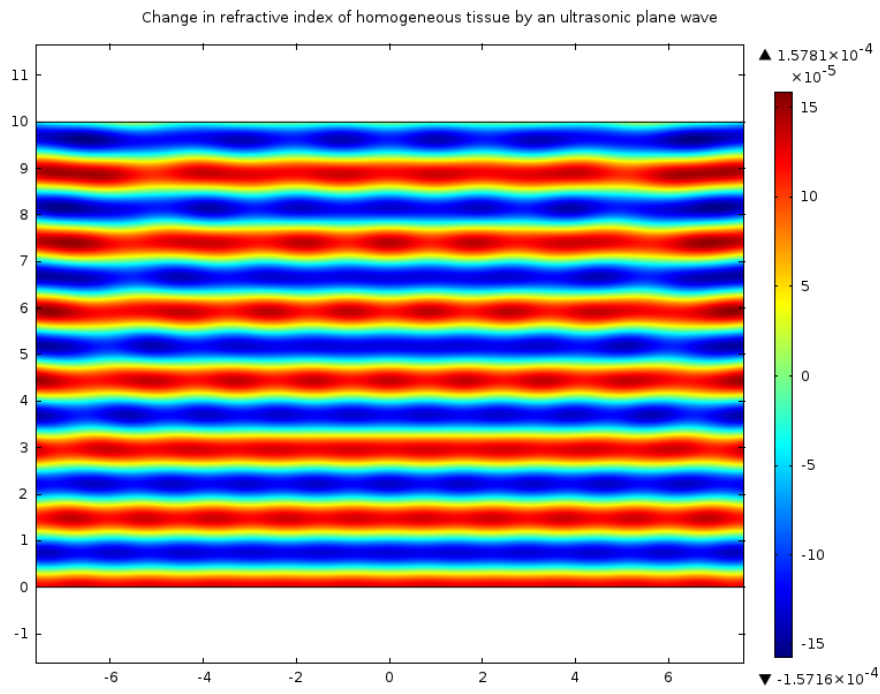


(a) Refractive index of homogeneous medium used in simulations

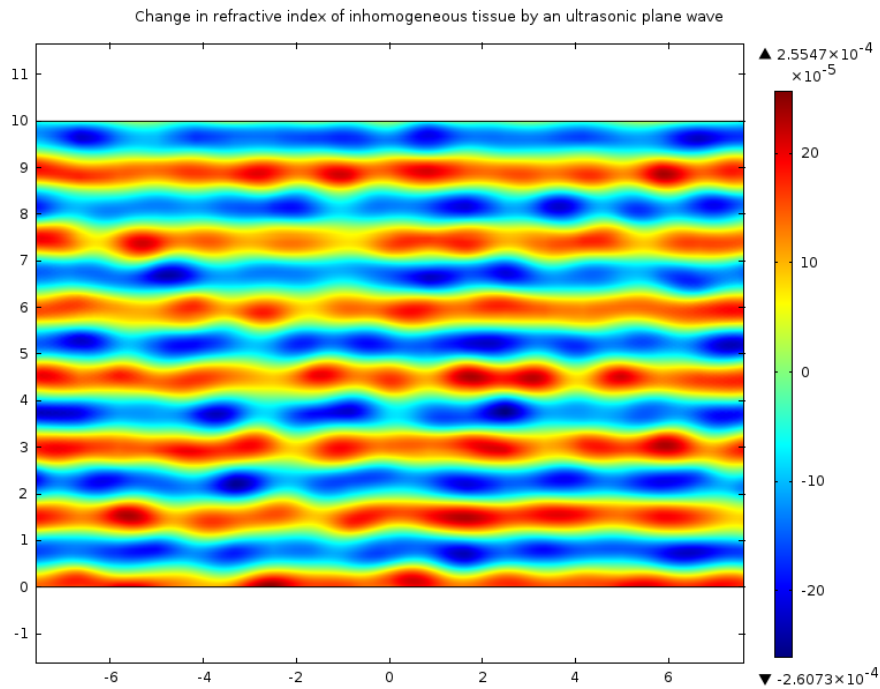


(b) Refractive index of inhomogeneous medium used in simulations

Figure 5.3. Refractive index of mediums used in simulations. Homogeneous medium with $n = 1.33$ on top (fig. (a)), and inhomogeneous medium with n varying between 1.3 to 1.5 bottom (fig. (b)).

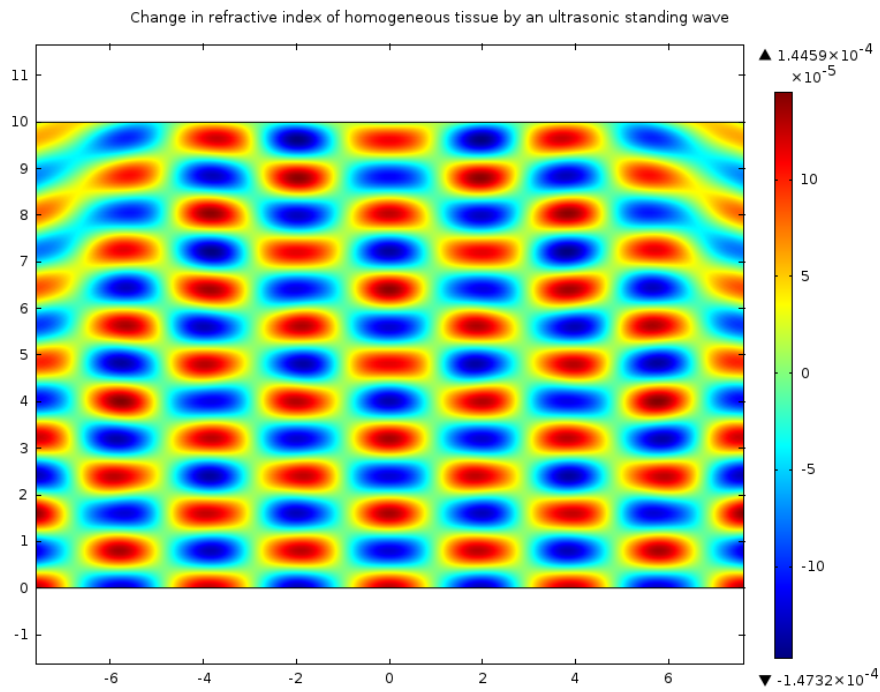


(a) Homogeneous medium

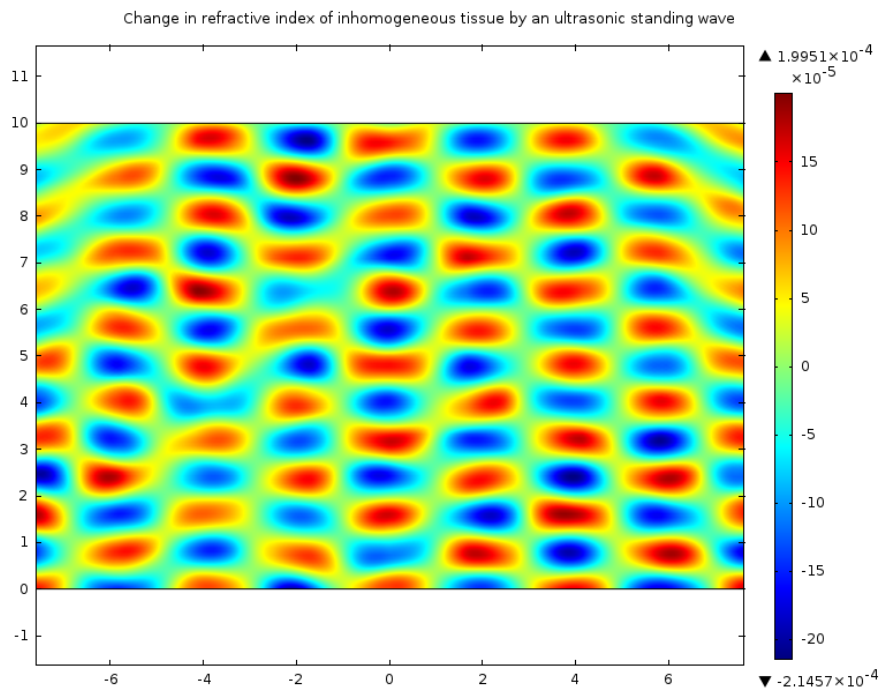


(b) Inhomogeneous medium

Figure 5.4. Fig. (a) shows the effect of an ultrasound plane wave on the refractive index of the homogeneous medium. Fig. (b) shows the effect of an ultrasound plane wave on the refractive index of the inhomogeneous medium.

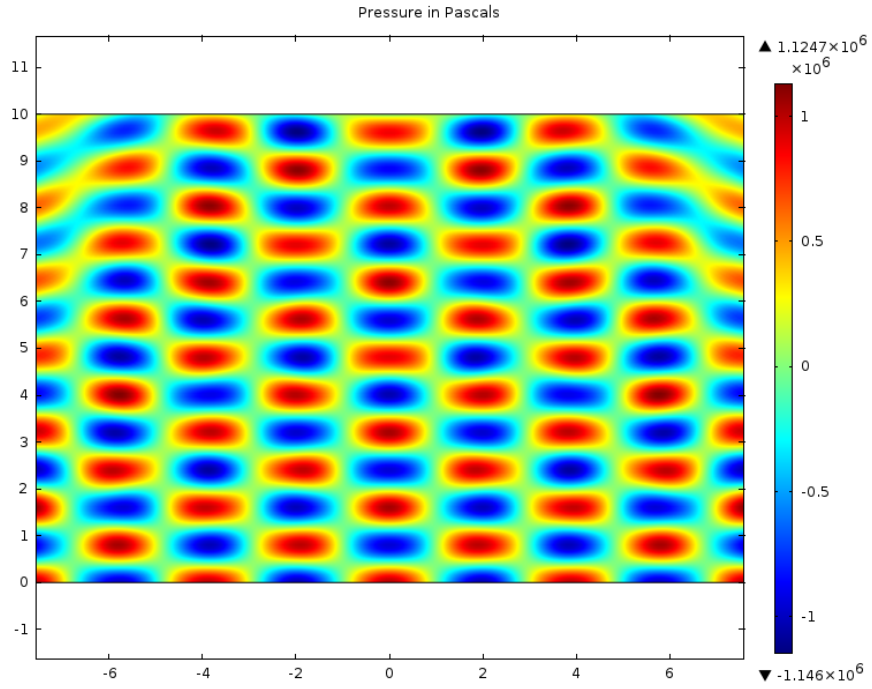


(a) Homogeneous medium

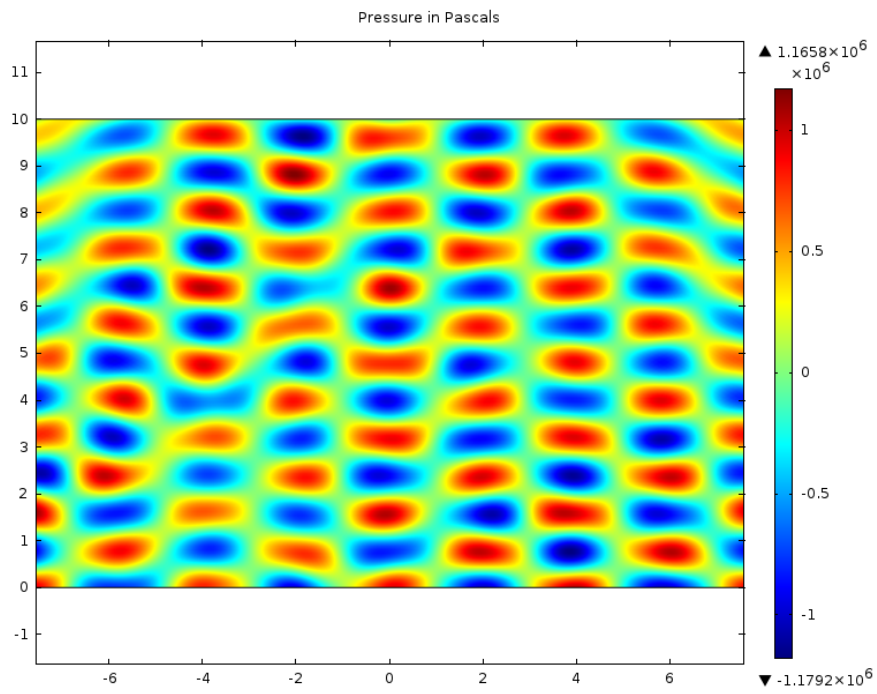


(b) Inhomogeneous medium

Figure 5.5. Effect of an ultrasound standing-wave pattern on the refractive index. Fig. (a) shows the effect on the homogeneous medium, while fig. (b) shows the effect on the inhomogeneous medium.



(a) Homogeneous medium



(b) Inhomogeneous medium

Figure 5.6. Pressure applied to the mediums to generate the ultrasound standing-wave patterns shown in fig. 5.5

approximately $\pm 2.5 \times 10^{-4}$. This is a significant change in the refractive index of the medium, corresponding to a propagation phase change $\Delta\phi = (2\pi/\lambda)\Delta n z$. If $\lambda = 1 \mu m$ then $\Delta\phi = 2\pi \times 10^6 \times 2.5 \times 10^{-4} z \approx 1571z$. Such change in phase should have a significant effect on light propagation in the medium.

5.5 Conclusions

The most important conclusion we can obtain from the simulation results shown is that we can design the ultrasound patterns assuming an homogeneous acoustic medium. We can see in Figs. 5.4 and 5.5 that even though there is some change in the ultrasound patterns due to acoustic scattering this change is small. And the ultrasound patterns remain similar to those obtained in the homogeneous medium. Even more, in these simulations the refractive index of the inhomogeneous medium varies between 1.3 and 1.5 randomly. In tissue variations of the refractive index this large are not common. It is much more common to have tissue of similar refractive index grouped together.

Another important conclusion can be made by looking at the scale of change in the refractive index in Figs. 5.4 and 5.5. I emphasize again that the applied pressure was adjusted to comply with the limits of safe application of ultrasound in tissue. Within these limits the change in refractive index obtained was $\Delta n \approx \pm 2.5 \times 10^{-4}$. This is a significant change in the refractive index of the medium, corresponding to a change in propagation phase $\Delta\phi = (2\pi/\lambda)\Delta n z$. If $\lambda = 1 \mu m$ then $\Delta\phi = 2\pi \times 10^6 \times 2.5 \times 10^{-4} z \approx 1571z$. Such change in phase should have a significant effect on light propagation in the medium. Therefore, ultrasound patterns can be designed assuming an homogeneous acoustic medium. This was also shown with an extreme case of optical inhomogeneous medium in the simulations. These results show that an integrated computational OCT system, that uses ultrasound patterns to modulate the refractive index of tissue, is feasible.

Chapter 6

Monte Carlo Methods

6.1 Introduction

So far in this thesis we have shown that it is feasible to modulate the refractive index of tissue using ultrasound patterns generated by a combination of acoustic beams, and that these patterns can be designed inside an homogeneous medium and still be applied to an inhomogeneous one. To assess the effectiveness of these patterns in minimizing multiple-scattering, we need to simulate the propagation of light within tissue modulated by ultrasound patterns. The simulation of light propagation involves the solution of an integral equation. Though there are many methods for solving integral equations we opted for a Monte Carlo method. We based this decision in the fact that Monte Carlo methods are efficient in high dimensions, and can solve the problem just at the domain of interest.

In the present chapter we present the basics of Monte Carlo methods. Then in chapter 7 we present an approach to use Reversible jump Markov chain Monte Carlo for solving Fredholm integral equations of the second kind. We implement the presented approach in MATLAB and test it with integral equations in 1-D and 2-D. We then present the results of our tests.

6.2 Why use a Monte Carlo approach?

The first question we have to answer is certainly, why Monte Carlo methods? There are a number of methods available to solve integral equations, e.g., method of moments, finite element method. These methods require some discretization of the problem, in the form of a grid or mesh, which when dealing with optical wavelengths can be quite big. Consider the region of $1 \text{ cm} \times 1 \text{ cm}$ used in our ultrasound simulations and a light source with wavelength of $1,300 \mu\text{m}$. Mesh elements should be smaller than the wavelength to obtain accurate results. Consider grid elements of half the wavelength in size, that would amount to approximately 237 million elements in the grid. If we want to solve in a volume of 1 cm^3 , then 3.64×10^{12} elements would be required if we use just two grid elements per wavelength. The amount of elements required though is usually more than 2, so the number of elements in the could quickly become impractical.

Another issue with these type of methods is that they are more suited to the global problem. That is, the problem has to be solved at all points in the domain, even if one is only interested in the solution at one point. As we mentioned before, the number of points is very high when dealing with optical wavelengths, so the size of the matrices generated by these methods is huge. The storage requirements for these matrices could also be impractical.

Our choice of a Monte Carlo method to solve our integral equation of interest is related to these issues. Monte Carlo methods are very computationally efficient, they are considered the best algorithm for evaluation of multidimensional integrals in cases of high dimension [26]. We note here that solving a Fredholm integral equation of second kind, as we want to do, requires the evaluation of integrals of increasing dimension. Monte Carlo methods have a convergence rate of \sqrt{N} regardless of dimension, with N being the number of points being evaluated. That is the Monte Carlo estimate converges to the exact value with a rate of \sqrt{N} . If one wants to reduce the standard deviation of a Monte Carlo algorithm by half, four times the number of samples are required [27]. Most quadrature rules

have their convergence rates reduced by a factor of $1/d$ as the dimension d increases [28], see table 6.1. Since the convergence of Monte Carlo is independent of the dimension there is always a value of d above which Monte Carlo converges to the solution faster than the other methods.

Table 6.1. Comparison of convergence rate of Monte Carlo with different quadrature rules

Convergence as function of n	
Monte Carlo	\sqrt{n}
Trapezoidal rule	$n^{-2/d}$
Simpson's rule	$n^{-4/d}$
Gauss' rule	$n^{-(2m-1)/d}$

Another very attractive characteristic of Monte Carlo methods is that they are well suited to solve both the local problem and the global problem. The global problem involves finding the solution on the whole domain, while the local problem means finding the solution at just one point in the domain or just one part of the domain. There are times when one is only interested in the solution of the problem at some region in the domain, in these cases Monte Carlo methods present the advantage of being able to solve the problem only at the points of interest; as opposed to solving the whole domain to obtain the solution for just those points. This in itself allows a simplification of the problem as the number of points in which the solution needs to be calculated is reduced. Also the storage required for a Monte Carlo method is minimal, as a minimum only the summation of the estimate has to be stored. From a more computational side Monte Carlo presents the advantage that the algorithm is naturally suited to run in parallel; thus reducing the amount of time required to obtain solutions of complicated problems.

In summary, we chose Monte Carlo as the algorithm to solve the integral equation of scattering because of the following reasons,

1. The Monte Carlo algorithm is computationally efficient, especially in high dimensions.
2. Solving an integral equation involves evaluating a sum of integrals of increasing dimensionality.

3. It is relatively easy to parallelize the algorithm.
4. Monte Carlo methods can solve the local problem, that means the problem can be solved only in the domain of interest.
5. It can solve the global problem as well, that is it can solve the problem on the whole domain.

6.3 History of Monte Carlo methods

Monte Carlo methods are a group of algorithms in which the solution to a particular problem is found through repeated random sampling. The essence of the method is to use distributions of random numbers that reflect a particular process in a series of processes, such as the diffusion of neutrons in various materials, to calculate samples that approximate the real diffusion history [29]. Although the name Monte Carlo method dates back to the 1950's, statistical sampling have been known for a long time. It was not very popular due to the complications of calculations involving random numbers. However, the development of powerful computers helped to greatly increase the number of applications for the Monte Carlo method.

It is generally considered that the Monte Carlo method was officially introduced in 1949, the year on which Metropolis and Ulam published their paper on the subject [26]. However there are earlier examples of what could be called Monte Carlo method. In 1777 George-Louis Leclerc posed the following problem: given a needle of length l dropped on a plane ruled with parallel lines t units apart, what is the probability P that the needle will cross a line? He established that the probability was $P = 2l/(\pi t)$. In 1866 Pierre Simon de Laplace showed that this result can be used to approximate the value of the number π by repeatedly throwing a needle onto a lined sheet of paper and counting the number of intersected lines [26, 27, 30].

The development and the intensive application of the Monte Carlo method are related to the Manhattan Project. That is the project by the United States to develop the first

nuclear weapons during World War II. This project brought together such people as John von Neumann, Enrico Fermi, Stan Ulam and Nick Metropolis. It also give them access to some of the most advanced computers of that time. In order to study neutron transport, what we now call Monte Carlo methods were developed.

6.4 Random number generation

The application of Monte Carlo methods usually require the generation of an uniform random number ξ between (0,1). Once this random number can be generated, any other type of probability distribution, or interval can be sampled from it. Before the advent of computers statistical sampling implied the use of random number tables, or some other methods, to obtain the random samples required. Nowadays Monte Carlo calculations in computers are the standard. A source of random numbers capable of delivering them at a rate adequate for a machine capable of billions of floating point operations per second is thus required. Natural sources of random events don't supply the events at such a rate. And it would require an enormous table of previously obtained random numbers to complete a large calculation [30]. Furthermore, repeatability of the results is desired in order to debug and verify our algorithms. Thus, pseudo-random number generators are used.

A pseudo-random number generator is an algorithm which creates a sequence of numbers with statistical properties similar to those of random numbers. One of the most important of this properties is called the period. The period of a pseudo-random number generator is the length of the sequence it creates.

6.4.1 Classes of pseudo-random number generators

The earliest, purely deterministic method of generating pseudo-random numbers was the mid-square method proposed by von Neumann. We take a number x_n that has $2a$ digits and square it. The result has $4a$ digits, of which the a most significant and the a least significant

digits are discarded. Then x_{i+1} is formed from the remaining digits as

$$x_{i+1} = \left[\frac{x_i^2}{b^a} \right] - \left[\frac{x_i^2}{b^{3a}} \right] \cdot b^{2a} \quad (6.1)$$

where b is the base of the number representation. This method was soon abandoned because of its short period.

Linear recurrence methods

Many pseudo-random number generators used are special cases of

$$x_{i+1} = a_0x_i + a_1x_{i-1} + \cdots + a_jx_{i-j} + b \pmod{P} \quad (6.2)$$

The generator is initialized with a vector of numbers with length $j + 1$. This generator has a period τ which cannot exceed $P^{j+1} - 1$. The statistical properties of the generated sequences depend on the selection of a_j , b , and P .

Multiple recursive generator

A multiple recursive generator uses the following recursion

$$x_{i+1} = a_0x_i + a_1x_{i-1} + \cdots + a_jx_{i-j} \pmod{P} \quad (6.3)$$

Multiple recursive generators can be combined to create a generator with larger a modulus and a longer period. If each individual modulus $P_\ell = 1, \dots, L$ is prime and each component generator has a period τ_ℓ , then the period of the combined generator is $\tau_1 \times \cdots \times \tau_L / 2^{L-1}$. The coefficients of each individual generator should be selected such that the combined generator has a good lattice structure in higher dimensions.

MATLAB random number generators

Monte Carlo simulations in this work were done in MATLAB. The default (pseudo)-random generator in this software is the Mersenne twister algorithm. However, since a number of independent random number streams were needed, the combined multiple recursive generator was used in the simulations. This generator has a period of 2^{127} and allows the creation of multiple independent streams [31].

6.5 Monte Carlo integration

Let us now discuss the solution of integrals via a Monte Carlo quadrature. Consider an integral of the form

$$G = \int_D g(x) f(x) dx \quad (6.4)$$

where

$$f(x) \geq 0 \quad \int_D f(x) dx = 1 \quad (6.5)$$

We must note here that the integral to be evaluated need not explicitly contain the function $f(x)$. It is possible to assume in those cases that $f(x) = 1/D$. That is $f(x)$ is a probability density function uniformly distributed in the domain D .

To evaluate such an integral the following procedure could be used. Draw samples x_1, x_2, \dots, x_N from the distribution $f(x)$, and form the estimate

$$G_N = \frac{1}{N} \sum_i g(x_i) \quad (6.6)$$

G_N is an estimator for G , and given a large enough number of samples N it is guaranteed that $\langle G_N \rangle$ will converge to the value of G , in other words

$$\langle G_N \rangle = G \quad \text{as} \quad N \rightarrow \infty \quad (6.7)$$

At this point we should make the interesting question, how big should N be? In order

to answer that we need to somehow evaluate the quality of the estimate constructed.

6.5.1 Assessment of the estimate

It is usually the case when performing Monte Carlo calculations that one would need evaluate the quality of the estimate. This quality evaluation can be needed to decide on a particular Monte Carlo algorithm to use. Or more importantly, to decide the number of samples required to obtain an estimate with a particular level of error. Since Monte Carlo methods are normally used to tackle problems that would require a great amount of time to be solved by other methods, it is not efficient to continue the calculation once the estimate is within a reasonable level of error.

Normally, to assess the quality of the estimate one would keep track of some other estimates during the Monte Carlo simulation. Let us call the accumulator for the estimate S_1 . Then an estimate of $z(x)$ is given by,

$$\bar{z} = \int_a^b z(x)f(x)dx \approx \frac{1}{N} \sum_{i=1}^N z(x_i) = \frac{S_1}{N}, \quad (6.8)$$

where x_i is a given random sample.

Another accumulator can provide an estimate of z^2 . Let us call it S_2 , then

$$\bar{z}^2 = \frac{1}{N} \sum_{i=1}^N z^2(x_i) = \frac{S_2}{N} \quad (6.9)$$

with this an estimate of the standard deviation of \bar{z} can be obtained,

$$s(\bar{z}) \approx \frac{1}{\sqrt{N}} \sqrt{\frac{N}{N-1} (\bar{z}^2 - \bar{z}^2)} = \frac{1}{\sqrt{N-1}} \sqrt{(\bar{z}^2 - \bar{z}^2)}, \quad (6.10)$$

We can notice in Eq. (6.10) that the standard deviation is proportional to $1/\sqrt{N}$ for large N. This means that to reduce the standard deviation by half, four times the number of samples would have to be recorded [27]. To improve the estimate without the need of an

incredible amount of samples being recorded we resort to variance reduction techniques.

6.6 Importance Sampling

The most basic variance reduction technique is importance sampling. Importance sampling involves choosing a sampling distribution that favours important samples [32]. That is samples concentrated in an area of greater contribution to the integral being solved. See Fig. 6.1 for an example of what this means.

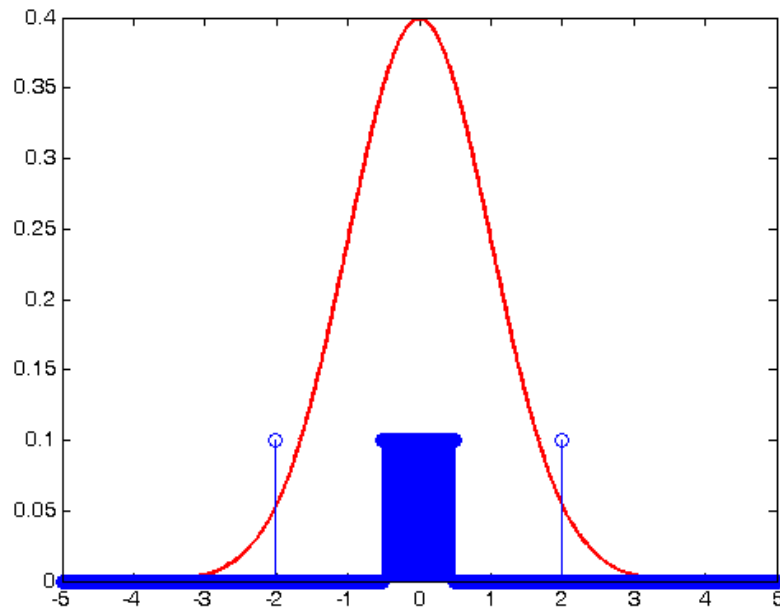


Figure 6.1. An example of importance sampling. The samples are concentrated in the area of higher value of the distribution. There are still a few samples in areas of lower value.

To explain the concept of importance sampling, let us suppose we want to evaluate integral in Eq. (6.4). Let us further assume that it is rather difficult to sample from the probability density distribution $f(x)$. In this case it is possible to introduce a different probability distribution function (pdf), $\tilde{f}(x)$, into the integral and sample from it as follows.

$$G = \int g(x) \frac{f(x)}{\tilde{f}(x)} \tilde{f}(x) dx \quad (6.11)$$

$\tilde{f}(x)$ should have the same characteristics of $f(x)$, as described in Eq. (6.5). Also, $\tilde{f}(x)$ should dominate $f(x)$. That is, $\tilde{f}(x) > 0$ wherever $f(x) > 0$. Taking random samples x_1, \dots, x_N from $\tilde{f}(x)$ it is possible to construct the unbiased estimator

$$\tilde{G}_N = \frac{1}{N} \sum_i g(x_i) \frac{f(x_i)}{\tilde{f}(x_i)} \quad (6.12)$$

where $\tilde{f}(x)$ is called importance sampling density, importance distribution or proposal density. The ratio of densities,

$$W(x) = \frac{f(x)}{\tilde{f}(x)} \quad (6.13)$$

is called the likelihood ratio, or just the weight of the samples. The likelihood ratio need only be known up to a constant, that is $W(x) = c w(x)$ [32].

We would like to select an $\tilde{f}(x)$ that minimizes the variance of the estimate. It is possible to prove (see [30]) that one can obtain zero variance if one selects

$$\tilde{f}(x) = \frac{g(x)f(x)}{G} \quad (6.14)$$

of course, this means that the correct answer G is known. In other words, if we know the correct answer we can always perform a Monte Carlo calculation that will return an estimate with zero variance. This does not seem like a very useful conclusion. Yet we can expect that if we select $\tilde{f}(x)$ somehow similar to Eq. (6.14) we should obtain estimates of very low variance.

6.7 Sequential Importance sampling

An extension of importance sampling is called *sequential importance sampling* or *dynamic importance sampling*. It is useful when one wants to sample from very complex distributions. The technique works as follows. First, we assume we can decompose \mathbf{X} as a vector $\mathbf{X} = (X_1, \dots, X_n)$, and it is possible to sample from a density $g(\mathbf{X})$ sequentially.

Specifically we assume,

$$g(\mathbf{X}) = g_1(x_1)g_2(x_2 | x_1) \dots g_n(x_n | x_1, \dots, x_{n-1}) \quad (6.15)$$

where we generate x_1 from density $g_1(x_1)$, x_2 is obtained from $g_2(x_2 | x_1)$, and so on until one obtains a random vector \mathbf{X} from $g(\mathbf{X})$. This is repeated N times obtaining vectors $\mathbf{X}_1, \dots, \mathbf{X}_N$ and then an estimate can be calculated using Eq. (6.12).

The importance weights will be represented as,

$$W(\mathbf{X}) = \frac{f(x_1)f(x_2 | x_1) \dots f(x_n | x_{1:n-1})}{g_1(x_1)g_2(x_2 | x_1) \dots g_n(x_n | x_{1:n-1})} \quad (6.16)$$

where we represent $f(\mathbf{X})$ sequentially and abbreviate (x_1, \dots, x_n) as $x_{1:n}$ [32]. In this way, importance sampling allows one to sample from complex distributions as long as those distributions can be represented as a joint distribution.

6.8 Markov Chain Monte Carlo

So far all the random samples we considered were independent and identically distributed (i.i.d.) random variables. There could be cases in which it is very difficult to obtain i.i.d. samples, e.g.: high-dimensional distributions. In these cases it is still possible to use the Monte Carlo method, by sampling correlated samples. These correlated samples are obtained using the sampling method known as Markov chain Monte Carlo.

Markov Chain Monte Carlo (MCMC) allows to approximately generate samples from any arbitrary distribution. An advantage of MCMC is that it only requires specification of the target probability distribution function (pdf) up to a normalization constant. MCMC was developed in Los Alamos in the decade of 1950's by Metropolis *et al.*. They were doing simulations of thermodynamic equilibrium of a liquid with its gas phase, and realized that they did not need to simulate the exact dynamics, but only some Markov chain with the same equilibrium distribution. Hastings generalized the the Metropolis algorithm in the

1970's, and simulations using his scheme are said to use the Metropolis-Hastings algorithm [33].

The usual approach to Markov chain theory is to start with the transition kernel $P(x, A)$. An important concern of Markov chain theory is to determine the conditions under which there exist an invariant distribution π^* and the conditions under which iterations of the transition kernel converge to the invariant distribution. MCMC methods turn the theory around. The invariant density is known, this is the target density $\pi(\cdot)$ from which the samples are desired, but the transition kernel is unknown. To generate samples from $\pi(\cdot)$ the methods find and utilize a transition kernel $P(x, dy)$ whose n th iterate converges to $\pi(\cdot)$ for a large n [34].

To do this we propose a generating density $q(x, y)$ where $\int q(x, y)dy = 1$. This density is to be interpreted as saying that when a process is at x , the density generates a value y from $q(x, y)$. The distribution $q(x, y)$ should satisfy the reversibility condition. That is,

$$\pi(x)q(x, y) = \pi(y)q(y, x) \quad (6.17)$$

then, we would have generated a proper Markov chain. To ensure this, we introduce a probability $\alpha(x, y) < 1$. We call $\alpha(x, y)$ the probability of move.

Transitions for x to y are given by

$$p_{MH}(x, y) = q(x, y)\alpha(x, y), \quad x \neq y$$

With $\alpha(x, y)$ being a probability its upper limit will be 1.

Since $\alpha(x, y)$ should satisfy the reversibility condition, we can state that

$$\pi(x)q(x, y)\alpha(x, y) = \pi(y)q(y, x)\alpha(y, x) \quad (6.18)$$

The transition kernel of the Metropolis-Hastings chain created by the algorithm is given by

$$P_{MH}(x, dy) = q(x, y)\alpha(x, y)dy + \left[1 - \int_R q(x, y)\alpha(x, y)dy\right] \delta_x(dy) \quad (6.19)$$

In summary, the Metropolis-Hastings algorithm works as follows;

- Set x_0 to randomly or deterministically
- Repeat for $k=1,2,\dots,N$
- Generate y from $q(x_k, \cdot)$ and u from $U(0,1)$
- If $u \leq \alpha(x_k, y)$

set $x_{k+1} = y$

- Else

set $x_{k+1} = x_k$

- Return the values $\{x_0, x_1, \dots, x_N\}$

The Metropolis-Hastings algorithm provides a method to sample from essentially any distribution regardless of its analytic complexity or its number of dimensions. However, sampling is correct only asymptotically. Meaning that the algorithm has to converge to the stationary distribution before providing correct samples. And as mentioned before the samples obtained are correlated, sometimes strongly.

6.8.1 Evaluating performance of MCMC

Evaluation of the performance of an MCMC sampler is related to the convergence of the Markov chain to its stationary distribution. Before we can do any correct estimation the Markov chain should have converged to this distribution. Note, that by the theory of MCMC the chain is constructed in such a way that this convergence is guaranteed, but there is no indication of the rate of convergence.

There are many methods of monitoring stationarity of the Markov chain. The simplest method is to plot the evolution of the chain in time and visually check for stationarity. A more elaborate method is the application of a test of fit, like the Kolmogorov-Smirnov test. This test can be applied to two different runs of an MCMC sampler or to two halves of one chain from one run. Since convergence is not usually exact, test of fit are not that used by practitioners of MCMC. A more common way to assess convergence of an MCMC sampler is to run several chains in parallel and perform an analysis of variance [35].

We can evaluate the variance of the estimate after we are sure of convergence to the stationary distribution. Since samples are now correlated, we can't proceed exactly as in the i.i.d. Monte Carlo case. We cite here a couple of methods described in [36], window estimators and specialized estimators for Markov chains. In his paper Geyer consider an integral,

$$\mu = \int g(x) dP(x) \quad (6.20)$$

For window estimators the lag t autocovariance is estimated first as,

$$\hat{\gamma}_{n,t} = \frac{1}{n} \sum_{i=1}^{n-t} [g(x_i) - \hat{\mu}_n][g(x_{i+t}) - \hat{\mu}_n] \quad (6.21)$$

where $\hat{\mu}_n$ is the Monte Carlo estimator of the integral using n samples. The variance can be estimated as

$$\sigma_n^2 = \sum_{-\infty}^{\infty} w_n(t) \hat{\gamma}_{n,t} \quad (6.22)$$

where $w_n(t)$ is some weight function called *lag window*.

The specialized method described by Geyer, assumes that $\Gamma_m = \gamma_{2m} + \gamma_{2m+1}$, a sum of adjacent pairs of autocovariance, is an strictly positive, decreasing function of m . A possible estimator of the variance is

$$\sigma^2 = -\hat{\gamma}_0 + 2 \sum_{i=1}^m \Gamma_{n,i} \quad (6.23)$$

where m is the largest integer such that $\Gamma_{n,i} > 0$.

We end this section by emphasizing that there are more methods to check convergence of MCMC samplers. There is not a method generally accepted as the most accurate or convenient, and all methods can fail to assess convergence in some cases. Thus, it is good practice to use more than one of the suggested methods when evaluating an MCMC sampler.

6.9 Reversible Jump Markov Chain Monte Carlo

Reversible Jump Markov chain Monte Carlo allows to generalize MCMC simulation to cases when the dimension of the parameter space can vary between iterates of the Markov chain. The reversible jump sampler can be viewed as an extension of the Metropolis-Hastings algorithm onto more general spaces [37]. The method was proposed by Green [38] as a framework to apply Markov chain Monte Carlo to Bayesian model determination problems where the dimensionality of the parameter vector is typically not fixed.

To explain the reversible jump sampler we first assume a number of candidate models $\mathcal{M} = \{\mathcal{M}_1, \mathcal{M}_2, \dots\}$ indexed by a parameter $k \in \mathcal{K}$. Each model \mathcal{M}_k has a n_k -dimensional vector of unknown parameters $\theta_k \in \mathbb{R}^{n_k}$, where n_k can take different values for different models $k \in \mathcal{K}$. As with standard Metropolis-Hastings transitions from a state $\theta = (k, \theta'_k)$ in model \mathcal{M}_k are done by first proposing a new state $\theta' = (k', \theta'_{k'})$ in model $\mathcal{M}'_{k'}$ from a proposal distribution $q(\theta, \theta')$, as before we must ensure the reversibility condition. The reversibility condition is ensured through the acceptance probability $\alpha(\theta, \theta')$, where the proposed move is accepted with probability $\alpha(\theta, \theta')$. If rejected the chain stays in the previous state θ in model $(M)_k$ [37].

To implement this sampler we first define a set of moves M . We sample the next move from $q(\cdot, (k, \theta)) = p_\theta(m)q_m(\theta, \theta')$. That is we select a move type from $p_\theta(m)$, which means the dimension k is increased, decreased or left unchanged. Then a new θ is selected from a transition probability $q_m(\theta, \theta')$ [32]. We can summarize the reversible jump sampler as follows,

Given the current state θ_k

1. Generate $m \sim p_{\theta_k}(m)$
2. Generate $\theta' \sim q_m(\theta, \theta')$. Let m' be the reverse move.
3. Generate $U \sim U(0, 1)$ and deliver

$$\theta_{k'} = \begin{cases} \theta' & \text{if } U \leq \alpha \\ \theta_k & \text{otherwise,} \end{cases}$$

where

$$\alpha = \min \left\{ \frac{f(\theta')p_{\theta'}(m')q_{m'}(\theta', \theta_k)}{f(\theta_k)p_{\theta_k}(m)q_m(\theta_k, \theta')}, 1 \right\} \quad (6.24)$$

6.9.1 Evaluating performance of RJMCMC

RJMCMC is a generalization of MCMC where the dimensions of the parameters studied can vary between iterations. Taking this into consideration the methods discussed for MCMC could be applied in this case too, with proper care taken about the change of dimensionality. To use the methods for fixed-dimension MCMC one has to perform the diagnostics on parameters whose interpretation does not change when there is a change of dimension. However, such parameters could be difficult to identify. There are some methods suggested specifically for the RJMCMC case, but they require special constraints on the formulation of the problem [39]. All things considered, evaluating performance of RJMCMC is a difficult matter and one which requires careful consideration of the problem being solved.

6.10 Summary

The present chapter introduced the Monte Carlo method. We make emphasis on importance sampling as a variance reduction technique and Markov chain Monte Carlo as a very

powerful sampling method. In the next chapter we show the application of Monte Carlo methods to solve Fredholm integral equations of the second kind, of which the equation of scalar optical scattering is an example.

Chapter 7

Implementation of a RJMCMC solver for Fredholm integral equations of the second kind

7.1 Introduction

In the previous chapter we discussed the basics of Monte Carlo methods. Our interest in Monte Carlo methods come from our wish to use them to solve integral equations. In particular we want to solve the equation for optical scattering.

$$U^t(r) = U^i(r) + \int_{vol} O(r')G(r - r')U^t(r') dr' \quad (7.1)$$

where $U^t(r)$ is the total optical field, $U^i(r)$ is the incident field, $G(r - r')$ is a Green's function, and $O(r')$ is the object. Eq. (7.1) is what we call a Fredholm integral equation of the second kind. Finding analytical solutions for Fredholm equations of the second kind is usually difficult, and can be done only for the simplest of problems. Therefore, much work has been devoted to developing numerical methods to obtain approximate solutions of these systems.

Numerical approaches for the solution of integral equations can be broadly categorized as either deterministic techniques; or Monte Carlo techniques [40]. Deterministic techniques typically depend upon quadrature or finding a suitable finite representation of the system and using numerical techniques to solve for that representation. However, obtaining a good finite dimensional representation of the system can be complicated, in particular for high dimensional problems. Therefore, we chose to concentrate in Monte Carlo methods.

7.2 Linear integral equations

An equation in which an unknown function appears in a linear fashion under one or more signs of integration is called a *linear integral equation*. The study of such equations is mostly of interest because they appear rather naturally in various practical problems [41]. Examples of integral equations are shown below;

$$g(x) = \int_a^b K(x, y)f(y) dy \quad x \in [a, b] \quad (7.2)$$

$$f(x) = g(x) + \int_a^b K(x, y)f(y) dy \quad x \in [a, b] \quad (7.3)$$

Eq. (7.2) is normally known as a *Fredholm integral equation of the first kind*; and Eq. (7.3) is normally known as a *Fredholm integral equation of the second kind*. Equation 7.3 is the main concern of this chapter. The integral in Eq. (7.2) and Eq. (7.3) is also called an *integral operator* A . That is, the integral operator is defined by

$$(A\varphi)(x) := \int_D K(x, y)\varphi(y) dy, \quad x \in D,$$

This operator is bounded with norm [42]

$$\|A\|_\infty = \int_D |K(x, y)| dy. \quad (7.4)$$

7.3 Monte Carlo methods to solve integral equations

We are interested in solving Eq. (7.3) using Monte Carlo methods. The first step in looking for the solution through Monte Carlo methods is to expand the solution of Eq. (7.3) into a Neumann series [40, 43],

$$\begin{aligned} f(x) &= g(x_0) + \int_{vol} K(x_0, x_1) f(x_1) dx_1 \\ &= g(x_0) + \int_{vol} K(x_0, x_1) g(x_1) dx_1 + \int_{vol} \int_{vol} K(x_0, x_1) K(x_1, x_2) g(x_2) dx_1 dx_2 + \dots \end{aligned} \quad (7.5)$$

existence and uniqueness of a solution can be established by the Neumann series provided that the operator A , as defined by Eq. (7.4), is a contraction; i.e. $\|A\| < 1$ [42].

Eq. (7.5) can be expressed as an iterative process like,

$$f(x) = g(x_0) + \sum_{n=1}^{\infty} \int_{vol} \left(\prod_{m=1}^n K(x_m, x_{m-1}) \right) g(x_n) dx_{1:M} \quad (7.6)$$

The integral on the right-hand side of Equation 7.5 resembles a random walk process with a transition kernel $K(x_m, x_{m-1})$. At each step of the random walk x_m , one samples $K(x_m, x_{m-1})$ to obtain the next point of the walk. The process is repeated till the walk is terminated at some point x_M . This requires that $K(x_m, x_{m-1}) \leq 1$, so that the walk will terminate at some point. This last condition was already established as a condition for the existence and uniqueness of the solution of Eq. (7.3). Thus we have that Eq. (7.6) can be approached as a random walk.

An approach to the solution based on random walks could be summarized as follows,

- Generate N random walks of length $M + 1$, $\mathbf{X}_{0:M}^n$
- Average over the weights W^n of N random walks $\mathbf{X}_{0:M}^1, \dots, \mathbf{X}_{0:M}^N$ generated before

where we note that W^n means the weight associated with the random walk n of the N

random walks generated. And M^n indicates the maximum length M for random walk n . This notation is maintained through this chapter, a super-script n indicates that the quantity is from the Markov chain n , it does not indicate a power.

7.4 A solution based on sequential importance sampling

Computing Eq. (7.6) involves an infinite sum of integrals of increasing dimensions. As has been stated an approach based in random walks seems reasonable. One of the earliest methods to solve integral equations via Monte Carlo methods was proposed by Albert in the 1950's [44]. This standard approach is based on sequential importance sampling, and we will summarize it here.

Consider a Markov chain with an initial probability distribution given by $\mu(x)$ and a transition probability $P(y, x)$ that gives the probability to move to state y when the current state is x . We have to select $\mu(x)$ such that the initial point can be anywhere in the domain of interest. That is $\mu(x) > 0$ in the whole domain. $P(y, x)$ should be selected such that it dominates $K(y, x)$, that is $P(y, x) > 0$ whenever $K(y, x) \neq 0$. We must also select a probability P_d that the chain will be terminated at a certain step. That is the chain possess an absorbing state. We note that instead of choosing a probability for the termination of the chain, we could also sample from a distribution of possible lengths of the chain at the beginning of the algorithm.

Having defined these probabilities we can now proceed as follows,

1. Simulate N independent random walks $\mathbf{X}_{0:M^n}^n$ with $n = 1, \dots, N$, from x_0 selected from the distribution $\mu(x)$ till absorption at length M .
2. Assign each of the simulated random walks a weight given by

$$W^n(\mathbf{X}_{0:M^n}^n) = \frac{g(x_M)}{\mu(x)} \prod_{m=1}^M \left(\frac{K(x_m, x_{m-1})}{P(x_m, x_{m-1})} \right), \quad (7.7)$$

3. Then the average

$$f(x) = \frac{1}{N} \sum_{n=1}^N W^n(\mathbf{X}_{0:K}^n) \delta(x - \mathbf{X}_0^n) \quad (7.8)$$

gives an unbiased estimate of 7.6.

The estimate provided by Equation 7.8 is for the global case. That is it estimate the value of the full function $f(x)$. Sometimes it is just required to obtain a solution $f(x_0)$ at a certain point x_0 . This is normally called the local problem, and it can be solved very efficiently via Monte Carlo methods. Using the sequential importance sampling method we are describing we proceed as follows to solve the local problem,

1. Select the starting point for the random walks x_0 . We are solving for this point.
2. Generate N independent random paths from x_0 to x_M .
3. Assign each of the generated random walks an importance weight given by,

$$W^n(\mathbf{X}_{0:M}^n) = g(x_M) \prod_{m=1}^M \left(\frac{K(x_m, x_{m-1})}{P(x_m, x_{m-1})} \right), \quad (7.9)$$

4. Then the following is an unbiased estimate of $f(x_0)$,

$$f(x_0) = \frac{1}{N} \sum_{n=1}^N W^n(\mathbf{X}_{0:M}^n) \quad (7.10)$$

Equation 7.10 provides an estimate of the function at the point x_0 . Thus, we could use Monte Carlo methods to solve the integral only on the points of our interest.

7.4.1 Limitation of the sequential importance sampling approach

The estimates provided by Eq. (7.8) and Eq. (7.10) will have reasonable variance if the variance of the importance weights remains small. However this can be difficult to assure. First, we are imposing an arbitrary distribution to the path lengths, through P_d which might not always be appropriate. And also, if $P(y, x) \neq K(y, x)$, or at least very close, the

variance of the weights will increase exponentially fast [40]. In many applications the kernel of the integral equation can not be selected as the transition probability.

These performance problems of sequential importance sampling are usually solved using a resampling step. The resampling consist in monitoring the variance of the weights during execution, and when it becomes too large discard those paths with small weights and multiply those with high weights. Then all the remaining weights are set to the same value in a way that retains the expectation of the estimator. However, when solving integral equations the starting point of the paths is very important. And every time a resampling step is used it decreases the diversity of the paths starting from time 0. Thus, the resampling step should be avoided when using sequential importance sampling to solve integral equations. In the next section we present an alternative method to solving integral equations using Monte Carlo methods.

7.5 A solution based on RJMCMC

We noted that a problem of sequential importance sampling is related to the selection of a transition probability $P(y, x)$ that is close enough to the kernel of the integral equation. This difficulty increases the variance of the importance weights. It would be very interesting to be able to sample from an optimal importance distribution. One that minimizes the variance of the importance weights. In [40] it was shown that this is possible using trans-dimensional Monte Carlo, like the reversible jump Markov chain Monte Carlo sampler.

We define this optimal importance distribution as

$$\pi_M(\mathbf{X}_{0:M}) = \frac{1}{c_M} \left| g(x_M) \prod_{m=1}^M (K(x_m, x_{m-1})) \right| \quad (7.11)$$

where π_M is a joint posterior distribution for a certain path of length $M + 1$, and c_M is a

normalization factor defined as

$$c_M = \int_D \left| g(x_M) \prod_{m=1}^M (K(x_m, x_{m-1})) \right| dx_{0:M}$$

We also consider that,

$$p_M = \frac{c_M}{C}$$

with C defined as

$$C = \sum_{m=1}^{\infty} c_m$$

We note that if the series converges then $C < \infty$.

Taking all these considerations, we can sample from the optimal importance distribution using RJMCMC as follows.

1. Define two sampling distributions, for update and birth, q_u, q_b
2. We sample from $\pi(\mathbf{X}_{0:M})$ using a reversible jump sampler as follows,
 - (a) Set an starting vector $X_{0:M}^0$ of arbitrary length $M + 1$
 - (b) Select uniformly among the move types, birth, death and update.
 - i. In case of an update move,
 - A. Sample J uniformly from $0, \dots, M$
 - B. Sample a new value for x_J from a density $q_u(X_J, \cdot)$
 - C. Accept the new value with probability,¹

$$\min \left\{ 1, \frac{\pi(M, (\mathbf{X}_{0:J-1}, x_J, \mathbf{X}_{J+1:M})) q_u(x_J, x_j)}{\pi(M, (\mathbf{X}_{0:M})) q_u(x_j, x_J)} \right\} \quad (7.12)$$

If the new value for X_J is not accepted the chain does not change from the previous iteration

- ii. In case of a birth move,

¹In these equations the index J represents the new value at the point and the index j the previous value

- A. Increase the dimension of the chain to $M + 1$
- B. Sample J uniformly from $0, \dots, M$
- C. Sample a new value for x_J from a density $q_b(X_J)$
- D. Accept the new value with probability,

$$\min \left\{ 1, \frac{\pi(M + 1, (\mathbf{X}_{0:j-1}, x_J, \mathbf{X}_{j:M}))d_{m+1}}{\pi(M, (\mathbf{X}_{0:M}))q_b(x_J)b_m} \right\} \quad (7.13)$$

where d_k is the probability of obtaining a death move and b_k is the probability of obtaining a birth move. In our described method, they are equal. If the new value X_J and dimension $M + 1$ are not accepted the chain does not change from the previous iteration

iii. In case of a death move,

- A. Decrease the dimension of the chain to $M - 1$
- B. Sample J uniformly from $0, \dots, M$
- C. Eliminate the value at x_J from the chain,
- D. Accept the new value with probability,

$$\min \left\{ 1, \frac{\pi(M - 1, (\mathbf{X}_{0:j-1}, \mathbf{X}_{j+1:M}))q_b(x_J)b_{m-1}}{\pi(M, (\mathbf{X}_{0:M}))d_m} \right\} \quad (7.14)$$

where d_m is the probability of obtaining a death move and b_m is the probability of obtaining a birth move. In our described method, they are equal. If the elimination of the value X_J and dimension decrease to $m - 1$ are not accepted the chain does not change from the previous iteration

3. At this point, assuming

$$c_0 = \int_D |g(x) dx|$$

is known or can be estimated we can use the following estimation of the solution

$$\hat{f}(x) = \frac{C}{N} \sum_{n=1}^N \text{sgn} \left(g(x_M^n) \prod_{m=1}^M K(x_m^n, x_{m-1}^n) \right) \delta(x - \mathbf{X}_0^n) \quad (7.15)$$

where we can use an estimate of the normalization factor C obtained from

$$\hat{C} = \frac{c_0}{\hat{p}_0}$$

where \hat{p}_0 is the number of random samples with length $M = 0$.

Equation 7.15 provides a global estimate Equation 7.6.

We note here that to calculate the ratios in (7.12)-(7.14) we need only to be able to compute

$$\frac{\pi(M, \mathbf{X}_{0:M})}{\pi(L, \mathbf{X}_{0:L})} = \frac{c_M \pi(\mathbf{X}_{0:M})}{c_L \pi(\mathbf{X}_{0:L})} = \frac{\left| g(x_M) \prod_{m=1}^M (K(x_m, x_{m-1})) \right|}{\left| g(x_L) \prod_{l=1}^L (K(x_l, x_{l-1})) \right|}$$

7.5.1 Local Solution through RJMCMC

A local estimate of Eq. (7.6) can also be obtained. In this case we adjust the algorithm to sample from a distribution $\pi(M, (x_0 : \mathbf{X}_{1:M}))$, defined as

$$\pi_K(x_0 : \mathbf{X}_{1:M}) = \frac{1}{c_M} \left| g(x_M) \prod_{m=1}^M (K(x_m, x_{m-1})) \right| \quad (7.16)$$

that is x_0 is a fixed point and we sample chains $\mathbf{X}_{1:M}$ of length M through RJMCMC.

In this case we need to be able to estimate the constant c_1 , defined as

$$c_1(x) = \int_D |g(y)K(x, y)| dy$$

and we can get an estimation of C from

$$\hat{C} = \frac{c_1}{\hat{p}_1}$$

where \hat{p}_1 is the proportion of samples of length $M = 1$. Finally, we can find an estimate of the solution through

$$\hat{f}(x_0) = g(x_0) + \frac{\hat{C}}{N} \sum_{n=1}^N \text{sgn} \left(g(x_M^n) \prod_{m=1}^M K(x_m^n, x_{m-1}^n) \right) \quad (7.17)$$

7.6 Implementation of a RJMCMC solver for Fredholm integral equations of second kind

We implemented the described RJMCMC solver for Fredholm integral equations of second kind as a MATLAB function. To test our function we use it to solve integral equations with known analytical solutions. The analytical solutions of the test problems can either be found in the literature or found via straight application of a formula. For each one of these examples we ran the algorithm ten times. We report averaged results, relative error compared to analytical result and variance between the different runs.

Example 1. *The integral equation of the second kind*

$$f(x) = .9x^2 + \int_0^1 .5x^2y^2 f(y) dy \quad (7.18)$$

whose exact solution is given by $f(x) = x^2$ [43].

Results are shown in Table 7.1 and Fig. 7.1.

Example 2. *We solve the equation*

$$f(x) = \frac{2}{3}e^x + \int_0^1 \frac{1}{3}e^{(x-y)} f(y) dy \quad (7.19)$$

which has the known exact solution $f(x) = e^x$ [40].

Results for this example are shown in Table 7.2 and Fig. 7.2.

Table 7.1. Results obtained for Eq. (7.18) with 10 runs of 300,000 samples.

x	Exact	Estimate	% Error	Variance
0.1	0.01	0.0100	-0.0005	1.5×10^{-12}
0.2	0.04	0.0400	-0.0009	4.9×10^{-11}
0.3	0.09	0.0900	-0.0001	1.1×10^{-10}
0.4	0.16	0.1600	-0.0019	2.4×10^{-10}
0.5	0.25	0.2500	-0.0023	1.17×10^{-9}
0.6	0.36	0.3599	-0.0019	4.3×10^{-9}
0.7	0.49	0.4900	-0.0013	1.3×10^{-9}
0.8	0.64	0.6400	-0.0036	9.3×10^{-9}
0.9	0.81	0.8100	-0.0010	6.4×10^{-9}
1	1	1.0000	-0.0060	1.7×10^{-8}

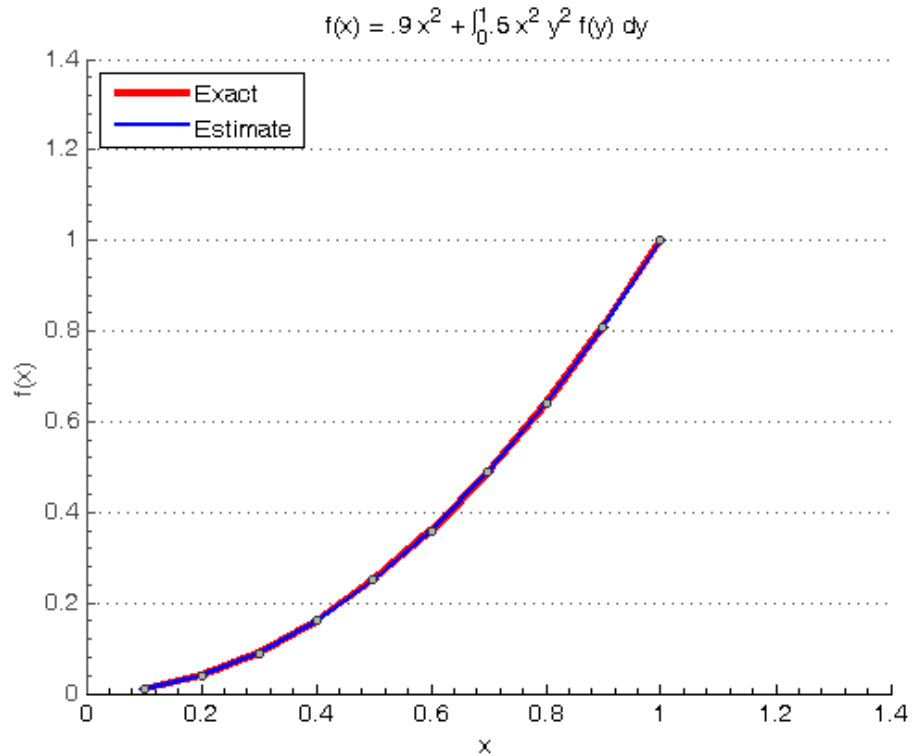


Figure 7.1. The curves of the analytical solution and the estimates for Eq. (7.18) are on top of each other in the graph. Points of evaluation are shown.

Table 7.2. Results obtained for Eq. (7.19) with 10 runs of 300,000 samples used for the estimation.

x	Exact	Estimate	% Error	Variance
0.1	1.1052	1.1051	0.0081	1.6×10^{-6}
0.2	1.2214	1.2226	-0.0969	1.1×10^{-6}
0.3	1.3499	1.3490	0.0634	5.2×10^{-6}
0.4	1.4918	1.4919	-0.0028	1.7×10^{-6}
0.5	1.6487	1.6495	-0.0491	6.9×10^{-6}
0.6	1.8221	1.8218	0.0149	1.6×10^{-6}
0.7	2.0138	2.0136	0.0072	8.2×10^{-7}
0.8	2.2255	2.2256	-0,0027	8.5×10^{-6}
0.9	2.4596	2.4587	0.0381	1.2×10^{-5}
1	2.7183	2.7197	-0.0534	7.6×10^{-6}

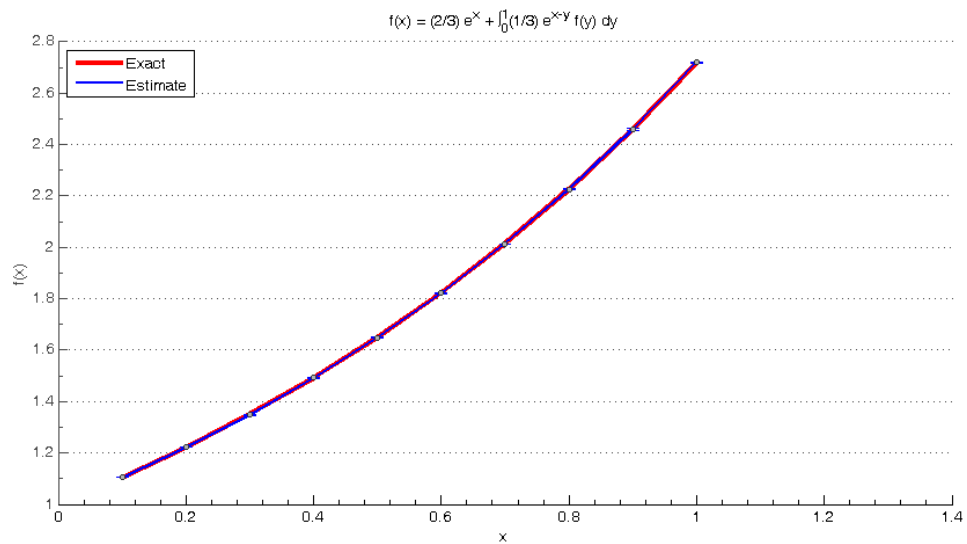


Figure 7.2. Figure shows the curves of the estimate and analytical solution on top of each other.

Example 3. *The equation*

$$f(x) = x + (1 - x)e^x + \int_0^1 x^2 e^{y(x-1)} f(y) dy \quad (7.20)$$

has the solution $f(x) = e^x$ [43].

We show our results in Table 7.3.

Table 7.3. Results obtained for Eq. (7.20) with 10 runs of 300,000 samples used for the estimation.

x	Exact	Estimate	% Error	Variance
0.1	1.1052	1.1053	-0.0091	5.2×10^{-10}
0.2	1.2214	1.2217	-0.0287	1.3×10^{-8}
0.3	1.3499	1.3507	-0.0588	7.3×10^{-8}
0.4	1.4918	1.4929	-0.0736	3.8×10^{-7}
0.5	1.6487	1.6504	-0.1045	1.6×10^{-6}
0.6	1.8221	1.8237	-0.0863	1.1×10^{-6}
0.7	2.0138	2.0153	-0.0772	5.1×10^{-6}
0.8	2.2255	2.2257	-0.0088	6.1×10^{-6}
0.9	2.4596	2.4590	0.0222	2.6×10^{-5}
1	2.7183	2.7132	0.1884	2.6×10^{-5}

The next example has separable kernels as described in [45]. Their analytical solution is given by,

$$f(x) = g(x) + \lambda h j(x) \quad \text{where, } h = \left(1 - \lambda \int_a^b j(y)k(y)dy\right)^{-1} \int_a^b k(y)g(y)dy$$

Example 4.

$$f(x) = \frac{1}{8}e^x + \frac{1}{4} \int_0^1 x e^y f(y) dy \quad (7.21)$$

We show our results for this example in Table 7.4.

Example 5. *This example shows another integral equation with a separable kernel,*

$$f(x) = 2 + x + \int_0^1 x^2 e^{(y-x)} f(y) dy \quad (7.22)$$

Result for this example are in Table 7.5.

Table 7.4. Results obtained for Eq. (7.21) with 10 runs of 300,000 samples use for the estimation.

x	Exact	Estimate	% Error	Variance
0.1	0.1515	0.1514	0.0036	7.7×10^{-10}
0.2	0.1793	0.1792	-0.0189	3.4×10^{-9}
0.3	0.2087	0.2087	0.0260	1.7×10^{-8}
0.4	0.2397	0.2398	0.0432	2.2×10^{-8}
0.5	0.2726	0.2727	0.0265	3.3×10^{-8}
0.6	0.3076	0.3076	0.0053	2.3×10^{-8}
0.7	0.3449	0.3449	0.0180	1.0×10^{-7}
0.8	0.3847	0.3847	0.0152	8.6×10^{-8}
0.9	0.4272	0.4274	0.0444	6.3×10^{-8}
1	0.4729	0.4729	0.0023	7.7×10^{-8}

Table 7.5. Results obtained for Eq. (7.22) averaging 10 runs of 300,000 samples.

x	Exact	Estimate	% Error	Variance
0	2	2	0	0
0.1	2.1602	2.1602	-0.0015	-0.0001
0.2	2.4179	2.4183	0.0135	0.0016
0.3	2.7437	2.7436	-0.0031	-0.0005
0.4	3.1137	3.1153	0.0486	0.0094
0.5	3.5091	3.5094	0.0092	0.0023
0.6	3.9148	3.9142	-0.0147	-0.0045
0.7	4.3193	4.3238	0.1043	0.0390
0.8	4.7137	4.7153	0.0325	0.0145
0.9	5.0916	5.0908	-0.0151	-0.0083
1	5.4482	5.4469	-0.0226	-0.0134

The next problem shows that our implementation can be extended to higher dimensions, in this case two. To extend our solver to 2-D we sample a vector (x, y) instead of an scalar at each step of the random walk.

Example 6. Consider the 2-D integral equation,

$$f(x, y) = g(x, y) + \int_0^1 \int_0^1 \frac{x}{(8+y)(1+t+s)} f(t, s) dt ds \quad (7.23)$$

where

$$g(x, y) = \frac{1}{(1+x+y)^2} - \frac{x}{6(8+y)}$$

its exact solution is $f(x) = (1 + x + y)^{-2}$ [46].

We show the results obtained in Table 7.6.

Table 7.6. Results obtained for equation 7.23 with 10 runs of 4 million samples used for the estimation.

(x, y)	Exact	Estimate	% Error	Variance
(0,0)	1	1	0	0
(0.1,0.1)	0.6944	0.6958	0.2041	-0.2117
(0.2,0.2)	0.5102	0.5140	0.7539	-0.2498
(0.3,0.3)	0.3906	0.3957	1.3065	-0.2391
(0.4,0.4)	0.3086	0.3167	2.6326	-0.2164
(0.5,0.5)	0.2500	0.2579	3.1727	-0.1914
(0.6,0.6)	0.2066	0.2188	5.9076	-0.1709
(0.7,0.7)	0.1736	0.1888	8.7659	-0.1531
(0.8,0.8)	0.1479	0.1604	8.4254	-0.1345
(0.9,0.9)	0.1275	0.1387	8.7813	-0.1194

We consider the results obtained to be accurate. There is still adjustments to be done when increasing the dimensionality of the problem, although the 2-D problem solved shows that the method can be extended to higher dimensions and still provide good estimates of the solution. An analysis of the variance between the different runs was performed to evaluate the quality of the estimation in these problems. A method to assess the convergence of the algorithm will be required for problems with unknown analytical solutions. Many methods to assess convergence of RJMCMC have been proposed in the literature. However, none of them works in all cases, and some of them only work under certain constraints [36, 39, 47]. We can say that convergence analysis on RJMCMC should be considered in an application by application basis. In future work we shall develop a suitable convergence analysis for our integral equation solver.

7.7 Conclusion

The standard method for solving integral equations of the second kind via Monte Carlo methods is called *Sequential Importance Sampling*. It has difficulties related to the proper

selection of a transition kernel to generate the required random walks, and in setting the length of said random walks. An alternative, to bypass the limitations of sequential importance sampling, was proposed in [40]. That is using *Reversible Jump Markov chain Monte Carlo*.

By using RJMCMC to generate the random walks required in the sequential importance sampling methods, one can obtain samples from the optimal importance distribution. Thus obtaining an estimate of the solution of the integral equation that has very small variance. We presented in this chapter the estimate obtained when using RJMCMC to sample the random walks, for both the local and the global case. We also presented a number of examples that show the accuracy of the obtained estimate.

Chapter 8

Conclusions and future work

8.1 Conclusions

OCT is an imaging technique that provides very-high resolution sub-surface imaging. The technique is already an established diagnostic tool in ophthalmology and many other fields of applications are being studied. Most of them in biomedical imaging. In this field the technique has been demonstrated for optical biopsy, cancer detection, coronary disease assessment, gastrointestinal imaging, and even caries detection. Indeed many interesting applications are possible by this technique. The major limitation for many applications of OCT imaging is its very shallow imaging depth of only a few millimetres. The most important factor limiting OCT imaging depth is multiple scattering.

In this thesis we presented a novel approach to reduce the effect of multiple scattering in OCT imaging, and in that way increase imaging depth of this technique. This approach is based on integrated computational imaging, also called hybrid imaging. Integrated computational imaging modifies the optical system, and accounts for this modification in a posterior digital post-processing step. Post-processing and modified optical acquisition are part of a single system and are designed likewise. The proposed hybrid OCT system modulates the sample by using ultrasound patterns.

There is anecdotal evidence that the presence of ultrasound in tissue improves OCT

imaging depth. Experiments showing this improvement just applied an ultrasound beam to tissue at the same time that OCT imaging was performed and found that multiple scattering was reduced. We proposed the design of an ultrasound pattern to obtain this reduction of multiple scattering by performing simulations of the interaction of light with tissue modulated by ultrasound. One such design studied in this thesis is based on standing-wave ultrasound patterns. The use of these ultrasound patterns to “improve” OCT is approached as an integrated computational OCT system.

The goal of this thesis was to show the feasibility of this integrated computational OCT system. We performed simulations in COMSOL Multiphysics of the tissue interaction of ultrasound causing a change in refractive index. These simulations show that the ultrasound patterns can produce enough of a change in the refractive index without exceeding the safety limits of ultrasound application. These simulations also show that acoustic scattering does not have much effect on the ultrasound patterns. Even in the presence of a highly inhomogeneous medium the pattern still resemble the one applied to an homogeneous medium. This result tells us that the patterns can be designed by considering the medium homogeneous on the refractive index, with small effect on the later application to a medium of inhomogeneous refractive index. These results are promising results as they show that an integrated computational OCT system that uses ultrasound to modulate tissue could be viable. But to assess how much reduction in multiple-scattering can be obtained with this system we need to simulate the interaction of light with tissue that have been modulated by ultrasound. However, there does not exist a simulation software capable of providing information on the optical field within a volume as big as we need to consider. Thus, we decided to develop our own simulator. Since we were interested in obtaining the whole optical field, intensity and phase, we needed to solve the integral equation of scalar scattering. For this task we decided to use Monte Carlo methods.

The second part of this thesis deals with the implementation of a Monte Carlo algorithm to solve the Fredholm integral equations of the second kind. The algorithm is based on the Reversible Jump Markov chain Monte Carlo. The most interesting property of this method

is that it allows to sample from a distribution of variable dimensions. This is of importance as the solution of a Fredholm integral equation of second kind can be expressed as a sum of integrals of increasing dimensionality. We note that with a RJMCMC approach it is not required for us to truncate this sum of integrals at any order, the sampler selects by itself which orders of the sum contribute the most to the solution. The Monte Carlo based integral equation solver was implemented and tested on 1-D integrals. We also present an example of solving a 2-D integral with our implementation of the algorithm, to show that our solver can be extended to higher dimensions. Results of these tests show good accuracy when compared to the known analytical solution of the test problems. However, we haven't still tested the solver on the 3-D integral equation of scattering which is of our interest. To use the solver on the 3-D integral equation of scattering we need to develop an adequate method to lift the singularity in the equation. The method should be adequate for the RJMCMC approach we are using. Also, since it is very rare to have known solutions of integral equations in higher dimensions, that would allow for one to prove the implementation is working, we need to develop a method to measure the quality of the estimation. This method should assess the convergence of the estimate toward the unknown solution, and needs to be developed specifically for our problem of solving the integral equation of scalar scattering.

8.2 Future work

The implementation the Monte Carlo solver on the 3-D integral equation of scalar scattering is probably the most important goal of any future work based on this thesis, as it would open many avenues of research. To apply the solver to 3-D integral equation of scalar scattering requires the development of a method to analyze the convergence of the RJMCMC algorithm to the solution. In chapter 7 the error of the Monte Carlo solver could be obtained by comparing with the exact analytical solution, which is not possible in any practical problem. Many methods have been proposed in the Monte Carlo literature to es-

estimate the variance of MCMC, but there is no method that works for all cases and some only work under specific constraints. Reversible Jump MCMC is particularly complicated because of the possible change in dimension involved. The convergence analysis to be developed should consider specifically the problem of solving the integral equation of scalar scattering, and also trying to adapt some of the methods available in the literature to this particular problem is an obvious option in this case.

After the solver is implemented for the 3-D integral equation of scalar scattering, it could be used to run simulations of light propagation on tissue. One possible avenue of research is to run simulations in tissue modulated by ultrasound patterns. These simulations could allow one to obtain an optimum ultrasound pattern to minimize the backscattered multiple-scattered light. It is highly likely that the optimum ultrasound pattern will be different depending on the features of each tissue and their particular distribution of refractive index. So another possible research involves a classification of tissue types according to their refractive index and how they are affected by different ultrasound patterns. Of course the implemented solver can be used to simulate other cases of light propagation in tissue without any type of ultrasound pattern involved.

Appendix A

MATLAB code for Monte Carlo simulations

Shown below is the MATLAB code for the function that calculates 2-D Fredholm integral equations of the second kind using RJMCMC.

```
%INTEGRAL_EQ2 calculates the solution for a
%Fredholm integral equation of the second kind,
%at a point x0 using Reversible Jump Markov chain
%Monte Carlo.

%(a,b) : interval of evaluation
% x0 : vector (x,y) at which integral is evaluated
% sigma: std deviation for gaussian used in random walk
% g : input function
% k : kernel function

function [estimate b_length_pdf] = ...
    integral_eq2(a,b,c,d,x0,sigma,g,k,seed)
%% Randon walk parameters
```

```

kmax = 100; % maximum length possible for chains
ki = 20; %initial length
max_iteration = 3e5; %number of chains
current_chain = zeros(2,kmax);
%chain from previous iteration
burnin_end = 1e4; %number of samples to drop

%% Set counters for simulation
aupdate = 0; %update moves accepted
rupdate = 0; %update moves rejected
abirth = 0; %birth move accepted
rbirth = 0; %birth move rejected
adeath = 0; %death move accepted
rdeath = 0; %death move rejected
length_pdf = zeros(kmax,1);
%pdf of length of all chains
b_length_pdf = zeros(kmax,1);
%pdf of length of chains after burnin

%% Set the probabilities for each type of move
bi = 1/3 * ones(1,kmax); %vector of birth probabilities
de = 1/3 * ones(1,kmax); %vector of death probabilities
%Special considerations at boundaries of chain length
%Minumum useful chain length is 1 plus x0
bi(1) = .5; de(1) = 0;
bi(kmax) = 0; de(kmax) = .5;

%% Initialization

```

```

estimate = 0;

%% Set independent random streams
% s1: Stream for move selection
% s2: Stream for acceptance / rejection
% s3: Stream for birth candidates
% s4: Stream for update candidate
% s5: Stream to select position to change
[s1,s2,s3, s4, s5,s6,s7] = RandStream.create('mrg32k3a',...
      'NumStreams',7,'Seed',seed);

%% Loop for RJMCMC algorithm
for i = 1:max_iteration
    u = rand(s1,1);
    %-----DEATH MOVE-----%
    if u <= de(ki)
        ku = ceil(ki*rand(s5,1));
        potential_chain = cat(2,current_chain(:,1:ku-1),...
            current_chain(:,ku+1:kmax),[0;0]);
        pi_current = g(current_chain(1,ki),...
            current_chain(2,ki)) * k(x0(1),x0(2),...
            current_chain(1,1), current_chain(2,1));
        %pi_current: value of joint pdf for current chain
        pi_potential = g(potential_chain(1,ki),...
            potential_chain(2,ki)) * k(x0(1),x0(2),...
            potential_chain(1,1), potential_chain(2,1));
        %pi_potential: value of joint pdf of potential chain
        for j = 2:ki

```

```

pi_current = pi_current *...
    k(current_chain(1,j-1),...
      current_chain(2,j-1),current_chain(1,j),...
      current_chain(2,j));
%ki is included in current chain
if j < ki
    pi_potential = pi_potential *...
        k(potential_chain(1,j-1),...
          potential_chain(2,j-1),...
          potential_chain(1,j),...
          potential_chain(2,j));
    %ki excluded in potential chain
end
end
% Calculate the acceptance ratio alpha
ratio = abs(exp(log(pi_potential)-...
log(pi_current)))...
    *(bi(ki-1)/de(ki));
alpha = min(1,ratio);
% Accept the move if random roll < alpha
% otherwise reject
if rand(s2,1) < alpha
    adeath = adeath + 1;
    ki = ki - 1;
    current_chain = potential_chain;
    pi_current = pi_potential;
else
    rdeath = rdeath + 1;

```

```

end
%-----BIRTH MOVE-----%
elseif u <= bi(ki) + de(ki)
    ku = ceil(ki*rand(s5,1));
    xnew = a+(b-a)*rand(s3,1); %new value for x
    ynew = a+(b-a)*rand(s6,1); %new value for y
    potential_chain = cat(2,current_chain(:,1:ku-1),...
        [xnew;ynew],current_chain(:,ku:kmax-1));
    pi_current = g(current_chain(1,ki),...
        current_chain(2,ki)) * k(x0(1),x0(2),...
        current_chain(1,1),current_chain(2,1));
    pi_potential = g(potential_chain(1,ki)...
        ,potential_chain(2,ki)) * k(x0(1),x0(2),...
        potential_chain(1,1),potential_chain(2,1));
    for j =2:ki+1
        if j < ki+1
            pi_current = pi_current * ...
                k(current_chain(1,j-1),...
                current_chain(2,j-1),...
                current_chain(1,j),...
                current_chain(2,j));
        end
        pi_potential = pi_potential * ...
            k(potential_chain(1,j-1),...
            potential_chain(2,j-1),...
            potential_chain(1,j),...
            potential_chain(2,j));
    end
end

```

```

ratio = abs(exp(log(pi_potential)-...
    log(pi_current))) * ...
    (de(ki+1)/bi(ki));
alpha = min(1, ratio);
if rand(s2,1) < alpha
    abirth = abirth + 1;
    ki = ki + 1;
    current_chain = potential_chain;
    pi_current = pi_potential;
else
    rbirth = rbirth + 1;
end
%-----UPDATE MOVE-----%
else %update move
    ku = ceil(ki*rand(s5,1));
    xnew=current_chain(ku) + sigma*randn(s4,1);
    ynew=current_chain(ku) + sigma*randn(s7,1);
    potential_chain = cat(2,current_chain(:,1:ku-1),...
        [xnew;ynew],current_chain(:,ku+1:kmax));
    pi_current = g(current_chain(1,ki),...
        current_chain(2,ki)) * k(x0(1),x0(2),...
        current_chain(1,1),current_chain(2,1));
    pi_potential = g(potential_chain(1,ki),...
        potential_chain(2,ki)) * k(x0(1),x0(2),...
        potential_chain(1,1),potential_chain(2,1));
    for j =2:ki
        pi_current = pi_current *...
            k(current_chain(1,j-1),...

```

```

        current_chain(2, j-1), ...
        current_chain(1, j), ...
        current_chain(2, j));
    pi_potential = pi_potential * ...
        k(potential_chain(1, j-1), ...
        potential_chain(2, j-1), ...
        potential_chain(1, j), ...
        potential_chain(2, j));
end
if ((xnew <= a) || (xnew >= b) && ...
    (ynew <= c) || (ynew >= d))
    alpha = 0;
else
    ratio = abs(exp(log(pi_potential)-...
        log(pi_current)));
    alpha = min(1, ratio);
end
if rand(s2,1) < alpha
    aupdate = aupdate + 1;
    current_chain = potential_chain;
    pi_current = pi_potential;
else
    rupdate = rupdate + 1;
end
end
%-----%
length_pdf(ki) = length_pdf(ki) + 1;
%bin pdf of chain length

```

```

%-----Update estimate-----%
if i > burnin_end
    b_length_pdf(ki) = length_pdf(ki) + 1;
    %bin pdf of chain length after burnin
    estimate = estimate + sign(pi_current);
end
end

%% print counters
disp('Update moves')
fprintf('Accepted: %d\n', aupdate);
fprintf('Rejected: %d\n', rupdate);
disp('Birth moves')
fprintf('Accepted: %d\n', abirth);
fprintf('Rejected: %d\n', rbirth);
disp('Death moves')
fprintf('Accepted: %d\n', adeath);
fprintf('Rejected: %d\n', rdeath);

%% Estimate the result
myfun = @(y, z) abs(g(y, z) .* k(x0(1), x0(2), y, z));
q = quad2d(myfun, a, b, c, d);
norm_c = q / (b_length_pdf(1) / ...
    (max_iteration - burnin_end));

estimate = g(x0(1), x0(2)) + ...
    (norm_c / (max_iteration - burnin_end)) * estimate;
fprintf('Estimate: %d\n', estimate);

```


Bibliography

- [1] J. Fujimoto and W. Drexler, “Introduction to Optical Coherence Tomography,” in *Optical Coherence Tomography Technology and applications*, W. Drexler and J. G. Fujimoto, Eds. Springer, 2008, no. 1, ch. 1, pp. 1–45.
- [2] L. Meng, B. Lv, S. Zhang, and B. Yv, “In vivo optical coherence tomography of experimental thrombosis in a rabbit carotid model.” *Heart (British Cardiac Society)*, vol. 94, no. 6, pp. 777–80, Jun. 2008. [Online]. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2564841&tool=pmcentrez&rendertype=abstract>
- [3] T. Kubo, Y. Ino, T. Tanimoto, H. Kitabata, A. Tanaka, and T. Akasaka, “Optical coherence tomography imaging in acute coronary syndromes.” *Cardiology research and practice*, vol. 2011, p. 312978, Jan. 2011. [Online]. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3177459&tool=pmcentrez&rendertype=abstract>
- [4] J. Fujimoto, M. Brezinski, and G. Tearney, “Optical biopsy and imaging using optical coherence tomography,” *Nature*, 1995. [Online]. Available: http://biophotonics.illinois.edu/pubs/biophotonics_current/opticalbiopsyandimagingusingoet.pdf
- [5] B. Colston, U. Sathyam, L. Dasilva, M. Everett, P. Stroeve, and L. Otis, “Dental OCT.” *Optics express*, vol. 3, no. 6, pp. 230–8, Sep. 1998. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19384365>
- [6] D. Stifter, “Beyond biomedicine: a review of alternative applications and developments for optical coherence tomography,” *Applied Physics B*, vol. 88, no. 3, pp. 337–357, Aug. 2007. [Online]. Available: <http://www.springerlink.com/index/10.1007/s00340-007-2743-2>
- [7] A. Fercher and W. Drexler, “Optical coherence tomography-principles and applications,” *Reports on Progress in Physics*, vol. 66, pp. 239–303, 2003. [Online]. Available: <http://iopscience.iop.org/0034-4885/66/2/204>
- [8] J. De Boer, “Spectral / Fourier Domain Optical,” in *Optical Coherence Tomography*, W. Drexler and J. G. Fujimoto, Eds. Springer, 2008, pp. 147–175.
- [9] J. A. Izatt and M. A. Choma, “Theory of Optical Coherence Tomography,” in *Optical Coherence Tomography Technology and applications*, W. Drexler and J. G. Fujimoto, Eds. Springer, 2008, ch. 2, pp. 47–72. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-77550-8_2http://www.springerlink.com.proxy1.lib.umanitoba.ca/content/k5v88927g0317560/?p=16bb26e795944f698fa473f4a71d1593&pi=1
- [10] B. E. Bouma, G. J. Tearney, B. J. Vakoc, and S. H. Yun, “Optical Frequency Domain Imaging,” in *Optical Coherence Tomography Technology and applications*,

- W. Drexler and J. G. Fujimoto, Eds. Springer, 2008, pp. 209–237. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-77550-8_7<http://www.springerlink.com.proxy1.lib.umanitoba.ca/content/n87ru7440q28jwu1/?p=029c9c366e7e44b6b5f6a15dd5731ca6&pi=6>
- [11] J. G. Fujimoto and M. E. Brezinski, “Optical Coherence Tomography Imaging,” in *Biomedical photonics Handbook*, T. Vo-Dinh, Ed. CRC Press LLC, 2003, ch. 13.
- [12] M. E. Brezinski, G. J. Tearney, S. A. Boppart, E. A. Swanson, J. F. Southern, and J. G. Fujimoto, “Optical Biopsy with Optical Coherence Tomography: Feasibility for Surgical Diagnostics,” *Journal of Surgical Research*, vol. 71, no. 1, pp. 32–40, Jul. 1997. [Online]. Available: <http://www.sciencedirect.com/science/article/B6WM6-45K149J-1N/2/1dc2947a96eb080272c95341ab416914>http://www.sciencedirect.com/science?_ob=ArticleURL&.udi=B6WM6-45K149J-1N&.user=1068138&.coverDate=07/15/1997&.rdoc=1&.fmt=high&.orig=search%&.sort=d&.docanchor=&.view=c&.acct=C000051258&.version=1&.urlVersion=0&.userid=1068138&.md5=658c64df2ad57e0b520b04c4dec672e4
- [13] G. Yao and L. V. Wang, “Monte Carlo simulation of an optical coherence tomography signal in homogeneous turbid media,” *Physics in medicine and biology*, vol. 44, no. 9, pp. 2307–2320, 1999. [Online]. Available: <http://iopscience.iop.org/0031-9155/44/9/316>
- [14] I. T. Lima, A. Kalra, H. E. Hernández-Figueroa, and S. S. Sherif, “Fast calculation of multipath diffusive reflectance in optical coherence tomography.” *Biomedical optics express*, vol. 3, no. 4, pp. 692–700, Apr. 2012. [Online]. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3345799&tool=pmcentrez&rendertype=abstract>
- [15] S. S. Sherif and T. Cathey, “Depth of field control in incoherent hybrid imaging systems,” in *Optical Imaging and Microscopy*, P. Török and F.-J. Kao, Eds. Springer, 2007, ch. 6, pp. 137–167.
- [16] P. a. Edney and J. T. Walsh, “Acoustic modulation and photon-phonon scattering in optical coherence tomography.” *Applied optics*, vol. 40, no. 34, pp. 6381–8, Dec. 2001. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/18364947>
- [17] J. O. Schenk and M. E. Brezinski, “Ultrasound induced improvement in optical coherence tomography (OCT) resolution.” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 15, pp. 9761–4, Jul. 2002. [Online]. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=125006&tool=pmcentrez&rendertype=abstract>
- [18] C. Huang, B. Liu, and M. E. Brezinski, “Ultrasound-enhanced optical coherence tomography: improved penetration and resolution.” *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 25, no. 4, pp. 938–46, Apr. 2008. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/18382493>
- [19] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, 1st ed. John Wiley & sons, 1991.
- [20] P. M. Morse and K. U. Ingard, *Theoretical Acoustics*. New York: McGraw-Hill, 1968.
- [21] J. D. Jackson, *Classical Electrodynamics*. New York: John Wiley & sons, 1962.

- [22] D. Pollock, *Physical Properties of Materials for Engineers*, 2nd ed. New York: CRC Press, 1993.
- [23] K. Vedam and P. Limsuwan, "Piezo- and elasto-optic properties of liquids under high pressure. II. Refractive index vs density," *The Journal of Chemical Physics*, vol. 69, no. 11, p. 4772, 1978. [Online]. Available: http://jcp.aip.org.proxy1.lib.umanitoba.ca/resource/1/jcpsa6/v69/i11/p4772_s1
- [24] S. Sakadžić and L. Wang, "Modulation of multiply scattered coherent light by ultrasonic pulses: An analytical model," *Physical Review E*, vol. 72, no. 3, Sep. 2005. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevE.72.036620>
- [25] W. L. Nyborg, "Safety of medical diagnostic ultrasound." *Seminars in ultrasound, CT, and MR*, vol. 23, no. 5, pp. 377–86, Oct. 2002. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/12509107>
- [26] I. T. Dimov, *Monte Carlo methods for applied scientists*. World Scientific, 2008.
- [27] W. Dunn and J. K. Shultis, *Exploring Monte Carlo Methods*. Academic Press, 2011.
- [28] F. James, "Monte Carlo theory and practice," *Reports on Progress in Physics*, vol. 43, no. February, 1980. [Online]. Available: <http://iopscience.iop.org/0034-4885/43/9/002>
- [29] H. Anderson, "Metropolis, Monte Carlo, and the MANIAC," *Los Alamos Science*, pp. 96–108, 1986. [Online]. Available: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Metropolis,+Monte+Carlo,+and+the+MANIAC#0>
- [30] M. H. Kalos and P. A. Whitlock, *Monte Carlo methods*, 2nd ed. WILEY-VCH Verlag GmbH & Co. KGaA, 2004. [Online]. Available: <http://onlinelibrary.wiley.com/book/10.1002/9783527617395>
- [31] "Creating and Controlling a Random Number Stream :: Random Numbers (MATLAB)." [Online]. Available: <http://www.mathworks.com/help/techdoc/math/brn4ixh.html>
- [32] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo method*, 2nd ed. Hoboken: John Wiley & sons, 2008.
- [33] C. J. Geyer, "Introduction to Markov Chain Monte Carlo," in *Handbook of Markov chain Monte Carlo*, S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng, Eds. Chapman & Hall/CRC, 2010, ch. 1, p. 619.
- [34] S. Chib and E. Greenberg, "Understanding the metropolis-hastings algorithm," *American Statistician*, pp. 327–335, 1995. [Online]. Available: <http://www.jstor.org/stable/2684568>
- [35] G. Casella and C. P. Robert, *Introducing Monte Carlo methods with R*. Springer, 2010.
- [36] C. Geyer, "Practical markov chain monte carlo," *Statistical Science*, vol. 7, no. 4, pp. 473–483, 1992. [Online]. Available: <http://projecteuclid.org/euclid.ss/1177011137>
- [37] Y. Fan and S. A. Sisson, "Reversible Jump MCMC," in *Handbook of Markov chain Monte Carlo*, S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng, Eds. Chapman & Hall/CRC, 2010, ch. 3, p. 619.

- [38] P. J. Green, "Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination," *Biometrika*, vol. 82, no. 4, p. 711, Dec. 1995. [Online]. Available: <http://www.jstor.org/stable/2337340?origin=crossref>
- [39] S. A. Sisson, "Trans-dimensional Markov chains : A decade of progress and future perspectives," University of New South Wales, Tech. Rep., 2004.
- [40] A. Doucet, A. M. Johansen, and V. B. Tadić, "On solving integral equations using Markov chain Monte Carlo methods," *Applied Mathematics and Computation*, vol. 216, no. 10, pp. 2869–2880, Jul. 2010. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0096300310004042>
- [41] J. A. Cochran, *The analysis of linear integral equations*. McGraw-Hill, 1972.
- [42] R. Kress, *Linear Integral Equations*, 2nd ed. New York: Springer, 1999.
- [43] R. Farnoosh and M. Ebrahimi, "Monte Carlo method for solving Fredholm integral equations of the second kind," *Applied Mathematics and Computation*, vol. 195, no. 1, pp. 309–315, Jan. 2008. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0096300307005541>
- [44] R. Farnoosh and E. Morteza, "Monte Carlo simulation for solving Fredholm integral equations," *Kybernetes*, vol. 38, no. 9, pp. 1621–1629, 2009. [Online]. Available: <http://www.emeraldinsight.com/10.1108/03684920910991577>
- [45] A. D. Polyanin and A. V. Manzhirov, *Handbook of integral equations*, 2nd ed. Chapman & Hall/CRC, 2008.
- [46] F. Mirzaee and S. Piroozfar, "Numerical solution of the linear two-dimensional Fredholm integral equations of the second kind via two-dimensional triangular orthogonal functions," *Journal of King Saud University - Science*, vol. 22, no. 4, pp. 185–193, Oct. 2010. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S1018364710000364>
- [47] L. Tierney, "Markov chains for exploring posterior distributions," *The Annals of Statistics*, vol. 22, no. 4, pp. 1701–1728, 1994. [Online]. Available: <http://www.jstor.org/stable/10.2307/2242477>

Subject Index

Bayesian model, 50

Classius-Mossotti, 19, 20, 26, 28

COMSOL, 30

Diabetes, 8

Glaucoma, 8

Hybrid imaging, 12

Importance density, 44
 optimal, 45, 57

Importance sampling, 43
 Sequential, 45

Integral equation
 Fredholm
 first kind, 53
 second kind, 53
 linear, 53

Integral operator, 53

Markov chain Monte Carlo, 46
 convergence, 48
 Reversible Jump, 49, 57

MATLAB, 41

Monte Carlo method, 38

Optical Coherence Tomography, 1
 applications, 2, 8
 Non-destructive testing, 9
 Fourier domain, 2, 5
 Time domain, 2, 3

Polarization, 21
 atomic, 21
 molecular, 21

Random numbers, 39
 pseudo, 40

Random walk, 54

Safety
 Ultrasound, 32

Scattering, 1

variance reduction, 43