

Portland State University

PDXScholar

Dissertations and Theses

Dissertations and Theses

1-1-2010

Model-Based Material Parameter Estimation for Terahertz Reflection Spectroscopy

Gabriel Paul Kniffin

Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Let us know how access to this document benefits you.

Recommended Citation

Kniffin, Gabriel Paul, "Model-Based Material Parameter Estimation for Terahertz Reflection Spectroscopy" (2010). *Dissertations and Theses*. Paper 241.

<https://doi.org/10.15760/etd.241>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Model-Based Material Parameter Estimation for Terahertz Reflection Spectroscopy

by

Gabriel Paul Kniffin

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Electrical and Computer Engineering

Thesis Committee:
Lisa M. Zurk, Chair
Donald Duncan
Andres H. La Rosa

Portland State University
©2011

Abstract

Many materials such as drugs and explosives have characteristic spectral signatures in the terahertz (THz) band. These unique signatures imply great promise for spectral detection and classification using THz radiation. While such spectral features are most easily observed in transmission, real-life imaging systems will need to identify materials of interest from reflection measurements, often in non-ideal geometries. One important, yet commonly overlooked source of signal corruption is the etalon effect – interference phenomena caused by multiple reflections from dielectric layers of packaging and clothing likely to be concealing materials of interest in real-life scenarios.

This thesis focuses on the development and implementation of a model-based material parameter estimation technique, primarily for use in reflection spectroscopy, that takes the influence of the etalon effect into account. The technique is adapted from techniques developed for transmission spectroscopy of thin samples and is demonstrated using measured data taken at the Northwest Electromagnetic Research Laboratory (NEAR-Lab) at Portland State University. Further tests are conducted, demonstrating the technique’s robustness against measurement noise and common sources of error.

Acknowledgments

I would like to thank my advisor Dr. Lisa M. Zurk for her guidance and support during this research. I am also grateful to all my fellow students at the Northwest Electromagnetics and Acoustics Research Laboratory at Portland State University for their help and support. Additional thanks are due to Dr. Alla Timchenko, Dr. Jian Chen, and Scott Schecklman for their assistance, as well as to Dr. Don Duncan and Dr. Andres La Rosa for serving on my committee, and to Dr. John Wilkinson for providing samples measured for this thesis. Finally, I would like to acknowledge the National Science Foundation for funding this work.

Contents

Abstract	i
Acknowledgments	ii
List of Tables	v
List of Figures	vii
1 Introduction	1
1.1 Terahertz Spectroscopy	3
1.2 Thesis Work	6
1.3 Contributions	7
2 Theory	9
2.1 Material Properties	9
2.1.1 Classical Dispersion: The Lorentz Oscillator Model	12
2.2 Waves in Layered Media	13
3 Terahertz Time Domain Spectroscopy	17
3.1 NEAR-Lab Terahertz Time Domain Spectroscopy System	18

3.2	Material Parameter Estimation: Non-Parametric Techniques	21
3.2.1	Transmission Mode	22
3.2.2	Reflection Mode	33
3.3	Material Parameter Estimation: Parametric Technique	39
4	Results and Discussion	41
4.1	Comparison of Non-Parametric and Parametric Techniques in Trans- mission	42
4.2	Comparison of Parametric Technique in Transmission and Reflection	52
4.3	Simulations for Sensitivity Analysis	59
4.3.1	Initial Guesses for Numerical Optimization Algorithms	60
4.3.2	Effect of System Noise	67
4.3.3	Effect of Thickness Error	72
4.3.4	Effect of Positioning Error	74
4.4	Discussion	76
5	Conclusions	77
5.1	Future Work	80
	References	84
	Appendix A Derivation of Lorentz Dispersion Model	90
	Appendix B Nelder-Mead Simplex Algorithm	94

List of Tables

4.1	Initial Guess for Lorentz model parameters used in initializing the Nelder-Mead algorithm to solve (4.1).	47
4.2	Lorentz model parameters in $\boldsymbol{\theta}^{fit}$ resulting from solving (4.1) to fit \tilde{n}_1^{np} to the Lorentz model as shown in Step 2 of Figure 4.5.	47
4.3	Lorentz model parameters in $\boldsymbol{\theta}^p$ resulting from minimizing S_t using the parametric method as shown in Step 3 of Figure 4.7.	49
4.4	Quantitative comparison of \tilde{n}_1^{fit} and \tilde{n}_1^p to \tilde{n}_1^{np} in terms of squared error norm S_n from (4.1).	50
4.5	Difference in Lorentz parameters between those in $\boldsymbol{\theta}^{fit}$ and those in $\boldsymbol{\theta}^p$. Absolute differences are given for $\omega_p/2\pi$ and $\gamma_p/2\pi$ while percent differences are given for ϵ_∞ and $\Delta\epsilon_p$	51
4.6	Lorentz model parameters in $\boldsymbol{\theta}^r$ resulting from minimizing S_r using the parametric method applied to reflection mode as shown in Step 4 of Figure 4.13.	56
4.7	Difference in Lorentz parameters between those in $\boldsymbol{\theta}^r$ and those in $\boldsymbol{\theta}^p$. Absolute differences are given for $\omega_p/2\pi$ and $\gamma_p/2\pi$ while percent differences are given for ϵ_∞ and $\Delta\epsilon_p$	58

4.8	Percent difference between ω_p values in $\boldsymbol{\theta}^r$ and $\boldsymbol{\theta}^p$	59
4.9	Quantitative comparison of \tilde{n}_1^p and \tilde{n}_1^r to \tilde{n}_1^{np} in terms of S_n from (4.1).	59
4.10	Lorentz model parameters for hypothetical test material with a single resonance used in evaluation of numerical optimization techniques.	60
4.11	Lorentz model parameters for explosive composition 4 (C4) from Ya- mamoto [16].	67

List of Figures

1.1	The “Terahertz Gap” shown within the greater electromagnetic spectrum.	1
1.2	Terahertz science and engineering journal articles published by year. Data taken from the Compendex search engine.	2
1.3	Absorption spectra of four common explosives illustrate the potential of THz as a means of fingerprinting materials [3].	3
2.1	Wave vector $\bar{k} = k_x\hat{x} + k_y\hat{y} + k_z\hat{z}$	14
2.2	Geometry for reflection and transmission of incident plane wave with amplitude E_i from a layer of thickness d with refractive index \tilde{n}_1 against a semi-infinite half space of background material with refractive index \tilde{n}_2	15
3.1	(a) A THz Pulse measured with the Picometrix T-Ray 4000 at the NEAR-Lab. (b) Spectrum of THz pulse in obtained via fast Fourier transform.	19

3.2	(a) T-Ray 4000 time domain spectroscopy system shown configured for reflection measurements with collinear head. (b) Closeup of reflection sample stage. (c) Closeup of transmit and receive heads configured for transmission measurement in purge chamber.	20
3.3	Typical transmission measurement configuration: (a) Sample measurement. (b) Reference measurement required for deconvolution.	23
3.4	Transmission measurement of a sample of polyethylene showing multiple pulses due to reverberation within the sample as shown in Figure 3.3 and described in Section 2.2.	24
3.5	(a) THz pulse transmitted through a lactose sample. Blue: 1000 point (≈ 78 ps) ‘Long’ window, Green: 500 point (≈ 39 ps) ‘Medium’ window, Red: 250 point (≈ 20 ps) ‘Short’ window. (b) Shortening the time window reduces frequency resolution, smoothing the spectral features in resulting FFT spectrum.	27
3.6	Squared transmission mode objective function (3.11) plotted as a function of refractive index, n_1 , and extinction coefficient, κ_1 , at four frequencies. In (3.11), $E_{data,j}$ was calculated from (3.9) using $d = 3.4$ mm and $\tilde{n}_1 = 1.5 - i0.001$ for all frequencies.	30
3.7	Ideal reflection measurement setup.	34
3.8	Positioning error ΔL between sample and reference in reflection mode introduces linear phase shift expressed in (3.19).	36

3.9	Squared reflection mode objective function (3.21) as a function of refractive index, n_1 , and extinction coefficient, κ_1 , at four frequencies. In (3.21), $E_{data,j}$ was calculated from (3.18) using $d = 3.4$ mm and $\tilde{n}_1 = 1.5 - i0.001$ for all frequencies.	38
4.1	Lactose transmission data (a) in the time domain and (b) after FFT. Second reflected pulse visible at ≈ 132 ps in the time domain. . . .	43
4.2	Deconvolved lactose (a) transmission spectrum and (b) phase. . . .	44
4.3	Total variation V of estimated real refractive index and extinction coefficient as a function of assumed sample thickness.	45
4.4	Estimated (a) refractive index n_1^{np} and (b) extinction coefficient κ_1^{np} using non-parametric methods described in Section 3.2.1 and [10]. . .	45
4.5	Flowchart showing the estimation of the complex refractive index \tilde{n}_1^{np} from deconvolved data E_{data} using the conventional non-parametric method of Dorney as Step 1. Step 2 consists of fitting of the Lorentz model (2.7) to \tilde{n}_1^{np} , which yields the vector $\boldsymbol{\theta}^{fit}$ of Lorentz parameters. Using $\boldsymbol{\theta}^{fit}$ in the Lorentz model yields \tilde{n}_1^{fit}	46
4.6	(a) Refractive index n_1^{np} estimated using non-parametric method of Dorney [10], and refractive index n_1^{fit} calculated using Lorentz parameters in $\boldsymbol{\theta}^{fit}$ (shown in Table 4.2) in the Lorentz model (2.7), and (b) corresponding extinction coefficients, κ_1^{np} and κ_1^{fit}	48

4.7	Flowchart showing the parametric method of Section 3.3 in Step 3. Fitting the measured data E_{data} by minimizing S_t in (3.10) using tER_j^p from (3.23) yields the Lorentz parameter vector θ^p . Using θ^p as the input to the Lorentz model yields \tilde{n}_1^p	48
4.8	Results of parametric fitting of transmission data by minimizing the squared error norm S_t as described in Section 3.3: (a) transmittance spectrum and (b) phase.	49
4.9	(a) Refractive index and (b) extinction coefficient estimated using para- metric method described in Section 3.3 applied in transmission mode (\tilde{n}_1^p) compared to results of non-parametric method of Dorney [10] (\tilde{n}_1^{np}) and fitting of non-parametric result with Lorentz model (\tilde{n}_1^{fit}).	50
4.10	Measured lactose reflection data (a) in the time domain and (b) after FFT.	53
4.11	Deconvolved lactose (a) reflection spectrum and (b) phase.	53
4.12	(a) Linear fit of unwrapped phase. (b) Result of phase correction using slope of fit line.	56
4.13	Flowchart showing the parametric method of Section 3.3 applied to reflection mode in Step 4. Steps 1-3 are also shown for reference. Fitting the measured data E_{data} by minimizing S_r in (3.20) using rER_j^p from (3.24) yields the Lorentz parameter vector θ^r . Using θ^r as the input to the Lorentz model yields \tilde{n}_1^r	57

4.14	Results of parametric fitting of reflection data by minimizing the squared error norm S_r as described in Section 3.3: (a) reflection spectrum and (b) phase.	57
4.15	(a) Refractive index and (b) extinction coefficient estimated using parametric method in transmission \tilde{n}_1^p and reflection \tilde{n}_1^r	58
4.16	(a) refractive index and (b) extinction coefficient corresponding to Lorentz parameters in Table 4.10.	61
4.17	Calculated (a) reflectance and (b) unwrapped phase using assuming $d = 1.5$ mm in (3.18) and the refractive index and extinction coefficient values shown in Figure 4.16.	62
4.18	Effect of perturbing initial guess of ϵ_∞ on Nelder-Mead and Active-Set algorithms on (a) iterations required for convergence and (b) squared difference norm S_n between complex refractive index values at convergence and correct values shown in Figure 4.16. While less efficient, the Nelder-Mead algorithm converges consistently to the correct value (global minimum) for perturbations ranging from -6% to 18% of the correct value, while the more efficient active-set method converges to the correct value for perturbations within only $\pm 3\%$ of the correct value. Outside these ranges, both algorithms are thrown off by local minima.	63

4.19	Effect of perturbing initial guess of $\Delta\epsilon$ on Nelder-Mead and Active-Set algorithms on (a) iterations required for convergence and (b) squared difference norm S_n between complex refractive index values at convergence and correct values shown in Figure 4.16. The more efficient active-set algorithm converges consistently to the correct value (global minimum) for perturbations ranging from -45% to 39% of the correct value, while the less efficient Nelder-Mead method converges to the correct value for perturbations from -22% to 17% of the correct value. Outside these ranges, both algorithms are thown off by local minima.	64
4.20	Effect of perturbing initial guess of γ on Nelder-Mead and Active-Set algorithms on (a) iterations required for convergence and (b) squared difference norm S_n between complex refractive index values at convergence and correct values shown in Figure 4.16. The algorithms converge to comparable solutions when initial guesses are perturbed from the correct value.	65

4.21	Effect of perturbing initial guess of ω on Nelder-Mead and Active-Set algorithms on (a) iterations required for convergence and (b) squared difference norm S_n between complex refractive index values at convergence and correct values shown in Figure 4.16. While less efficient, the Nelder-Mead algorithm converges consistently to the correct value (global minimum) for perturbations ranging from -18% to 15% of the correct value with a single exception at -3%, while the more efficient active-set method converges to the correct value for perturbations between -11% and 23% of the correct value with two exceptions at 18% and 22%. Outside these ranges, both algorithms are consistently thrown off by local minima.	66
4.22	Refractive index (a) and extinction coefficient (b) corresponding to Lorentz parameters for C4 shown in Table 4.11 [16].	68
4.23	Calculated (a) reflectance and (b) unwrapped phase using assuming $d = 1.5$ mm in (2.15) and the refractive index and extinction coefficient values shown in Figure 4.22.	69
4.24	Effect of (a) 20 dB, (b) 40 dB, and (c) 60 dB SNR on mean μ_r and confidence interval $\mu_r \pm \sigma_r$ of reflectance based on an ensemble of 1000 noise realizations, and (d) the case of infinite SNR.	70

4.25	Effect of SNR on variance of estimated C4 Lorentz parameters as a percentage of the ensemble mean for all six molecular modes (see Table 4.11). (a) ϵ_∞ (b) $\Delta\epsilon_p$ (c) γ_p (d) ω_p	71
4.26	Results of solving (4.6) with Nelder-Mead algorithm starting from initial thickness guesses ranging from -15% to 10% of correct value. (a) Percent error between actual thickness and estimated thickness at convergence. (b) Squared difference norm S_n between correct complex refractive index and complex refractive index calculated from estimated θ' at convergence.	73
4.27	Calculated (a) reflectance and (b) unwrapped phase of simulated data from (3.19) in red, shown with unaltered data from Figure 4.23 in blue for comparison.	75
4.28	Results of solving (4.8) with Nelder-Mead algorithm starting from initial guesses for ΔL ranging from -200% to 350% of correct value. (a) Percent error between estimated thickness and actual thickness at convergence. (b) Squared difference norm S_n between actual complex refractive index and complex refractive index calculated from θ'' at convergence.	75

5.1	Time domain reflection waveforms of 10% lactose sample with no tape compared with waveforms from the same sample covered with a single layer of tape and two layers of tape. Waveforms are offset vertically by 0.1 a.u. for clarity.	82
5.2	Reflectance spectra of 10% lactose sample with no tape compared to spectra from the same sample covered with a single layer of tape and two layers of tape. Reflections from the back surface of the sample shown in Figure 5.1 were excluded from the Fourier transform windows.	83

Chapter 1

Introduction

The “Terahertz Gap” is a relatively unexplored band of the electromagnetic spectrum. Shown in Figure 1.1, it is commonly defined as residing between 0.3×10^{12} and 3.0×10^{12} Hz [1]. Historically, the development of sufficiently powerful terahertz

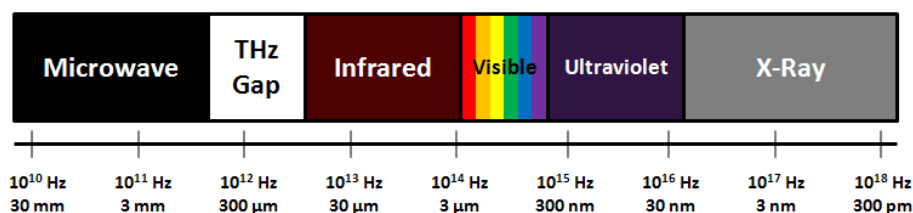


Figure 1.1: The “Terahertz Gap” shown within the greater electromagnetic spectrum.

(THz) sources and sufficiently sensitive THz receivers has lagged behind that of the neighboring microwave and infrared bands. This lack of instrumentation has left the region relatively underutilized. However, recent advances in both microwave/RF and optical technologies have begun to fill in the gap, unlocking a host of new and potentially revolutionary technologies and applications in a wide variety of scientific and engineering disciplines.

Terahertz is a growing area of research in modern electromagnetics, as evident in Figure 1.2, which shows the increasing number of published articles on the subject since 1984. Part of the appeal is due to the combination of properties exhibited by THz radiation. Like in the infrared band, many materials have unique characteristic absorption spectra in the THz band, facilitating the spectroscopic “fingerprinting” of

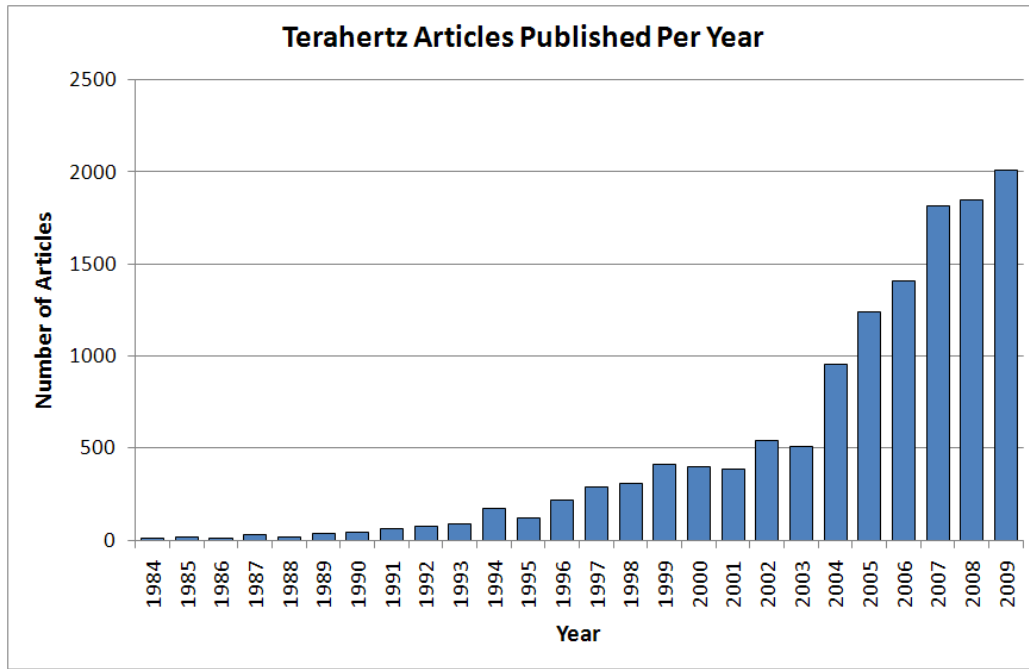


Figure 1.2: Terahertz science and engineering journal articles published by year. Data taken from the Compendex search engine.

compounds such as drugs and explosives [2]. Examples of such absorption features present in military-grade explosive compounds are shown in Figure 1.3 [3]. In addition, non-polar materials such as clothing, paper, and plastic are transparent to THz, just as they are to microwaves and millimeter waves. While many technical challenges still remain – such as mitigating the influence of atmospheric water vapor absorption lines on measured spectra – the aforementioned combination of properties, combined with the fact that THz waves are non-ionizing, makes the technology well suited for screening mail or luggage and passengers in airports.

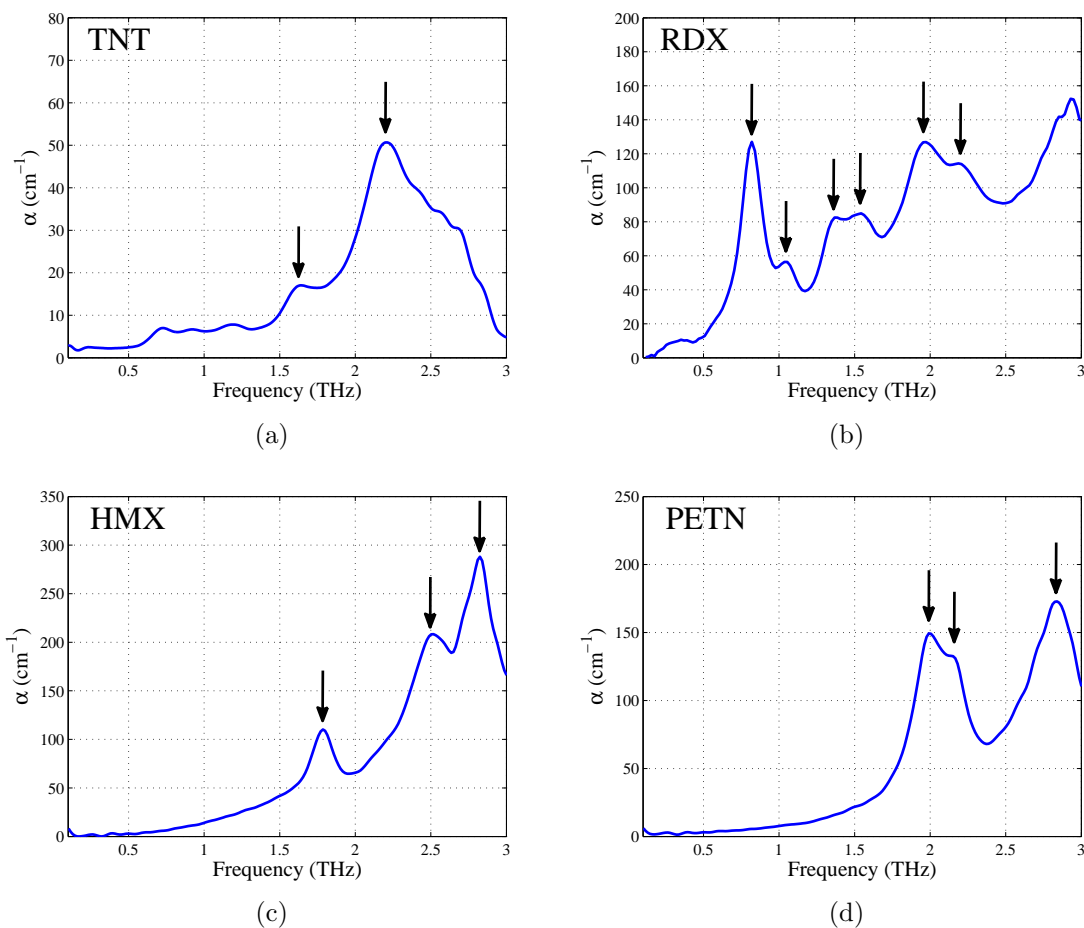


Figure 1.3: Absorption spectra of four common explosives illustrate the potential of THz as a means of fingerprinting materials [3].

1.1 Terahertz Spectroscopy

THz spectral features are detectable in both transmission mode, in which THz radiation propagates through a target sample, and reflection mode, in which THz radiation is reflected off of a target. These absorption features are primarily determined using transmission measurements, usually with carefully prepared samples in controlled laboratory settings. To date, many absorption spectra, including those shown in Figure 1.3 have been determined this way and subsequently published. However, most

real-world targets are opaque to THz waves, necessitating the use of reflection measurements for standoff detection. Most work on identifying materials based on their reflection spectra has been carried out using terahertz pulsed imaging (TPI) [4, 5, 6]. In TPI, a target is probed using broadband THz pulses in a configuration similar to pulse-echo ultrasound. For a multiply-layered target with near-parallel boundaries, such as a drug in tablet form or an explosive concealed in the sole of a shoe or beneath layers of clothing or packaging, reflections are caused by refractive index discontinuities at the boundaries between materials. These reflected echo pulses are recorded coherently in the time domain – their times of flight proportional to the optical thicknesses of each layer of material – thereby yielding information on the 3D structure of the target [7]. The echoes also contain spectral information, which can be used to estimate and classify the chemical compositions of the various layers using both spectral magnitude and phase information from the pulses [7].

To date, most of the work in material identification via reflection TPI assumes the layers of material are optically thick enough that reflected pulses do not overlap in time and can therefore be treated separately. If any of the layers are optically thin, the resulting pulses may not be separable or may be so close together that narrow time windows that limit spectral resolution are required to isolate them. When such multiple pulses are included in a time window, interference patterns in the observed spectra result due to the Fabry-Pérot or etalon effect. In the field of THz spectroscopy, the etalon effect has been treated primarily in the literature on material

parameter estimation from transmission measurements of thin samples [8, 9, 10, 11]. In such a scenario, the thin sample itself acts as the layer; the first transmitted pulse is followed in the time domain by multiple pulses from reflections within it. The basic method outlined in those works involved constructing a theoretical model for the transmission measurement that took these multiple pulses into account, and then solving the inverse problem using a numerical algorithm to fit the measured data to the model. This is typically done in a non-parametric fashion in which the number of complex refractive index data points is equal to the number of measured frequency domain data points. Problems arise using this method due to the multimodal solution space caused by ambiguity in the phase of the measured and modeled data [8]. Dorney [10] solved this problem by unwrapping the phase of both the measured and modeled data in the estimation algorithm. Unfortunately, applying this same technique in reflection is not straightforward, as the phase of a properly-aligned reflection measurement does not change linearly in frequency as does the phase from a transmission measurement. The unwrapping step therefore has no effect on the phase of a reflection measurement. The relationship between a reflection measurement's magnitude and phase and the real and imaginary parts of the sample's complex refractive index is also more complicated than it is the case of a transmission measurement [12]. Applying the numerical inversion technique used by Pupeza [11] to the case of reflection measurements typically results in discontinuities and other non-physical artifacts in the estimated material parameter curves due to the multimodal solution

space, especially when combined with noise and error in nuisance parameters such as sample thickness [13]. These difficulties associated with non-parametric model-based techniques suggest a parametric model-based technique which incorporates *a priori* information – the assumption that the complex refractive index behaves consistently with a dispersion model – may be preferable.

Such a parametric technique was recently developed by Ahmed [14], who parameterized the complex refractive index of various sample materials using a variety of dispersion models. The sample thickness was also parameterized, allowing its efficient estimation along with the dispersion model parameters. However, Ahmed’s formulation was only for transmission mode, and as previously mentioned, reflection mode is more of interest for standoff detection. Ahmed also only treated materials with single resonances, not materials with several distinct resonances such as the explosives listed in Figure 1.3.

1.2 Thesis Work

The focus of this thesis was the development of a model-based approach to material parameter estimation from layered materials, primarily for use in reflection mode. In this method, the complex refractive index is parameterized using the Lorentz dispersion model, allowing the absorption fingerprints to be described by a relatively small number of parameters that specify the number of absorption peaks, their individual

strengths, spectral locations, and spectral widths. The benefits of this parameterization are threefold: It simplifies the inversion process, increasing efficiency and robustness against the influence of measurement noise; allows the simultaneous estimation of sample thickness and displacement error between sample and reference mirror in reflection; and provides a concise description of a material's absorption features for the purpose of material classification. As reflection mode is more important for practical applications, it is emphasized over transmission mode in this thesis.

Findings from this research were presented at The International Society for Optics and Photonics (SPIE) Defense, Security, and Sensing conference in Orlando, FL on April 8, 2010 and published in Proceedings of SPIE - The International Society for Optical Engineering [13].

1.3 Contributions

- Developed MATLAB models of THz wave interactions with stratified media and a non-parametric material parameter estimation routine based on these models and techniques from literature.
- Developed parametric material parameter inversion method based on Lorentz dispersion model and implemented it in MATLAB.
- Validated parametric technique in transmission mode with results of conventional non-parametric technique based on method from the literature using measured THz transmission data.

- Tested parametric technique for consistency between results attained from transmission mode and reflection mode data.
- Quantified parametric technique's sensitivity to initial parameter estimates, measurement noise, and ability to estimate nuisance parameters including sample thickness and displacement error between sample and reference.

Chapter 2

Theory

As discussed previously, materials of interest such as drugs and explosives have characteristic absorption spectra in the THz band. These spectra arise from the excitation of molecular vibrational modes. This section discusses the parameters used to describe wave interactions with dielectric materials and introduces the Lorentz model – a classical model of the fluctuations in these parameters due to such vibrational modes. Mathematical descriptions of wave interactions with layered materials are then presented.

2.1 Material Properties

Wave propagation in a source-free region is described by the homogeneous wave equation,

$$\nabla^2 \mathcal{E} - \frac{1}{v_p^2} \frac{\partial^2}{\partial t^2} \mathcal{E} = 0, \quad (2.1)$$

where

$$v_p = \frac{1}{\sqrt{\mu\epsilon}}$$

is the phase velocity of the wave in the medium and, following the convention of Balanis [15], \mathcal{E} is the time-varying electric field vector. In time harmonic form, in which

$$\mathcal{E}(x, y, z; t) = \Re [\mathbf{E}(x, y, z) e^{i\omega t}],$$

the wave equation (2.1) simplifies to

$$\nabla^2 \mathbf{E} + k^2 \mathbf{E} = 0, \quad (2.2)$$

where the wavenumber

$$k = \frac{\omega}{v_p} = \omega \sqrt{\mu \epsilon}$$

encapsulates the characteristics of the medium in terms of its reaction to an oscillating electromagnetic wave of angular frequency ω . The magnetic permeability, μ , is the degree to which a material becomes magnetized in reaction to an applied magnetic field. For most dielectric materials, the permeability is approximately equal to that of free space, $\mu_0 = 4\pi \times 10^{-7}$ H/m. The electric permittivity ϵ describes a material's polarizability in response to an electric field. In general, the permittivity is frequency dependent and complex,

$$\epsilon = \epsilon' - i\epsilon'',$$

with absorption in the material described by the imaginary part of the permittivity, ϵ'' .

The dielectric constant or relative permittivity, ϵ_r , is defined as the permittivity of a substance normalized by the permittivity of free space, $\epsilon_0 = 8.85418782 \times 10^{-12}$ F/m,

$$\epsilon_r = \frac{\epsilon}{\epsilon_0}.$$

A material's properties can also be described in terms of the refractive index \tilde{n} which offers a more physically intuitive description of the way a material slows and attenuates a wave propagating through it. The refractive index is related to the permeability and permittivity by

$$\tilde{n} = \sqrt{\frac{\mu}{\mu_0} \frac{\epsilon}{\epsilon_0}},$$

which for most non-magnetic dielectrics simplifies to

$$\tilde{n} = \sqrt{\epsilon_r}. \quad (2.3)$$

The refractive index is therefore also complex and frequency dependent, and is often expressed as

$$\tilde{n} = n - i\kappa, \quad (2.4)$$

where the real part of the refractive index,

$$n = \frac{c}{v_p}, \quad (2.5)$$

describes how the phase velocity v_p of a wave of a given frequency is reduced relative to the speed of light in vacuum, c . The imaginary part of the refractive index, κ , is referred to as the extinction coefficient, as it describes the degree to which a wave is attenuated as it propagates through a medium. The absorption coefficient α is the power absorption per unit path length at a given frequency f . It is related to the

extinction coefficient by

$$\alpha = \frac{4\pi f}{c} \kappa. \quad (2.6)$$

2.1.1 Classical Dispersion: The Lorentz Oscillator Model

The Lorentz oscillator model of classical dispersion theory is often used to describe a material's dielectric constant in the terahertz regime [16, 17, 18]. In this model, a medium is described as being composed of atoms that act as atomic dipole (Lorentz) oscillators. The material's spectral absorption lines are then described as resonant modes of these oscillators [19, 12, 15]. A derivation of the Lorentz model is given in Appendix A. The resulting expression,

$$\epsilon_r(\omega) = \epsilon_\infty + \sum_{p=1}^P \frac{\Delta\epsilon_p \omega_p^2}{\omega_p^2 - \omega^2 - i\gamma_p \omega}, \quad (2.7)$$

is a description of the frequency (ω) dependent dielectric constant of a material with P molecular resonant modes in terms of $3P + 1$ parameters:

- ϵ_∞ , the dielectric constant in the high frequency limit which sets a baseline for the real part of the dielectric constant across frequency,
- ω_p , the angular frequency of the p -th molecular resonant mode in radians per second,
- $\Delta\epsilon_p$, the change in the relative permittivity due to the p -th molecular resonant mode, equivalent to the strength of the mode, and

- γ_p , the p -th damping coefficient which determines the full width at half maximum of the p -th resonant mode.

2.2 Waves in Layered Media

The basic principle of modeling the propagation of waves through layered media involves applying appropriate boundary conditions to solutions of the electromagnetic wave equation (2.2). A convenient solution for this analysis is that for the case of rectangular symmetry – namely, the plane wave. The general case of such a wave \mathbf{E} propagating in an unbounded medium can be expressed as

$$\mathbf{E} = \hat{a} E e^{-i\bar{k} \cdot \bar{r}}, \quad (2.8)$$

where \hat{a} is the unit polarization vector, E is the amplitude, $\bar{k} = k_x \hat{x} + k_y \hat{y} + k_z \hat{z}$ is the propagation vector shown in Figure 2.1, and $\bar{r} = x \hat{x} + y \hat{y} + z \hat{z}$ is the position vector in 3-dimensional space.

The specific case of a plane wave \mathbf{E}_i with amplitude E_i propagating through a semi-infinite half space of air with refractive index $\tilde{n}_0 = 1$ (medium 0) before impinging at normal incidence on a layer of thickness d with refractive index \tilde{n}_1 (medium 1) is depicted in Figure 2.2. Behind the layer lies a semi-infinite half space of background medium with refractive index \tilde{n}_2 (medium 2). All interfaces between layers are assumed to be planar and parallel and all media are assumed to be homogeneous, isotropic, and nonmagnetic ($\mu = \mu_0$).

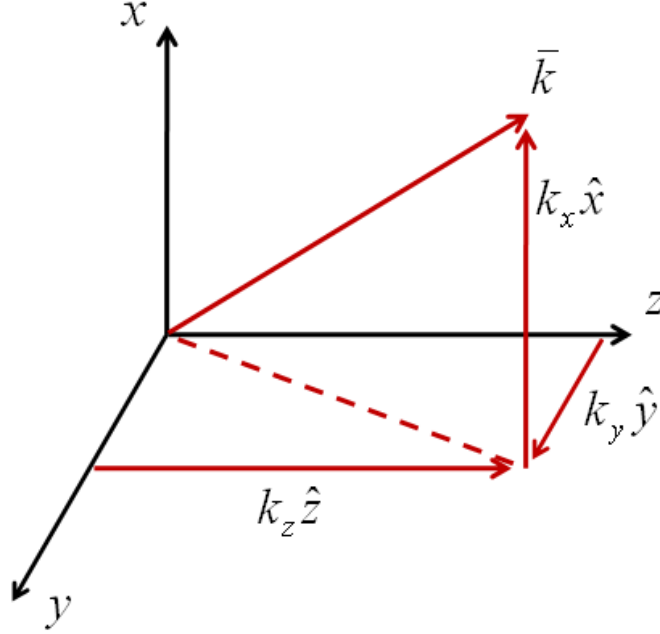


Figure 2.1: Wave vector $\bar{k} = k_x \hat{x} + k_y \hat{y} + k_z \hat{z}$.

Upon encountering the boundary at normal incidence, a portion of E_i is reflected while the remainder is transmitted. The transmitted wave then encounters the back surface of the sample, whereupon the wave is again split into a transmitted and reflected portion. The reflected portion from the back surface then encounters the sample-air boundary and is again split. The resulting reflected field E_r from successive reflections can be expressed as

$$\begin{aligned}
 E_r &= r_{01}E_i + t_{01}t_{10}r_{12}e^{-i2k_1d}E_i + t_{01}t_{10}r_{10}r_{12}^2e^{-i4k_1d}E_i + \dots \\
 &= r_{01}E_i + t_{01}t_{10}r_{12}e^{-i2k_1d}E_i \sum_{q=0}^Q (r_{10}r_{12}e^{-i2k_1d})^q, \tag{2.9}
 \end{aligned}$$

where $k_1 = \frac{2\pi f}{c}\tilde{n}_1$ is the wavenumber in medium 1, Q is the number of reflections, and f and c are the frequency and speed of light in vacuum, respectively. The

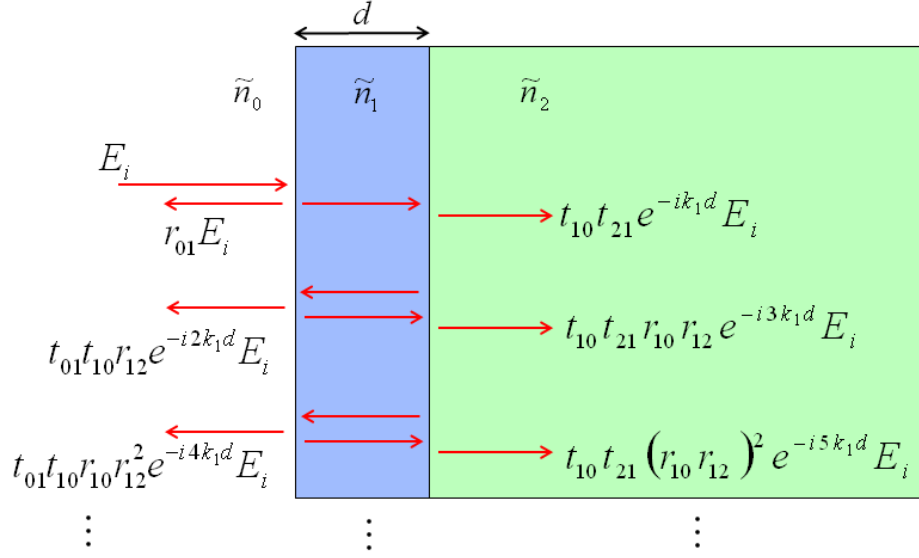


Figure 2.2: Geometry for reflection and transmission of incident plane wave with amplitude E_i from a layer of thickness d with refractive index \tilde{n}_1 against a semi-infinite half space of background material with refractive index \tilde{n}_2 .

Fresnel reflection coefficient r_{ab} , where $a, b = 0, 1$, or 2 in this case, corresponds to a wave reflected from medium b back into medium a while t_{ba} refers to the Fresnel transmission coefficient of the corresponding wave transmitted from medium a into medium b :

$$r_{ab} = \frac{\tilde{n}_b - \tilde{n}_a}{\tilde{n}_b + \tilde{n}_a}. \quad (2.10)$$

$$t_{ba} = \frac{2\tilde{n}_b}{\tilde{n}_a + \tilde{n}_b}. \quad (2.11)$$

Similarly, the resulting transmitted field E_t can be expressed as

$$\begin{aligned} E_t &= t_{10}t_{21}e^{-ik_1d}E_i + t_{10}t_{21}r_{10}r_{12}e^{-i3k_1d}E_i + \dots \\ &= t_{10}t_{21}e^{-ik_1d}E_i \sum_{q=0}^Q (r_{10}r_{12}e^{-i2k_1d})^q. \end{aligned} \quad (2.12)$$

Taking the limit as $Q \rightarrow \infty$, the sums in (2.9) and (2.12) converge in a geometric series, allowing the total reflected and transmitted fields to be expressed as

$$E_r = \frac{r_{01} + r_{12}e^{-i2k_1d}}{1 + r_{01}r_{12}e^{-i2k_1d}}E_i \quad (2.13)$$

and

$$E_t = \frac{t_{01}t_{12}e^{-ik_1d}}{1 + r_{01}r_{12}e^{-i2k_1d}}E_i, \quad (2.14)$$

respectively [15, 12]. Dividing both sides of (2.13) and (2.14) by the incident field E_i yield quantities similar to Fresnel reflection and transmission coefficients. These quantities,

$$r_{eff} = \frac{r_{01} + r_{12}e^{-i2k_1d}}{1 + r_{01}r_{12}e^{-i2k_1d}}, \quad (2.15)$$

and

$$t_{eff} = \frac{t_{01}t_{12}e^{-ik_1d}}{1 + r_{01}r_{12}e^{-i2k_1d}}, \quad (2.16)$$

constitute “effective” reflection and transmission coefficients, respectively, as they encapsulate the total reflection and transmission responses of the layer structure in Figure 2.2, including contributions from all orders of internally reflected waves in medium 1 [15, 12].

Chapter 3

Terahertz Time Domain Spectroscopy

In THz time domain spectroscopy (THz-TDS), broadband THz pulses are used to estimate the complex refractive index of a sample from measurements made in either transmission or reflection mode. TPI, mentioned previously, uses these same THz pulses, typically in reflection mode, in a similar fashion to pulse-echo ultrasound. In TPI, THz-TDS measurements are typically taken along a 2D grid to generate a 3D dataset of time-domain waveforms from which echo pulse time of flight yields depth information and short-time Fourier transforms are used to acquire spectral information [4, 5, 7].

This chapter provides a description of the Picometrix T-Ray 4000 THz-TDS system at the Northwest Electromagnetics and Acoustics Research Lab (NEAR-Lab) at Portland State University (PSU), which was used in all measurements in this thesis. An overview of its principles of operation is presented as well as a comparison of three common material parameter estimation methods using transmission THz-TDS measurements. Problems associated with adapting these non-parametric transmission mode methods for reflection mode are then discussed. Finally, a parametric method based on theory introduced in Chapter 2 is presented as a potential solution.

3.1 NEAR-Lab Terahertz Time Domain Spectroscopy System

Many methods exist for generating and detecting terahertz signals. One of the most widely used techniques involves the use of ultrafast lasers in conjunction with non-linear crystals or semiconductor antennas. The Picometrix T-Ray 4000, shown in Figure 3.2, is an example of such a system. It operates by splitting pulses from a mode-locked 100 femtosecond (fs) fiber laser operating at $1.064\text{ }\mu\text{m}$ wavelength into a pump beam and a probe beam. The pump beam is used in the transmitter to excite carriers in a photoconductive bow tie antenna held under a DC bias. The excited carriers create brief pulses in current across the antenna with each incident laser pulse. These short current pulses have a rise time on the order of picoseconds, resulting in a 0.2-3 THz frequency spectrum with average power less than 10 - 500 μW . An example of a pulse and its frequency spectrum obtained using a fast Fourier transform (FFT) is shown in Figure 3.1. The antenna is affixed to a hyper-hemispherical silicon lens, which directs the energy toward a polyethylene collimating or focusing lens. The receiver consists of a similar set of lenses that focus the incoming THz pulse into another photoconductive antenna. After passing through a delay line, the probe beam excites carriers in the receiving antenna. The carriers pass through the antenna in the presence of a THz pulse's electric field, inducing a photocurrent proportional to the strength of the field. Sweeping the delay of the probe beam with respect to the pump beam allows coherent sampling the incoming THz pulse incident on the antenna. As the electric field of the THz pulse is sampled directly in the time domain, the recorded

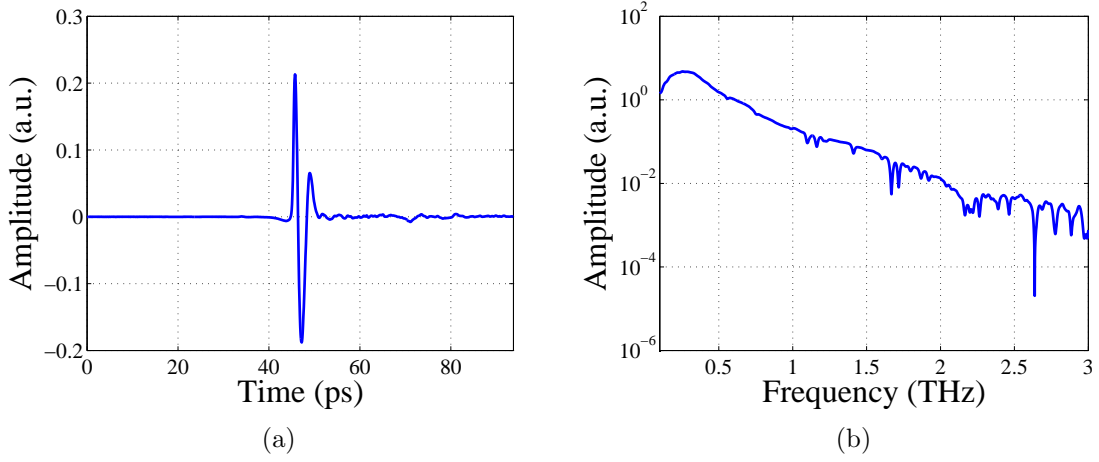
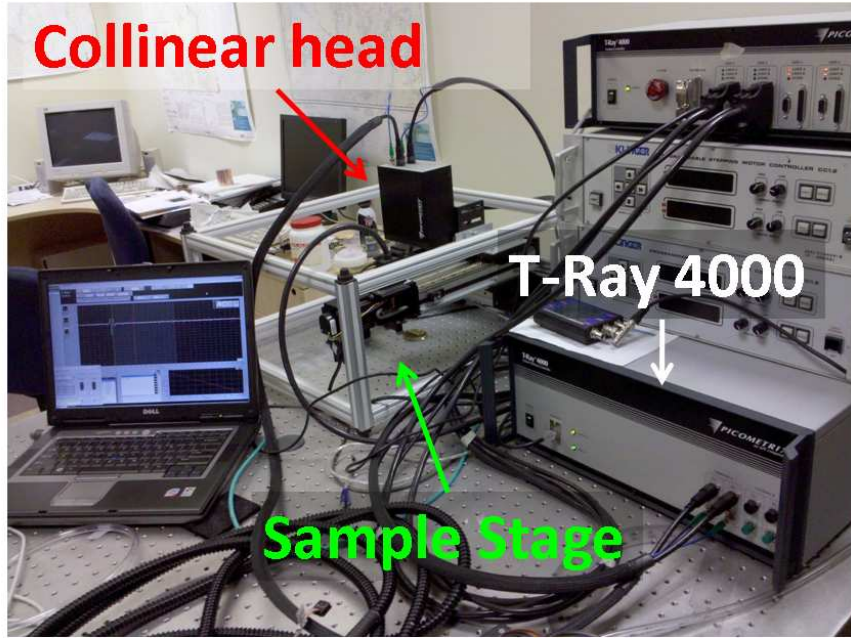


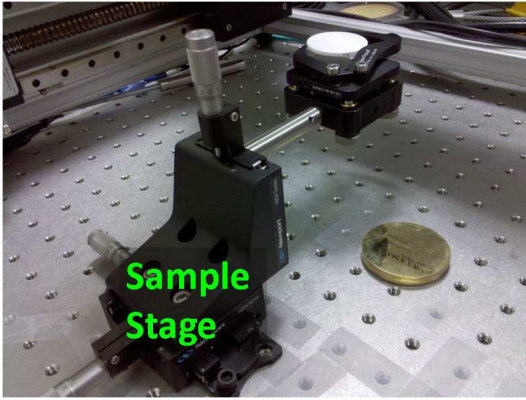
Figure 3.1: (a) A THz Pulse measured with the Picometrix T-Ray 4000 at the NEAR-Lab. (b) Spectrum of THz pulse in obtained via fast Fourier transform.

waveform contains both magnitude and phase information. The system records a new waveform every 10 ms and signal-to-noise ratio (SNR) is typically improved through coherent averaging of multiple waveforms.

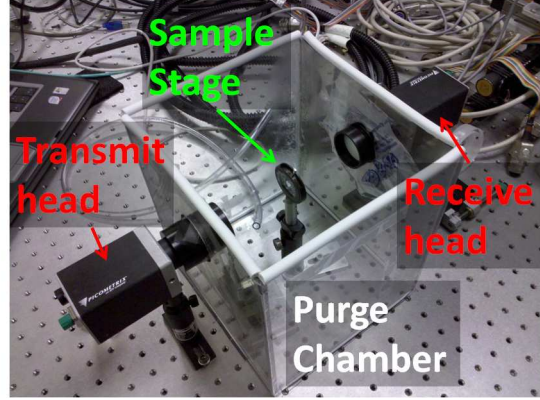
The TDS system can be configured for either monostatic measurements (in which the transmitter and receiver are collocated) or bistatic measurements (in which the transmitter and receiver are separate). The collinear head is shown in a reflection measurement configuration in Figure 3.2(a) along with a closeup of the sample stage shown in Figure 3.2(b). The collinear head uses a beam splitter so that a single polyethylene lens can be used for transmitting and receiving. As a result, SNR is reduced by approximately 10 dB across the measurement band as compared to the separate transmit and receive heads. This also reduces the maximum detectable bandwidth to approximately 2 THz, depending on the number of waveforms averaged. The separate transmit and receive heads are shown in Figure 3.2(c) configured



(a)



(b)



(c)

Figure 3.2: (a) T-Ray 4000 time domain spectroscopy system shown configured for reflection measurements with collinear head. (b) Closeup of reflection sample stage. (c) Closeup of transmit and receive heads configured for transmission measurement in purge chamber.

for transmission mode spectroscopy. The purge chamber is filled with dry air or nitrogen during measurement to reduce the influence of absorption lines in the measured spectra due to ambient water vapor in the air.

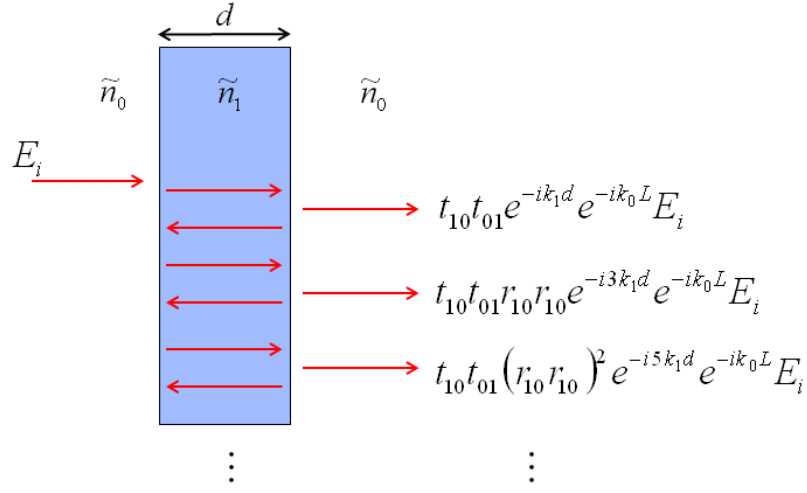
3.2 Material Parameter Estimation: Non-Parametric Techniques

To date, the bulk of the research on material parameter estimation has been conducted using non-parametric techniques in which each measured frequency domain data point yields a corresponding complex refractive index value, making the number of parameters required to describe a material's response to THz radiation equal to the number of measured data points [8, 9, 10, 11, 20, 3]. This section reviews three such commonly used non-parametric approaches for estimating material parameters in transmission spectroscopy; one which requires an optically thick sample and uses a single transmitted pulse, and two that solve the inverse problem including multiple pulses in the transmitted waveform. Starting with a description of the measurement procedure common to all three, the single pulse method is presented and its limitations are discussed. Attention then turns to the two methods of solving the inverse problem for multiple pulses in the transmitted waveform. The associated complications of the two methods are compared and one method is chosen for use in subsequent transmission mode analysis. Finally, the adaptation of the two multiple-pulse transmission mode techniques for reflection mode is discussed, starting with a description of the reflection mode measurement procedure. Complications in both reflection measurement and inversion processing are then presented before introducing a parametric technique as a potential solution.

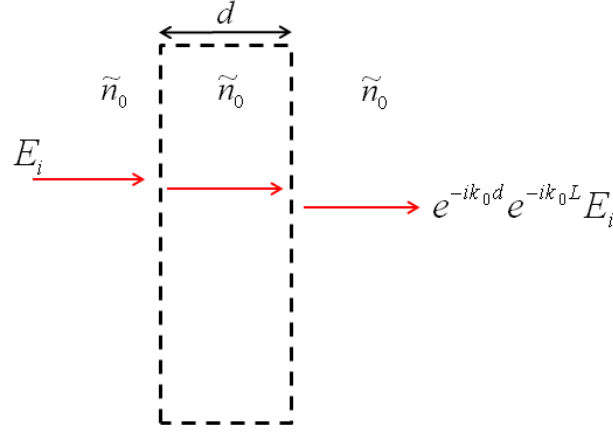
3.2.1 Transmission Mode

As shown in Figure 3.1(b), the measurement system has a frequency dependent input spectrum. To isolate the effect of a sample material on the received frequency spectrum of a transmission measurement, the measured spectrum from the sample is typically normalized by the spectrum from a reference measurement – a process often referred to as “deconvolution” [1]. The sample and reference measurements are illustrated in Figure 3.3. In the sample measurement shown in Figure 3.3(a), the incident wave E_i with spectrum shown in Figure 3.1(b) passes through the sample with complex refractive index \tilde{n}_1 and thickness d and is reflected within it as described in Section 2.2. These multiple reflections are typically visible as multiple pulses in the received time domain signal as shown in the transmitted waveform from a sample of polyethylene in Figure 3.4. The wave also accumulates phase as it propagates a total distance L through the air from the transmitter to the front surface of the sample and from the back surface of the sample to the receiver. This phase is represented by the complex exponential term, e^{-ik_0L} , where $k_0 = \frac{2\pi f}{c}\tilde{n}_0$ and $\tilde{n}_0 \cong 1$ are the wavenumber in and refractive index in medium 0 (which is always assumed to be free space), respectively. Similarly, Figure 3.3(b) shows how E_i accumulates phase along its path L through the air as well through a path of air of equal length to the sample thickness d .

Once the sample and reference measurements are made, two main approaches can be used to estimate \tilde{n}_1 from the data; one method involves only the first transmitted



(a) Sample measurement



(b) Reference measurement

Figure 3.3: Typical transmission measurement configuration: (a) Sample measurement. (b) Reference measurement required for deconvolution.

pulse, the other involves additional pulses.

Material Parameter Estimation from a Single Pulse

The method employed by [20], [3], and [4] involves assuming $n_1 \gg \kappa_1$ so that \tilde{n}_1 can be approximated as $\tilde{n}_1 \cong n_1$, allowing the transmission coefficients t_{01} and t_{10}

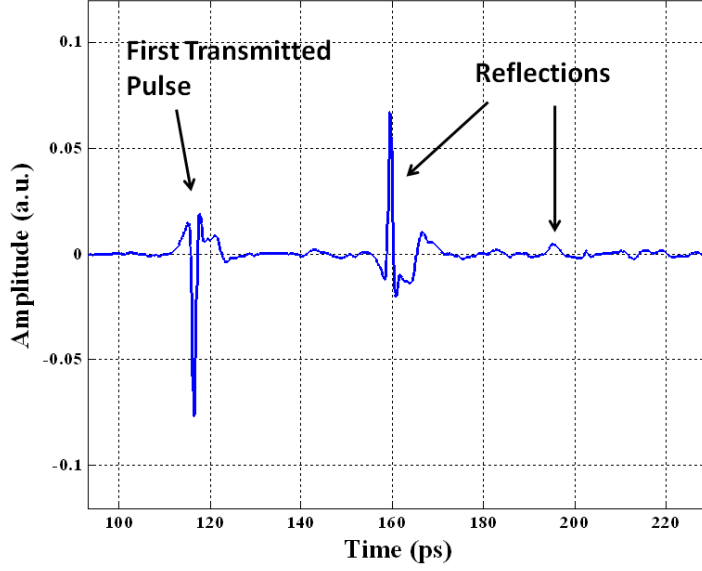


Figure 3.4: Transmission measurement of a sample of polyethylene showing multiple pulses due to reverberation within the sample as shown in Figure 3.3 and described in Section 2.2.

to be approximated by

$$t_{01} \cong \frac{2n_0}{n_1 + n_0} \quad (3.1)$$

and

$$t_{10} \cong \frac{2n_1}{n_0 + n_1}. \quad (3.2)$$

For a sample with sufficient optical thickness, the first transmitted pulse E_{samp} can be isolated in the time domain. Using (3.1) and (3.2) in (2.12) for the case of $Q = 0$ and including the phase shift e^{-ik_0L} introduced by the wave's propagation a distance

L through the air, E_{samp} becomes

$$\begin{aligned} E_{samp} &= t_{10}t_{01}e^{-ik_1d}e^{-ik_0L}E_i \\ &= \frac{4n_0n_1}{(n_0 + n_1)^2}e^{-ik_1d}e^{-ik_0L}E_i. \end{aligned} \quad (3.3)$$

Similarly, the reference measurement E_{ref} can be expressed as

$$E_{ref} = e^{-ik_0d}e^{-ik_0L}E_i. \quad (3.4)$$

After deconvolution with the reference measurement, the measured quantity E_{dec} becomes

$$\begin{aligned} E_{dec} &= \frac{E_{samp}}{E_{ref}} \\ &= \frac{t_{10}t_{01}e^{-ik_1d}e^{-ik_0L}E_i}{e^{-ik_0d}e^{-ik_0L}E_i} \\ &= \frac{4n_0n_1}{(n_0 + n_1)^2}e^{-ik_1d}e^{ik_0d} \\ &= \frac{4n_0n_1}{(n_0 + n_1)^2}e^{-ik_0(n_1-n_0)d}. \end{aligned} \quad (3.5)$$

Expressing the measured data in terms of its frequency dependent phase and magnitude yields

$$E_{data} = me^{i\phi}. \quad (3.6)$$

Rearranging (3.5), this measured magnitude m and phase ϕ in (3.6) can then be used to calculate the refractive index and extinction coefficient using

$$n_1 = \frac{\phi c}{2\pi f d} + n_0 \quad (3.7)$$

and

$$\kappa = \frac{c}{2\pi f d} \ln \left(\frac{4n_0 n_1}{m (n_0 + n_1)^2} \right). \quad (3.8)$$

While this method is simple and effective, there are practical limits on the thickness of samples on which it can be used [21]. Samples must be sufficiently thick for transmitted pulses not to overlap in time yet thin enough to allow a detectable amount of signal to pass through the sample. If the sample is too thin, a narrow time window must be used to isolate the first transmitted pulse. This reduces the resolution in the frequency domain, resulting in smoothing of the spectral features as illustrated in Figure 3.5 for the case of a transmission measurement through a sample of lactose. The portion of the measured waveform included in a short, 250 point FFT window centered about the peak value in the THz pulse and spanning ≈ 20 ps is shown in red in Figure 3.5(a). The corresponding FFT amplitude spectrum is shown in red in Figure 3.5(b). The effect of lengthening the FFT window to 500 points (‘medium’ window, spanning ≈ 39 ps) and 1000 points (‘long’ window, spanning ≈ 78 ps) are also shown in green and blue, respectively. The shorter FFT windows shown in Figure 3.5(a) act to reduce the frequency domain resolution, smoothing the

sharp spectral absorption features shown in Figure 3.5(b). On the other hand, if the

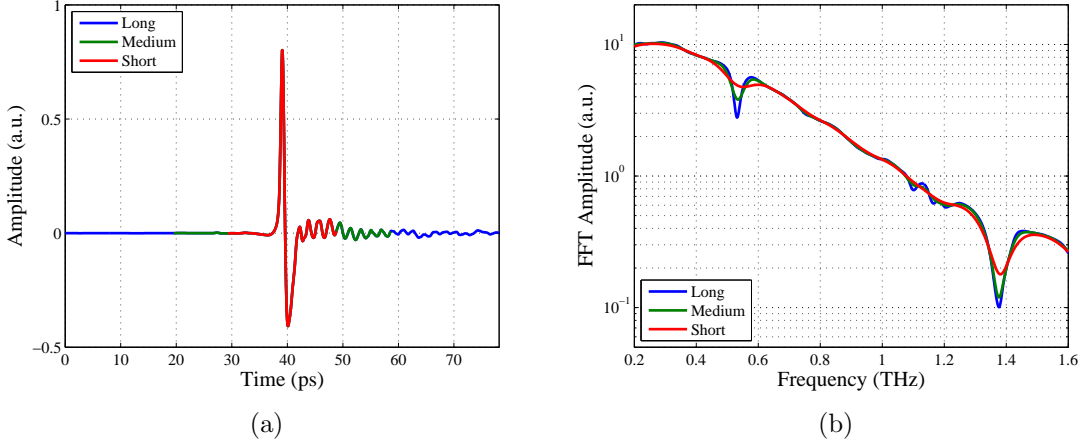


Figure 3.5: (a) THz pulse transmitted through a lactose sample. Blue: 1000 point (≈ 78 ps) ‘Long’ window, Green: 500 point (≈ 39 ps) ‘Medium’ window, Red: 250 point (≈ 20 ps) ‘Short’ window. (b) Shortening the time window reduces frequency resolution, smoothing the spectral features in resulting FFT spectrum.

sample is too thick, the amount of attenuation may exceed the dynamic range of the measurement system, causing the measured spectrum to reach the system’s noise floor. As such, these tradeoffs must be considered carefully during sample preparation [21].

Material Parameter Estimation from Multiple Pulses

Accounting for the influence of the multiple reflections in the time domain allows a longer time domain window to be used for measuring thinner samples, tying the lower limit of sample thickness to the measurement system’s SNR-limited bandwidth rather than the time delay of the first echo pulse as discussed previously [22]. While several variations of this technique have been developed [8, 9, 10, 11], most start from

a description of the measurement much like (3.5). The influence of the etalon effect can be accounted for using (2.16) with $\tilde{n}_2 = \tilde{n}_0$, resulting in the expression

$$\begin{aligned}
E_{dec} &= \frac{E_{samp}}{E_{ref}} \\
&= \frac{t_{eff} e^{-ik_0 L} E_i}{e^{-ik_0 d} e^{-ik_0 L} E_i} \\
&= t_{eff} e^{ik_0 d}
\end{aligned} \tag{3.9}$$

for the measured quantity. As (3.9) is a nonlinear function of the material's complex refractive index \tilde{n}_1 , the inverse problem cannot be solved analytically. Instead, it must be solved numerically by minimizing the squared error norm S_t between the measured data E_{data} and the model E_{dec} across frequency using \tilde{n}_1 . Representing the finite sets of measured frequency f , complex refractive index \tilde{n}_1 , model data E_{dec} , and measured data E_{data} as the discrete sets, $\{f_j\}$, $\{\tilde{n}_{1,j}\}$, $\{E_{dec,j}\}$, and $\{E_{data,j}\}$, respectively, with frequency domain index j , the squared error norm can be calculated using

$$S_t = \sum_j tER_j^2, \tag{3.10}$$

where the elements of the set $\{tER_j\}$ are given by

$$\begin{aligned}
tER_j &= |E_{data,j} - E_{dec,j}| \\
&= |E_{data,j} - t_{eff}(f_j, \tilde{n}_{1,j}) e^{ik_0(f_j)d}|.
\end{aligned} \tag{3.11}$$

The non-parametric nature of this method is indicated by the index j of $\tilde{n}_{1,j}$ in (3.11): Each individual data point in the measured data set $\{E_{data,j}\}$ has a corresponding $\tilde{n}_{1,j}$ value.

The minimization of S_t can be done in a variety of ways. One approach is to treat the frequency domain data points in $\{E_{data,j}\}$ separately and use a numerical algorithm at each individual frequency f_j in the dataset to search the complex plane of \tilde{n}_1 values for values of the real refractive index and extinction coefficient that minimize the corresponding tER_j . This is equivalent to solving (3.10), and is similar to the method used by Pupeza [11]. This approach will therefore be referred to hereafter as the Pupeza method. One main disadvantage of the Pupeza method is the multimodal solution space that results from the complex exponentials in (2.16) and (3.9) [10, 8]. This is illustrated in the ambiguity surfaces shown in Figure 3.6, generated by evaluating $(tER_j)^2$ on a 2D complex plane of \tilde{n}_1 values at four different frequencies for the simulated case of a material with a constant complex refractive index,

$$E_{data,j} = t_{eff}(f_j, \tilde{n}_1 = 1.5 - i0.001)e^{ik_0(f_j)d},$$

and thickness $d = 3.4$ mm. The blue regions in Figure 3.6 indicate minima in the solution space with the true solution, $\tilde{n}_1 = 1.5 - i0.001$, as labeled. As frequency increases, the minima move through the solution space, sometimes taking nonphysical values, with $n < 1$ and $\kappa < 0$. The solution space also scales such that the local minima become smaller and closer together with increasing frequency. These multiple

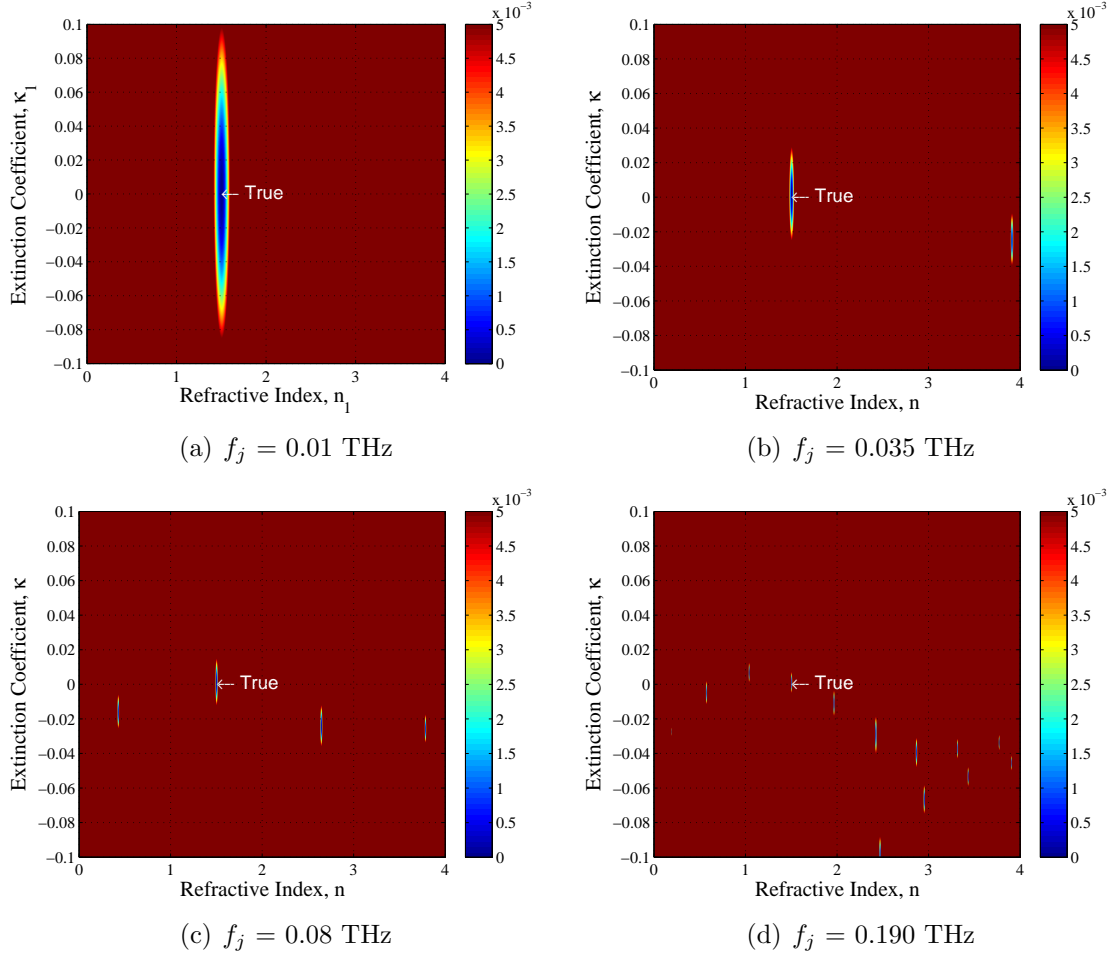


Figure 3.6: Squared transmission mode objective function (3.11) plotted as a function of refractive index, n_1 , and extinction coefficient, κ_1 , at four frequencies. In (3.11), $E_{data,j}$ was calculated from (3.9) using $d = 3.4$ mm and $\tilde{n}_1 = 1.5 - i0.001$ for all frequencies.

solutions act to throw off the numerical solver, resulting in discontinuous jumps in the estimated real refractive index and extinction coefficient.

This ambiguity in the solution space can be eliminated by unwrapping the phase of the measurement data E_{data} and model $t_{eff}e^{ik_0d}$ in a consistent manner before minimizing S_t in (3.10) [10, 8]. Using the unwrapped phase is facilitated by considering

the entire frequency series of the data and model rather than treating each frequency bin separately. This also changes the nature of the numerical algorithm required to perform the minimization, as will be shown.

In transmission mode, the THz pulse travels through the sample, causing the first transmitted pulse to arrive later in time than the reference pulse, which travels the same distance d through free space. This time delay results in a linear phase shift in the frequency domain, the slope of which is proportional to the difference between the optical path length through the sample and through free space,

$$(n_1 - n_0) d. \quad (3.12)$$

Localized variations in the sample's frequency dependent refractive index, n_1 are observable in localized changes in the frequency dependent slope of the unwrapped phase, ϕ . Similarly, localized changes in the extinction coefficient κ_1 result in localized absorption observable in the magnitude, m , of the transmitted spectrum. This results in straightforward relationships between n_1 and the unwrapped phase and κ_1 and the transmitted spectrum;

$$\begin{aligned} n_1(f) &\leftrightarrow \phi(f) \\ \kappa_1(f) &\leftrightarrow m(f). \end{aligned} \quad (3.13)$$

Dorney [10] used a variation on the gradient descent algorithm from [23] to exploit

this relationship. Starting with an initial guess for n_1 based on the sample thickness, time delay between the reference pulse and the first pulse transmitted through the sample, and assuming $\kappa_1 = 0$, the error in magnitude and unwrapped phase between measurement and model at frequency f_j ,

$$\begin{aligned} mER_j &\equiv |E_{data,j}| - |t_{eff}(f_j, \tilde{n}_{1,j})e^{ik_0(f_j)d}| \\ \phi ER_j &\equiv \angle E_{data,j} - \angle t_{eff}(f_j, \tilde{n}_{1,j})e^{ik_0(f_j)d}, \end{aligned} \quad (3.14)$$

can be used in a recursive update scheme, where values of the refractive index and extinction coefficient are updated using (3.14) according to

$$\begin{aligned} n_{1,j}^{new} &= n_{1,j}^{old} + \lambda \phi ER_j \\ \kappa_{1,j}^{new} &= \kappa_{1,j}^{old} + \lambda mER_j, \end{aligned} \quad (3.15)$$

where λ is the update step size. A step size of $\lambda = 0.01$ gives good results [10]. This numerical inversion method, hereafter referred to as the Dorney method, will be used in all subsequent non-parametric material parameter estimation due to its robustness (as compared to the method of Pupeza) against the influence of the multimodal solution space.

Whichever numerical inversion method is chosen, a good estimate of the sample thickness d is also required. If the incorrect thickness is used, the resulting estimated n_1 and κ_1 exhibit oscillatory behavior with frequency, [9, 10, 11]. This occurs because

the period of the oscillations in the frequency spectrum is a function of the optical path length $n_1 d$ in the sample. If the optical path length differs significantly between measurement and model, their oscillations in the frequency domain will have different spacing. In order for the oscillations to line up, the optical path length must change, which means either n_1 or d must be modified. To deal with this, a measure of the total variation in the estimated material parameters is used to optimize the unknown sample thickness [10, 11]. The total variation is determined by first calculating the set of absolute differences $\{D_j\}$ between adjacent material parameter values, given by

$$D_j = |n_{1,j-1} - n_{1,j}| + |\kappa_{1,j-1} - \kappa_{1,j}|. \quad (3.16)$$

The total variation V is simply the sum of these differences over frequency,

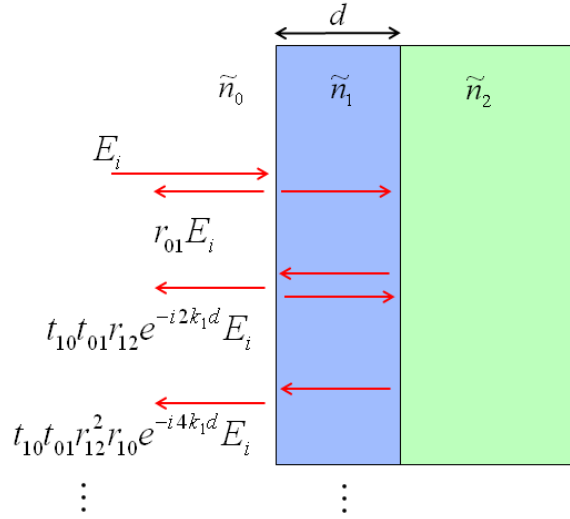
$$V = \sum_j D_j. \quad (3.17)$$

Repeating the inversion for a range of sample thickness values and calculating V for each yields the thickness \hat{d} that minimizes V , which constitutes the best estimate of the sample thickness.

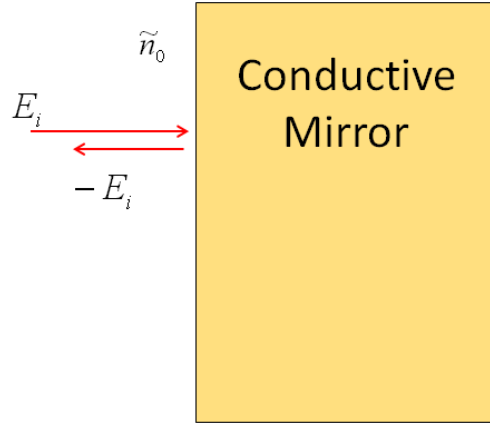
3.2.2 Reflection Mode

The measurement process in reflection mode differs from that in transmission mode in that a conductive mirror with reflection coefficient $r \cong -1$ is used as the reference as illustrated in Figure 3.7 for the case of a monostatic measurement (in which the

transmitter and receiver are collocated) made at normal incidence. Unlike in trans-



(a) Sample measurement



(b) Reference measurement

Figure 3.7: Ideal reflection measurement setup.

mission mode, the first received pulse has not penetrated the sample; only the trailing pulses are influenced (attenuated and shifted in phase) by propagation through the sample. This makes reflection mode more desirable for highly attenuating, opaque, and/or optically dense samples.

Ideally, the front surface of the reference mirror is placed exactly where the front

surface of the sample was during measurement as shown in Figure 3.7. If this is the case, the expression for the deconvolved measurement becomes simply

$$\begin{aligned}
E_{dec} &= \frac{E_{samp}}{E_{ref}} \\
&= \frac{r_{eff} e^{-i2k_0 L} E_i}{-e^{-i2k_0 L} E_i} \\
&= -r_{eff},
\end{aligned} \tag{3.18}$$

where r_{eff} is given by (2.15) and L is the distance from the transceiver to the front surface of both the sample and the reference. Unlike in the case of transmission mode, no time delay will occur between the reference pulse and the first reflected pulse from the sample in such a properly aligned reflection mode measurement. If the sample and reference are not placed the exact same distance from the transceiver, as depicted in Figure 3.8, a linear phase shift is introduced, modifying (3.18) and resulting in

$$\begin{aligned}
E_{dec} &= \frac{E_{samp}}{E_{ref}} \\
&= \frac{r_{eff} e^{-i2k_0 L} E_i}{-e^{-i2k_0 (L+\Delta L)} E_i} \\
&= -r_{eff} e^{i2k_0 \Delta L},
\end{aligned} \tag{3.19}$$

where ΔL is the difference in position between the reference and sample shown in Figure 3.8 [4]. Aligning reference and sample such that $\Delta L \cong 0$ is possible, but quite difficult in practice. Furthermore, unlike in a transmission measurement in which the measured phase is proportional to the optical thickness of the sample, the phase

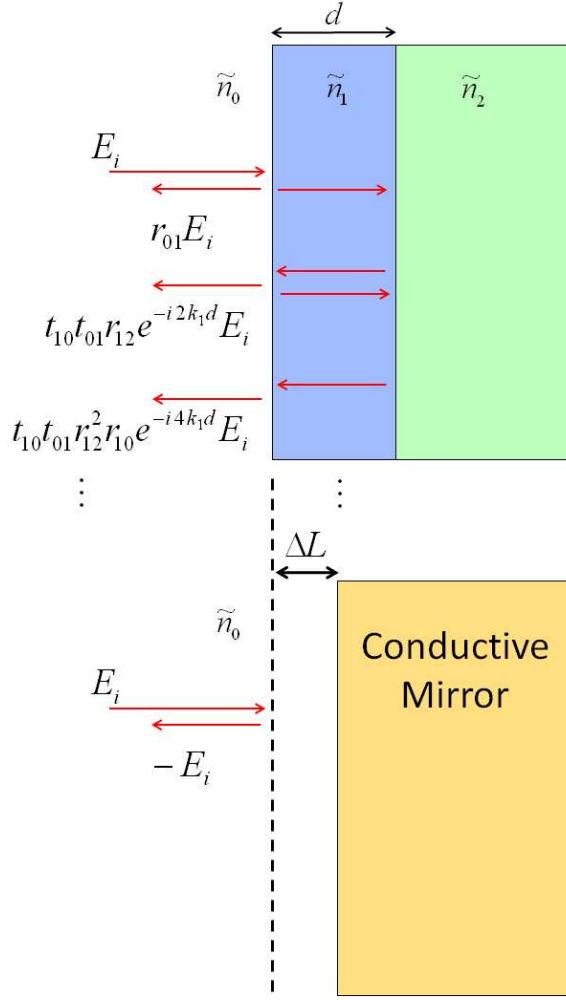


Figure 3.8: Positioning error ΔL between sample and reference in reflection mode introduces linear phase shift expressed in (3.19).

shift in the reflection measurement due to the sample's complex refractive index is much smaller than the influence of even a small misplacement error [24]. Methods have been developed to address this sensitivity, such as using the second derivative of the reflected phase with no reference measurement [25] or discarding the phase completely and using the first derivative of the reflected amplitude spectrum [6] to recover a qualitative spectral signature. However, it is unlikely that these techniques

will give useful results in the presence of interference phenomena from thin materials measured in reflection.

Instead, the slope of the linear trend in the unwrapped phase introduced by alignment error is used to estimate ΔL . This estimated ΔL is then used in (3.19) to apply a phase correction to the measured data [4]. The process then proceeds in a fashion similar to that from transmission mode; starting from the description of the measurement in (3.18) and using a numerical method to find the \tilde{n}_1 that minimizes the squared error norm S_r over frequency,

$$S_r = \sum_j rER_j^2, \quad (3.20)$$

where

$$rER_j = |E_{data,j} - r_{eff}(f_j, \tilde{n}_{1,j})|. \quad (3.21)$$

Just as in the case of transmission, the solution space of rER_j in the complex plane of n_1 and κ_1 values is multimodal, as shown in Figure 3.9, which shows values of $(rER_j)^2$ again evaluated on a complex plane of \tilde{n}_1 values at four different frequencies for the simulated case of a material with a constant complex refractive index,

$$E_{data,j} = -r_{eff}(f_j, \tilde{n}_1 = 1.5 - i0.001).$$

Unlike in transmission mode, in which a large linear phase shift is introduced by the wave's propagation through the sample, the main contribution to the reflection

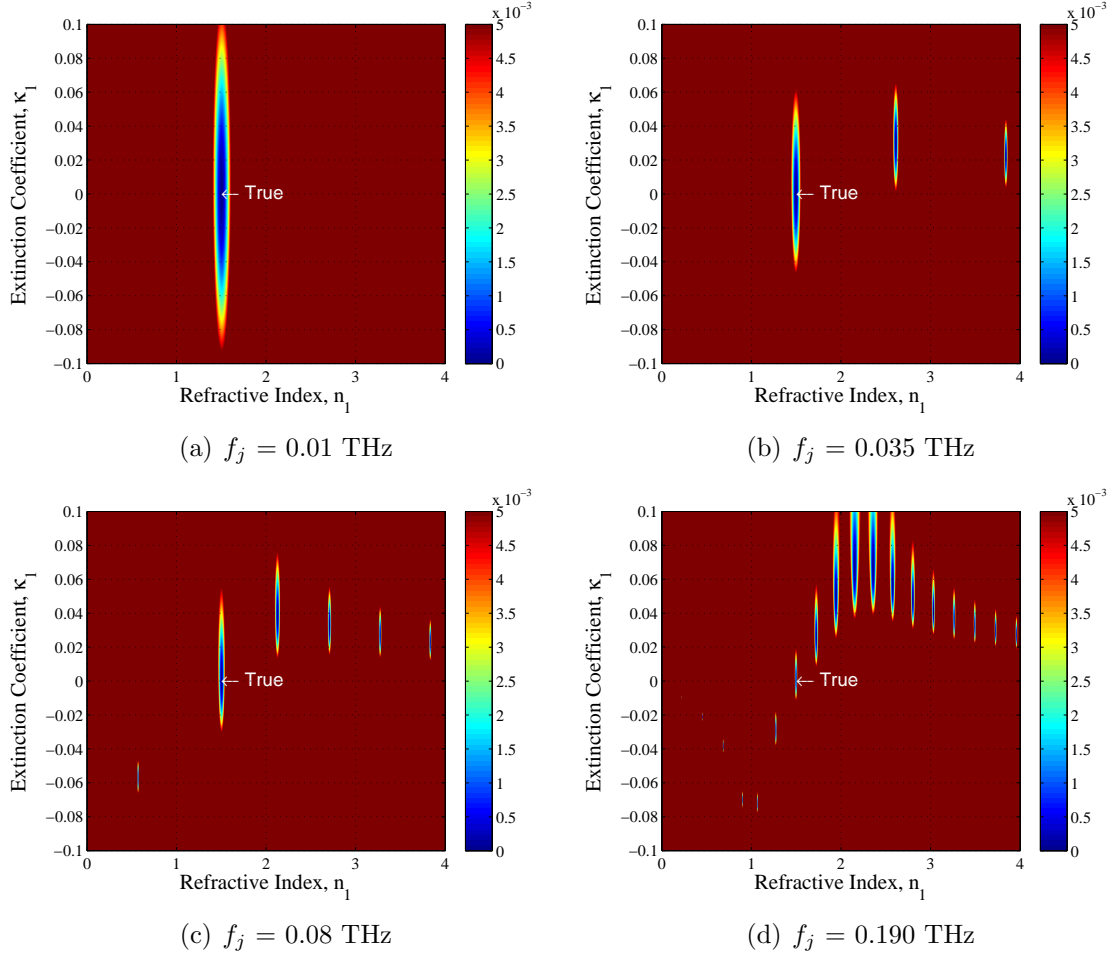


Figure 3.9: Squared reflection mode objective function (3.21) as a function of refractive index, n_1 , and extinction coefficient, κ_1 , at four frequencies. In (3.21), $E_{data,j}$ was calculated from (3.18) using $d = 3.4$ mm and $\tilde{n}_1 = 1.5 - i0.001$ for all frequencies.

response does not propagate through the sample in a properly aligned reflection measurement (depicted in Figure 3.7). Such a measurement will therefore have no linear phase shift. As a result, unwrapping the phase of a reflection measurement has no effect. In addition, the relationship between n_1 and κ_1 and the magnitude and phase of a reflection measurement are not nearly as straightforward as those in (3.13). These complications suggest an alternative approach – such as a parametric method – may

be preferable.

3.3 Material Parameter Estimation: Parametric Technique

If some assumptions are made about the complex refractive index, i.e. that it can be described by a parametric model such as the Lorentz model described in Section 2.1.1, the problems of a multimodal solution space and a complicated relationship between n_1 and κ_1 measured phase and magnitude can be mitigated. This method involves parameterizing \tilde{n}_1 using (2.7) and (2.3) and calculating (3.9) or (3.18) directly from the Lorentz parameters.

In a recent paper, Ahmed [14] introduced a similar approach. That work involved modeling the dispersion of a variety of samples with a variety of models, including Lorentz, Drude, Debye, and Cole-Cole. In contrast to this work, the Ahmed paper only discussed materials with single resonant modes and only treated the case of transmission mode measurement. It also used a slightly different objective function for fitting the model to the measured data.

As described in Section 2.1.1, the complex refractive index \tilde{n}_1 of a material with P molecular resonant modes can be described concisely by $3P + 1$ Lorentz parameters using (2.7) and (2.3). The material parameter estimation problem is then solved by finding the vector of $3P + 1$ Lorentz parameters,

$$\boldsymbol{\theta} = [\epsilon_\infty, \omega_1, \omega_2, \dots, \omega_P, \Delta\epsilon_1, \Delta\epsilon_2, \dots, \Delta\epsilon_P, \gamma_1, \gamma_2, \dots, \gamma_P], \quad (3.22)$$

that minimize the squared error norms S_t and S_r in (3.10) and (3.20), respectively. This requires slight modifications in tER_j from (3.11) and rER_j from (3.21), resulting in

$$tER_j^p = |E_{data,j} - t_{eff}(f_j, \boldsymbol{\theta})e^{ik_0(f_j)d}| \quad (3.23)$$

for transmission mode, and

$$rER_j^p = |E_{data,j} - r_{eff}(f_j, \boldsymbol{\theta})| \quad (3.24)$$

for reflection mode. The superscript “ p ” in (3.23) and (3.24) indicates the parametric nature of this method, wherein the complex refractive index is described completely by the $3P+1$ Lorentz parameters in $\boldsymbol{\theta}$ as opposed to a single complex refractive index value for each data point as indicated by $\tilde{n}_{1,j}$ in (3.11) and (3.21).

Such nonlinear optimization problems can be solved using a variety of techniques. In later sections, two such methods including the Nelder-Mead simplex described in Appendix B will be compared in their efficiency and accuracy in minimizing S_r in (3.20) using (3.24) for a simple test case.

Chapter 4

Results and Discussion

In this chapter, the parametric material parameter estimation technique from Section 3.3 applied in transmission mode is validated by comparison to the conventional, non-parametric technique of Dorney [10] described in Section 3.2.1 using measurement data collected at the NEAR-Lab. The parametric inversion technique is then applied to reflection data from the same sample as was used in transmission, the results of which are compared to results from transmission mode. Simulations are then run to evaluate the performance of two different optimization algorithms, including the Nelder-Mead method described in Appendix B, applied to the parametric method in reflection mode. The algorithms are assessed in terms of their efficiency and robustness against poor initial guesses. Additional simulations were then run to quantify the parametric method's sensitivity to differing levels of noise and its ability to determine sample thickness and displacement error between reference and sample in addition to the Lorentz parameters of the sample material.

4.1 Comparison of Non-Parametric and Parametric Techniques in Transmission

Transmission TDS measurements of a 13 mm diameter sample pellet of 10% (by weight) α -lactose monohydrate in polyethylene (PE) powder prepared by John Wilkinson of the Naval Surface Warfare Center, Indian Head, MD [26, 27] were carried out at normal incidence using focusing lenses with a 3" focal length. Lactose was chosen as a sample material due to its sharp spectral features at 0.527, 1.19, and 1.378 THz [17, 18]. The measurements were taken in a dry nitrogen atmosphere to minimize the influence of water vapor absorption lines on the measured spectra. A measurement with the sample at the focal point was taken along with a reference measurement without the sample as described in Section 3.2.1. While the focused beam was not characterized completely, the focal depth was calculated to be 10.2 mm assuming the wavefront incident on the 1.5" diameter 3" focal lens was a Gaussian beam [28] with a frequency of 0.3 THz (corresponding to the frequency component with the highest FFT amplitude). For simplicity, the focused THz beam incident on the sample was approximated as a plane wave in a manner consistent with the literature [8, 9, 29, 17, 3, 27, 11]. The average of 10,000 time-domain waveforms was used for a peak SNR of approximately 80 dB at 0.3 THz in the reference measurement. Figure 4.1 shows the reference and sample waveforms in the time domain along with their respective FFT amplitude spectra. In transmission mode, usable bandwidth is set by the sample measurement, as it is usually lower in power than the reference due

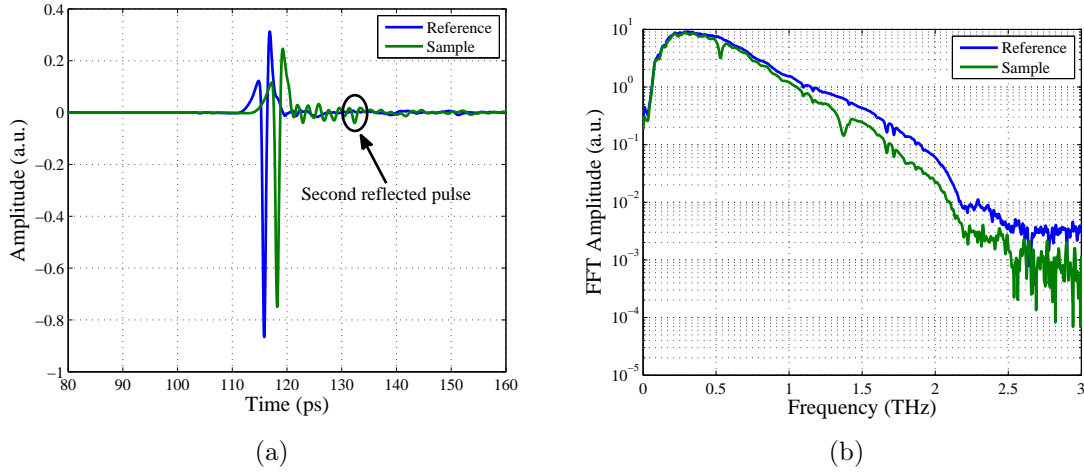


Figure 4.1: Lactose transmission data (a) in the time domain and (b) after FFT. Second reflected pulse visible at ≈ 132 ps in the time domain.

to reflection loss and sample attenuation. The upper limit of usable bandwidth was approximately 2 THz in this dataset. While frequency domain data below around 0.2 THz is typically unreliable, it is necessary to include in order to ensure the phase of both the model and measured data are unwrapped consistently to 0 at DC. Deconvolution in the frequency domain results in the transmissivity and phase shown in Figure 4.2. Frequency domain oscillations due to the etalon effect are clearly visible in the deconvolved transmittance spectrum in Figure 4.2(a). A corresponding second pulse is visible in the sample waveform at ≈ 132 ps in Figure 4.1(a).

As discussed in Section 3.2.1, an accurate sample thickness is required for non-parametric inversion. The sample was therefore measured with a micrometer with a tolerance of $\pm 1 \mu\text{m}$ at three points along the outside edge and one point in the center. These thickness values were found to be $d = 1.417$ mm, 1.450 mm, 1.426 mm, and

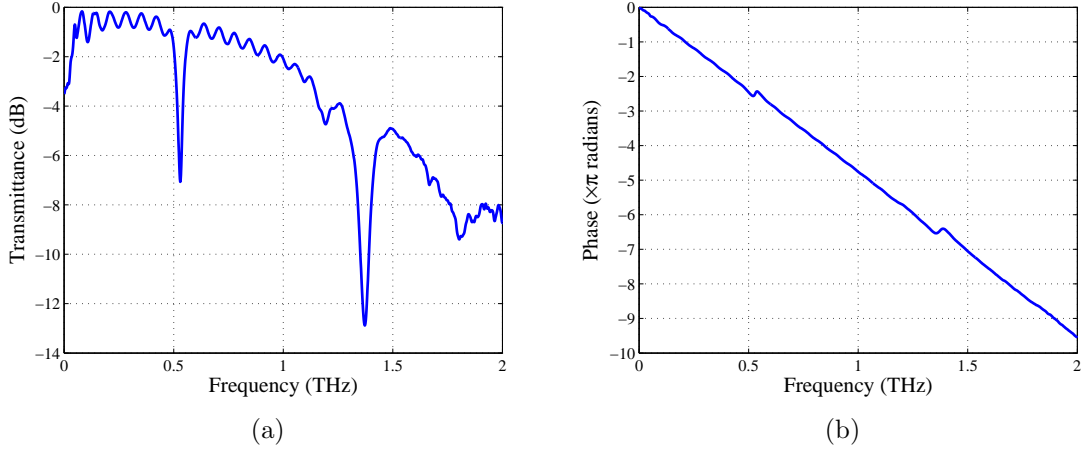


Figure 4.2: Deconvolved lactose (a) transmission spectrum and (b) phase.

1.418 mm, respectively, for a mean thickness $\mu_d = 1.428$ mm and standard deviation of $\sigma_d = 0.015$ mm. Using the non-parametric material parameter estimation technique of Dorney [10] described in Section 3.2.1, (3.14), and (3.15), were used to estimate \tilde{n}_1 for 101 values of thickness d in the range $\mu_d - 2\sigma_d \leq d \leq \mu_d + 2\sigma_d$. The total variation V was calculated for each assumed thickness using (3.17). Results of these calculations are shown in Figure 4.3, indicating a minimum corresponding to $\hat{d} \cong 1.420$ mm. This thickness was assumed in all subsequent analysis. The estimated real and imaginary parts of the complex refractive index, $\tilde{n}_1^{np} = n_1^{np} - i\kappa_1^{np}$, corresponding to this assumed thickness are shown in Figure 4.4. The superscript “ np ” indicates these material parameter curves are the result of the conventional, non-parametric method. While a small amount of oscillation is still present in the estimated n_1^{np} and κ_1^{np} curves in Figure 4.4, it is consistent with the results in the literature [18, 9] and probably arises from the slight suppression of the reflected

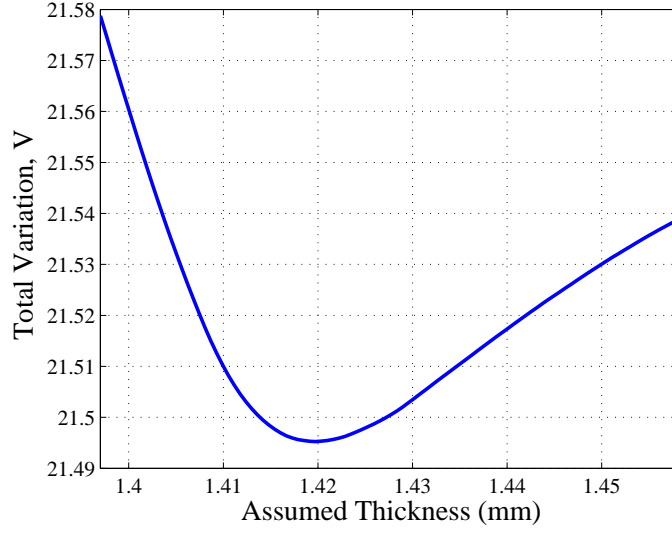


Figure 4.3: Total variation V of estimated real refractive index and extinction coefficient as a function of assumed sample thickness.

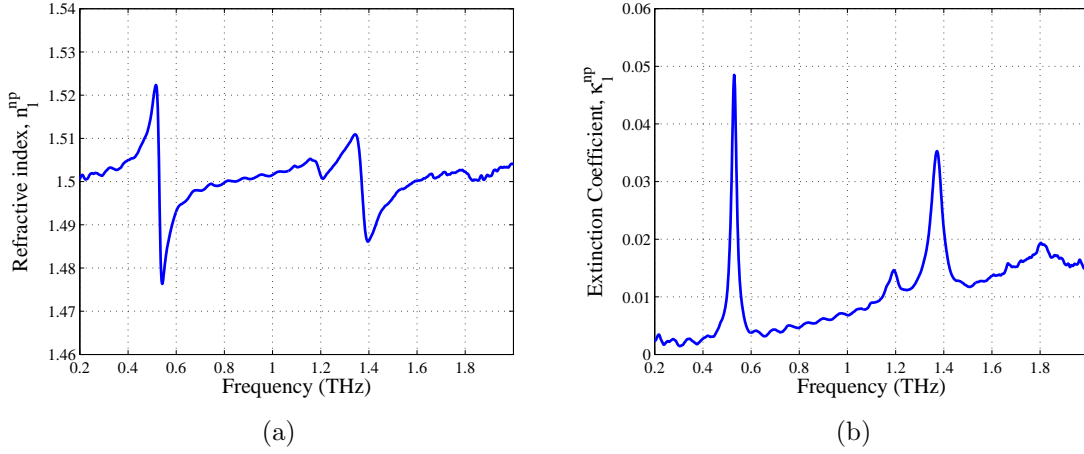


Figure 4.4: Estimated (a) refractive index n_1^{np} and (b) extinction coefficient κ_1^{np} using non-parametric methods described in Section 3.2.1 and [10].

pulses by the Hamming window applied to the time domain data prior to the Fourier transform.

To better compare the performance of the parametric inversion method to the conventional, non-parametric method in terms of Lorentz parameters, the complex

refractive index \tilde{n}_1^{np} resulting from the non-parametric method (shown in Figure 4.4) was fit with the Lorentz model (2.7). This process is illustrated in the flowchart in Figure 4.5, which shows the aforementioned estimation of \tilde{n}_1^{np} using the conventional non-parametric method of Dorney as Step 1. The Lorentz fitting shown in Step 2

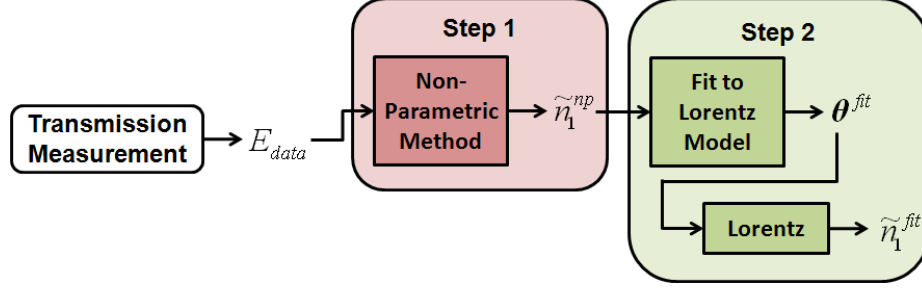


Figure 4.5: Flowchart showing the estimation of the complex refractive index \tilde{n}_1^{np} from deconvolved data E_{data} using the conventional non-parametric method of Dorney as Step 1. Step 2 consists of fitting of the Lorentz model (2.7) to \tilde{n}_1^{np} , which yields the vector θ^{fit} of Lorentz parameters. Using θ^{fit} in the Lorentz model yields \tilde{n}_1^{fit} .

was carried out using the Nelder-Mead algorithm described in Appendix B to find the Lorentz parameter vector θ^{fit} that minimizes the squared error norm S_n between \tilde{n}_1^{np} and the output of the Lorentz model, $\tilde{n}_1^{Lorentz}$, across frequency;

$$S_n = \sum_j nER_j^2, \quad (4.1)$$

where

$$nER_j = |\tilde{n}_{1,j}^{np} - \tilde{n}_1^{Lorentz}(f_j, \theta)|^2. \quad (4.2)$$

As indicated in Figure 4.5, \tilde{n}_1^{fit} is the output of the Lorentz model that results from using the parameters in θ^{fit} as inputs. Later sections will illustrate why the Nelder-Mead method was the algorithm of choice for this type of minimization problem,

including the parametric method described in Section 3.3. The initial guesses for the Lorentz parameters required to initialize the Nelder-Mead algorithm are shown in Table 4.1. It should be noted that while these initial guess values were chosen by

Mode	ϵ_∞	$\omega_p/2\pi$ (THz)	$\Delta\epsilon_p \times 10^3$	$\gamma_p/2\pi$ (GHz)
-	2.15	-	-	-
1	-	0.5300	8.00	30.0
2	-	1.194	0.900	40.0
3	-	1.370	3.10	48.0
4	-	1.800	0.900	200
5	-	3.680	87.0	1800

Table 4.1: Initial Guess for Lorentz model parameters used in initializing the Nelder-Mead algorithm to solve (4.1).

trial and error in this analysis, a practical system will likely have initial guess values stored in a database of parameters for known materials of interest. The corresponding

Mode	ϵ_∞	$\omega_p/2\pi$ (THz)	$\Delta\epsilon_p \times 10^3$	$\gamma_p/2\pi$ (GHz)
-	2.08	-	-	-
1	-	0.5292	6.54	25.2
2	-	1.187	1.24	92.9
3	-	1.370	3.06	55.1
4	-	1.791	0.808	135
5	-	4.570	161	2310

Table 4.2: Lorentz model parameters in θ^{fit} resulting from solving (4.1) to fit \tilde{n}_1^{np} to the Lorentz model as shown in Step 2 of Figure 4.5.

refractive index \tilde{n}_1^{fit} is compared to \tilde{n}_1^{np} in Figure 4.6. The results of this fitting will be treated as a best-case scenario – the closest the Lorentz model can get to \tilde{n}_1^{np} – and will constitute the standard to which results of the parametric method for transmission will be compared.

The parametric method for transmission mode is shown in Step 3 of the flowchart

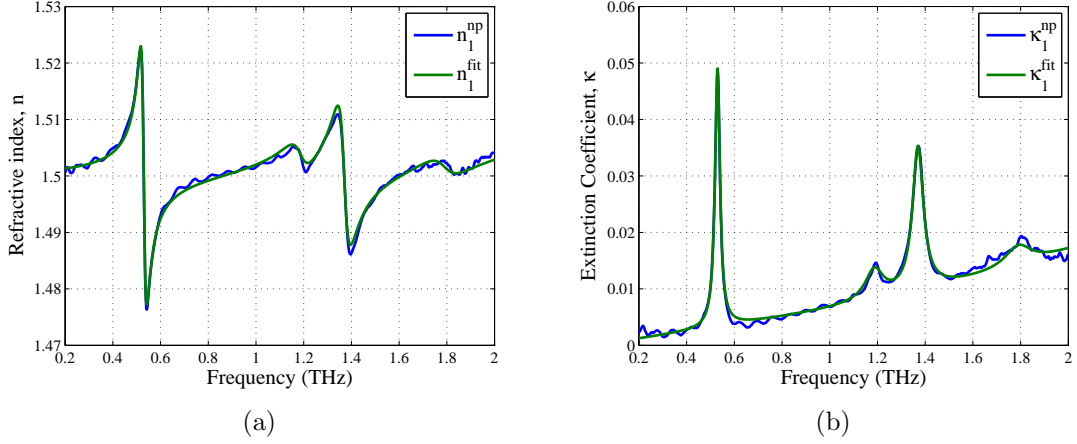


Figure 4.6: (a) Refractive index n_1^{np} estimated using non-parametric method of Dorney [10], and refractive index n_1^{fit} calculated using Lorentz parameters in θ^{fit} (shown in Table 4.2) in the Lorentz model (2.7), and (b) corresponding extinction coefficients, κ_1^{np} and κ_1^{fit} .

in Figure 4.7. The resulting vector of Lorentz parameters θ^p and corresponding

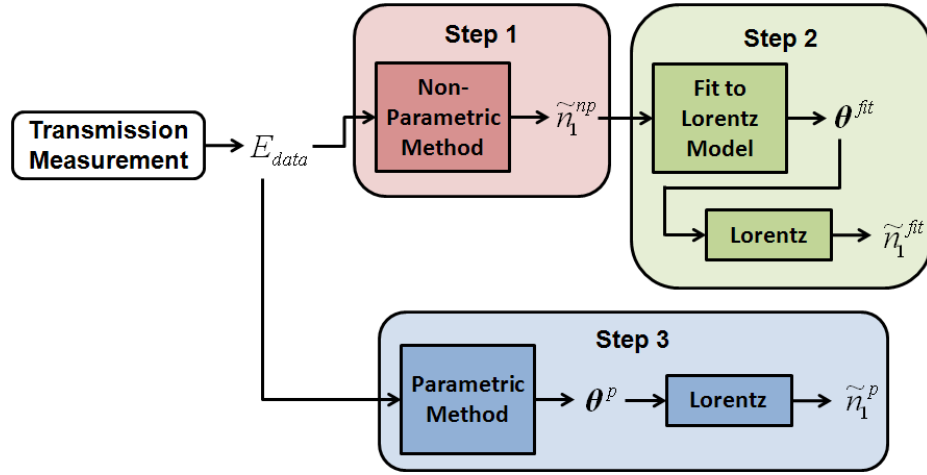


Figure 4.7: Flowchart showing the parametric method of Section 3.3 in Step 3. Fitting the measured data E_{data} by minimizing S_t in (3.10) using tER_j^p from (3.23) yields the Lorentz parameter vector θ^p . Using θ^p as the input to the Lorentz model yields \tilde{n}_1^p .

complex refractive index \tilde{n}_1^p are given the superscript “ p ” to signify the results of the parametric method, which in this case consisted of fitting the transmitted data shown

in Figure 4.2 by minimizing S_t in (3.10) using Lorentz model parameters as described in Section 3.3. The Nelder-Mead algorithm was initialized using the same initial guess parameters (given in Table 4.1) as were used in minimizing (4.1) to fit \tilde{n}_1^{np} to \tilde{n}_1^{fit} . Results of the fitting are shown in Figure 4.8 and corresponding Lorentz parameters in θ^p are given in Table 4.3. The resulting refractive index \tilde{n}_1^p is compared to \tilde{n}_1^{np}

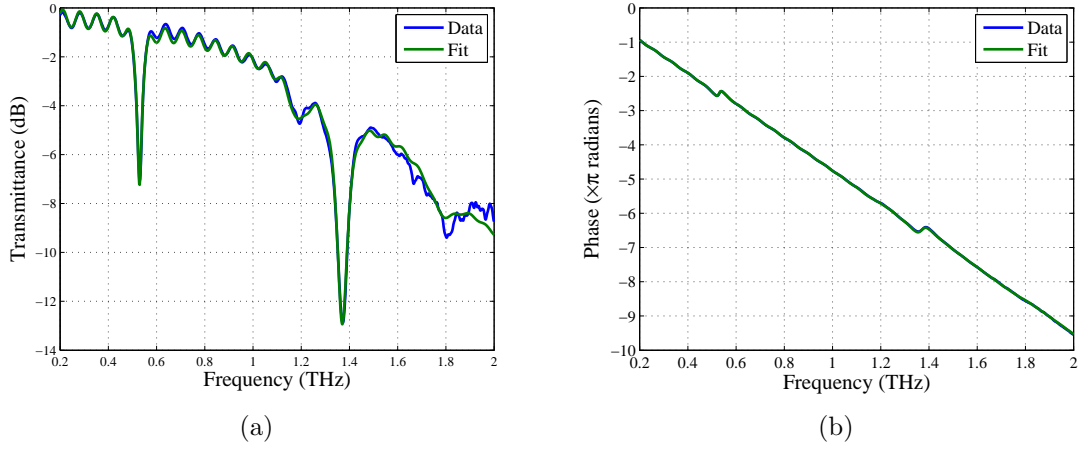


Figure 4.8: Results of parametric fitting of transmission data by minimizing the squared error norm S_t as described in Section 3.3: (a) transmittance spectrum and (b) phase.

Mode	ϵ_∞	$\omega_p/2\pi$ (THz)	$\Delta\epsilon_p \times 10^3$	$\gamma_p/2\pi$ (GHz)
-	2.06	-	-	-
1	-	0.5290	6.40	24.2
2	-	1.188	1.09	81.1
3	-	1.370	3.05	54.7
4	-	1.787	0.867	150
5	-	4.831	179	2380

Table 4.3: Lorentz model parameters in θ^p resulting from minimizing S_t using the parametric method as shown in Step 3 of Figure 4.7.

and \tilde{n}_1^{fit} in Figure 4.9.

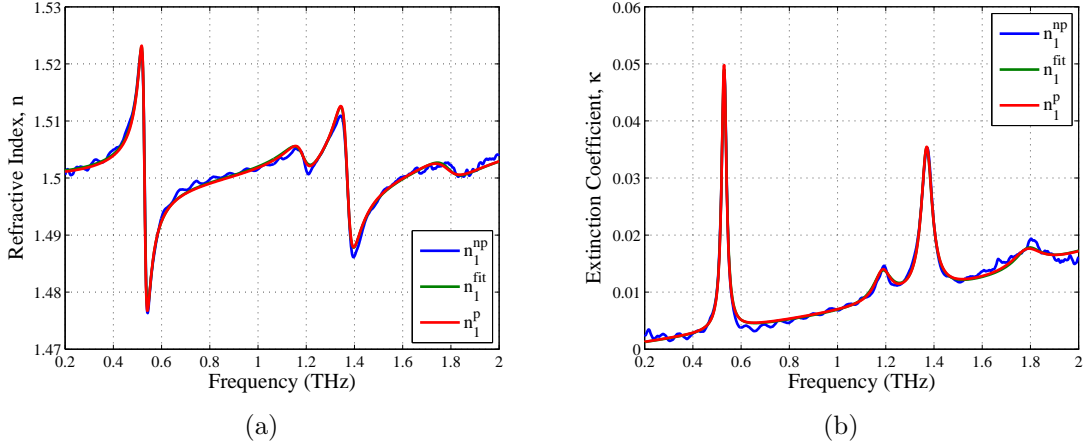


Figure 4.9: (a) Refractive index and (b) extinction coefficient estimated using parametric method described in Section 3.3 applied in transmission mode (\tilde{n}_1^p) compared to results of non-parametric method of Dorney [10] (\tilde{n}_1^{np}) and fitting of non-parametric result with Lorentz model (\tilde{n}_1^{fit}).

Figure 4.9 shows close qualitative agreement between \tilde{n}_1^{fit} and the results of the parametric method, \tilde{n}_1^p . The agreement is close enough that the curve of \tilde{n}_1^{fit} is almost indistinguishable from that of \tilde{n}_1^p . Both of these curves are also in agreement with the results of the conventional method, \tilde{n}_1^{np} . The degree of this agreement was quantified by comparing S_n from (4.1) to the same calculation with \tilde{n}_1^p substituted for $\tilde{n}_1^{Lorentz}$ in (4.2). The comparison is shown in Table 4.4. The close values of S_n show that

	S_n
\tilde{n}_1^{fit}	0.00111
\tilde{n}_1^p	0.00117

Table 4.4: Quantitative comparison of \tilde{n}_1^{fit} and \tilde{n}_1^p to \tilde{n}_1^{np} in terms of squared error norm S_n from (4.1).

fitting the Lorentz model to the conventional non-parametric estimate of the complex refractive index and using the parametric method in transmission yield comparable

results.

Table 4.5 shows the difference between the Lorentz parameters in $\boldsymbol{\theta}^{fit}$ and those in $\boldsymbol{\theta}^p$. Absolute differences in GHz are given for spectral locations ω_p and widths γ while percentage differences are given for the unitless quantities ϵ_∞ and $\Delta\epsilon_p$. The most

Mode	% Difference in ϵ_∞	Difference in $\omega_p/2\pi$ (GHz)	% Difference in $\Delta\epsilon_p$	Difference in $\gamma_p/2\pi$ (GHz)
-	0.903	-	-	-
1	-	0.2330	2.23	0.965
2	-	0.2760	12.0	11.8
3	-	0.1710	0.390	0.341
4	-	4.745	7.30	15.1
5	-	260.4	11.4	73.6

Table 4.5: Difference in Lorentz parameters between those in $\boldsymbol{\theta}^{fit}$ and those in $\boldsymbol{\theta}^p$. Absolute differences are given for $\omega_p/2\pi$ and $\gamma_p/2\pi$ while percent differences are given for ϵ_∞ and $\Delta\epsilon_p$.

relevant information for the purpose of material identification or characterization of lactose are the Lorentz parameters of the lowest three resonant modes. As Table 4.5 shows, the difference in spectral location between the lowest three modes of \tilde{n}_1^{fit} and \tilde{n}_1^p are all within 0.3 GHz. In terms of the spectral width, modes 1 and 3 (the two most prominent modes) are within 1 GHz of each other. The relative resonant strengths of the two most prominent modes are also in agreement, differing by less than 3%. This suggests that the accuracy in estimating spectral width and resonant strength increases with the prominence of the mode. In contrast, the spectral locations seem to be less dependent on mode prominence. The higher frequency modes are less consistent due to their lack of prominence as well as the reduced SNR and higher frequencies. The consistency in the estimates of the highest mode is further reduced

due to the fact that it resides outside of the system's measurement bandwidth.

Having established that the parametric method yields results consistent with the conventional non-parametric method in transmission mode, the parametric method used in the more practical case of reflection mode is now treated.

4.2 Comparison of Parametric Technique in Transmission and Reflection

Reflection measurements were conducted on the same lactose sample as was used in Section 4.1 at normal incidence with the 3" focal length lens in the collinear head in a dry air environment to reduce the effects of water vapor absorption lines on the measured reflection spectra. A measurement was also taken of a flat, polished metal reference mirror. As with the transmission mode measurements, the focused THz beam was approximated as a plane wave in a manner consistent with the literature [8, 9, 29, 17, 3, 27, 11]. The sample and reference mirror were mounted using a Newport Opti-Claw mount attached to a 3-axis translation stage during measurement, allowing precise adjustment of tip, tilt, and 3D translation. The first reflected pulse arrivals in the time domain waveforms of both reference and sample were aligned using the translation stage to reduce positioning error ΔL between the reference and sample. The average of 10,000 time-domain waveforms was used for a peak SNR of approximately 70 dB at 0.3 THz, allowing an upper limit of 2 THz on the usable bandwidth, consistent with transmission measurements. Figure 4.10 shows the reference and sample waveforms in the time domain along with their respective FFT

amplitude spectra. The deconvolved reflectance spectrum and unwrapped phase in

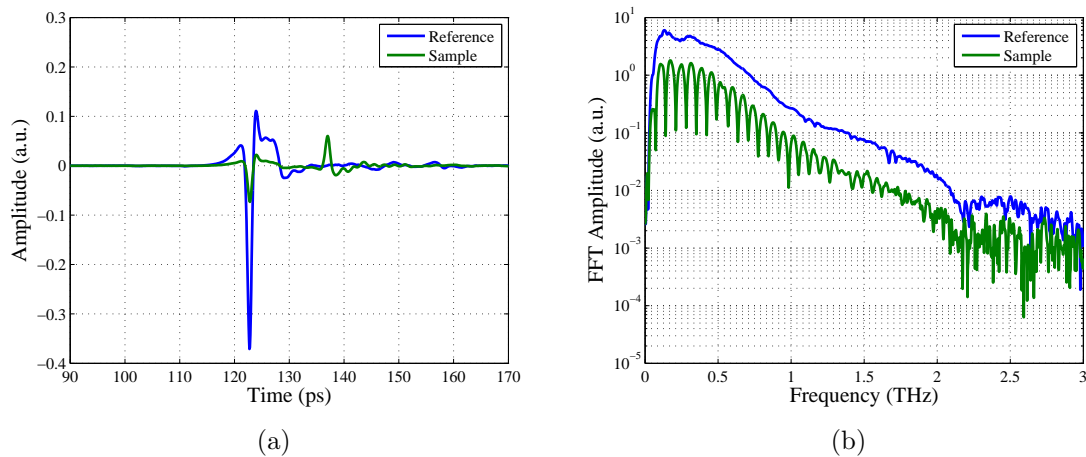


Figure 4.10: Measured lactose reflection data (a) in the time domain and (b) after FFT.

the frequency domain are shown in Figure 4.11. The aforementioned unreliability of

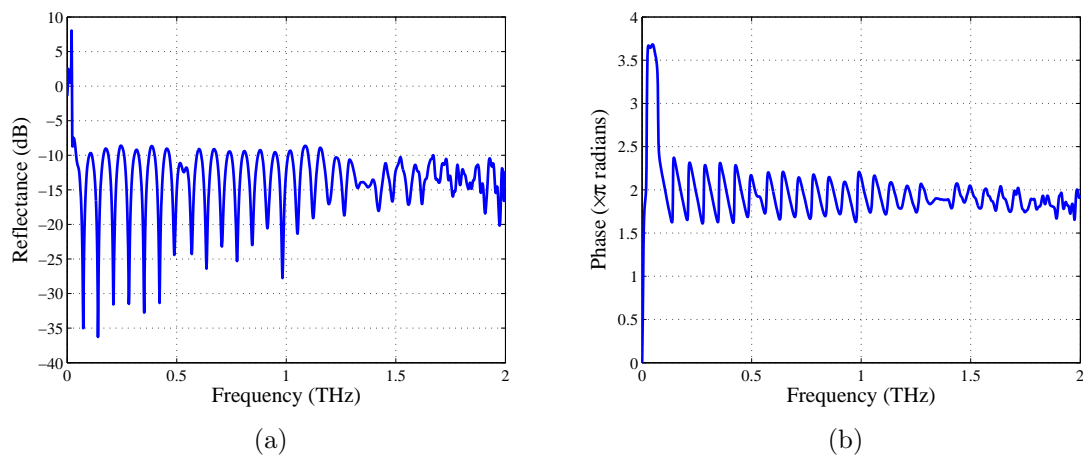


Figure 4.11: Deconvolved lactose (a) reflection spectrum and (b) phase.

the frequency domain data below around 0.1 THz is evident in Figure 4.11, but is again necessary to include when unwrapping the phase to ensure the phase of both the

model and measured data are unwrapped consistently to 0 at DC. In contrast to the second transmitted pulse in Figure 4.1(a), the second reflected pulse in Figure 4.10(a) is comparable in amplitude to the first reflected pulse. This results in the greater amplitude of frequency domain oscillations due to the etalon effect in reflection mode, as shown in Figure 4.11(a), as compared to transmission mode, shown in Figure 4.2(a).

Also notable is the influence of the spectral features on the etalon interference pattern shown in Figure 4.11(a) in the regions around 0.53 and 1.4 THz. The narrowband attenuation in those regions and increasing, wider band attenuation in the higher frequencies (shown in Figure 4.9(b)) result in selective attenuation of those frequency components as they pass through the sample. These frequency components then contribute less to the interference pattern, resulting in localized damping of the interference structure and causing the reflection from the front surface to dominate the total reflection response. In addition, the associated dispersion in these regions (shown in Figure 4.9(a)) result in slight dilations in the frequency spacing of the interference pattern due to the local increase in n_1 slightly below the resonant frequency and subsequent contractions due to the local decrease in n_1 slightly above the resonant frequency. This can be explained by referring to the complex exponential terms in the numerator and denominator of the expression (2.15) for r_{eff} , which give rise to its periodic behavior. The complex exponentials go through a complete cycle when the real parts of their arguments,

$$\Re \{2k_1 d\} = \frac{4\pi f n_1}{c} d,$$

change from 0 to 2π . These real parts change linearly with frequency according to

$$\frac{\partial}{\partial f} \frac{4\pi f n_1}{c} d = \frac{4\pi n_1}{c} d.$$

The change in the cycling rate of the complex exponentials in (2.15) with respect to frequency is therefore proportional to the real refractive index n_1 . Increases or decreases in n_1 act to increase or decrease the cycling rate, respectively, which changes the frequency spacing of the oscillations for frequencies close to a vibrational mode.

While alignment of the reference and first reflected sample pulse in Figure 4.10(a) suggests the sample and reference mirror were placed with very little displacement ΔL between them, the phase in Figure 4.11(b) shows a slight linear trend with negative slope in frequency, suggesting a small amount of positioning error in the data. This small displacement ΔL was estimated using a linear fit of the unwrapped phase as described in Section 3.2.2 and shown in Figure 4.12(a). In this case, the corresponding $\Delta L = -4.829 \mu\text{m}$, and the resulting corrected phase is shown in Figure 4.12(b).

Following the phase correction, the reflection data is ready for use in the parametric method from Section 3.3 to estimate the Lorentz parameters of the sample. The process is outlined in Step 4 of the flowchart in Figure 4.13. The resulting vector of Lorentz parameters $\boldsymbol{\theta}^r$ and corresponding complex refractive index \tilde{n}_1^r are given the superscript “ r ” to signify that they are the results of the parametric method applied in reflection mode. Again initializing the Nelder-Mead algorithm with the parameters in Table 4.1, the resulting fit of reflection data and corresponding Lorentz parameters

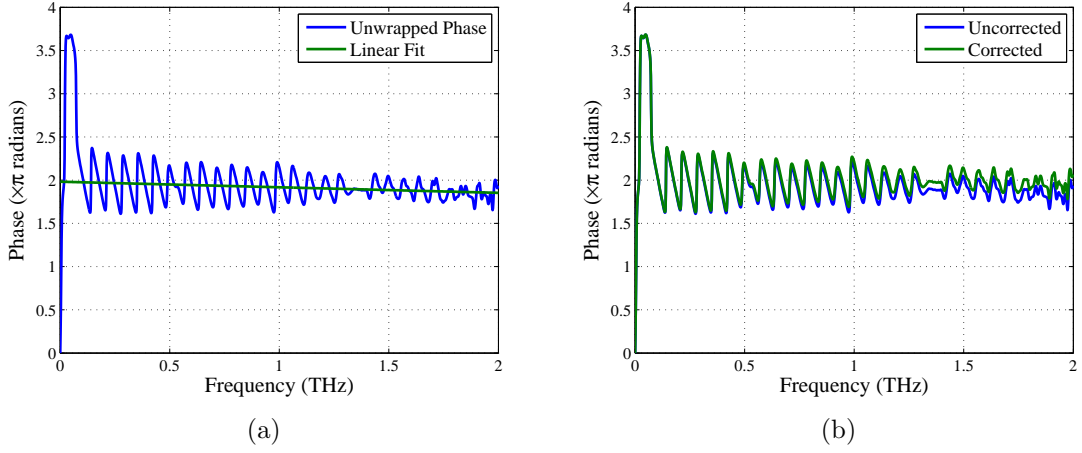


Figure 4.12: (a) Linear fit of unwrapped phase. (b) Result of phase correction using slope of fit line.

in θ^r are shown in Figure 4.14 and Table 4.6, respectively. The data and parametric

Mode	ϵ_∞	$\omega_p/2\pi$ (THz)	$\Delta\epsilon_p \times 10^3$	$\gamma_p/2\pi$ (GHz)
-	2.16	-	-	-
1	-	0.5300	7.00	35.5
2	-	1.192	0.620	34.3
3	-	1.366	4.18	50.9
4	-	1.094	0.000562	212
5	-	3.930	82.6	2940

Table 4.6: Lorentz model parameters in θ^r resulting from minimizing S_r using the parametric method applied to reflection mode as shown in Step 4 of Figure 4.13.

fit show strong qualitative agreement, especially in the lower frequencies (less than ≈ 950 GHz) where the SNR is higher.

The resulting complex refractive index \tilde{n}_1^r is compared to the transmission result \tilde{n}_1^p (from Figure 4.9) in Figure 4.15. Figure 4.15 shows fairly close qualitative agreement between the results of the parametric method in transmission and reflection, although not as close as that exhibited by the curves in Figure 4.9. While the resonant strengths

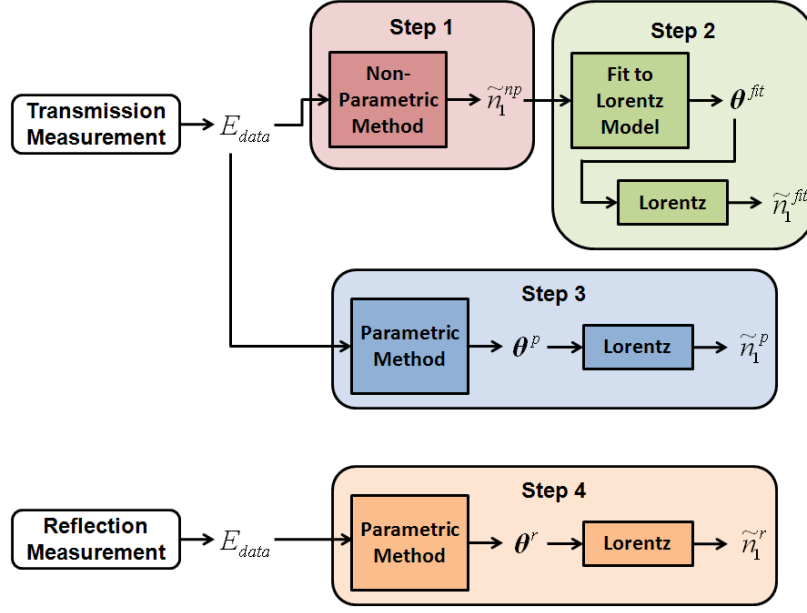


Figure 4.13: Flowchart showing the parametric method of Section 3.3 applied to reflection mode in Step 4. Steps 1-3 are also shown for reference. Fitting the measured data E_{data} by minimizing S_r in (3.20) using rER_j^p from (3.24) yields the Lorentz parameter vector θ^r . Using θ^r as the input to the Lorentz model yields \tilde{n}_1^r .

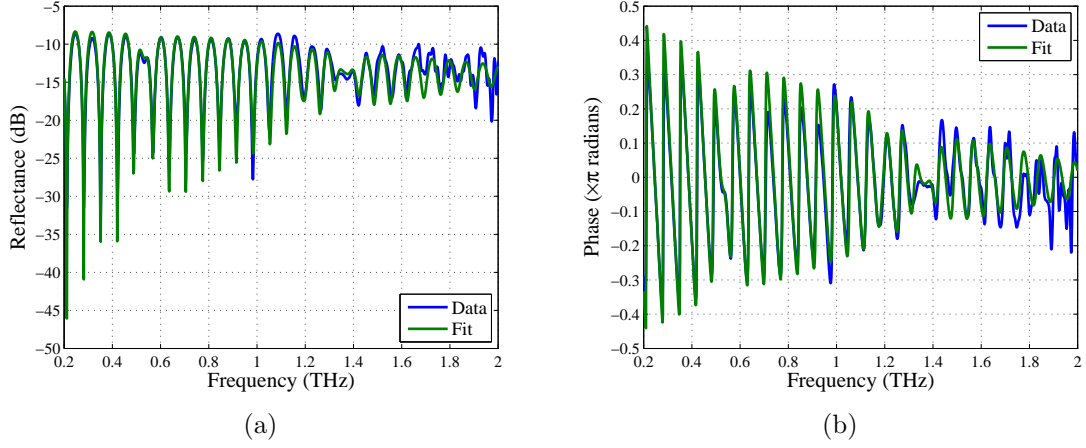


Figure 4.14: Results of parametric fitting of reflection data by minimizing the squared error norm S_r as described in Section 3.3: (a) reflection spectrum and (b) phase.

and widths of the lowest three modes appear to differ between \tilde{n}_1^p and \tilde{n}_1^r , their spectral locations appear to be consistent. This is verified by the difference values between

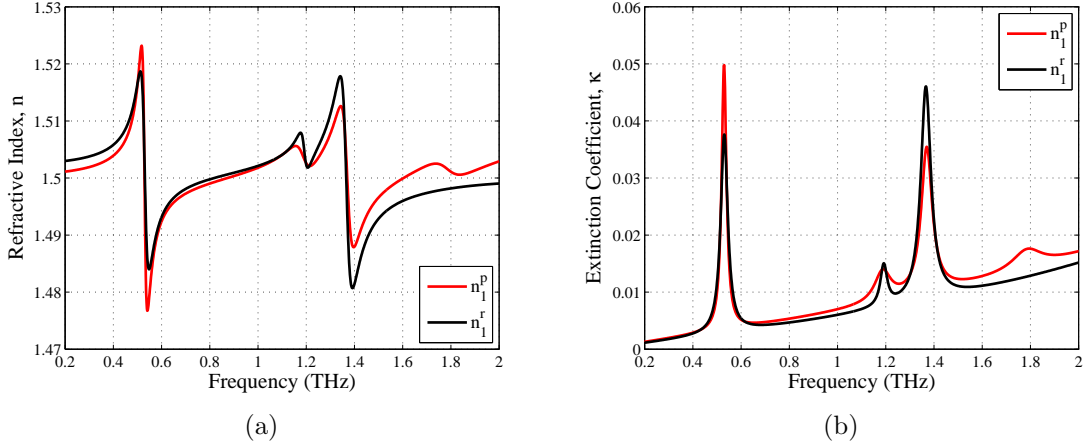


Figure 4.15: (a) Refractive index and (b) extinction coefficient estimated using parametric method in transmission \tilde{n}_1^p and reflection \tilde{n}_1^r .

Lorentz parameters in θ^r and θ^p shown in Table 4.7, which shows $\omega_p/2\pi$ varying by less than 5 GHz for these three modes. This is consistent with the observations from

Mode	% Difference in ϵ_∞	Difference in $\omega_p/2\pi$ (GHz)	% Difference in $\Delta\epsilon_p$	Difference in $\gamma_p/2\pi$ (GHz)
-	4.95	-	-	-
1	-	1.007	9.42	11.3
2	-	4.035	43.2	46.7
3	-	3.626	37.2	3.83
4	-	692.5	99.9	62.5
5	-	900.6	53.9	557

Table 4.7: Difference in Lorentz parameters between those in θ^r and those in θ^p . Absolute differences are given for $\omega_p/2\pi$ and $\gamma_p/2\pi$ while percent differences are given for ϵ_∞ and $\Delta\epsilon_p$.

Section 4.1 that suggest the relative robustness of the parametric method's ability to provide good estimates of spectral locations of the resonant modes in comparison to the other Lorentz parameters. Table 4.8 shows the percent difference in the resonant frequency ω_p of the lowest three modes estimated parametrically in reflection to the

corresponding values estimated in transmission. The values are within 0.4% of each other, indicating consistency between the values of ω_p estimated in transmission and reflection.

mode	1	2	3
% Difference	0.190	0.340	0.265

Table 4.8: Percent difference between ω_p values in $\boldsymbol{\theta}^r$ and $\boldsymbol{\theta}^p$.

The consistency of the results of the parametric method in both reflection (\tilde{n}_1^r) and transmission (\tilde{n}_1^p) with the result of the conventional, non-parametric, transmission mode method (\tilde{n}_1^{np}) was quantified by substituting \tilde{n}_1^r and \tilde{n}_1^p for $\tilde{n}_1^{Lorentz}$ in the expression for nER_j (4.2) and calculating S_n from (4.1) for each case. The results of the comparison, shown in Table 4.9, indicate that the transmission case gives a closer result than does the case of reflection mode. This is to be expected given the lower SNR of the reflection measurement due to the collinear head and the additional complications inherent in reflection measurement.

	S_n
\tilde{n}_1^p	0.00117
\tilde{n}_1^r	0.01729

Table 4.9: Quantitative comparison of \tilde{n}_1^p and \tilde{n}_1^r to \tilde{n}_1^{np} in terms of S_n from (4.1).

4.3 Simulations for Sensitivity Analysis

In this section, simulated data from a simple test case are first used to evaluate the performance of two different optimization algorithms, including the Nelder-Mead method described in Appendix B, applied to the parametric method in reflection

mode in terms of efficiency and robustness against poor initial guesses. Next, the effects of system noise on estimated Lorentz parameters is investigated. Finally, the ability of the parametric method to determine sample thickness and positioning error between sample and reference measurements is tested.

4.3.1 Initial Guesses for Numerical Optimization Algorithms

The parametric method described in Section 3.3 requires the use of a numerical optimization routine that should be capable of handling 4 to 30 or so unknowns and possibly as many constraints. Such routines typically require an initial guess for initialization. In future deployable systems, initial guesses may be stored in a database of known Lorentz parameter values for various materials of interest. Variations in concentration and other factors may complicate this process, so a robust system will need to be able to minimize S_r in (3.20) using rER_j^p in (3.24) with as little *a priori* information as possible. In this analysis, two optimization routines are compared in their ability to efficiently converge to the correct solution if given poor initial guesses. For simplicity, this was done using a test case of a hypothetical material with a single resonance, the Lorentz parameters for which are given in Table 4.10. Corresponding refractive index and extinction coefficient curves are shown in Figure 4.16. These

ϵ_∞	$\omega/2\pi$ (THz)	$\Delta\epsilon \times 10^3$	$\gamma/2\pi$ (GHz)
2	1	250	1

Table 4.10: Lorentz model parameters for hypothetical test material with a single resonance used in evaluation of numerical optimization techniques.

complex refractive index values were then used in (3.18) assuming a thickness of

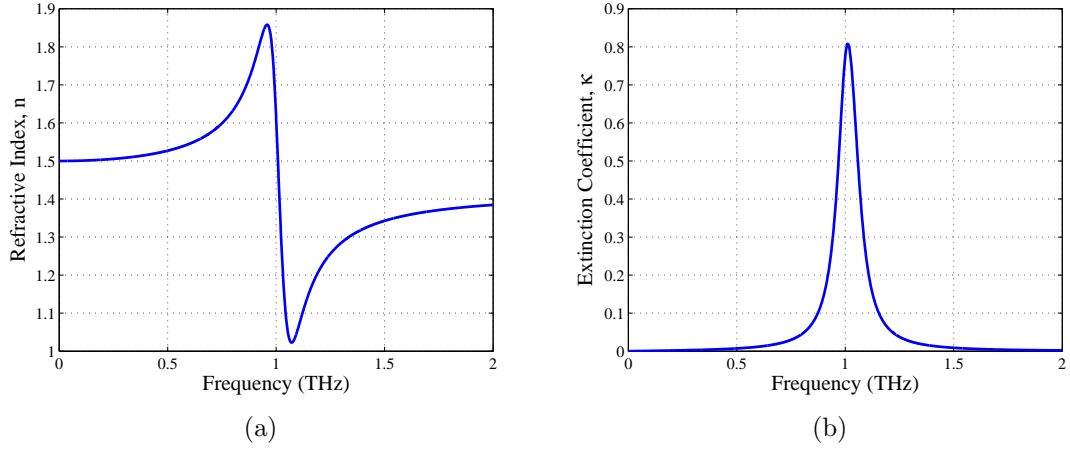


Figure 4.16: (a) refractive index and (b) extinction coefficient corresponding to Lorentz parameters in Table 4.10.

$d = 1.5$ mm to generate simulated reflection data shown in Figure 4.17. The two routines tested were the Nelder-Mead simplex algorithm described in detail in Appendix B and the medium-scale sequential quadratic programming (SQP) method implemented as the “Active-Set” algorithm in `fmincon.m` in MATLAB’s optimization toolbox. The Active-Set algorithm uses a quasi-Newton method to approximate the Hessian matrix of the Lagrangian function. This approximate Hessian is then used to obtain the search direction for a line search procedure. Convergence is attained when the magnitude of the search direction or its derivative fall below a given threshold while simultaneously, any constraints supplied to the optimizer are satisfied to within a given tolerance. Convergence criteria for the Nelder-Mead algorithm are given in Appendix B.

The algorithms were used to minimize S_r using the simulated data shown in Figure 4.17 as E_{data} after being initialized with each Lorentz parameter in Table 4.10

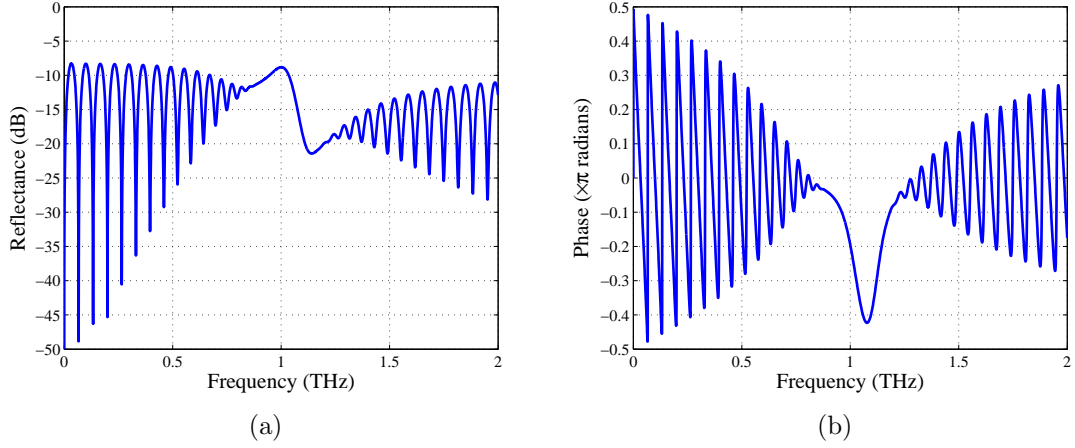


Figure 4.17: Calculated (a) reflectance and (b) unwrapped phase using assuming $d = 1.5$ mm in (3.18) and the refractive index and extinction coefficient values shown in Figure 4.16.

perturbed, one by one by a range of percentages from the correct value. Each algorithm's performance was quantified using two metrics, one of which being the number of iterations required for convergence. Each algorithm terminates if either is termination criteria are satisfied or after a specified maximum number of algorithm iterations and objective function evaluations. These upper limits on iterations and function evaluations were set high enough to ensure the algorithms terminated only due to convergence to a solution. Once convergence was reached, the squared difference norm S_n in (4.1) was calculated by substituting the correct complex refractive index (shown in Figure 4.16) and the estimated complex refractive index at convergence in place of \tilde{n}_1^{np} and $\tilde{n}_1^{Lorentz}$ in (4.2), respectively. The resulting S_n was used as the second performance metric for each algorithm. Figure 4.18 shows the result for perturbing the initial guess of ϵ_∞ . While fewer iterations of the active-set method were

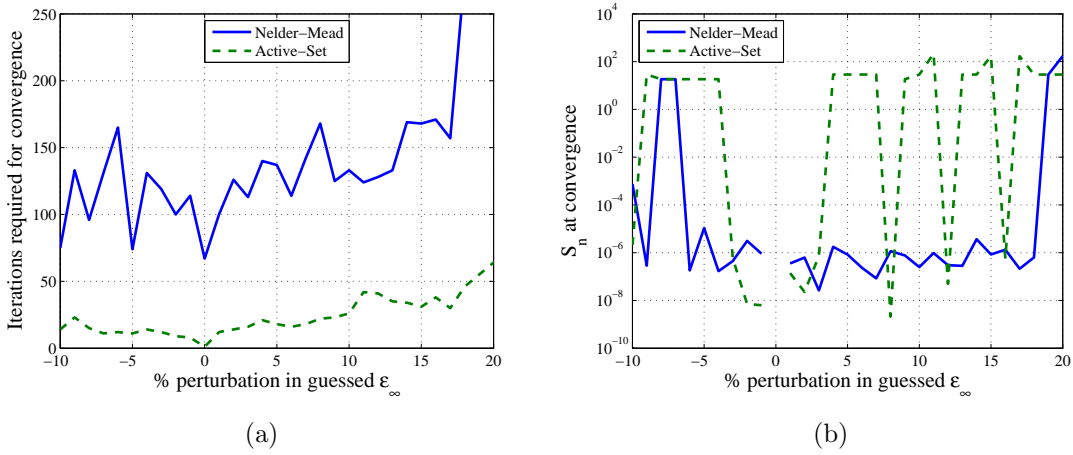


Figure 4.18: Effect of perturbing initial guess of ϵ_∞ on Nelder-Mead and Active-Set algorithms on (a) iterations required for convergence and (b) squared difference norm S_n between complex refractive index values at convergence and correct values shown in Figure 4.16. While less efficient, the Nelder-Mead algorithm converges consistently to the correct value (global minimum) for perturbations ranging from -6% to 18% of the correct value, while the more efficient active-set method converges to the correct value for perturbations within only $\pm 3\%$ of the correct value. Outside these ranges, both algorithms are thrown off by local minima.

required for convergence than were required of the Nelder-Mead method (as shown in Figure 4.18(a)), the Nelder-Mead algorithm consistently converged to values within $S_n < 10^{-4}$ for perturbations ranging from -6% to 18% of the correct value (as shown in Figure 4.18(b)). Outside this range, the Nelder-Mead method was attracted to local minima different from the global minimum corresponding to the correct value. In contrast, the active-set algorithm consistently converged to the correct solution for initial guess perturbations within $\pm 3\%$ of the correct value. Outside of this narrow range, the active-set algorithm converged to local minima distinct from the correct value (global minimum), as evident by the higher S_n values in this range. This suggests that while less efficient, the Nelder-Mead algorithm can be expected to be less

sensitive to poor initial guesses for ϵ_∞ than the active-set algorithm. This also suggests good results can be obtained using the Nelder-Mead method provided the initial guess is within -6% to 18% of the correct value.

The same perturbation test was run on initial guesses of $\Delta\epsilon$, the results of which are shown in Figure 4.19. In contrast to the results from perturbing ϵ_∞ , the active-set

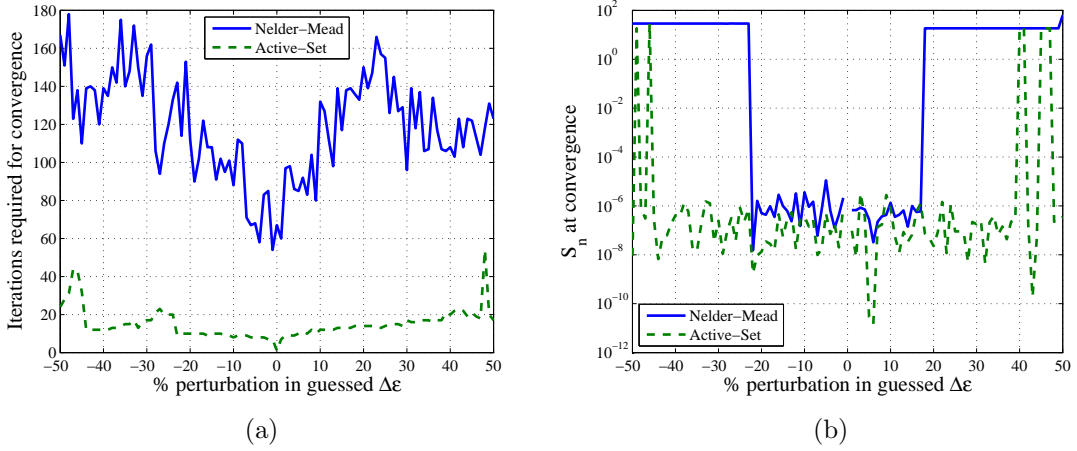


Figure 4.19: Effect of perturbing initial guess of $\Delta\epsilon$ on Nelder-Mead and Active-Set algorithms on (a) iterations required for convergence and (b) squared difference norm S_n between complex refractive index values at convergence and correct values shown in Figure 4.16. The more efficient active-set algorithm converges consistently to the correct value (global minimum) for perturbations ranging from -45% to 39% of the correct value, while the less efficient Nelder-Mead method converges to the correct value for perturbations from -22% to 17% of the correct value. Outside these ranges, both algorithms are thrown off by local minima.

method was found not only to be more efficient than the Nelder-Mead algorithm (as demonstrated in Figure 4.19(a)), but converged to the correct value (global minimum) if given initial guess values of $\Delta\epsilon$ perturbed from -45% to 39% of the correct value, as compared to -22% to 17% with the Nelder-Mead algorithm (shown in Figure 4.19(b)). Outside these ranges, the algorithms converged to local minima distinct from the

correct value (global minimum).

Results of perturbing initial guesses of γ are shown in Figure 4.20. Like for the

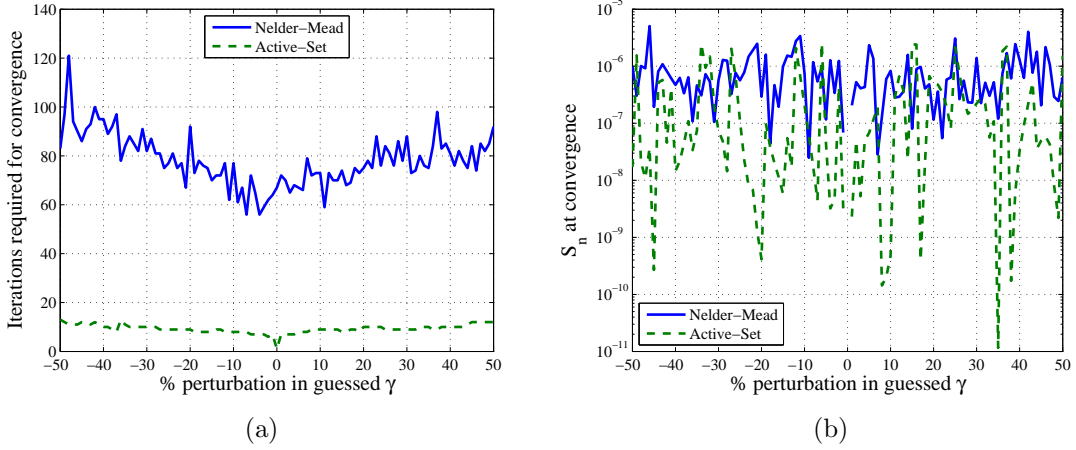


Figure 4.20: Effect of perturbing initial guess of γ on Nelder-Mead and Active-Set algorithms on (a) iterations required for convergence and (b) squared difference norm S_n between complex refractive index values at convergence and correct values shown in Figure 4.16. The algorithms converge to comparable solutions when initial guesses are perturbed from the correct value.

cases of ϵ_∞ and $\Delta\epsilon$, Figure 4.20(a) shows the active-set method is more efficient than the Nelder-Mead method. In contrast to the perturbation tests of the other two parameters however, the two algorithms are comparable in their ability to converge to the correct solution (global minimum) in the presence of poor initial guesses, as shown in Figure 4.20(b).

Finally, results of perturbing initial guesses of ω are shown in Figure 4.21. As before, the active-set method is more efficient than the Nelder-Mead method, as shown in Figure 4.21(a). Figure 4.21(b) shows the Nelder-Mead method converged to the correct solution (global minimum) for starting ω values from -18% to 15% of

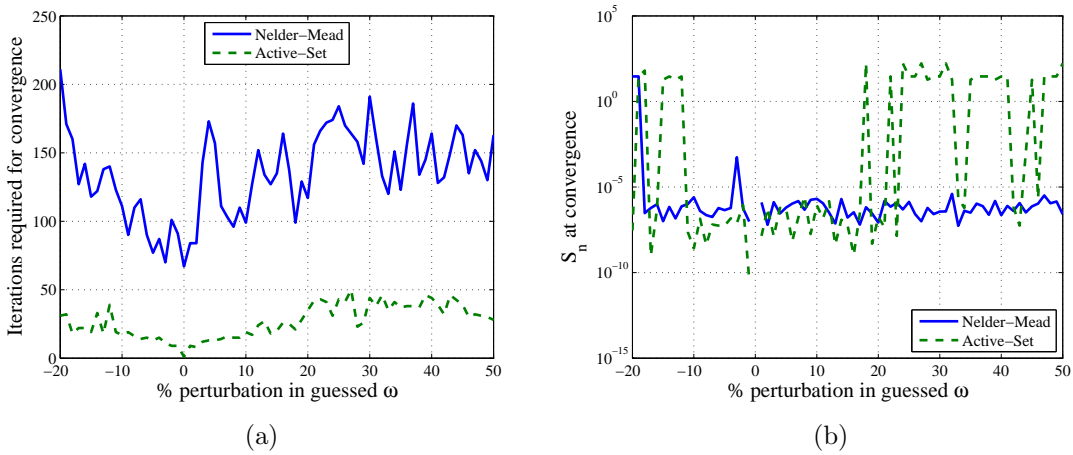


Figure 4.21: Effect of perturbing initial guess of ω on Nelder-Mead and Active-Set algorithms on (a) iterations required for convergence and (b) squared difference norm S_n between complex refractive index values at convergence and correct values shown in Figure 4.16. While less efficient, the Nelder-Mead algorithm converges consistently to the correct value (global minimum) for perturbations ranging from -18% to 15% of the correct value with a single exception at -3%, while the more efficient active-set method converges to the correct value for perturbations between -11% and 23% of the correct value with two exceptions at 18% and 22%. Outside these ranges, both algorithms are consistently thrown off by local minima.

the correct value with a single exception at -3% of the correct value. In contrast, the active set only converged consistently to the correct solution for perturbations between -11% and 23% of the correct value with two exceptions at 18% and 22%. This suggests the Nelder-Mead algorithm is more robust against poor initial guesses of the resonant frequency than the active-set algorithm.

The preceding results suggest that there is a tradeoff between efficiency and initial guess sensitivity between the two methods. The deciding factor for this work was the decreased sensitivity to poor initial guesses as provided by the Nelder-Mead algorithm. If give wholly incorrect guesses, both algorithms are likely to either not converge, or

converge to a local minimum. As mentioned before, a real-world system may draw initial guesses from a database of known values for a variety of materials of interest. This suggests a matched filter like approach to material identification in which the measured data set is tested using a range of initial guesses corresponding to several such materials. It should also be noted that the Nelder-Mead algorithm has been shown to become less efficient as the scale of the problem increases [30], suggesting that increasing the number of resonant frequencies will slow its convergence.

4.3.2 Effect of System Noise

Complex refractive index data for composition 4 (C4) explosive was generated using the Lorentz model (2.7) and the parameters from Yamamoto [16] shown in Table 4.11. The resulting refractive index and extinction coefficient curves are shown in

Mode	ϵ_∞	$\omega_p/2\pi$ (THz)	$\Delta\epsilon_p \times 10^3$	$\gamma_p/2\pi$ (GHz)
-	2.87	-	-	-
1	-	0.807	263	219
2	-	1.065	14.0	108
3	-	1.356	34.0	189
4	-	1.530	32.0	258
5	-	1.971	50.0	324
6	-	2.244	19.0	327

Table 4.11: Lorentz model parameters for explosive composition 4 (C4) from Yamamoto [16].

Figure 4.22. These refractive index and extinction coefficient curves then were used in (3.18) assuming a thickness of $d = 1.5$ mm to generate the idealized noiseless sample reflectance and phase curves shown in Figure 4.23. System noise was modeled in the

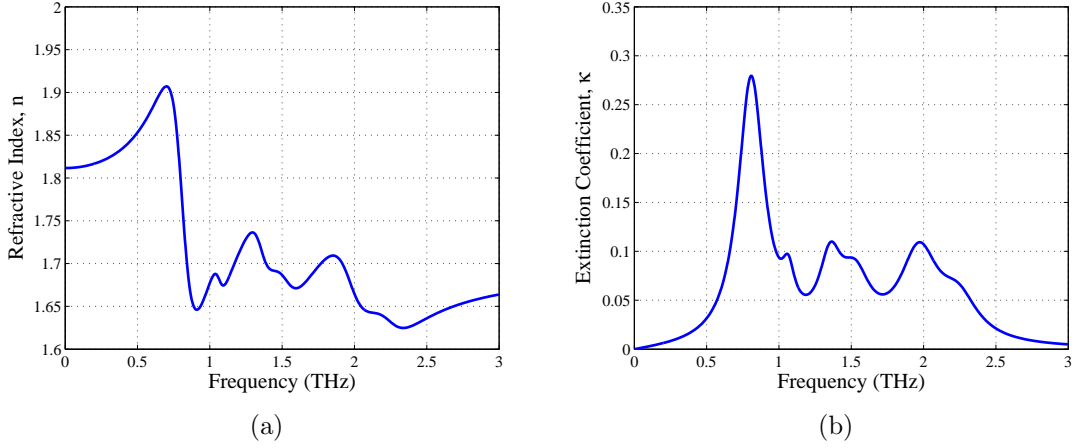


Figure 4.22: Refractive index (a) and extinction coefficient (b) corresponding to Lorentz parameters for C4 shown in Table 4.11 [16].

frequency domain as an ensemble of random phasors,

$$\tilde{a} = ae^{i\phi}, \quad (4.3)$$

where the amplitude a and phase ϕ are random variables with $a \sim \mathcal{N}(0, \sigma_a^2)$ and $\phi \sim \mathcal{U}(0, 2\pi)$. The standard deviation σ_a of the phasor amplitude was varied to adjust the SNR of the simulated data, relative to $E_{ref} = 1$ in (3.18), using

$$\begin{aligned} \text{SNR} &= -20 \log_{10} \left| \frac{\sigma_a}{E_{ref}} \right| \\ &= -20 \log_{10} |\sigma_a|. \end{aligned} \quad (4.4)$$

Differing levels of this system noise were then added to the noiseless curve shown in Figure 4.23 to generate simulated noisy data with adjustable SNR. The effect of

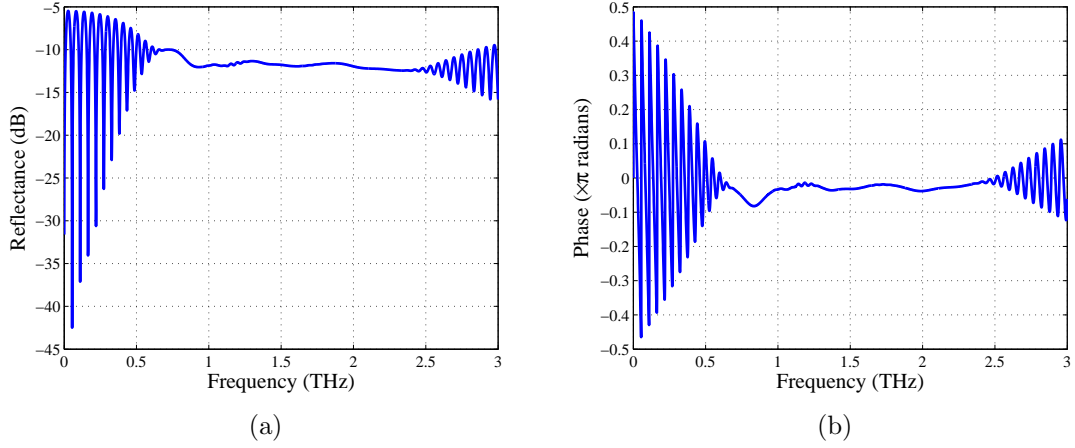


Figure 4.23: Calculated (a) reflectance and (b) unwrapped phase using assuming $d = 1.5$ mm in (2.15) and the refractive index and extinction coefficient values shown in Figure 4.22.

differing SNR on the simulated data is illustrated in Figure 4.24. In Figures 4.24(a), 4.24(b), and 4.24(c), 1000 realizations of simulated noisy data for each given SNR were generated. The ensemble mean μ_r and standard deviation σ_r of the reflectance were then calculated for each SNR value and used to generate the curves corresponding to the ensemble mean μ_r and confidence intervals $\mu_r \pm \sigma_r$ shown on a linear scale. Figure 4.24(d) shows the case of infinite SNR on the same axes for comparison.

The Nelder-Mead algorithm was then used to minimize S_r in (3.20) using rER_j^p from (3.24) for 200 realizations each of 20 dB, 40 dB, and 60 dB SNR to quantify the effect of noise on the estimated Lorentz parameters. Since the objective is to quantify the effect of noise on the recovered Lorentz parameters rather than inversion performance, the the correct C4 Lorentz parameters shown in Table 4.11 were used to initialize the Nelder-Mead algorithm. For each SNR level, the ensemble mean μ and

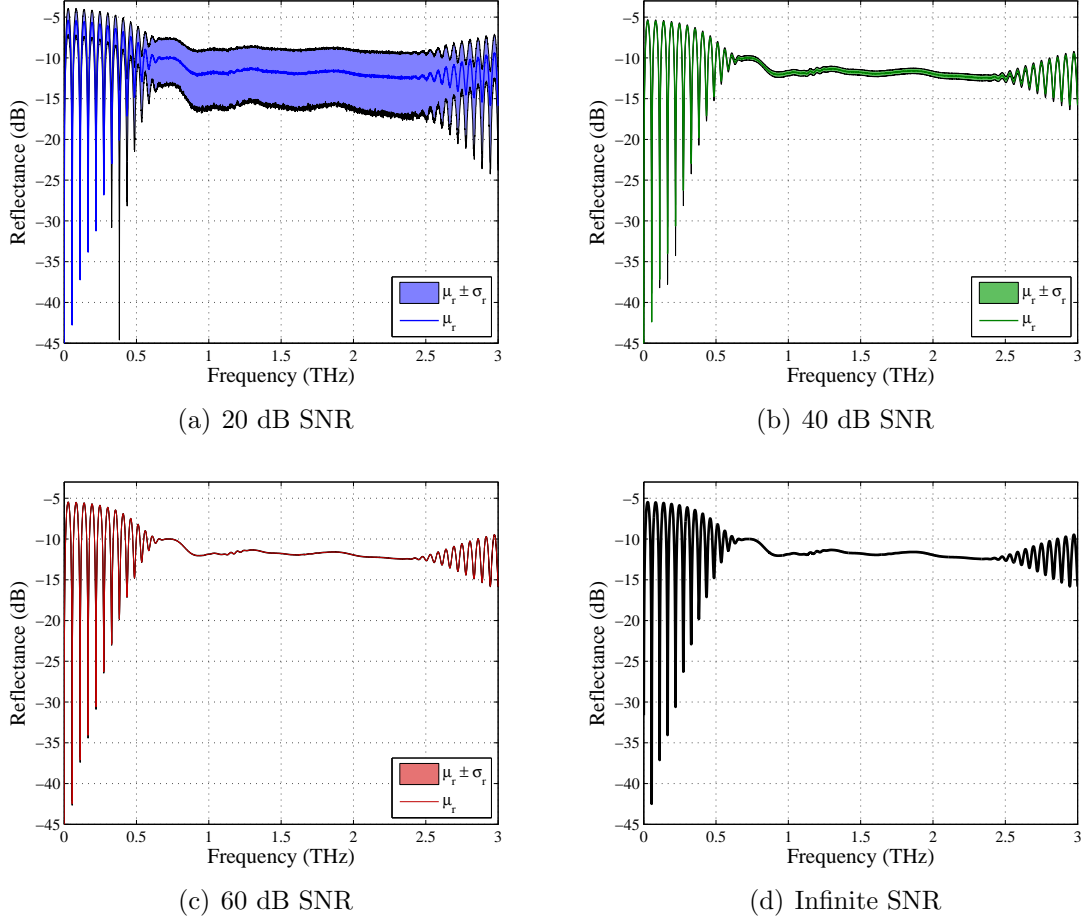


Figure 4.24: Effect of (a) 20 dB, (b) 40 dB, and (c) 60 dB SNR on mean μ_r and confidence interval $\mu_r \pm \sigma_r$ of reflectance based on an ensemble of 1000 noise realizations, and (d) the case of infinite SNR.

standard deviation σ of each estimated Lorentz parameter were calculated. To get an idea of the expected variability of each set of Lorentz parameters for each of C4's six molecular vibrational modes when subject to noise, the standard deviation was taken as a percentage of the mean before plotting in Figure 4.25. The results in Figures 4.25(a) and 4.25(d) suggest ϵ_∞ and ω_p , whose standard deviations are less than 0.1% and 6% of the mean, respectively, are much less vulnerable to the influence of

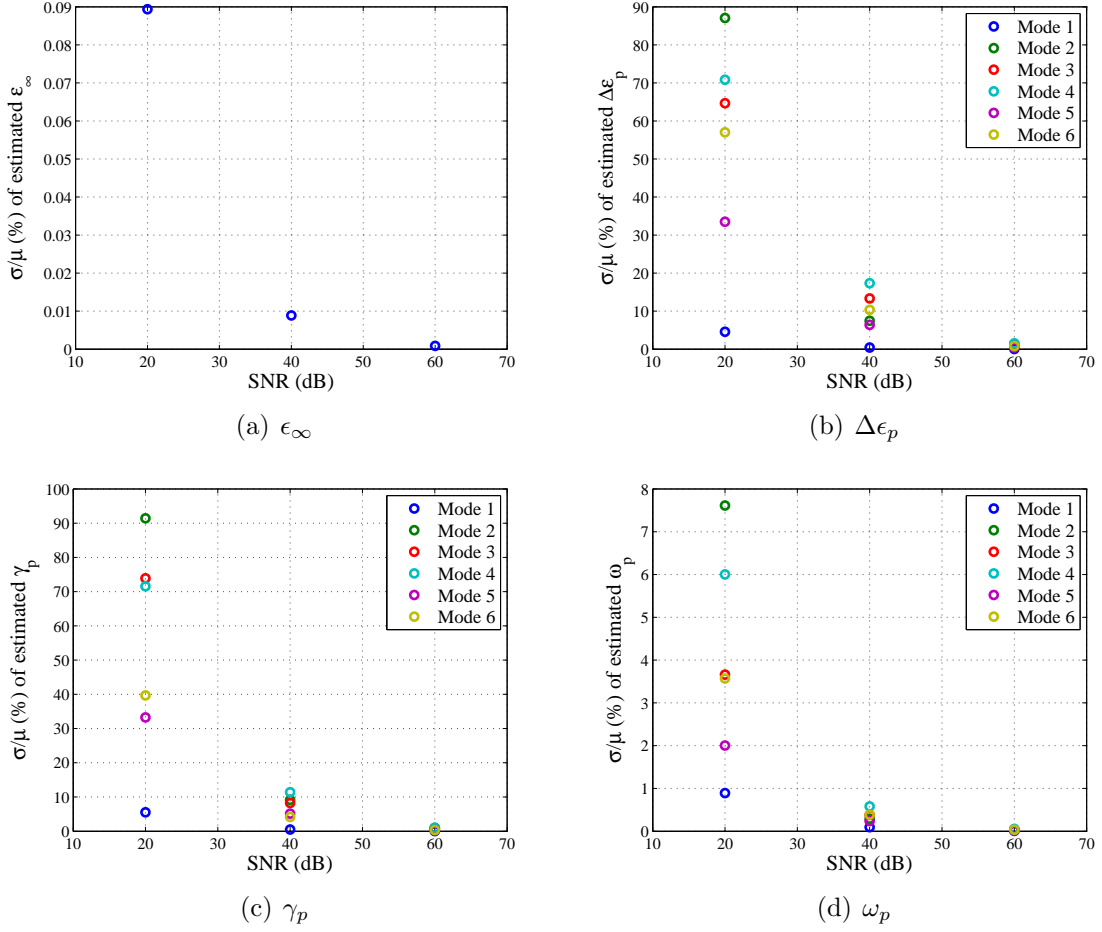


Figure 4.25: Effect of SNR on variance of estimated C4 Lorentz parameters as a percentage of the ensemble mean for all six molecular modes (see Table 4.11). (a) ϵ_∞ (b) $\Delta\epsilon_p$ (c) γ_p (d) ω_p

reduced SNR than $\Delta\epsilon_p$ and γ_p , as shown in Figures 4.25(b) and 4.25(c). Also noteworthy is the relatively low variability of the material parameters corresponding to modes 1 and 5 as compared with the other modes illustrated in Figures 4.25(d), 4.25(c), and 4.25(b). This is to be expected as modes 1 and 5 are the two strongest according to their $\Delta\epsilon_p$ values in Table 4.11, and most prominent according to Figure 4.22.

While better SNR will increase performance for any technique, these results suggest the inclusion of *a priori* information in the form of stored initial guess Lorentz parameters make the parametric approach effective in the presence of noise if given enough measurements to mitigate the effects of low SNR. In a real-world system, a tradeoff may be necessary between the number of measurements (increased measurement time) and convergence to a useful solution.

4.3.3 Effect of Thickness Error

The parametric approach of Section 3.3 potentially can be extended to include an optimization over sample thickness d as well as Lorentz parameters. In such an extension, the parameter vector $\boldsymbol{\theta}$ in (3.24) is modified so that it includes the thickness d , resulting in

$$\boldsymbol{\theta}' = [\epsilon_\infty, \omega_1, \omega_2, \dots, \omega_P, \Delta\epsilon_1, \Delta\epsilon_2, \dots, \Delta\epsilon_P, \gamma_1, \gamma_2, \dots, \gamma_P, d], \quad (4.5)$$

and changing rER_j^p in (3.24) to

$$rER_j^{p'} = |E_{data,j} - r_{eff}(f_j, \boldsymbol{\theta}')|. \quad (4.6)$$

While Ahmed [14] recently demonstrated this process for transmission mode, it has yet to be demonstrated in reflection mode. The feasibility of this process was assessed using the simulated reflection response shown in Figure 4.23 from a 1.5 mm

thick C4 sample as E_{data} in (4.6). The Nelder-Mead algorithm was used to solve (4.6) for a range of initial thickness guesses from -15% to 10% of the correct value, 1.5 mm. For each initial guess of the thickness, the percent difference between the estimated thickness value at convergence and the correct value was calculated as well as the squared difference norm S_n between the correct complex refractive index (shown in Figure 4.22) and the complex refractive index calculated from the estimated θ' at convergence. Results are shown in Figure 4.26. Figure 4.26(a) shows the algorithm

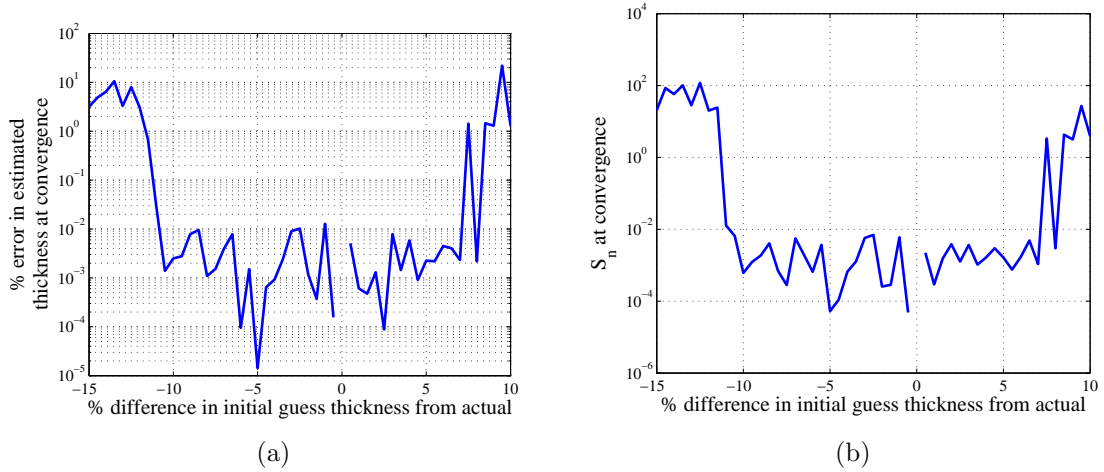


Figure 4.26: Results of solving (4.6) with Nelder-Mead algorithm starting from initial thickness guesses ranging from -15% to 10% of correct value. (a) Percent error between actual thickness and estimated thickness at convergence. (b) Squared difference norm S_n between correct complex refractive index and complex refractive index calculated from estimated θ' at convergence.

was able to correct initial guesses that deviate from the correct value by -11% to 7%, bringing them to less than 0.02% of the correct value. This is consistent with Figure 4.26(b), which shows complex refractive index values in this range are also close to the correct value. These results suggest a good initial guess will allow the

Nelder-Mead algorithm to optimize the sample thickness in reflection, eliminating the need for the use of the total variation method of Dorney [10].

4.3.4 Effect of Positioning Error

In addition to thickness error, the phase shift resulting from positioning error, ΔL in (3.19), can also be estimated parametrically by including ΔL in the modified parameter vector, resulting in

$$\boldsymbol{\theta}'' = [\epsilon_\infty, \omega_1, \omega_2, \dots, \omega_P, \Delta\epsilon_1, \Delta\epsilon_2, \dots, \Delta\epsilon_P, \gamma_1, \gamma_2, \dots, \gamma_P, \Delta L], \quad (4.7)$$

and changing rER_j^p in (3.24) to

$$rER_j^{p''} = |E_{data,j} - r_{eff}(f_j, \boldsymbol{\theta}'')e^{i2k_0(f_j)\Delta L(\boldsymbol{\theta}'')}|. \quad (4.8)$$

The feasibility of this process was assessed by introducing a phase shift corresponding to $\Delta L = 5 \mu\text{m}$ into the simulated reflection response shown in Figure 4.23, resulting in the simulated data shown in Figure 4.27. The Nelder-Mead algorithm was then used to solve (4.8) for a range of initial guesses for ΔL . For each initial guess, the percent difference between the estimated thickness value at convergence and the correct value was calculated as well as the squared difference norm S_n between the correct complex refractive index (shown in Figure 4.22)) and the complex refractive index calculated from the estimated $\boldsymbol{\theta}''$ at convergence. Results are shown

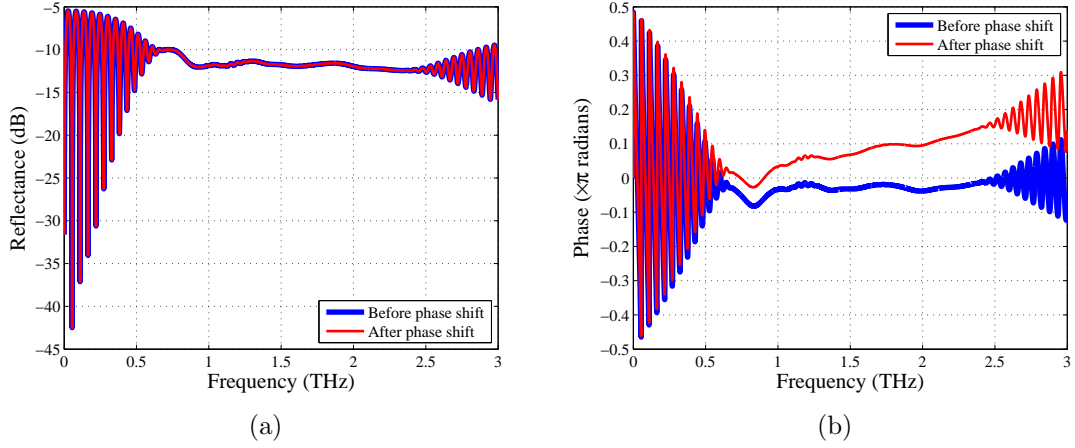


Figure 4.27: Calculated (a) reflectance and (b) unwrapped phase of simulated data from (3.19) in red, shown with unaltered data from Figure 4.23 in blue for comparison.

in Figure 4.28. Figure 4.28(a) shows the algorithm was able to reliably correct initial

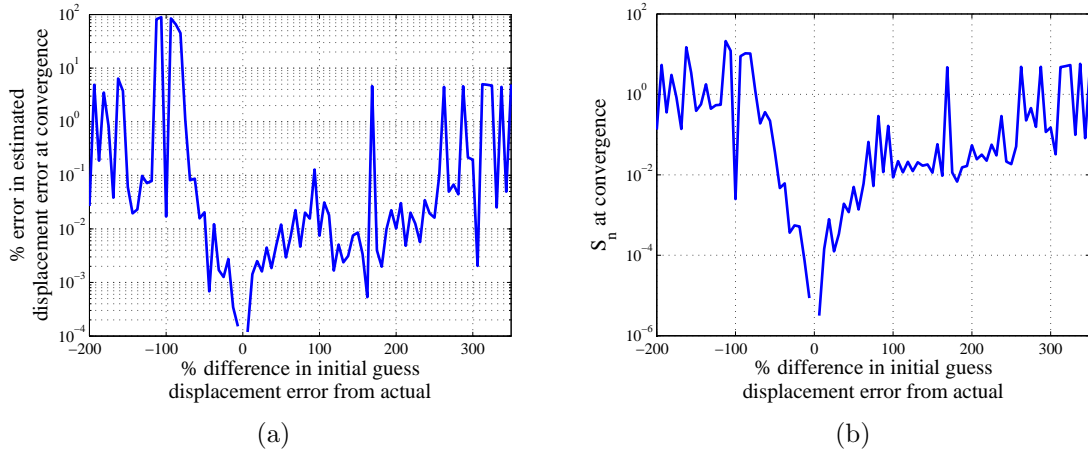


Figure 4.28: Results of solving (4.8) with Nelder-Mead algorithm starting from initial guesses for ΔL ranging from -200% to 350% of correct value. (a) Percent error between estimated thickness and actual thickness at convergence. (b) Squared difference norm S_n between actual complex refractive index and complex refractive index calculated from θ'' at convergence.

guesses to within 2% of the actual value from about -69% to about 162.5% of the correct value. However, Figure 4.28(b) indicates the estimated refractive index is more

sensitive to bad initial guesses of the displacement error. The squared difference norm S_n between the estimated and correct values was found to be below 0.01 a narrower range of initial guesses; only -37.5% to 62.5% of the correct value.

4.4 Discussion

The results in this chapter suggest that the parametric approach to material parameter estimation may be a useful tool for material identification and/or classification from reflection measurements of thin samples. The parametric method used in transmission was shown to compare well to results from a conventional non-parametric method. Applying the method in reflection yielded useful results, in that the method revealed the sample's most prominent resonant frequencies to within a few tenths of one percent. This was found to be consistent with results of sensitivity testing, which suggest ϵ_∞ and ω_p are the Lorentz parameters with the least susceptibility to the influence of measurement noise. Sensitivity testing also revealed the potential of the parametric method for simultaneously estimating layer thickness and misplacement error between sample and reference.

Chapter 5

Conclusions

Hailed as a new frontier of electromagnetics research, terahertz spectroscopy has gained recent attention for its potential applications in security screening, biomedical sensing, and nondestructive evaluation. While THz spectral “fingerprints” can be detected in both transmission and reflection modes, the opacity of many targets, such as human passengers in the case of airport screening, limit future THz scanners to reflection geometry. In practical scenarios, the spectral response may be affected by such uncontrollable factors as rough surface and volume scattering. Another complication worth considering is the case of layered structures with multiple boundaries, such as a target material under one or more layers of clothing or packaging. While such materials are usually transparent, if they present a significant impedance mismatch between the air and target material, the layers may cause multiple reflections within them that complicate the detected spectra by introducing interference patterns in the frequency domain due to the etalon effect. Most work to date in identifying layered materials from reflection measurements has ignored the etalon effect, using samples with layers that were sufficiently optically thick that reflected THz pulses did not overlap in the time domain. However, little attention has been paid to the case of optically thin layers, from which reflected THz pulses cannot be separated in time and for which the etalon effect cannot be ignored.

This thesis investigated techniques commonly used to estimate material parameters from transmission measurements of optically thin samples that take the etalon effect into account and adapted them for use in reflection geometry. The Lorentz oscillator model of classical dispersion was first introduced and derived, as were the effective reflection and transmission coefficients from a single layer of dielectric material between two semi-infinite half spaces of air and another dielectric. Conventional non-parametric material parameter estimation algorithms based on these effective reflection and transmission coefficients were introduced and discussed before introducing a parametric technique based on these coefficients and the Lorentz dispersion model.

Both the non-parametric and parametric techniques were implemented in MATLAB and tested using data measured using the Picometrix T-Ray 4000 time-domain spectroscopy system at Portland State University's Northwest Electromagnetics Research Laboratory (NEAR-Lab). Transmission measurements of a lactose sample were first presented and its complex refractive index was estimated from the data using a conventional non-parametric inversion technique from the literature. Results from this non-parametric technique were used to validate the parametric inversion technique in transmission. Reflection measurements were then presented and results from the parametric technique applied in reflection mode were compared to the validated results from the transmission mode data. The spectral locations of the three

most prominent modes of the sample material were found to be most consistent between the transmission and reflection results, suggesting the potential utility of the parametric method in material classification.

As the parametric technique requires the use of a numerical optimization routine, two such routines were compared in their efficiency and sensitivity to initial estimates for the Lorentz parameters required for algorithm initialization. Next, the sensitivity of the parametric technique to measurement noise was tested using simulated noisy data. Results of these analyses suggest the spectral locations of strong resonant modes are least susceptible to the influence of measurement noise. The parametric technique was then evaluated on its ability to correct for the effects of imprecise knowledge of sample thickness and positioning error between sample and reference measurements - two common sources of error.

This work established the utility of parametric material parameter estimation techniques for reflection spectroscopy and suggests their potential for future use in material classification algorithms based on estimated Lorentz parameters.

Findings from this research were published in the Proceedings of SPIE - The International Society for Optical Engineering [13], and were presented at the SPIE's annual Defense, Security, and Sensing conference in Orlando, Florida in April, 2010.

5.1 Future Work

Future work in this area will address the case in which a target material is covered by one or more layers of an optically thin barrier material – a situation likely to be encountered in real life scenarios. This section shows preliminary measurement results illustrating the complication of the reflected spectrum due to a single thin layer.

Vinyl electrical tape was chosen as the material for the layer as it is of practical importance and it fits the assumptions of a homogeneous medium with parallel planar boundaries made in the model in Section 2.2. The layer of adhesive on the tape was assumed to be thin enough to have a negligible effect on the measurement while providing a mechanically stable, airtight boundary between the layer and target material, as assumed in the model. The thickness of the tape was taken to be $1/3$ of the thickness (measured with a micrometer) of a sample consisting of three layers of the tape. In preparing the sample, care was taken not to change the layer thickness by stretching the tape or introduce scatterers by allowing pockets of air or other noticeable particles to be trapped between the tape layers before measurement. The sample was measured at three points along the outside edge and one point in the center. These thickness values were $d = 0.518$ mm, 0.519 mm, 0.527 mm, and 0.527 mm, respectively, for a mean thickness $\mu_d = 0.523$ mm and standard deviation of $\sigma_d = 0.005$ mm. This corresponds to a thickness of $d = 0.174$ mm for a single layer of tape.

A thick sample of target material is required for a good approximation to a semi-infinite half-space. Such a thick sample was prepared by mixing lactose 10% (w/w) in extra fine PE powder and pressing 5 g of the mixture into a 4 cm diameter pellet at 10 tons of pressure using a hydraulic press. The lactose powder was first sifted using a standard No. 270 sieve to ensure the particles were smaller than $53\text{ }\mu\text{m}$ before mixing and pressing to reduce volume scattering within the pellet. A micrometer was then used to measure the pellet at three points along the outside edge and one point in the center. These thickness values were $d = 4.006\text{ mm}$, 4.005 mm , 4.026 mm , and 4.013 mm , respectively, for a mean thickness $\mu_d = 4.013\text{ mm}$ and standard deviation of $\sigma_d = 0.010\text{ mm}$.

Reflection measurements were taken of the 10% lactose sample covered by a single layer of tape, two layers of tape, and no tape at normal incidence using the T-Ray 4000's collinear head equipped with a 6" focusing lens. The time domain waveforms are shown in Figure 5.1. The etalon reflections from within the tape are nearly invisible in the waveform from the single layer sample. They are only revealed in the slight difference in shape of the waveforms in the approximately 8 ps following the first reflected pulse. The waveform from the sample covered with two layers of tape clearly shows the influence of the thicker 0.3486 mm tape layer, but even so, the pulses overlap to such a degree that they cannot be separated by time domain windowing.

To study the effect of the tape layer on the returned spectra, the waveforms were

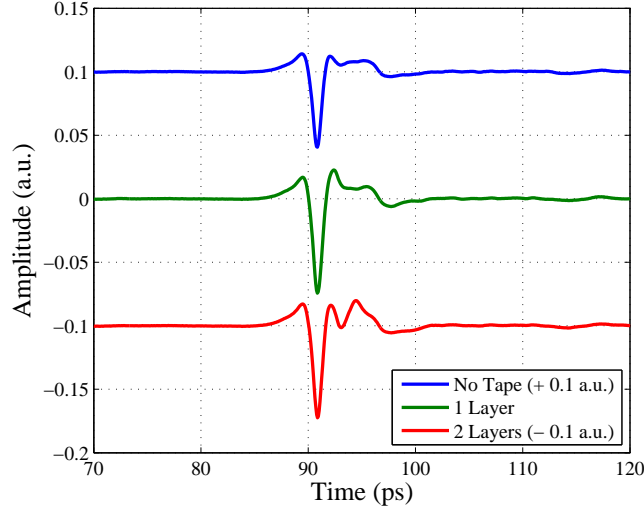


Figure 5.1: Time domain reflection waveforms of 10% lactose sample with no tape compared with waveforms from the same sample covered with a single layer of tape and two layers of tape. Waveforms are offset vertically by 0.1 a.u. for clarity.

time gated to exclude the reflection from the back surface of the sample (not shown in Figure 5.1) before being Fourier transformed and deconvolved with a reference measurement as described in Section 3.2.2. Results are shown in Figure 5.2. The reflectance curves from the pellets with tape show oscillatory behavior consistent with the etalon effect. The spacing of the oscillations in the frequency domain should be inversely proportional to the spacing between pulses in the time domain [22], which suggests the spectrum from the sample with the thicker tape layer should have more closely spaced oscillations than that from the sample with the thinner tape layer. This is consistent with the results in Figure 5.2.

These results demonstrate that optically thin layers covering materials of interest are likely to be detrimental to methods of material feature estimation in reflection

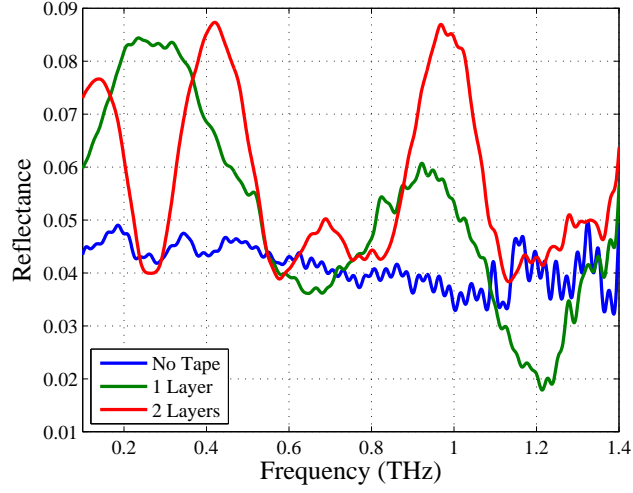


Figure 5.2: Reflectance spectra of 10% lactose sample with no tape compared to spectra from the same sample covered with a single layer of tape and two layers of tape. Reflections from the back surface of the sample shown in Figure 5.1 were excluded from the Fourier transform windows.

mode commonly found in the literature that rely on derivatives of measured magnitude or phase [4, 5, 7, 6, 25]. Model-based techniques such as the parametric approach explored in this thesis may eventually prove useful in the future development of practical, deployable systems for THz spectroscopic sensing.

References

- [1] S.P. Micken and X.-C. Zhang. T-ray sensing and imaging. *International Journal of High Speed Electronics and Systems*, 13(2):601 – 676, 2003.
- [2] M.C. Kemp, P.F. Taday, B.E. Cole, J.A. Cluff, A.J. Fitzgerald, and W.R. Tribe. Security applications of terahertz technology. volume 5070, pages 44 – 52, Orlando, FL, United states, 2003.
- [3] Jian Chen, Yunqing Chen, Hongwei Zhao, Glenn J. Bastiaans, and X.-C. Zhang. Absorption coefficients of selected explosives and related compounds in the range of 0.1-2.8 thz. *Optics Express*, 15(19):12060 – 12067, 2007.
- [4] Yaochun Shen, Philip F. Taday, and Michael C. Kemp. Terahertz spectroscopy of explosive materials. *Proceedings of SPIE - The International Society for Optical Engineering*, 5619:82 – 89, 2004.
- [5] Y.C. Shen, P.F. Taday, D.A. Newnham, and M. Pepper. Chemical mapping using reflection terahertz pulsed imaging. *Semiconductor Science and Technology*, 20(7):S254 – S257, 2005.

- [6] Hua Zhong, Albert Redo-Sanchez, and X.-C. Zhang. Identification and classification of chemicals using terahertz reflective spectroscopic focalplane imaging system. *Optics Express*, 14(20):9130 – 9141, 2006.
- [7] Yao-Chun Shen, Philip F. Taday, David A. Newnham, Michael C. Kemp, and Michael Pepper. 3d chemical mapping using terahertz pulsed imaging. *Proceedings of SPIE - The International Society for Optical Engineering*, 5727:24 – 31, 2005.
- [8] Lionel Duvillaret, Frederic Garet, and Jean-Louis Coutaz. Reliable method for extraction of material parameters in terahertz time-domain spectroscopy. *IEEE Journal on Selected Topics in Quantum Electronics*, 2(3):739 – 746, 1996.
- [9] Lionel Duvillaret, Frederic Garet, and Jean-Louis Coutaz. Highly precise determination of optical constants and sample thickness in terahertz time-domain spectroscopy. *Applied Optics*, 38(2):409 – 415, 1999.
- [10] Timothy D. Dorney, Richard G. Baraniuk, and Daniel M. Mittleman. Material parameter estimation with terahertz time-domain spectroscopy. *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, 18(7):1562 – 1571, 2001.
- [11] Ioachim Pupeza, Rafal Wilk, and Martin Koch. Highly accurate optical material parameter determination with thz time-domain spectroscopy. *Optics Express*, 15(7):4335 – 4350, 2007.

- [12] Max Born and Emil Wolf. *Principles of Optics*. Pergamon Press, 1980.
- [13] G.P. Kniffin, S. Schecklman, J. Chen, S.C. Henry, L.M. Zurk, B. Pejcinovic, and A.I. Timchenko. Measurement and modeling of terahertz spectral signatures from layered material. *Proceedings of SPIE - The International Society for Optical Engineering*, 7687:The Society of Photo-Optical Instrumentation Engineers (SPIE) –, 2010.
- [14] Osman S. Ahmed, Mohamed A. Swillam, Mohamed H. Bakr, and Xun Li. Efficient optimization approach for accurate parameter extraction with terahertz time-domain spectroscopy. *Journal of Lightwave Technology*, 28(11):1685 – 1692, 2010.
- [15] Constantine A. Balanis. *Advanced Engineering Electromagnetics*. John Wiley and Sons, 1989.
- [16] Kohji Yamamoto, Mariko Yamaguchi, Fumiaki Miyamaru, Masahiko Tani, Masanori Hangyo, Takeshi Ikeda, Akira Matsushita, Kenji Koide, Michiaki Tatsuno, and Yukio Minami. Noninvasive inspection of c-4 explosive in mails by terahertz time-domain spectroscopy. *Japanese Journal of Applied Physics, Part 2: Letters*, 43(3 B):L414 – L417, 2004.
- [17] Jon E. Bjarnason, Elliott R. Brown, and Timothy M. Korter. Comparison of the thz absorption feature in lactose to related saccharides. *Proceedings of SPIE - The International Society for Optical Engineering*, 6549:SPIE –, 2007.

- [18] A. Roggenbuck, H. Schmitz, A. Deninger, I. Camara Camara Mayorga, J. Hemberger, R. Gusten, and M. Gruninger. Coherent broadband continuous-wave terahertz spectroscopy on solid-state samples. *New Journal of Physics*, 12, 2010.
- [19] Mark Fox. *Optical Properties of Solids*. Oxford University Press, 2 edition, 2010.
- [20] Yunqing Chen, Haibo Liu, Yanqing Deng, Dunja Schauki, Michael J. Fitch, Robert Osiander, Caroline Dodson, James B. Spicer, Michael Shur, and X.-C. Zhang. Thz spectroscopic investigation of 2,4-dinitrotoluene. *Chemical Physics Letters*, 400:357–361, 2004.
- [21] Withawat Withayachumnankul, Bernd M. Fischer, and Derek Abbott. Material thickness optimization for transmission-mode terahertz time-domain spectroscopy. *Optics Express*, 16(10):7382 – 7396, 2008.
- [22] Maik Scheller, Christian Jansen, and Martin Koch. Analyzing sub-100-m samples with transmission terahertz time domain spectroscopy. *Optics Communications*, 282(7):1304 – 1306, 2009.
- [23] Simon Haykin. *Adaptive Filter Theory*. Prentice-Hall, 1996.
- [24] A. Pashkin, M. Kempa, H. Nemec, F. Kadlec, and P. Kuzel. Phase-sensitive time-domain terahertz reflection spectroscopy. *Review of Scientific Instruments*, 74(11):4711 – 4717, 2003.

- [25] Hua Zhong, Cunlin Zhang, Liangliang Zhang, Yuejin Zhao, and Xi-Cheng Zhang. A phase feature extraction technique for terahertz reflection spectroscopy. *Applied Physics Letters*, 92(22), 2008.
- [26] John Wilkinson and Mike Kemp. Thz simulants for explosives. Technical report, NATO SET-124 Research Task Group, 2009.
- [27] Christopher Konek, John Wilkinson, Okan Esenturk, Edwin Heilweil, and Michael Kemp. Terahertz spectroscopy of explosives and simulants - rdx, petn, sugar and l-tartaric acid. *Proceedings of SPIE - The International Society for Optical Engineering*, 7311:The International Society for Optical Engineering (SPIE) —, 2009.
- [28] J. Van Rudd and Daniel M. Mittleman. Influence of substrate-lens design in terahertz time-domain spectroscopy. *Journal of the Optical Society of America B: Optical Physics*, 19(2):319 – 329, 2002.
- [29] B. Fischer, M. Hoffmann, H. Helm, G. Modjesch, and P. Uhd Jepsen. Chemical recognition in terahertz time-domain spectroscopy and imaging. *Semiconductor Science and Technology*, 20(7):S246 – S253, 2005.
- [30] Lixing Han and Michael Neumann. Effect of dimensionality on the nelder-mead simplex method. *Optimization Methods and Software*, 21(1):1 – 16, 2006.
- [31] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.

- [32] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical recipes: the art of scientific computing*. Cambridge University Press, 3 edition, 2007.
- [33] J.C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9:112–147, 1998.

Appendix A

Derivation of Lorentz Dispersion Model

Derivations of the Lorentz model can be found in [12], [15], and [19]. This derivation follows closely that presented in Fox [19]. The derivation starts with an atom modeled as a simple electric dipole consisting of a positively charged nucleus with mass M and a negatively charged electron with mass m held together with a restorative force equal to $m_r\omega_0^2$, where the reduced mass m_r is given by

$$\frac{1}{m_r} = \frac{1}{m} + \frac{1}{M}.$$

Since $m \ll M$, the reduced mass can be approximated as $m_r \cong m$.

The motion of a Lorentz dipole oscillator in response to an applied alternating electric field with amplitude E can then be modeled as that of a damped harmonic oscillator,

$$m \frac{d^2x}{dt^2} + m\gamma \frac{dx}{dt} + m\omega_0^2 x = -qE, \quad (\text{A.1})$$

where x is the displacement of the electron from its equilibrium position, ω_0 is the resonant frequency, γ is the damping coefficient, q is the electric charge of an electron. Substituting the time-harmonic solutions $x(t) = X_0 e^{i\omega t}$ and $E = E_0 e^{i\omega t}$ into (A.1) yields

$$X_0 = \frac{-qE_0/m}{\omega_0^2 - \omega^2 - i\gamma\omega}. \quad (\text{A.2})$$

This time-harmonic motion of charge induces a time-varying dipole moment $p(t) = -qx(t)$ in the atom. For a bulk material made up of N atoms per unit volume, the resonant polarization amplitude P becomes

$$\begin{aligned}
P &= Np \\
&= -Nqx \\
&= -NqX_0e^{i\omega t} \\
&= \frac{Nq^2}{m} \frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega} E.
\end{aligned} \tag{A.3}$$

This amplitude is that of the the resonant polarization described by the vector \mathbf{P} . The electric flux density of a material in response to an applied electric field \mathbf{E} is then given in terms of \mathbf{P} by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}. \tag{A.4}$$

In (A.4), \mathbf{P} can be expressed in as a sum of a baseline component \mathbf{P}_b and a component due to resonance \mathbf{P}_r , which results in

$$\begin{aligned}
\mathbf{D} &= \epsilon_0 \mathbf{E} + \mathbf{P}_b + \mathbf{P}_r \\
&= \epsilon_0 \mathbf{E} + \epsilon_0 \chi \mathbf{E} + \mathbf{P}_r \\
&= \epsilon_0 (1 + \chi) \mathbf{E} + \mathbf{P}_r,
\end{aligned} \tag{A.5}$$

where χ is the electric susceptibility. Combining (A.5) with

$$\mathbf{D} = \epsilon_0 \epsilon_r \mathbf{E} \quad (\text{A.6})$$

and (A.3) gives an expression for the effective frequency dependent dielectric constant,

$$\epsilon_r(\omega) = 1 + \chi + \frac{Nq^2}{\epsilon_0 m} \frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega}. \quad (\text{A.7})$$

Using the low frequency (DC) limit,

$$\epsilon_{DC} \equiv \lim_{\omega \rightarrow 0} \epsilon_r(\omega) = 1 + \chi + \frac{Nq^2}{\epsilon_0 m \omega_0^2}, \quad (\text{A.8})$$

and the high frequency limit,

$$\epsilon_\infty \equiv \lim_{\omega \rightarrow \infty} \epsilon_r(\omega) = 1 + \chi, \quad (\text{A.9})$$

allows the simplification of (A.7) as

$$\epsilon_r(\omega) = \epsilon_\infty + \frac{\Delta\epsilon\omega_0^2}{\omega_0^2 - \omega^2 - i\gamma\omega}, \quad (\text{A.10})$$

where $\Delta\epsilon = \epsilon_{DC} - \epsilon_\infty$.

In general, a material may contain several such resonances. Assuming the resonant modes are uncoupled, the contribution to the complex dielectric constant from each

oscillator can be summed, yielding the final expression, (2.7), in Section 2.1.1, for the case of P resonances.

Appendix B

Nelder-Mead Simplex Algorithm

The algorithm developed in 1965 by Nelder and Mead [31] for unconstrained nonlinear optimization has gone through several stages of development since its inception. It is sometimes also referred to as the “Amoeba Method” [32]. The refined version of Lagarias [33] is implemented natively as the function `fminsearch.m` in MATLAB and is presented here.

The Nelder-Mead method is a simple derivative-free, direct search optimization method for solving the problem,

$$\min_{\boldsymbol{\theta}} f(\boldsymbol{\theta}), \quad (\text{B.1})$$

where $\boldsymbol{\theta} \in \mathbb{R}^n$ and $f(\boldsymbol{\theta}) \in \mathbb{R}$. At the each iteration, a simplex consisting of $n + 1$ vertices, P_1, P_2, \dots, P_{n+1} , in \mathbb{R}^n is generated and the objective function $f(P_i)$, where $1 \leq i \leq n + 1$, is evaluated at each vertex. The function values are then sorted and vertices renumbered such that

$$f(P_1) \leq f(P_2) \leq \dots \leq f(P_{n+1}). \quad (\text{B.2})$$

As the objective function is to be minimized, the point P_1 and function value $f(P_1)$ constitute the ‘best’ point and function value, respectively, for that iteration. Similarly, the P_{n+1} and $f(P_{n+1})$ constitute the ‘worst’ point and function value, and P_n

and $f(P_n)$ constitute the ‘next-worst’ point and function value, for the iteration. The next iteration generates a new and distinct simplex by either replacing the worst point, P_{n+1} with a new point using one of three operations – reflection, expansion, or contraction – or replacing all but the best point P_1 via a shrinkage operation.

The initial simplex is generated using an initial guess P_0 from which the other n points P_i of the simplex are generated using [32]

$$P_i = P_0 + \Delta \hat{a}_i, \quad (\text{B.3})$$

where Δ is a small deviation from P_0 in the directions specified by the n unit vectors \hat{a}_i in \mathbb{R}^n . In `fminsearch.m`,

$$\Delta = \begin{cases} 0.00025 & : P_0 = 0 \\ 0.05 P_0 \hat{a}_i^T & : P_0 \neq 0. \end{cases}$$

Each iteration begins with the evaluation of the objective function at the $n + 1$ points of the simplex, followed by the sorting and renaming of the vertices according to (B.2). After this initial sort operation, the centroid \bar{P} of the first n points (all but the worst point) is calculated using

$$\bar{P} = \frac{1}{n} \sum_{i=1}^n P_i. \quad (\text{B.4})$$

The algorithm then begins with a reflection in which a new point P_r is generated by

reflecting the worst point P_{n+1} about the centroid \bar{P} as illustrated in Figure B.1 for the 2D case. The distance $\|\bar{P}P_r\|$ along the line $\overline{P_{n+1}\bar{P}}$ in \mathbb{R}^n is set by the coefficient

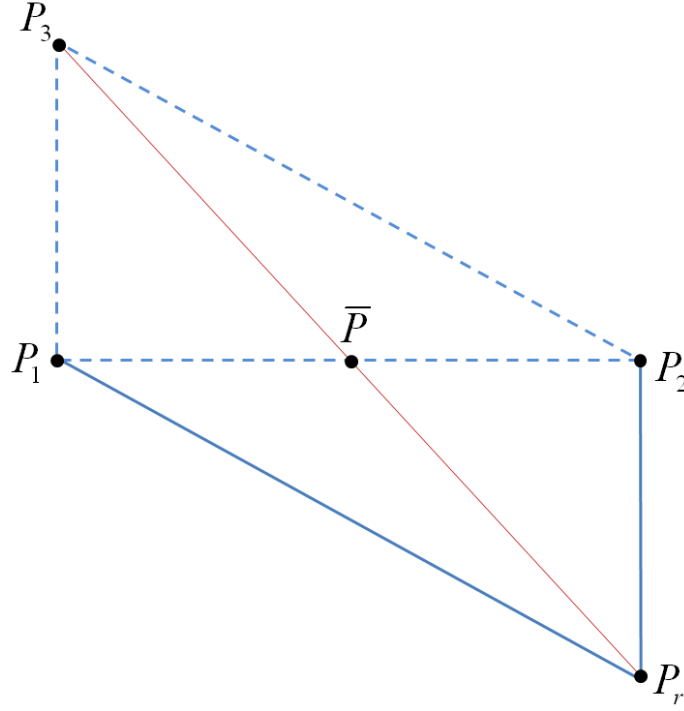


Figure B.1: Nelder-Mead reflection operation illustrated for the 2D case ($n = 2$). The new point P_r is generated by reflecting $P_{n+1} = P_3$ about the centroid \bar{P} . Original simplex indicated by dashed lines.

of reflection, ρ according to

$$\rho = \frac{\|\bar{P}P_r\|}{\|\overline{P_{n+1}\bar{P}}\|}. \quad (\text{B.5})$$

In general, $\rho > 0$, and the standard value is $\rho = 1$. Using ρ , \bar{P} , and P_{n+1} , the point

P_r is generated as follows;

$$\begin{aligned} P_r &= \bar{P} + \rho (\bar{P} - P_{n+1}) \\ &= (1 + \rho) \bar{P} - \rho P_{n+1}. \end{aligned} \tag{B.6}$$

If the function value at the reflected point, $f(P_r)$, is found to be between that at the best point, $f(P_1)$, and at the next-worst point, $f(P_n)$; that is, if

$$f(P_1) \leq f(P_r) < f(P_n), \tag{B.7}$$

the worst point P_{n+1} is replaced with the reflected point, the current iteration is terminated, and the next iteration begins. If (B.7) is not satisfied, then either

$$f(P_r) < f(P_1) \tag{B.8}$$

or

$$f(P_n) \leq f(P_r). \tag{B.9}$$

If (B.8) is satisfied, an expansion operation is performed in which another new point P_e is generated on the line $\overline{P_{n+1}\bar{P}}$ in \mathbb{R}^n at a distance $\|\overline{PP_e}\| > \|\overline{PP_r}\|$ as illustrated in Figure B.2 for the 2D case. The distance $\|\overline{PP_e}\|$ relative to $\|\overline{PP_r}\|$ is

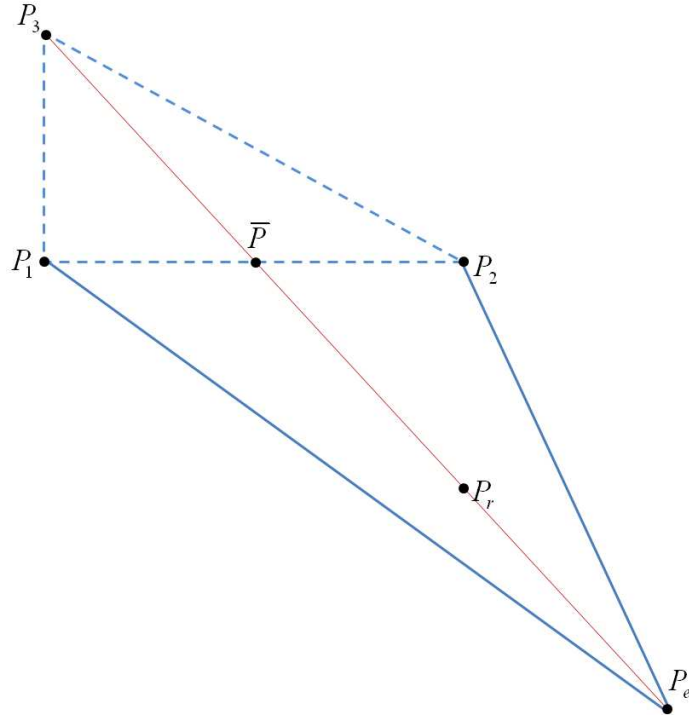


Figure B.2: Nelder-Mead expansion operation illustrated for the 2D case ($n = 2$). Original simplex indicated by dashed lines.

set by the coefficient of expansion, χ according to

$$\begin{aligned}
 \chi &= \frac{\|\overline{P}P_e\|}{\|\overline{P}P_r\|} \\
 &= \frac{\|\overline{P}P_e\|}{\rho \|\overline{P}_{n+1}\overline{P}\|}.
 \end{aligned} \tag{B.10}$$

In general, $\chi > 1$, and a standard value is $\chi = 2$. Using χ and ρ , the point P_e is

generated by

$$\begin{aligned}
P_e &= \bar{P} + \chi (P_r - \bar{P}) \\
&= \bar{P} + \rho\chi (\bar{P} - P_{n+1}) \\
&= (1 + \rho\chi) \bar{P} - \rho\chi P_{n+1}.
\end{aligned} \tag{B.11}$$

The function at this expansion point $f(P_e)$ is then evaluated and compared to the function value at the reflected point $f(P_r)$. If

$$f(P_e) < f(P_r), \tag{B.12}$$

the expansion point is chosen over the reflection point, otherwise the reflection point is chosen. The chosen point then replaces the worst point P_{n+1} , the current iteration terminates, and the next iteration begins.

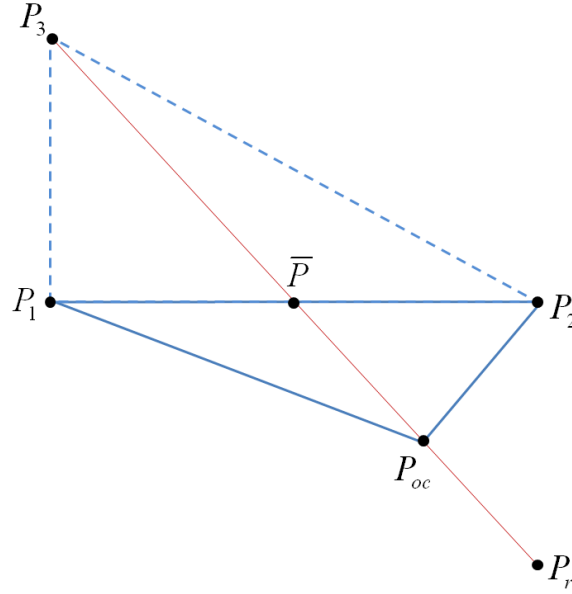
If (B.8) is not satisfied then (B.9) is the case and the function evaluated at the reflected point $f(P_r)$ either lies between that of the worst and next-worst points,

$$f(P_n) \leq f(P_r) < f(P_{n+1}), \tag{B.13}$$

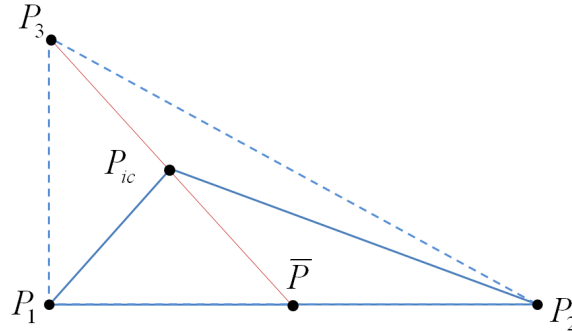
or it lies above the worst point,

$$f(P_{n+1}) \leq f(P_r). \tag{B.14}$$

In either case, a contraction operation is performed. The two possible contraction operations are illustrated in Figure B.3 for the 2D case. If (B.13) is satisfied, an



(a)



(b)

Figure B.3: Nelder-Mead contraction operations illustrated for the 2D case ($n = 2$). (a) Outside contraction, in which the new point P_{oc} is generated between \bar{P} and P_r along the line $\overline{P_{n+1}\bar{P}}$. (b) Inside contraction, in which the new point P_{ic} is generated between P_{n+1} and \bar{P} along the line $\overline{P_{n+1}\bar{P}}$. Original simplices indicated by dashed lines.

outside contraction is carried out in which a new point P_{oc} is generated on the line $\overline{P_{n+1}\bar{P}}$ between \bar{P} and P_r as shown in Figure B.3(a). The distance $\left\| \overline{\bar{P}P_{oc}} \right\|$ relative to

the distance $\left\| \overline{\bar{P}P_r} \right\|$ along the line $\overline{P_{n+1}\bar{P}}$ is set by the the coefficient of contraction, γ according to

$$\begin{aligned}\gamma &= \frac{\left\| \overline{\bar{P}P_{oc}} \right\|}{\left\| \overline{\bar{P}P_r} \right\|} \\ &= \frac{\left\| \overline{\bar{P}P_{oc}} \right\|}{\rho \left\| \overline{P_{n+1}\bar{P}} \right\|}.\end{aligned}\tag{B.15}$$

In general, $0 < \gamma < 1$, and a standard value is $\gamma = \frac{1}{2}$. Using γ and ρ , the point P_{oc} is generated by

$$\begin{aligned}P_{oc} &= \bar{P} + \gamma (P_r - \bar{P}) \\ &= \bar{P} + \gamma \rho (\bar{P} - P_{n+1}) \\ &= (1 + \rho\gamma) \bar{P} - \rho\gamma P_{n+1}.\end{aligned}\tag{B.16}$$

Similarly, if (B.14) is satisfied, an inside contraction is carried out in which a new point P_{ic} is generated on the line $\overline{P_{n+1}\bar{P}}$ between P_{n+1} and \bar{P} as shown in Figure B.3(b) for the 2D case. The distance $\left\| \overline{P_{ic}\bar{P}} \right\|$ is again set by γ , according to

$$\gamma = \frac{\left\| \overline{P_{ic}\bar{P}} \right\|}{\left\| \overline{P_{n+1}\bar{P}} \right\|}.\tag{B.17}$$

Using γ , P_{ic} is generated by

$$\begin{aligned} P_{ic} &= \bar{P} - \gamma (\bar{P} - P_{n+1}) \\ &= (1 - \gamma) \bar{P} + \gamma P_{n+1}. \end{aligned} \quad (\text{B.18})$$

If either

$$f(P_{oc}) \leq f(P_r) \quad (\text{B.19})$$

or

$$f(P_{ic}) \leq f(P_{n+1}), \quad (\text{B.20})$$

the result of the contraction operation replaces the worst point P_{n+1} , the current iteration terminates, and the next iteration begins.

If neither (B.19) nor (B.20) is satisfied, a shrink step is performed in which P_1 is retained and all other points $P_{i \neq 1}$ are replaced with new points $P'_{i \neq 1}$ as illustrated in Figure B.4. The degree to which the distances $\|\overline{P_1 P'_{i \neq 1}}\|$ are reduced relative to the corresponding distances $\|\overline{P_1 P_{i \neq 1}}\|$ is given by the coefficient of shrinkage, σ according to

$$\sigma = \frac{\|\overline{P_1 P'_{i \neq 1}}\|}{\|\overline{P_1 P_{i \neq 1}}\|}. \quad (\text{B.21})$$

In general, $0 < \sigma < 1$, and a standard value is $\sigma = \frac{1}{2}$. Using σ , the points $P'_{i \neq 1}$ are generated by

$$P'_i = P_1 + \sigma (P_i - P_1). \quad (\text{B.22})$$

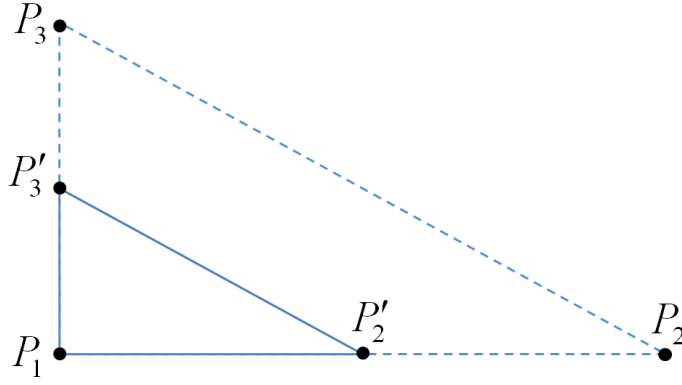


Figure B.4: Nelder-Mead shrink operation illustrated for the 2D case ($n = 2$). All points but P_1 are replaced with values $P'_{i \neq 1}$ closer to P_1 . Original simplex indicated by dashed lines.

After the shrink step, the current iteration is terminated and the next iteration begins.

The iterations continue until both the maximum distance in \mathbb{R}^n between the best point P_1 and all the other points $P_{i \neq 1}$ falls below a clustering threshold and the difference in objective function values $f(P_i)$ corresponding to these points falls to within a specified tolerance. The algorithm also terminates if it does not converge before it reaches an upper limit on either the number of iterations or the number of function evaluations.