

## ABSTRACT

KENZ, ZACKARY ROMAN. Stenosis-Driven Acoustic Wave Propagation in Biotissue: Modeling and the Inverse Problem. (Under the direction of H.T. Banks.)

A current goal in medical research is the development of a non-invasive method for detection, localization, and characterization of an arterial stenosis (a blockage or partial blockage in an artery). Partial blockage stenoses are known to cause disturbances in blood flow which generate pressure shear waves in the chest cavity. A method has been proposed to detect shear waves in the chest cavity.

In order to provide theoretical guidance and an understanding of the wave propagation process, we develop and validate a physics-based model for wave propagation. To this end, we first conduct proof-of-concept investigations using a preliminary one-dimensional viscoelastic model of pressure waves that incorporates Kelvin-Voigt damping and internal variables. We develop an estimation procedure for the material parameters and demonstrate its ability to recover the parameters which were used to create the simulated data. We then determine confidence intervals (using both bootstrapping and asymptotic error theory) for the estimated parameters, which indicates the efficacy of finding parameter estimates in practice. We then develop a model comparison test to be used in determining if a particular data set came from a low input amplitude or a high input amplitude; this we anticipate will aid in determining when stenosis is present.

To apply this methodology, we use one-dimensional pressure and shear wave experimental data from novel acoustic phantoms to validate an updated viscoelastic mathematical model. We estimate model parameters which give a good fit (in a sense to be precisely defined) to the experimental data, and at first use asymptotic error theory to provide confidence intervals for parameter estimates. Since a robust error model is necessary for accurate parameter estimates and confidence analysis, we include a comparison of absolute and relative models for measurement error.

We then move to comparing the performance of three methods for quantifying uncertainty in model parameters: asymptotic theory, bootstrapping, and Bayesian estimation. The first two methods are frequentist, meaning they assume a true value exists and provide uncertainty information about the parameter estimator; Bayesian estimation assumes the parameter is a density to be estimated, and the estimated density then provides uncertainty information. We study these methods on the pressure model, due to faster model computation times and the fact that the absolute error model (which possibly works with pressure data) makes comparison across the methods easier. In addition to parameter estimation, we use the results from the three algorithms to quantify complex correlations between our model parameters, which are

best seen using the more computationally expensive bootstrapping or Bayesian methods. We also hold constant the parameter causing the most complex correlation, obtaining results from all three methods which are more consistent than when estimating all parameters. Concerns regarding computational time and algorithm complexity are incorporated into discussion on differences between the frequentist and Bayesian perspectives.

Stenosis-Driven Acoustic Wave Propagation in Biotissue:  
Modeling and the Inverse Problem

by  
Zackary Roman Kenz

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the Degree of  
Doctor of Philosophy

Applied Mathematics

Raleigh, North Carolina

2013

APPROVED BY:

---

Ralph Smith

---

Robert Martin

---

Hien Tran

---

H.T. Banks  
Chair of Advisory Committee

## DEDICATION

This work is dedicated to my parents, brother, and sister, for their support and inspiration.

## BIOGRAPHY

The author was born in Fargo, ND on February 5, 1986. After a short period there (and much to the disappointment of his future PhD advisor), the family moved across the river to Moorhead, MN which the author would call home until graduate school. Later joined by a sister and brother, the author grew up in a quiet neighborhood where summers were filled with games like kickball and later long bicycle rides with friends as well as many hours spent golfing. After his elementary school years at St. Joseph's, the author entered Moorhead Junior High School (the Spuds, first of two food-related mascots much to the author's chagrin). It was there that the author first realized that a mathematical career might be in his future. The author participated in a number of activities like Knowledge Bowl, Science Olympiad, and MathCounts. This continued at Moorhead High School, where the author explored many interesting areas like history, the physical sciences, and economics, but mathematics still seemed to be the author's academic foundation. His time in Math League helped hone skills, and also provided some great memories traveling to competitions. The author also held jobs during high school at McDonald's and land surveying with Robert J. Roberts, both of which were instructive and formative experiences.

After high school, the author enrolled at Concordia College (the [corn] Cobbers), in Moorhead, MN. In what retrospectively seems quite brazen, the author sought research opportunities before starting his first undergraduate course. Fortunately, the mathematics faculty was willing to indulge the author's desire to start learning about mathematics research while also allowing him space to pursue his student government ambitions. Throughout undergrad, mathematics and student government competed for the author's time and attention, but mathematics again was the core foundation. A key moment in the author's undergraduate experience was a summer Research Experience for Undergraduates in the Los Angeles, CA, area after his sophomore year. This summer at the Applied Mathematical Sciences Summer Institute was the author's introduction to applied mathematics. Upon returning for his junior year, the author changed his majors to mathematics and computer science with an economics minor. Further coursework and another research summer (this time the Director's Summer Program) led the author to seek out graduate education in applied mathematics.

When the author added NC State to his list of possible graduate schools, he never imagined he would move as far south as Raleigh, NC. However, the opportunity available at NC State was a great fit for the author, so he moved in the summer of 2008 to start research work that summer before classes began. Though at times quite difficult, the graduate experience was ultimately successful for the author, not only through the production of this dissertation but also in gaining the tools and confidence needed to be a successful applied mathematician.

Though originally intending to pursue an academic career, in his final year of graduate school

the author discovered MIT Lincoln Labs and has taken a position there as a technical staff member. After defending this dissertation but before beginning at Lincoln Labs, the author will be fulfilling one of his life goals, working abroad. He will stay at Brunel University in Uxbrige, UK, for one month while working with the collaborators on this dissertation project. The author is excited to enter this next chapter of his life and to continue embracing unexpected opportunities.

## ACKNOWLEDGEMENTS

This dissertation could only have been completed with the support of many people throughout my 27 years. At the core has always been the support of my family. My parents instilled in me the drive to grow and succeed, and have consistently sacrificed for myself and my siblings to ensure we have the foundation to succeed. Though they probably don't know it, my brother and sister have helped keep me grounded and by being themselves remind me that while my academic work may often be completed individually I always need to keep in mind the impact I can have on others. My grandparents, in very different ways, complemented the life lessons of my parents and supported my growth as a person and student, for which I am always grateful. I love you all, and I hope I continue to apply the lessons learned from you.

My friends have also always been a great source of support. Though there are too many to list, I want to acknowledge some people specifically. Brad, Jordan, Ben, Steve, and Sam have been outstanding friends and have helped me be a better person for many years now. I'm truly fortunate to know them. Matt, the times at the Fort are always great. Molly, I don't think anyone outside my family has been supportive for as long as you have and I'm very grateful. Paul, I'm glad we were in the same Orientation group and for all the discussions we've had since. Breanna, I'm also fortunate we were clubbies, worked together on the Peace Prize Forum, and hang out together with Ben our senior year. Lucas, I don't remember everything you put up with as my freshman roommate, but I'm glad you were there to occasionally bump me back on track and for our continued friendship. Andy, I really appreciate the support during the student government years; you were an outstanding Executive Hand Holder and Chaplain. Kristin, Jordan, Whitney, and Sara, it was great running for office and serving with you both. Nate, I enjoyed our time in PBL and am thankful for the mentoring. Lacey, Travis, Jess, Kari, Kristle, Katie, and Kim, I'm thankful we had the year working together in the dorms and appreciate all you did to help me grow personally that year. Keri and Ralph, I have no idea how I would have survived my first few years without the support of you both. Our Friday dinners and time gaming was always a highlight during those weeks. Clay, I appreciate all your help during my first two years of coursework and for all the great discussions we had on mathematics research – I learned a great deal working with you. Danielle, Kathleen, Karen, Karyn, Ellen, Dustin, Kris, Amanda, Katie, and Jared, it's been great working with you mathematical siblings. Nathaniel, I've really appreciated all our conversation. Justin, thanks for a great four years as roommates and for getting to know you and Emma as friends. Lise, Anne, Nick, Deena, Anna, Susan, Mary Bridget, and Matt, among many others, thanks for hanging out and making sure I had some fun during grad school.

Of course, I wouldn't be in this position without all the teachers and mentors I have had

throughout the years. Lois Brown, the gifted and talented coordinator at Moorhead Junior High, was instrumental in making sure I participated in activities like Knowledge Bowl and signing up for the proper courses. It was in Ken Welken's pre-algebra class in 7th grade when I first thought I might have a future in mathematics, so I'm thankful for that class and for his mentoring in MathCounts. Missy Eidsness had a huge impact over my 8th-12th grade years, as the coach for Math League and as my calculus teacher. Her passion for the field and for student learning had a strong effect on how I want my work to impact others. I also enjoyed my time watching Jim Duncan and Jace Lahlum make algebra/geometry jokes, along with their quite different but still great teaching styles. Charles Fisher had a great impact on my appreciation for and understanding of history, how to process dense text, and how to communicate my understanding of ideas. Doug Johnson was an outstanding economics teacher, sparking a passion for that field, but also a great life mentor. I'm glad we have continued to meet for coffee when I am in Moorhead. During undergrad, Chelle Lyons Hanson and Paul Wraalstad, my student government advisors, encouraged me to grow personally and to have a positive impact on the world around me (in general, not just through student government). Chelle really helped me learn a good deal about myself, helping me to mature and learn to work with others in ways that result in action while maintaining respect for everyone with whom I work. I only hope that I did not cause her too much trouble during my two terms. The late Dr. Pamela Jolicouer demonstrated how to make hard choices that will be beneficial in the long run, and showed me that people in positions like college president are still people. I'm also thankful for the entire Cabinet that I worked with during my time in student government, and enjoy every time I see you when I am back on campus. Drs. Forde, Zeng, and Vanderschoot helped me transition to higher mathematics while demonstrating their infectious love of the field through teaching. Dr. Doug Anderson had a major impact on my life and my mathematical pursuits. I describe him as my advisor, mentor, teacher, research collaborator, and travel companion. I'm not sure it's possible for an undergrad professor to have more of an impact in a student's life as Doug has had on mine.

During my time at NC State, I'm glad to have had the chance to get to work with Lesa and Teresa in CRSC. Special thanks to Drs. Simon Shaw and Carola Kruse, who I worked most often with on this project. I also want to thank my committee, Drs. Ralph Smith, Hien Tran, Robert Martin, as well as my GSR Dr. John Meyer. Ralph, thanks also for the times when I stopped by and shoot the breeze. I also want to thank Dr. Shuhua Hu for all her mentoring, especially during my first summer project. I learned a great deal about being careful and diligent from Shuhua, and the success of this dissertation is due in no small part to her guidance. Finally, I want to thank my academic and thesis advisor Dr. Banks. I somehow ended up working with an advisor who not only showed me a number of tools that will be useful in my future career but who also helped me figure out the area and applications of mathematics that fit my interests.



It's hard to imagine anyone else being such a great fit and pushing me in the right mathematical directions. I'm also very appreciative of the Raleigh family he and Susie have been for me, especially so far from home.

I also want to acknowledge my support from the Department of Education through my four years of support as a GAANN Fellow (under grant number P200A070386), and thank Todd Marcks for his help administering my GAANN personal development funds. The GAANN allowed me to pursue research immediately and also afforded me the chance to participate in Preparing the Professoriate, a great teaching experience. In addition, funding for various projects came from the Air Force Office of Scientific Research (under grant numbers FA9550-09-1-0226 and AFOSR FA9550-12-1-0188) as well as the National Institute of Allergy and Infectious Diseases (under grant number R01AI071915-09).

# TABLE OF CONTENTS

<b>LIST OF TABLES</b>	<b>x</b>
<b>LIST OF FIGURES</b>	<b>xii</b>
<b>CHAPTER 1 Introduction</b>	<b>1</b>
<b>CHAPTER 2 Lab Experiment</b>	<b>5</b>
<b>CHAPTER 3 Proof of Concept and Methodology</b>	<b>8</b>
3.1 Viscoelastic Model	9
3.1.1 Existence and Uniqueness of a Weak Solution	10
3.1.2 Internal Variable Form of the Stress Relaxation Function	12
3.1.3 Parameter Values Used in Simulations	13
3.1.4 Direct (Forward) Problem	14
3.2 Estimation of Material Parameters: Preliminaries	15
3.2.1 Study of Effects of Changing Material Parameters	16
3.2.2 Sensitivity of Model Output with Respect to Material Parameters	18
3.3 Statistical Model and Inverse Problem	19
3.3.1 Simulated Data Generation	22
3.3.2 Parameter Estimates Obtained Using Different Routines	22
3.4 Asymptotic Error Analysis	25
3.4.1 Model Selection Criteria	29
3.5 Bootstrapping Error Analysis	31
3.6 Model Comparison and Hypothesis Testing on Amplitude	35
3.6.1 Setup	35
3.6.2 Hypothesis Testing Methodology	37
3.7 Proof of Concept Summary	39
<b>CHAPTER 4 Primary Model Derivation and Constitutive Relationship</b>	<b>40</b>
4.1 Basic Introduction to Viscoelasticity	40
4.2 Equations of Motion and General Constitutive Relationship	42
4.2.1 One-dimensional Pressure Equation	45
4.2.2 One-dimensional Shear Equation	46
4.3 Initial and Boundary Conditions	48
4.4 Constitutive Equation	49
4.4.1 Fung Quasi-Linear Model	49
4.4.2 Relating the Quasi-Linear and Strain Energy Function Formulations	50
4.4.3 Macroscopic Damping	51
4.4.4 Existence and Uniqueness for Pressure and Shear Models	51
4.4.5 Form for Constitutive Equation Kernel $G(t)$	53
4.4.6 Approximating the Loading Process	54
4.4.7 Internal Variable Formulation	56
4.4.8 Final Pressure and Shear PDE Models	59

4.5	Numerical Method . . . . .	60
<b>CHAPTER 5 Parameter Estimation and Results . . . . .</b>		<b>62</b>
5.1	Inverse Problem Preliminaries and Definitions . . . . .	62
5.1.1	Sensitivity of Model Output to Parameters . . . . .	64
5.2	Statistical Models and Parameter Estimators . . . . .	67
5.2.1	Optimization Considerations . . . . .	69
5.3	Results Using a Single Relaxation Time ( $N_p = 1$ ) . . . . .	70
5.3.1	Pressure Data Results . . . . .	71
5.3.2	Shear Data Results . . . . .	80
5.4	Model Development Summary . . . . .	86
<b>CHAPTER 6 Further Uncertainty Quantification Investigation . . . . .</b>		<b>87</b>
6.1	Mathematical and Statistical Models . . . . .	89
6.2	Methods for Studying Confidence in Parameters . . . . .	90
6.2.1	Frequentist Methods . . . . .	90
6.2.2	Bayesian Parameter Estimation and Confidence Analysis . . . . .	91
6.2.3	General Comparison Between Methods . . . . .	95
6.3	Results Based on Parameter Values From Chapter 5 . . . . .	97
6.3.1	Comparison of Preliminary Results Between Methods . . . . .	100
6.4	Updating Parameter Estimates . . . . .	104
6.4.1	OLS Inverse Problem Results . . . . .	104
6.4.2	Initial Comparison Between Methods . . . . .	104
6.5	Method Comparison When Holding $E$ Constant . . . . .	109
6.6	Uncertainty Quantification Analysis Summary . . . . .	113
<b>CHAPTER 7 Conclusion and Future Work . . . . .</b>		<b>115</b>
7.1	Future Investigation . . . . .	117
7.1.1	Relative versus Absolute Error for Restricted Parameter Set . . . . .	117
7.1.2	Parameter Correlation . . . . .	117
7.1.3	Piecewise Model Parameters . . . . .	117
7.1.4	Models with $N_p > 1$ ; Medium Inhomogeneities . . . . .	118
7.1.5	Modeling Input into the Body from the Arterial Wall . . . . .	118
7.1.6	Larger Phantom and Tissue Experiments . . . . .	119
7.1.7	Two-dimensional Model . . . . .	119
<b>REFERENCES . . . . .</b>		<b>121</b>
<b>APPENDICES . . . . .</b>		<b>128</b>
APPENDIX A Sensitivity Equations for Proof of Concept Model . . . . .		129
APPENDIX B Sensitivity Equations for Main Stenosis Wave Propagation Model . . . . .		134
B.1	Pressure Model . . . . .	134
B.2	Shear Model . . . . .	137

## LIST OF TABLES

Table 3.1	Estimation of material parameters at low, medium and high noise levels: Comparison between optimization routines ( <code>lsq-TRR=lsqnonlin</code> with the trust-region-reflective option, <code>lsq-LM=lsqnonlin</code> with the Levenburg-Marquardt option). . . . .	23
Table 3.2	<code>fmincon</code> : Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at low noise level. . . . .	26
Table 3.3	<code>fmincon</code> : Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at medium noise level. . . . .	26
Table 3.4	<code>fmincon</code> : Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at high noise level. . . . .	27
Table 3.5	TRR <code>lsqnonlin</code> : Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at low noise level. . . . .	27
Table 3.6	TRR <code>lsqnonlin</code> : Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at medium noise level. . . . .	28
Table 3.7	TRR <code>lsqnonlin</code> : Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at high noise level. . . . .	28
Table 3.8	<code>fmincon</code> : Residual sum of squares (RSS), AIC values, AIC difference ( $\Delta$ ) and AIC weights for zero-relaxation-time model (model 0), one-relaxation-time model (model 1) and two-times-relaxation model (model 2) obtained at low, medium and high noise levels. . . . .	30
Table 3.9	<code>lsqnonlin</code> : Residual sum of squares (RSS), AIC values, AIC differences ( $\Delta$ ) and AIC weights for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-relaxation-times model (model 2) obtained at low, medium and high noise levels. . . . .	31
Table 3.10	<code>fmincon</code> : Parameter estimates, bootstrap standard errors (SE) and confidence intervals obtained at low, medium and high noise levels for zero-relaxation-time model. . . . .	33
Table 3.11	TRR <code>lsqnonlin</code> : Parameter estimates, bootstrap standard errors (SE) and confidence intervals obtained at low, medium and high noise levels for one-relaxation-time model. . . . .	33

Table 3.12	Sample $\chi^2(1)$ values. . . . .	38
Table 3.13	Model comparison test results using (3.6.1) on low, medium, and high noise data sets generated with both high and low input amplitude parameter $A$ values. . . . .	39
Table 5.1	Pressure optimization results and confidence analysis for OLS on a 264 g data set using every data point. . . . .	72
Table 5.2	Pressure optimization results and confidence analysis for GLS on a 264 g data set using every data point. . . . .	72
Table 5.3	Pressure optimization results and confidence analysis for OLS on a 264 g data set using every other data point. . . . .	73
Table 5.4	Pressure optimization results and confidence analysis for GLS on a 264 g data set using every other data point. . . . .	73
Table 5.5	Pressure optimization results and confidence analysis for GLS (half error) on a 264 g data set, using every data point. . . . .	74
Table 5.6	Pressure optimization results and confidence analysis for GLS (half error) on a 264 g data set, using every other data point. . . . .	75
Table 5.7	Shear optimization results and confidence analysis for OLS on a 264 g data set using every data point. . . . .	80
Table 5.8	Shear optimization results and confidence analysis for GLS on a 264 g data set using every data point. . . . .	81
Table 5.9	Shear optimization results and confidence analysis for OLS on a 264 g data set using every other data point. . . . .	81
Table 5.10	Shear optimization results and confidence analysis for GLS on a 264 g data set using every other data point. . . . .	81
Table 6.1	Pressure bootstrap results. . . . .	98
Table 6.2	Pressure asymptotic results (using MCMC parameter means from the chains in Figure 6.2 as the initial guess). . . . .	106
Table 6.3	Pressure bootstrap results (using MCMC parameter means from the chains in Figure 6.2 as the initial guess). . . . .	106
Table 6.4	Pressure optimization results for the following cases: OLS, bootstrapping mean using OLS value to initiate, and DRAM parameter means; data frequency 1024 Hz. $RSS=\mathcal{J}_{ols}(q)$ . . . . .	106
Table 6.5	Pressure optimization results for the following cases (with $E$ fixed): OLS, bootstrapping mean using OLS value to initiate, and DRAM parameter means; data frequency 1024 Hz. $RSS=\mathcal{J}_{ols}(q)$ . . . . .	110
Table 6.6	Pressure OLS asymptotic results, with $E$ fixed. . . . .	110
Table 6.7	Pressure bootstrap results, with $E$ fixed. . . . .	110

## LIST OF FIGURES

Figure 1.1	Conceptual schematic for an <i>in vivo</i> coronary artery stenosis detection device based on wave propagation through the chest cavity. The vessel on the left is a magnified coronary artery, showing turbulent flow resulting from a stenosis. Graphic courtesy of Dr. S. Greenwald. . . . .	2
Figure 2.1	Pressure configuration, where TMM denotes the tissue mimicking material and LDT denotes the laser displacement transducer measurement device. (a) Experimental setup of agar phantom. (b) Schematic with one-dimensional domain denoted. . . . .	6
Figure 2.2	Shear configuration, where again TMM denotes the tissue mimicking material and LDT denotes the laser displacement transducer measurement device. (a) Experimental setup of agar phantom. (b) Schematic with one-dimensional domain denoted. . . . .	6
Figure 2.3	Sample one-dimensional data. Loading of the material (initially at rest) begins at $t = \Gamma_1$ , and the material is loaded and continuing to relax for $t \in (\Gamma_2, \Gamma_3)$ . At time $\Gamma_3$ the load is cut which takes roughly 10 ms–15 ms. The gel is then freely oscillating at $\Gamma_4 = 0$ , and oscillations continue for a period of time dependent on the loading weight and wave type (pressure/shear). The value $A$ is the displacement of the material at the beginning of free oscillations. The overall displacement scale of the data is on the order of $10^{-4}$ m, while the oscillations immediately after the weight release are on the order of $10^{-5}$ m. . . . .	7
Figure 3.1	Van Bladel function (c.f. (3.1.9)) with $A = 6 \times 10^3$ , $a = 6 \times 10^{-3}$ , and $b = 20 \times 10^{-3}$ . . . . .	14
Figure 3.2	Solution of (3.1.1) with (3.1.7)-(3.1.8) using the “true” parameters (3.1.10) and forcing function with parameters $a = 6 \times 10^{-3}$ , $b = 20 \times 10^{-3}$ , and $A = 6 \times 10^3$ (depicted by the solid line), alongside solutions using $E_0 = 2 \times 10^5$ (a) and $E_1 = 60$ (b) which are represented with dashed lines in their respective graphs. . . . .	16
Figure 3.3	Solution of (3.1.1) with (3.1.7)-(3.1.8) using the “true” parameters (3.1.10) and forcing function with parameters $a = 6 \times 10^{-3}$ , $b = 20 \times 10^{-3}$ , and $A = 6 \times 10^3$ (depicted by the solid line), alongside the dashed line using $\tau_1 = 0.5$ with the remaining parameters the same in (a). In (b), the solution is zoomed in for $t \in [0.02, 0.04]$ . . . . .	17
Figure 3.4	(a) Sensitivity of model output with respect to $\log_{10}(E_0)$ ; (b) Sensitivity of model output with respect to $\log_{10}(E_1)$ ; (c) Sensitivity of model output with respect to $\log_{10}(\tau_1)$ ; and (d) Sensitivity of model output with respect to $\log_{10}(\tau_2)$ . All sensitivities are around the baseline parameters (3.1.10) and (3.1.11). . . . .	19

Figure 3.5	(a) Sensitivity of model output with respect to $\log_{10}(p_1)$ . (b) Sensitivity of model output with respect to $\log_{10}(p_2)$ . Both sensitivities around the baseline parameters (3.1.10) and (3.1.11). . . . .	20
Figure 3.6	Simulated noisy data around the true parameter values. (a) Low noise level. (b) Medium noise level. (c) High noise level. . . . .	22
Figure 3.7	Data and two-relaxation-time model solutions at parameter estimates obtained using the <code>lsqnonlin</code> routine, trust-region-reflective method, at different noise levels (see Table 3.1). (a) Low noise. (b) Medium noise. (c) High noise. . . . .	24
Figure 3.8	Histograms of bootstrap estimates $q^m$ for a model with no relaxation times in the case of low noise data set (upper row), medium noise data set (middle row) and high noise data set (bottom row). (left column) Estimates for $\log_{10}(E_0)$ . (right column) Estimates for $\log_{10}(E_1)$ . . . . .	34
Figure 3.9	Histograms of bootstrap estimates $\hat{q}^m$ for a one-relaxation-time model obtained at low noise (upper row), medium noise (middle row) and high noise (bottom row) levels. (left column) Estimates for $\log_{10}(E_0)$ . (middle column) Estimates for $\log_{10}(E_1)$ . (right column) Estimates for $\log_{10}(\tau_1)$ . . . . .	35
Figure 3.10	Sensitivity of model with respect to Van Bladel input parameter $A$ around the baseline parameters (3.1.10). (a) High forcing function amplitude $A = 6 \times 10^3$ . (b) Low forcing function amplitude $A = 6 \times 10^2$ . Note the order of magnitude difference between (a) and (b). . . . .	36
Figure 3.11	Simulated low amplitude noisy data around the true parameter values. (a) Low noise level, $\sigma_0^2 = 5 \times 10^{-7}$ . (b) Medium noise, $\sigma_0^2 = 10 \times 10^{-7}$ . (c) High noise level, $\sigma_0^2 = 20 \times 10^{-7}$ . . . . .	37
Figure 4.1	Stress components in Cartesian coordinates. . . . .	41
Figure 5.1	Pressure sensitivity equation solutions in the $N_p = 1$ case. Solved using the following parameter values: $E = 5 \times 10^4$ Pa, $E_1 = 30$ Pa·s, $\gamma_1 = 1.2 \times 10^5$ Pa, $\tau = 0.02$ s, $A = -1.65 \times 10^{-4}$ m, and $\Upsilon = -0.01$ s. (top) Resulting pressure model solution. (bottom six) Sensitivity equation solutions for each parameter. . . . .	65
Figure 5.2	Shear sensitivity equation solutions in the $N_p = 1$ case. Solved using the following parameter values: $G = 6.5 \times 10^3$ Pa, $G_1 = 5.6$ Pa·s, $\zeta_1 = 8 \times 10^3$ Pa, $\tau = 0.07$ s, $A = -1.5 \times 10^{-4}$ m, and $\Upsilon = -0.01$ s. (top) Resulting shear model solution. (bottom six) Sensitivity equation solutions for each parameter. . . . .	66
Figure 5.3	Pressure data fit using every data point with absolute error model, $N_p = 1$ , weight 264 g. (a) Model fit to data. (b) Absolute residuals vs time. (c) Absolute residuals vs model. . . . .	74
Figure 5.4	Pressure data fit using every data point with relative error model, $N_p = 1$ , weight 264 g. (a) Model fit to data. (b) Relative residuals vs time. (c) Relative residuals vs model. . . . .	75

Figure 5.5	Pressure data fit using every other data point with absolute error model, $N_p = 1$ , weight 264 g. (a) Model fit to data. (b) Absolute residuals vs time. (c) Absolute residuals vs model. . . . .	76
Figure 5.6	Pressure data fit using every other data point with relative error model, $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Relative residuals vs time. (bottom right) Relative residuals vs model. . . . .	77
Figure 5.7	Pressure data fit using every data point with relative half error model, $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Relative residuals vs time. (bottom right) Relative residuals vs model. . . . .	78
Figure 5.8	Pressure data fit using every other data point with relative half error model, $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Relative residuals vs time. (bottom right) Relative residuals vs model. . . . .	79
Figure 5.9	Shear data fit using every data point with absolute error model, $N_p = 1$ , weight 264 g. (a) Model fit to data. (b) Absolute residuals vs time. (c) Absolute residuals vs model. . . . .	82
Figure 5.10	Shear data fit using every data point with relative error model, $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Relative residuals vs time. (bottom right) Relative residuals vs model. . . . .	83
Figure 5.11	Shear data fit using every other data point with absolute error model, $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Absolute residuals vs time. (bottom right) Absolute residuals vs model. . . . .	84
Figure 5.12	Shear data fit using every other data point with relative error model, $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Relative residuals vs time. (bottom right) Relative residuals vs model. . . . .	85
Figure 6.1	Pressure bootstrap histograms, with data sampled at 1024 Hz, $N_p = 1$ . The corresponding OLS estimate from Table 5.1 is shown as the dashed line. All have a mostly normal shape, with some small peaks but mainly one large peak for each parameter. . . . .	98
Figure 6.2	Pressure parameter chain results using DRAM, with data sampled at 1024 Hz, $N_p = 1$ . The corresponding OLS estimate from Table 5.1 is shown as the horizontal dashed line. The vertical dashed line represents the iteration after which we consider the chains to be burned in. . . . .	100
Figure 6.3	Pressure parameter density results using DRAM, with data sampled at 1024 Hz, $N_p = 1$ . The corresponding OLS estimate from Table 5.1 is shown as the dashed line. . . . .	101
Figure 6.4	Pressure parameter pairwise parameter comparisons using DRAM, with data sampled at 1024 Hz, $N_p = 1$ , using every 50th chain value. Noticeable patterns in each sub-graph indicate correlation between parameters (e.g., the upper left corner between $E$ and $E_1$ ) whereas roughly random relationships indicate less correlations (e.g., the parameter $A$ with any other parameter). . . . .	102



Figure 6.5	Comparison of density results between the three methods, for each parameter. The MCMC density is computed from the chain values using KDE; for bootstrapping and asymptotic analysis, the mean and variance for each parameter are used to create a normal probability density function which is plotted. . . . .	103
Figure 6.6	Pressure data fit using OLS, with initial guess set to be the mean of the MCMC parameter chains, with data sampled at 1024 Hz, $N_p = 1$ . (a) OLS model fit to data. (b) Model vs absolute residuals. (c) Time versus absolute residuals. . . . .	104
Figure 6.7	Comparison of density results between the three methods, for each parameter. The MCMC density is computed from the chain values using KDE; for bootstrapping and asymptotic analysis, the mean and variance for each parameter are used to create a normal probability density function which is plotted. . . . .	107
Figure 6.8	Pressure parameter chain results using DRAM, with data sampled at 1024 Hz, $N_p = 1$ . The corresponding OLS estimate from Table 6.4 is shown as the dashed line. . . . .	108
Figure 6.9	Pressure parameter pairwise comparisons using DRAM, with data sampled at 1024 Hz, $N_p = 1$ , using every 50th chain value. Noticeable patterns in a sub-graph indicate correlation between parameters (e.g., the middle left comparison between $\gamma_1$ and $E$ ) whereas roughly random relationships indicate less correlations (e.g., the parameter $A$ with any other parameter). . . . .	108
Figure 6.10	Comparison of density results between the three methods, for each parameter (with $E$ fixed). The MCMC density is computed from the chain values using KDE; for bootstrapping and asymptotic analysis, the mean and variance for each parameter are used to create a normal probability density function which is plotted. . . . .	111
Figure 6.11	Pressure parameter (with $E$ fixed) chain results using DRAM, with data sampled at 1024 Hz, $N_p = 1$ . The vertical dashed line indicates the point at which we consider the run to be burned in. . . . .	112
Figure 6.12	Pressure parameter (with $E$ fixed) pairwise comparisons using DRAM, with data sampled at 1024 Hz, $N_p = 1$ , using every 50th chain value. Note that most relationships are significantly more random than in the previous pairwise plot results, indicating that these parameters are roughly uncorrelated, although some linear correlation may still exist (e.g., $\tau_1$ and $\Upsilon$ ) . . . . .	112

Coronary artery disease (CAD) is an increasingly prevalent medical condition, often a precursor to and cause of a patient experiencing cardiac arrest. Current methods for detection of arterial stenoses (blocked arteries) include angiograms and CT scans. Angiograms are viable but quite invasive, while CT scans are expensive, introduce radiation into the patient, and can only detect hard plaques (blockages). A desirable new detection method would be noninvasive and less expensive but still effective. To this end, using acoustic waves generated by stenoses has been proposed. This would place sensors on the surface of the chest to listen for sounds from coronary arteries, with the hope of detecting and then localizing any blockages. A conceptual schematic is shown in Figure 1.1.

The current understanding (see, e.g., [9, 32, 73]) of the process is that turbulent flow produces normal forces on the vessel walls at and downstream of a stenosis, which then exert pressure on the vessels wall causing a small displacement in the surrounding soft tissue. Previous work (e.g. [2, 40, 74, 70, 71, 82, 92]) has demonstrated the existence of such sounds, and that they are possible to detect on the surface of the chest. The system couples two processes:

- (1) the generation of pressure and shear waves transmitted into the body by the arterial wall as a result of the turbulent blood flow generated by a stenosis, and
- (2) the propagation of pressure and shear waves through the chest to sensors attached to the chest wall.

The first process is not completely understood, though some ideas are present in the literature. Various researchers [30, 31, 32, 55, 64, 68, 73, 76, 77, 89, 98] have directly examined modeling

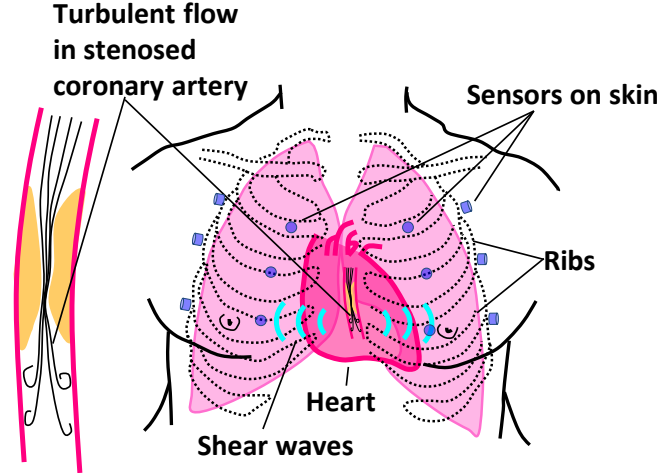


Figure 1.1: Conceptual schematic for an *in vivo* coronary artery stenosis detection device based on wave propagation through the chest cavity. The vessel on the left is a magnified coronary artery, showing turbulent flow resulting from a stenosis. Graphic courtesy of Dr. S. Greenwald.

blood flow through arteries. All have attempted to characterize the turbulence in the flow, which some then used to examine the sound field propagated into the chest. In this dissertation, we will not focus on aspects of turbulent flow, leaving that instead as an input to be later properly characterized when the properties of the stenosed artery itself are modeled. We will focus on the second process, understanding the propagation of sounds through the chest cavity which result from stenosed coronary arteries.

The modeling and detection of waves transmitted through the body has been approached in different ways. One approach has focused on characterizing properties of the sounds detected on the surface of the chest, characterizing aspects such as primary frequencies that these sounds exhibit. This line of work has been studied by Semmlow, et al., [2, 3, 4, 5, 6, 75, 84, 85, 100], as well as by other groups [30, 31, 32, 47, 74, 92, 93, 95]. Their methods are based on general sound features and detection through statistical methods, rather than modeling the underlying physics of sound transmission through the body. These methods have the benefit of being fast and fairly simple to implement, but do not provide a characterization for the mechanisms of wave transmission.

In another direction, more relevant to the situation this dissertation will study, researchers have worked to model the physics of sound wave transmission through the body. As is common in many physical wave phenomena, pressure and shear waves both propagate into the body as a result of stenosed coronary arteries. Since shear waves in general have a lower amplitude than pressure waves, intuition might suggest detecting pressure waves should be our focus; past research indicates the opposite is true when seeking to detect stenoses. Various groups

[36, 38, 65, 73] have demonstrated that shear waves should be the focus of detection efforts. The frequency range for shear waves resulting from coronary stenoses are below 2000 Hz [31, 36, 37, 44, 81, 84]. In the range of these frequencies, the pressure waves propagate very quickly while shear waves propagate much more slowly, which in practice means that pressure waves are difficult to measure. Also, in the context of waves propagating in tissue or tissue-mimicking materials, shear waves are measurable at greater distances from the source of the disturbance (see, e.g., Figures 9b and 10 of [36]). Though we will examine both models in this dissertation, the shear results in Chapter 5 are of greatest practical interest.

The benefits of using a viscoelastic wave propagation model in various contexts have been studied by [36, 38, 39, 44, 59, 65, 81]. Therein, the authors focused on determining the elastic modulus and viscoelastic parameter based on the shear wave speed and attenuation in either a gel mold or physical tissue, in both a stenosis context and general tissue shear wave propagation. The models were developed using plane waves in such a way that algebraic expressions were derived for shear wave speed as a function of frequency, elastic modulus, and the viscoelastic parameter. These demonstrate that modeling the underlying physics is not only possible, but quite beneficial in understanding shear wave propagation. These investigators also showed that a Kelvin-Voigt damping model is most appropriate for the situation, which we will incorporate into our model (and will discuss more later).

In this dissertation, we will take the physical models further, developing a dynamic model of the shear waves propagating through a tissue-mimicking material. Our goal will be to use this viscoelastic model with data from a tissue-mimicking homogeneous gel mold to validate the model and understand the uncertainty in the model parameters. The model here will incorporate the standard elastic/shear modulus and bulk viscoelastic parameter, as well as *internal variables* governed by *relaxation times* which can be used to model how different portions of the medium relax in different ways from being stressed. We will develop the model in such a way as to allow for multiple internal variables. In future studies these can be used to model the bulk effects of different types of tissue, which would be closer to the *in vivo* detection problem. This is also key to moving beyond just the *detection* of a stenosis to actually *localizing* the blockage. Our efforts hope to provide theoretical guidance for localization, something that models which only characterize received frequencies or even just basic properties based on attenuation do not promise to provide.

After some proof-of-concept investigation on a related model, the main model developed in this dissertation will pick up from a previous line of work by Banks, et al., [9, 20, 21, 22, 23, 67, 80]. These models allow for a characterization of shear waves resulting from coronary stenoses, which will assist in uncovering the coronary artery sounds from the noisy background in the body. Initial experiments were conducted where a gel mold was built with a tube running through the middle; cases where the tube was unblocked were compared to those with partial blockages, and

the results suggested that there were significant differences in sound generation between the blocked and unblocked cases. Unfortunately, this line of study ended before experimental data could be incorporated and fitted to models. This dissertation returns to a one-dimensional model and experimental setup, developing a model closely related to that in [9] which we validate with lab data from a homogeneous gel phantom and use to examine uncertainty in our parameter estimates. This dissertation will examine confidence in the relaxation times, demonstrating that the addition of internal variables into the model is viable.

In Chapter 2, we discuss the development of the one-dimensional lab experiments. Novel acoustic agar gel phantoms and holding rigs were developed at Queen Mary, University of London (QMUL) and Barts Health Trust (BHT) in England in order to experimentally generate one-dimensional pressure and shear waves. The experiments are a drop test, whereby a weight is attached to an apparatus that stresses the phantom. This is allowed to settle for a period of time, and then the weight is released nearly instantaneously. The material undergoes oscillation, which then damps out and the material settles toward some equilibrium. From a model validation standpoint this data is quite useful since it provides frequency and damping information, both of which are crucial when attempting to estimate model parameters. This experiment also demonstrates the response of the tissue-mimicking gel to an impact which is similar to the blood impacting the vessel wall downstream of a blockage.

In Chapter 3, we will use a preliminary viscoelastic model in order to build an inverse problem methodology using simulated data. The inverse problem is tested under different levels of noise, and we compare and contrast the performance of using two frequentist methods for quantifying the uncertainty in the sampling distribution for the parameters. We also develop a model comparison test that could serve as a basis for deciding if a stenosis is present or not. After that, in Chapter 4 we derive the one-dimensional equations of motion for pressure and shear wave propagation from the general three-dimensional equations of motion and specify a slightly more general constitutive relationship than was used in the preliminary model. These models are then used in Chapter 5 with experimental data from QMUL in order to estimate the model parameters which best describe the data. Different models for the measurement error are studied, and uncertainty information on the sampling distribution for the parameters is provided. This uncertainty analysis is expanded in Chapter 6, where we compare the performance of the two frequentist methods used in previous chapters with a Bayesian approach to understanding uncertainty in the parameter values. We summarize all the results in Chapter 7, and discuss ideas for future work on the coronary artery stenosis problem.

## CHAPTER 2

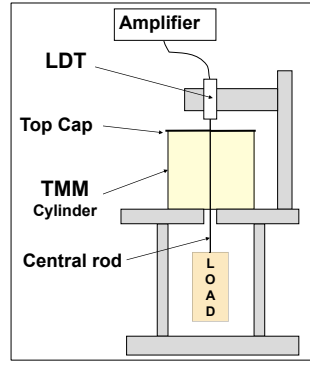
---

### Lab Experiment

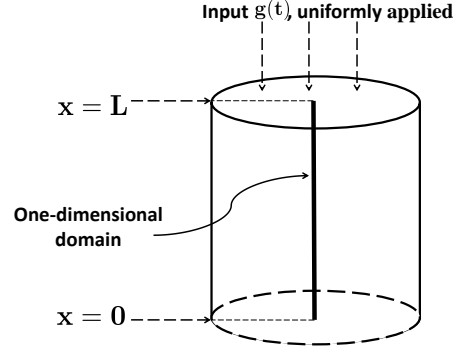
---

In order to develop an experiment and corresponding methodology, work was completed in collaboration with the teams at the Center for Research in Scientific Computation at North Carolina State University (H.T. Banks, Shuhua Hu, and the author); the Brunel Institute of Computational Mathematics (BICOM) at Brunel University in England (Carola Kruse, Simon Shaw, and John Whiteman); and Queen Mary, University of London (QMUL) and Barts & London NHS Trust (BLT) in England (Malcolm Birch, Mark Brewin, and Steve Greenwald) to devise and build a novel acoustic phantom and testing rig. These were designed to be drop tests which would excite many modes of vibration in the phantom, ensuring that the damping envelope was clear as well as the oscillation frequency (which encodes the elastic modulus or shear modulus). Describing the damping properties of the material is an important aspect of the viscoelastic wave propagation model we will build.

Two separate experiments were devised in consultation with QMUL to gather pressure and shear data; though we will tend to discuss them together, the experiments are run at completely different times and with slightly different phantoms. Devices have been designed (see left panes of Figures 2.1-2.2), in which an agar gel mold phantom (homogeneous, 97% water, density  $\rho = 1010 \text{ kg/m}^3$ ) is loaded into the rig, a weight is attached applying stress to the phantom, and then the weight is released, causing the material to oscillate. The displacement motion of the material throughout the experiment is measured with a laser device. The choice of loading and a quick release produces dynamic data; the idea was inspired in part by the impacts the stenosed vessel wall experiences with each heartbeat and also by past success in gathering shear data for filled rubber elastomers using an initially loaded rubber sample which then underwent

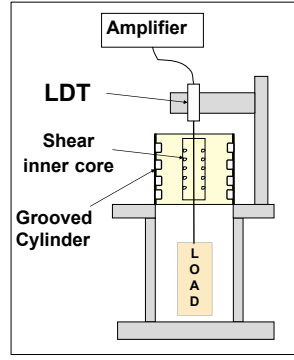


(a)

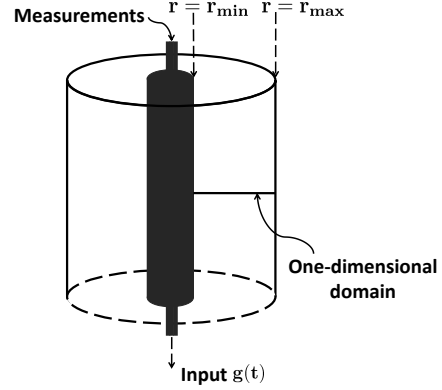


(b)

Figure 2.1: Pressure configuration, where TMM denotes the tissue mimicking material and LDT denotes the laser displacement transducer measurement device. (a) Experimental setup of agar phantom. (b) Schematic with one-dimensional domain denoted.



(a)



(b)

Figure 2.2: Shear configuration, where again TMM denotes the tissue mimicking material and LDT denotes the laser displacement transducer measurement device. (a) Experimental setup of agar phantom. (b) Schematic with one-dimensional domain denoted.

an impulsive hammer hit (see e.g. [20, 21]). This yields one-dimensional pressure data along the vertical axis in the pressure case (right pane of Figure 2.1) and, in the second experiment, shear data in the radial direction perpendicular to the vertical axis (right pane of Figure 2.2).

In order to test the phantom response to different stress levels, weights of 66 g, 132 g, 198 g, and 264 g were used in the experiment. The gel phantoms were stored in water when not in use, which keeps the gel at the desired 97% water composition. We will only use data from the 264 g

tests later when demonstrating results. These data provide the strongest oscillatory signal and thus are the best option for model validation. The data sets were generally similar otherwise, with the amplitude and number of oscillations decreasing as the weight level decreases.

When the experiment is conducted, data like those depicted in Figure 2.3 are produced. The material is at rest, a weight is added and allowed to settle, then the string holding the weight is rapidly cut with a flame to allow the material to freely oscillate. Once oscillations have died out the material relaxes back toward a stable state. The key pieces that will be modeled are the loading profile (loading begins at  $t = \Gamma_1$  and lasts until the weight begins to be released at  $t = \Gamma_3$ ), which we will model as instantaneous loading to position  $A$ , and the oscillations after weight release (free oscillations begin at  $t = \Gamma_4 = 0$ ) which are the main object of investigations here.

With the setup of the experiment in mind, we next turn to an examination of the inverse problem (parameter estimation) procedure on a preliminary viscoelastic wave propagation model. This proof of concept discussion will use simulated data, while later discussion on the inverse problem with a revised viscoelastic model will use data from the lab experiments. For more information on the experimental setup, interested readers may refer to the forthcoming report from QMUL [34].

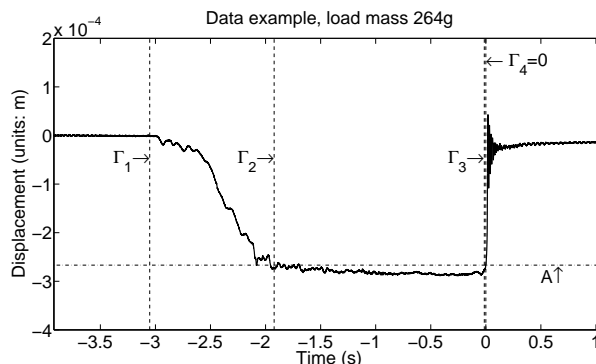


Figure 2.3: Sample one-dimensional data. Loading of the material (initially at rest) begins at  $t = \Gamma_1$ , and the material is loaded and continuing to relax for  $t \in (\Gamma_2, \Gamma_3)$ . At time  $\Gamma_3$  the load is cut which takes roughly 10 ms–15 ms. The gel is then freely oscillating at  $\Gamma_4 = 0$ , and oscillations continue for a period of time dependent on the loading weight and wave type (pressure/shear). The value  $A$  is the displacement of the material at the beginning of free oscillations. The overall displacement scale of the data is on the order of  $10^{-4}$  m, while the oscillations immediately after the weight release are on the order of  $10^{-5}$  m.



---

## Proof of Concept and Methodology

---

Our goals for this proof of concept chapter are twofold. We first focus on *developing methodologies for determining material parameters and analyzing data using a viscoelastic model*, as well as also *quantifying the uncertainty in the estimation procedure* through both bootstrapping and asymptotic error theories. As part of this latter goal, we also conduct model comparison testing to examine the viability of determining if data originated from a low-amplitude traction (e.g., resulting from normal blood flow) or high-amplitude traction (e.g., resulting from abnormal blood flow caused by a stenosis). This notion uses and enhances previous work (see, e.g., [9, 70, 71, 73, 74, 85, 91, 92]) which discussed the compression and shear waves which result from a stenosed vessel and some methods for measuring these waves, in particular the shear waves. An ultimate goal of our wider research project will be a methodology to decide if a vessel is stenosed or not, and if so, possibly the extent and location of the stenosis. However, at this point we are still in the process of carrying out experiments to determine these differences in either test devices or live subjects. Thus, in these preliminary efforts we make the (very) tentative assumption of representing the difference between normal vessels and stenosed vessels as a comparison between low and high shear input amplitudes, leaving the specifics of the actual system inputs to future work. Overall, then, these two thrusts of material parameter estimation and model comparison tests represent a proof-of-concept for our future data-driven inverse problem efforts. Note that much of this chapter has been published as [14].

### 3.1 Viscoelastic Model

We will examine a one-dimensional viscoelastic model for an agar phantom in the pressure case, as described in Chapter 2 and pictured in Figure 2.1. If we assume a uniform force applied along the top of the phantom and radial symmetry within the phantom (in part to closely match the symmetrically constructed phantoms used at QMUL), then we can simplify the cylindrical physical domain to a one-dimensional domain and to finding the function  $u(x, t)$  which represents the material response to, in this case, an applied stenosis-generated like force. In Chapter 4, we will go into more detail on how the one-dimensional equation is derived from the three-dimensional equations of motion. For this chapter, we use a general acoustic pressure viscoelastic wave equation on a one-dimensional domain  $\Omega = [0, L]$  (c.f. Figure 2.1). For the purposes of our initial investigation here, all parameters will be considered constant (i.e., a homogeneous medium). This is not necessary but significantly simplifies our initial computations in the methodology development. Choosing a material initially at rest with a reflecting boundary at  $x = 0$ , an applied force  $g(t)$  at the  $x = L$  boundary, and no additional forcing terms we obtain the system for the displacement  $u$  given by

$$\begin{aligned} \rho u_{tt} - \sigma_x &= 0 \\ u(0, t) &= 0, \quad \sigma(L, t) = -g(t), \\ u(x, 0) &= 0, \quad u_t(x, 0) = 0. \end{aligned} \tag{3.1.1}$$

Here the stress  $\sigma(x, t)$  is assumed to be described by

$$\sigma(x, t) = E_1 u_{xt}(x, t) + E_0 \int_0^t P(t-s) \frac{d}{ds} u_x(x, s) ds, \tag{3.1.2}$$

where the  $E_1 u_{xt}$  term is the so-called Kelvin-Voigt damping, and  $P$  is some stress relaxation function. Normally the integral in (3.1.2) must be considered for all previous time for a viscoelastic material, since the material responds differently based on its history. Our assumption that the material is at rest, coupled with assuming none of its history before the experiment begins affects the state at  $t = 0$ , allows us to consider the history integral from  $t = 0$  forward.

Development of this model is described in [13], as well as in standard viscoelastic theory [41, 48, 49, 50, 54, 63]. It is worth noting that the linear system (3.1.1) with (3.1.2) was found to give a reasonable preliminary approximation to the experimental data provided by QMUL. Thus, this linear model is deemed adequate for our investigation for the current experimental setting.

### 3.1.1 Existence and Uniqueness of a Weak Solution

Let  $\mathbb{H} = \mathcal{L}^2(0, L)$ ,  $\mathbb{V} = \{\phi \mid \phi \in \mathcal{H}^1(0, L), \phi(0) = 0\}$ , and  $\mathbb{V}^*$  denote the topological dual space of  $\mathbb{V}$ . If we identify  $\mathbb{H}$  with its topological dual  $\mathbb{H}^*$ , then  $\mathbb{V} \hookrightarrow \mathbb{H} = \mathbb{H}^* \hookrightarrow \mathbb{V}^*$  is a Gelfand triple [8, 99]. Throughout this presentation  $\langle \cdot, \cdot \rangle$  denotes the inner product in  $\mathbb{H}$ , and  $\langle \cdot, \cdot \rangle_{\mathbb{V}^*, \mathbb{V}}$  represents the duality pairing between  $\mathbb{V}^*$  and  $\mathbb{V}$  (again see [8, 99] for details).

Let  $\mathcal{C}_w(0, T; \mathbb{V})$  denote the set of weakly continuous functions in  $\mathbb{V}$  on  $[0, T]$ , and  $\mathcal{L}_T = \{v : [0, T] \rightarrow \mathbb{H} \mid v \in \mathcal{C}_w(0, T; \mathbb{V}) \cap \mathcal{L}^2(0, T; \mathbb{V}) \text{ and } v_t \in \mathcal{C}_w(0, T; \mathbb{H}) \cap \mathcal{L}^2(0, T; \mathbb{V})\}$ . The weak solution for system (3.1.1) with (3.1.2) is defined in the following way.

**Definition 3.1.1.** We say that  $u \in \mathcal{L}_T$  is a weak solution of the system (3.1.1) and (3.1.2) if it satisfies

$$\begin{aligned} & \rho \langle u_t(t), \eta_t(t) \rangle - \rho \int_0^t \langle u_s(s), \eta_s(s) \rangle ds + \int_0^t g(s) \eta(L, s) ds + E_1 \int_0^t \langle u_{sx}(s), \eta_x(s) \rangle ds \\ & + E_0 \int_0^t \left\langle \int_0^s P(s - \xi) \frac{d}{d\xi} u_x(\xi) d\xi, \eta_x(s) \right\rangle ds = 0 \end{aligned} \quad (3.1.3)$$

for any  $t \in [0, T]$  and  $\eta \in \mathcal{L}_T$ . Here and elsewhere in this section  $u(t)$  and  $\eta(t)$  denote the functions  $u(\cdot, t)$  and  $\eta(\cdot, t)$ , respectively.

As in [8, 22, 25, 99] we remark that this notation of the weak solution for system (3.1.1)-(3.1.2) agrees with the usual one in that it yields  $u_{tt} \in \mathcal{L}^2(0, T; \mathbb{V}^*)$  with equation (3.1.1) holding in the sense of  $\mathcal{L}^2(0, T; \mathbb{V}^*)$ . To ensure the existence and uniqueness of a weak solution to system (3.1.1) with (3.1.2), we make the following assumptions on the force function  $g$  and stress relaxation function  $P$ :

(A1)  $g \in L^2(0, T)$ .

(A2) The function  $P$  is differentiable with respect to  $t \in \mathbb{R}^+$ , and there exist constants  $c_1$  and  $c_2$  such that  $|P(t)| \leq c_1$  and  $|\dot{P}(t)| \leq c_2$  for all  $t \in \mathbb{R}^+$ , where  $\dot{P}$  denotes the differentiation of  $P$  with respect to  $t$ .

**Theorem 3.1.2.** *Under assumptions (A1) and (A2), the system (3.1.1) with (3.1.2) has a unique weak solution on any finite interval  $[0, T]$ .*

We remark that system (3.1.1) with (3.1.2) is a special case of the viscoelastic model presented in [22] if we change the left boundary condition  $u(0, t) = 0$  in (3.1.1) to  $\sigma(0, t) = g_l(t)$  with  $g_l(t)$  being some force function and change the right boundary condition to be stress free (i.e.,  $g \equiv 0$ ). Specifically, the stress-strain relationship in [22] is described by the nonlinear form

$$\sigma(x, t) = E_1 u_{xt}(x, t) + \int_0^t \Gamma(t - s; \mathcal{F}) \frac{d}{ds} \sigma^e(u_x(x, s)) ds. \quad (3.1.4)$$

Here  $\sigma^e$  is the elastic response function, and  $\Gamma$  is defined by

$$\Gamma(t; \mathcal{F}) = \int_{\mathcal{T}} \gamma(t, \tau) d\mathcal{F}(\tau),$$

where  $\mathcal{T} \subset (0, \infty)$ ,  $\gamma$  is a function of  $t$  and relaxation time  $\tau$ , and  $\mathcal{F}$  is a probability distribution function of relaxation time  $\tau$ . The existence and uniqueness of weak solution for this general viscoelastic model was given in [22], and the continuous dependence of the weak solution on the probability distribution function  $\mathcal{F}$  was also given in [22] under a Prohorov metric framework [8, 10, 17, 22, 78] on the space of probability distributions. These results on continuous dependence can be readily extended to the other parameters (e.g.,  $E_1$ ) of interest here.

Note that assumptions (A1) and (A2) conform with the ones made in [22]. Hence, the arguments for Theorem 3.1.2 are similar to those given in [22, 25]. Here we only sketch the ideas, referring readers to [8, 22, 25, 99] for further details. Let  $\{\psi_j\}_{j=1}^{\infty}$  be any linearly independent total subset of  $\mathbb{V}$ . We define the Galerkin approximation  $u^m(t) = \sum_{j=1}^m \beta_j^m(t) \psi_j$  as the unique solution of

$$\rho \langle u_{tt}^m, \psi_j \rangle_{\mathbb{V}^*, \mathbb{V}} + E_1 \langle u_{tx}^m, \psi_j' \rangle + E_0 \left\langle \int_0^t P(t-s) \frac{d}{ds} u_x^m(s) ds, \psi_j' \right\rangle + g(t) \psi_j(L) = 0, \quad j = 1, 2, \dots, m,$$

on the interval  $[0, T]$ , where  $\psi_j'$  denotes the derivative of  $\psi_j$  with respect to  $x$ ,  $j = 1, 2, \dots, m$ . We can argue that  $\{u^m\}$  and  $\{u_t^m\}$  are bounded uniformly in  $\mathcal{L}^2(0, T; \mathbb{V})$ . Then by the Banach-Alaoglu theorem we know that there exists a function  $u \in \mathcal{L}^2(0, T; \mathbb{V})$  such that

$$u^m \rightarrow u \text{ weakly in } \mathcal{L}^2(0, T; \mathbb{V}),$$

$$u_t^m \rightarrow u_t \text{ weakly in } \mathcal{L}^2(0, T; \mathbb{V}).$$

In addition, the following convergence results can be proven by using the Ascoli-Arzelà theorem (e.g., see [51, Theorem 3.6.4]) and Aubin's lemma (e.g., see [42, Lemma 8.4])

$$u^m \rightarrow u \text{ weakly in } \mathbb{V} \text{ uniformly in } t \in [0, T], \text{ i.e., } u^m \rightarrow u \text{ in } \mathcal{C}_w(0, T; \mathbb{V}),$$

$$u_t^m \rightarrow u_t \text{ weakly in } \mathbb{H} \text{ uniformly in } t \in [0, T], \text{ i.e., } u_t^m \rightarrow u_t \text{ in } \mathcal{C}_w(0, T; \mathbb{H}),$$

$$u_t^m \rightarrow u_t \text{ in } \mathcal{L}^2(0, T; \mathbb{H}).$$

We can then show that  $u_x^m(t) \rightarrow u_x(t)$  in  $\mathbb{H}$ . Based on these convergence results,  $u$  can be easily shown to be a weak solution of system (3.1.1)-(3.1.2). The uniqueness of the weak solution can be established by using a standard technique which demonstrates that the difference between any two possible solutions must be zero [99].

### 3.1.2 Internal Variable Form of the Stress Relaxation Function

Note that the origin for time is assumed at the beginning of motion and loading. Hence, we can rewrite (3.1.2) into the following form using integration by parts

$$\sigma(x, t) = E_1 u_{xt}(x, t) + E_0 \left( P(0) u_x(x, t) - \int_0^t \left( \frac{d}{ds} P(t-s) \right) u_x(x, s) ds \right). \quad (3.1.5)$$

For the remainder of this chapter, the form of the stress relaxation function  $P(t)$  is assumed to be a Prony series

$$P(t) = p_0 + \sum_{j=1}^{N_p} p_j e^{-t/\tau_j}, \quad (3.1.6)$$

where all the  $p_j$  are nonnegative numbers and the  $\tau_j$  values are positive, and with  $N_p$  being a positive integer. This series is based on the assumption that relaxation in a viscoelastic material can be well represented by a discrete number of relaxation times  $\tau_j$ . Without loss of generality, we will also enforce  $P(0) = 1$ . A result of this constraint is that  $\sum_{j=0}^{N_p} p_j = 1$ . It is worth noting here that this special form of  $P$  satisfies assumption (A2). Hence, system (3.1.1) and (3.1.5) with  $P$  given by (3.1.6) also has a unique weak solution on any finite time interval  $[0, T]$ .

If we replace  $\frac{d}{ds} P(t-s)$  in (3.1.5) with the  $s$ -derivative of the Prony series at  $t-s$ , we obtain

$$\sigma(x, t) = E_1 u_{xt}(x, t) + E_0 \left( u_x(x, t) - \sum_{j=1}^{N_p} \int_0^t \frac{p_j}{\tau_j} e^{-(t-s)/\tau_j} u_x(x, s) ds \right).$$

We can reformulate the integrals related to each internal variable as differential equations which we can solve simultaneously with the main system (3.1.1). To this end, we define the “internal variables”

$$\epsilon^j = \int_0^t \frac{p_j}{\tau_j} e^{-(t-s)/\tau_j} u_x(s) ds.$$

Then the time derivative of  $\epsilon^j$  is given by

$$\epsilon_t^j = \frac{p_j}{\tau_j} u_x(t) - \frac{1}{\tau_j} \int_0^t \frac{p_j}{\tau_j} e^{-(t-s)/\tau_j} u_x(s) ds.$$

Relating  $\epsilon^j$  and  $\epsilon_t^j$  allows us to model the internal variables dynamically as

$$\begin{aligned} \tau_j \epsilon_t^j + \epsilon^j &= p_j u_x, \\ \epsilon^j(0) &= 0. \end{aligned} \quad (3.1.7)$$

for  $j = 1, 2, \dots, N_p$ . We can then write the overall stress-strain relationship (3.1.5) as

$$\sigma = E_1 u_{xt} + E_0 \left( u_x - \sum_{j=1}^{N_p} \epsilon^j \right). \quad (3.1.8)$$

Note that even though  $p_0$  is an element in the Prony series for  $P(t)$ , once the series is substituted in the model the constant  $p_0$  no longer appears. However,  $p_0$  is still present in the sum-to-one constraint on *all*  $p_i$  values, but we can easily work with the alternate constraint that the remaining  $p_j$  terms must satisfy  $\sum_{j=1}^{N_p} p_j \leq 1$ .

The damping and internal variables provide us the future flexibility to match the model to data from the experimental devices, and also present an interesting question of identifiability of the damping and internal variable parameters which we will discuss in depth. Note also that the authors in [9, 67, 80] give computational results (using essentially an equivalent model from a slightly different conceptual formulation involving a distribution of relaxation mechanisms-see (3.1.4)) showing that discrete relaxation times can model well the viscoelastic material responses of the type we consider in this work (namely, attempting to approximate the response of biological soft tissue as characterized in [9, 52]). In fact, in previous work no more than two discrete relaxation times were used, which has informed our decision to allow a maximum of two relaxation times.

Since the ultimate goal of the wider research project will be examining the traction into the chest cavity that results from a healthy artery experiencing a heartbeat as compared with an artery containing a stenosis experiencing a heartbeat, our nonzero boundary input  $g(t)$  will here be represented by an approximation to a pulse traction. In order to ensure a smooth, compactly supported input, we implement the input function as a Van Bladel function which is a good approximation to expected perturbation inputs to our system [9, 91]. This smoothness is useful in order to get the maximum benefit from using high order numerics. The function used is

$$g(t) = \begin{cases} A \cdot \exp\left(\frac{|ab|}{t(t+a-b)}\right) & \text{if } t \in (0, b-a), \\ 0 & \text{otherwise,} \end{cases} \quad (3.1.9)$$

where  $A$  is some positive constant,  $a$  and  $b$  are some constants with  $b > a$ .

### 3.1.3 Parameter Values Used in Simulations

Motivated by the experimental data to which we intend to apply this methodology, we choose values for the system parameters which simulate low-amplitude (on the order of  $0.1mm$ ) oscillatory motions. For data generation, we will use two internal variables ( $N_p = 2$ ). The weights  $p_i$  for our two relaxation time model will be fixed as given below. The baseline material

parameter values chosen for this chapter are as follows:

$$\begin{aligned} E_0 = 2.2 \times 10^5 \text{ Pa}, \quad E_1 = 40 \text{ Pa} \cdot \text{s}, \quad \rho = 1010 \text{ kg/m}^3, \quad L = 0.053 \text{ m} \\ \tau_1 = 0.05 \text{ s}, \quad \tau_2 = 10 \text{ s}, \quad p_1 = 0.3, \quad p_2 = 0.55. \end{aligned} \quad (3.1.10)$$

Note that the density  $\rho = 1010 \text{ kg/m}^3$  is the true density of the agar gel that is used in the medium for our experiments at QMUL, and  $L = 0.053 \text{ m}$  is the true height of the phantom. These are parameter values which are directly taken to approximate the experimental device. The values for  $E_0$  and  $E_1$  and for the relaxation times are physically reasonable parameters based on a perusal of the viscoelastic materials research literature and are also informed by our early experiments with the agar gels. In the Van Bladel function, the values of  $a$  and  $b$  have an effect on pulse width as well as the amplitude, while the value of  $A$  only has an effect on the amplitude. Their baseline values are chosen as follows:

$$a = 6 \times 10^{-3}, \quad b = 20 \times 10^{-3}, \quad A = 6 \times 10^3, \quad (3.1.11)$$

where the values  $a$  and  $b$  allow for an effective pulse application time of 14 ms. We show the resulting function in Figure 3.1.

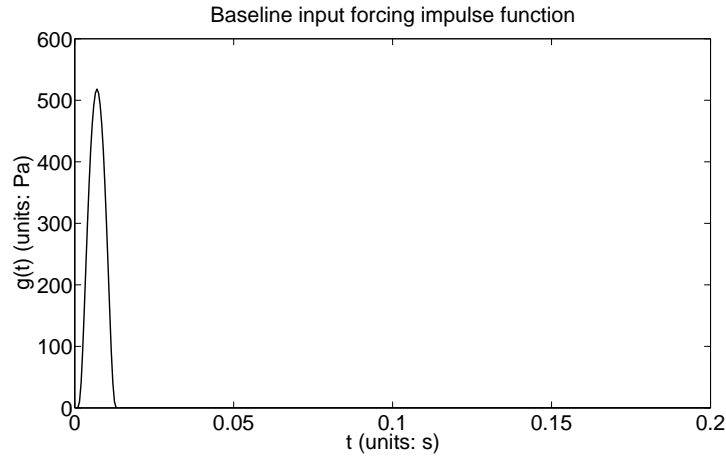


Figure 3.1: Van Bladel function (c.f. (3.1.9)) with  $A = 6 \times 10^3$ ,  $a = 6 \times 10^{-3}$ , and  $b = 20 \times 10^{-3}$ .

### 3.1.4 Direct (Forward) Problem

Altogether, this description (3.1.1) along with (3.1.7)-(3.1.9) encompasses the one-dimensional model for the displacement  $u(x, t)$  that will be studied in both major sections of this chapter. In

other words, system (3.1.1) along with (3.1.7)-(3.1.9) with given parameter values is the *direct or forward problem*. Throughout this chapter, we solve (3.1.1), (3.1.7)-(3.1.9) for  $u(x, t)$  with given parameter values using a spectral continuous finite element method in space (Gauss-Lobatto nodes) and a discontinuous Galerkin method in time. The numerical scheme is specially tailored to allow for high order space-time discretization in order to control dispersion errors and will be documented fully in [61] along with its convergence properties.

## 3.2 Estimation of Material Parameters: Preliminaries

In this section, we examine an inverse problem methodology for estimating material parameters (and thus gain a sense for our ability to characterize an individual's material properties) with given simulated observations of displacement at the  $x = L$  position (where the simulated data is generated under various measurement noise conditions, see details below). In addition to determining an *estimate* for material parameters, we also need to determine our *confidence in the estimation procedure*. To this end, we will compare two techniques for determining confidence intervals, specifically the asymptotic theory discussed in [12, 24] versus using bootstrapping as discussed in [12, 45].

In practice, one will obtain a set of experimental data and then one needs to determine how many (if any) relaxation times are required to represent well the data. Thus, we will want to compare the performance of our estimation procedures on three competing models. In each model, we will always estimate  $E_0$  and  $E_1$  (assuming given values for  $\rho$  and  $L$  in (3.1.10) along with given values of  $a$ ,  $b$  and  $A$  in (3.1.11)), but we will vary the number of relaxation times incorporated into the model. The three models to be used in inverse problems are as follows:

1. For a model with no relaxation times, we do not include any  $\tau_i$  or corresponding  $p_i$  in the model. Thus, we estimate only  $\bar{q} = (E_0, E_1)^T$ .
2. In the case with one relaxation time, we incorporate a single internal variable (i.e.,  $N_p = 1$ ). For this case,  $\bar{q} = (E_0, E_1, \tau_1)^T$  will be estimated, but the corresponding material weight  $p_1$  is fixed to be  $p_1 = 0.3$ .
3. For the case of two relaxation times (i.e.,  $N_p = 2$ ), we will estimate  $\bar{q} = (E_0, E_1, \tau_1, \tau_2)^T$  with the corresponding material weights fixed to be the values in (3.1.10), that is,  $p_1 = 0.3$  and  $p_2 = 0.55$ .

Note that for the models with one and two relaxation times the corresponding weights  $p_i$  are fixed. Though in reality one would certainly need to estimate the weights  $p_i$ , we take the liberty here of assuming them to be known so we can focus on the general methodology and in particular the reliability in identifying the relaxation times.



Considering this set of models will allow us to follow what one would consider in practice, examining the results of adding or subtracting model features. It is worth noting here that for this particular set of models, the one-relaxation-time model (the second case) is not a special case of the two-relaxation-time model (the third case) as the material weight  $p_2$  in the two-relaxation-time model is fixed, and that the zero-relaxation-time model (the first case) is not a special case of the one-time-relaxation model as the material weight  $p_1$  in the one-relaxation-time model is fixed. However, if we allow the corresponding material weights to be free (i.e., to be estimated along with relaxation times), then the zero-relaxation-time model is indeed a special case of the one-relaxation-time model, and the one-relaxation-time model is a special case of the two-relaxation-time model. We will therefore use the sensitivity equations and parameter estimation results as well as model selection criterion to suggest the number of relaxation times needed in practice.

### 3.2.1 Study of Effects of Changing Material Parameters

Before discussing simulated data and actually solving the inverse problem, we wish to complete some analysis on the model around the true material parameter values (3.1.10). It is clear that changing the amplitude factor  $A$  for the Van Bladel input will change the resulting amplitude of the system. Hence, we consider here changes in the material parameters  $E_0$ ,  $E_1$ , and  $\tau_j$  values.

We first consider changes in the stiffness  $E_0$  and damping factor  $E_1$ . As an example of typical effects of changing parameters, we show the effects of reducing stiffness to  $E_0 = 200,000$  in the left pane of Figure 3.2 and in the right pane the effects of increasing the damping to  $E_1 = 60$ . Changes in  $E_0$  are shown to have a significant effect on the oscillation frequency, as well as a

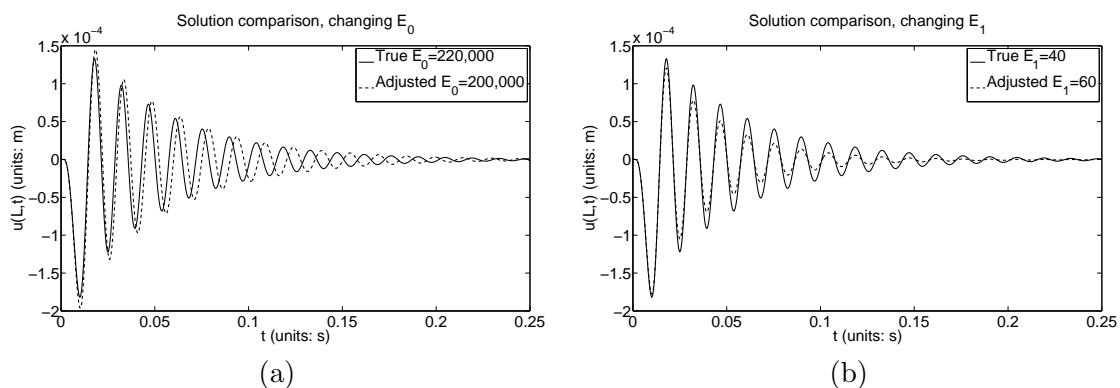


Figure 3.2: Solution of (3.1.1) with (3.1.7)-(3.1.8) using the “true” parameters (3.1.10) and forcing function with parameters  $a = 6 \times 10^{-3}$ ,  $b = 20 \times 10^{-3}$ , and  $A = 6 \times 10^3$  (depicted by the solid line), alongside solutions using  $E_0 = 2 \times 10^5$  (a) and  $E_1 = 60$  (b) which are represented with dashed lines in their respective graphs.

minor effect on peak heights. This makes sense – a more stiff material will propagate waves more quickly and will dissipate energy less slowly. Changes in damping,  $E_1$ , lead to the very expected effects that the energy dissipates faster in the material, so the oscillation peak heights become smaller and the material experiences fewer small oscillations in later simulation times. So, these two parameters seem to govern the major properties of how the material responds to an impulse response traction. Also, changes to the model solution induced by changes in  $E_0$  and  $E_1$  are generally on the same scale as the model solution itself (i.e., the changes are on the order of  $10^{-4}$ ).

Relaxation times can allow the model flexibility in matching the periodic local “peaks” and “troughs” in the oscillating solution. For example, if we change from the baseline  $\tau_1 = 0.05$  to  $\tau_1 = 0.5$ , the system experiences the changes shown in Figure 3.3. This response to changing

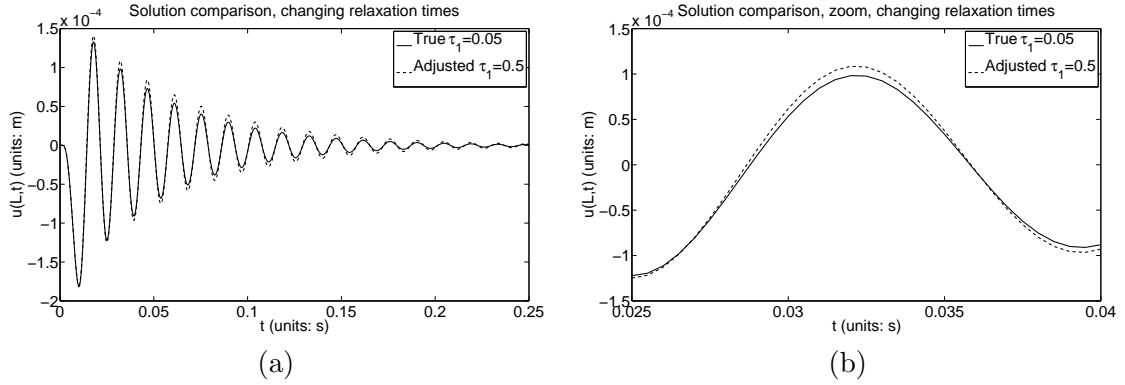


Figure 3.3: Solution of (3.1.1) with (3.1.7)-(3.1.8) using the “true” parameters (3.1.10) and forcing function with parameters  $a = 6 \times 10^{-3}$ ,  $b = 20 \times 10^{-3}$ , and  $A = 6 \times 10^3$  (depicted by the solid line), alongside the dashed line using  $\tau_1 = 0.5$  with the remaining parameters the same in (a). In (b), the solution is zoomed in for  $t \in [0.02, 0.04]$ .

relaxation times represents a typical example of changing either  $\tau_1$  or  $\tau_2$ . However, note the scale of the changes: the maximum difference between the solutions shown in Figure 3.3 is  $1.0996 \times 10^{-5}$ . As we will see later when adding noise, the noise itself is on the scale of  $10^{-5}$ . This foreshadows the difficulties in properly estimating relaxation times that we will see going forward. This will be evident also both in a discussion on using different optimization routines and in a discussion of model sensitivity with respect to parameters.

### 3.2.2 Sensitivity of Model Output with Respect to Material Parameters

In order to further quantify the model response to changes in parameters around the baseline values of (3.1.10), we will examine the sensitivity of the model output  $u(L, t)$  with respect to material parameters. Note that since the values of parameters are on such a varying scale, we will actually work with the log-scaled versions of the material parameters we are attempting to estimate. In other words, if  $\bar{q} = (E_0, E_1, \tau_1, \tau_2)^T$  is the vector of parameters to be estimated, we define  $q = \log_{10}(\bar{q})$ .

Sensitivity analysis has been widely used in inverse problem investigations (e.g., see [24] and the references therein for details) to identify the model parameters and/or initial conditions to which the model outputs are most sensitive and for which one can readily construct confidence intervals when they are estimated (i.e., which are the most reliably estimated values). To compute the sensitivity of the model output to each parameter, one needs to find sensitivity equations which describe the time evolution of the partial derivatives of the model state with respect to each parameter. Sensitivity equations in terms of the non-log-scaled parameters  $\bar{q}$  are derived in Appendix A.

We can use the sensitivity of model output to the non-log-scaled parameters to find the sensitivity of model output with respect to the log-scaled parameters, which will be of interest here. Using the chain rule, we find that

$$\frac{\partial u(L, t; 10^q)}{\partial q_i} = \bar{q}_i \ln(10) \frac{\partial u(L, t; \bar{q})}{\partial \bar{q}_i},$$

where  $q_i$  and  $\bar{q}_i$  are the  $i$ th elements of  $q$  and  $\bar{q}$ , respectively.

The sensitivities of model output with respect to parameters  $(\log_{10}(E_0), \log_{10}(E_1), \log_{10}(\tau_1), \log_{10}(\tau_2))$  are depicted in Figure 3.4. From this figure we see that model output is most sensitive to  $\log_{10}(E_0)$ , sensitive to  $\log_{10}(E_1)$ , less sensitive to  $\log_{10}(\tau_1)$ , and only minimally sensitive to  $\log_{10}(\tau_2)$ . The most interesting feature related to our study is the fact that the scale of sensitivity of model output to the first relaxation time is on the order of  $10^{-5}$  whereas the sensitivity of model output to the second relaxation time is roughly two orders of magnitude smaller on the order of  $10^{-7}$ . We will later see that, while we have difficulty estimating both relaxation times due to the small changes they induce in the model solution (as previously discussed), we especially have difficulty obtaining a reasonable estimate for  $\tau_2$  because the model is so much less sensitive to the second relaxation time than to the first.

Figure 3.5 demonstrates the sensitivities of model output with respect to material proportions  $\log_{10}(p_1)$  and  $\log_{10}(p_2)$ . From this figure we see that the model is less sensitive to the second proportion than to the first one. Though these will be fixed and not estimated in this chapter, when we update the model in Chapter 4 we will estimate the parameters analogous to the

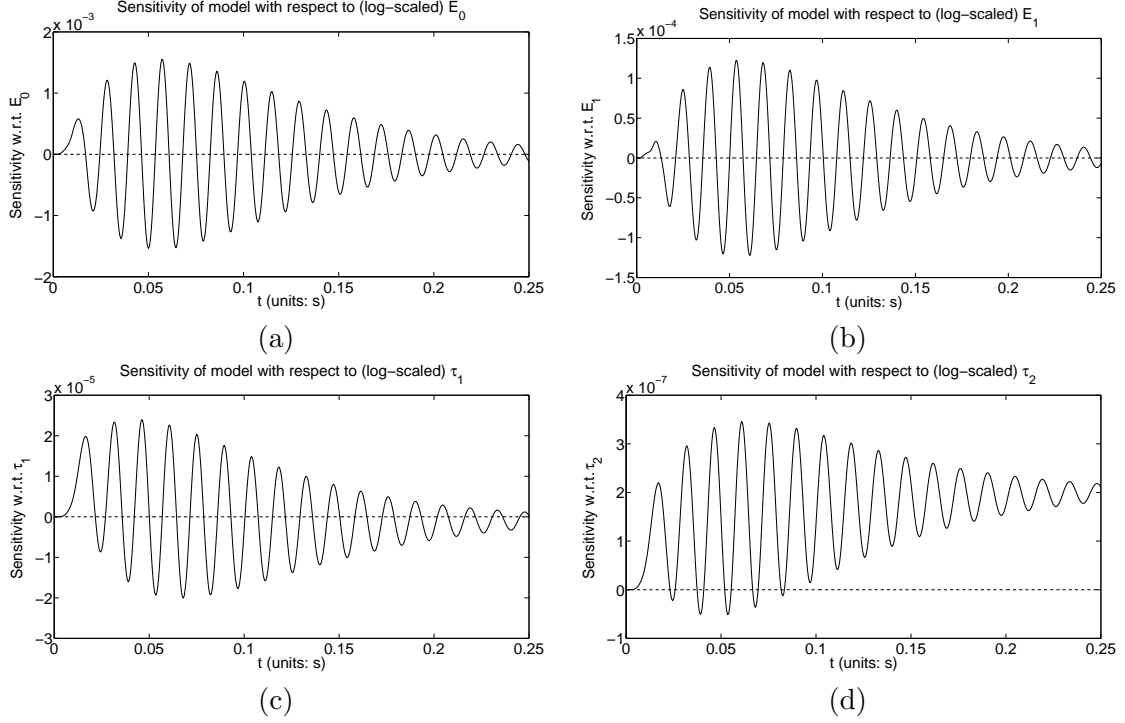


Figure 3.4: (a) Sensitivity of model output with respect to  $\log_{10}(E_0)$ ; (b) Sensitivity of model output with respect to  $\log_{10}(E_1)$ ; (c) Sensitivity of model output with respect to  $\log_{10}(\tau_1)$ ; and (d) Sensitivity of model output with respect to  $\log_{10}(\tau_2)$ . All sensitivities are around the baseline parameters (3.1.10) and (3.1.11).

proportions in Chapter 5. The results here indicate we can predict that accurately estimating a second material proportion will be much more difficult than estimating the first one.

Armed with our knowledge of sensitivities of model output with respect to the material parameters around the true parameter values (3.1.10), and our knowledge of effects on the model solution of changing the parameters, we next describe the generation of our simulated data and discuss solving the inverse problem.

### 3.3 Statistical Model and Inverse Problem

We will work with simulated data for various noise levels generated at position  $x = L$ , namely data  $u_j$  corresponding to the model solution  $u(L, t_j)$  at measurement time points  $t_j, j = 0, 1, \dots, n-1$ . Then the *statistical model* (a model used to describe the observation process) is assumed to take the following form

$$U_j = u(L, t_j; 10^{q_0}) + \mathcal{E}_j, \quad j = 0, 1, \dots, n-1, \quad (3.3.1)$$

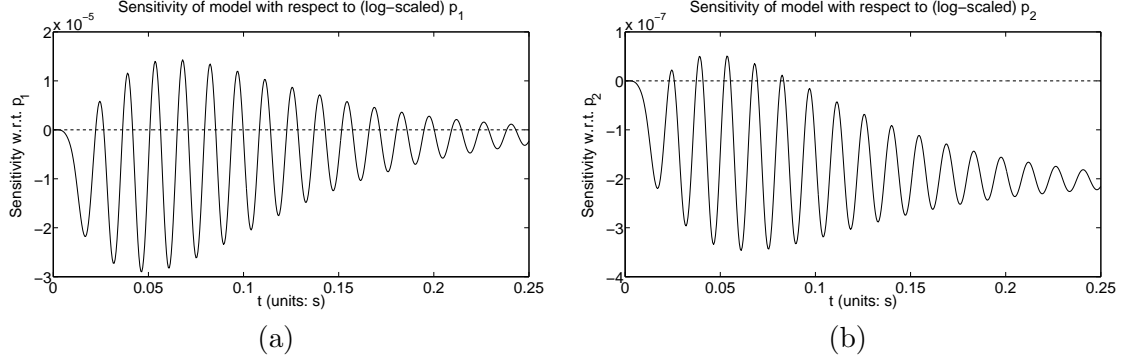


Figure 3.5: (a) Sensitivity of model output with respect to  $\log_{10}(p_1)$ . (b) Sensitivity of model output with respect to  $\log_{10}(p_2)$ . Both sensitivities around the baseline parameters (3.1.10) and (3.1.11).

where  $u(L, t_j; 10^{q_0})$  is the solution to (3.1.1) along with (3.1.7)-(3.1.9) at  $x = L$  with a given set of “true” material parameter values  $q_0$  and the values of the rest of parameters given in (3.1.10) and (3.1.11). Here  $\mathcal{E}_j$  denotes the measurement error (a random variable) at measurement time point  $t_j$ ,  $j = 0, 1, \dots, n-1$ . It is worth noting that  $U_j$ ,  $j = 0, 1, 2, \dots, n-1$ , are also random variables due to the randomness of measurement errors. For the current proof of concept discussion, we will assume the measurement errors  $\mathcal{E}_j$ ,  $j = 0, 1, \dots, n-1$ , are independent and identically distributed with mean zero ( $E(\mathcal{E}_j) = 0$ ) and constant variance  $\text{var}(\mathcal{E}_j) = \sigma_0^2$ . We thus are assuming absolute additive error; this is reasonable as an initial error model for our proof of concept investigations. We do not make further assumptions about the distributions of the  $\mathcal{E}_j$  in order to carry out the inverse problem methodology or the asymptotic analysis below. When we later apply the AIC comparison methodology there is the tacit assumption of normality on the  $\mathcal{E}_j$ .

Under these assumptions for the measurement errors in the statistical model (3.3.1), the estimator  $\mathcal{Q}$  of  $q$  can be obtained by using the ordinary least squares method

$$\mathcal{Q} = \arg \min_{q \in Q} \sum_{j=0}^{n-1} [U_j - u(L, t_j; 10^q)]^2, \quad (3.3.2)$$

where  $Q \subset \mathbb{R}^\kappa$  is some viable admissible parameter set, assumed compact in  $\mathbb{R}^\kappa$  with  $\kappa$  being the number of parameters requiring estimation. Thus,  $\mathcal{Q}$  can be viewed as a minimizer that minimizes the distance between the data and the model where all observations are treated with equal importance. Note that under different error assumptions, one would need to modify the cost function in (3.3.2) (a topic discussed in, e.g., [17, 24, 43] as well as Chapter 5) for an appropriate asymptotic parameter distribution theory to be valid.

Since  $\mathcal{Q}$  is a random variable (inherited from the fact that  $U_j$  are random variables), we can define its corresponding realizations  $\hat{q}$  by minimizing the cost function

$$\mathcal{J}(q) = \sum_{j=0}^{n-1} [u_j - u(L, t_j; 10^q)]^2$$

over the set  $\mathcal{Q}$ . That is,  $\hat{q}$  is obtained by solving the following inverse problem

$$\hat{q} = \arg \min_{q \in \mathcal{Q}} \mathcal{J}(q) = \arg \min_{q \in \mathcal{Q}} \sum_{j=0}^{n-1} [u_j - u(L, t_j; 10^q)]^2, \quad (3.3.3)$$

Here  $u_j$  is a corresponding realization of  $U_j$ , and it is given by

$$u_j = u(L, t_j; 10^{q_0}) + \epsilon_j, \quad j = 0, 1, \dots, n-1. \quad (3.3.4)$$

with  $\epsilon_j$  being realizations of  $\mathcal{E}_j$ ,  $j = 0, 1, \dots, n-1$ . Note that the model solution  $u(x, t)$  is continuously dependent on the model parameter  $q$  (recall the remarks in Section 3.1.1). Hence,  $\mathcal{J}$  is a continuous function of  $q$ . Since  $\mathcal{Q}$  is assumed to be compact, the inverse problem (3.3.3) has a solution. Since for our studies we will be interested in perturbations around nominal values  $10^{q_0}$  of parameters and the corresponding solutions, our test problems will not in general suffer from serious ill-posedness and some type of stability or regularization techniques (Tikhonov regularization, regularization by discretization, etc., [18]) are not required for our studies. This will not necessarily be the case when using the inverse problem methods studied here with experimental data.

Estimating material parameters  $\hat{q}$  from given sets of data with different noise levels, as well as *quantifying uncertainty* in our estimates, will be the key focus of our work in this section. We use the values for  $E_0$ ,  $E_1$ ,  $\tau_1$ , and  $\tau_2$  in (3.1.10) in their log scaled form as the true values  $q_0$  used to simulate data. That is,

$$q_0 = (5.3424, 1.6021, -1.3010, 1)^T = (\log_{10}(2.2 \times 10^5), \log_{10}(40), \log_{10}(.05), \log_{10}(10))^T.$$

As previously discussed, we will find parameter estimates for models with zero, one, and two relaxation times in the model itself (and thus the number of parameters estimated changes). In all cases, the parameters belong to a viable compact set  $\mathcal{Q}$  with the upper and lower bounds on parameters being taken (in educated guesses) as  $q_{lb} = (-15, -15, -15, -15)^T$ ,  $q_{ub} = (7.3010, 2.3010, 2, 2)^T$  for two relaxation times estimation,  $q_{lb} = (-15, -15, -15)^T$ ,  $q_{ub} = (7.3010, 2.3010, 2)^T$  for the one relaxation time estimation, and  $q_{lb} = (-15, -15)^T$ ,  $q_{ub} = (7.3010, 2.3010)^T$  for the no relaxation time estimation.

### 3.3.1 Simulated Data Generation

We will simulate data using two relaxation times (and a question of interest later will be how many of those relaxation times we can recover) with the values of parameters given in (3.1.10) and (3.1.11). The measurement time points are taken at  $t_j = 0.001j$ ,  $j = 0, 1, \dots, 250$ . Thus, there are a total of  $n = 251$  data points. We note that noiseless data has maximum amplitude on the order of  $10^{-4}$  (depicted by the solid line in Figure 3.6), which was again motivated by the anticipated scale of results from the experimental device. This level informs the magnitude we choose for the additive noise. We represent a “low” noise level with  $\sigma_0^2 = 5 \times 10^{-6}$ , a “medium” noise level by  $\sigma_0^2 = 10 \times 10^{-6}$ , and a “high” noise level by taking  $\sigma_0^2 = 20 \times 10^{-6}$ . In Figure 3.6, we show plots corresponding to the three levels of noisy simulated data against the system dynamics

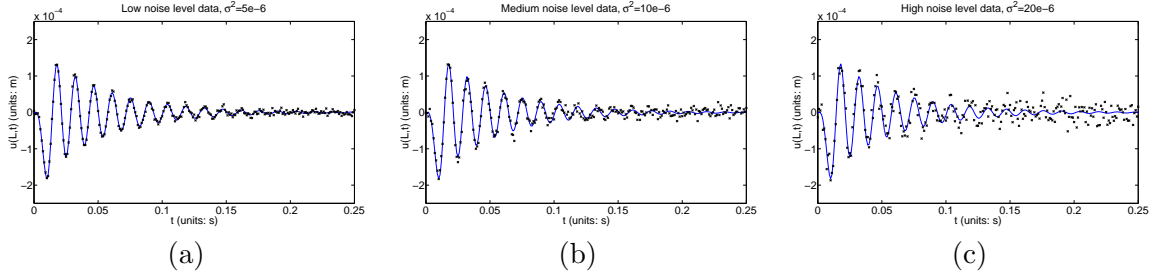


Figure 3.6: Simulated noisy data around the true parameter values. (a) Low noise level. (b) Medium noise level. (c) High noise level.

corresponding to the true parameters  $q_0$ . Noise is assumed absolute for our initial investigations (in Chapter 5 we will later also explore a relative noise error model for our experimental data), and is added according to the error model (3.3.1). Low noise results in data mostly along the trajectory of the true model. Medium noise begins to obfuscate the later-time oscillations which have lost much of their earlier energy. High noise significantly affects the level of peaks and troughs from  $t = 0.05$  forward. We thus obtain a series of increasingly difficult problems in obtaining material parameter estimates, though entirely expected since higher noise tends to significantly affect data features and presents a more difficult parameter estimation problem.

### 3.3.2 Parameter Estimates Obtained Using Different Routines

In this section, we discuss different options for the optimization routine used to solve the inverse problem (3.3.3), and begin to gain a sense of the robustness of parameter estimation with respect to the optimization routine. Note that we expect to have some difficulty in relaxation time estimation, based on our earlier discussion on the model response to changes in relaxation times

as well as model sensitivities. We do expect to obtain more accurate estimates for  $E_1$ , and very good estimates for  $E_0$ . To begin this discussion, we will examine parameter estimates for a model which incorporates two relaxation times.

The optimization routines we compare are all built-in Matlab Optimization Toolbox routines. We use `fmincon` with active-set optimization, which treats the optimization as constrained nonlinear programming with our cost function  $\mathcal{J}(q)$ . We also examine the use of `lsqnonlin`, which is designed for nonlinear least squares data-fitting problems; our cost function is exactly the form of a nonlinear least squares function. We test both the Levenburg-Marquardt (LM) option and the trust-region-reflective (TRR) option. Note that the Levenburg-Marquardt algorithm does *not* allow bound constraints; we tried the routine out of curiosity, to see if it would produce unrealistic estimates of any parameters (it does at high noise levels).

Results from optimizing for  $q$  by using different optimization routines are shown in Table 3.1. All optimization runs used the initial guess

$$q_{init} = (\log_{10}(1.8 \times 10^5), \log_{10}(60), \log_{10}(0.5), \log_{10}(20))^T = (5.2553, 1.7782, -0.3010, 1.3010)^T.$$

This table includes the parameter estimates  $\hat{q}$ , computation time (CPU) in seconds for that

Table 3.1: Estimation of material parameters at low, medium and high noise levels: Comparison between optimization routines (`lsq-TRR=lsqnonlin` with the trust-region-reflective option, `lsq-LM=lsqnonlin` with the Levenburg-Marquardt option).

Noise level	Routines	Estimated parameter values $\hat{q}$	CPU (s)	RSS
Low	<code>fmincon</code> :	$(5.3422, 1.6581, -0.3000, 1.3012)^T$	194.16	$6.7618 \times 10^{-9}$
	<code>lsq-TRR</code> :	$(5.3425, 1.6046, -1.2297, 1.0046)^T$	347.09	$6.2458 \times 10^{-9}$
	<code>lsq-LM</code> :	$(5.3425, 1.6044, -1.2309, 1.0316)^T$	613.47	$6.2458 \times 10^{-9}$
Medium	<code>fmincon</code> :	$(5.3430, 1.6583, -0.2998, 1.3012)^T$	203.75	$2.4435 \times 10^{-8}$
	<code>lsq-TRR</code> :	$(5.3433, 1.5889, -1.3269, 2.0000)^T$	241.51	$2.3647 \times 10^{-8}$
	<code>lsq-LM</code> :	$(5.3433, 1.5893, -1.3252, 5.9303)^T$	608.27	$2.3646 \times 10^{-8}$
High	<code>fmincon</code> :	$(5.3433, 1.6380, -0.2995, 1.3012)^T$	238.58	$1.03291 \times 10^{-7}$
	<code>lsq-TRR</code> :	$(5.3433, 1.6361, -1.990, 0.2496)^T$	606.36	$1.03257 \times 10^{-7}$
	<code>lsq-LM</code> :	$(5.3433, 1.6351, -0.02324, 0.0003611)^T$	1110.83	$1.03248 \times 10^{-7}$
true values $q_0$ :		$(5.3424, 1.6021, -1.3010, 1.0000)^T$		

particular optimization run, and the residual sum of squares (RSS) defined as

$$\text{RSS} = \sum_{j=0}^{n-1} \left[ u_j - u(L, t_j; 10^{\hat{q}}) \right]^2.$$



Overall, we observe from Table 3.1 that the routines do a good job of estimating  $E_0$  (as we expected). The `lsqnonlin` routines tend to better estimate  $E_1$ . As for relaxation times, we begin to see a major flaw in use of the `fmincon` routine. It does not seem particularly sensitive to the relaxation times, and the resulting estimates of the relaxation times stay near the initial guess. The `fmincon` routine produced similar non-responsive results for different initial guesses. The `lsqnonlin` routines estimate both the relaxation times well in the presence of low noise. At medium noise, the routines estimate  $\tau_1$  well but not  $\tau_2$ . At high noise, relaxation time estimation is poor. This will be quantified further in the following sections on error analysis.

Even though there might be some spurious computation times on desktop machines (due to other background programs), we still include them here in Table 3.1 to demonstrate typical optimization routine performance. Consistently, `fmincon` was the fastest routine. This is in part due to the fact that this routine alone of the three supports parallel computation, so on our multi-core desktop machines we were able to see a speed-up. However, the computation times for the trust-region-reflective `lsqnonlin` algorithms are reasonable. Using Levenburg-Marquardt consistently is the slowest method, and the results are not better than those using trust-region-reflective `lsqnonlin` algorithm.

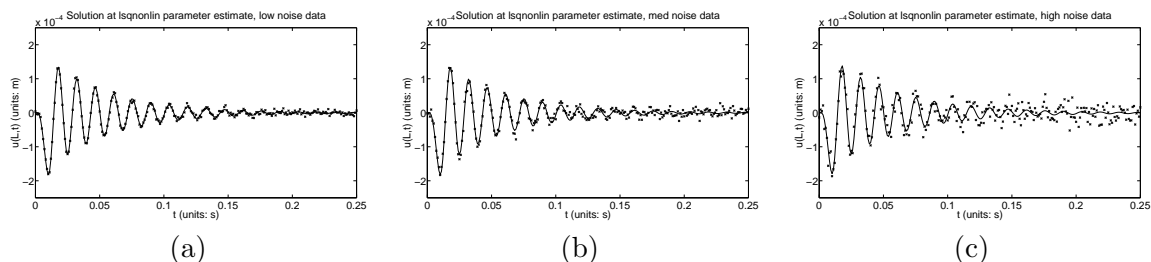


Figure 3.7: Data and two-relaxation-time model solutions at parameter estimates obtained using the `lsqnonlin` routine, trust-region-reflective method, at different noise levels (see Table 3.1). (a) Low noise. (b) Medium noise. (c) High noise.

As a result, we recommend using the trust-region-reflective `lsqnonlin` algorithm when trying to estimate relaxation times (this is also the routine that we use in the remainder of this dissertation). If the model does not contain relaxation times (i.e., only estimating  $E_0$  and  $E_1$ ), the speedup afforded by using `fmincon` may make that algorithm the one of choice. For reference, Figure 3.7 illustrates model fits to the data at different noise levels, where the model solution is calculated with the values of model parameters obtained through `lsqnonlin` TRR routine. We see in all cases that the model solution provides reasonable fits to the data.

### 3.4 Asymptotic Error Analysis

Most asymptotic error theory [12, 24, 83] is described in the context of an ODE model example  $\dot{z}(t) = f(z(t); q)$ . However, we can use the PDE sensitivities of the model output with respect to each parameter in  $q$ , namely  $\frac{\partial u(L, t; 10^q)}{\partial q_i}$ , in a similar manner to the ODE sensitivities in the asymptotic theory. The steps of the asymptotic theory error analysis are as follows (the theory for the following steps is described in [12, 24, 83]).

1. Determine  $\hat{q}$  by solving the inverse problem (3.3.3).
2. Compute the sensitivity equations to obtain  $\frac{\partial u(L, t; 10^{\hat{q}})}{\partial q_i}$  (as discussed in Section 3.2.2) for  $i = 1, \dots, \kappa$  where  $\kappa$  is the number of parameters being estimated. The sensitivity matrix  $\chi(\hat{q})$  can then be calculated with its entries

$$\chi_{j,i}(\hat{q}) = \frac{\partial u(L, t_j; 10^{\hat{q}})}{\partial q_i}, \quad j = 0, 1, \dots, n-1, \text{ and } i = 1, \dots, \kappa.$$

Note that  $\chi(\hat{q})$  is then an  $n \times \kappa$  matrix. We can also obtain an estimate for the constant variance  $\sigma_0^2$  as

$$\hat{\sigma}^2 = \frac{1}{n - \kappa} \sum_{j=0}^{n-1} \left( u_j - u(L, t_j; 10^{\hat{q}}) \right)^2.$$

3. Asymptotic theory yields that the estimator  $\mathcal{Q}$  is asymptotically (as sample size  $n \rightarrow \infty$ ) normal with mean approximated by  $\hat{q}$  and the covariance matrix approximated by

$$Cov(\mathcal{Q}) \approx \hat{\Sigma} = \hat{\sigma}^2 [\chi^T(\hat{q})\chi(\hat{q})]^{-1}.$$

4. The standard errors for each element in the parameter estimator  $\mathcal{Q}$  can be approximated by

$$SE(\mathcal{Q}_i) = \sqrt{\hat{\Sigma}_{ii}}, \quad i = 1, 2, \dots, \kappa,$$

where  $\mathcal{Q}_i$  is the  $i$ th element of  $\mathcal{Q}$ , and  $\hat{\Sigma}_{ii}$  is the  $(i, i)$ th entry of the matrix  $\hat{\Sigma}$ . Hence, the endpoints of the confidence intervals for  $\mathcal{Q}_i$  are given by

$$\hat{q}_i \pm t_{1-\alpha/2}^{n-\kappa} SE(\mathcal{Q}_i)$$

for  $i = 1, 2, \dots, \kappa$ . Here  $t_{1-\alpha/2}^{n-\kappa}$  is a distribution value that is determined from a statistical table for Student's t-distribution based on the level of significance  $\alpha$  (i.e.,  $\alpha = .05$  for a 95% confidence interval) as well as the number of degrees of freedom (the number of data points  $n$  less parameters  $\kappa$ ).

We will present results below in Tables 3.2-3.4 and 3.5-3.7 on the low, medium, and high noise data sets using zero, one, and two relaxation times, and using the routines `fmincon` and `lsqnonlin` (trust-region-reflective only, as we cannot enforce the bound constraints with Levenburg-Marquardt algorithm).

Table 3.2: `fmincon`: Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at low noise level.

Model	Params	True Value	Estimate	SE	95% Confidence Interval
0	$\log_{10}(E_0)$	5.3424	5.3422	0.01373	(5.3223, 5.3622)
	$\log_{10}(E_1)$	1.6021	1.6651	0.1052	(1.4579, 1.8723)
1	$\log_{10}(E_0)$	5.3424	5.3422	0.01012	(5.3223, 5.3622)
	$\log_{10}(E_1)$	1.6021	1.584	0.1436	(1.3757, 1.9412)
	$\log_{10}(\tau_1)$	-1.3010	-0.3002	6.9746	(-1.4037, 1.3437)
2	$\log_{10}(E_0)$	5.3424	5.3422	0.1014	(5.3223, 5.3622)
	$\log_{10}(E_1)$	1.6021	1.6581	0.1463	(1.3700, 1.9463)
	$\log_{10}(\tau_1)$	-1.3010	-0.3000	8.2619	(-16.5727, 15.9828)
	$\log_{10}(\tau_2)$	1	1.3012	51.0143	(-99.1775, 101.7798)

Table 3.3: `fmincon`: Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at medium noise level.

Model	Params	True Value	Estimate	SE	95% Confidence Interval
0	$\log_{10}(E_0)$	5.3424	5.3264	0.01409	(5.3087, 5.3542)
	$\log_{10}(E_1)$	1.6021	1.7641	0.01165	(1.5346, 1.9936)
1	$\log_{10}(E_0)$	5.3424	5.3430	0.01032	(5.3226, 5.3633)
	$\log_{10}(E_1)$	1.6021	1.6586	0.1462	(1.3706, 1.9467)
	$\log_{10}(\tau_1)$	-1.3010	-0.3001	7.1076	(-1.4299, 1.3699)
2	$\log_{10}(E_0)$	5.3424	5.3430	0.01034	(5.3226, 5.3633)
	$\log_{10}(E_1)$	1.6021	1.6583	0.1490	(1.3648, 1.9519)
	$\log_{10}(\tau_1)$	-1.3010	-0.2998	8.4206	(-16.8851, 16.2855)
	$\log_{10}(\tau_2)$	1	1.3012	51.9907	(-101.1004, 103.7028)

We see throughout the tables that the problem of estimating the second relaxation time is fraught with difficulty (the standard error is significantly higher than its estimated value), even

Table 3.4: `fmincon`: Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at high noise level.

Model	Params	True Value	Estimate	SE	95% Confidence Interval
0	$\log_{10}(E_0)$	5.3424	5.3433	0.01046	(5.3227, 5.3639)
	$\log_{10}(E_1)$	1.6021	1.6452	0.1136	(1.4214, 1.8691)
1	$\log_{10}(E_0)$	5.3424	5.3433	0.01045	(5.3227, 5.3639)
	$\log_{10}(E_1)$	1.6021	1.6382	0.1600	(1.3232, 1.9534)
	$\log_{10}(\tau_1)$	-1.3010	-0.2995	7.7134	(-15.4916, 14.8927)
2	$\log_{10}(E_0)$	5.3424	5.3433	0.01047	(5.3227, 5.3639)
	$\log_{10}(E_1)$	1.6021	1.6380	0.1633	(1.3164, 1.9596)
	$\log_{10}(\tau_1)$	-1.3010	-0.2995	9.1380	(-18.2979, 17.6989)
	$\log_{10}(\tau_2)$	1	1.3012	56.4222	(-109.8289, 112.4314)

Table 3.5: `TRR lsqnonlin`: Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at low noise level.

Model	Params	True Value	Estimate	SE	95% Confidence Interval
0	$\log_{10}(E_0)$	5.3424	5.3422	$4.9498 \times 10^{-4}$	(5.3413, 5.3432)
	$\log_{10}(E_1)$	1.6021	1.6651	0.005434	(1.6544, 1.6758)
1	$\log_{10}(E_0)$	5.3424	5.3425	0.01011	(5.3226, 5.3624)
	$\log_{10}(E_1)$	1.6021	1.6050	0.3167	(0.9811, 2.2288)
	$\log_{10}(\tau_1)$	-1.3010	-1.2317	2.2200	(-5.6041, 3.1407)
2	$\log_{10}(E_0)$	5.3424	5.3425	0.0101	(5.3225, 5.3624)
	$\log_{10}(E_1)$	1.6021	1.6046	0.3202	(0.9738, 2.2353)
	$\log_{10}(\tau_1)$	-1.3010	-1.2297	2.2369	(-5.635, 3.1761)
	$\log_{10}(\tau_2)$	1	1.0046	16.0237	(-30.5560, 32.5651)

though we know the simulated data came from a model incorporating two relaxation times. This could be predicted from our earlier examination of the sensitivities with respect to the second relaxation time, as well as the results for relaxation times seen when using different optimization routines. In addition, when estimating two relaxation times on high noise data (shown in Table 3.7) we see that the estimates for  $\tau_1$  and  $\tau_2$  are not close to the true parameter values; also, the standard error for  $\tau_1$  is much larger than in any other case. Thus, instead of merely having difficulty estimating a second relaxation time, in this estimation we now additionally have less confidence in the estimate of  $\tau_1$ .

From Tables 3.2-3.4 we see that standard errors of  $E_0$  and  $E_1$  for model with no relaxation time is comparable with those for model with one relaxation time when using `fmincon`, but the standard error for  $\tau_1$  is fairly large (around 7, as compared with the estimated parameter value

Table 3.6: TRR `lsqnonlin`: Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at medium noise level.

Model	Params	True Value	Estimate	SE	95% Confidence Interval
0	$\log_{10}(E_0)$	5.3424	5.3429	$9.2836 \times 10^{-4}$	(5.3411, 5.3448)
	$\log_{10}(E_1)$	1.6021	1.6653	0.0102	(1.6452, 1.6854)
1	$\log_{10}(E_0)$	5.3424	5.3433	0.01042	(5.3228, 5.3638)
	$\log_{10}(E_1)$	1.6021	1.6050	0.3167	(0.9811, 2.2288)
	$\log_{10}(\tau_1)$	-1.3010	-1.2317	2.2200	(-5.6041, 3.1407)
2	$\log_{10}(E_0)$	5.3424	5.3433	0.01045	(5.3227, 5.3639)
	$\log_{10}(E_1)$	1.6021	1.5889	0.3717	(0.8567, 2.3211)
	$\log_{10}(\tau_1)$	-1.3010	-1.3269	2.0383	(-5.3415, 2.6878)
	$\log_{10}(\tau_2)$	1	2.0000	156.5630	(-306.369, 310.369)

Table 3.7: TRR `lsqnonlin`: Parameter estimates, asymptotic standard errors (SE) and confidence intervals for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-times-relaxation model (model 2) obtained at high noise level.

Model	Params	True Value	Estimate	SE	95% Confidence Interval
0	$\log_{10}(E_0)$	5.3424	5.3433	$1.7767 \times 10^{-3}$	(5.3398, 5.3468)
	$\log_{10}(E_1)$	1.6021	1.6452	0.0204	(1.6050, 1.6855)
1	$\log_{10}(E_0)$	5.3424	5.3433	0.01046	(5.3227, 5.3639)
	$\log_{10}(E_1)$	1.6021	1.6397	0.1526	(1.3391, 1.9403)
	$\log_{10}(\tau_1)$	-1.3010	-1.3253	2.0328	(-5.3291, 2.6785)
2	$\log_{10}(E_0)$	5.3424	5.3433	0.01046	(5.3227, 5.3639)
	$\log_{10}(E_1)$	1.6021	1.6361	0.1591	(1.3227, 1.9496)
	$\log_{10}(\tau_1)$	-1.3010	-0.1990	15.8821	(-31.4806, 31.0827)
	$\log_{10}(\tau_2)$	1	0.2496	12.0051	(-23.3958, 23.8951)

$\log_{10}(\hat{\tau}_1) \approx -0.3$ ). When using `lsqnonlin` (see Tables 3.5-3.7), the standard errors for  $E_0$  and  $E_1$  increase significantly at all noise levels when moving from the no relaxation time model to the one relaxation time model, but the standard error for  $\tau_1$  is closer to 2 rather than the 7 for `fmincon`. This may not seem significant, but if we recall that these are log-scaled parameter values, then the difference between standard errors of 2 and 7 is fairly large.

We also found that at all noise levels the difference for the residual sum of squares is small among the zero-relaxation-time, one-relaxation-time, and two-relaxation-time models (see the third columns of Table 3.9). In addition, for each level noise data set, when we plot the model solutions corresponding to the zero, one, and two relaxation time models, we found that they approximately lie on top of each other, and give good fits to the data. To gain further insight into which model should be chosen, we turned to some model selection criterion analysis.

### 3.4.1 Model Selection Criteria

There are numerous model selection criteria in the literature that can be used to select a best approximating model from a prior set of candidate models. These criteria are based on hypothesis testing, mean squared error, Bayes factors, or information theory, and they all are based to some extent on the principle of parsimony (see [35]). It should be noted that some of these criteria can only be used for nested models (e.g., two models are said to be nested if one model is a special case of the other), but others can be used for both nested models and non-nested models.

Here we employ one of the most widely used model selection criteria – the Akaike Information Criterion (AIC). The AIC was developed by Akaike (in 1973) who formulated a relationship between the Kullback-Leibler information (used to measure the information lost when a model is used to approximate the true model) and the maximum value of the log likelihood function of the approximating model. As might be expected we find that the AIC value depends on the data set used. Thus, when we try to select a best model from a set of candidate models, we must use the same data set to calculate AIC values for each of the models. One of the advantages of the AIC is that it can be used to compare non-nested models (which is our case here). For the least squares case, it can be found (e.g., see [35, Section 2.2]) that if the observation errors are i.i.d normally distributed, then the AIC is given by

$$AIC = n \log \left( \frac{RSS}{n} \right) + 2(\kappa + 1). \quad (3.4.1)$$

Here  $\kappa + 1$  is the total number of estimated parameters including  $q$  and the observation error variance. Given a prior set of candidate models, we can calculate the AIC value for each model, and the best approximating model is the one with minimum AIC value. It should be noted that the AIC may perform poorly if the sample size  $n$  is small relative to the total number of estimated parameters (it is suggested in [35] that the sample size  $n$  should be at least 40 times the total number of estimated parameters ( $\kappa + 1$ ); this is true for our investigations here).

In practice, the absolute size of the AIC value may have limited use in supporting the chosen best approximating model, and one may often employ other related values such as Akaike differences and Akaike weights to further compare models. The Akaike difference is defined by

$$\Delta_i = AIC_i - AIC_{\min}, \quad i = 1, 2, \dots, R, \quad (3.4.2)$$

where  $AIC_i$  is the AIC value of the  $i$ th model in the set,  $AIC_{\min}$  denotes the AIC value for the best fit model in the set, and  $R$  is the total number of models in the set. The larger  $\Delta_i$ , the less plausible it is that the  $i$ th model is a good approximating model for given the data set. The

Akaike weights are defined by

$$w_i = \frac{\exp(-\frac{1}{2}\Delta_i)}{\sum_{r=1}^R \exp(-\frac{1}{2}\Delta_r)}, \quad i = 1, 2, \dots, R. \quad (3.4.3)$$

These Akaike weights  $w_i$  can then be interpreted heuristically as the probability that  $i$ th model is the best approximating model (see [35]).

Tables 3.8 and 3.9 demonstrate residual sum squares (RSS), AIC values, AIC difference, and AIC weights obtained for the two-relaxation-time model, one-relaxation-time model and no-relaxation-time model at low, medium and high noise levels using `fmincon` and `lsqnonlin`, respectively.

Table 3.8: `fmincon`: Residual sum of squares (RSS), AIC values, AIC difference ( $\Delta$ ) and AIC weights for zero-relaxation-time model (model 0), one-relaxation-time model (model 1) and two-times-relaxation model (model 2) obtained at low, medium and high noise levels.

noise level	model	RSS	AIC	$\Delta$	AIC weights
low noise	0	$7.0368 \times 10^{-9}$	$-6.0927 \times 10^3$	7.5867	$1.5257 \times 10^{-2}$
	1	$6.7731 \times 10^{-9}$	$-6.1003 \times 10^3$	0	$6.7748 \times 10^{-1}$
	2	$6.7618 \times 10^{-9}$	$-6.0987 \times 10^3$	1.5814	$3.0726 \times 10^{-1}$
medium noise	0	$3.6421 \times 10^{-8}$	$-5.6800 \times 10^3$	98.1093	$3.5908 \times 10^{-22}$
	1	$2.4442 \times 10^{-8}$	$-5.7782 \times 10^3$	0	$7.2334 \times 10^{-1}$
	2	$2.4435 \times 10^{-8}$	$-5.7762 \times 10^3$	1.9222	$2.7666 \times 10^{-1}$
high noise	0	$1.0337 \times 10^{-7}$	$-5.4182 \times 10^3$	0	$6.4532 \times 10^{-1}$
	1	$1.0330 \times 10^{-7}$	$-5.4164 \times 10^3$	1.8267	$2.5889 \times 10^{-1}$
	2	$1.0329 \times 10^{-7}$	$-5.4144 \times 10^3$	3.8151	$9.5794 \times 10^{-2}$

From these two tables, we see that on low and medium level noise data sets the one-relaxation-time model is the best model with the probability to be chosen as the best model being more than 0.7 (see Akaike weights in the last column of these two tables), and the no-relaxation time model has almost no chance to be selected as the best model. On the high noise data set the no-relaxation-time model is the best one with the probability chosen as the best model being more than 0.6, and two-relaxation-time model has almost no chance to be selected as the best model.

### Asymptotic Analysis Summary Remark

Based on our analysis to this point, we can conclude that estimating two relaxation times is likely to be difficult. Adopting a model with zero or one relaxation times may be the most

Table 3.9: `lsqnonlin`: Residual sum of squares (RSS), AIC values, AIC differences ( $\Delta$ ) and AIC weights for the zero-relaxation-time model (model 0), the one-relaxation-time model (model 1) and the two-relaxation-times model (model 2) obtained at low, medium and high noise levels.

Noise level	Model	RSS	AIC	$\Delta$	AIC weights
low noise	0	$7.0368 \times 10^{-9}$	$-6.0927 \times 10^3$	27.8791	$6.4125 \times 10^{-7}$
	1	$6.2470 \times 10^{-9}$	$-6.1206 \times 10^3$	0	$7.2595 \times 10^{-1}$
	2	$6.2458 \times 10^{-9}$	$-6.1186 \times 10^3$	1.9483	$2.7405 \times 10^{-1}$
medium noise	0	$2.4674 \times 10^{-8}$	$-5.7778 \times 10^3$	8.6863	$9.4255 \times 10^{-3}$
	1	$2.3646 \times 10^{-8}$	$-5.7865 \times 10^3$	0	$7.2527 \times 10^{-1}$
	2	$2.3647 \times 10^{-8}$	$-5.7845 \times 10^3$	2.0113	$2.6531 \times 10^{-1}$
high noise	0	$1.0337 \times 10^{-7}$	$-5.4182 \times 10^3$	0	$6.4303 \times 10^{-1}$
	1	$1.0330 \times 10^{-7}$	$-5.4164 \times 10^3$	1.8299	$2.5756 \times 10^{-1}$
	2	$1.0326 \times 10^{-7}$	$-5.4145 \times 10^3$	3.7340	$9.9406 \times 10^{-2}$

feasible approach. However, until we confirm this approach by examining these methods on experimental data we believe that attempting all three options for including relaxation times in the viscoelastic model (zero, one, or two times) may be advisable.

### 3.5 Bootstrapping Error Analysis

For ease of presentation, we reiterate here the algorithm described in [12], in the context of the current viscoelastic model under study (additional theory is available in, e.g., [45]).

1. Determine  $\hat{q}^0$  by solving the inverse problem (3.3.3).
2. Define the *standardized residuals* (recall  $n$  is the number of data points, and  $\kappa$  is the number of parameters under consideration) to be

$$\bar{r}_j = \sqrt{\frac{n}{n - \kappa}} \left( u_j - u(L, t_j; 10^{\hat{q}^0}) \right)$$

for  $j = 0, 1, \dots, n - 1$ . Set  $m = 0$ .

3. Create a sample of size  $n$  by randomly sampling, with replacement, from the standardized residuals  $\bar{r}_j$  to form a bootstrap sample  $\{r_0^m, \dots, r_{n-1}^m\}$ .
4. Create bootstrap sample points

$$u_j^m = u(L, t_j; 10^{\hat{q}^0}) + r_j^m, \quad j = 0, 1, \dots, n - 1.$$



5. Solve the OLS minimization problem (3.3.3) with the bootstrap-generated data  $\{u_j^m\}$  to obtain a new estimate  $\hat{q}^{m+1}$  which we store.
6. Increase the index  $m$  by 1 and repeat steps 3-5. This iterative process should be carried out for  $M$  times where  $M$  is large (we used  $M = 1000$ , as suggested in [12, 46]). This will give  $M$  estimates  $\{\hat{q}^m\}_{m=1}^M$ .

Upon completing all  $M$  simulation runs, the following will give the mean and covariance matrix for the bootstrap estimator  $\mathcal{Q}_{boot}$  of  $q_0$ :

$$\begin{aligned}\hat{q}_{boot} &= \frac{1}{M} \sum_{m=1}^M \hat{q}^m, \\ \hat{\Sigma}_{boot} &= \frac{1}{M-1} \sum_{m=1}^M (\hat{q}^m - \hat{q}_{boot})(\hat{q}^m - \hat{q}_{boot})^T.\end{aligned}\tag{3.5.1}$$

Then the standard errors for the bootstrap estimator  $\mathcal{Q}_{boot}$  are given by

$$(SE_{boot})_i = \sqrt{\left(\hat{\Sigma}_{boot}\right)_{ii}}, \quad i = 1, 2, \dots, \kappa,$$

where  $\left(\hat{\Sigma}_{boot}\right)_{ii}$  is the  $(i, i)$ th entry of covariance matrix  $\hat{\Sigma}_{boot}$ . Hence, the endpoints of the confidence intervals for  $(\mathcal{Q}_{boot})_i$  (the  $i$ th element of  $\mathcal{Q}_{boot}$ ) are given by

$$(\hat{q}_{boot})_i \pm t_{1-\alpha/2}^{n-\kappa} (SE_{boot})_i \tag{3.5.2}$$

for  $i = 1, 2, \dots, \kappa$ .

Note that bootstrapping requires solving the inverse problem 1000 times. Even for a model that is solved in a short time (e.g., less than one minute), bootstrapping takes a significant time to compute (as we must solve the inverse problem many times and each inverse problem involves solving the model multiple times). Due to long computational times (e.g., one week for bootstrapping versus minutes for the asymptotic theory), we report here results for a case using `fmincon` to estimate  $E_0$  and  $E_1$  in a zero-relaxation-time model and a case using `lsqnonlin`, the trust-region-reflective option, to estimate  $E_0$ ,  $E_1$ , and  $\tau_1$  in a one-relaxation-time model. It is worth noting that even though the bootstrapping algorithm can be implemented in parallel, this requires a considerable amount of computing resources (unavailable to most investigators) to achieve computational times comparable to that attained in using the asymptotic theory. For our purposes, the bootstrap results we provide are sufficient to indicate that the less conservative asymptotic error analysis yields a reasonable uncertainty measure in the inverse problem we investigate.

For the model with no relaxation times, we see from Table 3.10 that the confidence intervals for  $E_0$  and  $E_1$  at all noise levels are more conservative than those obtained using the asymptotic theory (shown in Tables 3.2, 3.3 and 3.4), especially for the cases of medium and high noise level. However, this table still indicates that reasonable parameter estimates are obtained.

Table 3.10: **fmincon**: Parameter estimates, bootstrap standard errors (SE) and confidence intervals obtained at low, medium and high noise levels for zero-relaxation-time model.

Noise Level	Params	True Value	$\hat{q}_{boot}$	SE	95% Confidence Interval
Low noise	$\log_{10}(E_0)$	5.3424	5.3422	0.01557	(5.3115, 5.3729)
	$\log_{10}(E_1)$	1.6021	1.6649	0.1716	(1.3270, 2.0029)
Medium noise	$\log_{10}(E_0)$	5.3424	5.3429	0.03038	(5.2831, 5.4028)
	$\log_{10}(E_1)$	1.6021	1.6653	0.3155	(1.0438, 2.2867)
High noise	$\log_{10}(E_0)$	5.3424	5.3432	0.05714	(5.2306, 5.4557)
	$\log_{10}(E_1)$	1.6021	1.6451	0.6603	(0.3446, 2.9456)

In Figure 3.8, we show the bootstrap estimates obtained for  $E_0$  and  $E_1$  for this no relaxation time model. Note that each estimator tends to have the shape of a normal distribution, which we would expect if our formulation of bootstrapping is to work properly.

The bootstrapping results for a one-relaxation-time model obtained by using **lsqnonlin** trust-region-reflective routine are summarized in Table 3.11. We see that confidence intervals

Table 3.11: **TRR lsqnonlin**: Parameter estimates, bootstrap standard errors (SE) and confidence intervals obtained at low, medium and high noise levels for one-relaxation-time model.

Noise level	Params	True Value	$\hat{q}_{boot}$	SE	95% Confidence Interval
Low noise	$\log_{10}(E_0)$	5.3424	5.3425	0.01547	(5.3120, 5.3730)
	$\log_{10}(E_1)$	1.6021	1.6025	0.5937	(0.4332, 2.7719)
	$\log_{10}(\tau_1)$	-1.3010	-1.2294	3.8697	(-8.8510, 6.3923)
Medium noise	$\log_{10}(E_0)$	5.3424	5.3434	0.03136	(5.2816, 5.4052)
	$\log_{10}(E_1)$	1.6021	1.5852	1.4590	(-1.2884, 4.4589)
	$\log_{10}(\tau_1)$	-1.3010	-1.2079	16.9971	(-34.6849, 32.2692)
High noise	$\log_{10}(E_0)$	5.3424	5.3434	0.061762	(5.2218, 5.4651)
	$\log_{10}(E_1)$	1.6021	1.6029	2.4545	(-3.2314, 6.4372)
	$\log_{10}(\tau_1)$	-1.3010	-0.1592	40.8381	(-80.5930, 80.2746)

for all parameters are wider than those obtained using the asymptotic error theory, especially for the cases of medium and high noise level. However, this is expected. At the low noise level,

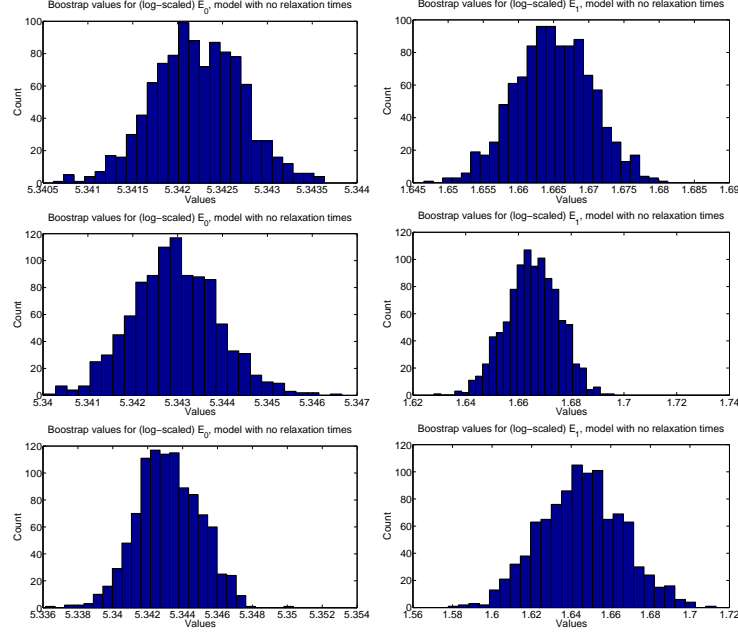


Figure 3.8: Histograms of bootstrap estimates  $q^m$  for a model with no relaxation times in the case of low noise data set (upper row), medium noise data set (middle row) and high noise data set (bottom row). (left column) Estimates for  $\log_{10}(E_0)$ . (right column) Estimates for  $\log_{10}(E_1)$ .

we obtained fairly good results for  $E_0$  and  $E_1$  but the standard error for the relaxation time is larger in magnitude than the relaxation time value itself. This is even more prominent at higher noise levels – the results in the table indicate that on medium and high noise data sets, the estimation of  $\tau_1$  is not very robust. Note also that the estimation of  $E_1$  begins to suffer as well, resulting in a higher standard error than its own value on the high noise data set. This is a further indication that we may have problems in the future estimating even the single relaxation time.

We depict histograms of the estimates in Figure 3.9. We see on a low noise data set that each parameter estimator appears to be mostly normally distributed. This begins to break down for the case of middle noise level data set (shown in the middle row of Figure 3.9), where we begin to see some outliers at the  $\log_{10}(\hat{\tau}_1) = 2$  level (which means the estimates were converging to our upper bound on that parameter) and also some more pronounced skewness in the count levels. Finally, on the high noise level (shown in bottom row of Figure 3.9) we have the distribution for the  $E_1$  estimates skewed, and we also observe a clear proliferation of estimates of the first relaxation time approaching the bounding value 2. This further supports the expectation of difficulty in estimating relaxation times, particularly when the noise level is high.

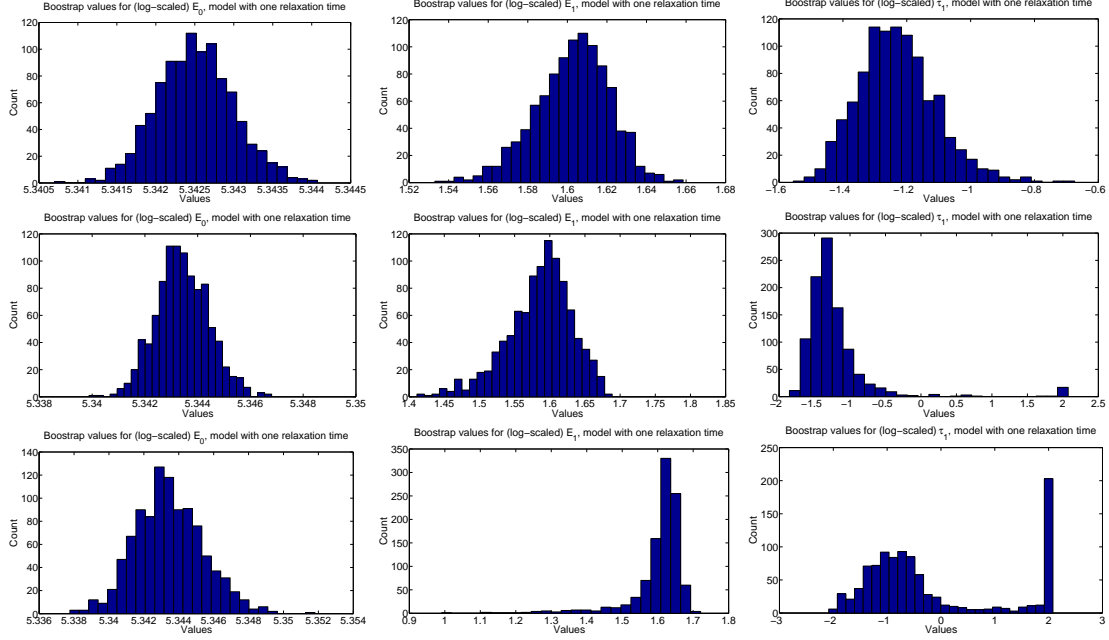


Figure 3.9: Histograms of bootstrap estimates  $\hat{q}^m$  for a one-relaxation-time model obtained at low noise (upper row), medium noise (middle row) and high noise (bottom row) levels. (left column) Estimates for  $\log_{10}(E_0)$ . (middle column) Estimates for  $\log_{10}(E_1)$ . (right column) Estimates for  $\log_{10}(\tau_1)$ .

### 3.6 Model Comparison and Hypothesis Testing on Amplitude

In this section, we develop a methodology for determining whether or not data came from a low-amplitude input traction. This simulates the problem of determining if the data came from a vessel experiencing a normal heartbeat or not. We will ultimately run the inverse problem without amplitude restrictions and use a scoring function to compare results with the score of the model solved at a low amplitude. A model comparison test will be implemented to determine if there is statistical significance in the differences between the model solved with the unrestricted estimate and the model solved using the restricted amplitude value.

#### 3.6.1 Setup

We first examine the sensitivity of the model with respect to the Van Bladel input amplitude parameter  $A$ , to insure that an estimation procedure is reasonable (if the model were insensitive to  $A$  then the results from the optimization routine would be suspect). The form of the sensitivity equation is nearly identical to that of the actual model, just with a lower amplitude (see Appendix A). This is seen in Figure 3.10, which has a form similar to that of the model solution (depicted

by the solid line in Figure 3.6).

In both the low and high amplitude cases, the sensitivity with respect to amplitude is most marked during early times and less so at later times; this makes perfect sense, as the amplitude is greater early on before being damped out. In the problem below, we will take data throughout the full time frame  $t \in [0, 0.25]$  so with our sensitivity results we can be assured that the early data will drive estimation of the amplitude parameter.

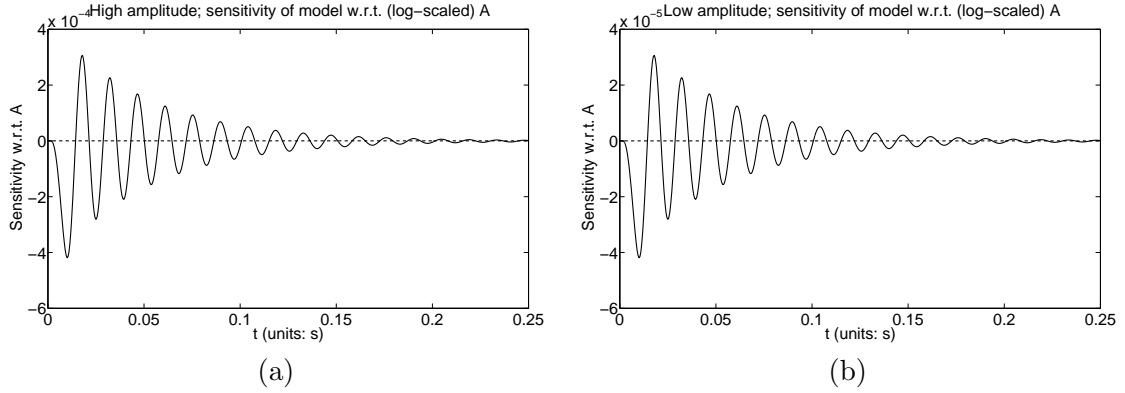


Figure 3.10: Sensitivity of model with respect to Van Bladel input parameter  $A$  around the baseline parameters (3.1.10). (a) High forcing function amplitude  $A = 6 \times 10^3$ . (b) Low forcing function amplitude  $A = 6 \times 10^2$ . Note the order of magnitude difference between (a) and (b).

### Low Amplitude Simulated Data Generation

For the high amplitude data, we use the same low, medium, and high noise data sets as described in Section 3.3.1 and shown above in Figure 3.6. We form the low amplitude data by taking  $A_{low} = A/10$  as our Van Bladel input amplitude parameter. Thus, the dynamics are roughly 10% the magnitude of the high amplitude data. This means the corresponding noise for the low noise, low amplitude data set will be generated with variance  $\sigma^2 = 5 \times 10^{-7}$ , medium noise with  $\sigma^2 = 10 \times 10^{-7}$ , and high noise with  $\sigma^2 = 20 \times 10^{-7}$ . The low amplitude input data set then is supposed to represent a normal heartbeat and the high amplitude data set then is meant to represent the input shear for a heartbeat in the presence of a stenosis in the vessel. Note that we are not yet exactly certain regarding the difference between these effects in an actual patient, so the data sets here are truly for a proof-of-concept investigation. The low amplitude data sets are depicted in Figure 3.11.

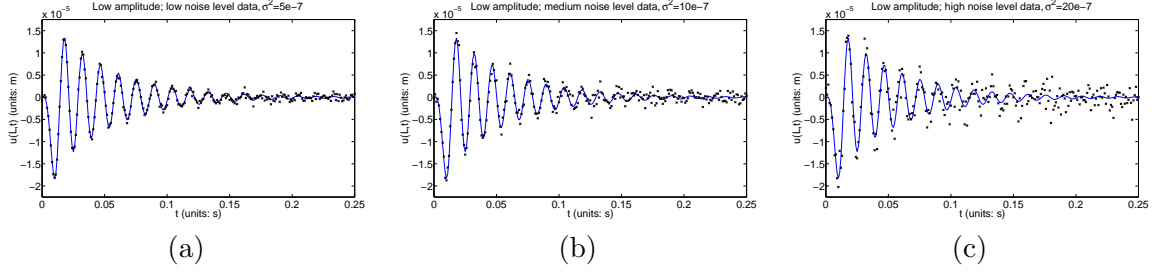


Figure 3.11: Simulated low amplitude noisy data around the true parameter values. (a) Low noise level,  $\sigma_0^2 = 5 \times 10^{-7}$ . (b) Medium noise,  $\sigma_0^2 = 10 \times 10^{-7}$ . (c) High noise level,  $\sigma_0^2 = 20 \times 10^{-7}$ .

### 3.6.2 Hypothesis Testing Methodology

We can now begin to discuss the approach to model comparison and hypothesis testing that we will use by defining a model comparison test statistic. The work here follows the development in [11, 24]. The framework as developed in [11] and used here requires an absolute error model; though we will not use it here, this framework was expanded in [16] to allow for relative error models. The performance criterion for hypothesis testing will be

$$J(\vec{U}, q) = \sum_{j=0}^{n-1} [U_j - u(L, t_j; q)]^2.$$

For the purposes of this paper, we postulate that a normal (non-stenosed) vessel corresponds with a low amplitude input parameter  $A \leq 6 \times 10^2$ . Then, a stenosed vessel would have a high input amplitude parameter with  $A > 6 \times 10^2$ . The hypothesis test we use requires a set benchmark value for  $A$ , so we choose that benchmark to be  $A_0 = 6 \times 10^2$ . Then, we define the restricted parameter set

$$\mathbb{A}_H = \{A \in \mathbb{A} | A = A_0 = 6 \times 10^2\},$$

where  $\mathbb{A} = [A_0, \infty)$  is the larger set of unrestricted admissible amplitudes.

Our null hypothesis  $H_0$  is that the amplitude is a low amplitude, represented by  $A \in \mathbb{A}_H = \{A_0\}$ . The unrestricted amplitude model would then represent the amplitude parameter as  $A = A_0 + \tilde{A}$  where  $\tilde{A} \in [0, \infty)$ . This framework will allow us to develop a test statistic to determine the confidence level of accepting or rejecting  $H_0$  for a given data set. In other words, we will develop a test to determine if the data is statistically better represented by the benchmark  $A_0$  than the unrestricted amplitude.

The first step is to determine the performance criterion at the benchmark amplitude  $\hat{q}_H = 6 \times 10^2$ , which we will denote  $J(\vec{u}, \hat{q}_H)$  (Since the value  $\hat{q}_H$  is fixed in our case, no optimization problem is needed to compute these values). We then run an optimization routine

to determine an unrestricted input amplitude parameter estimate  $\hat{q}$ , which we then use to compute  $J(\vec{u}, \hat{q})$ . The value for  $\hat{q}$  comes from solving the unrestricted optimization problem (3.3.3). As discussed in [11, 16, 24], the model comparison statistic is defined as

$$\hat{V} = n \frac{J(\vec{U}, \mathcal{Q}_H) - J(\vec{U}, \mathcal{Q})}{J(\vec{U}, \mathcal{Q})}$$

with realization

$$\hat{v} = n \frac{J(\vec{u}, \hat{q}_H) - J(\vec{u}, \hat{q})}{J(\vec{u}, \hat{q})}. \quad (3.6.1)$$

If our null hypothesis  $H_0$  were true, the model comparison statistic  $\hat{V}$  converges in distribution to  $V$  as  $n \rightarrow \infty$  where  $V \sim \chi^2(r)$  is a chi-square distribution with  $r$  degrees of freedom ( $r$  is the number of constraints in  $\mathbb{A}_H$ ). For our problem,  $r = 1$ . Given the significance level  $\alpha$ , we can obtain a threshold value  $\nu$  such that the probability that  $V$  will take on a value greater than  $\nu$  is  $\alpha$ . In other words,  $\text{Prob}(V > \nu) = \alpha$ . In our context, if the test statistic  $\hat{v} > \nu$  we reject  $H_0$  as false with confidence level  $(1 - \alpha)100\%$ . Otherwise we do not reject  $H_0$  as false, at the specified confidence level. In Table 3.12 we include sample values from the  $\chi^2(1)$  distribution for reference (table repeated from [24]).

Table 3.12: Sample  $\chi^2(1)$  values.

$\alpha$	$\nu$	confidence
0.25	1.32	75%
0.1	2.71	90%
0.05	3.84	95%
0.01	6.63	99%
0.001	10.83	99.9%

We summarize in Table 3.13 the results of computing the OLS performance criterion for the low amplitude and high amplitude data each with the restricted/unrestricted parameters. Based on this table and Table 3.12, we see for both the low and medium noise cases with data generated with a low  $A$  value that we do not reject  $H_0$  with high degrees of confidence. However, the case with high noise is somewhat less certain, though we would still likely not reject  $H_0$  with a fairly high degree of confidence. The results are more stark in the cases where the data was generated from a high amplitude. Given that the magnitude of  $\hat{v}$  is greater than 900 at all noise levels, we would reject  $H_0$  as false on these data sets with confidence level more than 99.9%. Altogether, these results suggest robustness in our methodology for determining whether the data came from a normal vessel experiencing a heartbeat (low input amplitude) or from a

Table 3.13: Model comparison test results using (3.6.1) on low, medium, and high noise data sets generated with both high and low input amplitude parameter  $A$  values.

	$J(\vec{u}, \hat{q})$	$J(\vec{u}, \hat{q}_H)$	$\hat{v}$
Low $A$ , low noise	6.3846e-11	6.3887e-11	0.1609
Low $A$ , medium noise	2.6872e-10	2.6896e-10	0.2258
Low $A$ , high noise	9.8836e-10	9.9658e-10	2.0878
High $A$ , low noise	$6.6812 \times 10^{-9}$	$3.5229 \times 10^{-7}$	1.2984e+04
High $A$ , medium noise	$3.1016 \times 10^{-8}$	$3.4730 \times 10^{-7}$	2.5596e+03
High $A$ , high noise	$9.9737 \times 10^{-8}$	$4.6015 \times 10^{-7}$	907.0283

abnormal (stenosed) response. Though this is a first pass investigation, tests like we examined here should be useful in practice, particularly once models of sounds generated from a stenosis have been incorporated as inputs into the viscoelastic wave propagation model.

### 3.7 Proof of Concept Summary

In this chapter we have carried out proof-of-concept investigations for estimating material parameters and created a model comparison test as a basis for distinguishing between data that comes from a normal or from a stenosed blood vessel. We found that the model was less sensitive to a second viscoelastic relaxation time than to the other parameters, and this was manifested as a difficulty in recovering two relaxation times. On the other hand, models with zero or one relaxation time allowed for more confidence in the estimation procedure (i.e., smaller standard errors). We compared asymptotic error analysis with bootstrapping, and found (as expected) that bootstrapping gives more conservative confidence intervals but not so much so that the asymptotic theory cannot be profitably used for uncertainty quantification in models with large computational costs rendering bootstrapping less desirable. In terms of the model comparison on the input amplitude parameter  $A$ , we were able to develop a successful methodology for statistically determining whether or not data came from a low amplitude input force. This will form the basis of a model comparison test that can later be used with experimental data sets.

In the upcoming chapters, we will examine the possibility of relative error instead of absolute error, which will necessitate a generalized least squares (GLS) cost function in our inverse problems due to changes in the error process. This will be coupled with a study of a statistical model for the measurement processes being used in the experiments at QMUL. It is to this end that we return to the general equations of motion in the next chapter, carefully deriving pressure and shear wave models and a constitutive relationship similar to (3.1.2) that we will call the *primary model* for this wave propagation problem.



---

## Primary Model Derivation and Constitutive Relationship

---

In this chapter, we return to the basics of solid mechanics. We state the equations of motion, and also postulate a general form for the constitutive relationship so that we might later be able to properly make simplifying assumptions to obtain one-dimensional pressure and shear models. The general constitutive relationship will be based on a strain-energy formulation [13, 29, 52] which is fairly general and also provides a framework for future development of two-dimensional or even three-dimensional wave propagation models. This framework was used in the previous incarnation of this project by [67, 80]. The latter parts of this chapter follow the development of the model as discussed in [15] (which has been accepted for publication).

### 4.1 Basic Introduction to Viscoelasticity

We first introduce some basics regarding how one models displacements in material. First, we use three directions of motion  $(r, \theta, z)$  in cylindrical coordinates. There are corresponding displacements, denoted with subscripts (e.g.,  $u_z$  is the displacement in the  $z$ -direction). In order to denote changes in quantities, derivatives will be indicated in this section with appropriate partial derivative notation (e.g.,  $\frac{\partial u_z}{\partial r}$  is the change in the displacement in the  $z$ -direction with respect to  $r$ ). In later sections, once we have the one-dimensional models in hand, we will no longer require subscripts to denote directions (since each one-dimensional model will describe a single directional displacement component) and will thus use the partial derivative and subscript notation interchangeably (i.e., later we will allow notation like  $u_x = \frac{\partial u}{\partial x}$  to both represent the change in displacement with respect to the spatial variable). Also, an equivalent form for

subscripts often used in the literature is to denote the (in our model, cylindrical) components by  $1 = r$ ,  $2 = \theta$ , and  $3 = z$ . For example, the three displacement components can be written  $u_1 = u_r$ ,  $u_2 = u_\theta$ , and  $u_3 = u_z$ . From this point onward, we will use subscripts interchangeably as just defined.

In addition to displacements, we incorporate the concepts of stresses (directional forces acting on and within the material) and strains (descriptions of relative displacements). In the finite strain theory case, we will denote the tensor of strains by  $\bar{E} = \{E_{ij}\}$  and the tensor of stresses by  $\bar{S} = \{S_{ij}\}$ , for  $i, j = 1, 2, 3$ . The strain tensor  $\bar{E}$  is called Green's strain tensor, and the stress tensor  $\bar{S}$  is the Second Piola-Kirchhoff stress tensor [13, 52, 53]. In the infinitesimal strain case, the tensor of strains  $\bar{\epsilon} = \{\epsilon_{ij}\}$  is called the Cauchy strain tensor and the Cauchy stress tensor is denoted  $\bar{\sigma} = \{\sigma_{ij}\}$ , again for  $i, j = 1, 2, 3$ . In the infinitesimal case, we have  $\bar{\epsilon} = \bar{E}$  and  $\bar{\sigma} = \bar{S}$ . In our case, we have only small displacements, so we will use the infinitesimal strain theory. We also note that modeling nonlinear behavior between stress and the finite strains  $\bar{E}$  can be equivalently formulated as nonlinear relationships between the stress and infinitesimal strains  $\bar{\epsilon}$ , as discussed in [20]. Thus even though we consider the small-strain framework here, this applies equally well to finite strains.

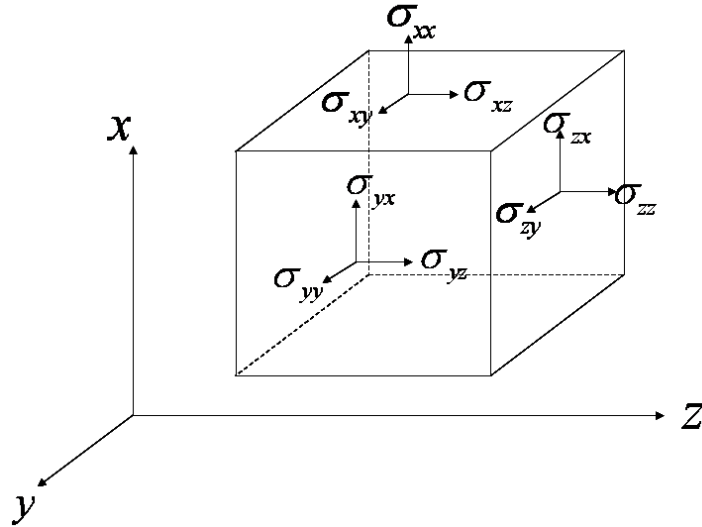


Figure 4.1: Stress components in Cartesian coordinates.

We briefly describe the meaning of the notation used when discussing stresses and strains. Though our later discussion will use a cylindrical geometry, intuition for the notation is easiest in Cartesian  $(x, y, z)$  coordinates. Figure 4.1 shows a sample (infinitesimal) cube with the stress

components indicated. The first subscript denotes the face of the cube, which is perpendicular to the direction indicated. The second subscript denotes the direction of the stress. The normal stresses are  $\sigma_{xx}$ ,  $\sigma_{yy}$ , and  $\sigma_{zz}$ ; the remaining components are all shear stresses. Components are positive if they act in the positive direction of the coordinate axis (i.e., if  $\sigma_{xx} > 0$  that normal stress is acting in the direction of the  $x$ -axis). Further discussion is available in most texts on elasticity and viscoelasticity [29, 41, 48, 49, 50, 52, 53, 63, 66, 69, 72] as well as an elasticity survey paper [13].

## 4.2 Equations of Motion and General Constitutive Relationship

We now turn to the problem at hand. Based on the experimental setup, we have a cylindrical geometry. This means there are three coordinate directions,  $r$ ,  $\theta$ , and  $z$ . Displacements in each direction are denoted  $u_r$ ,  $u_\theta$ , and  $u_z$ , respectively. In the full three-dimensional case, for an isotropic solid we have six strains  $\epsilon_{ij}$  for  $i, j = r, \theta, z$ , where  $\epsilon_{ij} = \epsilon_{ji}$  for  $j \neq i$ . We also have six stresses,  $\sigma_{ij}$  for  $i, j = r, \theta, z$ , with symmetry when  $i \neq j$ . The equations of motion in cylindrical coordinates are

$$\rho \ddot{u}_r = \frac{\partial \sigma_{rr}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{r\theta}}{\partial \theta} + \frac{\partial \sigma_{rz}}{\partial z} + \frac{1}{r} (\sigma_{rr} - \sigma_{\theta\theta}), \quad (4.2.1)$$

$$\rho \ddot{u}_\theta = \frac{\partial \sigma_{r\theta}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{\theta\theta}}{\partial \theta} + \frac{\partial \sigma_{\theta z}}{\partial z} + \frac{2}{r} \sigma_{r\theta}, \quad (4.2.2)$$

$$\rho \ddot{u}_z = \frac{\partial \sigma_{rz}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{\theta z}}{\partial \theta} + \frac{\partial \sigma_{zz}}{\partial z} + \frac{1}{r} \sigma_{rz}. \quad (4.2.3)$$

These equations are commonly known; one may see [29, 66, 80] for the derivation of the equations of motion in cylindrical coordinates. In the infinitesimal strain theory case (which describes our situation), we have the following forms for the strains in terms of displacements [29, p. 64]:

$$\begin{aligned} \epsilon_{rr} &= \frac{\partial u_r}{\partial r}, & \epsilon_{\theta\theta} &= \frac{1}{r} \left( u_r + \frac{\partial u_\theta}{\partial \theta} \right), & \epsilon_{zz} &= \frac{\partial u_z}{\partial z}, \\ \epsilon_{r\theta} &= \frac{1}{2} \left( \frac{1}{r} \frac{\partial u_r}{\partial \theta} + \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} \right), & \epsilon_{rz} &= \frac{1}{2} \left( \frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r} \right), & \epsilon_{\theta z} &= \frac{1}{2} \left( \frac{\partial u_\theta}{\partial z} + \frac{\partial u_z}{\partial \theta} \right). \end{aligned} \quad (4.2.4)$$

We now need some general description of the *constitutive relationship* describing the components of stress  $\bar{\sigma}$  in terms of strain  $\bar{\epsilon}$ , so that we might be able to properly reduce (4.2.1) in the pressure and shear cases.

The constitutive relationship discussion must begin with some general notions used to describe the relationships between each stress component  $\sigma_{ij}$  with the strains  $\epsilon_{ij}$ . We begin by noting, as discussed in [13, 52], that one may define the stress components as derivatives of a strain energy function  $W(\bar{\epsilon})$ . The strain energy function is a measure of the internal energy stored in a strained or stressed material, and its existence can be justified in some cases by the

laws of thermodynamics (see, e.g., [29, Sec. 3.1]). The following relationship between  $\bar{\sigma}$ ,  $W$ , and  $\bar{\epsilon}$  then holds:

$$\sigma_{ij} = \frac{\partial W(\bar{\epsilon})}{\partial \epsilon_{ij}}.$$

This relationship was given in [13, 52], and an explicit justification is described in [29, Sec. 3.1]. Fung in [52] (which [67, 80] follow) expands this idea to the viscoelastic case. Defining a *pseudo-strain energy function*, denoted  $\rho_0 W^{(2)}(\bar{\epsilon})$  where  $\rho_0$  is the density of the material in its initial configuration, and defining the stress reached instantaneously when strains are suddenly increased as  $\sigma_{ij}^e$ , the following relationship holds:

$$\sigma_{ij}^e = \frac{\partial \rho_0 W^{(2)}(\bar{\epsilon})}{\partial \epsilon_{ij}}. \quad (4.2.5)$$

These components are then incorporated into a viscoelastic relationship for  $\sigma_{ij}$ , defined to be

$$\sigma_{ij} = \int_{-\infty}^t \mathcal{G}_{ijkl}(t-s) \frac{\partial \sigma_{kl}^e(\bar{\epsilon}(s))}{\partial s} ds. \quad (4.2.6)$$

This form was generalized from a one-dimensional quasi-linear viscoelastic relationship for soft tissues developed by Fung [52, Sec. 7.6]. Also, this form was originally developed by Fung in the finite strain case, so it can apply equally well for general stress-strain laws even though our case is small-strain (as previously mentioned, finite strains can be incorporated as nonlinear infinitesimal stress-strain relationships). The infinite lower bound of the integral in (4.2.6) signifies the fact that in a general viscoelastic material the entire material history effects the present material responses and as such must be considered. Usually, if the material is initially at rest (as in our case) we can assume some beginning time for the history without loss of generality. The development in this section will be kept general, but the assumption of finite history will be used later in this chapter.

We see from (4.2.6) that complex relationships exist between the stress and strain components which would make further reduction impossible, in general. However, our material is isotropic (invariant under rotation and reflection), which means (see, e.g., [13, Rem. 3.2]) we can write the form for  $\mathcal{G}$  as

$$\mathcal{G}_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk})$$

where  $\delta_{ij}$  is the Dirac delta ( $\delta_{ij} = 1$  if  $i = j$ ,  $\delta_{ij} = 0$  otherwise). If we use this form for  $\mathcal{G}$  in (4.2.6), we obtain

$$\sigma_{ij} = \int_{-\infty}^t \left[ \lambda(t-s) \frac{\partial \sigma_{kk}^e}{\partial s} \delta_{ij} + 2\mu(t-s) \frac{\partial \sigma_{ij}^e}{\partial s} \right] ds. \quad (4.2.7)$$

Note that repeated subscripts are Einstein notation, which is a compact way to represent

summation. This means that in (4.2.7) we read the double subscript as  $\frac{\partial \sigma_{kk}^e}{\partial s} = \frac{\partial \sigma_{11}^e}{\partial s} + \frac{\partial \sigma_{22}^e}{\partial s} + \frac{\partial \sigma_{33}^e}{\partial s}$ . The assumption of isotropy means that we will only need to characterize two functions  $\lambda(t)$  and  $\mu(t)$  (or a combination of the two functions) later in order to complete the constitutive relationship description.

We now define a general form for  $\rho_0 W^{(2)}(\bar{\epsilon})$  as that will ultimately dictate how the form of the  $\sigma_{ij}$  components simplify as we reduce to the one-dimensional pressure and physical models. Following [67, 80] which used the Fung pseudo-energy relationship [52, Sec. 7.12], we define

$$\rho_0 W^{(2)}(\bar{\epsilon}) = \frac{1}{2} f(\vec{\alpha}, \bar{\epsilon}) + \frac{c}{2} \exp(F(\vec{a}, \bar{\epsilon})) \quad (4.2.8)$$

where

$$\begin{aligned} f(\vec{\alpha}, \bar{\epsilon}) = & \alpha_1 \epsilon_{11}^2 + \alpha_2 \epsilon_{22}^2 + \alpha_3 \epsilon_{33}^2 + \alpha_4 \epsilon_{12}^2 + \alpha_4 \epsilon_{21}^2 + \alpha_5 \epsilon_{13}^2 + \alpha_5 \epsilon_{31}^2 \\ & + \alpha_6 \epsilon_{23}^2 + \alpha_6 \epsilon_{32}^2 + 2\alpha_7 \epsilon_{11} \epsilon_{22} + 2\alpha_8 \epsilon_{11} \epsilon_{33} + 2\alpha_9 \epsilon_{22} \epsilon_{33}, \end{aligned}$$

and

$$\begin{aligned} F(\vec{a}, \bar{\epsilon}) = & a_1 \epsilon_{11}^2 + a_2 \epsilon_{22}^2 + a_3 \epsilon_{33}^2 + a_4 \epsilon_{12}^2 + a_4 \epsilon_{21}^2 + a_5 \epsilon_{13}^2 + a_5 \epsilon_{31}^2 \\ & + a_6 \epsilon_{23}^2 + a_6 \epsilon_{32}^2 + 2a_7 \epsilon_{11} \epsilon_{22} + 2a_8 \epsilon_{11} \epsilon_{33} + 2a_9 \epsilon_{22} \epsilon_{33}. \end{aligned}$$

Note that we have  $\vec{\alpha} = (\alpha_1, \dots, \alpha_9)$  and  $\vec{a} = (a_1, \dots, a_9)$  where  $\alpha_j \geq 0$  and  $a \geq 0$  for  $j = 1, \dots, 9$ .

As noted in [52, 67, 80], the first term in (4.2.8) is more sensitive to small strains while the second term is more sensitive to large strains, based on the fact that the second term involves an exponential function whereas the first term does not. In this case, “small” and “large” are relative terms, as for the situation here we are still in an overall small-scale case (though we note that this form for  $\rho_0 W^{(2)}$  could also be used in a finite-strain case). This form for  $\rho_0 W^{(2)}$  was originally chosen in [52] to model skin tissue, was adopted in a two-dimensional form by [67] when creating that particular body tissue propagation model for a related stenosis problem, and then was used in the three-dimensional formulation used by [80] when extending the two-dimensional model of [67].

Using (4.2.8) in (4.2.5), we find the elastic response forms:

$$\begin{aligned}
\sigma_{11}^e &= \frac{1}{2}(2\alpha_1\epsilon_{11} + 2\alpha_7\epsilon_{22} + 2\alpha_8\epsilon_{33}) + \frac{c}{2}(2a_1\epsilon_{11} + 2a_7\epsilon_{22} + 2a_8\epsilon_{33}) \exp(F(\vec{a}, \vec{e})) \\
\sigma_{22}^e &= \frac{1}{2}(2\alpha_2\epsilon_{22} + 2\alpha_7\epsilon_{11} + 2\alpha_9\epsilon_{33}) + \frac{c}{2}(2a_2\epsilon_{22} + 2a_7\epsilon_{11} + 2a_9\epsilon_{33}) \exp(F(\vec{a}, \vec{e})) \\
\sigma_{33}^e &= \frac{1}{2}(2\alpha_3\epsilon_{33} + 2\alpha_8\epsilon_{11} + 2\alpha_9\epsilon_{22}) + \frac{c}{2}(2a_3\epsilon_{33} + 2a_8\epsilon_{11} + 2a_9\epsilon_{22}) \exp(F(\vec{a}, \vec{e})) \\
\sigma_{12}^e &= \frac{1}{2}(2\alpha_4\epsilon_{12}) + \frac{c}{2}(2a_4\epsilon_{12}) \exp(F(\vec{a}, \vec{e})) \\
\sigma_{13}^e &= \frac{1}{2}(2\alpha_5\epsilon_{13}) + \frac{c}{2}(2a_5\epsilon_{13}) \exp(F(\vec{a}, \vec{e})) \\
\sigma_{23}^e &= \frac{1}{2}(2\alpha_6\epsilon_{23}) + \frac{c}{2}(2a_6\epsilon_{23}) \exp(F(\vec{a}, \vec{e}))
\end{aligned} \tag{4.2.9}$$

We thus see two general cases. For  $\sigma_{ii}^e$  where  $i = 1, 2, 3$ , the elastic response components depend on  $\epsilon_{11}$ ,  $\epsilon_{22}$ , and  $\epsilon_{33}$ . Each of the remaining components  $\sigma_{ij}^e$  where  $i \neq j$  depends on a single corresponding component  $\epsilon_{ij}$ . Thus, for example, if we find that  $\epsilon_{12} = 0$ , we would necessarily have that  $\sigma_{12}^e = 0$ . Based on the earlier assumptions on the form for  $\mathcal{G}$ , we would then also have  $\sigma_{12} = 0$ . We will refer back to (4.2.9) as we reduce the three-dimensional equations of motion (4.2.1) in the pressure and shear wave cases.

#### 4.2.1 One-dimensional Pressure Equation

Recall the setup of the pressure configuration in Figure 2.1. The experimental setup is such that the stress is imparted to the gel uniformly along the top  $x = L$  edge. This results in displacements only in the  $z$ -direction. Thus, we use the fact that the gel is uniform in the  $r$  and  $\theta$  directions as well as the fact that there are no displacements in the  $r$  and  $\theta$  directions to set  $u_r = u_\theta = 0$  and  $\frac{\partial(\cdot)}{\partial r} = \frac{\partial(\cdot)}{\partial \theta} = 0$ . Thus, from the equations in (4.2.1) we retain only the PDE for  $u_z$ , which is then simplified to

$$\rho \ddot{u}_z = \frac{\partial \sigma_{zz}}{\partial z} \tag{4.2.10}$$

since  $\epsilon_{rz} = \frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r} = 0$  and because  $\sigma_{rz}^e$  being dependent solely on  $\epsilon_{rz}$  by (4.2.9) results in  $\epsilon_{rz} = 0$ , which implies  $\sigma_{rz}^e = 0$ , which then implies  $\sigma_{rz} = 0$ . This also means that we have only the strain variable  $\epsilon_{zz}$  and the stress variable  $\sigma_{zz}^e$ , and thus the single quasi-linear stress-strain component  $\sigma_{zz} = \int_{-\infty}^t [\lambda(t-s) \frac{\partial}{\partial s} \sigma_{kk}^e(s) + 2\mu(t-s) \frac{\partial}{\partial s} \sigma_{zz}^e(s)] ds$ . Without further simplification to the forms for  $\sigma_{rr}^e$ ,  $\sigma_{\theta\theta}^e$ , and  $\sigma_{zz}^e$ , one must implement the constitutive relation as currently stated. This was done in [67, 80].

However, for our purposes, we assume that in our small-strain case we can treat the exponential terms in  $\sigma_{rr}^e$ ,  $\sigma_{\theta\theta}^e$ , and  $\sigma_{zz}^e$  as negligible (i.e., we set  $a_8 = a_9 = 0$ ). As a result, we

can simplify the form of the constitutive relationship for  $\sigma_{zz}$  as

$$\begin{aligned}
\sigma_{zz} &= \int_{-\infty}^t \left[ \lambda(t-s) \frac{\partial}{\partial s} (\sigma_{rr}^e(s) + \sigma_{\theta\theta}^e(s) + \sigma_{zz}^e(s)) + 2\mu(t-s) \frac{\partial}{\partial s} \sigma_{zz}^e(s) \right] ds \\
&= \int_{-\infty}^t \left[ \lambda(t-s) \frac{\partial}{\partial s} (\alpha_8 \epsilon_{zz}(s) + \alpha_9 \epsilon_{zz}(s) + \alpha_3 \epsilon_{zz}(s) + 2\mu(t-s) \frac{\partial}{\partial s} (\alpha_3 \epsilon_{zz}(s))) \right] ds \\
&= \int_{-\infty}^t [\lambda(t-s)(\alpha_8 + \alpha_9 + \alpha_3) + 2\mu(t-s)\alpha_3] \frac{\partial}{\partial s} \epsilon_{zz}(s) ds \\
&= \gamma \int_{-\infty}^t G(t-s) \frac{\partial}{\partial s} \epsilon_{zz}(s) ds,
\end{aligned} \tag{4.2.11}$$

where  $\gamma$  is a constant and we will later define the form for  $G(t)$  (which represents the combined effects of  $\lambda(t)$  and  $\mu(t)$  in this case) to complete the description of the constitutive equation (which we will later take on as a topic of consideration). We justify the factor  $\gamma$  by noting that in the development of (4.2.11) both  $\lambda(t)$  and  $\mu(t)$  are being multiplied by combinations of the constants  $\alpha_3$ ,  $\alpha_8$ , and  $\alpha_9$ , so it is reasonable to incorporate the overall effect by a single constant  $\gamma$ . We will later see that this formulation is adequate to describe experimental one-dimensional pressure data.

Later on in this work, we will use the independent variable  $x$  in the pressure case, a notation choice due in part to how the pressure independent variable was denoted in some of the preceding work completed by the author on this topic [14, 15], as well as being reflected in the variable used in Figure 2.1. Using this variable, the pressure PDE is

$$\rho \ddot{u} = \frac{\partial \sigma}{\partial x} \tag{4.2.12}$$

where  $\epsilon = \frac{\partial u}{\partial x}$  and  $\sigma = \gamma \int_{-\infty}^t G(t-s) \frac{\partial}{\partial s} \epsilon(s) ds$ . The change of independent variable to  $x$  will ensure we are consistent with the rest of this dissertation.

### 4.2.2 One-dimensional Shear Equation

Recall the setup of the shear configuration in Figure 2.2. We induce shear into the experimental phantom from a central rod moving in the  $z$ -direction, which produces shear waves that have displacement in the  $z$ -direction but move along the  $r$  axis. The shear input is then represented by the  $\sigma_{rz}$  stress on the boundary. We also then expect that we are interested in the  $u_z$  component, but solved along a radial axis. For the shear case, the experiment was designed such that the phantom was symmetric in  $\theta$  and uniform in  $z$ . We will work through these simplifying assumptions one at a time.

First, we will incorporate uniformity in  $\theta$  so that  $u_\theta = 0$  and  $\frac{\partial(\cdot)}{\partial \theta} = 0$ . This means that  $\frac{\partial u_\theta}{\partial r} = \frac{\partial u_\theta}{\partial z} = \frac{\partial u_\theta}{\partial \theta} = \frac{\partial u_r}{\partial \theta} = \frac{\partial u_z}{\partial \theta} = 0$ . As a result,  $\epsilon_{r\theta} = \epsilon_{\theta z} = 0$ . Given the relationships described

in (4.2.9), we then have that  $\sigma_{r\theta} = \sigma_{\theta z} = 0$ . This leaves us with the two simplified PDEs

$$\begin{aligned}\rho\ddot{u}_r &= \frac{\partial\sigma_{rr}}{\partial r} + \frac{\partial\sigma_{rz}}{\partial z} + \frac{1}{r}(\sigma_{rr} - \sigma_{\theta\theta}), \\ \rho\ddot{u}_z &= \frac{\partial\sigma_{rz}}{\partial r} + \frac{\partial\sigma_{zz}}{\partial z} + \frac{1}{r}\sigma_{rz}.\end{aligned}\tag{4.2.13}$$

Then, from simplifying and reducing (4.2.4) we have the following forms for the strains in terms of displacements:

$$\epsilon_{rr} = \frac{\partial u_r}{\partial r}, \quad \epsilon_{\theta\theta} = \frac{1}{r}u_r, \quad \epsilon_{zz} = \frac{\partial u_z}{\partial z}, \quad \epsilon_{rz} = \frac{1}{2}\left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r}\right).\tag{4.2.14}$$

The typical assumption used to further reduce the model dimension is that we have an “infinite” cylinder in the  $z$ -direction; this corresponds with assuming there are no changes in the  $z$ -direction so that  $\frac{\partial(\cdot)}{\partial z} = 0$ . This means that  $\frac{\partial u_r}{\partial z} = \frac{\partial u_z}{\partial z} = 0$ , and thus that  $\epsilon_{zz} = 0$ . We apply these simplifications to (4.2.13) in order to obtain

$$\begin{aligned}\rho\ddot{u}_r &= \frac{\partial\sigma_{rr}}{\partial r} + \frac{1}{r}(\sigma_{rr} - \sigma_{\theta\theta}), \\ \rho\ddot{u}_z &= \frac{\partial\sigma_{rz}}{\partial r} + \frac{1}{r}\sigma_{rz}.\end{aligned}\tag{4.2.15}$$

We also update the relations in (4.2.14):

$$\epsilon_{rr} = \frac{\partial u_r}{\partial r}, \quad \epsilon_{\theta\theta} = \frac{1}{r}u_r, \quad \epsilon_{rz} = \frac{1}{2}\frac{\partial u_z}{\partial r}.\tag{4.2.16}$$

From (4.2.16) we see that  $\sigma_{rr}$  and  $\sigma_{\theta\theta}$  depend on both  $\epsilon_{rr}$  and  $\epsilon_{\theta\theta}$ , while  $\sigma_{rz}$  depends only on  $\epsilon_{rz}$ . Note that a shear input implies that only (4.2.15) has a nonzero boundary condition. As we assume the phantom is at rest initially, this means that  $u_r = 0$  since there is no displacement or stress input in the  $r$ -direction and thus no dynamics. Accordingly, the single shear governing equation of interest is

$$\rho\ddot{u}_z = \frac{\partial\sigma_{rz}}{\partial r} + \frac{1}{r}\sigma_{rz},$$

with the single quasi-linear stress-strain component  $\sigma_{rz} = \int_{-\infty}^t \mu(t-s) \frac{\partial}{\partial s} \sigma_{rz}^e(s) ds$ .

As in the pressure case we assume the exponential term in  $\sigma_{rz}^e$  is negligible (i.e.,  $a_5 = 0$ ).



Thus, we simplify

$$\begin{aligned}
\sigma_{rz} &= \int_{-\infty}^t \mu(t-s) \frac{\partial}{\partial s} \sigma_{rz}^e(s) ds \\
&= \int_{-\infty}^t \mu(t-s) \frac{\partial}{\partial s} (\alpha_5 \epsilon_{rz}(s)) ds \\
&= \gamma \int_{-\infty}^t G(t-s) \frac{\partial}{\partial s} \epsilon_{rz}(s) ds,
\end{aligned} \tag{4.2.17}$$

where  $\gamma = \alpha_5$  in this case. Again,  $\gamma$  and  $G(t)$  will be later be defined to provide an appropriate constitutive relationship. The reuse of  $G(t)$  for the shear case is intentional; the form for  $G(t)$  for shear will be the same as that for the pressure case (4.2.12) though the parameter values will differ.

As with the pressure case, later on in this dissertation we will use a simplified form for the shear equation. We retain the independent variable  $r$ , and write the shear PDE as

$$\rho \ddot{u} = \frac{\partial \sigma}{\partial r} + \frac{1}{r} \sigma \tag{4.2.18}$$

where  $\epsilon = \frac{\partial u}{\partial r}$  and  $\sigma = \gamma \int_{-\infty}^t G(t-s) \frac{\partial}{\partial s} \epsilon(s) ds$ .

### 4.3 Initial and Boundary Conditions

Using the simplified pressure and shear equations, (4.2.12) and (4.2.18), we now specify initial conditions. In both the pressure and shear cases, the phantom will be initially at rest. In the pressure case, a uniform force is applied to the top of the device (c.f. Figure 2.1). The bottom of the phantom is set on the rig, so that no movement occurs along the  $x = 0$  edge. Thus, the governing partial differential equation becomes

$$\begin{aligned}
\rho \frac{\partial^2}{\partial t^2} u(x, t) - \frac{\partial}{\partial x} \sigma(x, t) &= 0 \\
u(0, t) = 0, \quad \sigma(L, t) &= -g(t) \\
u(x, \Gamma_1) = 0, \quad u_t(x, \Gamma_1) &= 0
\end{aligned} \tag{4.3.1}$$

where  $\rho$  is the density of the material, the stress tensor  $\sigma$  is given by the constitutive relationship for the material (the form of which will be discussed later),  $g(t)$  is a function that describes the loading process (again, to be discussed later), and the material is initially at rest. The value  $u(x, t)$  represents the displacement of the material at position  $x$  and time  $t$ , with  $x \in (0, L)$  and  $t > \Gamma_1$ . The time  $\Gamma_1$  is chosen as the beginning of any stress-strain history in the material; we are assuming the material has been at rest long enough that it is only affected by displacements

for  $t > \Gamma_1$ , where  $\Gamma_1$  is the time when we start modeling the material history. For our device,  $L = 0.0518$  m is the height of the phantom.

In the shear displacement case, the phantom is again at rest. The force is applied along the inner radius  $r_{min}$  (c.f. Figure 2.2), and the outer surface at  $r_{max}$  is fixed. Thus, the governing equation becomes

$$\begin{aligned}\rho \frac{\partial^2}{\partial t^2} u(r, t) - \frac{\partial}{\partial r} \sigma(r, t) - \frac{\sigma(r, t)}{r} &= 0 \\ \sigma(r_{min}, t) &= g(t), \quad u(r_{max}, t) = 0 \\ u(r, \Gamma_1) &= 0, \quad u_t(r, \Gamma_1) = 0\end{aligned}\tag{4.3.2}$$

where  $\rho$ ,  $\sigma$ , and  $g(t)$  are analogous to the pressure case and where  $r \in (r_{min}, r_{max})$  for  $t > \Gamma_1$ . For our device,  $r_{min} = 0.0105$  m and  $r_{max} = 0.054$  m. Throughout the remainder of this dissertation, we will use  $r$  as the spatial variable when the model is for shear displacement and  $x$  as the spatial variable for pressure displacement unless otherwise stated

In order to complete these models, we must provide a form for  $\sigma$ . This is the constitutive relationship, also called the stress-strain law since it relates strain ( $\frac{\partial u}{\partial x}$  or  $\frac{\partial u}{\partial r}$ ) and/or the strain rate to stress  $\sigma$ . The next sections discuss this aspect of the model.

## 4.4 Constitutive Equation

We incorporate the previous modeling ideas together into a new constitutive equation for the pressure (4.3.1) and shear (4.3.2) wave PDEs. The constitutive relationship form is the same for the pressure and shear cases, so  $x$  and  $r$  are interchangeable unless otherwise noted; for notational convenience, we use  $x$  as the spatial variable in the discussion which follows. The variable  $\epsilon$  then will denote  $\epsilon = u_x$  for the pressure case or  $\epsilon = u_r$  for the shear case. We also assume the material is initially at rest and has no relevant history, so that the integrals can be taken from the initial loading time denoted  $\Gamma_1$ .

### 4.4.1 Fung Quasi-Linear Model

Some of the initial investigation into the viscoelastic nature of tissue was completed by Fung (see [13] and the references therein). His work is of particular interest because it was validated in actual tissue. Fung developed a “quasi-linear” model

$$\sigma(t) = \int_{\Gamma_1}^t G(t-s) \frac{d\sigma^e(\lambda(s))}{ds} ds\tag{4.4.1}$$

with a kernel of the form

$$G(t) = \frac{1 + c \int_{\tau_1}^{\tau_2} \frac{1}{\tau} \exp(-t/\tau) d\tau}{1 + c \ln(\tau_2/\tau_1)}.\tag{4.4.2}$$

Within (4.4.1),  $\lambda$  represents the stretch of a material ( $\lambda = 1 + \epsilon$ ) and  $\sigma^e$  describes the elastic response to the elongation  $\lambda$ , given by (see [9])

$$\sigma^e(\lambda) = -\beta + \beta \exp(\alpha\epsilon)$$

where  $\alpha$  and  $\beta$  are constants to be estimated (and where the derivation of  $\sigma^e$  will be discussed in the next section). The parameters  $\tau_i$  are lower and upper bounds on *relaxation times*, which describe the ways in which the material responds to imposed stresses and strains. This model incorporates a continuum  $\tau \in [\tau_1, \tau_2]$  of relaxation times, which Fung found to be necessary in order for his model to match the response of tissue, as well as a constant term in the kernel. This Fung kernel will serve as a baseline which we will refer back to when developing the model for this paper.

#### 4.4.2 Relating the Quasi-Linear and Strain Energy Function Formulations

We have already seen that the assumptions of isotropy, the form (4.2.8) for the pseudo-strain energy function  $\rho_0 W^{(2)}$ , and the assumption that the exponential terms in  $\rho_0 W^{(2)}$  are negligible led to the one-dimensional stress-strain relationship

$$\sigma = \gamma \int_{-\gamma_1}^t G(t-s) \frac{\partial}{\partial s} \epsilon(s) ds$$

for both the pressure and shear cases. We will now relate this to another development of the one-dimensional constitutive relationship discussed by Fung and described above as (4.4.1)-(4.4.2).

In a separate one-dimensional development of his quasi-linear constitutive relationship in [52, Sec. 7.6], Fung related the instantaneous stress  $\sigma^e$  and the stretch ratio  $\lambda$  by the relationship

$$\frac{d\sigma^e}{d\lambda} = \alpha(\sigma^e + \beta)$$

where  $\alpha$  and  $\beta$  are constants used to fit the experimental results. This was based on a linear approximation of an experimental curve of  $\sigma^e$  against  $d\sigma^e/d\lambda$  (i.e., the instantaneous load versus how that load changes with respect to the strain ratio  $\lambda$ ). Solving that relationship yields  $\sigma^e = -\beta + (\sigma_0^e + \beta) \exp(\alpha(\lambda - \lambda_0))$ . Assuming we have no stress initially ( $\sigma_0^e = 0$ ) and that there is no initial strain ( $\lambda_0 = 1$ ), and noting that the relationship between the one-dimensional strain ratio and infinitesimal strain is  $\lambda = 1 + \epsilon$ , we find that

$$\sigma^e = -\beta + \beta \exp(\alpha\epsilon).$$

We again assume we are in the small-strain case, which we incorporate by linearizing the

exponential term  $\exp(\alpha\epsilon) \approx 1 + \alpha\epsilon + \dots$ , a linearization done in this case since the exponential term contains the strain variable. This leads to the simplified form

$$\sigma^e \approx \gamma\epsilon$$

where  $\gamma = \alpha\beta$  and again  $\epsilon = u_x$  in the pressure case or  $\epsilon = u_r$  in the shear case. This was used directly in the one-dimensional quasi-linear formulation

$$\sigma = \int_{-\Gamma_1}^t G(t-s) \frac{\partial}{\partial s} \sigma^e(s) ds \approx \gamma \int_{-\Gamma_1}^t G(t-s) \frac{\partial}{\partial s} \epsilon(s) ds.$$

This is now the exact same form as we obtained using the pseudo-strain energy approach.

We have now shown that in the case of one-dimensional small strains, the pseudo-strain energy approach to determining  $\sigma$  and Fung's one-dimensional approach to forming  $\sigma$  are equivalent. If one were to expand to multiple dimensions, the pseudo-strain energy function approach should be considered as in [67, 80].

#### 4.4.3 Macroscopic Damping

Though the form  $\gamma \int_{\Gamma_1}^t G(t-s) \frac{d\epsilon(s)}{ds} ds$  naturally allows us to describe some types of damping (see the later discussion on internal variables), we will also incorporate macroscopic damping into the model. This will be done in a phenomenological way, by adding a Kelvin-Voigt term to describe the overall nature of the damping present in the material (thus making the constitutive equation strain-rate dependent). This is a common damping model which has been used in previous work [20, 21, 22], and was preferred over the Maxwell model in the work by [36, 38, 39, 44, 59, 65, 81]. After incorporating Kelvin-Voigt damping, we obtain the constitutive relationship

$$\sigma(t) = E_1 u_{xt} + \gamma \int_{\Gamma_1}^t G(t-s) \frac{d\epsilon(s)}{ds} ds \quad (4.4.3)$$

where  $G(t)$  is still a kernel to be specified.

#### 4.4.4 Existence and Uniqueness for Pressure and Shear Models

Before moving on to the specific form of the constitutive equation kernel, we first establish existence and uniqueness for the pressure (4.3.1) and shear (4.3.2) equations with the constitutive equation (4.4.3). To that end, we set up a similar framework as in Chapter 3 and connect those results to the current model. We will require that the following assumptions hold:

(A1) The boundary condition function satisfies  $g \in L^2(\Gamma_1, T)$ ;

(A2) The kernel  $G$  is differentiable with respect to  $t \in \mathbb{R}^+$  and with constants  $c_1$  and  $c_2$  such that  $|G(t)| \leq c_1$  and  $|\dot{G}(t)| \leq c_2$  for all  $t \in \mathbb{R}^+$ .

### **Pressure case**

The pressure PDE (4.3.1) with constitutive equation (4.4.3) are of the same form as those in Section 3.1.1, except that here we have the initial time denoted as  $t = \Gamma_1$  instead of  $t = 0$  and slightly different variable names (inconsequential changes).

Let  $\mathbb{H} = \mathcal{L}^2(0, L)$ ,  $\mathbb{V} = \{\phi | \phi \in \mathcal{H}^1(0, L), \phi(0) = 0\}$ , and  $\mathbb{V}^*$  denote the topological dual space of  $\mathbb{V}$ . We identify  $\mathbb{H}$  with its topological dual  $\mathbb{H}^*$  and thus obtain  $\mathbb{V} \hookrightarrow \mathbb{H} = \mathbb{H}^* \hookrightarrow \mathbb{V}^*$  as a Gelfand triple [8, 99]. The notation  $\langle \cdot, \cdot \rangle$  denotes the inner product in  $\mathbb{H}$ , and  $\langle \cdot, \cdot \rangle_{\mathbb{V}^*, \mathbb{V}}$  represents the duality pairing between  $\mathbb{V}^*$  and  $\mathbb{V}$ . Let  $\mathcal{C}_w(\Gamma_1, T; \mathbb{V})$  denote the set of weakly continuous functions in  $\mathbb{V}$  on  $[\Gamma_1, T]$ , and  $\mathcal{L}_T = \{v : [\Gamma_1, T] \rightarrow \mathbb{H} \mid v \in \mathcal{C}_w(\Gamma_1, T; \mathbb{V}) \cap \mathcal{L}^2(\Gamma_1, T; \mathbb{V}) \text{ and } v_t \in \mathcal{C}_w(\Gamma_1, T; \mathbb{H}) \cap \mathcal{L}^2(\Gamma_1, T; \mathbb{V})\}$ . The notion of weakly continuous (i.e.,  $u^m \rightarrow u$  in  $\mathcal{C}_w(\Gamma_1, T; \mathbb{V})$ ) means that  $u^m \rightarrow u$  weakly in  $\mathbb{V}$  and uniformly in  $t \in [\Gamma_1, T]$ . Then a weak solution  $u \in \mathcal{L}_T$  for the pressure equation must satisfy

$$\begin{aligned} 0 = & \rho \langle u_t(t), \eta_t(t) \rangle - \rho \int_{\Gamma_1}^t \langle u_s(s), \eta_s(s) \rangle ds + \int_{\Gamma_1}^t g(s) \eta(L, s) ds + E_1 \int_{\Gamma_1}^t \langle u_{sx}(s), \eta_x(s) \rangle ds \\ & + \gamma \int_{\Gamma_1}^t \left\langle \int_{\Gamma_1}^s G(s - \xi) \frac{d}{d\xi} u_x(\xi) d\xi, \eta_x(s) \right\rangle ds \end{aligned} \quad (4.4.4)$$

for any  $t \in [\Gamma_1, T]$  and  $\eta \in \mathcal{L}_T$ . Here and elsewhere  $u(t)$  and  $\eta(t)$  denote the functions  $u(\cdot, t)$  and  $\eta(\cdot, t)$ , respectively. With these definitions, we still have that the following theorem (a restatement of Theorem 3.1.2) holds:

**Theorem 4.4.1.** *Assuming (A1) and (A2), the pressure equation (4.3.1) with the constitutive relation (4.4.3) has a unique weak solution on any finite interval  $[\Gamma_1, T]$ .*

### **Shear case**

This requires a bit more consideration. The shear domain is  $\Omega = [r_{min}, r_{max}]$ , and is solved on the time frame  $t \in [\Gamma_1, T]$ . We must slightly redefine the spaces from above to fit the shear model. Let  $\mathbb{H} = \mathcal{L}^2(r_{min}, r_{max})$ ,  $\mathbb{V} = \{\phi | \phi \in \mathcal{H}^1(r_{min}, r_{max}), \phi(r_{max}) = 0\}$ , and  $\mathbb{V}^*$  denote the topological dual space of  $\mathbb{V}$ . We identify  $\mathbb{H}$  with its topological dual  $\mathbb{H}^*$  and thus again obtain  $\mathbb{V} \hookrightarrow \mathbb{H} = \mathbb{H}^* \hookrightarrow \mathbb{V}^*$  as a Gelfand triple. Let  $\mathcal{C}_w(\Gamma_1, T; \mathbb{V})$  denote the set of weakly continuous functions in  $\mathbb{V}$  on  $[\Gamma_1, T]$ , and  $\mathcal{L}_T = \{v : [\Gamma_1, T] \rightarrow \mathbb{H} \mid v \in \mathcal{C}_w(\Gamma_1, T; \mathbb{V}) \cap \mathcal{L}^2(\Gamma_1, T; \mathbb{V}) \text{ and } v_t \in \mathcal{C}_w(\Gamma_1, T; \mathbb{H}) \cap \mathcal{L}^2(\Gamma_1, T; \mathbb{V})\}$ . Then a weak solution  $u \in \mathcal{L}_T$  for the shear equation must

satisfy

$$\begin{aligned}
0 = & \rho \langle u_t(t), \eta_t(t) \rangle - \rho \int_{\Gamma_1}^t \langle u_s(s), \eta_s(s) \rangle ds + \int_{\Gamma_1}^t g(s) \eta(r_{min}, s) ds + E_1 \int_{\Gamma_1}^t \langle u_{sr}(s), \eta_r(s) \rangle ds \\
& + \gamma \int_{\Gamma_1}^t \left\langle \int_{\Gamma_1}^s G(s - \xi) \frac{d}{d\xi} u_r(\xi) d\xi, \eta_r(s) \right\rangle ds - E_1 \int_{\Gamma_1}^t \int_{r_{min}}^{r_{max}} \frac{u_{rt}(r, s)}{r} \eta(r, s) dr ds \\
& - \gamma \int_{\Gamma_1}^t \int_{r_{min}}^{r_{max}} \left( \int_{\Gamma_1}^s \frac{1}{r} G(s - \xi) \frac{du_r(r, s)}{d\xi} d\xi \right) \eta(r, s) dr ds
\end{aligned} \tag{4.4.5}$$

for any  $t \in [\Gamma_1, T]$  and  $\eta \in \mathcal{L}_T$  and where  $\langle \cdot, \cdot \rangle$  is the usual inner product. Since  $r_{min} > 0$ , there are no singularities in the final term in (4.4.5), and the kernel integral in the numerator of that term will converge in the same manner as the preceding kernel integral. Thus, the arguments from the pressure case apply in the shear case, and we have the following theorem:

**Theorem 4.4.2.** *Assuming (A1) and (A2), the shear equation (4.3.2) with constitutive relationship (4.4.3) has a unique weak solution on any finite interval  $[\Gamma_1, T]$ .*

#### 4.4.5 Form for Constitutive Equation Kernel $G(t)$

We will now state the particular kernel used in the rest of this dissertation, and then manipulate it into a form that gives more physical insight and which will later allow for a conceptual framework using *internal variables*. We develop this kernel from a different perspective than that given in Chapter 3, but the resulting form will be quite similar. Using the notation and parameter conventions of [13], we define the kernel in this work to be

$$G(t; P) = \kappa_r + K(t; P) \tag{4.4.6}$$

where  $\kappa_r$  is a constant representing an instantaneous relaxation modulus (justified by the fact that our gel phantom acts partly as a solid) and  $K(t; P) = \int_{\mathcal{T}} \exp(-t/\tau) dP(\tau)$  represents a continuum of distributed relaxation times with  $\mathcal{T} = [\tau_1, \tau_2] \subset (0, \infty)$  and where  $P(\tau)$  is a probability measure on  $\mathcal{T}$ . Note that this form for  $G$  satisfies  $|G(t)| \leq c_1$  with  $G$  clearly differentiable and  $|\dot{G}(t)| \leq c_2$  for some constants  $c_1, c_2$  so that assumption (A2) is satisfied. It is also worth noting here that our proposed kernel form (4.4.6) is similar to that in Fung's model (4.4.2), as we see that  $\kappa_r$  serves as an analog to the constant portion of Fung's kernel (i.e.,  $\frac{1}{1+c \ln(\tau_1/\tau_1)}$ ) and the  $K(t; P)$  portion is similar to the the continuous relaxation spectrum in Fung's model (i.e.,  $\frac{c \int_{\tau_1}^{\tau_2} \frac{1}{\tau} \exp(-t/\tau) d\tau}{1+c \ln(\tau_2/\tau_1)}$ ). It is also similar to the Prony series (3.1.6) used in the preliminary model of Chapter 3.

We substitute (4.4.6) into (4.4.3) and manipulate the form of the stress, noting that  $u_x(\Gamma_1) =$

0 since the material is initially at rest and using the fact that  $K(0; P) = 1$ :

$$\begin{aligned}
\sigma(t; P) &= E_1 u_{xt} + \gamma \int_{\Gamma_1}^t G(t-s) \frac{du_x(s)}{ds} \\
&= E_1 u_{xt} + \gamma \int_{\Gamma_1}^t (\kappa_r + K(t-s; P)) \frac{du_x(s)}{ds} \\
&= E_1 u_{xt} + \kappa_r \gamma \int_{\Gamma_1}^t \frac{d}{ds} u_x(s) ds + \gamma \int_{\Gamma_1}^t K(t-s; P) \frac{d}{ds} u_x(s) ds \\
&= E_1 u_{xt} + \kappa_r \gamma \left( u_x(t) - \underbrace{u_x(\Gamma_1)}_0 \right) + \gamma \int_{\Gamma_1}^t K(t-s; P) \frac{d}{ds} u_x(s) ds \\
&= E_1 u_{xt} + \kappa_r \gamma u_x(t) + \gamma \int_{\Gamma_1}^t K(t-s; P) \frac{d}{ds} u_x(s) ds \\
&= E_1 u_{xt} + E u_x(t) + \gamma \left( \underbrace{K(0; P) u_x(t) - K(t; P) u_x(\Gamma_1)}_0 - \int_{\Gamma_1}^t \frac{\partial K(t-s; P)}{\partial s} u_x(s) ds \right) \\
&= (E + \gamma) u_x(t) + E_1 u_{xt}(t) - \gamma \int_{\Gamma_1}^t \frac{\partial K(t-s; P)}{\partial s} u_x(s) ds, \tag{4.4.7}
\end{aligned}$$

where  $E = \kappa_r \gamma$ . This equation (4.4.7) is the general form of the constitutive equation used here. The value  $E_0 = E + \gamma$  can be considered to be a dynamic analog to the static Young's modulus in the pressure case or the static shear modulus in the shear case; this also makes clear the fact that Hooke's Law is incorporated into our model. We have already discussed that  $E_1$  is the bulk damping parameter for the Kelvin-Voigt damping term. The final integral represents a history term which describes the relaxation of the material in response to an applied stress/strain.

We will ultimately turn to a discretized distribution model (using a discrete measure  $P(\tau)$ ), and connect it to the continuum model through a probability measure approximation as in [10]. This will allow us to develop a computationally feasible inverse problem, and also give insight into the underlying material mechanics. But first we briefly discuss a method for approximating the loading process.

#### 4.4.6 Approximating the Loading Process

Recalling Figure 2.3, the loading profile is relatively long compared with the oscillatory period; since our concern is with modeling the oscillations, solving the model from  $\Gamma_1$  is much longer than necessary. Also, early experimentation with the model indicated that the parameters governing the loading and resting process may differ from those governing the very dynamic post-release oscillatory process.

We address these concerns by modeling the loading as instantaneous from at rest to a

displacement of  $A$  at position  $x = L$  or  $r = r_{min}$ . Since the material is linear, this would then mean the phantom has the profile  $u(x, t) = \frac{A}{L}x$  in the pressure case and  $u(r, t) = \frac{A(r_{max}-r)}{r_{max}-r_{min}}$  in the shear case, up until the time of the weight release. Since this is an approximation, we will neglect the times  $t \in (\Gamma_3, \Gamma_4)$ , the weight release time period, since that time frame is small relative to the loading and settling time from  $\Gamma_1$  to  $\Gamma_3$ . We also incorporate a time parameter  $\Upsilon$  which will represent our approximation of the time when loading begins. In the formulation here we will use the same relaxation times during the loading process as during the oscillation period, which means that  $\Upsilon$  has no meaning other than as a tuning parameter that we must estimate. Thus, we assume the given loading profiles for  $t \in (\Upsilon, 0)$  since  $\Gamma_4 = 0$  in our convention. This also means that  $\Upsilon < 0$ .

We incorporate this loading approximation into our model by manually integrating the constitutive relationship (4.4.7). For the purposes here, we will call  $\hat{\sigma}$  the full constitutive relationship for  $t > \Upsilon$  that is described by (4.4.7) (where we now use  $\Upsilon$  in the place of  $\Gamma_1$ ), and  $\sigma$  the constitutive relationship for  $t > 0$ . We do this for notational simplicity in the final model, at the expense of some minor notational confusion at the current stage.

For the pressure case, we compute (noting that  $u(x, t) = \frac{A}{L}x$  implies  $u_x(x, t) = A/L$ , for  $\Upsilon < t < 0$ )

$$\begin{aligned}
\hat{\sigma}(t; P) &= (E + \gamma) u_x(t) + E_1 u_{xt}(t) - \gamma \int_{\Upsilon}^t \frac{\partial K(t-s; P)}{\partial s} u_x(s) ds \\
&= (E + \gamma) u_x(t) + E_1 u_{xt}(t) - \gamma \int_{\Upsilon}^0 \frac{\partial K(t-s; P)}{\partial s} u_x(s) ds - \gamma \int_0^t \frac{\partial K(t-s; P)}{\partial s} u_x(s) ds \\
&= (E + \gamma) u_x(t) + E_1 u_{xt}(t) - \gamma \int_{\Upsilon}^0 \frac{\partial K(t-s; P)}{\partial s} \frac{A}{L} ds - \gamma \int_0^t \frac{\partial K(t-s; P)}{\partial s} u_x(s) ds \\
&= (E + \gamma) u_x(t) + E_1 u_{xt}(t) - \gamma \frac{A}{L} (K(t; P) - K(t - \Upsilon; P)) - \gamma \int_0^t \frac{\partial K(t-s; P)}{\partial s} u_x(s) ds \\
&= \sigma(t; P) - \mathcal{F}(t; \Upsilon, A, P)
\end{aligned}$$

where  $\mathcal{F}(t; \Upsilon, A, P) = \gamma \frac{A}{L} (K(t; P) - K(t - \Upsilon; P))$  and

$$\sigma = (E + \gamma) u_x(t) + E_1 u_{xt}(t) - \gamma \int_0^t \frac{\partial K(t-s; P)}{\partial s} u_x(s) ds \tag{4.4.8}$$

incorporates the remaining terms and represents the constitutive relationship for  $t > 0$ . For the pressure setup, we then have the following:

- $\hat{\sigma}_x = \sigma_x$
- The original stress boundary condition is  $\hat{\sigma}(L, t; P) = 0$ . Using the preceding development,



this corresponds with

$$0 = \sigma(L, t; P) - \mathcal{F}(t; \Upsilon, A, P)$$

which allows us to write the boundary condition for a model solved for  $t > 0$  as

$$\sigma(L, t; P) = \mathcal{F}(t; \Upsilon, A, P).$$

Since  $A, \Upsilon < 0$  and  $\gamma > 0$ , we know that  $K(t; P) - K(t - \Upsilon; P) > 0$ . Hence, since  $L > 0$ , we have a compressive boundary stress, which is what we would expect in the pressure case.

The shear case is similar. For the loading profile  $u(r, t) = \frac{A(r_{max}-r)}{r_{max}-r_{min}}$  for  $t \in (\Upsilon, 0)$ , we have  $u_r(r, t) = -\frac{A}{r_{max}-r_{min}}$  which is incorporated when integrating the history in the same way as the pressure case. We then find the corresponding loading stress to be  $\mathcal{F}(t; \Upsilon, A, P) = -\zeta \frac{A}{r_{max}-r_{min}} (K(t; P) - K(t - \Upsilon; P))$ , where  $\zeta$  is the shear analog to  $\gamma$ . Also, we have  $\hat{\sigma}_r = \sigma_r$  as in the pressure case. However, we have the term  $\frac{\hat{\sigma}}{r} = \frac{\sigma}{r} - \frac{\mathcal{F}(t; \Upsilon, A, P)}{r}$ , which will result in a time-dependent forcing term in the shear PDE.

We make two comments before discussing the internal variable forms. First, if we assume, for example, a single relaxation time and that its value is small, say on the order of  $10^{-1}$ , then the term  $K(t - \Upsilon; P) = \exp(-(t - \Upsilon)/\tau_1) \approx \exp(-10(t - \Upsilon))$  attains its maximum value  $\exp(10\Upsilon)$  when  $t = 0$ . Note that for, say,  $\Upsilon < -1$ , this term is negligible. Relaxation times on this order are what we can later obtain in the inverse problem, which would imply that in our case the material is at rest after being loaded sufficiently long that it “forgets” its loading history by the time the weight is released. This is good from an experimental standpoint, since the loading process will never be quite uniform. It is also good to know from a computational perspective; we can limit  $\Upsilon$  to being greater than some value, such as  $-20 < \Upsilon < 0$ , which will keep the optimization algorithm from marching off unnecessarily (which occurred in some of our early inverse problem tests). Second, since we have integrated out the loading history, we now start the model at the time of weight release,  $t = 0$ . This means that the material is considered at rest just prior to the release; thus, in the history integrals we will discuss in the next section, all the history now starts at  $t = 0$  since the history before that point will be incorporated into the initial loading profile and an initial stress condition.

#### 4.4.7 Internal Variable Formulation

In the previously noted work on this stenosis problem, the double integrals that resulted from using the continuum of relaxation times in the stress equation were computationally intractable so another approach was required. The idea was to use a discrete number of internal variables. As will be noted, these gave rise to a differential form which was an improvement computationally

since it led to purely differential equations in the model rather than inclusion of an integro-differential equation. With the advances in desktop computation abilities since that time, the integral form is now reasonable to use in a dynamic model. However, internal variables are still attractive in that they provide a formulation that indicates some of the internal material dynamics. Physically, if we assume that the molecules within the biological tissue are on a microscopic scale then the portion of the material which is represented by each internal variable or internal strain  $\epsilon_j$  is being driven by the overall strain and has a response that varies depending on the value of the corresponding relaxation time  $\tau_j$ .

One of the earlier constitutive relationship formulations, in [9, 67, 80], approximates the Fung kernel as a finite sum of exponential functions

$$G(t) = \sum_{j=1}^N C_j \exp(-t/\tau_j)$$

where  $C_j$  are weights and  $\tau_j$  are *relaxation times* that describe how the material relaxes after undergoing deformation. If one uses this form for  $G(t)$  in the constitutive relationship (4.4.1), one still must compute the integral in

$$\sigma(t) = \int_0^t \sum_{j=1}^N C_j \exp(-(t-s)/\tau_j) \frac{d}{ds} \sigma^e(\lambda(s)) ds = \sum_{j=1}^N C_j \epsilon_j(t)$$

where  $\epsilon_j(t) = \int_0^t \exp(-(t-s)/\tau_j) \frac{d}{ds} \sigma^e(\lambda(s)) ds$ . We can instead compute each  $\epsilon_j(t)$  as a dynamic internal variable following the differential equation (for  $j = 1, 2, \dots, N$ )

$$\frac{d\epsilon_j(t)}{dt} + \frac{1}{\tau_j} \epsilon_j(t) = \frac{d\sigma^e(\lambda(t))}{dt}, \quad \epsilon_j(0) = 0.$$

Note that these are then linear differential equations for  $\epsilon_j$ . One could introduce nonlinearities, which is discussed in [9, 67] and was found to be equivalent to assuming multiple relaxation times (an example of where nonlinearities were necessary is the work on modeling wave propagation in filled rubber in [20, 21]). We do not consider nonlinear internal dynamics since we shall see later that the linear constitutive relationship (4.4.7) with a discrete measure gives a reasonable approximation to the data provided by QMUL and BHT.

Note, however, that the kernel here is composed of a discrete sum of exponentials. At first glance, this appears to run counter to the Fung results which point toward a continuum of times being important. The results in [9, 67, 80] demonstrate that the internal variable approach is valid and does appear to work as well as the continuum of times in the Fung kernel, but we would like to put this on firmer ground. A connection between the Fung continuum model and

the discrete kernel is provided by the work in [22]. The authors there form the kernel

$$G(t) = \int_{\mathcal{T}} R(t; \tau) dP(\tau)$$

where  $\mathcal{T} = [\tau_1, \tau_2] \subset (0, \infty)$  is the set of admissible relaxation times,  $P(\tau)$  is a probability measure on  $\mathcal{T}$ , and  $R(t; \tau)$  is a continuous function of relaxation times. If we take  $R(t; \tau) = \exp(-t/\tau)$ , this corresponds with the kernels previously discussed. The authors showed existence and uniqueness results for this kernel in the nonlinear constitutive equation (4.4.1). Though this framework is back to the continuous relaxation time case, a result from [10] allows one to approximate any measure  $P(\tau)$  with a discrete measure. This discrete measure leads us back to the previous case with a sum of exponentials, but from the probabilistic framework we know conclusively that we are approximating the continuous spectrum of Fung and from the results of [9, 67, 80] we know that this approximation has been viable when implemented.

With this understanding of previous work using internal variables, we move forward by modifying our current model. We manipulate the form of Equation (4.4.8) as follows:

$$\begin{aligned} \sigma(t; P) &= (E + \gamma) u_x(t) + E_1 u_{xt}(t) - \gamma \int_0^t \frac{\partial K(t-s; P)}{\partial s} u_x(s) ds \\ &= (E + \gamma) u_x(t) + E_1 u_{xt}(t) - \gamma \int_0^t \frac{\partial}{\partial s} \left( \int_{\mathcal{T}} \exp(-(t-s)/\tau) dP(\tau) \right) u_x(s) ds \\ &= (E + \gamma) u_x(t) + E_1 u_{xt}(t) - \gamma \int_{\mathcal{T}} \int_0^t \frac{\partial}{\partial s} (\exp(-(t-s)/\tau)) u_x(s) ds dP(\tau) \\ &= (E + \gamma) u_x(t) + E_1 u_{xt}(t) - \gamma \int_{\mathcal{T}} \epsilon_1(t; \tau) dP(\tau), \end{aligned} \quad (4.4.9)$$

where in the final step  $\epsilon_1(t; \tau) = \int_0^t \frac{\partial}{\partial s} (\exp(-(t-s)/\tau)) u_x(s) ds$ . Rather than the integral form for  $\epsilon_1$ , we can use the differential form

$$\tau \frac{d}{dt} \epsilon_1(t; \tau) + \epsilon_1(t; \tau) = u_x(t), \quad \epsilon_1(0; \tau) = 0 \quad (4.4.10)$$

which is then solved simultaneously with the rest of the model dynamics. This is then an *internal variable* or *internal strain*, driven by the overall strain  $u_x(t)$ , which is the continuous form of the internal variable formulation.

We now may finally make the discrete assumption

$$P(\tau) = \sum_{j=1}^{N_p} p_j \Delta_{\tau_j}$$

where  $\Delta_{\tau_j}$  is the Heaviside function with step at  $\tau_j$  and  $p_j$  are the proportions of the material

subject to relaxation time  $\tau_j$ . By substituting this discrete  $P$  into the form for  $\sigma$  as developed in (4.4.9), we obtain the discrete, internal variable form of the constitutive relationship

$$\sigma(t) = \left( E + \sum_{j=1}^{N_p} \gamma_j \right) u_x(t) + E_1 u_{xt}(t) - \sum_{j=1}^{N_p} \gamma_j \epsilon^j(t), \quad (4.4.11)$$

with internal variables obeying (for  $j = 1, 2, \dots, N_p$ )

$$\tau_j \frac{d}{dt} \epsilon^j(t) + \epsilon^j(t) = u_x(t), \quad \epsilon^j(0) = 0, \quad (4.4.12)$$

and where we have defined  $\gamma_j = \gamma p_j$  so that  $\gamma = \sum_{j=1}^{N_p} \gamma_j$ . Note that we assume  $E > 0$ , since the agar gel acts at least partly as a solid, and that  $\epsilon^j = \epsilon_1(\cdot; t_j)$ .

#### 4.4.8 Final Pressure and Shear PDE Models

We now put together the pressure (4.3.1) and shear (4.3.2) wave equations, using the constitutive equation (4.4.11)-(4.4.12) but with the loading history approximation integrated out as discussed in Section 4.4.6. Recall also that the discrete assumption for  $P$  and the form of  $K$  gives us  $\gamma K(t; P) = \gamma \sum_{j=1}^{N_p} p_j \exp(-t/\tau_j) = \sum_{j=1}^{N_p} \gamma_j \exp(-t/\tau_j)$  where  $\gamma_j = \gamma p_j$ . These equations are just manipulated versions of the general equations of Theorems 4.4.1-4.4.2, so we still know a unique weak solution exists on any finite time interval.

##### Pressure Model

We first define

$$\mathcal{F}_p(t; A, \Upsilon, \gamma_1, \dots, \gamma_{N_p}, \tau_1, \dots, \tau_{N_p}) = \sum_{j=1}^{N_p} \gamma_j \exp(-t/\tau_j) - \sum_{j=1}^{N_p} \gamma_j \exp(-(t - \Upsilon)/\tau_j).$$

Then the pressure equations, solved for  $t > 0$  which is the release time, are

$$\begin{aligned} \rho \frac{\partial^2}{\partial t^2} u(x, t) - \frac{\partial}{\partial x} \sigma(x, t) &= 0 \\ u(0, t) &= 0, \quad \sigma(L, t) = \frac{A}{L} \mathcal{F}_p(t) \\ u(x, 0) &= \frac{A}{L} x, \quad u_t(x, 0) = 0, \end{aligned} \quad (4.4.13a)$$

where

$$\sigma(t) = \left( E + \sum_{j=1}^{N_p} \gamma_j \right) u_x(t) + E_1 u_{xt}(t) - \sum_{j=1}^{N_p} \gamma_j \epsilon^j(t) \quad (4.4.13b)$$

with the internal variables subject to (for  $j = 1, 2, \dots, N_p$ )

$$\tau_j \frac{d}{dt} \epsilon^j(t) + \epsilon^j(t) = u_x, \quad \epsilon^j(0) = 0. \quad (4.4.13c)$$

The parameter  $\rho$  still represents the material density and  $E_1$  the damping parameter.  $E$  represents an instantaneous relaxation modulus. The  $\gamma_j$  values are weightings on relaxation times  $\tau_j$ ; also, we can write  $E_0 = E + \sum_{j=1}^{N_p} \gamma_j$  as the viscoelastic analog to Young's modulus.

### Shear Model

We next present the shear equations. In order to more easily distinguish between pressure and shear model parameters, we will use  $G$  and  $G_1$  in place of  $E$  and  $E_1$  and  $\zeta_j$  instead of  $\gamma_j$ . We define

$$\mathcal{F}_s(t; A, \Upsilon, \zeta_1, \dots, \zeta_{N_p}, \tau_1, \dots, \tau_{N_p}) = \sum_{j=1}^{N_p} \zeta_j \exp(-t/\tau_j) - \sum_{j=1}^{N_p} \zeta_j \exp(-(t - \Upsilon)/\tau_j).$$

Then we have

$$\begin{aligned} \rho \frac{\partial^2}{\partial t^2} u(r, t) - \frac{\partial}{\partial r} \sigma(r, t) - \frac{\sigma(r, t)}{r} &= \frac{1}{r} \frac{A}{r_{max} - r_{min}} \mathcal{F}_s(t) \\ \sigma(r_{min}, t) &= \frac{-A}{r_{max} - r_{min}} \mathcal{F}_s(t), \quad u(r_{max}, t) = 0 \\ u(r, 0) &= \frac{A(r_{max} - r)}{r_{max} - r_{min}}, \quad u_t(r, 0) = 0, \end{aligned} \quad (4.4.14a)$$

where

$$\sigma(t) = \left( G + \sum_{j=1}^{N_p} \zeta_j \right) u_r(t) + G_1 u_{rt}(t) - \sum_{j=1}^{N_p} \zeta_j \epsilon^j(t) \quad (4.4.14b)$$

with the internal variables subject to (for  $j = 1, 2, \dots, N_p$ )

$$\tau_j \frac{d}{dt} \epsilon^j(t) + \epsilon^j(t) = u_r, \quad \epsilon^j(0) = 0. \quad (4.4.14c)$$

We note that  $G_0 = G + \sum_{j=1}^{N_p} \zeta_j$  is the dynamic analog of the shear modulus.

## 4.5 Numerical Method

We use the same numerical implementation for both the pressure (4.4.13) and shear (4.4.14) models, which is also the same as in Chapter 3. In time, we use a discontinuous Galerkin method composed of normalized Legendre polynomials (of order 4). In space, we use a continuous spectral finite element method composed of Lagrange basis functions on Gauss-Lobatto nodes (also of order 4). This allows the higher order (4-5th order) elements in space while controlling dispersion

error. Under this scheme, the system matrices are diagonalizable and hence the time-coupled computations within a time step can be decoupled. This makes the reasonably high order finite element time discretizations feasible. The reason for adopting this higher order scheme instead of the lower order ones is that higher order spatial discretizations have been found highly desirable for wave equations in terms of the control of dispersion errors (e.g., see [1, 26]). In addition, higher order schemes are capable of providing higher fidelity solutions than lower order schemes for the same amount of computational work. This is especially important to make the inverse problem practical (as each inverse problem may require solving the forward problem hundreds of times). Further details on this numerical method are in a forthcoming BICOM report [61], wherein an extensive set of numerical results are given to demonstrate the favorable effect on the numerical error and computational work of the higher order temporal and spatio-temporal approximations.

---

## Parameter Estimation and Results

---

With models in hand, we now turn to matching the model output to data. We will use two common methods in order to estimate model parameters. One is ordinary least squares (OLS) and the other is generalized least squares (GLS). These will be defined later in Section 5.2. Also, note that the majority of this chapter has been submitted for publication ([15]).

### 5.1 Inverse Problem Preliminaries and Definitions

As discussed in Chapter 2, separate novel acoustic phantom experiments have been designed to gather one-dimensional pressure and shear data. Measurements in our experiment are taken at  $x = L$  for the pressure case and  $r = r_{min}$  for the shear case, and will be denoted  $u_j$ . Corresponding pressure or shear model solutions at the same spatial location will be denoted  $u(t_j; 10^q)$ , where the measurement location has been suppressed so we can retain a general pressure/shear model solution notation and where  $q$  represents a vector of the base-10 logarithm of each parameter (the same idea used in Chapter 3 to reduce parameter scaling issues). Each data set has been trimmed to the dynamic oscillations after the release, and thus the time frame for pressure data is roughly 150ms while that for shear is 200ms, with a data sampling rate of 2.048kHz. Using the full set of data points proved to make the inverse problem difficult and computationally intractable, as that many data points made the inverse problem too overdetermined. Thus, we use the data with a sampling rate of 1.024kHz instead as the full data set. We take  $n$  to be the total number of data points for a particular data set, and thus can describe the measurement time points for the full “every data point” set as  $t_j = j/1024$  where  $j = 0, 1, \dots, n - 1$ . There

will also be a reduced data set where we take every other data point starting with  $t_0 = 0$ .

Since some of the data points were near zero in absolute value, we found that those points resulted in scaling problems when using the GLS model to estimate model parameters (since the corresponding cost functional divides by the model value as we will see later when this method is defined). To account for this, we removed from consideration any data points  $u_j$  (and their corresponding model solutions at that time point) where  $|u_j| < 5 \times 10^{-6}$ . This value was chosen by examining the data, noting that the data is on the order of  $10^{-5}$  and that the “jitter” one can see in Figure 2.3 has a magnitude of roughly  $5 \times 10^{-6}$  during the times before loading up to  $\Gamma_1$ , then during the settling period from  $\Gamma_2$  to  $\Gamma_3$ , and again in the settling period after the oscillations have died out. Thus, our threshold level is below the level of noise in the data. This level also eliminated only a few data points, while providing significantly improved GLS robustness. The number of data points  $n$  is then reduced according to how many thresholded data points were removed.

Before going into the setup and results for the inverse problem, we note that the forward (i.e., direct) problems where we solve for displacement (using the method discussed in Section 4.5) are as follows:

- **Pressure forward problem:** Given  $E$ ,  $E_1$ ,  $\tau_j$  and  $\gamma_j$  for  $j = 1, 2, \dots, N_p$ ,  $\Upsilon$ ,  $A$ ,  $L$ , and  $\rho$ , solve model (4.4.13) for displacement  $u(x, t)$  at each position  $x \in [0, L]$  for  $t \in [0, T]$ .
- **Shear forward problem:** Given  $G$ ,  $G_1$ ,  $\tau_j$  and  $\zeta_j$  for  $j = 1, 2, \dots, N_p$ ,  $\Upsilon$ ,  $A$ ,  $r_{min}$ ,  $r_{max}$ , and  $\rho$ , solve model (4.4.14) for displacement  $u(r, t)$  at each position  $x \in [r_{max}, r_{min}]$  for  $t \in [0, T]$ .

The inverse problems we will develop here are as follows:

- **Pressure inverse problem:** Given pressure displacement data at  $x = L$  and a corresponding forward problem solver for displacement, along with specified values for  $\rho$  and  $L$ , find values for the constants  $E$ ,  $E_1$ ,  $\tau_j$  and  $\gamma_j$  (for  $j = 1, 2, \dots, N_p$ ),  $A$ , and  $\Upsilon$  which provide the best fit to the data (in a manner which will be defined shortly).
- **Shear inverse problem:** Given shear displacement data at  $r = r_{min}$  and a corresponding forward problem solver for displacement, along with specified values for  $\rho$ ,  $r_{min}$ , and  $r_{max}$ , find values for the constants  $G$ ,  $G_1$ ,  $\tau_j$  and  $\zeta_j$  (for  $j = 1, 2, \dots, N_p$ ),  $A$ , and  $\Upsilon$  which provide the best fit to the data (again, in a manner which will be defined shortly).

We assume for both the pressure and shear cases that the parameters lie in some admissible set  $Q \subset \mathbb{R}^\kappa$ , where  $Q$  is assumed to be compact and  $\kappa$  is the number of parameters requiring estimation. We will the log-scaled parameter vector for pressure (for  $N_p = 1$ ) as

$$q = (\log_{10}(E), \log_{10}(E_1), \log_{10}(\gamma_1), \log_{10}(\tau_1), \log_{10}(-A), \log_{10}(-\Upsilon)) \quad (5.1.1)$$



with a similar vector for shear where we use the shear parameters  $G$ ,  $G_1$ , and  $\zeta_1$  in place of  $E$ ,  $E_1$ , and  $\gamma_1$ , respectively. Thus, as long as we define our cost function to be a continuous function of the parameters, we know the inverse problem has a solution (minimizing a continuous function on a compact parameter space). One could broaden this parameter estimation framework to the distributional case if desired, taking an admissible parameter space as a compact subset of Euclidean space (including all parameters excluding relaxation times) along with the space of probability measures, and use the Prohorov metric framework (see, e.g., [17, Sec. 4]) and the approximation results of [10]. This again leads to minimizing a continuous function of the parameters over a compact space. Either way, the inverse problems we will shortly define will have solutions.

### 5.1.1 Sensitivity of Model Output to Parameters

We consider the sensitivity of the model output to the parameters (equations are derived in Appendix B). For our examinations here, we will look at both the pressure and shear parameter sensitivities. The weight level is 264g, and we will show results for one relaxation time ( $N_p = 1$ ). Since we estimate the log-scaled parameter values (due to the varying scales of the parameters), we depict here the sensitivities with respect to those log-scaled parameters in Figures 5.1-5.2. The particular parameter values at which we solved the sensitivities are located in the figure captions, and are parameters which produce a model solution with roughly the same features (e.g., overall amplitude, wave frequency, damping envelope) as the experimental data.

In the pressure case, the model is most sensitive to  $E$ ,  $\gamma_1$ ,  $A$ , and  $\Upsilon$  and less sensitive to  $E_1$  and  $\tau_1$  (as can be seen in Figure 5.1). In the shear case, the model is most sensitive to  $G$ ,  $\zeta_1$  and  $A$ , and less sensitive to  $G_1$ ,  $\Upsilon$  and  $\tau_1$  (as can be seen in Figure 5.2). This lower sensitivity to  $\tau_1$  in both cases is fully consistent with the results in Chapter 3.

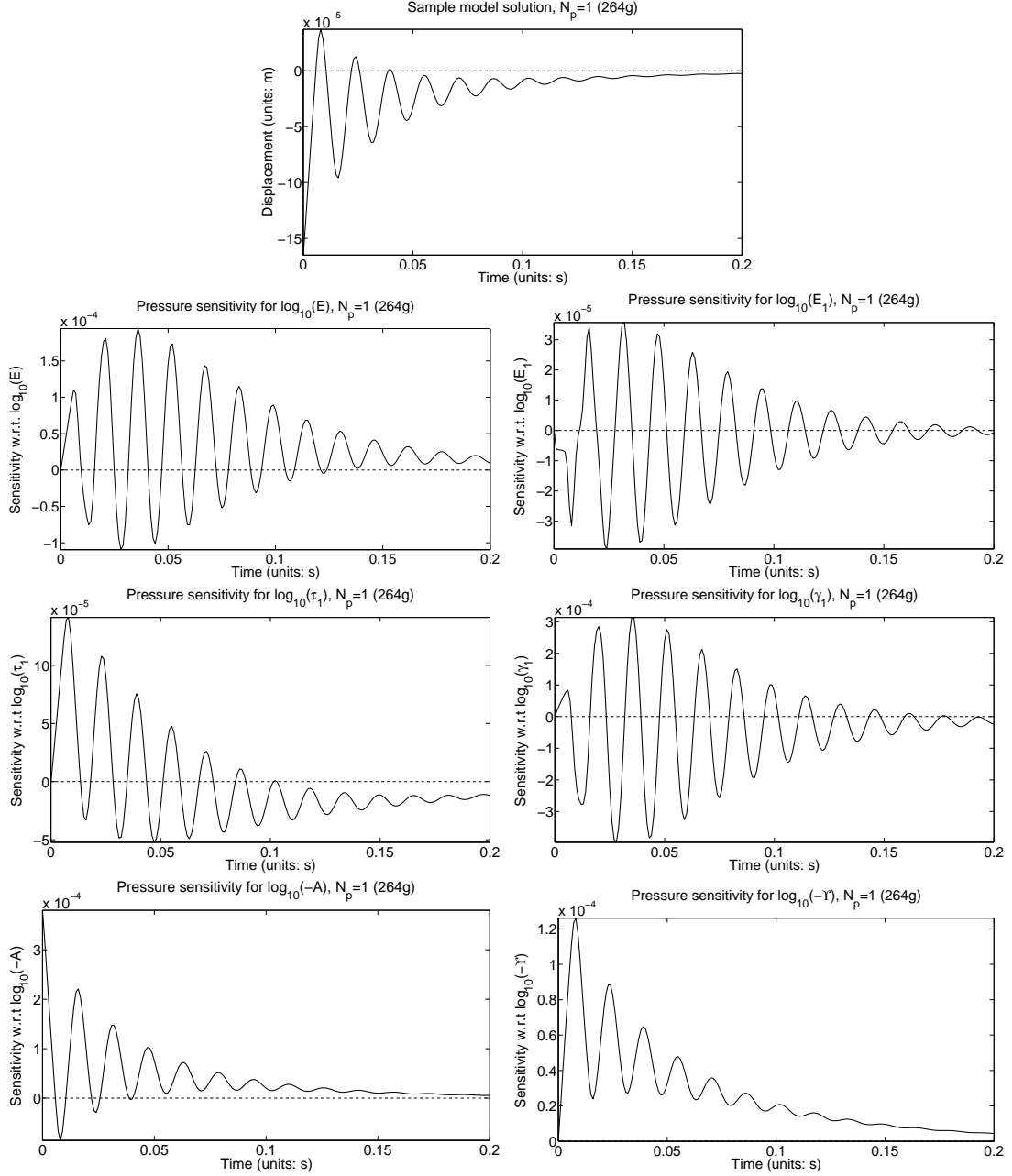


Figure 5.1: Pressure sensitivity equation solutions in the  $N_p = 1$  case. Solved using the following parameter values:  $E = 5 \times 10^4$  Pa,  $E_1 = 30$  Pa·s,  $\gamma_1 = 1.2 \times 10^5$  Pa,  $\tau = 0.02$  s,  $A = -1.65 \times 10^{-4}$  m, and  $\Upsilon = -0.01$  s. (top) Resulting pressure model solution. (bottom six) Sensitivity equation solutions for each parameter.

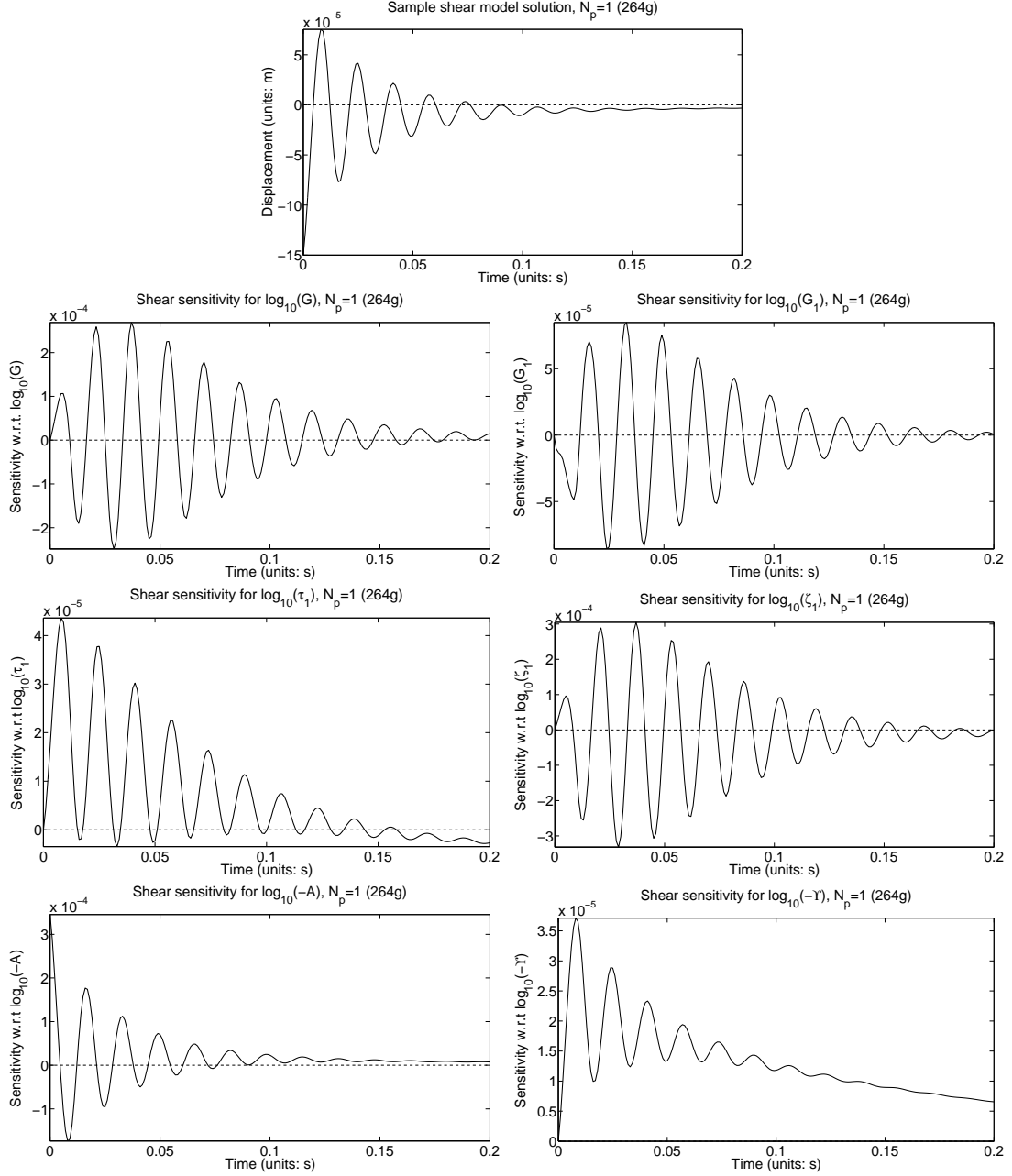


Figure 5.2: Shear sensitivity equation solutions in the  $N_p = 1$  case. Solved using the following parameter values:  $G = 6.5 \times 10^3$  Pa,  $G_1 = 5.6$  Pa·s,  $\zeta_1 = 8 \times 10^3$  Pa,  $\tau = 0.07$  s,  $A = -1.5 \times 10^{-4}$  m, and  $\Upsilon = -0.01$  s. (top) Resulting shear model solution. (bottom six) Sensitivity equation solutions for each parameter.

## 5.2 Statistical Models and Parameter Estimators

In order to carefully define the way in which we will measure the closeness of the data to model values, we must first discuss underlying statistical models for the error present in the data. A proper error model is also key to correctly determining parameter confidence intervals. Much of the discussion here is similar to that in Chapter 3, with background on ordinary least squares (OLS) and generalized/weighted least squares (GLS or WLS) given in [16, 24], for example.

We will assume the errors  $\mathcal{E}_j$  are independent, identically distributed with mean zero ( $E[\mathcal{E}_j] = 0$ ) and constant variance  $\text{var}(\mathcal{E}_j) = \sigma_0^2$ ; this process has realizations  $\varepsilon_j$ . Note that we do not assume we know the underlying distributions from which the errors come; we only know the first two central moments as specified. We use this error process in proposing two error models and corresponding parameter estimators.

- **Absolute error:** Here we have the error process  $U_j = u(t_j; 10^{q_0}) + \mathcal{E}_j$ , with realizations

$$u_j = u(t_j; 10^{q_0}) + \varepsilon_j, \quad (5.2.1)$$

where  $q_0$  is some hypothesized “true” parameter value (see [24]). We use the ordinary least squares cost function

$$\mathcal{J}_{ols}(q) = \sum_{j=0}^{n-1} [u_j - u(t_j; 10^q)]^2.$$

The corresponding inverse problem for the logged parameters is then

$$\hat{q}_{ols} = \arg \min_{q \in Q} \mathcal{J}_{ols}(q) = \arg \min_{q \in Q} \sum_{j=0}^{n-1} [u_j - u(t_j; 10^q)]^2. \quad (5.2.2)$$

This function minimizes the distance between the data and model where all observations are considered to have equal importance (weight). Since  $u(t_j; 10^q)$  is a continuous function of  $q$ ,  $\mathcal{J}_{ols}$  is also a continuous function of  $q$ , which means we are minimizing a continuous function of  $q$  over a compact set  $Q$ , and thus this inverse problem has a solution.

- **Relative error:** Here we have the error process  $U_j = u(t_j; 10^{q_0}) + u(t_j; 10^{q_0})\mathcal{E}_j$  with realizations

$$u_j = u(t_j; 10^{q_0}) + u(t_j; 10^{q_0})\varepsilon_j. \quad (5.2.3)$$

For this case, we construct the generalized (weighted) least squares cost function (as per, e.g., [24])

$$\mathcal{J}_{gls}(q) = \sum_{j=0}^{n-1} w_j^2 [u_j - u(t_j; 10^q)]^2$$

where we define the weights  $w_j = u(t_j; 10^q)^{-1}$ . In this case, since we are examining a relative error model (5.2.3), these weights take into account the unequal quality of observations; dividing by the function value has a “normalizing” effect on the errors, accounting for the scale differences which may be present in the errors at larger versus smaller model values.

We now wish to find  $q$  such that  $\mathcal{J}_{gls}(q)$  is minimized. We can either solve this directly, or by using an iterative procedure in order to estimate  $\hat{q}_{gls}$  (since the weights must also be estimated). We will use an iterative method, described as follows (see [24] and references therein for convergence details):

1. Define  $\hat{q}^0 = \hat{q}_{ols}$ , and set  $k = 0$ .
2. Form the weights  $\hat{w}_j = u(t_j; 10^{\hat{q}^k})^{-1}$ , using weight thresholding (described below).
3. Re-estimate  $\hat{q}_{gls}$  by solving

$$\hat{q}^{k+1} = \arg \min_{q \in Q} \sum_{j=0}^{n-1} \hat{w}_j^2 [u_j - u(t_j; 10^q)]^2$$

to obtain the  $(k + 1)$ th estimate  $\hat{q}^{k+1}$  for  $\hat{q}_{gls}$ .

4. Set  $k = k + 1$  and return to Step 2. Terminate when successive estimates for  $\hat{q}_{gls}$  are sufficiently close, or when one has iterated 20 times. For this problem, the “sufficiently close” criterion was determining if  $\|\hat{q}^{k+1} - \hat{q}^k\|_\infty \leq 10^{-3}$ , where  $\|q\|_\infty$  is the maximum component of the given vector  $q$ . The parameter values being estimated are all log-scaled, and are thus on the order of  $[10^{-1}, 10^1]$ . This puts the stopping criterion at two orders of magnitude less than the smallest log-scaled parameter value, which is sufficient in our problem.

Even though we have removed all data points with absolute value under  $5 \times 10^{-6}$ , we still account for the (now unlikely) possibility that some model values may still end up small in absolute value. Thus, we incorporate thresholding on the weights to keep from dividing by zero. We take a weight threshold value of  $1 \times 10^{-10}$ , as this is almost certainly below the threshold of significance in terms of the model displacements. Then, for all indices  $\bar{j} \in \{k \mid |u(t_k; 10^{\hat{q}})| < 1 \times 10^{-10}\}$ , we set  $\hat{w}_{\bar{j}} = 1 \times 10^{10}$ . This is done each time the weights are re-estimated in Step 2 of the iterative process.

With weight thresholding, we are assured that the iterative process is possible numerically. Thus, similar to the ordinary least squares case, at each step  $k$  in the iterative GLS estimation process we are minimizing a continuous function of  $q$  over a compact parameter space  $Q$ , and thus the inverse problem in each iteration will have a solution. Also, as long

as the iterative process is carried out sufficiently many times, under certain conditions the weights will converge  $\hat{w}_j \rightarrow u(t_j; 10^{\hat{q}_{gls}})^{-1}$  (see, e.g., [24]).

### Residual Definitions

We will also include residual plots to assist in analysis of the model fit to data, and to indicate which error model best describes the error in the data. Residuals give a sense for the model fit to data, but more importantly the residuals can give an indication [24] regarding the appropriateness of our error model. If the absolute residuals seem to be randomly dispersed around the horizontal axis and form a horizontal band around that axis, then the absolute error model may be correct. On the other hand, if the (modified) relative residuals seem to be randomly dispersed, then the relative error model may be correct. We define the following:

- **Absolute residuals** are computed as  $r_j = u_j - u(t_j; 10^{\hat{q}})$ , where  $\hat{q}$  is the particular parameter estimate being considered.
- **Relative residuals** are computed as  $r_j = \hat{w}_j(u_j - u(t_j; 10^{\hat{q}}))$  where  $\hat{w}_j = u(t_j; 10^{\hat{q}})^{-1}$  and the  $\hat{w}_j$  are thresholded in the same manner as discussed earlier.

### Asymptotic Error Analysis

In addition to determining the parameter estimates  $\hat{q}_{ols}$  and  $\hat{q}_{gls}$  for pressure and shear data, we will provide confidence interval information for each entry in the parameter vectors. For the absolute error model, the process is the same as that which we used in Chapter 3, and is also described in [24]; for the relative error model, the corresponding asymptotic error methodology is discussed in [24, Ch. 3]. Since the theory is common enough, we do not reiterate it here and refer interested readers to the aforementioned references.

#### 5.2.1 Optimization Considerations

We use the built-in Matlab Optimization Toolbox routine `lsqnonlin` for our optimization routine to solve for  $\hat{q}_{ols}$  and  $\hat{q}_{gls}$ . We used the trust-region-reflective (TRR) algorithm that is built in; as our previous efforts in Chapter 3 demonstrated, the Levenburg-Marquardt option was slower than TRR and did not give us better results. Since we are using at least one relaxation time, we do not consider `fmincon` which we have shown to be ineffective in estimating relaxation times.

In order to start the optimization routines for computing  $\hat{q}_{ols}$ , we must provide initial parameter values (for  $\hat{q}_{gls}$  we use the estimated value for  $\hat{q}_{ols}$  as our initial guess). From a perusal of the viscoelastic materials literature, our experience from the previous conceptual work, and

from some manual examination on the current data sets, we developed pressure and shear initial values.

**Pressure initial guess:** For  $N_p = 1$ , the initial values we use are  $E = 4.5 \times 10^4$  Pa,  $E_1 = 55$  Pa·s,  $\gamma_1 = 1.9 \times 10^5$  Pa,  $\tau_1 = 0.05$  s,  $A = -1.75 \times 10^{-4}$  m, and  $\Upsilon = -0.01$  s. As log-scaled values (c.f. (5.1.1)), this gives us

$$q_{ols}^0 = (4.6532, 1.7404, 5.2788, -1.3010, -3.7570, -2)^T.$$

**Shear initial guess:** For  $N_p = 1$ , the values we use are  $E = 4.5 \times 10^3$  Pa,  $E_1 = 5$  Pa·s,  $\gamma_1 = 2.8 \times 10^4$  Pa,  $\tau_1 = 0.06$  s,  $A = -1.7 \times 10^{-4}$  m, and  $\Upsilon = -0.01$  s. As log-scaled values (c.f. (5.1.1)), this gives us

$$q_{gls}^0 = (3.6532, 0.6990, 4.4472, -1.2218, -3.7696, -2)^T.$$

### 5.3 Results Using a Single Relaxation Time ( $N_p = 1$ )

We now demonstrate the ability of our model to match data. For this purpose, we will take a single relaxation time ( $N_p = 1$ ). We run both the absolute (OLS) and relative (GLS) error models on a sample data set using a 264 g loading weight, separately for both pressure and shear data. We will report parameter estimates, standard errors, plots of model fits to data, plots of residuals versus time, and plots of residuals versus model values. We use these elements in order to recommend error models for the pressure and shear cases.

We will also examine parameter estimation using data sampled at different rates. This will allow for a study of whether the parameter estimates stay consistent as the number of data points is reduced, and will also allow us to examine issues of independence between measurements (data points). It is expected that if the data points are sampled too frequently, nearby measurements are more likely to be dependent. This is due to the inherent limitations in hardware; too frequent sampling may not allow the measurement device to return to its resting state between measurements. Using fewer data points is a way of increasing the likelihood that neighboring measurements are independent. We run the inverse problem on each data set and using each error model with the following two options:

1. Using all the data points (1.024 kHz), and
2. Using every other data point (512 kHz).

Before discussing results, we recall the earlier discussion in Section 5.1.1 on the sensitivity of model output to the parameters. The pressure model output was most sensitive to  $E$ ,  $\gamma_1$ ,  $A$ , and  $\Upsilon$  and less sensitive to  $E_1$  and  $\tau_1$ . The shear model output was most sensitive to  $G$ ,  $\zeta_1$ ,

and  $A$  and less sensitive to  $G_1$ ,  $\tau_1$ , and  $\Upsilon$ . Thus, throughout the results, we are likely to see larger standard errors relative to the parameter magnitude for  $E_1$  ( $G_1$ ) and  $\tau_1$ . Intuitively, this is due in part to the fact that in the mathematical model the relaxation times influence how the damping properties of a material are described, and thus there is likely some interplay between the bulk damping parameter  $E_1$  ( $G_1$ ) and the relaxation time  $\tau_1$  in the pressure and shear cases. The parameter  $\Upsilon$  is a special case. As discussed in Section 4.4.6, if the relaxation times are small (which they will be in our results) then  $\Upsilon$  will not have much of an effect on the model output once it becomes sufficiently negative. Thus, even though the model output is at least somewhat sensitive to  $\Upsilon$ , particularly for the pressure model, we may still obtain large confidence intervals for this parameter once sufficiently negative. We will see these larger confidence intervals for  $\Upsilon$  in the results, and it should be noted that this is not a major concern since it is an artificial parameter designed to approximate the loading process.

### 5.3.1 Pressure Data Results

The results of the pressure parameter estimates and confidence intervals (see Chapter 3 or [11, 12, 17, 24] for information on computing confidence intervals) are shown in Tables 5.1-5.4, and the model fits as well as residuals are shown in Figures 5.3-5.6. In all cases, model fits to data are good. Comparing the OLS results in Tables 5.1 and 5.3 with their GLS counterparts in Tables 5.2 and 5.4, we see that the parameter estimates for OLS are generally more consistent between the full and reduced data sets than those for GLS. We also see that the standard errors for OLS are generally smaller than those for GLS for the results using all the data. Even though the OLS standard errors increase slightly when we cut the number of data points used in half, these comparisons give an initial indication that OLS is better than GLS. However, these results are only valid if the error model is correct. To that end, the residuals versus time plots in Figures 5.3-5.6 all have some patterns in the residuals but those do not substantially change when going between OLS and GLS. Additionally, these plots show less of a pattern in the residuals (i.e., more random) as the data sampling frequency is reduced. In the same figures, the residuals versus model plots are also not much different when comparing the OLS cases to the GLS cases. From a residual analysis standpoint, then, either model appears reasonable. Thus, since OLS is a simpler error model and since we have a higher degree of confidence in the parameter estimates due to their consistency for different data sampling rates and smaller standard errors, we recommend the OLS model when using pressure data. We do note that the standard error for  $\log_{10}(\tau_1)$  and  $\log_{10}(E_1)$  for both OLS and GLS cases is in general larger relative to the parameter estimate itself than for the other parameters; this is expected since the pressure model output is less sensitive to  $\tau_1$  and  $E_1$ . Thus, we do not have as much confidence in the estimate for  $\tau_1$  and  $E_1$  as in the other parameter estimates.



Since there was not a significant difference between OLS and GLS in terms of residual analysis, we examined another error model,

$$U_j = u(t_j; 10^{q_0}) + \text{sign}(u(t_j; 10^{q_0}))|u(t_j; 10^{q_0})|^{1/2}\mathcal{E}_j,$$

which we call the “half error” model (for lack of a better phrase) due to the  $1/2$  exponent in the final term. This is solved in a manner similar to the “full” GLS, the difference being that here we use the weights  $w_j = |u(t_j; 10^q)|^{-1/2}$ . Results of the corresponding inverse problem are shown in in Tables 5.5-5.6 and Figures 5.7-5.8. There does not appear to be any improvement over the previous OLS and GLS results in terms of residual randomness, and the parameter values are more consistent than those from GLS but less consistent than those from OLS. Additionally, the standard errors when using the half error model are generally larger than those when using OLS. Thus, we continue to conclude that the OLS model is most appropriate for the pressure data.

Table 5.1: Pressure optimization results and confidence analysis for OLS on a 264 g data set using every data point.

Param.	Estimate	SE	CI95
$\log_{10}(E)$	4.6164	0.5071	(3.6140, 5.6188)
$\log_{10}(E_1)$	1.7385	0.2180	(1.3076, 2.1694)
$\log_{10}(\tau_1)$	-1.3365	0.5089	(-2.3425, -0.3306)
$\log_{10}(\gamma_1)$	5.2748	0.1096	(5.0581, 5.4914)
$\log_{10}(-A)$	-3.7520	0.0061	(-3.7641, -3.7399)
$\log_{10}(-\Upsilon)$	-1.8549	0.6463	(-3.1326, -0.5773)

Young’s modulus dynamic analog  $E_0 = 229.604$  kPa

Table 5.2: Pressure optimization results and confidence analysis for GLS on a 264 g data set using every data point.

Param.	Estimate	SE	CI95
$\log_{10}(E)$	5.0523	2.1445	(0.8131, 9.2915)
$\log_{10}(E_1)$	1.8025	0.2192	(1.3692, 2.2358)
$\log_{10}(\tau_1)$	-0.7878	2.1350	(-5.0082, 3.4326)
$\log_{10}(\gamma_1)$	5.0664	2.0697	(0.9750, 9.1578)
$\log_{10}(-A)$	-3.8031	0.0124	(-3.8276, -3.7786)
$\log_{10}(-\Upsilon)$	-1.0152	4.9889	(-10.8772, 8.8468)

Young’s modulus dynamic analog  $E_0 = 229.323$  kPa

Table 5.3: Pressure optimization results and confidence analysis for OLS on a 264 g data set using every other data point.

Param.	Estimate	SE	CI95
$\log_{10}(E)$	4.6051	0.8396	(2.9302, 6.2800)
$\log_{10}(E_1)$	1.7426	0.3850	(0.9745, 2.5107)
$\log_{10}(\tau_1)$	-1.3661	0.8407	(-3.0433, 0.3111)
$\log_{10}(\gamma_1)$	5.2775	0.1755	(4.9274, 5.6277)
$\log_{10}(-A)$	-3.7442	0.0080	(-3.7601, -3.7284)
$\log_{10}(-\Upsilon)$	-1.8921	1.0604	(-4.0075, 0.2233)

Young's modulus dynamic analog  $E_0 = 229.749$  kPa

Table 5.4: Pressure optimization results and confidence analysis for GLS on a 264 g data set using every other data point.

Param.	Estimate	SE	CI95
$\log_{10}(E)$	4.1586	0.1827	(3.7942, 4.5231)
$\log_{10}(E_1)$	1.2236	0.4868	(0.2525, 2.1948)
$\log_{10}(\tau_1)$	-1.6888	0.1550	(-1.9981, -1.3794)
$\log_{10}(\gamma_1)$	5.3313	0.0134	(5.3045, 5.3580)
$\log_{10}(-A)$	-3.8057	0.0169	(-3.8393, -3.7721)
$\log_{10}(-\Upsilon)$	-2.2436	0.1726	(-2.5879, -1.8992)

Young's modulus dynamic analog  $E_0 = 228.827$  kPa

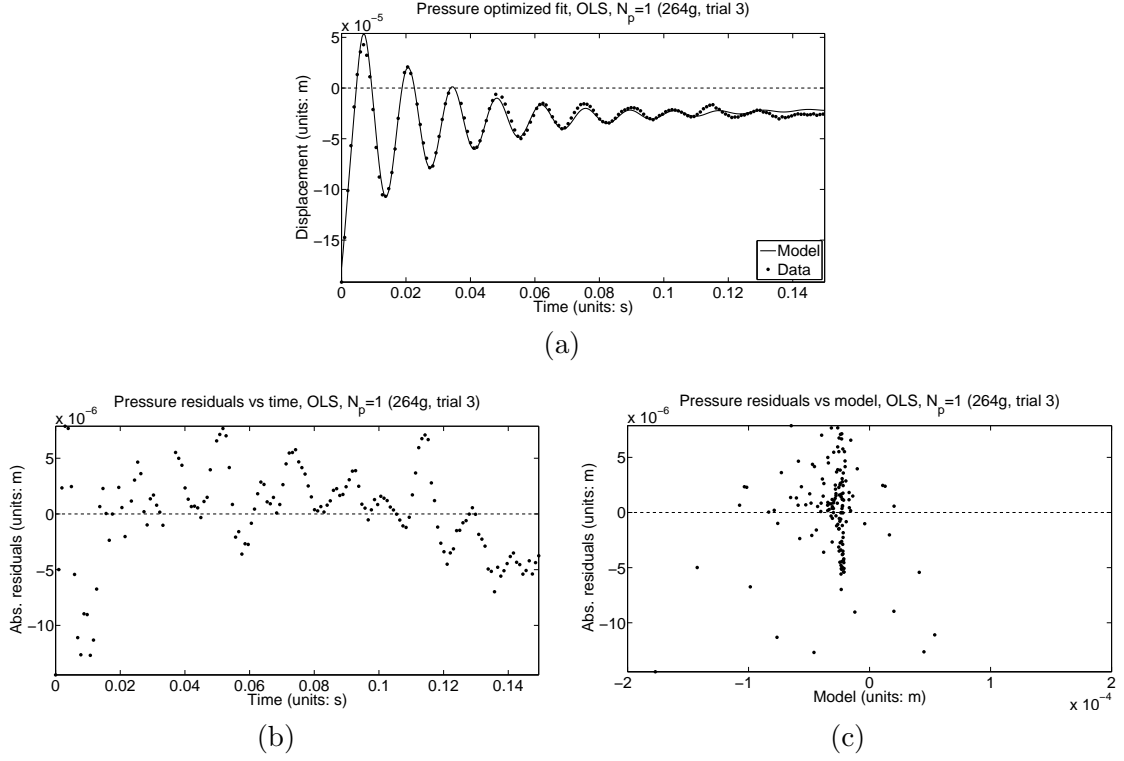


Figure 5.3: Pressure data fit using every data point with absolute error model,  $N_p = 1$ , weight 264 g. (a) Model fit to data. (b) Absolute residuals vs time. (c) Absolute residuals vs model.

Table 5.5: Pressure optimization results and confidence analysis for GLS (half error) on a 264 g data set, using every data point.

Param.	Estimate	SE	CI95
$\log_{10}(E)$	4.7747	0.9355	(2.9254, 6.6240)
$\log_{10}(E_1)$	1.7594	0.2289	(1.3069, 2.2120)
$\log_{10}(\tau_1)$	-1.1169	0.9304	(-2.9562, 0.7223)
$\log_{10}(\gamma_1)$	5.2302	0.3247	(4.5883, 5.8721)
$\log_{10}(-A)$	-3.7831	0.0091	(-3.8011, -3.7651)
$\log_{10}(-\Upsilon)$	-1.5638	1.3384	(-4.2095, 1.0818)

Young's modulus dynamic analog  $E_0 = 229.421$  kPa

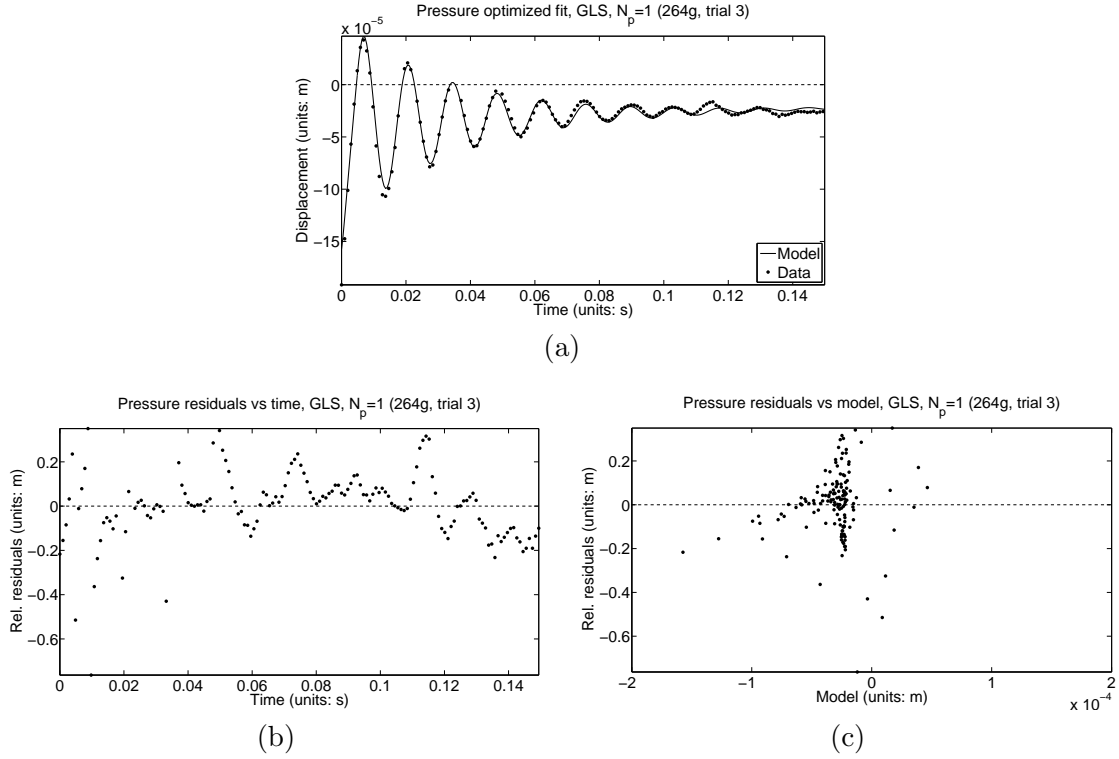


Figure 5.4: Pressure data fit using every data point with relative error model,  $N_p = 1$ , weight 264 g. (a) Model fit to data. (b) Relative residuals vs time. (c) Relative residuals vs model.

Table 5.6: Pressure optimization results and confidence analysis for GLS (half error) on a 264 g data set, using every other data point.

Param.	Estimate	SE	CI95
$\log_{10}(E)$	4.5630	0.9095	(2.7486, 6.3774)
$\log_{10}(E_1)$	1.6866	0.4513	(0.7863, 2.5869)
$\log_{10}(\tau_1)$	-1.3489	0.9004	(-3.1452, 0.4473)
$\log_{10}(\gamma_1)$	5.2848	0.1699	(4.9458, 5.6238)
$\log_{10}(-A)$	-3.7797	0.0128	(-3.8052, -3.7542)
$\log_{10}(-\Upsilon)$	-1.8602	1.1158	(-4.0862, 0.3657)

Young's modulus dynamic analog  $E_0 = 229.224$  kPa

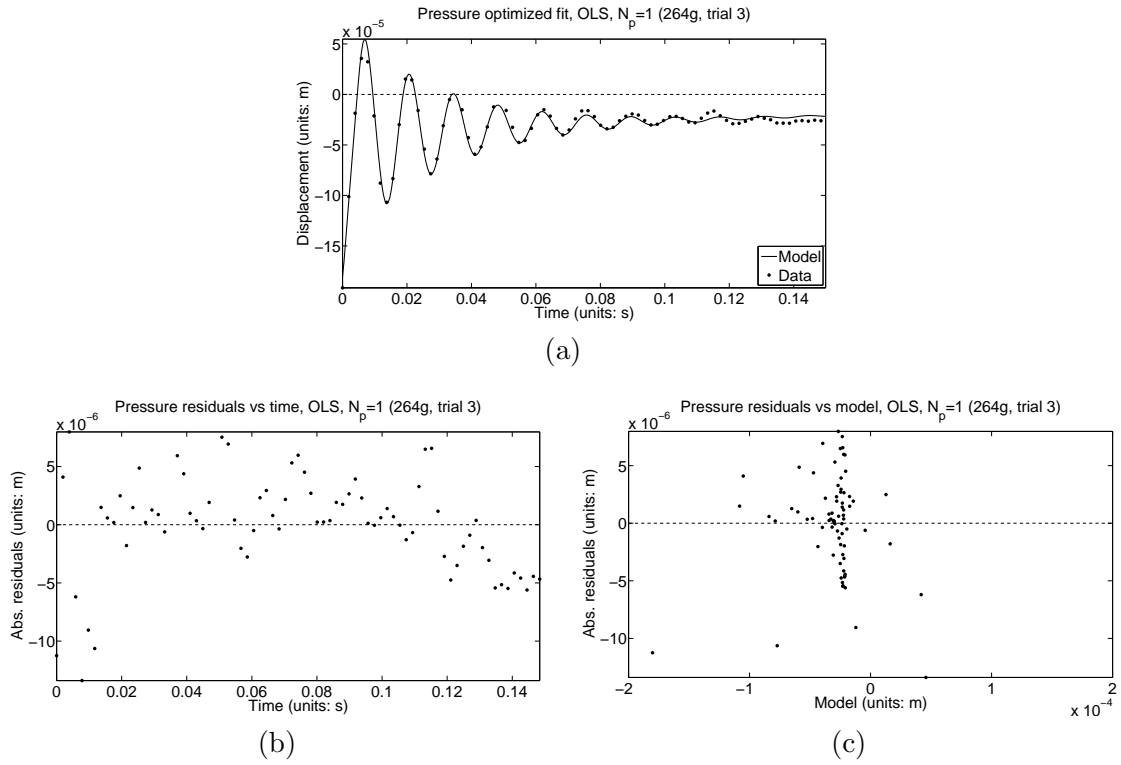


Figure 5.5: Pressure data fit using every other data point with absolute error model,  $N_p = 1$ , weight 264 g. (a) Model fit to data. (b) Absolute residuals vs time. (c) Absolute residuals vs model.

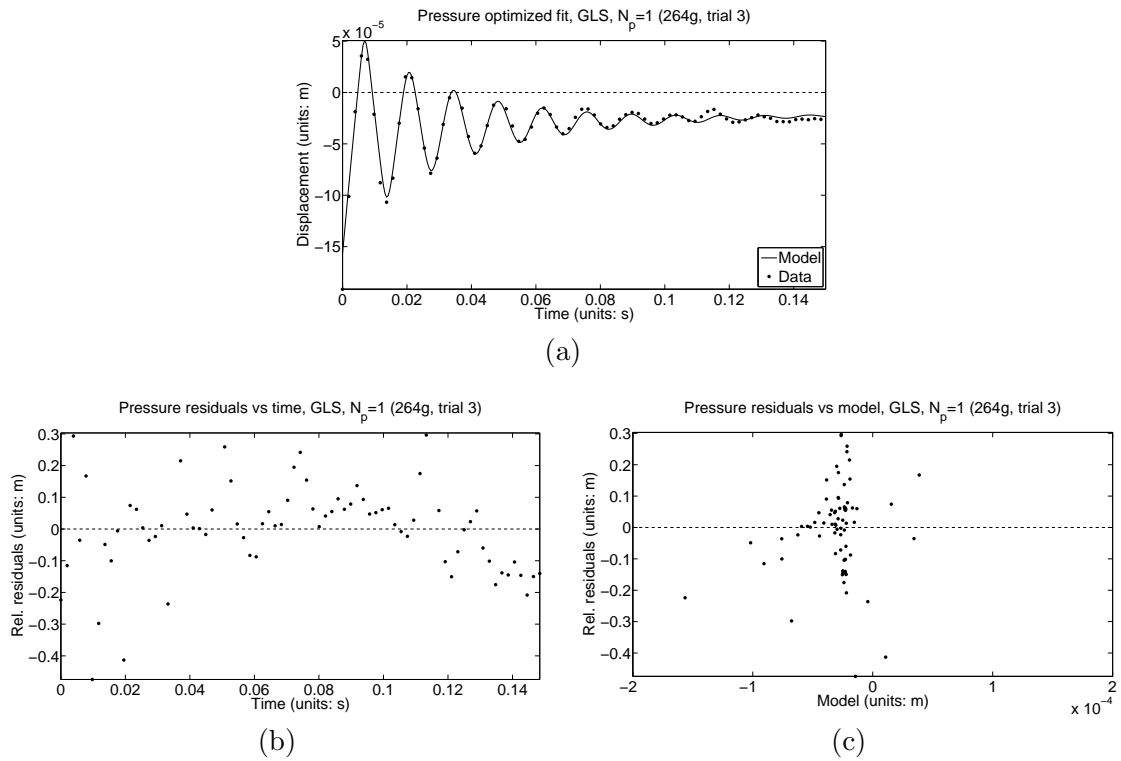


Figure 5.6: Pressure data fit using every other data point with relative error model,  $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Relative residuals vs time. (bottom right) Relative residuals vs model.

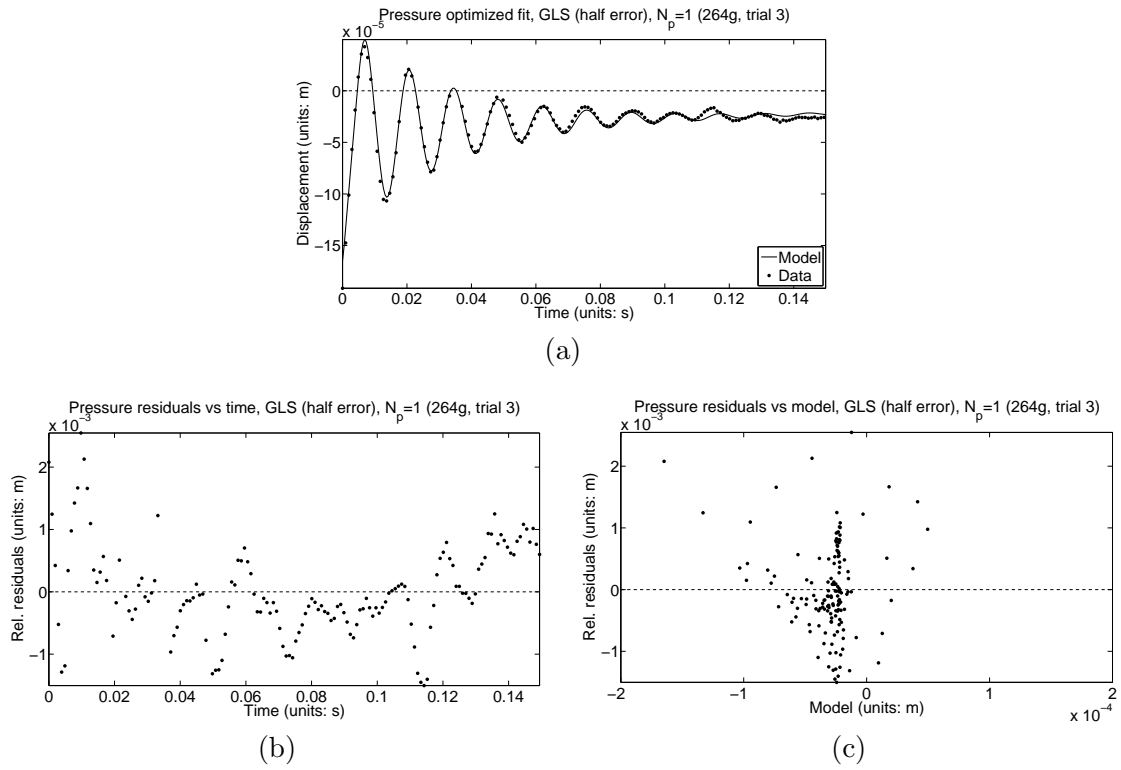
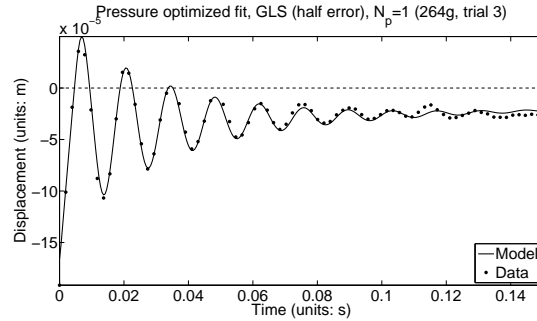
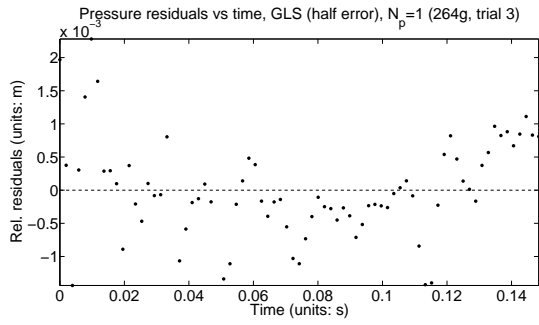


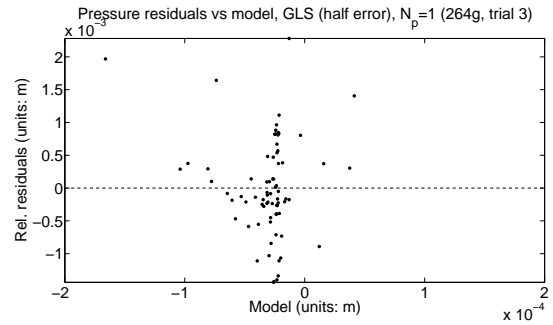
Figure 5.7: Pressure data fit using every data point with relative half error model,  $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Relative residuals vs time. (bottom right) Relative residuals vs model.



(a)



(b)



(c)

Figure 5.8: Pressure data fit using every other data point with relative half error model,  $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Relative residuals vs time. (bottom right) Relative residuals vs model.



### 5.3.2 Shear Data Results

Shear results are shown in Tables 5.7-5.10 and Figures 5.9-5.12. In all cases, the model fits to data are good. Comparing the OLS results in Tables 5.7 and 5.9 with their GLS counterparts in Tables 5.8 and 5.10 we see that the parameter estimates for OLS demonstrate similar consistency when using the full and reduced data sets as the GLS estimates. The standard errors for the GLS cases are larger than the OLS cases. In Figures 5.9-5.12, the residuals versus model plots are again not noticeably different for the OLS and GLS cases. The initial indication is that we have more confidence in the OLS results. However, the time versus residual plots raise cause for concern. In the OLS residual versus time plots, there is a noticeable “fan” structure for early times. However, for the GLS error model, the residual versus time plots do not show a fan structure and are fairly randomly distributed. Since this indicates that the OLS error model may not be correct, we are inclined to recommend the GLS error model in the shear case so that we do not mistakenly overstate our confidence in the parameter estimates, which we could do if we used the parameter estimates from the possibly-wrong OLS case. For the shear case we see that the standard error for  $G_1$  is on the same order of magnitude as the parameter estimate itself for both the OLS and GLS results, and the standard error for  $\tau_1$  in the GLS case is also on the same order of magnitude as the estimate for  $\tau_1$ . This is consistent with the sensitivity results, where in the shear case the model is less sensitive to  $G_1$  and  $\tau_1$  than to  $G$ ,  $\zeta_1$ , and  $A$ .

Table 5.7: Shear optimization results and confidence analysis for OLS on a 264 g data set using every data point.

Param.	Estimate	SE	CI95
$\log_{10}(G)$	3.6362	0.2465	(3.1499, 4.1225)
$\log_{10}(G_1)$	0.4725	0.2025	(0.0730, 0.8719)
$\log_{10}(\tau_1)$	-1.3433	0.2455	(-1.8276, -0.8589)
$\log_{10}(\zeta_1)$	4.4637	0.0366	(4.3915, 4.5358)
$\log_{10}(-A)$	-3.7543	0.0054	(-3.7649, -3.7436)
$\log_{10}(-\Upsilon)$	-2.0632	0.2862	(-2.6278, -1.4985)

Shear modulus dynamic analog  $G_0 = 33.411$  kPa

Table 5.8: Shear optimization results and confidence analysis for GLS on a 264 g data set using every data point.

Param.	Estimate	SE	CI95
$\log_{10}(G)$	4.1059	1.0993	(1.9371, 6.2747)
$\log_{10}(G_1)$	0.5925	0.2149	(0.1686, 1.0165)
$\log_{10}(\tau_1)$	-0.8318	1.1018	(-3.0054, 1.3418)
$\log_{10}(\zeta_1)$	4.3150	0.6781	(2.9774, 5.6527)
$\log_{10}(-A)$	-3.7990	0.0093	(-3.8173, -3.7807)
$\log_{10}(-\Upsilon)$	-1.3449	1.8982	(-5.0898, 2.4000)

Shear modulus dynamic analog  $G_0 = 33.416$  kPa

Table 5.9: Shear optimization results and confidence analysis for OLS on a 264 g data set using every other data point.

Param.	Estimate	SE	CI95
$\log_{10}(G)$	3.5431	0.2498	(3.0469, 4.0393)
$\log_{10}(G_1)$	0.3753	0.3227	(-0.2657, 1.0164)
$\log_{10}(\tau_1)$	-1.4450	0.2474	(-1.9364, -0.9535)
$\log_{10}(\zeta_1)$	4.4761	0.0294	(4.4178, 4.5345)
$\log_{10}(-A)$	-3.7501	0.0071	(-3.7643, -3.7360)
$\log_{10}(-\Upsilon)$	-2.1813	0.2779	(-2.7334, -1.6293)

Shear modulus dynamic analog  $G_0 = 33.423$  kPa

Table 5.10: Shear optimization results and confidence analysis for GLS on a 264 g data set using every other data point.

Param.	Estimate	SE	CI95
$\log_{10}(G)$	3.8649	1.0435	(1.7922, 5.9377)
$\log_{10}(G_1)$	0.5449	0.3561	(-0.1625, 1.2523)
$\log_{10}(\tau_1)$	-1.0217	1.0543	(-3.1161, 1.0726)
$\log_{10}(\zeta_1)$	4.4171	0.2918	(3.8375, 4.9967)
$\log_{10}(-A)$	-3.8026	0.0119	(-3.8263, -3.7789)
$\log_{10}(-\Upsilon)$	-1.6729	1.3806	(-4.4153, 1.0695)

Shear modulus dynamic analog  $G_0 = 33.454$  kPa

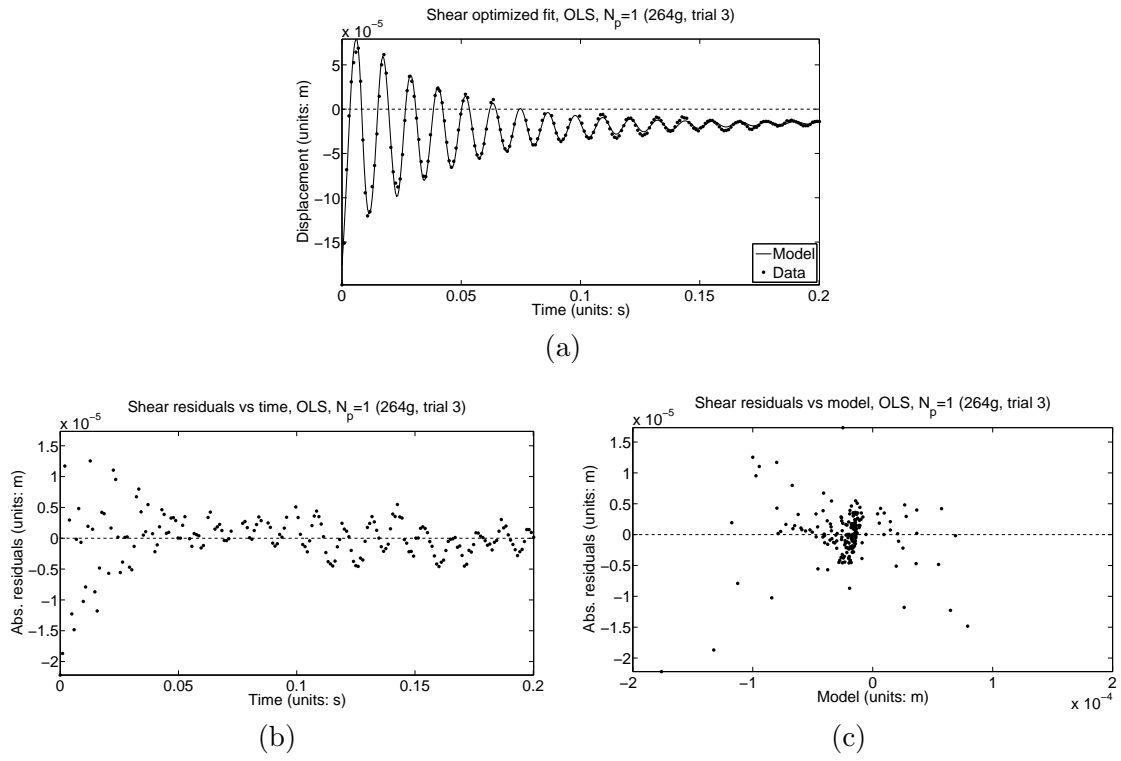


Figure 5.9: Shear data fit using every data point with absolute error model,  $N_p = 1$ , weight 264 g. (a) Model fit to data. (b) Absolute residuals vs time. (c) Absolute residuals vs model.

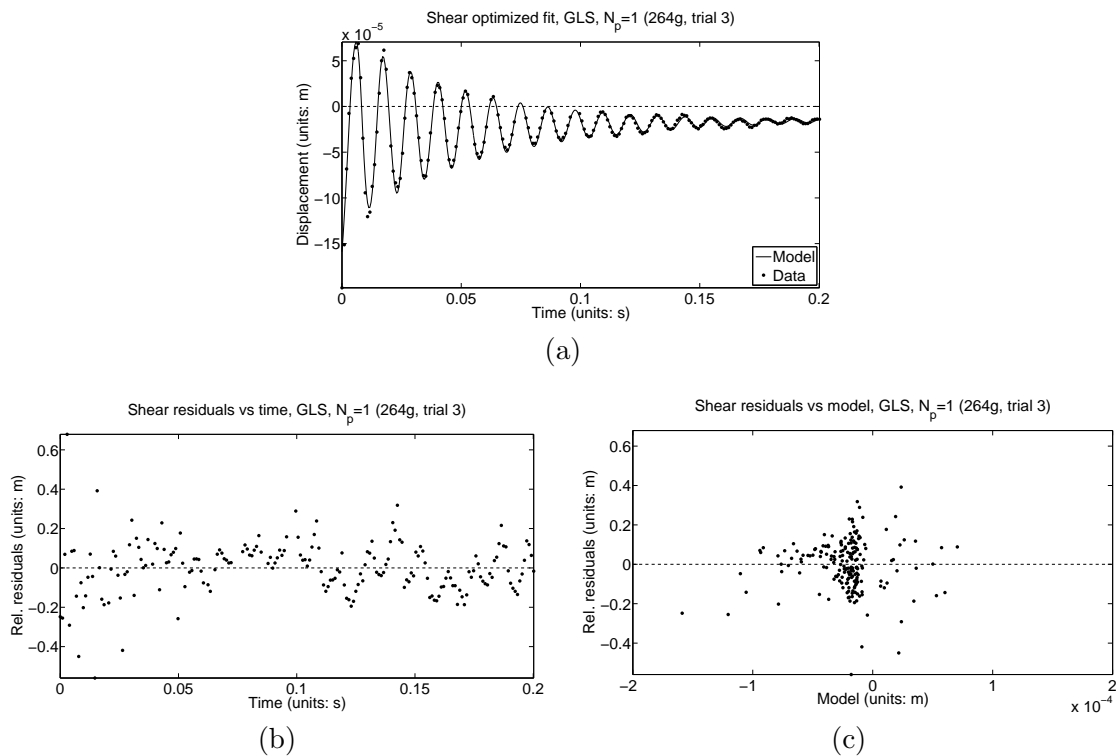


Figure 5.10: Shear data fit using every data point with relative error model,  $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Relative residuals vs time. (bottom right) Relative residuals vs model.

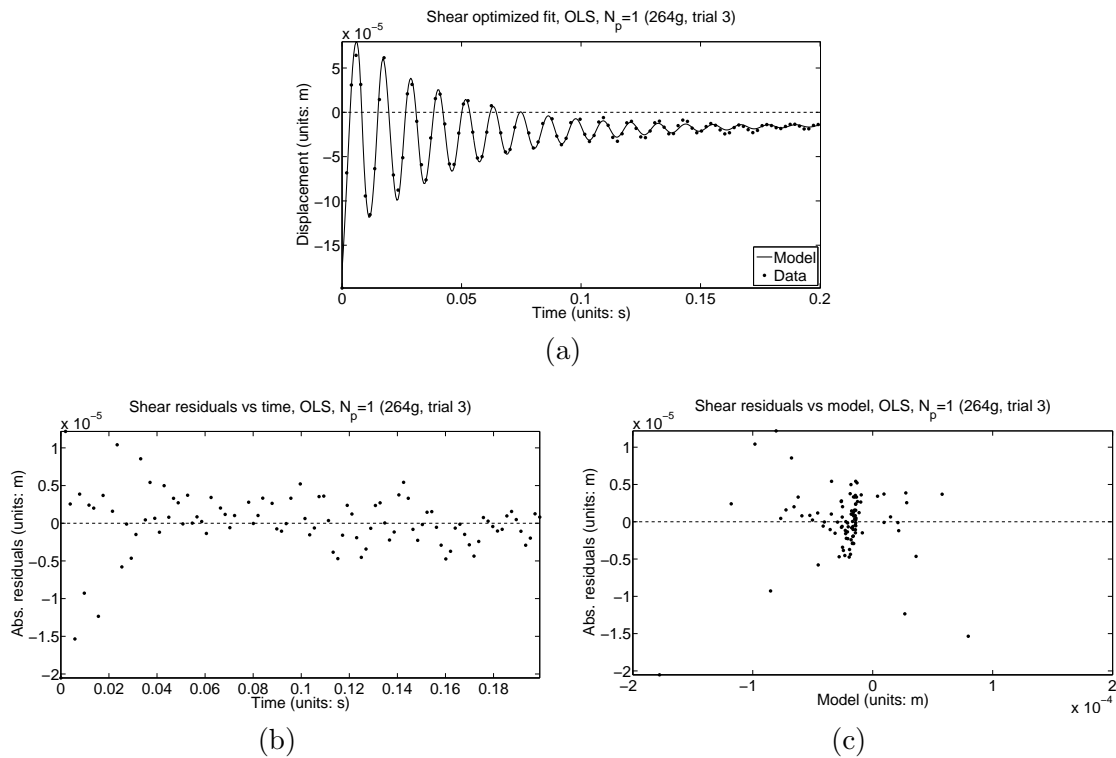


Figure 5.11: Shear data fit using every other data point with absolute error model,  $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Absolute residuals vs time. (bottom right) Absolute residuals vs model.

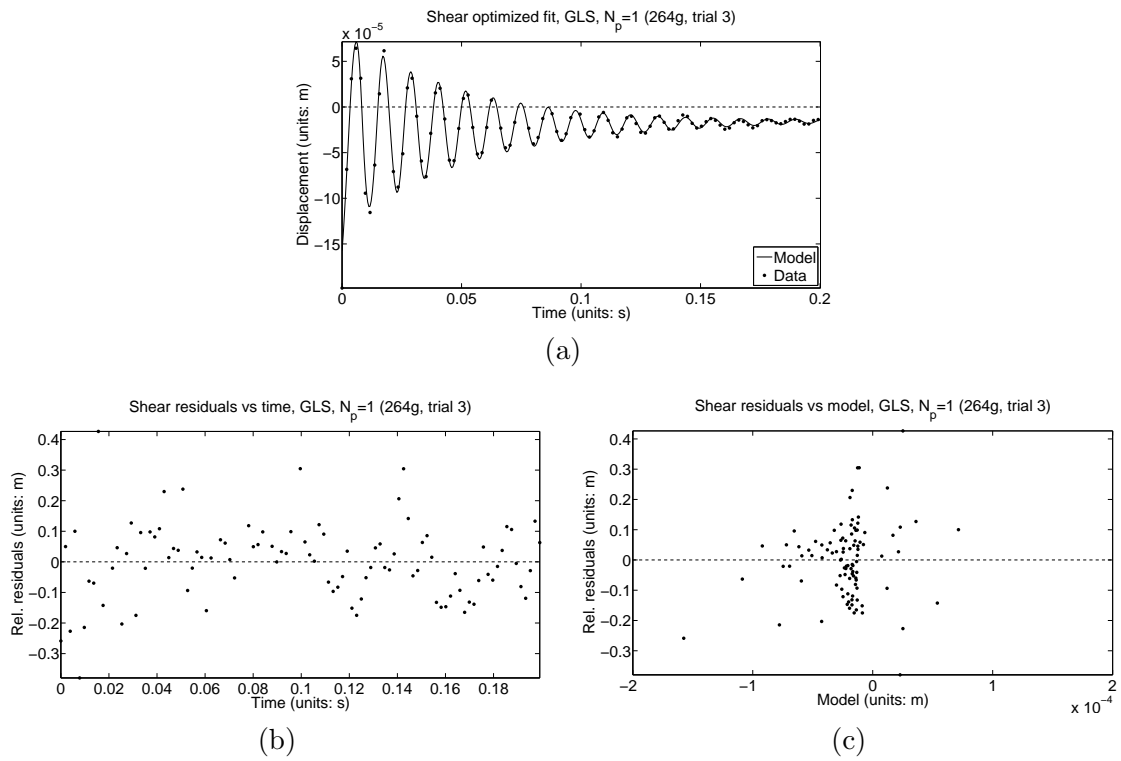


Figure 5.12: Shear data fit using every other data point with relative error model,  $N_p = 1$ , weight 264 g. (top) Model fit to data. (bottom left) Relative residuals vs time. (bottom right) Relative residuals vs model.

## 5.4 Model Development Summary

We have developed an updated one-dimensional viscoelastic model for tissue and have used experimental data from a simple homogeneous gel phantom to test the ability of our model to describe wave propagation in the medium. The data were generated from a drop experiment designed to mimic the disturbance into the chest cavity produced by blood flow in a stenosed coronary artery impacting the vessel wall, a disturbance which results in pressure and shear waves propagating away from the vessel walls downstream of the blockage. In our inverse problem results as discussed in Section 5.3, we have shown an ability to consistently model the wave propagation using different error models and at different data frequencies, obtaining good fits to data in all of our inverse problems. In addition to a good fit, though, we also examined statistical properties of the parameter estimators as well as residual plots to gain more insight into the proper error model for the pressure and shear data sets. This is necessary, since a correct error model is essential in order to apply the asymptotic error theory properly and thus obtain correct confidence intervals. For the pressure case, we prefer the absolute error model (OLS) over the relative error model (GLS) since the residual plots indicated no noticeable difference between the two models while the OLS parameter estimates were more consistent and had generally smaller corresponding standard errors. For the shear case, we recommend taking the more conservative route and using the GLS parameter estimates; even though the GLS estimates had larger standard errors, there were indications from the residual versus time plots for OLS that the OLS model is not correct. Overall, we have successfully demonstrated the ability of the pressure and shear mathematical models to accurately describe the data from laboratory experiments.

---

## Further Uncertainty Quantification Investigation

---

The need to quantify the accuracy of model parameter values determined during model calibration has become increasingly important. In practice, this is done in the frequentist approach by quantifying the sampling distribution (which is an uncertainty statement about the estimator for an assumed “true” parameter vector) or in the Bayesian approach by estimating the parameter posterior density which provides uncertainty information (which is a statement about the estimation of an assumed underlying parameter density). The methods used to understand parameter uncertainty make fundamentally different assumptions, which we will later examine in more detail. Differences between model prediction and measurements result from the fact that not only are models not perfectly descriptive of underlying phenomena (due to modeling error), but data (e.g., from lab experiments) inherently has measurement error. Thus, when working with mathematical models and the model parameters used in attempting to describe real-life phenomenon, we need to consider the methods by which we describe the uncertainty in the parameter estimates which “best” match the data. To this end, we will examine different inverse uncertainty quantification methods in the context of the viscoelastic wave propagation model developed in Chapter 5. This provides us with a nonlinear, partial differential equation model of a complex phenomenon whereby we can examine the performance of the different methods on an active research problem.

In this chapter, we take the uncertainty analysis another step beyond the asymptotic analysis of Chapter 4, comparing three different methods. The previously used asymptotic error theory is a frequentist method, meaning we have assumed that a single true parameter value exists and are trying to estimate that value and then study uncertainty in the estimator. Another frequentist



method (used on the proof-of-concept model in Chapter 3) is bootstrapping, which tests the robustness of the parameter values by taking the residuals between the model predictions and data points, randomly mixing the residuals across the time points to create new “simulated” data sets, then solving the inverse problem on the simulated data sets in order to obtain new parameter estimates. These are compiled, and statistics such as mean and variance are computed which then describe the uncertainty in the parameter estimates. In both frequentist cases, the uncertainty is related to the so-called “sampling distribution,” which is a measure of how well the estimation method performs in quantifying the parameter values for which the corresponding model solution fits the data under consideration.

In a fundamentally different approach, Bayesian methods regard the underlying parameters as random variables with associated densities and attempt to construct these parameter densities directly. A random, Markovian walk is built which steps through the parameter space (often initialized with a least squares or maximum likelihood estimate) which accepts parameter values based on their closeness to the data. This approach will be discussed in this chapter; recall the frequentist methods were described in Chapter 3.

Though one may be concerned with issues like propagation of parameter uncertainty through the model solution and subsequent model output predictive intervals, we are concerned primarily here with the uncertainty in the parameter values themselves. Thus, when comparing algorithms we will focus on the following considerations:

1. Complexity of the algorithm;
2. Computational time considerations (including parallelization);
3. Insight into correlation between parameters; and
4. Ability to provide a density that can be subsequently propagated through models.

The asymptotic, bootstrapping, and Bayesian methods will be compared and contrasted in this chapter. The viscoelastic model and data provide an example of where these methods can be successful and also reveal some of the various drawbacks to each of the methods. Since these statistical methods are developed from very different theory and under different basic assumptions about the parameter being estimated, they are hard to compare in the abstract; hence, results from the wave propagation modeling problem will be used to provide an example problem in active research for which we demonstrate the performance of the methods. Note that a version of this chapter with slightly different but corresponding results is being prepared for publication [60].

## 6.1 Mathematical and Statistical Models

The example problem we examine in this chapter is the inverse problem of estimating  $\bar{q} = (E, E_1, \tau_1, \gamma_1, A, \Upsilon)$  in the pressure wave propagation model (4.4.13). This is due to the fact that the pressure model forward solve is many times faster than the shear model solver. Since bootstrapping and Bayesian estimation are quite computationally intensive, the faster pressure model was the more reasonable choice. These results may hold similarly for shear, though that would need to be studied in the future. In all the discussion which follows, we will still be estimating the log-scaled versions of the parameters, since the parameter values are on quite varying scales. This means that we will use as our sought-after parameters

$$q = (\log_{10}(E), \log_{10}(E_1), \log_{10}(\tau_1), \log_{10}(\gamma_1), \log_{10}(-A), \log_{10}(-\Upsilon)).$$

The notation  $\bar{q} = 10^q$  signifies the fact that the parameter values enter the model in their original scale and with proper sign so that  $A, \Upsilon < 0$ . We assume  $q \in Q$ , where  $Q$  is some compact set of admissible parameters. We enforce the bounds  $-15 \leq q_i \leq 15$  for  $i = 1, \dots, 5$  and  $-15 \leq q_6 \leq \log_{10}(20)$ . As discussed earlier, the upper bound on  $q_6 = \log_{10}(-\Upsilon)$  is due to modeling considerations.

The specific inverse problem method we again use is ordinary least squares (OLS). Based on an analysis of residual plots, the lab measurements are considered the most significant source of error. We assume these are additive, independent, and identically distributed errors, which leads to realizations

$$u_j = u(L, t_j; 10^{q_0}) + \varepsilon_j.$$

With this error model, we use the ordinary least squares cost function

$$\mathcal{J}_{ols}(q) = \sum_{j=0}^{n-1} [u_j - u(L, t_j; 10^q)]^2,$$

and the usual inverse problem

$$\hat{q}_{ols} = \arg \min_{q \in Q} \mathcal{J}_{ols}(q) = \arg \min_{q \in Q} \sum_{j=0}^{n-1} [u_j - u(L, t_j; 10^q)]^2.$$

In the frequentist case, we assume that the errors are iid, have mean zero (i.e.,  $E[\mathcal{E}_j] = 0$ ), and constant variance  $\text{var}(\mathcal{E}_j) = \sigma_0^2$ , but do not need to assume a particular distributional form. In the Bayesian case, we will make the stronger assumption of normally distributed errors, each with the same mean and variance so that they are also identically distributed.

Here we use additive error because it has been shown to be a reasonable assumption for the

problem (as shown in Chapter 5) and also because it simplifies the resulting algorithms so we can focus more on comparisons across methods rather than being concerned with notation and more complex estimation algorithms. We will use a data measurement frequency of 1024 Hz in the present chapter, which should be sufficiently frequent to provide useful information for the inverse problem but sufficiently infrequent that the data points can be assumed independent.

## 6.2 Methods for Studying Confidence in Parameters

In Chapter 3 we studied and compared the results of the frequentist methods. We found that the asymptotic theory was comparable to the bootstrap method, and thus the asymptotic theory was preferable since it requires significantly less computational time. In addition to using the updated model (4.4.13), we now consider Bayesian methods, since this approach directly provides densities for the parameters rather than for the sampling distribution. This distinction is particularly important if one is concerned with propagating parameter uncertainty through the model in order to provide solution confidence and/or predictive intervals; in that case, the parameter densities obtained from Bayesian methods must be used since propagating uncertainty requires direct knowledge of the parameters rather than just knowledge of the sampling distribution. We will give descriptions of the algorithms for each method below (with some additional theoretical background for the Bayesian methods), referring to specific references which provide more detailed information on the development of each method.

Note that in all cases, we assume that we have already computed the OLS estimate  $\hat{q}_0 = \hat{q}_{ols}$  by solving (5.2.2). We have computed the estimate using the Matlab Optimization Toolbox routine `lsqnonlin`. This initial OLS estimate is common to all three methods, and as such this initial step will *not* be considered in later reporting on computational times. We comment that a common, pre-computed initial OLS estimate is not necessary for all Bayesian methods one might consider. In particular, parallel methods such as DREAM [90, 96, 97] are designed to be global optimizers as well as quantify the parameter densities. A different comparison of computational times would be needed when using these methods.

### 6.2.1 Frequentist Methods

We will again compare asymptotic analysis and bootstrapping, two frequentist methods. For the asymptotic algorithm, refer to Section 3.4; for the bootstrapping algorithm, refer to Section 3.5. As the name implies, the theory is asymptotic meaning it is only guaranteed to hold for large values of  $n$ . This may be quite suspect in many real-life applications since  $n$  can be somewhat small depending on data availability. Also, in the development of the theory, linearizations are made which then makes asymptotic theory invalid in regions of high nonlinearity. Finally,

asymptotic theory constructs confidence intervals as combinations of Gaussians, which is a possible limitation.

Bootstrapping requires solving the inverse problem  $M$  times. This is significantly more computationally expensive than solving the asymptotic sensitivity equations. Depending on how easily the inverse problem is solved, we will not know *a priori* whether bootstrapping or the to-be-discussed Bayesian methods will end up computing more forward solves. Regardless, asymptotic theory will be significantly faster to implement. Methods also exist [45, 46] whereby one can obtain confidence intervals directly from the bootstrap distribution *without* needing to assume that the form of the sampling distribution is normal; however, here we use the form in (3.5.2) since it is simple and effective for our case.

### 6.2.2 Bayesian Parameter Estimation and Confidence Analysis

In a Bayesian framework, one assumes that the parameters are random variables with associated densities. Initially, the parameters are described by a prior density  $\pi_0(q)$ ; we will use a “noninformative” prior (of course, we still implement bounds on parameter values). We then use the data realizations  $\vec{u}$  to compute a posterior density  $\pi(q|\vec{u})$ . The data values are incorporated through a likelihood function  $\pi(\vec{u}|q)$ . We will assume the measurement errors are normally distributed, so that the likelihood function becomes the multivariate normal density. Note that this assumes a specific form for the measurement error process density; the frequentist methods only required we specify the first two moments, instead of the entire error distribution. To solve the inverse problem in this framework, we use “Bayes’ theorem for inverse problems” which is given by the following statement:

**Definition 6.2.1** (Bayes’ theorem for inverse problems, referred to as such by [58]). We assume that the parameter vector  $q$  is a random variable which has a known prior density  $\pi_0(q)$  (possibly noninformative), and corresponding realizations  $\vec{u}$  of the random variable  $\vec{U}$  associated with the measurement process. The posterior density of  $q$ , given measurements  $\vec{u}$ , is then

$$\pi(q|\vec{u}) = \frac{\pi(\vec{u}|q)\pi_0(q)}{\pi(\vec{u})} = \frac{\pi(\vec{u}|q)\pi_0(q)}{\int_{\mathbb{R}^p} \pi(\vec{u}|q)\pi_0(q)dq} \quad (6.2.1)$$

where we have assumed that the marginal density  $\pi(\vec{u}) = \int_{\mathbb{R}^p} \pi(q, \vec{u})dq = \int_{\mathbb{R}^p} \pi(\vec{u}|q)\pi_0(q)dq \neq 0$  (a normalizing factor) is the integral over all possible joint densities  $\pi(q, \vec{u})$ . Note here that  $\pi(\vec{u}|q)$  is a likelihood function which describes how likely a data set  $\vec{u}$  is when given a model solution at the parameter value  $q$ .

We could solve the inverse problem directly, using cubature or Monte Carlo techniques to compute the integral in the definition. This is computationally viable only for a small number of parameters – as the number of parameters becomes larger, the space over which the integral

must be evaluated becomes so large as to be computationally prohibitive. The alternate method, which we discuss here, is to build a Markov chain whose stationary distribution is the posterior density  $\pi(q|\vec{u})$  of (6.2.1). This is a method in which we sample the parameter space, accepting the parameters based on closeness of the model solution at the parameters to the data. It is known (see, e.g., [87]) that a Markov chain defined by the random walk Metropolis algorithm will converge if the chain is run sufficiently long (i.e., if we allow the algorithm to sample the parameter space a large number of times). A chain is a set of parameter values, begun with an initial guess  $q_0$ , that results from the random walk. There are no analytic results to confirm one has run the chain sufficiently long for convergence, so one must run chains long enough to ensure that they are sampling the posterior density (this is called *mixing*). Early chain values are considered a *burn-in* phase, which are not included in the parameter density results. The burn-in and total chain values are highly problem-dependent; we will run  $M = 50,000$  chain values with varying lengths of burn-in depending on how quickly the chain seems to visually settle into a random walk. For each individual problem, one must balance between long enough chains to hopefully ensure that the posterior density is adequately sampled while also keeping the runtime reasonable. More discussion on this issue can be found throughout the literature, for example in [7, 86, 87].

We briefly discuss the meaning of the prior term,  $\pi_0(q)$ . If any information is known about the parameters beforehand, perhaps if they are known to follow a particular distribution, this information can be incorporated via the prior and result in faster convergence. However, it is known (and demonstrated in, for example, the forthcoming manuscript [86]) that a poor choice for the prior can result in much slower convergence of the posterior. Besides placing bounds on the parameter values, we know nothing else about the parameter densities, so we take a noninformative prior  $\pi_0(q) = H(0)$  where  $H(0)$  is the Heaviside step function with step at 0. For discussion purposes, though, we will leave  $\pi_0$  in the formulations below.

We now turn to a discussion of the Metropolis algorithm which is used to create the Markov chain of parameters sampled from the posterior distribution. We will describe and implement here a delayed rejection adaptive Metropolis (DRAM) algorithm; for more general details on the background and algorithm see [7, 57, 56, 87]. We must begin the chain at some initial parameter value with positive likelihood; this will be  $q_0 = q_{OLS}$ , the solution to the ordinary least squares pressure inverse problem (5.2.2). To step through the parameter space, we require a *proposal distribution*  $J(q^*|q^{k-1})$  which provides a new sample parameter  $q^*$  that depends only on the previously sampled parameter  $q^k$ . We also must form a probability  $\alpha(q^*|q^{k-1})$ , dependent on the prior density  $\pi_0(q)$  and likelihood function  $\pi(\vec{u}|q)$ , with which we accept or reject the new parameter value.

We will use a normally distributed proposal function  $J(q^*|q^{k-1}) = \mathcal{N}(q^{k-1}, V_0)$ , where  $V_0$  is an estimate for the covariance matrix at  $q_0$ . In this case we use the standard asymptotic

theory estimate for the covariance matrix, namely  $V_0 = \sigma_{ols}^2 [\chi^T(q_{ols})\chi(q_{ols})]^{-1}$ , with  $\sigma_{ols}^2 = \frac{1}{n-\kappa} \sum_{j=0}^{n-1} [u_j - u(L, t; 10^{q_{ols}})]^2$  and  $\chi_{jk}(q) = \frac{\partial u(L, t; 10^q)}{\partial q_k}$  where  $n$  is the number of data points and  $\kappa = 6$  is the number of parameters being estimated. This choice for  $V_0$  hopefully ensures that the shape of the proposal function distribution matches, to some extent, the shape of the posterior distribution. Adaptive methods change the proposal covariance  $V_k$  (where  $V_0$  is the initial covariance and later matrices are denoted  $V_k$ ) in a prescribed manner depending on the previous states. Since a general adaptation breaks the Markov property of the chain (the adaptive  $V_k$  will depend on more than just the preceding state), it must be carefully constructed and abide by a condition that requires the adaptation to decrease as the chain progresses in order to retain convergence of the chain to the posterior density. Particular definitions of adaptation are beyond the scope of this dissertation and more information can be found in [7, 57, 56, 79, 87, 94].

As for the acceptance of a candidate  $q^*$ , we focus first on defining the acceptance probability  $\alpha$ . We form the ratio of the likelihoods of the new parameter  $q^*$  with that of the preceding parameter  $q^{k-1}$  as

$$r(q^*|q^{k-1}) = \frac{\pi(\vec{u}|q^*)\pi_0(q^*)}{\pi(\vec{u}|q^{k-1})\pi_0(q^{k-1})}.$$

We then define  $\alpha = \min(1, r)$  and set  $q^k = q^*$  with probability  $\alpha$ ; otherwise we reject  $q^*$  and set  $q^k = q^{k-1}$ . Rejections occur when the parameter is less likely (quantified through  $r$ ). This form for  $\alpha$  is known to provide the properties necessary for the chain to properly converge. However, we know that the proposed  $q^*$  is dependent on the choice of  $V$ . Even when adaptive methods are considered, it may sometimes be better to delay rejection and construct an alternative candidate. In algorithms using delayed rejection, if  $q^*$  is rejected a second stage candidate  $q^{*2}$  is found using the proposal function  $J(q^{*2}|q^{k-1}, q^*) = \mathcal{N}(q^{k-1}, \gamma_2^2 V_k)$ . Here,  $V_k$  is the current adaptive covariance matrix and  $\gamma_2 < 1$  ensures that the second stage proposal function is narrower than the original. This can be carried on for as many stages as desired, though we worked with just two stages based on the default in the software we use. Thus, delayed rejection is an open loop mechanism that alters the proposal function in a predetermined manner to improve mixing in the parameter space. Recall that mixing is the ability of the algorithm to properly, randomly sample the parameter space. The results in the next sections will demonstrate that adaptation was necessary, as the posterior results do not resemble the prior  $V_0$  that we specify. Delayed rejection was also key in producing well-mixed chains, so both these fixes were necessary.

At this point, we have described the way in which we construct a (local) random walk using a proposal function  $J$  and acceptance probability  $\alpha$  to sample the posterior density and thus solve (6.2.1). We introduced adaptation, which provides feedback to the proposal function based on the chain to that point, and delayed rejection. We now assume that the measurement errors

are iid, additive, and normally distributed. This means that the likelihood function is defined as

$$\pi(\vec{u}|q, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{\sum_{j=0}^{n-1} [u_j - u(L, t_j; 10^q)]^2}{2\sigma^2}}.$$

Note that this likelihood is dependent on  $\sigma^2$ , for which we have an estimate  $\sigma_{ols}^2$ . However, this parameter can also be treated in a Bayesian framework, namely by assuming it has a density which we sample through realizations of another Markov chain. The particular likelihood we have defined means  $\sigma^2$  is in the inverse-gamma family (the reason is discussed in, e.g., [86, 87]), so that we have

$$\sigma^2 | (\vec{u}, q) \sim \text{Inv-gamma} \left( \frac{n_s + n}{2}, \frac{n_s \sigma_s^2 + \sum_{j=0}^{n-1} [u_j - u(L, t_j; 10^q)]^2}{2} \right)$$

where  $n_s$  represents the number of observations encoded in the prior (which for us means  $n_s = 1$  since we have the single estimate for  $\sigma_{ols}^2$ ) and  $\sigma_s^2$  is the mean squared error of the data observations (for which a reasonable choice [86] is just  $\sigma_s^2 = \sigma_{ols}^2$ ). In the version under review for publication [60], we present results for  $n_s = 0$  (where the prior for  $\sigma^2$  does not depend on the data). The results are similar to those presented here.

The combined algorithm we use is delayed rejection adaptive Metropolis, or DRAM, with the assumption of iid, additive, normally distributed measurement errors. We now summarize the DRAM algorithm (with  $SS_q = \mathcal{J}_{ols}(q)$  throughout):

**Algorithm 6.2.2** (Delayed Rejection Adaptive Metropolis).

1. Set design parameters (e.g., the adaptation interval).
2. Determine  $q^0 = q_{ols} = \arg \min_q SS_q$ , (i.e., using (5.2.2)).
3. Set the initial sum of squares  $SS_{q^0}$ .
4. Compute the initial variance estimate  $\sigma_0^2 = \frac{1}{n-p} SS_{q^0}$ .
5. Construct the covariance estimate  $V_0 = \sigma_0^2 [\chi^T(q^0) \chi(q^0)]^{-1}$  and the corresponding Cholesky decomposition  $R_0 = \text{chol}(V_0)$ .
6. For  $k = 1, \dots, M$ 
  - (a) Sample  $z_k \sim \mathcal{N}(0, 1)$ .
  - (b) Construct candidate  $q^* = q^{k-1} + R_k z_k$ .
  - (c) Sample  $u_\alpha \sim \mathcal{U}(0, 1)$ , where  $\mathcal{U}(0, 1)$  is the uniform distribution on  $(0, 1)$ .
  - (d) Compute  $SS_{q^*}$ .

- (e) Compute  $\alpha(q^*|q^{k-1}) = \min \left( 1, e^{-[SS_{q^*} - SS_{q^{k-1}}]/2\sigma_{k-1}^2} \right)$ .
- (f) If  $u_\alpha < \alpha$ , set  $q^k = q^*$ ,  $SS_{q^k} = SS_{q^*}$ . Otherwise, enter the delayed rejection procedure ([86, Algorithm 5.10]):
  - i. Set the design parameter  $\gamma_2$ .
  - ii. Sample  $z_k \sim \mathcal{N}(0, 1)$ .
  - iii. Construct second-stage candidate  $q^{*2} = q^{k-1} + \gamma_2 R_k z_k$ .
  - iv. Sample  $u_\alpha \sim \mathcal{U}(0, 1)$ .
  - v. Compute  $SS_{q^{*2}}$ .
  - vi. Compute  $\alpha_2(q^{*2}|q^{k-1}, q^*)$ , where  $\alpha_2$  is a modified acceptance criterion described in Sec. 5.6.2 of [86].
  - vii. If  $u_\alpha < \alpha_2$ , set  $q^k = q^{*2}$ ,  $SS_{q^k} = SS_{q^{*2}}$ . Otherwise, set  $q^k = q^{k-1}$ ,  $SS_{q^k} = SS_{q^{k-1}}$ .
- (g) Update  $\sigma_k^2 \sim \text{Inv-gamma}((n_s + n)/2, (n_s \sigma_{k-1}^2 + SS_{q^k})/2)$ .
- (h) If at an adaptation interval (for us, every 10 chain values), compute adaptive update for proposal covariance  $V_k$ , then set  $R_k = \text{chol}(V_k)$ ; the adaptive update is computed recursively and using the preceding chain values.

We will later depict the MCMC chains (the sequence of parameter values accepted by the algorithm) which result from Algorithm 6.2.2, along with plots of the (posterior) parameter density. We used the DRAM options of the MCMC toolbox for Matlab, available from Marko Laine at [62]. The density plots are created using kernel density estimation (KDE), which constructs a density without a specified structure by weighted sums of a particular defined density. This is not unlike a finite element or polynomial approximation to a function. KDE software for Matlab is available from the Mathworks File Exchange at [33].

### 6.2.3 General Comparison Between Methods

We first discuss the complexity of the algorithms. The asymptotic algorithm is by far the least complex. At its core, the asymptotic theory linearizes about the true parameter value  $q_0$  to obtain an estimate for the covariance matrix. The estimate  $\hat{q}$  for  $q_0$  obtained by (5.2.2) and a corresponding approximation  $\hat{\sigma}^2$  for  $\sigma_0^2$  are then used, along with the sensitivity equations, to provide an estimate for the covariance matrix for  $\hat{q}$ . One important point of consideration is that the asymptotic theory is only able to construct confidence intervals that are combinations of Gaussian densities. Thus, if the density is multimodal or non-Gaussian, asymptotic theory will not be successful; there is no indication from asymptotic results that the form is anything other than Gaussian. The simplicity and speed of the asymptotic theory comes with the price of possibly reduced accuracy.



Bootstrapping does not make the linearization assumption. There, we resample from the error distribution (as approximated by the standardized residuals), generate simulated data, and obtain the  $M$  bootstrap estimates for  $\hat{q}$ . Nowhere in this process do we linearize; hence, compared with asymptotic theory, bootstrapping generally is better able to accurately incorporate correlation between parameters, particularly nonlinear correlation.

The MCMC DRAM algorithm also does not make linearization assumptions, using the likelihood function in order to accept or reject parameter candidates and construct the posterior parameter density. With DRAM, we can create pairwise plots of the components of the accepted parameter estimates to directly examine parameter correlations. Thus, we obtain the most information about the parameter estimate when using the Bayesian method since it constructs a density of the parameter itself rather than a sampling distribution. If one wishes to remain in a frequentist context, bootstrapping gives more accurate information about the confidence intervals than asymptotic theory. However, as one can see from the preceding algorithm descriptions, asymptotic theory is less complex than the other two methods. Thus, if one needs a quick estimate for the confidence in the parameter identification procedure, asymptotic theory may be the best choice. If one needs more or better information about the confidence in parameters, bootstrapping or Bayesian methods may be superior.

In terms of global optimization (finding the smallest possible cost function value), the Bayesian algorithm is naturally a (crude) global optimizer. The OLS estimation procedure we described earlier can possibly (and in practice is likely to) pick out a local minimum in the cost function which may not be the global minimum. Methods exist to make this process more global, including the obvious notion of starting the OLS procedure from different parameter values and selecting the result which has the lowest cost value. Though this is not necessarily of concern here, we mention this global property of the Bayesian algorithm as a reminder that the more complex algorithm may be able to reduce some of the work one may have otherwise done manually in order to find a good initial parameter guess to compute the OLS estimate.

In terms of computational time, asymptotic theory is significantly faster than the other two methods. One must be sure to have an accurate and reasonably fast way of solving for the sensitivity equations when using asymptotic theory, but that is the case in most problems of interest. Though we will later report that bootstrapping is faster than the Bayesian results for our model, this is not necessarily the case in general. First, bootstrapping can easily be parallelized by creating the  $M$  bootstrap samples and then splitting up the  $M$  inverse problems across multiple processor cores. Thus, if given enough processors, bootstrapping could be as fast as the asymptotic theory (of course, with significantly higher hardware requirements than the asymptotic theory). As for the Bayesian method, DRAM (Algorithm 6.2.2) is inherently serial since Markov chains in general are a serial process. However, successful methods for parallel Bayesian estimation have been developed and implemented including parallel DRAM [88] and

Differential Evolution Adaptive Metropolis (DREAM) [90, 96, 97]. The DREAM algorithm employs multiple MCMC chains and combines them using differential evolution to improve the overall estimation procedure. Thus, though we will later report times for the serial bootstrapping and Bayesian estimation procedures, these times will be different when using more sophisticated (and complex) implementations.

Finally, regarding correlation between parameters, asymptotic theory performs acceptably well as long as there is minimal or linear correlation between parameters which can be represented by a multivariate normal. Due to the linearizations made when developing the theory, any more complicated correlation may result in spurious conclusions. We will see this effect later;  $E$  and  $\gamma_1$  in our model are known to be correlated and together influence the frequency of the oscillations in the wave. These parameters turn out to be correlated in a nonlinear way, which the asymptotic theory is unable to handle. Bootstrapping and Bayesian methods, on the other hand, are able to properly estimate the confidence in parameters even with some nonlinear correlations. Also, the Bayesian method allows us examine the correlations directly through pairwise plots of parameter components. As we will later see, the DRAM results clearly show nonlinear correlation between  $E$  and  $\gamma_1$ .

Overall, the asymptotic analysis, though simple to implement, has its drawbacks in terms of being less able to properly describe complex and/or non-Gaussian relationships between model parameters. Bootstrapping and Bayesian methods, though more intensive to implement, are able to better describe more complex behavior. If one wishes to know the most about the model parameters, the Bayesian approach may be the most successful as it directly estimates the posterior parameter density (and is in fact the only method to do so, as frequentist methods only estimate the sampling distribution). The parameter density can then be used to make the most accurate predictions of model solution behavior. The sampling distribution can only be used to predict the behavior of the model solution if one is convinced that the sampling distribution and the parameter distribution are the same for a particular problem. In general, there is no one superior method for all problems; rather, one must choose the method which best fits the situation at hand and the goals of a particular problem.

### 6.3 Results Based on Parameter Values From Chapter 5

We begin our discussion using the parameter values obtained in the previous chapter. These were obtained through the OLS estimation procedure, and so may be located in a local minimum. In Chapter 5, we only examined the asymptotic standard errors; here we expand that analysis to include bootstrapping and Bayesian methods, which provide some novel insight into the problem and spur on the later sections in the current chapter. Recall the results of OLS estimation from Chapter 5 as shown in Figure 5.3. Recall also the confidence analysis for the OLS estimate

in Table 5.1. This is the baseline with which we next compare the bootstrapping and MCMC DRAM results.

### Bootstrapping confidence intervals

Table 6.1: Pressure bootstrap results.

Param.	Estimate	SE	CI95
$\log_{10}(E)$	4.5792	0.1613	(4.2604, 4.8979)
$\log_{10}(E_1)$	1.6897	0.1325	(1.4277, 1.9517)
$\log_{10}(\tau_1)$	-1.3719	0.1638	(-1.6957, -1.0480)
$\log_{10}(\gamma_1)$	5.2757	0.0330	(5.2105, 5.3409)
$\log_{10}(-A)$	-3.7516	0.0060	(-3.7635, -3.7396)
$\log_{10}(-\Upsilon)$	-1.8911	0.2011	(-2.2886, -1.4935)

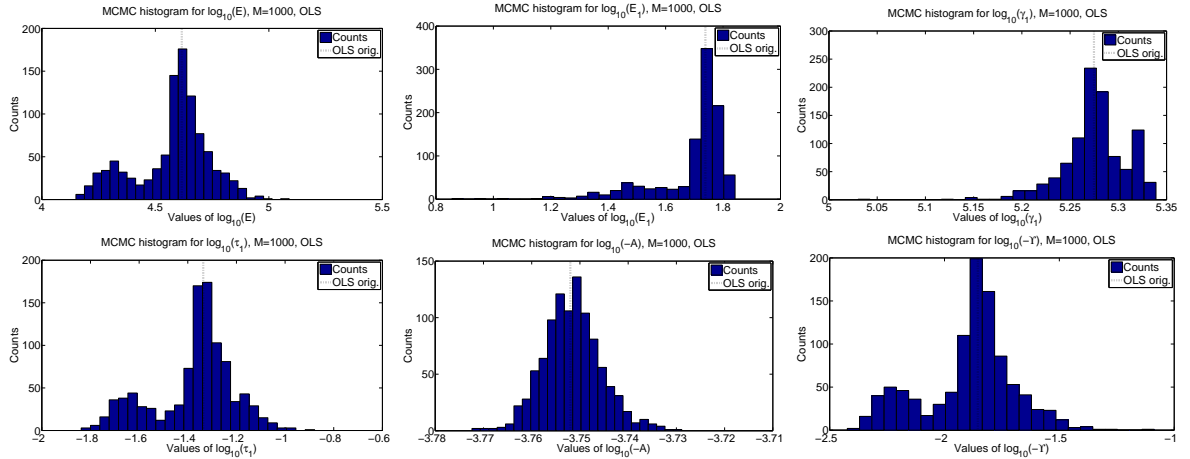


Figure 6.1: Pressure bootstrap histograms, with data sampled at 1024 Hz,  $N_p = 1$ . The corresponding OLS estimate from Table 5.1 is shown as the dashed line. All have a mostly normal shape, with some small peaks but mainly one large peak for each parameter.

We have also computed bootstrap results around the parameter values in Table 5.1. One way of confirming that the bootstrapping procedure is working properly (particularly that the normal assumption (3.5.2) holds) is to examine the histogram of all  $M = 1000$  bootstrap parameter estimates. We show these histograms in Figure 6.1. Some of the histograms for parameters show

some small peaks away from the larger peak, though the counts are largely centered in a large peak around the respective OLS parameter values (shown as the dashed line). This is the first indication that there may be more underlying complex relationships between the parameters than the asymptotic results showed. In Table 6.1 we show the results from computing the bootstrap mean and confidence intervals. In the table, we see the standard errors are generally smaller than those from the asymptotic results. However, based on experience with the methods this is not an expected result, as bootstrapping makes fewer restrictive assumptions in the development of the theory and is thus more likely to provide wider (generally more conservative) confidence intervals. This is another indication that complex behavior is being exhibited in this problem, which asymptotic theory is unable to handle with the problem as stated of finding all six parameter values. We will defer further discussion until after examining the Bayesian results.

### Bayesian estimation results

Using the DRAM method previously described, we computed parameter densities. We found  $M = 50,000$  parameter values in the chain. The MCMC chains are shown in Figure 6.2, and densities in Figure 6.3. We also include a pairwise plot of the parameters in Figure 6.4, which allows us to examine correlation between parameters.

There are two main results we see from Figure 6.2. First, the chains are mixing reasonably well; they are fairly randomly walking through the parameter space and do not get stuck on particular values for long. Second, the DRAM procedure seems to have found another local minimum, as evidenced by the jumps (between the 15,000 and 25,000 iteration marks) in the apparent chain mean for nearly all parameters except  $\log_{10}(-A)$ . The jump is most clear in the plot for  $\log_{10}(\gamma_1)$ . This shows up in the density plots in Figure 6.3 as the peaks of the densities for all parameters except  $\log_{10}(-A)$  being away from the OLS estimate (denoted with the dashed line). We examine this further in Section 6.4, where we run the OLS estimation procedure but use the mean of the MCMC parameter results as the initial guess for OLS; in this way, we hope to provide evidence that a new local cost minimum was found (along with a corresponding new parameter estimate), and then recompute the uncertainty analysis for each method.

Before proceeding, though, we discuss the relationships between parameters in the pairwise plots of Figure 6.4. We see strong correlations between all parameters except  $A$ . The strong correlation between  $E$  and  $\gamma_1$  is not surprising. These parameters together describe the overall frequency of the resulting wave propagation through the viscoelastic medium, so if one decreases the other must increase to compensate. The parameter  $E$  describes the exclusively solid response to stresses while  $\gamma_1$  represents the remaining response factors. It is also not surprising that  $\tau_1$  and  $E_1$  are correlated;  $E_1$  determines the overall damping envelope for the wave while  $\tau_1$

(in part) describes smaller damping effects, effects which we might expect to be related. Given these correlations between  $(E, \gamma_1)$  and  $(\tau_1, E_1)$ , it then makes sense that these four parameters are correlated with each other. Since  $\Upsilon$  is a nonphysical parameter which is heavily dependent on the values of  $\gamma_1$  and  $\tau_1$ , it is expected that it is correlated highly with those two parameters and as a result also with  $E$  and  $E_1$ . Finally, since  $A$  represents the initial load position, it is the parameter best known directly from the data and thus does not change much even as the other parameters change. In order to ensure that we make proper conclusions about the relationships between parameters, we will recompute the DRAM results from a new starting point and reexamine these correlations again.

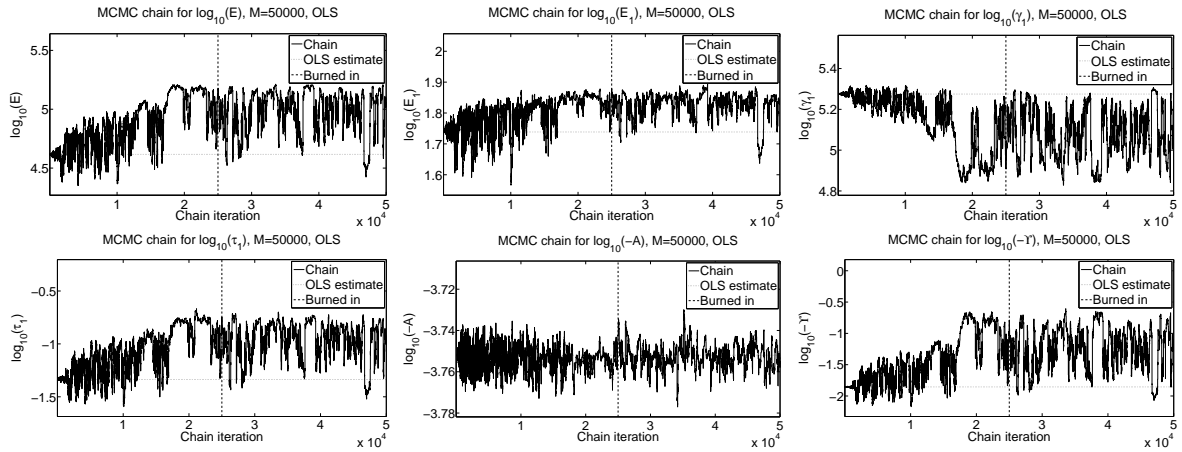


Figure 6.2: Pressure parameter chain results using DRAM, with data sampled at 1024 Hz,  $N_p = 1$ . The corresponding OLS estimate from Table 5.1 is shown as the horizontal dashed line. The vertical dashed line represents the iteration after which we consider the chains to be burned in.

### 6.3.1 Comparison of Preliminary Results Between Methods

In order to best compare the methods, for each parameter we plot in Figure 6.5 the asymptotic and bootstrap normal densities (which visually show the mean and variance estimates obtained previously) on top of the Bayesian posterior density. Though the asymptotic and bootstrapping densities have peaks near the same value (which is expected, as both are used to construct confidence bands around the OLS result, in this case), the Bayesian densities are quite different. Most are skewed to the left (which was nearer the location of the original OLS parameter estimate), and are not centered around the same parameter value as the frequentist confidence analysis methods. This is due to the fact that the Bayesian process also searches the parameter

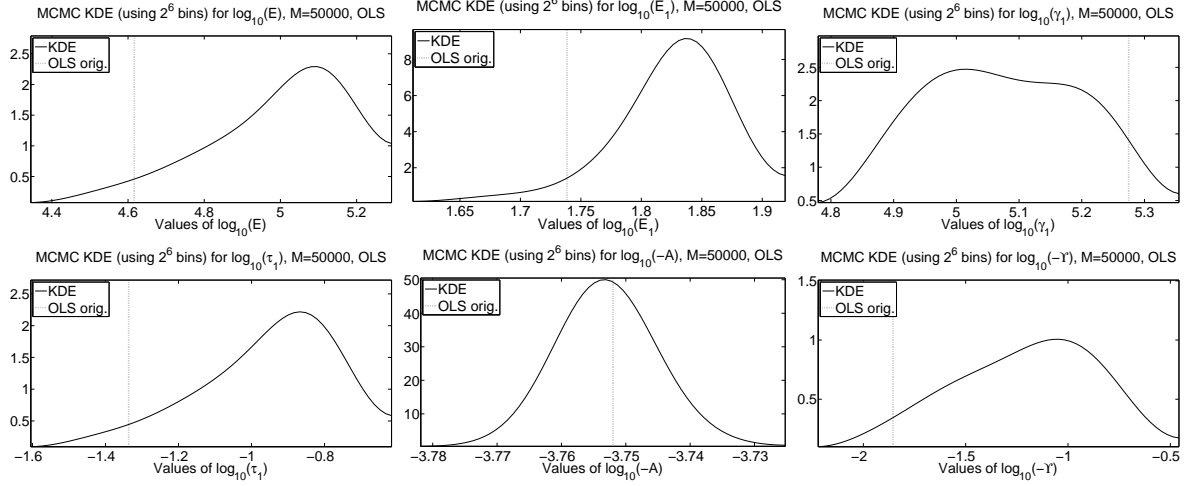


Figure 6.3: Pressure parameter density results using DRAM, with data sampled at 1024 Hz,  $N_p = 1$ . The corresponding OLS estimate from Table 5.1 is shown as the dashed line.

space for a smaller cost value, which it found (as evidenced earlier in the parameter chains of Figure 6.2). The resulting difference between the frequentist and Bayesian results will be remedied in the next section by using the MCMC results to restart the OLS procedure, finding the OLS frequentist estimate for the new parameter values which lies in this new cost function local minimum and then re-computing the asymptotic, bootstrapping, and Bayesian confidence analysis from that new estimate. We turn to these computations next.

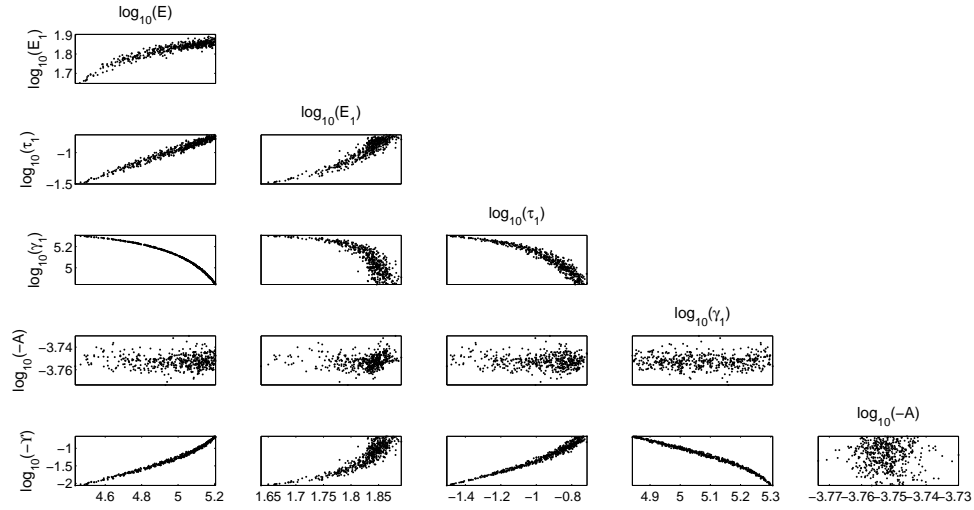


Figure 6.4: Pressure parameter pairwise parameter comparisons using DRAM, with data sampled at 1024 Hz,  $N_p = 1$ , using every 50th chain value. Noticeable patterns in each sub-graph indicate correlation between parameters (e.g., the upper left corner between  $E$  and  $E_1$ ) whereas roughly random relationships indicate less correlations (e.g., the parameter  $A$  with any other parameter).

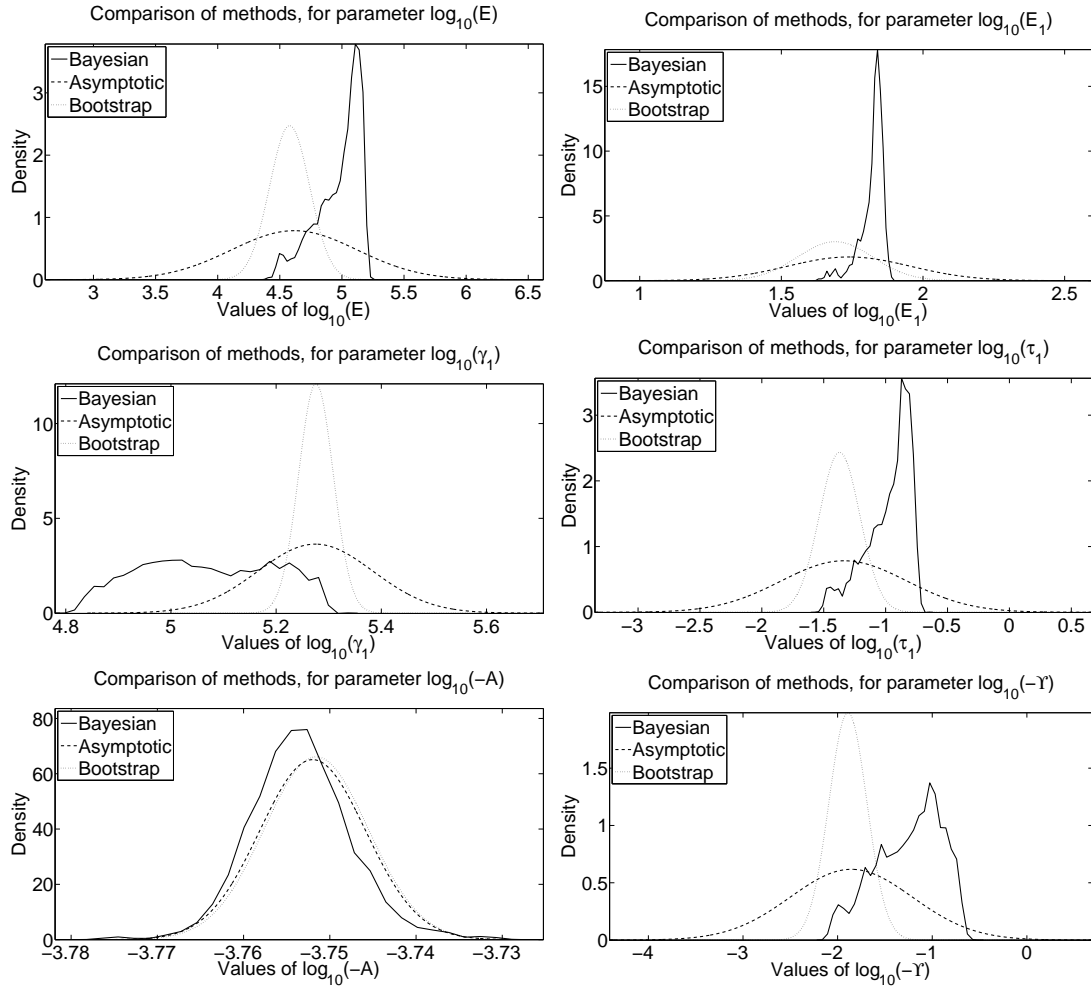


Figure 6.5: Comparison of density results between the three methods, for each parameter. The MCMC density is computed from the chain values using KDE; for bootstrapping and asymptotic analysis, the mean and variance for each parameter are used to create a normal probability density function which is plotted.



## 6.4 Updating Parameter Estimates

We now turn to re-estimating the parameter values in order to find an OLS estimate in the new local minimum found by the DRAM algorithm. At this new value, we will re-compute the asymptotic, bootstrapping, and Bayesian algorithms to update the results. This will allow a better comparison between methods.

### 6.4.1 OLS Inverse Problem Results

We begin by solving the inverse problem (5.2.2) using the means of the MCMC results obtained in the previous section. The results are shown in the left column of Table 6.2. Note that we have refined the OLS results from what was presented in the previous chapter, though the values here differ from the previously reported results only slightly. We show the model solution, as well as residual plots, in Figure 6.6. The residuals appear random, which indicates the validity of our original assumption of iid errors due primarily to measurements. Our next step will be to run an MCMC chain using the OLS results as a starting place. We will also compute the asymptotic and bootstrap confidence analysis around this new OLS parameter value.

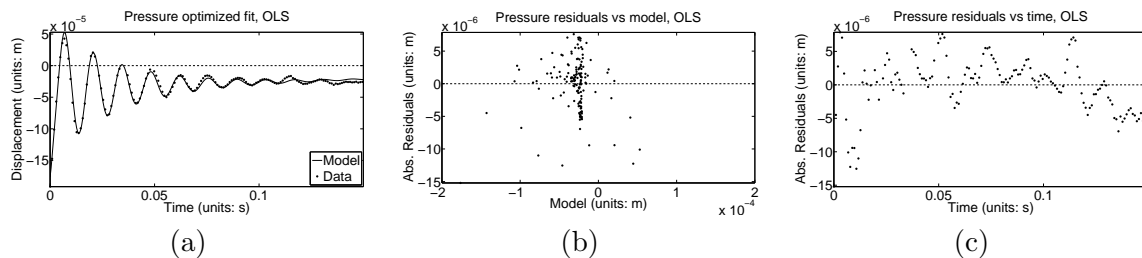


Figure 6.6: Pressure data fit using OLS, with initial guess set to be the mean of the MCMC parameter chains, with data sampled at 1024 Hz,  $N_p = 1$ . (a) OLS model fit to data. (b) Model vs absolute residuals. (c) Time versus absolute residuals.

### 6.4.2 Initial Comparison Between Methods

The results of computing the asymptotic analysis using the OLS parameter values are reported in Table 6.2. We also compute the bootstrap confidence intervals around that OLS parameter estimate (using  $M = 1000$  bootstrap samples). The bootstrapping results are shown in Table 6.3. Finally, we run the Bayesian DRAM algorithm (with  $M = 50,000$ ), reporting the chains in Figure 6.8. A comparison of the resulting densities is shown in Figure 6.7, and a comparison of the parameter estimate (OLS) and parameter means (bootstrapping and Bayesian) is shown in

Table 6.4 along with computational times. Finally, note that we have run chains for longer which produced similar results, indicating that we have converged to the posterior at  $M = 50,000$ ; thus we are confident that the results presented here are from the posterior.

We see that the densities for all three methods are largely centered at the same location. The asymptotic theory gives the widest confidence band, which is actually quite unexpected. Since the asymptotic theory holds only for large numbers of data points (we only use  $n = 251$  data points here) and incorporates a linearization, the use of asymptotic theory can result in too-small confidence intervals. The wider confidence intervals we found may be due to complex nonlinearities between parameters, which we will soon examine further. For the damping parameter  $E_1$ , bootstrapping and Bayesian methods give largely comparable results, with the asymptotic theory being wider. For the parameter  $A$ , the results are roughly the same between methods – this is expected, again since that is the parameter most readily verified by directly from the data. This parameter is also largely uncorrelated with the remaining parameters, providing an indication that correlations may be causing problems with the asymptotic theory. For the remaining four parameters, the bootstrapping confidence bands were the narrowest, though more comparable to the Bayesian results than the asymptotic analysis. We also point out the ability of the DRAM algorithm to clearly show the nonlinear correlations, particularly between  $E$  and  $\gamma_1$ , is a significant achievement. It indicates that the DRAM algorithm is effective on this problem. Also, the chains in Figure 6.8 are mixing reasonably well after the burn-in phase, an indication that we are likely sampling adequately from the posterior density.

In this case of estimating all six parameters, in order to gain a sense of the uncertainty we would need to use either bootstrapping or Bayesian methods due to the parameter correlation which asymptotic theory is unable to accommodate. Even though the asymptotic results are much faster to compute, this situation necessitates the use of the more complex algorithms to obtain correct results. Whether one should use bootstrapping or Bayesian methods in practice on this problem depends on situational considerations (e.g., if the parameter uncertainty needs to be propagated through the model or whether we are just concerned with uncertainty in the estimation procedure). It is also of interest to see how these results change when the nonlinear correlation is no longer present. Without the nonlinear correlation, asymptotic theory may give improved performance; bootstrapping and asymptotic theory are then likely to appear comparable to the Bayesian densities if the parameter densities are normal. We examine this in the next section.

Table 6.2: Pressure asymptotic results (using MCMC parameter means from the chains in Figure 6.2 as the initial guess).

Param.	Estimate	SE	CI95
$\log_{10}(E)$	5.0708	1.8364	(1.4404, 8.7011)
$\log_{10}(E_1)$	1.8412	0.2179	(1.4104, 2.2719)
$\log_{10}(\tau_1)$	-0.8804	1.8394	(-4.5166, 2.7558)
$\log_{10}(\gamma_1)$	5.0510	1.9143	(1.2669, 8.8352)
$\log_{10}(-A)$	-3.7539	0.0062	(-3.7662, -3.7416)
$\log_{10}(-\Upsilon)$	-1.1109	4.4702	(-9.9477, 7.7259)

Table 6.3: Pressure bootstrap results (using MCMC parameter means from the chains in Figure 6.2 as the initial guess).

Param.	Estimate	SE	CI95
$\log_{10}(E)$	5.0672	0.0432	(4.9817, 5.1526)
$\log_{10}(E_1)$	1.8399	0.0250	(1.7904, 1.8893)
$\log_{10}(\tau_1)$	-0.8829	0.0484	(-0.9786, -0.7873)
$\log_{10}(\gamma_1)$	5.0518	0.0346	(4.9834, 5.1201)
$\log_{10}(-A)$	-3.7538	0.0057	(-3.7651, -3.7425)
$\log_{10}(-\Upsilon)$	-1.1123	0.0889	(-1.2882, -0.9365)

Table 6.4: Pressure optimization results for the following cases: OLS, bootstrapping mean using OLS value to initiate, and DRAM parameter means; data frequency 1024 Hz.  $RSS = \mathcal{J}_{ols}(q)$

	Method		
	OLS	Bootstrap mean	MCMC mean
$\log_{10}(E)$	5.0708	5.0672	5.0164
$\log_{10}(E_1)$	1.8412	1.8399	1.8327
$\log_{10}(\tau_1)$	-0.8804	-0.8829	-0.9300
$\log_{10}(\gamma_1)$	5.0510	5.0518	5.0696
$\log_{10}(-A)$	-3.7539	-3.7538	-3.7539
$\log_{10}(-\Upsilon)$	-1.1109	-1.1123	-1.1765
RSS	2.538e-009	2.569e-009	6.136e-009
Comp. Time	0.1662 mins	25.99 hrs*	52.82 hrs*

*\*Note: Computations here are serial. Bootstrapping may be easily parallelized, and parallelized versions of MCMC exist; see the discussion in Section 6.2.3.*

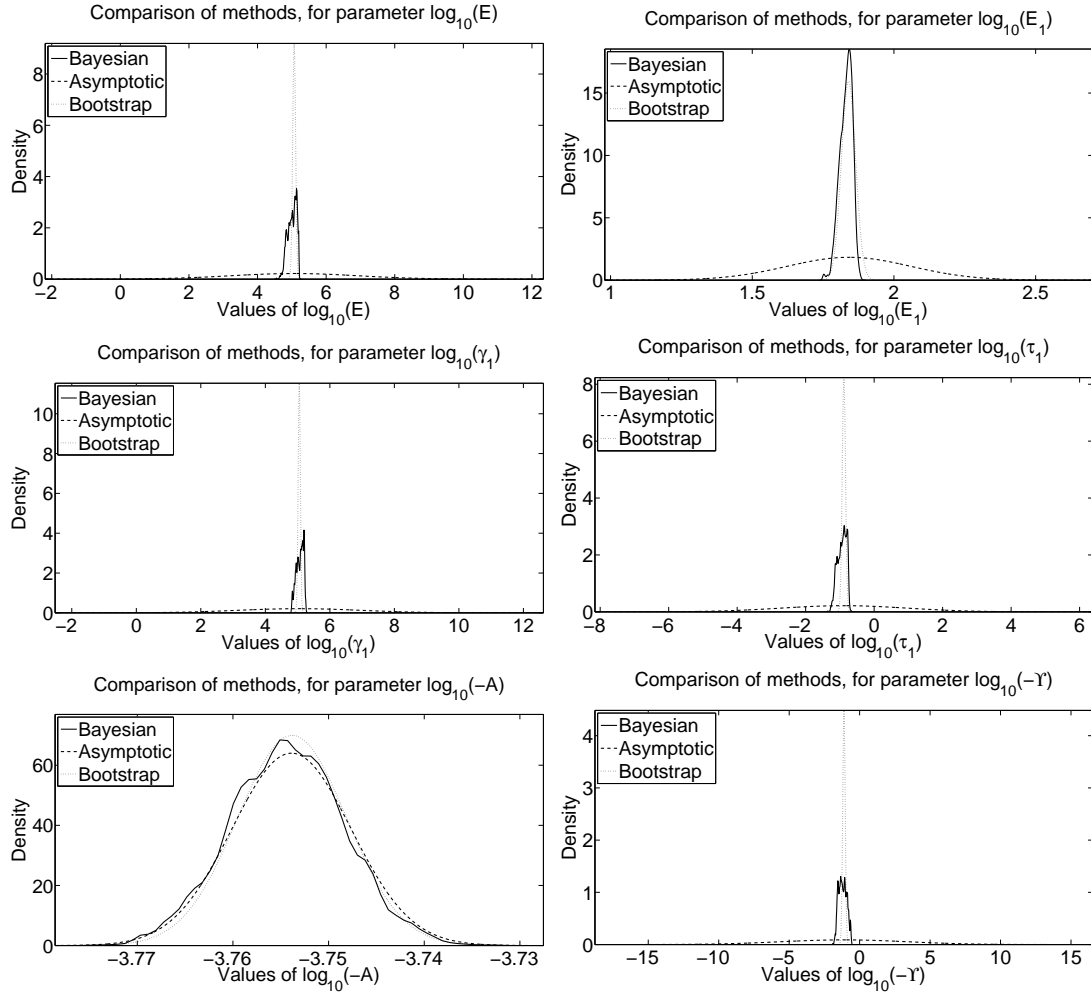


Figure 6.7: Comparison of density results between the three methods, for each parameter. The MCMC density is computed from the chain values using KDE; for bootstrapping and asymptotic analysis, the mean and variance for each parameter are used to create a normal probability density function which is plotted.

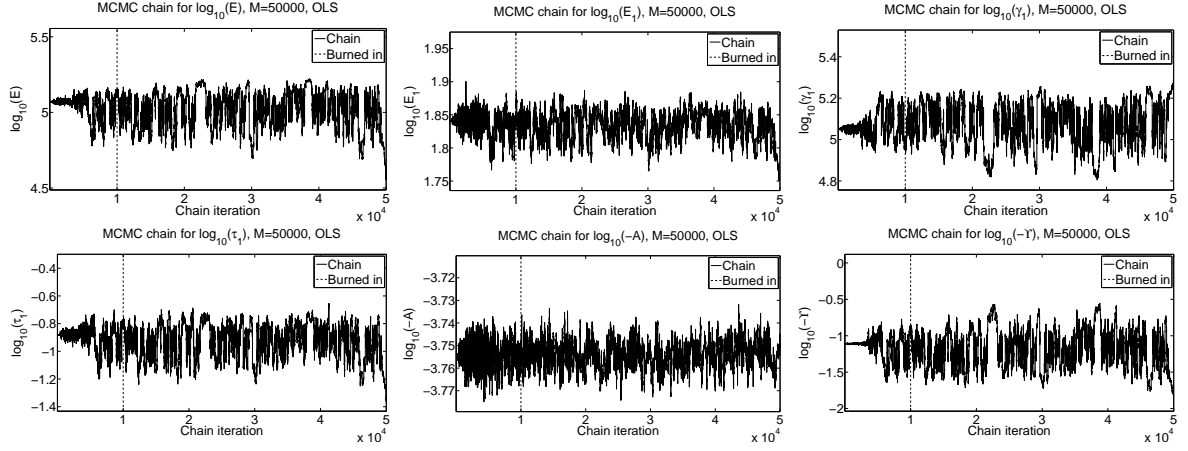


Figure 6.8: Pressure parameter chain results using DRAM, with data sampled at 1024 Hz,  $N_p = 1$ . The corresponding OLS estimate from Table 6.4 is shown as the dashed line.

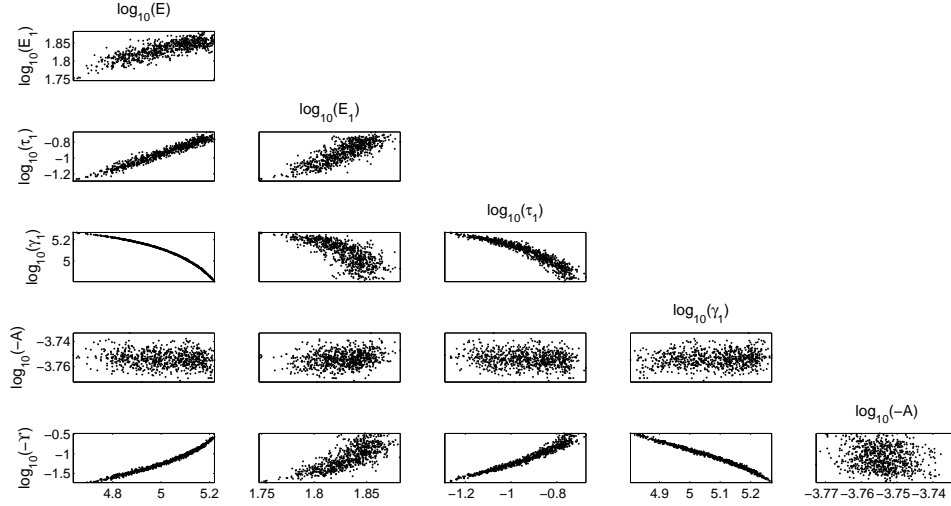


Figure 6.9: Pressure parameter pairwise comparisons using DRAM, with data sampled at 1024 Hz,  $N_p = 1$ , using every 50th chain value. Noticeable patterns in a sub-graph indicate correlation between parameters (e.g., the middle left comparison between  $\gamma_1$  and  $E$ ) whereas roughly random relationships indicate less correlations (e.g., the parameter  $A$  with any other parameter).

## 6.5 Method Comparison When Holding $E$ Constant

Given the strong correlation between  $E$  and the other parameters, particularly  $\gamma_1$ , we next examine what happens when we hold  $E$  constant and estimate the remaining parameters. This is interesting from the pedagogical standpoint, but also from a modeling standpoint. We know from the development of the model that  $E$  and  $\gamma_1$  are highly correlated (both contribute to modeling the oscillation frequency). For this problem, it might be possible to design an experiment to separately estimate the value of  $E$  which could then be held fixed when estimating the remaining parameters which are difficult to physically measure. In terms of how this affects computations here, now  $E$  will not be an element in the covariance matrix, and the strong correlation between it and other parameters will not affect the asymptotic computation (which we suspect does not handle the nonlinear correlations well). Estimated parameter values are shown in Table 6.5. Asymptotic results are in Table 6.6, bootstrap results in Table 6.7, the MCMC chains in Figure 6.11 (which are mixing even better than before), and densities (sampling for asymptotic and bootstrapping, posterior for Bayesian) in Figure 6.10.

We now see much more consistent relationships between the different methods. Asymptotic standard errors, though not smaller than bootstrap standard errors, are nearly identical; this is fine, and serves to indicate that the linearizations made when deriving the asymptotic theory are applicable for the restricted inverse problem. Without the highly nonlinear correlation between  $E$  and  $\gamma_1$ , the sampling distribution of the OLS estimate in the asymptotic and bootstrapping results lies very near the posterior parameter density found by the Bayesian method. There are slight differences in the apparent mean of the distribution for each parameter (see Table 6.5), but the differences are fairly insignificant. The variances of the sampling distributions and of the posterior density are nearly identical across the methods. Thus, now that the parameters are either uncorrelated or only linearly correlated (see plots of Figure 6.12), the asymptotic theory provides results that are comparable to the other two methods and which are obtained with significantly less computational effort. However, the need to fix either  $E$  or  $\gamma_1$  would not have been apparent had we not appealed directly to the physical meanings of these parameters or had we not computed the bootstrap confidence intervals and, particularly, the Bayesian posterior densities which demonstrate the strong nonlinear correlation in these parameters. We will comment further on this point in the conclusion.

Table 6.5: Pressure optimization results for the following cases (with  $E$  fixed): OLS, bootstrapping mean using OLS value to initiate, and DRAM parameter means; data frequency 1024 Hz.  $RSS = \mathcal{J}_{ols}(q)$ .

	Method		
	OLS	Bootstrap mean	MCMC mean
$\log_{10}(E_1)$	1.8406	1.8403	1.8408
$\log_{10}(\tau_1)$	-0.8818	-0.8809	-0.8776
$\log_{10}(\gamma_1)$	5.0510	5.0511	5.0511
$\log_{10}(-A)$	-3.7540	-3.7540	-3.7543
$\log_{10}(-\Upsilon)$	-1.1117	-1.1097	-1.1081
RSS	2.538e-009	2.539e-009	2.538e-009
Comp. Time	0.1691 mins	11.33 hrs*	46.70 hrs*

*\*Note: Here again the computations are serial.*

Table 6.6: Pressure OLS asymptotic results, with  $E$  fixed.

Param.	Estimate	SE	CI95
$\log_{10}(E_1)$	1.8406	0.0122	(1.8164, 1.8648)
$\log_{10}(\tau_1)$	-0.8818	0.0311	(-0.9433, -0.8203)
$\log_{10}(\gamma_1)$	5.0510	0.0026	(5.0459, 5.0562)
$\log_{10}(-A)$	-3.7540	0.0057	(-3.7654, -3.7427)
$\log_{10}(-\Upsilon)$	-1.1117	0.0248	(-1.1606, -1.0627)

Table 6.7: Pressure bootstrap results, with  $E$  fixed.

Param.	Estimate	SE	CI95
$\log_{10}(E_1)$	1.8403	0.0127	(1.8153, 1.8654)
$\log_{10}(\tau_1)$	-0.8809	0.0310	(-0.9423, -0.8196)
$\log_{10}(\gamma_1)$	5.0511	0.0026	(5.0459, 5.0562)
$\log_{10}(-A)$	-3.7540	0.0057	(-3.7652, -3.7427)
$\log_{10}(-\Upsilon)$	-1.1097	0.0247	(-1.1586, -1.0608)

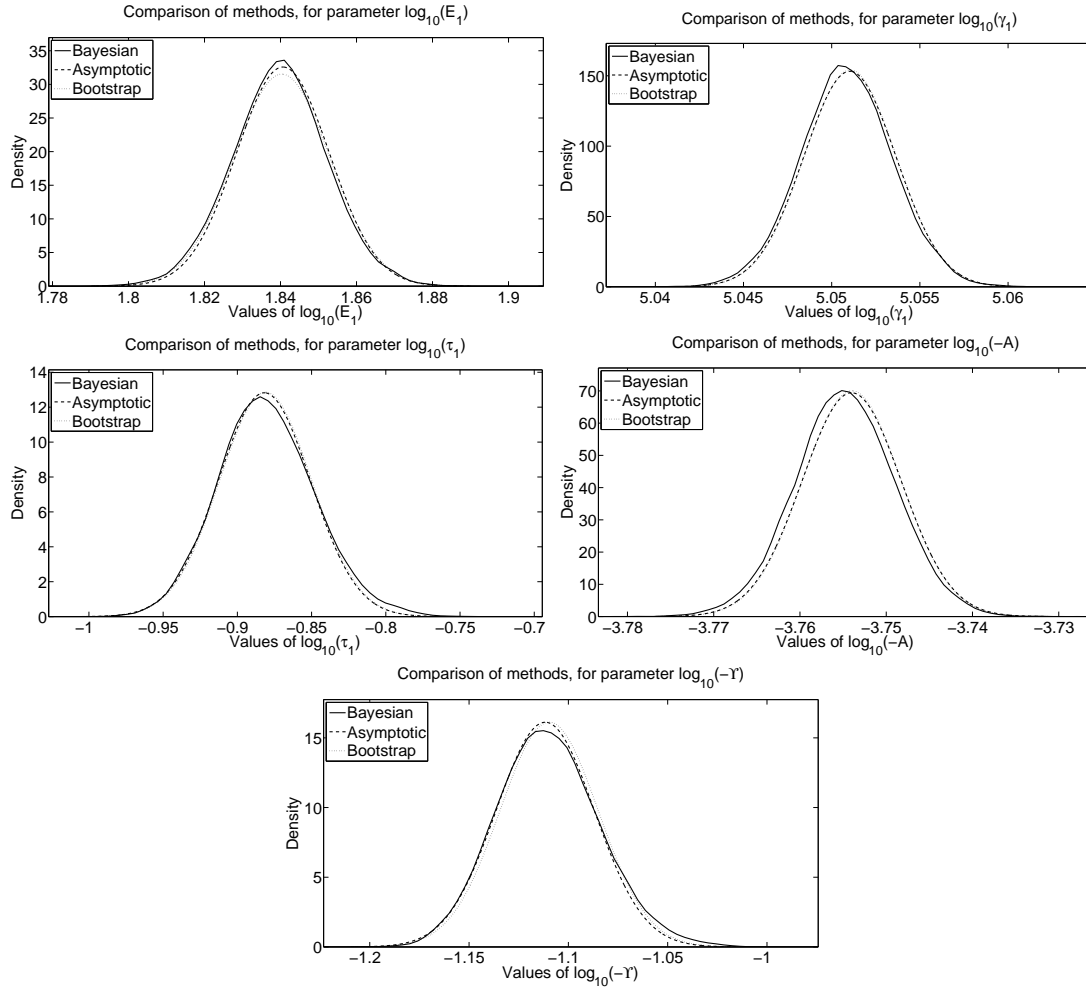


Figure 6.10: Comparison of density results between the three methods, for each parameter (with  $E$  fixed). The MCMC density is computed from the chain values using KDE; for bootstrapping and asymptotic analysis, the mean and variance for each parameter are used to create a normal probability density function which is plotted.



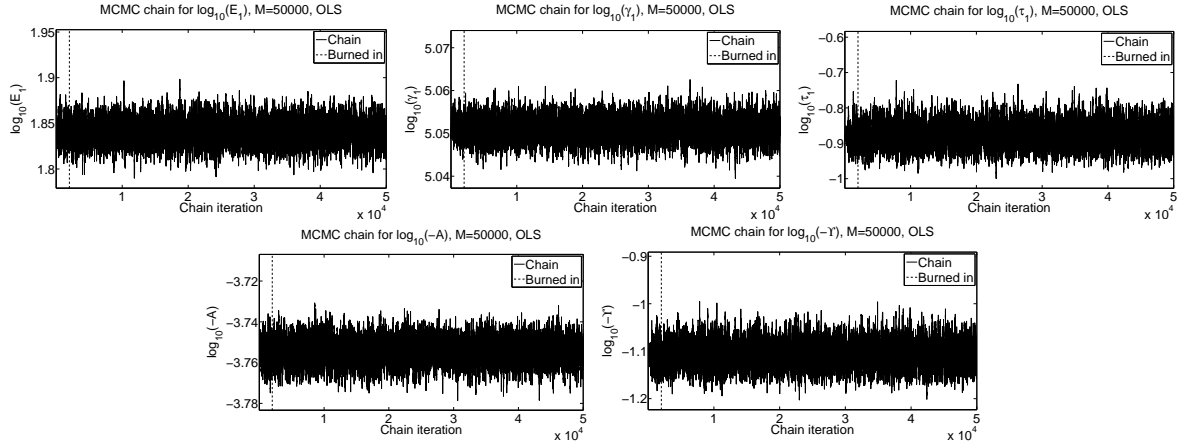


Figure 6.11: Pressure parameter (with  $E$  fixed) chain results using DRAM, with data sampled at 1024 Hz,  $N_p = 1$ . The vertical dashed line indicates the point at which we consider the run to be burned in.

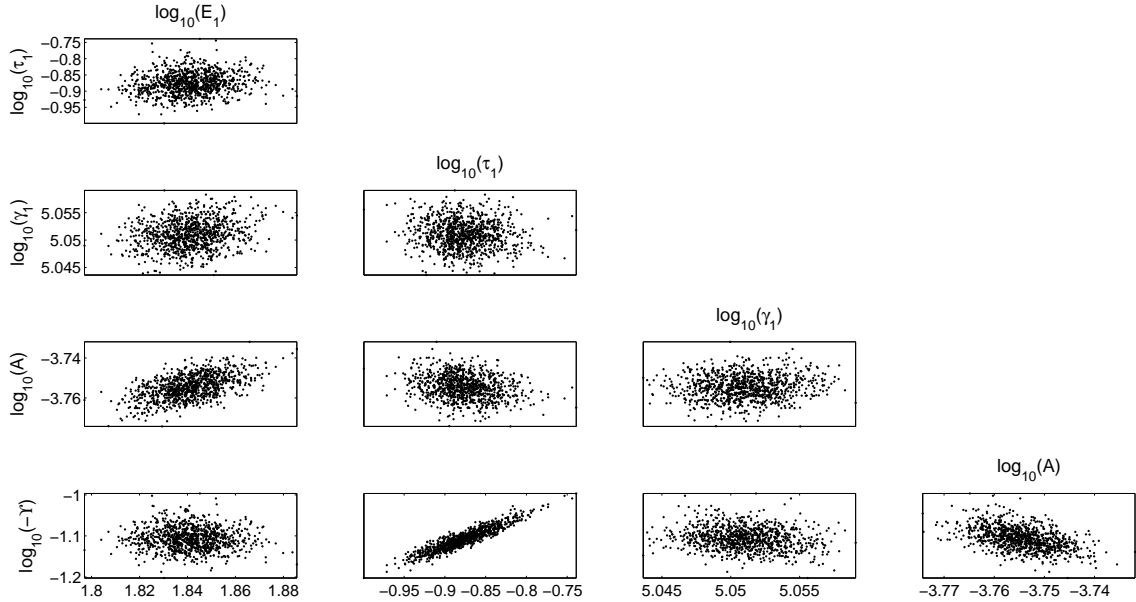


Figure 6.12: Pressure parameter (with  $E$  fixed) pairwise comparisons using DRAM, with data sampled at 1024 Hz,  $N_p = 1$ , using every 50th chain value. Note that most relationships are significantly more random than in the previous pairwise plot results, indicating that these parameters are roughly uncorrelated, although some linear correlation may still exist (e.g.,  $\tau_1$  and  $\Upsilon$ )

## 6.6 Uncertainty Quantification Analysis Summary

We have seen that each of the three inverse uncertainty quantification methods can be effective on a problem. If parameters are highly correlated in a nonlinear way, asymptotic analysis alone may provide at best unclear or at worst wrong results. However, one may not suspect the nonlinear correlations without additional computations like the bootstrap or Bayesian methods. The Bayesian approach naturally searches not only for a global cost minimizer, but more importantly, it provides a joint density for the parameters. We found that leaving  $E$  in the model resulted in strong correlation between parameters, which asymptotic theory was unable to discern, bootstrapping suggested, and Bayesian methods clearly described.

If we assumed  $E$  was a fixed value and estimated the remaining five parameters, the nonlinear relationships between parameters were reduced or eliminated. At that point, all three methods gave results that were similar. The amount of data  $n$  required to accurately appeal to the limit statements in the asymptotic theory is problem-dependent and difficult to quantify in a rigorous manner. For this problem, the consistency of the asymptotic and Bayesian results when  $E$  is fixed indicates that  $n$  is sufficiently large to permit asymptotic analysis. Recalling that the two frequentist methods describe the sampling distribution while the Bayesian method describes the parameter densities directly, this is a fairly convincing indication that the sampling distribution is an acceptable approximation to the parameter distribution if we fix  $E$ , reducing complex correlation between parameters. Thus, if our goal had been to propagate parameter uncertainty through the model, in this case we could use the asymptotic results (which, as has been discussed, is not true unless we show the sampling and posterior parameter densities are approximately the same). Note that in order to propagate uncertainty using the asymptotic results, one would need to require that the joint asymptotic distribution is similar enough to the Bayesian posterior; in this case, the marginal distributions (Figure 6.10) are very similar so one would need to examine the correlations in each method in order to determine if the joint distributions are also similar between methods. However, without resorting to the physical meaning of these parameters, this may not have been found (and was not found in our initial calculations) until we used the more computationally expensive bootstrapping and Bayesian techniques. If one wishes to propagate uncertainty through the model, either the parameter densities from the Bayesian method must be used or one must be convinced that the frequentist sampling distribution is an adequate approximation of the parameter density (which requires the Bayesian results for comparison). If one is just concerned with uncertainty in the parameter estimation procedure, then the frequentist methods may be adequate.

While no formal definitive conclusions can be made, the results from this chapter suggest that some serious consideration should be given before choosing any single one of these three methods. Each method has strengths and weaknesses, which must be considered and balanced for

a problem at hand. Since the ultimate goal of this project is a diagnostic device, the particular method of choice will be subject to time restrictions (i.e., asymptotic may be the only feasible choice due to a need for quickly obtained uncertainty results) and accuracy requirements. One possible compromise would be fixing  $E$ , as we have studied; this allows for a parameter identification problem that is still broad enough to incorporate the remaining model parameters (broad in the sense of our ability to match the model to different patients) while indicating that asymptotic results are likely equivalent to the more computationally intensive methods in this case. These considerations may be carried over to a two- or three-dimensional model that is closer to modeling the *in vivo* stenosis problem; the results discussed here will be beneficial in directing the uncertainty studies needed when using a higher-dimensional model.

---

## Conclusion and Future Work

---

In this dissertation we first introduced a novel acoustic phantom experimental rig. One-dimensional pressure data is available from one setup, while one-dimensional shear is available from another. Both setups provide displacement data from a weight-drop experiment designed to produce oscillations in the gel of a magnitude comparable to that produced by blood flow in a stenosed coronary artery impacting the vessel wall, a disturbance which results in shear wave propagating away from the vessel walls downstream of the blockage. As our model validation has shown, this data encodes information about the elastic or shear modulus as well as information about bulk damping and the relaxation times.

In terms of modeling, we first carried out proof-of-concept investigations for estimating material parameters and created a model comparison test as a basis for distinguishing between data that comes from a normal or from a stenosed blood vessel. We found that the model was less sensitive to a second viscoelastic relaxation time than to the other parameters, and this was manifested as a difficulty in recovering two relaxation times. On the other hand, models with zero or one relaxation time allowed for more confidence in the estimation procedure (i.e., smaller standard errors). We compared asymptotic error theory with bootstrapping error theory, and found (as expected) that bootstrapping gives more conservative (larger) confidence intervals but not so much so that the asymptotic theory cannot be profitably used for uncertainty quantification in models with large computational costs rendering bootstrapping less desirable. However, we later saw in Chapter 6 that similar confidence analysis results did not initially hold for experimental data (we needed to fix  $E$  in order for asymptotic and bootstrapping results to be comparable). In terms of the model comparison on the input amplitude parameter  $A$ , we

were able to develop a successful methodology for statistically determining whether or not data came from a low amplitude input force. This will form the basis of a model comparison test we can use on experimental data sets.

We then developed an updated one-dimensional viscoelastic model for tissue and have used the experimental data to test the ability of our model to describe wave propagation in the medium. In the inverse problem results as discussed in Section 5.3, we have shown an ability to consistently model the wave propagation using different error models and at different data sampling frequencies, obtaining good fits to data in all of our inverse problems. In addition to a good fit, we also examined statistical properties of the parameter estimators as well as residual plots to gain more insight into the proper error model for the shear data set. This is necessary, since a correct error model is essential in order to apply the asymptotic error theory properly and thus obtain correct confidence intervals. For the pressure case, the results were inconclusive. For the shear case, we recommend taking the more conservative route and using the GLS parameter estimates; even though the GLS estimates had larger asymptotic theory standard errors, there were indications from the residual versus time plots for OLS that the OLS model is not correct. In these modeling efforts with data we have successfully demonstrated the ability of mathematical model to accurately describe the data from laboratory experiments using a homogenous tissue-mimicking material gel phantom. A linear viscoelastic constitutive relationship, i.e., (4.4.14b), was adequate. This is a significant achievement, as all the work previously discussed was limited to inverse problems on simulated data or data that was not from the impulse-type experiments. For the shear model, which is the focus of detection efforts in practice, we have indicated that a relative error model is likely the best option, which is what one would intuitively believe when measuring an oscillating phenomenon.

We then used the pressure wave propagation model (5.2.2) and corresponding inverse problem for the model parameters to illustrate how to provide more complete information about parameter uncertainty in our model as well as demonstrate the pitfalls and/or advantages of each of the three uncertainty methods (asymptotic theory, bootstrapping, and Bayesian methods). For the stenosis problem (particularly when moving to a two-dimensional setup), the results point to the need for an initial bootstrap or Bayesian estimation in order to ascertain the extent to which nonlinear correlation or non-Gaussian uncertainty is present. However, computational considerations may make this impossible. Our investigation here suggests that if asymptotic theory is the only reasonable computational option, one must fix the corresponding  $E$  and  $G$  parameters and only estimate the  $\gamma_i$  and  $\zeta_i$  values. Of course, these results are restricted to the pressure model, so the corresponding shear results are still a topic for further consideration.

## 7.1 Future Investigation

The avenues for future investigation are quite extensive. After defending this dissertation, the author will be traveling to BICOM to work on the two-dimensional inverse problem for one month. Though this short duration will only allow for some basic consideration, we hope to carefully define that inverse problem and begin a proof-of-concept process similar to the one-dimensional case discussed in Chapter 3. Besides the two-dimensional model, there are extensions of the model and experimental setup that should still be investigated in the one-dimensional case which we will detail shortly. Though we will not discuss it further here, signal processing methods also need to be developed to pick out from the body background noise the appropriate shear waves that correspond with an arterial stenosis.

### 7.1.1 Relative versus Absolute Error for Restricted Parameter Set

The results of Chapter 6 indicated that removing the nonlinearly correlated parameter  $E$  from the estimation process significantly improved the asymptotic results. A logical step would be to return to the comparison of absolute and relative error models when not estimating  $E$ . This would likely provide better information on the standard errors obtained by each theory (though this will not change the residual analysis already conducted).

### 7.1.2 Parameter Correlation

We saw in Chapter 6 that the full set of six parameters in the pressure model exhibited nonlinear correlations which the asymptotic theory was unable to describe. However, it is still an open question as to whether the off-diagonal elements of the covariance matrix  $\hat{\Sigma} = \hat{\sigma}^2[\chi^T(\hat{q})\chi(\hat{q})]^{-1}$  may indicate the extent to which these nonlinear correlations are present. Also, instead of using the linearized theory, one might be able to use the “exact” confidence intervals described in [28, 83] might be able to describe complex parameter relationships in a way similar to the bootstrapping or Bayesian methods. This might allow one to retain the speed and relative simplicity of asymptotic computations while still indicating any complex correlations between parameters. Also, examining ways of re-parameterizing in more complex ways than fixing a parameter is a possibility, as discussed in [27]. As we saw in Chapter 6, reducing nonlinear correlations has a direct effect on the confidence interval computations.

### 7.1.3 Piecewise Model Parameters

One obvious model extension would be using piecewise model parameters. This could be done in a proof-of-concept case, by creating simulated data using piecewise parameters (representing different materials in the domain) and trying to estimate the known parameters. Using data, a

phantom could be created using gel of different properties either vertically (pressure case) or radially (shear case); piecewise parameters could be then estimated. However, this approach requires not only estimating multiple sets of material parameters but also determining where in the spatial domain those variables apply. Since the internal variable framework was designed to incorporate this sort of situation in terms of the bulk effect of the different materials on the overall response (not requiring information about *where* the material properties change, just that they change), the author’s instinct is that examining multiple internal variables is likely to be more productive.

#### 7.1.4 Models with $N_p > 1$ ; Medium Inhomogeneities

Though we examined this situation on the proof-of-concept model in Chapter 3, when we applied the main model to data in Chapter 5 only one relaxation time was used. This was partly due to the fact that the results in Chapter 3 indicated that a second relaxation time made the estimation problem more difficult due to decreased model solution sensitivity to the parameters, and also to the fact that the gel phantom was linear and homogeneous. Since internal variables describe the way different portions of the material respond to stress, and the material was the same throughout, it is not surprising that one relaxation time was adequate.

There are a few ways one could attempt to validate the main model with  $N_p > 1$ . A proof of concept investigation could entail creating simulated data using multiple relaxation times and attempting to estimate those times, as we did for the preliminary model in Chapter 3. However, we saw there that estimating a second relaxation time was difficult. Another option would entail generating simulated data using piecewise model parameters. Then an inverse problem using internal variables could be studied to determine if estimating a second relaxation time (or more) is feasible. This is closer to the problem in reality, where we try to estimate the overall response while taking different materials making up the chest (muscle, lung, bone, etc.) into account without necessarily knowing their exact location. This idea could be easily applied to lab data, where “ribs” or other inhomogeneities could be added to the gel phantom, which would then provide displacement data from that experimental setup for the inverse problem with multiple relaxation times. In all of these cases, one would likely use some version of the Akaike Information Criterion to compare between models, as was done in Section 3.4.1.

#### 7.1.5 Modeling Input into the Body from the Arterial Wall

In Chapter 1 we discussed past research on modeling turbulent blood flow and the resulting impacts on the vessel wall (e.g., [30, 31, 32, 55, 64, 68, 73, 76, 77, 89, 98]) that would be propagated into the body as the pressure and shear waves this dissertation has sought to model. In addition to those works, Chapter 8 of [52] studied the properties of arterial walls. All of these

could be used to provide the basis of a model for the process generating the pressure and shear waves that we have modeled. It is not yet certain exactly what properties of turbulent blood flow and arterial walls are needed to characterize appropriate inputs that generate detectable shear waves. Perhaps our work could proceed with a description of the range of possible inputs, which we could then use Monte-Carlo sampling to propagate through the viscoelastic models in order to obtain the characteristics of displacements one could find on the surface of the chest. The work of [30, 31, 32] provides one guide for how we can use our viscoelastic model to examine the “noise field” in the chest cavity and on the chest surface that results from a stenosed coronary artery.

Our collaborators at QMUL have been examining a rectangular phantom with a plastic pipe incorporated through the center which includes a constriction. This could be used to examine the noises introduced into the phantom from a viscous fluid being pushed past the constriction. The researchers at QMUL have produced some data varying the flow rate and size of the constriction, which changes the level of sounds produced from the constricted flow. This will likely be a subject of discussion when the author visits BICOM.

#### **7.1.6 Larger Phantom and Tissue Experiments**

Future efforts will also involve scaling up all these experiments to larger phantoms and then to some sort of actual tissue sample experiments. A key point of investigation with a larger phantom is the attenuation of the shear signals from the point of generation to the surface of the material. Our current one-dimensional model should provide some guidance on the level of signal attenuation, which we can then test on the actual phantom or tissue. Some of the work discussed in Chapter 1 did examine properties of tissue, so their work would provide a starting point when designing an experiment for examining the response of actual tissue. The work of [67, 80] was intended to be applied to pigs since their chest cavity is similar to humans. After larger gel phantoms, and possibly some experiments with muscle tissue, pigs may be the next logical focus for studying wave propagation.

#### **7.1.7 Two-dimensional Model**

As mentioned, the author will be traveling to BICOM to examine the inverse problem with a two-dimensional model and (ideally) corresponding experimental configurations. Experiments are currently being designed to produce a two-dimensional wave from different points in the medium and with different detection points along the outer wall of the phantom. It is conceivable that the one-dimensional parameters could be used as a rough first approximation in a corresponding two-dimensional code, which would allow us to focus on trying to determine the location of the wave generation in the medium. Also, these parameter values could be used in a model of



wave propagation in another conceptual device designed to mimic a constricted artery and the waves that result from passing fluid through a constricted pipe in the center of the medium. Therefore in the slightly longer term, we will also likely need to conduct an inverse problem using a two-dimensional model and corresponding data. These one-dimensional results will provide a starting point for parameters in that inverse problem, hopefully decreasing runtime and the time it takes to find viable parameters. The same issues discussed here (sensitivity to parameters, data frequency, number of relaxation times) will again be of concern for the two-dimensional problem.

Mathematically, an existence, uniqueness, and continuous dependence framework needs to be developed for any two-dimensional model. The derivation in Chapter 4 in a strain-energy function framework should be amenable to the work on mathematical questions. The work in [22] may provide a starting point.

## REFERENCES

- [1] M. Ainsworth, P. Monk, and W. Muniz, Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation, *Journal of Scientific Computing*, **27** (2006), 5–40.
- [2] M. Akay, *Noninvasive Detection of Coronary Artery Disease Using Advanced Signal Processing Methods*, PhD Dissertation, Rutgers University, Piscataway, NJ, 1990.
- [3] M. Akay, Y. Akay, W. Welkowitz, J. Semmlow, and J. Kostis, Application of adaptive filters to noninvasive acoustical detection of coronary occlusions before and after angioplasty, *IEEE Trans. on Biomed. Eng.*, **39** (1992), 176–184.
- [4] M. Akay, Y. Akay, W. Welkowitz, J. Semmlow, and J. Kostis, Noninvasive detection of coronary artery disease using neural networks, *Proc. IEEE Eng. in Med. & Biol. Soc.*, (1991), 1434–1435.
- [5] M. Akay, M. Bauer, J. Semmlow, W. Welkowitz, and J. Kostis, Analysis of diastolic heart sounds before and after angioplasty, *Proc. IEEE Eng. in Med. & Biol. Soc.*, (1988), 257–260.
- [6] M. Akay, W. Welkowitz, J. Semmlow, and J. Kostis, Application of the ARMA method to acoustic detection of coronary artery diseases, *Med. & Biol. Eng. & Comput.*, **29** (1991), 365–372.
- [7] C. Andrieu and J. Thoms, A tutorial on adaptive MCMC, *Stat. and Comp.*, **18** (2008), 343–373.
- [8] H.T. Banks, *A Functional Analysis Framework for Modeling, Estimation and Control in Science and Engineering*, CRC Press/Taylor-Francis, Boca Raton London New York, 2012.
- [9] H.T. Banks, J.H. Barnes, A. Eberhardt, H. Tran, and S. Wynne, Modeling and computation of propagating waves from coronary stenoses, *Comput. and Appl. Math.*, **21** (2002), 767–788.
- [10] H.T. Banks and K.L. Bihari, Modeling and estimating uncertainty in parameter estimation, *Inverse Problems*, **17** (2001), 1–17.
- [11] H.T. Banks and B.G. Fitzpatrick, Inverse problems for distributed systems: statistical tests and ANOVA, LCDS/CCS Rep. 88-16, Brown University, July, 1988; *Proc. International Symposium on Math. Approaches to Envir. and Ecol. Problems*, Springer Lecture Notes in Biomath., **81** (1989), 262–273.
- [12] H.T. Banks, K. Holm and D. Robbins, Standard error computations for uncertainty quantification in inverse problems: Asymptotic theory vs. Bootstrapping, CRSC-TR09-13, North Carolina State University, June 2009; Revised May 2010; *Math. and Comp. Model.*, **52** (2010), 1610–1625.
- [13] H.T. Banks, S. Hu and Z.R. Kenz, A brief review of elasticity and viscoelasticity for solids, *Adv. in Appl. Math. and Mech.*, **3** (2011), 1–51.

- [14] H.T. Banks, S. Hu, Z.R. Kenz, C. Kruse, S. Shaw, J.R. Whiteman, M.P. Brewin, S.E. Greenwald, and M.J. Birch, Material parameter estimation and hypothesis testing on a 1D viscoelastic stenosis model: Methodology, CRSC-TR-12-09, Center for Research in Scientific Computation, North Carolina State University, April, 2012; *J. Inverse and Ill-Posed Problems*, **21** (2013), 25–57.
- [15] H.T. Banks, S. Hu, Z.R. Kenz, C. Kruse, S. Shaw, J.R. Whiteman, M.P. Brewin, S.E. Greenwald and M.J. Birch, Model validation for a noninvasive arterial stenosis detection problem, CRSC-TR12-22, Center for Research in Scientific Computation, North Carolina State University, December, 2012; *Math. Biosci. and Eng.*, to appear.
- [16] H.T. Banks, Z.R. Kenz and W.C. Thompson, An extension of RSS-based model comparison tests for weighted least squares, CRSC-TR12-18, Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC, August, 2012; *Intl. J. Pure and Appl. Math.*, **79** (2012), 155–183.
- [17] H.T. Banks, Z.R. Kenz and W. Clayton Thompson, A review of selected techniques in inverse problem nonparametric probability distribution estimation, CRSC-TR12-13, Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC, May, 2012; *J. Inverse and Ill-posed Problems*, **20** (2012), 429–460.
- [18] H.T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, Boston, 1989.
- [19] H.T. Banks and N. Luke, Modeling of propagating shear waves in biotissue employing an internal variable approach to dissipation, *Comm. in Comp. Phys.*, **3** (2008), 603–640.
- [20] H.T. Banks, N. Medhin, and G. Pinter, Multiscale considerations in modeling of nonlinear elastomers, *Int. J. for Comp. Meth. in Eng. Sci. and Mech.*, **8** (2007), 53–62.
- [21] H.T. Banks, N. Medhin, and G. Pinter, Nonlinear reptation in molecular based hysteresis models for polymers, *Quart. of Applied Math.*, **62** (2004), 767–779.
- [22] H.T. Banks and G.A. Pinter, A probabilistic multiscale approach to hysteresis in shear wave propagation in biotissue, *Multiscale Model. and Simul.*, **3** (2005), 395–412.
- [23] H.T. Banks and J.R. Samuels, Jr., Detection of cardiac occlusions using viscoelastic wave propagation, *Adv. in Appl. Math. and Mech.*, **1** (2009), 1–28.
- [24] H.T. Banks and H.T. Tran, *Mathematical and Experimental Modeling of Physical and Biological Processes*, CRC Press, Boca Raton, FL, 2009.
- [25] H.T. Banks, H. Tran and S. Wynne, A well-posedness result for a shear wave propagation model, *Intl. Series Num. Math.*, **143** (2002), Birkhäuser Verlag, Basel, 25–40.
- [26] J.D. De Basabe, M.K. Sen, and M.F. Wheeler, The interior penalty discontinuous Galerkin method for elastic wave propagation: grid dispersion, *Geophys J. Int.*, **175** (2008), 83–93.
- [27] D.M. Bates and D.G. Watts, Parameter transformations for improved approximate confidence regions in nonlinear least squares, *The Annals of Sci.*, **9** (1981), 1152–1167.

- [28] E.M.L. Beale, Confidence regions in non-linear estimation, *J. of the Royal Stat. Soc. B*, **22** (1960), 41–88.
- [29] A.P. Boresi and R.J. Schmidt, *Advanced Mechanics of Materials*, John Wiley & Sons, New York, 2003.
- [30] A.O. Borisyuk, Noise field in the human chest due to turbulent flow in a larger blood vessel, *Flow, Turbulence and Combust.*, **61** (1999), 269–284.
- [31] A.O. Borisyuk, Experimental study of voice produced by steady flow through a simulated vascular stenosis, *J. of Sound and Vib.*, **256.3** (2002), 475–498.
- [32] A.O. Borisyuk, Model study of noise field in the human chest due to turbulent flow in a larger blood vessel, *J. of Fluids and Struct.*, **17** (2003), 1095–1110.
- [33] Z. Botev, Kernel Density Estimator, <http://www.mathworks.com/matlabcentral/fileexchange/14034>.
- [34] M.P. Brewin, M.J. Birch, S.E. Greenwald, et. al., Characterization of the uniaxial elastic properties of an agar-based tissue mimicking material, *in preparation*.
- [35] K.P. Burnham and D.R. Anderson, *Model Selection and Inference: A Practical Information-Theoretical Approach*, Springer-Verlag, New York, 1998.
- [36] S. Catheline, J.-L. Gennisson, G. Delon, M. Fink, R. Sinkus, S. Abouelkaram, and J. Culioli, Measurement of viscoelastic properties of homogeneous soft solid using transient elastography: an inverse problem approach, *J. Acoust. Soc. Am*, **116.6** (2004), 3734–3741.
- [37] S. Catheline, L. Sandrin, J.-L. Gennisson, M. Tanter, and M. Fink, Ultrasound-based noninvasive shear elasticity probe for soft tissues, *IEEE Ultrason. Symp.*, **2** (2000), 1799–1801.
- [38] S. Chen, M. Fatemi, and J. Greenleaf, Quantifying elasticity and viscosity from measurement of shear wave speed dispersion, *J. Acoust. Soc. Am.*, **115.6** (2004), 2781–2785.
- [39] S. Chen, M. Urban, C. Pislaru, R. Kinnick, Y. Zheng, A. Yao, and J. Greenleaf, Shearwave dispersion ultrasound vibrometry (SDUV) for measuring tissue elasticity and viscosity, *IEEE Trans. on Ultrason., Ferroelectrics, and Freq. Contr.*, **56** (2009), 55–62.
- [40] T. Cheng, Diastolic murmur caused by coronary artery stenosis, *Ann. Int. Med*, **72** (1970), 543.
- [41] B.D. Coleman and W. Noll, Foundations of linear viscoelasticity, *Reviews of Modern Physics*, **33** (1961), 239–249.
- [42] P. Constantin and C. Foias, *Navier-Stokes Equations*, University of Chicago Press, Chicago, 1988.
- [43] M. Davidian and D. Giltinan, *Nonlinear Models for Repeated Measurement Data*, Chapman and Hall, London, 1995.

- [44] T. Deffieux, G. Montaldo, and M. Tanter, Shear wave spectroscopy for in vivo quantification of human soft tissues visco-elasticity, *IEEE Trans. on Medical Imag.*, **28.3** (2009), 313–322.
- [45] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, New York, 1993.
- [46] B. Efron and R. Tibshirani, Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Stat. Sci.*, **1.1** (1986), 54–75.
- [47] B. El-Asir, L. Khadra, A. Al-Abbasi, and M. Mohammed, Time-frequency analysis of heart sounds, *Proc. IEEE TENCON*, (1996), 553–558.
- [48] M. Fabrizio and A. Morro, *Mathematical Problems in Linear Viscoelasticity*, Studies in Applied Mathematics, SIAM, Philadelphia, 1992.
- [49] J.D. Ferry, *Viscoelastic Properties of Polymers*, John Wiley and Sons Inc., New York, 1970.
- [50] W.N. Findley, J.S. Lai, and K. Onaran, *Creep and Relaxation of Nonlinear Viscoelastic Materials (With an Introduction to Linear Viscoelasticity)*, Dover Publications Inc., New York, 1989.
- [51] A. Friedman, *Foundations of Modern Analysis*, Dover Publications, Inc., New York, 1982.
- [52] Y.C. Fung, *Biomechanics: Mechanical Properties of Living Tissues*, Springer-Verlag, New York, 1993.
- [53] Y.C. Fung, *A First Course in Continuum Mechanics*, Prentice Hall, New Jersey, 1994.
- [54] J.M. Golen and G.A.C. Graham, *Boundary Value Problems in Linear Viscoelasticity*, Springer-Verlag, New York, 1988.
- [55] A. Góral-Wójcicka, W. Borgiel, Z. Małota, and Z. Nawrat, One the acoustic phenomena produced by turbulence in the flowing blood, *Polish J. Med. Phys. & Eng.*, **8.1** (2002), 29–35.
- [56] H. Haario, E. Saksman, and J. Tamminen, An adaptive Metropolis algorithm, *Bernoulli*, **7.2** (2001), 223–242.
- [57] H. Haario, M. Laine, A. Mira, and E. Saksman, DRAM: Efficient adaptive MCMC, *Stat. and Comp.*, **16.4** (2006), 339–354.
- [58] J. Kaipio and E. Somersalo, *Stat. and Comp. Inv. Prob.*, Springer, New York, 2005.
- [59] A. Karpouk, S. Alglyamov, Y. Illinskii, E. Zabolotskaya, and S. Emelianov, Assessment of shear modulus of tissue using ultrasound radiation force acting on a spherical acoustic inhomogeneity, *IEEE Trans. on Ultrason., Ferroelectrics, and Freq. Contr.*, **56** (2009), 2380–2387.

- [60] Z.R. Kenz, H.T. Banks, and R.C. Smith, Comparison of frequentist and Bayesian confidence analysis methods on a viscoelastic stenosis model, CRSC-TR13-05, Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC, April 2013; *SIAM J. on Uncert. Quant.*, submitted.
- [61] C. Kruse, M. Maischak, S. Shaw, J.R. Whiteman, et al., High order space-time finite element schemes for acoustic and viscodynamic wave equations with temporal decoupling, BICOM Technical Report 12/3, 2012; in preparation.
- [62] M. Laine, MCMC toolbox for MATLAB, <http://helios.fmi.fi/~lainema/mcmc/>.
- [63] R. Lakes, *Viscoelastic Materials*, Cambridge Univ. Press, Cambridge, 2009.
- [64] T.S. Lee, W. Liao, and H.T. Low, Numerical simulation of turbulent flow through series stenoses, *Inter. J. for Numer. Meth. in Fluids*, **42** (2003), 717–740.
- [65] S. Levinson, M. Shinagawa, and T. Sato, Sonoelastic determination of human skeletal muscle elasticity, *J. Biomech.*, **28.10** (1995), 1145–1154.
- [66] A.E.H. Love, *A Treatise on the Mathematical Theory of Elasticity*, Cambridge Univ. Press, London, 1927.
- [67] N. Luke, *Modeling Shear Wave Propagation in Biotissue: An Internal Variable Approach to Dissipation*, PhD Dissertation, NCSU, Raleigh, 2006.
- [68] S. Lundin, R. Metcalf, C. Hartley, Effects of severity and eccentricity of carotid stenosis on pulsatile blood flow, *Proc. Joint EMBS/BMES*, (2003), 1311–1312.
- [69] J. E. Marsden and T. J. R. Hughes, *Mathematical Foundations of Elasticity*, Dover Publications, Inc., Mineola, NY, 1994.
- [70] S.E. Nissen, Application of intravascular ultrasound to characterize coronary artery disease and assess the progression or regression of atherosclerosis, *Am. J. Cardiol.*, **89** (2002), 24B–31B.
- [71] S.E. Nissen and P. Yock, Intravascular ultrasound: novel pathophysiological insights and current clinical applications, *Circulation*, **103** (2001), 604–616.
- [72] R.W. Ogden, *Non-Linear Elastic Deformations*, Dover Publications, 1984
- [73] N.L. Owsley and A.J. Hull, Beamformed nearfield imaging of a simulated coronary artery containing a stenosis, *IEEE Trans. on Med. Imag.*, **17** (1998), 900–909.
- [74] N.L. Owsley, A.J. Hull, M.H. Ahmed, and J. Kassal, A proof of concept experiment for the detection of occluded coronary arteries using array sensor technology, *Engr. in Medicine and Biol. Society*, IEEE 17th Annual Conf., **1** (1995), 145–146.
- [75] V. Padmanabhan and J. Semmlow, A dedicated system for acoustic detection of coronary artery disease, *Proc. Eng. in Med. & Biol. Soc.*, (1992), 457–458.

- [76] T. Pedley, Mathematical modelling of arterial fluid dynamics, *J. of Eng. Math.*, **47** (2003), 419–444.
- [77] C. Prado, S. Ramos, J. Elias, Jr., and M. Rossi, Turbulent blood flow plays an essential localizing role in the development of atherosclerotic lesions in experimentally induced hypercholesterolaemia in rats, *Int. J. Exp. Path.*, **89** (2008), 72–80.
- [78] Y.V. Prohorov, Convergence of random processes and limit theorems in probability theory, *Theor. Prob. Appl.*, **1** (1956), 157–214.
- [79] G.O. Roberts and J.S. Rosenthal, Examples of adaptive MCMC, *Journal of Computational and Graphical Statistics*, **18.2** (2009), 349–367.
- [80] J.R. Samuels, Jr., *Inverse Problems and Post Analysis Techniques for a Stenosis-driven Acoustic Wave Propagation Model*, PhD Dissertation, NCSU, Raleigh, 2008.
- [81] L. Sandrin, S. Catheline, M. Tanter, X. Hennequin, and M. Fink, Time-resolved pulsed elastography with ultrafast ultrasonic imaging, *Ultrason. Imag.*, **21** (1999), 259–272.
- [82] J. Sangster and C. Oakley, Diastolic murmur of coronary artery stenosis, *Brit. Heart J.*, **35** (1973), 840–844.
- [83] G. Seber and C. Wild, *Nonlinear Regression*, J. Wiley & Sons, Hoboken, NJ, 2003.
- [84] J. Semmlow and K. Rahalkar, Acoustic detection of coronary artery disease, *Annu. Rev. Biomed. Eng.*, **9** (2007), 449–469.
- [85] J. Semmlow, J. Welkowitz, W. Kostis, and J.W. Mackenzie, Coronary artery disease – correlates between diastolic auditory characteristics and coronary artery stenoses, *IEEE Trans. on Biomed. Engr.*, **30** (1983), 136–139.
- [86] R. Smith, *Uncertainty Quantification for Predictive Estimation*, SIAM, in preparation.
- [87] A. Solonen, *Monte Carlo Methods in Parameter Estimation of Nonlinear Models*, MS Thesis, Lappeenranta University of Technology, Lappeenranta, Finland, 2006.
- [88] A. Solonen, P. Ollinaho, M. Laine, H. Haario, J. Tamminen and H. Järvinen, “Efficient MCMC for climate model parameter estimation: parallel adaptive chains and early rejection,” *Bayesian Analysis*, **7.3** (2012), 715–736.
- [89] C. Taylor, T. Hughes, and C. Zarins, Finite element modeling of three-dimensional pulsatile flow in the abdominal aorta: relevance to atherosclerosis, *Annals of Biomed. Eng.*, **26** (1998), 975–987.
- [90] C.J.F. Ter Braak, A Markov chain Monte Carlo version of the genetic algorithm differential evolution: easy Bayesian computing for real parameter spaces, *Stat. and Comp.*, **16.3** (2006), 239–249.
- [91] J. Verburg, Transmission of vibrations of the heart to the chestwall, *Adv. Cardiovasc. Phys.*, **5** (1983), 84–103.

- [92] J. Verburg and E. van Vollenhoven, Phonocardiography: Physical and technical aspects and clinical uses, in *Noninvasive Physiological Measurements* (ed. P Rolfe), Academic Press, London (1979).
- [93] H. Vermarien and E. van Vollenhoven, The recording of heart vibrations: a problem of vibration measurement on soft tissue, *Med. & Biol. Eng. & Comput.*, **22** (1984), 168–178.
- [94] M. Vihola, Robust adaptive Metropolis algorithm with coerced acceptance rate, *Statistics and Computing*, **22.5** (2012), 997–1008.
- [95] A. Voss, A. Mix, and T. Hübner, Diagnosing aortic valve stenosis by parameter extraction of heart sound signals, *Annals of Biomed. Eng.*, **33.9** (2005), 1167–1174.
- [96] J.A. Vrugt and C.J.F. Ter Braak, DREAM<sub>(D)</sub>: an adaptive Markov Chain Monte Carlo simulation algorithm to solve discrete, noncontinuous, and combinatorial posterior parameter estimation problems, *Hydrology and Earth Sys. Sci.*, **15** (2011), 3701–3713.
- [97] J.A. Vrugt, C.J.F. Ter Braak, C.G.H. Diks, B.A. Robinson, J.M. Hyman, and D. Higdon, Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *Int. J. of Nonlinear Sci. and Numer. Sim.*, **10.3** (2009), 273–290.
- [98] J.-Z. Wang, B. Tie, W. Welkowitz, J. Semmlow, and J. Kostis, Modeling sound generation in stenosed coronary arteries, *IEEE Trans. on Biomed. Eng.*, **37.11** (1990), 1087–1094.
- [99] J. Wloka, *Partial Differential Equations*, Cambridge University Press, Cambridge, 1987.
- [100] M. Zia, B. Griffel, V. Fridman, C. Saponieri, and J. Semmlow, Noise detection in heart sound recordings, *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, (2011), 5880–5883.



## APPENDICES

# APPENDIX A

---

## Sensitivity Equations for Proof of Concept Model

---

In this appendix, we derive the sensitivity equations used in Chapter 3. Let

$$\bar{q} = (\rho, E_0, E_1, \tau_1, \dots, \tau_{N_p}, p_1, \dots, p_{N_p}, A)^T$$

be the vector of length  $\kappa$  for the parameters needed to be estimated, and  $s^{\bar{q}_i}(t, x; \bar{q}) = \frac{\partial u(t, x; \bar{q})}{\partial \bar{q}_i}$ ,  $i = 1, 2, \dots, \kappa$  with  $\bar{q}_i$  being the  $i$ th component of  $\bar{q}$ . We first find the equation for  $s^\rho(t, x; \bar{q})$ . To do that, we take the partial derivatives of all parts of (3.1.1) with respect to  $\rho$  and obtain

$$\begin{aligned} \rho u_{\rho tt} + u_{tt} - \sigma_{\rho x} &= 0 \\ u_\rho(t, 0) &= 0, \quad \sigma_\rho(t, L) = 0, \\ u_\rho(0, x) &= 0, \quad u_{\rho t}(0, x) = 0. \end{aligned} \tag{A.0.1}$$

(Note that the chain rule on the first term resulted in  $\frac{\partial}{\partial \rho}(\rho u_{tt}) = \rho u_{\rho tt} + u_{tt}$ , which means the sensitivity partial differential equation (PDE) will be driven by the original system values.) By changing the order of differentiation and then substituting  $s^\rho$  for  $\frac{\partial u}{\partial \rho}$ , equation (A.0.1) can be rewritten as

$$\begin{aligned} \rho(s^\rho)_{tt} - \frac{\partial}{\partial x}\sigma_\rho &= -u_{tt} \\ s^\rho(t, 0) &= 0, \quad \sigma_\rho(t, L) = 0, \\ s^\rho(0, x) &= 0, \quad (s^\rho)_t(0, x) = 0, \end{aligned} \tag{A.0.2}$$

In the above equation, the sensitivity of stress with respect to  $\rho$  (i.e.,  $\sigma_\rho$ ) can be obtained by differentiating both sides of (3.1.8) with respect to  $\rho$

$$\sigma_\rho = E_1(s^\rho)_{xt} + E_0 \left( (s^\rho)_x - \sum_{j=1}^{N_p} \epsilon_\rho^j \right). \quad (\text{A.0.3})$$

Here the sensitivity of internal variable  $\epsilon^j$  with respect to  $\rho$  (i.e.,  $\epsilon_\rho^j$ ) satisfies the following equation

$$\begin{aligned} \tau_j(\epsilon_\rho^j)_t + \epsilon_\rho^j &= p_j(s^\rho)_x, \\ \epsilon_\rho^j(0) &= 0, \end{aligned} \quad (\text{A.0.4})$$

where  $j = 1, \dots, N_p$ . The above equation is obtained by differentiating both sides of (3.1.7) with respect to  $\rho$ . Thus, we see the sensitivity PDE (A.0.2)-(A.0.4) for  $\rho$  has the same form as the original system PDE (3.1.1)-(3.1.7), except with zero initial/boundary conditions and now being driven by the solution to the original system PDE. Thus, the sensitivity PDE is coupled to the original system PDE.

We now state the remaining sensitivity PDEs. We highlighted terms in red that correspond with values from the solution to the original system PDE.

#### 1. Sensitivity PDE for $E_0$ :

Note that solving this sensitivity PDE requires knowledge of the internal variables, in particular the first derivative with respect to  $x$  (i.e., we need to know  $\epsilon_x^j$  since we have the  $(\sigma_{E_0})_x$  term in (A.0.5), and  $\sigma_{E_0}$  incorporates  $\epsilon^j$ ).

$$\begin{aligned} \rho(s^{E_0})_{tt} - (\sigma_{E_0})_x &= 0 \\ s^{E_0}(t, 0) &= 0, \quad \sigma_{E_0}(t, L) = 0, \\ s^{E_0}(0, x) &= 0, \quad (s^{E_0})_t(0, x) = 0. \end{aligned} \quad (\text{A.0.5})$$

In the above equation, the sensitivity of stress with respect to  $E_0$  (i.e.,  $\sigma_{E_0}$ ) is given by

$$\sigma_{E_0} = E_1(s^{E_0})_{xt} + E_0 \left( (s^{E_0})_x - \sum_{j=1}^{N_p} \epsilon_{E_0}^j \right) + \left( u_x - \sum_{j=1}^{N_p} \epsilon^j \right). \quad (\text{A.0.6})$$

Here the sensitivity of internal variables  $\epsilon^j$  with respect to  $E_0$  (i.e.,  $\epsilon_{E_0}^j$ ) satisfies the following equation

$$\begin{aligned} \tau_n(\epsilon_{E_0}^j)_t + \epsilon_{E_0}^j &= p_j(s^{E_0})_x, \\ \epsilon_{E_0}^j(0) &= 0, \end{aligned} \quad (\text{A.0.7})$$

where  $j = 1, \dots, N_p$ .

#### 2. Sensitivity PDE for $E_1$ :

$$\begin{aligned}
\rho(s^{E_1})_{tt} - (\sigma_{E_1})_x &= 0 \\
s^{E_1}(t, 0) &= 0, \quad \sigma_{E_1}(t, L) = 0, \\
s^{E_1}(0, x) &= 0, \quad (s^{E_1})_t(0, x) = 0.
\end{aligned} \tag{A.0.8}$$

In the above equation, the sensitivity of stress with respect to  $E_1$  (i.e.,  $\sigma_{E_1}$ ) is given by

$$\sigma_{E_1} = E_1(s^{E_1})_{xt} + E_0 \left( (s^{E_1})_x - \sum_{j=1}^{N_p} \epsilon_{E_1}^j \right) + u_{xt}. \tag{A.0.9}$$

Here the sensitivity of internal variables  $\epsilon^j$  with respect to  $E_1$  (i.e.,  $\epsilon_{E_1}^j$ ) satisfies the following equation

$$\begin{aligned}
\tau_j(\epsilon_{E_1}^j)_t + \epsilon_{E_1}^j &= p_j(s^{E_1})_x, \\
\epsilon_{E_1}^j(0) &= 0,
\end{aligned} \tag{A.0.10}$$

where  $j = 1, \dots, N_p$ .

3. Sensitivity PDE for  $\tau_k$ ,  $k = 1, \dots, N_p$ :

$$\begin{aligned}
\rho(s^{\tau_k})_{tt} - (\sigma_{\tau_k})_x &= 0 \\
s^{\tau_k}(t, 0) &= 0, \quad \sigma_{\tau_k}(t, L) = 0, \\
s^{\tau_k}(0, x) &= 0, \quad (s^{\tau_k})_t(0, x) = 0.
\end{aligned} \tag{A.0.11}$$

In the above equation, the sensitivity of stress with respect to  $\tau_k$  (i.e.,  $\sigma_{\tau_k}$ ) is given by

$$\sigma_{\tau_k} = E_1(s^{\tau_k})_{xt} + E_0 \left( (s^{\tau_k})_x - \sum_{j=1}^{N_p} \epsilon_{\tau_k}^j \right). \tag{A.0.12}$$

Here the sensitivity of internal variables  $\epsilon^j$  with respect to  $\tau_k$  (i.e.,  $\epsilon_{\tau_k}^j$ ) is given as follows:

- if  $j = k$ , then we have

$$\begin{aligned}
\tau_k(\epsilon_{\tau_k}^k)_t + \epsilon_{\tau_k}^k &= p_k(s^{\tau_k})_x, \\
\epsilon_{\tau_k}^k(0) &= 0.
\end{aligned} \tag{A.0.13}$$

- if  $j \neq k$ , then we have

$$\begin{aligned}
\tau_j(\epsilon_{\tau_k}^j)_t + \epsilon_{\tau_k}^j &= p_j(s^{\tau_k})_x, \\
\epsilon_{\tau_k}^j(0) &= 0.
\end{aligned} \tag{A.0.14}$$

4. Sensitivity PDE for  $p_k$ ,  $k = 1, \dots, N_p$ :

$$\begin{aligned}
\rho(s^{p_k})_{tt} - (\sigma_{p_k})_x &= 0 \\
s^{p_k}(t, 0) &= 0, \quad \sigma_{p_k}(t, L) = 0, \\
s^{p_k}(0, x) &= 0, \quad (s^{p_k})_t(0, x) = 0.
\end{aligned} \tag{A.0.15}$$

In the above equation, the sensitivity of stress with respect to  $p_k$  (i.e.,  $\sigma_{p_k}$ ) is given by

$$\sigma_{p_k} = E_1(s^{p_k})_{xt} + E_0 \left( (s^{p_k})_x - \sum_{j=1}^{N_p} \epsilon_{p_k}^j \right). \tag{A.0.16}$$

Here the sensitivity of internal variables  $\epsilon^j$  with respect to  $p_k$  (i.e.,  $\epsilon_{p_k}^j$ ) is given as follows:

- if  $j = k$ , then we have

$$\begin{aligned}
\tau_k(\epsilon_{p_k}^k)_t + \epsilon_{p_k}^k &= p_k(s^{p_k})_x + u_x, \\
\epsilon_{p_k}^k(0) &= 0.
\end{aligned} \tag{A.0.17}$$

- if  $j \neq k$ , then we have

$$\begin{aligned}
\tau_j(\epsilon_{p_k}^j)_t + \epsilon_{p_k}^j &= p_j(s^{p_k})_x, \\
\epsilon_{p_k}^j(0) &= 0.
\end{aligned} \tag{A.0.18}$$

## 5. Sensitivity PDE for $A$ :

To some extent, this is an “outlier” as compared with the preceding sensitivity PDEs. The coefficient  $A$  only appears in the (right) boundary condition and not explicitly in the PDE itself; thus, the sensitivity PDE for  $A$  will be nearly the same as the original system PDE except with a different (right) boundary condition. The sensitivity equation is given as follows.

$$\begin{aligned}
\rho(s^A)_{tt} - (\sigma_A)_x &= 0, \\
s^A(t, 0) &= 0, \quad \sigma_A(t, L) = g_A(t), \\
s^A(0, x) &= 0, \quad (s^A)_t(0, x) = 0.
\end{aligned} \tag{A.0.19}$$

In the above equation,  $g_A$  is given by

$$g_A(t) = \begin{cases} \exp\left(\frac{|ab|}{t(t+a-b)}\right) & \text{if } t \in (0, b-a) \\ 0 & \text{otherwise} \end{cases}.$$

The sensitivity of stress with respect to  $A$  (i.e.,  $\sigma_A$ ) is given by

$$\sigma_A = E_1(s^A)_{xt} + E_0 \left( (s^A)_x - \sum_{j=1}^{N_p} \epsilon_A^j \right). \tag{A.0.20}$$

Here the sensitivity of internal variables  $\epsilon^j$  with respect to  $A$  (i.e.,  $\epsilon_A^j$ ) is given by

$$\begin{aligned}\tau_j(\epsilon_A^j)_t + \epsilon_A^j &= p_j(s^A)_x, \\ \epsilon_A^j(0) &= 0,\end{aligned}\tag{A.0.21}$$

where  $j = 1, \dots, N_p$ .

Based on the above discussion, we see that all the sensitivity equations are coupled to the original system except the one for  $A$ . However, it is important to note that the sensitivity PDEs are *not* coupled to each other.

---

## Sensitivity Equations for Main Stenosis Wave Propagation Model

---

### B.1 Pressure Model

We derive the sensitivity equations for the pressure model 4.4.13.

1. Sensitivity PDE for  $\rho$  (where  $s^\rho = \frac{\partial u}{\partial \rho}$ ):

$$\begin{aligned} \rho(s^\rho)_{tt} - (\sigma_\rho)_x &= -u_{tt} \\ s^\rho(0, t) &= 0, \quad \sigma_\rho(L, t) = 0, \\ s^\rho(x, 0) &= 0, \quad (s^\rho)_t(x, 0) = 0, \end{aligned} \tag{B.1.1a}$$

In the above equation, the sensitivity of stress with respect to  $\rho$  (i.e.,  $\sigma_\rho$ ) is given by

$$\sigma_\rho = E_1(s^\rho)_{xt} + \left(E + \sum_{j=1}^{N_p} \gamma_j\right) (s^\rho)_x - \sum_{j=1}^{N_p} \gamma_j \epsilon_\rho^j. \tag{B.1.1b}$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $\rho$  (i.e.,  $\epsilon_\rho^j = \frac{\partial \epsilon^j}{\partial \rho}$ ) satisfies (for  $j = 1, \dots, N_p$ ).

$$\tau_j(\epsilon_\rho^j)_t + \epsilon_\rho^j = (s^\rho)_x, \quad \epsilon_\rho^j(0) = 0. \tag{B.1.1c}$$

2. Sensitivity PDE for  $E$  (where  $s^E = \frac{\partial u}{\partial E}$ ):

$$\begin{aligned}\rho(s^E)_{tt} - (\sigma_E)_x &= 0 \\ s^E(0, t) &= 0, \quad \sigma_E(L, t) = 0, \\ s^E(x, 0) &= 0, \quad (s^E)_t(x, 0) = 0,\end{aligned}\tag{B.1.2a}$$

In the above equation, the sensitivity of stress with respect to  $E$  (i.e.,  $\sigma_E$ ) is given by

$$\sigma_E = E_1(s^E)_{xt} + \left(E + \sum_{j=1}^{N_p} \gamma_j\right) (s^E)_x - \sum_{j=1}^{N_p} \gamma_j \epsilon_E^j + u_x.\tag{B.1.2b}$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $E$  (i.e.,  $\epsilon_E^j = \frac{\partial \epsilon^j}{\partial E}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

$$\tau_j(\epsilon_E^j)_t + \epsilon_E^j = (s^E)_x, \quad \epsilon_E^j(0) = 0.\tag{B.1.2c}$$

3. Sensitivity PDE for  $E_1$  (where  $s^{E_1} = \frac{\partial u}{\partial E_1}$ ):

$$\begin{aligned}\rho(s^{E_1})_{tt} - (\sigma_{E_1})_x &= 0 \\ s^{E_1}(0, t) &= 0, \quad \sigma_{E_1}(L, t) = 0, \\ s^{E_1}(x, 0) &= 0, \quad (s^{E_1})_t(x, 0) = 0,\end{aligned}\tag{B.1.3a}$$

In the above equation, the sensitivity of stress with respect to  $E_1$  (i.e.,  $\sigma_{E_1}$ ) is given by

$$\sigma_{E_1} = E_1(s^{E_1})_{xt} + \left(E + \sum_{j=1}^{N_p} \gamma_j\right) (s^{E_1})_x - \sum_{j=1}^{N_p} \gamma_j \epsilon_{E_1}^j + u_{xt}.\tag{B.1.3b}$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $E_1$  (i.e.,  $\epsilon_{E_1}^j = \frac{\partial \epsilon^j}{\partial E_1}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

$$\tau_j(\epsilon_{E_1}^j)_t + \epsilon_{E_1}^j = (s^{E_1})_x, \quad \epsilon_{E_1}^j(0) = 0.\tag{B.1.3c}$$

4. Sensitivity PDE for  $\tau_k$  (where  $s^{\tau_k} = \frac{\partial u}{\partial \tau_k}$ ):

$$\begin{aligned}\rho(s^{\tau_k})_{tt} - (\sigma_{\tau_k})_x &= 0 \\ s^{\tau_k}(0, t) &= 0, \quad \sigma_{\tau_k}(L, t) = \frac{A}{L} \gamma_k \left( \frac{t}{\tau_k^2} e^{-t/\tau_k} - \frac{t-\Upsilon}{\tau_k^2} e^{-(t-\Upsilon)/\tau_k} \right), \\ s^{\tau_k}(x, 0) &= 0, \quad (s^{\tau_k})_t(x, 0) = 0,\end{aligned}\tag{B.1.4a}$$



In the above equation, the sensitivity of stress with respect to  $\tau_k$  (i.e.,  $\sigma_{\tau_k}$ ) is given by

$$\sigma_{\tau_k} = E_1(s^{\tau_k})_{xt} + \left(E + \sum_{j=1}^{N_p} \gamma_j\right) (s^{\tau_k})_x - \sum_{j=1}^{N_p} \gamma_j \epsilon_{\tau_k}^j. \quad (\text{B.1.4b})$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $\tau_k$  (i.e.,  $\epsilon_{\tau_k}^j = \frac{\partial \epsilon^j}{\partial \tau_k}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

- if  $j = k$  then

$$\tau_k(\epsilon_{\tau_k}^k)_t + \epsilon_{\tau_k}^k = (s^{\tau_k})_x - \epsilon_t^k, \quad \epsilon_{\tau_k}^k(0) = 0, \quad (\text{B.1.4c})$$

- if  $j \neq k$  then

$$\tau_j(\epsilon_{\tau_k}^j)_t + \epsilon_{\tau_k}^j = (s^{\tau_k})_x, \quad \epsilon_{\tau_k}^j(0) = 0. \quad (\text{B.1.4d})$$

5. Sensitivity PDE for  $\gamma_k$  (where  $s^{\gamma_k} = \frac{\partial u}{\partial \gamma_k}$ ):

$$\begin{aligned} \rho(s^{\gamma_k})_{tt} - (\sigma_{\gamma_k})_x &= 0 \\ s^{\gamma_k}(0, t) &= 0, \quad \sigma_{\gamma_k}(L, t) = \frac{A}{L} (e^{-t/\tau_k} - e^{-(t-\Upsilon)/\tau_k}), \\ s^{\gamma_k}(x, 0) &= 0, \quad (s^{\gamma_k})_t(x, 0) = 0, \end{aligned} \quad (\text{B.1.5a})$$

In the above equation, the sensitivity of stress with respect to  $\gamma_k$  (i.e.,  $\sigma_{\gamma_k}$ ) is given by

$$\sigma_{\gamma_k} = E_1(s^{\gamma_k})_{xt} + \left(E + \sum_{j=1}^{N_p} \gamma_j\right) (s^{\gamma_k})_x - \sum_{j=1}^{N_p} \gamma_j \epsilon_{\gamma_k}^j + u_x - \epsilon^k. \quad (\text{B.1.5b})$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $\gamma_k$  (i.e.,  $\epsilon_{\gamma_k}^j = \frac{\partial \epsilon^j}{\partial \gamma_k}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

$$\tau_j(\epsilon_{\gamma_k}^j)_t + \epsilon_{\gamma_k}^j = (s^{\gamma_k})_x, \quad \epsilon_{\gamma_k}^j(0) = 0. \quad (\text{B.1.5c})$$

6. Sensitivity PDE for  $A$  (where  $s^A = \frac{\partial u}{\partial A}$ ):

$$\begin{aligned} \rho(s^A)_{tt} - (\sigma_A)_x &= 0 \\ s^A(0, t) &= 0, \quad \sigma_A(L, t) = \frac{1}{L} \left( \sum_{j=1}^{N_p} \gamma_j e^{-t/\tau_j} - \sum_{j=1}^{N_p} \gamma_j e^{-(t-\Upsilon)/\tau_j} \right), \\ s^A(x, 0) &= x/L, \quad (s^A)_t(x, 0) = 0, \end{aligned} \quad (\text{B.1.6a})$$

In the above equation, the sensitivity of stress with respect to  $A$  (i.e.,  $\sigma_A$ ) is given by

$$\sigma_A = E_1(s^A)_{xt} + \left(E + \sum_{j=1}^{N_p} \gamma_j\right) (s^A)_x - \sum_{j=1}^{N_p} \gamma_j \epsilon_A^j. \quad (\text{B.1.6b})$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $A$  (i.e.,  $\epsilon_A^j = \frac{\partial \epsilon^j}{\partial A}$ ) satisfies, for

$$j = 1, \dots, N_p,$$

$$\tau_j(\epsilon_A^j)_t + \epsilon_A^j = (s^A)_x, \quad \epsilon_A^j(0) = 0. \quad (\text{B.1.6c})$$

7. Sensitivity PDE for  $\Upsilon$  (where  $s^\Upsilon = \frac{\partial u}{\partial \Upsilon}$ ):

$$\begin{aligned} \rho(s^\Upsilon)_{tt} - (\sigma_\Upsilon)_x &= 0 \\ s^\Upsilon(0, t) &= 0, \quad \sigma_\Upsilon(L, t) = \frac{A}{L} \left( -\sum_{j=1}^{N_p} \frac{\gamma_j}{\tau_j} e^{-(t-\Upsilon)/\tau_j} \right), \\ s^\Upsilon(x, 0) &= 0, \quad (s^\Upsilon)_t(x, 0) = 0, \end{aligned} \quad (\text{B.1.7a})$$

In the above equation, the sensitivity of stress with respect to  $\Upsilon$  (i.e.,  $\sigma_\Upsilon$ ) is given by

$$\sigma_\Upsilon = E_1(s^\Upsilon)_{xt} + \left( E + \sum_{j=1}^{N_p} \gamma_j \right) (s^\Upsilon)_x - \sum_{j=1}^{N_p} \gamma_j \epsilon_\Upsilon^j. \quad (\text{B.1.7b})$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $\Upsilon$  (i.e.,  $\epsilon_\Upsilon^j = \frac{\partial \epsilon^j}{\partial \Upsilon}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

$$\tau_j(\epsilon_\Upsilon^j)_t + \epsilon_\Upsilon^j = (s^\Upsilon)_x, \quad \epsilon_\Upsilon^j(0) = 0. \quad (\text{B.1.7c})$$

## B.2 Shear Model

We derive the sensitivity equations for the shear model 4.4.14.

1. Sensitivity PDE for  $\rho$  (where  $s^\rho = \frac{\partial u}{\partial \rho}$ ):

$$\begin{aligned} \rho(s^\rho)_{tt} - (\sigma_\rho)_r - \frac{\sigma_\rho}{r} &= -u_{tt} \\ \sigma_\rho(r_{min}, t) &= 0, \quad s^\rho(r_{max}, t) = 0, \\ s^\rho(r, 0) &= 0, \quad (s^\rho)_t(r, 0) = 0, \end{aligned} \quad (\text{B.2.1a})$$

In the above equation, the sensitivity of stress with respect to  $\rho$  (i.e.,  $\sigma_\rho$ ) is given by

$$\sigma_\rho = G_1(s^\rho)_{rt} + \left( G + \sum_{j=1}^{N_p} \zeta_j \right) (s^\rho)_r - \sum_{j=1}^{N_p} \zeta_j \epsilon_\rho^j. \quad (\text{B.2.1b})$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $\rho$  (i.e.,  $\epsilon_\rho^j = \frac{\partial \epsilon^j}{\partial \rho}$ ) satisfies (for  $j = 1, \dots, N_p$ ).

$$\tau_j(\epsilon_\rho^j)_t + \epsilon_\rho^j = (s^\rho)_r, \quad \epsilon_\rho^j(0) = 0. \quad (\text{B.2.1c})$$

2. Sensitivity PDE for  $G$  (where  $s^G = \frac{\partial u}{\partial G}$ ):

$$\begin{aligned}\rho(s^G)_{tt} - (\sigma_G)_r - \frac{\sigma_G}{r} &= 0 \\ \sigma_G(r_{min}, t) &= 0, \quad s^G(r_{max}, t) = 0, \\ s^G(r, 0) &= 0, \quad (s^G)_t(r, 0) = 0,\end{aligned}\tag{B.2.2a}$$

In the above equation, the sensitivity of stress with respect to  $G$  (i.e.,  $\sigma_G$ ) is given by

$$\sigma_G = G_1(s^G)_{rt} + \left(G + \sum_{j=1}^{N_p} \zeta_j\right) (s^G)_r - \sum_{j=1}^{N_p} \zeta_j \epsilon_G^j + u_r.\tag{B.2.2b}$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $G$  (i.e.,  $\epsilon_G^j = \frac{\partial \epsilon^j}{\partial G}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

$$\tau_j(\epsilon_G^j)_t + \epsilon_G^j = (s^G)_r, \quad \epsilon_G^j(0) = 0.\tag{B.2.2c}$$

3. Sensitivity PDE for  $G_1$  (where  $s^{G_1} = \frac{\partial u}{\partial G_1}$ ):

$$\begin{aligned}\rho(s^{G_1})_{tt} - (\sigma_{G_1})_r - \frac{\sigma_{G_1}}{r} &= 0 \\ \sigma_{G_1}(r_{min}, t) &= 0, \quad s^{G_1}(r_{max}, t) = 0, \\ s^{G_1}(r, 0) &= 0, \quad (s^{G_1})_t(r, 0) = 0,\end{aligned}\tag{B.2.3a}$$

In the above equation, the sensitivity of stress with respect to  $G_1$  (i.e.,  $\sigma_{G_1}$ ) is given by

$$\sigma_{G_1} = G_1(s^{G_1})_{rt} + \left(G + \sum_{j=1}^{N_p} \zeta_j\right) (s^{G_1})_r - \sum_{j=1}^{N_p} \zeta_j \epsilon_{G_1}^j + u_{rt}.\tag{B.2.3b}$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $G_1$  (i.e.,  $\epsilon_{G_1}^j = \frac{\partial \epsilon^j}{\partial G_1}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

$$\tau_j(\epsilon_{G_1}^j)_t + \epsilon_{G_1}^j = (s^{G_1})_r, \quad \epsilon_{G_1}^j(0) = 0.\tag{B.2.3c}$$

4. Sensitivity PDE for  $\tau_k$  (where  $s^{\tau_k} = \frac{\partial u}{\partial \tau_k}$ ):

$$\begin{aligned}\rho(s^{\tau_k})_{tt} - (\sigma_{\tau_k})_r - \frac{\sigma_{\tau_k}}{r} &= \frac{1}{r} \frac{A}{r_{max} - r_{min}} \left( t \frac{\zeta_k}{\tau_k^2} e^{-t/\tau_k} - (t - \Upsilon) \frac{\zeta_k}{\tau_k^2} e^{-(t-\Upsilon)/\tau_k} \right) \\ \sigma_{\tau_k}(r_{min}, t) &= -\frac{A}{r_{max} - r_{min}} \left( t \frac{\zeta_k}{\tau_k^2} e^{-t/\tau_k} - (t - \Upsilon) \frac{\zeta_k}{\tau_k^2} e^{-(t-\Upsilon)/\tau_k} \right), \quad s^{\tau_k}(r_{max}, t) = 0, \\ s^{\tau_k}(r, 0) &= 0, \quad (s^{\tau_k})_t(r, 0) = 0,\end{aligned}\tag{B.2.4a}$$

In the above equation, the sensitivity of stress with respect to  $\tau_k$  (i.e.,  $\sigma_{\tau_k}$ ) is given by

$$\sigma_{\tau_k} = G_1(s^{\tau_k})_{rt} + \left(G + \sum_{j=1}^{N_p} \zeta_j\right) (s^{\tau_k})_r - \sum_{j=1}^{N_p} \zeta_j \epsilon_{\tau_k}^j. \quad (\text{B.2.4b})$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $\tau_k$  (i.e.,  $\epsilon_{\tau_k}^j = \frac{\partial \epsilon^j}{\partial \tau_k}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

- if  $j = k$  then

$$\tau_k(\epsilon_{\tau_k}^k)_t + \epsilon_{\tau_k}^k = (s^{\tau_k})_r - \epsilon_t^k, \quad \epsilon_{\tau_k}^k(0) = 0, \quad (\text{B.2.4c})$$

- if  $j \neq k$  then

$$\tau_j(\epsilon_{\tau_k}^j)_t + \epsilon_{\tau_k}^j = (s^{\tau_k})_r, \quad \epsilon_{\tau_k}^j(0) = 0. \quad (\text{B.2.4d})$$

5. Sensitivity PDE for  $\zeta_k$  (where  $s^{\zeta_k} = \frac{\partial u}{\partial \zeta_k}$ ):

$$\begin{aligned} \rho(s^{\zeta_k})_{tt} - (\sigma_{\zeta_k})_r - \frac{\sigma_{\zeta_k}}{r} &= \frac{1}{r} \frac{A}{r_{max} - r_{min}} (e^{-t/\tau_k} - e^{-(t-\Upsilon)/\tau_k}) \\ \sigma_{\zeta_k}(r_{min}, t) &= -\frac{A}{r_{max} - r_{min}} (e^{-t/\tau_k} - e^{-(t-\Upsilon)/\tau_k}), \quad s^{\zeta_k}(r_{max}, t) = 0 \\ s^{\zeta_k}(r, 0) &= 0, \quad (s^{\zeta_k})_t(r, 0) = 0, \end{aligned} \quad (\text{B.2.5a})$$

In the above equation, the sensitivity of stress with respect to  $\zeta_k$  (i.e.,  $\sigma_{\zeta_k}$ ) is given by

$$\sigma_{\zeta_k} = G_1(s^{\zeta_k})_{rt} + \left(G + \sum_{j=1}^{N_p} \zeta_j\right) (s^{\zeta_k})_r - \sum_{j=1}^{N_p} \zeta_j \epsilon_{\zeta_k}^j + u_r - \epsilon^k. \quad (\text{B.2.5b})$$

The sensitivity of the internal variable  $\epsilon^j$  with respect to  $\zeta_k$  (i.e.,  $\epsilon_{\zeta_k}^j = \frac{\partial \epsilon^j}{\partial \zeta_k}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

$$\tau_j(\epsilon_{\zeta_k}^j)_t + \epsilon_{\zeta_k}^j = (s^{\zeta_k})_r, \quad \epsilon_{\zeta_k}^j(0) = 0. \quad (\text{B.2.5c})$$

6. Sensitivity PDE for  $A$  (where  $s^A = \frac{\partial u}{\partial A}$ ):

$$\begin{aligned} \rho(s^A)_{tt} - (\sigma_A)_r - \frac{\sigma_A}{r} &= \frac{1}{r} \frac{1}{r_{max} - r_{min}} \left( \sum_{j=1}^{N_p} \zeta_j e^{-t/\tau_j} - \sum_{j=1}^{N_p} \zeta_j e^{-(t-\Upsilon)/\tau_j} \right) \\ \sigma_A(r_{min}, t) &= -\frac{1}{r_{max} - r_{min}} \left( \sum_{j=1}^{N_p} \zeta_j e^{-t/\tau_j} - \sum_{j=1}^{N_p} \zeta_j e^{-(t-\Upsilon)/\tau_j} \right), \quad s^A(r_{max}, t) = 0 \\ s^A(r, 0) &= \frac{r_{max} - r}{r_{max} - r_{min}}, \quad (s^A)_t(r, 0) = 0, \end{aligned} \quad (\text{B.2.6a})$$

In the above equation, the sensitivity of stress with respect to  $A$  (i.e.,  $\sigma_A$ ) is given by

$$\sigma_A = G_1(s^A)_{rt} + \left(G + \sum_{j=1}^{N_p} \zeta_j\right) (s^A)_r - \sum_{j=1}^{N_p} \zeta_j \epsilon_A^j. \quad (\text{B.2.6b})$$

The sensitivity of the internal variables  $\epsilon^j$  with respect to  $A$  (i.e.,  $\epsilon_A^j = \frac{\partial \epsilon^j}{\partial A}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

$$\tau_j(\epsilon_A^j)_t + \epsilon_A^j = (s^A)_r, \quad \epsilon_A^j(0) = 0. \quad (\text{B.2.6c})$$

7. Sensitivity PDE for  $\Upsilon$  (where  $s^\Upsilon = \frac{\partial u}{\partial \Upsilon}$ ):

$$\begin{aligned} \rho(s^\Upsilon)_{tt} - (\sigma_\Upsilon)_r - \frac{\sigma_\Upsilon}{r} &= \frac{1}{r} \frac{A}{r_{max} - r_{min}} \left( - \sum_{j=1}^{N_p} \frac{\zeta_j}{\tau_j} e^{-(t-\Upsilon)/\tau_j} \right) \\ \sigma_\Upsilon(r_{min}, t) &= - \frac{A}{r_{max} - r_{min}} \left( - \sum_{j=1}^{N_p} \frac{\zeta_j}{\tau_j} e^{-(t-\Upsilon)/\tau_j} \right), \quad s^\Upsilon(r_{max}, t) = 0 \\ s^\Upsilon(r, 0) &= 0, \quad (s^\Upsilon)_t(r, 0) = 0, \end{aligned} \quad (\text{B.2.7a})$$

In the above equation, the sensitivity of stress with respect to  $\Upsilon$  (i.e.,  $\sigma_\Upsilon$ ) is given by

$$\sigma_\Upsilon = G_1(s^\Upsilon)_{rt} + \left( G + \sum_{j=1}^{N_p} \zeta_j \right) (s^\Upsilon)_r - \sum_{j=1}^{N_p} \zeta_j \epsilon_\Upsilon^j. \quad (\text{B.2.7b})$$

The sensitivity of the internal variables  $\epsilon^j$  with respect to  $\Upsilon$  (i.e.,  $\epsilon_\Upsilon^j = \frac{\partial \epsilon^j}{\partial \Upsilon}$ ) satisfies, for  $j = 1, \dots, N_p$ ,

$$\tau_j(\epsilon_\Upsilon^j)_t + \epsilon_\Upsilon^j = (s^\Upsilon)_r, \quad \epsilon_\Upsilon^j(0) = 0. \quad (\text{B.2.7c})$$