CHARACTERIZATION OF NONLINEAR MATERIAL RESPONSE IN THE PRESENCE OF LARGE UNCERTAINTIES – A BAYESIAN APPROACH

A Dissertation

by

SRIKRISHNA DORAISWAMY

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Arun R. Srinivasa
Co-Chair of Committee,	Krishna R. Narayanan
Committee Members,	Junuthula N. Reddy
	John C. Criscione
	Sivakumar Rathinam
Head of Department,	Andreas A. Polycarpou

December 2013

Major Subject: Mechanical Engineering

ABSTRACT

The aim of the current work is to develop a Bayesian approach to model and simulate the behavior of materials with nonlinear mechanical response in the presence of significant uncertainties in the experimental data as well as the applicability of models. The core idea of this approach is to combine deterministic approaches by the use of physics based models, with ideas from Bayesian inference to account for such uncertainties.

Traditionally, parameters of models in mechanics have been identified through deterministic approaches to obtain single point estimates. Such methods perform very well for linear models and are the preferred approach in identifying model parameters, especially for precisely engineered systems such as structures and machinery. But in the presence of large variations such as in the response of biological materials, such deterministic approaches do not sufficiently capture the uncertainty in the response. We propose that the model parameters need to encode the spread that is observed in the data in addition to modeling the physics of the system. To this end, we propose the idea of probability distributions for model parameters in order to incorporate the uncertainty in the data.

We demonstrate this probabilistic approach to identifying model parameters with the example of two problems: the characterization of sheep arteries using data from inflation experiments and the problem of detecting an inhomogeneity in a cantilever beam. The parameters in the artery characterization problem are the model parameters in the constitutive models and in the cantilever problem the parameters are the stiffnesses of the inhomogeneity and the material of the beam. For each of these problems, we compute the probability distribution of the parameters using Bayesian

inference.

We show that the probability distributions of parameters can be used towards two kinds of diagnostics: assigning probability to a hypothesis (inhomogeneity detection problem) and using the probability distribution for classifying newly obtained data (characterization of artery data). For the inhomogeneity detection problem, the hypothesis is a statement on the ratio of the stiffnesses and it is observed that the probability of the hypothesis matches well with the data. In the case of the artery characterization problem, new data was successfully classified using the probability distributions computed with training data.

To my *Amma* and *Appa*

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Prof. Arun R. Srinivasa, for the continuous support and meticulous guidance throughout my PhD study and research. He has been a continuous source of inspiration and I am glad that I had the opportunity to know him the last seven years. I would also like to thank Prof. Krishna R. Narayanan, my co-advisor, for guiding me in my research through several fruitful discussions. Without the support and help of my advisors, this dissertation would not have been possible.

I would like to thank Dr. John C. Criscione for providing me with the experimental data that proved critical for me to present my dissertation. I would also like to thank Prof. J. N. Reddy, Dr. John C. Criscione and Dr. Sivakumar Rathinam for serving on my committee and giving valuable comments during my proposal.

I would like to thank all the members of my research group: Pritha Ghosh, Ashwin Rao, Jayavel Arumugam, Shreyas Balachandran and Jayadurga Iyer Ganapathi. I had a really good time picking their brains and learned a lot more than research from them.

My time in College Station has been made special and extremely memorable by my friends. Thanks to all of them: Harsha Nagarajan, Bhavana Iyer, Atul Narayan, Krishnakumar Ravikumar, Shriram Srinivasan, Kaarthik Sundar, Adithya Ganapathi and Nitin Suares. I would also like to thank the Indian classical music student organization SPICMACAY TAMU, and Prof. Shankar Bhattacharyya and Prof. Mysore Mohan, who gave me an opportunity to cultivate a taste in classical music and be a part of the organization. My special thanks go to Dr. N. Sivakumar who has been a mentor of sorts. His words of solace, his positive attitude, his taste in

music and his sense of humor have brought me great comfort at times I most needed it. I am honored to have him as a friend.

I would also like to thank my family for all their love and encouragement: my brother, Sriram Doraiswamy, who has never missed an opportunity to help; my father, Doraiswamy Natesan and my mother, Neela Doraiswamy, who have always supported me in all of my pursuits. Most of all, I would like to thank my loving wife, Ratipriya, who has been extremely patient, supporting and encouraging, through the toughest of times during my PhD. Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE Page

1. INTRODUCTION

The aim of the current work is to develop a general procedure based on Bayesian techniques to model and simulate the behavior of materials with nonlinear response, when there is significant uncertainty associated both with the applicability of the model as well as variability in experimental data. The idea behind this approach is to "embrace the uncertainty" and replace the traditional purely deterministic outlook of mechanics that seeks to eliminate variability by creating ever more complex models with an approach that augments the deterministic approach of mechanics with ideas from Bayesian inference to account for and propagate the uncertainty.

1.1 Motivation

Physics based models are the principal means of predicting the behavior of a wide range of systems for purposes of design and failure analysis and avoidance. Traditionally such approaches have been used for the modeling of precisely engineered systems such as structures and machinery where there is substantial control over dimensions, properties etc. Furthermore, most of these applications have been in bodies that are considered "almost rigid" so that linearized models have been sufficient. But with the recent interest in extending these ideas to biological systems, a number of new challenges have arisen. For example, biological systems are not controllably engineered, with huge variance across samples. Furthermore, unlike engineered systems that can be monitored in situ by embedding sensors during manufacture, *in-vivo* monitoring is rather minimal. The significant non-linearity in the response is another factor that compounds the difficulties. A key step in the development of models is the estimation of model parameters.

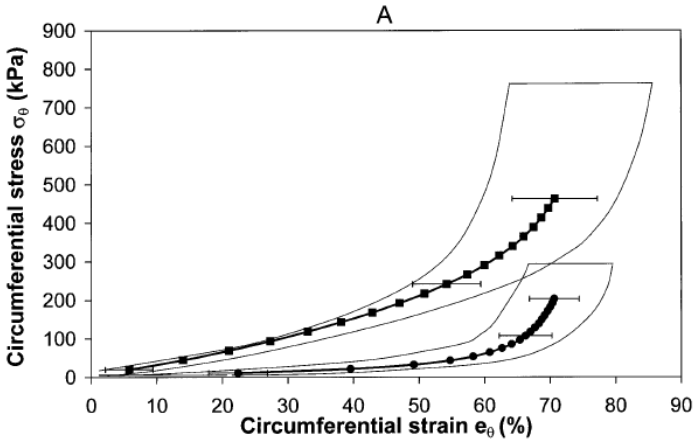
The estimation of model parameters in mechanics is traditionally done using

deterministic approaches to compute single point estimates. The core idea behind computing such point estimates is to optimize the value of the model parameters so that the "distance" between the data and model predictions is minimized. Least squares estimates and maximum likelihood estimates (in the context of statistical models) are examples of a larger class of point estimators known as M-estimators. M-estimators are extreme values of some chosen function of the data and model. (The reader is referred to [2] for detailed analysis of such M-estimators, with special reference to the robustness of such approaches.) If this function is chosen as the norm of the error (difference between data and model predictions), then the point estimate is the least squares estimate and if the function is a particular choice of likelihood function, then the estimate is known as the maximum likelihood estimates. Both the maximum likelihood estimates and least square approaches are popular in mechanics and it is well known that for linear models, such approaches perform very well and are the *de facto* choice (see for eg. [3, 4, 5, 6, 7]).

In the presence of large variations in measurements, as is observed in the mechanical response of biological materials (see for example in figure 1.1 (a)-(e) from [8, 9, 10, 11, 12] respectively), it is not sufficient to use point estimates for average response because, due to the spread in data, there is no reason to believe that the averaged response represents a "typical" response. Moreover, fitting parameters for the average response results in losing any information that may be contained in the spread (see Chapter 37 [13]). On the other hand, computing model parameters for each (or some chosen) experiment is inadequate and incomprehensive since the point estimates do not provide any information about the robustness of the fit or the ability to predict any future experiments. To put it differently, a deterministic approach, while capable of representing each experiment with a point estimate for the model parameters, does not contain any information to predict the possible response of a

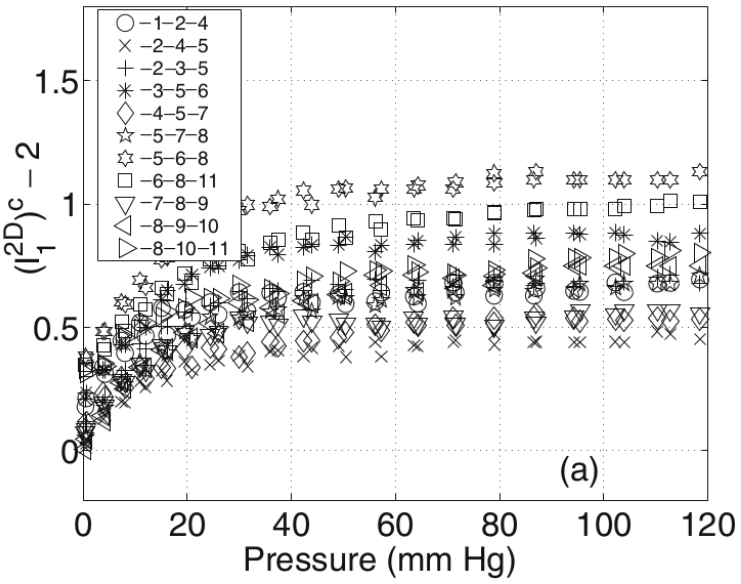
new sample.

The issues discussed above are exacerbated if the models are nonlinear. In particular, the response of biological materials, the mechanical response is known to be nonlinear [14, 15, 16, 17]. Since, for example, the least squares approach solves an optimization problem (minimizing "distance" between the data and model predictions), the use of nonlinear models, may result in a non-convex optimization problem, and therefore a unique value of the parameters for a given data set cannot be guaranteed. Furthermore, a nonlinear least squares estimate for the parameters does not provide any measure of robustness and hence errors in such estimates propagate into the characterization of the system (for details see for example chapters 10 and 11 in [18]).

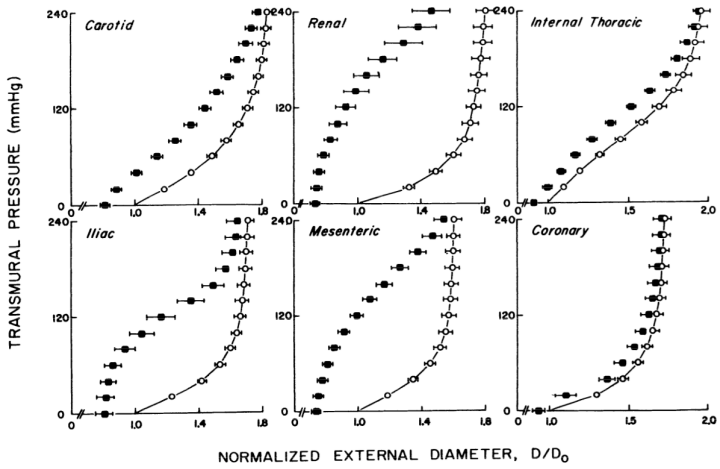


(a) The data presented in [12] – circumferential stress vs. strain of porcine coronary artery. The envelopes include all individual response curves at *in situ* axial stretch ($\lambda_z = 1.3$)

Figure 1.1: A sample of experimental data from the literature on biomaterials, showing large variations

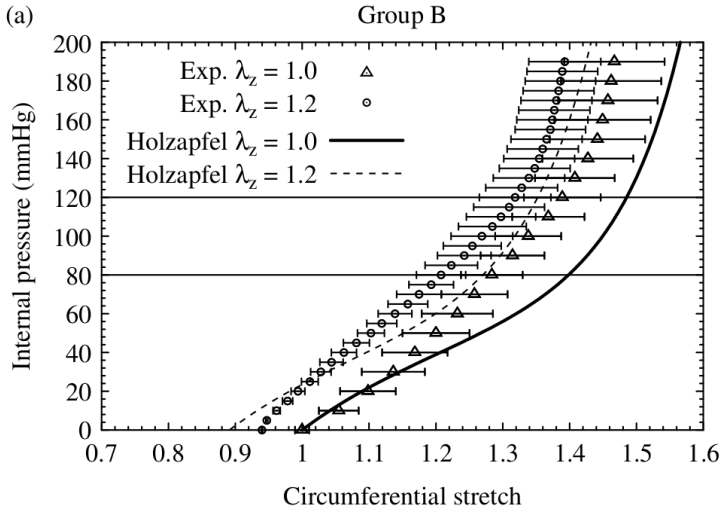


(b) The data shows the plot of pressure vs tr($\mathbf{C}$) for the inflation under constant stretch for a circumflex coronary artery [8]. Notice that there is a spread in the data (measured at different locations on the same sample) while the correlation within each curve is clear.

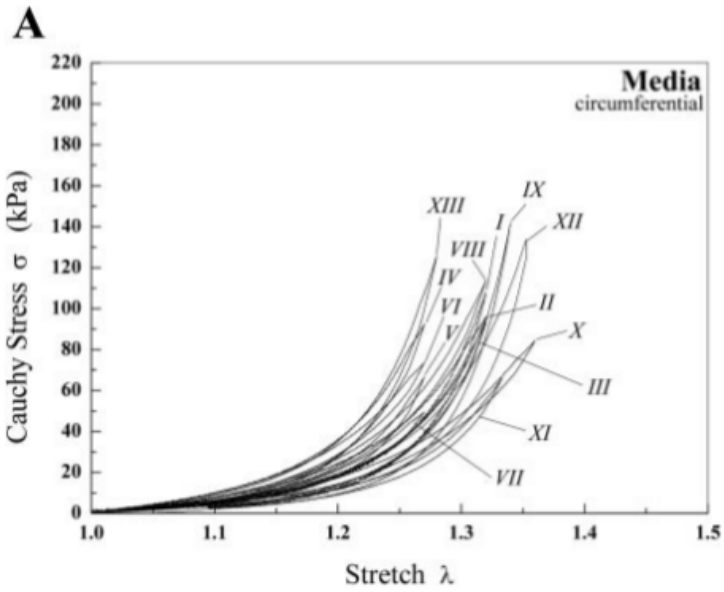


(c) The figure shows the pressure–diameter experimental data of canine arteries reported in Cox et al. [9]. Notice that the X-axis shows the normalized diameter plot, and the error bars are reported for the diameter variable only.

Figure 1.1: Continued.



(d) The data from [10] shows the internal pressure vs circumferential stretch at constant axial stretch for of human aorta and compared with model predictions (with the averaged data) using the Holzapfel model [19]



(e) The data from [11] shows the uniaxial response of human aorta layers with the roman numerals labeling the samples tested. Note the variations in the response across samples

Figure 1.1: Continued.

In summary, when the data shows large variations, point estimates for model parameters represent only an averaged response and are not sufficient to capture the variations. We propose that the model parameters need to encode the spread that is observed in the data in addition to capturing the correlation between the measured variables. To this end, we propose the idea of probability distributions for model parameters in order to incorporate the uncertainty in the data.

1.2 Advantages and consequences of using a probabilistic approach

A probabilistic approach to characterizing the system allows us to specify the probability that a parameter takes a particular value within a broad domain. In other words, the primary focus in the probabilistic approach to characterizing data is to answer the following question : "What is the probability distribution of the parameters of a model, conditioned on the observations (data)?". In this process, we circumvent issues of uniqueness, since the inverse problem reduces to a probability assignment over the parameter space. The probability that the parameter takes a particular value asserts the state of knowledge about the parameter, and thus provides an idea of the uncertainty in the value of the parameters. This approach is different in perspective from conventional methods of parameter estimation in that it does not seek to obtain a single set of parameter values that fit the data the "best", but rather characterize the parameter space as a probability distribution which is dependent on the quality (i.e. noise) or quantity of the data points. The probabilistic approach to identifying and representing model parameters is not limited to incorporating uncertainty in the data but also represents the uncertainty in models. The uncertainty arising from the lack of complete knowledge about the system and hence leading to an inaccurate representation of the system is known as *epistemic uncertainty* [20]. This uncertainty may be due to certain simplifying assumptions

on the system, or lack of accurate information on experiments or inability to obtain such information (e.g. *in vivo* material response for biological material needs to be modeled, but experiments can only be conducted on dead tissue). By representing model parameters of such models as probability distribution, the inability to match the data perfectly (due to this epistemic uncertainty) is also incorporated. A relevant discussion involving models, their complexity and the relation to the quality of available data is presented in section 1.3.

When the response of the system is represented through a physics based model, and probability distributions for parameters of the model represent the spread in the observations, the question to answer now becomes : "How likely is it that newly obtained data is different from the data used to calibrate the model?". Such questions are of primary importance in diagnostics, for example,

- Detect nature of an inhomogeneity : Based on observations of the system and physically motivated models of the system, is it possible to identify the nature of inhomogeneity?

  The presence of aneurysms [21, 22] in arteries is known to change the stiffness of the arteries and detecting them is a vital step in medical diagnostics. Detecting variations in tissue stiffness is also the central focus of imaging techniques such as elastography, particularly used for liver screening [23, 24] and breast cancer screening [25, 26].

- Detect if the material property of a biological material has changed over time: Using the newly obtained data, with the model and corresponding parameter probability distributions, can we detect if there is a change in material property over time?

  For example, atherosclerotic plaques are formed due to depositions of fatty

materials over time and the changes in arterial stiffness measured through pulse wave speed or ultrasonic measurements provide an estimate of damage in the tissue [27, 28].

Primary motivation for addressing this issue arises out of the unreliability on criteria for diagnostics, especially in the case of biomaterials. For example, existing maximum diameter criterion for prevention of rupture of an abdominal aortic aneurysm is unreliable, primarily due to the large variations observed in measured data [29]. Since the correlation between mechanical properties and diagnostics has been well studied in biomaterials [23, 25, 27, 30, 31], the diagnosis may be performed using material properties (such as stiffness) instead of using the measurements directly.

1.3 Complexity of models vs. quality of data – A qualitative overview

In the previous section, we presented the case for a probabilistic approach in the presence of uncertainty in data and models. In this section, we present a qualitative overview of when and why the probabilistic approach should be used. Figure 1.2 is a graphical representation of this overview.

Regions I through IV represent broadly the combinations of "complexity" in models vs. "quality" of data[1]. By complexity we mean the sophistication of explanation of reality. In other words, a more complex model, is able to explain the response of the system to a larger extent and under more conditions than a less complex model. The quality of the data is associated with the uniformity in experimental protocols, consistency across specimens and repeatability in obtaining data.

Based on the above classification, models which are sufficiently complex and have access to high quality of data are able to present accurate "predictions" of the

[1]It should be noted that the words complexity and quality here are used in an informal sense and no quantitative measure is attributed

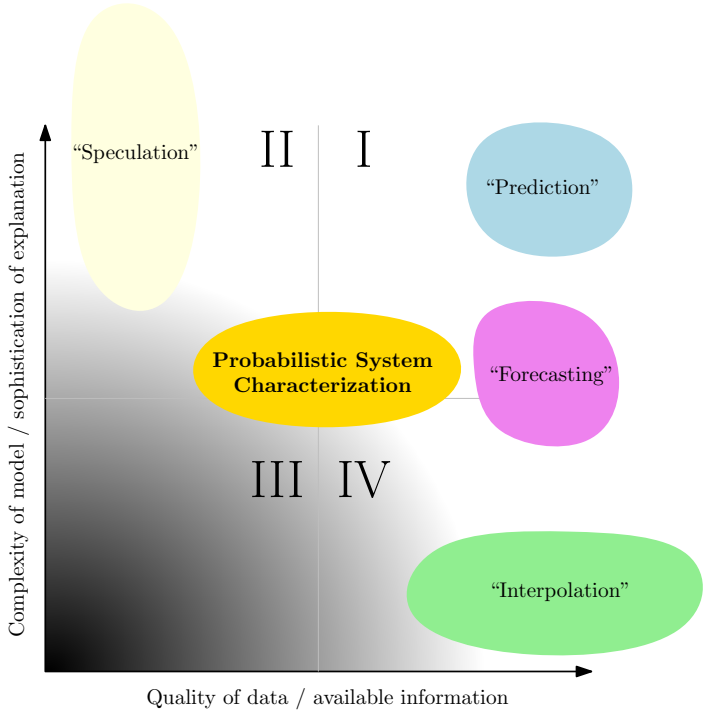


Figure 1.2: A qualitative representation of how the approach presented in this work is placed in relation to the existing approaches in representing data using models.

system's behavior. For such systems the least squares is the favored parameter estimation approach. Metallic materials and manufactured materials have been studied extensively and we can say with certainty that these systems are well characterized and predictions using models are accurate. An evidence to this fact is the establishment of multiple standardized protocols (such as ASTM standards) which enable obtaining material properties (i.e. model parameters) directly from the data.

On the other hand, region IV represents theories and statistical models associated with data obtained by social scientists, for example, where access to data is readily available, whereas the epistemic uncertainty is very high due to a lack of clear understanding of the underlying correlations. In such cases, only the data is used to develop statistical correlation models. With the correlation models thus developed,

"interpolations" may be made as to the probability of outcomes. For example, forecasting the results of an election involve a large amount of polling data combined with historical correlation models to obtain the probability of victory for a candidate [32, 33]

When the ability of collect data is limited, theories to explain reality supersede observations. Such theories lead to "speculations" (region II), which may be confirmed or rejected when the required data is available. Examples of speculative theories can be seen in particle physics and celestial physics where theories built on string theory and relativity generate "speculations", which are yet to be tested.

The region marked "forecasting" denotes the approach followed for financial and weather forecasts. Typically, the data (historical and current) is available, but developing models for such complex phenomenon is still a major challenge. The models used are thus simple and establish weak correlations, allowing the data to guide the "forecasts" which are presented as a series of possible outcomes rather than an accurate 'prediction' [34, 35, 36]. A relevant discussion on the different usages of the terms *prediction* and *forecasting*in the literature is presented in appendix B.

Given the above classification[2], a probabilistic approach of the type presented in this dissertation aims to focus on the area where models for the systems under consideration exist, but are not completely developed and the data that is obtained is not of very high quality, but does contain significant information. In this case, our approach suggests that both uncertainty in data and models need to be accounted for, and **probability distributions for model parameters** is one way to achieve this goal.

[2]Region III of figure 1.2 is not of scientific interest since neither data nor a physical understanding of such systems exist.

1.4 Bayesian inference

Probabilistic inference on parameters (i.e. parameter estimation) using data is typically performed in two different ways : the frequentist approach, and the Bayesian approach [18, 37]. The frequentist approach defines probability of an event as frequency of its occurrence in a large number of trials, whereas, Bayesian probability is a measure of state of knowledge or "belief". The frequentist approach to inference is primarily through hypothesis testing [38], where two competing hypothesis are tested against the data, to infer which hypothesis is more likely to explain the observations. Note that, while performing such tests, the hypotheses are only tested against the data, and that no prior available information is considered. Moreover, in hypothesis testing, the data is not used to update any existing belief on either hypothesis, but only to accept or reject it. Bayesian inference [39], on the other hand allows us to incorporate any prior available information, as well as use data to update the beliefs on the hypotheses. Given the fact that all models are themselves a priori beliefs about correlation in the variables of interest, it is natural to use Bayesian inference to update our state of knowledge of the parameters.

Bayesian inference is a method of statistical inference that uses the Bayes rule to compute the probability of a hypothesis in the presence of evidence. In other words, if E is the evidence, H is the hypothesis then

$$P(H \mid E) = \frac{P(E \mid H)P(H)}{P(E)}. \tag{1.1}$$

If we are comparing experimental measurements to update a hypothesis about a model, then if $\mathbf{z}$ is a list of measurements (evidence), and $\mathbf{y}$ the corresponding list of model predictions (model is the hypothesis on the response of the system) and $\mathbf{y}$ is a function of the parameters $\boldsymbol{\theta}$ the model, the Bayes rule for obtaining the probability

distribution on the parameters (for a given model, the hypothesis is on the values of the parameters) can be written as,

$$P(\boldsymbol{\theta} \mid \mathbf{z}) = \frac{P(\mathbf{z} \mid \boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\mathbf{z})} \tag{1.2}$$

$$= \frac{P(\mathbf{z} \mid \mathbf{y}(\boldsymbol{\theta}))P(\mathbf{y}(\boldsymbol{\theta}))}{P(\mathbf{z})} \tag{1.3}$$

since the model predictions $\mathbf{y}$ are a function of the parameters $\boldsymbol{\theta}$. In the above equation, $P(\mathbf{z} \mid \boldsymbol{\theta})$ is called the *likelihood* factor, $P(\boldsymbol{\theta})$, is called the *prior* probability and the denominator $P(\mathbf{z})$ is called the evidence. This approach allows one to incorporate prejudices and prior information into the inference procedure, while at the same time use the data to update these prior beliefs.

1.4.1 Features of the Bayesian inference procedure

There are several features of the Bayesian inference that are particularly useful for solving a probabilistic system characterization problem. Firstly, the prior probability distribution on the model parameters $\boldsymbol{\theta}$, encodes any prior knowledge that may be available on the model parameter values as a probability distribution. This prior information may range from a support constraint on the parameter space, to knowing the average values of the parameters for a set of specimens. It is important to note that if the prior probability assigns zero probability to a region of the probability space, the posterior probability *will* assign zero probability to this region irrespective of the information contained in data.

The likelihood factor ($P(\mathbf{z} \mid \boldsymbol{\theta})$) is a measure of influence that the data has on the model. A high likelihood factor for a particular choice of model parameters $\boldsymbol{\theta}_0$ implies that the model is significantly influenced by the data at this particular value of the

model parameters. Furthermore, the likelihood also allows us to enforce precision on the model parameters, i.e. a high precision requirement implies a correspondingly high belief in the quality of data, and hence the tolerance in the values that the model parameters can take is decreased. If the data requires such high tolerance, then the posterior probability distribution of the model parameters will reflect this through a decreased variance.

Having computed the posterior probability distribution, we have a comprehensive representation of the uncertainty in the data and the model. Several avenues of further analysis may be pursued with the posterior distribution. If the posterior distributions need to be directly compared, a "distance" measure between probability distributions such as the K-L divergence (Kullback-Leibler) [40] may be used. If the hypothesis on the model parameters is stated as a probability, i.e. "What is the probability that $\boldsymbol{\theta} \in \Theta$?", where Θ is a subset of the parameter space, then such a probability may be directly computed from the posterior probability distribution. Alternatively, such a hypothesis may be posed as a classification problem (i.e. class C_1 is defined as $\boldsymbol{\theta} \in \Theta$ and class C_2 is the complement of this set), then a consistent use of Bayesian inference to compute the probability $P(C_1 \mid \boldsymbol{\theta})$ may also be computed.

1.5 Probabilistic approaches for parameter estimation in mechanics – a brief survey

Recent work indicates an increased interest in applying probabilistic parameter estimation techniques to problems in mechanics. Mehrez et al. [41, 42] use a polynomial chaos (PC) expansion to represent the probability distribution for the spatial variation of elastic modulus along the length of a composite beam. The authors assume a random field over the modulus and use a Bayesian inference technique to obtain point estimates (maximum a posteriori) estimates for the coefficients of

the polynomial chaos expansion. The probability distribution thus obtained for the elastic modulus is used to compute maximum a posteriori (MAP) estimates. These estimates are used in the PC expansion again to validate the results across different sets of calibration and validation datasets. A similar estimation problem has also been solved by Zhang et al. [43] for linear models applied to using a combination of the PC expansion and Bayesian inference.

The works of Beck, Katafygiotis and co. [44, 45, 46] deal with probabilistic system identification approaches for dynamical systems. The authors apply probabilistic inference to identify the stiffnesses for multiple degree of freedom systems. The estimates obtained are typically MAP estimates, and are used towards structural health monitoring and reliability [47] applications. Another recent work (2013) in health monitoring and damage detection using elastic waves, by Yan, [48], explores the possibility of using sensors to measure the excitation at points on an elastic plate and use the measurements to obtain probability distributions on the damage location.

A very recent work by Crews and Smith [49] employs the Bayesian inference approach for estimating heat transfer parameters of a shape memory alloy bending actuator. The aim of the work is to obtain parameters towards use in design of robust controllers. The data used is the measurement of bending angle of SMA beam actuator on applying a voltage. The authors use the mechanical and the SMA model parameters from an earlier work [50], and for the heat transfer parameters, use Bayesian inference to present probability distribution. Since the approach is directed towards design of controllers, the mean value and 95% confidence intervals are presented.

In the field of geology and geomechanics, Cividini et al. [51] present a Bayesian inference approach to estimate the size, location and the elastic moduli governing

the geometry of an inclusion in a rock. The experimental data are given by the displacements and certain known loads. The authors use a linear elastic model to represent the system and apply the data in conjunction with a priori assumptions on the parameters to a Bayesian framework. The a prior information in this problem is the knowledge of mean values and variances of the parameters, and a normal distribution is thus assumed for these parameters. Using a Gaussian likelihood function therefore results in a Gaussian distribution for the posteriors as well. The results are then reported as the MAP estimates and their variances.

The above examples give a brief overview of probabilistic approaches used in the mechanics literature. It can be seen that probabilistic approaches are used mainly to obtain point estimates, as mean of the resultant distribution [49] or as maximum a posteriori estimates [41, 42, 47]. This is clearly due to the motivation of the estimation procedure, which is typically tailored towards use in design, system identification or health monitoring. Moreover, the uncertainty observed in the data in each of the listed works, considers uncertainty due to measurement errors only. As mentioned earlier, such an assumption is sufficient for engineered systems, but given the variations and spread in the data from biological materials (see figure 1.1), the probability distributions for model parameters need to incorporate the information in spread. With this spread information incorporated in the probability distribution, it is not appropriate to obtain point estimates, such as maximum a posterior estimates. Furthermore, a common feature across papers in the mechanics literature that use probabilistic approach is their application to linear models. While the probabilistic approaches are the same for nonlinear models, the interpretation of the distribution of parameters of nonlinear models is not the same. For example, choosing maximum a posterior estimates (i.e. choosing the maximum of the posterior distribution) ignores the presence of other modes in the probability distribution. Therefore, in this

dissertation we present techniques to use the probability distributions for the model parameters in their entirety instead of using only a point estimate.

1.6 Objective and scope

In this dissertation, we will focus our attention on a Bayesian inference approach to compute probability distributions for model parameters for nonlinear mechanical response from data that show large variations. The key difference in the perspective presented in the Bayesian approach is that we *do not* assume that the model parameters are distributed according to a particular family of distributions. This perspective leads to probability distributions that do not assume any specific features of a particular family (for example, models on correlation). Moreover, due to this freedom, the priors need not be chosen in any particular compliance with the likelihood factor, as is typically done with the use of conjugate priors [52]. On the other hand, this does include the possibility that the parameter probability distributions do not have analytically tractable forms, but this is handled by using numerical sampling techniques. The main algorithm used in this work for this purpose is an example of the Markov-Chain Monte-Carlo [53] techniques known as the Metropolis–Hastings algorithm [54, 55].

We will demonstrate the probabilistic approach with the example of two problems: the first one lays out the approach towards characterizing material properties with the example of artery data and a continuum mechanics model for the behavior of the artery and the second problem presents an example in structural mechanics of identifying an inhomogeneity in a cantilever beam. The following briefly outlines the problem statements for the two problems.

Probabilistic characterization and classification of sheep aorta inflation data

This problem is posed as the characterization of sheep aorta from inflation data.

In order to demonstrate and maintain focus on the inference approach, three simple hyperelastic models are used to model the inflation of the artery. The probability distributions for the parameters of each of these models is computed using the Metropolis–Hastings algorithm. Only a portion of the experimental data is used in computing the probability distribution. The remainder is used as "new" data and the problem of classifying the new data into predefined classes is shown, as an application of the posterior probability distribution.

Inhomogeneity detection in a cantilever beam

We consider the problem of detecting an inhomogeneity in a cantilever beam modeled as a nonlinear elastica using measurements of deflections along the beam. The measurements used are synthetically generated. This allows us to analyze the presented approach in the context of near-perfect models. The probability distribution is computed using Bayesian inference on the stiffnesses of the material of the beam and the material of the inhomogeneity. The probability distributions are computed under different assumptions on the noise in observations. Prior information on the stiffnesses is encoded using the Maximum Entropy principle. An application of the posterior distribution to assign probabilities to hypotheses is also presented in this work.

1.7 Structure of the dissertation

The structure of the dissertation is as follows:

1. Section 1 presents a introduction to the probabilistic approach applied to parameter estimation in mechanics models in the context of data that shows large variations. A detailed motivation for this approach is presented in comparison to the popular least squares approach to parameter identification. This section also lays down the particular scope of this approach and the current work.

2. In section 2, sheep aorta inflation data from experiments recently conducted [1] is used in conjunction with hyperelastic models. The use of the Bayesian inference approach to this problem and an application to a probabilistic classification problem are presented. The problem of classifying newly obtained data into predefined classes is directly related to medical diagnosis and the results are presented in this context.

3. In section 3, the problem of identifying an inhomogeneity in a cantilever beam is presented. The cantilever is modeled as a nonlinear elastica with an inclusion of a different stiffness and the data (deflection measurement) is synthetically generated. The use of maximum entropy principle to obtain a prior probability distribution is also demonstrated in this work. Posterior probability distribution of the stiffnesses (beam and inclusion) is used to find the probability of hypothesis as an application towards diagnostics.

4. We close in section 4, with observations on the advantages of the various aspects of this work and a discussion of future research directions.

2. PROBABILITY DISTRIBUTION OF MODEL PARAMETERS FOR ARTERY DATA – A COMPARATIVE STUDY OF THREE DIFFERENT MODELS

2.1 Introduction

In this section, we present the model-based Bayesian inference approach to characterizing and predicting the behavior of arteries. We will use data from experiments on inflation of arteries to compute the probability distribution on model parameters. Thus the uncertainty associated with the data as well as model is captured not through just a mean least squares estimate and its variance, but a probability distribution of the model parameters. This approach thus also circumvents issues of non-uniqueness in least squares estimates (see [56] for a discussion on non-uniqueness in least squares estimates for hyperelastic models). Moreover this approach paves the way for further analysis, particularly for classification towards diagnosis. For example, a measure such as the Kullback-Leibler divergence [40] may be used to obtain a measure of the difference between two probability distribution. In this section, we demonstrate the use of Bayes rule with the probability distribution of the model parameters to compute class membership probabilities i.e., the probability that a set of observations belong to a particular class. Another use towards diagnostics using the model parameter probability distributions is towards detecting inhomogeneities in diseased arteries (e.g. atherosclerotic arteries). A simplified version of this problem is presented in section 3 of this dissertation.

2.2 Experimental data

The experiments used for demonstrating the approach presented in this work were recently conducted [1] at Texas A&M University. Three inflation experiments

were performed on 5 samples of sheep aorta at the axial stretch corresponding to *in vivo* axial stretch. The data was reported as applied pressure (in mmHg) vs. volume (in mL). This data was reduced to pressure vs. internal radius (in mm). The data, in the reduced form is shown in figure 2.1.

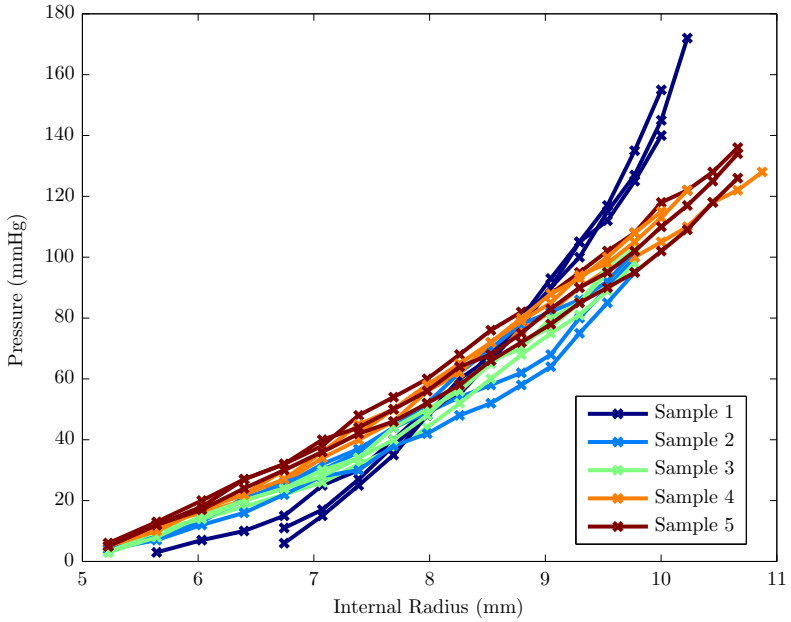


Figure 2.1: Inflation experimental data from 5 samples of sheep aorta – Pressure vs. Internal radius [1]. Note that in addition to a significant variance in the reported measurements of different experiments for a given sample, there is also a variance in the behavior across samples.

2.3 Model for artery inflation

For the purposes of this work the artery is modeled as a thick-walled tube made of a homogeneous, incompressible, isotropic and hyperelastic material. This choice is motivated by the fact that the exact geometry of the arteries that were tested

is not reported, and furthermore the data is presented as (section 2.2) the applied pressure vs. internal volume (and reduced to pressure vs. internal radius). The model that we develop is motivated towards obtaining the model predictions of the applied pressure $p_{\text{mod}}(\boldsymbol{\theta}, r_i)$ as a function of the internal radius r_i and the model parameters $\boldsymbol{\theta}$ of the strain energy chosen. The problem statement for the inflation problem and the solution technique used for the solving this problem are explained below.

2.3.1 Problem statement

Consider a thick walled cylinder B that occupies the reference configuration $k_r(B)$. Let $\mathbf{X} \in k_r(B)$ denote the position vector of a typical particle in the body. At time t, the body occupies a configuration $k_t(B)$ such that position of the same particle is now $\mathbf{x}$. The motion of the body can be described using the function $\boldsymbol{\chi}_{k_r}$

$$\mathbf{x} = \boldsymbol{\chi}_{k_r}(\mathbf{X}, t) \tag{2.1}$$

The deformation gradient $\mathbf{F}$ is given by

$$\mathbf{F} = \frac{\partial \boldsymbol{\chi}_{k_r}}{\partial \mathbf{X}}, \tag{2.2}$$

and the right Cauchy-Green stretch tensor is

$$\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F} \tag{2.3}$$

As mentioned earlier body B is assumed to be a thick walled tube,

$$\Omega_0 = \{(R, \Theta, Z\} \mid A \leq R \leq B, 0 \leq \Theta \leq 2\pi, 0 \leq Z \leq L\}. \tag{2.4}$$

Consider the inflation of this annular region under an applied pressure on the surface $R = A$. The solution to the following semi-inverse problem is sought:

$$r = f(R) \tag{2.5a}$$

$$\theta = \Theta \tag{2.5b}$$

$$z = \lambda_z Z \tag{2.5c}$$

where it is assumed that the current radius of the cylinder is independent of θ and z, since the data is reduced from pressure vs. volume assuming it is a uniform cylinder and that the inflation is independent of θ and z.

2.3.2 Constitutive relation

As mentioned earlier,the data is presented as pressure vs. internal radius. Therefore, we desire a similar form for the constitutive relation. We will use the principle of minimum potential energy to arrive at such a relation.

The deformation gradient in cylindrical coordinates is given by,

$$F = \begin{bmatrix} \frac{\mathrm{d}r}{\mathrm{d}R} & \frac{1}{R}\frac{\mathrm{d}r}{\mathrm{d}\Theta} & \frac{\mathrm{d}r}{\mathrm{d}Z} \\ r\frac{\mathrm{d}\theta}{\mathrm{d}R} & \frac{r}{R}\frac{\mathrm{d}\theta}{\mathrm{d}\Theta} & r\frac{\mathrm{d}\theta}{\mathrm{d}Z} \\ \frac{\mathrm{d}z}{\mathrm{d}R} & \frac{1}{R}\frac{\mathrm{d}z}{\mathrm{d}\Theta} & \frac{\mathrm{d}z}{\mathrm{d}Z} \end{bmatrix} \tag{2.6}$$

and therefore, for the deformation given in equation (2.5),

$$F = \begin{bmatrix} \frac{\mathrm{d}f(R)}{\mathrm{d}R} & 0 & 0 \\ 0 & \frac{f(R)}{R} & 0 \\ 0 & 0 & \lambda_z \end{bmatrix} \tag{2.7}$$

Since the cylinder is assumed to be incompressible,

$$\det\mathbf{F} = 1 \tag{2.8}$$

$$\lambda_z \times \frac{f(R)}{R} \times \frac{\mathrm{d}f(R)}{\mathrm{d}R} = 1 \tag{2.9}$$

The above differential equation gives,

$$f(R) = \sqrt{\frac{R^2 + 2c\lambda_z}{\lambda_z}}, \tag{2.10}$$

The function $f(R)$ contains one constant of integration, c. Traditionally this constant is evaluated by using information from the geometry in reference and current configurations (e.g. $R = A$, $\lambda_z = 1$, $r = a$, $\Rightarrow$ $c = \frac{a^2 - A^2}{2}$). Following this, one would solve the equilibrium equations to obtain the pressure – radius relations. We will depart from this approach, and retain this constant of integration c. We will instead use the principle of minimum potential energy and obtain the pressure vs. c relation. Such a relation, combined with equation (2.10) leads to a constitutive relation between the applied pressure p and the internal radius r_i (i.e. $f(A)$).

Remark. Note that c is a constant only for a particular deformation. In other words, if the applied boundary conditions are changed, then the corresponding value of c will also change.

The principal stretches for this deformation 2.5 are given by

$$\lambda_1 = \frac{\mathrm{d}f(R)}{\mathrm{d}R} = \frac{R}{\lambda_z\sqrt{2c + \frac{R^2}{\lambda_z}}} \tag{2.11a}$$

$$\lambda_2 = \frac{f(R)}{R} = \frac{\sqrt{2c + \frac{R^2}{\lambda_z}}}{R} \tag{2.11b}$$

$$\lambda_3 = \lambda_z \tag{2.11c}$$

and the invariants $\mathrm{I}^{\mathrm{C}}, \mathrm{II}^{\mathrm{C}}, \mathrm{III}^{\mathrm{C}}$ of the right Cauchy-Green stretch tensor are given as

$$\mathrm{I}^{\mathrm{C}} = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \qquad = \lambda_z^2 + \frac{R^2}{\lambda_z\left(R^2 + 2c\lambda_z\right)} + \frac{R^2 + 2c\lambda_z}{R^2\lambda_z} \tag{2.12a}$$

$$\mathrm{II}^{\mathrm{C}} = \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2 \qquad = \frac{1}{\lambda_z^2} + \frac{R^2\lambda_z}{R^2 + 2c\lambda_z} + \frac{\lambda_z\left(R^2 + 2c\lambda_z\right)}{R^2} \tag{2.12b}$$

$$\mathrm{III}^{\mathrm{C}} = \lambda_1^2\lambda_2^2\lambda_3^2 = 1 \tag{2.12c}$$

The principle of minimum potential energy states that admissible deformations are those that render the potential energy functional to be a minimum.

For the problem under consideration, the potential energy ψ is,

$$\psi = \int_{\Omega_0} W\,\mathrm{d}V - p\Delta V \tag{2.13}$$

where W is the strain energy , P is the pressure applied to the surface $R = A$ of the cylinder and ΔV is the change in volume enclosed by this surface.

Rewriting equation (2.13),

$$\psi = \int_0^L \int_0^{2\pi} \int_A^B W\, R\, \mathrm{d}R\, \mathrm{d}\Theta\, \mathrm{d}Z - p \times \pi(f(A)^2 - A^2)L\,. \tag{2.14}$$

Since the strain energy is typically a function of either the invariants of $\mathbf{C}$, ($W = \hat{W}(\mathrm{I^C}, \mathrm{II^C}, \mathrm{III^C})$), or the principal stretches ($W = \tilde{W}(\lambda_1, \lambda_2, \lambda_3)$), and from equations (2.11) and (2.12) it can be see that the invariants and stretches are independent of Θ and Z, the integrand is a function of R and c only. Therefore the equation (2.14) simplifies as

$$\begin{aligned}\psi &= \int_0^L \int_0^{2\pi} \int_A^B W\, R\, \mathrm{d}R\, \mathrm{d}\Theta\, \mathrm{d}Z - p \times \pi(\frac{A^2 + 2c\lambda_z}{\lambda_z} - A^2)L \\ &= 2\pi L \int_A^B W\, R\, \mathrm{d}R - p \times \pi(\frac{A^2(1-\lambda_z) + 2c\lambda_z}{\lambda_z})L\end{aligned}$$

Given that the experiments were all conducted at a constant axial stretch ($\lambda_z = 1$), the only variable in the above equation is c. Therefore, since the minimum of the potential energy is attained when the gradient is zero, i.e.,

$$\begin{aligned}\frac{\mathrm{d}\psi}{\mathrm{d}c} = 0 \Rightarrow 2\pi L \int_A^B \frac{\mathrm{d}W}{\mathrm{d}c}\, R\, \mathrm{d}R - 2p\pi L = 0 \\ p = \int_A^B \frac{\mathrm{d}W}{\mathrm{d}c}\, R\, \mathrm{d}R\end{aligned} \tag{2.15}$$

Equation (2.15) is the constitutive equation that relates the pressure p to the constant of integration c . The relation between the pressure and the (current) internal radius is sought since the data is also in this format. Since the relation between c and the internal radius can be easily obtained from equation (2.10), the

constitutive relation can be rewritten when required.

The specific form of equation (2.15) depends on the choice of the strain energy function. In the following section, three different strain energies will be used to get the corresponding constitutive equations. These are the power-law neo-Hookean model, a Criscione-type model and the Ogden model.

2.3.3 Forms of constitutive relations for specific strain energies

For each of the models presented, the form of equation (2.15) is presented below. To compute the pressure p, we will use the internal radius to obtain the value of c, from (2.10) (with $\lambda_z = 1$),

$$c = \frac{r_i^2 - A^2}{2} \tag{2.16}$$

and the for each value of c thus computed, we can compute the pressure from each of the models.

2.3.3.1 Power-law neo-Hookean model

The strain energy density for the incompressible power-law neo-Hookean model (from [57] with $b = 1$) is given by

$$W = \frac{\mu}{2}\left\{\left[1 + \frac{1}{n}(\mathrm{I}^{\mathrm{C}} - 3)\right]^n - 1\right\}, \tag{2.17}$$

where $\mu > 0$ is the shear modulus and n is a positive number. Note that when $n = 1$, the strain energy reduces to that of neo-Hookean model. The constitutive relation from equation (2.15) for this strain energy is

$$p = \int_A^B \frac{4\mu cn\left(c + R^2\right)\left(1 + \frac{4c^2}{2cnR^2 + nR^4}\right)^n}{\left(2c + R^2\right)\left(4c^2 + nR^2\left(2c + R^2\right)\right)} R\,\mathrm{d}R \tag{2.18}$$

The model parameters for this strain energy are μ and n.

2.3.3.2 Ogden model

The strain energy density for the incompressible Ogden model is chosen as

$$W = \frac{2\mu}{\alpha^2}\left(\lambda_1^\alpha + \lambda_2^\alpha + \lambda_3^\alpha + \frac{1}{\lambda_1^\alpha} + \frac{1}{\lambda_2^\alpha} + \frac{1}{\lambda_3^\alpha} - 6\right) \tag{2.19}$$

where μ and α are material constants. The constitutive relation using this strain energy is ,

$$p = \int_A^B -\left(\frac{2\mu\left[\left(\frac{R}{\sqrt{2c+R^2}}\right)^\alpha - \left(\frac{\sqrt{2c+R^2}}{R}\right)^\alpha\right]}{(2c+R^2)\alpha} - \frac{2\mu\left[\left(\frac{\sqrt{2c+R^2}}{R}\right)^{-\alpha} - \left(\frac{R}{\sqrt{2c+R^2}}\right)^{-\alpha}\right]}{(2c+R^2)\,\alpha}\right) R\,\mathrm{d}R \tag{2.20}$$

2.3.3.3 Criscione-type model with logarithmic strain attributes

The strain energy $W(\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6)$ is a function of the 6 strain attributes suggested by Criscione [58] and ,

$$\gamma_1 = \log J \qquad \gamma_2 = (3/2)\log\lambda_3 \qquad \gamma_3 = 2\log\lambda_2 + \log\lambda_3$$

$$\gamma_4 = \phi_{RZ} \qquad \gamma_5 = \phi_{\Theta Z} \qquad \gamma_4 = \phi_{\Theta R}$$

where, $J = \det F$, and ϕ_{RZ}, $\phi_{\Theta Z}$, $\phi_{\Theta R}$ are the shear strains in the planes denoted by the corresponding subscripts.

For the problem at hand, γ_3 is the only nonzero attribute and so, the chosen form

of W is

$$W(\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6) = \hat{W}(\gamma_3) = a\cosh(b\gamma_3) \tag{2.21}$$

The pressure, from (2.15) is therefore,

$$p = \int_A^B \frac{2ab\sinh\left[2b\log\left[\frac{\sqrt{2c+R^2}}{R}\right]\right]}{2c+R^2} R\,\mathrm{d}R \tag{2.22}$$

The model parameters in this model are a and b.

With the pressure p expressed as a function of the model parameters, the conventional approach to 'fit' the model to data would be to solve a non-linear least squares problem. We will depart from this approach and seek probability distribution for the reasons explained in section (2.1).

2.4 Probabilistic framework – Bayesian inference

Bayesian inference is the method of using the Bayes' rule to update the state of information using observations. The state of information before the observations were made is known as the *prior* and the updated state of information is known as the *posterior.* The following notation is used in the remainder of the section: let $\mathbf{p}_{\text{exp}}$ be the vector of measured pressure values corresponding to the vector of internal radius values $\mathbf{r}_i$. Let $\mathbf{p}_{\text{mod}}(\boldsymbol{\theta})$ be the vector of model predictions corresponding to $\mathbf{r}_i$ as a function of the model parameters $\boldsymbol{\theta}$. We are interested in computing the probability distribution $P(\boldsymbol{\theta} \mid \mathbf{p}_{\text{exp}})$, i.e. the probability distribution on the model parameters $\boldsymbol{\theta}$ given the data $\mathbf{p}_{\text{exp}}$.

2.4.1 Probability distribution of model parameters

As mentioned in the previous section, the classification problem requires the computation of the probability distribution of the model parameters for each of the

classes. As mentioned in section 2.2 the data is reported as pressure, $\mathbf{p}_{\text{exp}}$ vs internal radius $\mathbf{r}_i$. The probability distribution of the parameters $\boldsymbol{\theta}$, given the observations $\mathbf{p}_{\text{exp}}$ is computed using Bayes' rule:

$$P(\boldsymbol{\theta} \mid \mathbf{p}_{\text{exp}}) = \frac{P(\mathbf{p}_{\text{exp}} \mid \boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\mathbf{p}_{\text{exp}})} \tag{2.23}$$

Let the model predictions be $\mathbf{p}_{\text{mod}}(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the vector of parameters. Since the model predictions is a function of the parameter vector $\boldsymbol{\theta}$, equation (2.23) is rewritten as,

$$P(\boldsymbol{\theta} \mid \mathbf{p}_{\text{exp}}) = \frac{P(\mathbf{p}_{\text{exp}} \mid \mathbf{p}_{\text{mod}}(\boldsymbol{\theta}))P(\boldsymbol{\theta})}{P(\mathbf{p}_{\text{exp}})}$$

where

$$P(\mathbf{p}_{\text{exp}} \mid \mathbf{p}_{\text{mod}}(\boldsymbol{\theta})) = \frac{1}{Z} \exp\left(- \frac{(\mathbf{p}_{\text{mod}}(\boldsymbol{\theta}) - \mathbf{p}_{\text{exp}})^2}{\sigma^2} - \frac{\beta}{\sigma^2} \left[\frac{\mathrm{d}\mathbf{p}_{\text{mod}}(\boldsymbol{\theta})}{\mathrm{d}r} - \frac{\mathbf{p}_{\text{exp}}(i+1) - \mathbf{p}_{\text{exp}}(i-1)}{r(i+1) - r(i-1)} \right]^2 \right) \tag{2.24}$$

where Z is the appropriate normalizing factor. The exact value of this normalizing factor is not explicitly computed since the algorithms used in this work use only the ratio of likelihoods. The probability in equation (2.24) is from the assumption on the form of the likelihood function in equation (2.28).

The prior probability distribution for the parameters $P(\boldsymbol{\theta})$ may be chosen according to any available prior information on the value of the parameters. If there is no distinguishing prior information, then a *non-informative* prior may be chosen so as to not bias the inference process. For the work presented in this section, a

non-informative prior is assumed.

2.4.2 Likelihood function and its interpretation

In equation (2.23), $P(\mathbf{p}_{\text{exp}} \mid \boldsymbol{\theta})$ is called the likelihood function. An alternate interpretation of the inference presented can be motivated through the variational free energy formulation. The variational free energy is central to the set of techniques known as the variational Bayesian inference. We will show here how the variational free energy formulation is related to the Bayesian inference [59].

Let $\mathbf{z}$ be the data and $y(\boldsymbol{\theta})$ be the model predictions dependent on the parameters $\boldsymbol{\theta}$. Let $P(\boldsymbol{\theta})$ be the prior probability distribution on the model parameters and the $P(\mathbf{z} \mid \boldsymbol{\theta})$ be the likelihood function (which is related to the the model $y(\boldsymbol{\theta})$). The variational Bayesian problem is stated as the following optimization problem,

$$\max_{Q(\boldsymbol{\theta})} \quad \mathbb{E}_{Q(\boldsymbol{\theta})}\left[\log P(\mathbf{z} \mid \boldsymbol{\theta})\right] - D_{\text{KL}}\left(Q(\boldsymbol{\theta}) \| P(\boldsymbol{\theta})\right) \tag{2.25a}$$

$$\text{such that} \quad \int Q(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} = 1 \tag{2.25b}$$

In words, the above maximization problem seeks that probability distribution $Q(\boldsymbol{\theta})$ that simultaneously maximizes the likelihood while ensuring that the "distance" between the prior distribution and $Q(\boldsymbol{\theta})$ is minimum.

Rewriting equation (2.25),

$$\max_{Q(\boldsymbol{\theta})} \quad \mathbb{E}_{Q(\boldsymbol{\theta})}\left[\log P(\mathbf{z} \mid \boldsymbol{\theta})\right] - D_{\text{KL}}\left(Q(\boldsymbol{\theta}) \| P(\boldsymbol{\theta})\right) - \lambda\left(\int Q(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} - 1\right)$$

$$\max_{Q(\boldsymbol{\theta})} \quad \mathcal{I}[Q(\boldsymbol{\theta})] := \int Q(\boldsymbol{\theta}) \log P(\mathbf{z} \mid \boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{\theta} - \int Q(\boldsymbol{\theta}) \log \frac{Q(\boldsymbol{\theta})}{P(\boldsymbol{\theta})} \, \mathrm{d}\boldsymbol{\theta} - \lambda \left(\int Q(\boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{\theta} - 1 \right) \quad (2.26)$$

The problem can be solved by setting the variation of the objective with respect to $Q(\boldsymbol{\theta})$ to zero. This procedure leads to the corresponding Euler-Lagrange equation,

$$\begin{aligned} \frac{\mathrm{d}\mathcal{I}}{\mathrm{d}Q(\boldsymbol{\theta})} &= 0 \\ \log P(\mathbf{z} \mid \boldsymbol{\theta}) - \log Q(\boldsymbol{\theta}) - 1 + \log P(\boldsymbol{\theta}) - \lambda &= 0 \\ \log \left(P(\mathbf{z} \mid \boldsymbol{\theta}) P(\boldsymbol{\theta}) \right) &= \log(e^{\lambda+1} Q(\boldsymbol{\theta})) \\ Q(\boldsymbol{\theta}) &\propto P(\mathbf{z} \mid \boldsymbol{\theta}) P(\boldsymbol{\theta}) \end{aligned} \quad (2.27)$$

where, $e^{\lambda+1}$ is the normalizing factor for the probability distribution $Q(\boldsymbol{\theta})$. Equation (2.27) is essentially the statement of the Bayes theorem sans the normalization. This implies that the variational problem equation (2.25) is equivalent to computing the posterior distribution. Indeed, the variational Bayesian methods use this variational free energy to compute approximate posterior distributions, especially in the case of analytically intractable integrals [59].

As mentioned earlier, the posterior arises as the distribution that is a balance of maximizing the average log-likelihood function and minimizing the KL divergence between the prior and this distribution. In other words, since the likelihood function relates the data to the model and the prior distribution encodes prior beliefs/information, the variational problem seeks to find a balance between how much the data influences our updated belief (i.e. posterior distribution) while also accounting for any pre-existing belief. This interpretation of *influence* for the likelihood function

allows for choices of the likelihood function to be motivated on a physics basis, rather than a the absolute error between the model predictions and the data. In particular, for the problem at hand, for the pressure vs. internal radius data, two kinds of "distances" will be used to define the log likelihood function.

Let $\mathbf{p}_{\text{mod}}(\boldsymbol{\theta})$ be the model prediction for parameter vector $\boldsymbol{\theta}$, then the absolute error between the data and the model is given be $\mathbf{p}_{\text{mod}}(\boldsymbol{\theta}) - \mathbf{p}_{\text{exp}}$. The error between the tangent stiffnesses in the model and the data is given by $\frac{\mathrm{d}\mathbf{p}_{\text{mod}}(\boldsymbol{\theta})}{\mathrm{d}r} - \frac{\mathbf{p}_{\text{exp}}(i+1) - \mathbf{p}_{\text{exp}}(i-1)}{r(i+1) - r(i-1)}$ where $\mathbf{p}_{\text{exp}}(i)$ refers to the ith data point. The likelihood function proposed is given as

$$L(\boldsymbol{\theta}; \mathbf{p}_{\text{exp}}) = \exp\left(-\frac{(\mathbf{p}_{\text{mod}}(\boldsymbol{\theta}) - \mathbf{p}_{\text{exp}})^2}{\sigma^2} - \frac{\beta}{\sigma^2}\left(\frac{\mathrm{d}\mathbf{p}_{\text{mod}}(\boldsymbol{\theta})}{\mathrm{d}r} - \frac{\mathbf{p}_{\text{exp}}(i+1) - \mathbf{p}_{\text{exp}}(i-1)}{r(i+1) - r(i-1)} \right)^2 \right), \quad (2.28)$$

where $\frac{1}{\sigma^2}$ is the *precision* associated with the errors and β is the factor the controls the precision of the error in slope relative to the absolute error. The precision is inversely related to the tolerance in difference between model and data, i.e. lower tolerance for errors $\Rightarrow$ higher precision.

This likelihood measure includes the slope error in addition to absolute error,since the slope of the pressure - radius plot is directly related to the tangential stiffness of the artery. Clinically, the tangential stiffness of the artery is inferred from the pulse wave velocity. The pulse wave velocity is defined as the speed of a pulse in an arterial segment, and is thus directly related to the stiffness of the artery. It is known that this arterial stiffness is pressure-dependent [60, 61] i.e. the arterial stiffness is the local tangential stiffness. Therefore the slope is a quantity of interest, and errors in

slope will be penalized in addition to the errors in the values of the pressure.

2.4.3 Posterior distributions – Markov chain Monte Carlo sampling

The denominator of equation (2.23) is known as the probability of the evidence, and is given by

$$P(\mathbf{p}_{\text{exp}}) = \int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} P(\mathbf{p}_{\text{exp}} \mid \boldsymbol{\theta}) P(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta} \tag{2.29}$$

Since this integral is typically very hard to compute, we will use a Monte Carlo algorithm to solve the inference problem, without having to explicitly compute the above integral. The algorithm used is known as the Metropolis–Hastings algorithm [53, 54, 55]which is used to compute the posterior distribution $P(\boldsymbol{\theta} \mid \mathbf{p}_{\text{exp}})$.

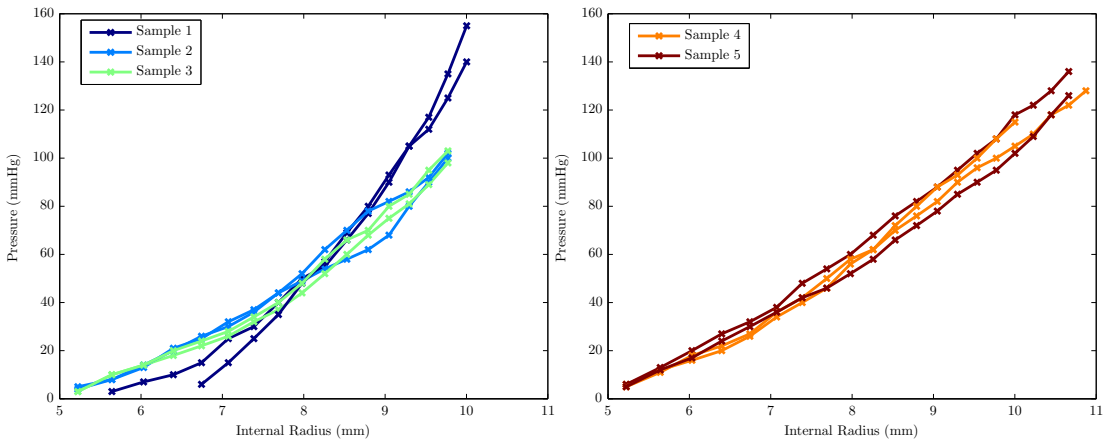
2.5 Results

The probability distribution of the model parameters are presented for a subset of the data that is shown in figure 2.1. This choice is in order to demonstrate an application of the probability distributions to a classification problem. Towards this, the samples are grouped into two classes - samples 1,2 and 3 into class C_1 and samples 4 and 5 into class C_2. For each of the classes, the data $\mathbf{p}_{\text{exp}}$ is chosen as two of the three experiments per sample. Therefore, for class C_1, $\mathbf{p}_{\text{exp}}$ is the set of six experiments (shown in figure 2.2 (a)) and for class C_2, $\mathbf{p}_{\text{exp}}$ is the set of four experiments (shown in figure 2.2(b)). The probability distribution for the parameters for the two classes corresponding to each of the strain energy functions from section (2.3.3) are shown in figure 2.3.

A cursory inspection of the probability distributions show that the regions of the parameter space with high probability are not the same for the two classes. Qualitatively, this observation shows that the two classes that we considered are

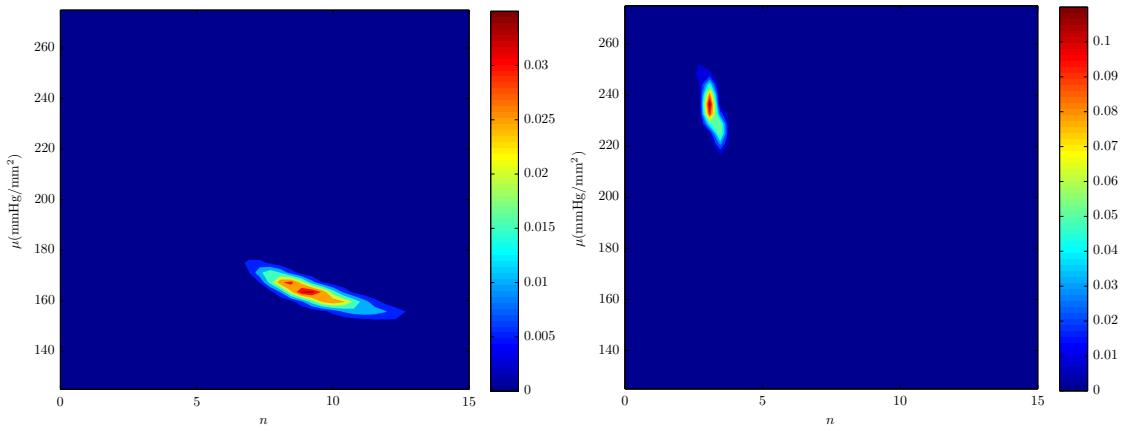
'different'. Indeed, a quantitative measure of this difference, as mentioned earlier, could be obtained by computing a distance such as the K-L divergence, between the probability distributions.

Another noteworthy feature in the model parameter probability distributions in figure 2.3 is that, for the Ogden model and the Criscione-type model, the probability distributions are bimodal. For the Ogden model, this is due to the fact that in equation 2.19, the strain energy is symmetric with respect to the sign of the exponent parameter α. It can be seen that this symmetry is reflected in the two modes in each of the contour plots in figure 2.3 (c) and (d). Similarly, for the Criscione-type model, the modes are due to the hyperbolic cosine function in the strain energy, since cosh is an even function. This is reflected in the symmetry in parameter b in figure 2.3 (e) and (f).



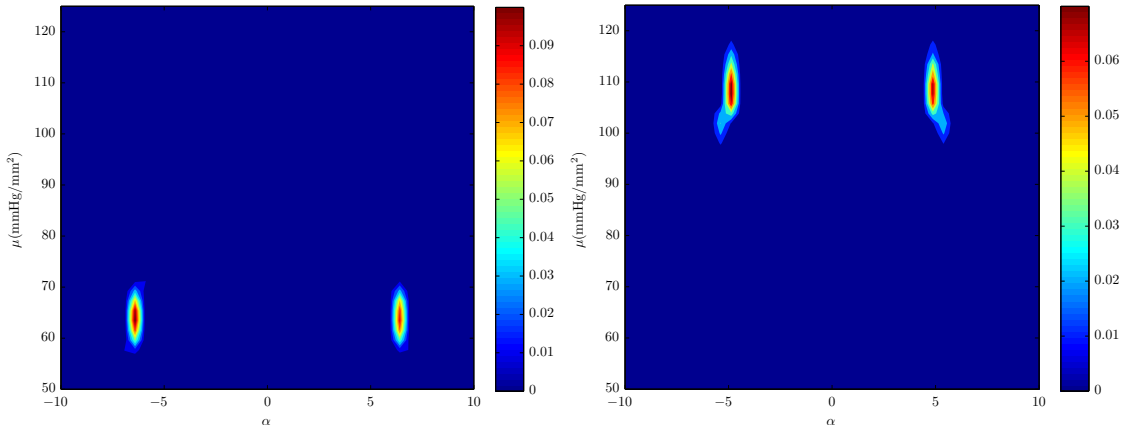
(a) Class C_1 (b) Class C_2

Figure 2.2: The experimental data shown as (a) class C_1 and class (b) class C_2. These two datasets are used to compute the probability distributions shown in figure 2.3.



(a) Samples 1,2,3 – Power-law neo-Hookean

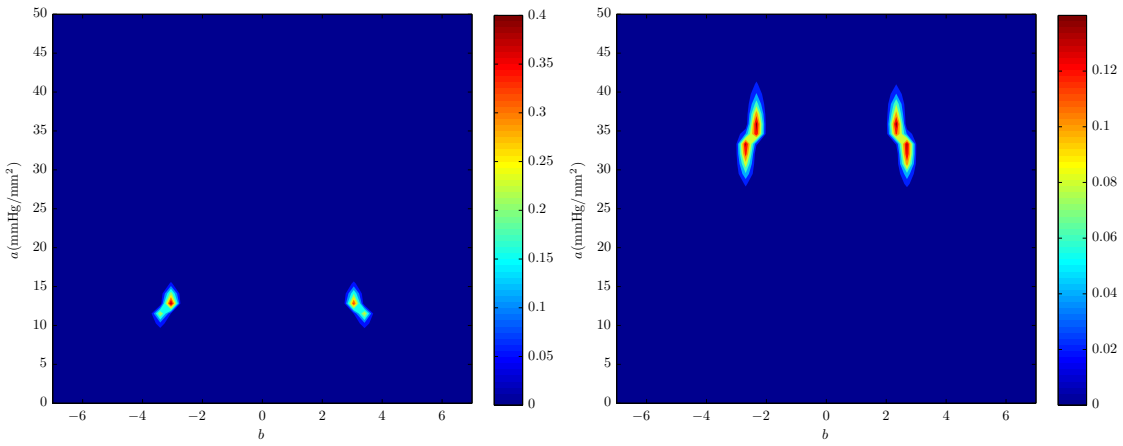
(b) Samples 4,5 – Power-law neo-Hookean



(c) Samples 1,2,3 – Ogden strain energy

(d) Samples 4,5 – Ogden strain energy

Figure 2.3: The probability distributions of the model parameters of the three strain energies using two out of three experiments for each class (see section 2.5) as data for each are shown.



(a) Samples 1,2,3 – Criscione-type strain energy (b) Samples 4,5 – Criscione-type strain energy

Figure 2.3: Continued

2.6 Using model parameter distributions for classification

An interesting application of the probability distributions for the model parameters is for classification problems. The techniques used for classification are typically of two types: deterministic or probabilistic. Deterministic algorithms describe hard boundaries between groups of observations, thus dividing the entire set of observations into clearly structured groups. Examples of such techniques are hypothesis testing, k -means clustering and computing separating hyperplanes using support vector machines. On the other hand, probabilistic algorithms relax such hard boundaries and instead assign the probability that a particular observation belonging to a certain class. Probabilistic techniques thus utilize a "fuzzy" classification approach. Examples of probabilistic algorithms for classification include logistic regression [62] and neural networks [63]. Probabilistic classification allows for naturally incorporating any uncertainty thus allowing for making decision only if sufficiently confi-

dent. This feature is especially critical in medical applications and thus probabilistic techniques have been the favored approach for applications in medical diagnosis [64, 65, 66, 67, 68].

The Bayesian framework naturally allows for assigning probabilities to each of the classes considered using the probability distributions computed earlier. Classifying the data *not* used as part of $\mathbf{p}_{\text{exp}}$ into classes C_1 and C_2 (see figure 2.2) is demonstrated as an example of the approach.

2.6.1 Class membership probabilities

Let classes $C_1, \ldots, C_n$ be classes into which a newly obtained data set $\mathbf{p}_{\text{new}}$ is to be classified. The classes C_i are each associated with a probability distribution for the model parameter $\boldsymbol{\theta}$, $P(\boldsymbol{\theta} \mid C_i)$. The probability distributions $P(\boldsymbol{\theta} \mid C_i)$ are obtained through a Bayesian inference procedure on the training data $\mathbf{p}_{\text{exp}}$ (see section 2.4.1).

For the classification problem, given a newly obtained data set $\mathbf{p}_{\text{new}}$, the probability that it belongs to a class C_i, $P(C_i \mid \mathbf{p}_{\text{new}})$ can be computed using the Bayes' rule as,

$$P(C_i \mid \mathbf{p}_{\text{new}}) = \frac{P(\mathbf{p}_{\text{new}} \mid C_i)P(C_i)}{P(\mathbf{p}_{\text{new}})} \tag{2.30}$$

In equation 2.30, the probability on the left hand side $P(C_i \mid \mathbf{p}_{\text{new}})$ is the *posterior* probability, $P(\mathbf{p}_{\text{new}} \mid C_i)$ is called the *likelihood* function and $P(C_i)$ is the *prior* probability.

The likelihood, $P(\mathbf{p}_{\text{new}} \mid i)$ is also known as the *marginal likelihood* due to the marginalization of $P(\mathbf{p}_{\text{new}} \mid i, \boldsymbol{\theta})$ over the parameter space $\boldsymbol{\Theta}$ (i.e.)

$$P(\mathbf{p}_{\text{new}} \mid C_i) = \int_{\boldsymbol{\Theta}} P(\mathbf{p}_{\text{new}}|\boldsymbol{\theta}, C_i) P(\boldsymbol{\theta} \mid C_i)\, \mathrm{d}\boldsymbol{\theta}$$

where, $P(\mathbf{p}_{\text{new}}|\boldsymbol{\theta}, C_i)$ is the likelihood associated with observing the data $\mathbf{p}_{\text{new}}$ given that the model parameters are $\boldsymbol{\theta}$ (see section 2.4.2 for details) and $P(\boldsymbol{\theta} \mid C_i)$ is the probability distribution on the parameter $\boldsymbol{\theta}$ associated with the class C_i.

Equation 2.30 now becomes,

$$P(C_i \mid \mathbf{p}_{\text{new}}) = \frac{\int_{\boldsymbol{\Theta}} P(\mathbf{p}_{\text{new}}|\boldsymbol{\theta}, C_i) P(\boldsymbol{\theta} \mid C_i)\, \mathrm{d}\boldsymbol{\theta}\, P(C_i)}{P(\mathbf{p}_{\text{new}})} \tag{2.31}$$

The results in figure 2.5 show the results of this classification applied to the data set shown in figure 2.4. The bar plot for each color represents the probabilities $P(C_1 \mid \mathbf{p}_{\text{new}})$ and $P(C_2 \mid \mathbf{p}_{\text{new}})$, and each color represents the data set that is used as $\mathbf{p}_{\text{new}}$ from 2.4.

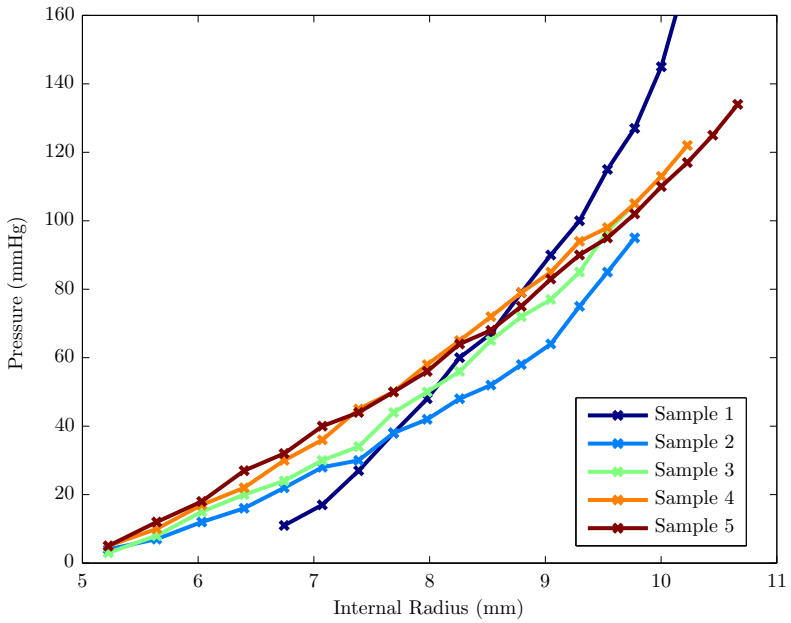
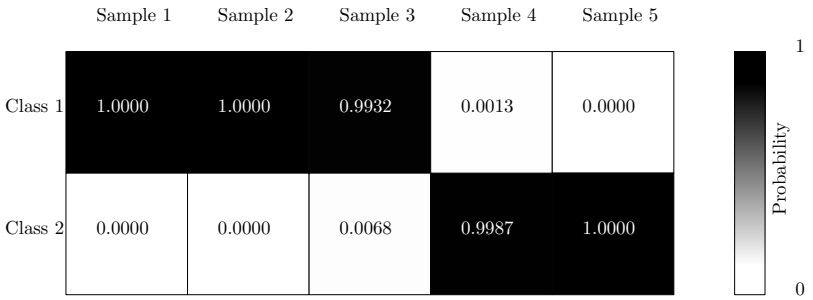
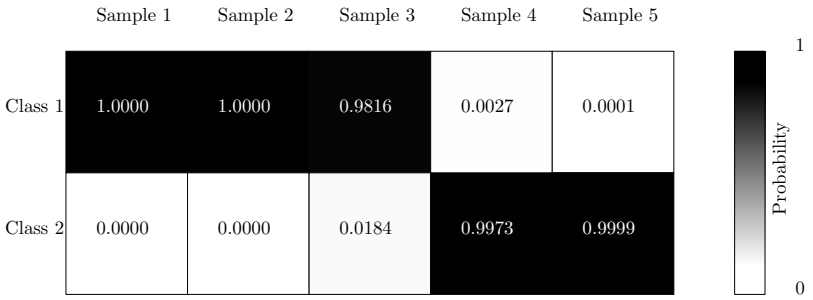


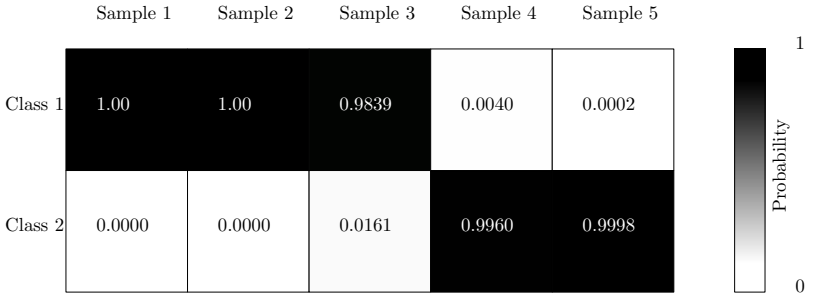
Figure 2.4: The data, $\mathbf{p}_{\text{new}}$, used for classification (see section 2.6.1). Each of the classes C_1 and C_2 experiments plotted here are assigned a probability (see figure 2.5).



(a) Power law neo-Hookean model



(b) Ogden model



(c) Criscione-type model

Figure 2.5: The class probabilities for the data set $\mathbf{p}_{\text{new}}$ and classified between class C_1 (class of samples 1,2,3) and C_2(class of samples 4,5). The row headings indicate the source of the data $\mathbf{p}_{\text{new}}$, although this information is *not used* in the classification procedure (section 2.6.1). Each of the figures correspond to the choice of model (strain energy functions) used in computing the marginal likelihood.

3. INHOMOGENEITY DETECTION IN CANTILEVER BEAMS

3.1 Introduction

In section 2, we presented a Bayesian approach to obtaining model parameter probability distributions as a means to characterize the material properties from arterial data. We showed that the probability distributions thus obtained may be used towards diagnostics and an example classifying (e.g. "healthy" vs. "diseased") newly obtained data was discussed in section 2.6.

On the other hand, if we are interested in performing diagnostics to identify diseased *regions* of a tissue sample, this may be addressed as an inhomogeneity detection problem. The presence of lower stiffness regions (in the case of aneurysms [21]) or higher stiffer regions (in the case of atherosclerotic plaques [27]) in arteries is well studied and thus serves as a motivation to perform such diagnostics. Since section 2 presents a probabilistic method to characterize the constitutive model's parameters from the data, characterizing the inhomogeneity can be performed using the model parameter probability distributions for different classes (e.g. healthy and diseased tissues). In this section, we show that this inhomogeneity detection problem can be solved as a parameter identification problem, by considering the model parameters for each of the regions as parameters of a structural mechanics problem. In other words, the two "stiffnesses" of the healthy and diseased regions need to be identified to detect the presence of inhomogeneities. We use the same approach presented in section 2 to solve this parameter identification problem. Since we do not have data from atherosclerotic arteries, we will present a simple non-trivial example to illustrate the approach.

We choose as an example, the large deformation bending of a cantilever beam

modeled as an elastica (see section 3.2) with an inclusion at a particular location to demonstrate the Bayesian inference approach introduced in section 1. The parameters in this problem are the bending stiffnesses of the inclusion and the material in the rest of the beam. In the approach presented here, we do not solve the inverse problem to obtain these parameters, but instead compute the likelihood of different values of parameters to explain the data. As in section 2, we will use Bayesian inference to obtain the joint probability distribution of these parameters.

This cantilever beam example is a simple, non-trivial example of parameter estimation for a nonlinear problem. Since the boundary value problem for the cantilever beam is straightforward to solve, it allows us to focus on the parameter identification approach rather than the solution of the nonlinear problem. Moreover, the approach presented here is modular in nature (i.e. independent of the solution of the nonlinear problem) and could be adapted for a variety of problems (provided a means can be found for solving the nonlinear boundary value problem efficiently), especially for materials which show nonlinear behavior and where the detection of such inhomogeneities are critical to characterizing the system.

3.2 Problem statement

Consider a cantilever beam AB fixed at end A and free at B, with an end load F_y at B. The beam is assumed to be an inextensible elastica made of two different materials, with bending stiffness k_1 and k_2. The beam is made of three sections, AC and DB with stiffness k_1 and CD with stiffness k_2.

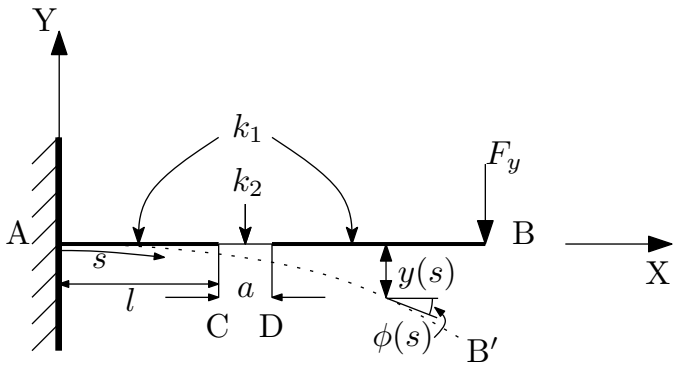


Figure 3.1: Cantilever beam AB considered for problem presented in section 3.2. Notice that the inhomogeneity in the beam is represented as a region of length a with stiffness k_2 at a distance l from A, while the remaining sections of the beam have a stiffness k_1. The dotted line represents the deformed shape, showing the arc-length coordinate s and Y-displacements $y(s)$.

The beam deforms to the curve AB′ under the action of the end load. Let s be the arc length coordinate, $y(s)$ be the Y-coordinate of a point on the curve AB′ and $\phi(s)$ be the angle made by the tangent to the current shape of the curve as shown in figure 3.1. If the stiffnesses k_1, k_2 are known, the shape $y(s)$ is computed by solving the following differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{k(s)\mathrm{d}\phi}{\mathrm{d}s}\right) + F_y \cos\phi = 0; \tag{3.1a}$$

$$\phi(0) = 0, \qquad \phi'(L) = 0 \tag{3.1b}$$

$$\text{where } k(s) = \begin{cases} k_1 & \text{if } 0 < s < l \text{ and } l + a < s < L, \\ k_2 & \text{if } l \le s \le l + a. \end{cases} \tag{3.1c}$$

$$\text{and } \frac{\mathrm{d}y}{\mathrm{d}s} = \sin\phi \tag{3.1d}$$

For the inference approach presented in this section, this differential equation

needs to be solved in order to compute the likelihood value, as detailed in section 3.3.

Let us assume that $\tilde{y}_i$, are the measured Y-deflections of the beam at uniformly distributed finite number of points $s_i, i = 1, \ldots, n$, along the length of the beam. Let the actual Y-deflections at these points, $y(s_i)$ be represented by y_i.

The following assumption regarding $\tilde{y}_i$ and y_i is made:

- The discrepancy between measurements, $\tilde{y}_i$, and the computed deflections, y_i, is attributed to a *measurement error* ε_i. We assume that the error ε_i are random variables that are independent and identically distributed according to a normal distribution[1] , i.e.

$$\tilde{y}_i = y_i + \varepsilon_i$$
$$\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$$

In order to now find the stiffnesses, we will treat k_1, k_2 as random variables, K_1, K_2 and compute the probability distribution of these random variables, given that the measurements $\tilde{y}_i$ were observed (i.e.) $P_{K_1,K_2}(k_1, k_2 \mid \tilde{\mathbf{y}})$.

Notice that this task is significantly different from querying those values of k_1, k_2, that best "fit" the data. The best fit parameters are those that are obtained by typically solving the following minimization problem,

$$(\hat{k}_1, \hat{k}_2) = \arg\min_{k_1,k_2} \|\tilde{\mathbf{y}} - \mathbf{y}(k_1, k_2)\| \tag{3.2}$$

[1]Since the sources of error are assumed to be external factors such as measuring instruments, independent of the beam itself, such an assumption is justified by the central limit theorem. Moreover, if the sources of error can be quantified more accurately, the probability distribution could be chosen accordingly without any change in the approach.

where the vectors $\tilde{\mathbf{y}}$ and $\mathbf{y}(k_1, k_2)$ are the measured Y-displacements and Y-displacements obtained from the solution of equation (3.1). Such a technique (or a variant thereof) is common in parameter estimation problems. When the model, $\mathbf{y}(k_1, k_2)$ is nonlinear, the minimization problem (3.2) is not guaranteed to have unique minimum, and the estimates thus obtained (as local minima) have issues of robustness.

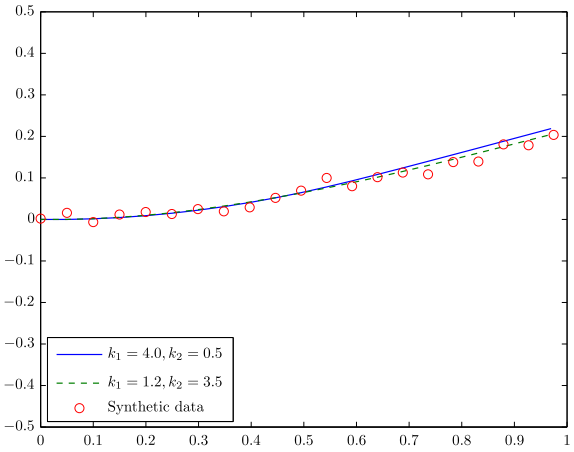


Figure 3.2: Comparison of shapes of two beams : a stiff beam with a compliant inclusion, and a compliant beam with stiff inclusion. Note that the beams show very similar deflections for widely different values of the parameters. The synthetic data was generated with $(k_1, k_2) = (1.4, 1.8)$.

Consider, for example, the computed deflections of the beam for two significantly different sets of stiffnesses (k_1, k_2) as shown in figure 3.2. Without any prior knowledge, both the stiffnesses seem a likely match to the data (which was generated with $(k_1, k_2) = (1.4, 1.8)$). This likelihood is the measure represented by the probability distribution. It is to be noted that such information is impossible to obtain from a

simple curve fit, and hence large errors may be committed in the process of finding single "best fit" parameters for such problems.

In adopting a probabilistic approach, we circumvent such issues of uniqueness by assigning a probability to values of the parameters. The reformulation of the inverse problem as a probabilistic Bayesian inference problem, and the details of the algorithm used to numerically compute this probability distribution is given in the following section.

3.3 Inference problem

The goal of the inference problem is to obtain the probability distribution of the parameters given the observations, i. e. $P_{K_1,K_2}(k_1, k_2 \mid \tilde{\mathbf{y}})$, which gives the probability that random variables K_1, K_2 take the values k_1, k_2 given the observations $\tilde{\mathbf{y}}$. In this section, we will state this inference problem in a Bayesian inference framework as well as explain the Metropolis–Hastings algorithm.

3.3.1 Bayesian inference

Bayesian inference is a method of statistical inference that uses the Bayes rule to compute the probability of a hypothesis in the presence of evidence. For the cantilever problem, the Bayes rule can be written as

$$P(k_1, k_2 \mid \tilde{\mathbf{y}}) = \frac{P(\tilde{\mathbf{y}} \mid k_1, k_2)P(k_1, k_2)}{\int\limits_{k_1,k_2} P(\tilde{\mathbf{y}} \mid k_1, k_2)P(k_1, k_2)} \tag{3.3}$$

The left hand side of eqn. (3.3) is the probability distribution that we seek and is known as the *posterior distribution*. The denominator on the right hand side is a normalizing constant and is independent of the parameters K_1, K_2. The numerator on the right hand side is the product of the *likelihood function,* $P(\tilde{\mathbf{y}} \mid K_1, K_2)$ and the *prior distribution* $P(K_1, K_2)$. The likelihood function is the probability of obtaining

the observations $\tilde{\mathbf{y}}$, for a given set of values of k_1, k_2. We know from equation (3.1) that given a set of values k_1, k_2, we can obtain the shape of the beam $y(s)$. Therefore, the likelihood function is the probability of observing the observations $\tilde{\mathbf{y}}$ given that the actual shape is $y(s)$. In other words, since the measurements were made at points $s_1, \ldots, s_n$ along the beam, the likelihood function is $P(\tilde{\mathbf{y}} \mid \mathbf{y}) = P(\tilde{y}_1, \ldots, \tilde{y}_n \mid y_1, \ldots, y_n)$.

Since we have assumed that any difference between the observations and the model is only due to an additive noise, the likelihood function is thus directly related to the errors in measurement, ε_i, mentioned in the assumptions under section 3.2. Indeed,

$$P(\tilde{\mathbf{y}} \mid \mathbf{y}) = \prod_i P(\tilde{y}_i \mid y_i) = \prod_i P(\varepsilon_i) \tag{3.4}$$

$$\text{where, } P(\varepsilon_i) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(\tilde{y}_i - y_i)^2}{\sigma^2}\right). \tag{3.5}$$

The solution of the boundary value problem (3.1), is required to compute y_i, which in turn is used to compute $P(\varepsilon_i)$ and hence the likelihood, $P(\tilde{\mathbf{y}} \mid \mathbf{y})$.

3.3.2 Prior probabilities and the maximum entropy principle

In equation (3.3) $P(k_1, k_2)$ is the *prior probability density function* and this is the probability that encodes our beliefs about the values of the parameters as a distribution, before the experiment was conducted. For example, if we choose to incorporate no prior information or very minimal information such as the fact that the two stiffnesses are in a certain domain, we would choose the prior probability density to be a uniform distribution in this domain. On the other hand, if there is some prior information that is not just the domain of the parameters, but indeed that

which relates the parameters, it is advantageous to incorporate it into our inference process.

In order to emphasize the role of prior information and their significance, we reproduce an illustrative example from L. J. Savage (from one of his communications to Jerome Cornfield, a biostatistician) [69, 70]:

> ... Imagine the following three experiments: 1. Fisher's lady has correctly dealt with ten pairs of cups of tea. 2. The professor of 18th century musicology at the University of Vienna has correctly decided for each of 10 pairs of pages of music which was written by Mozart and which by Hayden. 3. A drunk in a parlor car has succeeded 10 times in correctly calling a coin secretly tossed by you. These three experiments all have the same mathematical structure and the same high significance level. Can there, however, be any question that your reaction to them is justifiably different? My own would be: 1. I am still skeptical of the lady's claim, but her success in her experiment has definitely opened my mind. 2. I would originally have expected the musicologist to make this discrimination; I would even expect some success in making it myself; he, an expert in the matter, felt sure that he could make it. His success in 10 correct trials confirms my original judgment and leaves no practical doubt that he would be correct in substantially more than half of future trials, though I would not be surprised if he made occasional errors. 3. My original belief in clairvoyance was academic, if not utterly nonexistent. I do not even believe that the trial was conducted in such a way that trickery is a plausible hypothesis, and feel sure that the drunk simply had an unusual run of luck.[2] ...

[2]This paragraph has been reproduced from [70] section 2.4

In summary, it is evident from Savage's explanation, that any prior belief affects the inference made from the observations, and it would be a more accurate inference problem to incorporate such prior beliefs.

In cases where the prior information about the parameters can be expressed in the form of a distribution, such as, 'k_1, k_2 are each distributed according to an exponential distribution', then one could obtain closed-form solutions for the posterior distribution. In typical engineering scenarios, information such as the expected values of the parameters or their variances is more common. Jaynes [71, 72, 73] laid out a systematic way of incorporating such information into the Bayesian inference process through the *Maximum Entropy principle.*

The details of computing a probability distribution through the maximum entropy principle (i.e. the MaxEnt distribution) are given in appendix A.

For the cantilever problem at hand, we will compare two prior probability distributions:

1. The prior distribution is a uniform distribution over a domain Ω_k, (i.e)

$$P(k_1, k_2) = \begin{cases} c, & k_1, k_2 \in \Omega_k \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

   where c is a constant, such that

$$\iint_{\Omega_k} P(k_1, k_2)\, \mathrm{d}k_1 \mathrm{d}k_2 = 1,$$

2. The prior probability is a MaxEnt distribution, where we assume that the expected value and variance of k_1 is known, but no prior information on k_2 is available. This assumption essentially means that we are trying to identify

the stiffness of an inclusion of unknown material, in a beam of a some known material. In this case, k_1 and k_2 are independent and the MaxEnt distribution is only on k_1 since we assume k_2 has a uniform distribution. We further assume that both k_1 and k_2 lie in an interval $[0, \hat{k}]$. Therefore,

$$P(k_1, k_2) = P(k_1)P(k_2)$$

where,

$$P(k_1) = \arg\max_{P(k_1)} \int H(k_1)\,\mathrm{d}k_1 \tag{3.7}$$

$$\text{subject to} \int_0^{\hat{k}} k_1\, P(k_1)\,\mathrm{d}k_1 = E[K_1] = \langle k_1 \rangle, \tag{3.8}$$

$$\int_0^{\hat{k}} k_1^2\, P(k_1)\,\mathrm{d}k_1 - \langle k_1 \rangle^2 = \mathrm{Var}[K_1] \tag{3.9}$$

$$\text{and} \int_0^{\hat{k}} P(k_1)\,\mathrm{d}k_1 = 1$$

and,

$$P(k_2) = \begin{cases} \frac{1}{\hat{k}} & 0 \le k_2 \le \hat{k} \\ 0 & \text{otherwise} \end{cases} \tag{3.10}$$

3.3.3 Markov chain Monte Carlo method – Metropolis–Hastings sampling

We now know the prior probability distribution and the likelihood function, but the computation of the integral in the denominator of (3.3), is computationally expensive and analytically intractable. To circumvent this we will use the Metropolis–

Hastings algorithm to efficiently sample the parameter space and build a population from the posterior distribution.

Rewriting (3.3), the posterior distribution $P(k_1, k_2 \mid \tilde{\mathbf{y}})$ is

$$P(k_1, k_2 \mid \tilde{\mathbf{y}}) = \frac{P(\tilde{\mathbf{y}} \mid \mathbf{y}(k_1, k_2))P(k_1, k_2)}{\int\limits_{k_1,k_2} P(\tilde{\mathbf{y}} \mid \mathbf{y}(k_1, k_2))P(k_1, k_2)}, \tag{3.11}$$

where the prior probability distribution is $P(k_1, k_2)$ and the likelihood term $P(\tilde{\mathbf{y}} \mid \mathbf{y}(k_1, k_2))$ is computed by first solving the differential equation (3.1) and then using this in equation (3.5). We will use a numerical sampling method known as Metropolis–Hastings sampling [54] to populate a sample such that it has the distribution $P(k_1, k_2 \mid \tilde{\mathbf{y}})$. This method is a very well-established technique and popularly used in applications of Bayesian inference.

Metropolis–Hastings method belongs to a larger class of methods known as Markov Chain Monte Carlo methods, which construct samples by constructing a Markov Chain whose stationary distribution is the target distribution that is sought. The criterion to accept a 'candidate' $\theta_{\text{cand}} = [k_1, k_2]_{\text{cand}}$, given that the previous sample (i.e. i^{th} sample) is $\theta_i = [k_1, k_2]_i$, is defined as,

$$\alpha(\theta_i \mid \theta_{\text{cand}}) = \min\left(1, \frac{P(\theta = \theta_{\text{cand}} \mid \tilde{\mathbf{y}})q(\theta_i \mid \theta_{\text{cand}})}{P(\theta = \theta_i \mid \tilde{\mathbf{y}})q(\theta_{\text{cand}} \mid \theta_i)}\right) \tag{3.12}$$

Notice that this iterative sampling requires evaluating the value of the target distribution. Using Bayes rule from equation (3.3), we can rewrite the acceptance criterion as,

$$\alpha(\theta_i \mid \theta_{\text{cand}}) = \min\left(1, \frac{P(\tilde{\mathbf{y}} \mid \theta = \theta_{\text{cand}})P(\theta = \theta_{\text{cand}})q(\theta_i \mid \theta_{\text{cand}})}{P(\tilde{\mathbf{y}} \mid \theta = \theta_i)P(\theta = \theta_i)q(\theta_{\text{cand}} \mid \theta_i)}\right) \tag{3.13}$$

where $q(\theta_{\text{cand}} \mid \theta_i)$, known as the proposal distribution, is the probability of choosing θ_{cand} as the candidate sample when the current iterate is θ_i. Notice now that the integral in the denominator of the Bayes rule does not feature in this acceptance criterion. This avoids expensive computations and is thus an efficient way to build the posterior probability distribution.

It is common to choose the proposal distribution to be symmetric [74], i.e. $q(\theta_i \mid \theta_{\text{cand}}) = q(\theta_{\text{cand}} \mid \theta_i)$. The acceptance criterion now reduces to

$$\alpha(\theta_i \mid \theta_{\text{cand}}) = \min\left(1, \frac{P(\tilde{\mathbf{y}} \mid \theta = \theta_{\text{cand}})P(\theta = \theta_{\text{cand}})}{P(\tilde{\mathbf{y}} \mid \theta = \theta_i)P(\theta = \theta_i)}\right) \tag{3.14}$$

Note here that for every candidate, the boundary value problem in equation (3.1) needs to be solved to compute the likelihood value, $P(\tilde{\mathbf{y}} \mid \theta = \theta_{\text{cand}})$. The pseudo-code for this sampling technique is shown as Algorithm 1

The samples thus obtained are used to construct histograms and using these histograms as empirical probability distributions, we can compute the expectations or the probabilities associated with hypotheses for diagnostics. The following sections detail the particular aspects of the simulations performed to demonstrate the inference technique hitherto presented.

3.4 Results and discussion

As detailed in section 3.2, the cantilever is assumed to be a nonlinear elastica and the boundary value problem in equation 3.1 is solved by a numerical shooting method to obtain the deflections of the beam. This problem is chosen as an illustra-

Algorithm 1 Metropolis–Hastings algorithm

Initialize $\theta_0 = 1$; **set** $i = 0$
while $i <$ MaxIter **do**
  Sample a point θ_{cand} according to $q(\cdot \mid \theta_i)$
  Solve boundary value problem (3.1) with θ_{cand} as the stiffnesses and compute y_i
  Compute likelihood values $P(\tilde{\mathbf{y}} \mid \theta = \theta_{\text{cand}})$, $P(\tilde{\mathbf{y}} \mid \theta = \theta_i)$ (eqns. (3.4),(3.5))
  Compute priors $P(\theta = \theta_{\text{cand}})$, $P(\theta = \theta_i)$ (eqns (3.6) or (3.8),(3.10))
  Compute value of acceptance criterion $\alpha(\theta_i \mid \theta_{\text{cand}})$ (eqn (3.14))
  Sample a uniform random variable $u \sim \mathrm{U}(0, 1)$
  if $u \leq \alpha(\theta_i \mid \theta_{\text{cand}})$ **then**
    $\theta_{i+1} = \theta_{\text{cand}}$
  else
    $\theta_{i+1} = \theta_i$
  set $i = i + 1$

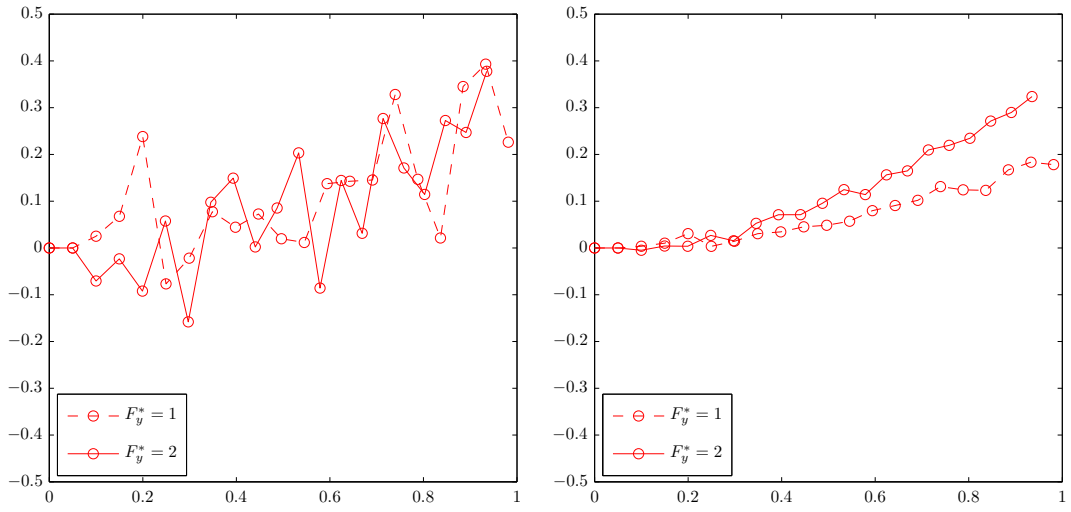
tive example to show the use of Bayesian inference in order to build a probabilistic representation of the stiffnesses for data and also address absolute criteria for diagnostics of the type detailed in section 3.1. The data for this problem is synthetically generated by adding Gaussian noise to the solution of the differential equation for a particular value of stiffnesses. This synthetic data is used in the Bayesian inference procedure to obtain a joint probability distribution for the stiffnesses as explained in section 3.3.

3.4.1 Simulation details

The inference procedure was carried out for a cantilever beam of length 1 unit with an inhomogeneity of width 0.2 unit between 0.5 and 0.7 units. Two sets of observations where the tip load on the beam is $F_y^* = 1$ (non-dimensional, $F_y^* = \frac{F_y L^2}{k}$) and $F_y^* = 2$ are used for the inference process. The beam is discretized into 20 grid points and the "noisy" data $\tilde{y_i}$ was synthetically generated by adding a noise $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$, to the solution of the differential equation with the stiffnesses set to

$k_1 = 2.5$ and $k_2 = 1$. The inference procedure was carried out for two values of the noise variance $\sigma^2 = 10^{-2}$ and $\sigma^2 = 10^{-4}$. Two sets of data points were used to perform the inference procedure : (1) the measurements of Y-deflections along the entire beam and (2) the measurements only at the free end of the beam. In the computation of the prior probability distribution, the stiffnesses are assumed to be within a square region $\Omega_k = [0, \hat{k}] \times [0, \hat{k}]$ where $\hat{k} = 5$.

The joint probability distribution of k_1, k_2 is compiled using the samples generated through the Metropolis–Hastings algorithm. As an example of diagnostics, we will use this joint probability density function to compute the probability $P(\frac{k_1}{k_2} \geqslant \alpha | \tilde{y}_i)$. The criterion $\frac{k_1}{k_2} \geqslant \alpha$ could be thought of as a diagnostics criterion for rejection/acceptance.



(a) $\sigma^2 = 10^{-2}, F_y^* = 1$ and $F_y^* = 2$ (b) $\sigma^2 = 10^{-4}, F_y^* = 1$ and $F_y^* = 2$

Figure 3.3: The figure shows the the set of observations at a finite number of points along the beam (in red circles) for different values of the error variance, σ^2 and for two experiments, one with applied tip load $F_y^* = 1$ and the other with $F_y^* = 2$.

Figure 3.3 shows the measurements from two sets of experiments for each error variance value. The probability distributions that we will compute will use data from either only one experiment, or both experiments, to observe the influence of the number of experiments on the inference process. Further, we will assume two cases, one where the dataset is the list of all deflections, and the other, with only the tip deflections.

For each of the cases chosen, the uniform prior and the MaxEnt prior distributions (with the constraints on k_1 imposed as $E[K_1] = 2.5$ and $\mathrm{Var}[K_1] = 0.5$) are used to compute the posterior distribution. The empirical probability mass distributions (histograms) are shown as contour maps below.

Figures 3.4 and 3.5 show the posterior distributions when only one of the two datasets (the one corresponding to $F_y^* = 1$) is used as the data for the inference procedure. Figure 3.6 and 3.7 show the posterior distributions when both datasets ($F_y^* = 1$ and $F_y^* = 2$) are used.

The key point to note is that the distributions shown in figures 3.4, 3.5, 3.6 and 3.7, do not have any similarity in features to a normal distribution or any other conventional heavy-tailed distributions. This suggests that assumptions on the distribution of the stiffnesses (normal distributions, for example) are not appropriate, and thus inference using such assumptions will lead to incorrect results.

As we mentioned earlier, the posterior distributions obtained through the proposed approach, allow us to circumvent any issues of uniqueness in parameter values and rather assigns probabilities to the parameter values indicating relative likelihood of explaining the data. As opposed to best-fit parameter estimates which do not encode this information, a probability distribution assigns higher probabilities to the stiffness values which show better fit to the data. For example, figure 3.8 shows the comparison of two samples, both chosen from a high probability region,

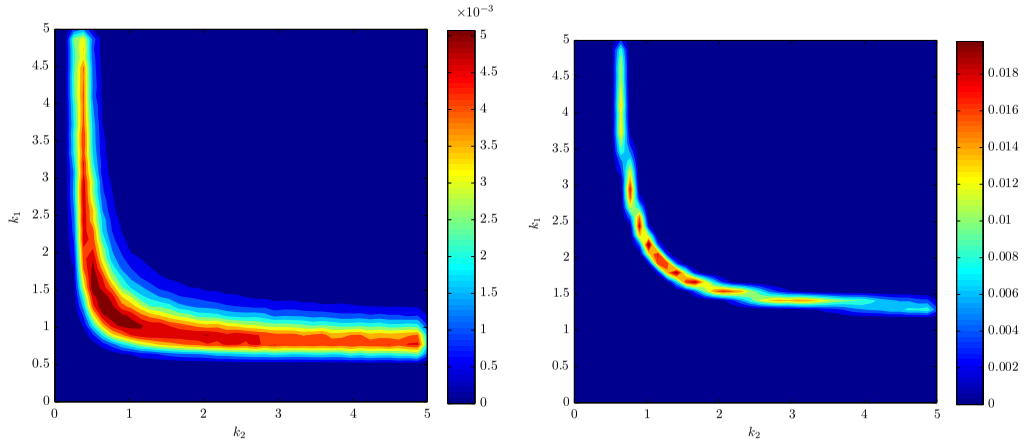
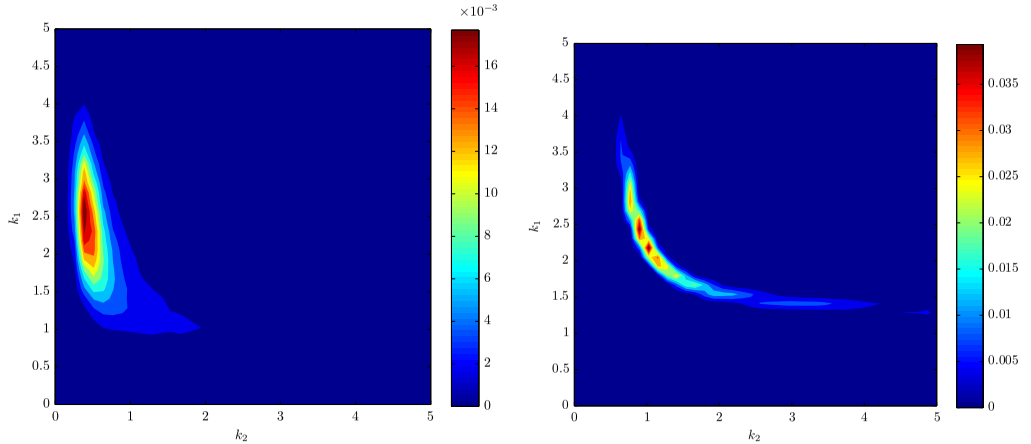
One Dataset	All measurements
Uniform Prior	 (a) $\sigma^2 = 10^{-2}$ (b) $\sigma^2 = 10^{-4}$
MaxEnt Prior	 (c) $\sigma^2 = 10^{-2}$ (d) $\sigma^2 = 10^{-4}$

Figure 3.4: The figure shows the (empirical) joint probability mass distribution for the case when all the deflection measurements are used and the data used is from one bending experiment (i.e. figure 3.3 (a)).

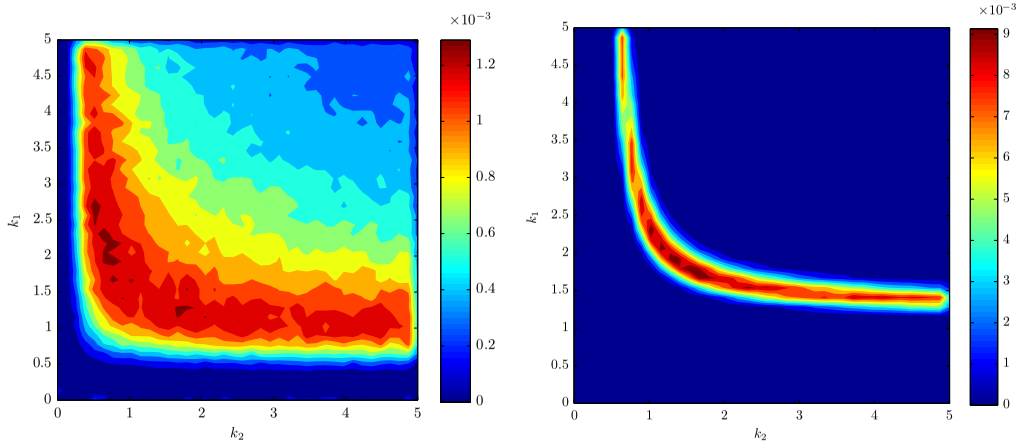
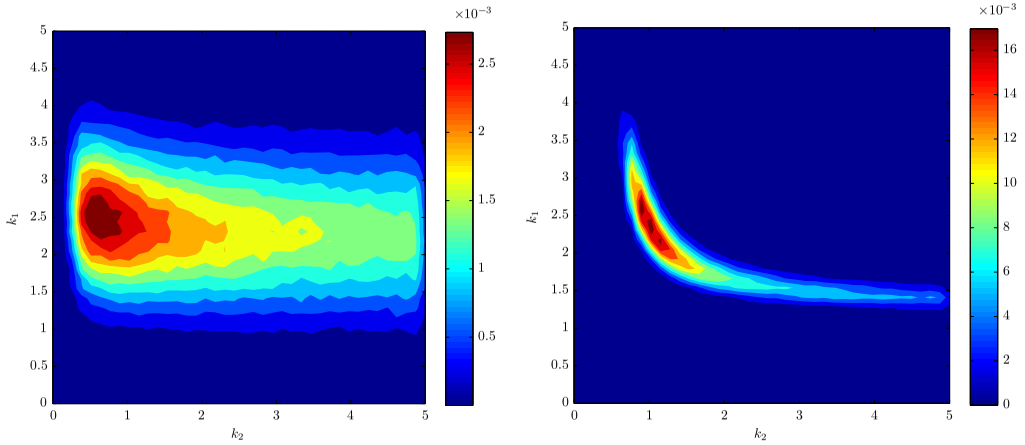
One Dataset	Tip deflections only
Uniform Prior	 (a) $\sigma^2 = 10^{-2}$ (b) $\sigma^2 = 10^{-4}$
MaxEnt Prior	 (c) $\sigma^2 = 10^{-2}$ (d) $\sigma^2 = 10^{-4}$

Figure 3.5: The figure shows the (empirical) joint probability mass distribution for the case when only the tip deflection measurements are used and the data used is from one bending experiment (i.e. figure 3.3 (a)).

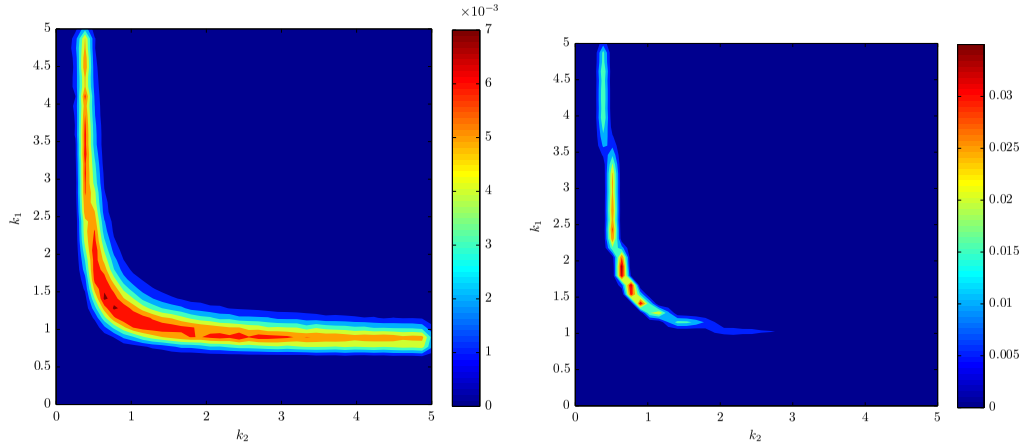
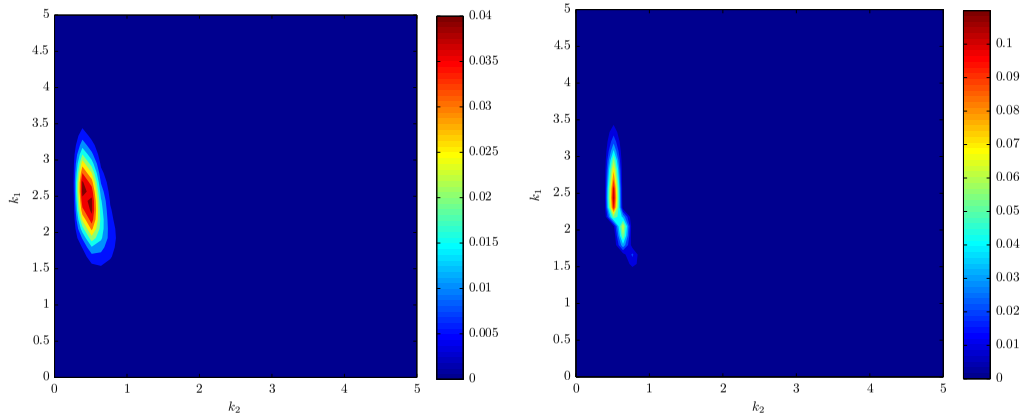
Two Datasets	All measurements
Uniform Prior	 (a) $\sigma^2 = 10^{-2}$ (b) $\sigma^2 = 10^{-4}$
MaxEnt Prior	 (c) $\sigma^2 = 10^{-2}$ (d) $\sigma^2 = 10^{-4}$

Figure 3.6: The figure shows the (empirical) joint probability mass distribution for the case when all the deflection measurements are used and the data used is from two bending experiments (i.e. figure 3.3 (a) and (b)).

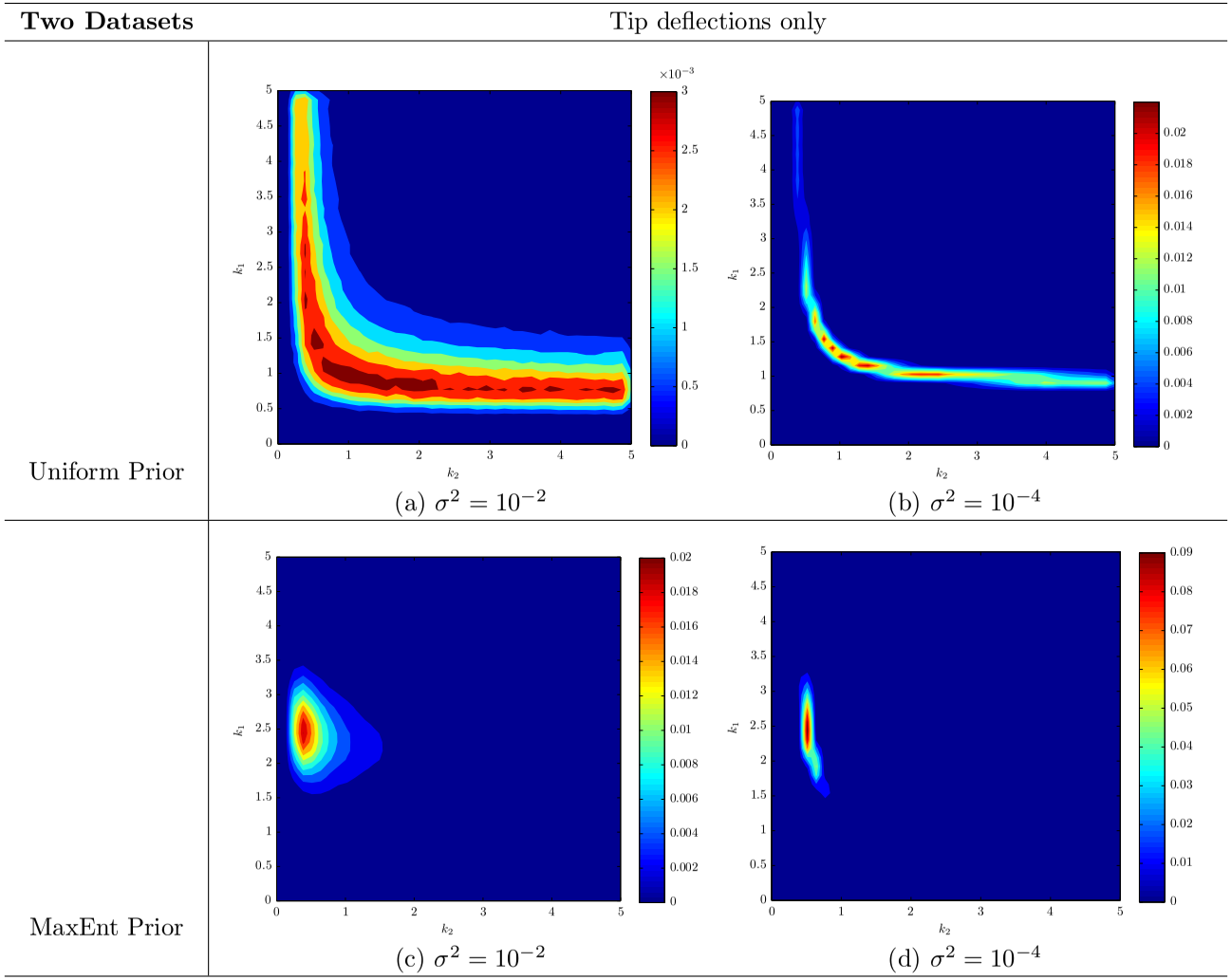
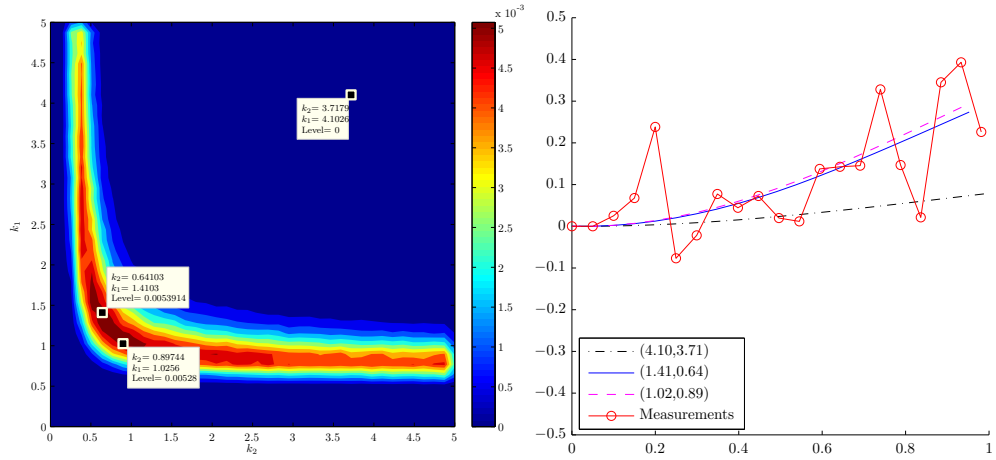


Figure 3.7: The figure shows the (empirical) joint probability mass distribution for the case when only the tip deflection measurements are used and the data used is from one bending experiment (i.e. figure 3.3 (a) and (b)).



(a) (b)

Figure 3.8: Comparison of shapes with parameters having similar posterior probability values. The figure on the left shows a choice of two samples from the higher probability region and one sample from the lower probability region, and the figure on the right shows the comparison of the deformed shapes of the corresponding beams : pink and blue shapes are for the two samples from the higher probability region and the black shape corresponds to the sample from the lower probability region.

with a sample chosen from a low probability region. The deformed shapes for these samples were computed using equation (3.1). Comparing the shapes with the data, it can be seen that the match between the samples which have higher probabilities assigned is much better than the low probability sample. More importantly, by assigning probabilities that correspond to the match with the data, we have obtained a "full-field" solution over the parameter space as opposed to picking a single best fit set of parameters.

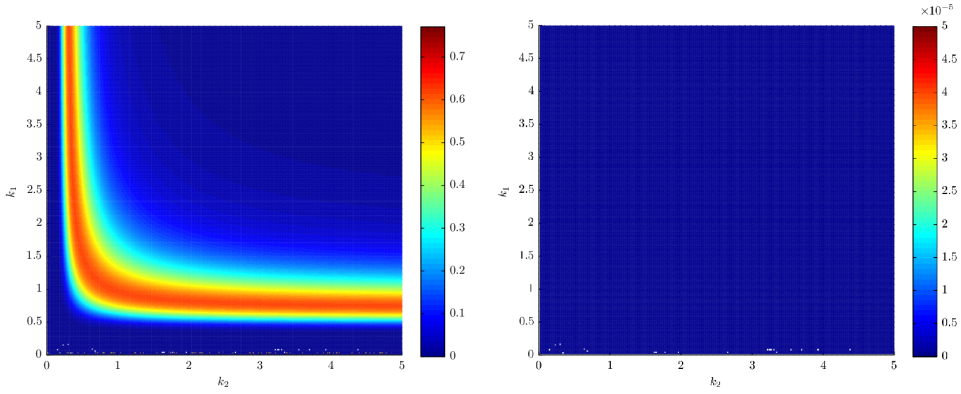
Furthermore, even if a weak prior information is available on one of the materials on the beam, it can be seen that the posterior distributions are centered around a smaller region in the parameter space. This reiterates the point made by L. J. Savage about prior information, and directly applies to such Bayesian inference problems.

It is clear that the role of the prior knowledge is to assign higher probability in favor of known information and cut out those portions of the posterior distribution (i.e. the one obtained using a uniform prior), that do not support the prior knowledge.

In terms of the number of data points used, it can be seen that accurate measurements of just the tip deflections from two experiments is comparable to having the full field displacements from one experiment. This can be explained by the fact that in general, tip deflects more than the root of the beam, and hence the measurements closer to the root are not as critical as those near the tip.

Remark. At this point we would like to make a remark about the use of the Metropolis–Hastings algorithm to sample the posterior as opposed to numerically evaluating the value of the posterior (i.e. the numerator on the right hand side of equation (3.11)) and subsequently normalizing it. Such an approach is possible, if the regions of non-zero probabilities are significantly large in the domain of interest. When the regions are small (as for example, in figure 3.6 (b)) this approach fails even when the number of grid points are comparable to the number of samples with the MH algorithm. As an illustration of this issue, the posterior distributions (unnormalized) are shown in figure 3.9. The corresponding distributions from using the sampling technique are figures 3.6 (a) and (b).

In addition to making such inferences, the probability distributions obtained using Bayesian inference could be used to test against diagnostic criteria, as discussed earlier, such as $\frac{k_1}{k_2} \geqslant \alpha$. Such diagnostic criteria are commonplace in medicine, for example, to diagnose the health of blood vessels. In particular, aneurysms and atherosclerotic arteries are known to have a significantly different stiffness from that of healthy arteries [21, 22, 27, 28]. We illustrate such a procedure using a diagnostic criterion $\frac{k_1}{k_2} \geqslant \alpha$ for $\alpha = 2$, i.e. we seek to compute the probability that the inclusion



(a) $\sigma^2 = 10^{-2}$ (b) $\sigma^2 = 10^{-4}$

Figure 3.9: The figure shows the unnormalized posterior distributions computed numerically (i.e. without sampling). Compare (a) and (b) with figures 3.6(a) and 3.6(b). Notice that the region of non-zero probability is much smaller in figure 3.6(b) than figure 3.6(a), and while the sampling method samples this region, simple numerical evaluation of the posterior on a grid is not able to represent this, as seen in (b).

is half as stiff as the rest of the beam. The results are tabulated in table 3.1. These results reiterate the observations made earlier, such as the benefit of using an appropriate prior distribution (marked by the increased probability for the criterion) and the fact that good quality (i.e. noise-free) tip deflection data contains as much information as all the deflections combined.

The use of such a criterion, especially in the presence of good prior knowledge, allows us to eliminate false positives, as can be seen from the tabulated results. It is also important to note that while there are other methods to obtain probability distributions for parameters such as the polynomial chaos expansion (PCE) [75], it is possible only to obtain the mean and variance of the random variables involved.

The use of a probability distribution to compute the probability and test it against this criterion is only a sample of the use of such distributions. We believe that this perspective allows one to quantify uncertainty in a systematic way. As discussed

Prior	σ^2	$P\left(\frac{k_1}{k_2} \geqslant 2 \mid \tilde{\mathbf{y}}_{\text{tip}}\right)_1$	$P\left(\frac{k_1}{k_2} \geqslant 2 \mid \tilde{\mathbf{y}}_{\text{tip}}\right)_2$	$P\left(\frac{k_1}{k_2} \geqslant 2 \mid \tilde{\mathbf{y}}\right)_1$	$P\left(\frac{k_1}{k_2} \geqslant 2 \mid \tilde{\mathbf{y}}\right)_2$
Uniform Prior	10^{-2}	0.2782	0.3171	0.3411	0.3525
	10^{-4}	0.3251	0.2659	0.3836	0.6274
Max Ent Prior	10^{-2}	0.2735	0.6840	0.6708	0.9243
	10^{-4}	0.3572	0.9258	0.4661	0.9764

Table 3.1: Probability of the criterion $P\left(\frac{k_1}{k_2} \geqslant 2 \mid \tilde{\mathbf{y}}\right)$ calculated from the probability mass distributions shown in figures 3.4, 3.5 and 3.6, 3.7. For both the prior distributions, the value of $P\left(\frac{k_1}{k_2} \geqslant 2\right) = 0.25$.

earlier the probability distribution of parameters paves for a new approach to diagnostics and allows one to incorporate not just the "best" fit values of parameters for comparison but also the inherent uncertainty, be it through measurement errors or otherwise, into the diagnostics process.

4. SUMMARY

This dissertation presents an alternate approach to identifying model parameters for data that shows large variations in response. It can be seen (figures 1.1 and 2.1) that such variations are not only due to measurement errors but such variability might well be a feature of the response. Given the uncertainty in the data and the induced uncertainty in developing models, we proposed a probabilistic approach in this work. We represent model parameters as probability distributions instead of point estimates. We use a Bayesian inference procedure to obtain these probability distributions

In this dissertation, we apply this approach to two example problems : detecting the an inhomogeneity in a nonlinear cantilever beam and characterizing sheep aorta using inflation data.

- The inhomogeneity detection problem is presented mainly as a simple non-trivial example to demonstrate the core philosophy of this work. The probability distribution of the stiffnesses in the cantilever beam serves as more than just a characterization of the system, providing insight into the effect of noise, amount of data and the use of any available prior information. We use Jaynes' [73] principle of maximum entropy to encode the prior information as a prior probability distribution.It is important to note that incorporating such prior information as constraints into the least squares problem is non-trivial.

- To characterize the inflation data from sheep aorta we use the Bayesian inference approach to obtain the model parameter probability distributions for a choice of three constitutive models. In the Bayesian inference procedure we modify the likelihood factor by introducing a distance measure to compare the

slope of data and model predictions in addition to the absolute error. For nonlinear models, such a measure penalizes severe departure in slope, and furthermore in the context of biological materials, this penalty has a physical meaning since the slope of the force–displacement response is directly related to the wave speed in the material. The probability distribution on the model parameters are subsequently used to classify "new" data into predefined classes. This procedure is also performed using Bayesian inference, thus evolving into a consistent Bayesian approach to characterize and classify experimental data.

4.1 Directions for future research

The following research topics are possible directions to continue the work presented in this dissertation:

1. The literature lists several hyperelastic models developed for biological materials, with particular reference to the anisotropy in observed response [19, 76, 77]. Moreover the modeling of biomaterials using finite element method is also presented in the literature [78, 79, 80]. Such models can be used in conjunction with the approach presented, although an efficient implementation of complex models (due to either FE or several model parameters) is required for use within the proposed framework. Variational methods for solving the boundary value problem such as the one presented in chapter 2, may reduce computational complexity significantly. In particular, if relevant geometry information is available, semi-inverse solution techniques may be an approach that may be used in conjunction with these models.

2. Variational methods may also be used in tandem with FE solutions. For example, a general solution may be obtained as a weighted linear interpolation of FE solutions computed for a sample of material properties. The weights of the

interpolation may then be obtained as the solution of a variational problem. For example, if the solutions of FE model for k sets of model parameters $\boldsymbol{\theta}_i$ are $\mathbf{y}_i$ then the variational problem may be posed as the minimization of the potential energy $\psi(\sum_i \beta_i \mathbf{y}_i)$. To interpolate between the FE solutions for a particular set of model parameters $\tilde{\boldsymbol{\theta}}$, one may compute the minimum over the space of β_i to find the solution $\tilde{\beta}_i \mathbf{y}_i$. Such an approach may combine the advantages of obtaining complete descriptions using FE methods, while still using a variational method over a significantly lower dimensional space to perform fast computations.

3. Metropolis–Hastings algorithm is one of several algorithms in the class Markov Chain Monte Carlo algorithms to sample the posterior distribution. Novel techniques such as approximate variational Bayes [59], delayed rejection adaptive Metropolis [81] may allow for computing posterior distributions of a larger number of model parameters.

REFERENCES

[1] A. Paul Andersohn, "Modeling frameworks for representing the mechanical behavior of tissues with a specific look at vasculature," Master's thesis, Texas A&M University, College Station, Dec. 2013.

[2] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions (Wiley Series in Probability and Statistics)*. New York: John Wiley & Sons, Inc., 1986.

[3] R. Mahnken, M. Johansson, and K. Runesson, "Parameter estimation for a viscoplastic damage model using a gradient-based optimization algorithm," *Engineering Computations*, vol. 15, no. 7, pp. 925–955, 1998.

[4] R. Butcher, C.-E. Rousseau, and H. Tippur, "A functionally graded particulate composite: preparation, measurements and failure analysis," *Acta Materialia*, vol. 47, no. 1, pp. 259–268, 1998.

[5] J. Kajberg and B. Wikman, "Viscoplastic parameter estimation by high strain-rate experiments and inverse modelling – speckle measurements and high-speed photography," *International Journal of Solids and Structures*, vol. 44, no. 1, pp. 145–164, 2007.

[6] H. Imai, C.-B. Yun, O. Maruyama, and M. Shinozuka, "Fundamentals of system identification in structural dynamics," *Probabilistic Engineering Mechanics*, vol. 4, no. 4, pp. 162–173, 1989.

[7] R. J. Hoeksema and P. K. Kitanidis, "An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling," *Water Resources Research*, vol. 20, no. 7, pp. 1003–1020, 1984.

[8] U. Saravanan, S. Baek, K. Rajagopal, and J. Humphrey, "On the deformation of the circumflex coronary artery during inflation tests at constant length," *Experimental Mechanics*, vol. 46, no. 5, pp. 647–656, 2006.

[9] R. H. Cox, "Regional variation of series elasticity in canine arterial smooth muscles," *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 234, pp. H542–H551, May 1978.

[10] C. M. García-Herrera, D. J. Celentano, M. A. Cruchaga, F. J. Rojo, J. M. Atienza, G. V. Guinea, and J. M. Goicolea, "Mechanical characterisation of the human thoracic descending aorta: experiments and modelling," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 15, no. 2, pp. 185–193, 2012.

[11] G. A. Holzapfel, G. Sommer, C. T. Gasser, and P. Regitnig, "Determination of layer-specific mechanical properties of human coronary arteries with nonatherosclerotic intimal thickening and related constitutive modeling," *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 289, no. 5, pp. H2048–H2058, 2005.

[12] C. J. van Andel, P. V. Pistecky, and C. Borst, "Mechanical properties of porcine and human arteries: implications for coronary anastomotic connectors," *The Annals of Thoracic Surgery*, vol. 76, no. 1, pp. 58–64, 2003.

[13] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge: Cambridge University Press, first ed., Oct. 2003.

[14] F. Mollica, L. Preziosi, and K. R. Rajagopal, *Modeling of Biological Materials (Modeling and Simulation in Science, Engineering and Technology)*. Boston: Birkhäuser, Apr. 2007.

[15] Y. C. Fung, *Biomechanics: Mechanical Properties of Living Tissues, Second Edition*. New York: Springer, second ed., June 1993.

[16] G. A. Holzapfel and R. W. Ogden, *Mechanics of Biological Tissue*. New York: Springer, Apr. 2006.

[17] J. D. Humphrey, *Cardiovascular Solid Mechanics: Cells, Tissues, and Organs*. New York: Springer, Jan. 2002.

[18] R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems*. Waltham: Academic Press, Jan. 2005.

[19] G. A. Holzapfel, T. C. Gasser, and R. W. Ogden, "A new constitutive framework for arterial wall mechanics and a comparative study of material models," *Journal of Elasticity*, vol. 61, pp. 1–48, July 2000.

[20] H. G. Matthies, "Quantifying uncertainty: modern computational representation of probability and applications," in *Extreme Man-Made and Natural Hazards in Dynamics of Structures* (A. Ibrahimbegovic and I. Kozar, eds.), NATO Security through Science Series, pp. 105–135, Springer Netherlands, Jan. 2007.

[21] B. Sonesson, F. Hansen, and T. Länne, "Abdominal aortic aneurysm: A general defect in the vasculature with focal manifestations in the abdominal aorta?," *Journal of Vascular Surgery*, vol. 26, no. 2, pp. 247–254, 1997.

[22] J. M. Dijk, Y. v. d. Graaf, D. E. Grobbee, J. D. Banga, and M. L. Bots, "Increased arterial stiffness is independently related to cerebrovascular disease and aneurysms of the abdominal aorta the second manifestations of arterial disease (SMART) study," *Stroke*, vol. 35, no. 7, pp. 1642–1646, 2004.

[23] N. Ganne-Carrié, M. Ziol, V. de Ledinghen, C. Douvin, P. Marcellin, L. Castera, D. Dhumeaux, J.-C. Trinchet, and M. Beaugrand, "Accuracy of liver stiffness

measurement for the diagnosis of cirrhosis in patients with chronic liver diseases," *Hepatology*, vol. 44, no. 6, pp. 1511–1517, 2006.

[24] M. Charbit, E. Angelini, and S. Audiere, "Maximum likelihood estimation of youngs modulus in transient elastography with unknown line-of-sight orientation," in *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1108–1111, May 2012.

[25] P. Wellman, R. D. Howe, E. Dalton, and K. A. Kern, "Breast tissue stiffness in compression is correlated to histological diagnosis," Technical Report, Harvard Biorobotics Laboratory, 1999.

[26] D. O. Cosgrove, W. A. Berg, C. J. Dor, D. M. Skyba, J.-P. Henry, J. Gay, C. Cohen-Bacrie, and B. S. Group, "Shear wave elastography for breast masses is highly reproducible," *European Radiology*, vol. 22, no. 5, pp. 1023–1032, 2012.

[27] N. M. v. Popele, D. E. Grobbee, M. L. Bots, R. Asmar, J. Topouchian, R. S. Reneman, A. P. G. Hoeks, D. A. M. v. d. Kuip, A. Hofman, and J. C. M. Witteman, "Association between arterial stiffness and atherosclerosis the rotterdam study," *Stroke*, vol. 32, no. 2, pp. 454–460, 2001.

[28] T. Hirai, S. Sasayama, T. Kawasaki, and S. Yagi, "Stiffness of systemic arteries in patients with myocardial infarction. a noninvasive method to predict severity of coronary atherosclerosis," *Circulation*, vol. 80, no. 1, pp. 78–86, 1989.

[29] D. A. Vorp, "Biomechanics of abdominal aortic aneurysm," *Journal of Biomechanics*, vol. 40, no. 9, pp. 1887–1902, 2007.

[30] T. Weber, J. Auer, M. F. O'Rourke, E. Kvas, E. Lassnig, R. Berent, and B. Eber, "Arterial stiffness, wave reflections, and the risk of coronary artery disease," *Circulation*, vol. 109, no. 2, pp. 184–189, 2004.

[31] G. Y. Lee and C. T. Lim, "Biomechanics approaches to studying human diseases," *Trends in Biotechnology*, vol. 25, no. 3, pp. 111–118, 2007.

[32] D. Rothschild, "Forecasting elections comparing prediction markets, polls, and their biases," *Public Opinion Quarterly*, vol. 73, no. 5, p. 895916, 2009.

[33] A. Gelman, N. Silver, and A. Edlin, "What is the probability your vote will make a difference?," *Economic Inquiry*, vol. 50, no. 2, pp. 321–326, 2012.

[34] T. Bollerslev, R. Y. Chou, and K. F. Kroner, "Arch modeling in finance: A review of the theory and empirical evidence," *Journal of Econometrics*, vol. 52, no. 1–2, pp. 5–59, 1992.

[35] T. N. Krishnamurti, C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, "Improved weather and seasonal climate forecasts from multimodel superensemble," *Science*, vol. 285, no. 5433, pp. 1548–1550, 1999.

[36] A. C. Lorenc, "Analysis methods for numerical weather prediction," *Quarterly Journal of the Royal Meteorological Society*, vol. 112, no. 474, pp. 1177–1194, 1986.

[37] Y. Bard, *Nonlinear Parameter Estimation*. New York: Academic Press, June 1973.

[38] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*. New York: Springer, third ed., Apr. 2005.

[39] P. Congdon, *Bayesian Statistical Modelling*. West Sussex: John Wiley & Sons, Inc., 2 ed., Jan. 2007.

[40] S. Kullback and R. A. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[41] L. Mehrez, D. Moens, and D. Vandepitte, "Stochastic identification of composite material properties from limited experimental databases, Part I: Experimental database construction," *Mechanical Systems and Signal Processing*, vol. 27, pp. 471–483, Feb. 2012.

[42] L. Mehrez, A. Doostan, D. Moens, and D. Vandepitte, "Stochastic identification of composite material properties from limited experimental databases, Part II: Uncertainty modelling," *Mechanical Systems and Signal Processing*, vol. 27, pp. 484–498, Feb. 2012.

[43] E. Zhang, P. Feissel, and J. Antoni, "A comprehensive Bayesian approach for model updating and quantification of modeling errors," *Probabilistic Engineering Mechanics*, vol. 26, no. 4, pp. 550–560, 2011.

[44] J. Beck and L. Katafygiotis, "Updating models and their uncertainties. I: Bayesian statistical framework," *Journal of Engineering Mechanics*, vol. 124, no. 4, pp. 455–461, 1998.

[45] L. Katafygiotis and J. Beck, "Updating models and their uncertainties. II: model identifiability," *Journal of Engineering Mechanics*, vol. 124, no. 4, pp. 463–467, 1998.

[46] C. Papadimitriou, L. Katafygiotis, and J. Beck, "Approximate analysis of response variability of uncertain linear systems," *Probabilistic Engineering Mechanics*, vol. 10, no. 4, pp. 251–264, 1995.

[47] C. Papadimitriou, J. Beck, and L. Katafygiotis, "Updating robust reliability using structural test data," *Probabilistic Engineering Mechanics*, vol. 16, no. 2, pp. 103–113, 2001.

[48] G. Yan, "A Bayesian approach for damage localization in plate-like structures using lamb waves," *Smart Materials and Structures*, vol. 22, no. 3, p. 035012, 2013.

[49] J. H. Crews and R. C. Smith, "Quantification of parameter and model uncertainty for shape memory alloy bending actuators," *Journal of Intelligent Material Systems and Structures*, p. 1045389X13490842, July 2013.

[50] J. H. Crews, R. C. Smith, K. M. Pender, J. C. Hannen, and G. D. Buckner, "Data-driven techniques to estimate parameters in the homogenized energy model for shape memory alloys," *Journal of Intelligent Material Systems and Structures*, vol. 23, no. 17, pp. 1897–1920, 2012.

[51] A. Cividini, G. Maier, and A. Nappi, "Parameter estimation of a static geotechnical model using a Bayes' approach," *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, vol. 20, no. 5, pp. 215–226, 1983.

[52] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*. Chapman & Hall/CRC Texts in Statistical Science, Boca Raton: Chapman and Hall/CRC, second ed., July 2003.

[53] W. R. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC Interdisciplinary Statistics, Boca Raton: Chapman and Hall/CRC, first ed., Dec. 1995.

[54] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[55] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[56] R. W. Ogden, G. Saccomandi, and I. Sgura, "Fitting hyperelastic models to experimental data," *Computational Mechanics*, vol. 34, no. 6, pp. 484–502, 2004.

[57] K. R. Rajagopal and L. Tao, "On an inhomogeneous deformation of a generalized neo-hookean material," *Journal of Elasticity*, vol. 28, no. 2, pp. 165–184, 1992.

[58] J. C. Criscione, "A constitutive framework for tubular structures that enables a semi-inverse solution to extension and inflation," *Journal of Elasticity*, vol. 77, no. 1, pp. 57–81, 2004.

[59] M. J. Beal, *Variational Algorithms for Approximate Bayesian Inference*. PhD thesis, University College London, 2003.

[60] M. F. O'Rourke, J. A. Staessen, C. Vlachopoulos, D. Duprez, and G. E. Plante, "Clinical applications of arterial stiffness; definitions and reference values," *American Journal of Hypertension*, vol. 15, no. 5, pp. 426–444, 2002.

[61] R. G. Gosling and M. M. Budge, "Terminology for describing the elastic behavior of arteries," *Hypertension*, vol. 41, no. 6, pp. 1180–1182, 2003.

[62] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*. Wiley Series in Probability and Statistics, Hoboken: John Wiley & Sons, Inc., third ed., Apr. 2013.

[63] G. Zhang, "Neural networks for classification: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 30, no. 4, pp. 451–462, 2000.

[64] D. J. Croft and R. E. Machol, "Mathematical methods in medical diagnosis," *Annals of Biomedical Engineering*, vol. 2, no. 1, pp. 69–89, 1974.

[65] I. Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective," *Artificial Intelligence in Medicine*, vol. 23, no. 1, pp. 89–109, 2001.

[66] R. S. Ledley and L. B. Lusted, "Reasoning foundations of medical diagnosis symbolic logic, probability, and value theory aid our understanding of how physicians reason," *Science*, vol. 130, no. 3366, pp. 9–21, 1959.

[67] P. Szolovits and S. G. Pauker, "Categorical and probabilistic reasoning in medical diagnosis," *Artificial Intelligence*, vol. 11, no. 1-2, pp. 115–144, 1978.

[68] Warner HR, Toronto AF, Veasey L, and Stephenson R, "A mathematical approach to medical diagnosis: Application to congenital heart disease," *Journal of Americal Medical Association*, vol. 177, no. 3, pp. 177–183, 1961.

[69] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag, second ed., Aug. 1985.

[70] J. B. Greenhouse, "On becoming a Bayesian: Early correspondences between J. Cornfield and L. J. Savage," *Statistics in Medicine*, vol. 31, no. 24, pp. 2782–2790, 2012.

[71] E. T. Jaynes, "Information theory and statistical mechanics," *Physical Review*, vol. 106, no. 4, pp. 620–630, 1957.

[72] E. T. Jaynes, "Information theory and statistical mechanics. II," *Physical Review*, vol. 108, no. 2, pp. 171–190, 1957.

[73] E. Jaynes, "Prior probabilities," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 3, pp. 227 –241, 1968.

[74] S. Chib and E. Greenberg, "Understanding the metropolis-hastings algorithm," *The American Statistician*, vol. 49, no. 4, pp. 327–335, 1995.

[75] F. Perrin, B. Sudret, G. Blatman, and M. Pendola, "Use of polynomial chaos expansions and maximum likelihood estimation for probabilistic inverse problems," in *18ème Congrès Français de Mécanique (Grenoble 2007)*, 2007.

[76] T. C. Gasser, R. W. Ogden, and G. A. Holzapfel, "Hyperelastic modelling of arterial layers with distributed collagen fibre orientations," *Journal of The Royal Society Interface*, vol. 3, no. 6, pp. 15–35, 2006.

[77] C. O. Horgan and G. Saccomandi, "A description of arterial wall mechanics using limiting chain extensibility constitutive models," *Biomechanics and Modeling in Mechanobiology*, vol. 1, no. 4, pp. 251–266, 2003.

[78] E. Kuhl, R. Maas, G. Himpel, and A. Menzel, "Computational modeling of arterial wall growth," *Biomechanics and Modeling in Mechanobiology*, vol. 6, no. 5, pp. 321–331, 2007.

[79] K. Perktold and G. Rappitsch, "Computer simulation of local blood flow and vessel mechanics in a compliant carotid artery bifurcation model," *Journal of Biomechanics*, vol. 28, no. 7, pp. 845–856, 1995.

[80] S. Zhao, B. Ariff, Q. Long, A. Hughes, S. Thom, A. Stanton, and X. Xu, "Inter-individual variations in wall shear stress and mechanical stress distributions at the carotid artery bifurcation of healthy humans," *Journal of Biomechanics*, vol. 35, no. 10, pp. 1367–1377, 2002.

[81] H. Haario, M. Laine, A. Mira, and E. Saksman, "DRAM: Efficient adaptive MCMC," *Statistics and Computing*, vol. 16, no. 4, pp. 339–354, 2006.

[82] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. University of Illinois Press, 1971.

[83] Prediction., "Merriam-Webster." http://www.merriam-webster.com/, Oct. 2013.

[84] Forecast., "Merriam-Webster." http://www.merriam-webster.com/, Oct. 2013.

[85] J. R. Freeman and B. L. Job, "Scientific forecasts in international relations: Problems of definition and epistemology," *International Studies Quarterly*, vol. 23, no. 1, pp. 113–143, 1979.

[86] J. D. Simon, "Political risk forecasting," *Futures*, vol. 17, no. 2, pp. 132–148, 1985.

[87] M. S. Lewis-Beck, "Election forecasting: Principles and practice," *The British Journal of Politics & International Relations*, vol. 7, no. 2, pp. 145–164, 2005.

[88] G. Todini, R. Rizzi, and E. Todini, "Detecting precipitating clouds over snow and ice using a multiple sensors approach," *Journal of Applied Meteorology and Climatology*, vol. 48, no. 9, pp. 1858–1867, 2009.

[89] P. Saffo, "Six rules for effective forecasting," *Harvard Business Review*, vol. 85, no. 7/8, p. 122, 2007.

[90] F. Evison, "Long-range synoptic earthquake forecasting: an aim for the millennium," *Tectonophysics*, vol. 338, no. 34, pp. 207–215, 2001.

APPENDIX A

MAXIMUM ENTROPY DISTRIBUTION

The idea of entropy in the sense of information theory is due to Claude Shannon[82], introduced as a measure to quantify uncertainty or "choice in selecting an outcome" in a set of events. Entropy $H(X)$ of a random variable X is defined as

$$H(X) = -\int P(x) \log p(x) \, \mathrm{d}x \tag{A.1}$$

The principle of maximum entropy was first proposed by Jaynes. In his first paper [71] on maximum entropy principle, he proposed that the probability distribution that is "maximally noncommittal" to the unavailable information can be obtained by maximizing the entropy.

Jaynes used this to obtain prior probability distributions for Bayesian inference [73]. The entropy maximization problem, for a random variable when the expectation of the set of functions $f_i(x)$, is known to be $\langle f_i(x) \rangle$, is stated as follows:

$$\underset{P(x)}{\text{maximize}} \; H(X) = \int P(x) \log P(x) \, \mathrm{d}x \tag{A.2}$$

$$\text{subject to} \int P(x) f_i(x) \, \mathrm{d}x = E[f_i(x)] = \langle f_i(x) \rangle \tag{A.3}$$

$$\int P(x) \, \mathrm{d}x = 1 \tag{A.4}$$

For the above maximization problem, the solution is given as

$$P(x) = \frac{1}{Z(\lambda_1, \ldots, \lambda_n)} \exp\left[\lambda_1 f_1(x) + \ldots + \lambda_n f_n(x)\right] \tag{A.5}$$

$$Z(\lambda_1, \ldots, \lambda_n) = \int \exp\left[\lambda_1 f_1(x) + \ldots + \lambda_n f_n(x)\right] \mathrm{d}x \tag{A.6}$$

$$\frac{\partial}{\partial \lambda_i} \log Z(\lambda_1, \ldots, \lambda_n) = \langle f_i(x) \rangle \tag{A.7}$$

The function $Z(\lambda_1, \ldots, \lambda_n)$ is known as the partition function. The distribution thus obtained is known as the MaxEnt distribution and is a probability distribution which incorporates the information in equation (A.3). This distribution can now be used as a prior probability distribution for the Bayesian inference process.

APPENDIX B

A NOTE ON PREDICTION AND FORECAST

This appendix explains the difference in usage of the terms "prediction" and "forecast" in this dissertation (particularly in section 1. The dictionary definitions [83, 84] of these words are

Prediction (noun)

1. a statement about what will happen or might happen in the future,
2. the act of saying what will happen in the future : the act of predicting something,

Forecasting 1. a statement about what you think is going to happen in the future.

Given that the definitions are quite close to each other, it is not surprising that the scientific community has used these words interchangeably. Indeed, a review paper on forecasts in international relations by Freeman and Job [85] observes the difference in the usage of the terms prediction and forecasts (see [85, pp. 114–117]). The authors lay down that at least three different usages exist in the literature, viz. (1) either they are used interchangeably, or (2) it is *forecast* if it is a probabilistic statement and *prediction* if the statement is about a single outcome, without any mention of probabilities or (3) *prediction* is a concrete statement which is logically equivalent to *explanation* whereas a *forecast* is a combination of several predictions and is hence a "weaker" statement on the outcomes (see [85] and the references thereof for a detailed discussion). Based on these observations, the authors define

that "... a *forecast* is a statement about unknown phenomena based upon known or accepted generalizations and uncertain conditions ..." and "*prediction* involves the linkage of known or accepted generalizations with certain conditions (knowns) to yield a statement about unknown phenomena."

Simon [86] differentiates the terms by saying that prediction is a statement about the future with certainty and does not involve probabilities and forecasts are collections of statements and that they are issued with a probability range. This definition falls in the category (2) of Freeman and Job [85].

On the other hand, Lewis-Beck [87] defines forecasts as a statement about the future, whereas prediction is purely a reconstruction of a past outcome. A similar definition given by Todini et. al. [88] from the weather forecast literature is that predictions make statements about time t based on information up until time t whereas forecasts allow one to estimate events at $t + \Delta t$ based on information until time t.

In his article "Six Rules for Effective Forecasting", Paul Saffo [89], declares that

> "Prediction is concerned with future certainty; forecasting looks at how hidden currents in the present signal possible changes in direction for companies, societies, or the world at large."

He also states that predictions deal with "preordained events" and hence current events cannot influence future events and even associates it with the world of "myth and superstition". Such a definition does not even admit the usage of the term prediction in a scientific context.

As a further illustration of the variety in usage, a paper on earthquake forecasting by F. F. Evison [90] explicitly states that prediction and forecast will not be differentiated and that they will be used interchangeably throughout.

Given this contrasting definitions used in the literature, we would like to state that, in this dissertation (particularly in section 1.3 and figure 1.2), the terms prediction and forecast are used in the sense of category (2) of Freeman and Job [85], in that prediction is a statement made with certainty about an outcome of a system, whereas forecast is a collective statement about possible outcomes, assigning probabilities to each of them.

APPENDIX C

SOFTWARE CODE

Fortran code to samples stiffness space using Metropolis – Hastings algorithm

```
program main
use dispmodule
implicit none
character(len=20) :: tempstr
real,dimension(:),allocatable :: likelihood_cand,
   likelihood_cur,alpha
real :: datavari,assumedvari
integer :: Nlinks,NumSamp,Numforce
real,dimension(:,:),allocatable :: obs,theta,y,mean_k,
   std_k,k
real :: kact(2)
real,dimension(20) :: inputdata
real,allocatable :: u(:)
real :: T_y(3)
integer :: i,j
real :: k_cand(2)
real :: sig
real :: timestart,timefinish
real,allocatable :: randomnumbervec(:)
real :: priorprob(2)
integer:: k1num,k2num
real,allocatable ::k1vec(:),k2vec(:),priorprobfromfile
   (:),ppm(:,:),ppderi(:,:),temprand(:)
integer :: outputdetail,numlines,ios
integer :: maxlines=6400
character(len=1)::junk
call cpu_time(timestart)
open(unit=100,file='inputdata.data')
do i=1,11
read(100,*)
read(100,*) inputdata(i)
enddo
close(100)
Nlinks = inputdata(1)
T_y(1) = inputdata(2)
T_y(2) = inputdata(3)
```

```
T_y(3) = inputdata(4)
datavari = inputdata(5)
assumedvari=inputdata(6)
NumSamp = inputdata(7)
kact(1)=inputdata(8)
kact(2)=inputdata(9)
sig=inputdata(10)
outputdetail=inputdata(11)

Numforce=2  ! Edit this to change number of experiments
    used


allocate(obs(Nlinks+1,Numforce),y(Nlinks+1,Numforce),
   theta(Nlinks,Numforce),randomnumbervec(2*NumSamp),u(
   NumSamp),temprand(Numforce*Nlinks+Numforce))
allocate(mean_k(NumSamp,2),std_k(NumSamp,2),k(NumSamp
   ,2),alpha(NumSamp),likelihood_cand(NumSamp),
   likelihood_cur(NumSamp))

do i=1,Numforce
call FiniteDifference(theta(:,i),Nlinks,T_y(i),kact)
enddo

do j=1,Numforce
        write(tempstr,*) j
        open(unit=200,file='theta_'//trim(adjustl(
           tempstr))//'.dat')
        do i=1,Nlinks
                write(200,*) theta(i,j)
        enddo
        close(200)


        y(1,j)=0.0
        do i=1,Nlinks
                y(i+1,j)=y(i,j)+(1./real(Nlinks))*sin(
                   theta(i,j))
        enddo

enddo

call grand(temprand,Numforce*Nlinks+Numforce,0.0,sqrt(
   datavari),1)
do j=1,Numforce
```

```
obs(:,j) = y(:,j)+temprand((j-1)*(Nlinks+1)+1:j*(Nlinks+1))
obs(1,j)=y(1,j)
obs(2,j)=y(2,j)
write(tempstr,*) j
open(unit=222,file='obs_'//trim(adjustl(tempstr))//'.dat')
do i=1,Nlinks+1
        write(222,*) obs(i,j)
enddo
close(222)
enddo


k(1,1) = 1.0
k(1,2) =1.0
mean_k(1,1) = k(1,1)
mean_k(1,2) = k(1,2)
std_k(1,1) = 0.0
std_k(1,2) = 0.0

call grand(randomnumbervec,2*NumSamp,0.0,1.0,1)
call grand(u,NumSamp,0.0,1.0,2)
do i=2,NumSamp
if (ceiling(real(i)/5e4).eq.real(i)/5e4) then
        write(*,*) i
endif

k_cand(1) = sig*randomnumbervec(2*i-1)+k(i-1,1)
k_cand(2) = sig*randomnumbervec(2*i)+k(i-1,2)
call LikelihoodFunction(likelihood_cur(i),obs,k(i-1,1),k(i-1,2),Nlinks,T_y,assumedvari,Numforce)
call LikelihoodFunction(likelihood_cand(i),obs,k_cand(1),k_cand(2),Nlinks,T_y,assumedvari,Numforce)

!!!!!!!!!!!!!! Uniform Prior !!!!!!!!!!!!!!!!!!!!!
if (k_cand(1)>5.or.k_cand(2)>5.or.k_cand(1)<0.0.or.k_cand(2)<0.0) then
        alpha(i)=0.0
else
        alpha(i) = min(1.0,likelihood_cand(i)/
          likelihood_cur(i))
endif
!!!!!!!!!!!!!! Uniform Prior !!!!!!!!!!!!!!!!!!!!!
```

```
!!!!!!!!!!!!!!!! MaxEntPrior !!!!!!!!!!!!!!!!!!!!!!!!
!call PriorProbDist(priorprob(1),k_cand(1),k_cand(2))
!call PriorProbDist(priorprob(2),k(i-1,1),k(i-1,2))

!alpha(i) = min(1.0,(likelihood_cand(i)*priorprob(1))/(
   priorprob(2)*likelihood_cur(i)))
!!!!!!!!!!!!!!!! MaxEntPrior !!!!!!!!!!!!!!!!!!!!!!!!
if (alpha(i).gt.u(i)) then
        k(i,1)=k_cand(1)
        k(i,2)=k_cand(2)

else
        k(i,1) = k(i-1,1)
        k(i,2) = k(i-1,2)

endif


mean_k(i,1) = (1./i)*((i-1)*mean_k(i-1,1)+k(i,1))
mean_k(i,2) = (1./i)*((i-1)*mean_k(i-1,2)+k(i,2))
!
std_k(i,1) = (1./(i-1))*((i-2)*std_k(i-1,1)+(k(i,1)-
   mean_k(i,1))**2)
std_k(i,2) = (1./(i-1))*((i-2)*std_k(i-1,2)+(k(i,2)-
   mean_k(i,2))**2)

enddo

open(unit=300,file='output.dat')

do i=1,NumSamp
write(300,'(D,1X,D,1X,D,1X,D)') k(i,1),k(i,2),
   likelihood_cur(i),likelihood_cand(i) !mean_k(i,1),
   mean_k(i,2)

enddo

close(300)
call cpu_time(timefinish)
write(*,"(A,F10.3,1X,A)"), "Elapsed time is",
   timefinish - timestart, "seconds"

deallocate(likelihood_cand,likelihood_cur,obs,theta,
   alpha,y,mean_k,std_k,k,u,randomnumbervec)

end program
```

```
subroutine linspace(vecout,start,endpt,numpoint)
implicit none
real :: start, endpt
integer :: numpoint, i
real :: d
real :: vecout(numpoint)
d = (endpt-start)/(real(numpoint-1))
do i = 1,numpoint
vecout(i) = start +real(i-1)*d
enddo
end subroutine linspace


!============================ Finite Difference
   solution ======================
subroutine FiniteDifference(theta,Nlinks,Fy,k)
implicit none
real :: root,x0,x1,func,Fy,delh
integer :: Nlinks,i
real :: theta(Nlinks),k(2),kvec(Nlinks)
x0=5.0
x1=3.0
do i=1,Nlinks
kvec(i)=-(k(2)*(0.5.lt.real(i)/Nlinks.and.0.7.gt.real(i
   )/Nlinks)+k(1)*(0.5.ge.real(i)/Nlinks.or.0.7.le.real
   (i)/Nlinks))
enddo
call SecantMethod(root,x0,x1,Nlinks,Fy,kvec)
forall (i=1:Nlinks) theta(i)=0.0
delh=1.0/(Nlinks-1)
theta(2)=delh*root+theta(1)
do i=2,Nlinks-1
        theta(i+1) = -(delh**2/kvec(i))*Fy*cos(theta(i)
           )-(theta(i-1)-2*theta(i))
enddo
end subroutine

subroutine SecantMethod(root,x0,x1,Nlinks,Fy,k)
implicit none
real :: root,x0,x1,func,Fy,k(Nlinks)
integer :: i,Nlinks
real,dimension(1000) :: x
forall (i=1:1000) x(i)=0.0
x(1)=x0
x(2)=x1
```

```
i=3
do while (abs(x(i-2)-x(i-1))>1e-4.and.i<1000)
        x(i) = x(i-1) -func(x(i-1),Nlinks,Fy,k)*(x(i-1)
           -x(i-2))/(func(x(i-1),Nlinks,Fy,k)-func(x(i
           -2),Nlinks,Fy,k))
        i=i+1
enddo
root = x(i-1)
end subroutine

real function func(x,Nlinks,Fy,k)
implicit none
real ::x,Fy,delh
integer :: Nlinks,i
real,dimension(Nlinks) :: theta,k
forall (i=1:Nlinks) theta(i)=0.0
delh=1.0/(Nlinks-1)
theta(2)=delh*x
!write(*,*) x, Nlinks, Fy, k
do i=2,Nlinks-1
        theta(i+1) = -(delh**2/k(i))*Fy*cos(theta(i))-(
           theta(i-1)-2*theta(i))
enddo
func=theta(Nlinks)-theta(Nlinks-1)
return
end function
!============================ Finite Difference
   solution ======================
subroutine GProb(probability,y,s,var)
implicit none
real::probability
real::y,s,var
real :: pi=atan(1.0)*4.0
probability = (1/sqrt(2*pi*var))*exp((-(y-s)**2)/(2*var
   ))
end subroutine

subroutine LikelihoodFunction(likelihood,obs,k1,k2,
   Nlinks,T_y,vari,Numforce)
implicit none
integer :: i,j,Nlinks,Numforce
real,dimension(Nlinks,Numforce) :: theta
real :: likelihood,k1,k2,T_y(Numforce),vari
real :: pi=atan(1.0)*4.0
real :: prob
real :: y(Nlinks+1,Numforce),obs(Nlinks+1,Numforce)
```

```
do j=1,Numforce
call FiniteDifference(theta(:,j),Nlinks,T_y(j),(/k1,k2
   /))
enddo
do j=1,Numforce
y(1,j)=0.0
do i=1,Nlinks
y(i+1,j)=y(i,1)+(1./real(Nlinks))*sin(theta(i,j))
enddo
enddo
likelihood=1.0
do i=Nlinks,Nlinks+1 ! Use if only tip deflections are
   required
!do i=1,Nlinks+1      ! Use if all deflections are used
do j=1,Numforce
call GProb(prob,obs(i,j),y(i,j),vari)
likelihood=prob*likelihood
enddo
enddo
end subroutine


include 'mkl_vsl.fi'
subroutine grand(r,numrand,mu,sigma,typ)
use MKL_VSL_TYPE
USE MKL_VSL
implicit none
integer :: method,numrand,seed,brng
real :: mu,sigma
real :: r(numrand)
integer :: stat,typ

type(VSL_STREAM_STATE) :: stream
brng=VSL_BRNG_MT19937
if(typ.eq.1) then
        method=VSL_RNG_METHOD_GAUSSIAN_ICDF
else
        method =VSL_RNG_METHOD_UNIFORM_STD
endif
seed=secnds(0.0)
open(112,file='seed.data')
read(112,*) seed
close(112)
stat =vslnewstream(stream,brng,seed)
if(typ.eq.1) then
```

```
        stat=vsrnggaussian(method,stream,numrand,r,mu,
           sigma)
else
        stat=vsrnguniform(method,stream,numrand,r,mu,
           sigma)
endif


stat=vsldeletestream(stream)
end subroutine

!!!!!!!!!!!!!!!!!!!!!!!! MAXENT DENSITY
   !!!!!!!!!!!!!!!!!!!!!!!!
subroutine PriorProbDist(pprob,k1,k2)
implicit none
real :: k1,k2,pprob
!~~ 2.5,0.2
if (k1.le.5.and.k1.ge.0.and.k2.le.5.and.k2.ge.0) then
pprob =exp(-15.7392+12.5*k1-2.5*k1**2)*0.2
else
        pprob=0
endif
!~~ 2.5,0.2


end subroutine
!!!!!!!!!!!!!!!!!!!!!!!! MAXENT DENSITY
   !!!!!!!!!!!!!!!!!!!!!!!!
```

Matlab Code to compute MaxEnt Prior Probability Distribution

```
x=linspace(0,5,50);
xvar=0.2;
xmean=2.5;
linit=-ones(1,3);
xlim=[0,5];
options=optimset('MaxFunEvals',10000,'Display','Iter');
lambda_opt=fsolve(@(lambda) ToSolve(lambda,xmean,xvar,
   xlim),linit,options)
ppd=PriorProbabilityDistribution(x,lambda_opt);
plot(x,ppd)

F(1) = trapz(x,ppd);
F(2) = trapz(x,x.*ppd);
```

```
F(3)=trapz(x,(x-xmean).^2.*ppd);

integral(@(x) -PriorProbabilityDistribution(x,
   lambda_opt).*log(PriorProbabilityDistribution(x,
   lambda_opt)),0,5)
y=x;
[xx,yy]=meshgrid(x,y);

ppd=PriorProbabilityDistribution(xx,lambda_opt)*0.2;



function F =  ToSolve(lambda,xmean,xvar,xlim)
F(1) = integral(@(x) PriorProbabilityDistribution(x,
   lambda),xlim(1),xlim(2))-1;
F(2) = integral(@(x) x.*PriorProbabilityDistribution(x,
   lambda),xlim(1),xlim(2))-xmean;
F(3) = integral(@(x) ((x-xmean).^2).*
   PriorProbabilityDistribution(x,lambda),xlim(1),xlim
   (2))-xvar;


function ppd=PriorProbabilityDistribution(x,lambda)

ppd=exp(-lambda(1)-lambda(2).*x-lambda(3).*x.^2);
```

Fortran Code to sample parameter space for different hyperelastic models using data from sheep aorta

```
program main
use dispmodule
implicit none

real*8 :: mu,alpha
integer,parameter :: NumProp=2
integer ::propiter
real*8 :: temp_l
real*8 :: data11(n11data,2),data12(n12data,2),data13(
   n13data,2),data21(n21data,2),data22(n22data,2),
   data23(n23data,2),data31(n31data,2),data32(n32data
   ,2),data33(n33data,2)
real*8 :: data41(n41data,2),data42(n42data,2),data43(
   n43data,2),data51(n51data,2),data52(n52data,2),
```

```
   data53(n53data,2)
real*8 :: sig_MCMC,assumedvari,slopefactor,
   likelihood_cand,likelihood_cur
integer :: i,NumSamp
real*8,allocatable :: propvec(:,:),randomnumbervec(:),u
   (:)
real*8 :: propcand(NumProp),inputdata(7),inipropvec(
   NumProp)
real*8 ::Cylinder1Geometry(2), Cylinder2Geometry(2),
   Cylinder3Geometry(2),Cylinder4Geometry(2),
   Cylinder5Geometry(2)
integer::j
call timestamp()
Cylinder1Geometry=(/Cyl1i,Cyl1o/)
Cylinder2Geometry=(/Cyl2i,Cyl2o/)
Cylinder3Geometry=(/Cyl3i,Cyl3o/)
Cylinder4Geometry=(/Cyl4i,Cyl4o/)
Cylinder5Geometry=(/Cyl5i,Cyl5o/)

open(file='Aorta1_1.dat',unit=11)
open(file='Aorta1_2.dat',unit=12)
open(file='Aorta1_3.dat',unit=13)
do i=1,n11data
read(11,*) data11(i,:)
enddo
close(11)
do i=1,n12data
read(12,*) data12(i,:)
enddo
close(12)
do i=1,n13data
read(13,*) data13(i,:)
enddo
close(13)

open(file='Aorta2_1.dat',unit=21)
open(file='Aorta2_2.dat',unit=22)
open(file='Aorta2_3.dat',unit=23)
do i=1,n21data
read(21,*) data21(i,:)
enddo
close(21)
do i=1,n22data
read(22,*) data22(i,:)
enddo
close(22)
```

```
do i=1,n23data
read(23,*) data23(i,:)
enddo
close(23)

open(file='Aorta3_1.dat',unit=31)
open(file='Aorta3_2.dat',unit=32)
open(file='Aorta3_3.dat',unit=33)
do i=1,n31data
read(31,*) data31(i,:)
enddo
close(31)
do i=1,n32data
read(32,*) data32(i,:)
enddo
close(32)
do i=1,n33data
read(33,*) data33(i,:)
enddo
close(33)
open(file='Aorta4_1.dat',unit=41)
open(file='Aorta4_2.dat',unit=42)
open(file='Aorta4_3.dat',unit=43)
do i=1,n41data
read(41,*) data41(i,:)
enddo
close(41)
do i=1,n42data
read(42,*) data42(i,:)
enddo
close(42)
do i=1,n43data
read(43,*) data43(i,:)
enddo
close(43)
open(file='Aorta5_1.dat',unit=51)
open(file='Aorta5_2.dat',unit=52)
open(file='Aorta5_3.dat',unit=53)
do i=1,n51data
read(51,*) data51(i,:)
enddo
close(51)
do i=1,n52data
read(52,*) data52(i,:)
enddo
close(52)
```

```
do i=1,n53data
read(53,*) data53(i,:)
enddo
close(53)


!!!!!!!!!!!!! INPUT DATA !!!!!!!!
open(unit=200,file='inputdata.dat')
do i=1,7
read(200,*)
read(200,*) inputdata(i)
enddo
close(200)
NumSamp = inputdata(1)
sig_MCMC = inputdata(2)
assumedvari = inputdata(3)
slopefactor=inputdata(4)

do i=1,NumProp
inipropvec(i)=inputdata(i+4)
end do

!!!!!!!!!!!!! INPUT DATA !!!!!!!!
allocate(propvec(NumSamp,NumProp),randomnumbervec(
   NumProp*(NumSamp-1)),u(NumSamp))
call grand(randomnumbervec,NumProp*NumSamp,0.0D0,1.0D0
   ,1)
call grand(u,NumSamp,0.0D0,1.0d0,2)

do i=1,NumProp
propvec(1,i)=inipropvec(i)
enddo

!!! START MCMC !!!!
do i=2,NumSamp
if (ceiling(real(i)/1d6)==real(i)/1d6) then
        call disp(i)
endif
do propiter=1,NumProp
propcand(propiter)=sig_MCMC*randomnumbervec((i-2)*
   NumProp+propiter)+propvec(i-1,propiter)
enddo

if (i>2.and.all(propvec(i-1,:)==propvec(i-2,:))) then
call LikelihoodFunction(temp_l,propcand,NumProp,data11,
   data12,data13,n11data,n12data,n13data,assumedvari,
```

```
   slopefactor,Cylinder1Geometry)
likelihood_cand=temp_l
call LikelihoodFunction(temp_l,propcand,NumProp,data21,
   data22,data23,n21data,n22data,n23data,assumedvari,
   slopefactor,Cylinder2Geometry)
likelihood_cand=temp_l*likelihood_cand
call LikelihoodFunction(temp_l,propcand,NumProp,data31,
   data32,data33,n31data,n32data,n33data,assumedvari,
   slopefactor,Cylinder3Geometry)
likelihood_cand=temp_l*likelihood_cand
!call LikelihoodFunction(temp_l,propcand,NumProp,data41
   ,data42,data43,n41data,n42data,n43data,assumedvari,
   slopefactor,Cylinder4Geometry)
!likelihood_cand=temp_l*likelihood_cand
!call LikelihoodFunction(temp_l,propcand,NumProp,data51
   ,data52,data53,n51data,n52data,n53data,assumedvari,
   slopefactor,Cylinder5Geometry)
!likelihood_cand=temp_l*likelihood_cand
else
call LikelihoodFunction(temp_l,propcand,NumProp,data11,
   data12,data13,n11data,n12data,n13data,assumedvari,
   slopefactor,Cylinder1Geometry)
likelihood_cand=temp_l
call LikelihoodFunction(temp_l,propcand,NumProp,data21,
   data22,data23,n21data,n22data,n23data,assumedvari,
   slopefactor,Cylinder2Geometry)
likelihood_cand=temp_l*likelihood_cand
call LikelihoodFunction(temp_l,propcand,NumProp,data31,
   data32,data33,n31data,n32data,n33data,assumedvari,
   slopefactor,Cylinder3Geometry)
likelihood_cand=temp_l*likelihood_cand
!call LikelihoodFunction(temp_l,propcand,NumProp,data41
   ,data42,data43,n41data,n42data,n43data,assumedvari,
   slopefactor,Cylinder4Geometry)
!likelihood_cand=temp_l*likelihood_cand
!call LikelihoodFunction(temp_l,propcand,NumProp,data51
   ,data52,data53,n51data,n52data,n53data,assumedvari,
   slopefactor,Cylinder5Geometry)
!likelihood_cand=temp_l*likelihood_cand

call LikelihoodFunction(temp_l,propvec(i-1,:),NumProp,
   data11,data12,data13,n11data,n12data,n13data,
   assumedvari,slopefactor,Cylinder1Geometry)
likelihood_cur=temp_l
call LikelihoodFunction(temp_l,propvec(i-1,:),NumProp,
   data21,data22,data23,n21data,n22data,n23data,
```

```
   assumedvari,slopefactor,Cylinder2Geometry)
likelihood_cur=temp_l*likelihood_cur
call LikelihoodFunction(temp_l,propvec(i-1,:),NumProp,
   data31,data32,data33,n31data,n32data,n33data,
   assumedvari,slopefactor,Cylinder3Geometry)
likelihood_cur=temp_l*likelihood_cur
!call LikelihoodFunction(temp_l,propvec(i-1,:),NumProp,
   data41,data42,data43,n41data,n42data,n43data,
   assumedvari,slopefactor,Cylinder4Geometry)
!likelihood_cur=temp_l*likelihood_cur
!call LikelihoodFunction(temp_l,propvec(i-1,:),NumProp,
   data51,data52,data53,n51data,n52data,n53data,
   assumedvari,slopefactor,Cylinder5Geometry)
!likelihood_cur=temp_l*likelihood_cur
endif


alpha = min(1.0,likelihood_cand/likelihood_cur)

if (alpha.gt.u(i)) then
        do propiter=1,NumProp
        propvec(i,propiter)=propcand(propiter)
        enddo
else
        do propiter=1,NumProp
        propvec(i,propiter)=propvec(i-1,propiter)
        enddo
endif

enddo

open(500,file='MCMCSamples.dat')
do i =1,NumSamp
write(500,'(D12.5,X,D12.5)') propvec(i,:)
enddo
close(500)
call timestamp()
end program

subroutine LikelihoodFunction(likelihood,propvector,
   numprop,data1,data2,data3,n1data,n2data,n3data,
   assumedvari,beta,CylinderGeometry)
use dispmodule
implicit none
integer :: i,j,numprop
integer :: n1data,n2data,n3data
```

```
real*8 :: likelihood,propvector(numprop),data1(n1data
   ,2),data2(n2data,2),data3(n3data,2),assumedvari,beta
   ,pi=datan(1.0d0)*4d0
real*8 :: grad_data1(n1data),grad_mod1(n1data)
real*8 :: grad_data2(n2data),grad_mod2(n2data)
real*8 :: grad_data3(n3data),grad_mod3(n3data)
real*8 :: CylinderGeometry(2),gprob
real*8 :: pressure1(n1data),pressure2(n2data),pressure3
   (n3data)

likelihood=1d0

!CylinderGeometry=(/Cyl1,Cyl2/)

!~ call PressureFromRadiusOgden(pressure,
   CylinderGeometry,lambda,mu,alpha,data_r,len_p)
call PressureFromRadiusOgden(pressure1,CylinderGeometry
   ,propvector,numprop,data1(:,1),n1data)
call PressureFromRadiusOgden(pressure2,CylinderGeometry
   ,propvector,numprop,data2(:,1),n2data)
call PressureFromRadiusOgden(pressure3,CylinderGeometry
   ,propvector,numprop,data3(:,1),n3data)

call gradient(grad_data1,data1(:,2),data1(:,1),n1data)
call gradient(grad_data2,data2(:,2),data2(:,1),n2data)
call gradient(grad_data3,data3(:,2),data3(:,1),n3data)

call gradient(grad_mod1,pressure1,data1(:,1),n1data)
call gradient(grad_mod2,pressure2,data2(:,1),n2data)
call gradient(grad_mod3,pressure3,data3(:,1),n3data)




!!! Experiment 1
do i=1,n1data
gprob = exp(-(pressure1(i)-data1(i,2))**2/(2d0*
   assumedvari**2)-beta*(grad_mod1(i)-grad_data1(i))
   **2/(2d0*assumedvari**2))
likelihood=likelihood*gprob;
end do

!!! Experiment 2
do i=1,n2data
gprob = exp(-(pressure2(i)-data2(i,2))**2/(2d0*
   assumedvari**2)-beta*(grad_mod2(i)-grad_data2(i))
```

```
   **2/(2d0*assumedvari**2))
likelihood=likelihood*gprob;
end do

!!! Experiment 3
!do i=1,n3data
!gprob = exp(-(pressure3(i)-data3(i,2))**2/(2d0*
   assumedvari**2))-beta*(grad_mod3(i)-grad_data3(i))
   **2/(2d0*assumedvari**2))
!likelihood=likelihood*gprob;
!end do

end subroutine

subroutine gradient(grad,y,x,len_x)
implicit none
integer :: len_x,i
real*8 :: y(len_x),x(len_x),grad(len_x)
grad(1) = (y(2)-y(1))/(x(2)-x(1))
grad(len_x) = (y(len_x)-y(len_x-1))/(x(len_x)-x(len_x
   -1))

do i=2,len_x-1
grad(i) = (y(i+1)-y(i-1))/(x(i+1)-x(i-1))
enddo
endsubroutine

subroutine linspace(vecout,start,endpt,numpoint)
implicit none
real*8 :: start, endpt
integer :: numpoint, i
real*8 :: d
real*8 :: vecout(numpoint)
d = (endpt-start)/(real(numpoint-1))
do i = 1,numpoint
vecout(i) = start +real(i-1)*d
enddo
end subroutine linspace

include 'mkl_vsl.fi'
subroutine grand(r,numrand,mu,sigma,typ)
use MKL_VSL_TYPE
USE MKL_VSL
implicit none
integer :: method,numrand,seed,brng
real*8 :: mu,sigma
```

```
real*8 :: r(numrand)
integer :: stat,typ

type(VSL_STREAM_STATE) :: stream
brng=VSL_BRNG_MT19937
if(typ.eq.1) then
        method=VSL_RNG_METHOD_GAUSSIAN_ICDF
else
        method =VSL_RNG_METHOD_UNIFORM_STD
endif
open(112,file='seed.dat')
read(112,*) seed
close(112)
stat =vslnewstream(stream,brng,seed)
if(typ.eq.1) then
        stat=vdrnggaussian(method,stream,numrand,r,mu,
           sigma)
else
        stat=vdrnguniform(method,stream,numrand,r,mu,
           sigma)
endif


stat=vsldeletestream(stream)
end subroutine

subroutine PressureFromRadiusOgden(pressure,
   CylinderGeometry,PropVector,NumProp,DataRadius,
   DataLength)
use dispmodule
implicit none
integer :: DataLength,i,NumProp
real*8 :: CylinderGeometry(2),PropVector(NumProp),
   DataRadius(DataLength),Ro,Ri,pressure(DataLength),ci
real*8 :: f0,f1,f2,f3,f4,outval

Ri=CylinderGeometry(1)
Ro=CylinderGeometry(2)


do i=1,DataLength
ci = (DataRadius(i)**2-Ri**2)/(2D0)
call IntegrandStrainEnergy(f0,ci,PropVector,NumProp,Ri)
call IntegrandStrainEnergy(f1,ci,PropVector,NumProp,Ri
   +(Ro-Ri)/4d0)
```

```
call IntegrandStrainEnergy(f2,ci,PropVector,NumProp,Ri
   +2d0*(Ro-Ri)/4d0)
call IntegrandStrainEnergy(f3,ci,PropVector,NumProp,Ri
   +3d0*(Ro-Ri)/4d0)
call IntegrandStrainEnergy(f4,ci,PropVector,NumProp,Ro)
call NewtonCoates(outval,f0,f1,f2,f3,f4,Ri,Ro)
pressure(i) = outval
enddo

endsubroutine

subroutine NewtonCoates(outval,f0,f1,f2,f3,f4,a,b)
implicit none
real*8 ::f0,f1,f2,f3,f4,a,b,outval
outval =(1.d0/90.d0)* (b-a)*(7.d0*f0+32.d0*f1+12.d0*f2
   +32.d0*f3+7.d0*f4)
end subroutine

subroutine SimpsonInt(outval,f0,f1,f2,f3,a,b)
implicit none
real*8 ::f0,f1,f2,f3,a,b,outval
outval =(1.d0/8.d0)* (b-a)*(f0+3.d0*f1+3.d0*f2+f3)
end subroutine

subroutine IntegrandStrainEnergy(integrandvalue,c,
   propvector,numprop,R)
use dispmodule
implicit none
integer :: numprop
real*8 :: term1,term2
real*8 :: integrandvalue,propvector(numprop),c,R

term1=R/sqrt(2*c+R**2)
term2=1/term1



select case (model)

case (1)


integrandvalue=2*propvector(1)*propvector(2)*sinh(2*
   propvector(2)*log(term2))/(2*c+R**2) !! Criscione
   type a cosh(b\gamma_{3})
```

```
case (2)

integrandvalue=-2*( (propvector(1)*(term1**propvector
   (2)-term2**propvector(2)))+(propvector(1)*(term2**(-
   propvector(2))-term1**(-propvector(2)))) )/((2*c+R
   **2)*propvector(2)) !!! Ogden W= 2*mu/a**2 (li^a-3)
   +2*mu/b**2*(li**-b-3)



case (3)

integrandvalue=(4*c*propvector(1)*propvector(2)*(R**2+c
   )*(1+(4*c**2)/(2*c*propvector(2)*R**2+propvector(2)*
   R**4))**propvector(2))/( (2*c+R**2)*(4*1*c**2+
   propvector(2)*R**2*(2*c+R**2)) ) !!! Generalized Neo
    Hookean with b=1


end select

integrandvalue=R*integrandvalue
end subroutine
```

Matlab Code to compute the class membership probabilities

```
sigma=10;

prior =0.5*ones(1,2);

load MCMC_CrisData_Two
model =1;
NumChoices=100;
NumChoices = ceil((5e6-2e6)/NumChoices);
A1_A2_A3_samp  =A1_A2_A3(2e6:NumChoices:end,:);
A4_A5_samp = A4_A5(2e6:NumChoices:end,:);


for aortanum=1:5
    switch aortanum
        case 1
            CylGeom=[5.2234,6.4154]; % Aorta1
            AD=load('../../FourAortaMCMC/WAorta1/
               Aorta1_3.dat');
```

```
18
        case 2
            CylGeom=[5.0010,6.3110]; % Aorta2
            AD=load('../../FourAortaMCMC/WAorta1/
               Aorta2_3.dat');

        case 3
            CylGeom=[5.0911,6.3211];% Aorta3
            AD=load('../../FourAortaMCMC/WAorta1/
               Aorta3_3.dat');

        case 4
28          CylGeom=[5.1356,6.3106];% Aorta4
            AD=load('../../FourAortaMCMC/WAorta1/
               Aorta4_3.dat');

        case 5
            CylGeom=[5.2234,6.5434];% Aorta5
            AD=load('../../FourAortaMCMC/WAorta1/
               Aorta5_3.dat');

    end

    AortaData=AD;
38

    NumSamples=length(A1_A2_A3_samp);
    likelihood=zeros(1,2);
    for i=1:NumSamples
        if (i/500 == ceil(i/500))
            disp(i)
        end
        likelihood(1) = likelihood(1)+ComputeLikelihood
           (AortaData,CylGeom,A1_A2_A3_samp(i,:),model,
           sigma);
        likelihood(2) = likelihood(2)+ComputeLikelihood
           (AortaData,CylGeom,A4_A5_samp(i,:),model,
           sigma);
    end
48  likelihood=likelihood/NumSamples;
    posterior = likelihood.*prior;
    posterior=posterior./sum(posterior);
    postmat(aortanum,:) = posterior


end
```