MATTI NISKANEN

Porous materials find applications in a wide range of fields from medicine to engineering. There is a need to reliably characterise these materials, preferably in a non-destructive manner. In this thesis, new methods based on acoustic measurements and statistical inversion are developed for the non-destructive characterisation of porous materials. The developed methods also provide quantitative information on the uncertainties in the characterisation. The methods are tested with both simulated and experimental data.
Matti Niskanen

BAYESIAN INVERSION METHODS FOR ACOUSTICAL CHARACTERISATION OF POROVISCOELASTIC MEDIA

ACADEMIC DISSERTATION

To be presented by the permission of the Faculty of Science and Forestry for public examination in the SN201 in Snellmania Building at the University of Eastern Finland, Kuopio, on January 17th, 2020, at 12 o’clock.

University of Eastern Finland
Department of Applied Physics
Kuopio 2020
Author’s address:
University of Eastern Finland
Department of Applied Physics
P.O.Box 1627
70211 Kuopio
Finland
email: matti.niskanen@uef.fi

Supervisors:
Docent Timo Lähivaara
University of Eastern Finland
Department of Applied Physics
P.O.Box 1627
70211 Kuopio
Finland
email: timo.lahivaara@uef.fi

Professor Olivier Dazel
Le Mans Université
Laboratoire d’Acoustique de l’Université du Mans
LAUM - UMR CNRS 6613
Avenue Olivier Messiaen
72085 Le Mans cedex 9
France
email: olivier.dazel@univ-lemans.fr

M. Jean-Philippe Groby, CNRS Researcher, Habilitation à Diriger des Recherches
Le Mans Université
Laboratoire d’Acoustique de l’Université du Mans
LAUM - UMR CNRS 6613
Avenue Olivier Messiaen
72085 Le Mans cedex 9
France
email: jean-philippe.groby@univ-lemans.fr

M. Aroune Duclos, Maître de Conférences
Le Mans Université
Laboratoire d’Acoustique de l’Université du Mans
LAUM - UMR CNRS 6613
Avenue Olivier Messiaen
72085 Le Mans cedex 9
France
email: aroune.duclos@univ-lemans.fr

Professor Jari Kaipio
University of Auckland
Department of Applied Mathematics
38 Princes Street
Auckland CBD
New Zealand
email: j.kaipio@auckland.ac.nz
Reviewers:
Professor Daniela Calvetti
Case Western Reserve University
Department of Mathematics,
Applied Mathematics and Statistics
10900 Euclid Avenue
Cleveland, OH 44106
United States
daniela.calvetti@case.edu

Professor Stuart Bolton
Purdue University
School of Mechanical Engineering
585 Purdue Mall
West Lafayette, IN 47907
United States
bolton@purdue.edu

Opponents:
Professor Nuutti Hyvönen
Aalto University
Department of Mathematics and Systems Analysis
P.O.Box 11000
00076 AALTO
Finland
nuutti.hyvonen@aalto.fi

M. Jean-Daniel Chazot, Maître de Conférences
Université de Technologie Compiègne
Laboratoire Roberval UMR CNRS 7337
CS 60319
60203 Compiègne cedex
France
jean-daniel.chazot@utc.fr
Matti Niskanen  
Bayesian inversion methods for acoustical characterisation of poroviscoelastic media  
Kuopio: University of Eastern Finland, 2020  
Publications of the University of Eastern Finland  
Dissertations in Forestry and Natural Sciences; 369

**ABSTRACT**

Natural and artificially made porous materials are encountered and used in numerous fields of engineering, medicine, and physics. One of the challenges is to be able to reliably characterise the pore structure of these materials, as it has a great influence on their physical properties. Characterising porous materials using acoustical methods has become a common procedure in many laboratories, in part because the methods are non-destructive and can be performed on relatively cheap instruments. The characterisation is commonly carried out using deterministic inversion methods, where a cost functional of model residuals is minimised, typically with a least squares (LS) approach. These approaches can provide reliable parameter estimates when the number of parameters is small, but with an increasing number of parameters the LS parameter estimates and, in particular, error estimates often become unreliable.

In this thesis, we study the characterisation of porous materials in the Bayesian framework for inverse problems, where uncertainties can be treated and quantified naturally. We first examine air-saturated porous materials whose frame can be modelled as rigid, which has been often the assumption in inverse characterisation methods. Next, we focus on water-saturated materials and the more general case where the frame is elastic. A key contribution of this thesis is to develop inversion methods with which we can reliably characterise these poroviscoelastic materials.

Each proposed method is tested with real measurements, carried out either in an impedance tube using audible frequencies or in a water tank with ultrasound. To compute the Bayesian parameter estimates and assess the uncertainties related to them, we employ Markov chain Monte Carlo algorithms. The results of this thesis indicate that the Bayesian approach for porous material characterisation can provide robust estimates for the parameters and their uncertainties.
RÉSUMÉ

Les matériaux poreux naturels et artificiels sont rencontrés et utilisés dans de nombreux domaines de l’ingénierie, de la médecine et de la physique. L’un des défis actuel consiste à caractériser de manière fiable la structure des pores de ces matériaux car elle influe grandement sur leurs propriétés physiques. Dans ce contexte, la caractérisation des matériaux poreux à l’aide de méthodes acoustiques est devenue une procédure courante dans de nombreux laboratoires. Cela se justifie en particulier en raison du caractère non destructif de ces méthodes associé à un faible coût de l’équipement nécessaire. La caractérisation est généralement effectuée à l’aide de méthodes d’inversion déterministe basées sur la minimisation d’une fonction coût par les moindres carrés. Ces approches fournissent des informations relativement fiables lorsque le nombre de paramètres recherché est faible. Lorsque ce nombre devient important, ces approches montrent une limite qui se traduit par une mauvaise estimation des paramètres de minimisation par les moindres carrés.

Dans cette thèse, la caractérisation des matériaux poreux est abordée par une approche bayésienne du problème inverse. Cette approche présente l’intérêt de pouvoir considérer quantitativement les incertitudes sur les paramètres recherchés. Dans les approches bayésiennes proposées, l’estimation des paramètres et l’évaluation des incertitudes sont obtenues en utilisant des algorithmes de Monte-Carlo avec des chaînes de Markov. Dans un premier temps, cette approche est appliquée au cas de matériaux poreux saturés d’air. Dans ce cas, la modélisation est réalisée en supposant le squelette comme étant rigide. Des expériences sont ensuite réalisées dans un tube d’impédance à des fréquences audibles. Les mesures ainsi réalisées permettent de tester cette première application. L’application au cas de matériaux poreux saturés par un fluide lourd est abordée dans une deuxième partie où la modélisation permet de considérer l’élasticité du squelette et son effet sur la propagation. De nouvelles mesures sont réalisées dans une cuve à eau à des fréquences ultrasonores afin de tester cette deuxième application de l’approche bayésienne. Le développement de méthodes d’inversion permettant de caractériser de manière fiable ce type de matériaux poro-visco-élastiques constitue une contribution majeure de ces travaux de thèses. Les résultats de cette thèse indiquent que l’approche bayésienne pour la caractérisation des matériaux poreux fournie des estimations robustes pour les paramètres des modèles tout en y associant leur incertitude.

**Universal Decimal Classification:** 519.217.2, 534.6, 539.217, 539.3, 620.179.16

**INSCPEC Thesaurus:** porous materials; porosity; viscoelasticity; acoustic analysis; acoustic measurement; ultrasonic measurement; acoustic impedance; nondestructive testing; inverse problems; parameter estimation; Bayes methods; Markov processes; Monte Carlo methods

**Yleinen suomalainen ontologia:** materiaalit; materiaalitutkimus; huokoisuus; joustavuus; kimmoisuus; akustinen analyysi; ultraäänitutkimus; rikkomaton aineenkuoetus; inversioongelmat; parametrit; estimointi; bayesilainen menetelmä; Markovin ketjut; Monte Carlo -menetelmät
ACKNOWLEDGEMENTS

This thesis was carried out as a jointly supervised thesis between the Department of Applied Physics in the University of Eastern Finland and the Laboratoire d’Acoustique de l’Université du Mans in the Le Mans University during the years 2015-2019. The study received financial support from the Academy of Finland, Jenny and Antti Wihuri Foundation, Saastamoinen Foundation, the RFI Le Mans Acoustique, and the DENORMS COST action.

I would first like to thank my supervisor in Finland, Docent Timo Lähivaara, for your guidance, encouragement, and help over these years. I would also like to thank Docent Tomi Huttunen for guiding me in the beginning of the project. It has been a privilege to work with both of you. In addition, I am grateful for the patience and understanding you showed as my plans were changing and I went on a rather long research visit to New Zealand.

I wish to express my deepest gratitude to my supervisors in France, Professor Olivier Dazel, HDR Jean-Philippe Groby, and MCF Aroune Duclos. You made me feel very welcome to Le Mans, and your enthusiasm towards acoustics and porous materials provided an exciting atmosphere to work in. During my 1.5 years in Le Mans I learned a lot about acoustics, but more importantly about the French way of life, and I will never forget the raclette nights we shared together.

I would also like to express my sincerest thanks to my New Zealand supervisor Professor Jari Kaipio, who welcomed me into his group in the University of Auckland when life was taking me to the other side of the world. My research visit there turned into almost two years, during which you were always there to offer me help, guidance, encouragement, and discouragement, in the correct proportions, when it was needed. It truly was an honour to work with you and your group.

In addition, I would like to thank the official pre-examiners Professor Daniela Calvetti and Professor J. Stuart Bolton for taking the time to review my thesis and providing excellent feedback.

During these four years I have had the pleasure to live and work in three different countries. I have met lots of amazing people, but moving around has meant I have had to say a lot of goodbyes as well. First I would like to thank my good friends Elie and Valentin for always being willing to act as my French-English translators, even when we had just barely met. When I moved to France I did not speak a word of French, but you helped me navigate through the French administration and most importantly, showed me which ones of the hundreds of different cheeses were worth trying (all of them). Thanks for all the gigs and all the evenings at Le Lézard. Also thanks to Charlotte, Mathieu, Quentin, Nico, Dimitri, Weichun, and many others for making my stay in Le Mans unforgettable. I want to also thank my Kuopio friends Antti, Timo, Petteri, Arttu, Johanna, Simo, and Pekka, for the friendship, housing assistance, and discgolf company. Further, I want to thank all my Auckland office mates, Ru, Owen, Pascal, Hwan, Susana, Vincent, Scarlett, and Annabelle, for the laughs and evenings we spent together. You were always down (up?) for some bananagrams to liven up the work day.

Of course, the longest lasting support for my scientific endeavours comes from my family. I owe a huge thanks to my parents Sirpa and Jorma, and my brothers Timo and Tuomas for always encouraging and supporting me.

Finally, I would like to thank my partner Lily. You have travelled across the
world with me and have shown endless love and support during these years, for which I cannot thank you enough.

Kuopio, January 17th, 2020

Matti Niskanen
LIST OF PUBLICATIONS

This thesis consists of the present review of the author’s work in the field of acoustics and the following selection of the author’s publications:


Throughout the overview, these papers will be referred to by Roman numerals.

AUTHOR’S CONTRIBUTION

The publications selected in this dissertation are original research papers on statistical characterisation of porous materials using acoustical methods. All publications are results of joint work with the author of this thesis, the supervisors, and the co-authors. The author of this thesis has been the principal writer and the corresponding author in all of the publications but the help, guidance, and contribution of the thesis supervisors has been significant. The author has also been responsible for conducting the experiments, developing the codes for the forward models and the inversion, and producing the numerical results.
# TABLE OF CONTENTS

1 Introduction

2 Wave propagation in porous media
   2.1 Rigid frame porous media ................................................................. 6
      2.1.1 The Johnson-Champoux-Allard-Lafarge model ......................... 7
   2.2 Poroelastic media ............................................................................. 10
      2.2.1 The Biot model ..................................................................... 11
      2.2.2 Losses in water-saturated materials ........................................ 12
   2.3 Models to compute the acoustic response of a multilayer system ...... 14
      2.3.1 Waves in layered media .......................................................... 14
      2.3.2 The transfer matrix method ................................................... 16
      2.3.3 The global matrix method ..................................................... 17

3 Bayesian methods for parameter estimation 19
   3.1 Parameter estimation in the Bayesian framework .............................. 19
      3.1.1 Likelihood ............................................................................. 21
      3.1.2 Prior models .......................................................................... 22
   3.2 Markov chain Monte Carlo ............................................................... 22
      3.2.1 Metropolis-Hastings update ................................................... 23
      3.2.2 Efficiency measures ................................................................ 24
      3.2.3 Stopping criteria and diagnosing convergence ......................... 25
   3.3 MCMC Samplers .............................................................................. 28
      3.3.1 Overview ............................................................................... 28
      3.3.2 Adaptive algorithms ................................................................ 29
      3.3.3 Parallel tempering .................................................................. 30

4 Review of thesis results 33
   4.1 Publication I: Characterisation of rigid frame porous materials from
      impedance tube measurements ................................................................. 33
      4.1.1 Measurement setup ............................................................... 33
      4.1.2 Inversion ............................................................................... 34
      4.1.3 Results .................................................................................. 35
   4.2 Publication II: Computational methods for estimating the material
      parameters of poroelastic media ............................................................ 37
      4.2.1 The simulation model ............................................................ 37
      4.2.2 Inversion ............................................................................... 38
      4.2.3 Results .................................................................................. 39
   4.3 Publication III: Estimating the material parameters of poroelastic
      objects from ultrasound measurements in water ................................ 42
      4.3.1 The measurement setup ......................................................... 42
      4.3.2 Inversion ............................................................................... 43
      4.3.3 Results .................................................................................. 43


1 Introduction

Nearly all materials are porous to some degree. In some, the porosity occurs natu-
really, others are manufactured to be porous, and in fact it is rather difficult to find
or manufacture a genuinely non-porous solid [1]. In this thesis, we examine mate-
rials whose pores form a fluid-filled interconnected network inside a solid frame.
Such materials commonly around us include the human body with its bones and
tissues, plants, rocks, soils, concrete, plaster, fabrics, wools, and so on [2]. It is well
known that the structure of the pores greatly affects the materials’ physical proper-
ties, such as density, thermal conductivity, and strength. In addition, the connected
pore structure allows fluid to flow through the materials, which gives them special
properties and enables countless applications. These include insulation and dissi-
pation of thermal energy, absorption of acoustic and mechanical energy, filtration,
diffusion, and mixing of flow, which is why porous materials find uses in fields
such as civil engineering, geophysical exploration, food industry, and chemical in-
dustry to name a few [2]. It is therefore essential that we have reliable methods to
characterise the pore structure of materials.

Due to the practical importance of porous material characterisation, numerous
methods have been proposed for it [1,3–5]. These include measurements of fluid
flow, gas adsorption, and X-ray and neutron scattering. One class of methods in
particular, based on acoustic probing of the medium, has received a lot of interest
in the recent years. Acoustic characterisation methods are based on the fact that
pore parameters have a strong influence on the response of porous media to sound
waves [6,7]. In reverse, the acoustic response carries lots of information on the pore
structure, which we can attempt to extract using inversion methods. Acoustical
characterisation methods are non-destructive, relatively low-cost, and have the po-
tential to provide information on a multitude of pore structure parameters from a
single measurement.

The first step in acoustical characterisation is to solve the forward problem, that
is, develop methods to model the acoustic response of porous materials that we can
then compare to measurements. Mathematically, we can represent porous media
to consist of two phases, a compressible fluid and an elastic solid phase, which
both contribute to wave propagation as a coupled fluid-solid system. We call these
kind of materials poroelastic. A generally accepted theory to model the deformations
in this bi-phasic media is the Biot theory [8–12], which was published by M. A.
Biot in 1956 [8,9]. The Biot theory predicts the existence of three distinct waves in
an isotropic poroelastic medium, all of which have been experimentally observed
[13,14]. In some cases, such as with many air-saturated porous media, the coupling
between the solid and the fluid phases is weak. In such a case, an incident acoustic
wave may not transfer enough energy to the solid phase to cause it to vibrate, and
we can model the porous medium as having a rigid frame. Only one type of wave,
the longitudinal acoustic wave, can propagate in this medium. We can then simplify
our models by replacing the poroelastic material with an equivalent fluid model
whose properties are defined by the saturating fluid and the solid matrix.

The inverse problem in acoustical characterisation has been traditionally carried
out in the deterministic sense. The most straightforward approach is to minimise a cost functional, such as the least squares functional, of the model residuals. Studies that use a least squares minimisation approach to estimate the pore structure parameters include [15–27]. Approaches where the parameters are connected to the data using an analytical inverse mapping have also been proposed [28–31]. When the number of estimated parameters is small, deterministic approaches can provide reliable parameter estimates for porous media, as shown by the review articles [3,5]. However, with a growing number of estimated parameters, the parameter estimates and especially possible uncertainty estimates often become unreliable.

Recent studies [32,33] have shown that the reproducibility of deterministic characterisation methods between different acoustic laboratories is quite poor, with the exception of a few parameters. Some of the poor reproducibility can be explained by inhomogeneities in the measured samples and by a lack of standardised measurement and calibration procedures. On the other hand, acoustic models for porous media include nonlinear interactions between the model parameters, the pore parameters may be strongly correlated with each other, and the cost functions often exhibit several minima. These factors present considerable difficulties for minimising the cost function, and the point estimate provided by the optimisation may not summarise the complex parameter space structures very well. In addition, there are underlying uncertainties associated with the measurements that can be difficult to model, and are therefore often neglected. In the context of inverse problems, however, neglecting the modelling errors may yield meaningless estimates [34–36].

Another approach to the inverse problem is to solve it in the Bayesian framework [34, 37–39], which has several pragmatic advantages compared to the deterministic solutions. For example, it allows us to be explicit in defining prior information, and provides a way to directly quantify uncertainty. In addition, the Bayesian approach can naturally handle inaccuracies between the models and actual experiments. In the Bayesian framework, all unknowns are modelled as random variables, and the inverse problem takes the form of statistical inference. The solution of the inverse problem is then the posterior probability density (ppd), which combines our knowledge on the measurement and model uncertainties, possible prior information, and the measurement data. Once the posterior model has been derived, we can compute point and interval estimates to summarise the information in an easily understandable form. The maximum a posteriori (MAP) and conditional mean (CM) estimates, defined in Chapter 3, are examples of Bayesian point estimates. The MAP estimate can be solved with an optimisation approach, but finding the CM estimate, as well as computing the interval estimates, is an integration problem. In the recent years, Bayesian methods have started to gain popularity for characterising porous materials, especially ones that can be modelled as an equivalent fluid [40–44], but methods for poroelastic inversion have also been proposed [44, 45].

In this thesis, the aim is to develop robust acoustical characterisation methods for both rigid and elastic frame porous materials, using measurements done with typical existing instruments. In particular, we are interested in quantifying the uncertainty in the parameter estimates so that the repeatability and reproducibility of the characterisation methods could be improved. To this end, we adopt the Bayesian framework for inverse problems. To explore the true posterior density, we then use Markov chain Monte Carlo (MCMC) [46, 47] methods, which are often the only practical way to draw samples from a high-dimensional and possibly multi-modal posterior. However, MCMC methods can be computationally very demanding, because often to obtain a representative sample of the posterior the forward problem
needs to be solved hundreds of thousands of times, possibly even more. This is why
we pay close attention to the speed of the forward models used, and use fast matrix
methods to solve the Biot equations. In addition, we have been able to take advan-
tage of the numerous advances to MCMC in the recent years [48–51], which have
made it possible to construct efficient samplers for poroelastic wave propagation
problems.

The results of this thesis have been published in I–III.

In publication I, we examine rigid frame porous media that can be measured
in a standard transmission impedance tube [52]. These materials include acoustic
foams, automotive felt materials, and different kinds of wools. We solve the inverse
problem in the Bayesian framework, but in addition compare the results to inversion
from a least-squares minimisation approach. We make the comparison to highlight
the differences between the traditional way of solving the characterisation inverse
problem, and the Bayesian approach.

In publications II and III, we focus on porous materials with an elastic frame.
First, publication II provides a numerical study into the feasibility of estimating
basically all Biot input parameters, just from acoustic reflection and transmission
data in the ultrasonic frequencies. We consider some stability issues of the forward
model, and then construct an efficient MCMC sampler to explore the posterior. We
use methods such as Adaptive Metropolis (AM) [48] and Parallel Tempering (PT)
for the MCMC sampler and find that it can explore the peaky posterior density
reasonably efficiently.

Publication III is a study into the characterisation of a poroelastic ceramic ma-
terial, submerged in water so that we can excite the solid waves. We examine nu-
umerous measurements from several parts of a object made from the porous ceramic,
and carry out the inversion separately in each measurement spot. This highlights
the heterogeneity of the sample, and provides a way of quantifying it.

The introduction to this thesis is organised as follows. Chapter 2 presents briefly
the basics of acoustic wave propagation in rigid and elastic frame porous media, and
matrix methods for their numerical simulation. Chapter 3 introduces the Bayesian
framework for inverse problems, along with the sampling methods used in this
thesis. Chapter 4 reviews and studies the results of the publications I–III, and a
conclusion is given in Chapter 5.


2 Wave propagation in porous media

In this chapter, we will review the basics of sound propagation in rigid and elastic frame porous media. Rigid frame porous media were considered in publication I, and elastic frame materials in publications II and III. We first consider mathematical models that describe the acoustically relevant properties of porous media, and define the parameters that are related to the pore properties. We also state the acoustic and poroelastic wave equations. At the end of the chapter we discuss the construction of the forward models and their computational implementation, and especially some fast matrix methods for solving the wave equations.

Porous materials consist of a continuous solid skeleton that forms the frame of the material and of a fluid phase that fills the pores in the solid, such as in the porous rock of Figure 2.1. For most applications, we are interested in materials whose pores form an interconnected network, which allows acoustic propagation within the porous frame. We call these kind of materials open-cell, as opposed to closed-cell porous materials where each pore is isolated from their neighbour. In addition, we can define a saturated porous medium, where all the pores are filled with the same fluid. This simplifies their modelling considerably. An example of an unsaturated porous medium would be a porous rock submerged in water where air is trapped inside some of the pores. In this thesis, we will consider saturated open-cell porous media, which we can model as a superimposition of two continua (the frame and the fluid) [53].

![Figure 2.1: An example of a naturally occurring porous material, scoria, which is volcanic rock solidified from foaming lava. Scoria typically has a high porosity and pore connectivity [54]. The rock is from the Maungauika volcano in Auckland, New Zealand.](image)
Modelling attenuation mechanisms is of prime importance when modelling wave propagation in porous media. Acoustic waves in porous media are mainly attenuated through viscous and thermal losses that result from the interaction between the solid frame and the fluid. In the free field and at short distances, attenuation can be usually neglected without introducing large errors. However, when waves are propagating in small ducts (which porous media closely resemble), the effects of viscosity and heat conduction become important. These effects were understood already in 1868, when Kirchhoff published a general theory for sound propagation in tubes with an arbitrary cross section [55]. In his work, Kirchhoff considered the viscous and thermal effects together in the same equations [6]. In the case of cylindrical tubes, Zwikker and Kosten [56] worked out simpler expressions, where the viscous and thermal losses are treated separately. Stinson [57] later generalised these equations for an arbitrary cross section, and the use of separate expressions for viscous and thermal losses for sound propagation in ducts is now commonly used.

Analytical description of sound propagation in porous media is, in general, impossible due to the complex geometry of their microstructure. Therefore, we have to model porous media in the macroscopic sense, relying on homogenised representations of the material. Such simplifications are possible when the long-wavelength condition is met, that is, acoustic wavelengths are much larger than the homogenised representative volume of the material, which in turn is much larger than the average pore size [6].

2.1 RIGID FRAME POROUS MEDIA

Often in cases where a porous material is saturated with air and excited by an airborne wave, we can assume the solid frame is motionless. In general, the frame can be assumed motionless when the solid-fluid coupling is weak, and little energy is transferred to the solid. The strength of this coupling is a function of frequency, and it gets weaker as the frequency increases [56]. A porous medium with a motionless solid phase can be modelled as a fluid whose apparent properties are a function of the solid properties [6]. We call such a medium an equivalent fluid.

Time-harmonic wave propagation in an isotropic equivalent fluid medium can be described by the Helmholtz-like equation

\[
\left( \nabla^2 + \tilde{k}_{eq}^2 \right) p = 0, \tag{2.1}
\]

where \( \nabla^2 \) is the Laplace operator, \( p \) denotes the acoustic pressure, and \( \tilde{k}_{eq} \) is the equivalent wave number in the fluid. We adopt the \( \exp(-i\omega t) \) time convention, where \( i = \sqrt{-1}, \omega = 2\pi f \) is the angular frequency, and \( f \) is the frequency. The wave number can be defined with the equivalent density \( \tilde{\rho}_{eq} \) and the equivalent adiabatic bulk modulus \( \tilde{K}_{eq} \) as

\[
\tilde{k}_{eq} = \omega \sqrt{\frac{\tilde{\rho}_{eq}}{\tilde{K}_{eq}}}. \tag{2.2}
\]

The tilde over the quantities denotes that they are complex-valued (due to attenuation) and frequency-dependent (because the governing physics is different at low and high frequencies). The acoustic behaviour of a fluid medium is completely described by \( \tilde{\rho}_{eq} \) and \( \tilde{K}_{eq} \), and from these parameters we can derive interesting
properties such as the fluid’s equivalent speed of sound, \( \tilde{c}_{eq} \), and the equivalent characteristic impedance \( \tilde{Z}_{eq} \) as

\[
\tilde{c}_{eq} = \sqrt{\frac{\tilde{K}_{eq}}{\tilde{\rho}_{eq}}} \quad \text{and} \quad \tilde{Z}_{eq} = \sqrt{\tilde{\rho}_{eq}\tilde{K}_{eq}}.
\]

(2.3)

Any two of the above parameters can be used to describe the medium’s acoustic behaviour.

Several models exist to express the acoustic properties \( \tilde{\rho}_{eq} \) and \( \tilde{K}_{eq} \) of an equivalent fluid. The simplest of these are empirical. They depend on very few macroscopic parameters, and are derived based on fitting suitable simple equations to a set of measurements. One of the first and the most well-known empirical models is the Delany-Bazley (D-B) model [58], where a material’s characteristic impedance and wave number are stated as a function of only one parameter, the flow resistivity. Some limitations of the D-B model are that it is only suitable for fibrous porous materials (such as various glass fibres), and that it is valid only for a restricted frequency range. In the low frequencies, the model returns non-causal solutions [59] that violate the Kramers-Kröning relations [60]. Since their introduction, several improvements to the D-B model have been made, which address the causality requirement but also widen the frequency range and types of applicable materials [59, 61–65].

Empirical models are simple, but clearly not well suited for use in material characterisation due to the fact that they contain very few physical material parameters, and do not model the materials with enough accuracy to do inversion. For this reason, we will now consider semi-phenomenological models that are at least partly based on the physics of sound propagation in porous media. Again, there are many models available, such as the Wilson model [66] where the thermal and viscous effects are modelled as relaxation processes, or the pore size distribution model [67,68] which assumes that the pore sizes are log-normally distributed. In this thesis, we will use the Johnson-Champoux-Allard-Lafarge (JCAL) model, which is currently perhaps the most commonly used model and has been found to represent porous media accurately in many studies.

2.1.1 The Johnson-Champoux-Allard-Lafarge model

The JCAL model [69–71] is a popular semi-phenomenological model that describes the equivalent density and bulk modulus of rigid frame porous media based on six physical parameters, defined in the following, that are independent of the saturating fluid. The JCAL model is directly used as the forward model in publication I, since our measurement procedure provides the materials’ densities and bulk moduli. The model is based on considering the asymptotic behaviour of the density and bulk modulus at low and high frequencies, and joining the asymptotes together with simple expressions. In the JCAL model, viscous losses are treated in the expression of the density and thermal losses are treated in the expression of bulk modulus. Let us first define the physical parameters, while keeping in mind that they are defined generally for any porous media, and are not dependent on the JCAL, or any other model, that uses them.

**Open porosity \( \phi \).** Porosity is one of the most important parameters describing a porous medium. It is defined as the volume of the fluid \( V_f \) divided by the total volume of the material \( V_t \):
\[ \phi = \frac{V_f}{V_t}, \]  

(2.4)

However, for most applications we are mainly interested in open porosity, the fraction of pores that are accessible from the outside. Incidentally, this is also the pore space that acoustic methods are sensitive to, and so we use \( \phi \) to denote the open porosity. Typical values for sound absorbing materials are \( 0.95 \leq \phi \leq 0.99 \), but in general \( \phi \in [0, 1] \).

**Static viscous permeability** \( k_0 \). Alongside porosity, viscous permeability \( k(\omega) \) is the other main parameter describing porous media, and the most important parameter for describing sound absorbing materials. Viscous permeability describes the ease at which fluid can flow through a porous material, and is defined by Darcy’s law:

\[ k(\omega) = -\frac{\eta \langle v \rangle}{\nabla p}, \]  

(2.5)

where \( \nabla p \) is the pressure gradient, \( \langle v \rangle \) is the average fluid velocity in a porous medium, and \( \eta \) is the dynamic viscosity of the saturating fluid. The static limit \( k_0 = \lim_{\omega \to 0} k(\omega) \) depends only on the structure of the porous frame. Flow resistivity \( \sigma \) is the reciprocal of permeability \( \sigma = \eta / k_0 \). The values of permeability can span several orders of magnitude, from practically impermeable \( 10^{-20} \) m\(^2\) to highly permeable \( 10^{-8} \) m\(^2\) [72].

**Geometrical tortuosity** \( \alpha_\infty \). When the pore network is not perfectly straight and parallel to the direction of wave propagation, the waves travel along longer paths. This increase in the path length results in an apparent increase of the density of the material compared to the density of the saturating fluid \( \rho_f \). At high frequencies the viscous skin depth \( \delta = \left(\frac{2\eta}{\rho_f \omega}\right)^{\frac{1}{2}} \) becomes tiny compared to the pore size, and the fluid behaves approximately as an ideal, non-viscous, fluid [69]. In this limit, we can define the geometrical tortuosity \( \alpha_\infty \), which is a real-valued parameter that relates the effective density to the density of the fluid as \( \rho_e = \alpha_\infty \rho_f \). Theoretically, tortuosity is also connected to \( k(\omega) \) [69], and in fact \( k_0 \) and \( \alpha_\infty \) can be seen as parameters describing the effective density of the material at low and high frequencies, respectively.

Geometrical tortuosity can be defined with the help of fluid velocities. Let us denote the microscopic fluid velocity of nonviscous fluid flow at \( r \) by \( v_m(r) \). Tortuosity can then be defined as

\[ \alpha_\infty = \frac{\int_V v_m^2(r) dV}{\left(\int_V v_m(r) dV\right)^2}, \]  

(2.6)

where the integrals are carried out over the homogenized representative volume \( V \). We can also define tortuosity with a simpler concept, related to the path lengths. Let \( L \) be the length of a straight path through a porous material, and \( L' \) the actual tortuous path (see Figure 2.2). Then
Figure 2.2: Left: Illustration of a tortuous path. Right: The open pore space regions have the greatest influence on \( \Lambda' \), and the constricted regions of the pore space have the greatest influence on \( \Lambda \).

By definition, \( \alpha_\infty \geq 1 \), and typical values of tortuosity range between 1 and 3.

**Characteristic viscous length \( \Lambda \).** In many applications, we are interested in the ratio of pore volume to surface area. In the JCAL model, the characteristic viscous and thermal lengths are pore shape factors and provide an indication of this ratio. The characteristic viscous length \( \Lambda \) describes viscous effects at middle and high frequencies, and is defined as [69]

\[
\Lambda = 2 \frac{\int_V |v_i(r)|^2 dV}{\int_A |v_i(r_w)|^2 dA},
\]

where the integrals are carried over the pore space, \( v_i(r) \) is the fluid velocity in the pore, and \( v_i(r_w) \) is the fluid velocity on the pore wall. The definition of \( \Lambda \) is such that constrictions in the pore network (see Figure 2.2), where the velocities are greater, contribute to the integral the most. It can therefore be thought of as an indication of the size of the constrictions. Typically, the viscous characteristic length ranges from 10 \( \mu \)m to 500 \( \mu \)m.

**Characteristic thermal length \( \Lambda' \).** Thermal effects at high frequencies are described by the characteristic thermal length \( \Lambda' \). It is defined in a similar way to \( \Lambda \), except that it does not consider fluid velocity in the pores, but only the ratio of average pore volume to average pore surface area [70]:

\[
\Lambda' = 2 \frac{\int_V dV}{\int_A dA}.
\]

As the exchange area between the solid and the fluid is greater in the open regions (see Figure 2.2), these areas contribute to Eq. (2.9) the most. Therefore, thermal
characteristic length can be seen as an indication on the size of the open pore areas. By definition, $\Lambda' \geq \Lambda$, and typically for fibrous sound absorbing materials $\Lambda' = 2\Lambda$.

**Static thermal permeability $k'_0$.** Analogously to the dynamic viscous permeability, the dynamic thermal permeability $k' (\omega)$ relates the pressure gradient to the macroscopic excess temperature $\langle \tau \rangle$ as [71]

$$k'(\omega) = -\frac{\kappa \phi \langle \tau \rangle}{i\omega p} ,$$

where $\kappa$ denotes thermal conductivity. The static thermal permeability $k'_0 = \lim_{\omega \to 0} k' (\omega)$, describes the thermal effects at low frequencies.

**Expressions for $\tilde{\rho}_{eq}$ and $\tilde{K}_{eq}$**

Now we can state the equations for the density and bulk modulus in the JCAL model. The equivalent density $\tilde{\rho}_{eq}$ of the material is defined

$$\tilde{\rho}_{eq} = \frac{\rho_f}{\phi} \tilde{\alpha} (\omega) ,$$

and the dynamic tortuosity is written as

$$\tilde{\alpha} (\omega) = \alpha_\infty + \frac{i\nu}{\omega k_0} \sqrt{1 - \frac{i\omega}{\nu} \left( \frac{2\alpha_\infty k_0}{\phi \Lambda} \right)^2} ,$$

where $\nu = \eta/\rho_f$ is the kinematic viscosity.

The expression for the equivalent bulk modulus $\tilde{K}_{eq}$ is given by

$$\tilde{K}_{eq} = \frac{\gamma P_0}{\phi} \left( \gamma - \frac{\gamma - 1}{\tilde{\alpha}' (\omega)} \right)^{-1} ,$$

where $\gamma$ is the specific heat ratio, $P_0$ is the static fluid pressure, and $\tilde{\alpha}' (\omega)$ is called the thermal tortuosity, as a homologue to $\tilde{\alpha} (\omega)$. The quantity $\tilde{\alpha}' (\omega)$ is written as

$$\tilde{\alpha}' (\omega) = 1 + \frac{i\nu'}{\omega k_0} \sqrt{1 - \frac{i\omega}{\nu'} \left( \frac{2k'_0}{\phi \Lambda} \right)^2} ,$$

where $\nu' = \nu/\text{Pr}$, and Pr is the Prandtl number. The Prandtl number is an approximately constant property of fluids, describing the ratio of momentum diffusivity to thermal diffusivity.

### 2.2 POROELASTIC MEDIA

In general, the frame of a porous medium is not motionless and vibrates in response to an acoustic excitation. This happens especially when the fluid-solid coupling is strong, for example when the saturating fluid is dense, such as water or oil, when the material the frame is made of is easily deformed, or when the acoustic frequencies are low. Obviously, the solid will vibrate from direct mechanical excitation as well. Accurate prediction of sound propagation in these cases can only be achieved if we
consider a model where the fluid and the frame move simultaneously. The Biot theory \[8, 9\] describes this situation, and is commonly considered to be the most accurate model for poroelastic media.

### 2.2.1 The Biot model

Biot’s poroelastic wave equations can be thought of as coupling the linear elastodynamic equation for wave propagation in solids, and the Helmholtz equation for wave propagation in fluids. Biot derived his equations from the principles of continuum mechanics, assuming the existence of potentials and the principle of least action \[73\]. Later it has been shown that the same equations can be derived by a homogenisation approach starting from the microstructure \[74\].

The original Biot model \[8, 9\], published in 1956, represents a porous medium with six parameters, which are the displacements of the solid phase \(u\) and the fluid phase \(U\). Some assumptions of the theory are isotropy and homogeneity of the medium, and that the long-wavelength condition is fulfilled. In 1962, Biot published an alternative formulation of the theory \[11\], which describes the medium with the solid displacement and a fluid/solid relative displacement \(w = \phi(U - u)\). This formulation is more flexible because it allows the treatment of inhomogeneous materials and simplifies some boundary conditions, and we therefore use it in this thesis. Other representations of the Biot equations have also been proposed \[75, 76\].

We will next describe the basic equations and parameters in the Biot theory. Further information, and detailed derivation of the theory, can be found for example in \[2, 6, 73\]. In Biot’s alternative formulation \[11\], we can write the equations of motion in the frequency domain as

\[
\begin{align*}
\omega^2 \rho_f w + \omega^2 \rho u &= -\nabla \cdot \sigma, \\
\omega^2 \rho_f u + \omega^2 \rho_{eq} w &= \nabla \cdot p,
\end{align*}
\]

(2.15)

where \(\sigma\) is the total stress tensor, \(\rho = (1 - \phi)\rho_s + \phi\rho_f\) is the bulk density of the medium, \(\rho_s\) is the density of the solid, and the rest of the parameters are defined in the same way as in the rigid frame case. In particular, we use the Johnson model of dynamical tortuosity \(2.11\)–\(2.12\) in the expression of equivalent density \(\rho_{eq}\). The stress-strain relations in the alternative formulation are of the form

\[
\sigma = 2N\epsilon + (\lambda_c \nabla \cdot u + a_B M \nabla \cdot w) I,
\]

\[
p = -M (\nabla \cdot w + a_B \nabla \cdot u).
\]

(2.16)

Here \(\epsilon = \frac{1}{2} (\nabla u + (\nabla u)^T)\) is the strain tensor, and \(I\) denotes an identity tensor. The quantities \(N, a_B, \lambda_c,\) and \(M\) are elastic coefficients. The parameter \(N\) is the classic shear modulus and \(a_B\) is the Biot-Willis coefficient \[10\]. The Biot-Willis coefficient controls the rigidity of the frame and is bounded by \(\phi \leq a_B \leq 1\), where the lower limit corresponds to a rigid frame and the upper limit to a completely limp frame. In addition, \(\lambda_c = \lambda + a_B^2 M\), where \(\lambda\) and \(M\) are other elastic parameters. It can be shown that when the parameters in the model are that of a perfect fluid \((\phi = 1, a_\infty = 1, N = 0, k_0 \to \infty)\) the Biot equations reduce to the linear lossless acoustic wave equation, and when the parameters are that of a pure solid \((\phi = a_B = 0, a_\infty \to \infty)\), we get the elastic wave equation \[73\].
Rather than use the elastic parameters defined by Biot, we would like to characterise materials with more commonly used parameters. We therefore relate the $\alpha_B, M$, and $\lambda_c$ to porosity, shear modulus, and the bulk moduli $K_f, K_s$ (the bulk modulus of the elastic solid from which the frame is made of), and $K_b$ (the bulk modulus of the porous frame in a vacuum), with the following relations

\[
\alpha_B = 1 - \frac{K_b}{K_s}, \quad (2.17)
\]

\[
M = \left[ \alpha_B + \left( \frac{K_s}{K_f} - 1 \right) \phi \right]^{-1} K_s, \quad (2.18)
\]

\[
\lambda_c = \left[ (1 + \phi)\alpha_B + \left( \frac{K_b}{K_f} - 1 \right) \phi \right] M - \frac{2}{3} N, \quad (2.19)
\]

see [6, 10].

**Waves in a poroelastic medium**

A fluid medium can only support the compressional wave whereas an isotropic elastic medium can support both compressional and shear wave propagation. We can therefore expect that these three waves will also propagate in a poroelastic medium, a consolidation of a solid and a fluid phase. Indeed, in a homogeneous isotropic media the Biot equations admit three different kinds of solutions which are the three types of waves that can propagate in porous media, called the fast and slow longitudinal waves and the shear wave. However, the waves do not propagate in a single phase in isolation, but all waves propagate both in the solid and the fluid phase and can be distinguished by their velocity and mode of propagation. The faster of the two compressional waves corresponds to vibration in which overall and fluid displacements are nearly in phase, and the slower one to vibration where they are out of phase. For the shear wave the movements are in phase since the fluid does not exert shear force and influences the propagation only through inertial effects [73].

Figure 2.3 depicts the generation of the three Biot waves by mode conversion at the poroelastic boundary. A compressional wave is initially propagating in a fluid medium, meets a poroelastic slab at an angle and transmits a part of its energy into the poroelastic material (reflections are not shown) where the propagation continues as Biot waves. The waves refract at different angles depending on their velocity, and can eventually be recorded after they are converted back to a compressional wave in the fluid. Note that at normal incidence, a plane wave excitation does not generate a shear wave because there is no sideways excitation at the interface.

**2.2.2 Losses in water-saturated materials**

Let us now consider attenuation of the waves in a poroelastic material. Like in the rigid frame situation, viscous effects in the fluid are an important loss factor, and we can use the dynamical tortuosity [69] model (2.11)–(2.12) to account for them. Possible thermal losses in the saturating fluid can also be accounted for with the Champoux-Allard-Lafarge model [70, 71] (2.13)–(2.14). However, thermal effects can often be neglected when the medium is saturated with a liquid due to liquids' compressibility being much lower than that of gases.
Another source of attenuation in porous media are the viscoelastic losses in the solid frame. A viscoelastic material responds to an applied stress with a small delay, which means that the behaviour is time-dependent and the material can be said to have a memory [77]. In the frequency domain, we can model linear viscoelastic wave propagation with a complex elastic modulus, where the real part determines the phase velocity and the imaginary part determines the damping. When damping in the solid is small, we can use real and imaginary parts that are constant with frequency as the simplest attenuation model, but this is in general not valid [78]. The requirement of causality, which in the frequency domain is expressed with the Kramers-Kröning relations [60, 73], necessitates that the real and imaginary parts are frequency dependent.

Let us consider a general complex elastic modulus $M(\omega) = M_R(\omega) + iM_I(\omega)$, which consists of frequency-dependent real and imaginary parts. We can now define a dimensionless quality factor $Q$, which describes the dissipation in the medium:

$$Q(\omega) = \frac{M_R(\omega)}{M_I(\omega)}. \quad (2.20)$$

A highly dissipative medium has a small $Q$, whereas for a purely elastic medium $Q \rightarrow \infty$ and the elastic modulus is real and constant. Note that the quality factor is always defined as (2.20), but different viscoelasticity models give different values of $Q$.

Models for describing the behaviour of $Q$ with respect to frequency are often derived based on simple mechanical models, springs and dampers. The simplest ones are the Maxwell and Kelvin-Voigt models, in which a spring and a damper are connected in series and in parallel, respectively [77]. These models are not very realistic, though, because the Maxwell model predicts infinite damping at zero frequency and no damping when $\omega \rightarrow \infty$, and the Kelvin-Voigt model the opposite.
A more realistic model is the Zener or standard linear solid model, which combines a spring and a Kelvin-Voigt model. Attenuation in the Zener model is again frequency dependent, but now has a peak at a central frequency $\omega_m$, and approaches zero as $\omega \to 0$ and $\omega \to \infty$. It is also possible to construct an attenuation profile that has multiple peaks by connecting several Zener elements in parallel. This model is called the generalised Zener model.

The Zener model and especially its generalised version give a lot of freedom to represent an attenuation profile of a material. However, the downside is that each Zener element adds three free parameters that we need to set (or estimate during inversion). Often the frequency dependence of the quality factor is not known, but some experiments have shown that $Q$ stays nearly constant over a wide range of frequencies [73]. We can then make use of the constant $Q$ model, derived by Kjartansson in 1979 [79], which is a causal model where the quality factor is constant over the whole frequency range. The advantages of using this model are that it only includes two parameters, and is mathematically very simple. A complex modulus in the constant $Q$ model is expressed as

$$M(\omega) = M_0 \left(\frac{i\omega}{\omega_0}\right)^{\frac{2}{3}} \arctan\left(1/Q\right),$$

where $\omega_0$ is a reference frequency, $M_0$ is the modulus of $M(\omega)$ at $\omega_0$, and $Q$ is the quality factor.

### 2.3 MODELS TO COMPUTE THE ACOUSTIC RESPONSE OF A MULTILAYER SYSTEM

Let us now consider the construction of a forward model for the characterisation problem in publications II and III. This means constructing a model that predicts an acoustic response of a layered system that we can measure (in our case the reflection and transmission coefficients), when the equations governing wave propagation and the physical parameters of the system are known.

In this thesis, motivated by computational efficiency, we use plane waves to model the wave propagation. Sound propagation in complex-shaped structures has to be usually solved with element-based methods such as the Finite Element [80], Ultra Weak Variational Formulation [81], or Discontinuous Galerkin [82] methods. These methods require a certain number of elements per acoustic wavelength and with high frequencies can become intolerably complex computationally, especially with inverse problems that we wish to solve by sampling methods where we have to repeatedly solve the forward problem. In this case the computational efficiency of the forward model can determine whether or not the problem is solvable at all in practice. A considerable saving of computational effort can be made if we have a layered system. This allows us to use a number of matrix methods to analytically represent the wave propagation [83], and solve the motion equations in a fraction of the time it takes with element-based methods.

#### 2.3.1 Waves in layered media

Let us consider a layered system as in Figure 2.4, where a plane wave initially propagating in medium $\Omega^{[0]}$ is incident on the system consisting of $n$ layers. The layers are assumed to be flat, stacked along the $x_2$-dimension, and to extend to infinity.
in the other directions. The layers can consist of a fluid (or an equivalent fluid), an elastic solid, or a saturated poroelastic material. The termination condition can either be a rigid backing, in which case there is no transmission, or another layer extending to infinity in the $x_2$-direction. Assuming isotropy, the coordinate axis can always be rotated so that there is no propagation in the $x_3$ direction, and the problem reduces to two dimensions [6]. Let us represent an arbitrary field in the frequency domain as $\chi(x_1, x_2; \omega)$. Because of the plane wave assumption, we can perform a Fourier transform along $x_1$ and write the field as $\hat{\chi}(k_1, x_2; \omega) = \hat{\chi}(k_1, x_2; \omega)e^{i k_1 x_1}$, where $k_1 = -k^0 \sin(\varphi_{inc})$, and $k^0 = \omega/c_f$. We will omit the $\omega$ and $k_1$ dependence for clarity, and write $\hat{\chi}(k_1, x_2; \omega) \equiv \hat{\chi}(x_2)$.

In each layer, equations for the displacements, stresses, and pressure can be stated as the superposition of the wave fields in the layer. These physical fields are solutions to the wave equation (acoustic, elastic, or poroelastic depending on the material) in an infinite medium, and can be expressed as a function of the complex amplitude $A$ and wave number $k_2^{type}$ of each wave, where $type$ indicates the mode of the wave. The idea of multilayer matrix methods is then to combine these solutions of an infinite medium together by introducing boundary conditions at interfaces between the layers. In each layer we can have a backward and a forward propagating wave, so in a fluid layer we can have two, in an elastic layer four, and in a poroelastic layer six wave fields. We can thus write an expression of, for example, the $x_2$ solid displacement $\hat{u}_2$ in a poroelastic material as

$$\hat{u}_2(x_2) = A_1 e^{-ik_{2}^{fast}x_2} + A_2 e^{ik_{2}^{fast}x_2} + A_3 e^{-ik_{2}^{shear}x_2} + A_4 e^{ik_{2}^{shear}x_2} + A_5 e^{-ik_{2}^{slow}x_2} + A_6 e^{ik_{2}^{slow}x_2},$$

(2.22)

where amplitudes 1 and 2 correspond to the fast longitudinal wave, amplitudes 3 and 4 correspond to the shear wave, and amplitudes 5 and 6 correspond to the slow longitudinal wave. We can state the other displacements and stresses as a function of the wave amplitudes $\mathbf{a} = [A_1, \ldots, A_6]^T$ as well, using the stress-strain relations (2.16). This derivation can be quite cumbersome, however, and in practice we can use the following simpler method based on state vector formalism.

Let us combine the displacements, stresses, and pressure into a state vector $\mathbf{s}(x_2)$. The state vector components can be selected arbitrarily as long as they are continuous at the interfaces, and it can be shown that for a given layer the number of the state vector components is equal to the number of the forward and backward propagating waves in the layer [84]. For example for fluids it is convenient to select

---

**Figure 2.4:** Representation of a multilayered system. Different media are denoted with $\Omega[i]$ and the locations of the interfaces by $I_i$. 
pressure and velocity normal to the interface since they are continuous over boundaries. For a poroelastic material we can choose \( \hat{s}(x_2) = [\hat{\sigma}_{12}, \hat{\sigma}_{22}, \hat{p}, \hat{u}_1, \hat{u}_2, \hat{v}_2] \), i.e. the normal components of the total stress tensor, \( \hat{\sigma}_{12} \) and \( \hat{\sigma}_{22} \), the pressure \( \hat{p} \), the solid displacements \( \hat{u}_1 \) and \( \hat{u}_2 \), and the normal component of the total velocity \( \hat{v}_2 = -i\omega(\hat{w}_2 + \hat{u}_2) \).

Using the state vector, we can write the equations of motion for any type of layer (see for example [85]) as a system of ordinary differential equations as

\[
\frac{d\hat{s}(x_2)}{dx_2} + A\hat{s} = 0, \tag{2.23}
\]

where \( A \) is the state matrix, which depends on the components selected for the state vector, and is a function of the material parameters, properties of the incident wave, and frequency. We can use this approach to model inhomogeneous materials too, in which case \( A \) will also be a function of \( x_2 \). It can be shown (see [84]) that we can now write a matrix equation where the state vector is given as a function of the wave amplitudes \( a \) as

\[
\hat{s}(x_2) = [M(x_2)]a. \tag{2.24}
\]

To find the field matrix \([M]\), we calculate the eigendecomposition of \( A = \Phi \Gamma \Phi^{-1} \), where \( \Phi \) and \( \Gamma \) are matrices that include the eigenvectors and the eigenvalues, respectively. Then \([M(x_2)] = \Phi \exp\{-i(x_2 - x_{\text{ref}})\Gamma\}, \) where \( x_{\text{ref}} \) is a reference location.

### 2.3.2 The transfer matrix method

The transfer matrix method (TMM), also known as the Thomson-Haskell method [86, 87], is a well-known general purpose multilayer modelling tool which can be used to calculate the acoustic response of layered systems. The method assumes that the propagating waves are plane waves, but finite fields can be studied by superposition of the plane wave solutions [88]. The basic idea in the TMM is to represent the propagation in each layer with a matrix that relates the state vector of the system at the back of the layer to that at the front of the layer [83]. In the following, the notations are as in Figure 2.4.

The goal in the TMM is to find a matrix \( T \) which relates the state vector \( \hat{s}^{[1]}(I_1) \) of layer \( \Omega^{[1]} \) at the first interface to the state vector \( \hat{s}^{[n-1]}(I_n) \) of layer \( \Omega^{[n-1]} \) at the last interface, i.e.

\[
\hat{s}^{[n-1]}(I_n) = T \hat{s}^{[1]}(I_1). \tag{2.25}
\]

The state vectors at the beginning and end of the layers can then be connected to the known incident, unknown reflected, and unknown transmitted (if no rigid termination) waves, which lets us calculate the acoustic response of the system by solving a matrix equation.

We can find the matrix \( T \) in the following way. The amplitudes of the waves inside the same layer are obviously the same near both interfaces, so based on (2.24) we can write for the layer \( \Omega^{[1]} \):

\[
[M^{[1]}(I_1)]^{-1} \hat{s}^{[1]}(I_1) = a^{[1]} = [M^{[1]}(I_2)]^{-1} \hat{s}^{[1]}(I_2). \tag{2.26}
\]

This means we can find a single layer transfer matrix:
\[ \mathbf{s}^{[1]}(I_2) = [\mathbf{T}^{[1]}]\mathbf{s}^{[1]}(I_1), \]  
\[ (2.27) \]

where \( [\mathbf{T}^{[1]}] = [\mathbf{M}^{[1]}(I_2)][\mathbf{M}^{[1]}(I_1)]^{-1}. \) Now, if layer \( \Omega^{[2]} \) is made of the same type of material as layer \( \Omega^{[1]} \), and the components of the state vectors are continuous at the boundary, we must have \( \mathbf{s}^{[2]}(I_2) = \mathbf{s}^{[1]}(I_2) \) and therefore

\[ \mathbf{s}^{[2]}(I_2) = [\mathbf{T}^{[1]}]\mathbf{s}^{[1]}(I_1). \]  
\[ (2.28) \]

If the layers are of different materials we have to include interface matrices, but for simplicity we do not consider them here and direct the reader to [6, 89]. The same process can be repeated until all layers have been considered:

\[ \mathbf{s}^{[n-1]}(I_n) = [\mathbf{T}^{[n-1]}][\mathbf{T}^{[n-2]}] \cdots [\mathbf{T}^{[1]}]\mathbf{s}^{[1]}(I_1). \]  
\[ (2.29) \]

One of the strengths of the TMM is that we can describe multilayered systems by a single transfer matrix, simply by multiplying together the matrices of the single layers:

\[ \mathbf{T} = [\mathbf{T}^{[n-1]}][\mathbf{T}^{[n-2]}] \cdots [\mathbf{T}^{[1]}]. \]  
\[ (2.30) \]

Finally, to find the response of the layered system, we express it in terms of the waves in the first and the last medium:

\[ \mathbf{a}^{[n]} = [\mathbf{M}^{[n]}(I_n)]^{-1}\mathbf{T}[\mathbf{M}^{[0]}(I_1)\mathbf{a}^{[0]}]. \]  
\[ (2.31) \]

In a usual case, the first and last layers are fluids (this corresponds to a layered system in air or water for example), and we can then have a longitudinal wave propagating towards and away from the system at both ends. To find the reflection and transmission coefficients \( R \) and \( T \), respectively, we set the amplitude of the incoming wave to 1 at layer \( \Omega^{[0]} \), and to 0 at layer \( \Omega^{[n]} \) so there is not a wave propagating towards the system from the back. Then the amplitude of the reflected wave corresponds to \( R \) and the transmitted wave to \( T \), i.e. \( \mathbf{a}^{[0]} = [1, R]^T \) and \( \mathbf{a}^{[n]} = [0, T]^T. \) Now we have two equations and two unknowns so the coefficients can be solved.

An important issue with the TMM is that it is known to exhibit instability at high frequencies and/or large layer thicknesses. This is known as the “large \( f d \) problem” [83]. The problem is of numerical origin (finite precision computers) and is caused mainly by the poor condition number of the transfer matrix, due to the combination of both exponentially decaying and growing components. We observed this instability in our simulations, which meant the TMM approach could not be used. Even though the cause is well-known, the proposed solutions to this differ, and are also dependent on the application. One option is the information vector method [84]. Another solution is to use the so-called global matrix method [90], which was adopted in this thesis.

### 2.3.3 The global matrix method

The global matrix method (GMM) [83,90–92] is an alternative matrix formulation to the TMM, originally proposed in [90]. The main idea in the GMM is to write the forward and backward propagating waves inside the material explicitly as a function of their amplitudes, and choose the origin of the waves to be on the interface.
they originate from. We then assemble the equations for these waves in a single global matrix that represents the whole system. A disadvantage of the GMM is that the resulting matrix can be quite big, which can lead to longer calculation times. A significant advantage, however, is that this method is stable for all frequencies and thicknesses [91, 92]. The stability comes from the freedom to choose the origin of each wave on the interface where they originate, and the exponentially growing and decaying components are not coupled from one interface to the other.

To show the principle of the GMM, let us again consider the layered system of Figure 2.4. Assuming layers of the same type, we can use the continuity of the state vector components on the first interface:

\[
[M^0(I_1)]a^0 = [M^1(I_1)]a^1. \tag{2.32}
\]

This can be also stated as

\[
[[M^0(I_1)][-M^1(I_1)]]a^0 = [0]. \tag{2.33}
\]

We can repeat the process for all interfaces and assemble the equations into a global matrix:

\[
\begin{bmatrix}
[M^0(I_1)] & [-M^1(I_1)] & 0 & \cdots & 0 & 0 \\
0 & [M^1(I_2)] & [-M^2(I_2)] & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & [M^{n-1}(I_n)] & [-M^n(I_n)]
\end{bmatrix}
\begin{bmatrix}
a^0 \\
a^1 \\
a^2 \\
\vdots \\
a^{n-1} \\
a^n
\end{bmatrix} = 0. \tag{2.34}
\]

If adjacent layers are not of similar media (fluid-porous for example), the state vectors are not of the same size and we have to again use interface matrices [6].

Once the global matrix (2.34) is assembled, just like with the TMM, we have to know some of the wave amplitudes in order to solve the response of the system. Typically, we set one incoming wave amplitude to one and the rest to zero, and in the case of a layered system surrounded by fluid we then have \(a^0 = [1, R]^T\) and \(a^n = [0, T]^T\). The unknowns in the problem then consist of the amplitudes within the layers \((a^1, \ldots, a^{n-1})\), as well as of \(R\) and \(T\). The known parts of \(a^0\) and \(a^n\), along with the parts of the global matrix corresponding to them, can now be moved to the right hand side of equation (2.34), and the unknown amplitudes can be solved by a single matrix inversion.
3 Bayesian methods for parameter estimation

In this chapter, we discuss the inverse problem in the characterisation of porous media, namely estimating model parameters from acoustic measurements. Parameter estimation is a topic in publications I–III, but computational methods for the inversion are developed especially in publication II. We will treat the parameter estimation problem in the Bayesian framework, and this chapter starts by giving a brief review of Bayesian statistics. For a more comprehensive treatment we refer the reader to the following textbooks on Bayesian statistics and inversion: [34,37–39,93]. In the latter part of this chapter, we concentrate on methods of computational inference, mainly using sampling methods.

3.1 PARAMETER ESTIMATION IN THE BAYESIAN FRAMEWORK

In the Bayesian approach, all unknowns are modelled as random variables, and we express the information about them as probability densities. The solution of the inverse problem is the posterior probability density (ppd), which combines the information that we learn from measurement data and what we know from other sources (the prior). The ppd can then be used to answer all questions that we can pose in terms of probabilities. These include “What are the parameter values that maximise the posterior probability?”, “What is the mean value of the posterior?”, and “What are the intervals that include most of the posterior probability?”.

The posterior is constructed using the famous Bayes’ formula:

\[ \pi(\theta | y) = \frac{\pi(y | \theta) \pi(\theta)}{\pi(y)} \propto \pi(y | \theta) \pi(\theta), \]  

which is an expression of the conditional probability of the unknown parameters \( \theta \) given the measurement data \( y \). In order to find the posterior we have to consider the probability distribution of \( y \) in terms of \( \theta \), which is done by constructing a model for the likelihood \( \pi(y | \theta) \). Likelihood tells us how the measurements would be distributed if we knew the unknown \( \theta \). In addition, we need to include the possible prior information on \( \theta \), denoted by the density \( \pi(\theta) \), in which we express what we know about the unknowns prior to taking the measurement data into account. Lastly, in formulation (3.1) the term \( \pi(y) \) is fixed, and only acts as a normalising constant. We can therefore ignore it and consider the unnormalised posterior that is proportional to the product of the likelihood and the prior.

If the inverse problem has more than two unknowns, visualising the full posterior is difficult. Typically we then summarise the information contained in the posterior with point and interval estimates. Two perhaps the most common statistical point estimates are the maximum a posteriori (MAP) and the conditional mean (CM) estimates. The MAP estimate is the point which maximises the posterior probability

\[ \theta_{\text{MAP}} = \arg \max_{\theta} \pi(\theta | y). \]
Note that this maximiser may not exist or be unique [38]. Computing the MAP estimate requires the solution of an optimisation problem, and the procedure of computing the MAP estimates is similar to the one of the least squares.

The CM is defined as the expected value of the posterior, that is, the center-of-mass

\[
\theta_{CM} = \int_{\mathbb{R}^{n_\theta}} \theta \pi(\theta | y) d\theta,
\]

(3.3)

where \( \theta \in \mathbb{R}^{n_\theta} \) and \( n_\theta \) denotes the number of the unknowns. The conditional mean is a conditionally biased but unconditionally unbiased estimate [94], and can be also shown to minimise the trace of the posterior covariance [38]. Computing the CM estimate is an integration problem. Usually the integral (3.3) is high-dimensional which renders quadrature integration approaches infeasible. In this case we typically approximate the integral using Markov chain Monte Carlo techniques, described later in this chapter.

Examples of the spread estimates include the conditional covariance

\[
cov(\theta | y) = \int_{\mathbb{R}^{n_\theta}} (\theta - \theta_{CM})(\theta - \theta_{CM})^T \pi(\theta | y) d\theta,
\]

(3.4)

if it exists, and the Bayesian (marginal) credible intervals \( CI_k(p) = [a_I, b_I] \subset \mathbb{R} \), defined as

\[
\int_{a_I}^{b_I} \pi(\theta_k | y) d\theta_k = p / 100,
\]

(3.5)

where the marginal posterior density of \( \theta_k \) is defined as

\[
\pi(\theta_k | y) = \int_{\mathbb{R}^{n_{\theta_k}}} \pi(\theta | y) d\theta_1 \cdots d\theta_{k-1} d\theta_{k+1} \cdots d\theta_{n_\theta},
\]

(3.6)

and the interval \( CI_k(p) \) contains \( p \% \) of the marginal posterior mass. Definition of the \( p \% \) credible interval is not unique. One way to define the interval is so that the tails of the density function contain an equal amount of posterior mass

\[
\int_{-\infty}^{a_I} \pi(\theta_k | y) d\theta_k = \int_{b_I}^{\infty} \pi(\theta_k | y) d\theta_k = (100 - p) / 200.
\]

(3.7)

Another common way to define the interval is to minimise the distance between the endpoints

\[
[a_I, b_I] = \arg \min_{a_I, b_I} \left\{ b_I - a_I \left| \int_{a_I}^{b_I} \pi(\theta_k | y) d\theta_k = p / 100 \right. \right\}.
\]

(3.8)

Credible intervals can be directly interpreted as statements on the probabilities of the parameter values [39], i.e. they represent the uncertainty in the estimates.

It is also possible to define a credible set, where instead of the endpoints of the marginal densities we are looking for an equiprobability hypersurface in \( \mathbb{R}^{n_\theta} \) which encloses \( p \% \) of the posterior mass.
### 3.1.1 Likelihood

Let us now describe the construction of the likelihood model. The likelihood is a function that describes the distribution of the measurements when we know what $\theta$ is, and what the measurements conditions are [34]. Likelihood thus contains the forward model, and information on measurement noise and some of the model uncertainties.

We will start by considering the measurement noise. Most often, as is done in this thesis as well, measurement noise is modelled as additive and mutually independent of the unknowns. Let $y \in \mathbb{R}^{n_y}$, where $n_y$ denotes the length of the measurement vector. With the additive noise assumption, an observation model for the measurements can be written

$$y = h(\theta) + e,$$

where $h : \mathbb{R}^{n_\theta} \to \mathbb{R}^{n_y}$ is the forward model that maps the unknown parameters to measurable data, and $e \in \mathbb{R}^{n_y}$ denotes the additive measurement noise component.

Because $\theta$ and $e$ are mutually independent, the expression for the distribution of the measurements when $\theta$ is fixed can be stated as

$$\pi(y|\theta) = \pi_e(y - h(\theta)), \quad (3.10)$$

where $\pi_e(\cdot)$ denotes the distribution of the noise. The noise is often assumed to be normally distributed, which means that the likelihood can be written as

$$\pi(y|\theta) = \frac{1}{\sqrt{(2\pi)^{n_y}\det\Gamma_e}} \exp\left\{ -\frac{1}{2} (y - h(\theta) - e_\ast)^T\Gamma_e^{-1}(y - h(\theta) - e_\ast) \right\}, \quad (3.11)$$

where $\Gamma_e$ is the noise covariance matrix, and $e_\ast$ the noise mean. If the noise covariance is known, the term in front of the exponential is constant and only acts as a normalising factor. In this thesis we assume that $e_\ast = 0$ and that all covariances are positive definite. Then, by denoting the square root of the inverse of the noise covariance with $L_e = \Gamma_e^{-1/2}$, we can write the likelihood as

$$\pi(y|\theta) \propto \exp\left\{ -\frac{1}{2} \|L_e(y - h(\theta))\|^2 \right\}. \quad (3.12)$$

If the noise covariance is unknown, we can model it as another random variable that we estimate, and write it inside the exponential term

$$\pi(y|\theta, \Gamma_e) \propto \exp\left\{ -\frac{1}{2} \|L_e(y - h(\theta))\|^2 + \log(\det\Gamma_e) \right\}. \quad (3.13)$$

In publications I—III we model the noise as independent and identically distributed (i.i.d.), so that the general form of the noise covariance matrix is

$$\Gamma_e = \sigma^2_e I,$$ 

where $\sigma^2_e$ is the noise variance and $I$ denotes a $n_y \times n_y$ identity matrix.
3.1.2 Prior models

The prior density $\pi(\theta)$ is used to represent information about the unknowns that we have prior to taking the measurement data into account. In principle, constructing the prior is simple since the main requirement is that it gives a high probability to parameter values we expect to see, and a low probability to values we do not expect to see. The prior density should also constrain the unknowns to feasible values. Since the posterior is formed as a product of the likelihood and the prior, the prior has the biggest effect along the posterior dimensions that the data is not very informative. The exact analysis can be carried out using the eigenvalue analysis of posterior covariance. In parameter estimation this means that the prior is highly influential for parameters of which the data does not contain much information about. One difficulty in designing the prior may arise when the prior knowledge is qualitative in nature, because it can be hard to translate into a quantitative representation.

The simplest form of the prior is a uniform density

$$
\pi(\theta) = \begin{cases} 
1, & \theta \in D \\
0, & \text{otherwise,}
\end{cases}
$$

(3.15)

where $D$ is the domain (of feasible values) to which we want to restrict $\theta$. We do not normalise the density because we are only interested in the relative probabilities between parameter values, and the constant normalising factor disappears in the proportionality (3.1). Priors of the form (3.15) are often used to specify positivity constraints, as was done in I, and they can also be used together with other types of priors, as was done in II and III.

Perhaps the most commonly used form of prior is the multivariate normal distribution, where the probability density is written as

$$
\pi(\theta) \propto \exp\left\{ -\frac{1}{2} \|L_\theta(\theta - \theta_*)\|^2 \right\}.
$$

(3.16)

Here $\theta_*$ denotes the expected value of the prior, and $L_\theta L_\theta = \Gamma^{-1}_\theta$ is a matrix square root of the inverse of the prior covariance matrix. We use multivariate normal priors (truncated with in indicator function of the form (3.15)) in publications II and III.

3.2 Markov Chain Monte Carlo

Computation of the conditional mean and interval estimates requires integrating over the posterior density. The integrals are nearly always approximated numerically since an analytical representation of the posterior is rarely obtainable. In low dimensions we can use deterministic integration methods such as Gauss quadratures, but as the number of unknowns grows, the so-called “curse of dimensionality” renders such approaches infeasible. The alternative is to approximate the high-dimensional integrals by statistical methods, which are based on obtaining representative samples from the probability density of interest, and using the samples to approximate the expectations and other summaries. A common approach is Monte Carlo integration [95], where the samples are drawn randomly around the parameter space. The basic Monte Carlo method can be made more efficient with methods such as importance sampling, where the samples are still drawn at random but more often in regions of high probability.
Markov chain Monte Carlo (MCMC) is a class of particularly useful sampling methods, where samples from the parameter space are obtained by simulating random walks in it [96]. The random walk must form a Markov chain, which by definition means that the next step in the chain can depend only on the chain’s previous position. The basic idea behind a MCMC simulation is that we start at a point (or in general, a state) $X_1$ in the parameter space, do a random change to it that satisfies certain conditions, and repeat enough times. In doing so we generate a large sample of points that are distributed according to a given target distribution $\pi(\cdot)$, such as the posterior. MCMC differs from ordinary Monte Carlo methods in that the samples are no longer independent. However, despite the dependence of the samples, running a MCMC simulation may be significantly more efficient, and often the only feasible solution, for integrating complicated probability densities.

A strength of the sampling methods is that once enough samples have been obtained, we can easily approximate unknown quantities of $\pi(\cdot)$, such as the expected value of a function $g$:

$$\mathbb{E}_\pi g = \int_{\mathcal{X}} g(\theta) \pi(\theta) d\theta \approx \frac{1}{n} \sum_{i=1}^{n} g(\theta_i), \quad (3.17)$$

where $\mathcal{X}$ is the support of $\pi(\theta)$, and $n$ is the sample size. For example the CM estimate is computed by taking the mean value of the obtained posterior samples. In addition, computing the posterior marginal densities and credible intervals from the samples is trivial.

The reason we can make the approximation in (3.17) is the ergodic theorem, which implies that if $\Theta = (\theta_1, \theta_2, \theta_3, \ldots)$ is a Harris ergodic Markov chain with $\pi(\cdot)$ as its limiting distribution, then with probability 1,

$$\bar{g}_n := \frac{1}{n} \sum_{i=1}^{n} g(\theta_i) \to \mathbb{E}_\pi g \quad \text{as} \quad n \to \infty, \quad (3.18)$$

if $\mathbb{E}_\pi |g| < \infty$ [97]. For the precise definitions, see [95,96,98]. To construct such a Markov chain we use a transition kernel $K$, which is a conditional probability density such that $\theta_{k+1} \sim K(\theta_k, \theta_{k+1})$. A simple example is the random walk process, $\theta_{k+1} = \theta_k + \epsilon_k$, where the random term $\epsilon_k$ does not depend on the previous samples. Obviously, when $n < \infty$ the approximation $\bar{g}_n$ is not exact. We will discuss the error variance later in this chapter.

### 3.2.1 Metropolis-Hastings update

The transition kernel, or the update mechanism, with which the chain is generated is a central factor in the behaviour and efficiency of the MCMC sampler. The Metropolis-Hastings (MH) accept-reject update [46,47] forms the basis of most of the update mechanisms. Let us again denote the target probability density with $\pi(\cdot)$, and the current location of the chain in the parameter space by $\theta_k$. Then the steps for a single MH update are:

1. Propose a candidate $\theta'$ for the next element in the chain by drawing it from the proposal density $q(\theta_k, \theta')$.
2. Calculate the acceptance ratio

$$a(\theta_k, \theta') = \min \left( 1, \frac{\pi(\theta')q(\theta', \theta_k)}{\pi(\theta_k)q(\theta_k, \theta')} \right). \quad (3.19)$$
3. With probability \( a(\theta_k, \theta') \) set \( \theta_{k+1} = \theta' \) (accept proposal), and otherwise set \( \theta_{k+1} = \theta_k \) (reject proposal).

Note that the chain is propagated even if the proposed candidate is rejected. In addition, when sampling from posterior densities, the normalising factor \( (\pi(y) \text{ in } (3.1)) \) is cancelled out. The MH update is, in principle, very simple to implement. However, in practice there are enormous differences in the efficiency depending on the choice of the proposal density.

As stated above, a special case of the MH algorithm is the random walk Metropolis (RWM) algorithm, where the proposal is of the form \( \theta' = \theta_k + \epsilon_k \), where \( \epsilon_k \) is stochastically independent of \( \theta_k \). One of the most common methods of constructing a RWM proposal is to use a Gaussian distribution, so the proposal has the form and \( \epsilon_k \sim \mathcal{N}(0, \Sigma) \), where \( \Sigma \) is the proposal covariance matrix. A useful property of the RWM proposals is that they are symmetric, i.e. \( q(\theta_k, \theta') = f(||\theta' - \theta_k||) \). With symmetric proposal densities the acceptance ratio \((3.19)\) simplifies to

\[
a(\theta_k, \theta') = \min \left(1, \frac{\pi(\theta')}{\pi(\theta_k)}\right). \tag{3.20}
\]

This is called the Metropolis acceptance ratio and it is the original MCMC algorithm that was proposed in 1953 [46]. In the Metropolis algorithm, if \( \pi(\theta') \) is greater than or equal to \( \pi(\theta_k) \) the proposal is always accepted, whereas if \( \pi(\theta') \) is smaller than \( \pi(\theta_k) \) the proposal is accepted with the probability \( a(\theta_k, \theta') \).

With Gaussian proposal densities, a big question is how to choose the proposal covariance \( \Sigma \). As already noted, the proposal has a huge effect on the efficiency of the algorithm, and different ways to choose \( \Sigma \) has been an area of active research.

### 3.2.2 Efficiency measures

Although in the limit \( n \to \infty \), any correctly implemented MCMC sampler will generate a set of samples that are distributed according to the limiting distribution, an efficient sampler produces usable estimates quickly, whereas an inefficient one may not converge in the lifetime of the universe. We need therefore ways to assess the efficiency of the samplers.

**Autocorrelation** An efficient MCMC sampler will mix well, i.e. it moves quickly all around the relevant parts of the parameter space, and will achieve convergence quickly (as opposed to pseudo-convergence where the chain seems to have converged but is actually stuck in a local maximum). Unfortunately, no diagnostic can reliably detect the convergence of a Markov chain. What we can do, however, is to check the autocorrelation of the chain, which gives an indication on how many iterations we need to produce one independent sample, and thus says something about the efficiency of the sampler. The ideal Markov chain would have an integrated autocorrelation time of one, \( \tau = 1 \), equivalent to i.i.d. sampling.

Samples in a Markov chain are roughly independent when we consider a subsampled chain, where we take every \( \tau \)-th sample. To measure the efficiency of the sampler, we are mainly interested in the number of independent samples, or the effective sample size (ESS), of the chain. The ESS can be estimated by considering the autocorrelation time for one variable at a time, but this ignores the correlations between parameters and gives an inaccurate picture of the sample quality. Vats et
al. \[99\] have proposed a multivariate version of the ESS, which considers the covariance structure of the sample. They define it as

\[
m\text{ESS} = n \left( \frac{|\Lambda|}{|\Sigma_\infty|} \right)^{1/n_\theta},
\]

(3.21)

where \(n\) is the sample size, \(n_\theta\) is the dimension of the parameter space, \(\Lambda\) is the sample covariance matrix, \(\Sigma_\infty\) is the asymptotic covariance matrix in the Markov chain central-limit theorem (see eq. (3.24)), and \(|\cdot|\) represents the matrix determinant. An estimate \(\Sigma_n\) for \(\Sigma_\infty\) can be found by a multivariate batch means estimator:

\[
\Sigma_n = \frac{b_n}{a_n - 1} \sum_{k=0}^{a_n-1} (\bar{\theta}_k - \bar{\theta})(\bar{\theta}_k - \bar{\theta})^T.
\]

(3.22)

To calculate \(\Sigma_n\) we need to cut the chain into \(a_n\) batches that have the length \(b_n\) (i.e. \(n = a_nb_n\)), \(\bar{\theta}_k\) is the mean value of the batch \(k\):

\[
\bar{\theta}_k = \frac{1}{b_n} \sum_{i=(k-1)b_n+1}^{kb_n} \theta_i,
\]

(3.23)

and \(\bar{\theta}\) is the mean value of all of the samples (i.e. the CM estimate). There are several ways to decide the batch size, but in general it should grow with the chain length, for example as \(b_n = n^{1/2}\) or \(b_n = n^{1/3}\). Also higher correlation in the parameters calls for a larger batch size, suggesting that \(b_n \propto \tau\).

### 3.2.3 Stopping criteria and diagnosing convergence

We need to have some criteria to stop the MCMC simulation. The simplest one is to run the chain for \(N\) predetermined steps and then stop, but this is far from ideal. To make meaningful inferences we need to be relatively sure that the chain has first of all converged and secondly had enough time to sample the target distribution. Autocorrelation times together with looking at the time series of the individual chains provide a tool to assessing whether we can stop the simulation or not.

A more objective stopping criteria can also be constructed. For example, we can decide to stop the simulation once a predetermined accuracy of a statistic is reached. Commonly we look at the estimate of the mean, \(\bar{\theta}\), and its estimation error, the Monte Carlo standard error (MCSE), defined in the following.

### Burn-in

Before it is sensible to start estimating properties such as the mESS and MCSE, we need to make sure the chain has reached the steady-state. In the beginning of the MCMC run, especially if the starting location of the chain was far away from the main support of the pdf, the values do not represent the target density very well and it can be good to remove them so that they do not cause bias in the estimates. This initial period is called the burn-in period. Removing the burn-in is important especially when we are using adaptive MCMC methods that we will describe later. However, note that removing a burn-in phase from the beginning of the simulation is sometimes completely unnecessary \[96\]. This is the case when we do not use adaptation, and do sufficiently long runs so that the effect caused by the initial locations is negligible.
Figure 3.1: The first 45,000 iterations of one parameter of an example MCMC run. During the first 15,000 iterations the chain is drifting towards the region of the main probability mass, but after that it stays more or less stationary.

Figure 3.1 shows an example of a MCMC run, where the chain was not started in the region where the most of the probability mass is, and therefore in the beginning we can see the chain drifting towards this area. After about 15,000 iterations, the drifting seems to stop and we have reached the steady state. We could then remove these 15,000 iterations from the beginning as the burn-in.

Monte Carlo standard error

The difference $\bar{g}_n - E_\pi g$ is called the Monte Carlo error, and although in practice we can never know its true value, we can approximate it by estimating the variance of its asymptotic distribution. Let us examine the unidimensional case, i.e. $\theta$ is a single unknown. The central limit theorem (CLT) states that [97]:

$$\sqrt{n}(\bar{g}_n - E_\pi g) \to N\left(0, \sigma^2_g\right), \quad (3.24)$$

as $n \to \infty$, where $\sigma^2_g := \text{var}_\pi \{g(\theta)\} + 2 \sum_{i=2}^{\infty} \text{cov}_\pi \{g(\theta_1), g(\theta_i)\}$, and we assume that $\theta_1, \theta_2, \theta_3, \ldots$ are identically distributed and $\theta_1$ is distributed according to the target distribution $\pi(\cdot)$. The CLT can be stated in the multivariate setting too, where we replace $\sigma^2_g$ with its multivariate counterpart $\Sigma_g$. The CLT shows that the accuracy of a Monte Carlo simulation increases as $\sqrt{n}$.

To estimate the Monte Carlo error we need an estimator $\hat{\sigma}^2_g$ for the asymptotic variance $\sigma^2_g$. Non-overlapping batch means are one of several ways to do this. If we denote the mean of the $k$:th batch $[g(\theta_{(k-1)b_n+1}), \ldots, g(\theta_{kb_n})]$ with $\bar{Y}_k$ (see Eq. (3.23)), the batch means estimate for $\sigma^2_g$ is

$$\hat{\sigma}^2_g = \frac{b_n}{a_n - 1} \sum_{k=1}^{a_n} (\bar{Y}_k - \bar{g}_n)^2, \quad (3.25)$$

where again $a_n$ is the number of batches and $b_n$ is the batch length. Now we can estimate the Monte Carlo standard error of $\bar{g}_n$ [97]:

$$\text{MCSE} = \frac{\hat{\sigma}_g}{\sqrt{n}}, \quad (3.26)$$
and we can also calculate the \((1 - \alpha) \cdot 100\%\) confidence interval for the mean. We know that the error of the mean has a \(t\)-distribution with \(n - 1\) degrees of freedom (the CLT tells us that asymptotically, the error is normally distributed). So we are looking for values \(q_1\) and \(q_2\) that contain the true value with probability \((1 - \alpha)\). From probability theory we know that \[ P\left(q_1 < \frac{\bar{g}_n - E\hat{g}}{\hat{\sigma}_g/\sqrt{n}} < q_2\right) = 1 - \alpha . \] (3.28)

For example the half width of the \((1 - \alpha) \cdot 100\%\) confidence interval for \(Eg\) is then given by

\[ CI_{1-\alpha} = t_{a_n-1}\left(1 - \frac{\alpha}{2}\right) \frac{\hat{\sigma}_g}{\sqrt{n}}, \] (3.29)

where \(t_{a_n-1}\) denotes the inverse cumulative distribution function of the \(t\)-distribution with \(a_n - 1\) degrees of freedom.

### Stopping criteria

Let us now examine a few criterion for stopping the simulation.

**Fixed-length** The standard way to run MCMC is to iterate for \(N\) steps, check visually if the run has “converged” and either stop or run for longer. This is not very accurate, though, and better ways have been proposed.

**Fixed-width** In the fixed-width criterion the desired level of accuracy is predefined and the simulation is stopped once it is reached. One example of such a stopping rule is given in [97], where the stopping criterion is

\[ t_{a_n-1}\hat{\sigma}_g + \varepsilon I(n < n^*) < \varepsilon , \] (3.30)

where \(\varepsilon\) is the maximum error level, \(I\) is an indicator function, and the role of \(n^*\) is to prevent a premature stopping due to a bad estimate of \(\hat{\sigma}_g\). In practice, this is calculated for each parameter separately and the simulation is terminated when the uncertainties in all of the parameters are under the error limit.

**Fixed-width relative to standard deviation** Another way to stop the simulation is to base the maximum error level \(\varepsilon\) on a percentage of the standard deviation of the parameters, instead of a predefined level of accuracy. This allows for some flexibility in the stopping criterion since the method allows for more absolute error in parameters that have inherently more variability while keeping the relative error the same. This is the method we chose to use in publications II and III.

The stopping criterion is
\[ t_{\alpha} - 1 \frac{\hat{\sigma}_g}{\sqrt{n}} + \varepsilon_{SD} I(n < n^*) < \varepsilon_{SD} , \]  

(3.31)

where \( \varepsilon_{SD} \) represents now the allowed percentage error in the standard deviation \( \sigma_\theta \) of the actual parameter, for example \( \varepsilon_{SD} = 0.10 \cdot \sigma_\theta \).

**Multivariate fixed-width** It can be shown [99] that stopping based on the effective sample size is equivalent to the fixed-width relative to standard deviation rule. Calculating the multivariate ESS would take the correlations between parameters into account, making the stopping rule even more robust.

### 3.3 MCMC SAMPLERS

In this section, we will describe some commonly used updating mechanisms for the Markov chain. We call the update mechanisms ‘samplers’ since they determine how the MCMC code samples the target density.

A robust sampler is one that can deal with probability densities that are challenging to sample, such as multimodal or strongly non-normal ones. We need a robust sampler so we can be reasonably confident that we are indeed sampling from the full distribution instead of just a local maximum. Unfortunately, the robustness of a sampler also depends on the application, and hence universal recommendations that work in all cases are not possible.

#### 3.3.1 Overview

Since the “second-generation MCMC revolution” starting approximately in the 1990’s [96], a huge number of MCMC samplers have been proposed. These include adaptive Metropolis [48, 50, 101–106], Metropolis-adjusted Langevin [107–109], Hamiltonian Monte Carlo [110–113], Gibbs [114, 115], reversible jump MCMC [41, 116–120], delayed rejection [121, 122], randomize-then-optimize [123, 124], parallel tempering [49, 51, 125–129], affine invariant ensemble sampling [130–132], elliptical slice sampling [133], stochastic Newton [134, 135], and likelihood-informed subspace [136, 137] algorithms. This list is by no means exhaustive, and in the following we will only describe the samplers that were used in I–III.

**Random walk Metropolis and acceptance rate** First let us comment on the RWM sampler, which most of the more advanced samplers are based on. The proposal density in RWM is usually Gaussian, \( \epsilon \sim \mathcal{N}(0, \Sigma) \), or uniform, \( \epsilon \sim \mathcal{U}([-\sigma, \sigma]) \), where \( \Sigma \) denotes the proposal covariance matrix, and \( \sigma \) a vector of proposal widths. However, other types of proposals have also been successfully used [138]. In the following we will use the term “proposal (co)variance” to mean generally the width of the proposal density. The RWM algorithm is very simple to implement, but can be difficult to tune properly, the reason being that no matter what kind of proposal we are using, we have to fix its variance before the simulation.

If the variance of the proposal is set too large, the proposals \( \theta' \) are far away from the current location \( \theta_k \), and likely the value of the density at \( \pi(\theta') \) is much lower than at \( \pi(\theta_k) \). Then the acceptance ratio (3.20) is small, the proposal will be rejected often, and the chain will not move. This wastes computation resources and leads to long autocorrelation times. On the other hand, if the proposal variance is set too
small, virtually all proposals will be accepted but the chain still moves very slowly around the parameter space. Again we have long autocorrelation times. Clearly there must exist better values for the proposal variance between these two extremes, where the proposals are sometimes accepted and sometimes rejected, and the chain moves around the parameter space more efficiently. However, finding the efficiency-wise optimal, or even a good, proposal covariance by trial and error when little is known about the target distribution, is problematic to say the least.

This behaviour of the chain can be monitored with the acceptance rate, which is defined simply as the long-term average of the acceptance ratios (3.20). Again, a general optimal value for the acceptance rate does not exist, but as a rule of thumb as long as the acceptance rate is between 10 and 60 % the extremes of bad sampling are avoided. However, it has been shown [139, 140] that efficiency-wise the optimal acceptance rate for uncorrelated Gaussian target distributions is 0.44 in the one-dimensional case, declining to 0.23 when $n_\theta \to \infty$. In addition, the highest efficiency is achieved when the proposal covariance mimics the covariance of the target distribution, scaled by approximately $2.4^2 / n_\theta$.

### 3.3.2 Adaptive algorithms

Even if the size and spatial orientation of the proposal covariance could be tuned by trial and error in low-dimensional cases, in high dimensions it is virtually impossible. Fortunately, the proposal covariance can be adapted automatically as the simulation goes on and we learn more about the target distribution.

One of the first adaptive algorithms was the adaptive Metropolis (AM) [48, 101], where it was proposed to adapt the size and orientation of the proposal distribution on the fly as the simulation progresses. The basic idea is to approximate the target density with a multivariate Gaussian and use the approximation as the proposal distribution:

$$
\theta' = \theta_k + \epsilon_k, \quad \epsilon_k \sim N(0, \Sigma_k),
$$

(3.32)

where

$$
\Sigma_k = \begin{cases} 
\Sigma_0, & k \leq k_0 \\
 s_d \text{cov} (\theta_1, ..., \theta_k) + s_d I, & k > k_0 
\end{cases}
$$

(3.33)

We need to select a starting proposal covariance $\Sigma_0$, but usually it is enough to choose something small so that proposals are accepted and a first estimate of the target distribution covariance can be calculated after $k_0$ iterations. The scaling parameter $s_d$ is typically set to $s_d = 2.4^2 / n_\theta$, which is a good choice especially if the target density is close to Gaussian [139]. The parameter $\epsilon$ is introduced to prevent the covariance $\Sigma_k$ from becoming a singular matrix, and is selected to be very small compared to the size of the support of the target density. The proposal covariance $\Sigma_k$ can be updated iteratively, so that computing it induces practically no extra computational cost.

It is clear that with adaptive methods the resulting chain is no longer strictly Markovian, since $\theta_n$ now depends on all the previous states. However, it is shown in [48] that under certain conditions the chain is still ergodic and the adaptive algorithm can be used. Later it was realised that the AM algorithm can be viewed as belonging to a more general set of stochastic control algorithms [141] and convergence of more general adaptive algorithms was also proved [102, 103].
requirement for the adaptive algorithms to be ergodic is that the adaptation is “vanishing”, i.e. that for every successive iteration we adapt less and less.

The AM algorithm conditions hold also for a case where we do not consider all the past states of the chain, but only an increasing part of the near history. Calculating the covariance from \((\theta_{n/2}, \theta_{n/2+1}, \ldots, \theta_n)\) instead of all the samples can be better in some situations. This can be the case for example when the starting location is bad and the first samples would orientate the covariance unfavourably.

### 3.3.3 Parallel tempering

The AM algorithm works very well in unimodal and convex cases, even if there are strong correlations between the parameters. However, it can be slow to converge and travel around the parameter space in more complicated situations. Multimodal distributions are notoriously hard to sample efficiently, because the individual peaks are usually separated by large areas of almost zero probability. Even if the target density is not multi-modal as such, it can include narrow peaks of high probability where the chain is likely to get stuck for long periods of time, especially if using adaptive methods.

*Parallel tempering* is one way to make sampling such distributions easier \([49, 125, 126]\). It is based on raising the target density to a power, i.e. tempering, as

\[
\pi(\theta)^{1/T}
\]

where \(T\) can be interpreted as the temperature. At high temperatures the possible peaks of the density are flattened out, making it easier for the chain to travel between them. For a simple example with a one-dimensional density, see Fig. 3.2. At \(T = 1\) the target density consists of two peaks separated by a wide zero probability area. This is a case where it would be difficult to construct an efficient sampler, because to sample within the peaks the proposal needs to be narrow but to jump between the peaks it needs to be wide. When the temperature is increased, the peaks start to flatten out. At \(T = 100\) the modes have mixed into one, making the density easier to sample.
Naturally, tempering changes the target density into something else that we are not primarily interested in. The idea in parallel tempering is to sample in parallel \( n_T \) Markov chains at temperatures \( 1 = T_1 < T_2 < \cdots < T_{n_T} \), so we always keep sampling the density we are interested in, but also densities in higher temperatures. Then we construct a mechanism that couples the different temperature chains together, so they can exchange information on their location in the parameter space. Chains in densities with higher temperatures can in effect tell the chains sampling the colder densities where they should move. In practice, this exchange of information is done by swapping the locations of two chains.

### Between-chain Metropolis update

Swapping the locations must preserve the distributions of the respective tempered chains, and this is achieved by once again using the Metropolis rule (3.20). This gives a general name for parallel tempering and other similar auxiliary variable methods, Metropolis-coupled Markov chain Monte Carlo, abbreviated sometimes as MCMCMC or MC\(^3\).

The Metropolis acceptance ratio can be derived as follows in the between-temperature swapping case. First, let us define the inverse temperature \( \beta_i \equiv 1/T_i \), and consider the joint density of two chains, \( \pi(\theta_i, \theta_j) \), that are at temperatures \( \beta_i \) and \( \beta_j \):

\[
\pi(\theta_i, \theta_j) = \pi(\theta_i)^{\beta_i} \pi(\theta_j)^{\beta_j},
\]

(3.34)

where \( \theta_i \) is the current location of the Markov chain that samples from the distribution \( \pi(\theta_j)^{\beta_i} \) (respectively \( \theta_j \) for \( \pi(\theta_j)^{\beta_j} \)). Proposing to swap the locations is equal to proposing new values \( (\theta_i', \theta_j') = (\theta_j, \theta_i) \). Then we can write the Metropolis ratio:

\[
A_{j,i} = \frac{\pi(\theta_i', \theta_j')}{\pi(\theta_i, \theta_j)} = \frac{\pi(\theta_j)^{\beta_i} \pi(\theta_i)^{\beta_j}}{\pi(\theta_i)^{\beta_i} \pi(\theta_j)^{\beta_j}},
\]

(3.35)

and the acceptance probability of the swap is \( a = \min(1, A_{j,i}) \). Often for numerical reasons we want to deal with the logarithms of the proposal densities. In that case the Metropolis ratio simplifies to

\[
A_{j,i} = \log \left( \frac{\pi(\theta_i')}{\pi(\theta_i)} \right)^{\beta_i - \beta_j}.
\]

(3.36)

In Bayesian inference, we often do not want to temper the whole posterior density but only the likelihood:

\[
\pi_T(\theta|y) \propto \pi(y|\theta)^{\frac{1}{T_T}} \pi(\theta).
\]

(3.37)

Then if we set \( T_{n_T} = \infty \), at the highest temperature we are sampling only the prior, which is often easy to sample. In addition, values outside the prior support we have defined as infeasible, so restricting the hottest chain to search in the prior volume is a natural limit for the space.

The Metropolis ratio for the case (3.37) reduces to

\[
A_{j,i} = \log \left( \frac{\pi(y|\theta_i')}{\pi(y|\theta_i)} \right)^{\beta_i - \beta_j}.
\]

(3.38)
because the prior is cancelled out. In principle, swaps can be proposed between any two chains, but in practice it is enough to propose to swap adjacent chains since the swap acceptance ratio drops quickly with a growing temperature difference.

**Temperature ladder**

An essential factor affecting the efficiency of parallel tempering algorithms is the number and distribution of the temperatures $T_1 < T_2 < \cdots < T_{n_T}$ (the temperature ladder). Analogously to the MH acceptance rate issue, if the temperatures are too close to each other the distributions are nearly the same and thus almost all swaps are accepted. But there is no advantage to sampling the same distribution, so we want to spread the temperatures out. On the other hand, if the temperature difference is too big, none of the swaps are accepted and we are simulating multiple auxiliary chains for nothing. It turns out that a similar acceptance rate as the within-chain rate (23 % in the asymptotic limit) is optimal for the between-chain swaps as well [142]. Then our goal should be to first make sure the swapping rate between adjacent chains is equal for all pairs, and close to 23 %. A related topic is the maximum temperature we should use, if we do not use the tempered likelihood form (3.37) where it can be set to infinity. The higher $T_{n_T}$ is, the higher is the number of temperatures needed which increases the computational cost. However, for the sampling to be efficient $T_{n_T}$ must be high enough so that the peaks have flattened out.

It is not possible to choose a perfectly efficient temperature ladder before starting the simulation because we do not know the type of the target distribution. A good starting point for the temperature ladder is to distribute the temperatures geometrically [49, 143]. This provides a nearly constant acceptance rate for swaps in many systems. However, big changes may be still needed before the temperature ladder works optimally, and doing it by trial and error is very time-consuming. One option is to adapt the temperature values automatically during the simulation, where $T_1 = 1$ and $T_{n_T}$ are kept constant and the rest are adapted so that the acceptance rate for between-chain swaps is equal between all temperatures. This maximises the communication efficiency between the chains. For examples on the temperature adaptation, see [51, 127].
4 Review of thesis results

In this Chapter, we review the main results in publications I–III. Methods that we used to solve the forward and inverse problems in these publications were discussed in Chapters 2 and 3. We will therefore concentrate here mainly on describing the measurement setups and findings based on the parameter estimates.

4.1 PUBLICATION I: CHARACTERISATION OF RIGID FRAME POROUS MATERIALS FROM IMPEDANCE TUBE MEASUREMENTS

In publication I, we studied the characterisation of air-saturated porous media using the Bayesian approach, and compared the results to a least squares method. The main contribution of the paper was to provide a robust and easy-to-use characterisation method for rigid frame materials. In particular, we computed parameter uncertainty estimates and examined correlations between the parameters. The measurements were done in a standard transmission impedance tube (see for example [52]).

4.1.1 Measurement setup

We selected five materials, three polyurethane foams and two wools, that are commonly used, for example in sound and heat insulation applications in buildings and cars, see Fig. 4.1. Naturally, the proposed characterisation method extends beyond these materials and can be used with any air-saturated media that we can measure in an impedance tube and model as an equivalent fluid. These include glass beads, sandy soils and gravel, or even living plants [5]. The impedance tube where the samples were tested is shown in Fig. 4.2. The sound source is on the left, the sample in the middle, and there are two microphones on both sides of the sample. The tube is closed off by a semi-anechoic ending. By sending a logarithmic sine sweep, and using the scattering matrix formalism, we measured the waves reflected and transmitted through the tested sample at frequencies 300–5800 Hz. The reflection and transmission coefficients were then converted into the equivalent density $\tilde{\rho}_{\text{meas}}$ and the equivalent bulk modulus $\tilde{K}_{\text{meas}}$ of the material.

Figure 4.1: Samples measured in publication I: (1) soft polyurethane foam, (2) acoustic foam, (3) melamine foam, (4) felt, (5) glass wool.
4.1.2 Inversion

As mentioned earlier, we represented the porous materials using the JCAL model, which gives us directly the density $\tilde{\rho}_{\text{model}}^{\text{eq}}$ and bulk modulus $\tilde{K}_{\text{model}}^{\text{eq}}$. We write the observation model as

$$y = h_{\text{JCAL}}(\theta) + e,$$  \hspace{1cm} (4.1)

where $\theta = [\phi, a_\infty, \Lambda, \Lambda', k_0, k'_0]$, and $y = [\tilde{\rho}_{\text{meas}}^{\text{eq}}(\omega), \tilde{K}_{\text{meas}}^{\text{eq}}(\omega)]^T$.

The cost functional for the least squares solution was formed as

$$f(\omega; \theta) = W_1 \left\| \tilde{\rho}_{\text{meas}}^{\text{eq}}(\omega) - \tilde{\rho}_{\text{model}}^{\text{eq}}(\omega; \phi, a_\infty, \Lambda, k_0) \right\|^2 + W_2 \left\| \tilde{K}_{\text{meas}}^{\text{eq}}(\omega) - \tilde{K}_{\text{model}}^{\text{eq}}(\omega; \phi, \Lambda', k'_0) \right\|^2,$$  \hspace{1cm} (4.2)

where the weights $W_1 = 1/\rho_{\text{RMS}}^2$ and $W_2 = 1/K_{\text{RMS}}^2$ are added to compensate for the different amounts of influence the measurements of the density and bulk modulus have on the estimate of $\phi$. Note that the viscous and thermal effects are decoupled when we form the minimization problem this way, and only porosity is found in both expressions. The LS solution is then obtained as

$$\hat{\theta} = \arg \min_{\theta \in D} f(\omega; \theta),$$  \hspace{1cm} (4.3)

where $D$ is the domain to which we want to restrict the parameter values. In this paper, the only restriction is to bound the values of the unknowns to physically reasonable ones, i.e. $\phi \in [0, 1]$, $a_\infty \in [1, 10]$, and so on. The minimisation was carried out using the Nelder-Mead algorithm [144]. We found that with some materials the algorithm often converged to different minima depending on the initial values. To increase the chances of finding the global minimum we used a multistart approach, where the minimisation was started several times with the initial values drawn randomly around the parameter space.

Let us now describe the construction of the prior and likelihood densities so that we can compute the Bayesian estimates. To achieve results comparable to those given by the LS solution, we use a uniform prior that provides the same physical constraints as in the LS case. We assume that the measurement is corrupted by white noise, with the covariance
Figure 4.3: Measurements of density and bulk modulus, shown as real and imaginary parts, of two of the samples (black continuous lines), and the model prediction corresponding to the MAP estimate plotted on top (dotted lines).

\[ \Gamma_e = \begin{bmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{bmatrix}, \]

where \( \Gamma_1 = \sigma_{\rho e}^2 \rho_{\text{RMS}}^2 I \) and \( \Gamma_2 = \sigma_{K e}^2 K_{\text{RMS}}^2 I \). The quantities \( \sigma_{\rho e} \) and \( \sigma_{K e} \) denote the standard deviations of the measurement error, estimated from repeated measurements, and \( I \) represents an identity matrix. The posterior density can then be stated as

\[
\pi(\theta | y) \propto \exp \left\{ -\frac{1}{2} \| L_e (y - h_{\text{JCAL}}(\theta)) \|^2 \right\} \pi_{\text{unif}}(\theta),
\]

where \( L_e L_e^T = \Gamma_e^{-1} \). We used the adaptive Metropolis algorithm to compute statistics of the posterior.

4.1.3 Results

The fit of the JCAL model (with the MAP estimate as the input) to measurements of the density and bulk modulus of two example materials is shown in Fig. 4.3. The largest model discrepancy is observed at the solid frame resonance frequency (\( \sim 1 \) kHz for the foam and \( \sim 2 \) kHz for the wool), where the coupling between the air and the frame is higher and the frame vibrates in response to the acoustic excitation. The effect of this model discrepancy on the inversion is relatively small, however, because the frequency range where the resonance effects appear is narrow and the model errors exhibit symmetry around the reconstructed curves. We observed similar model fits to all measured samples (see publication I).

Figure 4.4 shows posterior summaries for the acoustic foam sample. Each subplot shows the joint marginal distributions of the estimated parameters, the one
Figure 4.4: Marginal posterior densities of the acoustic foam sample. The red curves on the bottom and left side of each plot are the 1D marginal densities, and the contour lines indicate the 50 and 95% regions of distribution. The dots represent a sample of the MCMC points, and $r$ is the Pearson correlation coefficient. The black cross denotes the MAP-estimate, and the red circle the LS-solution.

dimensional marginal distributions and the MAP and LS point estimates. The two point estimates coincide very accurately, as expected, and are always found within the 50% regions of distribution. We can observe a strong positive correlation between porosity $\phi$ and the thermal characteristic length $\Lambda'$. However, both parameters are still resolved well as shown by the narrow support of the posterior distribution.

The posterior densities of all the foam materials showed only one mode. They were thus easy to sample and we did not encounter identifiability issues. The posterior densities of the wool samples were still mostly unimodal, but especially the thermal parameters ($\Lambda'$ and $k_0'$) were not as well resolved as with the foams. One reason for this is that the wool samples showed more anisotropy, and consequently
4.2 PUBLICATION II: COMPUTATIONAL METHODS FOR ESTIMATING THE MATERIAL PARAMETERS OF POROELASTIC MEDIA

In the second publication we examined the estimation of the parameters of a poroelastic plate from simulated ultrasound measurements. The plate was assumed to be submerged in water so that acoustic waves can excite the frame of the material. The motivation for this study was to develop statistical inversion methods that work with the strongly non-linear poroelastic models, design a simple measurement that could be carried out in practice, and assess what kind of information can be extracted from this kind of measurements. A numerical study is needed so that we can compare the estimates and the true parameter values to evaluate possible estimation biases.

4.2.1 The simulation model

We considered a measurement setup shown in Fig. 4.5, where \( \Omega^{[1]} \) denotes the poroelastic layer, and \( \Omega^{[0]} \) and \( \Omega^{[2]} \) are semi-infinite halfspaces of water. The measurement data we wanted to obtain were the acoustic reflection and transmission coefficients \( R^{\text{exp}} \) and \( T^{\text{exp}} \), respectively. To simulate the measurements while avoiding the inverse crime [38], we used a finite-element model constructed in COMSOL Multiphysics® software v. 5.3. We truncated the computational domain by the following boundary conditions: a plane wave radiation condition on the left side (\( \Gamma_1 \)) to create the incident plane wave, an absorbing impedance boundary condition on the right (\( \Gamma_2 \)), and a periodic Floquet condition on the top and bottom to simulate a geometry extending to infinity in the \( x_2 \) direction. We considered frequencies between 100 and 400 kHz, with 5 kHz steps, and added Gaussian noise to the final measurements. The simulation was repeated for angles between 0° and 45°, with 5 degree steps.
4.2.2 Inversion

Our goal was to estimate all parameters included in the Biot model that are related to the poroelastic plate (i.e. we assumed that the parameters of the saturating fluid are known). The real and imaginary parts of the elastic parameters were estimated separately, and we also estimated the angle of the incoming wave. The vector of all unknowns thus reads

$$\theta = [\Lambda, \alpha_\infty, \log_{10}(k_0), \phi, \text{Re} \ K_b, \text{Im} \ K_b, \text{Re} \ K_s, \text{Im} \ K_s, \text{Re} \ N, \text{Im} \ N, \rho_s, L, \phi_{\text{inc}}].$$

We expressed permeability on a logarithmic scale because of the huge range of values it can take. In addition, we estimated the measurement noise level(s) simultaneously with the material parameters. We formed the measurement vector by stacking the reflection and transmission coefficients together as

$$y(\phi_{\text{inc}}) = [\text{Re} \exp(\phi_{\text{inc}}), \text{Im} \ exp(\phi_{\text{inc}})]^T$$

in the same vector, and we considered inversion from one angle at a time. The measurement noise was again assumed to be additive, and the noise covariance read

$$\Gamma_e = \begin{bmatrix} \sigma_{r_k}^2 \mathbf{I} & 0 \\ 0 & \sigma_{t_k}^2 \mathbf{I} \end{bmatrix},$$

(4.6)

where $\sigma_{r_k}^2$ and $\sigma_{t_k}^2$ are the noise variances in the reflection and transmission measurements, respectively. We can now denote the full vector of unknowns by $\tilde{\theta} = [\theta, \sigma_{r_k}, \sigma_{t_k}]$.

The prior was constructed using a multivariate Gaussian density, $\pi(\tilde{\theta}) \sim \mathcal{N}(\tilde{\theta}_*, \Gamma_{\tilde{\theta}})$, whose tails were truncated using an indicator function $B(\tilde{\theta})$:

$$B(\tilde{\theta}_k) = \begin{cases} 1, & \text{if } \tilde{\theta}_{\text{min},k} \leq \tilde{\theta}_k \leq \tilde{\theta}_{\text{max},k} \\ 0, & \text{otherwise}. \end{cases} \quad (4.7)$$

The indicator function was used to enforce hard boundaries such as positivity or negativity constraints, as well as the physical constraint $K_b \leq K_s(1 - \phi)$ [10].

Figure 4.6: An example of the simulated noisy measurement (at normal incidence) and the fit of the plane wave Biot model.
Figure 4.7: The marginal prior and posterior distributions of tortuosity, permeability, porosity, and the real part of the bulk modulus, estimated from measurements at normal incidence. Also shown is are the true values, the CM estimates, and the 95% credible intervals.

The final step in the inverse problem was to select a suitable MCMC sampler to calculate posterior statistics. We could not make the AM algorithm, used in publication I, work satisfactorily in the present problem, because the sampler had a tendency to get stuck in a local maximum for long periods of time, which could lead to false diagnosis of convergence if we did now know the true parameter values. A solution we found to work very well was to use parallel tempering, especially of the form (3.37) where we could set the highest temperature to infinity. Especially when the PT algorithm was combined with an adapted proposal density and adapted temperature ladder, the sampler was easily able to jump from any starting point within the prior support to sampling close to the true parameter values.

An example of the simulated noisy measurement is shown in Fig. 4.6. The figure also shows the (noiseless) model prediction that corresponds to the CM estimate, and the posterior predictive distribution. We can observe that the measurement points are all found within the ±3 standard deviations of the posterior predictive distribution, which suggests that the noise levels were estimated correctly.

4.2.3 Results

The most interesting measurement to consider is the one at normal incidence, since it is simpler to carry out than measurements at oblique incidences. We found that many parameters can be identified from the normal incidence data, in the sense that
the posterior variance is much smaller than the prior variance. These parameters include porosity \( \phi \), tortuosity \( \alpha_\infty \), viscous characteristic length \( \Lambda \), and density of the solid \( \rho_s \).

The marginal posterior distributions of some of the parameters are shown in Fig. 4.7, which also shows the marginal prior density, point estimates, and the 95% credible intervals. An interesting observation is that while the marginal distribution of permeability \( k_0 \) has a peak at the location of the true value, the distribution is highly skewed towards higher values of permeability. It seems that there is a clear lower bound for \( k_0 \), but once past a certain value, changes in permeability do not show in the data any more. Apart from \( k_0 \), the marginal posterior densities seem to be centred around the true parameter value, and the CM estimate tends to be very close to the truth.

The joint marginals posterior distributions reveal more information on the parameters and their identifiability. Figure 4.8 shows the joint distributions of pa-
rameters that exhibited the most correlation. The strongest correlation is between the real parts of the bulk and the shear moduli, which are almost exactly linearly negatively correlated, and both parameters admit values from zero to way past the true values. However, the true values of the parameters are on the long and narrow support of the posterior which shows that the parameter estimates and their uncertainty bounds still contain the correct value. Another parameter that exhibits strong correlation with the elastic moduli is $\alpha_\infty$.

The strong correlation between $K_b$ and $N$ is a result of how the elastic moduli affect the wave speeds. The speed of the fast longitudinal wave is controlled by a linear combination of $K_b$ and $N$ (in addition to other factors such as density), whereas the only elastic modulus that controls the speed of the shear wave is $N$. Therefore we would need both types of waves present to resolve both moduli, which we do not have at normal incidence.

At oblique angles of measurement, the shear was is generated, and consequently we are able to estimate $K_b$ and $N$ with much better accuracy. This can be clearly seen in Fig. 4.9, which shows some of the parameter estimates, along with the estimate of the 95 % credible interval, as a function of the measurement angle. Unsurprisingly, the estimates of $K_b$ and $N$ get better as the angle increases, but also the estimate of $\alpha_\infty$ improves with a growing angle of incidence. This may be due to the fact that $\alpha_\infty$ is correlated with both elastic moduli at normal incidence, as Fig. 4.8 shows, and can be more accurately estimated as the estimates of the elastic moduli are more accurate. However, the accuracy of the estimates for parameters such as porosity and solid density does not change significantly with the incidence angle.
4.3 PUBLICATION III: ESTIMATING THE MATERIAL PARAMETERS OF POROELASTIC OBJECTS FROM ULTRASOUND MEASUREMENTS IN WATER

In publication III, we studied the applicability of the inversion method introduced in II to real measurements of a porous ceramic plate, a glass bonded silica, type QF-20 (Filtros Ltd., East Rochester, NY). The measurements were carried out in a water tank using focused ultrasound transducers, and in total 216 points on the plate were measured on a 80 × 230 mm regular grid, with 10 mm spacing between the measurement points. There were two main reasons for doing such a comprehensive set of measurements. First, estimating the material parameters independently at two or more locations very close to each other provides a sensitivity check for the results. Large sudden jumps in parameter values could indicate a problem with the inversion and/or measurement method. Secondly, the resulting numerous parameter (and uncertainty) estimates can be plotted as a map that shows the potential variability of the parameters as a function of the measurement location. This is one way to quantify the inhomogeneity of the object.

4.3.1 The measurement setup

The frequency range of the measurements depends on the validity of the model and practical constraints for producing ultrasound in water. The effects of the Biot waves can only be seen below frequencies where scattering effects start to dominate, which we found for the studied material submerged in water to start at around 400 kHz. To produce low-frequency ultrasound, we used Panametrics V389 (Olympus Scientific Solutions Americas Inc., MA) point-focused transducers whose central frequency was 500 kHz. The usable frequency range of these transducers goes down to 200 kHz. The transducers were immersed in the water tank and connected to a computer-controlled measurement system. The porous plate was placed in the focal zone of the transducers, and the -6 dB beam diameter of a pulse/echo signal at

![Figure 4.10: A schematic view of the measurement system and the poroelastic plate, a) reflection measurement, b) transmission measurement. The shaded zones indicate the -6 dB size of a pulse/echo signal on the surface of the plate.](image-url)
400 kHz at this distance was 10 mm. The emitting transducer was excited by a spike signal, which produced a short ultrasonic pulse in water. The received signals were then recorded on a computer, and 40 pulses were averaged for each measurement to reduce the effect of random noise.

We then computed the reflection and transmission coefficients of the plate, $R^{\text{exp}}$ and $T^{\text{exp}}$, respectively, using the spectral ratio technique. First, the reflected and transmitted pulses are recorded using the two configurations shown in Fig. 4.10, and then a reference signal for both configurations is also recorded. For reflection, a single transducer is used in the pulse/echo mode and the reference signal is obtained by pointing the transducer upwards to the water-air interface. In the transmission configuration, two transducers are used and the reference signal is obtained in water. Finally, we take the Fourier transform to represent the measurements in the frequency domain, and use for example the Wiener deconvolution filter [145] to obtain $R^{\text{exp}}$ and $T^{\text{exp}}$.

4.3.2 Inversion

The inversion method used here is similar to the one introduced in publication II, but a few modifications were made. In II, we modelled the visco-elastic losses in the solid with an imaginary part that was constant with respect to frequency. However, such an approach can be shown to be slightly non-causal [78], which is why we implemented the constant Q model [79] of Kjartansson to represent the losses with a causal model. To differentiate the elastic parameters from publication II, we denote their modulus here with a subscript of zero, like $K_{b,0}$ and the associated quality factor by $Q_{K_b}$ (respectively for $K_s$ and $N$).

In addition to the plate’s physical properties, we considered the measurement noise levels, thickness of the plate, and a distance mismatch term $\varepsilon_R$, as nuisance parameters that we are not primarily interested in but nevertheless improve the inversion results. The term $\varepsilon_R$ is defined as the difference in the length of the signal path between the pulse/echo reflection measurement and the reference measurement, and has an effect on the phase of $R^{\text{exp}}$.

The prior was constructed using truncated normal distributions, whose mean values were selected based on information that was available from the manufacturer of the plate and other studies. The prior variance was set to be large so that a wide range of parameter values were possible. In addition to the constraints employed in II using the indicator function, we included a constraint $2/3N_0 \leq K_{b,0}$ which corresponds to requiring that the Poisson’s ratio of the solid frame is positive.

Parameter inference at each measurement location was carried out independently of the other points.

4.3.3 Results

There is a large number of measurements and inversion results to examine, which is why we will summarise the results by showing only the point estimates and the width of their credible intervals. However, before presenting this summary let us concentrate on a randomly selected measurement location, and plot the model fit and marginal posterior distributions. This gives a more detailed view on the inversion, and lets us compare the findings in publication II.

Figure 4.11 shows the measurement of $R^{\text{exp}}$ and $T^{\text{exp}}$ from one of the measured locations, along with the model prediction that is computed using the CM estimate.
Figure 4.11: An example of typical measured reflection and transmission coefficients, and the fit of the Biot model. The shaded areas denote the one, two, and three standard deviations of the posterior predictive distribution.

The figure also shows the posterior predictive distribution using standard deviations. Perhaps the only clearly visible systematic discrepancy between the model and measurements can be seen at low frequencies of the transmission coefficient. Elsewhere the measurement points are found within plus or minus two standard deviations, and this was a trend observed in the other measurement locations as well.

Continuing with the same measurement location, the marginal posterior distributions of $a_\infty$, $\log_{10}(k_0)$, $\phi$, $K_{b,0}$, $N_0$, and $\rho_s$ are shown in Fig. 4.12. As was the case for the estimates of the numerical study, tortuosity, porosity, and solid density are well identified in the sense that their marginal posterior distributions are narrow compared to the prior distributions. Permeability also exhibits a trend similar to that in publication II, where it has a definite lower bound, but can admit values up to the highest allowed value of $10^{-8}$ m$^2$. However, as opposed to the finding of publication II, both $K_{b,0}$ and $N_0$ seem to be well identified even from the normal incidence data. One clear reason for this is the condition $2/3N_0 \leq K_{b,0}$ that was added to the prior. Indeed, the estimated values of $K_{b,0} \approx 13$ GPa and $N_0 \approx 18$ GPa are close to this relation, and the estimated Poisson’s ratio of the porous frame is very close to zero. In contrast, at this location the estimate for the bulk modulus of the solid is $K_{s,0} \approx 22$ GPa, and the Poisson’s ratio is about 0.18.

The effect of the prior on $K_{b,0}$ and $N_0$ can also be seen in the joint marginal densities, shown in Fig. 4.13, where the line $2/3N_0 = K_{b,0}$ is clearly visible. Other correlations can be seen between the solid density and the bulk and elastic moduli, as well as between tortuosity and the solid bulk modulus. However, the observed correlations are lower than in the numerical case. This may be explained by the fact that with the real measurements we have a lot more model error, and therefore the model together with the data behave in a different way.

Let us now summarise the results in Fig. 4.14, which shows the CM and 95% CI estimates, for selected parameters, of all 216 measurements. The estimates are plotted according to where the measurement data they are based on was taken at, and hence the images show how the parameter estimates and their uncertainties vary spatially. We can see immediately that overall the parameters show variability
over the scale of the plate, but the small-scale variability from one measurement location to the next is relatively small. Only at a very small point there is a sudden jump in the estimated values (visible especially in the estimates of $\Lambda$ and $\alpha_\infty$).

Because the inversion at each point is carried out independently of the other locations, the smooth varying of the parameters must be found in the measurement data. The smoothness in the data may be explained by two factors. The first possibility is that the resolution of the measurements is significantly lower than the spacing of the measured locations, i.e., the effective area seen by the transducers is much larger than the 10 mm pulse/echo beam diameter. If this were the case, however, we would see a low resolution in all parameters whereas now the estimates of $\phi$ and $\rho_s$ show structures much smaller than the ones seen in $\alpha_\infty$ and $K_{b,0}$. We can then conclude that the observed smooth variability in the parameters is a property of the object.

Finally, Fig. 4.15 shows the spatial variability of the estimated nuisance parameters. None of the nuisance parameters can be seen to vary in the same way as the Biot parameters, which further shows that the variability in the pore parameters comes from the measured plate. Another notable observation is that the nuisance parameters can be identified accurately, as shown by the relatively narrow widths of the 95% credible intervals. For example, the distance from the pulse/echo transducer to the plate is estimated to vary by up to 3 mm (the distance between the transducer and the water-air interface is constant in the reference measurement, so changes in $\varepsilon_R$ show directly the changes in distance between the transducer and the plate), but the associated uncertainty interval for the distance estimate is less than

Figure 4.12: The marginal prior and posterior distributions of selected parameters. Estimated from an example measurement in the water tank. Also shown is are the CM estimates and the 95% credible intervals.
Figure 4.13: The joint posterior distributions and the CM estimate of selected parameters, estimated from the water tank measurements.

0.06 mm. The estimates of $\varepsilon_R$ also show that the plate is not completely planar. In addition, the thickness of the plate can be seen to vary by almost a millimetre, with an uncertainty of less than 0.15 mm. These nuisance parameters are clearly identified independently, and accounting for them during the inversion thus improves the estimates of the Biot parameters.
Figure 4.14: The CM and 95 % CI estimates of selected parameters at every measured location (the pixelated images), and a schematic that shows the plate and the locations inside it that were measured (denoted by the crosses). The parameter estimates were calculated separately for each measurement. Therefore, the colour of the pixels denotes the value of the of the estimated quantity, and the location of the pixels corresponds to the location where the measurement was made.
<table>
<thead>
<tr>
<th>Param.</th>
<th>CM estimate</th>
<th>95 % CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L$</td>
<td>25.51</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>24.55</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td>1.76</td>
<td>0.057</td>
</tr>
<tr>
<td>$\epsilon_R$</td>
<td>-1.47</td>
<td>0.026</td>
</tr>
<tr>
<td>$\sigma_{e_T}$</td>
<td>0.14</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>0.011</td>
</tr>
<tr>
<td>$\sigma_{e_R}$</td>
<td>0.08</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>0.04</td>
<td>0.012</td>
</tr>
</tbody>
</table>

**Figure 4.15**: The CM and 95 % CI estimates of the nuisance parameters.
5 Summary and conclusions

In this thesis, we developed Bayesian characterisation methods for poroviscoelastic and rigid frame porous media. The motivation was to improve the reliability and repeatability of the characterisation, by considering measurement and model errors, incorporating prior information, and finally approximating the true posterior probability density of the parameters using sampling methods. In particular, our goal was to provide feasible error estimates for each unknown parameter. We considered materials that were saturated either with air or water, and tested the proposed methods with real measurement data.

In publication I, we proposed a Bayesian method for the characterisation of rigid frame materials, and compared the results to a least squares (LS) estimate, and other characterisation methods. We used a standard transmission impedance tube to measure the materials’ equivalent density and bulk modulus, and the Johnson-Champoux-Allard-Lafarge (JCAL) model [69–71] to connect the measurement data to six physical parameters. The parameter inference in the Bayesian framework was then carried out using an adaptive Markov chain Monte Carlo (MCMC) technique, and a point estimate was reported in the form of the maximum a posteriori (MAP). The results showed that the measurement data carried information on all JCAL parameters, and that the MAP estimates were within a few percent of those found using the reference methods. In addition, the LS estimates were found to be nearly identical to the MAP estimates, but only if the minimisation of the cost functional was done using a multistart approach. Otherwise the LS solution often returned a local minimum. The statistical approach also showed that most parameters are identified independently, but with sometimes certain parameters may exhibit strong correlation with one another. However, the correlations were found to be different for each material, which is why it is informative to summarise the posterior using MCMC each time a new material is measured.

Publication II was a numerical study in which we developed inversion methods for poroviscoelastic media, and designed an experimental setup where the characterisation could be also done in practice. To make the inversion using sampling methods possible, we needed to have a simple measurement geometry which could be modelled with fast numerical methods. For this reason, the measurement setup consisted of a flat slab of porous material submerged in a water tank. Such a setting can be considered as a layered system, and if we assume plane wave propagation, several fast methods are available to solve the equations. The porous medium was modelled using the Biot equations, with the Johnson et al. model [69] to account for viscous losses. The forward model was then constructed using a global matrix method, which is both a fast and stable way to solve wave propagation problems in layered media. We simulated the measurement data, which consisted of the acoustic reflection and transmission coefficients of the porous plate, with a finite-element model, and considered several angles of incidence.

A key contribution was to construct an adaptive parallel tempering MCMC algorithm, and to demonstrate that it is a viable method to compute the parameter estimates in the considered setting. The results showed that many interesting pa-
parameters, including porosity, can be estimated from just the measurements at normal incidence, and incorporating measurements from oblique angles would not improve the estimates much. These parameters include porosity and permeability. However, if accurate information on the elastic parameters is desired, measurements at oblique angles of incidence are recommended so that the effect of the shear waves can be seen.

In publication III, we implemented the experimental design from publication II in a water tank measurement system in the Laboratoire d’Acoustique de l’Université du Mans (LAUM), carried out extensive measurements of a porous ceramic plate, and applied the inversion method introduced in II to the measurements. The goal was to evaluate the method with real measurements and to estimate the possible spatial changes in the parameters. The actual measurement system consisted of focused ultrasound transducers connected to a computer-controlled positioning system, which was used to measure the reflection and transmission coefficients at several locations. However, due to the limitations of the system, we were able to make the measurements only at normal incidence. The inversion was carried out independently at each measurement location. In addition to the physical parameters of the plate, we estimated some nuisance parameters related to the measurement setup, such as the measurement noise level.

The measured reflection and transmission coefficients were found to vary significantly from location to location, suggesting that the plate is inhomogeneous. Nevertheless, the fit of the Biot model to the measurement data was equally good for nearly every measurement, and the measurements were contained within the envelope of the posterior predictive distribution, which indicates that the employed model is a feasible one. However, we found that the residuals between the model and the measurements were not distributed exactly as white noise. This implies that there are features of the measurement system and the measured object that are not accounted for in our forward model, such as possible depth-wise inhomogeneities of the plate, errors resulting from the plane wave approximation, and attenuation due to scattering effects. Modelling errors can obviously introduce bias into the parameter estimates, but quantifying this was not pursued in this thesis. We also found that the observed model residuals were significantly smaller than the variability in the measurements between different locations of the ceramic plate.

The inversion results for each unknown were plotted as two-dimensional graphs which showed the conditional mean (CM) estimate and the length of the associated 95 % credible interval at each measurement location. The parameter estimates were different depending on the location of the measurement, but at the same time the changes in the values were gradual and smooth. In addition, the parameters did not all vary in the same way, and in particular the pore structure parameters did not vary in the same way as the nuisance parameters, such as the plate thickness, which suggests that the observed variation is real and not an artefact of the model errors. In conclusion, the results indicate that the employed computational model can be used to feasibly estimate the parameters of viscoelastic porous media.

Although the results of this thesis have shown that the proposed methods are a promising step towards the robust characterisation of porous materials, a lot of future work remains. For instance, throughout the thesis we assumed that we have only plane wave propagation, which can be easily justified for an impedance tube, but is only an approximation in free-field measurements. Modelling the finite, non-planar, field of the ultrasound transducers would be an obvious improvement for the inversion, and would let us quantify possible biases the plane-wave approxima-
tion leads to in the parameter estimates. One way to model the true ultrasound field is to use element-based methods such as the finite element or the boundary element methods. This would also enable the characterisation of objects that are of irregular shape. However, modelling complex objects and acoustic fields using element-based approaches will substantially increase the computational burden, which may quickly render the proposed MCMC method infeasible. There is clearly a need for fast computational methods that can still account for a finite-sized acoustic field.

One possible approach to model more realistically the fields emitted by ultrasound transducers, without having to resort to full-wave solutions, is to use the superposition of harmonic plane waves. The idea is that the finite incident field is decomposed into infinite plane wave components using the Fourier transform, and then a plane wave solution in the layered media is computed for each component. Finally, to find the reflection and transmission coefficients, the responses of each plane wave component are summed together [88, 146]. The computational requirements of the plane wave decomposition approach is naturally directly proportional to the number of plane waves considered.

Another possibility in the Bayesian framework is to construct a statistical model for the errors that are induced when we use approximate forward models, and to marginalise over them before the inversion. This is called the Bayesian approximation error (BAE) approach [35, 147]. An advantage of using the BAE is that once the statistical model for the approximation errors (the discrepancies of the predictions between an approximative and a more accurate model) have been computed, we can use the fast approximative model in the inversion, while partly recovering from the errors. Applied to characterisation of porous media, we could consider a finite element model as the accurate one, and marginalise over the approximation errors when using a plane wave model.

Solving the inverse problem using MCMC methods is, however, always going to be relatively slow. If a significantly faster, or even real-time, characterisation method is needed, we have to compute the MAP estimate which can often be obtained with only a few iterations, and use the Laplace approximation to get an estimate of the posterior covariance.

During the thesis work, we implemented and tested numerous MCMC algorithms for the problem of poroelastic material characterisation, and found that adaptive Metropolis and parallel tempering worked efficiently. However, there are certainly other good algorithms that could work equally well or better, and new MCMC samplers are being developed all the time. While parallel tempering is one of the best ways to deal with multimodal densities, other methods such as Hamiltonian Monte Carlo [112] may be more efficient in different settings, especially if derivatives of the posterior are easily obtained. Additionally, nowadays there are many freely available MCMC software packages which can facilitate easy out-of-the-box sampling [132, 148–150].

The measurements of the porous ceramic layer carried out during this thesis were done only at normal incidence. This was mostly a limitation of the measurement system in the water tank, which did not allow for accurate measurements of the transducer angle that would be needed for the oblique incidence measurements. Recently, LAUM has acquired a new computer-controlled measurement system which would allow these measurements. An obvious extension would therefore be to test the proposed inversion method with measurements at oblique angles, which would give more information especially on the bulk and shear moduli that were found to be correlated at normal incidence.
The inversion method should also be tested with control materials whose properties are known. The problem of obtaining such porous materials is that porous materials are inherently heterogeneous, and reliable control materials manufactured using traditional methods are therefore hard to come by. On the other hand, because the pores need to be very small for the setup and frequencies we have considered (on the order of micrometers), 3D printing such structures is only now starting to become possible.

Another generalisation for the characterisation of porous materials would be to account for the possible macroscopic heterogeneity and anisotropy of the samples. Macroscopically heterogeneous porous materials offer, for example, more possibilities for the design of sound absorbing materials [151,152], and anisotropic materials are found in many porous structures such as fibrous sound and heat insulation materials and sediment structures, as well as acoustic metamaterials [153]. Gradually changing material profiles in one dimension have been estimated in the rigid frame case using the wave splitting and invariant imbedding techniques, with a LS minimisation approach [154]. Another possibility to model macroscopically inhomogeneous layered materials is to use the reversible jump MCMC methods [116], where the number of the layers can be another unknown. Such methods provide estimates with uncertainties for the parameters within the layers, but also include the uncertainty in the model parametrisation [155,156]. An approach for the characterisation of three-dimensional anisotropic materials has been proposed in [157].
BIBLIOGRAPHY


Porous materials find applications in a wide range of fields from medicine to engineering. There is a need to reliably characterise these materials, preferably in a non-destructive manner. In this thesis, new methods based on acoustic measurements and statistical inversion are developed for the non-destructive characterisation of porous materials. The developed methods also provide quantitative information on the uncertainties in the characterisation. The methods are tested with both simulated and experimental data.