Bayesian Inversion for Magnetic Resonance Elastography

Fabian Cejka

March 6, 2023

Abstract

In this thesis we study an inverse problem arising in Magnetic Resonance Elastography (MRE) which is a noninvasive method for quantifying soft tissue stiffness. We develop a Bayesian formulation of this problem which involves inferring (heterogeneous) elastic properties in the time-harmonic purely elastic or viscoelastic wave equation. We apply modern Ensemble Kalman Inversion (EKI) algorithms which are derivative-free and provide robust approximations of the Bayesian posterior in a computationally tractable manner. Moreover, we show how parametrisations of EKI can be used to design effective inversions of properties with complex geometries relevant to the detection of diseased tissue via MRE.

In *in-silico* experiments, we showcase under the viscoelastic and purely elastic modelling assumptions that EKI can provide accurate estimates of the unknown local tissue stiffness and we also discuss limitations of both models. In particular we test EKI using the viscoelastic model in virtual experiments with complex geometries and unknown elastic properties that occur, for example, in brain MRE. We demonstrate how our algorithms are able to successfully discover cancer tissue and provide confidence intervals for

the estimates and predictions of tissue stiffness, which can be diagnostically valuable for physicians. Furthermore, we analyse the influence of the prior, the amount of noise in the data and the ensemble size on posterior estimates provided by EKI and discuss the design of informative priors for EKI.

NOMENCLATURE

α_n regularisation parameter in EKI, page 88

$\|\cdot\|_B$ Operator norm, page 77

B Covariance of the noise, page 76

$\mathbb{C}$ Set of complex numbers

η Noise in measurement data, page 63

$\mathcal{F}$ Forward map from parameter space $\mathcal{K}$ to measurements, page 75

$\mathcal{G}$ Forward map from physical parameter space $\mathcal{M}$ to measurements, page 62

$H^{3/2}(\Omega)$ Fractional Sobolev space of order $3/2$, page 36

J Ensemble size in EKI, page 92

$\mathcal{K}$ Space of parameters for the parametrisation of the unknown, page 75

$k_n^{(j)}$ j-th particle of the ensemble after n iterations in the parameter space $\mathcal{K}$ in EKI, page 92

λ 1st Lamé Parameter, page 53

L_∞ Vector space of essentially bounded measurable functions

$\mathcal{M}$ Set of admissible physical parameters, page 62

m Number of measurement locations and dimension of the data $\mathbf{v}$, page 62

$\mathfrak{m}_n$ Mean of the Gaussian approximation after n iterations in EKI, page 89

μ Shear modulus, page 53

μ_L Loss modulus, page 53

μ_S Storage modulus, page 53

$\mu_n^{(j)}$ j-th sample of the physical ensemble after n iterations in the physical parameter space $\mathcal{M}$ in EKI, page 174

N_S Iteration after which EKI is stopped, page 96

ν Poisson's Ratio, page 62

Ω Domain of Interest, page 53

ω Frequency of the Sinusoidal Excitation, page 53

$\partial\Omega_D$ Dirichlet Boundary, page 54

$\partial\Omega_N$ Neumann Boundary, page 54

$\mathcal{P}$ parametrisation map $\mathcal{K} \to \mathcal{M}$, page 75

$\mathcal{P}_1$ Parametrisation used in experiments using the purely elastic model, page 104

$\mathcal{P}_2^{(2)}$ Second level of the two-level parametrisation, page 107

$\mathcal{P}_2$ Two-level parametrisation used in experiments using the viscoelastic model, page 107

$\mathcal{P}_2^{(1)}$ First level of the two-level parametrisation, page 107

$\mathbb{R}$ Set of real numbers

ρ_W Density of soft tissue in the domain of interest, page 53

σ_{el} Stress tensor for infinitesimal deformations in an isotropic linearly elastic solid, page 51

Σ_n Covariance of the Gaussian approximation after n iterations of EKI, page 89

σ_N Noise level in data v , page 118

$\mathbf{v}$ Measurement data, page 63

v Time-harmonic wave field, page 53

v_t Time-dependent wave field, page 54

x	Spatial Coordinate in Domain of Interest, page 54

CONTENTS

1

INTRODUCTION

Inverse problems arise when physical quantities are not directly accessible for measurements, and other quantities are connected with the unknown quantities by physical laws. Gaining indirect access to the unknown quantities by inverting these laws is an inverse problem [35, 83] .

Modelling a radio channel in telecommunications that introduces many modifications to the information signal is a direct problem that involves physical phenomena like reflection, interference, refraction, diffraction, absorption, polarisation, and scattering [79]. Reconstructing the information signal from the measurements made at the receiver is an inverse problem.

The physical laws that connect the unknown with the measurements are often expressed by a system of differential equations [45], which is a formal problem within the inverse problem setting. The unknowns are the boundary conditions, the initial conditions or the parameters of the differential equations. Even if the forward problem is well-posed and the solution of the differential

equations exists and is unique and changes continuously with its parameters or boundary conditions, the inversion is not necessarily well-posed [85].

The solution of an inversion can change drastically for small changes in the measurements. For example, in the just mentioned radio channel in telecommunication due to the modifications like reflection, interference, refraction, diffraction, absorption, polarisation, and scattering measurements can be similar even though the information signal is very different. This makes inverse problems ill-posed and difficult to solve, especially in the presence of noise in the measurements [45, 86].

The inverse problem in this thesis has measurements $\mathbf{v}$ which are assumed to be corrupted by additive noise η and the unknown μ is the parameter of a differential equation which is denoted as a forward operator $\mathcal{G}$. We have an equation of the form

$$\mathbf{v} = \mathcal{G}(\mu) + \eta$$

and the inversion problem is to find μ from given $\mathbf{v}$. In the inverse problem in this thesis the solution depends sensitively on $\mathbf{v}$ and therefore the problem is ill-posed.

This thesis approaches an inverse problem arising in magnetic resonance elastography (MRE). MRE is a non-invasive technology that quantitatively assesses the mechanical properties of soft tissue in vivo [69]. It can be regarded

as "virtual palpation" [55], an imaging-based equivalent to manual palpation, a technique used by physicians to determine location, shape, size and firmness of diseased tissue which is in many cases stiffer than healthy tissue [26]. The analysis of soft tissue stiffness can provide valuable diagnostic information about diseases that affect the tissue stiffness like cancer, fibrosis and inflammation, even at an early stage of the disease [59].

In MRE, the idea is to perturb tissue using mechanical vibration applied to the skin and infer mechanical properties from the displacements which are measured using Magnetic Resonance Imaging (MRI) [14]. The measured displacements are the consequence of the mechanical properties of the tissue, so, inferring them from a certain measured displacement is an inverse problem.

In this thesis, we will develop a Bayesian formulation of the inverse problem arising in MRE and approximate the Bayesian posterior on the elastic properties by applying Ensemble Kalman Inversion (EKI) algorithms which trace back to Ensemble Kalman Filters [36, 19, 1]. The Ensemble Kalman Filter is a derivative free sampling based method that gives Gaussian approximations of the Bayesian posterior in data assimilation [18, 44]. In recent years it has been applied to the parameter estimation in inverse problems [37, 78, 40, 38]. An ensemble of random samples is used in order to represent distributions and is updated by Ensemble Kalman Filter rules which makes EKI computationally cheap and

allows for high-dimensional unknowns. Furthermore, it allows for advanced black box models because it is derivative free.

In MRE, we usually model biological soft tissue as heterogeneous, isotropic and linearly viscoelastic [63, Chapter 3]. Many approaches simplify this model further and assume that there is no time lag between deforming forces and the induced deformation. In other words this simplification models soft tissue as a Hookian material which is purely elastic and not viscoelasticity [14]. In this thesis the Bayesian approach is applied to MRE for both purely elastic and viscoelastic model, and we show that for this approach the viscoelastic model is clearly preferred over the purely elastic model.

The development of a Bayesian formulation of the inverse problem arising in MRE and the application of an approach from Bayesian inversion theory [83] to time harmonic MRE in this thesis is a novelty.

The Bayesian approach not only takes the noise in the data into account but also provides uncertainty of the provided estimates and allows for straightforward incorporation of prior knowledge.

In this thesis we provide confidence intervals for the estimates and introduce a quantity called ”probability of tumour” which is useful measure to quantify uncertainty of estimates.

Another novelty in this thesis is incorporation of MRI data into inversion algorithms for MRE. In other approaches to the inverse problem for MRE, like direct algebraic inversion or the least squares approach [14, 32], incorporation of knowledge available prior to the inversion is difficult. One of the advantages of Bayesian approaches is an easy implementation of the prior knowledge about the unknown that is available before the inversion and that remains unused in most of the other approaches to MRE [14], especially in connection with the parametrisation used in this thesis.

We will show how the use of prior knowledge and parametrisation considerably impacts the quality of inversion results.

The parametrisation can be freely designed due to the fact that Ensemble Kalman Inversion allows for advanced black box models [12] and can therefore be used to incorporate knowledge about the hydrogen density in the tissue into the prior. Usually, in application of MRE, MRI data which estimates the hydrogen density are available.

The parametrisation of the unknown used in this thesis, the two-level parametrisation, is relatively new in application with EKI and has been used in this context in [61] before. The first level of this parametrisation allows for discontinuities in the unknown, the second level of parametrisation adds spatial variability of the unknown. While in [61] the authors parametrise a real valued

unknown, in MRE with a viscoelastic model we are required to parametrise a complex valued unknown with a comparatively difficult geometry.

In Chapter 2, we recall some basic mathematical definitions and theorems used throughout this thesis. After that, in Chapter 3, we introduce the notion of elastography, the inverse problem arising in MRE, the models we use and give an overview of recent approaches to the inverse problem for MRE. In Chapter 4 we describe EKI, the inversion method we apply to MRE. Chapter 5 and Chapter 6 showcase EKI in various experimental setups using both the purely elastic and viscoelastic model. Finally, in Chapter 7 we give a conclusion and describe potential extensions to the work done in this thesis.

2

PRELIMINARY MATHEMATICAL CONCEPTS

In this chapter we introduce basic notions and theorems used throughout this thesis. We will cover basic Calculus and Linear Algebra in Section 2.1, Integration Theory in Section 2.3, Probability Theory in Section 2.4 and Partial Differential Equations in Section 2.5. The variable names used in this Chapter should only be understood in the context of this chapter and have no connection to variable names in Chapter 3 and onward.

2.1 DIFFERENTIAL OPERATORS

In this Section we repeat basic definitions of differential operators used throughout this thesis. Apart from the definition of the curl, all definitions are taken from [3] and adapted to our less general setting.

It is worth noting at this point that in this thesis, functions $v : \mathbb{R}^n \to \mathbb{C}^m$, $m, n \in \mathbb{N}$ will be treated as functions $\mathbb{R}^n \to \mathbb{R}^{2m}$ and we then consider

differentiability of v in the way we introduced it in this section only and never complex-differentiability.

Definition 1. *Let $v : \mathbb{R}^n \to \mathbb{R}$ be a function, $n \in \mathbb{N}$, and $x_0 \in \mathbb{R}^n$. We write ∂_k for the derivatives in the direction of the standard basis vectors e_k for $k = 1, \dots, n$ and call*

$$\partial_k v(x_0) := \lim_{t \to 0} \frac{v(x_0 + te_k) - v(x_0)}{t}, \quad \text{for } k = 1, \dots, n,$$

*the **partial derivative** with respect to x^k of v at x_0.*

*Let $\Omega \subset \mathbb{R}^n$ be an open subset and $v : \Omega \to \mathbb{R}$ be (totally) differentiable [3] at $x_0 \in \Omega$. We define the **gradient** of v in x_0 to be*

$$\nabla v(x_0) = (\partial_1 v(x_0), \dots, \partial_2 v(x_0)) \in \mathbb{R}^n.$$

We generalise the notion of gradients to higher dimensions and also introduce the divergence.

Definition 2. *Let $\Omega \subset \mathbb{R}^n$ be an open subset and $v : \Omega \to \mathbb{R}^n$ be (totally) differentiable at $x_0 \in \Omega$. We define the gradient (or **Jacobian**) of v in x_0 to be*

$$\nabla v(x_0) = \begin{bmatrix} \partial_1 v_1(x_0) & \cdots & \partial_n v_1(x_0) \\ \vdots & \ddots & \vdots \\ \partial_1 v_n(x_0) & \cdots & \partial_n v_n(x_0) \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

*We define the **divergence** of v in x_0 to be*

$$\nabla \cdot v(x_0) = \partial_1 v_1(x_0) + \cdots + \partial_n v_n(x_0) \in \mathbb{R}^n.$$

For $v : \Omega \to \mathbb{R}^{n\times n}$ we generalise the notion of divergence to matrix valued functions in the following way:

$$\nabla \cdot v(x_0) = \begin{bmatrix} \partial_1 v_{11}(x_0) + \cdots + \partial_n v_{1n}(x_0) \\ \vdots \\ \partial_1 v_{n1}(x_0) + \cdots + \partial_n v_{nn}(x_0) \end{bmatrix} \in \mathbb{R}^n.$$

Finally, we introduce the definition of the curl function.

Definition 3. *For $v : \Omega \to \mathbb{R}^2$ we define the **curl** to be*

$$\nabla \times v(x_0) = \partial_1 v_2(x_0) - \partial_2 v_1(x_0) \in \mathbb{R}.$$

We generalise the curl to higher dimensions and matrix valued functions. For the case $v : \Omega \to \mathbb{R}^4$ we define

$$\nabla \times v(x_0) = \begin{bmatrix} \partial_1 v_2(x_0) - \partial_2 v_1(x_0) \\ \partial_3 v_4(x_0) - \partial_4 v_3(x_0) \end{bmatrix} \in \mathbb{R}^2.$$

If $v : \Omega \to \mathbb{R}^{2\times 2}$, we define

$$\nabla \times v(x_0) = \begin{bmatrix} \partial_1 v_{12}(x_0) - \partial_2 v_{11}(x_0) \\ \partial_3 v_{22}(x_0) - \partial_4 v_{21}(x_0) \end{bmatrix} \in \mathbb{R}^2.$$

2.2 NORMS AND SPACES

We recall some definitions of norms and spaces in this section. We follow mostly [2] in the first part of the section and [3] for the definition of (Fréchet) differentiability. In this section, let $\mathbb{K}$ be alternatively $\mathbb{R}$ or $\mathbb{C}$.

Definition 4. *A **metric space** X is a set with a function $d : X \times X \to \mathbb{R}_{\geq 0}$ defined on, satisfying*

- $d(x, y) = 0$, *if and only if* $x = y$.
- $d(x, y) = d(y, x)$, $x, y \in X$ *(symmetry).*
- $d(x, y) \leq d(x, z) + d(z, y)$, $x, y, z \in X$ *(triangle inequality).*

A metric space X is called **complete** if every Cauchy sequence in X converges. A sequence (x_n) in X is **Cauchy** if, for each $\varepsilon > 0$, there is some $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon$ for all $m, n \in \mathbb{N}$ with $m, n > N$. A sequence (x_n) **converges** in X if each neighbourhood of some $a \in X$ contains almost all terms of the sequence.

Definition 5. *A **vector space** over $\mathbb{K}$ is a nonempty set V with an inner operation $+$ on V and an outer operation*

$$\mathbb{K} \times V \to V, \quad (\lambda, v) \mapsto \lambda \cdot v,$$

satisfying

- *V and the inner operation $+$ is an Abelian group, i.e.:*
  - *$u + (v + w) = (u + v) + w$ for $u, v, w \in V$ (Associativity).*
  - *$u + v = v + u$ for $u, v \in V$ (Commutativity).*
  - *There is $0 \in V$ such that $v + 0 = v$ (Neutral Element).*
  - *For every $v \in V$ there exists an $-v \in V$ such that $v + (-v) = 0$ (Additive Inverse).*
- *The distributive law holds:*

$$\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w, \quad (\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v, \quad \lambda, \mu \in \mathbb{K},\ v, w \in V.$$

- *$\lambda \cdot (\mu v) = (\lambda \mu) \cdot v$ and $1 \cdot v = v$ for $\lambda, \mu \in \mathbb{K}$, $v \in V$.*

At this point we add the definition of dual spaces.

Definition 6. *Let V be a vector space over $\mathbb{K}$. The **dual space** V^* is defined as the set of all linear maps $\varphi : V \to \mathbb{K}$.*

Definition 7. *Let V be a vector space over $\mathbb{K}$ and let $\|\cdot\| : V \to \mathbb{R}^+$ be a function satisfying:*

- *$\|x\| = 0 \iff x = 0$.*
- *$\|\lambda x\| = |\lambda| \|x\|$, for $x \in V$ and $\lambda \in \mathbb{K}$ (positive homogeneity).*

- $\|x+y\| \leq \|x\| + \|y\|$, *for* $x, y \in V$ *(triangle inequality).*

Vector space V *together with function* $\|\cdot\|$ *is called a* ***normed vector space****.*

The function $\|\cdot\|_2 : \mathbb{R}^n \to \mathbb{R}$, $n \in \mathbb{N}$ defined by

$$x \mapsto \|x\|_2 := \sqrt{x_1^2 + \cdots + x_n^2},$$

defines a norm and is called **Euclidean norm**.

The norm $\|\cdot\|$ of a normed vector space $(V, \|\cdot\|)$ induces a metric $d : V \times V \to \mathbb{R}_{\geq 0}$ by $d(x, y) := \|x - y\|$.

Definition 8. *Let* V *be a vector space over* $\mathbb{K}$ *and let* $(\cdot|\cdot) : V \times V \to \mathbb{K}$, $(x, y) \mapsto (x|y)$ *a function satisfying:*

- $(x|y) = \overline{(y|x)}$ *for* $x, y \in V$*, where* $\overline{z}$ *denotes the complex conjugate of* $z \in \mathbb{C}$*.*
- $(\lambda x + \mu y|z) = \lambda(x|z) + \mu(y|z)$*, for* $x, y, z \in V$*,* $\lambda, \mu \in \mathbb{K}$*.*
- $(x|x) \geq 0$*,* $x \in V$ *and* $(x|x) = 0 \iff x = 0$*.*

Vector space V *together with function* $(\cdot|\cdot)$ *is an* ***inner product space****.*

A complete inner product space is called a **Hilbert space**. A complete normed vector space is called a **Banach space**.

Let $A, B \in \mathbb{C}^{m \times n}$ complex valued matrices, $m, n \in \mathbb{N}$. The **Frobenius inner product** $A : B$ is defined as

$$A : B = \sum_{i,j} \overline{A_{ij}} B_{ij} = \operatorname{tr}(\overline{A^T} B),$$

where tr is the trace, i.e. $\mathrm{tr}(C) = \sum_{i=1}^{n} c_{ii}$ for a square matrix $C \in \mathbb{C}^{n\times n}$.

We will need derivatives in a more general setting than introduced in Section 2.1.

Definition 9. *Let $(E, \|\cdot\|)$ and $(F, \|\cdot\|)$ be Banach spaces over $\mathbb{C}$ and X be an open subset of E. The function $v : X \to F$ is called **(Fréchet) differentiable** in $x_o \in X$ if there is bounded linear operator $A_{x_0} : E \to F$ such that*

$$\lim_{x\to x_0} \frac{v(x) - v(x_0) - A_{x_0}(x - x_0)}{\|x - x_0\|} = 0.$$

$A_{x_0} : E \to F$ is then called the (Fréchet) derivative and denoted by $Dv(x_0)$.

*An operator A is **bounded** if bounded subsets of its domain are mapped to bounded subsets of its image $Im(A)$.*

2.3 INTEGRATION THEORY

In this section we repeat the basic notions and definitions from integration theory. All definitions are taken from [47] and adapted to our setting. The central notion in this section is the integral of a function over a measure space. We begin with the definition of a σ-algebra.

Definition 10. *Let $\mathcal{X}$ be a set and let $\mathcal{P}(\mathcal{X})$ denote its power set. A subset $\Sigma \subset \mathcal{P}(\mathcal{X})$ is called **σ-algebra** if*

1. *If S is an element of Σ then so is its complement $\Sigma \setminus S$.*

2. *$\mathcal{X}$ is an element of Σ.*

3. *If $S_1, S_2, S_3, \ldots$ are elements of Σ then so is their union $\bigcup_{i=1}^{\infty} S_i \in \Sigma$.*

Definition 11. *A **measure** is a mapping $\mu : \Sigma \to [0, \infty]$ from a σ-algebra Σ to the positive extended real number line such that*

1. *The empty set has measure zero $\mu(\emptyset) = 0$.*

2. *For a countable collection $\{S_i\}_{i=1}^{\infty}$ of pairwise disjoint sets, i.e. $S_i \cap S_j = \emptyset$ for $i \neq j$ the measure μ is countable additive, that is $\mu\left(\bigcup_{i=1}^{\infty} S_i\right) = \sum_{i=1}^{\infty} \mu(S_i)$*

*A **measure space** is a triplet $(\mathcal{X}, \Sigma, \mu)$ consisting of set $\mathcal{X}$ called the **sample space**, a σ-algebra Σ containing sets $S \subset \mathcal{X}$ called **events** and the measure μ.*

Let $f, g : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ functions on a measure space $(\mathcal{X}, \Sigma, \mu)$. We write $f \leq g$ if $f(x) \leq g(x)$ for $x \in \mathcal{X}$.

Let $f\mathcal{X} \to \mathbb{R}$ functions on a measure space $(\mathcal{X}, \Sigma, \mu)$. We call f measurable, if $f^{-1}(S)$ is an element of Σ for every open set $S \in \mathbb{R}$.

Let $(\mathcal{X}, \Sigma, \mu)$ be a measure space and $f : \mathcal{X} \to [0, \infty]$ be measurable, i.e. the preimage of any open interval in $[0, \infty]$ is measurable. The **integral** of f with respect to μ is defined by

$$\int f \, \mathrm{d}\mu := \sup\{I(g) : g \in \mathbb{E}^+, g \leq f\}.$$

where $\mathbb{E}^+$ is the vector space of simple positive functions, i.e.

$$\mathbb{E}^+ := \{f : \mathcal{X} \to \mathbb{R} | f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{S_i}, \text{mutually disjoint } S_i \in \Sigma, \alpha_i \in \mathbb{R}, \alpha_i \geq 0\},$$

and $I : \mathbb{E}^+ \to [0, \infty]$ the map defined by

$$I(f) = \sum_{i=1}^{m} \alpha_i \mu(S_i).$$

More general, let $f : \mathcal{X} \to \mathbb{R}$ measurable and the integral $\int |f| \, \mathrm{d}\mu < \infty$. We call f μ-**integrable** and define the **integral** of f with respect to μ to be

$$\int f(x)\mu(\mathrm{d}x) := \int f \, \mathrm{d}\mu := \int f^+ \, \mathrm{d}\mu - \int f^- \, \mathrm{d}\mu, \tag{2.3.1}$$

where f^+ and f^- are signed functions defined by

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) > 0 \\ 0 & \text{if otherwise} \end{cases}, \qquad f^-(x) = \begin{cases} f(x) & \text{if } f(x) < 0 \\ 0 & \text{if otherwise} \end{cases}.$$

For $S \in \Sigma$ we define

$$\int_S f \, \mathrm{d}\mu := \int (f \mathbb{1}_S) \, \mathrm{d}\mu.$$

Definition 12. *Let μ and ν be measures on $(\mathcal{X}, \Sigma)$, where $\mathcal{X}$ is a set and Σ is a σ-algebra on $\mathcal{X}$. Let $f : \mathcal{X} \to [0, \infty)$ be a measurable function such that*

$$\nu(S) = \int f \mathbb{1}_S \, \mathrm{d}\mu \quad \textit{for all } S \in \Sigma.$$

*In this case f is called **density** of ν with respect to μ and we call $f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ the **Radon-Nikodym derivative**.*

For example the normal distribution $\nu = \mathcal{N}_{0,1}$ has density $f(x) = \frac{1}{2\pi}\exp(-x^2/2)$ with respect to Lebesgue measure $\mu = \lambda$ on $\mathbb{R}$, which will be introduced in the next section.

We also need the definition of a special type of measure, the Dirac measure δ_x. Let $\mathcal{X}$ be a set with a $\sigma-$algebra Σ defined and $x \in \mathcal{X}$. We define the **Dirac measure** $\delta_x : \Sigma \to [0, \infty]$ as

$$\delta_x(S) = \begin{cases} 0, & x \notin S \\ \\ 1, & x \in S \end{cases}.$$

Definition 13. *Let $\Omega \subset \mathbb{R}^n$ be measurable and p an integer with $1 \leq p < \infty$. The Lebesgue space $L_p(\Omega)$ is defined as*

$$L_p(\Omega) = \left\{ f : \Omega \to \mathbb{C} | \left(\int_\Omega |f(x)|^p \, \mathrm{d}x \right)^{1/p} < \infty \right\}.$$

The mapping $< \cdot, \cdot >: L_2(\Omega) \times L_2(\Omega) \to \mathbb{C}$ with

$$< f, g >= \int_\Omega f(x)\overline{g(x)} \, \mathrm{d}x,$$

defines an inner product on $L_2(\Omega)$.

Furthermore, we define the space of **locally integrable functions** on an open subset $\Omega \subset \mathbb{R}^n$ as

$$L_{1,\mathrm{Loc}}(\Omega) = \{ f : \Omega \to \mathbb{C} |\ f|_K \in L_1(K) \text{ for every compact } K \subset \Omega \}.$$

Finally, we define the Lebesgue space

$$L_\infty = \left\{ f : \Omega \to \mathbb{C} | \text{ess sup}_{x \in \Omega} |f(x)| < \infty \right\},$$

where ess $\sup_{x\in\Omega}|f(x)|$ denotes infinum of all $K \in \mathbb{R}$ such that $|f(x)| < K$ for almost every $x \in \Omega$, i.e. for all $x \in \Omega \setminus N$ where $N \subset \Omega$ is a set with $\mu(N) = 0$.

2.4 PROBABILITY

In this section we go through necessary definitions and theorems from the probability theory. We follow the introduction from [47] for all the definitions and [48] for Bayes Theorem.

Definition 14. *Let $(\mathcal{X}, \Sigma, \mu)$ be a measure space. Measures that satisfy $\mu(\mathcal{X}) = 1$ are called **probability measures** and in this case the triplet $(\mathcal{X}, \Sigma, \mu)$ is called **probability space**.*

Definition 15. *Let $(\mathcal{X}, \Sigma, \mu)$ be a probability space. A **real-valued random variable** $\mathbf{V}$ is a mapping*

$$\mathbf{V} : \mathcal{X} \to \mathbb{R}^n,$$

*that is **measurable**.*

A real-valued random variable $\mathbf{V}$ generates a probability measure $\mu_{\mathbf{V}} : \mathcal{B} \to \mathbb{R}$ through

$$\mu_{\mathbf{V}}(B) = \mu(\mathbf{V}^{-1}(B)),$$

and defines together with the sample space $\mathbb{R}^n$ and the σ-algebra $\mathcal{B}$ a probability space. Here, $\mathcal{B}$ is the smallest σ-algebra containing the open sets of $\mathbb{R}^n$, called the **Borel σ-algebra** and λ is the **Lebesgue measure** assigning to a set its volume [47]. For a cube $[a,b]^n$, $a < b$ the Lebesgue measure is given by $\lambda([a,b]^n) = (b-a)^n$. The measure $\mu_{\mathbf{V}}$ is called the **probability distribution** of $\mathbf{V}$.

Let $B \in \mathcal{B}$. We use the common notation $\{\mathbf{V} \in B\} := \mathbf{V}^{-1}(B)$, $\{\mathbf{V} \geq 0\} := \mathbf{V}^{-1}([0,\infty])$ and $\{\mathbf{V} \leq b\} := \mathbf{V}^{-1}([-\infty,b])$ and also $\mu(\mathbf{V} \in B) := \mu(\mathbf{V}^{-1}(B)$, $\mu(\mathbf{V} \geq 0) := \mu(\mathbf{V}^{-1}([0,\infty])$ and $\mu(\mathbf{V} \leq b) := \mu(\mathbf{V}^{-1}([-\infty,b])$.

Definition 16. *For a real-valued random variable $\mathbf{V}$, the map $F_{\mathbf{V}} : \mathbb{R}^n \to [0,1]$, $x \mapsto \mu(\mathbf{V} \leq x)$ is called the **distribution function of** $\mathbf{V}$.*

For any distribution function F, there exists a real random variable $\mathbf{V}$ with $F_{\mathbf{V}} = F$ [47]. Both distribution functions and - if existent - density functions are important functions that describe a random variable.

Definition 17. *Let $m \in \mathbb{R}^d$ and Σ_p be a positive definite symmetric $d \times d$ matrix. Let $\mathbf{V}$ be an $\mathbb{R}^d$-valued random variable such that*

$$\mu(\mathbf{V} \leq x) = det(2\pi\Sigma_p)^{-1/2} \int_{(-\infty,x]^d} \exp\left(-1/2\langle k-m, \Sigma_p^{-1}(k-m)\rangle\right) \lambda^d(\mathrm{d}k),$$

for $x \in \mathbb{R}^d$ where λ^d denotes the d-dimensional Lebesgue measure and $(-\infty,x]^d = \prod_{i=1}^d (-\infty,x]$ the d-dimensional interval. The distribution $\mathcal{N}(m,\Sigma_p) := \mu_{\mathbf{V}}$ is called

Gaussian normal distribution *with mean m and Covariance $\Sigma_p \in \mathbb{R}^{d\times d}$. The density of $\mathbf{V}$ with respect to Lebesgue measure is*

$$f(k) = det(2\pi\Sigma_p)^{-1/2} \exp\left(-1/2\langle k-m, \Sigma_p^{-1}(k-m)\rangle\right).$$

Definition 18. *Let $(\mathcal{X}, \Sigma, \mu)$ be a probability space and $\mathbf{V}$ and $\mathbf{W}$ be n- and m-dimensional real random variables. We define*

$$F_J : \mathbb{R}^{n+m} \to [0,1], \quad (v,w) \mapsto \mu\left(\mathbf{V} \leq v \text{ and } \mathbf{W} \leq w\right).$$

*F_J is called the **joint distribution function** of $\mathbf{V}$ and $\mathbf{W}$ the probability measure $\mu_{\mathbf{V},\mathbf{W}} : \mathbf{B} \to \mathbb{R}$ on $\mathbb{R}^{n+m}$ is called **joint distribution** of $\mathbf{V}$ and $\mathbf{W}$. The two random variables $\mathbf{V}$ and $\mathbf{W}$ are **independent** if*

$$F_J(x) = F_{\mathbf{V}}(x) \cdot F_{\mathbf{W}}(x).$$

*If F_J has a density π_J with respect to Lebesgue measure we call this density the **joint density**. If, in addition, $\mathbf{V}$ and $\mathbf{W}$ are independent and have densities $\pi_{\mathbf{V}}$ and $\pi_{\mathbf{W}}$ with respect to Lebesgue measure, then $\pi_J = \pi_{\mathbf{V}} \cdot \pi_{\mathbf{W}}$.*

Let $(\mathcal{X}, \Sigma, \mu)$ be a measure space. We denote for any $p \geq 1$

$$\mathcal{L}_p(\mu) := \{f : \mathcal{X} \to \mathbb{R} | f \text{ measurable and } \left(\int |f|^p \, \mathrm{d}\mu\right)^{1/p} < \infty\},$$

which is the Lebesgue space $L_p(\Omega)$ if $\mu = \lambda$ is the Lebesgue measure.

Definition 19. *Let $(\mathcal{X}, \Sigma, \mu)$ be a probability space and $\mathbf{V} \in \mathcal{L}_1(\mu)$ be a real-valued random variable. The **mean** or **expected value** of $\mathbf{V}$ is defined as*

$$\mathbf{E}(\mathbf{V}) = \int_{\mathcal{X}} \mathbf{V}(x)\, \mathrm{d}\mu.$$

*If $\mathbf{E}(V) = 0$, then V is called **centred**.*

Definition 20. *Let $\mathbf{V} \in \mathcal{L}_2(\mu)$, then $\mathbf{V}$ is called **square integrable** and*

$$Var(\mathbf{V}) := \mathbf{E}(\mathbf{V}^2) - \mathbf{E}(\mathbf{V})^2,$$

*is the **variance** of $\mathbf{V}$. The quantity $\sigma := \sqrt{Var(\mathbf{V})}$ is called the **standard deviation** of $\mathbf{V}$.*

Definition 21. *Let $\mathbf{V}, \mathbf{W} \in \mathcal{L}_2(\mu)$, then we call*

$$Cov(\mathbf{V}, \mathbf{W}) := \mathbf{E}((\mathbf{V} - \mathbf{E}(\mathbf{V}))(\mathbf{W} - \mathbf{E}(\mathbf{W}))).$$

*the **covariance**.*

Let $(\mathcal{X}, \Sigma, \mu)$ be a probability space. Let U and V be a pair of jointly distributed real-valued $n-$ and $m-$dimensional random variables and assume they have a density $\rho(U, V)$ with respect to Lebesgue measure. Then **Bayes formula** [48] describes the conditional density $\rho(U|V = v) = \rho(U|V) : \mathbb{R}^n \to [0, \infty)$ of U given $V = v$ in terms of the conditional density $\rho(V|U = u) =$

$\rho(V|U)$, an unconditioned density $\rho(U)$ and a normalisation that only depends on V, in the following way:

$$\rho(U|V) = \frac{\rho(V|U)\rho(U)}{\int_{\mathbb{R}^n} \rho(V|U)\rho(U)\lambda(\mathrm{d}u)}.$$

Definition 22. *A **stochastic process/random field** $f = \{f(x), x \in \mathcal{X}\}$ is a collection of random variables indexed by elements of a parameter set $\mathcal{X}$.*

If the parameter space $\mathcal{X}$ of a stochastic process is the Cartesian plane or a higher-dimensional Euclidean space the term random field is commonly used.

Definition 23. *A **Gaussian random field** is a random field $f = \{f(x), x \in \mathcal{X}\}$ with the property that for each $x_1, \ldots, x_d$ with some $d \in \mathbb{N}$, the vector $(f(x_1), \ldots, f(x_d))^T \sim \mathcal{N}_d(\mu, \Sigma)$ is Gaussian distributed with some mean μ and covariance Σ.*

If the distribution of a Gaussian random field $f = \{f(x), x \in \mathcal{X}\}$ is the normal distribution and $f(x_1)$ and $f(x_2)$ are independent for $x_1 \neq x_2$ the random field is called white Gaussian noise.

2.5 PARTIAL DIFFERENTIAL EQUATIONS

In this section we recall the relevant definitions and theorems from the partial differential equations theory and cover the definitions of Sobolev spaces, Lips-

chitz domains, weak and classical solutions and state the trace theorem. We follow [6] in this section.

Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{N}^d$ be a multi-index. We define the following differential operator

$$D^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_d}\right)^{\alpha_d},$$

and define $|\alpha| = \alpha_1 + \alpha_2 + \ldots \alpha_d$.

Definition 24. *Let $\Omega \subset \mathbb{R}^n$ be an open subset. We denote the **space of smooth functions with compact support** by*

$$C_0^\infty(\Omega) = \{f \in C^0(\Omega) : D^\alpha f \in C^0 \text{ for } \alpha \in \mathbb{N} \text{ and } supp(f) \text{ is compact}\},$$

where $C^0(\Omega)$ denotes the set of continuous functions on Ω and $supp(f) = \{x \in \Omega : |f(x)| > 0\}$ the support.

Definition 25. *Let $\Omega \subset \mathbb{R}^n$ be a domain and $m \geq 0$ be an integer. The m-th order **Sobolev space** in $L_2(\Omega)$ is defined by*

$$H^m(\Omega) = \{f \in L_2(\Omega) : D^\alpha f \in L_2(\Omega) \text{ for all multi-indices } \alpha \text{ such that } |\alpha| \leq m\}.$$

Definition 26. *Let $\Omega \subset \mathbb{R}^n$ and $u, v \in L_{1,Loc}(\Omega)$ and α some multi-index. We say v is the α^{th}-**weak derivative** of u if*

$$\int_U u D^\alpha \phi \, \mathrm{d}x = (-1)^{|\alpha|} \int_U v\phi \, \mathrm{d}x, \quad \text{for all } \phi \in C_0^\infty(\Omega).$$

Definition 27. *Let $\Omega \subset \mathbb{R}^n$ be a set. We call $\partial\Omega = \overline{\Omega} \setminus int(\Omega)$ the **boundary** of the set, where $\overline{\Omega}$ denotes the closure and $int(\Omega)$ the interior.*

Definition 28. *Let $\Omega \subset \mathbb{R}^n$ be a domain, i.e. a connected and open subset in $\mathbb{R}^n$. For $R > 0$ and $x_0 \in \mathbb{R}^n$ let us define $B_R(x_0) := \{x \in \mathbb{R}^n | \|x - x_0\| < r\}$. We say Ω is a **Lipschitz domain** or has a **Lipschitz boundary** if there exists sets $\Omega_j \subset \mathbb{R}^n$, $j = 1, \ldots, N$ with $N \in \mathbb{N} \cup \{\infty\}$, such that*

- *$\overline{\Omega_j} \subset \mathbb{R}^n$, $\partial\Omega \subset \bigcup_j \Omega_j$, and only finitely many of the sets intersect $B_R(0)$ for all $R > 0$.*
- *There are Lipschitz continuous functions $\phi_j : \Omega_j \to B_1(0)$ that are one-to-one and ϕ_j^{-1} is Lipschitz as well, i.e. there exists $C > 0$ such that $\|\phi_j^{-1}(x) - \phi_j^{-1}(y)\| \leq C\|x - y\|$ for all $x, y \in \Omega$.*
- *$\phi_j(\Omega_j \cap \Omega) = B_1(0) \cap \mathbb{R}^n_+$ and $\phi_j(\Omega_j \cap \partial\Omega) = B_1(0) \cap \mathbb{R}^n_+ \cap \partial\mathbb{R}^n_+$.*

The intuition behind this definition is that a Lipschitz boundary is locally the graph of a Lipschitz continuous function.

Let $\Omega \subset \mathbb{R}^n$ be some bounded Lipschitz domain. A **second order elliptic partial differential equation** is an equation of the form

$$-\nabla \cdot (a\nabla v + bv) + cv = f \quad x \in \Omega, \tag{2.5.1}$$

where $v : \Omega \to \mathbb{C}$ is an unknown function, a is a $n \times n$ uniformly positive definite matrix-valued function, b is a n-vector-valued function, and c and f are

functions with $c \geq 0$. A matrix function $a : \overline{\Omega} \to \mathbb{R}^{n\times n}$ is **uniformly positive definite** if there is some $a_* > 0$ such that for almost every $x \in \overline{\Omega}$,

$$\xi^T a(x)\xi \geq a_*|\xi|^2, \quad \forall \xi \in \mathbb{R}^n.$$

A description of a dynamic system requires the quantification of the external interaction with the world through its boundary $\partial\Omega$ in form of boundary conditions. More specifically let $\partial\Omega = (\partial\Omega)_D \cup (\partial\Omega)_N$ be a partition into **Dirichlet-** and **Neumann-**boundary of the boundary with $(\partial\Omega)_D \cap (\partial\Omega)_N = \emptyset$. The boundary conditions are given by

$$v = v_D, \qquad \text{on } (\partial\Omega)_D, \tag{2.5.2}$$

$$(a\nabla v + bv) \cdot \mathbf{n} = g_N \qquad \text{on } (\partial\Omega)_N, \tag{2.5.3}$$

where u_D and g_N are functions and $\mathbf{n}$ is the outer unit normal vector on Ω, i.e. $\|\mathbf{n}\| = 1$ and for a tangent vector w at a point $x_B \in \partial\Omega$ we have $\mathbf{n} \cdot w = 0$.

The problem of finding functions u that satisfy (2.5.1)-(2.5.3) is called a **boundary value problem (BVP)**. A **classical solution** to this BVP is a v that satisfies (2.5.1)-(2.5.3) in every $x \in \Omega$ which requires v to be twice differentiable. Also, we require v to be continuous in $\overline{\Omega}$, so classical solutions are $v \in C^2(\Omega) \cap C^0(\overline{\Omega})$. This then requires $f, c \in C^0(\Omega)$, $a \in (C^1(\Omega))^{n\times n}$, $b \in (C^1(\Omega))^n$ and $v_D, g_N \in C^0((\partial\Omega)_D)$.

These requirements are too strong in many applications. Therefore, we usually relax the requirements to $v \in H^1(\Omega)$ only which then require f, c, a, v_D

and g_N to be not continuous or continuously differentiable anymore. Also, we will see in Chapter 3.3 that this enables us to find a variational problem called the weak form that is equivalent to finding a solution $v \in H^1(\Omega)$ for the BVP (2.5.1)-(2.5.3). We do not derive the weak form for general elliptic PDEs here but we will derive it for the specific PDE we use in Section 3.3. The solution $v \in H^1(\Omega)$ is called a **weak solution** of (2.5.1).

Weak solutions $v \in H^1(\Omega)$ by definition only represent an equivalence class of functions. The boundary $\partial\Omega$ is a null set with respect to Lebesgue measure and therefore values of weak solutions v can be arbitrary. So, it is not apparent, in which way $v \in H^1(\Omega)$ should satisfy $v = v_D$ on $(\partial\Omega)_D$. We define an operator λ_0 that defines values on the boundary for a weak solution v, i.e. $\lambda_0 v = v_D$. The **trace theorem** states under which conditions this operator exists and extends the classical trace, i.e. how this operator extends $v \in H^1(\Omega)$ to the boundary of the domain. In order to state this theorem we need the following definition of fractional Sobolev spaces:

Definition 29. *Let $s \in \mathbb{R}$. We define*

$$H^s(\mathbb{R}^n) = \{u \in \mathcal{S}' : \Lambda^s u \in L_2(\mathbb{R}^n)\},$$

where $\mathcal{S}'(\mathbb{R}^n)$ is the dual space of $\mathcal{S}(\mathbb{R}^n)$ which is defined to be

$$\mathcal{S}(\mathbb{R}^n) = \{\varphi \in C^\infty(\mathbb{R}^n) : \sup_{x\in\mathbb{R}^n} |x^\alpha D^\beta \varphi(x)| < \infty \text{ for all multi-indices } \alpha \text{ and } \beta\}.$$

and Λ^s *is the* ***Bessel potential*** *of order* $s \in \mathbb{R}$*, i.e.*

$$(\Lambda^s u)^{(\xi)} = (1 + |\xi|^2)^{s/2} \hat{u}(\xi),$$

where $\hat{u}$ *is the Fourier transform of* u*, i.e.*

$$\hat{u}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} u(x) \exp(-x \cdot \xi) \, \mathrm{d}x.$$

Theorem 1. *Let* $\Omega \subset \mathbb{R}^d$ *be a Lipschitz domain. The trace operator* $\lambda_0 : C^0(\overline{\Omega}) \to C^0(\partial\Omega)$ *defined by* $(\lambda_0 u)(x) = u(x)$ *for* $x \in \partial\Omega$*, extends to a bounded linear map*

$$\lambda_0 : H^s(\Omega) \to H^{s-1/2}(\partial\Omega),$$

for any $s \geq 1$*, i.e. for all* $u \in H^s(\Omega) \cap C^0(\overline{\Omega})$ *we have* $(\lambda_0 u)(x) = u(x)$*. Here,* $H^{s-1/2}$ *or* H^s *can be fractional Sobolev spaces which was defined above.*

We will need these definitions and theorems in the description of our model in the following chapter.

3

PROBLEM DESCRIPTION

In this chapter we describe the idea of elastography, we introduce the model that is used and the inverse problem that arises in elastography. Central parts of this Chapter is Section 3.2 where we mathematically describe the model we use in MRE and Section 3.1 gives context and background information about MRE. Finally, in Section 3.5 we give a brief overview of popular methods that approximate solutions of the inverse problem arising in MRE.

3.1 ELASTOGRAPHY

In this section we introduce the concept of elastography. In Section 3.1.1 we describe the basic idea and the context in which elastography is usually used. Section 3.1.2 describes the fundamental idea of magnetic resonance elastography. We will only give a short qualitative description of this imaging modality as we will use MRI data only to provide a mask for one of our algorithms in Chapter 6. After that, we outline in Section 3.1.3 the acquisition of our displacement

data using a standard MRI machine. These displacements will be the input for our algorithms later. Finally, in Section 3.1.4 we deduce and explain the equations of motion we deploy in our approach to elastography.

3.1.1 *Introduction*

The idea of Elastography ([51], [14]) is to perturb the tissue using mechanical vibration and infer mechanical properties from the measured displacements. It follows the same principles as manual palpation which is used in medicine to determine location, size, shape and firmness of an object in the body (Figure 3.1.1). The difference is that MRE is using a pneumatic vibrator in order to trigger displacements of the tissue and a doctor is pushing his fingers into the tissue. Instead of measuring the displacements by using the sense of touch, MRE is using Magnetic Resonance Imaging (MRI) as imaging modality (Figure 3.1.1). In some applications of elastography, ultrasound is used as imaging modality [52] which is cheaper and easier to implement. Ultrasound imaging of some parts of the body, e.g. the brain, can be difficult whereas MR imaging can be applied to the whole body. Also, MR elastograms are operator independent whereas ultrasound elastography depends on the skill of

the person operating the ultrasound device. These are the reasons why in brain and liver elastography, MRE is preferred over ultrasound elastography [77].

The displacement data provided by MRI can be then used in order to infer mechanical properties of the tissue by either using a continuum mechanical model or a simplified model that relates deformations and mechanical properties.

The result of an MRE scan is an elastogram, i.e. a 2D or 3D map of mechanical properties of the tissue.

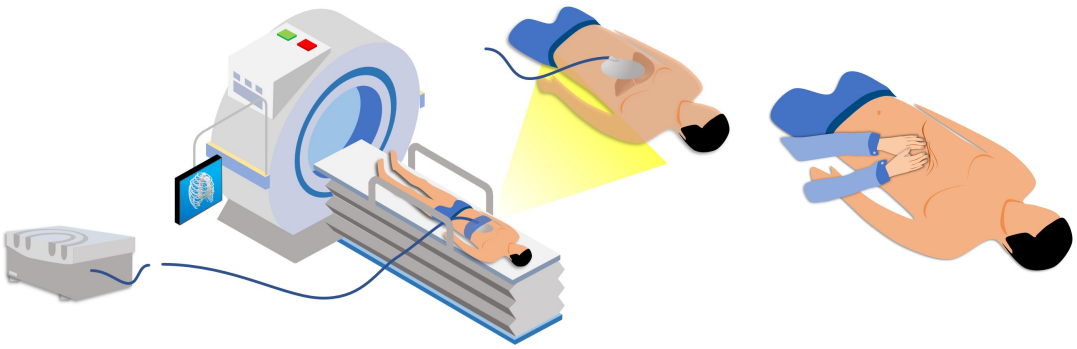


Figure 3.1.1: **Left:** Magnetic Resonance Elastography **Right:** Palpation

The big advantage of MRE over palpation is that one can obtain quantitative and objective measures with MRE as opposed to subjective and qualitative assessment from palpation. Furthermore, MRE allows exploration of sites deep inside the body which are impossible for palpation, such as the brain.

MRE can provide high resolution stiffness images of soft tissue areas and is therefore diagnostically useful in imaging of diseases that affect the tissues stiffness like cancer, fibrosis and inflammation [78]. Elastic properties of tumour change at an early state of the disease [90] and therefore, MRE can be used to detect and locate early cancer tissue.

MRE is a well-established technique or is being actively researched for application to the liver [90, 14], brain [84], breast [64, 77], skeletal muscle [15], heart [17], lung [27], kidney [75] cartilage [54] and prostate [46] tissue. In this thesis we will focus mainly on brain, liver and kidney tissues.

3.1.2 *Magnetic Resonance Imaging*

Magnetic Resonance Imaging makes use of the nuclear magnetic resonance of the hydrogen atom [63]. The nucleus of the hydrogen atom has an intrinsic angular momentum called spin. Moreover, the nucleus of the hydrogen atom is made up of one proton carrying a positive electric charge and the spin makes it a rotating electric charge which induces a magnetic field with a magnetic dipole moment. In the presence of an external magnetic field with magnetic field strength B_0, the magnetic dipole moment of the nucleus can align with or against the field and interacts like a bar magnet with the field.

Radio waves (electromagnetic energy waves) of the appropriate frequency induce the transition from being aligned with the field to being aligned against the field. The radiowave frequency ω_L that induces the transition is called **Larmor frequency** and is proportional to the magnetic field strength, i.e.

$$\omega_L = \gamma_G \cdot \mathcal{B}_0, \tag{3.1.1}$$

where γ_G is a (hydrogen-specific) constant called gyromagnetic ratio.

After this transition of the proton of the nucleus to a higher energy state where it is aligned against the magnetic field, it returns - after a few milliseconds [21] - to the lower energy state where it is aligned with the field. During this transition the nucleus emits electromagnetic radiation at the Larmor frequency ω_L of the current magnetic field. So, if the magnetic field $\mathcal{B}_0$ is not changed after transition of the proton of the nucleus, Eq. (3.1.1) holds for the radio waves emitted. If the magnetic field is changed after transition the radio waves resonate at frequency ω_R given by

$$\omega_R = \gamma_G \cdot \mathcal{B}, \tag{3.1.2}$$

where $\mathcal{B}$ is the magnetic field.

MRI uses this resonance effect in order to approximate the local hydrogen density. More specifically, on top of the (spatially) constant magnetic field B_0 one applies additional magnetic fields that vary locally, called gradients.

Usually, three gradients are applied: The slice selecting gradient, the phase encoding gradient and the frequency encoding gradient. These gradients are orthogonal to each other and the variation of field strength is linear. Furthermore, these gradients are applied in the following order from the time the transition inducing radio waves are sent out by the MRI machine until the time the receiver coils measure the incoming radiosignal emitted by the nuclei.

The slice selecting one is applied at the same time as the radio waves are sent out by the MRI machine. As a consequence, only the hydrogen nuclei in one slice of the object of interest resonate. While sampling the incoming radio signal another gradient orthogonal to the slice selecting one is applied. This is the frequency encoding gradient. Subsequently, due to (3.1.2) the frequency of the already resonating nuclei in the selected slice is changed depending on their location along the direction of the frequency encoding gradient. In other words, due to the frequency encoding gradient, there are lines in the selected slice, orthogonal to the frequency encoding gradient and nuclei along each of these lines emit radio waves at the same frequency and after Fourier transform of the incoming radio signal at a receiver coil, different frequencies can be assigned to locations along these lines. The frequency encoding gradient introduces equifrequent lines in the slice selected by the slice selecting gradient.

The third gradient applied is the phase encoding gradient. This gradient is turned on and off between the slice selecting and the frequency encoding gradient. As this gradient is orthogonal to the other two gradients the emitted radio waves from one equifrequent line have different phases when measured in the receiver coil.

This sequence of sending out radio waves while applying a slice selecting gradient, then turning on and off a phase-encoding gradient and measuring the incoming radio signal while applying a frequency encoding gradient is repeated several times with different slopes in the linear phase encoding gradients. By comparing the phase offsets induced by different (linear) phase-encoding gradients in a certain location, a nucleus can be located along the direction of the phase-encoding gradient. Big differences in phase-offsets correspond to a bigger distance to the isocentre of the phase-encoding gradient, i.e. the location where the phase encoding gradient is zero. A small difference corresponds to a small distance to the isocentre. The amplitude of the radio signal in a location corresponds to the density of emitting radiation.

By application of this sequence of radiowave pulses and time-dependant gradients, the density of hydrogen nuclei in a certain location in the object of interest can be estimated. The sequence of gradients and radiowave pulses in the description above only describes the basic idea of MRI. Recent MRI devices

use more advanced sequences of radiowave pulses and gradients in order to achieve images faster and with higher accuracy (Spin Echo Pulse, Gradient Echo Pulse, Fast Gradient Echo Pulse, Fast Spin Echo Pulse [89]). The puls-sequence diagram of a spin echo puls sequence is shown in Figure 3.1.2.

In order to create a map of hydrogen density, the incoming electromagnetic radiation at receive coils is described using a model. More specifically, measuring the incoming electromagnetic radiation $\mathfrak{y} \in \mathbb{R}^n$, $n \in \mathbb{N}$, at the boundary of the domain of interest leads to an inverse problem [21]:

$$\mathfrak{y} = \mathcal{A}\mathfrak{f} + \varepsilon_N,$$

where $\varepsilon_N \in \mathbb{R}^n$ is (usually Gaussian) noise and $\mathcal{A}$ is a known linear operator comprising the forward model for MR signals, $\mathfrak{f} \in \mathbb{R}^m$, $m \in \mathbb{N}$, is the unknown vector of magnetisation immediately after excitation at locations in the body. The magnitude of this magnetisation is proportional to the product of local strength of the magnetic field and the local density of Hydrogen atoms and therefore, $\mathfrak{f}$ can be used to obtain images of the hydrogen density distribution in the body [21].

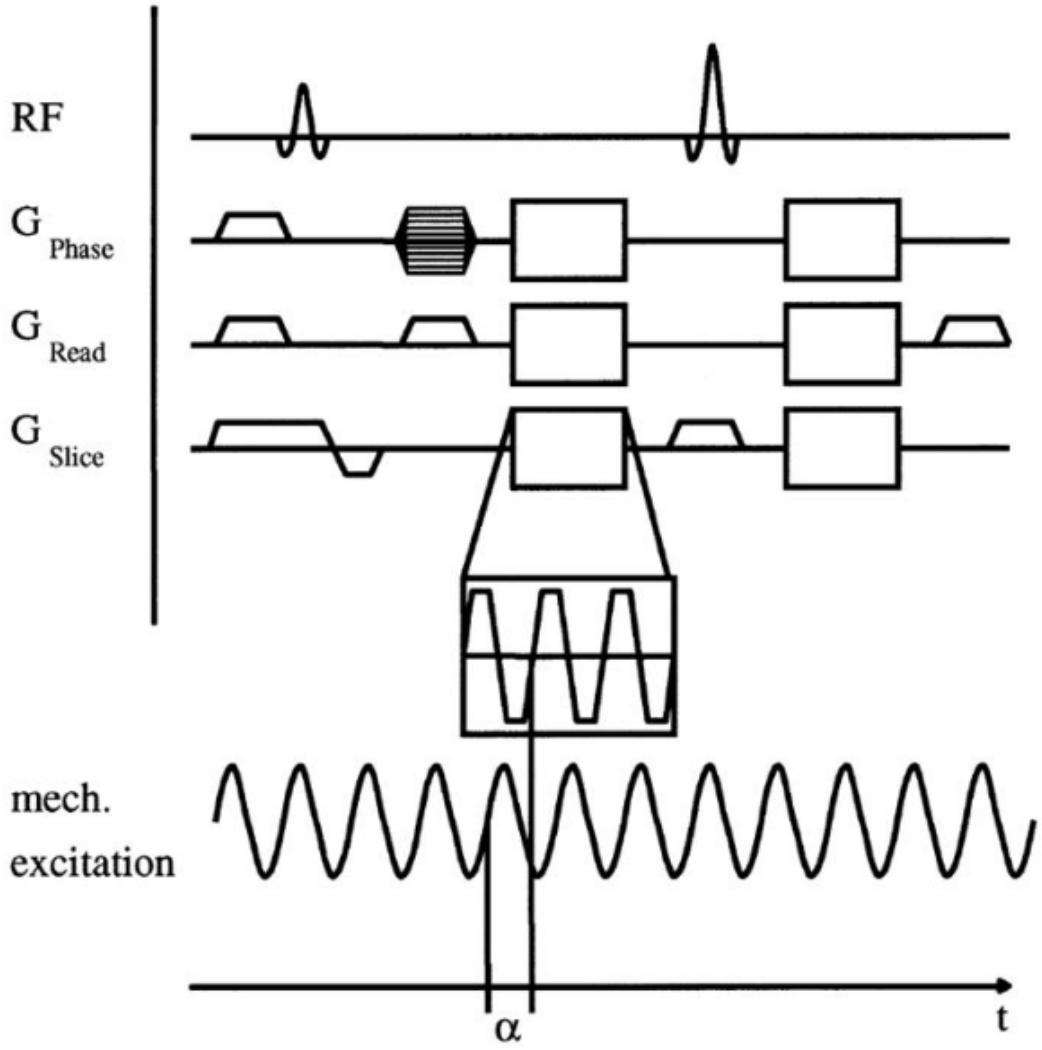


Figure 3.1.2: MRE pulse sequence diagram. RF = radiofrequency, GPhase = phase-encoding gradient, GRead = readout gradient or frequency-encoding gradient, Gslice = slice selective gradient. In a spin echo sequence motion sensitizing gradients (MSGs) are inserted before and after the second RF impulse. The phase offset between MSG and mechanical excitation is in this picture α (in this thesis the phase shift is denoted by β). (Reprinted with permission from U. Hamhaber, F. Grieshaber, J. Nagel, U. Klose, "Comparison of quantitative shear wave MR-elastography with mechanical compressiontests", Magn Reson Med. 49 no. 1 (2003):71-7.)

3.1.3 *Displacement Data from MRI*

In this subsection we give a description of the acquisition of displacement data of the tissue in MRE using a standard MRI machine. On top of the three gradients applied in standard MRI pulse sequences a slowly oscillating fourth gradient is applied. We start this subsection with the quantitative description of (general) gradients.

Let $G(t)$ be a magnetic field gradient that only modifies the x_1-component of a magnetic field $\mathcal{B}$, i.e.

$$G(t) = \begin{pmatrix} G_{x_1}(t) \\ G_{x_2}(t) \\ G_{x_3}(t) \end{pmatrix} = \begin{pmatrix} \frac{\partial \mathcal{B}_{x_1}(t)}{\partial x_1} \\ \frac{\partial \mathcal{B}_{x_1}(t)}{\partial x_2} \\ \frac{\partial \mathcal{B}_{x_1}(t)}{\partial x_3} \end{pmatrix}. \tag{3.1.3}$$

This magnetic field gradient is added to a constant magnetic field $\mathcal{B}_0$ which results in a total magnetic field at position $x(t) = (x_1(t), x_2(t), x_3(t))$ given by

$$\mathcal{B}(x,t) = (\mathcal{B}_0 + G(t)x(t)).$$

Let the position of the nucleus be $x(t) = x_0 + v_t(t)$ with x_0 being the position at rest and v_t the displacement. Due to (3.1.2), a resonating nucleus in magnetic field $\mathcal{B}$ is emitting radio waves of frequency

$$\omega_R(x,t) = \gamma_G \cdot (\mathcal{B}_0 + G(t) \cdot x(t)) = \omega_0 + \gamma_G G(t)x_0 + \gamma_G G(t)v_t(t),$$

where ω_0 is the frequency of radio waves emitted by the nucleus in magnetic field $\mathcal{B}_0$ defined in (3.1.1). Let

$$\omega_1(t) = \omega_0 + \gamma_G G(t) x_0,$$

the frequency emitted by the nucleus at the position at rest when the gradient G is applied.

The amount of accumulated phase of the nucleus resonating at frequency $\omega_R(x,t)$ due to displacement v_t relative to the frequency $\omega_1(t)$ at time t is given by [65]

$$\theta(x,t) = \int_{t_0}^{t} (\omega_R(x,t') - \omega_1(t'))\, \mathrm{d}t' = \gamma_G \int_{t_0}^{t} G(t') v_t(x,t') dt'. \tag{3.1.4}$$

Let us assume that a harmonic boundary excitation $\hat{v}(x)\cos(\omega t)$, $\omega \in \mathbb{R}$ generates a wave field in the object MRE is applied to and that the generated wave field is in stationary state after time t_0 and can be described by $v(x) = v_r(x) - \mathbf{j} v_c(x)$ for functions $v_r, v_c : \mathbb{R}^2 \to \mathbb{R}$ and complex quantity $\mathbf{j} := \sqrt{-1}$. We will later explain how these assumptions stem from the model we are using for MRE. The time-dependant wave field after time t is described by [14]

$$\begin{aligned} v_t(x,t) &= \mathrm{Re}\left[v(x)\exp(-\mathbf{j}\omega t)\right] \\ &= \mathrm{Re}\left[(v_r(x) - \mathbf{j} v_c(x))\left(\cos(\omega t) + \mathbf{j}\sin(\omega t)\right)\right] \\ &= v_r(x)\cos(\omega t) + v_c(x)\sin(\omega t). \end{aligned}$$

In MRE, we apply a motion-sensitizing gradient in the chosen pulse sequence of MRI (see Figure 3.1.2), i.e. a gradient of the form

$$G(t) = \mathbf{G_0} \cos(\omega t - \beta), \tag{3.1.5}$$

where ω is the angular frequency of the motion-sensitizing gradient which is chosen to be the same as that of the harmonic oscillation in the tissue, β is a phase shift between the motion-sensitizing gradient and the external excitation. This gradient is included into the normal pulse sequence of the MRI machine after the first radio frequency pulse. If two or more radio frequencies pulses are used in a pulse sequence the motion sensitizing gradient is added after the first and before the second radio frequency pulse [63].

Choosing the time $t = t_0 + 2\pi N_g/\omega$ and using trigonometric product-to-sum identities result in a phase shift $\theta_\beta(x)$:

$$\theta_\beta(x) = \gamma_G \int_{t_0}^{t_0+2\pi N_g/\omega} G(t') \cdot v(t')\, \mathrm{d}t' \tag{3.1.6}$$

$$= \gamma_G \int_{t_0}^{t_0+2\pi N_g/\omega} \mathbf{G_0} \cos(\omega t' - \beta) \left(v_r(x)\cos(\omega t') + v_c(x)\sin(\omega t')\right) \mathrm{d}t' \tag{3.1.7}$$

$$= \frac{\pi N_g \gamma_G}{\omega} \left[(v_r \cdot \mathbf{G_0})\cos\beta + (v_c \cdot \mathbf{G_0})\sin\beta\right]. \tag{3.1.8}$$

Here we denote by N_g the number of gradient cycles. By using motion-sensitizing gradients $\mathbf{G_0} = (G_0, 0, 0)$, $\mathbf{G_0} = (0, G_0, 0)$ or $\mathbf{G_0} = (0, 0, G_0)$, we

can obtain the phase shifts $\theta_\beta^{(i)}$ in each coordinate $i = 1, 2, 3$ in MRE measurements given by

$$\frac{\pi N_g \gamma_G G_0}{\omega}(v_{r_i}(x)\cos\beta + v_{c_i}(x)\sin\beta) = \theta_\beta^{(i)}(x) \quad i = 1, 2, 3.$$

Let $\beta = 0$ or $\beta = \pi/2$, we have

$$v_{r_i}(x) = \frac{\omega}{\pi N_g \gamma_G G_0}\theta_0^{(i)}, \quad v_{c_i}(x) = \frac{\omega}{\pi N_g \gamma_G G_0}\theta_{\pi/2}^{(i)} \quad i = 1, 2, 3.$$

We assume our measurements later to be of the form $v = v_r + \mathbf{j}v_c$ at some discrete measurement points and assume the measurements to be corrupted by noise. We will give a more detailed description of this in Section 3.4.

3.1.4 *Elasticity - Equations of Motion*

In this section we briefly describe the physics leading to the equations of motion of soft tissue under small harmonic deformation. In Section 3.2 we give a mathematical description of these equations and their boundary conditions.

The momentum transport in a continuum that is not in motion and has a constant density ρ, is described in the Lagrangian form by the following system of partial differential equations

$$\frac{1}{\rho}\nabla \cdot \sigma_{\text{el}} = \mathbf{f}, \tag{3.1.9}$$

where σ_{el} is the stress tensor and $\mathbf{f}$ is a vector containing all the accelerations caused by body forces. These equations are a special case of the Cauchy momentum equations [49] but in (3.1.9) we assume a vanishing flow velocity [8].

Stress expresses the loading in terms of the force applied to a certain cross-sectional area of an object which then causes the material to deform. Let us denote by v_t these displacements caused by stress. In the case of infinitesimal deformations in an isotropic linearly elastic solid we can relate the stress with the resulting deformations v_t in the following way that only involves the two mechanical parameters μ and λ [14]:

$$\sigma_{\text{el}} = \mu \left(\nabla v_t + \nabla v_t^T \right) + \lambda \left(\nabla \cdot v_t \right) I. \tag{3.1.10}$$

Elasticity describes the reversibility of deformations caused by stress, i.e. if stress components vanish, the displacement components drop to zero as well [66]. Linearly elastic materials deform proportional to applied load. Isotropic materials deform independent of direction.

We will later distinguish between purely elastic and viscoelastic material. For purely elastic materials the relation between stress and displacements is instantaneous, i.e. there is no time lag between a change in stress and the deformations caused by the stress. The response of viscoelastic materials on the

other hand is time dependent. Viscoelastic materials have a certain resistance to deform and this damping introduces a time dependency in the material's response. We will have a closer look at viscoelasticity in Section 3.2.1. Also, we will discuss in Section 3.5.1 how these underlying assumptions about soft tissue can influence elastography results.

Using $\mathbf{f} = (\partial^2 v_t / \partial t^2)$ [14] results in

$$\nabla \cdot \left(\mu \left(\nabla v_t + \nabla v_t^T \right) + \lambda \left(\nabla \cdot v_t \right) I \right) = \rho \frac{\partial^2 v_t}{\partial t^2},$$

and assuming that the displacements are time harmonic [14], i.e. there is a time independent $v(x)$ such that

$$v_t(x,t) = \mathrm{Re}\left[v(x) \exp(-\mathbf{j}\omega t) \right],$$

results in the equations of motion of soft tissue for harmonic displacements

$$\nabla \cdot \left(\mu \left(\nabla v + \nabla v^T \right) + \lambda \left(\nabla \cdot v \right) I \right) = -\rho \omega^2 v.$$

In the following section we will introduce boundary conditions and describe preconditions under which there exist unique weak solutions to the boundary value problem.

3.2 MODELLING ASSUMPTIONS

In this section we describe the model of harmonic mechanical deformations in the soft tissue of a living body and its underlying modelling assumptions. Let Ω be a non-empty bounded, open and connected set in $\mathbb{R}^2$ with Lipschitz continuous boundary $\partial\Omega$. Let $v(x) \in \mathbb{C}^2$ be the displacement field describing small mechanical deformations in the soft tissue. The displacement field under harmonic deformation satisfies the following system of elliptic partial differential equations (see [25], [41]):

$$\nabla \cdot [\sigma_{\text{el}}(v)] = -\rho_W \omega^2 v, \qquad \text{for } x \in \Omega, \tag{3.2.1}$$

$$v = \hat{v}, \qquad \text{for } x \in \partial\Omega_D, \tag{3.2.2}$$

$$[\sigma_{\text{el}}(v)] \cdot \mathbf{n} = 0, \qquad \text{for } x \in \partial\Omega_N, \tag{3.2.3}$$

with Cauchy stress tensor [14]

$$\begin{aligned} \sigma_{\text{el}}(v) &= \mu \left(\nabla v + \nabla v^T \right) + \lambda \left(\nabla \cdot v \right) I \\ &= \mu_S \left(\nabla v + \nabla v^T \right) + \lambda \left(\nabla \cdot v \right) I + \mathbf{j} \mu_L \left(\nabla v + \nabla v^T \right), \end{aligned} \tag{3.2.4}$$

where $\mathbf{j} := \sqrt{-1}$, $\nabla v^T = (\nabla v)^T$, $\mu = \mu_S + \mathbf{j}\mu_L \in L^\infty(\Omega, \mathbb{C})$ is the shear modulus, $I = \text{diag}(1, 1)$ the identity matrix, μ_S is the storage modulus , $\mu_L \in L^\infty(\Omega, \mathbb{R})$ is the loss modulus, $\lambda \in L^\infty(\Omega, \mathbb{R})$ is the first Lamé parameter, $\rho_W \in \mathbb{R}$, $\rho_W > 0$, is the density, $\omega \in \mathbb{R}$, $\omega > 0$ the frequency of the sinusoidal excitation , $x \in \Omega$

spatial coordinate and $\mathbf{n}$ the outer unit normal vector on body Ω. We denote by $\partial\Omega_D$ and $\partial\Omega_N$ the Dirichlet $(\partial\Omega)_D$ and Neumann $(\partial\Omega)_N$ part of the boundary $\partial\Omega$. Furthermore, we assume $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$ and $\partial\Omega_D \cap \partial\Omega_N = \varnothing$. We also assume

$$\mu_S, \mu_L > 0 \text{ and } \lambda + \mu_S > 0.$$

If the Dirichlet input $\hat{v}(x) \in H^{3/2}(\overline{\partial\Omega_D})$ there exists a unique weak solution $v \in H^1(\Omega)$ to the problem (3.2.1)-(3.2.3) (see [41]). Here $H^{3/2}(\overline{\partial\Omega_D})$ denotes the Sobolev space of fractional order $3/2$ [56].

Note that the problem (3.2.1)-(3.2.3) corresponds to the time-harmonic case where it is assumed that all waves propagating through the tissue are oscillating with angular frequency ω. The time-dependent wave field v_t is related to the time-harmonic v in the following way

$$v_t(x,t) = \mathrm{Re}\left[v(x)\exp(-\mathbf{j}\omega t)\right].$$

We call the set of points $x \in \partial\Omega_D$ on the Dirichlet boundary with $\hat{v}(x) \neq 0$ loading edge boundary. This models the sinusoidal excitation of angular frequency ω and amplitude $\hat{v}(x)$ applied by some external source. We call the set of points $x \in \partial\Omega_D$ on the Dirichlet boundary with $\hat{v}(x) = 0$ fixed edge boundary. This models fixed points like bones in the body where any incoming waves bounce off.

The set of points $x \in \partial\Omega_N$ with zero surface traction are called free edge boundary. Traction is defined to be $T_{\text{el}} = \sigma_{el} \cdot \mathbf{n}$. It can be shown that if no external forces are applied at a point $x \in \partial\Omega_N$ then the traction vanishes at this point. So, it is natural to impose this traction Neumann boundary condition if no external force is applied.

A free edge boundary assumes vacuum and a fixed edge boundary assumes infinitely stiff enclosure. This is not physical, however the classical boundary conditions used here are used as approximations of real world boundary conditions.

The shear modulus μ is a measure of the elastic shear stiffness. It describes how the material deforms under stress. The SI unit is Pascal. In the human body it varies on the order of 2000% under the presence of diseases like cancer or fibrosis and is therefore diagnostically useful [92].

The density ρ_W varies only on the order of 8% [59] and is therefore considered as a constant in our model.

Common frequencies used in MRE lie in range between 25Hz and 250Hz [43]. We will later also discuss low frequency MRE experiments with less common frequency < 4Hz.

In Figure 3.2.1 we illustrate the model given by (3.2.1)-(3.2.3).

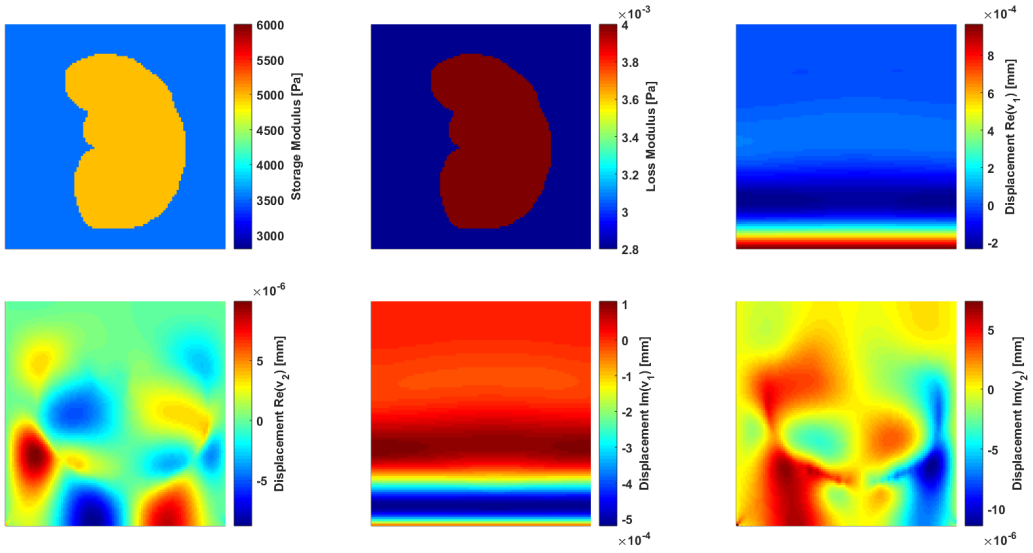


Figure 3.2.1: Example for a distribution of parameter μ across a square domain Ω of size 70mm by 70mm and corresponding wave field $v \in \mathbb{C}^2$ solving (3.2.1) - (3.2.3). The sinusoidal excitation has amplitude $\mathrm{Re}[\hat{v}_1] = 0.001$mm along the horizontal direction inducing shear waves. Frequency for the sinusoidal excitation is 60Hz ($\omega \approx 377$ rad/s) at the bottom edge. The top edge is a fixed edge. Left and right edge is a traction free Neumann boundary. Furthermore, $\rho_W = 1\mathrm{kg/l}$, $\nu = 0.499$. The amplitude of the displacement field is given in mm. **Top left:** Plot of the storage modulus distribution. The background value of μ_S is 3500Pa and the inclusion value is 5000Pa. **Top middle:** Plot of the loss modulus distribution. The background value of μ_L is 2800Pa and the inclusion value is 4000Pa. **Top right:** Horizontal displacement $\mathrm{Re}[v_1]$. **Bottom left:** Horizontal displacement $\mathrm{Im}[v_1]$. **Bottom middle:** Vertical displacement $\mathrm{Re}[v_2]$. **Bottom right:** Vertical displacement $\mathrm{Im}[v_2]$.

3.2.1 *Viscoelastic and Purely Elastic Model*

Viscosity of a fluid is a measure of its resistance to deformation at a given rate. In (3.2.1) - (3.2.3) the viscous contribution to the shear modulus μ is quantified by $\mu_L(x) \in \mathbb{R}$. For a purely elastic material, i.e. a material without any viscosity, the equations of motion are given by

$$\begin{aligned} \nabla \cdot \left[\mu_S \left(\nabla v + \nabla v^T\right) + \lambda \left(\nabla \cdot v\right) I\right] &= -\rho_W \omega^2 v, && \text{for } x \in \Omega, \\ v &= \hat{v}, && \text{for } x \in \partial\Omega_D \\ \left[\mu_S \left(\nabla v + \nabla v^T\right) + \lambda \left(\nabla \cdot v\right) I\right] \cdot \mathbf{n} &= 0, && \text{for } x \in \partial\Omega_N. \end{aligned}$$

So, in this viscosity free case we assume $\mu_L = 0$, the Cauchy stress tensor becomes real-valued

$$\sigma_{\text{el}}(v) = \mu_S \left(\nabla v + \nabla v^T\right) + \lambda \left(\nabla \cdot v\right) I. \tag{3.2.5}$$

In the viscosity free case we have $v(x) \in \mathbb{R}^2$ and we assume the measurements to be real-valued as well. We discuss assumptions on the measurements in MRE in Section 3.4.

In purely elastic materials both stress $\sigma_{\text{el}}(v)$ and strain $\mu_S \left(\nabla v + \nabla v^T\right) + \lambda \left(\nabla \cdot v\right) I$ (see 3.2.5), are in phase.

In viscoelastic fluids stress and strain are out of phase [57], i.e. there is a time lag between stress and displacements caused by the stress when the

material is subjected to oscillatory displacements. Due to the time-harmonic nature of stress and strain, this time lag can be described by a phase angle between stress and strain. Conveniently, this can be expressed using complex variables for strain and stress which then results in a complex shear modulus μ [57, Chapter 2].

The real part μ_S is negatively correlated with the phase angle between stress and strain and describes the elastic part, i.e. how much in-phase the stress and strain component is and is a measure of the stored energy. The complex part μ_L is correlated to the phase angle between stress and strain and describes the viscous part, i.e. how much out-of-phase the stress and strain component is and is a measure of the energy lost as heat. We have $\mu_L = \omega\mu_l$ where μ_l is the so called shear viscosity [42].

In Figure 3.2.2 we illustrate the purely elastic model.

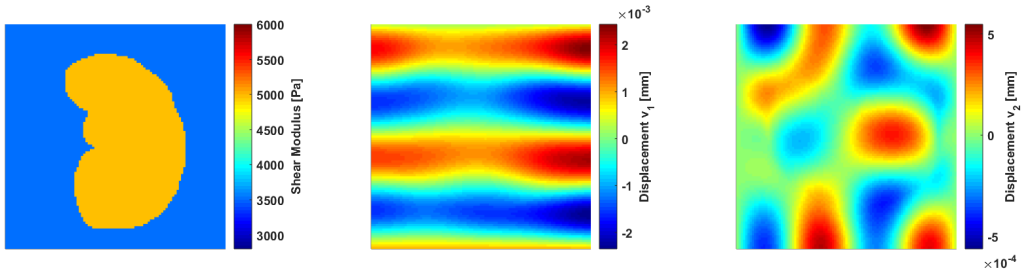


Figure 3.2.2: Example for a distribution of parameter μ across a square domain Ω of size 70mm by 70mm and corresponding wave field $v \in \mathbb{R}^2$ solving (3.2.1) - (3.2.3) for the purely elastic case, i.e. $\mu_L = 0$. The sinusoidal excitation has amplitude $\hat{v}_1 = 0.001$mm along the horizontal direction inducing shear waves. Frequency for the sinusoidal excitation is 60Hz ($\omega \approx 377$ rad/s) at the bottom edge. The top edge is a fixed edge. Left and right edge is a traction free Neumann boundary. Furthermore, $\rho_W = 1$kg/l, $\nu = 0.499$. The amplitude of the displacement field is given in mm. **Left:** Plot of the storage modulus distribution. The background value of μ_S is 3500Pa and the inclusion value is 5000Pa. **Middle:** Horizontal displacement v_1. **Right:** Vertical displacement v_2.

3.3 WEAK FORMULATION OF THE VISCOELASTIC HELMHOLTZ EQUATION

In this section we derive a weak formulation for problem (3.2.1)-(3.2.3) which is then used to describe how we can find approximations of the solutions using the standard Galerkin finite element method. We homogenise the problem

(3.2.1)-(3.2.3) (see for example [33]). We take $\tilde{\Phi} \in H^1(\Omega)$ such that $\tilde{\Phi}|_{\partial\Omega_D} = \hat{v}$ and seek $\tilde{v} := v - \tilde{\Phi}$ such that

$$\nabla \cdot \sigma_{\text{el}}(\tilde{v}) = -\rho_W \omega^2 \tilde{v} + \nabla \cdot \sigma_{\text{el}}(\tilde{\Phi}), \qquad \text{for } x \in \Omega, \tag{3.3.1}$$

$$\tilde{v} = 0, \qquad \text{for } x \in \partial\Omega_D \tag{3.3.2}$$

$$\sigma_{\text{el}}(\tilde{v}) \cdot \mathbf{n} = -\sigma_{\text{el}}(\tilde{\Phi}) \cdot \mathbf{n}, \qquad \text{for } x \in \partial\Omega_N. \tag{3.3.3}$$

In order to get the weak form we multiply the residual by a test function $w \in H_0^1(\Omega)$ defined by

$$H_0^1(\Omega) := \left\{ w \in H^1(\Omega); w|_{\partial\Omega_D} = 0 \right\},$$

and integrate over the domain Ω

$$\int_\Omega w \left(\nabla \cdot \sigma_{\text{el}}(\tilde{v}) + \rho_W \omega^2 \tilde{v} - \nabla \cdot \sigma_{\text{el}}(\tilde{\Phi}) \right) \mathrm{d}\Omega = 0. \tag{3.3.4}$$

By applying integration by parts [50], i.e.

$$-\int_\Omega w \nabla \cdot \sigma_{\text{el}}(\tilde{v}) \, \mathrm{d}\Omega = -\int_{\partial\Omega} w \sigma_{\text{el}}(\tilde{v}) \cdot \mathbf{n} \, \mathrm{d}(\partial\Omega) + \int_\Omega \nabla w : \sigma_{\text{el}}(\tilde{v}) \, \mathrm{d}\Omega,$$

where $A : B$, the Frobenius inner product for matrices A and B defined in Chapter 2. As the test function is zero on $\partial\Omega_D$, we get

$$-\int_\Omega w \nabla \cdot \sigma_{\text{el}}(\tilde{v}) \, \mathrm{d}\Omega = -\int_{\partial\Omega_N} w \sigma_{\text{el}}(\tilde{v}) \cdot \mathbf{n} \, \mathrm{d}(\partial\Omega_N) + \int_\Omega \nabla w : \sigma_{\text{el}}(\tilde{v}) \, \mathrm{d}\Omega,$$

We note that [50]

$$\int_\Omega \nabla w : \sigma_{\text{el}}(\tilde{v}) \, \mathrm{d}\Omega = \int_\Omega 2\mu \cdot \varepsilon(\tilde{v}) : \varepsilon(w) \, \mathrm{d}\Omega + \int_\Omega \lambda \left(\nabla \cdot \tilde{v} \right) \left(\nabla \cdot w \right) \mathrm{d}\Omega,$$

can be written symmetrically, where we used the notation

$$\varepsilon(\tilde{v}) := \frac{1}{2}\left(\nabla\tilde{v} + \nabla\tilde{v}^T\right).$$

We define the Hermitian forms $\hat{a}_{\text{her}}, \bar{a}_{\text{her}} : H_0^1 \times H_0^1 \to \mathbb{C}$ with

$$\hat{a}_{\text{her}}(w, \tilde{v}) := \int_\Omega 2\mu \cdot \varepsilon(\tilde{v}) : \varepsilon(w)\, \mathrm{d}\Omega + \int_\Omega \lambda\, (\nabla \cdot \tilde{v})\, (\nabla \cdot w)\, \mathrm{d}\Omega,$$

and

$$\bar{a}_{\text{her}}(w, \tilde{v}) := \hat{a}_{\text{her}}(w, \hat{v}) - \int_\Omega w \rho_W \omega^2 \tilde{v} \mathrm{d}\Omega.$$

We also define the linear form $l_{\text{lin}} : H_0^1 \to \mathbb{C}$ with

$$l_{\text{lin}}(w) = \int_{\partial\Omega_N} w \cdot \sigma_{\text{el}}(\tilde{\Phi}) \cdot \mathbf{n}\, \mathrm{d}(\partial\Omega_N).$$

We can rewrite (3.3.4) now in the following way

$$\bar{a}_{\text{her}}(w, \tilde{v}) = \hat{a}_{\text{her}}(w, \tilde{\Phi}) - 2l(w). \tag{3.3.5}$$

We call a function $\tilde{v} \in H_0^1(\Omega)$ satisfying this variational problem (3.3.5) for all $w \in H_0^1(\Omega)$ a weak solution of the elasticity problem.

In order to find approximations of the weak solutions, we consider a subspace $H_0^{\mathfrak{N}}$ of H_0^1 and try to find $\tilde{\tilde{v}} \in H_0^{\mathfrak{N}}(\Omega)$ satisfying this variational problem (3.3.5) for all $\tilde{w} \in H_0^{\mathfrak{N}}(\Omega)$. We assume $H_0^{\mathfrak{N}}$ to be finite-dimensional with linear tent [60] basis functions $\mathfrak{w}_i$, $i = 1, \ldots, N$. We can write $\tilde{w} = \sum_i^{\mathfrak{N}} w_i \mathfrak{w}_i$ and $\tilde{\tilde{v}} = \sum_i^{\mathfrak{N}} \tilde{v}_i \mathfrak{w}_i$ in (3.3.5) which then leads after some calculation (for example found in [9]) to a linear system of equations that can be solved for w_i.

3.4 INVERSE PROBLEM

In this subsection we describe the inverse problem arising in MRE and the modelling assumptions underlying this inversion. We go back to (3.2.1)-(3.2.3) in order to describe the inverse problem. We usually replace λ by an expression containing the dimensionless constant Poisson's ratio ν defined as follows [14]

$$\lambda = \frac{2\mu_S \nu}{1-2\nu}.$$

Then (3.2.1) can be written as

$$\nabla \cdot \left[\mu \left(\nabla v + \nabla v^T \right) + \frac{2\mu\nu}{1-2\nu} \left(\nabla \cdot v \right) I \right] = -\rho_W \omega^2 v, \qquad \text{for } x \in \Omega. \tag{3.4.1}$$

We assume Poisson's ratio $\nu = 0.499$ to be constant which models the near incompressibility of human tissue [62].

Let $\mathcal{M} := L^\infty(\Omega, \mathbb{C})$. The only remaining unknown variables in (3.2.1)-(3.2.3) are μ and v. Equation (3.4.1) defines a (nonlinear) parameter-to-output map assigning each $\mu \in \mathcal{M}$ the solution v of (3.4.1) with boundary conditions (3.2.2) and (3.2.3). Let $x_1, \ldots, x_m \in \Omega$ be measurement locations and let us define the map

$$\mathcal{G} : \mathcal{M} \to \mathbb{C}^m, \quad \mu \mapsto \mathcal{G}(\mu) := \left(v(x_1), \ldots, v(x_m) \right), \tag{3.4.2}$$

where v is the solution v of (3.4.1) with boundary conditions (3.2.2) and (3.2.3) evaluated at the measurement points.

Our modelling assumption is that our measurement $\mathbf{v} \in \mathbb{C}^m$ is a solution v of (3.2.1)-(3.2.3) for an unknown parameter μ evaluated at the measurement points $x_1, \ldots, x_m \in \Omega$. Furthermore, we assume that $\mathbf{v}$ is corrupted by noise and that the noise is an m-dimensional mean zero Gaussian random variable. In other words, when using the forward operator defined in (3.4.2), we assume

$$\mathbf{v} = \mathcal{G}(\mu) + \eta, \tag{3.4.3}$$

where η denotes the measurement noise. The inverse problem consists of finding $\mu \in \mathcal{M}$ from measurements $\mathbf{v}$.

This problem is ill-posed that means the solution μ does not depend continuously on the data $\mathbf{v}$. Small changes in $\mathbf{v}$ lead to big changes in the solution μ for this inversion [92].

We also assume that the noise $\eta \in \mathbb{C}^m$ in MRE is the sum of many, mostly independent sources of noise. The central limit theorem states that the normalised sum of independent random variables with some overall mean and finite variance tends to be Gaussian. So, the central limit theorem modifies our assumption of additive Gaussian noise in the model. However, we only have a finite number of sources of noise and therefore, the central limit theorem cannot applied exactly. Sources of noise in MRE might be, among others, motion of the body, respiration, movement of the table and signal noise in MRI, which goes

back to radiofrequency noise from external sources and increased temperature (thermal agitation) of the body and the electronics inside the MRI system [91].

3.5 COMMON APPROACHES TO INVERSION IN MRE

In this section we describe the two commonly applied approaches to the inverse problem in MRE (4.1.1). Both approaches do not take into account the statistics of the noise η unlike the approaches from the inversion theory discussed later in Section 4.1.5. The first one solves for μ after applying an operation of curl to (3.2.1). In Section 3.5.1 we describe this approach which is called direct inversion. The other approach will be described in Section 3.5.2 and tries to minimise the difference between the measured wave field and those wave fields computed by a finite element method. A common modification of this computationally expensive approach is to split the minimisation problem into smaller sub-problems. We describe this method in general in Section 3.5.3 and how to combine it with the least-squares optimisation in Section 3.5.4.

3.5.1 *Direct Algebraic Inversion*

A widely used approach for solving the MRE inverse problem is direct algebraic inversion [23, 14]. In order to derive the algebraic inversion formula for soft tissue stiffness from equation (3.2.1), we rewrite it using the fact that $\nabla \cdot \lambda (\nabla \cdot v) I = \nabla (\lambda \nabla \cdot v)$ which gives us

$$\nabla \cdot \mu \left(\nabla v + \nabla v^T \right) + \nabla (\lambda \nabla \cdot u) = -\rho_W \omega^2 v. \tag{3.5.1}$$

Now, an operation of curl performed on equation (3.5.1), which results in

$$\nabla \times \nabla \cdot \mu \left(\nabla v + \nabla v^T \right) + \nabla \times \nabla (\lambda \nabla \cdot u) = -\rho_W \omega^2 (\nabla \times v), \tag{3.5.2}$$

and then using the fact that the curl of the gradient of any smooth f is always zero [30], i.e. $\nabla \times \nabla f = 0$, simplifies (3.5.2) to

$$\nabla \times \nabla \cdot \mu \left(\nabla v + \nabla v^T \right) = -\rho_W \omega^2 (\nabla \times v).$$

In the next step of the algebraic direct inversion approach we assume that μ is a constant which is usually referred as the "local homogeneity assumption" [80]. It can be shown that $\nabla \times \nabla \left(\nabla v + \nabla v^T \right) = \nabla^2 (\nabla \times v)$, where $\nabla^2 = \nabla \cdot \nabla$ is the Laplace operator. We can now write

$$\mu \nabla^2 (\nabla \times v) = \rho_W \omega^2 \cdot (\nabla \times v). \tag{3.5.3}$$

We can interpret this as the wave equation for the propagating shear wave [14] and μ as a factor in the wave number. Rearranging 3.5.3 yields the direct inversion formula which is

$$\mu = \frac{\rho_W \cdot \omega^2(\nabla \times v)}{\nabla^2(\nabla \times v)}. \tag{3.5.4}$$

In algebraic direct inversion approaches like [23, 14] the measurement data $\mathbf{v} \in \mathbb{C}^m$ is then used to calculate the finite-dimensional approximation $\hat{\mu} \in \mathbb{C}^m$ of the shear modulus by

$$\hat{\mu} = \frac{\rho_W \cdot \omega^2(\hat{\nabla} \times \mathbf{v})}{\hat{\nabla}^2(\hat{\nabla} \times \mathbf{v})}, \tag{3.5.5}$$

where a discrete curl operator that uses finite differences is applied.

This approach is one of the computationally cheapest approaches to the inverse problem in MRE and inversions for real world data sets only take a few seconds [22]. However, there are several weaknesses of this approach: We assume $\mathbf{v}$ is corrupted by noise and gets differentiated twice in (3.5.5). In the presence of a high noise level the accuracy of direct inversion is poor. Usually, smoothing is applied to the data $\mathbf{v}$ before the inversion or after the inversion to $\hat{\mu}$.

Also, direct inversion is restricted to model (3.2.1) and suffers from the modelling errors attached to this model. The differential equation (3.2.1) is a reasonable choice for modelling soft human tissue. However, [42] discusses

various viscoelastic models for the soft tissue in a living body and each of them results in different equations of motion. In [70], [28] the authors show how the application of a poroelastic model instead of the linear elastic model (3.2.1)-(3.2.3) can improve inversion results in real world experimental setups with tofu. The poroelastic model models fluids inside organs, such as blood and interstitial fluid. The governing equations contain a term for pore water pressure and direct inversion cannot be applied anymore. Also, if MRE is applied to other materials direct inversion (3.5.5) might not be applied. In applications of MRE to materials that are incompressible, the model (3.2.1)-(3.2.3) changes and we have $\nabla \cdot v = 0$.

In the application of the discrete curl (3.5.5) the numerical derivatives contain errors which make approximations $\hat{\mu}$ less accurate. In regions with significant heterogeneity, the accuracy of direct inversion suffers [23]. We will see in the experiments in Section 6.4 that this is the main weakness of this approach.

3.5.2 *Least-Squares - General Introduction*

Another commonly used approach to the inverse problem in MRE is to find the optimiser of the least-squares function: [23]

$$\min_{\mu \in \mathcal{M}} \pi(\mu), \tag{3.5.6}$$

where

$$\pi(\mu) = \|(\mathcal{G}(\mu) - \mathbf{v})\|^2.$$

Algorithms that try to find solutions or approximations are usually iterative, the most common one is Gauss-Newton [31]. We discuss several weaknesses of least-squares approaches and the specific least-squares approach in [31] in the context of inverse problems later in this Chapter and in Section 4.1.

Finding an optimiser of (3.5.6) ignores the distribution of noise and therefore, the distribution of measurements $\mathbf{v}$. If we have a very high noise level at only a few measurement locations $\mathbf{v}_i$ and low noise level otherwise, the euclidean norm in (3.5.6) is dominated by these few locations and iterative methods tune updates to these outliers only. Even though this is an extreme example illustrating the weakness of this approach, it is realistic however, that displacement data in one direction in MRE might be more corrupted by noise than in other ones.

In the statistical approach explained below the precision operator coming from the distribution of the noise is used in the least square function which takes into account the distribution of noise in the measurements.

3.5.3 *Subzones*

In this subsection we describe the subzone inversion approach which was introduced by van Houten in [32] and [31] in the context of least-squares minimisation described above. We give a description of this idea in this subsection that enables us to use it later on in statistical inversion approaches. The key idea is to circumvent high computational costs by dividing the field-of-view into a series of overlapping subzones Ω_z.

More specifically, we define non-empty, open and connected subsets $\Omega_z \subset \Omega$ in the domain of interest Ω such that they cover the whole domain, i.e. $\Omega = \cup\Omega_z$. In each subzone Ω_z we define a boundary value problem

$$\nabla \cdot \sigma_{\mathrm{el}_z}(v) = -\rho_W \omega^2 v_z, \qquad \text{for } x \in \Omega_z, \tag{3.5.7}$$

$$v_z = \hat{v}, \qquad \text{for } x \in \partial\Omega_D \cap \partial\Omega_z, \tag{3.5.8}$$

$$v_z = \mathbf{v}_C, \qquad \text{for } x \in \partial\Omega_z \setminus (\partial\Omega_N \cup \partial\Omega_D), \tag{3.5.9}$$

$$\sigma_{\mathrm{el}_z}(v) \cdot \mathbf{n} = 0, \qquad \text{for } x \in \partial\Omega_N \cap \partial\Omega_z, \tag{3.5.10}$$

where $\sigma_{\mathrm{el}_z}(v) := \mu\left(\nabla v_z + \nabla v_z^T\right) + \lambda\left(\nabla \cdot v_z\right) I$ and $\mathbf{v}_C$ is obtained from measurements $\mathbf{v}$ by interpolation in Ω_z, e.g. bilinear interpolation. So, measurements

are used to pose boundary conditions on the subzone. Similarly to (3.4.2) we can define a subzone forward map

$$\mathcal{G}_z : L^\infty(\Omega_z, \mathbb{R}) \to \mathbb{R}^{m_z}, \quad \mu \mapsto (v_z(x_1), \dots, v_z(x_{m_z})), \tag{3.5.11}$$

where $\mathcal{G}_z(\mu)$ is the solution v_z of the boundary value problem above for μ and $x_1, \dots, x_{m_z}$ are measurement points inside subzone Ω_z. The inverse problem (4.1.1) can now be approximated in a subzone Ω_Z by using $\mathcal{G}_z$ instead of $\mathcal{G}|_{\Omega_Z}$. The noise in the data $\mathbf{v}$ and the fact that we use interpolation in order to get $\mathbf{v}_C$ introduce an error on the subzone boundary (3.5.9) which is carried further to an error between the subzone solution v_z using $\mathcal{G}_z$ and the global solution v when restricted to the subzone Ω_Z, in other words $v|_{\Omega_Z}$.

The quantification of this error as well as the development of methods to minimise the error is an open problem.

3.5.4 *Subzone-Least-Squares*

In this subsection we describe further details of one of the most commonly used approaches to the inverse problem in MRE next to direct inversion.

In the least squares approach in [31], [32] the authors consider the unknown μ in (4.1.1) and μ_z in (3.5.11) only on a discrete mesh in Ω even in the formulation of the inverse problem. Let points $\{\hat{x}_i\}_{i=1}^{\hat{n}} \in \Omega$ correspond to the chosen

computational mesh of the discretisation scheme used in the finite element method described at the end of Section 3.3 and define

$$\hat{\mu} := (\mu(\hat{x}_1), \dots, \mu(\hat{x}_{\hat{n}})) \,. \tag{3.5.12}$$

The least-squares function considered in [31] is

$$\min_{\hat{\mu} \in \hat{\mathcal{M}}} \hat{\pi}(\hat{\mu}), \tag{3.5.13}$$

where $\hat{\mathcal{M}} := \mathbb{R}^{\hat{n}}$ is the space of all $\hat{n}$-dimensional approximations of μ and

$$\hat{\pi}(\hat{\mu}) = \|(\hat{\mathcal{G}}(\hat{\mu}) - \mathbf{v})\|^2. \tag{3.5.14}$$

Here, we use $\hat{\mathcal{G}} : \mathbb{R}^{\hat{n}} \to \mathbb{R}^m$ that maps from the finite-dimensional approximation $\hat{\mu}$ of μ to the the finite-dimensional approximation $\tilde{v}^h$ of $\tilde{v}$ described at the end of Section 3.3, i.e. $\hat{\mathcal{G}}(\hat{\mu})$ is the finite element approximation of $G(\mu)$ for computational mesh $\{\hat{x}_i\}_{i=1}^{\hat{n}} \in \Omega$. The mapping $\hat{\pi}$ is mapping $\mathbb{R}^{\hat{n}} \to \mathbb{R}$.

In [31] the authors then use the subzone idea described in Section 3.5.3 and define non-empty, open and connected subsets $\Omega_z \subset \Omega$ in the domain of interest Ω such that they cover the whole domain, i.e. $\Omega = \cup \Omega_z$. Subzones might overlap. In Figure 3.5.1 we give an example on a domain division into several overlapping subzones on an rectangular grid of discretisation. In each subzone Ω_z we define an objective function

$$\hat{\pi}_Z(\hat{\mu}) := \sum_z \hat{\pi}_z(\hat{\mu}_z) := \sum_z \|\mathbf{v_z} - \hat{\mathcal{G}}_z(\hat{\mu}_z)\|^2, \tag{3.5.15}$$

where $\mathbf{v}_z := \mathbf{v}|_{\Omega_z} \in \mathbb{R}^{m_z}$ is the measurement data and $\hat{\mu}_z \in \mathbb{R}^{\hat{n}_z}$ the shear modulus in the corresponding subzone z. $\hat{\pi}_z$ denotes the least square functional (3.5.14) in subzone Ω_z. We denote the mapping $\hat{\mathcal{G}}_z : \mathbb{R}^{\hat{n}_z} \to \mathbb{R}^{m_z}$ maps the finite-dimensional approximation $\hat{\mu}$ of μ given on the nodes $\hat{x}_i \in \Omega_z$ of the chosen computational mesh inside the subzone Ω_z to the displacements v_z in points $x_1, \ldots, x_{m_z}$.

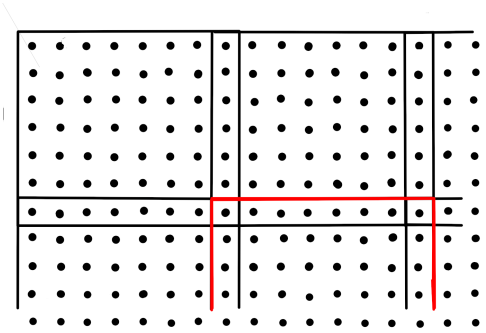


Figure 3.5.1: After discretisation, subzones are defined on the mesh. Subzones might overlap. In this picture subzones overlap by one point.

In [31] the authors use the assumption that the minimiser of the sum in equation (3.5.15) is equivalent to the sum over the minimisers in each sub-zones and we have :

$$\arg\min_{\hat{\mu}\in\hat{\mathcal{M}}} \hat{\pi}_Z(\hat{\mu}) = \arg\min_{\hat{\mu}\in\hat{\mathcal{M}}} \left\{ \sum_z \hat{\pi}_z(\hat{\mu}_z) \right\} = \sum_z \arg\min_{\hat{\mu}\in\hat{\mathcal{M}}} \hat{\pi}_z(\hat{\mu}_z) = \arg\min_{\hat{\mu}\in\hat{\mathcal{M}}} \hat{\pi}(\hat{\mu}),$$

for the minimiser. In [31] the authors then use the Levenberg–Marquardt method to update the parameter in the following way

$$\hat{\mu}_z^{i+1} = \hat{\mu}_z^i + \Delta\hat{\mu}_z,$$

where the update $\Delta\mu_z$ is the solution of the following matrix system:

$$[\mathbf{H}_z^i + \hat{\alpha} I][\Delta\hat{\mu}_z] = -\text{grad}_z \pi(\hat{\mu}^i),$$

where approximations of the gradient $\text{grad}_z \pi(\hat{\mu}^i)$ and the Hessian $\mathbf{H}_z^i$ which involve the derivative $\frac{\partial \hat{\mathcal{G}}_z(\hat{\mu}^i))_l}{\partial \hat{\mu}_{z_j}}$ can be found by approximating a modified version of (3.2.1)-(3.2.3) [32].

4

BAYESIAN INVERSION

In this section we describe the Bayesian inversion approach we use to solve the inverse problem arising in MRE. In Section 4.1 we introduce the two main approaches from inversion theory called Bayesian and deterministic inversion, and explain the properties and advantages of each approach. In Section 4.2 we explain the specific Bayesian approach that we have chosen which is EKI. The key part in this Chapter is Algorithm 1 which is the version of EKI that we use in this thesis. In Section 4.3 we show how we define the prior in our Bayesian approach. Finally, in Section 4.4 and Section 4.5 we show how our approach can be modified using a different parametrisations so that it can capture well complex geometries in the unknown.

4.1 BASIC APPROACHES FOR INVERSE PROBLEMS

In this section we describe the two main approaches to inversion and why we have chosen one over the other. In Section 4.1.3 we introduce deterministic

inversion and in Section 4.1.4 and 4.1.5 we describe Bayesian inversion. In Section 4.1.6 we explain why we follow a Bayesian approach. We start this section with an overview over inversion approaches in Section 4.1.2. but before we do that we give a more general formulation of the inverse problem (3.4.3) that includes a parametrisation of the unknown. This general framework will be useful later on when we use parametrisations in Sections 4.4 and4.5.

4.1.1 *Formulation of the Inverse Problem that includes a parametrisation*

Before we describe basic approaches from the inversion theory we give a more general formulation of the inverse problem (3.4.3) by allowing for a class of suitable parametrisations $\mathcal{P} : \mathcal{K} \to \mathcal{M}$ for our estimate of the unknown $\mu \in \mathcal{M}$. These parametrisations characterise the physical properties in $\mathcal{M}$ in terms of an input $k \in \mathcal{K}$. The space of parameters $\mathcal{K}$ will be specified later depending on the chosen parametrisation. For now, we assume $\mathcal{K}$ is a (infinite-dimensional) Hilbert space. When using a parametrisation, inversion approaches try to estimate the parameters $k \in \mathcal{K}$ instead of the physical property $\mu \in \mathcal{M}$. We define $\mathcal{F} := \mathcal{G} \circ \mathcal{P}$ and write down the general formulation of the inverse problem (3.4.3) as finding $k \in \mathcal{K}$ from measurements $\mathbf{v}$ with

$$\mathbf{v} = \mathcal{F}(k) + \eta. \tag{4.1.1}$$

Choosing $\mathcal{P} := \mathrm{id}$ we have (3.4.3) as special case of (4.1.1).

4.1.2 *Overview*

The inversion approaches described in Section 3.5 do not take into account the distribution of noise. Due to the aforementioned ill-posedness of inverse problems two outcomes of the same experiment can therefore lead to very different inversion results in their approaches. We want to deploy and develop methods that give similar inversion results if they are applied to the same experiment. Also, approaches in Section 3.5 assume the real world could be exactly represented by fairly simple models like (3.2.1) and do not allow for modelling errors.

Methods in inversion theory try to incorporate the statistics of the noise or modelling errors into their methods [45, 86]. More precisely, they consider the measurements $\mathbf{v}$ in (4.1.1) as realisation of a (finite-dimensional) random variable V. We assume the noise $\eta \sim \mathcal{N}(0, B)$ is centred Gaussian with covariance operator $B \in \mathbb{C}^{m\times m}$ and denote its probability density function with respect to Lebesgue measure as ρ_n. From (4.1.1) we conclude that the density ρ of $\mathbf{v}$ given k is given by

$$\rho(\mathbf{v}|k) := \rho_n(\mathbf{v} - \mathcal{F}(k)) = \frac{1}{\sqrt{(2\pi)^m |B|}} \exp\left(-\frac{1}{2}\|(\mathbf{v} - \mathcal{F}(k))\|_B^2\right),$$

and therefore, the data likelihood L_V for k given observation $V = \mathbf{v}$ is

$$L_V(k;\mathbf{v}) := \rho(\mathbf{v}|k) = \frac{1}{\sqrt{(2\pi)^m |B|}} \exp\left(-\frac{1}{2}\|(\mathbf{v} - \mathcal{F}(k))\|_B^2\right), \tag{4.1.2}$$

where $\|\cdot\|_B := \|B^{-1/2}\cdot\|$.

If $L_V(k_1;\mathbf{v}) > L_V(k_2;\mathbf{v})$ then the observation v is more likely to occur under k_1 than k_2 so that k_1 can be considered as more plausible than k_2.

Approaches from inversion theory can be split into two groups: Deterministic inversion and Bayesian inversion.

4.1.3 *Deterministic Approach to Inversion*

In the classical or deterministic approach to inversion k is a fixed unknown constant which is inferred utilising the distribution $\rho(\mathbf{v}|k)$ (or more precisely the covariance B of the noise) after $V = \mathbf{v}$ has been observed [45]. The centre of interest is to construct good methods of inference and the unknown k is not treated as a random variable. A familiar method of deterministic inversion is to find the $k^* \in \mathcal{M}$ that is maximising the likelihood and is therefore the most plausible estimate given the data $\mathbf{v}$. In other words we are solving for given observation $\mathbf{v}$ the optimisation problem of finding the optimiser

$$k^*_{\mathrm{MDL}} := \arg\max_{k\in\mathcal{K}} L_V(k;\mathbf{v}) = \arg\min_{k\in\mathcal{K}} \|\mathbf{v} - \mathcal{F}(k)\|_B^2, \tag{4.1.3}$$

which is in the case of Gaussian noise with covariance operator B a generalised (nonlinear) least square problem.

In (finite-dimensional) parameter estimation theory of deterministic inference the quantity (4.1.3) is called the maximum likelihood estimator [45].

Solving or approximating (4.1.3) is a difficult optimisation task since the functional that needs to be minimised is not a quadratic one with respect to k because $\mathcal{F}$ is nonlinear. If the minimisation problem is ill-posed, iterative numerical approaches may fail finding updates [72]. If maxima exist optimisation methods may only converge to a local maximum rather than the global maximum. Also, they might overfit the data $\mathbf{v}$ if the of the noise is not known exactly, i.e. $\mathcal{F}(k^*_{\mathrm{MDL}})$ is almost exactly $\mathbf{v}$. In this case, we try to explain the noise in the measurement which is not a goal of inversion. Therefore, it is common in deterministic inversion to replace (4.1.3) by similar optimisation problems with better properties, e.g. regularised optimisation problems before (Tikhonov) or iteratively while solving the problem (Levenberg-Marquardt, see Section 4.2.6 and [45]).

The Tikhonov optimisation problem for problem (4.1.3) is finding the optimiser [45]

$$k^*_{\mathrm{Tk}} := \arg\min_{k\in\mathcal{K}} \|\mathbf{v} - \mathcal{F}(k)\|^2_B + \delta_{\mathrm{Tk}} \cdot G_{\mathrm{TK}}(k), \tag{4.1.4}$$

where δ_{Tk} is a regularisation parameter and $G_{\text{Tk}} : \mathcal{M} \to \mathbb{R}$ some functional. Depending on the problem, some choices of G_{Tk} and δ_{Tk} can enforce existence, uniqueness and continuity with respect to the measurements of the solution of this problem [45].

Deterministic statistical inversion requires the knowledge of the distribution or the covariance of the noise ρ_n and many approaches like Tikhonov-regularisation allow further incorporation of prior knowledge through the choice of the regularisation parameter [11].

4.1.4 *Motivation for Bayesian Inversion*

The inversion approaches in Section 3.5 and deterministic inversion methods return a single estimate for the unknown without information about how certain these results are. Due to the presence of noise in the input of these algorithms it is clear that their output is fraught with uncertainty as well even if the distribution of the noise is taken into account like in deterministic inversion methods. It is therefore desirable to rather use an approach that attempts to attach a measure of uncertainty or a probability statement to the inversion result which is what Bayesian (or statistical) inversion does. The central idea

of statistical inversion is to consider data as partial information representing a larger context and to make conclusions that try to involve this context.

4.1.5 *Bayesian Inversion*

Bayesian inversion enables us to code prior knowledge directly into our algorithms. Not only measurements are treated as random variable but also k is treated as random. The uncertainty about the parameter k prior to observing $V = \mathbf{v}$ is represented by a prior distribution on k. This prior should incorporate all the structural knowledge we have about the problem prior the acquisition of the data. Let us assume the density of the prior with respect to Lebesgue measure exists and let us denote it by $\pi_0(k)$. Bayes' theorem states that the posterior distribution, that is, the conditional probability distribution $\pi^k(\mathbf{v})$ with respect to Lebesgue measure of k given $V = \mathbf{v}$ is given by [83]

$$\pi^k(\mathbf{v}) = \frac{\rho(\mathbf{v}|k)\pi_0(k)}{\int_{\mathcal{K}} \rho(\mathbf{v}|k)\pi_0(k)\mathrm{d}k}. \tag{4.1.5}$$

In Bayesian inversion, the inference statement about the unknown k in form of the posterior distribution condenses all that is known about k following the

measurement of $\mathbf{v}$. It is worth noting that for each choice of prior distribution π_0 the mapping

$$\mathbf{v} \mapsto \pi^k(\mathbf{v}),$$

that assigns the posterior for measurement $\mathbf{v}$ defined by (4.1.5), is different. We will see later in Chapter 5 in Figure 4.3.2 how different priors in the same experiment with the same measurement $\mathbf{v}$ yield different posteriors in a Bayesian inversion approach.

One of many challenges in Bayesian inversion is to specify a suitable prior for the problem and calculate the posterior in (4.1.5) involving the usually high-dimensional integral in the denominator which is both usually non-trivial.

Unlike the deterministic approach that assumes that the parameter k has one true value, Bayesian methods believe k is fixed but has been drawn from some probability distribution.

4.1.6 *Choice Between Deterministic Approach and Bayesian Approach*

The choice of school of inversion can be based on practical, theoretical, computational and sometimes even philosophical aspects. In the inverse problem arising in MRE we will apply algorithms that belong to Bayesian inversion. One of the main downsides and critiques of Bayesian inversion is that the

prior is subjective and different priors for the same experiment with the same measurements may produce different posteriors. In the Bayesian approach to MRE we have to specify a prior on the shear modulus μ which describes mechanical properties of the human tissue. The mechanics in human tissue follow physical laws and there is plenty of empirical data on how the shear modulus distributes across tissue. This enables us to give informative priors and also specify the assumptions for them. For example, one of the parameters k of the prior determines the range of values μ. Due to the fact that we apply MRE to human tissue, we are able to narrow down the possible values of μ in the prior to a relatively small range. So, it can be considered as a positive aspect of the Bayesian approach that we can make use of this detailed prior knowledge and that we can incorporate it into our algorithms in a such simple and direct way that even non-mathematicians are able to adapt it to different MRE-experiments.

Throughout this thesis we consider the inversion that we develop from a Bayesian viewpoint. However, it is important to note that many approaches from deterministic inversion can be interpreted from a Bayesian perspective and vice versa [11]. The Bayes formula should be merely seen as a starting point for the development of the algorithm in this thesis. Later, we will see how the Bayesian approach used in this work can be related to the Levenberg-Marquardt

approach and that iterative regularisation ideas from Levenberg-Marquardt are used as well in this approach.

4.2 ENSEMBLE KALMAN INVERSION

In this section we explain the Bayesian approach we follow in the inversion problem in MRE: Ensemble Kalman Inversion (EKI). In Section 4.2.2 and 4.2.3 we explain how the Ensemble Kalman update in Section 4.2.4 can be deduced from the Bayes formula (4.1.5). After that in Section 4.2.5 we introduce a loss function that is used in Section 4.2.6 to explain the connection between Levenberg-Marquardt method and EKI and in Section 4.2.7 to define the adaptive regularisation of EKI. We start with Section 4.2.1 which describes the idea of EKI.

4.2.1 *Overview*

We go back to the assumption that the noise in (4.1.1) is Gaussian $\eta \sim \mathcal{N}(0, B)$ and so the Likelihood for k given the observation $\mathbf{V} = \mathbf{v}$ is given by (4.1.2).

Furthermore, let $\pi_0(k)$ denote the prior density with respect to Lebesgue measure again. From Bayes formula (4.1.5) we have

$$\pi^k(\mathbf{v}) = \frac{\rho(\mathbf{v} - \mathcal{F}(k))\pi_0(k)}{\int_{\mathcal{K}} \rho(\mathbf{v} - \mathcal{F}(k))\pi_0(k)\mathrm{d}k} \propto \rho(v - \mathcal{F}(k))\pi_0(k),$$

and using the bespoken Gaussianity of Likelihood yields

$$\pi^k(\mathbf{v}) = \frac{1}{Z} \exp\left(-\frac{1}{2}\|\mathbf{v} - \mathcal{F}(k)\|_B^2\right) \cdot \pi_0(k), \tag{4.2.1}$$

with normalising factor Z given by

$$Z = \int_{\mathcal{K}} \exp\left(-\frac{1}{2}\|\mathbf{v} - \mathcal{F}(k)\|_B^2\right) \cdot \pi_0(k)\mathrm{d}k. \tag{4.2.2}$$

In (4.2.1) the posterior is known up to the constant Z which involves integration over $\mathcal{K}$. Considering the high dimensionality of the parameter space - in three-dimensional MRE the dimension of $\mathcal{K}$ can be of order 10^8 and greater - common quadrature methods are computationally intractable.

The Bayesian approach used in this work is EKI which is a derivative-free iterative regularisation algorithm that produces a sequence of Gaussian distributions approximating the posterior in (4.2.1) [40].

Being regularised iteratively means that regularisation is chosen while inverting which is different from approaches where regularisation is chosen before inverting, for example Tikhonov regularisation in deterministic inversion. Iterative regularisation is powerful because convergence speed is often coupled

with regularisation. Iterative regularisation algorithms can be efficient and accurate at the same time.

Many iterative deterministic inversion approaches (Levenberg-Marquardt, Newton) require Fréchet-derivatives and their associate adjoint of the (nonlinear) function $\mathcal{F}$ [40]. EKI, the iterative Bayesian approach used in this thesis, is derivative free and it is not necessary to numerically approximate derivatives or adjoints which demonstrates another advantage of some Bayesian inversion over deterministic inversion. Being derivative free can be handy when working with commercial software where only the forward solution $\mathcal{F}(k)$ but no derivatives are provided.

An underlying key idea [83] in the development of EKI is to avoid discretisation of $\mathcal{K}$ until the last possible moment. This results in a discretisation invariant model, i.e. the way how we determine the posterior is the same for all discretisations and is not changing under grid-refinement.

4.2.2 *Tempering Approach*

Rather than approximating the posterior in (4.2.1) in one big leap EKI follows the idea of tempering approaches [10] and defines a sequence of intermedi-

ate posterior distributions $\pi_0^k, \pi_1^k, \ldots, \pi_{N+1}^k$ that evolve from the prior to the posterior as

$$\pi_{n+1}^k(\mathbf{v}) = \frac{1}{Z_{n+1}} \exp\left(-\frac{\phi_{n+1}}{2}\|\mathbf{v} - \mathcal{F}(k)\|_B^2\right) \cdot \pi_0(k), \qquad n = 0, 1, \ldots N, \tag{4.2.3}$$

with

$$Z_{n+1} = \int_{\mathcal{K}} \exp\left(-\frac{\phi_{n+1}}{2}\|\mathbf{v} - \mathcal{F}(k)\|_B^2\right) \cdot \pi_0(k)\mathrm{d}k, \tag{4.2.4}$$

and tempering parameter $\{\phi_n\}_{n=1}^{N+1}$ satisfying

$$\phi_0 := 0 < \phi_1 < \phi_2 < \cdots < \phi_N < 1 =: \phi_{N+1}.$$

It is clear from the definition (4.2.3) that the first iterate is the prior and that the last iterate $\pi_{N+1}^k = \pi^k$ is the posterior. The bridging parameter ϕ_n reduces the ”peakiness” of the Likelihood and makes it flatter with larger variance in the first iterates. This can be seen in the definition (4.2.3). So, the tempering approach gradually includes more certainty about the data into the intermediate posterior distributions. Figure 4.2.1 illustrates the idea of tempering in a simple one-dimensional example and it can be seen how the intermediate posterior distributions evolve from the prior to the posterior.

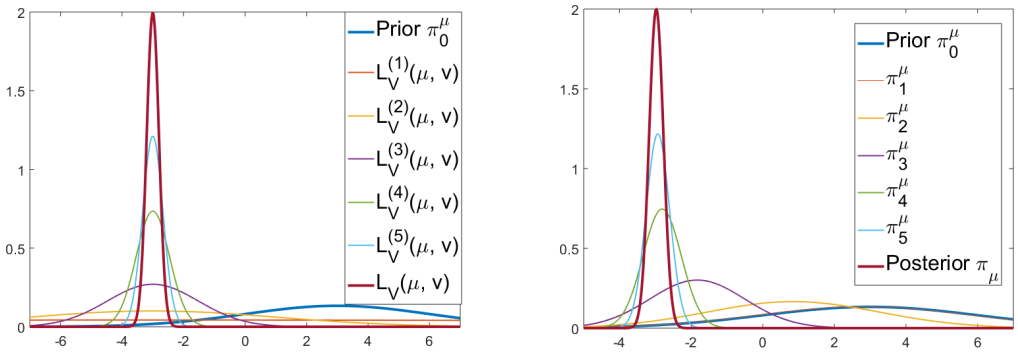


Figure 4.2.1: Illustration of tempering in a one-dimensional case. **Left:** Prior $\pi_0(k)$ and intermediate likelihoods $L_V^{(n)} \propto \exp\left(-\frac{\phi_n}{2}\|\mathbf{v} - \mathcal{F}(k)\|_B^2\right)$ for some bridging parameter ϕ_n have larger variance and are "flatter" than likelihood $L_V \propto \exp\left(-\frac{1}{2}\|\mathbf{v} - \mathcal{F}(k)\|_B^2\right)$. **Right:** The intermediate posterior distributions $\pi_{n+1}^k \propto L_V^{(n+1)} \cdot \pi_0(k)$ evolve from the prior to the posterior. In this illustration of tempering the six bridging parameters $\varphi_0, \ldots, \varphi_5$ split the interval between 0 and 1 into equi-distant intervals. This choice is only used in this example. The choice of tempering parameters for EKI is later discussed in Section 4.2.7.

The tempering distributions (4.2.3) can be transformed into an iteration:

$$\begin{aligned}
\pi_{n+1}^k(\mathbf{v}) &= \frac{1}{Z_{n+1}} \exp\left(-\frac{\phi_{n+1}}{2}\|\mathbf{v}-\mathcal{F}(k)\|_B^2\right)\cdot \pi_0(k) \\
&= \frac{Z_n}{Z_{n+1}}\cdot\frac{1}{Z_n}\exp\left(\left(-\frac{\phi_{n+1}-\phi_n}{2}-\frac{\phi_n}{2}\right)\cdot\|\mathbf{v}-\mathcal{F}(k)\|_B^2\right)\cdot\pi_0(k) \\
&= \frac{Z_n}{Z_{n+1}}\exp\left(-\frac{\phi_{n+1}-\phi_n}{2}\cdot\|\mathbf{v}-\mathcal{F}(k)\|_B^2\right) \\
&\qquad\qquad \cdot\frac{1}{Z_n}\exp\left(-\frac{\phi_n}{2}\cdot\|\mathbf{v}-\mathcal{F}(k)\|_B^2\right)\cdot\pi_0(k) \\
&= \frac{Z_n}{Z_{n+1}}\exp\left(-\frac{\phi_{n+1}-\phi_n}{2}\cdot\|\mathbf{v}-\mathcal{F}(k)\|_B^2\right)\cdot\pi_{n-1}^k(\mathbf{v}) \\
&= \frac{Z_n}{Z_{n+1}}\exp\left(-\frac{\alpha_n^{-1}}{2}\cdot\|\mathbf{v}-\mathcal{F}(k)\|_B^2\right)\cdot\pi_{n-1}^k(\mathbf{v}) \\
&= \frac{Z_n}{Z_{n+1}}\exp\left(-\frac{1}{2}\cdot\|\mathbf{v}-\mathcal{F}(k)\|_{(\alpha_n\cdot B)}^2\right)\cdot\pi_n^k(\mathbf{v}),
\end{aligned} \tag{4.2.5}$$

where we have defined

$$\alpha_n^{-1} = \phi_{n+1}-\phi_n. \tag{4.2.6}$$

4.2.3 *Gaussianisation and Linearisation*

A central idea of EKI is to approximate each distribution in the iterative scheme (4.2.5) by a Gaussian distribution with density φ_n^k [39]. This idea can be implemented by choosing a Gaussian prior and by linearising $\mathcal{F}$ in the Likelihood. So far, $\mathcal{F}$ is non-linear and therefore, the Likelihood in (4.2.5) is non-Gaussian.

Let us assume the density with respect to Lebesgue measure of the Gaussian approximation φ_n^k has mean $\mathfrak{m}_n$ and covariance Σ_n , i.e. φ_n^k is the density of some random variable $k_n \sim \mathcal{N}(\mathfrak{m}_n, \Sigma_n)$ and we have $\varphi_n^k \propto \exp\left(-\|k - \mathfrak{m}_n\|^2_{\Sigma_n}\right)$. By φ_0 we denote the density of the chosen Gaussian approximation of the prior $\pi_0(k)$ which we assume has mean m_0 and covariance operator Σ_0.

These quantitative approximations are useful as Gaussian densities are fully characterised by only the mean and the covariance. Later in Section 4.3, we will see that they are easy to construct and versatile, which will allow us to code our prior knowledge into the algorithm.

We linearise $\mathcal{F}$ in each iteration around the mean of φ_n^k, so [39]

$$\mathcal{F}(k) \approx \mathcal{F}(\mathfrak{m}_n) + \mathrm{D}\mathcal{F}(\mathfrak{m}_n)(k - \mathfrak{m}_n) = \mathcal{F}_n + \mathrm{D}\mathcal{F}_n(k - \mathfrak{m}_n), \tag{4.2.7}$$

where $\mathrm{D}\mathcal{F}(\mathfrak{m}_n) : \mathcal{K} \to \mathbb{R}^m$ denotes the Frechet derivative of $\mathcal{F}$ evaluated at the mean $\mathfrak{m}_n$. We use the notation $\mathcal{F}_n := \mathcal{F}(\mathfrak{m}_n)$ and $\mathrm{D}\mathcal{F}_n := \mathrm{D}\mathcal{F}(\mathfrak{m}_n)$.

Starting the iteration with a Gaussian approximation φ_0^k of the prior, we can express the update (4.2.5) as product of two Gaussian density functions for k if we use linearisation (4.2.7). The resulting sequence of Gaussian approximations of (4.2.5) given by [39]

$$\varphi_{n+1}^k(\mathbf{v}) = \frac{\tilde{Z}_n}{\tilde{Z}_{n+1}} \exp\left(-\frac{1}{2} \cdot \|\mathbf{v} - \mathcal{F}_n - \mathrm{D}\mathcal{F}_n(k - \mathfrak{m}_n)\|^2_{(\alpha_n \cdot B)}\right) \cdot \varphi_n^k(\mathbf{v}), \tag{4.2.8}$$

where

$$\tilde{Z}_n = \int_{\mathcal{K}} \exp\left(-\frac{\phi_n}{2}\|\mathbf{v} - \mathcal{F}_n - \mathrm{D}\mathcal{F}_n(k - \mathfrak{m}_n)\|^2_{(\alpha_n \cdot B)}\right) \cdot \varphi_0 \mathrm{d}k.$$

In (4.2.8) using $\varphi_n^k \propto \exp\left(-\|k - \mathfrak{m}_n\|^2_{\Sigma_n}\right)$ enables us to calculate the mean and the covariance of the product of these two Gaussian densities by adding the two inner products in the exponent. This calculation called completing the square can be found, for example, in [83]. The product $\varphi_{n+1}^k \propto \exp\left(-\|k - \mathfrak{m}_{n+1}\|^2_{\Sigma_{n+1}}\right)$ has mean and covariance given by

$$\begin{aligned}
\mathfrak{m}_{n+1} &= \mathfrak{m}_n + \Sigma_n \mathrm{D}\mathcal{F}_n^* \left(\mathrm{D}\mathcal{F}_n \Sigma_n \mathrm{D}\mathcal{F}_n^* + \alpha_n B\right)^{-1} (\mathbf{v} - \mathcal{F}_n), \\
\Sigma_{n+1} &= \Sigma_n + \Sigma_n \mathrm{D}\mathcal{F}_n^* \left(\mathrm{D}\mathcal{F}_n \Sigma_n \mathrm{D}\mathcal{F}_n^* + \alpha_n B\right)^{-1} \mathrm{D}\mathcal{F}_n \Sigma_n.
\end{aligned} \tag{4.2.9}$$

We used $\mathrm{D}\mathcal{F}_n^*$ which denotes the adjoint of $\mathrm{D}\mathcal{F}_n$ in $\mathfrak{m}_n$.

From (4.2.7) we conclude that the expected value is approximated by [39]

$$\begin{aligned}
\mathbb{E}_n\left[\mathcal{F}(k_n)\right] &\approx \mathbb{E}_n\left[\mathcal{F}_n\right] + \mathrm{D}\mathcal{F}_n(\mathbb{E}_n\left[k_n - \mathfrak{m}_n\right]) \\
&= \mathcal{F}_n + \mathrm{D}\mathcal{F}_n(\mathbb{E}_n\left[k_n\right] - \mathbb{E}_n\left[\mathfrak{m}_n\right]) \\
&= \mathcal{F}_n,
\end{aligned} \tag{4.2.10}$$

where $\mathbb{E}_n$ denotes the expected value with respect to the Gaussian measure $\mathcal{N}(\mathfrak{m}_n, \Sigma_n)$. Using this approximation and (4.2.7) again we also conclude [39]

$$\begin{aligned}
\mathcal{F}(k_n) - \mathbb{E}_n\left[\mathcal{F}(k_n)\right] &\approx \mathcal{F}(k_n) - \mathcal{F}_n \\
&\approx \mathrm{D}\mathcal{F}_n\left(k_n - \mathfrak{m}_n\right).
\end{aligned} \tag{4.2.11}$$

These two approximations can be used to replace derivatives in (4.2.9). From the definition of the covariance we know that [39]

$$\Sigma_n = \mathbb{E}_n\left[(k_n - \mathbb{E}_n(k_n)) \otimes (k_n - \mathbb{E}_n(k_n))\right] = \mathbb{E}_n\left[(k_n - \mathfrak{m}_n) \otimes (k_n - \mathfrak{m}_n)\right],$$

where $\otimes$ denotes the outer product. We can now conclude using (4.2.11) and the linearity of the expectation $\mathbb{E}_n$ that a cross-covariance can be used to approximate [36]

$$\begin{aligned}\Sigma_n D\mathcal{F}_n^* &= \mathbb{E}_n\left[(k_n - \mathfrak{m}_n) \otimes (k_n - \mathfrak{m}_n)\right] \cdot D\mathcal{F}_n^* \\ &= \mathbb{E}_n\left[(k_n - \mathfrak{m}_n) \otimes (D\mathcal{F}_n(k_n - \mathfrak{m}_n))\right] \\ &\approx \mathbb{E}_n\left[(k_n - \mathfrak{m}_n) \otimes (\mathcal{F}(k_n) - \mathbb{E}_n\left[\mathcal{F}(k_n)\right])\right] =: \mathrm{Cov}_n(k_n, \mathcal{F}(k_n)),\end{aligned} \tag{4.2.12}$$

and a covariance to approximate [36]

$$\begin{aligned}D\mathcal{F}_n\Sigma_n D\mathcal{F}_n^* &= D\mathcal{F}_n\mathbb{E}_n\left[(k_n - \mathfrak{m}_n) \otimes (k_n - \mathfrak{m}_n)\right] \cdot D\mathcal{F}_n^* \\ &= \mathbb{E}_n\left[(D\mathcal{F}_n(k_n - \mathfrak{m}_n)) \otimes (D\mathcal{F}_n(k_n - \mathfrak{m}_n))\right] \\ &\approx \mathbb{E}_n\left[(\mathcal{F}(k_n) - \mathbb{E}_n\left[\mathcal{F}(k_n)\right]) \otimes (\mathcal{F}(k_n) - \mathbb{E}_n\left[\mathcal{F}(k_n)\right])\right] =: \mathrm{Cov}_n(\mathcal{F}(k_n)).\end{aligned} \tag{4.2.13}$$

Thus, by substituting (4.2.13) for $D\mathcal{F}_n\Sigma_n D\mathcal{F}_n^*$ and (4.2.12) for $\Sigma_n D\mathcal{F}_n^*$ in (4.2.9), we can approximate the updates of the mean and covariance free of derivatives in the following way [36]

$$\begin{aligned}\mathfrak{m}_{n+1} &= \mathfrak{m}_n + \mathrm{Cov}_n(k_n, \mathcal{F}(k_n))\left(\mathrm{Cov}_n(\mathcal{F}(k_n)) + \alpha_n B\right)^{-1}\left(\mathbf{v} - \mathbb{E}_n\left[\mathcal{F}(k_n)\right]\right),\\ \Sigma_{n+1} &= \Sigma_n + \mathrm{Cov}_n(k_n, \mathcal{F}(k_n))\left(\mathrm{Cov}_n(\mathcal{F}(k_n)) + \alpha_n B\right)^{-1}\mathrm{Cov}_n(k_n, \mathcal{F}(k_n)).\end{aligned} \tag{4.2.14}$$

4.2.4 *Particle Approximation*

Although, (4.2.14) is derivative free both (4.2.12) and (4.2.13) cannot be calculated analytically. We use particle approximations, i.e. each k_n is approximated by [39]

$$k_n^J = \frac{1}{J}\sum_{j=1}^{J}\delta_{k_n^{(j)}}, \quad k_n^{(j)} \sim \mathcal{N}(\mathfrak{m}_n, \Sigma_n),$$

where we used the Dirac delta measure δ. The idea of using this particle approximation now is to update particles $\{k_n^{(j)}\}_{j=1,\dots,J}$, $J \in \mathbb{N}$ in such way that $\{k_{n+1}^{(j)}\}_{j=1,\dots,J}$ is distributed according to $\mathcal{N}(\mathfrak{m}_{n+1}, \Sigma_{n+1})$. Because of the Gaussianisation and the linearisation that we have already introduced this is straightforward. We update particles in the following way

$$k_{n+1}^{(j)} = k_n^{(j)} + \mathcal{C}_n^{k\mathcal{F}} \left(\mathcal{C}_n^{\mathcal{F}\mathcal{F}} + \alpha_n B \right)^{-1} \left(\mathbf{v} - \sqrt{\alpha_n} \xi_n - \mathcal{F}(k_n^{(j)}) \right), \tag{4.2.15}$$

where $\xi_n \sim \mathcal{N}(0, B)$ and we use the sample covariance and cross-covariance defined by

$$\mathcal{C}_n^{\mathcal{F}\mathcal{F}} = \frac{1}{J} \sum_{j=1}^{J} \left(\mathcal{F}(k_n^{(j)}) - \overline{\mathcal{F}}_n \right) \otimes \left(\mathcal{F}(k_n^{(j)}) - \overline{\mathcal{F}}_n \right) \tag{4.2.16}$$

$$\mathcal{C}_n^{k\mathcal{F}} = \frac{1}{J} \sum_{j=1}^{J} \left(k_n^{(j)} - \overline{k_n} \right) \otimes \left(\mathcal{F}(k_n^{(j)}) - \overline{\mathcal{F}}_n \right), \tag{4.2.17}$$

which uses the two definition of

$$\overline{k_n} = \frac{1}{J} \sum_{j=1}^{J} k_n^{(j)}, \quad \overline{\mathcal{F}}_n = \frac{1}{J} \sum_{j=1}^{J} \mathcal{F}(k_n^{(j)}). \tag{4.2.18}$$

We define the ensemble covariance

$$\mathcal{C}_n^{kk} = \frac{1}{J} \sum_{j=1}^{J} \left(k_n^{(j)} - \overline{k_n} \right) \otimes \left(k_n^{(j)} - \overline{k_n} \right). \tag{4.2.19}$$

It can be shown (informally or rigorously in finite dimensions in [58]) that the mean and covariance of the ensemble updated in (4.2.15) is converging to $\overline{k_{n+1}} \to \mathfrak{m}_{n+1}$ and $\mathcal{C}_n^{kk} \to \Sigma_{n+1}$ for $J \to \infty$ [39]. Hence the particle update introduced in (4.2.15) is actually approximating (4.2.14) which updates mean and covariance of the Gaussian approximation of the sequence π_n^k.

4.2.5 *Loss Function*

We deploy the negative log-likelihood (4.1.2) as loss function $\mathcal{L} : \mathcal{K} \times \mathbb{R}^m \to \mathbb{R}^+$, so [40]

$$\mathcal{L}(k, \mathbf{v}) := -\log L_V(k; \mathbf{v}) = \frac{1}{2}\|\mathbf{v} - \mathcal{F}(k)\|_B^2. \tag{4.2.20}$$

This function is used in application to monitor the progress of the ensemble mean $\bar{k}_n$ in EKI. From the fact that this is the negative log-likelihood we conclude that small values correspond to a high likelihood. Furthermore, the loss function is used in order to update regularisation parameter α_n which will be explained below. Also, it can be considered as objective function when putting EKI in the context of deterministic inversion with the Levenberg-Marquardt approach which is explained in the following section [39].

4.2.6 *Connection to Levenberg-Marquardt*

Although, we consider EKI from the Bayesian viewpoint in this thesis relating it to classic inversion can provide interesting insights. It can be shown [39] that $\mathfrak{m}_{n+1}$ in (4.2.9) is the minimiser of

$$\mathfrak{m}_{n+1} = \arg\min_{k \in \mathcal{K}} \|\mathbf{v} - \mathcal{F} - \mathrm{D}\mathcal{F}_n\,(k - \mathfrak{m}_n)\|_B^2 + \alpha_n \|\Sigma_n^{-1/2}\,(k - \mathfrak{m}_n)\|_{\mathcal{K}}^2. \tag{4.2.21}$$

If the infinite-dimensional space of the parameters $\mathcal{K}$ is finite-dimensional and if Σ_n is the identity operator and B a diagonal matrix, then (4.2.21) is the regularised optimisation task that iteratively replaces the least-squares functional (4.2.20) in the Levenberg-Marquardt approach [29]. The resulting iterative scheme is (4.2.9). Due to the fact that the ensemble mean is converging to $\overline{k_{n+1}} \to \mathfrak{m}_{n+1}$ the update for the ensemble mean which is given by

$$\overline{k_{n+1}} = \overline{k_n} + \mathcal{C}_n^{k\mathcal{F}} \left(\mathcal{C}_n^{\mathcal{F}\mathcal{F}} + \alpha_n B \right)^{-1} \left(\mathbf{v} - \sqrt{\alpha_n}\overline{\xi_n} - \mathcal{F}(\overline{k_n}) \right), \tag{4.2.22}$$

can be interpreted as derivative free approximation of the Levenberg-Marquardt scheme and can be therefore also be interpreted as derivative-free minimiser for the linear least-square functional (4.1.3). Convergence of Levenberg-Marquardt is well understood [94], convergence of EKI is an open problem. Therefore, the viewpoint of deterministic inversion is useful to explain convergence of EKI. In this context α_n can be interpreted as Tikhonov regularisation parameter and (4.2.21) can be interpreted as Tikhonov regularisation applied to the linearisation of $\mathcal{F}$ [38].

4.2.7 *Regularisation*

In this work, we continue considering EKI from a Bayesian viewpoint and α_n is not motivated by Levenberg-Marquardt regularisation but by some statistical discrepancy principle. We define α_n to be

$$\alpha_0 = \max\left\{\frac{2}{m}\cdot\Xi(\{k_0^{(j)}\}_{j=1,\dots,J},\mathbf{v}),1\right\}, \tag{4.2.23}$$

and

$$\alpha_n = \max\left\{\frac{2}{m}\cdot\Xi(\{k_n^{(j)}\}_{j=1,\dots,J},\mathbf{v}),\left[1-\sum_{j=1}^{n-1}\alpha_j^{-1}\right]^{-1}\right\}, \quad n>0, \tag{4.2.24}$$

where the definition of the average data misfit $\Xi : \mathcal{K}^J \times \mathbb{R}^m \to \mathbb{R}^+$ is used which is defined by

$$\Xi(\{k^{(j)}\}_{j=1,\dots,J},\mathbf{v}) = \frac{1}{J}\sum_{j=1}^{J}\mathcal{L}(k^{(j)},\mathbf{v}).$$

This adaptive choice of regularisation is introduced in [39] and does not need any additional tuning parameters. Also in [39] the authors show that this parameter makes EKI efficient and accurate and it comes with a natural stopping rule. EKI is stopped after N_S iterations if

$$\sum_{n=0}^{N_S}\alpha_n^{-1} = 1. \tag{4.2.25}$$

We can see that this definition is consistent with (4.2.6).

This choice of regularisation is motivated in [39] by requiring the "difference" between two consecutive distributions updates of the intermediate distributions $\pi_0^k, \pi_1^k, \ldots, \pi_{N+1}^k$ to be bounded above which ensure a certain level of smoothness of the transition. In order to define an appropriate upper limit for change of the distribution from one iterate to the next in [39] the authors apply Morozov's discrepancy principle [13].

It is worth noting that there are other ways to adaptively choose the tempering parameter in EKI [38].

We summarise the EKI method with adaptive regularisation in Algorithm 1.

Algorithm 1: Regularising Ensemble Kalman Algorithm

Require: $\mathcal{K}$ Hilbert Space, model $\mathcal{F} : \mathcal{K} \to \mathbb{R}^m$

Input: Initial ensemble $\{k_0^{(j)}\}_{j=1}^J$; error covariance Γ; data $\mathbf{v} \in \mathbb{R}^m$.

Output: Posterior Ensemble $\{k_n^{(j)}\}_{j=1}^J$

Set $\alpha_0 = 0$

repeat

1. **Prediction Step.** Define $\overline{\mathcal{F}}_n = \frac{1}{J}\sum_{j=1}^J \mathcal{F}_n^{(j)}$ where

$$\mathcal{F}_n^{(j)} = \mathcal{F}(k_n^{(j)}), \quad j \in \{1, \dots, J\}.$$

2. **Update Regularisation Parameter** α_{n+1}**.** Set

$$\alpha_{n+1}^* = \frac{1}{m}\frac{1}{J}\sum_{j=1}^J \|B^{-1/2}(\mathbf{v} - \mathcal{F}(k))\|^2.$$

if $\sum_{i=0}^n \alpha_i^{-1} + \alpha_{n+1}^{*-1} \geq 1$. **then** $\alpha_{n+1} = 1/\left(1 - \sum_{i=0}^n \alpha_i^{-1}\right)$ **else** $\alpha_{n+1} = \alpha_{n+1}^*$.

3. **Analysis Step.** Compute empirical covariance and cross-covariances

$$\mathcal{C}_n^{\mathcal{F}\mathcal{F}} = \frac{1}{J}\sum_{j=1}^J \left(\mathcal{F}(k_n^{(j)}) - \overline{\mathcal{F}}_n\right) \otimes \left(\mathcal{F}(k_n^{(j)}) - \overline{\mathcal{F}}_n\right)$$
$$\mathcal{C}_n^{k\mathcal{F}} = \frac{1}{J}\sum_{j=1}^J \left(k_n^{(j)} - \overline{k_n}\right) \otimes \left(\mathcal{F}(k_n^{(j)}) - \overline{\mathcal{F}}_n\right),$$

Update the ensemble:

$$k_{n+1}^{(j)} = k_n^{(j)} + \mathcal{C}_n^{k\mathcal{F}}(\mathcal{C}_n^{(\mathcal{F}\mathcal{F})} + \alpha_n B)^{-1}(\mathbf{v}^m + \sqrt{\alpha}\xi_n^{(j)} - \mathcal{F}(k_n^{(j)})).$$

with $\xi_n^{(j)} \sim \mathcal{N}(0, B)$.

$n + 1 \to n$

until $\sum_{i=0,\dots n+1} \alpha_i^{-1} = 1$;

4.3 PRIOR

The linear Gaussian update (4.2.15) together with regularisation α_n transforms a set of J particles into a new one until the stopping rule applies and a Gaussian approximation of the posterior is found. The user has to choose an informative prior to draw the initial set of particles from. This is an essential step in Ensemble Kalman and Bayesian inversion in general. Adequate priors improve accuracy of the estimate provided by EKI and improve convergence in EKI.

We draw this initial ensemble from some prior $\mathcal{N}(m_0, \Sigma_0)$ with some mean m_0 and covariance Σ_0 which comprises prior knowledge of the unknown. We can use the mean m_0 to inform our algorithms about prior knowledge about the values of the mechanical parameters and Σ_0 to incorporate knowledge about the regularity and the structural distribution of the mechanical parameters across the domain of interest. As Gaussian distributions are characterised by the mean and the covariance only, it is essential to define them in a way that allows the incorporation of prior knowledge.

4.3.1 *Definition of Prior Covariance and Mean*

For this purpose we use an Whittle-Matérn autocorrelation function defined by ([73], [82])

$$\mathrm{ACF}(x) = \frac{2^{1-r}}{\Gamma(r)} \left(\frac{|x|}{l} \right)^r K_r \left(\frac{|x|}{l} \right), \tag{4.3.1}$$

for $x \in \mathbb{R}^2$, where Γ is the gamma function, l is the characteristic length scale and K_r is the modified Bessel function of the second kind of order r.

It can be shown that [82]

$$\mathrm{Cov}(x_1, x_2) := \sigma^2 \cdot \mathrm{ACF}(x_1 - x_2), \quad x_1, x_2 \in \mathbb{R}^2, \tag{4.3.2}$$

defines a covariance operator. Here σ^2 is an amplitude scale.

The Whittle-Matérn autocorrelation function gives us the opportunity to design a wide range of Gaussian priors. We use this covariance operator to define our prior $\mathcal{N}(m_0, \Sigma_0)$. More specifically let $\Sigma_0 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ be the covariance operator defined by $\Sigma_0(x_1, x_2) = \mathrm{Cov}(x_1, x_2)$ for $x_1, x_2 \in \mathbb{R}^2$.

In order to draw samples from the prior we use Karhunen-Loève expansion ([87], [67]).In Section 4.5 we discuss another way of drawing samples from Gaussian distributions with Whittle-Matérn covariances. Using the computationally more expensive Karhunen-Loève expansion allows for non-integer orders r.

Let $(\lambda_{E_{(i)}}, \psi_{(i)})$ be the corresponding eigenpair of Σ that is

$$\Sigma_0 \psi_{(i)} = \lambda_{E_{(i)}} \psi_{(i)}. \tag{4.3.3}$$

Using the Karhunen-Loève expansion ([87], [67]), samples $k_0^{(j)}$ can be drawn from $\mathcal{N}(m_0, \Sigma_0)$ by using the following formula

$$k_0^{(j)} = m_0 + \sum_{i=1}^{\infty} \lambda_{E_{(i)}}^{1/2} \psi_{(i)} r^{(j)}, \tag{4.3.4}$$

with $r^{(j)} \sim \mathcal{N}(0, 1)$.

In the experiments in Chapter 5 we use the Karhunen-Loève expansion in order to define our prior sample. More precisely we choose points $\{\hat{x}_i\}_{i=1}^{\hat{n}} \in \Omega$ corresponding to the chosen computational mesh of the discretisation scheme. After choosing a mean m_0 and hyperparameters l, r and σ^2 of the autocorrelation function (4.3.1) we can define a covariance matrix using (4.3.2) and produce prior samples from (4.3.4).

4.3.2 *Parameters of Prior*

The parameters of the autocorrelation function (4.3.1) and the mean m_0 can now be used to form informative priors. Firstly, it is worth noting that the covariance operator (4.3.2) is a function of the distance $|x_1 - x_2|$. Therefore, a value of a sample drawn from $\mathcal{N}(m_0, \Sigma_0)$ is related to nearby values. The statistical

relationship of values decreases with increasing distance. The definition of 'nearby' in this context is determined by the length scale l. This length scale should be chosen such that areas of similar values in a sample drawn from $\mathcal{N}(m_0, \Sigma_0)$ form structures of similar size to the size of the structures we expect to explore in the tissue (cancer, fibrosis). In Figure 4.3.1 it can be seen how l affects the samples.

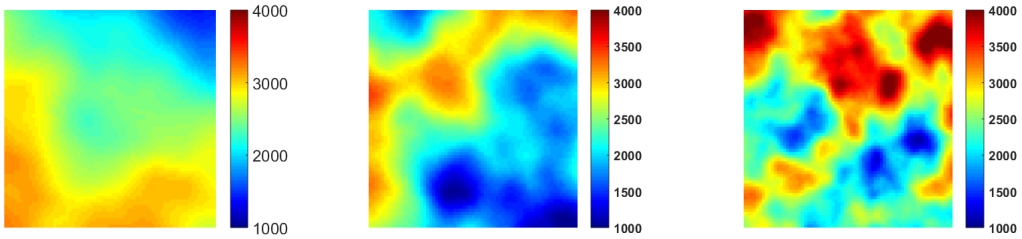


Figure 4.3.1: Samples drawn from $\mathcal{N}(0, \Sigma_0)$ with Σ_0 being a covariance operator that arises from the auto-correlation function (4.3.1) with $r = 2$ and different length scales l. **Left:** $l = 1/2$. **Mid:** $l = 1/4$. **Right:** $l = 1/8$

The parameter r in (4.3.1) controls the smoothness of the samples and can be used to further tune the occurring shapes and structures in the samples drawn from the prior such that they are similar to the expected unknown in size and smoothness. The amplitude scale σ^2 in (4.3.1) should be chosen such that samples drawn from $\mathcal{N}(m_0, \Sigma_0)$ cover the range of values the unknown k has.

Furthermore, the mean in (4.3.4) can be also used to inform the algorithm about expected values of the unknown. Using the result of other inversion approaches like direct inversion as (non-homogeneous) mean in the prior is a way to easily combine EKI with other approaches. Also, a function describing the elastic parameter of a healthy organ can be used as mean to form an informative prior for detection of diseased tissue within the organ.

There are other correlation functions that can be used to define useful covariances. The correlation function $\mathrm{ACF}(x) = \exp(-|x|^2/l^2)$ tends to be too smooth for the experiments run in Chapter 5.

In Figure 4.3.2 it can be seen how different priors affect the result of EKI. The size of the structures in Figure 4.3.1 match the size of the structures in the ground truth in Figure 3.2.2 for length scale $l = 1/4$. Indeed, the results of EKI is the best for this length scale in the prior.

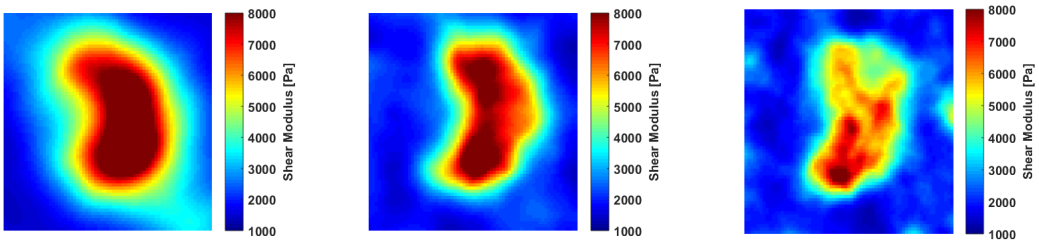


Figure 4.3.2: Inversions results for parameter k in Figure 3.2.2 of EKI using priors with different length scales l in Figure 4.3.1. The detailed problem description can be found in Section 5.2 and 5.4. **Left:** $l = 1/8$. **Mid:** $l = 1/4$. **Right:** $l = 1/2$

4.4 LEVEL-SET PARAMETRISATION

In this section we consider a class of level-set-based parametrisations $\mathcal{P}$ that we have already considered in the formulation of the inverse problem in Section 3.4. We call this parametrisation $\mathcal{P}_1 : \mathcal{K} \to \mathcal{M}$ and will introduce in Section 4.5 another class of parametrisation. Particles drawn from a Gaussian prior are continuous. EKI updates $k_n^{(j)} \mapsto k_{n+1}^{(j)}$ in (4.2.15) retain the continuity of the particles over the domain Ω after an ensemble update and pass on the regularity to the next iterate. Also, the fact that we are often reporting the sample mean $\overline{k_n}$ as the result of the inversion implies a high level of regularity in the outcome of our algorithm.

The level-set parametrisation [37] introduced in this section allows characterising discontinuous properties. This can be considered as a further way to incorporate prior knowledge into the Bayesian approach because the knowledge about discontinuities in the parameter is usually available prior to the inversion. For example, it is known that diseases like cancer and cirrhosis have relatively sharp interfaces between the stiffness of diseased and healthy tissue [59]. So, if the presence of these diseases is investigated, this prior knowledge about the nature of these diseases can be coded into the inversion using parametrisation $\mathcal{P}_1$.

We will only use this parametrisation in the context with the purely elastic model described in Section 3.2.1, so $\mu_L = 0$ and $\mu = \mu_S$ and introduce this parametrisation only for this case and do not parametrise the loss modulus. In order to define $\mathcal{P}_1$ we assume we have prior knowledge about the geometric structure of the data and incorporate that knowledge into a function $\mu \in L^\infty(\Omega)$ of the form

$$\mu(x) = \sum_{i=1}^{\mathfrak{l}} k_{\mathcal{S}}^{(i)} \mathbb{1}_{D^{(i)}}(x). \tag{4.4.1}$$

Here, $\{k_{\mathcal{S}}^{(i)}\}_{i=1}^{\mathfrak{l}}$ is a set of known constants in $\mathbb{R}$, $\mathbb{1}_D$ is the indicator function in Ω and $\{D^{(i)}\}_{i=1}^{\mathfrak{l}}$ is a partition of Ω defined by

$$D^{(i)} = \{x \in \Omega | c^{(i-1)} \leq k(x) < c^{(i)}\},$$

where $c^{(i)} \in \overline{\mathbb{R}}$ are constant thresholds with $-\infty = c^{(0)} < c^{(1)} < \cdots < c^{(\mathfrak{l})} = \infty$ for some $k \in C(\Omega, \mathbb{R})$. We define the level set map $\mathcal{P}_1 : \mathcal{K} \to \mathcal{M}$ which maps

$$(\mathcal{P}_1(k))(x) \mapsto \mu(x). \tag{4.4.2}$$

Note that the level set map (4.4.2) is discontinuous because (4.4.1) is and therefore also $\mathcal{G}$ is discontinuous.

Now, in (4.1.1) we consider

$$\mathcal{F} : \mathcal{K} \to \mathbb{R}^m, \quad k \mapsto \mathcal{F}(k) := (\mathcal{G} \circ \mathcal{P}_1)(k),$$

as forward map and solve the inverse problem of approximating $k \in \mathcal{K}$ for measurement $\mathbf{v} \in \mathbb{R}^m$ given by

$$\mathbf{v} = \mathcal{F}(k) + \eta. \tag{4.4.3}$$

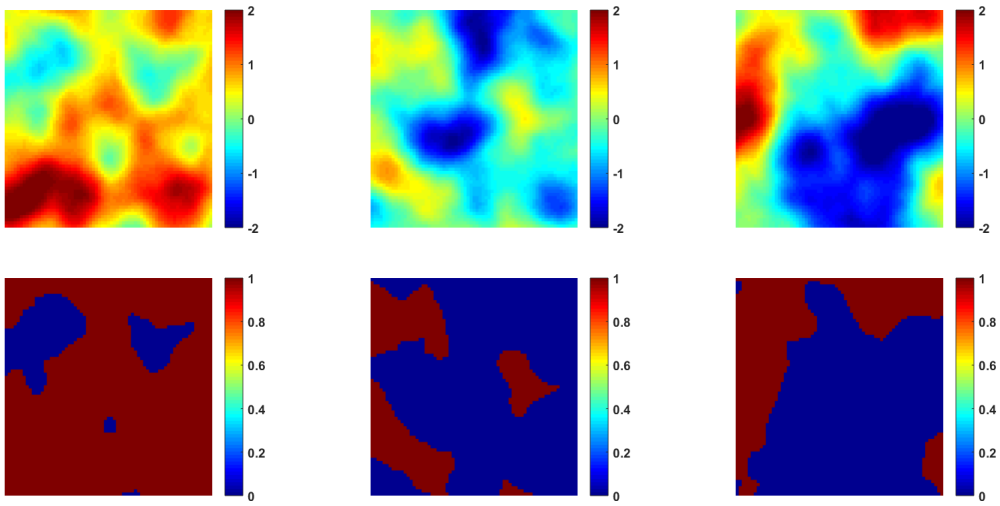


Figure 4.4.1: Illustration of level set map L. **Top:** Centred Gaussian random fields k. **Bottom:** Indicator function $\mathbb{1}_D$ of level set map $\mathcal{P}_1(k)$ with constants $c_0 = -\infty$, $c_1 = 0$ and $c_2 = \infty$ defining D.

Figure 4.4.1 illustrates how the level set map $\mathcal{P}_1$ enforces discontinuities using Gaussian random fields.

In Section 5.6 we give a detailed explanation how we use the level set parametrisation in MRE in order to parametrise the unknown geometry of shear modulus μ.

4.5 TWO-LEVEL-PARAMETRISATION

In this section we introduce a two-level parametrisation $\mathcal{P}^{(2)}$ and improve the level-set parametrisation of the unknown introduced in Section 4.4. Level $\mathcal{P}_2^{(2)}$ of the two-level parametrisation is a modified version of the level-set parametrisation introduced in Section 4.4. Also, we will be using a mask carrying information about different regions in the organ. This mask can be provided by MRI prior to inversion. From the definition of this parametrisation in Section 4.4 and also in Figure 4.4.1 it is clear that the shear modulus has no spatial variability inside the regions $D^{(i)}$.

We therefore add the other and first level $\mathcal{P}_2^{(1)}$ of parametrisation which parametrises a Gaussian random field.

The two-level approach makes us less dependant on the choice of the parameters for the prior compared to the approach described in Section 4.3.

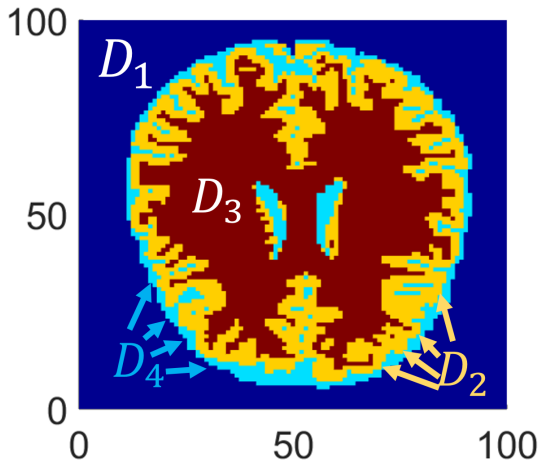


Figure 4.5.1: Mask provided by MRI that is used for the two-level parametrisation. The dark blue region D_1 of background matter surrounds the brain tissues cerebrospinal fluid (CSF) matter D_4 in light blue, grey matter D_2 in yellow and white matter D_3 in dark brown.

We use this parametrisation in context with the viscoelastic model and therefore, we also parametrise the loss modulus μ_L.

The second level of parametrisation is a map $\mathcal{P}_2^{(2)}$ that parametrises the partly unknown geometry of the discontinuous distribution of the unknown shear modulus $\mu \in \mathcal{M}$. Parts of the geometry is assumed to be known, which will be described below. It is a mapping

$$\mathcal{P}_2^{(2)} : \prod_{i=1}^{11} C(\Omega, \mathbb{R}) \to \mathcal{M}. \tag{4.5.1}$$

The eleven level set functions which are parameters for $\mathcal{P}_2$ are smooth fields with variability introduced by the first level of parametrisation explained below.

Five of the eleven smooth fields $\Bbbk_{\alpha,S}$, $\alpha \in \{1,\dots,5\}$ are used as parameters for the storage modulus μ_S. Then, we have another five smooth fields $\Bbbk_{\alpha,L}$, $\alpha \in \{1,\dots,5\}$ that are parameters for the loss modulus μ_L.

The last field $\Bbbk_C$ is used as a parameter for the unknown geometry of the cancer. Similar to Section 4.4, let $c \in \mathbb{R}$ be a parameter and define

$$D_{\Bbbk_C} = \{x \in \Omega | \Bbbk_C(x) > c\}, \quad \text{and} \quad \Omega \setminus D_{\Bbbk_C}.$$

Also, we define the indicator function

$$\mathbb{1}_{D_{\Bbbk_C}}(x) = \begin{cases} 1 & \text{if } x \in D_{\Bbbk_C}, \\ 0 & \text{if } x \notin D_{\Bbbk_C}. \end{cases} \tag{4.5.2}$$

We assume to have prior knowledge about regions $D_i \subset \Omega$, $i = 1,2,3,4$, in our domain of materials with different mechanical properties. So $\Omega = \bigcup_{i=1}^{4} D_i$ with and $D_i \cap D_j = \emptyset$ for $i \neq j$ is a (known) partition of Ω. In Figure 4.5.1 we give plots that show the partition we are using for the brain.

We assume that the storage and loss modulus μ_s and μ_L have a region D_1 (dark blue in Figure 4.5.1) where the storage and loss modulus have values $\exp(\Bbbk_{1,S}(x))$ and $\exp(\Bbbk_{1,L}(x))$. Outside this region, that is in $\Omega \setminus D_1$, storage and loss modulus have distribution $\mu_S^{(1)}$ and $\mu_L^{(1)}$. In other words, we assume

$$\mu_S(x) = \exp(\Bbbk_{1,S}(x))\mathbb{1}_{D_1}(x) + \mu_S^{(1)}(x)\mathbb{1}_{\Omega\setminus D_1}(x), \tag{4.5.3}$$

$$\mu_L(x) = \exp(\Bbbk_{1,L}(x))\mathbb{1}_{D_1}(x) + \mu_L^{(1)}(x)\mathbb{1}_{\Omega\setminus D_1}(x), \tag{4.5.4}$$

Furthermore, we assume that storage and loss modulus in $\Omega \setminus D_1$ have distribution $\exp(\mathbb{k}_{5,S}(x))$ and $\exp(\mathbb{k}_{5,L}(x))$ in an unknown area $D_{\mathbb{k}_C}$ within $\Omega \setminus D_1$. Outside $D_{\mathbb{k}_C}$ we assume the storage and loss modulus to be distributed $\mu_S^{(2)}$ and $\mu_L^{(2)}$. Note that $D_{\mathbb{k}_C}$ is the region of the brain, i.e. D_2, D_3 and D_4 and the unknown location of the tumour inside the brain $D_{\mathbb{K}_C}$. In other words, we assume

$$\mu_S^{(1)}(x) = \exp(\mathbb{k}_{5,S}(x))\mathbb{1}_{D_{\mathbb{k}_C}} + \mu_S^{(2)}(x)\mathbb{1}_{\Omega\setminus(D_1\cup D_{\mathbb{k}_C})},$$
$$\mu_L^{(1)}(x) = \exp(\mathbb{k}_{5,L}(x))\mathbb{1}_{D_{\mathbb{k}_C}} + \mu_L^{(2)}(x)\mathbb{1}_{\Omega\setminus(D_1\cup D_{\mathbb{k}_C})}.$$

Finally, we assume the region $\Omega \setminus (D_1 \cup D_{\mathbb{k}_C})$ of healthy brain tissue is split into the three domains D_2, D_3 and D_4 with the following distributions for loss and storage modulus

$$\mu_S^{(2)}(x) = \exp(\mathbb{k}_{2,S}(x))\mathbb{1}_{D_2}(x) + \exp(\mathbb{k}_{3,S}(x))\mathbb{1}_{D_3}(x) + \exp(\mathbb{k}_{4,S}(x))\mathbb{1}_{D_4}(x),$$
$$\mu_L^{(2)}(x) = \exp(\mathbb{k}_{2,L}(x))\mathbb{1}_{D_2}(x) + \exp(\mathbb{k}_{3,L}(x))\mathbb{1}_{D_3}(x) + \exp(\mathbb{k}_{4,L}(x))\mathbb{1}_{D_4}(x).$$

Note, that $D_2 \cup D_3 \cup D_4$ is the region of brain tissue which is made up of grey matter, white matter and CSF matter, so the yellow, dark brown and the light blue regions in the mask in Figure 4.5.1.

So, (4.5.3)-(4.5.4) maps $\mathbb{k}_F := (\{\mathbb{k}_{\alpha,S}\}_{\alpha\in\{1,\dots,5\}}, \{\mathbb{k}_{\alpha,L}\}_{\alpha\in\{1,\dots,5\}}, \mathbb{k}_C)$ to a shear modulus distribution $\mu = \mu_S + \mathbf{j}\cdot\mu_L$. Mapping (4.5.1) is

$$\mathbb{k}_F \mapsto \mathcal{P}_2^{(2)}(\mathbb{k}_F) = \mu. \tag{4.5.5}$$

For the first level of parametrisation, let w be stochastic process, for example Gaussian white noise, and let us assume $r \in \mathbb{N}$. The function $\Bbbk$ is defined to be the solution of the fractional stochastic PDE [34]

$$\left[I - \nabla \cdot \begin{pmatrix} l_1^2 & 0 \\ 0 & l_2^2 \end{pmatrix} \nabla \right]^{(r+1)/2} (\Bbbk(x) - \log m) = \left[4\sqrt{l_1 l_2}\sigma^2 \pi \frac{\Gamma(r+1)}{\Gamma(r)} \right]^{1/2} w(x), \tag{4.5.6}$$

where I is the identity and Γ is the gamma function, $\Bbbk$'s smoothness is controlled by parameter r and it's amplitude by parameter σ^2. Furthermore, $l_1, l_2 > 0$ is the intrinsic length scale along the horizontal and vertical direction.

The solution $\Bbbk(x)$ of the stochastic PDE 4.5.6 is a Gaussian random field with mean $\log m$ and a covariance operator given by an Whittle-Matérn auto-correlation function that was already defined in Section 4.3.1 [61]. Instead of having only one lengthscale like in Section 4.3.1, we allow for more flexibility in the incorporation of prior knowledge by considering different length scales along vertical and horizontal directions, so

$$\mathrm{ACF}(x) = \frac{\|x\|_{l_1,l_2}}{2^{r-1}\Gamma(r)} K_r \left(\|x\|_{l_1,l_2} \right),$$

where K_r is the modified Bessel function of the second kind of order r, and

$$\|x\|_{l_1,l_2} = \sqrt{\frac{x_1^2}{l_1^2} + \frac{x_2^2}{l_2^2}}.$$

Equation (4.5.6) can be considered as mapping

$$(m, \sigma, r, l_1 l_2, w) \mapsto \Bbbk,$$

We use this mapping as parametrisation of smooth fields $\Bbbk_{\alpha,S}$, $\Bbbk_{\alpha,L}$ and $\Bbbk_C$ in (4.5.5) with parameters $(m, \sigma, r, l_1 l_2, w)$. More specifically, we define

$$\begin{aligned} k_{\alpha,S} &:= (m_{\alpha,S}, \sigma_{\alpha,S}, r_{\alpha,S}, l_{\alpha,S,1} l_{\alpha,S,2}, w_{\alpha,S}) \quad \alpha = 1, \dots, 5, \\ k_{\alpha,L} &:= (m_{\alpha,L}, \sigma_{\alpha,L}, r_{\alpha,L}, l_{\alpha,L,1} l_{\alpha,L,2}, w_{\alpha,L}) \quad \alpha = 1, \dots, 5, \\ k_C &:= (m_C, \sigma_C, r_C, l_{C,1} l_{C,2}, w_C). \end{aligned}$$

Let $k_F := (\{k_{\alpha,S}\}_{a\in\{1,\dots,5\}}, \{k_{\alpha,L}\}_{a\in\{1,\dots,5\}}, k_C)$. We define the first level of parametrisation $\mathcal{P}_2^{(1)}$ to be the mapping that maps the hyperparameters k_F to the Gaussian random fields $\Bbbk_F$, so

$$k_F \mapsto \mathcal{P}_2^{(1)}(k_F) = \Bbbk_F.$$

We concatenate $\mathcal{P}_2^{(1)}$ and $\mathcal{P}_2^{(2)}$ to define the following operator $\mathcal{P}_2$ mapping input k_F to loss and storage modulus (μ_L, μ_S). This two-level parametrisation of the shear modulus is defined a

$$k_F \mapsto \mathcal{P}_2(k_F) := \mathcal{P}_2^{(2)}(\mathcal{P}_2^{(1)}(k_F)) = \mu. \tag{4.5.7}$$

We use this two-level parametrisation in the context of the experiments done in Chapter 6. More precisely we choose points $\{\hat{x}_i\}_{i=1}^{\hat{n}} \in \Omega$ corresponding to the chosen computational mesh of the discretisation scheme. In order to define

a prior we choose for each sample of the prior ensemble the hyperparameters k_F of the three Gaussian random fields in $\mathbb{k}_F$ and draw white Gaussian noise samples $w_{\alpha,S}, w_{\alpha,L}, w_C \sim \mathcal{N}(0, \mathbb{1}_{\hat{n}\times\hat{n}})$, where $\mathbb{1}_{\hat{n}\times\hat{n}}$ is the $\hat{n}$ by $\hat{n}$ unit matrix.

When using the two-level parametrisation $\mathcal{P}_2$ in each update of EKI we need to calculate the Gaussian random fields $\mathbb{k}_F$ for hyperparameters k_F. Instead of using the Whittle-Matérn autocorrelation function in order to define a covariance matrix and the Karhunen-Loève expansion to define Gaussian random fields as described in Section 4.3.1, in Chapter 6 we use the solution of 4.5.6 which is a Gaussian random field. In practice, we do that by applying the finite difference method approximation of this differential equations which is described in [74] and [34]. This is computationally much cheaper than the approach with the Karhunen-Loève expansion in Section 4.3.1. However, the approach using the stochastic PDE 4.5.6 only allows for integer smoothness parameters r of the Gaussian random field.

4.6 SUMMARY

In this Chapter we have described EKI, the Bayesian inversion approach we apply to the inverse problem arising in MRE. We have discussed the properties and advantages of EKI and showed how we define the prior. Finally, we showed

the different parametrisation we use. In the remaining part of this thesis we will describe and analyse the application of EKI to MRE introduced in this Chapter.

5

NUMERICAL TESTING - PURELY ELASTIC MODEL

Algorithm 1 is implemented in MATLAB. For most of this Chapter we will be using the forward map $\mathcal{F} = \mathcal{G} \circ \mathrm{id}$, i.e. we are using no parametrisation and $\mathcal{K} = \mathcal{M}$ and $\mathrm{id} : \mathcal{K} \to \mathcal{M}$ is the identity on $\mathcal{M}$. In Section 5.6 we will then use $\mathcal{P}_1$ introduced in Section 4.4, where the details of the parametrisation will be explained.

Throughout this whole Chapter we work with the purely elastic model. That is, we assume $\mu_L = 0$ and we work with the stress tensor (3.2.5) when considering $\mathcal{G}$ or $\mathcal{F}$.

In this Chapter we use the Karhunen-Loève expansion to define the prior sample described in Section 4.3.1. We discussed the advantages and disadvantages of Karhunen-Loève expansion in Section 4.3.1.

5.1 IMPLEMENTATION

The forward solution $\mathcal{G}(\mu)$ which requires the solution of boundary value problem (3.2.1)-(3.2.3) is approximated using a finite element method provided by ABAQUS FEA using plane strain, bilinear, hybrid elements with 4 nodes and constant pressure and we use a direct solver. On a mesh with 490000 elements, synthetic data are generated in the square shaped domain of size 70mm by 70mm enclosing the domain Ω. This very dense mesh avoids any pollution which will be discussed in Section 5.7.1. For inversion, a mesh of 4900 elements is used. The inversion has high enough resolution and EKI has low computational costs. The measurements $\mathbf{v} \in \mathbb{R}^m$ are given in $70^2 = 4900$ nodes, so $m = 9800$ in (3.4.2).

5.2 BASIC SETUP AND SYNTHETIC DATA

We choose a distribution of the parameter μ across the domain with two values, a background value and inclusion value as shown in Figure 5.2.1 (left). The background value of the shear modulus μ to be 1000Pa and the inclusion value to be 4000Pa. The measured shear modulus in MRE depends on the frequency of the sinusoidal excitation [76]. The chosen background and inclusion values

could model the difference between healthy tissue and tissue that is affected by cancer for frequency 1Hz.

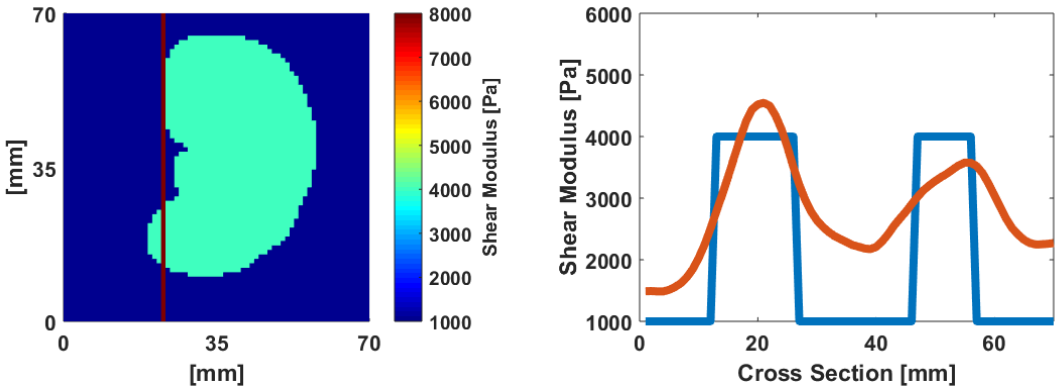


Figure 5.2.1: **Left:** Distribution of the parameter μ across the domain Ω. The red line indicates the location that is used for cross-sectional plots. **Right:** Distribution of the parameter μ in the cross-section and a mean of a posterior ensemble as an example.

The experiments were all run with a frequency of 1Hz ($\omega \approx 6.28\text{rad}/s$) in (3.2.1)-(3.2.3) which is also used in other studies ([95],[81]). Due to high pollution in standard Galerkin finite element method approximations in high frequency MRE which will be discussed in Section 5.7.1 we can only run inversion up to a frequency of 10Hz.

The top edge of the domain is the Dirichlet boundary with $\hat{v}_{\text{top}} = 0$. Harmonic deformation of frequency 1Hz and amplitude $\hat{v}_{\text{bot}} = 0.001\text{mm}$ is applied to the bottom edge. Left and right edge are traction free boundaries.

In all experiments Poisson's ratio is set to 0.499 and density $\rho_W = 1\text{kg/l}$ which is discussed in Section 3.

In (3.4.2) we defined $\mathcal{G}(\mu)$ to be the quantity that is obtained by evaluating the solution v of the boundary value problem (3.2.1)-(3.2.3) for some $\mu \in \mathcal{M}$ at 4900 measurement points. This density of measurements is one that can be found in many real-world applications [14]. We construct data by evaluating the finite element method approximation $v^\dagger$ of $F(\mu)$ at the measurement points and define measurements to be $\mathbf{v} = v^\dagger + \eta$. The noise η is drawn from a centred Gaussian distribution $\eta \sim \mathcal{N}(0, B)$, where we choose $B = \text{diag}(b_1, \ldots, b_m)$ with

$$b_i = \left(\sigma_N \cdot 10^{-2} |v_i^\dagger|\right)^2 + \left(10^{-3} |\max\{v_i^\dagger\}_{i=1}^m - \min\{v_i^\dagger\}_{i=1}^m|\right)^2, \quad i = 1, \ldots, m. \tag{5.2.1}$$

The first term on the right hand side of this definition tunes the noise at each point $i = 1, \ldots, m$ to the amplitude of the displacement [39]. Considering (4.3.3) and (4.3.4) it can be seen that the first term in (5.2.1) corresponds to adding σ_N% Gaussian noise at each point. The second term ensures that small variances are avoided which could result in numerical instabilities when using noise covariance B in EKI.

We sometimes refer to σ_N as the noise level and say the data has σ_N% noise level if the noise added to the data is drawn from centred Gaussian distribution with Covariance described in (5.2.1).

5.3 CONVERGENCE OF ENSEMBLE FOR SAMPLE SIZE $J = 150$

In the first series of experiments we apply EKI to the synthetic data described in Section 5.2. In the autocorrelation function (4.3.1) that is used to define the prior we choose the length scale to be $l = 1/4$ and smoothness $r = 1$. The autocorrelation function is evaluated on a 70 by 70 mesh on a unit square in order to obtain the prior covariance Σ_0. The prior mean is set to constant $m_0 = 2500$Pa which is the average of the inclusion and background value of the ground truth μ. The prior ensemble has ensemble size $J = 150$.

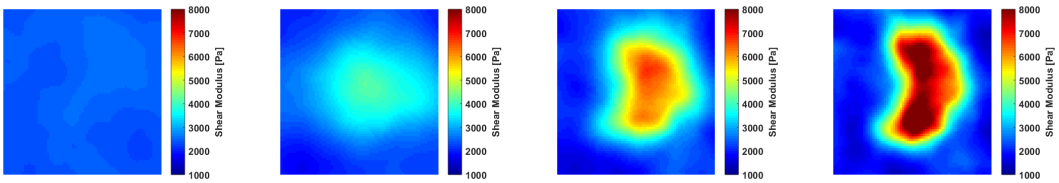


Figure 5.3.1: The evolution of the ensemble mean from the prior to the posterior. **Left:** Initial Ensemble mean. **Centre left:** Ensemble mean after one iteration. **Centre right:** Ensemble mean after three iterations. **Right:** Posterior ensemble mean.

In Figure 5.3.1 we show the ensemble mean for the prior, the ensemble after one iteration, after three iterations and for the posterior ensemble. In the left column of Figure 5.3.2 we show plots for five members of the initial ensemble. In the second and third column we show plots for these members after iteration

one and three. The right column shows plots for these members in the posterior ensemble.

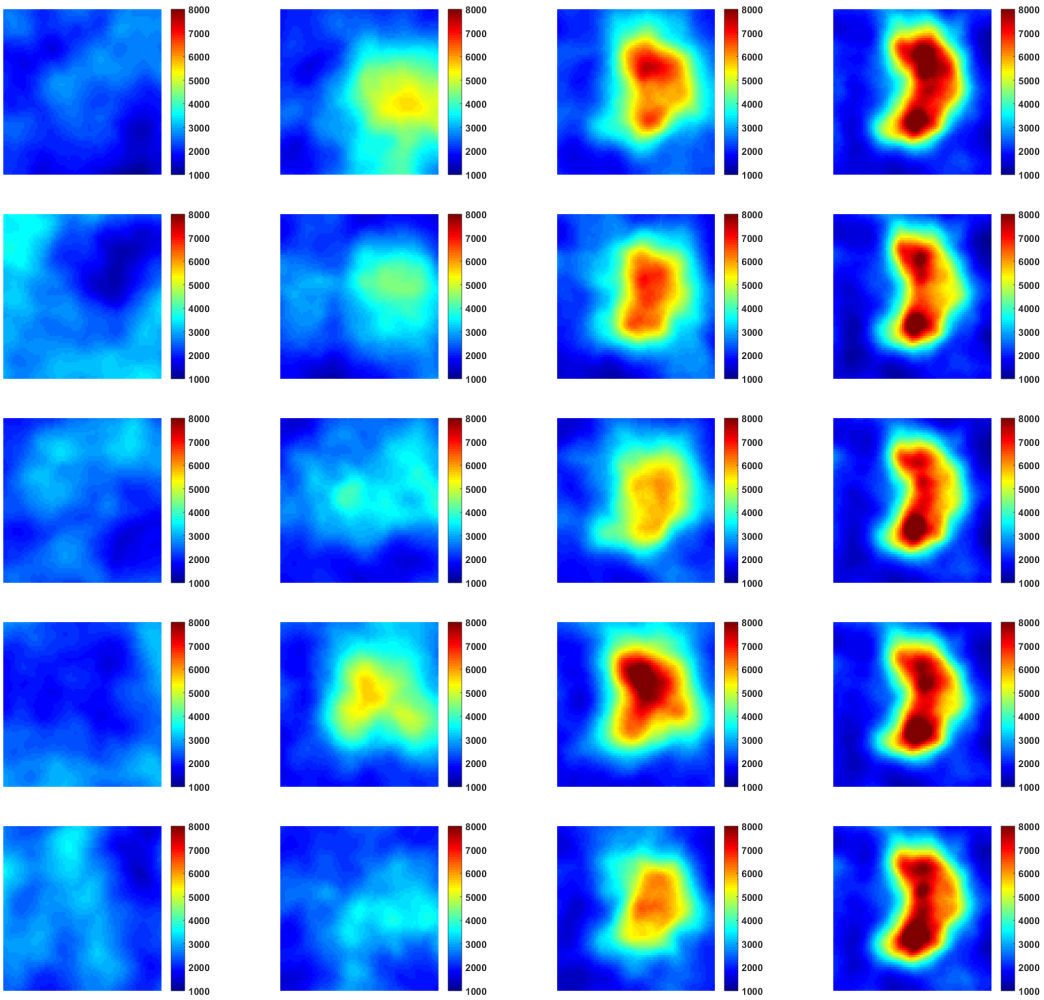


Figure 5.3.2: The evolution from the prior to the posterior of five exemplary samples in the ensemble. **First Column:** Initial Ensemble. **Second Column:** Ensemble after one iteration. **Third Column:** Ensemble after three iterations. **Fourth Column:** Posterior Ensemble.

The stopping criterion (4.2.25) is met after 11 iterations. On a standard office computer with a four core 3.4GHz CPU this takes around 2 hours. This is the case if B in the EKI ensemble update (4.2.15) is defined by (5.2.1) in the synthetic data. The exact knowledge of the noise and therefore B is unrealistic. Increasing the noise level - e.g. by multiplying B with a constant > 1 in (4.2.15) has the consequence that stopping criterion (4.2.25) is met after fewer iterations. However, the approximation of the ground truth by the posterior ensemble and the posterior ensemble mean is less accurate. Decreasing the noise level in EKI updates increases the amount of iterations until the stopping criterion is met.

Results can be compared to the ground truth in Figure 5.2.1 which has the same scaling.

The cross-sectional plots shown in Figure 5.3.3 reveal how the ensemble approaches the posterior over the iterates. The approximation of the ground truth by the mean of the ensemble gets better in every iteration. Also, it can be seen that the variance of the ensemble at each point is decreasing over the iterates.

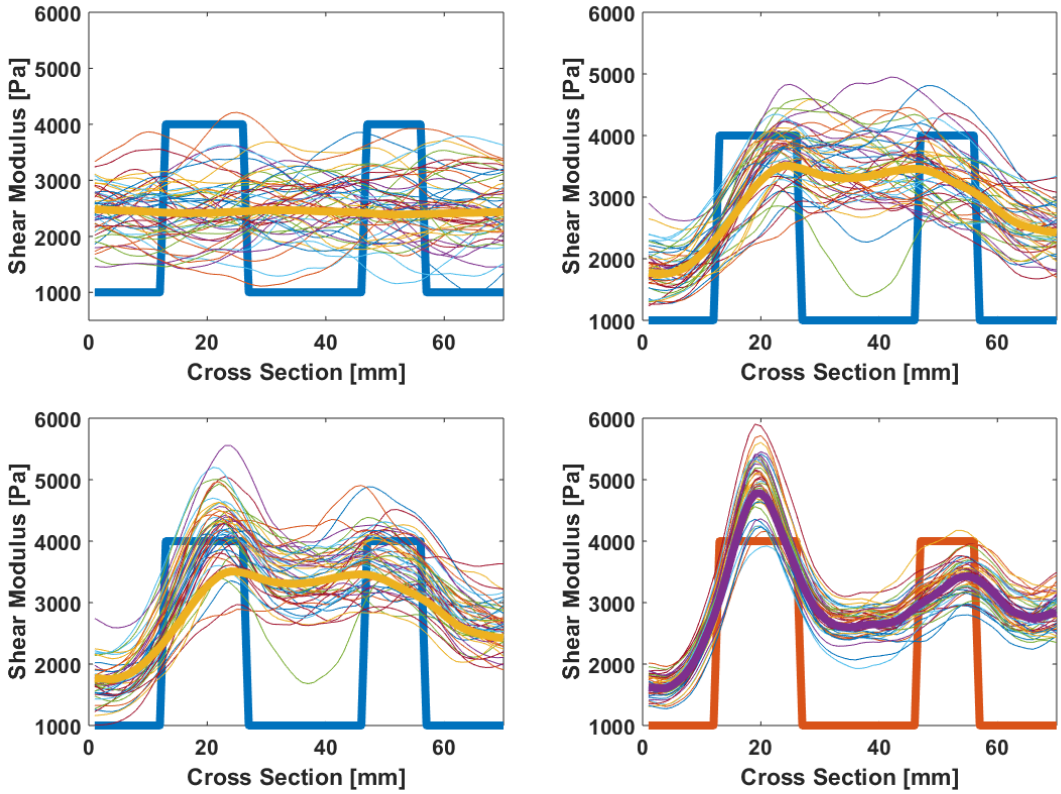


Figure 5.3.3: Cross-sectional plot in the location indicated in Figure 5.2.1 of 50 members of the ensemble after various iterations. The thick line indicates the sample mean. **Top left:** Prior ensemble. **Top right:** Ensemble after one iteration. **Bottom left:** Ensemble after three iterations. **Bottom right:** Posterior ensemble.

5.4 INVERSIONS USING DIFFERENT PRIORS

In this experiment we draw the prior ensemble from priors with different length scales and try to understand how this influences the posterior ensemble. An

ensemble of size $J = 150$ is drawn from a prior with smoothness $r = 2$ and length scale $l = 1/8$, $l = 1/4$ and $l = 1/2$ in autocorrelation function (4.3.1). In Figure 4.3.1 we show a plot for one member of the initial ensemble for each lenght scale.

In Figure 4.3.2 the corresponding plots of the posterior ensemble mean are shown. It can be seen that for length scale $l = 1/4$ in the prior the posterior ensemble mean captures well the overall shape of the inclusion in the distribution of μ across the domain. However, values of the posterior inside the area of inclusion are very high. The mean of the posterior ensemble for $l = 1/8$ in Figure 4.3.2 (left) is capturing the shape of the inclusion worse than for $l = 1/4$. However, values of the posterior are closer to the ground truth. The euclidean distance between the posterior ensemble mean and the ground truth is lower for the posterior obtained from length scale $l = 1/8$. However the distribution of the parameter μ across the domain in posterior is very heterogeneous and there are some peaks in the inclusion. These peaks might be misinterpreted as diseased tissue and therefore, the posterior for length scale $l = 1/4$ might be diagnostically more useful in an application.

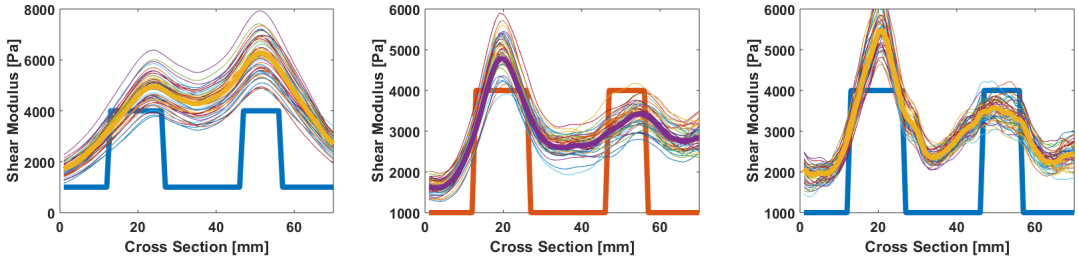


Figure 5.4.1: Cross-sectional plot in the location indicated in Figure 5.2.1 of 50 members of the posterior ensemble for different length scales l in the prior. Samples of the corresponding priors can be found in Figure 4.3.1. **Left:** $l = 1/8$. **Mid:** $l = 1/4$. **Right:** $l = 1/2$.

In the cross-sectional plot of this experiment shown in Figure 5.4.1 it can be seen in the left plot that the spread in the posterior ensemble is very high. The spread of the posterior ensemble is high in locations where the ensemble is further away from the ground truth. All plots show that the posterior ensemble is close to the ground truth between 10mm and 15mm in the cross-section (shown on the x-axis). The spread of all posterior ensembles is low there. The spread of all ensembles is high in the left peak of the plot where the ensemble is further away from the ground truth. This expresses the uncertainty about the inversion in this location.

5.5 INVERSIONS USING VARIOUS ENSEMBLE SIZES

We run the Algorithm 1 in the basic setup described in Section 5.2 for different choices of ensemble sizes J: 50, 150 and 250. For the prior described in Section 5.2 stopping criterion is usually met after 10 iterations for sample size $J = 50$, after 11 iterations for $J = 150$ and after 13 iterations for $J = 250$. The computational time grows approximately linearly with sample size J. In Figure 5.5.1 we show plots of the posterior ensemble mean.

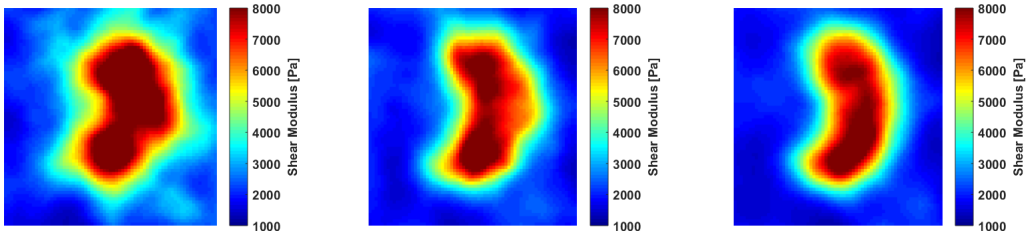


Figure 5.5.1: Posterior ensemble mean for different ensemble sizes J. **Left:** $J = 50$. **Centre:** $J = 150$. **Right:** $J = 250$

Also in the cross-sectional plots in Figure 5.5.2 it can be seen how the sample mean seems to give the best approximation for the ground truth distribution of μ for $J = 250$.

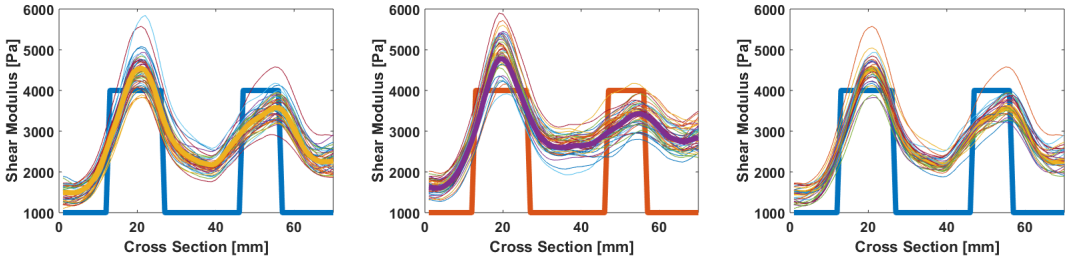


Figure 5.5.2: Cross-sectional plot in the location indicated in Figure 5.2.1 of the posterior ensemble for different ensemble sizes J. **Left:** $J = 50$. **Centre:** $J = 150$. **Right:** $J = 250$

5.6 INVERSIONS USING LEVEL SET PARAMETRISATION AND DIFFERENT PRIORS

In this series of experiments we apply EKI with level-set parametrisation to synthetic data coming from a ground truth distribution of μ across the domain shown in Figure 5.6.1.

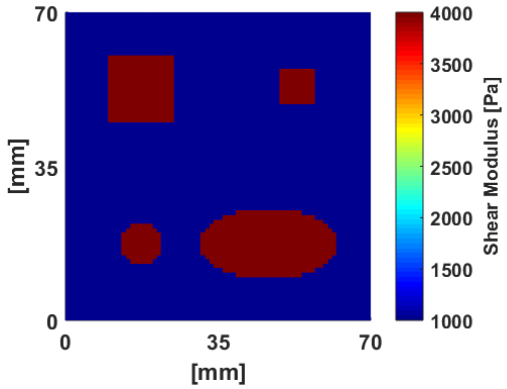


Figure 5.6.1: Ground Truth distribution of μ across the domain. The domain is of size 70mm by 70mm.

We wish to estimate the interface between the region $D^{(1)}$ with background value of μ and the region $D^{(2)}$ with inclusion value of μ. The distribution of shear modulus μ across the domain can be described by a discontinuous function $\mu(x) = k_S^{(1)} \mathbb{1}_{D^{(1)}} + k_S^{(2)} \mathbb{1}_{D^{(2)}}$. The level set map $\mathcal{P}_1$ is now used in order to parametrise the unknown geometry $D^{(1)}, D^{(2)}$ and the discontinuous function μ. EKI applied to (4.4.3) provides an approximation k of the posterior in the space $\mathcal{K}$ of the level set function. The field $\mathcal{P}_1(k)$ is giving us a quantity that could be called the "level-set posterior" approximating the discontinuous field μ and therefore the unknown geometry.

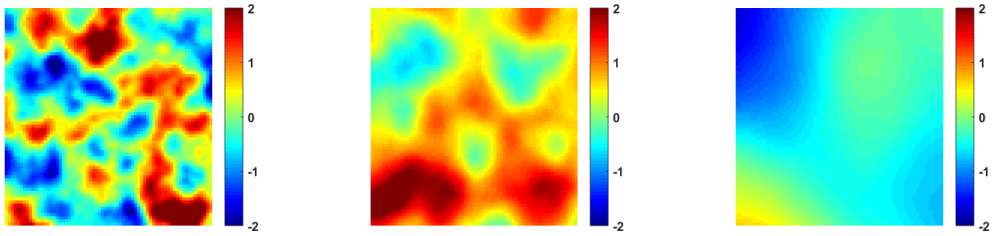


Figure 5.6.2: Samples drawn from $\mathcal{N}(0, \Sigma_0)$ (dimensionless) with Σ_0 being a covariance operator that arises from the autocorrelation function (4.3.1) with $r = 2$, amplitude scale $\sigma^2 = 1$ and different length scales l. **Left:** $l = 1/10$. **Mid:** $l = 1/5$. **Right:** $l = 1/2$

We define priors with different lengthscales in the space $\mathcal{K}$ of the parameter functions. In Figure 5.6.2 we show a plot for one member of the initial ensemble for lengthscale $l = 1/10$, $l = 1/5$ and $l = 1/2$ and smoothness $r = 2$ in the autocorrelation function (4.3.1) of the prior. We assume to have prior knowledge about $k_S^{(i)}$ and set $k_S^{(1)} = 1000$ and $k_S^{(2)} = 4000$ in the level set map $\mathcal{P}_1$. Also, we set $c^{(0)} = -\infty$, $c^{(1)} = 0$ and $c^{(2)} = \infty$. In this case the primary unknown is the geometry $D^{(1)}$ and $D^{(2)}$ and we are interested in $\mathbb{1}_{D^{(2)}}$.

In Figure 5.6.3 we can see plots of both the field k in the space of the level-set function and $\mathbb{1}_{D^{(2)}}$ of the level-set posterior for priors with different lengthscales. A lengthscale of $l = 1/5$ in the prior seems to be the best choice for this ground truth. The inversion is good for lengthscale $l = 1/5$ in the prior in this case.

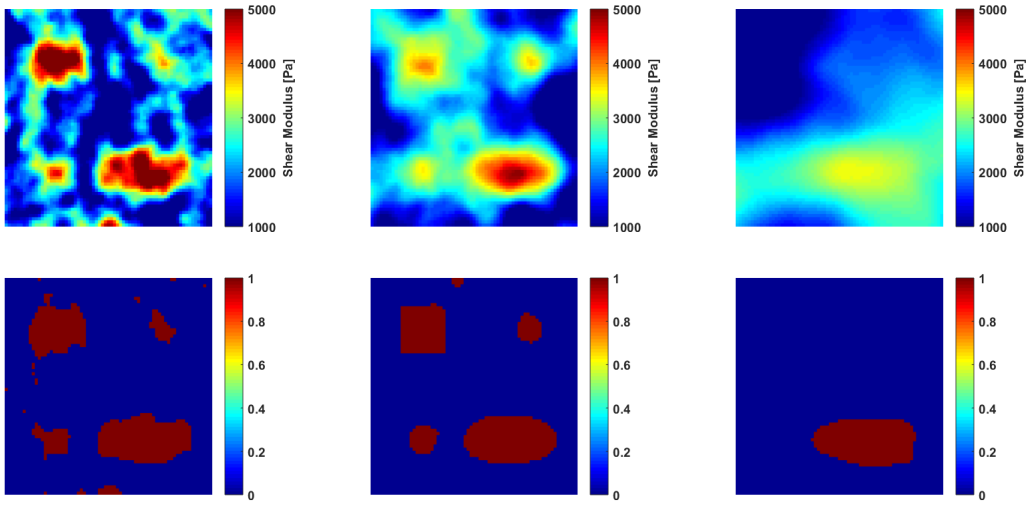


Figure 5.6.3: Inversions results for parameter μ in Figure 3.2.2 of EKI with level-set parametrisation and corresponding priors with different length scales l in Figure 4.3.1 . **Top:** EKI inversion $k \in \mathcal{K}$. **Bottom:** Indicator function $\mathbb{1}_{D^{(2)}}$ (dimensionless) of level set map $\mathcal{P}_1(k)$ with constants $c_0 = -\infty$, $c_1 = 0$ and $c_2 = \infty$. **Left:** $l = 1/10$. **Mid:** $l = 1/5$. **Right:** $l = 1/2$

In Figure 5.6.4 we provide plots for $\mathbb{1}_{D^{(2)}}$ of the level-set posterior in the case of different distributions of parameter μ in the domain. All of them use a prior with length scale $l = 1/5$.

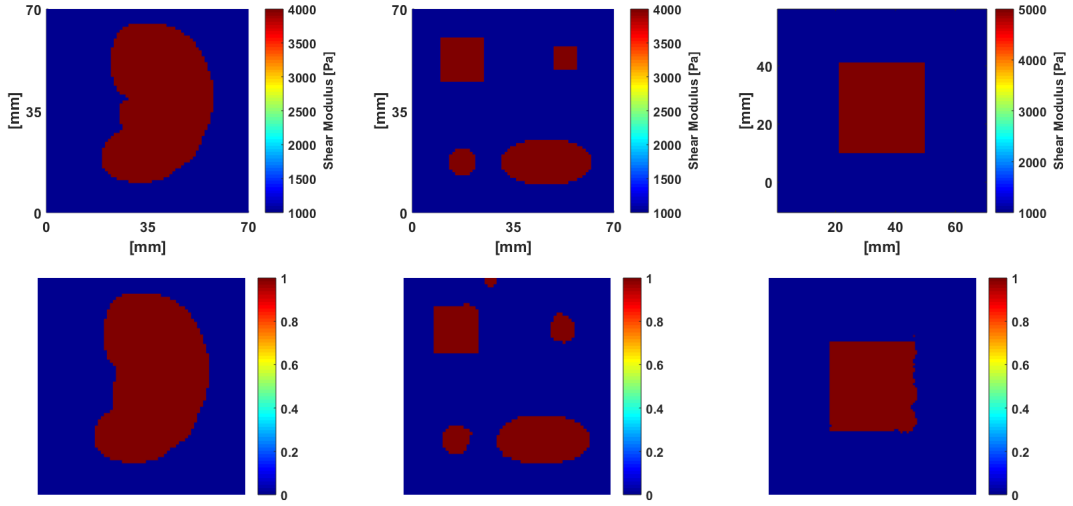


Figure 5.6.4: Inversions results using EKI with level-set parametrisation for various distributions of parameter μ across the domain. **Top:** Ground Truth distribution of μ. **Bottom:** Indicator function $\mathbb{1}_{D^{(2)}}$ (dimensionless) of level set map of the corresponding posterior ensemble mean after applying EKI with level set parametrisation.

5.7 CONCLUSION: LIMITATIONS AND PROBLEMS WITH A PURELY ELASTIC MODEL

In this section we want to draw interim conclusions from the experiments conducted in Section 5.1 to 5.6 and further discuss and analyse the limitations and problems of using the purely elastic model that became apparent by the experiments in this chapter.

The experiments used Algorithm 1 without any parametrisation or the parametrisation $\mathcal{P}_1$ introduced in Section 4.4. We could see how the choice of the prior has a crucial influence on the accuracy of the posterior. Also, we analysed the influence of ensemble size on convergence speed and accuracy of the posterior ensemble. Overall, without parametrisation, EKI estimates were too smooth and could not capture sharp interfaces. We therefore used the parametrisation $\mathcal{P}_1$ introduced in Section 4.4 which delivered highly accurate estimates. However, when using parametrisation $\mathcal{P}_1$ in EKI, prior knowledge about $k_S^{(1)}$ and $k_S^{(2)}$ is necessary, which cannot be assumed in application usually. Also, in all experiments, regardless of the parametrisation used, there is a high dependency of the posterior on the prior. We could demonstrate this dependency specifically for the length scale l. There is no straight-forward way for how to choose these parameters prior to inversion. In the experiments in

Chapter 6 we therefore use the two-level parametrisation introduced in Section 4.5.7. When using this parametrisation, the length scale and mean become part of the unknown in the inversion and we only need to choose an interval of values for the prior, not one value. If we are not sure about the length scales when choosing the prior, the interval can be chosen to be big.

High frequency MRE using a purely elastic model turned out to be more difficult than low frequency MRE. The numerical approximation of the forward map (3.4.2) using standard Galerkin finite element method is affected by pollution which is described in Section 5.7.1. A potential solution to this problem is described in Section 5.7.2.

Furthermore, the direct application of Algorithm 1 without a modification turned out to be unsuccessful because of resonating wave fields. This problem is described on the basis of a simple 1-D version of EKI and a solution to this is described in Section 5.7.3.

It will become obvious however that using a viscoelastic model is solving both the problem of pollution and resonating wave fields. Extended high frequency experiments using this model will be then shown in Chapter 6.

5.7.1 *Pollution*

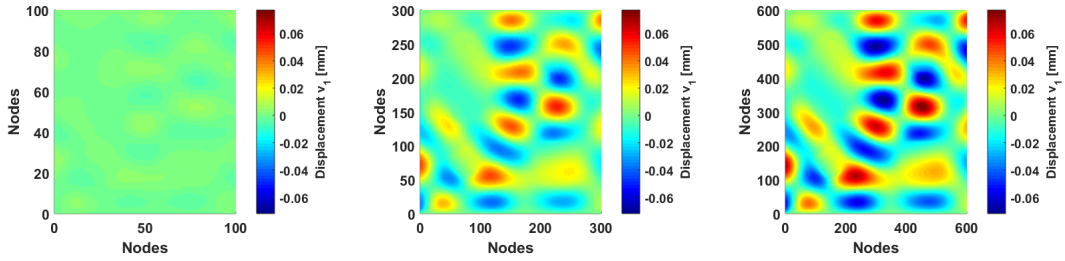


Figure 5.7.1: Plot for the finite element method approximation of the first component of the solution to boundary value problem (3.2.1)-(3.2.3) using various mesh sizes and high frequency of 60Hz. The setup described in Section 5.2 is used. The domain size is 70mm by 70mm and μ is distributed across the domain like shown in Figure 5.6.4 (top left). **Left:** 100 by 100 nodes. **Middle:** 300 by 300 nodes. **Right:** 600 by 600 nodes.

The simulation of the purely elastic Helmholtz equation (3.2.1), i.e. $\mu_L = 0$, at high frequencies suffers from the pollution effect [24] and it is proven that for standard Galerkin finite element method it cannot be removed completely [7]. In the finite element space, for low frequencies, the Galerkin approximation and the best approximation only differs by a constant factor. For increasing ω or wave number however, the ratio of the error of the Galerkin approximation and the error of the best approximation tends to infinity [7]. This lack of robustness is called the pollution effect.

In Figure 5.7.1 we show plots obtained using the standard Galerkin finite element method approximations implemented in ABAQUS for high frequency of 60Hz, various mesh sizes and a ground truth distribution of parameter μ across the domain is shown in Figure 5.6.4 (top left). A very small mesh size is necessary until the error between the Galerkin approximation and the best approximation has converged. This is the case in the right plot.

It is worth noting at this point that the pollution occurs for any set up of (3.2.1) that results in a high wave number. High densities ρ_W, small parameters μ and λ or solving (3.2.1) on a large domain result in the pollution effect. In fact, both

$$\nabla \cdot \left[\left(\mu/\omega^2 \right) \left(\nabla v + \nabla v^T \right) + \left(\lambda/\omega^2 \right) (\nabla \cdot v)\, I \right] = -\rho_W v, \qquad \text{for } x \in \Omega,$$

and

$$\nabla \cdot \left[\mu \left(\nabla v + \nabla v^T \right) + \lambda (\nabla \cdot v)\, I \right] = -\rho_W v, \qquad \text{for } x \in \omega \cdot \Omega,$$

result in the same wave field as (3.2.1) and therefore, contain the same wave number and the same amount of pollution. The solutions of the PDEs above have the same number of waves inside the domain. Here, we use $\omega \cdot \Omega = \{\omega \cdot x | x \in \Omega\}$.

In practice, the pollution effect requires the usage of very small mesh sizes for Galerkin approximations to be accurate for high frequencies [88]. In the

context of EKI this exceeds the computational capacity of today's computer. For EKI with sample size $J = 200$ and convergence after 11 updates, the solution of boundary value problem (3.2.1)-(3.2.3) needs to be approximated at least 2200 times.

A possible solution to this problem when using the purely elastic model is therefore, to split both the numerical forward solution method and the inversion into subzones, described in the following section.

However, pollution is very small for the viscoelastic model. In Figure 5.7.2 we show plots of the finite element method approximation for various mesh sizes at a high frequency and use a constant loss modulus of $\mu_L = 400$Pa in the ground truth distribution. It can be seen in Figure 5.7.2 that finite element method approximations are very accurate for coarse grids already.

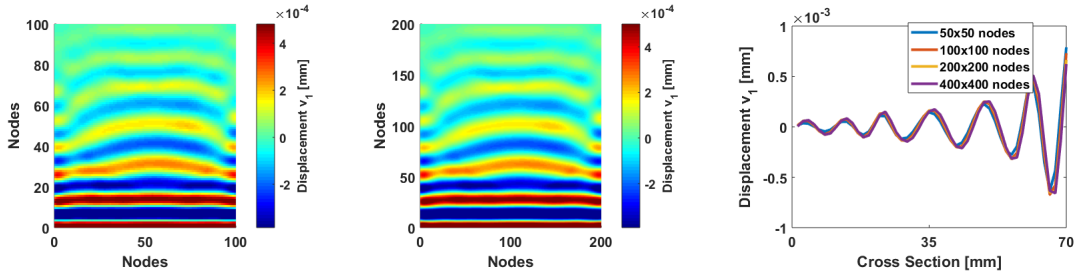


Figure 5.7.2: Plot for the finite element method approximation of the real part of the first component of the solution to boundary value problem (3.2.1)-(3.2.3) using various mesh sizes and a high frequency of 60Hz. The setup is similar to the basic setup described in Section 5.2, however, the shear modulus is complex and we assume to have a constant loss modulus of $\mu_L = 400$Pa across the domain. The domain size is 70mm by 70mm and μ is distributed across the domain like shown in Figure 5.6.4. **Left:** 100 by 100 nodes. **Middle:** 200 by 200 nodes. **Right:** Cross-sectional Plot.

5.7.2 *Subzone Inversion*

One way of avoiding the problem of pollution introduced in Section 5.7.1 is to split both the numerical forward solution method and the inversion into smaller sub-problems with very high resolution and thereby adapt the subzone approach discussed in Section 3.5.3 to EKI.

The application of EKI to inverse problem (4.1.1) on a subzone using (3.5.11) as a forward map turns out to be successful and robust regarding noise on

the boundary coming from measurements. In Figure 5.7.3 we show plots of subzone ensemble Kalman algorithm, i.e. EKI applied in four subzones using $\mathcal{G}_z$ instead of $\mathcal{G}$ in each subzone. In this experiment we used $\mathcal{P} = \mathrm{id}$, so we are not using any parametrisation. The plot at the top shows the inversion result that is obtained by applying EKI without subzones in this basic one-dimensional set up with a sinusoidal excitation of 1Hz, a noise level of 5% in the data and a mesh size of 1 mm. The bottom plot shows the result using four subzones which is about as accurate as the result obtained from EKI without subzones.

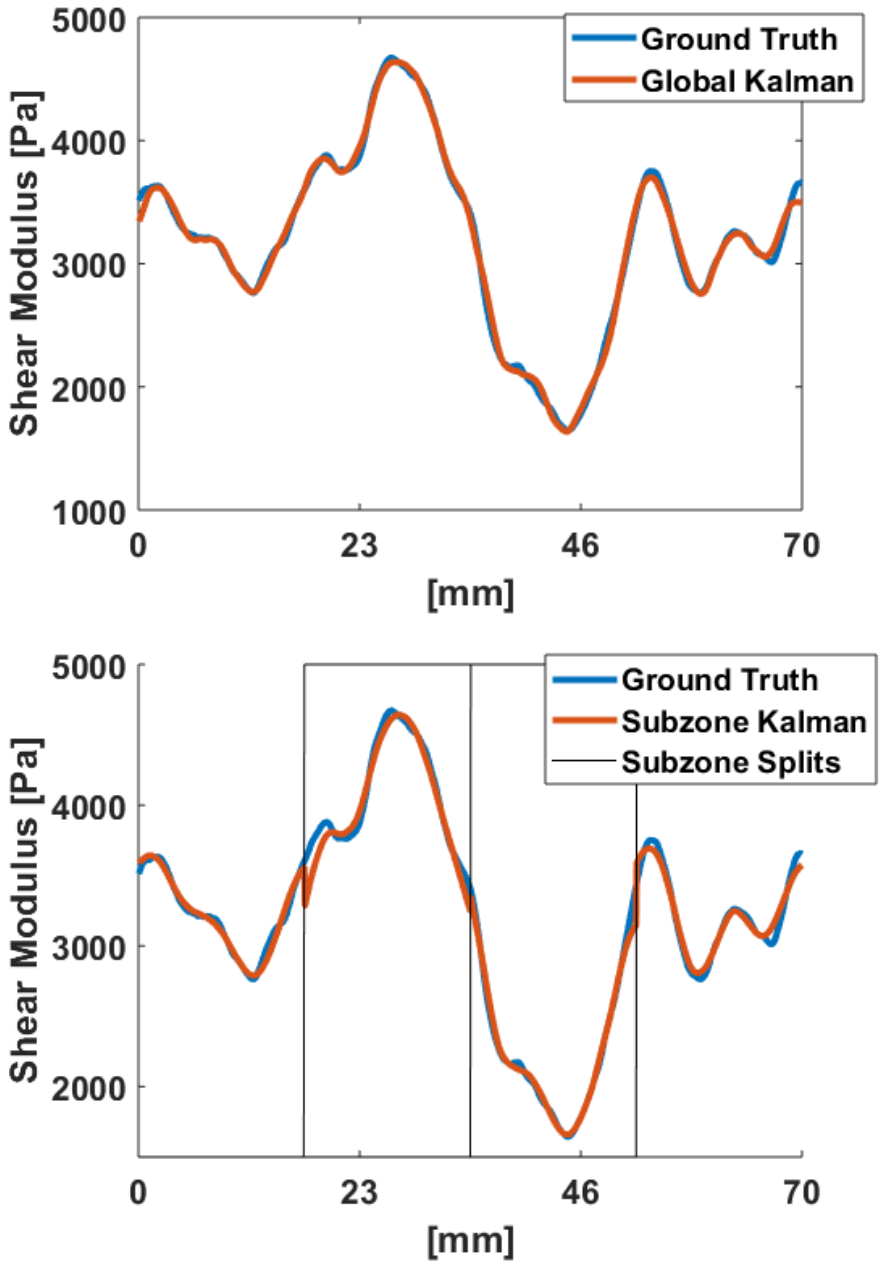


Figure 5.7.3: **Top:** EKI. **Bottom:** Subzone EKI.

5.7.3 *Resonance*

Another challenge that arises when applying EKI to high frequency MRE using the purely elastic model is described in this section. For $\mu_L = 0$ the model (3.2.1) does not take into account viscosity which describes the resistance of solid to deformation. Therefore, we model wave propagation without damping. As a consequence, we can show in experiments that the forward map $\mathcal{G}$ turns out to be not continuous with respect to parameter μ, which is required [83]. Some members of the ensemble $\mu^{(j)}$ produce wave fields that have very big amplitudes, even though they are very similar to other members in the ensemble.

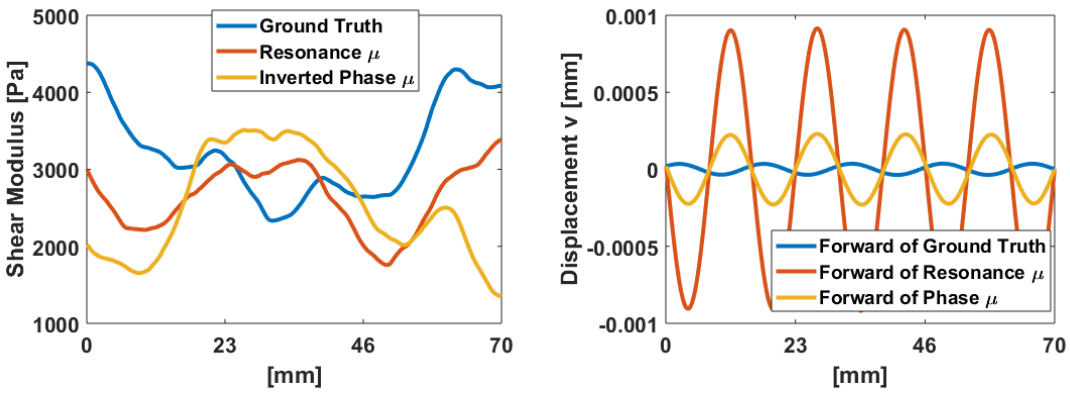


Figure 5.7.4: **Left:** Different distributions of μ across the domain. **Right:** Corresponding solutions of the boundary value problem (3.2.1)-(3.2.3) in the same colour.

In one dimension (3.2.1) is solved for constant μ on $\Omega = [x_0, x_J]$ with Dirichlet boundary $v(x_0) = v_0$ and $v(x_J) = v_J$ by

$$v(x) = \frac{v_J - v_0 \cdot \cos\left(\sqrt{\mathfrak{k} \cdot \mu^{-1}} \cdot (x_J - x_0)\right)}{\sin\left(\sqrt{\mathfrak{k} \cdot \mu^{-1}} \cdot (x_J - x_0)\right)} \cdot \sin\left(\sqrt{\frac{\mathfrak{k}}{\mu}}(x - x_0)\right) + v_0 \cdot \cos\left(\sqrt{\frac{\mathfrak{k}}{\mu}}(x - x_0)\right),$$

where $\mathfrak{k} := \frac{(1-2\nu)\cdot\rho_W\omega^2}{2-2\nu}$. For high frequencies, even a small change in μ can increase or decrease the amplitude of the wave field a lot. The same applies to non-constant μ, see Figure 5.7.4.

The same observation is made in the two-dimensional setup for MRE described in Section 5.2. For a high frequency of 60Hz, a cross-sectional plot in the location indicated in Figure 5.2.1 of the wave fields $\mathcal{G}(\mu_0^{(j)})$ of some prior ensemble $\{\mu_0^{(j)}\}_{j=1}^J$ can be seen in Figure 5.7.5. The left plot is showing the wave fields for an ensemble of size $J = 100$ drawn from a prior with variance $\sigma^2 = 10$ and the right with variance $\sigma^2 = 0.5$. It can be seen that the amount of members in the ensemble that produce resonating wave fields increases with variance.

One way to tackle the problem of resonating wave fields: At every iteration n after the forward solution $\mathcal{G}(\mu_n^{(j)})$ of the ensemble is approximated (this is step 3 in Algorithm 1), every member of the ensemble that is producing a wave field with more than twice the amplitude of the measured wave field $\mathbf{v}$

is removed from the ensemble. In Figure 5.7.6 we show a plot of wave fields $\mathcal{G}(\mu_0^{(j)})$ after all members $\mu_0^{(j)}$ that produced resonating wave fields in Figure 5.7.5 (left) have been removed.

However, because we are removing members of the ensemble in each iteration, this approach requires a prior ensemble of large size, especially if the prior has big variance. Therefore, the model (3.2.1)-(3.2.3) incorporating viscosity, i.e. $\mu_L \neq 0$ is used in Chapter 6, which does not produce resonating wave fields as described in the beginning of this subsection.

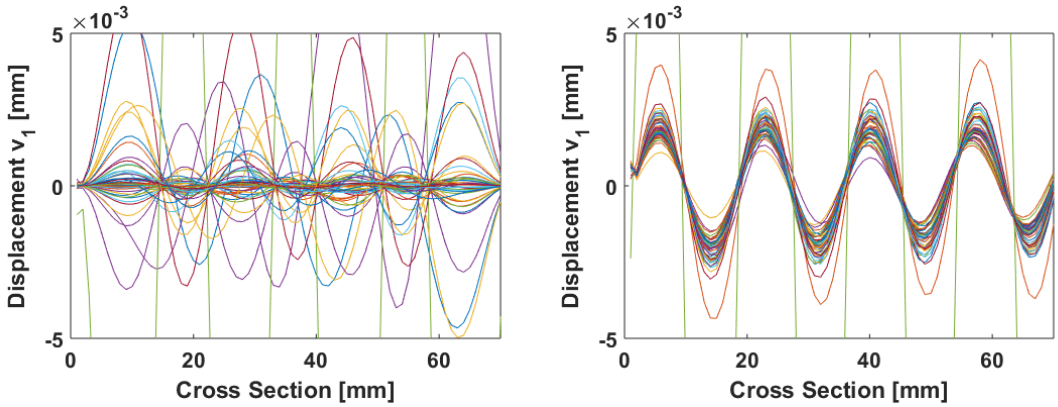


Figure 5.7.5: Cross-sectional plot in the location indicated in Figure 5.2.1 of the wave fields $\mathcal{G}(\mu_0^{(j)})$ of some prior ensemble $\mu_0^{(j)}$ with various variances. **Left:** Variance $\sigma^2 = 10$. **Right:** Variance $\sigma^2 = 0.1$

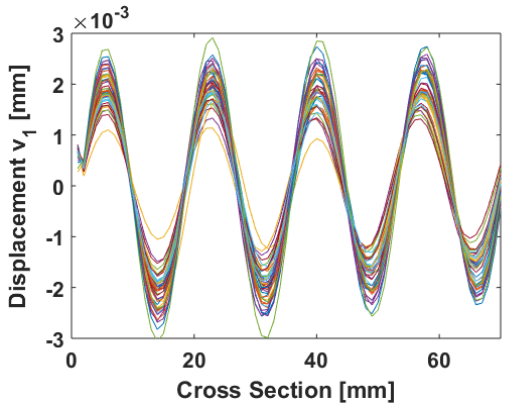


Figure 5.7.6: Wave fields $\mathcal{G}(\mu^{(j)})$ after all members $\mu^{(j)}$ that produced resonating wave fields in Figure 5.7.5 (left) have been removed.

5.7.4 *Viscoelastic Model*

Due to the limitations and problems discussed in Section 5.7 so far, we use a viscoelastic model for MRE in the following chapter. Viscoelasticity takes into account the resistance of the solid soft tissue to deformation and wave fields in this model are damped. A viscoelastic model is capturing more accurately the behaviour of real world soft tissue than the purely elastic model [57].

In Fig. 5.7.2 we show the real part of the displacement field in horizontal direction deploying the viscoelastic model and we can clearly see the damping in the wave field. As a result we do not have highly oscillating wave fields

described in Section 5.7 and also we do not observe pollution in the numerical forward solution.

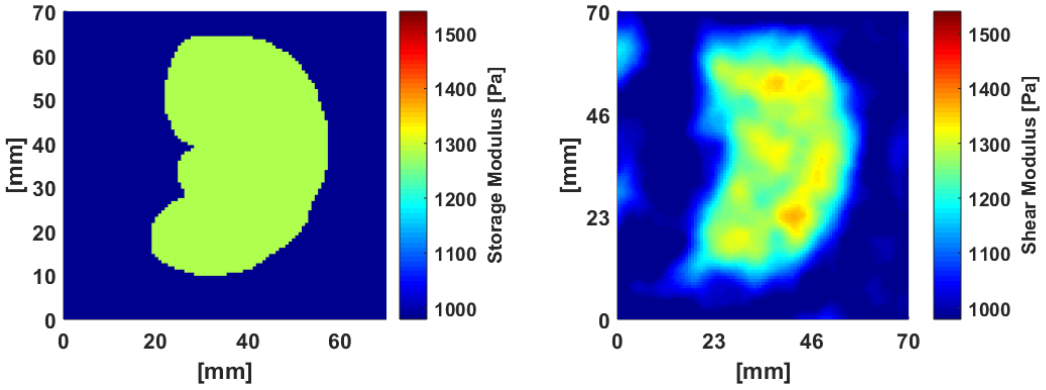


Figure 5.7.7: Setup for this experiment described in Section 5.1 at a frequency of 60Hz. This experiment however, was run with the viscoelastic model (3.2.1) - (3.2.3) with constant $\mu_L = 200$Pa throughout the domain. **Left:** Ground Truth Storage Modulus **Right:** Posterior ensemble mean.

The experiments described in Chapter 5 could be run at 60Hz when deploying this new model. In Figure 5.7.7 we show plots for the experiment described in Chapter 5 but we use the viscoelastic model with a constant loss modulus $\mu_L = 200$Pa throughout the domain and a frequency of 60Hz. In Figure 5.7.8 we show cross-sectional plots of the prior and the posterior ensemble. It can be seen that the values of the ground truth are within the values of the prior across the whole cross section. Also, it can be seen how the posterior ensemble gives a smooth approximation of the ground truth.

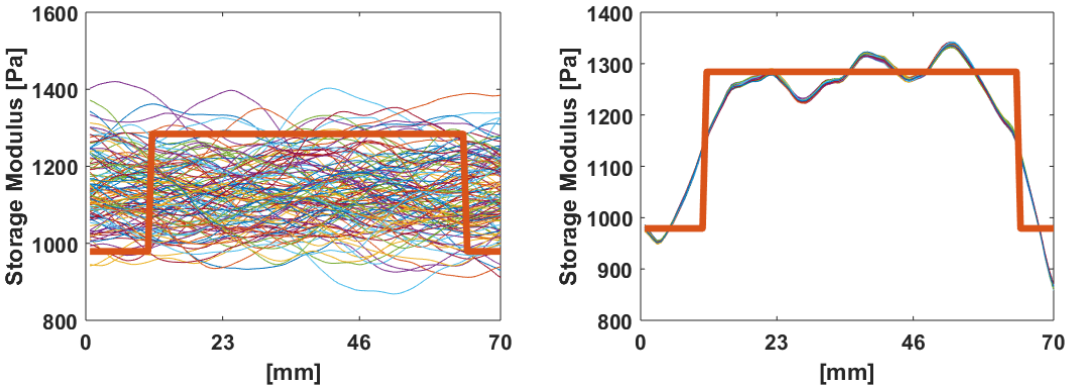


Figure 5.7.8: Cross-Sectional Plot of Prior Ensemble (**left)**) and Posterior Ensemble (**right**; after 3 iterations) for the experiment in Figure 5.7.7.

6

NUMERICAL TESTING - VISCOELASTIC MODEL

The application of a purely elastic model uses a real valued Cauchy stress tensor (3.2.5) and is therefore a relatively simple model. However, this is not a very accurate model of the real world where soft tissue exhibits viscoealstic behaviour [80]. Furthermore, in experiments in Section 5.7.1, we could show that in comparison with the purely elastic model the Galerkin method applied to a viscoelastic model is less polluted and does not allow for highly resonating wave fields. We show in this Chapter that the application of a viscoelastic model enables us to apply EKI to high frequency MRE.

Viscosity of a solid is a measure of its resistance to deformation at a given rate [57]. In the experiments in Chapter 5 we assume a viscosity free solid which simplifies the Cauchy stress tensor.

In this section we test the performance of EKI Algorithm 1. Throughout this Chapter we use the viscoelastic model, i.e. $\mu_L \neq 0$ in (3.2.1) - (3.2.3). Also, we use the two-level-parametrisation $\mathcal{P}_2$ introduced in Section 4.5. Furthermore,

in this Chapter we run experiments at common MRE frequencies [23] between 50Hz and 60Hz.

In Section 6.1 we discuss implementation aspects and how we measure performance of EKI. In Section 6.2 we discuss numerical results and analyse the performance of EKI for various ensemble sizes J, the influence of different noise levels σ_N in the data and influence of various priors. Section 6.3 looks at that posterior ensemble that we receive from the application of EKI and calculates sample variances, confidence intervals and probabilities of tumour. Section 6.4 compares the performance of EKI with direct Inversion. In Section 6.5 we show some ideas how to further improve the prior for EKI applied to MRE by combining it with direct inversion.

Throughout this chapter we use the two-level parametrisation introduced in Section 4.5 and we calculate Gaussian random fields using the stochastic PDE (4.5.6) and not Karhunen-Loève expansions.

6.1 IMPLEMENTATION

Algorithm 1 was implemented in MATLAB using the two-level-parametrisation $\mathcal{P}_2$ as defined in (4.5.7). So $\mathcal{F} = \mathcal{G} \circ \mathcal{P}_2$ is the formal problem used in Algorithm 1. The forward problem $\mathcal{F}(k)$ requires approximations of solutions of the

boundary value problem (3.2.1)-(3.2.3). These approximations are calculated by using MATLAB's inbuilt **Partial Differential Equation Toolbox** which is an implementation of the finite element method. The main reason for preferring this toolbox over ABAQUS FEA is that the toolbox is much faster than ABAQUS FEA in our application.

We use the function *generateMesh* to create a mesh with triangular elements with maximum edge length of 0.2cm and linear shape functions. The domain size is 18.5cm by 15.7cm for the brain experiment and 18cm by 18cm for the kidney experiment. *generateMesh* creates in this setup a mesh with 10502 nodes and 20622 elements for the brain and 5266 nodes and 10262 elements for the kidney.

The function **solvepde** returns the finite element approximation at the nodes and we interpolate this result onto a regular 100 by 100 rectangular grid in the domain Ω. Also, the centred Gaussian noise w_α in the ensemble $\{k_n^{(j)}\}$ inversion from Algorithm 1 is considered on the 100 by 100 rectangular mesh and interpolated onto the mesh generated by *generateMesh* for the forward solution. As shown in Section 5.7.1 we only observe little pollution for the viscoelastic model (3.2.1)-(3.2.3) for these resolutions. Measurements $\mathbf{v} \in \mathbb{R}^m$ are given on the same mesh of 10000 nodes, so $m = 40000$ in (3.4.2).

EKI Algorithm 1 gives us a sequence of ensembles $\{k_n^{(j)}\}$ where $n = 0$ corresponds to the prior ensemble and $n = n_S$ is defined to be the posterior ensemble, i.e. the ensemble that we get after the last iteration before the stopping criteria is met.

We are interested in approximating the shear modulus μ. We estimate this quantity by the physical ensemble mean or level set posterior mean given by

$$\overline{\mu_{N_S}} = \mathcal{P}_2\left(\frac{1}{J}\sum_{j=1}^{J} k_n^{(j)}\right).$$

We often use the level set posterior mean $\overline{\mu_{N_S}}$ in plots.

We measure the accuracy of our estimate by comparing it to the ground truth shear modulus $\mu^\dagger$ given by

$$\mu^\dagger = \mathcal{P}_2(k^\dagger),$$

where $k^\dagger$ is the ground truth that will describe in the next section. More precisely, we monitor the relative error with respect to the truth which is

$$\varepsilon_n = \frac{\|\overline{\mu_{N_S}} - \mu^\dagger\|_{L_2(\Omega)}}{\|\mu^\dagger\|_{L_2(\Omega)}}. \tag{6.1.1}$$

Also, we consider the average data misfit

$$\text{ADM} = \frac{1}{J}\|B^{-1/2}(\mathbf{v} - \mathcal{G}(\overline{k_n^{(j)}}))\|^2,$$

which indicates poor inversion results if the average data misfit is high above 1 after the stopping criterion of Algorithm 1 is met.

6.1.1 *Basic Setup and Synthetic Data (Brain)*

In the brain experiment we assume we have five types of tissue: Grey matter D_2, white matter D_3, background tissue D_1, cerebrospinal fluid matter (CSF) D_4 and tumour tissue D_5. We partition the domain into the four non-tumour tissues and add two tumours in random locations within grey, white and CSF matter. We use the two-level parametrisation introduced in Section 4.5 and the parameters as in Table 6.1.1

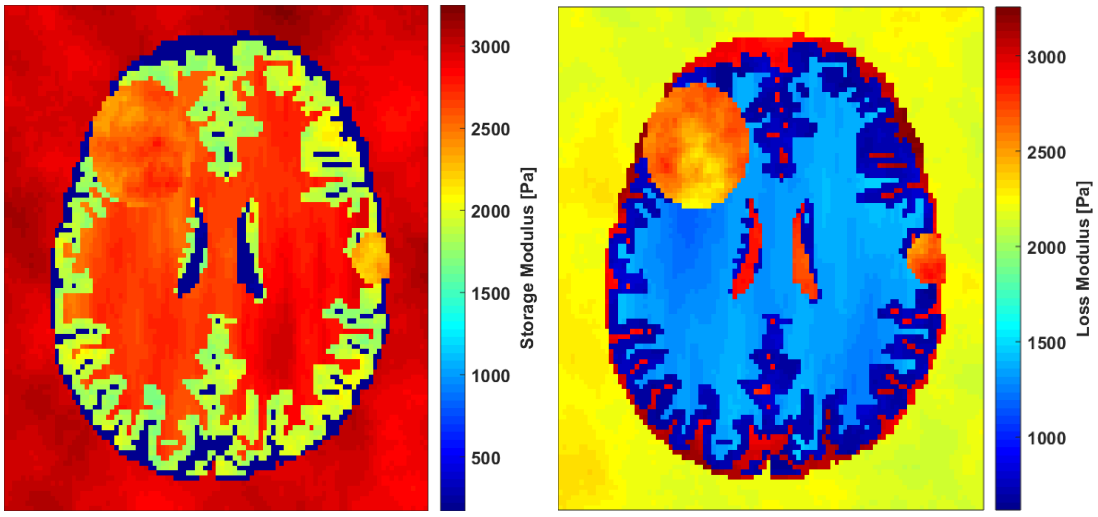
Table 6.1.1: Parameters of the ground truth for the brain experiment.

	$r_{\alpha,S}$, $r_{\alpha,L}$	$\sigma_{\alpha,S}$, $\sigma_{\alpha,L}$	$m_{\alpha,S}$	$m_{\alpha,L}$	$l_{\alpha,S,1}$, $l_{\alpha,L,1}$	$l_{\alpha,S,2}$, $l_{\alpha,L,2}$
Grey Matter	1	0.05	1865	723	0.035	0.035
White Matter	1	0.035	2688	1333	0.3	0.075
Background	1	0.025	3000	2992	0.1	0.1
CSF	1	0.05	212.2	2200	0.15	0.15
Tumour Tissue	1	0.05	2360	2553	0.05	0.05

For all types of tissue $\alpha \in \{1, \ldots, 5\}$, we choose $w_{\alpha S}, w_{\alpha,L} \sim \mathcal{N}(0,1)$. Note that smoothness r_α, amplitude scale σ_α, and length scale $l_{\alpha,1}$ and $l_{\alpha,2}$ are the same for both loss and storage modulus parametrisation and the parameters

only differ for the means $m_{\alpha,S}$, $m_{\alpha,L}$ of the fields. In Figure 6.1.1 (b) we show a plot of a resulting ground truth distribution of the loss modulus and in Figure 6.1.1 (a) of the distribution of the storage modulus.

It is worth noting at this point that our ground truth distributions of storage and loss moduli is only an approximation to real-world MRE. The human brain is usually encased in the skull which is surrounded by air. In brain MRE mechanical waves are delivered by vibrating the skull. Our setup models an *ex vivo* (i.e. outside an organism) brain that has been set into a gel (background tissue D_1). In this case the CSF D_4 would drain away, however.



(a) Ground truth storage modulus

(b) Ground truth loss modulus

Figure 6.1.1: Ground Truth Distribution of Shear Modulus

The examples in this chapter were all run with a frequency of 50Hz ($\omega \approx 314\text{rad}/s$). Due to the absence of high pollution levels for the viscoelastic

model discussed in Section 5.7.1, we are able to run experiments now at high frequencies, even higher than 50Hz.

The top edge of the domain Ω is a loading edge. Harmonic deformation of frequency 50Hz and amplitude $\hat{v}_{\text{top}} = 0.001\text{mm}$ is applied. The other three edges are traction free boundaries.

The domain is 157mm wide and 185mm high. Measurements $\mathbf{v} = v^\dagger + \eta$ are created by using the finite element method approximation $v^\dagger$ of $\mathcal{G}(k_F^\dagger)$ for ground truth $k_F^\dagger$ as defined above. The noise η is drawn again from a Gaussian distribution $\eta \sim \mathcal{N}(0, B)$ as defined in (5.2.1).

6.1.2 *Prior (Brain)*

In the first series of experiments we apply EKI to synthetic data described in Section 6.1.1. We leave $\sigma_{\alpha,S}$, $\sigma_{\alpha,L}$, $r_{\alpha,S}$ and $r_{\alpha,L}$ fixed to the values of the ground truth, i.e. both in the prior and in each update we use $\sigma_{\alpha,S} = \sigma_{\alpha,L} = 1$ and $r_{\alpha,S}$ and $r_{\alpha,L}$ as in the table for the ground truth in Section 6.1.1. For each of the five types of tissue the unknown consists of the two length scales $l_{\alpha,1}$ and $l_{\alpha,2}$, the mean m_α and the function w_α that we discretise on a 100×100 grid. Each of these parameters exists twice, once for the parametrisation of the loss modulus and once for the parametrisation of the storage modulus.

Also, we have the level set functions $\mathbb{k}_C$ as unknown which is used to parametrise the location of the cancer. For $\mathbb{k}_C$ only $l_{C,1}$, $l_{C,2}$ and the function w_C that we discretise on a 100×100 grid are unknown. The mean is chosen to be $m_C = 0$ for this random field.

The total dimension of the unknown is $\dim(\mathcal{K}) = 2 \cdot 5 \cdot 10003 + 10000 + 2 = 110032$.

We select an initial ensemble by drawing for each tissue type $\alpha = 1, \dots, 4$ the parameters for loss and storage modulus $k^{(j)}_{0,\alpha,S} := (m_{\alpha,S}, l_{\alpha,S,1}, l_{\alpha,S,2}, w_{\alpha,S})$ and $k^{(j)}_{0,\alpha,L} := (m_{\alpha,L}, l_{\alpha,L,1}, l_{\alpha,L,2}, w_{\alpha,L})$ from

$$\begin{aligned} k^{(j)}_{0,\alpha,S} &\sim \mathcal{U}_{[\underline{m_{\alpha,S}}, \overline{m_{\alpha,S}}]} \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1), \\ k^{(j)}_{0,\alpha,L} &\sim \mathcal{U}_{[\underline{m_{\alpha,L}}, \overline{m_{\alpha,L}}]} \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1), \end{aligned} \tag{6.1.2}$$

where we used the uniform distribution $\mathcal{U}_{[a,b]}$ on the interval $[a, b]$. Note that for loss and storage modulus for every tissue type α the length scale in x and y direction is drawn from the same uniform distribution. For the mean we choose the values from Table 6.1.2 depending on the tissue type and loss or storage modulus parametrisation

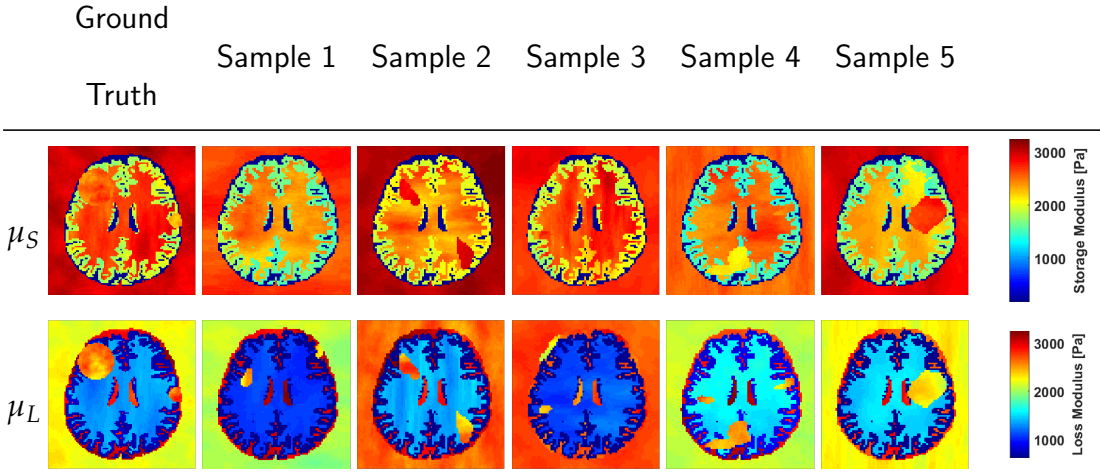
Table 6.1.2: Parameters of the prior for Brain MRE experiments

	$\underline{m_{\alpha,L}}$	$\overline{m_{\alpha,L}}$	$\underline{m_{\alpha,S}}$	$\overline{m_{\alpha,S}}$
Grey Matter	500	900	1500	2000
White Matter	1000	1800	2300	3000
Background Matter	2600	3300	2500	4000
CSF Matter	2000	2800	180	300
Tumour Tissue	2200	2700	2300	3000

Also, a part of the initial ensemble is $k_{0,C}^{(j)} \sim \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1)$ for the parametrisation of the location of the cancer. The mean for this random field is set to zero. We use $c' = 1.5$ in this parametrisation throughout the inversion. Let us denote the prior ensemble as $k_{0,F}^{(j)} := (\{k_{0,\alpha,S}^{(j)}\}_{\alpha\in\{1,\dots,5\}}, \{k_{0,\alpha,L}^{(j)}\}_{\alpha\in\{1,\dots,5\}}, k_{0,C}^{(j)})$.

In Table 6.1.3 we show plots of loss and storage moduli that correspond to five members of the resulting prior ensemble, i.e. we plot $\mu_0^{(j)} = \mathcal{P}(k_{0,F}^{(j)})$ for $j \in \{1,\dots,5\}$.

Table 6.1.3: Five members from the prior ensemble $\{\mu_0^{(j)}\}_{j=1}^{J}$ described in this section. Top row: storage modulus. Bottom row: loss modulus.



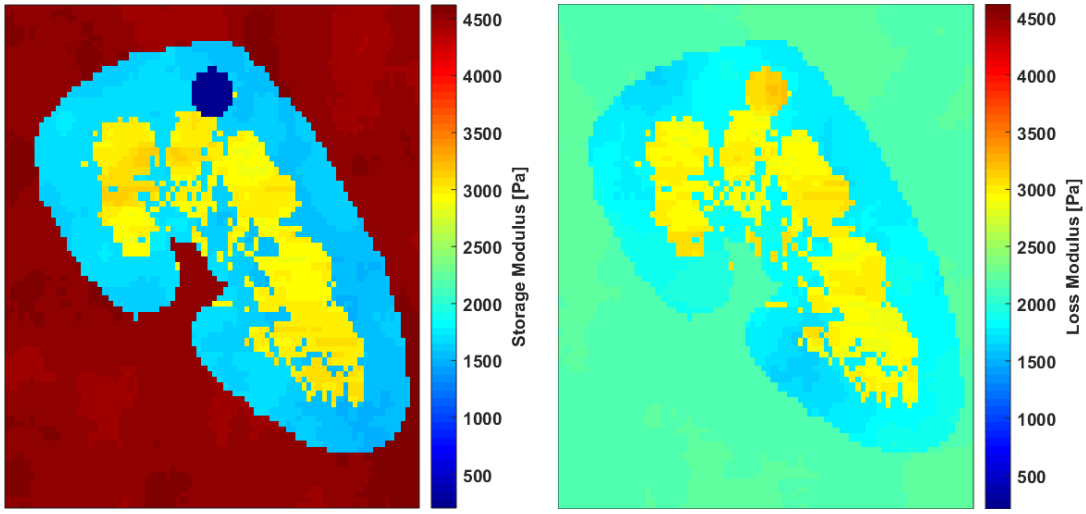
6.1.3 *Basic Setup and Synthetic Data (Kidney)*

In the kidney experiment we assume we have four types of tissue: Cortex tissue D_2, medulla tissue D_3, background tissue D_1 and tumour tissue D_5. We partition the domain into the three non-tumour tissues and include one tumour at a random location within cortex or medulla tissue. We use the two-level parametrisation introduced in Section 4.5, assume $D_4 = \emptyset$ and use the parameters from Table 6.1.4

Table 6.1.4: Parameters of the ground truth for the brain experiment.

	$r_{\alpha,S}$, $r_{\alpha,L}$	$\sigma_{\alpha,S}$, $\sigma_{\alpha,L}$	$m_{\alpha,S}$	$m_{\alpha,L}$	$l_{\alpha,S,1}$, $l_{\alpha,L,1}$	$l_{\alpha,S,2}$, $l_{\alpha,L,2}$
Cortex Tissue	1	0.05	3000	3000	0.15	0.15
Medulla Tissue	1	0.025	1600	1800	0.035	0.1
Background	1	0.01	4500	2200.4	0.1	0.1
Tumour Tissue	1	0.035	212	2292.5	0.1	0.05

For all types of tissue $\alpha \in \{1, \ldots, 4\}$, we choose $w_{\alpha,S}, w_{\alpha,L} \sim \mathcal{N}(0,1)$. Like in the ground truth for the brain, we note that smoothness r_α, amplitude scale σ_α, length scales $l_{\alpha,1}$ and $l_{\alpha,2}$ are the same for both loss and storage modulus parametrisation and the parameters only differ for the means $m_{\alpha,S}$, $m_{\alpha,L}$ of the fields. In Figure 6.1.2 (b) we show a plot of a resulting ground truth distribution of the loss modulus and in Figure 6.1.2 (a) of the distribution of the storage modulus.



(a) Ground truth storage modulus (b) Ground truth loss modulus

Figure 6.1.2: Ground Truth Distribution of Shear Modulus

The examples for the kidney were all run with a frequency of 60Hz ($\omega \approx 372\text{rad}/s$). Due to the absence of high pollution levels for the viscoelastic model discussed in Section 5.7.1, we are able to run experiments now at high frequencies, even higher than 60Hz.

The top edge of the domain Ω is a loading edge. Harmonic deformation of frequency 60Hz and amplitude $\hat{v}_{\text{top}} = 0.001\text{mm}$ is applied. The other free edges are traction free boundaries.

The domain is 120mm wide and 120mm high. Measurements $\mathbf{v} = v^\dagger + \eta$ are created by using the finite element approximation $v^\dagger$ of $\mathcal{G}(k_F^\dagger)$ for ground truths $k_F^\dagger$ as defined above. The noise η is drawn again from a Gaussian distribution $\eta \sim \mathcal{N}(0, B)$ as defined in (5.2.1).

6.1.4 *Prior (Kidney)*

In the second series of experiments we apply EKI to synthetic data described in Section 6.1.3. Although, EKI can be used to estimate smoothness parameter r_α [38] and amplitude scale σ_α, for simplicity we leave $\sigma_{\alpha,S}$, $\sigma_{\alpha,L}$ and $r_{\alpha,S}$, $r_{\alpha,L}$ fixed to the values of the ground truth, i.e. both in the prior and in each update we use $\sigma_{\alpha,S}, \sigma_{\alpha,L} = 1$ and $r_{\alpha,S}$, $r_{\alpha,S}$ as in Table 6.1.4 for the ground truth in Section 6.1.3. This is due to simplicity. In Chapter 7 we discuss that experiments might be interesting that show that spatial variability can be captured by w_α alone and analyse and quantify how various fixed choices of σ_α and r_α in the prior affect the posterior.

For each of the four types of tissue the unknown consists of the two length scales $l_{\alpha,1}$ and $l_{\alpha,2}$, the mean m_α and the function w_α that we discretise on a 100×100 grid. We need each of these parameters twice, once for the parametrisation of the loss modulus and once for the parametrisation of the storage modulus.

Also, we have the level set functions $\Bbbk_C$ as unknown which is used to parametrise the location of the cancer. For $\Bbbk_C$ only $l_{C,1}$, $l_{C,2}$ and the function w_C that we discretise on a 100×100 grid are unknown. The mean is chosen to be $m_C = 0$ for this random field.

The total dimension of the unknown is $\dim(\mathcal{K}) = 2 \cdot 4 \cdot 10003 + 10000 + 2 = 90032$.

In the experiments run with kidney MRE data, we select an initial ensemble by drawing for each tissue type $\alpha = 1, \ldots, 4$ the parameters $k^{(j)}_{0,\alpha,S} := (m_{\alpha,S}, l_{\alpha,S,1}, l_{\alpha,S,2}, w_{\alpha,S})$ and $k^{(j)}_{0,\alpha,L} := (m_{\alpha,L}, l_{\alpha,L,1}, l_{\alpha,L,2}, w_{\alpha,L})$ for loss and storage modulus from

$$k^{(j)}_{0,\alpha,S} \sim \mathcal{U}_{[\underline{m_{\alpha,S}}, \overline{m_{\alpha,S}}]} \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1), \tag{6.1.3}$$

$$k^{(j)}_{0,\alpha,L} \sim \mathcal{U}_{[\underline{m_{\alpha,L}}, \overline{m_{\alpha,L}}]} \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1), \tag{6.1.4}$$

where we used the uniform distribution $\mathcal{U}_{[a,b]}$ on the interval $[a, b]$. Note that for loss and storage modulus for every tissue type α the length scale in x and y direction is drawn from the same uniform distribution. For the mean we choose the values from Table 6.1.5 depending on the tissue type and loss or storage modulus parametrisation

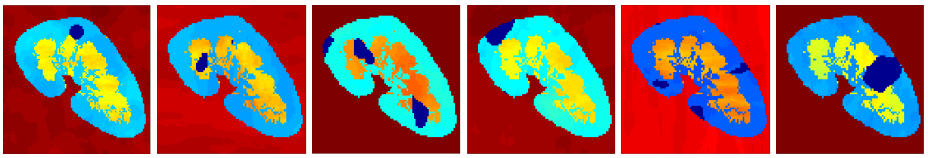
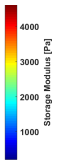
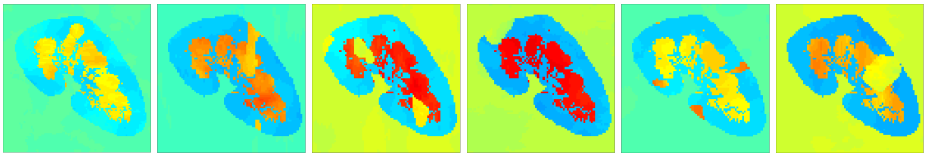
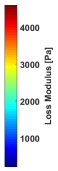
Table 6.1.5: Parameters of the prior for the Kidney MRE experiments

	$\underline{m_{\alpha,L}}$ $\overline{m_{\alpha,L}}$ $\underline{m_{\alpha,S}}$ $\overline{m_{\alpha,S}}$
Cortex Tissue	2500 3500 2500 3500
Medulla Tissue	1000 1800 1500 2200
Background Matter	4000 6200 2000 4000
Tumour Tissue	2000 2700 100 300

Also, a part of the initial ensemble is $k_{0,C}^{(j)} \sim \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1)$ for the parametrisation of the location of the cancer. The mean for this random field is set to zero. We use $c' = 1.5$ in this parametrisation throughout the inversion. Let us denote the prior ensemble as $k_{0,F}^{(j)} := (\{k_{0,\alpha,S}^{(j)}\}_{\alpha\in\{1,\dots,4\}}, \{k_{0,\alpha,L}^{(j)}\}_{\alpha\in\{1,\dots,4\}}, k_{0,C}^{(j)})$.

In Table 6.1.6 we show plots of loss and storage moduli that correspond to five members of the resulting prior ensemble, i.e. we plot $\mu_0^{(j)} = \mathcal{P}(k_{0,F}^{(j)})$ for $j \in \{1,\dots,5\}$.

Table 6.1.6: Five members from the prior ensemble $\{\mu_0^{(j)}\}_{j=1}^J$ described in this section. Top row: storage modulus. Bottom row: loss modulus.

	Ground Truth	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	
μ_S							
μ_L							

6.2 INFLUENCE OF PRIOR, NOISE AND ENSEMBLE SIZE ON THE POSTERIOR

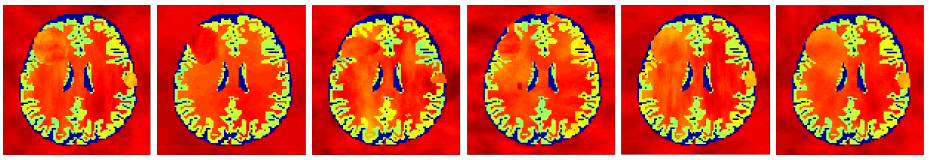
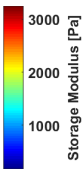
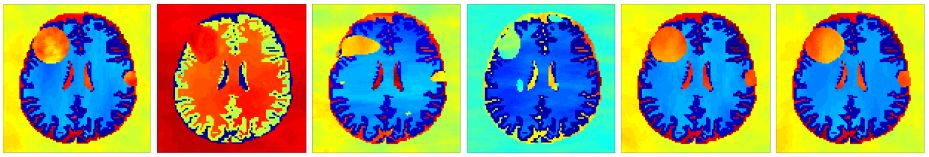
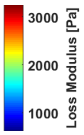
In this section we conduct experiments that analyse the affect of the prior, the noise and the ensemble size on the posterior ensemble mean in the brain MRE experimental setup described in Section 6.1.1. In Section 6.2.1 we run EKI for various ensemble sizes J. In Section 6.2.2 we determine the effect of the noise level σ_n in the data on the posterior ensemble mean. In Section 6.2.3 we examine performance of EKI for a range of $m_{\alpha,S}$ and $m_{\alpha,L}$ in the prior which entails a range of physical ensemble means. In Section 6.3 we will then also consider quantities like the sample variance and confidence intervals of the posterior.

6.2.1 *Ensemble Size*

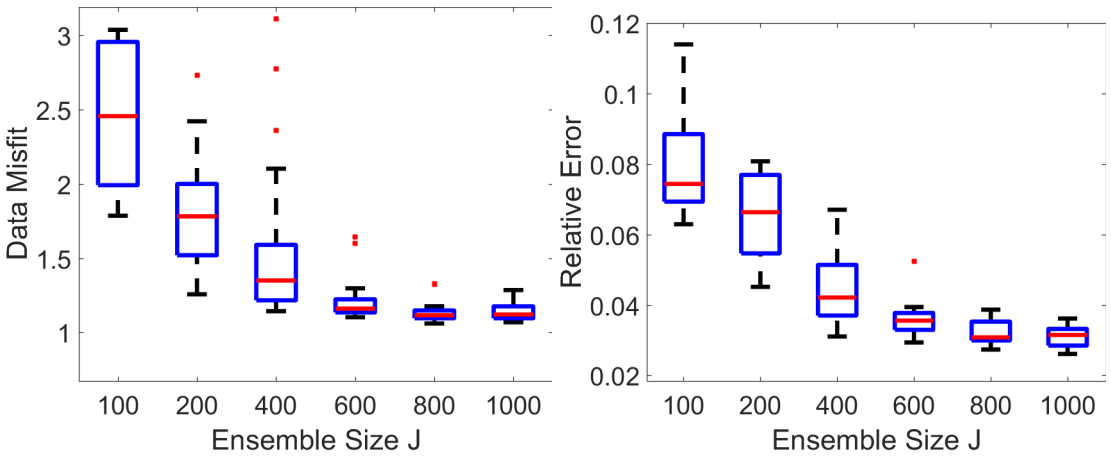
We run Algorithm 1 in the setup described in Section 6.1.1 with prior described in Section 6.1.2 for different choices of ensemble sizes J : 50, 100, 200, 400, 600, 800, 1000. For each choice of J, we run 20 experiments with different random selections of the initial ensemble.

In Table 6.2.1 we show plots of the physical posterior ensemble mean μ_{N_S} and the ground truth $\mu^\dagger$ for one of these experiments for the brain. A noise level of 5% is used in this series. In most experiments, stopping criterion is met after $N_S = 6$ iterations for sample size $J = 50$ and after $N_S = 4$ iterations for the other sample sizes.

Table 6.2.1: Physical ground truth shear modulus $\mu^\dagger$ and physical posterior ensemble means $\overline{\mu_{N_S}}$ for various ensemble sizes and a noise level $\sigma_N = 0.05$. Top row: Storage Modulus. Bottom row: Loss modulus.

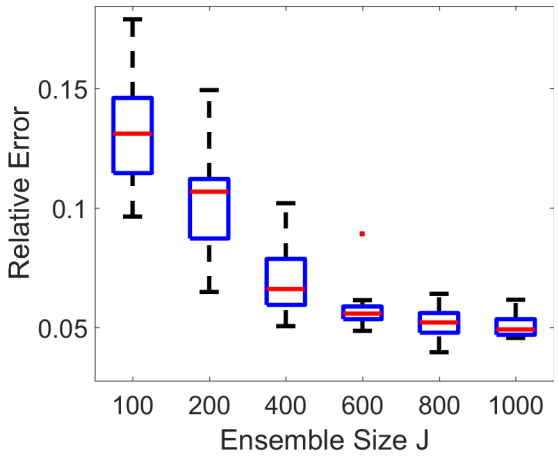
	Ground Truth	$J = 50$	$J = 100$	$J = 200$	$J = 400$	$J = 1000$	
μ_S							
μ_L							

It can be seen that for ensemble size $J = 50$ the second (smaller) tumour is not captured by the inversion. For $J = 200$ the small tumour is located at a wrong location. This accuracy of the EKI estimate for this ensemble size is clearly too poor. Also, the error in the EKI-reconstruction seems to be significant in areas of white matter and tumour. For $J = 100$ and $J = 200$, there seems to be a significant error in the shape of the big tumour. From these plots an ensemble size of at least $J = 400$ is necessary to capture all the details accurately.



(a) Brain: Average Data Misfit over Ensemble Size

(b) Brain: relative error of storage modulus w.r.t ground truth over Ensemble size



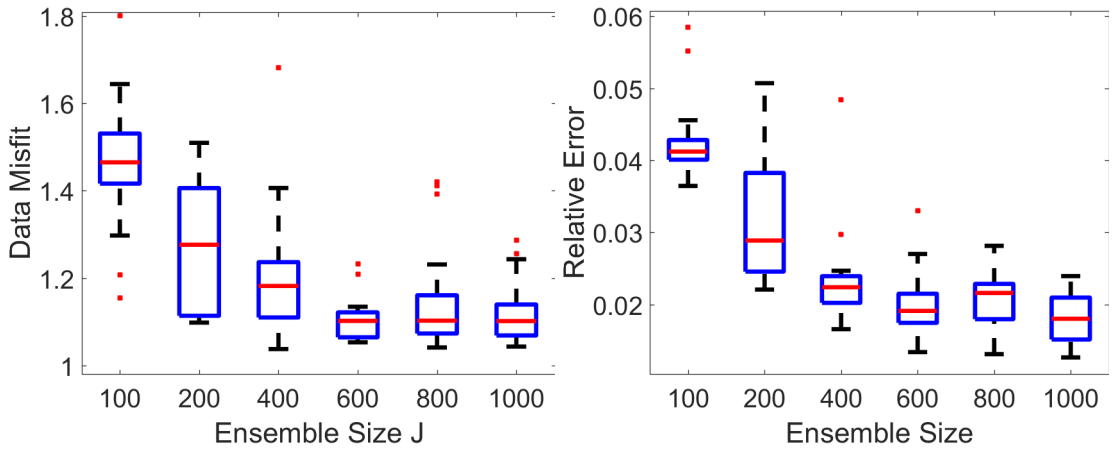
(c) Brain: relative error of loss modulus w.r.t ground truth over Ensemble size

Figure 6.2.1: Brain MRE. Noise level is 0.05. Boxplots for relative errors in 20 experiments.

In Figure 6.2.1 we show box plots of the data misfit and the relative error of the storage and loss modulus over the various ensemble sizes for the brain.

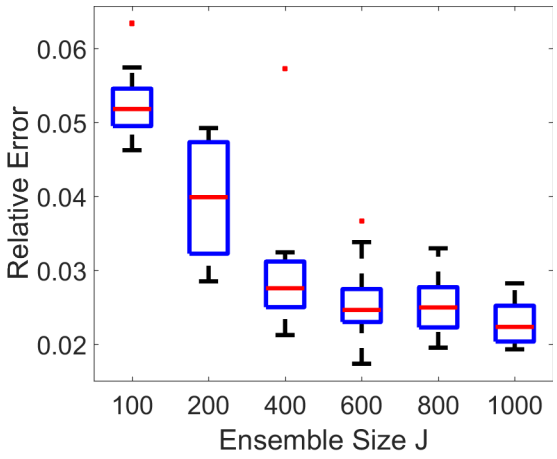
These plots confirm a large rate of error in the case of ensemble size $J = 100$ and $J = 200$. We can conclude that an ensemble size of $J = 600$ is a good trade-off between computational costs and accuracy. Also, the relative error with respect to (w.r.t) the truth for $J = 600$ barely improves as we increase J.

We run the same series of experiments with the same ensemble sizes J : 50, 100, 200, 400, 600, 800, 1000 and and noise levels 5% for the kidney. The setup is described Section 6.1.3 and we use the prior described in Section 6.1.4. We run 20 experiments with random selections of the prior ensemble.



(a) Kidney: Average Data Misfit over Ensemble Size

(b) Kidney: relative error of storage modulus with ground truth over Ensemble size



(c) Brain: relative error of loss modulus with ground truth over Ensemble size

Figure 6.2.2: Kidney MRE. Noise level is 0.05. Box-plots for relative errors and data misfits in 20 experiments.

In Figure 6.2.2 we show box plots of the data misfit and the relative error of the storage and loss modulus over the various ensemble sizes for the kidney.

The results are similar to the results in the brain experiments: These plots show large errors for ensemble size $J = 100$ and $J = 200$. Although, the accuracy for $J = 400$ is higher than for the brain experiment we can conclude that an ensemble size of $J = 600$ is a good trade-of between computational costs and accuracy. Also, the relative error with respect to the truth for $J = 600$ barely improves as we increase J.

6.2.2 *Results from Inversion Using Various Noise Levels (Brain)*

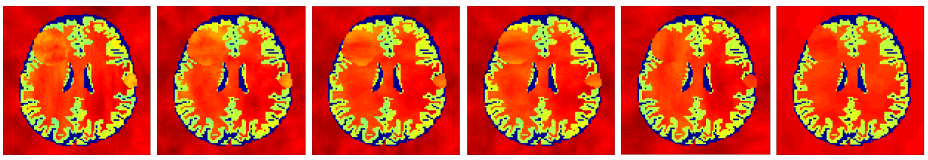
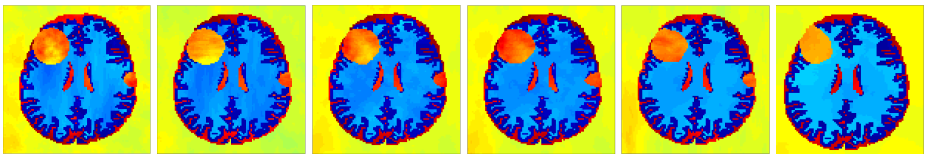
In this series of 20 experiments we run Algorithm 1 with data coming from the basic setup described in Section 6.1.1 for various levels of noise σ_n : 0.01, 0.05, 0.1, 1, 32 corresponding to a noise level of 1%, 5%, 10%, 100%, 3200%. A constant ensemble size of $J = 600$ is used throughout this series of experiments. We use an initial ensemble defined in Section 6.1.2 which is the same for all noise levels but is chosen randomly for each of the 20 experiments.

In Table 6.2.2 we show plots of the physical posterior ensemble mean $\overline{\mu_{N_S}}$ and the ground truth $\mu^\dagger$ for one of these experiments for the brain. Up to a noise level of 100% we can see a high accuracy in the reconstruction. For $\sigma_n = 32$ the shape of the big tumour is inaccurate and the small tumour cannot

be seen. We conclude that Algorithm 1 cannot cope with this noise level in this application.

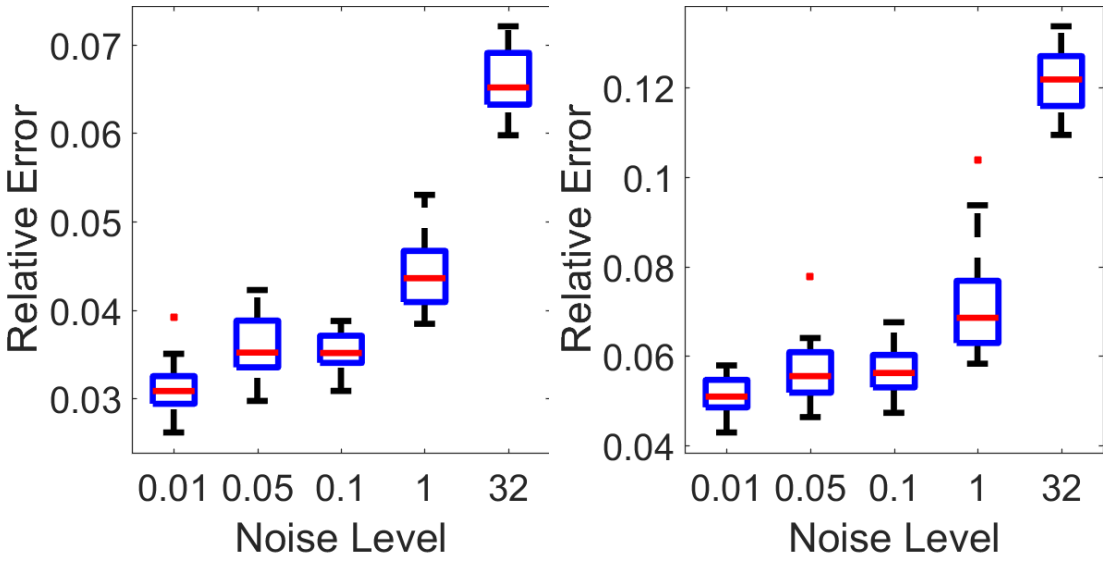
In most experiments, stopping criterion is met after $N_S = 6$ iterations for noise level $\sigma_n = 0.01$, after $N_S = 3$ iterations for $\sigma_n = 32$ and after $N_S = 4$ iterations for the other noise levels.

Table 6.2.2: Physical ground truth shear modulus $\mu^\dagger$ and physical posterior ensemble means $\overline{\mu_{N_S}}$ for various noise levels and constant ensemble size $J = 600$. Top row: Storage Modulus. Bottom row: Loss modulus.

	Ground Truth	$\sigma_n = 0.01$	$\sigma_n = 0.05$	$\sigma_n = 0.1$	$\sigma_n = 1$	$\sigma_n = 32$
μ_S						
μ_L						

In Figure 6.2.3 we show box plots of the data misfit and the relative error of the storage and loss modulus with respect to the ground truth over the various noise levels for the brain. These plots confirm a large rate of error for noise level $\sigma_n = 32$ and a small rate of error for noise levels $\sigma_N = 0.01, 0.05, 0.1$ in

the data. We can conclude that for a noise level below $\sigma_n = 1$ that is 100% or $SNR = 1$, EKI delivers accurate estimates.



(a) Brain: relative error of storage modulus with ground truth over Noise level

(b) Brain: relative error of loss modulus with ground truth over Noise level

Figure 6.2.3: Brain MRE. Ensemble Size is 600. Box-plots for relative errors in 20 experiments.

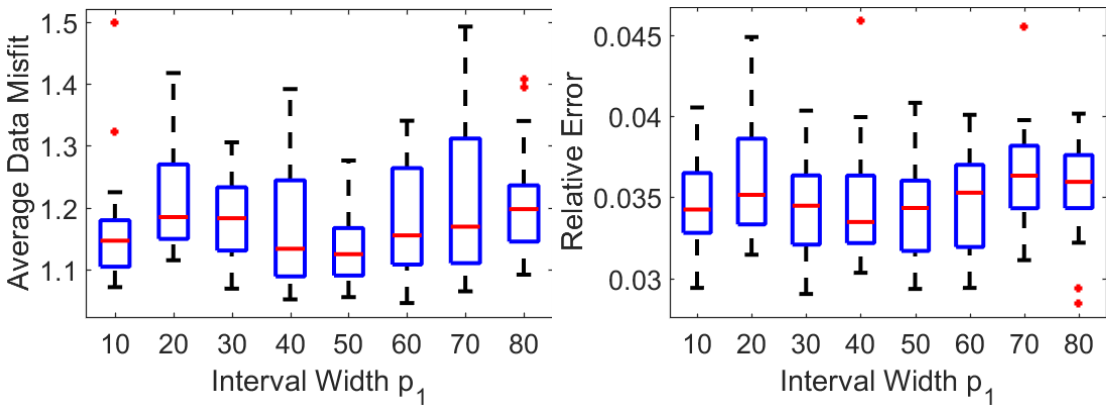
6.2.3 *Results from Inversion using Various Priors*

In this series of 20 experiments we run Algorithm 1 with data coming from the basic setup described in Section 6.1.1 for various priors. A constant ensemble size of $J = 600$ noise level $\sigma_n = 0.05$ in the data is used throughout this series of experiments.

For each of the 20 experiments we draw an initial ensemble as defined in (6.1.2) and in this series we do not use the fixed uniform distributions on the intervals defined in Table 6.1.2. Instead, we vary the size of the interval and define this interval with the ground truth value defined in Table 6.1.1 as the midpoint of the interval. More specifically in (6.1.2) we set

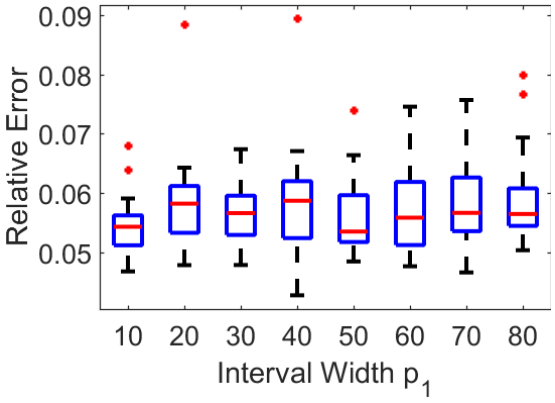
$$\underline{m_{\alpha,L}} = m_{\alpha,L} - \frac{p_1}{100} \cdot m_{\alpha,L}, \qquad \overline{m_{\alpha,L}} = m_{\alpha,L} + \frac{p_1}{100} \cdot m_{\alpha,L},$$
$$\underline{m_{\alpha,S}} = m_{\alpha,S} - \frac{p_1}{100} \cdot m_{\alpha,S}, \qquad \overline{m_{\alpha,S}} = m_{\alpha,S} + \frac{p_1}{100} \cdot m_{\alpha,S}.$$

In Figure 6.2.4 we show box plots of the data misfit and the relative error of the storage and loss modulus with respect to the ground truth over the various priors used. We can record only a very slight loss in accuracy for big intervals of the uniform distribution. In Figure 6.2.4 (d) an increase in number of iterations until the stopping criterion is met can be analysed for increasing sizes of the interval $[\underline{m_{\alpha,L}}, \overline{m_{\alpha,L}}]$ and $[\underline{m_{\alpha,S}}, \overline{m_{\alpha,S}}]$.



(a) Brain: Average Data Misfit over different priors.

(b) Brain: relative error of storage modulus with ground truth over different priors.



(c) Brain: relative error of loss modulus with ground truth over different priors.

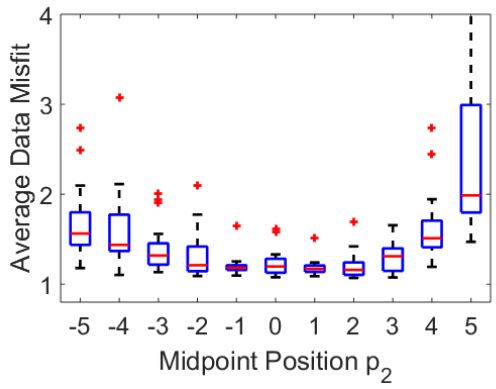
Figure 6.2.4: Performance of EKI applied to brain MRE for various priors defined in this chapter with $p_1 = 10, 20, 30, 40, 50, 60, 70, 80$. Noise level is 0.05, ensemble size is 600. Box-plots for relative errors and data misfits in 20 experiments.

We run this series of experiment also with other priors. For each of the 20 experiments, we draw an initial ensemble as defined in (6.1.2) and in this series

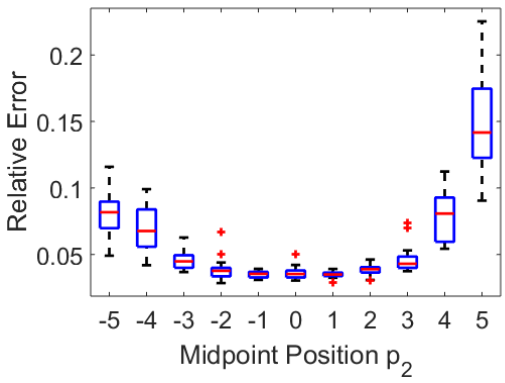
we do not use the fixed uniform distributions on the intervals defined in Table 6.1.2. Instead, we vary the midpoint of a fixed sized interval. More specifically in (6.1.2) we set

$$\begin{aligned}
\underline{m_{\alpha,L}} &= m_{\alpha,L} - \frac{20}{100} \cdot m_{\alpha,L} + \frac{p_2 \cdot 10}{100} \cdot m_{\alpha,L}, \\
\overline{m_{\alpha,L}} &= m_{\alpha,L} + \frac{20}{100} \cdot m_{\alpha,L} + \frac{p_2 \cdot 10}{100} \cdot m_{\alpha,L}. \\
\underline{m_{\alpha,S}} &= m_{\alpha,S} - \frac{20}{100} \cdot m_{\alpha,S} + \frac{p_2 \cdot 10}{100} \cdot m_{\alpha,S}, \\
\overline{m_{\alpha,S}} &= m_{\alpha,S} + \frac{20}{100} \cdot m_{\alpha,S} + \frac{p_2 \cdot 10}{100} \cdot m_{\alpha,S}.
\end{aligned} \tag{6.2.1}$$

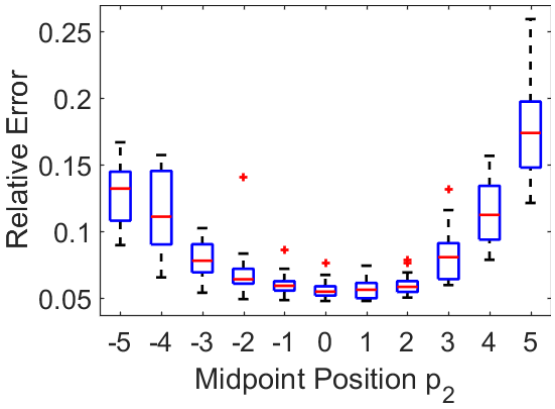
So, in this series we choose the fixed uniform distribution on an interval with $p_1 = 20$ and move the midpoint of the interval it further away from the ground truth value in order to get different priors. It can clearly be seen in Figure 6.2.5 how bad priors, i.e. priors that are far away from the ground truth, result in low accuracy in the posterior ensemble.



(a) Brain: Average Data Misfit over different priors



(b) Brain: relative error of storage modulus with ground truth over different priors



(c) Brain: relative error of loss modulus with ground truth over different priors

Figure 6.2.5: Performance of EKI applied to brain MRE for various priors defined in this chapter with $p_2 = -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5$. Noise level is 0.05, ensemble size is 600. Box-plots for relative errors and data misfits in 20 experiments.

The two series of experiments with results shown in Figure 6.2.4 and Figure 6.2.5 demonstrate that a prior with a big interval of the mean hyperprior should

be preferred over a prior with an interval of the mean hyperprior that does not contain the ground truth. The plots in Figure 6.2.4 show that estimates coming from EKI with a prior with a big interval of the mean hyperprior have similar relative errors as estimates that go back to a prior with small interval of the mean hyperprior. Plots in Figure 6.2.5 illustrate that the relative error drastically increases if the interval of the mean hyperprior does not the contain the unknown which is the case for $p_2 < -2$ or $p_2 > 2$ in (6.2.1).

6.2.4 *Conclusion*

EKI in application to MRE with *in silico* model data as input can provide robust approximations of shear moduli for high noise levels in the data. Prior ensembles with a mean far off the ground truth lead to high relative error between the unknown and the posterior mean. The two worst cases in our experiments in this chapter is if the interval of the mean hyperprior does not contain the unknown and small ensemble size. In these case the relative errors and data misfit have high values. In these cases estimates can not accurately recover the unknown and are not useful for the detection of cancer.

6.3 POSTERIOR VARIANCE AND CONFIDENCE INTERVALS

In this section we will have a closer look at the posterior distribution provided by EKI in the experimental set up for brain MRE described in Section 6.1.1. More precisely, we use the physical ensemble at the n-th iteration

$$\mu_n^{(j)} = \mathcal{P}_2\left(k_n^{(j)}\right),$$

and the **physical ensemble mean**

$$\overline{\mu_n} = \frac{1}{J}\sum_{j=1}^{J} \mu_n^{(j)},$$

in order to calculate the following **sample variance**

$$\mathfrak{s}_n^2 = \frac{1}{J-1}\sum_{j=1}^{J}\left(\mu_n^{(j)} - \overline{\mu_n}\right)^2,$$

and in order to approximate the following 95% **sample confidence intervals**

$$\mathrm{CI}_{95,n} \approx \left[\overline{\mu_n} - \frac{1.96\mathfrak{s}_n}{\sqrt{J}}, \overline{\mu_n} + \frac{1.96\mathfrak{s}_n}{\sqrt{J}}\right].$$

Furthermore, we define a **probability of cancer** which is similarly defined as the "probability to find a defect" in [61]. This quantity reflects the confidence for whether we have cancer tissue in a certain point $x \in \Omega$ or not. For a fixed $x \in \Omega$, the probability of $\Bbbk_C(x)$ in (4.5.2) to be above the user defined threshold c that defines the region of the cancer is

$$P_C(x) = \mathbb{P}(\Bbbk_C(x) > c).$$

We consider $\mathbb{k}_C(x)$ as a random variable with density $\pi_{\mathbb{k}_C}$ and can therefore rewrite

$$P_C(x) = \int_c^\infty \pi_{\mathbb{k}_C}(y)\,\mathrm{d}y = \int_{-\infty}^\infty \mathbb{1}_{y>c}\pi_{\mathbb{k}_C}(y)\,\mathrm{d}y,$$

Using the posterior ensemble $\left(k_n^{(j)}\right)_{j=1}^J$ we can now consider the ensemble with ensemble members defined by

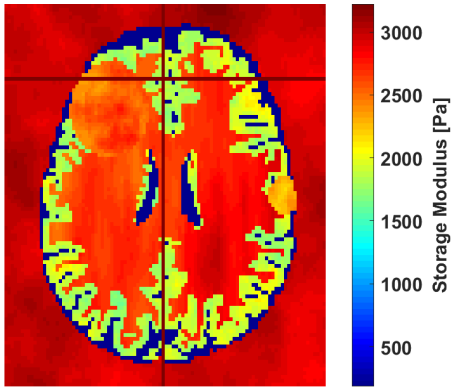
$$\mathbb{k}_C^{(j)} := \mathcal{P}_2^{(1)}(k_n^{(j)}).$$

The ensemble $\left(\mathbb{k}_C^{(j)}(x)\right)_{j=1}^J$ is a sample distribution of $\mathbb{k}_C(x)$ and can be used to approximate the probability of cancer by using (4.5.2) again

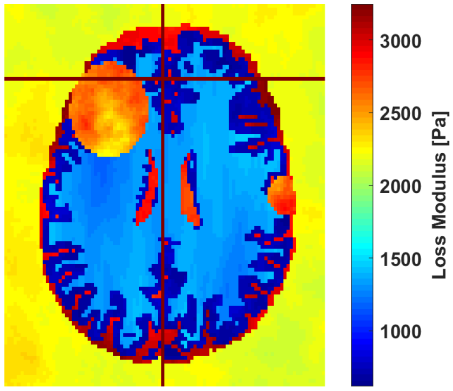
$$P_C(x) = \int_{-\infty}^\infty \mathbb{1}_{y>c}\pi_{\mathbb{k}_C}(y)\,\mathrm{d}y \approx \frac{1}{J}\sum_{j=1}^J \mathbb{1}_{\mathbb{k}_C^{(j)}(x)>c} = \frac{1}{J}\sum_{j=1}^J \mathbb{1}_{D_{\mathbb{k}_C^{(j)}}}(x).$$

We calculate all these quantities for posterior ensembles $\left(k_n^{(j)}\right)_{j=1}^J$ stemming from the application of EKI to synthetic data described in Section 6.1.1. We consider these quantities for different noise levels in the data and a number of different relatively good and bad priors.

In Figure 6.3.1 we give plots of the ground truth. In these plots we also indicate two locations with two lines in which we plot 95% sample confidence intervals. One line covers a location with tumour the other line not.



(a) Ground truth storage modulus $\mu_S^\dagger$



(b) Ground truth loss modulus $\mu_L^\dagger$

Figure 6.3.1: Ground Truth Distribution of Shear Modulus. The two lines indicate the location of the cross-sectional plots.

As before, we leave $\sigma_{\alpha,S}$, $\sigma_{\alpha,L}$, $r_{\alpha,S}$ and $r_{\alpha,L}$ fixed to the values of the ground truth, i.e. both in the prior and in each update we use $\sigma_{\alpha,S} = \sigma_{\alpha,L} = 1$ and $r_{\alpha,S}$ and $r_{\alpha,L}$ as in the table for the ground truth in Section 6.1.1. The unknown consists for each of the five types of tissue of the two length scales $l_{\alpha,1}$ and $l_{\alpha,2}$, the mean m_α and the function w_α that we discretise on a 100×100 grid. Each of these parameters exist twice, one for the parametrisation of the loss modulus and one for the parametrisation of the storage modulus.

Again, we have the level set functions $\Bbbk_C$ as unknown which is used to parametrise the location of the cancer. For $\Bbbk_C$ only $l_{C,1}$, $l_{C,2}$ and the function w_C that we discretise on a 100×100 grid are unknown. The mean is chosen to be $m_C = 0$ for this random field.

The total dimension of the unknown is, as before, $\dim(\mathcal{K}) = 2 \cdot 5 \cdot 10003 + 10000 + 2 = 110032$.

6.3.1 *Good Prior, Large Ensemble and Low Noise Level*

In this experiment we select an initial ensemble by drawing for each tissue type $\alpha = 1, \ldots, 4$ the parameters $k^{(j)}_{0,\alpha,S} := (m_{\alpha,S}, l_{\alpha,S,1}, l_{\alpha,S,2}, w_{\alpha,S})$ and $k^{(j)}_{0,\alpha,L} := (m_{\alpha,L}, l_{\alpha,L,1}, l_{\alpha,L,2}, w_{\alpha,L})$ for loss and storage modulus from

$$\begin{aligned} k^{(j)}_{0,\alpha,S} &\sim \mathcal{U}_{[\underline{m_{\alpha,S}}, \overline{m_{\alpha,S}}]} \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1), \\ k^{(j)}_{0,\alpha,L} &\sim \mathcal{U}_{[\underline{m_{\alpha,L}}, \overline{m_{\alpha,L}}]} \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1), \end{aligned} \tag{6.3.1}$$

where we used the uniform distribution $\mathcal{U}_{[a,b]}$ on the interval $[a, b]$. Note that for loss and storage modulus for every tissue type α the length scale in x and y direction is drawn from the same uniform distribution. For the mean we choose the values from Table 6.3.1 depending on the tissue type and loss or storage modulus parametrisation

Table 6.3.1: Parameters of the prior for Brain MRE experiments

	$\underline{m_{\alpha,L}}$	$\overline{m_{\alpha,L}}$	$\underline{m_{\alpha,S}}$	$\overline{m_{\alpha,S}}$
Grey Matter	508	945	1304	2423
White Matter	933	1732	2300	3000
Background Matter	2600	3300	2500	4000
CSF Matter	2000	2800	180	300
Tumour Tissue	2200	2700	2300	3000

Also, a part of the initial ensemble is $k_{0,C}^{(j)} \sim \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1)$ for the parametrisation of the location of the cancer. The mean for this random field is set to zero. We use $c' = 1.5$ in this parametrisation throughout the inversion. Let us denote the prior ensemble as $k_{0,F}^{(j)} := (\{k_{0,\alpha,S}^{(j)}\}_{\alpha\in\{1,\dots,5\}}, \{k_{0,\alpha,L}^{(j)}\}_{\alpha\in\{1,\dots,5\}}, k_{0,C}^{(j)})$.

We run Algorithm 1 in the setup described in Section 6.1.1 with prior described above with an ensemble sizes $J = 600$. A noise level of 5% is used in this series.

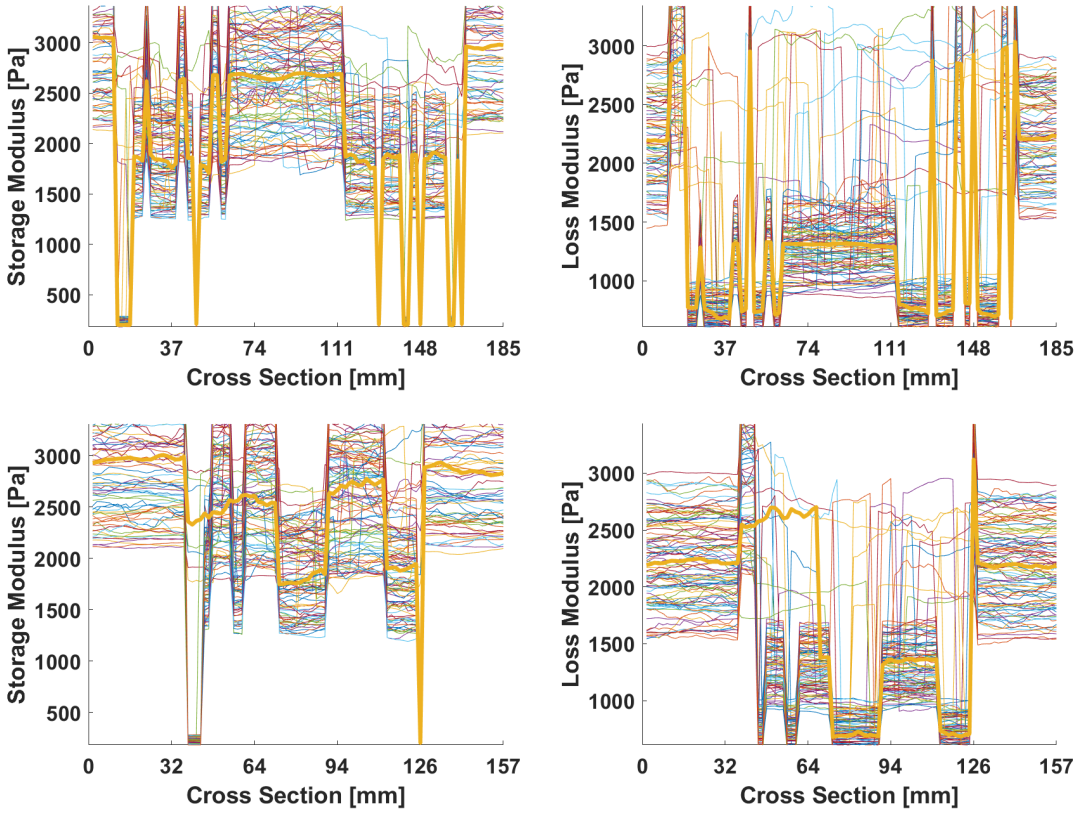


Figure 6.3.2: Prior ensemble $\mu_0^{(j)}$ in cross-sections indicated in Fig. 6.3.1. The yellow line indicates the ground truth $\mu^\dagger$. **Top left:** Prior ensemble for storage modulus in vertical cross-section. **Top right:** Prior ensemble for loss modulus in vertical cross-section. **Bottom left:** Prior ensemble for storage modulus in horizontal cross-section. **Bottom right:** Prior ensemble for loss modulus in horizontal cross-section.

In Figure 6.3.2 we can see the physical prior ensemble $\mu_0^{(j)}$ in the cross-sections indicated in the plots in Figure 6.3.1. In Figure 6.3.3 we show the 95% sample confidence intervals $\mathrm{CI}_{95,0}$ calculated from the physical prior ensemble

$\mu_0^{(j)}$. We will compare these confidence intervals with the confidence intervals in other experiments later.

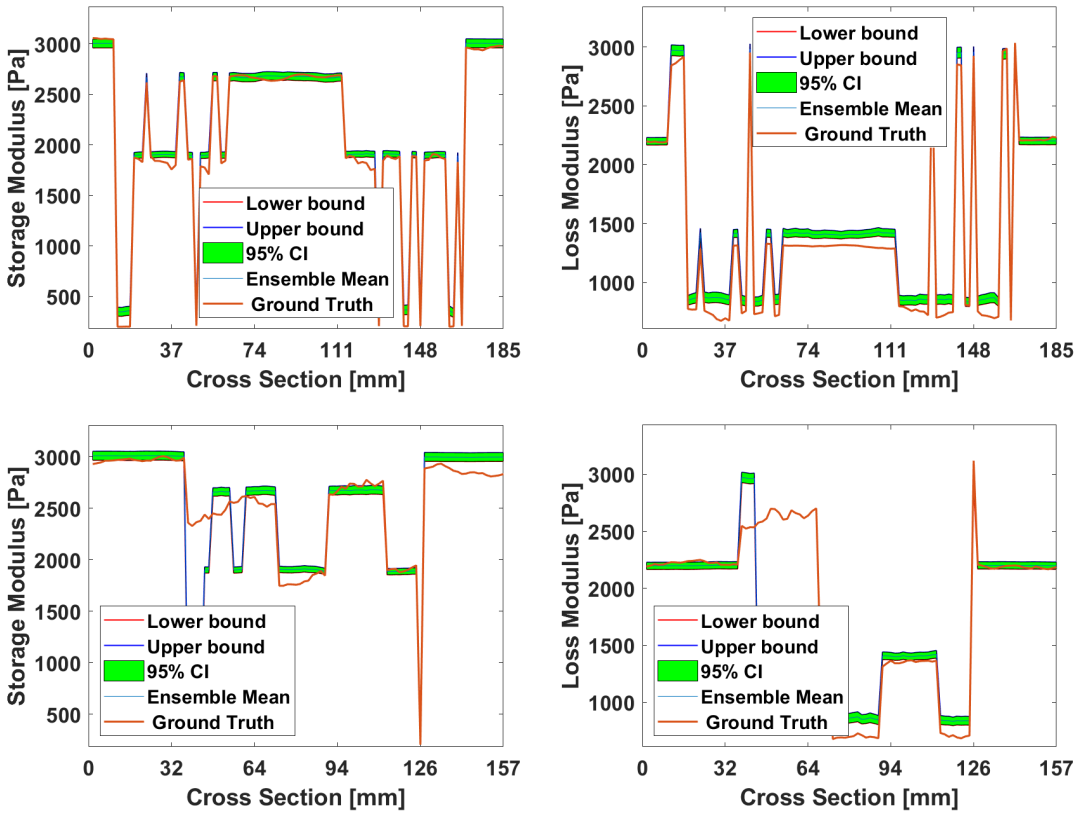


Figure 6.3.3: 95% Confidence Intervals $\mathrm{CI}_{95,0}$ in cross-sections indicated in Fig. 6.3.1 for the prior ensemble $\mu_0^{(j)}$. We plot the lower and the upper bound of the 95% confidence interval as well as the ensemble mean $\overline{\mu_0}$ and the ground truth $\mu^\dagger$. **Top left:** Storage modulus in vertical cross-section. **Top right:** Loss modulus in vertical cross-section. **Bottom left:** Storage modulus in horizontal cross-section. **Bottom right:** Loss modulus in horizontal cross-section.

In Figure 6.3.4 we show plots of the physical posterior ensemble mean $\overline{\mu_{N_S}}$ for the storage and loss modulus. It can be seen that the posterior ensemble mean shows cancer tissue in the locations of the cancer tissue in the ground truth. Also, the probability of cancer P_C shown in the plot in Figure 6.3.5 is close to 1 and the variance $\mathfrak{s}^2_{N_S}$ shown in plots also in Figure 6.3.5 is relatively low. Only around the interface between cancer tissue and no cancer tissue we see a high sample variance $\mathfrak{s}^2_{N_S}$.

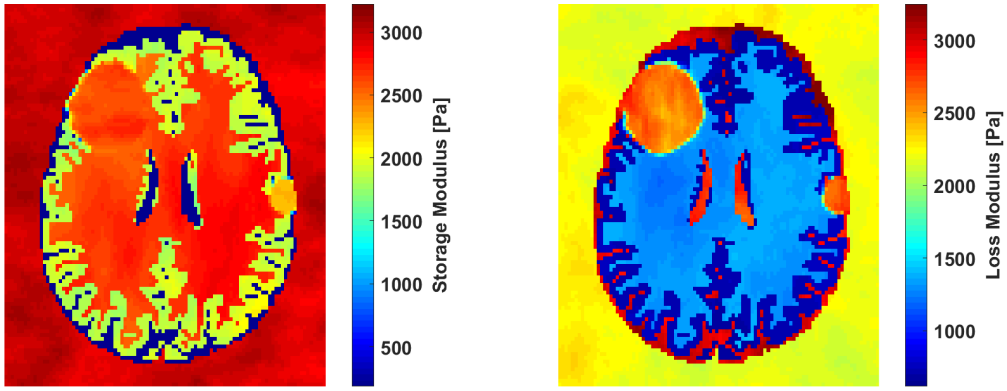


Figure 6.3.4: Physical Posterior Ensemble Mean $\overline{\mu_{N_S}}$. **Left:** Storage modulus. **Right:** Loss Modulus

This high sample variance in the posterior ensemble around the interface between cancer and healthy tissue can also be seen in plots in Figure 6.3.6 in form of a higher variability of loss and storage modulus values in posterior ensemble around the interface in the two cross-sections. In Figure 6.3.7 we give plots of the 95% sample confidence interval CI_{95,N_S} of the posterior ensemble

in the two cross-sections which can barely be seen because it is so small. Only around the interface of tumour tissue and healthy tissue the confidence interval is a bit larger. It can clearly be seen that the 95% sample confidence intervals of the posterior ensemble $\mu_{N_S}^{(j)}$. is smaller than the one of the prior $\mu_0^{(j)}$.

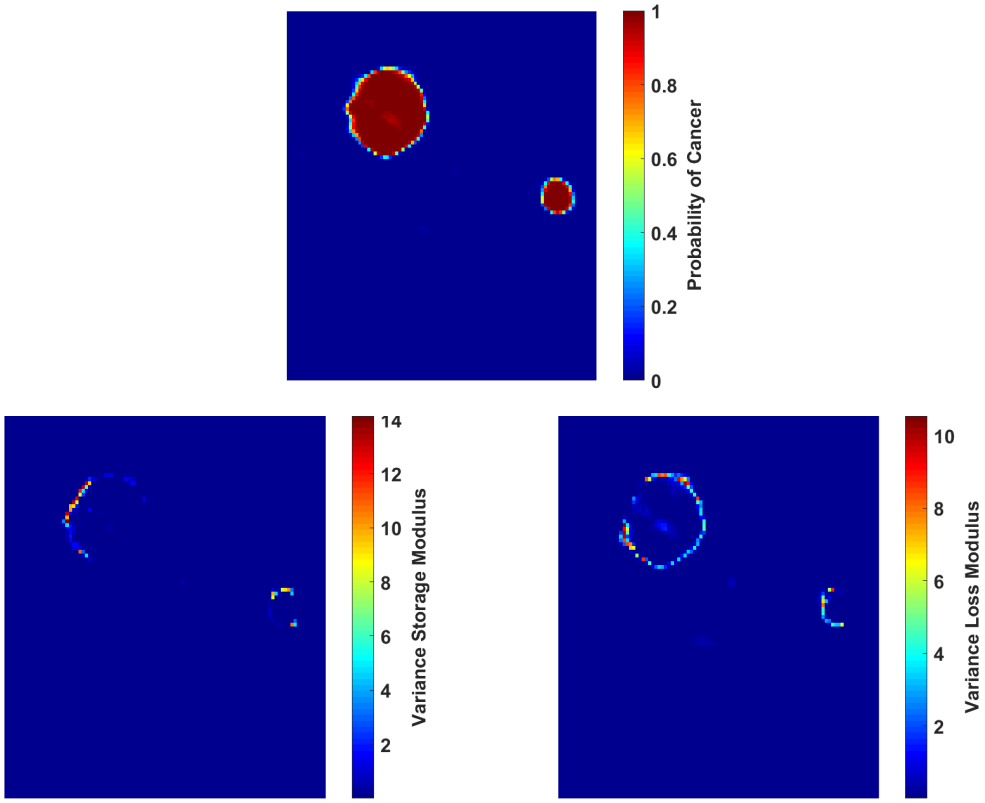


Figure 6.3.5: **Top:** Probability of Cancer P_C. **Bottom Left:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu_{N_S}^{(j)}$; storage modulus. **Bottom Right:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu_{N_S}^{(j)}$; loss modulus.

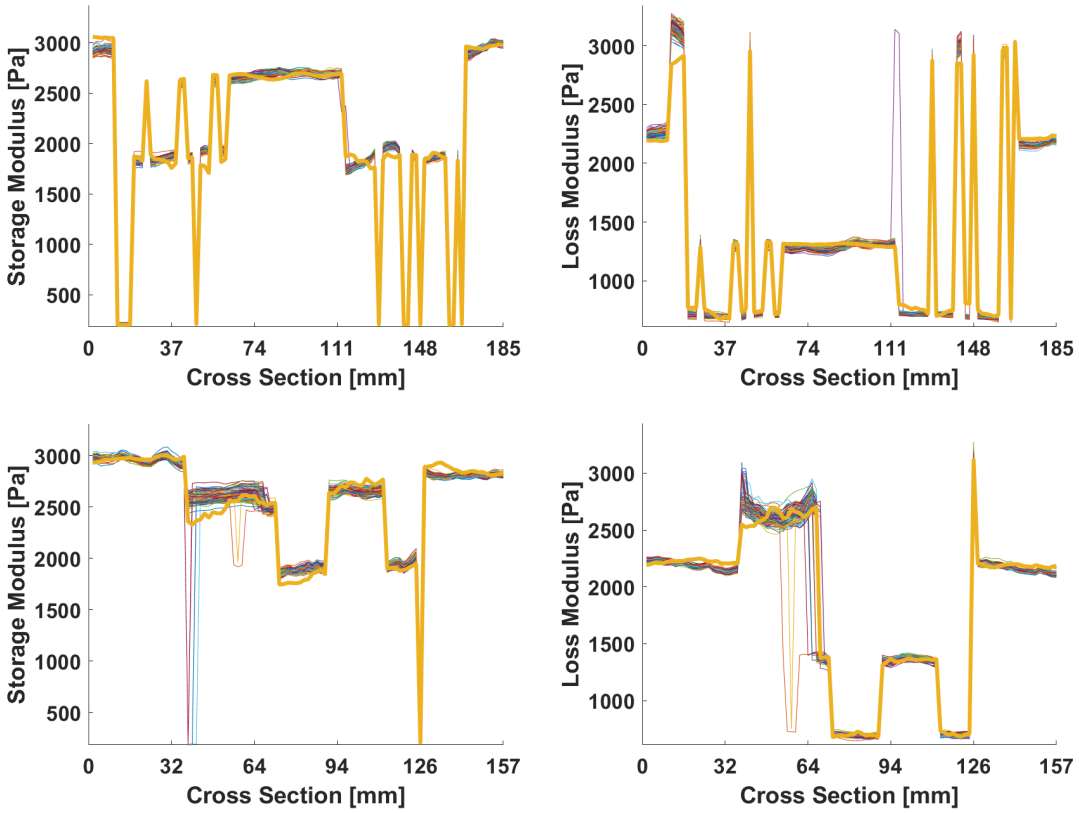


Figure 6.3.6: Posterior ensemble $\mu_{N_S}^{(j)}$ in cross-sections indicated in Fig. 6.3.1. The yellow line indicates the ground truth $\mu^\dagger$. **Top left:** Posterior ensemble for storage modulus in vertical cross-section. **Top right:** Posterior ensemble for loss modulus in vertical cross-section. **Bottom left:** Posterior ensemble for storage modulus in horizontal cross-section. **Bottom right:** Posterior ensemble for loss modulus in horizontal cross-section.

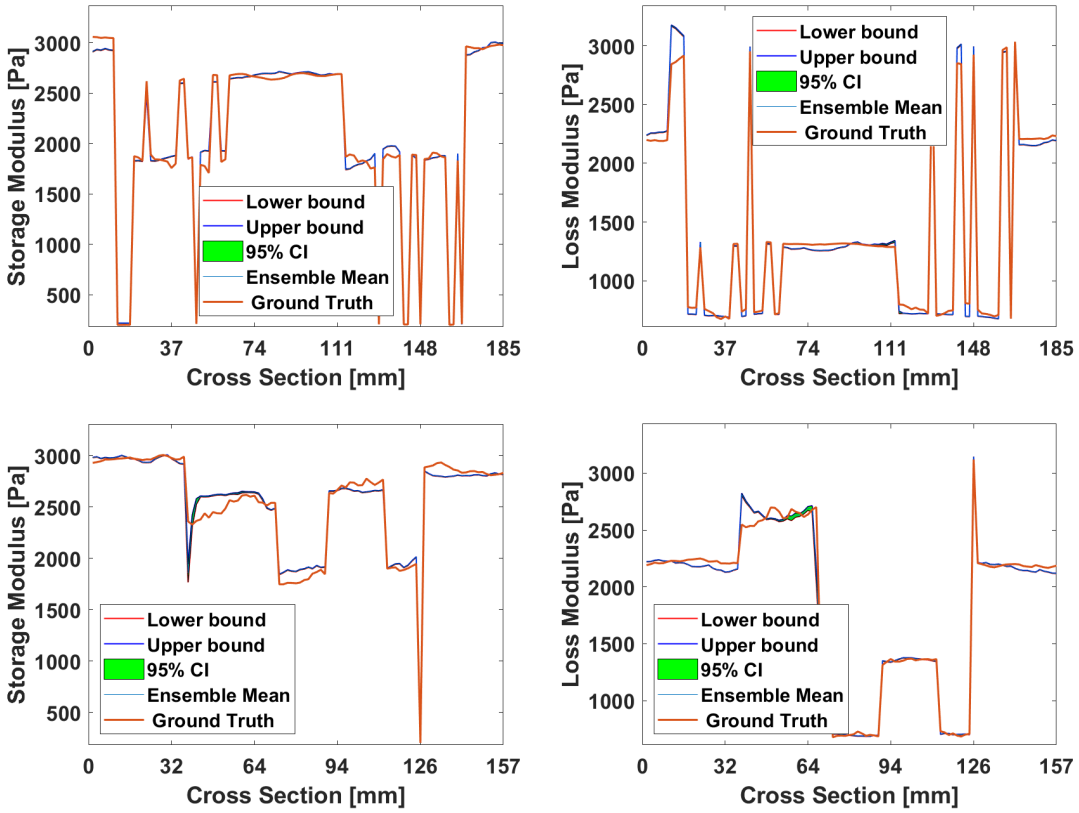


Figure 6.3.7: 95% Confidence Intervals CI_{95,N_S} in cross-sections indicated in Fig. 6.3.1 for the posterior ensemble $\mu_{N_S}^{(j)}$. We plot the lower and the upper bound of the 95% confidence interval as well as the ensemble mean $\overline{\mu_{N_S}}$ and the ground truth $\mu^\dagger$. **Top left:** Storage modulus in vertical cross-section. **Top right:** Loss modulus in vertical cross-section. **Bottom left:** Storage modulus in horizontal cross-section. **Bottom right:** Loss modulus in horizontal cross-section.

6.3.2 *Good Prior, Large Ensemble and High Noise Level*

In this section we run Algorithm 1 in the setup described in Section 6.1.1 with the prior described in Section 6.3.1 and ensemble size $J = 600$. However, this time we use a noise level of 40% which slightly exceeds the noise level of most real-world applications [14].

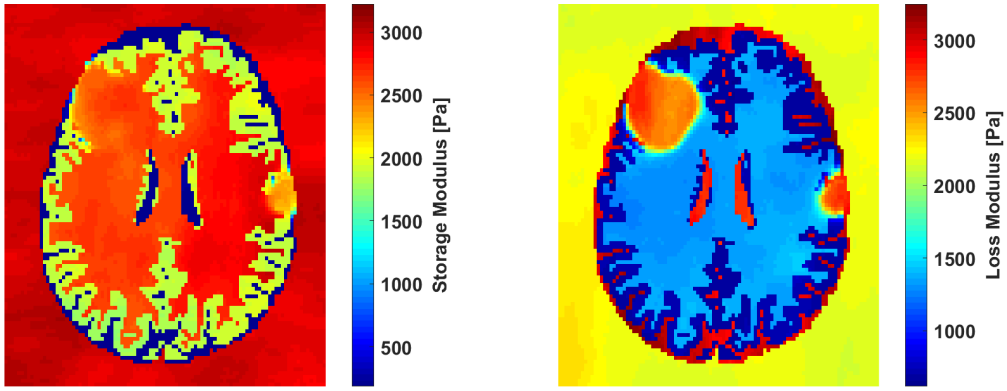


Figure 6.3.8: Physical Posterior Ensemble Mean $\overline{\mu_{N_S}}$. **Left:** Storage modulus. **Right:** Loss Modulus

In Figure 6.3.8 we show plots of the physical posterior ensemble mean $\overline{\mu_{N_S}}$ for the storage and loss modulus. It can be seen that despite the higher noise level the posterior ensemble mean shows cancer tissue in the locations of the cancer tissue in the ground truth. Also, the probability of cancer P_C shown in the plot in Figure 6.3.9 is close to 1 in the location of cancer in the ground truth and the variance $s^2_{N_S}$ also shown in Figure 6.3.9 is higher than the sample

variance of the posterior ensemble in Section 6.3.1 where the data was only corrupted by noise with noise level of 5%. Around the interface between cancer tissue and no cancer tissue we see a high sample variance $\mathfrak{s}^2_{N_S}$.

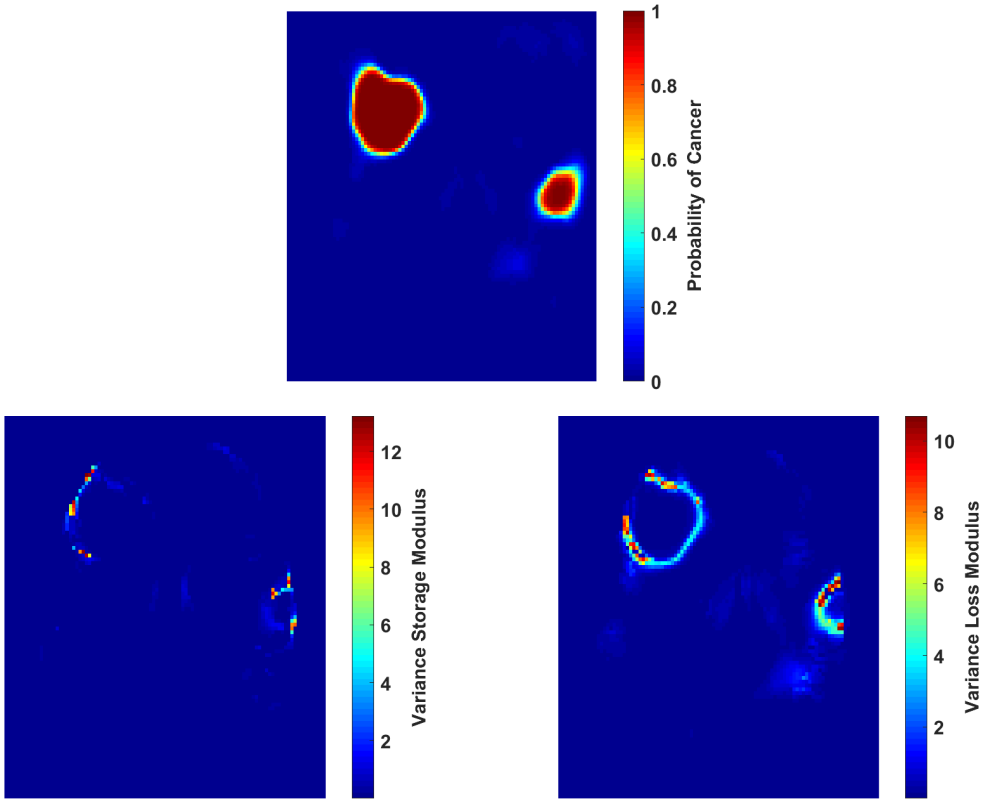


Figure 6.3.9: **Top:** Probability of Cancer P_C. **Bottom Left:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu^{(j)}_{N_S}$; storage modulus. **Bottom Right:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu^{(j)}_{N_S}$; loss modulus.

This high sample variance in the posterior ensemble around the interface between cancer and healthy tissue can also be seen in plots in Figure 6.3.10 in the form of a higher variability of loss and storage modulus values in posterior

ensemble around the interface in the two cross-sections. In Figure 6.3.11 we give plots of the 95% sample confidence interval CI_{95,N_S} of the posterior ensemble in the two cross-sections. The confidence intervals of the posterior ensemble are larger than the confidence intervals of the posterior ensemble in Section 6.3.1 where the data was corrupted by noise with a noise level of 5%.

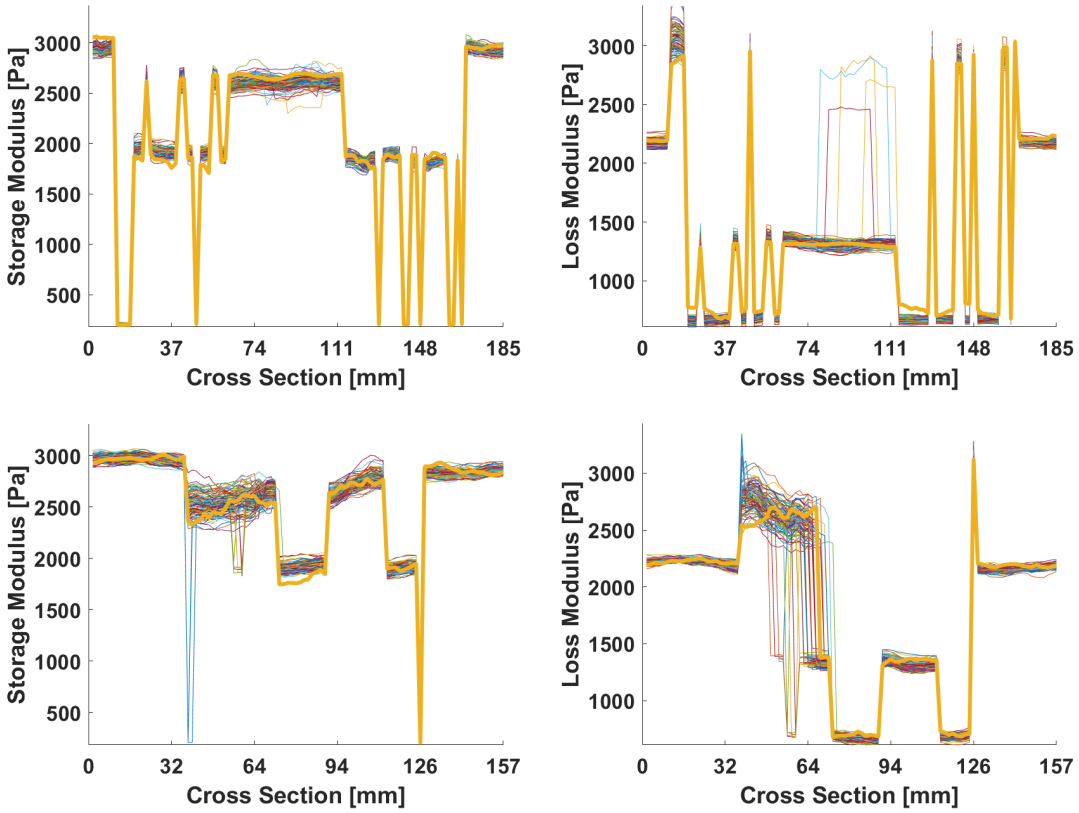


Figure 6.3.10: Posterior ensemble $\mu_{N_S}^{(j)}$ in cross-sections indicated in Fig. 6.3.1. The yellow line indicates the ground truth $\mu^\dagger$. **Top left:** Posterior ensemble for storage modulus in vertical cross-section. **Top right:** Posterior ensemble for loss modulus in vertical cross-section. **Bottom left:** Posterior ensemble for storage modulus in horizontal cross-section. **Bottom right:** Posterior ensemble for loss modulus in horizontal cross-section.

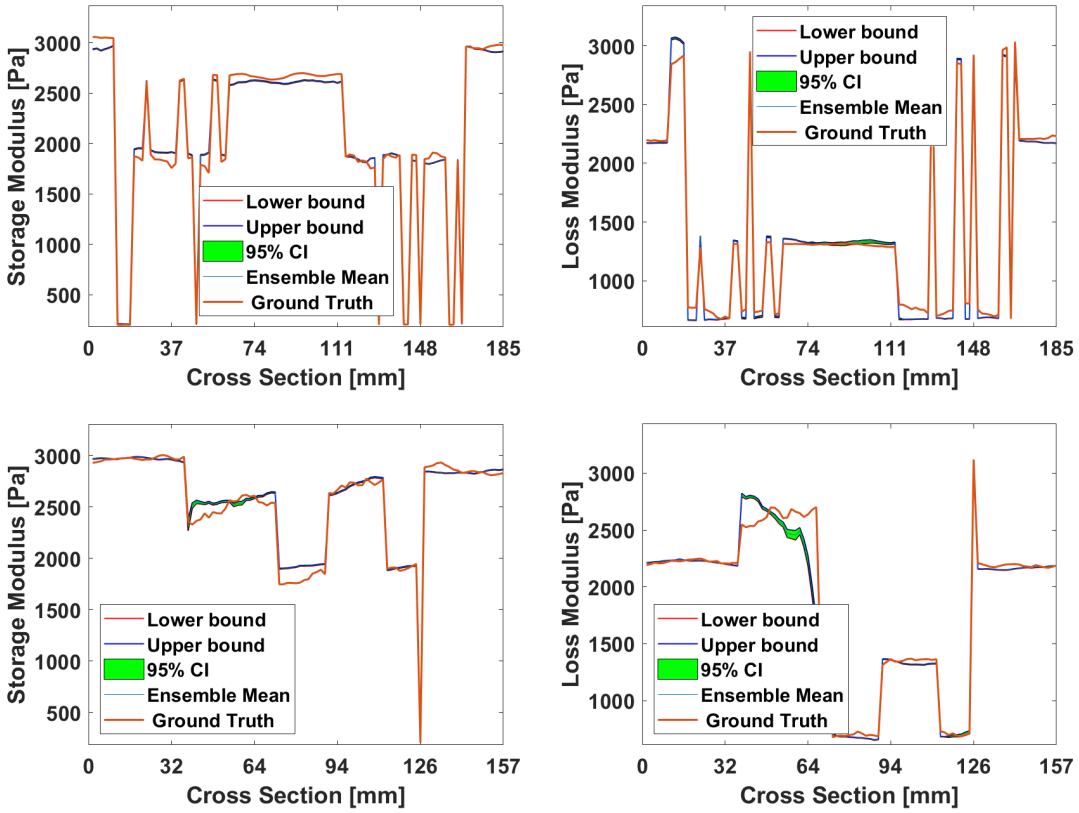


Figure 6.3.11: 95% Confidence Intervals CI_{95,N_S} in cross-sections indicated in Fig. 6.3.1 for the posterior ensemble $\mu_{N_S}^{(j)}$. We plot the lower and the upper bound of the 95% confidence interval as well as the ensemble mean $\overline{\mu_{N_S}}$ and the ground truth $\mu^\dagger$. **Top left:** Storage modulus in vertical cross-section. **Top right:** Loss modulus in vertical cross-section. **Bottom left:** Storage modulus in horizontal cross-section. **Bottom right:** Loss modulus in horizontal cross-section.

6.3.3 *Good Prior, Small Ensemble and Low Noise Level*

In this section we run Algorithm 1 in the setup described in Section 6.1.1 with prior described in Section 6.3.1 with an ensemble sizes $J = 50$. A noise level of 5% is used in this series.

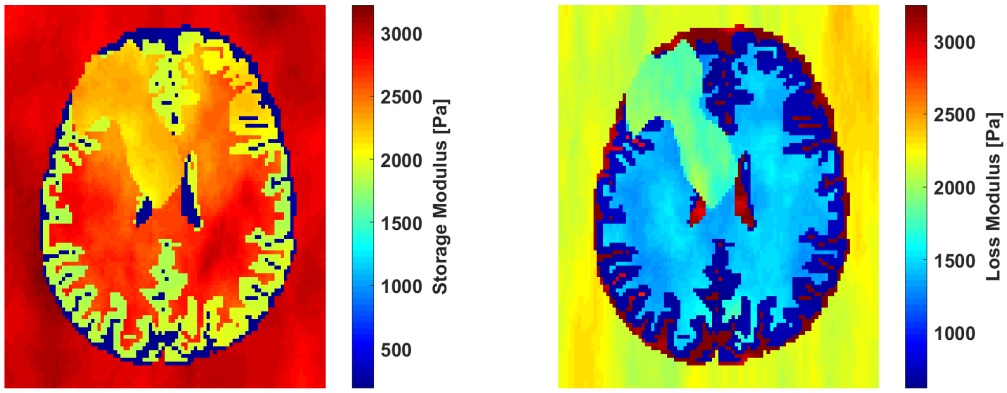


Figure 6.3.12: Physical Posterior Ensemble Mean $\overline{\mu_{N_S}}$. **Left:** Storage modulus. **Right:** Loss Modulus

In Figure 6.3.12 we show plots of the physical posterior ensemble mean $\overline{\mu_{N_S}}$ for the storage and loss modulus. It can be seen that the posterior ensemble mean does not show cancer tissue at the smaller cancer location. The posterior ensemble mean shows cancer tissue at the bigger cancer location but the shape, size and location does not resemble the ground truth. Also, the probability of cancer P_C shown in the plot in Figure 6.3.13 is 1 in areas where we do not have cancer tissue in the ground truth. The variance $\mathfrak{s}^2_{N_S}$ shown in Figure 6.3.13 is

very low compared to the sample variance of the posterior ensemble in Section 6.3.1 where the ensemble size was $J = 600$. Even around the interface between cancer tissue and no cancer tissue we see a low sample variance $\mathfrak{s}^2_{N_S}$.

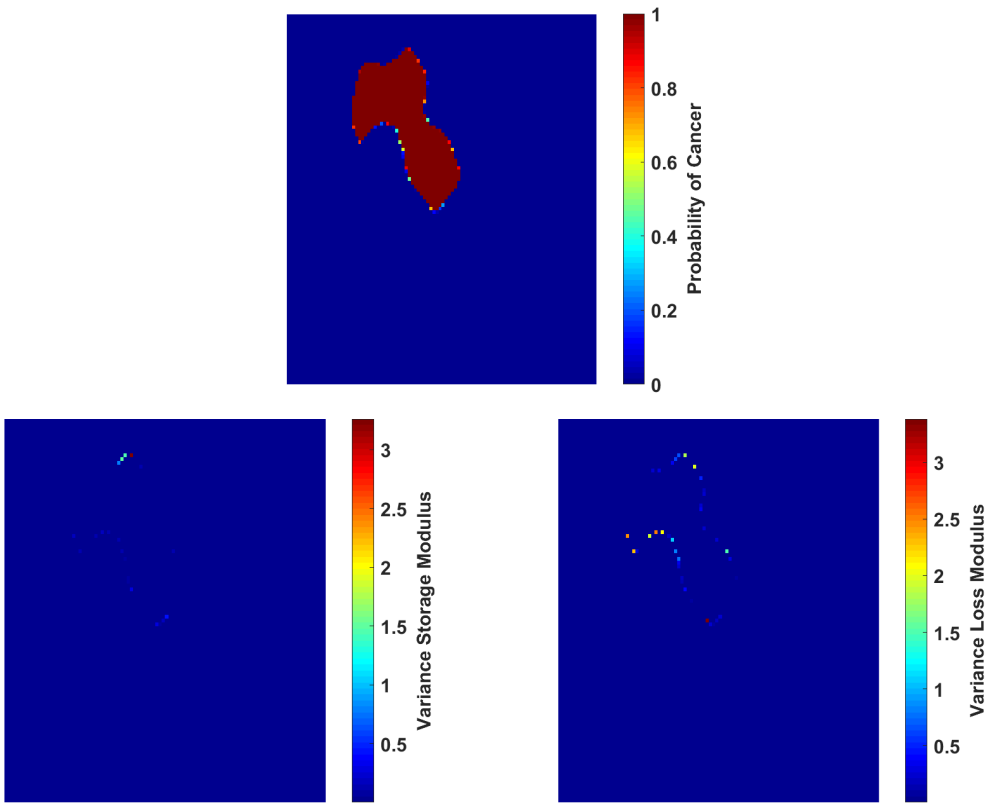


Figure 6.3.13: **Top:** Probability of Cancer P_C. **Bottom Left:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu^{(j)}_{N_S}$; storage modulus. **Bottom Right:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu^{(j)}_{N_S}$; loss modulus.

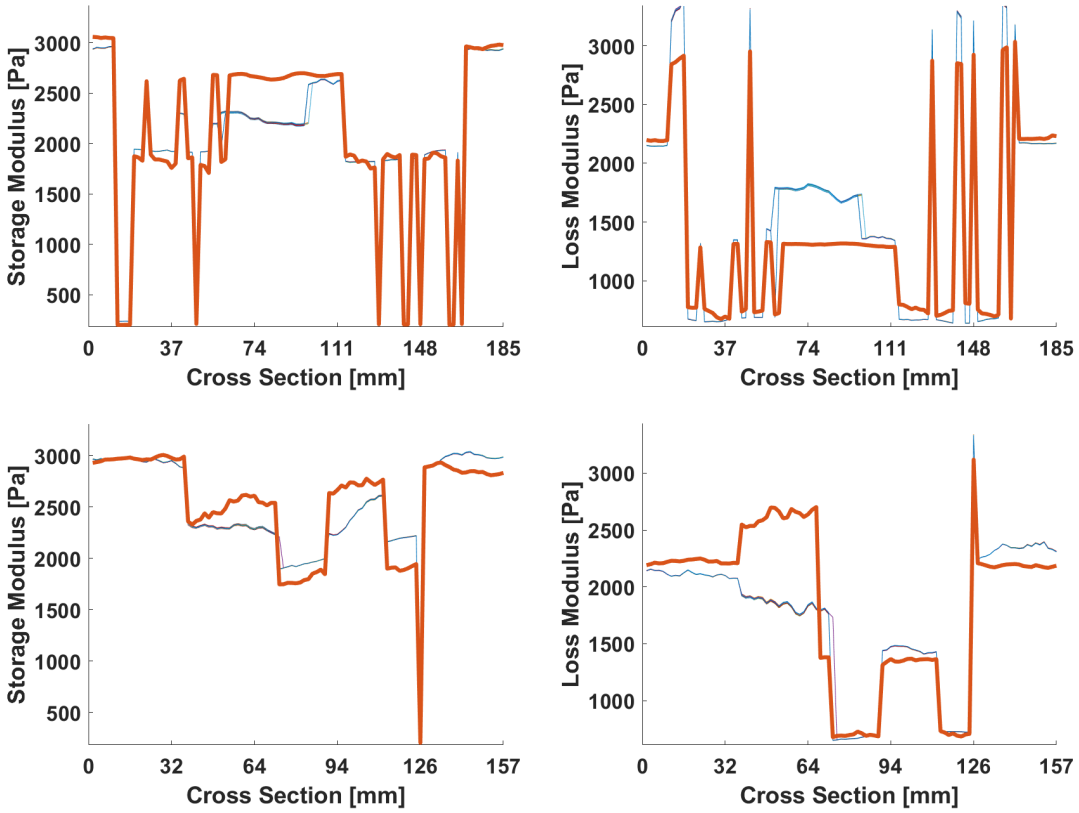


Figure 6.3.14: Posterior ensemble $\mu_{N_S}^{(j)}$ in cross-sections indicated in Fig. 6.3.1. The yellow line indicates the ground truth $\mu^{\dagger}$. **Top left:** Posterior ensemble for storage modulus in vertical cross-section. **Top right:** Posterior ensemble for loss modulus in vertical cross-section. **Bottom left:** Posterior ensemble for storage modulus in horizontal cross-section. **Bottom right:** Posterior ensemble for loss modulus in horizontal cross-section.

The low sample variance in the posterior ensemble around the interface between cancer and healthy tissue can also be seen in Figure 6.3.14 in the form of a no variability of loss and storage modulus values in posterior ensemble

around the interface in the two cross-sections. In Figure 6.3.15 we give plots of the 95% sample confidence interval CI_{95,N_S} of the posterior ensemble in the two cross-sections. The confidence intervals of the posterior ensemble are smaller than the confidence intervals of the posterior ensemble in Section 6.3.1 where the ensemble size was $J = 600$. In all the cross-sectional plots it can be seen that the posterior ensemble has values far off the ground truth. We will discuss and interpret these results in Section 6.3.8.

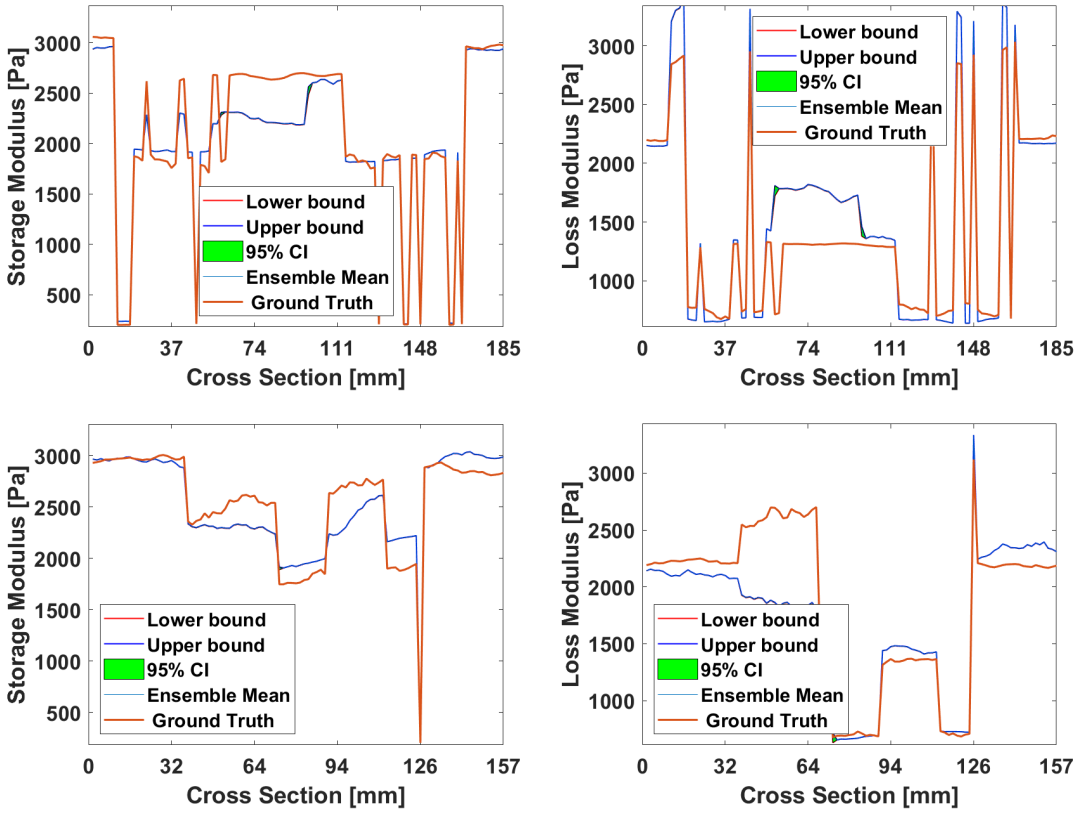


Figure 6.3.15: 95% Confidence Intervals CI_{95,N_S} in cross-sections indicated in Fig. 6.3.1 for the posterior ensemble $\mu_{N_S}^{(j)}$. We plot the lower and the upper bound of the 95% confidence interval as well as the ensemble mean $\overline{\mu_{N_S}}$ and the ground truth $\mu^\dagger$. **Top left:** Storage modulus in vertical cross-section. **Top right:** Loss modulus in vertical cross-section. **Bottom left:** Storage modulus in horizontal cross-section. **Bottom right:** Loss modulus in horizontal cross-section.

6.3.4 *Good Prior, Small Ensemble and High Noise Level*

In this section we run Algorithm 1 in the setup described in Section 6.1.1 with prior described in Section 6.3.1 with an ensemble sizes $J = 50$. However, this time we use a noise level of 40%.

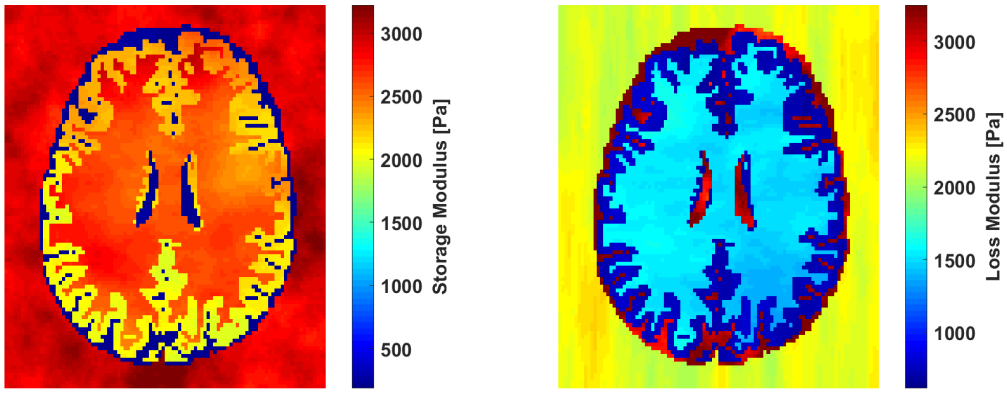


Figure 6.3.16: Physical Posterior Ensemble Mean $\overline{\mu_{N_S}}$. **Left:** Storage modulus. **Right:** Loss Modulus

In Figure 6.3.16 we show plots of the physical posterior ensemble mean $\overline{\mu_{N_S}}$ for the storage and loss modulus. In the posterior ensemble we cannot see any cancer tissue. Also, the probability of cancer P_C shown in Figure 6.3.17 is close to 0 everywhere. The variance $\mathfrak{s}^2_{N_S}$ shown in Figure 6.3.17 is very low compared to the sample variance of the posterior ensemble in Section 6.3.1 where the ensemble size was $J = 600$ and Section 6.3.3 where we used the same ensemble size $J = 50$ as in this section but used a smaller noise level of 5% in the data.

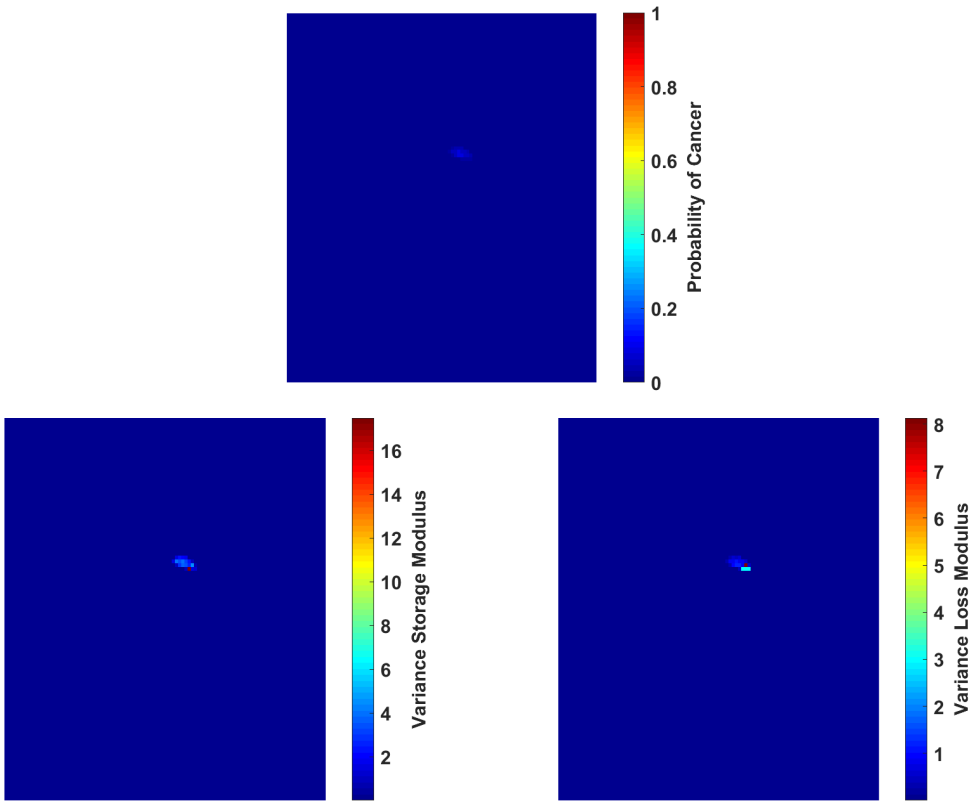


Figure 6.3.17: **Top:** Probability of Cancer P_C. **Bottom Left:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu_{N_S}^{(j)}$; storage modulus. **Bottom Right:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu_{N_S}^{(j)}$; loss modulus.

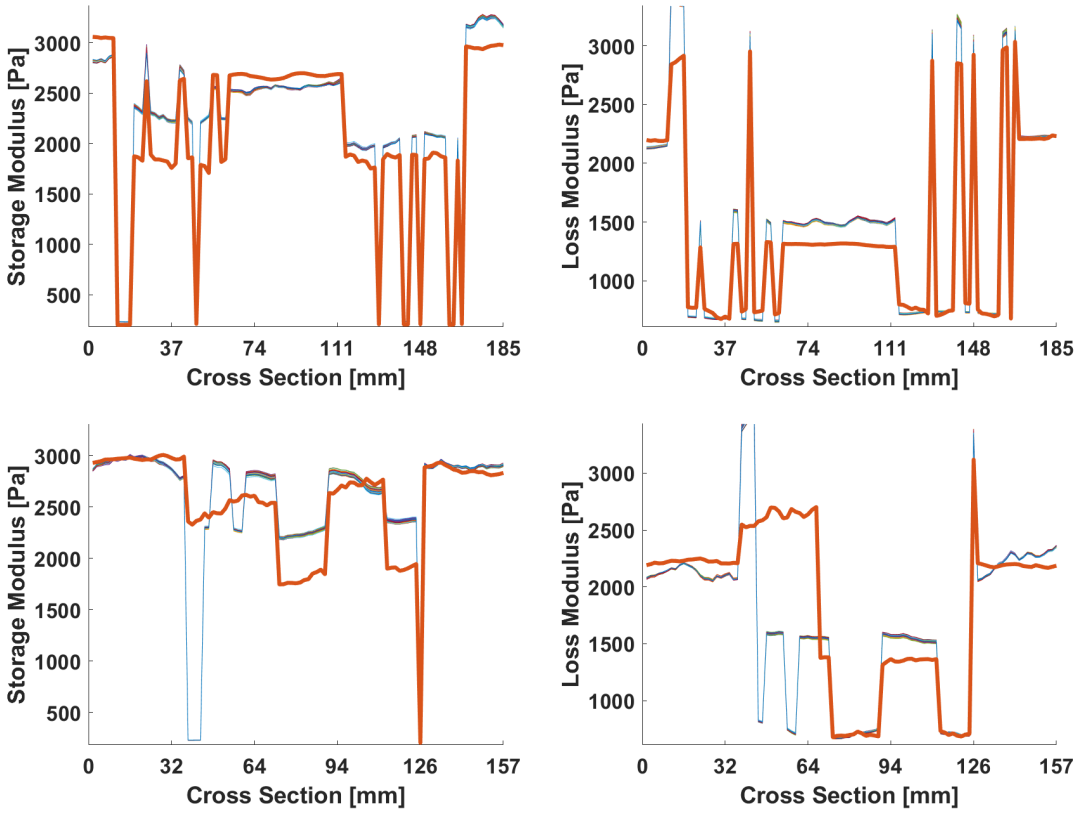


Figure 6.3.18: Posterior ensemble $\mu_{N_S}^{(j)}$ in cross-sections indicated in Fig. 6.3.1. The yellow line indicates the ground truth $\mu^\dagger$. **Top left:** Posterior ensemble for storage modulus in vertical cross-section. **Top right:** Posterior ensemble for loss modulus in vertical cross-section. **Bottom left:** Posterior ensemble for storage modulus in horizontal cross-section. **Bottom right:** Posterior ensemble for loss modulus in horizontal cross-section.

The low sample variance in the posterior ensemble around the interface between cancer and healthy tissue can also be seen in plots in Figure 6.3.18 in form of a no variability of loss and storage modulus values in posterior

ensemble around the interface in the two cross-sections. In Figure 6.3.19 we give plots of the 95% sample confidence interval CI_{95,N_S} of the posterior ensemble in the two cross-sections. The confidence intervals of the posterior ensemble are smaller than the confidence intervals of the posterior ensemble in Section 6.3.1 where the ensemble size was $J = 600$. In all the cross-sectional plots it can be seen that the posterior ensemble has values far off the ground truth.

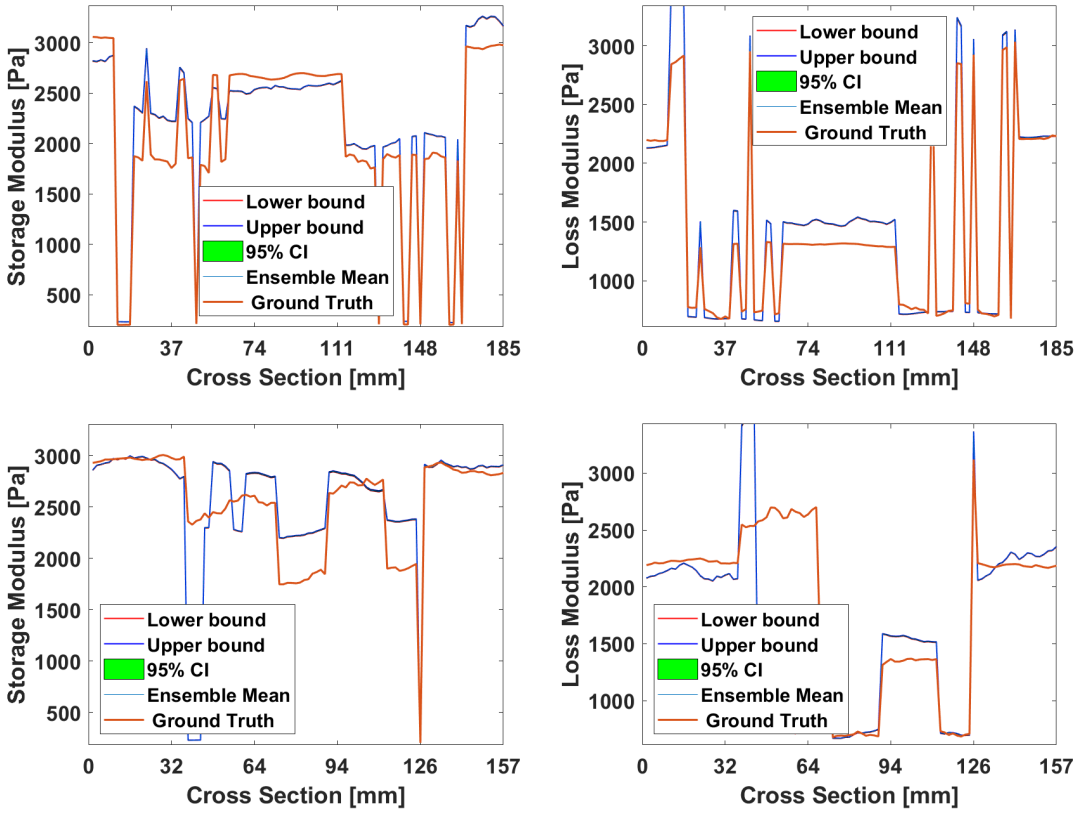


Figure 6.3.19: 95% Confidence Intervals CI_{95,N_S} in cross-sections indicated in Fig. 6.3.1 for the posterior ensemble $\mu_{N_S}^{(j)}$. We plot the lower and the upper bound of the 95% confidence interval as well as the ensemble mean $\overline{\mu_{N_S}}$ and the ground truth $\mu^\dagger$. **Top left:** Storage modulus in vertical cross-section. **Top right:** Loss modulus in vertical cross-section. **Bottom left:** Storage modulus in horizontal cross-section. **Bottom right:** Loss modulus in horizontal cross-section.

6.3.5 *Bad Prior, Large Ensemble Size and Low Noise Level*

Table 6.3.2: Parameters of the prior for Brain MRE experiments

	$\underline{m_{\alpha,L}}$	$\overline{m_{\alpha,L}}$	$\underline{m_{\alpha,S}}$	$\overline{m_{\alpha,S}}$
Grey Matter	218	654	560	1677
White Matter	400	1200	806	2419
Background Matter	897	2692	900	2700
CSF Matter	660	1960	63	190
Tumour Tissue	765	2297	708	2124

In this experiment we select an initial ensemble by drawing for each tissue type $\alpha = 1, \ldots, 4$ the parameters $k^{(j)}_{0,\alpha,S} := (m_{\alpha,S}, l_{\alpha,S,1}, l_{\alpha,S,2}, w_{\alpha,S})$ and $k^{(j)}_{0,\alpha,L} := (m_{\alpha,L}, l_{\alpha,L,1}, l_{\alpha,L,2}, w_{\alpha,L})$ for loss and storage modulus from the same uniform and normal distributions as defined in (6.3.1). However, for the uniform distribution for mean $m_{\alpha,S}$ and $m_{\alpha,L}$, this time we use the values from Table 6.3.2 depending on the tissue type and loss or storage modulus parametrisation. This choice allows us to track the performance of Ensemble Kalman Inversion applied to MRE for "bad" priors.

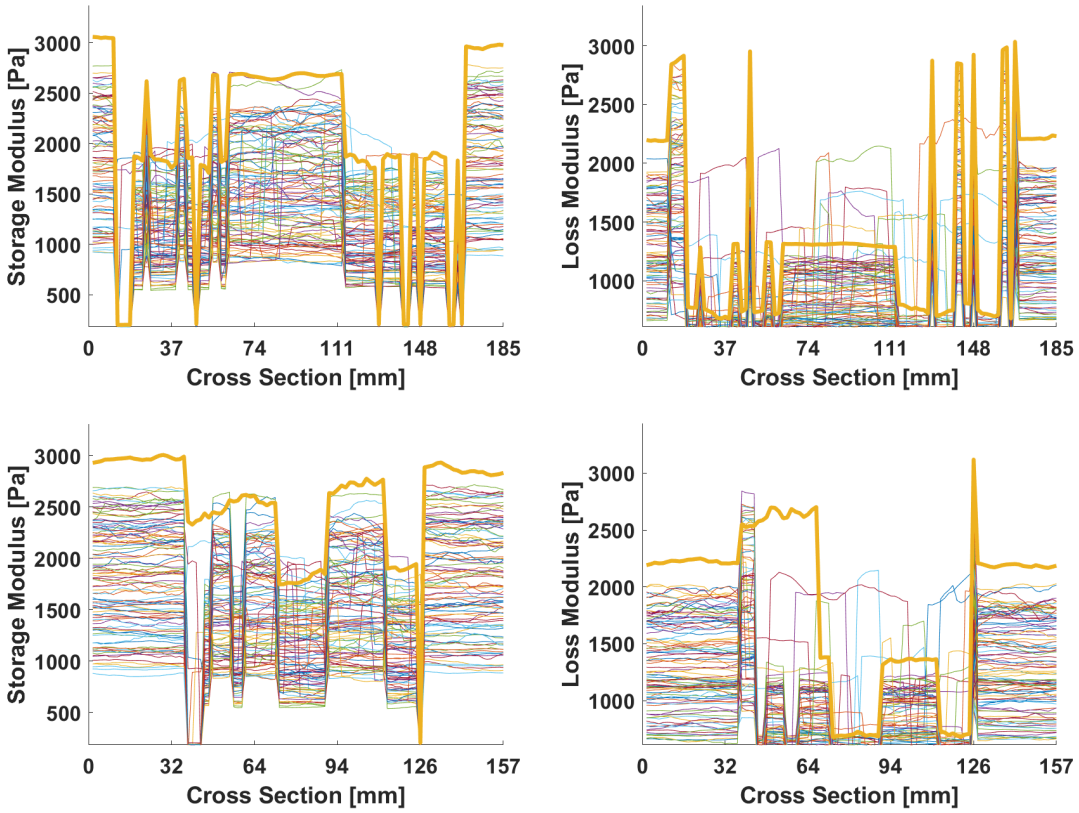


Figure 6.3.20: Prior ensemble $\mu_0^{(j)}$ in cross-sections indicated in Fig. 6.3.1. The yellow line indicates the ground truth $\mu^\dagger$. **Top left:** Prior ensemble for storage modulus in vertical cross-section. **Top right:** Prior ensemble for loss modulus in vertical cross-section. **Bottom left:** Prior ensemble for storage modulus in horizontal cross-section. **Bottom right:** Prior ensemble for loss modulus in horizontal cross-section.

Again, a part of the initial ensemble is $k_{0,C}^{(j)} \sim \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1)$ for the parametrisation of the location of the cancer. The mean for this random field is set to zero. We use $c' = 1.5$ in this parametrisation throughout the inversion. Let us denote the prior ensemble as $k_{0,F}^{(j)} := (\{k_{0,\alpha,S}^{(j)}\}_{\alpha\in\{1,\dots,5\}}, \{k_{0,\alpha,L}^{(j)}\}_{\alpha\in\{1,\dots,5\}}, k_{0,C}^{(j)})$.

We run Algorithm 1 in the setup described in Section 6.1.1 with prior described in this section above with an ensemble sizes $J = 600$. A noise level of 5% is used in this series.

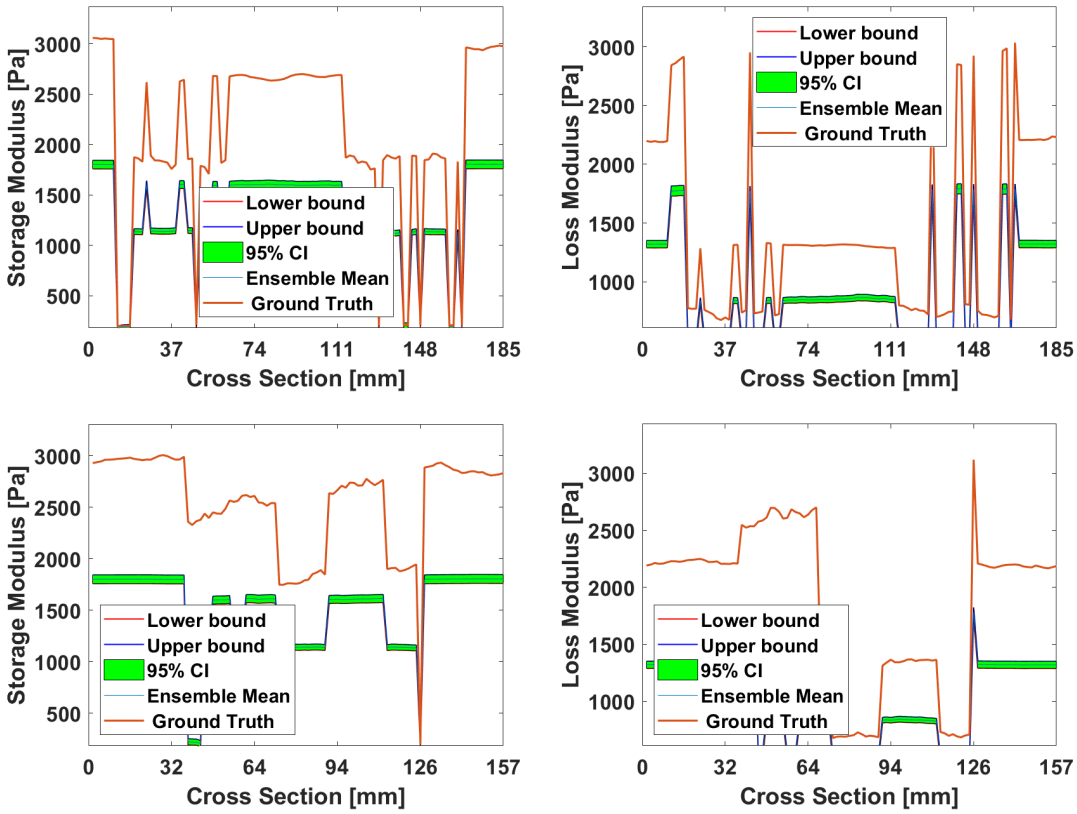


Figure 6.3.21: 95% Confidence Intervals $\mathrm{CI}_{95,0}$ in cross-sections indicated in Fig. 6.3.1 for the prior ensemble $\mu_0^{(j)}$. We plot the lower and the upper bound of the 95% confidence interval as well as the ensemble mean $\overline{\mu_0}$ and the ground truth $\mu^\dagger$. **Top left:** Storage modulus in vertical cross-section. **Top right:** Loss modulus in vertical cross-section. **Bottom left:** Storage modulus in horizontal cross-section. **Bottom right:** Loss modulus in horizontal cross-section.

In Figure 6.3.20 we can see the physical prior ensemble $\mu_0^{(j)}$ in the cross-sections indicated in the plots in Figure 6.3.1. In Figure 6.3.21 we show the 95% sample confidence intervals $\mathrm{CI}_{95,0}$ calculated from the physical prior ensemble $\mu_0^{(j)}$.

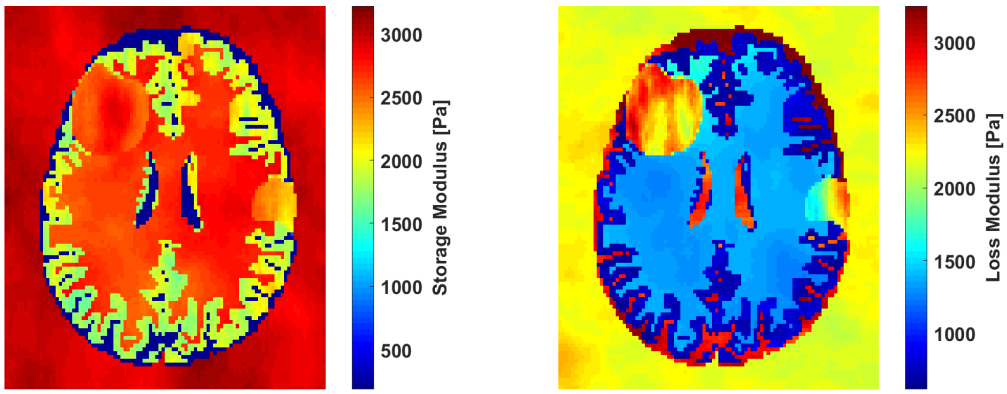


Figure 6.3.22: Physical Posterior Ensemble Mean $\overline{\mu_{N_S}}$. **Left:** Storage modulus. **Right:** Loss Modulus

In Figure 6.3.22 we show plots of the physical posterior ensemble mean $\overline{\mu_{N_S}}$ for the storage and loss modulus. It can be seen that despite the mean of the prior physical ensemble being far off the ground truth, the posterior ensemble mean has cancer tissue in the locations of the cancer tissue in the ground truth. Also, the probability of cancer P_C shown in Figure 6.3.23 is almost 1. However, around the location of the smaller cancer tissue the cancer probability is almost 1 even in some adjacent regions which have no cancer tissue. The variance $\mathfrak{s}^2_{N_S}$ shown in Figure 6.3.23 is high only around the interface between cancer

tissue and healthy tissue. There the sample variance is similar in amplitude to the sample variance in this location of the posterior ensemble in Section 6.3.1, where we used the same noise level and the same ensemble size but the mean of the prior physical ensemble was close to the ground truth.

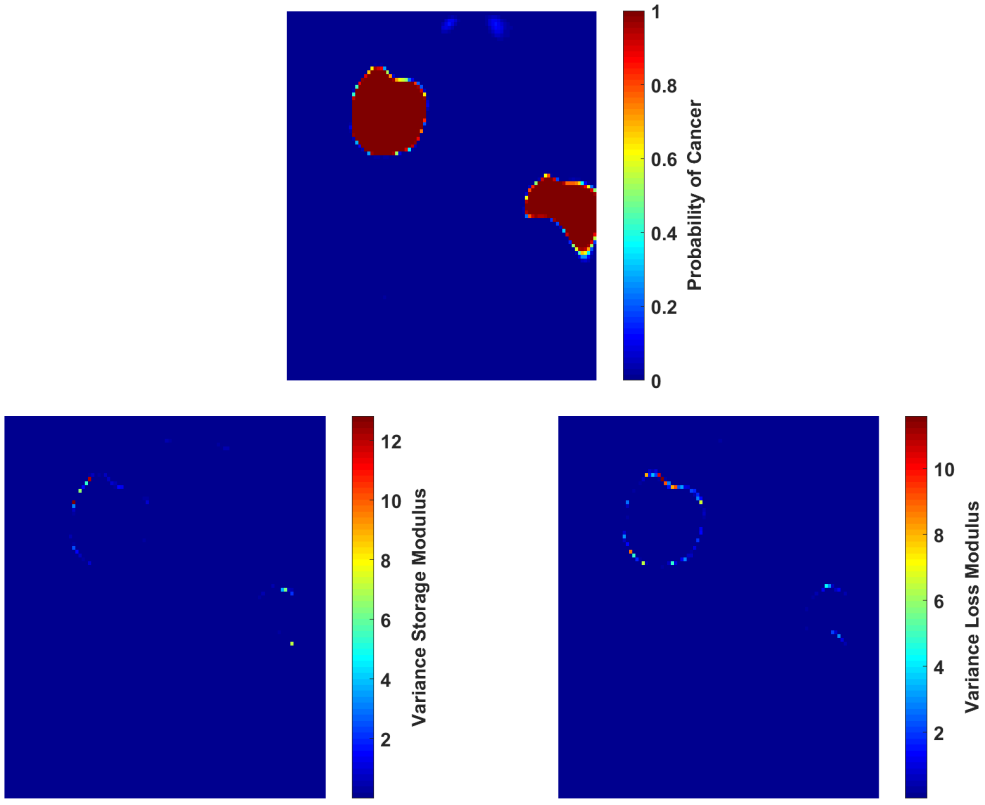


Figure 6.3.23: **Top:** Probability of Cancer P_C. **Bottom Left:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu_{N_S}^{(j)}$; storage modulus. **Bottom Right:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu_{N_S}^{(j)}$; loss modulus.

The low sample variance in the posterior ensemble around the interface between cancer and healthy tissue can also be seen in Figure 6.3.24 in form

of a no variability of loss and storage modulus values in posterior ensemble around the interface in the two cross-sections. In Figure 6.3.25 we give plots of the 95% sample confidence interval CI_{95,N_S} of the posterior ensemble in the two cross-sections. The confidence intervals of the posterior ensemble are about the same size as the confidence intervals of the posterior ensemble in Section 6.3.1 where the mean of the physical prior ensemble was close to the ground truth. In all the cross-sectional plots it can be seen that the posterior ensemble has values close to the ground truth.

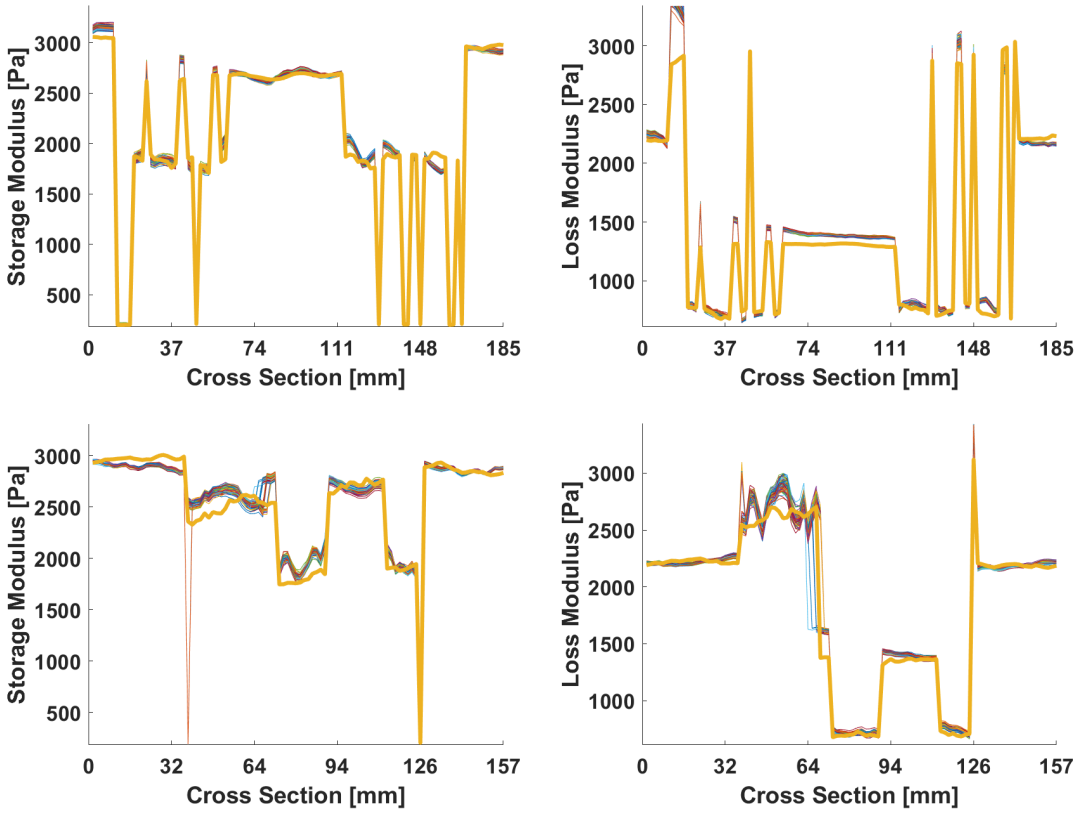


Figure 6.3.24: Posterior ensemble $\mu_{N_S}^{(j)}$ in cross-sections indicated in Fig. 6.3.1. The yellow line indicates the ground truth $\mu^\dagger$. **Top left:** Posterior ensemble for storage modulus in vertical cross-section. **Top right:** Posterior ensemble for loss modulus in vertical cross-section. **Bottom left:** Posterior ensemble for storage modulus in horizontal cross-section. **Bottom right:** Posterior ensemble for loss modulus in horizontal cross-section.

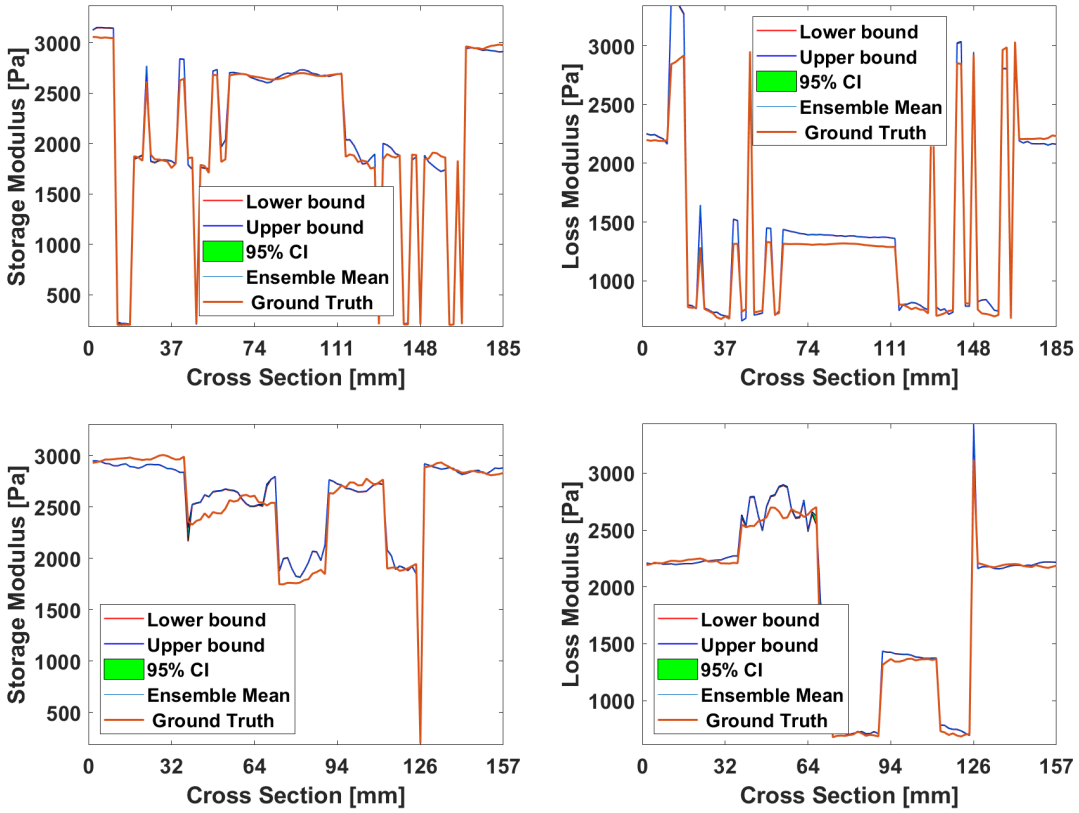


Figure 6.3.25: 95% Confidence Intervals CI_{95,N_S} in cross-sections indicated in Fig. 6.3.1 for the posterior ensemble $\mu_{N_S}^{(j)}$. We plot the lower and the upper bound of the 95% confidence interval as well as the ensemble mean $\overline{\mu_{N_S}}$ and the ground truth $\mu^\dagger$. **Top left:** Storage modulus in vertical cross-section. **Top right:** Loss modulus in vertical cross-section. **Bottom left:** Storage modulus in horizontal cross-section. **Bottom right:** Loss modulus in horizontal cross-section.

6.3.6 *Bad Prior, Large Ensemble and High Noise Level*

In this section we run Algorithm 1 in the setup described in Section 6.1.1 with prior described in Section 6.3.5 with an ensemble sizes $J = 600$. However, this time we use a noise level of 40%.

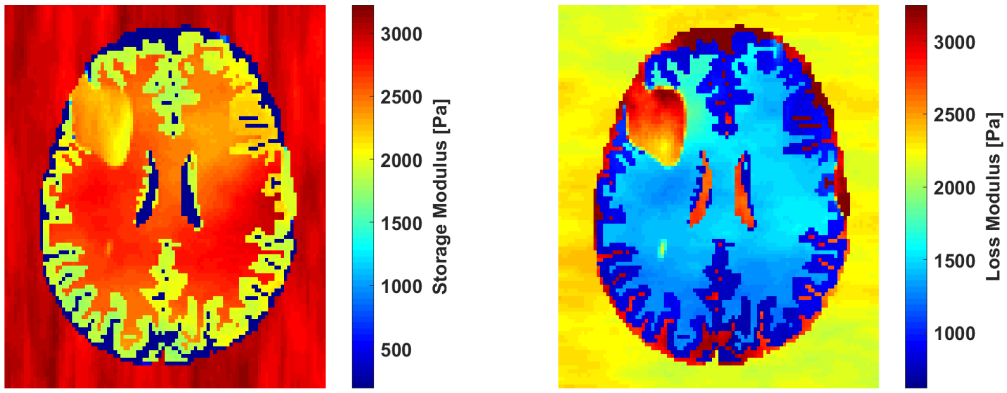


Figure 6.3.26: Physical Posterior Ensemble Mean $\overline{\mu_{N_S}}$. **Left:** Storage modulus. **Right:** Loss Modulus

In Figure 6.3.26 we show plots of the physical posterior ensemble mean $\overline{\mu_{N_S}}$ for the storage and loss modulus. It can be seen that despite the higher noise level in the data the posterior ensemble mean has cancer tissue in the bigger location of cancer tissue. However, the physical posterior ensemble mean does not have cancer tissue in the small cancer location. The probability of cancer P_C shown in the plot in Figure 6.3.27 is almost 1 in the large cancer location but also in many other regions without cancer. The sample variance $\mathfrak{s}^2_{N_S}$ of the

posterior ensemble shown in Figure 6.3.27 is higher than the sample variance of the posterior ensemble in Section 6.3.5, where the data were corrupted by noise with noise level of 5%. Around the interface between the cancer tissue and non cancer tissue we see a high sample variance $\mathfrak{s}^2_{N_S}$. The variance is low around the small cancer location where the posterior ensemble mean does not show cancer tissue.

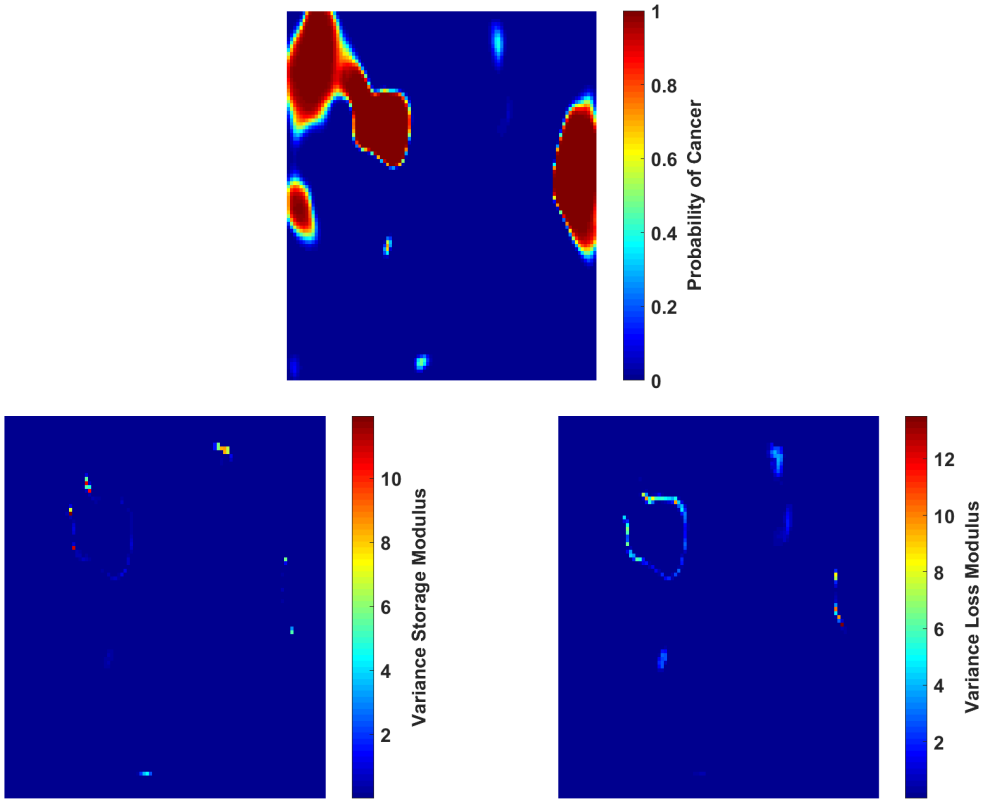


Figure 6.3.27: **Top:** Probability of Cancer P_C. **Bottom Left:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu^{(j)}_{N_S}$; storage modulus. **Bottom Right:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu^{(j)}_{N_S}$; loss modulus.

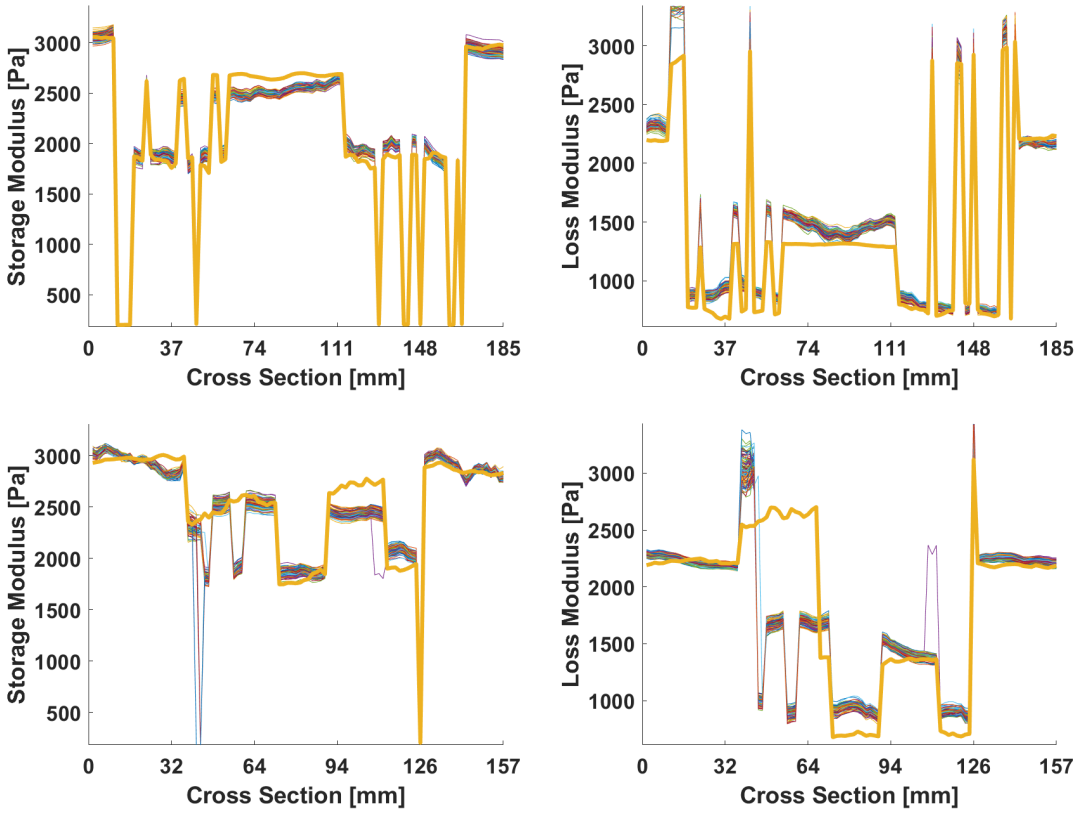


Figure 6.3.28: Posterior ensemble $\mu_{N_S}^{(j)}$ in cross-sections indicated in Fig. 6.3.1. The yellow line indicates the ground truth $\mu^\dagger$. **Top left:** Posterior ensemble for storage modulus in vertical cross-section. **Top right:** Posterior ensemble for loss modulus in vertical cross-section. **Bottom left:** Posterior ensemble for storage modulus in horizontal cross-section. **Bottom right:** Posterior ensemble for loss modulus in horizontal cross-section.

The slightly higher sample variance of the posterior ensemble compared to the sample variance in Section 6.3.5 around the interface between cancer and healthy tissue can also be seen in Figure 6.3.28 in the form of a higher

variability of loss and storage modulus values in posterior ensemble around the interface in the two cross-sections. In Figure 6.3.29 we give plots of the 95% sample confidence interval CI_{95,N_S} of the posterior ensemble in the two cross-sections. The confidence intervals of the posterior ensemble are larger than the confidence intervals of the posterior ensemble in Section 6.3.5, where the data was corrupted by noise with a noise level of 5%.

In general, however, the variance in the posterior ensemble is lower compared to variance in the posterior ensemble in Section 6.3.2 where the increase of the noise level in the data lead to considerable increase in variance compared to variance in the posterior ensemble in Section 6.3.1. In other words, the fact that the physical prior ensemble has a mean that is far off the ground truth entailed a smaller sample variance in the posterior ensemble for high noise levels.

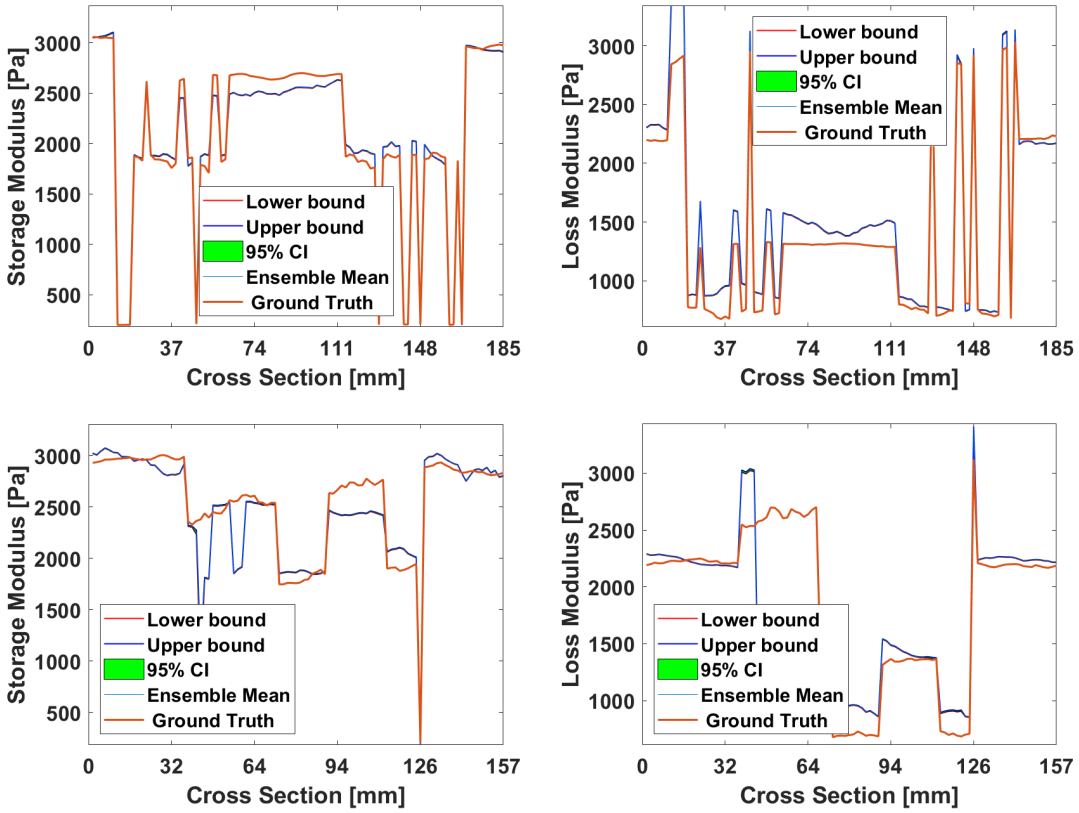


Figure 6.3.29: 95% Confidence Intervals CI_{95,N_S} in cross-sections indicated in Fig. 6.3.1 for the posterior ensemble $\mu_{N_S}^{(j)}$. We plot the lower and the upper bound of the 95% confidence interval as well as the ensemble mean $\overline{\mu_{N_S}}$ and the ground truth $\mu^\dagger$. **Top left:** Storage modulus in vertical cross-section. **Top right:** Loss modulus in vertical cross-section. **Bottom left:** Storage modulus in horizontal cross-section. **Bottom right:** Loss modulus in horizontal cross-section.

6.3.7 *Prior with Large Variance, Large Ensemble Size and High Noise Level*

Table 6.3.3: Parameters of the prior for Brain MRE experiments

	$\underline{m_{\alpha,L}}$	$\overline{m_{\alpha,L}}$	$\underline{m_{\alpha,S}}$	$\overline{m_{\alpha,S}}$
Grey Matter	1	867	1	2236
White Matter	1	1600	1	3255
Background Matter	1	3590	1	3600
CSF Matter	1	2640	1	254
Tumour Tissue	1	3062	1	2832

In this experiment we select an initial ensemble by drawing for each tissue type $\alpha = 1, \ldots, 4$ the parameters $k^{(j)}_{0,\alpha,S} := (m_{\alpha,S}, l_{\alpha,S,1}, l_{\alpha,S,2}, w_{\alpha,S})$ and $k^{(j)}_{0,\alpha,L} := (m_{\alpha,L}, l_{\alpha,L,1}, l_{\alpha,L,2}, w_{\alpha,L})$ for loss and storage modulus from the same uniform and normal distributions as defined in (6.3.1). However, for the uniform distribution for mean $m_{\alpha,S}$ and $m_{\alpha,L}$, this time we use the values from Table 6.3.3 depending on the tissue type and loss or storage modulus parametrisation. This choice allows us to track the performance of EKI applied to MRE for priors with high sample variance.

Again, a part of the initial ensemble is $k_{0,C}^{(j)} \sim \mathcal{U}_{[0.035,0.3]} \otimes \mathcal{N}(0,1)$ for the parametrisation of the location of the cancer. The mean for this random field is set to zero. We use $c' = 1.5$ in this parametrisation throughout the inversion. Let us denote the prior ensemble as $k_{0,F}^{(j)} := (\{k_{0,\alpha,S}^{(j)}\}_{\alpha\in\{1,\dots,5\}}, \{k_{0,\alpha,L}^{(j)}\}_{\alpha\in\{1,\dots,5\}}, k_{0,C}^{(j)})$.

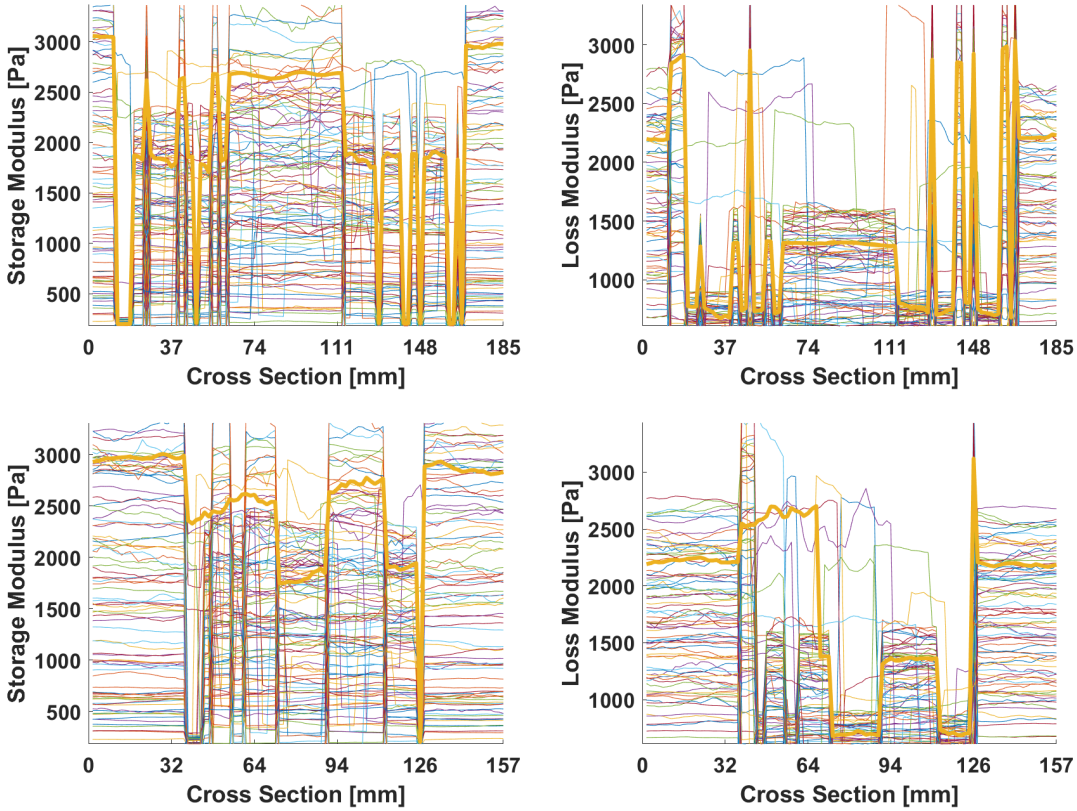


Figure 6.3.30: Prior ensemble $\mu_0^{(j)}$ in cross-sections indicated in Fig. 6.3.1. The yellow line indicates the ground truth $\mu^\dagger$. **Top left:** Prior ensemble for storage modulus in vertical cross-section. **Top right:** Prior ensemble for loss modulus in vertical cross-section. **Bottom left:** Prior ensemble for storage modulus in horizontal cross-section. **Bottom right:** Prior ensemble for loss modulus in horizontal cross-section.

We run Algorithm 1 in the setup described in Section 6.1.1 with prior described in this section above with an ensemble sizes $J = 600$. A noise level of 40% is used in this series.

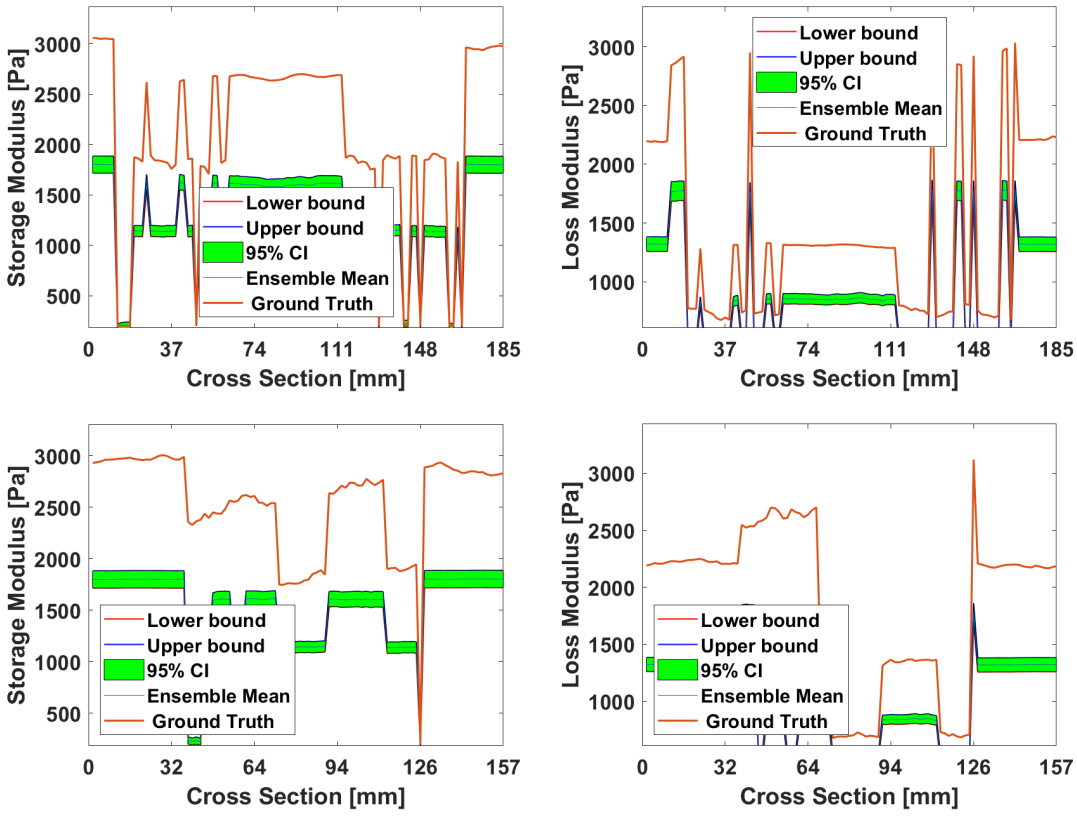


Figure 6.3.31: 95% Confidence Intervals $\mathrm{CI}_{95,0}$ in cross-sections indicated in Fig. 6.3.1 for the prior ensemble $\mu_0^{(j)}$. We plot the lower and the upper bound of the 95% confidence interval as well as the ensemble mean $\overline{\mu_0}$ and the ground truth $\mu^\dagger$. **Top left:** Storage modulus in vertical cross-section. **Top right:** Loss modulus in vertical cross-section. **Bottom left:** Storage modulus in horizontal cross-section. **Bottom right:** Loss modulus in horizontal cross-section.

In Figure 6.3.30 we can see the physical prior ensemble $\mu_0^{(j)}$ in the cross-sections indicated in the plots in Figure 6.3.1. In Figure 6.3.31 we show the 95% sample confidence intervals $\mathrm{CI}_{95,0}$ calculated from the physical prior ensemble $\mu_0^{(j)}$.

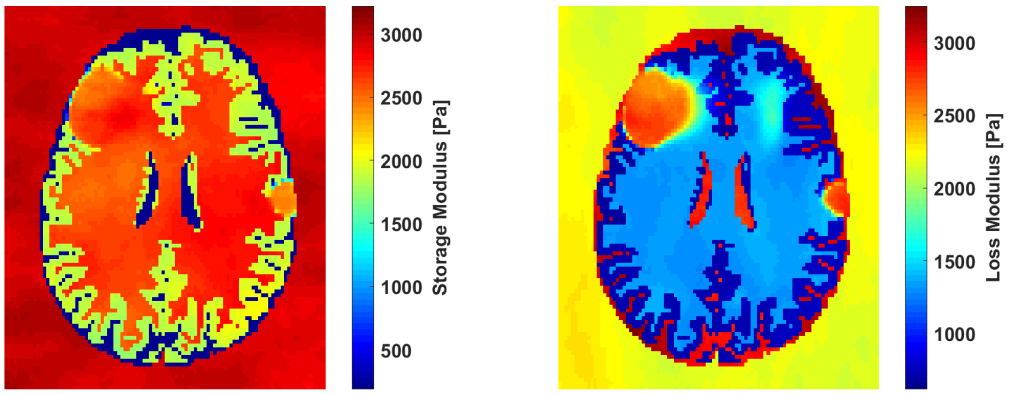


Figure 6.3.32: Physical Posterior Ensemble Mean $\overline{\mu_{N_S}}$. **Left:** Storage modulus. **Right:** Loss Modulus

It can clearly be seen that the variance in the prior ensemble $\mu_0^{(j)}$ in the sample is much larger than the variance in the prior ensemble in Section 6.3.5 and Section 6.3.1. Also, the 95% sample confidence intervals $\mathrm{CI}_{95,0}$ is larger than in Section 6.3.5 and Section 6.3.1. The prior physical ensemble mean is similar to the physical ensemble mean in Section 6.3.5.

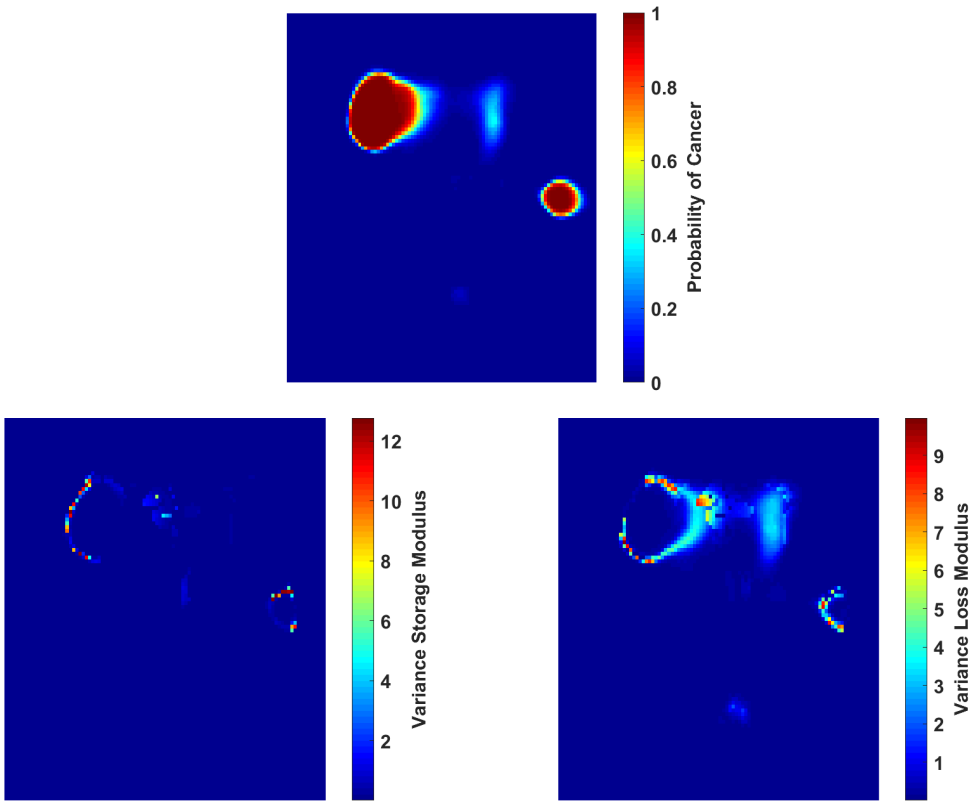


Figure 6.3.33: **Top:** Probability of Cancer P_C. **Bottom Left:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu_{N_S}^{(j)}$; storage modulus. **Bottom Right:** Sample variance $\mathfrak{s}_n$ for posterior ensemble $\mu_{N_S}^{(j)}$; loss modulus.

In Figure 6.3.32 we show plots of the physical posterior ensemble mean $\overline{\mu_{N_S}}$ for the storage and loss modulus. It can be seen that despite the mean of the prior physical ensemble being far off the ground truth, the posterior ensemble mean has both cancer tissue in the locations of the cancer tissue in the ground truth. Also, the probability of cancer P_C shown in the plot in Figure 6.3.33 is almost 1 in the two cancer regions. The variance $\mathfrak{s}_{N_S}^2$ shown in Figure 6.3.33 is

high around the interface between cancer tissue and healthy tissue but also has higher values in other areas. The sample variance $\mathfrak{s}^2_{N_S}$ in the posterior ensemble is in many regions higher than the sample variance in the posterior in Section 6.3.5 where we used the same noise level in the noise, the same ensemble size and a similar prior ensemble with a higher variance in the physical ensemble, however.

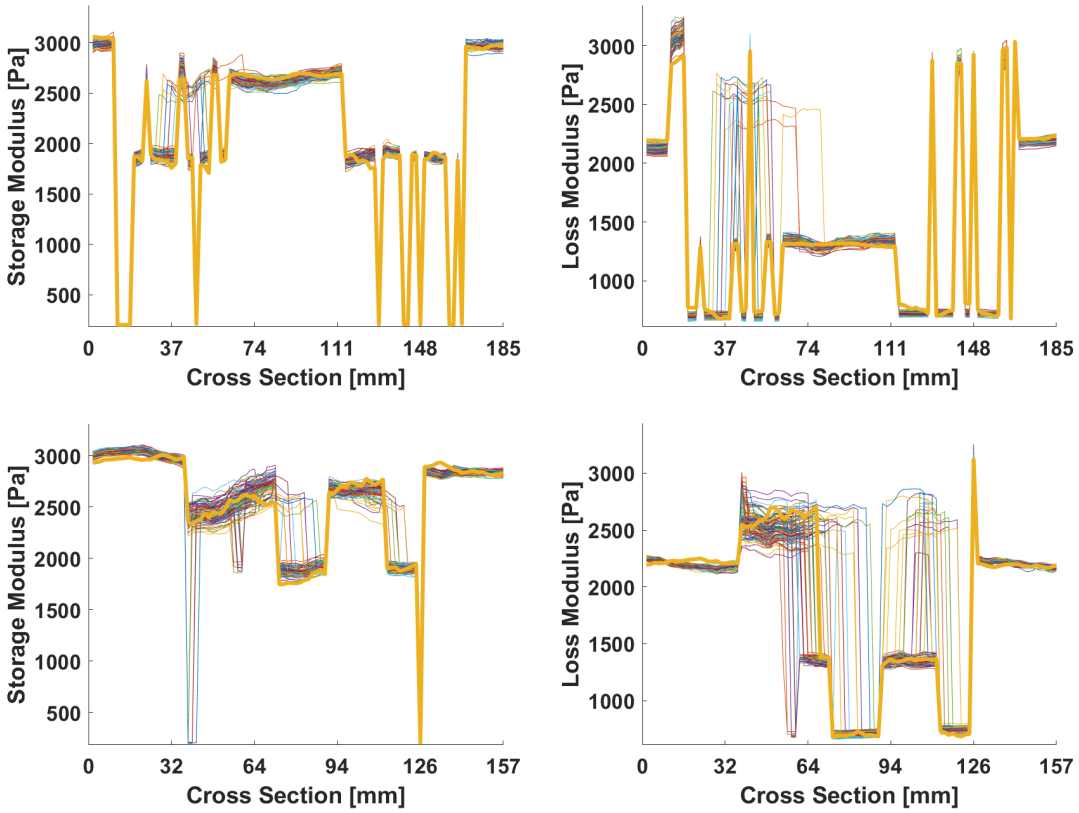


Figure 6.3.34: Posterior ensemble $\mu_{N_S}^{(j)}$ in cross-sections indicated in Fig. 6.3.1. The yellow line indicates the ground truth $\mu^{\dagger}$. **Top left:** Posterior ensemble for storage modulus in vertical cross-section. **Top right:** Posterior ensemble for loss modulus in vertical cross-section. **Bottom left:** Posterior ensemble for storage modulus in horizontal cross-section. **Bottom right:** Posterior ensemble for loss modulus in horizontal cross-section.

The variance in the posterior ensemble around the interface between cancer and healthy tissue can also be seen in plots in Figure 6.3.34 in form of variability of loss and storage modulus values in posterior ensemble around the interface

in the two cross-sections. In Figure 6.3.35 we give plots of the 95% sample confidence interval CI_{95,N_S} of the posterior ensemble in the two cross-sections. The confidence intervals of the posterior ensemble are larger than the confidence intervals of the posterior ensemble in Section 6.3.5 and Section 6.3.1, where the sample variance in the physical prior ensemble was smaller than in this section. In all the cross-sectional plots it can be seen that the posterior ensemble has values close to the ground truth.

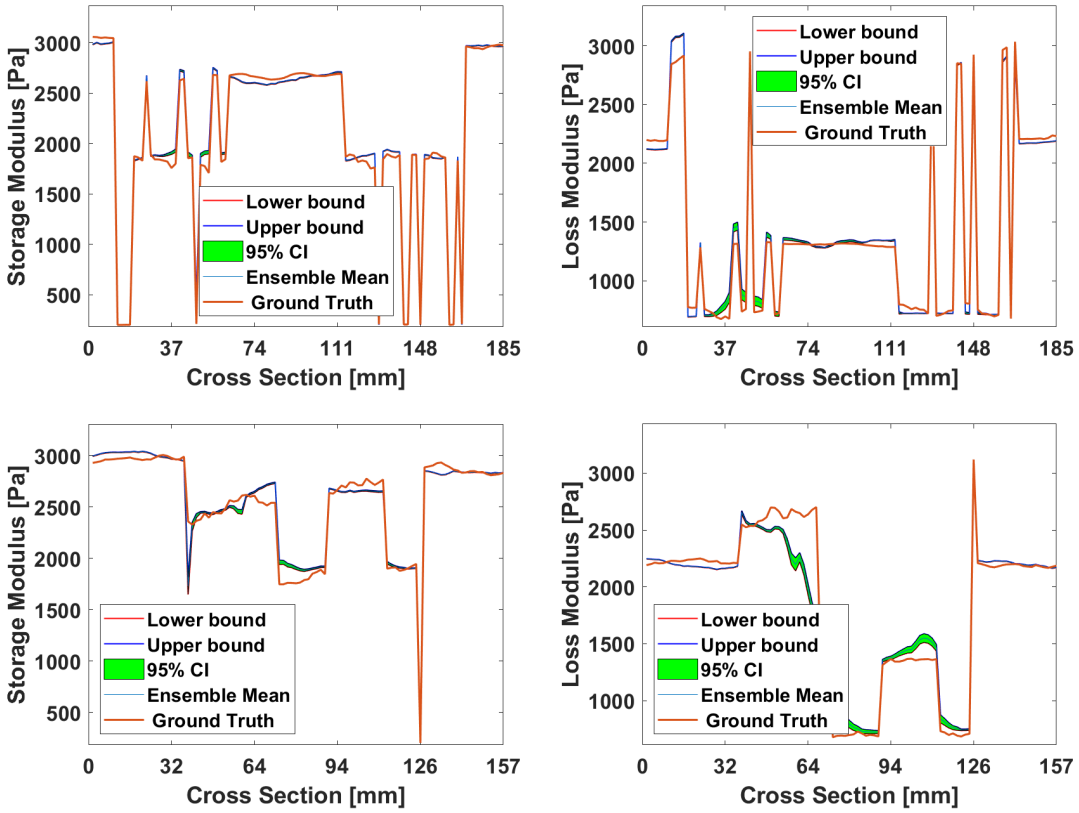


Figure 6.3.35: 95% Confidence Intervals CI_{95,N_S} in cross-sections indicated in Fig. 6.3.1 for the posterior ensemble $\mu_{N_S}^{(j)}$. We plot the lower and the upper bound of the 95% confidence interval as well as the ensemble mean $\overline{\mu_{N_S}}$ and the ground truth $\mu^\dagger$. **Top left:** Storage modulus in vertical cross-section. **Top right:** Loss modulus in vertical cross-section. **Bottom left:** Storage modulus in horizontal cross-section. **Bottom right:** Loss modulus in horizontal cross-section.

In general the prior in this section leads to a posterior ensemble that is closer to the ground truth than the posterior in Section 6.3.5.

6.3.8 *Conclusion Posterior Variance and Confidence Intervals*

The series of experiments with the focus on confidence intervals and variance conducted in Section 6.3.1-6.3.7 first of all confirmed some results seen in Section 6.2: EKI in application to MRE with *in silico* model data as input can provide robust approximations of shear moduli for high noise levels in the data. Prior ensembles with a physical ensemble mean far off the ground truth lead to posteriors that do not detect the cancer or its location. In every experiment a smaller ensemble sizes significantly increase the difference between the posterior ensemble mean and the ground truth and therefore, produce results that are not useful for the detection of cancer.

Additional to these findings that confirm results in Section 6.2, we can observe that the sample variance in the physical posterior ensemble increases when the noise level in the noisy data is increased. Furthermore, we can notice a dependency of the sample variance in the physical posterior ensemble and the ensemble size. We find a smaller sample variance in the physical posterior ensemble for smaller ensemble sizes. Moreover, we noticed that a higher sample variance in the physical prior ensemble leads to a higher variance in the physical posterior ensemble.

However, we cannot see a link between the variance and the quality of the posterior ensemble as approximation of the unknown. In cases where the posterior ensemble and the posterior ensemble mean is far off the unknown, the variance in the posterior ensemble is low and the probability of cancer is low even in areas with cancer tissue. We can see that the sample variance in the posterior ensemble is even lower if the posterior ensemble is far off. We show in Section 6.2 that when making a statement on the level of certainty of an approximation using the EKI we should also take the average data misfit into account.

The definition of the probability of cancer seems to be a useful summary of the results.

Finally, we have shown that a prior with a large sample variance leads to posterior ensembles that give accurate approximations of the unknown even if the mean of the physical prior ensemble is far off the unknown. This can be used in situations where we are uncertain about what shear modulus can be expected before the inversion in an experiment.

6.4 DIRECT INVERSION

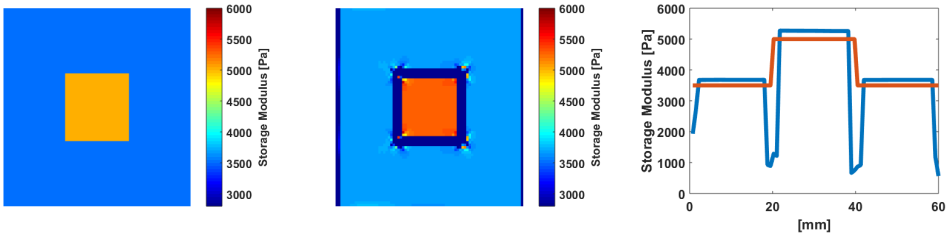
In this section we look at strength and weaknesses of direct inversion, a widely used approach to the inverse problem arising in MRE [14] which we described in Section 3.5. Furthermore, we compare the performance of direct inversion with EKI in different experimental setups.

Direct inversion is computationally the cheapest common inversion approach to MRE [14] and gives good approximations of storage and loss modulus distributions if the ground truth distribution of these quantities have areas of constant values. In Figure 6.4.1 we show plots of direct inversion applied to *in silico* model data coming from various ground truths in an experimental setup already described in Section 5.1. In Figure 6.4.1 (b) and (c) we can see that the shear modulus approximation provided by direct inversion is accurate in areas with constant ground truth distribution.

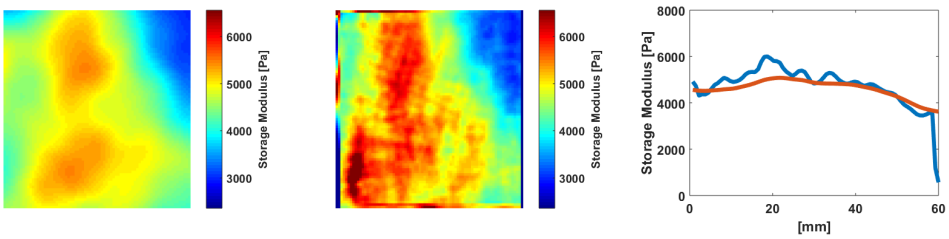
Direct inversion performs poorly if the ground truth distribution of shear modulus contains heterogeneity, i.e. these quantities vary in a local neighbourhood. This is due to the fact that the ”local heterogeneity assumption” discussed in Section 3.5.1 does not hold. Adding to this, the derivatives used in this approach contain errors when calculated using finite differences on a relatively coarse computational grid. The ground truth distribution of the shear

modulus of the experiment shown in the plot in Figure 6.4.1 (d) varies smoothly. The direct inversion estimates of the shear modulus shown in Figure 6.4.1 (e) and (f) contain significant inaccuracies.

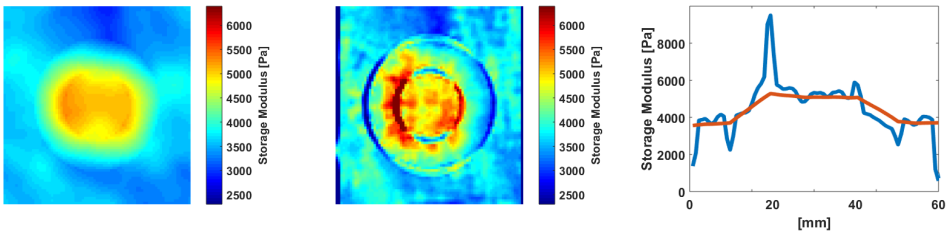
In the presence of discontinuities in the ground truth distribution of shear modulus we observe a considerable error in the inversion estimate provided by direct inversion. In Figure 6.4.1 (a) the ground truth distribution of the shear modulus has a discontinuous jump from a background value of 3500Pa to 5000Pa. The direct inversion estimate around this jump is as low as 800Pa. Also, we can see in Figure 6.4.1 (h) and (i) that around points of steep changes in the gradient of the ground truth distribution of the shear modulus lead to high errors in the estimate provided by direct inversion there. We can see considerable spikes in the direct inversion reconstruction around the beginning and the end of the slope between the background and the inclusion value of the shear modulus.



(a) Ground Truth (b) Direct Inversion (c) Cross-Sectional Plot



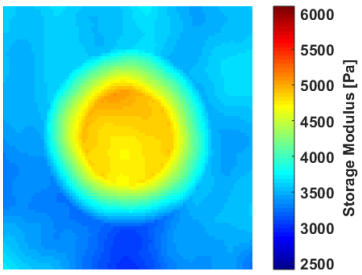
(d) Ground Truth (e) Direct Inversion (f) Cross-Sectional Plot



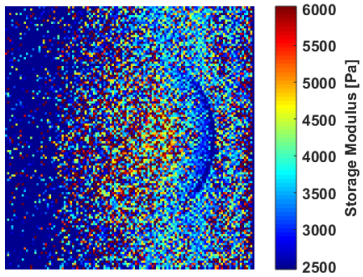
(g) Ground Truth (h) Direct Inversion (i) Cross-Sectional Plot

Figure 6.4.1: Direct Inversion of displacement fields without noise coming from different ground truth distributions of the shear modulus across the domain. The basic setup of this experiment is described in Section 5.1. Apart from the different ground truth distribution and the absence of noise, the domain size is changed to 60mm by 60 mm and displacements are given on an 80 by 80 grid. The frequency of the sinusoidal excitation is 60Hz. The cross-section of the cross-sectional plot is a horizontal line right in the middle of the domain

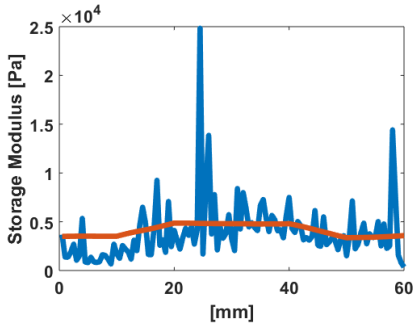
Plain algebraic inversion as described in Section 3.5 without modifications does not cope with the presence of noise in the data. In Figure 6.4.2 (b) and (c) we show plots of direct inversion approximations of the shear modulus for data corrupted by noise with a noise level of 5% which we defined in (5.2.1).



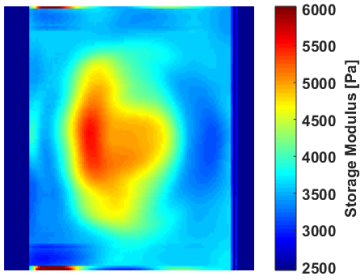
(a) Ground Truth μ



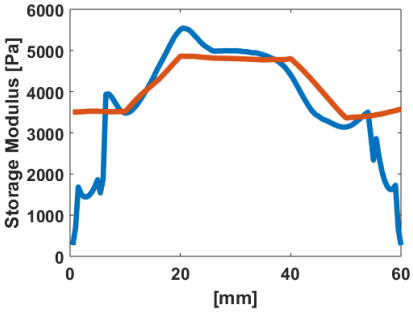
(b) Direct Inversion w/o filter



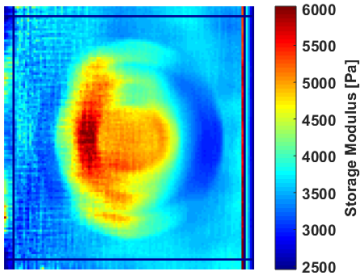
(c) Cross-Sectional Plot



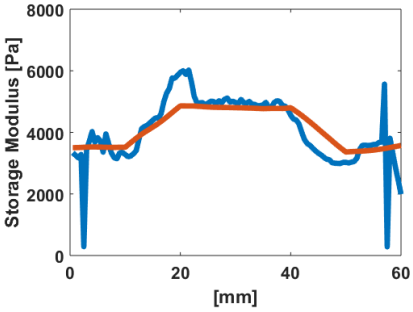
(d) *Imgaussfilt* before Direct Inversion



(e) Cross-Sectional Plot



(f) Meanfilter after curl



(g) Cross-Sectional Plot

Figure 6.4.2: Direct Inversion applied to displacement fields corrupted by noise with a noise level of 5% defined in (5.2.1). Basic setup is the same as in Figure 6.4.1 and is described there. The data is given on a 120 by 120 grid.

Direct inversion is often used in combination with filtering methods [68]. In Figure 6.4.2 (d)-(g) we show plots of direct inversion results where common filtering methods are combined with direct inversion. Both the filter using a Gaussian smoothing kernel with standard deviation 5 (Figure 6.4.2 (d); matlab inbuilt function *Imgaussfilt* [53]) and the mean filter (Figure 6.4.2 (f); [20]) are applied in a 10 by 10 nodes neighbourhood and successfully remove the noise. However, they introduce new errors due to filtering, especially around the boundary of the domain. The use of Gaussian smoothing kernels with other standard deviations does not change the result.

In Figure 6.4.3 we show plots of the posterior mean of the EKI that we get in experiments described in Chapter 6. We compare them to direct inversion results we get for these experimental setups. It is clear that even in the best case scenario for direct inversion, that is, without any noise in the data, EKI estimates outperform the accuracy of direct inversion. The main reason apart from the weaknesses of direct inversion described in this Section already is, that there is no straight forward way to code in prior knowledge into this approach unlike in EKI.

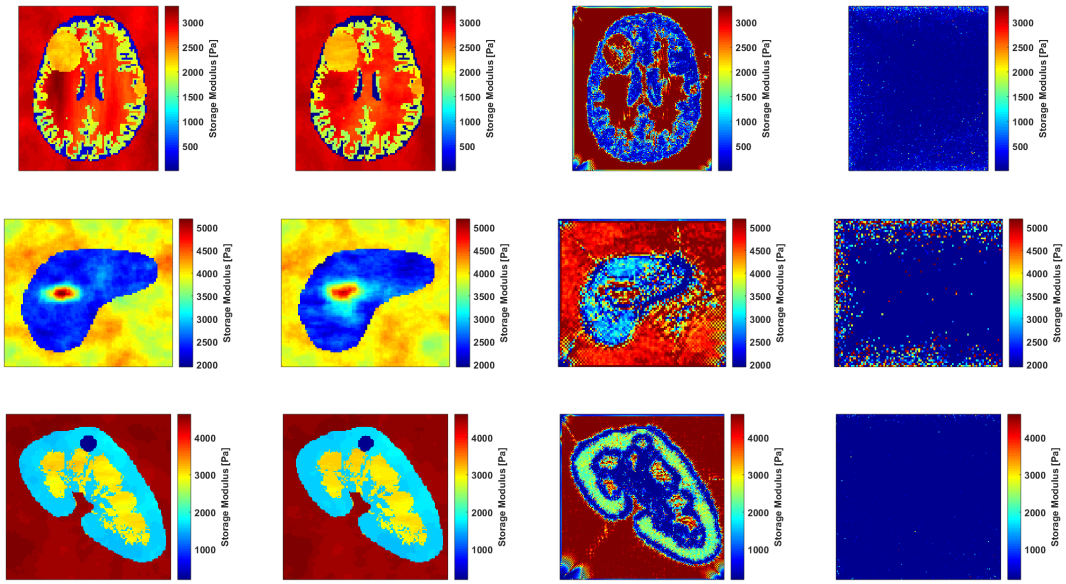


Figure 6.4.3: **First Column:** Ground truth distribution of storage modulus μ_S across domain. **Second Column:** Posterior mean of EKI (5 % noise in data). **Third Column:** Direct inversion (without filtering) applied to data without noise. **Fourth Column:** Direct Inversion (without filtering) applied to data with 5 % noise added.

6.5 FURTHER IMPROVEMENTS FOR THE PRIOR

In this section we describe two ideas to design informative priors for EKI by using direct inversion results. They both turned out to be less successful than expected. Nonetheless, we give a short description of them because it is interesting to see why they fail.

6.5.1 *Preconditioning on Direct Inversion*

In order to find good priors for EKI *gpstuff* was used in order to find maximum *a posteriori* estimates of the parameters of a Whittle-Matérn Gaussian random field approximating data coming from direct inversion. In Appendix A we describe how *gpstuff* finds *a posteriori* estimates of the parameters. The *a posteriori* estimates of the parameters are then used as parameters described in Section 4.3.2 in the Whittle-Matérn autocorrelation function defining the Covariance of the prior and the direct inversion result as the non-homogeneous mean of the prior.

In middle plot of Fig. 6.5.1 we can see a cross-sectional plot of the prior ensemble obtained by applying this hyperparameter estimation. In many cases the maximum *a posteriori* estimates of the parameters were similar to the hyperparameters we have chosen before. In almost every application of this way of finding good priors the direct inversion result was too poor. The prior conditioned on this result did not reduce the amount of iterations in EKI or increased the accuracy of the EKI-estimates. In the right plot in Fig. 6.5.1 we can see the EKI estimate using a prior from a Gaussian random field with hyperparameter estimation.

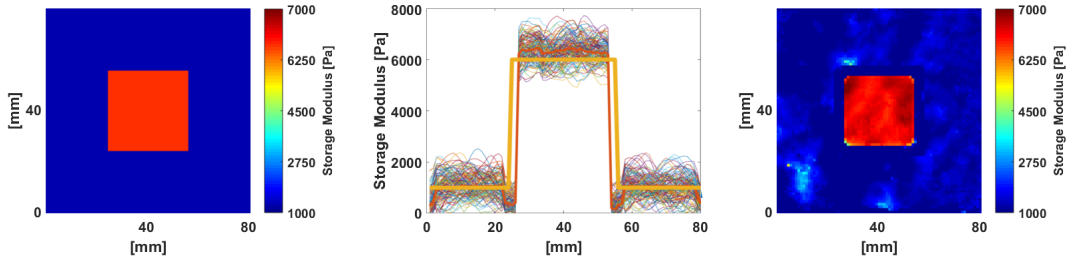


Figure 6.5.1: Set up of this experiment is the same as the one for Figure 5.7.7. **Left:** Ground truth storage modulus distribution **Mid:** Initial Ensemble, Direct Inversion (Orange) and Ground Truth(Yellow) in a vertical cross-section (drawn from GRF with hyperparameter estimation). **Right:** Posterior Ensemble Mean

The main problem with this approach of defining good priors is, that direct inversion results are not good enough for priors to be conditioned on. Even the direct inversion result for a relatively simple ground truth seen in Figure 6.5.1 has large errors around the interface and therefore, priors conditioned on that tend to be not useful.

6.5.2 *EKI applied to the Identity*

Due to the high sensitivity to noise of direct inversion, we apply Algorithm 1 to data $\mathbf{v}$ but we use the identity map $\mathcal{G} := \mathrm{id} : \mathbb{C}^m \to \mathbb{C}^m$ as forward operator. The idea is to use EKI in order to remove noise in the data and to get a prior ensemble for the application of EKI for estimating the shear modulus.

In the first application of EKI a prior for the displacement data is chosen using a Whittle-Matérn auto-correlation function described in Section 4.3. Then Algorithm 1 is run with this prior, data $\mathbf{v}$ and forward operator $\mathcal{G} = \mathrm{id}$. The posterior ensemble (which is estimating the displacement data) is then inverted by using direct inversion (3.5.5). The inverted posterior ensemble then becomes the prior ensemble for the second application of EKI that uses (3.4.2) as the forward map as before.

In fact, EKI is successfully removing noise from displacement data for various priors and noise levels. In Figure 6.5.2 we provide plots of an posterior ensemble mean that gives a good approximation of the ground truth displacement data.

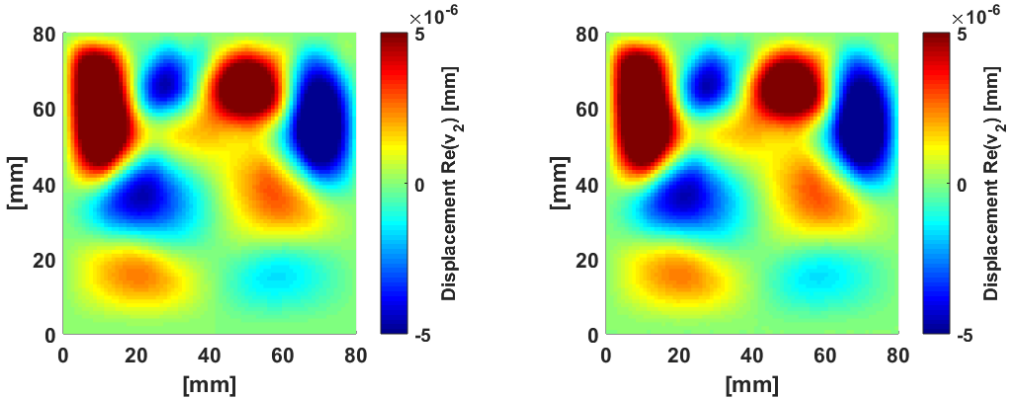


Figure 6.5.2: **Left:** Real part of the second component of the ground truth displacement data **Right:** Posterior Ensemble Mean of EKI applied to data with 5% noise level.

However, the posterior ensemble has a very small variance. In Figure 6.5.3 we provide a cross-sectional plot of the posterior and can see the small variance. There are ways to artificially add variance to the ensemble before inverting it using direct inversion, e.g. by stopping EKI early. However - as discussed in Section 6.4 - direct inversion does not only cope well with noise but also with heterogeneity. Noise could be removed from the data but direct inversion remains inaccurate for discontinuous or heterogeneous distributions of the shear modulus. Therefore, this method does not provide informative priors.

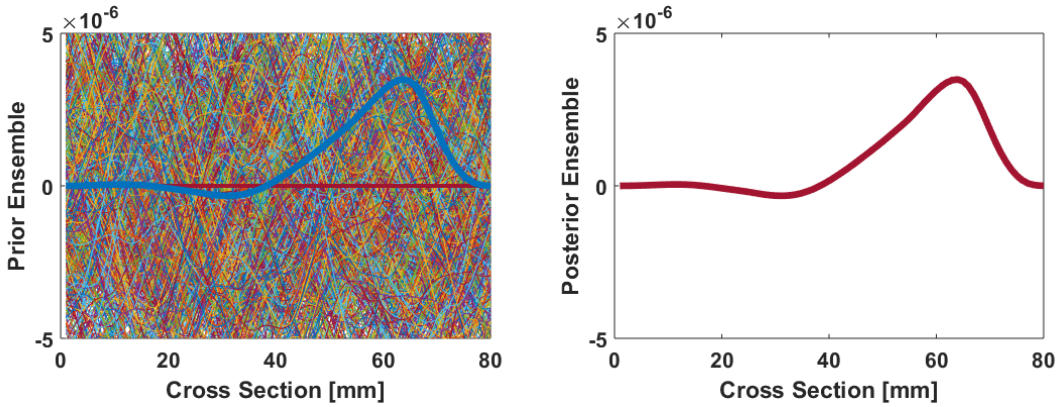


Figure 6.5.3: **Left:** Cross-sectional plot prior ensemble, ground truth (blue) and mean of the prior ensemble (red). **Right:** Cross-sectional plot posterior ensemble.

6.6 SUMMARY AND CONCLUSION

EKI was successfully applied to *in silico* model examples of Brain and Kidney MRE. The influence of ensemble size and noise in the data on accuracy of the estimate was quantified. It was shown that EKI is robust and computational cheap in this application. It outperformed direct inversion in accuracy of the estimates for most priors chosen, especially in the presence of noise and heterogeneity in the ground truth. This is not only due to the fact that EKI copes better with noisy data but also because EKI with two-level-parametrisation can capture sharp discontinuities and high levels of heterogeneity in the spatial distribution of the physical parameters. A key contribution to the high accuracy of EKI estimates is the fact that prior knowledge about the unknown can be incorporated into the inversion approach in a simple fashion. Non-Bayesian inversion approaches usually do not have the feature to incorporate detailed prior knowledge.

The idea to use direct inversion in order to provide informative priors for EKI and by that combine these two inversion approaches turned out to be unsuccessful, mainly because of the inaccuracy of direct inversion estimates in many scenarios.

7

CONCLUSION AND FUTURE WORK

In this Chapter's first section we summarise the main findings of this thesis and have a look at potential extensions and future work in the second section.

7.1 CONCLUSION

We showed in *in silico* model examples that EKI applied to the inverse problem arising in MRE can provide accurate estimates of the unknown local soft tissue stiffness when the viscoelastic model is used. For the viscoelastic model, EKI together with the two-level parametrisation is capable of giving estimates of unknowns with complex geometries found in, for example, brain MRE, and it can deliver confidence intervals for these estimates. In the virtual experiments with brain geometries in the unknown EKI could successfully discover cancer tissue with good accuracy.

Some other main findings in this thesis are:

- We showcased the advantages of various parametrisations of the unknown and the design of informative priors. The parametrisations enable us to fully utilise the MRI data that is available prior to inversion in EKI.
- The viscoelastic model should be preferred over the purely elastic model when EKI is applied to time-harmonic MRE at common frequencies.
- A sufficiently large ensemble size is essential. Then EKI with the two-level parametrisation can provide accurate estimates of unknowns with complex geometries even for high amounts of noise and priors relatively far off the ground truth.
- Direct inversion is unsuitable to design informative priors. This is due to the fact that direct inversions are too inaccurate if the unknown soft tissue has spatial variability.

The main reason why the viscoelastic model should be preferred over the purely elastic model is, that soft tissue exhibits viscoelastic behaviour and, in order to minimise modelling errors, it should therefore be modelled as such despite the purely elastic model being mathematically simpler. Besides, the efficient numerical approximation of solutions of the purely elastic wave equation for a high wave number is still an open problem and makes the application of EKI difficult because an accurate approximation of the forward

operator is needed. On top of this, we could show that when using the purely elastic wave equations in the ensemble of wave fields corresponding to the ensemble of elastic parameters there can be resonating wave fields with high amplitudes which lead to outliers in the ensemble of wave fields that make the whole ensemble diverge.

The level set parametrisation used in experiments in Chapter 5 enables EKI to approximate sharp interfaces between different tissue types for example diseased tissue and healthy tissue. However, this parametrisation does not allow for heterogeneous approximations within a tissue type and in the version used in this thesis it requires the prior knowledge of mechanical parameters of certain tissue types, i.e. EKI with this level set parametrisation only approximates the location and shape of certain tissue types, not the mechanical parameters of the tissue.

The two-level parametrisation used in Chapter 6 which extends the level set parametrisation by a second level, allows for heterogeneity within different tissue types. Also, EKI with the two-level parametrisation approximates the mechanical parameters of different tissue types and does not require exact prior knowledge of them. We use this parametrisation in EKI in order to approximate the unknown shape and location of the heterogeneous mechanical parameters of cancer tissue which has a sharp interface with healthy tissue.

7.2 FUTURE WORK

A natural extension to this work would be the analysis of noisy boundary conditions. In real-world applications we would define our domain of interest Ω to be a certain organ or part of the body, meaning that the boundary conditions would be Dirichlet boundaries given by the measurements along the boundary of Ω. These measurements are corrupted by noise which carries through to the approximated wave field for each sample in the ensemble which then affects the ensemble updates. The authors in [16] analyse EKI for another elliptic inverse problem with uncertain boundary conditions but not the inverse problem arising in MRE.

Another direction would be to further exploit the fact that EKI is computational cheap and extend the experiments to three-dimensions. The code could then be applied to real-world data as well.

A further extension to the experiments shown in Chapter 6 would be to run these experiments with unknown smoothness and amplitude hyperparameters. We have fixed them to the values of these hyperparameters in the ground truth because we believe that the lengthscale and mean hyperparameter is sufficient to capture spatial variability. However, this needs to be confirmed in experiments.

An interesting theoretical problem connected to this thesis is to study forward map $\mathcal{F}$ defined in (4.1.1) for continuity. More precisely, [83] reviews conditions under which the solution of Bayesian inversions is well posed. They show that the Bayesian posterior is well-posed if a certain continuity assumption on the forward map $\mathcal{F}$ holds, i.e. the forward map changes continuously with respect to parameter k. In experiments, we could observe that $\mathcal{F}$ in the purely elastic case behaves like a discontinuous map due to resonating wave fields. For the viscoelastic case we did not make this observation. So the natural question that arises is if these properties of the viscoelastic and purely elastic forward map can be rigorously proven.

Finally, a modification of EKI, the stochastic ensemble Kalman inversion could be considered instead. We could show in Chapter 6 that the posterior ensemble can lie far off the ground truth in cases with small ensemble size and large noise level and the posterior ensemble variance is very low. In fact, it is a known problem of Ensemble Kalman Inversion and Ensemble Kalman filters in general that in the case of a small ensemble size and a large amount of noise in the data the ensemble collapses, the ensemble variance becomes very small and iterates can end up in a deadlock [4, 5]. [93] shows how a variance inflation approach can reduce this problem.

A

GPSTUFF

In this Chapter of the appendix we want to give a short description of the matlab software *gpstuff* used in Section 6.5.1 in order estimate hyperparameters. We follow [71] in this Chapter.

In all our applications of EKI we need to somehow give an estimate of the hyperparameters lengthscale l, amplitude scale σ^2 and smoothness parameters r in the autocorrelation function (4.3.1) that is used to define the prior, either in form of a single estimate like in Chapter 5 or in form of an interval of a uniform distribution like in Chapter 6 for the lengthscale. In Section 6.5.1 we try to give estimates for these hyperparamters by using *gpstuff* with direct inversion results as inputs.

More precisely, let $\mu_{S_i}^{\mathrm{dI}}$ be the result from direct inversion given in points $x_i \in \Omega \subset \mathbb{R}^2$, $i = 1, \ldots, \hat{n}$. Here, μ_{S_i} is the storage modulus because we assume in Section 6.5.1 the loss modulus to be a known constant.

In GP regression observation pairs $\mathfrak{D} = \{(x_i, \mu_{S_i}), i = 1, \ldots, \hat{n}\}$ are assumed to be modelled by an unknown function $\mathfrak{g}$ in the following way

$$\mu_{S_i}^{\mathrm{dI}} = \mathfrak{g}(x_i) + \hat{\eta}_i, \tag{A.0.1}$$

with noise assumed to be $\hat{\eta}_i \sim \mathcal{N}(0, \hat{b})$ independent identically Gaussian distributed.

The idea of GP-regression is to make also the assumption that $\mathfrak{g} : \Omega \to \mathbb{R}^{\hat{n}}$ is a Gaussian process $\mathrm{GP}(m_{GP}^*, \Sigma_{GP}^*)$ with some mean $m_{GP} \in \mathbb{R}^{\hat{n}}$ and covariance $\Sigma_{GP} \in \mathbb{R}^{\hat{n}} \times \mathbb{R}^{\hat{n}}$, that is, for each $x_1, \ldots, x_{\hat{n}}$ we have $(\mathfrak{g}(x_1), \cdots \mathfrak{g}(x_{\hat{n}})) \sim \mathcal{N}(m_{GP}^*, \Sigma_{GP}^*)$. In addition, we follow an Bayesian approach, so we seek a posterior distribution of $\mathfrak{g}$ given data $\mathfrak{D}$. The likelihood based on $\mathfrak{D}$ in this scenario is $L_{GP}(\mathfrak{g}, \hat{b}) = \prod_{i=1}^{\hat{n}} \mathcal{N}(\mathfrak{g}(x_i), \hat{b})$. We choose a Gaussian process as prior of $\mathfrak{g}$ with mean function m_{GP} and covariance function Σ_{GP}. This makes the posterior a Gaussian process as well as mentioned before. The posterior mean in $x \in \Omega$ is given by

$$m_{\mathrm{GP}}^*(x) = \Sigma_{\mathrm{GP}}(x, \mathbf{x}) \left[\Sigma_{\mathrm{GP}}(\mathbf{x}, \mathbf{x}) + \hat{b}\mathbf{I}\right]^{-1} \mu_{S_i},$$

and the covariance in (x, y) by

$$s^*(x, y) = s(x, y) - \Sigma_{\mathrm{GP}}(x, \mathbf{x}) \left[\Sigma_{\mathrm{GP}}(\mathbf{x}, \mathbf{x}) + \hat{b}\mathbf{I}\right]^{-1} \Sigma_{\mathrm{GP}}(\mathbf{x}, x),$$

where $\Sigma_{\mathrm{GP}}(x, \mathbf{x}) = (s(x, x_1), \ldots, s(x, x_{\hat{n}}))^T$, $\mathbf{I}$ is the $\hat{n}$-dimensional identity matrix and $\Sigma_{\mathrm{GP}}(\mathbf{x}, \mathbf{x})$ is the covariance matrix based on evaluations of k at pairs (x_i, x_j), $i, j = 1, \ldots, \hat{n}$.

The covariance Σ_{GP} of a Gaussian process depends on parameters. In the case of Section 6.5.1 they are the parameters $\Theta_{GP} := (l, \sigma^2, r)$ of the Whittle-Matérn autocorrelation function.

Instead of finding the mean m^*_{GP} and covariance Σ^*_{GP} of the posterior Gaussian process $\mathrm{GP}(m^*_{GP}, \Sigma^*_{GP})$ we now try to estimate the parameter Θ^*_{GP} of the covariance Σ^*_{GP}. The posterior distribution of Θ^*_{GP} given data $(\mathbf{x}, \mu_S) = \{(x_i, \mu_{S_i}), i = 1, \ldots, \hat{n}\}$ is given by

$$\rho(\Theta_{GP}|x, \mu_S) = \frac{L(\mathfrak{g}, \Theta_{GP})\rho(\Theta_{GP})}{\int_{\Theta_{GP}} L(\mathfrak{g}, \Theta_{GP})\rho(\Theta_{GP})},$$

where $L(\mathfrak{g}, \Theta_{GP})$ is the likelihood function depending on Θ_{GP} and $\rho(\Theta_{GP}) = \rho_1(l)\rho_2(\sigma^2)\rho_3(r)$ is the prior. In section 6.5.1 we used uniform distributions ρ_1, ρ_2, ρ_3 on various intervals. We maximise $\rho(\Theta_{GP}|\mathbf{x}, \mu_S)$ with respect to Θ_{GP} in order to obtain a *maximum a posteriori estimate*. It can be shown [71] that maximising $\rho(\Theta_{GP}|\mathbf{x}, \mu_S)$ with respect to Θ_{GP} is equivalent to maximise

$$-\frac{1}{2}\mu^T\Theta_{GP}^{-1}\mu - \frac{1}{2}\log|\Theta_{GP}| - \frac{n}{2}\log 2\pi + \log\rho(\Theta_{GP})$$

for Θ_{GP} for chosen prior $\rho(\Theta_{GP})$ where $\Sigma_{\Theta_{GP}}$ is the Whittle-Matérn covariance using parameters $\Sigma_{\Theta_{GP}}$. Maximisation is carried out by *gpstuff* by gradient-based algorithms.

BIBLIOGRAPHY

[1] S. Aanonsen, G. Naevdal, D.Oliver, A. Reynolds, and B. Valles, *The ensemble kalman filter in reservoir engineering–a review*, SPE j., 14 (2009), pp. 393–412.

[2] H. Amman and J. Escher, *Analysis I*, Springer Berlin Heidelberg, 2008.

[3] H. Amman and J. Escher, *Analysis II*, Springer Berlin Heidelberg, 2008.

[4] J. Anderson, *An adaptive covariance inflation error correction algorithm for ensemble filters*, Tellus A, 59 (2007), pp. 210–224.

[5] J. Anderson, *Spatially and temporally varying adaptive covariance inflation for ensemble filters*, Tellus A, 61 (2009), pp. 72–83.

[6] T. Arbogast and J.Bona, *Methods of Applied Mathematics*, Lecture Notes, 2008.

[7] I. Babuska and S. Sauter, *Is the pollution effect of the fem avoidable for the helmholtz equation considering high wave numbers?*, SIAM Review, 42 (2000), pp. 451–484.

[8] R. H. Barnard, *Elementary fluid dynamics*, The Aeronautical Journal, 94 (1990), p. 265–265.

[9] T. Belytschko, *The finite element method: Linear static and dynamic finite element analysis: Thomas j. r. hughes*, Computer-Aided Civil and Infrastructure Engineering, 4 (2008), pp. 245–246.

[10] A. Beskos, D. Crisan, and A. Jasra, *On the stability of sequential monte carlo methods in high dimensions*, The Annals of Applied Probability, 24 (2014), pp. 1396–1445.

[11] D. Calvetti and E. Somersalo, *Inverse problems: From regularization to bayesian inference*, WIREs Computational Statistics, 10 (2018).

[12] N. Chada and X.Tong, *Convergence acceleration of ensemble kalman inversion in nonlinear settings*, 2019.

[13] D. Colton, M. Piana, and R. Potthast, *A simple method using morozov's discrepancy principle for solving inverse scattering problems*, Inverse Problems, 13 (1997), pp. 1477–1493.

[14] M. Doyley, *Model-based elastography: a survey of approaches to the inverse elasticity problem*, Physics in Medicine and Biology, 57 (2012), pp. R35–R73.

[15] M. Dresner, G. Rose, P. Rossman, R. Muthupillai, A. Manduca, and R. Ehman, *Magnetic resonance elastography of skeletal muscle*, Journal of Magnetic Resonance Imaging, 13 (2001), pp. 269–276.

[16] S. Dubinkina and S. Ruchi, *Comparison of regularized ensemble kalman filter and tempered ensemble transform particle filter for an elliptic inverse problem with uncertain boundary conditions*, Computational Geosciences, 24 (2019), pp. 149–160.

[17] T. Elgeti, J. Rump, U. Hamhaber, S. Papazoglou, B. Hamm, J. Braun, and I. Sack, *Cardiac magnetic resonance elastography. initial results*, Invest. Radiol., 43 (2008), pp. 762–772.

[18] G. Evensen, *Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics*, J. Geophys. Res., 99 (1994), p. 10143.

[19] G. Evensen, *Data Assimilation*, Springer, 2 ed., Jan. 2009.

[20] L. Fan, F. Zhang, H.Fan, and C. Zhang, *Brief review of image denoising techniques*, Visual Computing for Industry, Biomedicine, and Art, (2019).

[21] J. Fessler, *Model-based image reconstruction for MRI*, IEEE Signal Process. Mag., 27 (2010), pp. 81–89.

[22] D. Fovargue, S. Kozerke, R. Sinkus, and D. Nordsletten, *Robust MR elastography stiffness quantification using a localized divergence free finite element reconstruction*, Medical Image Analysis, 44 (2018), pp. 126–142.

[23] D. Fovargue, D. Nordsletten, and R. Sinkus, *Stiffness reconstruction methods for MR elastography*, NMR in Biomedicine, 31 (2018), p. e3935.

[24] D. Gallistl, D. Peterseim, and C. Carstensen, *Multiscale petrov-galerkin FEM for acoustic scattering*, PAMM, 16 (2016), pp. 745–746.

[25] H. Gimperlein and A. Waters, *Stability analysis in magnetic resonance elastography ii*, J. Mathematical Analysis and Applications, 434 (2016), p. 1801–1812.

[26] K. Glaser, A. Manduca, and R. Ehman, *Review of MR elastography applications and recent developments*, J. Magn. Reson. Imaging, 36 (2012), pp. 757–774.

[27] B. Goss, K. McGee, E. Ehman, A. Manduca, and R. Ehman, *Magnetic resonance elastography of the lung: technical feasibility*, Magn. Reson. Med., 56 (2006), pp. 1060–1066.

[28] L. Guo, J. Vardakis, D. Chou, and Y. Ventikos, *A multiple-network poroelastic model for biological systems and application to subject-specific modelling*

of cerebral fluid transport, International Journal of Engineering Science, 147 (2020), p. 103204.

[29] M. Hanke, *A regularizing levenberg - marquardt scheme, with applications to inverse groundwater filtration problems*, Inverse Problems, 13 (1997), pp. 79–95.

[30] M. Honarvar, R. Sahebjavaher, R. Sinkus, R. Rohling, and S. Salcudean, *Curl-based finite element reconstruction of the shear modulus without assuming local homogeneity: Time harmonic case*, IEEE Trans. Med. Imaging, 32 (2013), pp. 2189–2199.

[31] E. V. Houten, M. Michael, J. Weaver, F. Kennedy, and K. Paulsen, *Three-dimensional subzone-based reconstruction algorithm for mr elastography*, Magnetic Resonance in Medicine, 45 (2001), pp. 827–837.

[32] E. V. Houten, K. Paulsen, M. Miga, F. Kennedy, and J. Weaver, *An overlapping subzone technique for MR-based elastic property reconstruction*, Magnetic Resonance in Medicine, 42 (1999), pp. 779–786.

[33] S. Hubmer, E. Sherina, A. Neubauer, and O. Scherzer, *Lamé parameter estimation from static displacement field measurements in the framework of nonlinear inverse problems*, SIAM J. Imaging Sciences, 11 (2018), pp. 1268–1293.

[34] L. R. J. Huttunen and S. Lasanen, *Whittle-matérn priors for bayesian statistical inversion with applications in electrical impedance tomography*, Inverse Problems &amp Imaging, 8 (2014), pp. 561–586.

[35] J. Idier, ed., *Bayesian Approach to Inverse Problems*, ISTE, 2008.

[36] M. Iglesias, K. Law, and A. Stuart, *Ensemble kalman methods for inverse problems*, Inverse Probl., 29 (2013), p. 045001.

[37] M. Iglesias, Y. Lu, and A. Stuart, *A Bayesian Level Set Method for Geometric Inverse Problems*, arXiv e-prints, (2015).

[38] M. Iglesias, M. Park, and M. Tretyakov, *Bayesian inversion in resin transfer molding*, Inverse Problems, 34 (2018), p. 105002.

[39] M. Iglesias and Y. Yang, *Adaptive regularisation for ensemble kalman inversion*, Inverse Problems, 37 (2021), p. 25008.

[40] M. A. Iglesias, *A regularizing iterative ensemble kalman method for PDE-constrained inverse problems*, Inverse Problems, 32 (2016), p. 025002.

[41] Y. Jiang, H. Fujiwara, and G. Nakamura, *Approximate steady state models for magnetic resonance elastography*, SIAM J. Applied Mathematics, 71 (2011), pp. 1965–1989.

[42] Y. Jiang and G. Nakamura, *Viscoelastic properties of soft tissues in a living body measured by MR elastography*, Journal of Physics: Conference Series, 290 (2010), p. 012006.

[43] Y. Jiang, G. Nakamura, and K. Shirota, *Levenberg-Marquardt method for solving inverse problem of MRE based on the modified stationary stokes system*, Inverse Probl., (2021).

[44] C. Johns and J. Mandel, *A two-stage ensemble kalman filter for smooth data assimilation*, Environ. Ecol. Stat., 15 (2008), pp. 101–110.

[45] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, 2005.

[46] J. Kemper, R. Sinkus, J. Lorenzen, C. Nolte-Ernsting, A. Stork, and G. Adam, *MR elastography of the prostate: initial in-vivo application*, Rofo, 176 (2004), pp. 1094–1099.

[47] A. Klenke, *Probability Theory*, Springer London, 2014.

[48] A. Kolmogorov, *Foundations of the theory of probability*, Chealsea Publishing Company, 1950.

[49] L. Landau, L. Pitaevskii, E. Lifshitz, and A. Kosevich, *Theory of elasticity*, Butterworth-Heinemann, Oxford, England, 3 ed., Jan. 1984.

[50] M. Larson and F. Bengzon, *The Finite Element Method: Theory, Implementation, and Applications*, Springer Berlin Heidelberg, 2013.

[51] R. Lerner, S. Huang, and K. Parker, *"sonoelasticity" images derived from ultrasound signals in mechanically vibrated tissues*, Ultrasound in Medicine and Biology, 16 (1990), pp. 231 – 239.

[52] R. Lerner, K. Parker, J. Holen, R. Gramiak, and R. Waag, *Sono-Elasticity: Medical Elasticity Images Derived from Ultrasound Signals in Mechanically Vibrated Targets*, Springer US, 1988.

[53] Q. Li and J. Racine, *Nonparametric Econometrics: Theory and Practice*, Princeton University Press, 12 2006.

[54] O. Lopez, K. Amrami, A. Manduca, P. Rossman, and R. Ehman, *Developments in dynamic MR elastography for in vitro biomechanical assessment of hyaline cartilage under high-frequency cyclical shear*, J. Magn. Reson. Imaging, 25 (2007), pp. 310–320.

[55] G. Low, S. Kruse, and D. Lomas, *General review of magnetic resonance elastography*, World J. Radiol., 8 (2016), pp. 59–72.

[56] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Birkhäuser, 1995.

[57] K. C. M. Meyers, *Mechanical Behavior of Materials*, Cambridge University Press, 2009.

[58] J. Mandel, L. Cobb, and J. Beezley, *On the convergence of the ensemble kalman filter*, Applications of Mathematics, 56 (2011), pp. 533–541.

[59] Y. Mariappan, K. Glaser, and R. Ehman, *Magnetic resonance elastography: A review*, Clinical Anatomy, 23 (2010), pp. 497–511.

[60] MathWorks, *Partial differential equation toolbox™-user's guide*, MathWorks, (2016).

[61] M. Matveev, A. Endruweit, A. Long, M. Iglesias, and M. Tretyakov, *Bayesian inversion algorithm for estimating local variations in permeability and porosity of reinforcements using experimental data*, Composites Part A: Applied Science and Manufacturing, 143 (2021), p. 106323.

[62] M. McGarry, C. Johnson, B. Sutton, J. Georgiadis, E. V. Houten, A. Pattison, J. Weaver, and K. Paulsen, *Suitability of poroelastic and viscoelastic mechanical models for high and low frequency MR elastography*, Medical Physics, 42 (2015), pp. 947–957.

[63] D. McGrath, *Biomechanics of soft tissues: Principles and applications ed a. al mayah*, Boca Raton: CRC Press, Feb. 2018, pp. 55–94.

[64] A. McKnight, J. Kugel, P. Rossman, A. Manduca, L. Hartmann, and R. Ehman, *MR elastography of breast cancer: preliminary results*, AJR Am. J. Roentgenol., 178 (2002), pp. 1411–1417.

[65] J. Meakin, *Velocity Selective Preparations in Magnetic Resonance Imaging*, PhD thesis, University of Oxford, 2013.

[66] W. Means, *Hookean behavior*, in Stress and Strain, Springer New York, New York, NY, 1976, pp. 240–253.

[67] M.Loeève, *Probability Theory*, Springer, 1978.

[68] M. Murphy, J. Huston, C. Jack, K. Glaser, M. Senjem, J. Chen, A. Manduca, J. Felmlee, and R. Ehman, *Measuring the characteristic topography of brain stiffness with magnetic resonance elastography*, 8 (2013), p. e81668.

[69] R. Muthupillai, D. Lomas, P.Rossman, J. Greenleaf, a. Manduca, and R. Ehman, *Magnetic resonance elastography by direct visualization of propagating acoustic strain waves*, Science, 269 (1995), pp. 1854–1857.

[70] P. Perrinez, F. Kennedy, E. V. Houten, J. Weaver, and K. Paulsen, *Modeling of soft poroelastic tissue in time-harmonic MR elastography*, IEEE Transactions on Biomedical Engineering, 56 (2009), pp. 598–608.

[71] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, 11 2005.

[72] A. Rieder, *On the regularization of nonlinear ill-posed problems via inexact newton interactions*, (1999).

[73] L. Roininen, J. Huttunen, and S. Lasanen, *Whittle-matérn priors for bayesian statistical inversion with applications in electrical impedance tomography*, Inverse Problems & Imaging, 8 (2014), pp. 561–586.

[74] L. Roininen, P. Piiroinen, and M. Lehtinen, *Constructing continuous stationary covariances as limits of the second-order stochastic difference equations*, Inverse Problems &amp Imaging, 7 (2013), pp. 611–647.

[75] O. Rouvière, R. Souchon, G. Pagnoux, J. Ménager, and J. Chapelon, *Magnetic resonance elastography of the kidneys: feasibility and reproducibility in young healthy adults*, J. Magn. Reson. Imaging, 34 (2011), pp. 880–886.

[76] O. Rouvière, R. Souchon, G. Pagnoux, J. Ménager, and J. Chapelon, *Magnetic resonance elastography of the kidneys: Feasibility and reproducibility in young healthy adults*, Journal of Magnetic Resonance Imaging, 34 (2011), pp. 880–886.

[77] A. Sarvazyan, T. Hall, M. Urban, M. Fatemi, S. Aglyamov, and B. Garra, *An overview of elastography-an emerging branch of medical imaging*, Current Medical Imaging Reviews, 7 (2011), pp. 255–282.

[78] C. Schillings and A. Stuart, *Analysis of the ensemble kalman filter for inverse problems*, SIAM Journal on Numerical Analysis, 55 (2017), pp. 1264–1290.

[79] J. Seybold, *Introduction to RF Propagation*, John Wiley & Sons, Inc., 2005.

[80] R. Sinkus, M. Tanter, T. Xydeas, S. Catheline, J. Bercoff, and M. Fink, *Viscoelastic shear properties of in vivo breast lesions measured by MR elastography*, Magnetic Resonance Imaging, 23 (2005), pp. 159–165.

[81] L. Solamen, S. Gordon-Wylie, M. McGarry, J. Weaver, and K. Paulsen, *Phantom evaluations of low frequency MR elastography*, Physics in Medicine & Biology, 64 (2019).

[82] D. Stoyan, *Matérn, spatial variation*, Biometrical Journal, 30 (2007), pp. 594–594.

[83] A. Stuart, *Inverse problems: A bayesian perspective*, Acta Numerica, 19 (2010), pp. 451–559.

[84] S. Svensson, J. D. Arcos, O. Darwish, J. Fraser-Green, T. Storås, S. Holm, E. Vik-Mo, R. Sinkus, and K. Emblem, *Robustness of MR elastography in the healthy brain: Repeatability, reliability, and effect of different reconstruction methods*, Journal of Magnetic Resonance Imaging, (2021), pp. 1510–1521.

[85] A. Tarantola, *Inverse Problem Theory*, Elsevier, 1987.

[86] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, 2005.

[87] C. S. R. Todor, *Karhunen–loève approximation of random fields by generalized fast multipole methods*, Journal of Computational Physics, 217 (2006), pp. 100–122.

[88] E. Treister, *Shifted laplacian multigrid for the elastic helmholtz equation*, CoRR, abs/1806.11277 (2018).

[89] G. T. M. V. D. Velde, *Imaging techniques: Echocardiography, magnetic resonance imaging, and computerized tomography*, in Nadas' Pediatric Cardiology, Elsevier, 2006, pp. 183–211.

[90] S. Venkatesh, M. Yin, and R. Ehman, *Magnetic resonance elastography of liver: Technique, analysis, and clinical applications*, Journal of Magnetic Resonance Imaging, (2013), pp. 544–555.

[91] J. Veraart, D. Novikov, D. Christiaens, B. Ades-aron, J. Sijbers, and E. Fieremans, *Denoising of diffusion MRI using random matrix theory*, NeuroImage, 142 (2016), pp. 394–406.

[92] D. Wall and E. V. Houten, *On an inverse problem from magnetic resonance elastic imaging*, SIAM J. Applied Mathematics, (2011).

[93] S. Weissmann, N. Chada, C. Schillings, and X. Tong, *Adaptive tikhonov strategies for stochastic ensemble kalman inversion*, Inverse Problems, 38 (2022), p. 045009.

[94] N. Yamashita and M. Fukushima, *On the Rate of Convergence of the Levenberg-Marquardt Method*, Springer, 2001.

[95] W. Zeng, S. W. Gordon-Wylie, L. Tan, L. Solamen, M. D. J. McGarry, J. B. Weaver, and K. D. Paulsen, *Nonlinear inversion MR elastography with low-frequency actuation*, IEEE Transactions on Med. Imaging, 39 (2020), pp. 1775–1784.