Forward and Inverse Problems Under Uncertainty

Wenlong Zhang

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée dans le cadre d'une cotutelle entre
École normal Supérieure
et University of Chinese Academy of Sciences

Forward and Inverse Problems Under Uncertainty

Problèmes directets et inverses Sous incertitude

Ecole doctorale n°386

Sciences Mathématiques de Paris-Centre

Spécialité Mathématiques Appliquées

COMPOSITION DU JURY :

M. AMMARI Habib
ETH Zürich, Directeur de these

M. CHEN Zhiming
Chinese Academy of Science,
CoDirecteur de these

Mme. BONNAILLIE-NOEL Virginie
CNRS, École normale supérieure,
President

Mme. BORCEA Liliana
University of Michigan, Rapporteur

M. GARNIER Josselin
Ecole polytechnique, Rapporteur

Mme. LIM Mikyoung
Korea Advanced Institute of Science and Technology, Examinateur

M. POIGNARD Clair
Institut de Mathématiques de Bordeaux, Examinateur

Soutenue par Wenlong ZHANG
le 27 juin 2017

Dirigée par **Habib AMMARI**

Acknowledgements

In front of this thesis, I would like to give my warmest thanks to those who help me a lot during my PHD, including my two advisers, other collaborators and my family.

I will first thank Chen Zhiming in Beijing, a very nice and general professor, with whom I start my joint PHD for my first three years. During these three years in Chinese Academy of Sciences, he teaches and supports me a lot for my research and life, also, his team in Beijing gives a lot advises.

I would also like to thank my other adviser Habib Ammari in Paris. He leads me to the door of imaging research. Thanks so much that he offers the opportunity for me to finish my PHD in Paris for the last three years. He gives me many wonderful advises and helps a lot for my life in Paris. Also it's a great honour to work with his team.

During my PHD, I had the opportunity to work with several young collaborators: Tuo Rui, Jing Bangti, Thomas Widlak, Giovanni Alberti and Qiu Lingyun. I thank them for their sincere collaborations. We together discuss the problems we work on and I learn a lot from them.

Finally I would like to thank my family for their supports, I give my sincere thanks to my girl friend Hu Lin for her accompany and support.

Contents

Introduction

This thesis contains two different subjects. The first part introduces finite element methods to deal with the observational data with random noise arising from the thin plate spline problem and elliptic boundary problems. The goal is to investigate the stochastic convergence of the finite element method which characterizes the tail property of the probability distribution function of the finite element error. The second part introduces a mathematical and numerical framework for tissue property imaging from the cellular and the macroscopic scale. The mathematical models in this part help to understand the dependence of the conductivity of the tissue on the frequency and micro structure of the cells. A method for anisotropy imaging is also demonstrated in this Part. The imaging methods and reconstruction algorithms are developed for the corresponding models. Additional introductions for both subjects will be given respectively in this chapter.

0.1 Finite element methods with the observational data

In Part I, two cases are considered. One is the the thin plate spline smoother model and the other one is the elliptic boundary equations with uncertain boundary data. In this part, stochastic convergences of the finite element methods are proved for each problem.

The thin plate spline smoother is a classical model for finding a smooth function from the knowledge of its observation at scattered locations which may have random noises. In Chapter 1, a nonconforming Morley finite element method to approximate thin plate spline model is considered.

The spline model for scattered data has been extensively studied in the literature. In [27] it is proved that the model has a unique solution in $H^2(\mathbf{R}^d)$ under certian conditions. Explicit formula of the solution is constructed in [27] based on radial basis functions. In [111] the convergence rate for the expectation of the error is derived. Under the assumption that the additional random noises are also sub-Gaussion random variables, [32] proved the stochastic convergence of the error in terms of the empirical norm for $d = 1$. The stochastic convergence which provides additional tail information of the probability distribution function for the random error is very desirable for the approximation of random variables. We refer to [114] for a detailed analysis of the thin plate spline smoothers.

It is well-known that the numerical method based on radial basis functions to solve the thin plate spline smoother requires to solve a symmetric indefinite dense linear system of equations of the size $O(n)$, which is challenging for applications with very large data sets [92]. Conforming finite element methods for the solution of thin plate model are studied in [12] and the references therein. In [92] a mixed finite element method is proposed and the expectation of the finite element error is proved.

In this Chapter, we provide the optimal choice of smoothing parameter and propose a self-consistent iterative algorithm to determine the smoothing parameter based on our theoretical analysis. Numerical examples are presented to confirm the theoretical analysis and to show the competitive performance of the self-consistent algorithm for finding the smoothing parameter.

In Chapter 2, we propose a finite element method for solving elliptic equations with the observational Dirichlet boundary data which may subject to random noises. The method is based on the weak formulation of the Lagrangian multiplier. We show the convergence of the random finite element error in expectation and, when the noise is sub-Gaussian, in the Orlicz ψ_2-norm which implies the probability that the finite element error estimates are violated decays exponentially. Numerical examples are presented to support this result.

In many scientific and engineering applications involving partial differential equations, the input data such as sources or boundary conditions are usually given through the measurements which may be subject to random noises. A different perspective of solving partial differential equations with uncertain input data due to incomplete knowledge or inherent variability in the system has drawn considerable interests in recent years (see e.g. [101, 23, 40, 108] and the references therein). The goal of those studies is to learn about the uncertainties in system outputs of interest, given information about the uncertainties in the system inputs which are modeled as random fields. This goal usually leads to the mathematical problem of breaking the curse of dimensionality for solving partial differential equations having large number of parameters.

The classical problem of finding a smooth function from the knowledge of its observation at scattered locations subject to random noises is well studied in the literature [115]. One popular model to tackle this classical problem is to use the thin plate spline model [27, 111], which can be efficiently solved by using finite element methods [13, 92, 19]. But the method we propose in this Chapter is more efficient. One can combine the techniques developed in this chapter with the weak formulations given in [104] to deal with the observational Dirichlet boundary condition.

0.2 Two imaging methods and analysis of cell model for electropermeabilization

In Part II, we propose and analyze two imaging methods: the linearized model in multi-frequency electrical impedance tomography and the imaging of anisotropy conductivity using Diffusion Tensor. At the end of this Part, we analyze the well-posedness of the cell model for electropermeabilization and propose a dynamical homogenization scheme.

In Chapter 3, we provide a mathematical analysis of the linearized inverse problem in multifrequency electrical impedance tomography. We consider the isotropic conductivity distribution with a finite number of unknown inclusions with different frequency dependence.

Electrical impedance tomography (EIT) is a diffusive imaging modality in which the conductivity distribution of the concerned object is recovered from the electrode voltage measurements on the boundary, induced

by (multiple) known injected currents. The modality is safe, cheap and portable, and has the potential to be an established clinical imaging method in a multitude of applications [47]. However, the EIT inverse problem is severely ill-posed. An approach to improve this problem is multi-frequency electrical impedance tomography (mfEIT), also known as EIT spectroscopy, which is also attracting attention in recent years. There have been several studies on frequency-difference imaging, see [38, 94, 119, 96, 76, 60, 75].

Our main contributions are as follows. First, we systematically discuss mfEIT reconstruction in the following three different scenarios, i.e., known spectral profiles, partially known spectral profiles and unknown spectral profiles. This analysis generalizes the existing studies, especially [96]. Second, we provide a rigorous justification of mfEIT for handling geometrical errors. Third, we present a novel group sparse reconstruction algorithm of iterative shrinkage type, which is easy to implement and converges fast. The extensive numerical experiments fully confirm our discussions.

In Chapter 4, we present a mathematical and numerical framework for a procedure of imaging anisotropic electrical conductivity tensor using a novel technique called Diffusion Tensor Magneto-acoustography.

MAT-MI is a new noninvasive modality for imaging electrical conductivity distributions of biological tissue [118, 72, 77, 122, 71, 117, 74]. In the experiments, the biological tissue is placed in a static magnetic field. A pulsed magnetic field is applied to induce an eddy current inside the conductive tissue. Diffusion Tensor Imaging (DTI) is a non-invasive technique for characterizing the diffusion properties of water molecules in tissues (see e.g. [14] and the references therein). Imaging conductivity tensors in the tissue with DTI is based on the correlation property between diffusion and conductivity tensors [110]. This linear relationship can be used to characterize the conductivity tensor.

Once the conductivity directions of anisotropy are determined, one needs only to reconstruct a cross-property factor which is a scalar function. However, up to now, all techniques have assumed an isotropic conductivity distribution in the image reconstruction problem to simplify the underlying mathematical theory [5, 90]. In this Chapter, we firstly formulate a new image reconstruction method of an anisotropic conductivity tensor distribution by combining the MAT-MI and DTI techniques. We propose an optimal control approach for reconstructing the cross-property factor relating the diffusion tensor to the anisotropic electrical conductivity tensor. We prove convergence and Lipschitz type stability of the algorithm and present numerical examples to illustrate its accuracy.

In Chapter 5, the cell model for Electropermeabilization is demonstrated. The technique of electropermeabilization is employed to make the chemotherapeutical treatment of cancer more efficient and avoid side-effects. Instead of spreading out drugs over the whole body, electropermeabilization makes it possible to focus drug application on special areas. The mechanism of electropermeabilization relies on careful exposition of biological tissue to electrical fields: this changes the membrane properties of the cells such that treatment can enter more easily just at precisely defined areas of the tissue [54, 81].

For treatment planning in electropermeabilization, one is interested in the percentage of electroporated cells over the whole tissue to make decisions in the short term how to gear treatment [64, 26, 81]. We tackle the

next step in electropermeabilization monitoring and investigate the question of determining microscopic parameters from macroscopic measurements. The modelling used stems from general physiological tissue models for cells, asymptotically simplified by Neu and Krassowska [83]. Whereas the mathematical well-posedness of the model of that model is not available in the literature, there exists an investigation of well-posedness for a similar model in [58].

In order to describe the relation between macroscopic and microscopic quantities, we apply the homogenization scheme in [4] to the cell model of Neu and Krassowska [83]. This not only describes isotropic effective parameters such as classical theory [85], but includes also anisotropy.

In this Chapter, we study effective parameters in a homogenization model as the next step to monitor the microscopic properties in clinical practice. We start from a physiological cell model for electropermeabilization and analyze its well-posedness. For a dynamical homogenization scheme, we prove convergence and then analyze the effective parameters, which can be found by macroscopic imaging methods. We demonstrate numerically the sensitivity of these effective parameters to critical microscopic parameters governing electropermeabilization. This opens the door to solve the inverse problem of reconstructing these parameters.

The results of Chapters 1 and 2 are from [19] and [18], respectively. The results of Chapter 3, 4, and 5 are from [11], [8] and [6], respectively.

List of Figures

List of Tables

Part I

Finite element methods with the observational data

Chapter 1

FEM for the thin plate spline model

1.1 Introduction

In this chapter, we show the convergence analysis of finite element method for the thin plate model and the method to choose the optimal parameter.

The thin plate spline smoother is a classical mathematical model for finding a smooth function from the knowledge of its observation at scattered locations which may be the subject to random noises. Let Ω be a bounded Lipschitz domain in $\mathbf{R}^d$ ($d \leq 3$) and $u_0 \in H^2(\Omega)$ be the unknown smooth function. Let $\{x_i\}_{i=1}^n \subset \Omega$ be the scattered locations in the domain where the observation is taken. We want to approximate u_0 from the noisy data $y_i = u_0(x_i) + e_i,\ 1 \leq i \leq n$, where $\{e_i\}_{i=1}^n$ are independent and identically distributed random variables on some probability space $(\mathfrak{X}, \mathcal{F}, \mathbb{P})$ satisfying $\mathbb{E}[e_i] = 0$ and $\mathbb{E}[e_i^2] \leq \sigma^2$. Here and in the sequel $\mathbb{E}[X]$ denotes the expectation of the random variable X. The thin plate spline smoother, i.e., D^2-spline smoother to approximate u_0, is defined to be the unique solution of the following variational problem

$$\min_{u \in H^2(\Omega)} \frac{1}{n} \sum_{i=1}^{n} (u(x_i) - y_i)^2 + \lambda_n |u|^2_{H^2(\Omega)}, \tag{1.1}$$

where $\lambda_n > 0$ is the smoothing parameter.

The spline model for scattered data has been extensively studied in the literature. For $\Omega = \mathbf{R}^d$, it is proved in [27] that (1.1) has a unique solution $u_n \in H^2(\mathbf{R}^d)$ when the set $\mathbb{T} = \{x_i : i = 1, 2, \cdots, n\}$ is not collinear (i.e. the points in $\mathbb{T}$ are not on the same plane). An explicit formula of the solution is constructed in [27] based on radial basis functions. [111] derived the convergence rate for the expectation of the error $|u_n - u_0|^2_{H^j(\Omega)}$, $j = 0, 1, 2$. Under the assumption that e_i, $i = 1, 2, \cdots, n$, are also sub-Gaussion random variables, [32] proved the stochastic convergence of the error in terms of the empirical norm $\|u_n - u_0\|_n := (n^{-1} \sum_{i=1}^n |u_n(x_i) - u_0(x_i)|^2)^{1/2}$ when $d = 1$. The stochastic convergence which provides additional tail information of the probability distribution function for the random error is very desirable for the approximation of random variables. We refer to [114] for further information on the thin plate spline smoothers.

It is well-known that the numerical method based on radial basis functions for solving the thin plate spline smoother requires to solve a symmetric indefinite dense linear system of equations of the size $O(n)$, which is challenging for applications with very large data sets [92]. Conforming

finite element methods for the solution of thin plate model are studied in [12] and the references therein. In [92] a mixed finite element method for solving ∇u_n is proposed and the expectation of the finite element error is proved. The advantage of the mixed finite element method in [92] lies in that one can use simple $H^1(\Omega)$-conforming finite element spaces. The H^1 smoother in [92] that the mixed finite element method aims to approximate is not equivalent to the thin plate spline model (1.1).

In this chapter we consider the nonconforming finite element approximation to the problem (1.1). We use the Morley element [82, 99, 116] which is of particular interest for solving fourth order PDEs since it has the least number of degrees of freedom on each element. The difficulty of the finite element analysis for the thin plate smoother is the low stochastic regularity of the solution u_n. One can only prove the boundedness of $\mathbb{E}[|u_n|^2_{H^2(\Omega)}]$ (see Theorem 1.2.1 below). This difficulty is overcome by a smoothing operator based on the C^1-element for any Morley finite element functions. We also prove that the probability distribution function of the empirical norm of the finite element error has an exponentially decaying tail. For that purpose we also prove the convergence of the error $\|u_n - u_0\|_n$ in terms of the Orlicz ψ_2 norm (see Theorem 1.4.1 below) which improves the result in [32].

One of the central issues in the application of the thin plate model is the choice of the smoothing parameter λ_n. In the literature it is usually made by the method of cross validation [114]. The analysis in this chapter suggests that the optimal choice should be

$$\lambda_n^{1/2+d/8} = O(\sigma n^{-1/2}(|u_0|_{H^2(\Omega)} + \sigma n^{-1/2})^{-1}). \tag{1.2}$$

Since, in practical applications, one does not know u_0 and the upper bound of the variance σ, we propose a self-consistent algorithm to determine λ_n from the natural initial guess $\lambda_n = n^{-\frac{4}{4+d}}$. Our numerical experiments show this self-consistent algorithm performs rather well.

The layout of the chapter is as follows. In Section 1.2 we recall some preliminary properties of the thin plate model. In Section 1.3 we introduce the nonconforming finite element method and show the convergence of the finite element solution in terms of the expectation of Sobolev norms. In Section 1.4 we study the tail property of the probability distribution function for the finite element error based on the theory of empirical process for sub-Gaussion noises. In Section 1.5 we introduce our self-consistent algorithm for finding the smooth parameter λ_n and show several numerical examples to support the analysis in this chapter.

1.2 The thin plate model

In this section we collect some preliminary results about the thin plate smoother (1.1). In this chapter, we will always assume that Ω is a bounded Lipschitz domain satisfying the uniform cone condition. We will also assume that $\mathbb{T}$ are uniformly distributed in the sense that [111] there exists a constant $B > 0$ such that $\frac{h_{\max}}{h_{\min}} \leq B$, where

$$h_{\max} = \sup_{x\in\Omega} \inf_{1\leq i\leq n} |x - x_i|, \quad h_{\min} = \inf_{1\leq i\neq j\leq n} |x_i - x_j|.$$

It is easy to see that there exist constants B_1, B_2 such that $B_1 n^{-1/d} \le h_{\max} \le Bh_{\min} \le B_2 n^{-1/d}$.

We write the empirical inner product between the data and any function $v \in C(\bar{\Omega})$ as $(y, v)_n = \frac{1}{n}\sum_{i=1}^n y_i v(x_i)$. We also write $(u, v)_n = \frac{1}{n}\sum_{i=1}^n u(x_i)v(x_i)$ for any $u, v \in C(\bar{\Omega})$ and the empirical norm $\|u\|_n = (\frac{1}{n}\sum_{i=1}^n u^2(x_i))^{1/2}$ for any $u \in C(\bar{\Omega})$. By [111, Theorems 3.3-3.4], there exists a constant $C > 0$ depending only on Ω, B such that for any $u \in H^2(\Omega)$ and sufficiently small $h_{\max}$,

$$\|u\|_{L^2(\Omega)} \le C(\|u\|_n + h_{\max}^2 |u|_{H^2(\Omega)}), \quad \|u\|_n \le C(\|u\|_{L^2(\Omega)} + h_{\max}^2 |u|_{H^2(\Omega)}). \tag{1.3}$$

It follows from (1.3) and Lax-Milgram lemma that the minimization problem (1.1) has a unique solution $u_n \in H^2(\Omega)$. The following convergence result is proved in [111].

Lemma 1.2.1. *Let $U_n \in H^2(\mathbf{R}^d)$ be the solution of the following variational problem:*

$$\min_{u \in D^{-2}L^2(\mathbf{R}^d)} \|u - y\|_n^2 + \lambda_n |u|^2_{H^2(\mathbf{R}^d)}, \tag{1.4}$$

where $D^{-2}L^2(\mathbf{R}^d) = \{u | D^\alpha u \in L^2(\mathbf{R}^d),\ |\alpha| = 2\}$. Then there exist constants $\lambda_0 > 0$ and $C > 0$ such that for any $\lambda_n \le \lambda_0$ and $n\lambda_n^{d/4} \ge 1$,

$$\mathbb{E}\big[\|U_n - u_0\|_n^2\big] \le C\lambda_n |u_0|^2_{H^2(\Omega)} + \frac{C\sigma^2}{n\lambda_n^{d/4}}, \tag{1.5}$$

$$\mathbb{E}\big[|U_n|^2_{H^2(\Omega)}\big] \le C|u_0|^2_{H^2(\Omega)} + \frac{C\sigma^2}{n\lambda_n^{1+d/4}}. \tag{1.6}$$

Define the bilinear form $a : H^2(\Omega) \times H^2(\Omega) \to \mathbf{R}$ as

$$a_\Omega(u, v) = \sum_{1 \le i,j \le d} \int_\Omega \frac{\partial^2 u}{\partial x_i \partial x_j} \frac{\partial^2 v}{\partial x_i \partial x_j} dx, \quad \forall u, v \in H^2(\Omega). \tag{1.7}$$

It is obvious that $|u|^2_{H^2(\Omega)} = a(u, u)$ for any $u \in H^2(\Omega)$.

Theorem 1.2.1. *Let $u_n \in H^2(\Omega)$ be the unique solution of (1.1). Then there exist constants $\lambda_0 > 0$ and $C > 0$ such that for any $\lambda_n \le \lambda_0$ and $n\lambda_n^{d/4} \ge 1$,*

$$\mathbb{E}\big[\|u_n - u_0\|_n^2\big] \le C\lambda_n |u_0|^2_{H^2(\Omega)} + \frac{C\sigma^2}{n\lambda_n^{d/4}}, \tag{1.8}$$

$$\mathbb{E}\big[|u_n|^2_{H^2(\Omega)}\big] \le C|u_0|^2_{H^2(\Omega)} + \frac{C\sigma^2}{n\lambda_n^{1+d/4}}. \tag{1.9}$$

Proof. It is clear that $u_n \in H^2(\Omega)$ and $U_n \in H^2(\mathbf{R}^d)$ satisfy the following variational forms:

$$\lambda_n a_\Omega(u_n, v) + (u_n, v)_n = (y, v)_n,\ \forall v \in H^2(\Omega), \tag{1.10}$$

$$\lambda_n a_{\mathbf{R}^d}(U_n, w) + (U_n, w)_n = (y, w)_n,\ \forall w \in H^2(\mathbf{R}^d). \tag{1.11}$$

Let $F : H^2(\Omega) \to D^{-2}L^2(\mathbf{R}^d)$ be the extension operator defined by

$$Fu = \operatorname*{argmin}_{v \in D^{-2}L^2(\mathbf{R}^d), v|_\Omega = u} |v|_{H^2(\Omega)}.$$

It is known [27, 111] that $Fu = u$ in Ω and $|Fu|_{H^2(\mathbf{R}^d)} \le C|u|_{H^2(\Omega)}$ for some constant $C > 0$. We write $\tilde{u} = Fu$ in $\mathbf{R}^d$ in the following. Thus, it follows from (1.10)-(1.11) that

$$\lambda_n a_\Omega(u_n - U_n, v) + (u_n - U_n, v)_n = \lambda_n a_{\mathbf{R}^d\setminus\bar{\Omega}}(U_n, \tilde{v}), \quad \forall v \in H^2(\Omega),$$

which implies by taking $v = u_n - U_n|_\Omega \in H^2(\Omega)$ that

$$\begin{aligned} \lambda_n |u_n - U_n|^2_{H^2(\Omega)} + \|u_n - U_n\|^2_n &\le \lambda_n |U_n|_{H^2(\mathbf{R}^d)} |\tilde{u}_n - \tilde{U}_n|_{H^2(\mathbf{R}^d)} \\ &\le C\lambda_n |U_n|_{H^2(\mathbf{R}^d)} |u_n - U_n|_{H^2(\Omega)}, \end{aligned}$$

where $\tilde{U}_n = F(U_n|_\Omega)$. Therefore

$$|u_n - U_n|^2_{H^2(\Omega)} \le C|U_n|^2_{H^2(\mathbf{R}^d)}, \quad \|u_n - U_n\|^2_n \le \lambda_n |U_n|^2_{H^2(\mathbf{R}^d)}. \tag{1.12}$$

Since U_n is the solution of (1.4) and $\tilde{U}_n = U_n$ in Ω, we have $|U_n|_{H^2(\mathbf{R}^d)} \le |\tilde{U}_n|_{H^2(\mathbf{R}^d)} \le C|U_n|_{H^2(\Omega)}$. Therefore, $\mathbb{E}[|u_n|^2_{H^2(\Omega)}] \le C\mathbb{E}[|U_n|^2_{H^2(\Omega)}]$, which implies (1.9) by using (1.6). Similarly one obtains (1.8) from the second estimate in (1.12) and (1.5)-(1.6). This completes the proof. □

Theorems 1.2.1 and 1.2.1 suggest that an optimal choice of the parameter λ_n is such that $\lambda_n^{1+d/4} = O((\sigma^2 n^{-1})|u_0|^{-2}_{H^2(\Omega)})$.

1.3 Nonconforming finite element method

In this section we consider the nonconforming finite element approximation to the thin plate model (1.1) whose solution $u_n \in H^2(\Omega)$ satisfies the following weak formulation

$$\lambda_n a_\Omega(u_n, v) + (u_n, v)_n = (y, v)_n, \ \forall v \in H^2(\Omega). \tag{1.13}$$

We assume that Ω is a polygonal or polyhedral domain in $\mathbf{R}^d$ $(d = 2, 3)$ in the reminder of this chapter. Let $\mathcal{M}_h$ be a family of shape regular and quasi-uniform finite element meshes over the domain Ω. We will use the Morley element [82] for 2D, [116] for 3D to define our nonconforming finite element method. The Morley element is a triple (K, P_K, Σ_K), where $K \in \mathcal{M}_h$ is a simplex in $\mathbf{R}^d$, $P_K = P_2(K)$ is the set of second order polynomials in K, and Σ_K is the set of the degrees of freedom. In 2D, for the element K with vertices $a_i, 1 \le i \le 3$, and mid-points b_i of the edge opposite to the vertex a_i, $1 \le i \le 3$, $\Sigma_K = \{p(a_i), \partial_\nu p(b_i), 1 \le i \le 3, \forall p \in C^1(K)\}$. In 3D, for the element K with edges S_{ij} which connects the vertices a_i, a_j, $1 \le i < j \le 4$, and faces F_j opposite to a_j, $1 \le j \le 4$, $\Sigma_K = \{\frac{1}{|S_{ij}|}\int_{S_{ij}} p, 1 \le i < j \le 4, \frac{1}{|F_j|}\int_{F_j} \partial_\nu p, 1 \le j \le 4, \forall p \in C^1(K)\}$. Here $\partial_\nu p$ is the normal derivative of p of the edges (2D) or faces (3D) of the element. We refer to Figure 1.1 for the illustration of the degrees of freedom of the Morley element.

Let V_h be the Morley finite element space

$$V_h = \{v_h : v_h|_K \in P_2(K), \forall K \in \mathcal{M}_h, f(v_h|_{K_1}) = f(v_h|_{K_2}), \forall f \in \Sigma_{K_1} \cap \Sigma_{K_2}\}.$$

The functions in V_h may not be continuous in Ω. Given a set $G \subset \mathbf{R}^2$, let $\mathcal{M}_h(G) = \{K \in \mathcal{M}_h : G \cap K \ne \emptyset\}$ and $N(G)$ the number of elements in

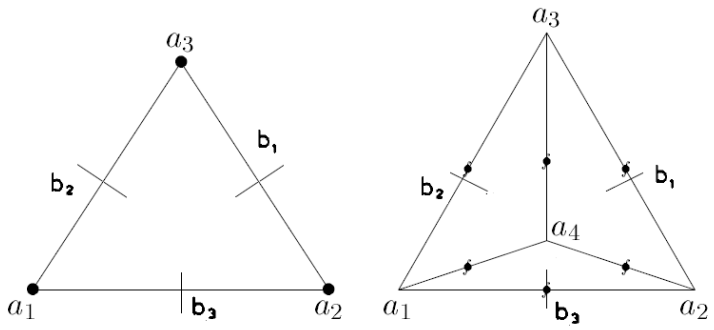


FIGURE 1.1: The degrees of freedom of 2D Morley (left) and 3D Morley (right) element.

$\mathcal{M}_h(G)$. For any $v_h \in V_h$, we define

$$\hat{v}_h(x_i) = \frac{1}{N(x_i)} \sum_{K' \in \mathcal{M}_h(x_i)} (v_h|_{K'})(x_i), \quad i = 1, 2, \cdots, n. \tag{1.14}$$

Notice that if x_i is located inside some element K, then $\mathcal{M}_h(x_i) = \{K\}$ and $\hat{v}_h(x_i) = v_h(x_i)$, $i = 1, 2, \cdots, n$. With this definition we know that $(\hat{v}_h, \hat{w}_h)_n$ and $(e, \hat{w}_h)_n$ are well-defined for any $v_h, w_h \in V_h$.

Let

$$a_h(u_h, v_h) = \sum_{K \in \mathcal{M}_h} \sum_{1 \le i,j \le d} \int_K \frac{\partial^2 u_h}{\partial x_i \partial x_j} \frac{\partial^2 v_h}{\partial x_i \partial x_j} dx, \quad \forall u_h, v_h \in V_h.$$

The finite element approximation of the problem (1.13) is to find $u_h \in V_h$ such that

$$\lambda_n a_h(u_h, v_h) + (\hat{u}_h, \hat{v}_h)_n = (y, \hat{v}_h)_n, \quad \forall v_h \in V_h. \tag{1.15}$$

Since the sampling point set $\mathbb{T}$ is not collinear, by Lax-Milgram lemma, problem (1.15) has a unique solution.

Let $I_K : H^2(K) \to P_2(K)$ be the canonical local nodal value interpolant of Morley element [99, 116] and $I_h : L^2(\Omega) \to V_h$ be the global nodal value interpolant such that $(I_h u)|_K = I_K u$ for any $K \in \mathcal{M}_h$ and piecewise $H^2(K)$ functions $u \in L^2(\Omega)$. We introduce the mesh dependent semi-norm $|\cdot|_{m,h}$, $m \ge 0$,

$$|v|_{m,h} = \left(\sum_{K \in \mathcal{M}_h} |v|^2_{H^m(K)} \right)^{1/2},$$

for any $v \in L^2(\Omega)$ such that $v|_K \in H^m(K), \forall K \in \mathcal{M}_h$.

Lemma 1.3.1. *We have*

$$|u - I_K u|_{H^m(K)} \le C h_K^{2-m} |u|_{H^2(K)}, \quad \forall u \in H^m(K), 0 \le m \le 2, \tag{1.16}$$

$$\|u - \widehat{I_h u}\|_n \le C h^2 |u|_{H^2(\Omega)}, \quad \forall u \in H^2(\Omega), \tag{1.17}$$

where h_K is the diameter of the element K and $h = \max_{K \in \mathcal{M}_h} h_K$.

Proof. Since $I_K p = p$ for any $p \in P_2(K)$ [116], estimate (1.16) follows from the standard interpolation theory for finite element method [21]. Moreover, we have, by local inverse estimates and the standard interpolation estimates

$$\begin{aligned}\|u - I_K u\|_{L^\infty(K)} &\leq \inf_{p\in P_2(K)} \Big[\|u-p\|_{L^\infty(K)} + |K|^{-1/2}\|I_K(u-p)\|_{L^2(K)}\Big] \\ &\leq Ch_K^{2-d/2}|u|_{H^2(K)}.\end{aligned}$$

Let $\mathbb{T}_K = \{x_i \in \mathbb{T} : x_i \in K, 1 \leq i \leq n\}$. By assumption, $\mathbb{T}$ is uniformly distributed and the mesh is quasi-uniform, and hence the cardinal $\#\mathbb{T}_K \leq Cnh^d$. Thus

$$\|u - \widehat{I_h u}\|_n^2 \leq \frac{1}{n}\sum_{K\in\mathcal{M}_h} \#\mathbb{T}_K \|u - I_K u\|_{L^\infty(K)}^2 \leq Ch^4 |u|_{H^2(\Omega)}^2.$$

This proves (1.17). □

The following property of Morley element will be used below.

Lemma 1.3.2. *Let $K, K' \in \mathcal{M}_h$ and $F = K \cap K'$. There exists a constant C independent of h such that for any $v_h \in V_h$, $|\alpha| \leq 2$,*

$$\|\partial^\alpha(v_h|_K - v_h|_{K'})\|_{L^\infty(F)} \leq Ch^{2-|\alpha|-d/2}(|v_h|_{H^2(K)} + |v_h|_{H^2(K')}).$$

Proof. By [116, Lemma 5] we know that

$$\|v_h|_K - v_h|_{K'}\|_{L^2(F)} \leq Ch^{3/2}(|v_h|_{H^2(K)} + |v_h|_{H^2(K')}).$$

By using the inverse estimate we then obtain

$$\begin{aligned}\|\partial^\alpha(v_h|_K - v_h|_{K'})\|_{L^\infty(F)} &\leq Ch^{-|\alpha|}\|v_h|_K - v_h|_{K'}\|_{L^\infty(F)} \\ &\leq Ch^{-|\alpha|-(d-1)/2}\|v_h|_K - v_h|_{K'}\|_{L^2(F)} \\ &\leq Ch^{2-|\alpha|-d/2}(|v_h|_{H^2(K)} + |v_h|_{H^2(K')}).\end{aligned}$$

This proves the lemma. □

Lemma 1.3.3. *There exists a linear operator $\Pi_h : V_h \to H^2(\Omega)$ such that for any $v_h \in V_h$,*

$$|v_h - \Pi_h v_h|_{m,h} \leq Ch^{2-m}|v_h|_{2,h}, \quad m = 0, 1, 2, \tag{1.18}$$

$$\|\hat{v}_h - \Pi_h v_h\|_n \leq Ch^2 |v_h|_{2,h}, \tag{1.19}$$

where the constant C is independent of h.

Proof. We will only prove the lemma for the case $d = 2$. The case of $d = 3$ will be briefly discussed in the Appendix A. We will construct $\Pi_h v_h$ by using the Agyris element. We recall [21, P.71] that for any $K \in \mathcal{M}_h$, Agyris element is a triple (K, P_K, Λ_K), where $P_K = P_5(K)$ and the set of degrees of freedom, with the notation in Figure 1.2, $\Lambda_K = \{p(a_i), Dp(a_i)(a_j - a_i), D^2p(a_i)(a_j - a_i, a_k - a_i), \partial_\nu p(b_i), 1 \leq i, j, k \leq 3, j \neq i, k \neq i, \forall p \in C^2(K)\}$. Let X_h be the Agyris finite element space

$$X_h = \{v_h : v_h|_K \in P_5(K), \forall K \in \mathcal{M}_h, f(v_h|_{K_1}) = f(v_h|_{K_2}), \forall f \in \Lambda_{K_1} \cap \Lambda_{K_2}\}.$$

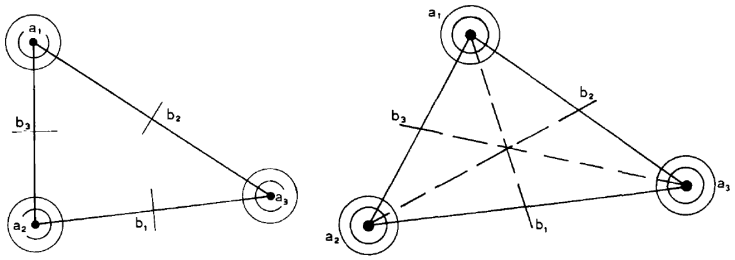


FIGURE 1.2: The degrees of freedom of Agyris element (left) and Hermite triangle of type (5) (right).

It is known that $X_h \subset H^2(\Omega)$.

We define the operator Π_h as follows. For any $v_h \in V_h$, $w_h := \Pi_h v_h \in X_h$ such that for any $K \in \mathcal{M}_h$, $w_h|_K \in P_5(K)$ and

$$\partial^\alpha(w_h|_K)(a_i) = \frac{1}{N(a_i)} \sum_{K' \in \mathcal{M}_h(a_i)} \partial^\alpha(v_h|_{K'})(a_i), \quad 1 \le i \le 3, \quad |\alpha| \le 2, \tag{1.20}$$

$$\partial_\nu(w_h|_K)(b_i) = \partial_\nu(v_h|_K)(b_i), \quad 1 \le i \le 3. \tag{1.21}$$

Here $\mathcal{M}_h(a_i)$ and $N(a_i)$ are defined above (1.14). To show estimate (1.18) we follow an idea in [21, Theorem 6.1.1] and use the element Hermite triangle of type (5) [21, P.102], which is a triple (K, P_K, Θ_K), where $P_K = P_5(K)$ and the set of degrees of freedom $\Theta_K = \{p(a_i), Dp(a_i)(a_j - a_i), D^2p(a_i)(a_j - a_i, a_k - a_i), Dp(b_i)(a_i - b_i), 1 \le i, j, k \le 3, j \ne i, k \ne i, \forall p \in C^2(K)\}$. The finite element space of Hermite triangle of type (5) is H^1 conforming and a regular family of Hermite triangle of type (5) is affine-equivalent. For any $K \in \mathcal{M}_h$, denote by $p_i, p_{ij}, p_{ijk}, q_i$ the basis functions associated with the degrees of freedom $p(a_i), Dp(a_i)(a_j - a_i), D^2p(a_i)(a_j - a_i, a_k - a_i), Dp(b_i)(a_i - b_i)$, $1 \le i, j, k \le 3, j \ne i, k \ne i$.

For any $v_h \in V_h$, we also define a linear operator $q_h := \Lambda_h v_h$ as follows: for any $K \in \mathcal{M}_h$, $q_h|_K \in P_5(K)$ and

$$\partial^\alpha(q_h|_K)(a_i) = \frac{1}{N(a_i)} \sum_{K' \in \mathcal{M}_h(a_i)} \partial^\alpha(v_h|_{K'})(a_i), \quad 1 \le i \le 3, \quad |\alpha| \le 2, \tag{1.22}$$

$$D(q_h|_K)(b_i)(a_i - b_i) = D(v_h|_K)(b_i)(a_i - b_i), \quad 1 \le i \le 3. \tag{1.23}$$

Then from the definition of Morley element and Hermite triangle of type (5), we know that $\phi_h|_K := (v_h - q_h)|_K \in P_5(K)$ satisfies

$$\begin{aligned} \phi_h(x) &= \sum_{i,j=1,2,3, j\ne i} D(\phi_h|_K)(a_i)(a_j - a_i) p_{ij}(x) \\ &+ \sum_{i,j,k=1,2,3, j\ne i, k\ne i} D^2(\phi_h|_K)(a_i)(a_j - a_i, a_k - a_i) p_{ijk}(x). \end{aligned}$$

Since a regular family of Hermite triangle of type (5) is affine-equivalent, by standard scaling argument [21, Theorem 3.1.2], we obtain easily $|q_i|_{H^m(K)} + |p_i|_{H^m(K)} + |p_{ij}|_{H^m(K)} + |p_{ijk}|_{H^m(K)} \le C h_K^{1-m}$, $m = 0, 1, 2$. Thus, for $m =$

$0, 1, 2$,

$$|\phi_h|_{H^m(K)} \leq C h_K^{1-m} \left(\sum_{i=1}^{3} \sum_{1 \leq |\alpha| \leq 2} h^{|\alpha|} |\partial^\alpha (v_h|_K)(a_i) - \partial^\alpha (q_h|_K)(a_i)|^2 \right)^{1/2}. \tag{1.24}$$

By Lemma 1.3.2 and the fact that $\partial^\alpha (q_h|_K)(a_i)$ is the local average of $\partial^\alpha v_h$ over elements around a_i in (1.22)

$$|\partial^\alpha (v_h|_K)(a_i) - \partial^\alpha (q_h|_K)(a_i)| \leq C h^{1-|\alpha|} \left(\sum_{K' \in \mathcal{M}_h(a_i)} |v_h|^2_{H^2(K')} \right)^{1/2}, \quad \forall 1 \leq |\alpha| \leq 2.$$

Inserting the above estimate into (1.24), we get

$$|v_h - q_h|_{H^m(K)} \leq C h^{2-m} \left(\sum_{K' \in \mathcal{M}_h(K)} |v_h|^2_{H^2(K')} \right)^{1/2}, \ m = 0, 1, 2. \tag{1.25}$$

By (1.20)-(1.23) we know that $q_h - w_h \in P_5(K)$ and satisfies

$$q_h(x) - w_h(x) \ = \ \sum_{i=1}^{3} D(q_h|_K - w_h|_K)(b_i)(a_i - b_i) q_i(x).$$

On the other hand, for $1 \leq i \leq 3$,

$$D(q_h|_K - w_h|_K)(b_i)(a_i - b_i) = \partial_\nu (q_h|_K - v_h|_K)(b_i)[(a_i - b_i) \cdot \nu],$$

since $\partial_\nu (w_h|_K)(b_i) = \partial_\nu (v_h|_K)(b_i)$ by (1.21) and the tangential derivative of $(q_h|_K - w_h|_K)$ vanishes as a consequence of (1.20) and (1.22). Since $|q_i|_{H^m(K)} \leq C h_K^{1-m}$ for $m = 0, 1, 2$, we then obtain that

$$\begin{aligned} |q_h - w_h|_{H^m(K)} &\leq C h^{2-m} \left(\sum_{i=1}^{3} |\partial_\nu (q_h|_K - v_h|_K)(b_i)|^2 \right)^{1/2} \\ &\leq C h^{2-m} \left(\sum_{K' \in \mathcal{M}_h(K)} |v_h|^2_{H^2(K')} \right)^{1/2}, \ m = 0, 1, 2, \end{aligned} \tag{1.26}$$

where in the second inequality we have used the fact that by the inverse estimate and (1.25),

$$\begin{aligned} |\partial_\nu (q_h|_K - v_h|_K)(b_i)| \leq |q_h - v_h|_{W^{1,\infty}(K)} &\leq C h_K^{-1} |q_h - v_h|_{H^1(K)} \\ &\leq C \left(\sum_{K' \in \mathcal{M}_h(K)} |v_h|^2_{H^2(K')} \right)^{1/2}. \end{aligned}$$

Combining (1.25) and (1.26) shows (1.18).

To show (1.19), we use the notation in the proof of Lemma 1.3.1, the inverse estimate and (1.18) to get

$$\|\hat{v}_h - w_h\|_n^2 \le \frac{C}{n} \sum_{K\in\mathcal{M}_h} \#\mathbb{T}_K \|v_h - w_h\|_{L^\infty(K)}^2 \le C\|v_h - w_h\|_{L^2(\Omega)}^2 \le Ch^4 |v_h|_{2,h}^2.$$

This completes the proof. □

For any function v which is piecewise in $C^2(K)$ for any $K \in \mathcal{M}_h$, we use the convenient energy norm

$$|||v|||_h = \left(\lambda_n |v|_{2,h}^2 + \|\hat{v}\|_n^2\right)^{1/2}.$$

Here $\hat{v}(x_i)$, $i = 1, 2, \cdots, n$, is defined as in (1.14), that is, $\hat{v}(x_i)$ is the local average of all $v|_{K'}(x_i)$, where $K' \in \mathcal{M}_h$ such that $x_i \in K'$.

Theorem 1.3.1. *Let $u_n \in H^2(\Omega)$ be the unique solution of* (1.13) *and $u_h \in V_h$ be the solution of* (1.15). *Then there exist constants $\lambda_0 > 0$ and $C > 0$ such that for any $\lambda_n \le \lambda_0$ and $n\lambda_n^{d/4} \ge 1$,*

$$\mathbb{E}\left[\|u_0 - \hat{u}_h\|_n^2\right] \le C(\lambda_n + h^4)|u_0|_{H^2(\Omega)}^2 + C\left[1 + \frac{h^4}{\lambda_n} + \left(\frac{h^4}{\lambda_n}\right)^{1-d/4}\right] \frac{\sigma^2}{n\lambda_n^{d/4}}. \tag{1.27}$$

In particular, if $h^4 \le C\lambda_n$, we have

$$\mathbb{E}\left[\|u_0 - \hat{u}_h\|_n^2\right] \le C\lambda_n |u_0|_{H^2(\Omega)}^2 + \frac{C\sigma^2}{n\lambda_n^{d/4}}. \tag{1.28}$$

Proof. We start by using the Strang lemma [21]

$$|||u_n - \hat{u}_h|||_h \le C \inf_{v_h\in V_h} |||u_n - \hat{v}_h|||_h + C \sup_{0\neq v_h\in V_h} \frac{|\lambda_n a_h(u_n, v_h) + (u_n - y, \hat{v}_h)_n|}{|||v_h|||_h}. \tag{1.29}$$

By Lemma 1.3.1 we have

$$\inf_{v_h\in V_h} |||u_n - \hat{v}_h|||_h \le C(\lambda_n^{1/2} + h^2)|u_n|_{H^2(\Omega)}. \tag{1.30}$$

Since for any $v_h \in V_h$, $\Pi_h v_h \in H^2(\Omega)$, by (1.13) and the fact that $y_i = u_0(x_i) + e_i$, $i = 1, 2, \cdots n$, we obtain

$$\begin{aligned} & \lambda_n a_h(u_n, v_h) + (u_n - y, \hat{v}_h)_n \\ = \; & \lambda_n a_h(u_n, v_h - \Pi_h v_h) + (u_n - y, \hat{v}_h - \Pi_h v_h)_n \\ \le \; & \lambda_n |u_n|_{H^2(\Omega)} |v_h - \Pi_h v_h|_{2,h} + \|u_n - u_0\|_n \|\hat{v}_h - \Pi_h v_h\|_n + (e, \hat{v}_h - \Pi_h v_h)_n. \end{aligned}$$

Now by using Lemma 1.3.3 we have

$$\sup_{0\neq v_h\in V_h} \frac{|\lambda_n a_h(u_n, v_h) + (u_n - y, \hat{v}_h)_n|}{|||v_h|||_h}$$

$$\le C\lambda_n^{1/2} |u_n|_{H^2(\Omega)} + C\frac{h^2}{\lambda_n^{1/2}} \|u_n - u_0\|_n + \sup_{0\neq v_h\in V_h} \frac{|(e, \hat{v}_h - \Pi_h v_h)_n|}{|||v_h|||_h}. \tag{1.31}$$

Since e_i, $i = 1, 2, \cdots, n$, are independent and identically distributed random variables, we have

$$\mathbb{E}\big[|(e, \hat{v}_h - \Pi_h v_h)_n|^2\big] = \sigma^2 n^{-1} \|\hat{v}_h - \Pi_h v_h\|_n^2 \le C\sigma^2 n^{-1} h^4 |v_h|_{2,h}^2,$$

where we have used Lemma 1.3.3 in the last inequality.

Let N_h be the dimension of V_h which satisfies $N_h \le Ch^{-d}$ since the mesh is quasi-uniform. Recall that if $\{X_i\}_{i=1}^{N_h}$ are random variables, $\mathbb{E}[\sup_{1\le i\le N_h} |X_i|] \le \sum_{i=1}^{N_h} \mathbb{E}[|X_i|]$. then we have

$$\mathbb{E}\left[\sup_{0\neq v_h\in V_h} \frac{|(e, v_h - \Pi_h v_h)_n|^2}{|||v_h|||_h^2}\right] \le N_h \cdot \sup_{0\neq v_h\in V_h} \mathbb{E}\left[\frac{|(e, v_h - \Pi_h v_h)_n|^2}{|||v_h|||_h^2}\right] \le C\frac{\sigma^2 h^{4-d}}{n\lambda_n}. \tag{1.32}$$

Combining (1.29)-(1.32) we obtain

$$\mathbb{E}\big[|||u_n - \hat{u}_h|||_h^2\big] \le C\lambda_n \mathbb{E}\big[|u_n|_{H^2(\Omega)}^2\big] + C\frac{h^4}{\lambda_n}\mathbb{E}\big[\|u_n - u_0\|_n^2\big] + C\frac{\sigma^2 h^{4-d}}{n\lambda_n}.$$

This completes the proof by using Theorem 1.2.1. □

1.4 Stochastic convergence

In this section we study the stochastic convergence of the error $\|u_0 - \hat{u}_h\|_n$ which characterizes the tail property of $\mathbb{P}(\|u_0 - \hat{u}_h\|_n \ge z)$ for $z > 0$. We assume the noises e_i, $i = 1, 2, \cdots, n$, are independent and identically distributed sub-Gaussian random variables with parameter $\sigma > 0$. A random variable X is sub-Gaussion with parameter σ if it satisfies

$$\mathbb{E}\left[e^{\lambda(X - \mathbb{E}[X])}\right] \le e^{\frac{1}{2}\sigma^2\lambda^2}, \quad \forall \lambda \in \mathbf{R}. \tag{1.33}$$

The probability distribution function of a sub-Gaussion random variable has a exponentially decaying tail, that is, if X is a sub-Gaussion random variable, then

$$\mathbb{P}(|X - \mathbb{E}[X]| \ge z) \le 2e^{-\frac{1}{2}z^2/\sigma^2}, \ \ \forall z > 0. \tag{1.34}$$

In fact, by Markov inequality, for any $\lambda > 0$,

$$\mathbb{P}(X - \mathbb{E}[X] \ge z) = \mathbb{P}(\lambda(X - \mathbb{E}[X]) \ge \lambda z) \le e^{-\lambda z}\mathbb{E}[e^{\lambda(X - \mathbb{E}[X])}] \le e^{-\lambda z - \frac{1}{2}\sigma^2\lambda^2}.$$

Taking $\lambda = z/\sigma^2$ yields $\mathbb{P}(X - \mathbb{E}[X] \ge z) \le e^{-\frac{1}{2}z^2/\sigma^2}$. Similarly, one can prove $\mathbb{P}(X - \mathbb{E}[X] \le -z) \le e^{-\frac{1}{2}z^2/\sigma^2}$. This shows (1.34).

1.4.1 Stochastic convergence of the thin plate splines

We will use several tools from the theory of empirical processes [112, 32] for our analysis. We start by recalling the definition of Orlicz norm. Let ψ be a monotone increasing convex function satisfying $\psi(0) = 0$. Then the Orilicz norm $\|X\|_\psi$ of a random variable X is defined by

$$\|X\|_\psi = \inf\left\{C > 0 : \mathbb{E}\left[\psi\left(\frac{|X|}{C}\right)\right] \le 1\right\}. \tag{1.35}$$

By using Jensen inequality, it is easy to check that $\|X\|_\psi$ is a norm. In the following we will use the $\|X\|_{\psi_2}$ norm with $\psi_2(t) = e^{t^2} - 1$ for any $t > 0$. By definition we know that

$$\mathbb{P}(|X| \geq z) \leq 2\, e^{-z^2/\|X\|_{\psi_2}^2}, \quad \forall z > 0. \tag{1.36}$$

The following lemma is from [112, Lemma 2.2.1]. It shows the inverse of this property.

Lemma 1.4.1. *If there exist positive constants C, K such that $\mathbb{P}(|X| > z) \leq Ke^{-Cz^2}$, $\forall z > 0$, then $\|X\|_{\psi_2} \leq \sqrt{(1+K)/C}$.*

Let T be a semi-metric space with the semi-metric d and let $\{X_t : t \in T\}$ be a random process indexed by T. Then the random process $\{X_t : t \in T\}$ is called sub-Gaussian if

$$\mathbb{P}(|X_s - X_t| > z) \leq 2e^{-\frac{1}{2}z^2/d(s,t)^2}, \quad \forall s, t \in T, \;\; z > 0. \tag{1.37}$$

For a semi-metric space (T, d), an important quantity to characterize the complexity of the set T is the entropy which we now introduce. The covering number $N(\varepsilon, T, d)$ is the minimum number of ε-balls that cover T. A set is called ε-separated if the distance of any two points in the set is strictly greater than ε. The packing number $D(\varepsilon, T, d)$ is the maximum number of ε-separated points in T. $\log N(\varepsilon, T, d)$ is called the covering entropy and $\log D(\varepsilon, T, d)$ is called the packing entropy. It is easy to check that [112, P.98]

$$N(\varepsilon, T, d) \leq D(\varepsilon, T, d) \leq N(\frac{\varepsilon}{2}, T, d). \tag{1.38}$$

The following maximal inequality [112, Section 2.2.1] plays an important role in our analysis.

Lemma 1.4.2. *If $\{X_t : t \in T\}$ is a separable sub-Gaussian random process, then*

$$\| \sup_{s,t \in T} |X_s - X_t| \|_{\psi_2} \leq K \int_0^{\operatorname{diam} T} \sqrt{\log D(\varepsilon, T, d)}\, d\varepsilon.$$

Here $K > 0$ is some constant.

The following result on the estimation of the entropy of Sobolev spaces is due to Birman-Solomyak [15].

Lemma 1.4.3. *Let Q be the unit square in $\mathbf{R}^d$ and $SW^{\alpha,p}(Q)$ be the unit sphere of the Sobolev space $W^{\alpha,p}(Q)$, where $\alpha > 0, p \geq 1$. Then for $\varepsilon > 0$ sufficient small, the entropy*

$$\log N(\varepsilon, SW^{\alpha,p}(Q), \|\cdot\|_{L^q(Q)}) \leq C\varepsilon^{-d/\alpha},$$

where if $\alpha p > d$, $1 \leq q \leq \infty$, otherwise if $\alpha p \leq d$, $1 \leq q \leq q^*$ with $q^* = p(1 - \alpha p/d)^{-1}$.

For any $\delta > 0, \rho > 0$, define

$$S_{\delta,\rho}(\Omega) := \{u \in H^2(\Omega) : \|u\|_n \leq \delta, |u|_{H^2(\Omega)} \leq \rho\}. \tag{1.39}$$

The following lemma estimates the entropy of the set $S_{\delta,\rho}(\Omega)$.

Lemma 1.4.4. *There exists a constant C independent of $\delta, \rho, \varepsilon$ such that*

$$\log N(\varepsilon, S_{\delta,\rho}(\Omega), \|\cdot\|_{L^\infty(\Omega)}) \leq C\left(\frac{\rho+\delta}{\varepsilon}\right)^{d/2}.$$

Proof. By (1.3) we have for any $u \in S_{\delta,\rho}(\Omega)$, $\|u\|_{H^2(\Omega)} \leq C(\|u\|_{L^2(\Omega)} + |u|_{H^2(\Omega)}) \leq C(\|u\|_n + |u|_{H^2(\Omega)}) \leq C(\delta+\rho)$, where we have used the fact that $h_{\max} \leq Cn^{-1/d} \leq C$. The lemma now follows from Lemma 1.4.3. □

The following lemma is proved by the argument in [112, Lemma 2.2.7].

Lemma 1.4.5. *$\{E_n(u) := (e,u)_n : u \in H^2(\Omega)\}$ is a sub-Gaussian random process with respect to the semi-distance $d(u,v) = \|u-v\|_n^*$, where $\|u\|_n^* := \sigma n^{-1/2}\|u\|_n$.*

Proof. By definition $E_n(u) - E_n(v) = \sum_{i=1}^n c_i e_i$, where $c_i = \frac{1}{n}(u-v)(x_i)$. Since e_i is a sub-Gaussion random variable with parameter σ and $\mathbb{E}[e_i] = 0$, by (1.33), $\mathbb{E}[e^{\lambda e_i}] \leq e^{\frac{1}{2}\sigma^2\lambda^2}, \forall \lambda > 0$. Thus, since e_i, $i = 1, 2, \cdots, n$, are independent random variables,

$$\mathbb{E}\left[e^{\lambda\sum_{i=1}^n c_i e_i}\right] \leq e^{\frac{1}{2}\sigma^2\lambda^2\sum_{i=1}^n c_i^2} = e^{\frac{1}{2}\sigma^2 n^{-1}\lambda^2\|u-v\|_n^2} = e^{\frac{1}{2}d(u,v)^2\lambda^2}.$$

This shows $E_n(u) - E_n(v)$ is a sub-Gaussion random variable with parameter $d(u,v)$. By (1.34) we have

$$\mathbb{P}(|E_n(u) - E_n(v)| \geq z) \leq 2e^{-\frac{1}{2}z^2/d(u,v)^2}, \ \ \forall z > 0.$$

This shows the lemma by the definition of sub-Gaussion random process (1.37). □

The following lemma which improves Lemma 1.4.1 will be used in our subsequent analysis.

Lemma 1.4.6. *If X is a random variable which satisfies*

$$\mathbb{P}(|X| > \alpha(1+z)) \leq C_1 e^{-z^2/K_1^2}, \ \ \forall \alpha > 0, z \geq 1,$$

where C_1, K_1 are some positive constants, then $\|X\|_{\psi_2} \leq C(C_1, K_1)\alpha$ for some constant $C(C_1, K_1)$ depending only on C_1 and K_1.

Proof. If $y \geq 2\alpha$, then $z = (y/\alpha) - 1 \geq 1$. Thus

$$\mathbb{P}(|X| > y) = \mathbb{P}(|X| > \alpha(1+z)) \leq C_1 \exp\left[-\frac{1}{K_1^2}\left(\frac{y}{\alpha} - 1\right)^2\right].$$

Since $(\frac{y}{\alpha} - 1)^2 \geq \frac{1}{2}(\frac{y}{\alpha})^2 - 1$ by Cauchy-Schwarz inequality, we obtain

$$\mathbb{P}(|X| > y) \leq C_1 e^{\frac{1}{K_1^2}} e^{-\frac{y^2}{2K_1^2\alpha^2}} = C_1 e^{\frac{1}{K_1^2}} e^{-\frac{y^2}{K_2^2}},$$

where $K_2 := \sqrt{2}\alpha K_1$. On the other hand, if $y < 2\alpha$, then

$$\mathbb{P}(|X| > y) \leq e^{\frac{y^2}{K_2^2}} e^{-\frac{y^2}{K_2^2}} \leq e^{\frac{2}{K_1^2}} e^{-\frac{y^2}{K_2^2}}.$$

Therefore, $\mathbb{P}(|X| > y) \le C_2 e^{-y^2/K_2^2}, \forall y > 0$, where $C_2 = \max(C_1 e^{1/K_1^2}, e^{2/K_1^2})$. This implies by Lemma 1.4.1,

$$\|X\|_{\psi_2} \le \sqrt{1+C_2}K_2 = C(C_1, K_1)\alpha, \text{ where } C(C_1, K_1) = \sqrt{2}K_1\sqrt{1+C_2},$$

which completes the proof. □

Theorem 1.4.1. *Let $u_n \in H^2(\Omega)$ be the solution of* (1.13). *Denote by $\rho_0 = |u_0|_{H^2(\Omega)} + \sigma n^{-1/2}$. If we take*

$$\lambda_n^{1/2+d/8} = O(\sigma n^{-1/2}\rho_0^{-1}), \tag{1.40}$$

then there exists a constant $C > 0$ such that

$$\| \|u_n - u_0\|_n \|_{\psi_2} \le C\lambda_n^{1/2}\rho_0, \quad \| |u_n|_{H^2(\Omega)} \|_{\psi_2} \le C\rho_0. \tag{1.41}$$

Proof. We will only prove the first estimate in (1.41) by the peeling argument. The other estimate can be proved in a similar way. From (1.10) it follows that

$$\|u_n - u_0\|_n^2 + \lambda_n |u_n|_{H^2(\Omega)}^2 \le 2(e, u_n - u_0)_n + \lambda_n |u_0|_{H^2(\Omega)}^2. \tag{1.42}$$

Let $\delta > 0,\ \rho > 0$ be two constants to be determined later, and

$$A_0 = [0, \delta), A_i = [2^{i-1}\delta, 2^i\delta), \ \ B_0 = [0, \rho), B_j = [2^{j-1}\rho, 2^j\rho), \quad i, j \ge 1. \tag{1.43}$$

For $i, j \ge 0$, define

$$F_{ij} = \{v \in H^2(\Omega) : \ \|v\|_n \in A_i \ , \ |v|_{H^2(\Omega)} \in B_j\}.$$

Then we have

$$\mathbb{P}(\|u_n - u_0\|_n > \delta) \le \sum_{i=1}^{\infty}\sum_{j=0}^{\infty} \mathbb{P}(u_n - u_0 \in F_{ij}). \tag{1.44}$$

Now we estimate $\mathbb{P}(u_n - u_0 \in F_{ij})$. By Lemma 1.4.5, $\{(e, v)_n : v \in H^2(\Omega)\}$ is a sub-Gaussion random process with respect to the semi-distance $d(u, v) = \sigma n^{-1/2}\|u - v\|_n$. It is easy to see that

$$\operatorname{diam} F_{ij} \le \sigma n^{-1/2} \sup_{u-u_0, v-u_0 \in F_{ij}} (\|u - u_0\|_n + \|v - u_0\|_n) \le 2\sigma n^{-1/2} \cdot 2^i\delta.$$

Then by (1.38) and the maximal inequality in Lemma 1.4.2 we have

$$\begin{aligned} \| \sup_{u-u_0 \in F_{ij}} |(e, u - u_0)_n| \|_{\psi_2} &\le K \int_0^{\sigma n^{-1/2} \cdot 2^{i+1}\delta} \sqrt{\log N\left(\frac{\varepsilon}{2}, F_{ij}, d\right)}\, d\varepsilon \\ &= K \int_0^{\sigma n^{-1/2} \cdot 2^{i+1}\delta} \sqrt{\log N\left(\frac{\varepsilon}{2\sigma n^{-1/2}}, F_{ij}, \|\cdot\|_n\right)}\, d\varepsilon. \end{aligned}$$

By Lemma 1.4.4 we have the estimate for the entropy

$$\begin{aligned}\log N\left(\frac{\varepsilon}{2\sigma n^{-1/2}}, F_{ij}, \|\cdot\|_n\right) &\leq \log N(\frac{\varepsilon}{2\sigma n^{-1/2}}, F_{ij}, \|\cdot\|_{L^\infty(\Omega)}) \\ &\leq C\left(\frac{2\sigma n^{-1/2}\cdot(2^i\delta+2^j\rho)}{\varepsilon}\right)^{d/2}.\end{aligned}$$

Therefore,

$$\begin{aligned}\|\sup_{u-u_0\in F_{ij}}|(e,u-u_0)_n\|_{\psi_2} &\leq K\int_0^{\sigma n^{-1/2}\cdot 2^{i+1}\delta}\left(\frac{2\sigma n^{-1/2}\cdot(2^i\delta+2^j\rho)}{\varepsilon}\right)^{d/4}d\varepsilon \\ &= C\sigma n^{-1/2}(2^i\delta+2^j\rho)^{d/4}(2^i\delta)^{1-d/4} \\ &\leq C\sigma n^{-1/2}[2^i\delta+(2^i\delta)^{1-d/4}(2^j\rho)^{d/4}]. \qquad (1.45)\end{aligned}$$

By (1.42) and (1.36) we have for $i,j\geq 1$:

$$\begin{aligned}\mathbb{P}(u_n-u_0\in F_{ij}) &\leq \mathbb{P}(2^{2(i-1)}\delta^2+\lambda_n 2^{2(j-1)}\rho^2\leq 2\sup_{u-u_0\in F_{ij}}|(e,u-u_0)_n|+\lambda_n\rho_0^2) \\ &= \mathbb{P}(2\sup_{u-u_0\in F_{ij}}|(e,u-u_0)_n|\geq 2^{2(i-1)}\delta^2+\lambda_n 2^{2(j-1)}\rho^2-\lambda_n\rho_0^2) \\ &\leq 2\exp\left[-\frac{1}{C\sigma^2 n^{-1}}\left(\frac{2^{2(i-1)}\delta^2+\lambda_n 2^{2(j-1)}\rho^2-\lambda_n\rho_0^2}{2^i\delta+(2^i\delta)^{1-d/4}(2^j\rho)^{d/4}}\right)^2\right].\end{aligned}$$

Now we take

$$\delta^2=\lambda_n\rho_0^2(1+z)^2,\ \rho=\rho_0,\ \text{ where } z\geq 1. \qquad (1.46)$$

Since by assumption $\lambda_n^{1/2+d/8}=O(\sigma n^{-1/2}\rho_0^{-1})$ and $\sigma n^{-1/2}\rho_0^{-1}\leq 1$, we have $\lambda_n\leq C$ for some constant. By some simple calculation we have for $i,j\geq 1$,

$$\mathbb{P}(u_n-u_0\in F_{ij}) \leq 2\exp\left[-C\left(\frac{2^{2(i-1)}z(1+z)+2^{2(j-1)}}{2^i(1+z)+(2^i(1+z))^{1-d/4}(2^j)^{d/4}}\right)^2\right].$$

By using the elementary inequality $ab\leq\frac{1}{p}a^p+\frac{1}{q}b^q$ for any $a,b>0,p,q>1,p^{-1}+q^{-1}=1$, we have $(2^i(1+z))^{1-d/4}(2^j)^{d/4}\leq(1+z)2^i+2^j$. Thus

$$\mathbb{P}(u_n-u_0\in F_{ij})\leq 2\exp\left[-C(2^{2i}z^2+2^{2j})\right].$$

Similarly, one can prove for $i\geq 1,j=0$,

$$\mathbb{P}(u_n-u_0\in F_{i0})\leq 2\exp\left[-C(2^{2i}z^2)\right].$$

Therefore, since $\sum_{j=1}^\infty e^{-C(2^{2j})}\leq e^{-C}<1$ and $\sum_{i=1}^\infty e^{-C(2^{2i}z^2)}\leq e^{-Cz^2}$, we finally obtain

$$\begin{aligned}\sum_{i=1}^\infty\sum_{j=0}^\infty\mathbb{P}(u_n-u_0\in F_{ij}) &\leq 2\sum_{i=1}^\infty\sum_{j=1}^\infty e^{-C(2^{2i}z^2+2^{2j})}+2\sum_{i=1}^\infty e^{-C(2^{2i}z^2)} \\ &\leq 4e^{-Cz^2}.\end{aligned}$$

Now inserting the above estimate into (1.44) we have

$$\mathbb{P}(\|u_n - u_0\|_n > \lambda_n^{1/2}\rho_0(1+z)) \le 4e^{-Cz^2}, \quad \forall z \ge 1. \tag{1.47}$$

This implies by using Lemma 1.4.6 that $\|\|u_n - u_0\|_n\|_{\psi_2} \le C\lambda_n^{1/2}\rho_0$, which completes the proof. □

We remark that (1.47) implies that

$$\lim_{z\to\infty} \overline{\lim_{n\to\infty}} \mathbb{P}(\|u_n - u_0\|_n > \lambda_n^{1/2}\rho_0(1+z)) = 0.$$

In terms of the terminology of the stochastic convergence order, we have $\|u_n - u_0\|_n = O_p(\lambda_n^{1/2})\rho_0$ which by assumption (1.40) yields

$$\|u_n - u_0\|_n = O_p(n^{-\frac{2}{4+d}})\sigma^{\frac{4}{4+d}}\rho_0^{-\frac{4}{4+d}}.$$

This estimate is proved in [32, Section 10.1.1] when $d = 1$. Our result in Theorem 1.4.1 is stronger in the sense that it also provides the tail property of the probability distribution function of the random error $\|u_n - u_0\|_n$.

1.4.2 Stochastic convergence of the finite element method

The following lemma provides the estimate of the entropy of finite dimension subsets [32, Corollary 2.6].

Lemma 1.4.7. *Let G be a finite dimensional subspace of $L^2(\Omega)$ of dimension $N > 0$ and $G_R = \{f \in G : \|f\|_{L^2(\Omega)} \le R\}$. Then*

$$N(\varepsilon, G_R, \|\cdot\|_{L^2(\Omega)}) \le (1 + 4R/\varepsilon)^N, \ \ \forall \varepsilon > 0.$$

Lemma 1.4.8. *Let $G_h := \{v_h \in V_h : |||v_h|||_h = (\lambda_n |v_h|_{2,h}^2 + \|\hat{v}_h\|_n^2)^{1/2} \le 1\}$. Assume that $h = O(\lambda_n^{1/4})$ and $n\lambda_n^{d/4} \ge 1$. Then*

$$\| \sup_{v_h \in G_h} |(e, \hat{v}_h - \Pi_h v_h)_n| \|_{\psi_2} \le C\sigma n^{-1/2}\lambda_n^{-d/8}.$$

Proof. Similar to the proof of Lemma 1.4.5 we know that $\{\hat{E}_n(v_h) := (e, \hat{v}_h - \Pi_h v_h)_n, \ \forall v_h \in G_h\}$ is a sub-Gaussion random process with respect to the semi-distance $\hat{d}(v_h, w_h) = \sigma n^{-1/2}\|(\hat{v}_h - \Pi_h v_h) - (\hat{w}_h - \Pi_h w_h)\|_n$. By Lemma 1.3.3, for any $v_h \in G_h$, $\|\hat{v}_h - \Pi_h v_h\|_n \le Ch^2|v_h|_{2,h} \le Ch^2\lambda_n^{-1/2} \le C$, where we have used the assumption $h = O(\lambda_n^{1/4})$ in the last inequality. This implies that the diameter of G_h is bounded by $C\sigma n^{-1/2}$. Now by the maximal inequality in Lemma 1.4.2

$$\| \sup_{v_h \in G_h} |(e, \hat{v}_h - \Pi_h v_h)_n| \|_{\psi_2} \le K \int_0^{C\sigma n^{-1/2}} \sqrt{\log N\left(\frac{\varepsilon}{2}, G_h, \hat{d}\right)}\, d\varepsilon. \tag{1.48}$$

For any $v_h \in V_h$, by Lemma 1.3.3, $\Pi_h v_h \in H^2(\Omega)$ and thus by (1.3)

$$\begin{aligned}\|\Pi_h v_h\|_{L^2(\Omega)} &\leq C(h_{\max}^2|\Pi_h v_h|_{H^2(\Omega)} + \|\Pi_h v_h\|_n) \\ &\leq C(n^{-2/d}\lambda_n^{-1/2} + \|\Pi_h v_h - \hat{v}_h\|_n + \|\hat{v}_h\|_n) \\ &\leq C(n^{-2/d}\lambda_n^{-1/2} + Ch^2\lambda_n^{-1/2} + 1) \\ &\leq C,\end{aligned}$$

where we have used $h = O(\lambda_n^{1/4})$ and $n\lambda_n^{d/4} \geq 1$ in the last inequality. Thus

$$\|v_h\|_{L^2(\Omega)} \leq \|v_h - \Pi_h v_h\|_{L^2(\Omega)} + \|\Pi_h v_h\|_{L^2(\Omega)} \leq Ch^2|v_h|_{2,h} + C \leq C, \quad \forall v_h \in G_h. \tag{1.49}$$

Moreover, by Lemma 1.3.3 and the inverse estimate,

$$\hat{d}(v_h, w_h) \leq C\sigma n^{-1/2}h^2|v_h - w_h|_{2,h} \leq C\sigma n^{-1/2}\|v_h - w_h\|_{L^2(\Omega)}, \quad \forall v_h, w_h \in V_h. \tag{1.50}$$

Now since the dimension of V_h is bounded by Ch^{-d}, Lemma 1.4.7 together with (1.49)-(1.50) implies

$$\begin{aligned}\log N\left(\frac{\varepsilon}{2}, G_h, \hat{d}\right) &= \log N\left(\frac{\varepsilon}{C\sigma n^{-1/2}}, G_h, \|\cdot\|_{L^2(\Omega)}\right) \\ &\leq Ch^{-d}(1 + \sigma n^{-1/2}/\varepsilon).\end{aligned}$$

Inserting this estimate to (1.48)

$$\begin{aligned}\| \sup_{v_h \in G_h} |(e, \hat{v}_h - \Pi_h v_h)_n| \, \|_{\psi_2} &\leq C\int_0^{C\sigma n^{-1/2}} \sqrt{Ch^{-d}(1 + \sigma n^{-1/2}/\varepsilon)}\, d\varepsilon \\ &\leq Ch^{-d/2}\sigma n^{-1/2}.\end{aligned}$$

This completes the proof since $h = O(\lambda_n^{1/4})$. □

The following theorem is the main result of this section.

Theorem 1.4.2. *Let $u_h \in V_h$ be the solution of* (1.15)*. Denote by $\rho_0 = |u_0|_{H^2(\Omega)} + \sigma n^{-1/2}$. If we take*

$$h = O(\lambda_n^{1/4}) \text{ and } \lambda_n^{1/2+d/8} = O(\sigma n^{-1/2}\rho_0^{-1}), \tag{1.51}$$

then there exists a constant $C > 0$ such that

$$\| \|\hat{u}_h - u_0\|_n \|_{\psi_2} \leq C\lambda_n^{1/2}\rho_0, \quad \| |u_h|_{H^2(\Omega)} \|_{\psi_2} \leq C\rho_0. \tag{1.52}$$

Proof. By (1.29)-(1.31) we have

$$\begin{aligned}&\lambda_n^{1/2}|u_h|_{H^2(\Omega)} + \|\hat{u}_h - u_0\|_n \\ \leq\ & C(1 + \frac{h^2}{\lambda_n^{1/2}})\|u_n - u_0\|_n + C(h^2 + \lambda_n^{1/2})(|u_n|_{H^2(\Omega)} + |u_0|_{H^2(\Omega)}) \\ +\ & C\sup_{0 \neq v_h \in V_h} \frac{|(e, \hat{v}_h - \Pi_h v_h)_n|}{|||v_h|||_h}.\end{aligned}$$

The theorem follows now from Theorem 1.4.1, Lemma 1.4.8 and the assumption $\sigma n^{-1/2} \leq C\lambda_n^{1/2+d/8}\rho_0$. □

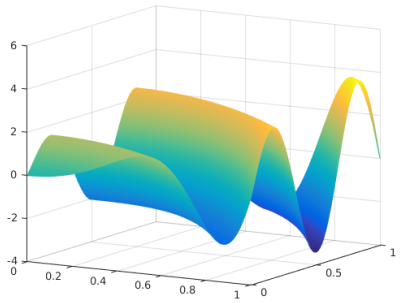


FIGURE 1.3: *The surface plot of the exact solution* u_0.

By (1.36), we know from Theorem 1.4.2 that

$$\mathbb{P}(\|\hat{u}_h - u_0\|_n \geq z) \leq 2e^{-z^2/(C\lambda_n \rho_0^2)}, \quad \forall z > 0,$$

that is, the probability density function of the random error $\|\hat{u}_h - u_0\|_n$ decays exponentially as $n \to \infty$.

1.5 Numerical examples

From Theorem 1.4.2 we know that the mesh size should be comparable with $\lambda_n^{1/4}$. The smoothing parameter λ_n is usually determined by the cross-validation in the literature [114]. Here we propose a self-consistent algorithm to determine the parameter λ_n based on $\lambda_n^{1/2+d/8} = \sigma n^{-1/2}(|u_0|_{H^2(\Omega)} + \sigma n^{-1/2})^{-1}$ as indicated in Theorem 1.4.2. In the algorithm we estimate $|u_0|_{H^2(\Omega)}$ by $|u_h|_{2,h}$ and σ by $\|u_h - y\|_n$ since $\|u_0 - y\|_n = \|e\|_n$ provides a good estimation of the variance by the law of large number.

Algorithm 1.5.1. (SELF-CONSISTENT ALGORITHM FOR FINDING λ_n)
1° *Given an initial guess of* $\lambda_{n,0}$;
2° *For* $k \geq 0$ *and* $\lambda_{n,k}$ *known, compute* u_h *with the parameter* $\lambda_{n,k}$ *over a quasi-uniform mesh of the mesh size* $h = \lambda_{n,k}^{1/4}$;
3° *Compute* $\lambda_{n,k+1}^{1/2+d/8} = \|u_h - y\|_n n^{-1/2}(|u_h|_{2,h} + \|u_h - y\|_n n^{-1/2})^{-1}$.

Now we show several examples to confirm our theoretical analysis. We will always take $\Omega = (0,1) \times (0,1)$ and $\{x_i\}_{i=1}^n$ being uniformly distributed over Ω. We take $u_0 = \sin(2\pi x^2 + 3\pi y)e^{\sqrt{x^3+y}}$, see Figure 1.3. The finite element mesh of Ω is construct by first dividing the domain into $h^{-1} \times h^{-1}$ uniform rectangles and then connecting the lower left and upper right vertices of each rectangle.

Example 1.5.1. *In this example we show that the choice of the smoothing parameter* λ_n *by (1.51) is optimal. We set* e_i, $i = 1, 2, \cdots, n$, *being independent normal random variables with variance* $\sigma = 1$ *and* $n = 2500$. *Since* $|u_0|_{H^2(\Omega)} \approx 200$, (1.51) *suggests the optimal choice of* $\lambda_n \approx 3 \times 10^{-6}$. *Figure 1.4 shows that* $\lambda_n = 1 \times 10^{-6}$ *is the best choice among 11 different choices* $\lambda_n = 10^{-k}$, $k = 1, 2, \cdots, 10$. *Here we also choose the mesh size* $h = \lambda_n^{1/4}$ *according to Theorem 1.4.2.*

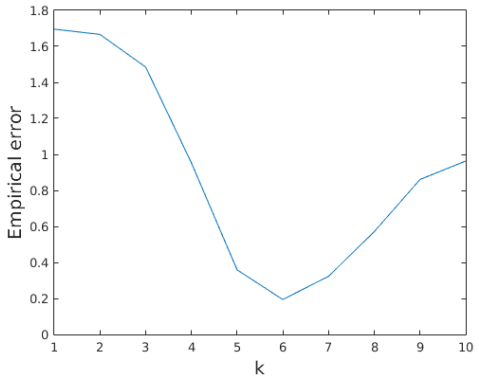
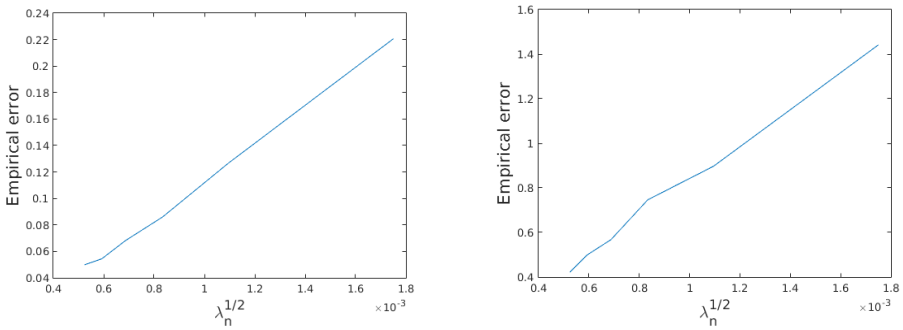


FIGURE 1.4: *The empirical error* $\|u_0 - u_h\|_n$ *for 11 different choices of* $\lambda_n = 10^{-k}, k = 0, 1, \cdots, 10$. *The mesh size* $h = \lambda_n^{1/4}$.



(a) (b)

FIGURE 1.5: *(a) The linear dependence of the empirical error* $\|u_0 - u_h\|_n$ *on* $\lambda_n^{1/2}$ *for* $\sigma = 1$. *(b) The linear dependence of the empirical error* $\|u_0 - u_h\|_n$ *on* $\lambda_n^{1/2}$ *for combined random noises.*

Example 1.5.2. *In this example we show that the empirical error* $\|u_0 - u_h\|_n$ *depends linearly on* $\lambda_n^{1/2}$ *to confirm* (1.52). *We set* $e_i,\ i = 1, 2, \cdots, n$, *to be independent normal random variables with variance* $\sigma = 1$. *We take* n *varying from* 2500 *to* 9×10^4. *In this test we use the optimal* λ_n *and take the mesh size* $h = \lambda_n^{1/4}$. *Figure 1.5 (a) shows clearly the linear dependence of the empirical error on* $\lambda_n^{1/2}$. *We also run the test for combined random errors, i.e.,* $e_i = \eta_i + \alpha_i$, *where* η_i *and* α_i *are independent normal random variables with variance* $\sigma_1 = 1$ *and* $\sigma_2 = 10$. *Figure 1.5 (b) shows also the linear dependence of the empirical error on* $\lambda_n^{1/2}$.

Example 1.5.3. *We test the efficiency of the Algorithm 1.5.1 to estimate the smoothing parameter* λ_n. *We will show two experiments of different noise levels. In the first test we set* $e_i,\ i = 1, 2, \cdots, n$, *being independent normal random variables with variance* $\sigma = 1$ *and* $n = 2500$. *Figure 1.6 (a) and (b) show clearly that the sequence of* $\{\lambda_{n,k}\}$ *generated by Algorithm 1.5.1 converges.* $\lambda_{n,16} = 4.12 \times 10^{-6}$ *agrees with the optimal choice* 3×10^{-6} *given by* (1.51). *Furthermore,* $\|u_h - y\|_n = 0.99$ *provides a good estimate of the variance* σ.

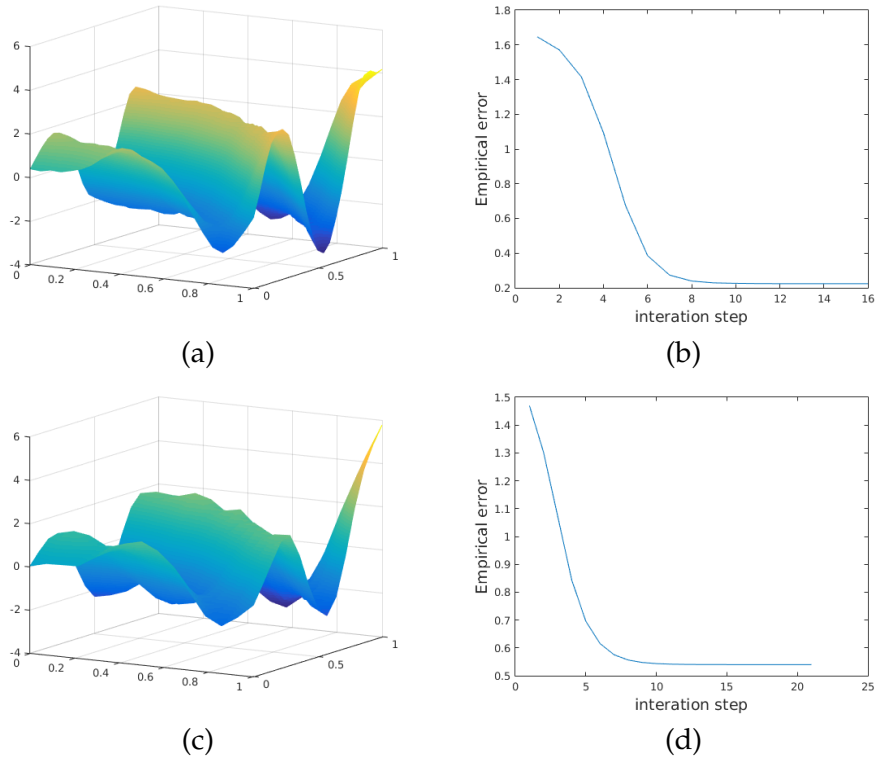


FIGURE 1.6: *(a) The solution u_h at the end of iteration for $\sigma = 1$. (b) The empirical error $\|u_0 - u_h\|_n$ of each iteration for $\sigma = 1$. (c) The solution u_h at the end of iteration for the combined random error $e_i = \eta_i + \alpha_i$. (d) The empirical error $\|u_0 - u_h\|_n$ of each iteration for the combined random error $e_i = \eta_i + \alpha_i$.*

We now consider the combined random noise. Let $e_i = \eta_i + \alpha_i$, $i = 1, 2, \cdots, n$, where η_i and α_i are independent normal random variables with variance $\sigma_1 = 1$ and $\sigma_2 = 10$. It is obvious that $\sigma^2 = \mathbb{E}e_i^2 = \sigma_1^2 + \sigma_2^2 = 101$. Let $n = 4 \times 10^4$. Again Figure 1.6 (c) and (d) show the sequence $\{\lambda_{n,k}\}$ generated by Algorithm 1.5.1 converges. Now $\lambda_{n,19} = 2.16 \times 10^{-5}$ which fits well the optimal choice 1.03×10^{-5} given by (1.51)*. Also $\|u_h - y\|_n = 10.07$ gives a good estimate of the variance σ.*

Chapter 2

Elliptic problems with observational boundary data

2.1 Introduction

In this chapter, we discuss another application of empirical process. Here, we demonstrate the Dirichlet problem with observational boundary data and analyze the convergence of FEM for this problem.

In many scientific and engineering applications involving partial differential equations, the input data such as sources or boundary conditions are usually given through the measurements which may subject to random noises. Let $\Omega \subset \mathbf{R}^2$ be a bounded domain with smooth boundary Γ. In this chapter we consider the problem to find $u \in H^1(\Omega)$ such that

$$-\Delta u = f \ \text{ in } \Omega, \quad u = g_0 \ \text{ on } \Gamma. \tag{2.1}$$

Here $f \in L^2(\Omega)$ is given but the boundary condition $g_0 \in H^2(\Gamma)$ is generally unknown. We assume that we know the measurements $g_i = g_0(x_i) + e_i$, $i = 1, 2, \cdots, n$, where $\mathbb{T} = \{x_i : 1 \leq i \leq n\}$ is the set of the measurement locations on the boundary Γ and e_i, $i = 1, 2, \cdots, n$, are independent identically distributed random variables over some probability space $(\mathfrak{X}, \mathcal{F}, \mathbb{P})$ satisfying $\mathbb{E}[e_i] = 0$ and $\mathbb{E}[e_i^2] = \sigma > 0$. In this chapter $\mathbb{P}$ denotes the probability measure and $\mathbb{E}[X]$ denotes the expectation of the random variable X. We remark that for simplicity we only consider the problem of observational Dirichlet boundary data in this chapter and the problem with observational sources f or other type of boundary conditions can be studied by the same method.

A different perspective of solving partial differential equations with uncertain input data due to incomplete knowledge or inherent variability in the system has drawn considerable interests in recent years (see e.g. [101, 23, 40, 108] and the references therein). The goal of those studies is to learn about the uncertainties in system outputs of interest, given information about the uncertainties in the system inputs which are modeled as random field. This goal usually leads to the mathematical problem of breaking the curse of dimensionality for solving partial differential equations having large number of parameters.

The classical problem to find a smooth function from the knowledge of its observation at scattered locations subject to random noises is well studied in the literature [114]. One popular model to tackle this classical problem is to use the thin plate spline model [27, 111] which can be efficiently solved by using finite element methods [12, 92, 19]. The scattered

data in our problem (2.1) are defined on the boundary of the domain and a straightforward application of the method developed in [27, 111, 12, 92, 19] would lead to solve a fourth order elliptic equation on the boundary which would be much more expansive than the method proposed in this chapter.

Our method is based on the following weak formulation of Lagrangian multiplier for (2.1) in [100]: Find $(u, \lambda) \in H^1(\Omega) \times H^{-1/2}(\Gamma)$ such that

$$\begin{aligned} (\nabla u, \nabla v) + \langle \lambda, v \rangle &= (f, v), \quad \forall v \in H^1(\Omega), && (2.2) \\ \langle \mu, u \rangle &= \langle \mu, g \rangle, \quad \forall \mu \in H^{1/2}(\Gamma), && (2.3) \end{aligned}$$

where $(\cdot, \cdot)$ is the duality pairing between $H^1(\Omega)$ and $H^1(\Omega)'$ which is an extension of the inner product of $L^2(\Omega)$ and $\langle \cdot, \cdot \rangle$ is the duality pairing between $H^{1/2}(\Gamma)$ and $H^{-1/2}(\Gamma)$ which is an extension of the inner product of $L^2(\Gamma)$. Let Ω_h be a polygonal domain which approximates the domain Ω. Let $V_h \subset H^1(\Omega_h)$ and $Q_h \subset L^2(\Gamma)$ be the finite element spaces for approximating the field variable and the Lagrangian multiplier. Our finite element method is defined as follows: Find $(u_h, \lambda_h) \in V_h \times Q_h$ such that

$$\begin{aligned} (\nabla u_h, \nabla v_h)_{\Omega_h} + \langle \lambda_h, v_h \rangle_n &= (I_h f, v_h)_{\Omega_h}, \quad \forall v_h \in V_h, \\ \langle \mu_h, u_h \rangle_n &= \langle \mu_h, g \rangle_n, \quad \forall \mu_h \in Q_h, \end{aligned}$$

where $(\cdot, \cdot)_{\Omega_h}$ is the inner product of $L^2(\Omega_h)$, $\langle \cdot, \cdot \rangle_n$ is some quadrature rule for approximating $\langle \cdot, \cdot \rangle$, and I_h is some finite element interpolation operator (we refer to Section 2.2 for the precise definitions). We remark that while the method of Lagrangian multiplier is one of the standard ways in enforcing Dirichlet boundary condition on smooth domains, it is essential here for solving the problem with Dirichlet observational boundary data even when the domain Ω is a polygon. One can also combine the techniques developed in this chapter with other weak formulations in [104] to deal with the observational Dirichlet boundary condition.

Our analysis in Section 2.3 shows that

$$\mathbb{E}\left[\|u - u_h \circ \Phi_h^{-1}\|_{L^2(\Omega)}\right] \leq Ch^2 |\ln h| \Delta(u, f, g_0) + C|\ln h|(\sigma n^{-1/2}), \quad (2.4)$$

where $\Delta(u, f, g_0) = \|u\|_{H^2(\Omega)} + \|f\|_{H^2(\Omega)} + \|g_0\|_{H^2(\Gamma)}$ and $\Phi_h : \Omega_h \to \Omega$ is the Lenoir homeomorphism defined in Section 2.3 . This error estimate suggests that in order to achieve the optimal convergence, one should take the number of sampling points satisfying $\sigma n^{-1/2} \leq Ch^2$ to compute the solution over a finite element mesh of the mesh size h. For problems having Neumann or Robin boundary conditions, the same method of the analysis in this chapter yields this relation should be changed to $\sigma n^{-1/2} \leq Ch$. This suggests the importance of appropriate balance between the number of measurements and the finite element mesh sizes for solving PDEs with random observational data.

If the random variables $e_i, 1 \leq i \leq n$, are also sub-Gaussian, we prove by resorting to the theory of empirical processes that for any $z > 0$,

$$\mathbb{P}\left(\|u - u_h \circ \Phi_h^{-1}\|_{L^2(\Omega)} \geq \left[h^2 |\ln h| \Delta(u, f, g_0) + |\ln h|(\sigma n^{-1/2})\right] z\right) \leq 2e^{-Cz^2}.$$

This implies that the probability of the random error $\|u - u_h\|_{L^2(\Omega)}$ violating the error estimate in (2.4) decays exponentially.

The layout of the chapter is as follows. In Section 2.2 we introduce our finite element formulation and derive an error estimate based on the Babuška-Brezzi theory. In Section 2.3 we study the random finite element error in terms of the expectation. In Section 2.4 we show the stochastic convergence of our method when the random noise is sub-Gaussion. In Section 2.5 we report some numerical examples to confirm our theoretical analysis.

2.2 The finite element method

We start by introducing the finite element meshes. Let $\mathcal{M}_h$ be a mesh over Ω consisting of curved triangles. We assume that each element $K \in \mathcal{M}_h$ has at most one curved edge and the edge of the element K is curved only when its two vertices all lie on the boundary Γ. For any $K \in \mathcal{M}_h$, we denote $\tilde{K}$ the straight triangle which has the same vertices as K. We set $\Omega_h = \cup_{K\in\mathcal{M}_h}\tilde{K}$ and assume that the mesh $\tilde{\mathcal{M}}_h = \{\tilde{K} : K \in \mathcal{M}_h\}$ over Ω_h is shape regular and quasi-uniform:

$$h_{\tilde{K}} \le C\rho_{\tilde{K}}, \ \ \forall K \in \mathcal{M}_h, \quad h_{\tilde{K}} \le Ch_{\tilde{K}'}, \ \ \forall K, K' \in \mathcal{M}_h, \tag{2.1}$$

where $h_{\tilde{K}}$ and $\rho_{\tilde{K}}$ are the diameter of $\tilde{K}$ and the diameter of the biggest circle inscribed in $\tilde{K}$. The finite element space for the field variable is then defined as

$$V_h = \{v_h \in C(\bar{\Omega}_h) : v_h|_{\tilde{K}} \in P_1(\tilde{K}), \forall \tilde{K} \in \tilde{\mathcal{M}}_h\},$$

where $P_1(\tilde{K})$ is the set of the linear polynomials on $\tilde{K}$. As usual, we demote $h = \max_{\tilde{K}\in\tilde{\mathcal{M}}_h} h_{\tilde{K}}$.

Let $\mathcal{E}_h = \{K \cap \Gamma : K \in \mathcal{M}_h\}$ be the mesh of Γ which is induced from $\mathcal{M}_h$. We assume that each element $E \in \mathcal{E}_h$ is the image of the reference element $\hat{E} = [0, 1]$ under a smooth mapping F_E. Since the boundary Γ is smooth, the argument in [21, Theorem 4.3.3] implies that if the diameter of the element h_E is sufficiently small,

$$\|\hat{D}F_E\|_{L^\infty(\hat{E})} \le Ch_E, \ \ \|D_T F_E^{-1}\|_{L^\infty(E)} \le Ch_E^{-1}, \ \ \forall E \in \mathcal{E}_h, \tag{2.2}$$

where $\hat{D}$ is the derivative in $\hat{E}$ and D_T is the tangential derivative on Γ. It is then obvious that there are constants C_1, C_2 independent of the mesh $\mathcal{M}_h$ such that $C_1 h \le h_E \le C_2 h, \ \ \forall E \in \mathcal{E}_h$. We use the following finite element space for the Lagrangian multiplier [104]:

$$Q_h = \{\mu_h \in C(\Gamma) : \mu_h|_E = \hat{\mu}_h \circ F_E^{-1} \text{ for some } \hat{\mu}_h \in P_1(\hat{E}), \forall E \in \mathcal{E}_h\}, \tag{2.3}$$

where $P_1(\hat{E})$ is the set of linear polynomials over $\hat{E}$.

We assume that the measurement locations $\mathbb{T}$ are uniformly distributed over Γ in the sense that [111] there exists a constant $B > 0$ such that $\frac{s_{\max}}{s_{\min}} \le B$, where

$$s_{\max} = \sup_{x\in\Gamma} \inf_{1\le i\le n} s(x, x_i), \quad s_{\min} = \inf_{1\le i\ne j\le n} s(x_i, x_j).$$

Here $s(x, y)$ is the arc length between $x, y \in \Gamma$. It is easy to see that there exist constants B_1, B_2 such that $B_1 n^{-1} \le s_{\max} \le Bs_{\min} \le B_2 n^{-1}$.

We introduce the empirical inner product between the data and any function $v \in C(\Gamma)$ as $\langle g, v\rangle_n = \sum_{i=1}^n \alpha_i g_i v(x_i)$. We also write $\langle u, v\rangle_n = \sum_{i=1}^n \alpha_i u(x_i)v(x_i)$ for any $u, v \in C(\Gamma)$ and the empirical norm $\|u\|_n = (\sum_{i=1}^n \alpha_i u^2(x_i))^{1/2}$ for any $u \in C(\Gamma)$. We remark that the empirical norm is in fact a semi-norm on $C(\Gamma)$. The weights α_i, $i = 1, 2\cdots, n$, are chosen such that $\langle u, v\rangle_n$ is a good quadrature formula for the inner product $\langle u, v\rangle$ that we describe now.

Let $\mathbb{T}_E = \mathbb{T}\cap E$ be the measurement points in $E \in \mathcal{E}_h$. Since the measurement locations are uniformly distributed, $n_E = \#\mathbb{T}_E \sim nh_E$. We further assume that $t_{j,E} = F_E^{-1}(x_j)$, $j = 1, 2\cdots, n_E$, are ordered as $0 = t_{0,E} \le t_{1,E} < t_{2,E} < \cdots < t_{n_E,E} \le t_{n_E+1,E} = 1$. We remark that the vertices of the element E need not be at the measurement locations. Denote $\Delta t_{j,E} = t_{j,E} - t_{j-1,E}$, $j = 1, 2, \cdots, n_E + 1$. We define the following quadrature formula

$$Q_{\mathbb{T}_E}(w) = \sum_{j=1}^{n_E} \omega_{j,E} w(t_{j,E}), \quad \forall w \in C(\hat{E}), \tag{2.4}$$

where $\omega_{1,E} = \Delta t_{1,E} + \frac{1}{2}\Delta t_{2,E}, \omega_{j,E} = \frac{1}{2}(\Delta t_{j,E} + \Delta t_{j+1,E}), j = 2, \cdots, n_E - 1, \omega_{n_E,E} = \frac{1}{2}\Delta t_{n_E,E} + \Delta t_{n_E+1,E}$.

Lemma 2.2.1. *There exists a constant C independent of $\mathbb{T}_E$ such that*

$$\begin{aligned}\left|\int_0^1 w(t)dt - Q_{\mathbb{T}_E}(w)\right| \;&\le\; C\int_0^1 |w''(t)|dt + \frac{1}{2}\Delta t_{1,E}\int_{t_{0,E}}^{t_{1,E}} |w'(t)|dt \\ &+\; \frac{1}{2}\Delta t_{n_E+1,E}\int_{t_{n_E,E}}^{t_{n_E+1,E}} |w'(t)|dt, \quad \forall w \in W^{2,1}(\hat{E}).\end{aligned}$$

Proof. We introduce the standard piecewise trapezoid quadrature rule

$$\tilde{Q}_{\mathbb{T}_E}(w) = \sum_{j=1}^{n_E+1} \Delta t_{j,E}\frac{w(t_{j-1,E}) + w(t_{j,E})}{2}, \tag{2.5}$$

which is exact for linear functions. By the Bramble-Hilbert lemma we know that there exists a constant C such that

$$\left|\int_0^1 w(t)dt - \tilde{Q}_{\mathbb{T}_E}(w)\right| \le C\int_0^1 |w''(t)|dt, \quad \forall w \in W^{2,1}(\hat{E}).$$

Now the lemma follows since

$$\begin{aligned}&Q_{\mathbb{T}_E}(w) - \tilde{Q}_{\mathbb{T}_E}(w) \qquad (2.6)\\ =\;& \frac{1}{2}\Delta t_{1,E}(w(t_{1,E}) - w(t_{0,E})) + \frac{1}{2}\Delta t_{n_E+1,E}(w(t_{n_E,E}) - w(t_{n_E+1,E})).\end{aligned}$$

This completes the proof. □

Now for any $v \in C(\Gamma)$ we can define the following quadrature rule

which defines the weights α_j, $j = 1, 2, \cdots, n$, in the empirical inner product,

$$\begin{aligned}\int_\Gamma v ds &= \sum_{E\in\mathcal{E}_h}\int_0^1 v(F_E(t))|F_E'(t)|dt \qquad (2.7)\\ &\approx \sum_{E\in\mathcal{E}_h}\sum_{j=1}^{n_E}\omega_{j,E}|F_E'(t_{j,E})|v(x_j)\\ &= \sum_{j=1}^{n}\alpha_j v(x_j), \quad \alpha_j = \sum_{E\in\mathcal{E}_h, x_j\in\mathbb{T}_E}\omega_{j,E}|F_E'(t_{j,E})|.\end{aligned}$$

Since $\Delta t_{1,E} \le C\Delta t_{2,E}, \Delta t_{n_E+1,E} \le C\Delta t_{n_E,E}$, and $\Delta t_{j,E}/\Delta t_{k,E} \le C$ for any $j, k = 2, \cdots, n_E$, because the points in $\mathbb{T}$ are uniformly distributed, we have $\omega_{j,E} \sim 1/n_E \sim 1/(nh_E)$. This implies by (2.2) that there exist constants B_3, B_4 such that

$$B_3 n^{-1} \le \alpha_j \le B_4 n^{-1}, \ \ j = 1, 2, \cdots, n. \qquad (2.8)$$

Let $y_j, j = 1, 2, \cdots, J$, be the nodes of the mesh $\mathcal{M}_h$ on Γ. For any $v_h \in V_h$, we define $\Pi_h v_h \in Q_h$ such that $\Pi_h v_h(y_j) = v_h(y_j), j = 1, 2, \cdots, J$. For any $E \in \mathcal{E}_h$, let $\tilde{E}$ be the segment connecting two vertices of E and denote $F_{\tilde{E}} : \hat{E} \to \tilde{E}$ the affine mapping from the reference element $\hat{E}$ to $\tilde{E}$. Then $(\Pi_h v_h)(F_E(t)) = v_h(F_{\tilde{E}}(t)), \forall t \in \hat{E}$.

Now we are in the position to define the finite element solution for the problem (2.2)-(2.3). Given $f \in H^2(\Omega)$ and the observation g_i at x_i of the boundary value $g_0(x_i)$, $i = 1, 2, \cdots n$, find $(u_h, \lambda_h) \in V_h \times Q_h$ such that

$$\begin{aligned}(\nabla u_h, \nabla v_h)_{\Omega_h} + \langle\lambda_h, \Pi_h v_h\rangle_n &= (I_h f, v_h)_{\Omega_h}, \ \ \forall v_h \in V_h, \qquad (2.9)\\ \langle\mu_h, \Pi_h u_h\rangle_n &= \langle\mu_h, g\rangle_n, \ \ \forall \mu_h \in Q_h, \qquad (2.10)\end{aligned}$$

where $(\cdot,\cdot)_{\Omega_h}$ is the inner product of $L^2(\Omega_h)$ and $I_h : C(\bar{\Omega}) \to V_h$ is the standard Lagrange interpolation operator. The interpolation operator I_h can be replaced by the Clément interpolant [22] if the source f is less regular. We remark that the computation in (2.9)-(2.10) does not involve any geometric representation of the boundary Γ due to the introduction of the quadrature.

Following [88, 104] we introduce the following mesh-dependent Sobolev norms

$$\|v\|_{1/2,h}^2 = \sum_{E\in\mathcal{E}_h} h_E^{-1}\|v\|_{L^2(E)}^2, \ \ \|v\|_{-1/2,h}^2 = \sum_{E\in\mathcal{E}_h} h_E\|v\|_{L^2(E)}^2, \ \ \forall v \in L^2(\Gamma).$$

We use the following norms for functions $v_h \in V_h$, $\mu_h \in Q_h$

$$\|v_h\|_{V_h} = \left(\|\nabla v_h\|_{L^2(\Omega_h)}^2 + \|\Pi_h v_h\|_{1/2,h}^2\right)^{1/2}, \ \ \|\mu_h\|_{Q_h} = \|\mu_h\|_{-1/2,h}.$$

We consider now the well-posedness of the discrete problem (2.9)-(2.10) in the framework of Babuška-Brezzi theory. We start from the following simple lemma.

Lemma 2.2.2. *There exists a constant C such that*

$$|\langle 1, v\rangle - \langle 1, v\rangle_n| \le C \sum_{E\in\mathcal{E}_h} \int_0^1 \left(h_E|\hat{v}_E''| + h_E^2|\hat{v}_E'| + h_E^3|\hat{v}_E|\right) dt, \ \ \forall v \in W^{2,1}(\Gamma),$$

where $\hat{v}_E(t) = v|_E(F_E(t))$ for any $E \in \mathcal{E}_h$.

Proof. We first note that since Γ is smooth, we have $|F_E''(t)| \le Ch_E^2, |F_E'''(t)| \le Ch_E^3$ for any $E \in \mathcal{E}_h$. Since

$$|\langle 1, v\rangle - \langle 1, v\rangle_n| \le \sum_{E\in\mathcal{E}_h} \left| \int_E v ds - Q_{\mathbb{T}_E}(\hat{v}_E(t)|F_E'(t)|) \right|,$$

the lemma follows easily from Lemma 2.2.1 by taking $w = \hat{v}_E(t)|F_E'(t)|$ in each element $E \in \mathcal{E}_h$. We omit the details. □

Lemma 2.2.3. *Let $K_h = \{v_h \in V_h : \langle \Pi_h v_h, \mu_h\rangle_n = 0, \ \ \forall \mu_h \in Q_h\}$. There exists a constant $\alpha > 0$ independent of h, n such that*

$$(\nabla v_h, \nabla v_h)_{\Omega_h} \ge \alpha \|v_h\|_{V_h}^2, \ \ \forall v_h \in K_h.$$

Proof. For simplicity we write $\tilde{v}_h = \Pi_h v_h \in Q_h$ for any $v_h \in V_h$. Then $\langle \tilde{v}_h, \tilde{v}_h\rangle_n = 0$ for any $v_h \in K_h$. By Lemma 2.2.2 for $v = \tilde{v}_h^2$ we obtain after some simple computations

$$|\langle \tilde{v}_h, \tilde{v}_h\rangle - \langle \tilde{v}_h, \tilde{v}_h\rangle_n| \le Ch(\|\nabla v_h\|_{L^2(\Omega_h)}^2 + Ch^{1/2}\|\nabla v_h\|_{L^2(\Omega_h)}\|\tilde{v}_h\|_{L^2(\Gamma)}).$$

Thus

$$\begin{aligned} \|\tilde{v}_h\|_{1/2,h}^2 \le Ch^{-1}\|\tilde{v}_h\|_{L^2(\Gamma)}^2 &= Ch^{-1}|\langle \tilde{v}_h, \tilde{v}_h\rangle - \langle \tilde{v}_h, \tilde{v}_h\rangle_n| \\ &\le C\|\nabla v_h\|_{L^2(\Omega_h)}^2 + Ch\|\nabla v_h\|_{L^2(\Omega_h)}\|\tilde{v}_h\|_{1/2,h}. \end{aligned}$$

This shows $\|\nabla v_h\|_{L^2(\Omega_h)}^2 \ge C\|\tilde{v}_h\|_{1/2,h}^2$ and completes the proof. □

Lemma 2.2.4. *There exists constants $C_1, C_2 > 0, h_0 > 0$ independent of h, n such that for $h \le h_0$,*

$$C_1\|\mu_h\|_{L^2(\Gamma)} \le \|\mu_h\|_n \le C_2\|\mu_h\|_{L^2(\Gamma)}, \ \ \forall \mu_h \in Q_h.$$

Proof. Since $\hat{\mu}_h(t) = \mu_h(F_E(t))$ is linear in $\hat{E}$ for any $E \in \mathcal{E}_h$, we use Lemma 2.2.2 for $v = \mu_h^2$ to obtain

$$|\langle \mu_h, \mu_h\rangle - \langle \mu_h, \mu_h\rangle_n| \le C \sum_{E\in\mathcal{E}_h} \int_{\hat{E}} h_E|\hat{\mu}_h|^2 dt \le C\|\mu_h\|_{L^2(\Gamma)}^2.$$

This shows the right inequality. Next by definition we have

$$\langle \mu_h, \mu_h\rangle_n = \sum_{E\in\mathcal{E}_h} Q_{\mathbb{T}_E}(\hat{\mu}_h^2|F_E'|). \tag{2.11}$$

From (2.5) and (2.2) we know that for any $E \in \mathcal{E}_h$,

$$\begin{aligned}\tilde{Q}_{\mathbb{T}_E}(\hat{\mu}_h^2|F_E'|) &= \frac{1}{2}\sum_{j=0}^{n_E}\int_{t_j}^{t_{j+1}} \left(\hat{\mu}_h(t_j)^2|F_E'(t_j)| + \hat{\mu}_h^2(t_{j+1})|F_E'(t_{j+1})|\right) dt \\ &\geq Ch_E\sum_{j=0}^{n_E}\int_{t_j}^{t_{j+1}} \left(\hat{\mu}_h(t_j)^2 + \hat{\mu}_h(t_{j+1})^2\right) dt \\ &\geq Ch_E\sum_{j=0}^{n_E}\int_{t_j}^{t_{j+1}} |\hat{\mu}_h(t)|^2 dt,\end{aligned}$$

where in the last inequality we have used the fact that $\hat{\mu}_h$ is linear in $\hat{E}$ and the Jensen inequality for convex functions. Thus $|\tilde{Q}_{\mathbb{T}_E}(\hat{\mu}_h^2|F_E'|)| \geq C\|\mu_h\|^2_{L^2(E)}, \forall E \in \mathcal{E}_h$. On the other hand, by (2.6) we have

$$|Q_{\mathbb{T}_E}(\hat{\mu}_h^2|F_E'|) - \tilde{Q}_{\mathbb{T}_E}(\hat{\mu}_h^2|F_E'|)| \leq Ch_E\|\mu_h\|^2_{L^2(E)}.$$

Therefore, by (2.11), $\|\mu_h\|_n \geq C\|\mu_h\|_{L^2(\Gamma)}$ for sufficiently small h. This completes the proof. □

We have the following inf-sup condition for the empirical inner product.

Lemma 2.2.5. *There exists a constant $h_0, \beta > 0$ independent of h, n such that for $h \leq h_0$,*

$$\sup_{v_h \in V_h \setminus \{0\}} \frac{\langle \Pi_h v_h, \mu_h\rangle_n}{\|v_h\|_{V_h}} \geq \beta\|\mu_h\|_{Q_h}, \quad \forall \mu_h \in Q_h.$$

Proof. The proof follows an idea in [87] where the inf-sup condition for the bilinear form $\langle v_h, \mu_h\rangle$ is proved. Let $y_j, j = 1, 2, \cdots, J$, be the nodes of the mesh $\mathcal{M}_h$ on Γ and denote $\psi_j, j = 1, 2, \cdots, J$, the corresponding nodal basis function of V_h.

For any $\mu_h \in Q_h$, we define $v_h(x) = \sum_{j=1}^J \mu_h(y_j)\psi_j(x) \in V_h$. It is easy to check that

$$\|v_h\|^2_{V_h} \leq C\sum_{j=1}^J |\mu_h(y_j)|^2 \leq Ch^{-1}\|\mu_h\|^2_{L^2(\Gamma)}. \tag{2.12}$$

From the definition of $\Pi_h v_h \in Q_h$ we know that $\Pi_h v_h = \mu_h$ on Γ. Thus by Lemma 2.2.4,

$$\langle \Pi_h v_h, \mu_h\rangle_n = \|\mu_h\|_n^2 \geq C\|\mu_h\|^2_{L^2(\Gamma)}.$$

This completes the proof by using (2.12). □

By Lemma 2.2.4 we know that for any $v_h \in V_h$, $\mu_h \in Q_h$

$$|\langle \Pi_h v_h, \mu_h\rangle_n| \leq C\|\Pi_h v_h\|_{L^2(\Gamma)}\|\mu_h\|_{L^2(\Gamma)} \leq C\|v_h\|_{V_h}\|\mu_h\|_{Q_h}.$$

Now by the standard Babuška-Brezzi theory (cf., e.g., [16, Proposition 5.5.4]) we obtain the following theorem.

Theorem 2.2.1. *There exists a constant $h_0 > 0$ independent of h, n such that for any $h \leq h_0$, the discrete problem* (2.9)-(2.10) *has a unique solution $(u_h, \lambda_h) \in$*

$V_h \times Q_h$. Moreover, for any $(u_I, \lambda_I) \in V_h \times Q_h$, we have

$$\|u_h - u_I\|_{V_h} + \|\lambda_h - \lambda_I\|_{Q_h} \leq C \sum_{i=1}^{3} M_{ih},$$

where the errors M_{1h}, M_{2h}, M_{3h} are defined by

$$M_{1h} = \sup_{v_h \in V_h \setminus \{0\}} \frac{|(\nabla u_I, \nabla v_h)_{\Omega_h} + \langle \lambda_I, \Pi_h v_h \rangle_n - (I_h f, v_h)_{\Omega_h}|}{\|v_h\|_{V_h}},$$

$$M_{2h} = \sup_{\mu_h \in Q_h \setminus \{0\}} \frac{|\langle \mu_h, \Pi_h u_I - g_0 \rangle_n|}{\|\mu_h\|_{Q_h}}, \quad M_{3h} = \sup_{\mu_h \in Q_h \setminus \{0\}} \frac{|\langle \mu_h, e \rangle_n|}{\|\mu_h\|_{Q_h}}.$$

2.3 Convergence of the finite element method

We will use the Lenoir homeomorphism $\Phi_h : \Omega_h \to \Omega$ [70]. The mapping Φ_h is defined elementwise: for any $\tilde{K} \in \tilde{\mathcal{M}}_h$, $\Phi_h|_{\tilde{K}} = \Psi_K$ is a C^2-diffeomorphism from $\tilde{K}$ to K. If no edge of K belongs to $\partial\Omega_h$, $\Psi_K = I$, the identity. If one edge $\tilde{E}$ of $\tilde{K}$ lies on $\partial\Omega_h$ which corresponds to the curved edge E of $K \in \mathcal{M}_h$, Ψ_K maps $\tilde{E}$ onto E and $\Psi_K = I$, the identity, along the other two edges of $\tilde{K}$. We need the following properties of Ψ_K from [70] in the following lemma.

Lemma 2.3.1. *The following assertions are valid for any $\tilde{K} \in \tilde{\mathcal{M}}_h$ and $K \in \mathcal{M}_h$.*
1° The mapping $\Psi_K : \tilde{K} \to K$ satisfies the following estimates

$$\|D^s(\Psi_K - I)\|_{L^\infty(\tilde{K})} \leq Ch^{2-s}, \ \forall s \leq 2, \quad \sup_{x \in \tilde{K}} |J(\Psi_K)(x) - 1| \leq Ch,$$

where $J(\Psi_K)$ denotes the modulus of the Jacobi determinant of Ψ_K.
2° The mapping $\Psi_K^{-1} : K \to \tilde{K}$ satisfies

$$\|D^s(\Psi_K^{-1} - I)\|_{L^\infty(K)} \leq Ch^{2-s}, \ \forall s \leq 2, \quad \sup_{x \in K} |J(\Psi_K^{-1})(x) - 1| \leq Ch.$$

Let $r_h : L^1(\Omega_h) \to V_h$ be the Clément interplant [21] which enjoys the following properties

$$|v - r_h v|_{H^j(\tilde{K})} \leq Ch^{m-j} |v|_{H^m(\Delta_{\tilde{K}})}, \ \forall \tilde{K} \in \tilde{\mathcal{M}}_h, 0 \leq j \leq m, m = 1, 2, \quad (2.1)$$

$$|v - r_h v|_{H^j(e)} \leq Ch^{m-j-1/2} |v|_{H^m(\Delta_e)}, \ \forall e \in \tilde{\mathcal{E}}_h, 0 \leq j < m, m = 1, 2, \quad (2.2)$$

where $\tilde{\mathcal{E}}_h$ is the set of all sides of the mesh $\tilde{\mathcal{M}}_h$, and for any set $A \subset \Omega_h$, Δ_A is the union of the elements surrounding A. We remark that (2.1) is proved in [21] and (2.2) is the consequence of (2.1) and the following scaled trace inequality

$$|v|_{L^2(e)} \leq Ch^{-1/2} \|v\|_{L^2(\Delta_e)} + Ch^{1/2} \|\nabla v\|_{L^2(\Delta_e)}, \ \forall v \in H^1(\Omega_h).$$

We will assume in this section that the solution $u \in H^2(\Omega)$ and thus $\lambda \in H^{1/2}(\Gamma)$. By the trace theorem, there exists a function $\tilde{\lambda} \in H^1(\Omega)$ such that $\tilde{\lambda} = \lambda$ on Γ and $\|\tilde{\lambda}\|_{H^1(\Omega)} \leq C \|\lambda\|_{H^{1/2}(\Gamma)}$. Now we define the following

interpolation operator $R_h : L^2(\Omega) \to L^2(\Omega)$

$$R_h v = [r_h(v \circ \Phi_h)] \circ \Phi_h^{-1}, \ \ \forall v \in L^2(\Omega).$$

We notice that similar interpolation functions are used in [70] where the Clément interpolation operator is replaced by the Lagrangian interpolation operator. The following theorem can be easily proved by using Lemma 2.3.1 and (2.1)-(2.2).

Lemma 2.3.2. *For any $v \in H^2(\Omega)$, we have $\|v - R_h v\|_{H^j(\Omega)} \le Ch^{m-j}\|v\|_{H^m(\Omega)}$, $\|v - R_h v\|_{H^j(\Gamma)} \le Ch^{m-j-1/2}\|v\|_{H^m(\Omega)}$, $0 \le j \le m-1, m = 1, 2$.*

For any $v_h \in V_h$, we denote $\check{v}_h = v_h \circ \Phi_h^{-1}$ which is a function defined in Ω. Let $\Omega^* = \cup_{K \in \mathcal{M}_h^*} K$, where $\mathcal{M}_h^*$ is the set of all elements having one curved edge. Obviously, $|\Omega^*| \le Ch$. By definition $\Phi_h = \Psi_K$ is identity for $K \in \mathcal{M}_h \backslash \mathcal{M}_h^*$. Then it is easy to check by using Lemma 2.3.1 that (cf. [70, Lemma 8]) for any $v_h, w_h \in V_h$,

$$|(\nabla v_h, \nabla w_h)_{\Omega_h} - (\nabla \check{v}_h, \nabla \check{w}_h)| \le Ch\|\check{v}_h\|_{H^1(\Omega^*)}\|\check{w}_h\|_{H^1(\Omega^*)}. \tag{2.3}$$

Now by the Poincáre inequality, it is easy to see that $\|v\|_{L^2(\Omega)} \le C\|\nabla v\|_{L^2(\Omega)} + C\|v\|_{1/2,h}, \forall v \in H^1(\Omega)$. Thus by (2.3)

$$\|\check{v}_h\|_{H^1(\Omega)} \le \|\nabla \check{v}_h\|_{L^2(\Omega)} + C\|\check{v}_h\|_{1/2,h} \le C\|v_h\|_{V_h} + Ch^{1/2}\|\check{v}_h\|_{H^1(\Omega)},$$

which implies, for sufficiently small h,

$$\|\check{v}_h\|_{H^1(\Omega)} \le C\|v_h\|_{V_h}, \ \ \forall v_h \in V_h. \tag{2.4}$$

Lemma 2.3.3. *Let $(u, \lambda) \in H^2(\Omega) \times H^{1/2}(\Gamma)$ be the solution of (2.2)-(2.3). We have*

$$\|u - u_h \circ \Phi_h^{-1}\|_{H^1(\Omega)} + \|\lambda - \lambda_h\|_{-1/2,h} \le Ch\|u\|_{H^2(\Omega)} + \sum_{i=1}^{3} M_{ih},$$

where M_{ih}, $i = 1, 2, 3$, are defined in Theorem 2.2.1 with $u_I = r_h(u \circ \Phi_h) \in V_h$ and $\lambda_I = R_h\tilde{\lambda} \in Q_h$.

Proof. We first observe that by Lemma 2.3.2

$$\|\lambda - \lambda_I\|_{L^2(\Gamma)} \le Ch^{1/2}\|\tilde{\lambda}\|_{H^1(\Omega)} \le Ch^{1/2}\|\lambda\|_{H^{1/2}(\Gamma)} \le Ch^{1/2}\|u\|_{H^2(\Omega)}.$$

Notice that $\check{u}_h = R_h u$, we obtain by Lemma 2.3.2, (2.4), and Theorem 2.2.1 that

$$\begin{aligned} & \|u - u_h \circ \Phi_h^{-1}\|_{H^1(\Omega)} + \|\lambda - \lambda_h\|_{-1/2,h} \\ \le \ & \|u - R_h u\|_{H^1(\Omega)} + \|\lambda - R_h\tilde{\lambda}\|_{-1/2,h} + C(\|u_h - u_I\|_{V_h} + \|\lambda_h - \lambda_I\|_{Q_h}) \\ \le \ & Ch\|u\|_{H^2(\Omega)} + C\sum_{i=1}^{3} M_{ih}. \end{aligned}$$

This completes the proof. □

Lemma 2.3.4. *We have $M_{1h} \le Ch|\ln h|^{1/2}(\|u\|_{H^2(\Omega)} + \|f\|_{H^2(\Omega)})$.*

Proof. We first note that by (2.2) we have

$$(\nabla u, \nabla \check{v}_h) + \langle \lambda, \check{v}_h \rangle = (f, \check{v}_h), \ \ \forall v_h \in V_h.$$

Now since $\Pi_h v_h = \check{v}_h$ on Γ, for any $v_h \in V_h$, we have

$$\begin{aligned} & |(\nabla u_I, \nabla v_h)_{\Omega_h} + \langle \lambda_I, \Pi_h v_h \rangle_n - (I_h f, v_h)_{\Omega_h}| \\ \leq \ & |(f, \check{v}_h) - (I_h f, v_h)_{\Omega_h}| + |(\nabla u_I, \nabla v_h)_{\Omega_h} - (\nabla u, \nabla \check{v}_h)| + |\langle \lambda, \check{v}_h \rangle - \langle \lambda_I, \check{v}_h \rangle_n|. \end{aligned}$$

Since $\Phi_h = \Psi_K$ is identity for $K \in \mathcal{M}_h \backslash \mathcal{M}_h^*$, we have

$$(f, \check{v}_h) - (I_h f, v_h)_{\Omega_h} = \sum_{K \in \mathcal{M}_h^*} \int_{\tilde{K}} ((f \circ \Psi_K) v_h J(\Psi_K) - I_h (f \circ \Psi_K) v_h) dx,$$

which implies by using Lemma 2.3.1 that

$$\begin{aligned} & |(f, \check{v}_h) - (I_h f, v_h)_{\Omega_h}| \\ \leq \ & Ch \|f\|_{L^2(\Omega^*)} \|\check{v}_h\|_{L^2(\Omega^*)} + Ch^2 \|f\|_{H^2(\Omega)} \|v_h\|_{L^2(\Omega)}. \end{aligned}$$

Obviously, $\|f\|_{L^2(\Omega^*)} \leq Ch^{1/2} \|f\|_{L^\infty(\Omega)} \leq Ch^{1/2} \|f\|_{H^2(\Omega)}$. Moreover, by the well-known embedding theorem [91]

$$\|v\|_{L^p(\Omega)} \leq C p^{1/2} \|v\|_{H^1(\Omega)}, \forall v \in H^1(\Omega), \forall p > 2,$$

we have

$$\|v\|_{L^2(\Omega^*)} \leq C |\Omega^*|^{\frac{1}{2} - \frac{1}{p}} p^{1/2} \|v\|_{H^1(\Omega)} \leq C h^{\frac{1}{2} - \frac{1}{p}} p^{1/2} \|v\|_{H^1(\Omega)}.$$

By taking $p = \ln(h^{-1})$ we obtain then

$$\|v\|_{L^2(\Omega^*)} \leq Ch^{1/2} |\ln h|^{1/2} \|v\|_{H^1(\Omega)}, \ \ \forall v \in H^1(\Omega). \tag{2.5}$$

This implies

$$|(f, \check{v}_h) - (I_h f, v_h)_{\Omega_h}| \leq Ch^2 |\ln h|^{1/2} \|f\|_{H^2(\Omega)} \|v_h\|_{V_h}. \tag{2.6}$$

By Lemma 2.3.2, (2.3) and (2.4) we have

$$\begin{aligned} & |(\nabla u_I, \nabla v_h)_{\Omega_h} - (\nabla u, \nabla \check{v}_h)| \\ \leq \ & |(\nabla u_I, \nabla v_h)_{\Omega_h} - (\nabla \check{u}_I, \nabla \check{v}_h)| + |(\nabla (u - \check{u}_I), \nabla \check{v}_h)| \\ \leq \ & Ch \|u\|_{H^2(\Omega)} \|v_h\|_{V_h}. \end{aligned} \tag{2.7}$$

By using Lemma 2.2.2 one can prove

$$|\langle \check{v}_h, \check{w}_h \rangle_n - \langle \check{v}_h, \check{w}_h \rangle| \leq Ch \|v_h\|_{H^1(\Omega_h)} \|w_h\|_{H^1(\Omega_h)}, \ \ \forall v_h, w_h \in V_h. \tag{2.8}$$

Thus

$$\begin{aligned} |\langle \lambda_I, \check{v}_h \rangle_n - \langle \lambda_I, \check{v}_h \rangle| \ & \leq \ Ch \|r_h(\tilde{\lambda} \circ \Phi_h)\|_{H^1(\Omega_h)} \|v_h\|_{H^1(\Omega_h)} \\ & \leq \ Ch \|u\|_{H^2(\Omega)} \|v_h\|_{V_h}, \end{aligned}$$

which implies by using Lemma 2.3.2 that

$$|\langle \lambda, \check{v}_h\rangle - \langle \lambda_I, \check{v}_h\rangle_n| \le Ch\|u\|_{H^2(\Omega)}\|v_h\|_{V_h}. \tag{2.9}$$

The estimate for M_{1h} now follows from (2.6), (2.7) and (2.9). □

Lemma 2.3.5. *We have* $M_{2h} \le Ch\|u\|_{H^2(\Omega)}$.

Proof. We first we observe that the argument in the proof of Lemma 2.2.1 implies that

$$\left|\int_0^1 w(t)dt - Q_{\mathbb{T}_E}(w)\right| \le C\|w'\|_{L^2(\hat{E})}, \ \ \forall w \in H^1(\hat{E}).$$

For any $v \in H^1(\Gamma)$, by taking $w(t) = \hat{v}_E(t)|F_E'(t)|$ in each element $E \in \mathcal{E}_h$, where $\hat{v}_E(t) = v|_E(F_E(t))$, we know that

$$|\langle 1, v\rangle - \langle 1, v\rangle_n| \le C\sum_{E\in\mathcal{E}_h}(h_E\|\hat{v}_E'\|_{L^2(\hat{E})} + h_E^2\|\hat{v}_E\|_{L^2(\hat{E})}).$$

We use the above inequality for $v = \mu_h\varphi$, where $\varphi = u - \check{u}_I$ in Γ, to obtain

$$|\langle \mu_h, \varphi\rangle - \langle \mu_h, \varphi\rangle_n| \le C\sum_{E\in\mathcal{E}_h}\|\hat{\mu}_h\|_{L^2(\hat{E})}(h_E\|\hat{\varphi}_E\|_{L^2(\hat{E})} + h_E^2\|\hat{\varphi}_E'\|_{L^2(\hat{E})}),$$

where we have used the fact $\|\hat{\mu}_h\|_{W^{1,\infty}(\hat{E})} \le C\|\hat{\mu}_h\|_{L^2(\hat{E})}$ since $\hat{\mu}_h \in P_1(\hat{E})$. This implies by using Lemma 2.3.2 again

$$\begin{aligned} & |\langle \mu_h, u - \check{u}_I\rangle - \langle \mu_h, u - \check{u}_I\rangle_n| \\ \le\ & C\|\mu_h\|_{L^2(\Gamma)}(\|u - R_hu\|_{L^2(\Gamma)} + h|u - R_hu|_{H^1(\Gamma)}) \\ \le\ & Ch^{3/2}\|u\|_{H^2(\Omega)}\|\mu_h\|_{L^2(\Gamma)}. \end{aligned}$$

This completes the proof. □

The following theorem shows the convergence of the finite element solution in the sense of expectation.

Theorem 2.3.1. *We have*

$$\begin{aligned} & \mathbb{E}\left[\|u - u_h\circ\Phi_h^{-1}\|_{H^1(\Omega)} + h^{1/2}\|\lambda - \lambda_h\|_{L^2(\Gamma)}\right] \\ \le\ & Ch|\ln h|^{1/2}(\|u\|_{H^2(\Omega)} + \|f\|_{H^2(\Omega)}) + Ch^{-1}(\sigma n^{-1/2}). \end{aligned}$$

Proof. By Lemmas 2.3.3-2.3.5 we are left to estimate $\mathbb{E}[M_{3h}]$. We first observe that

$$\mathbb{E}\left[\sup_{\mu_h\in Q_h\setminus\{0\}}\frac{|\langle \mu_h, e\rangle_n|^2}{\|\mu_h\|_{Q_h}^2}\right] \le Ch^{-1}\mathbb{E}\left[\sup_{\mu_h\in Q_h\setminus\{0\}}\frac{|\langle \mu_h, e\rangle_n|^2}{\|\mu_h\|_{L^2(\Gamma)}^2}\right]. \tag{2.10}$$

Let N_h be the dimension of Q_h and let $\{\psi_j\}_{j=1}^{N_h}$ be the orthonormal basis of Q_h in the $L^2(\Gamma)$ inner product. Then for any $\mu_h = \sum_{j=1}^{N_h}(\mu_h, \psi_j)\psi_j$, by

Cauchy-Schwarz inequality and (2.8)

$$|\langle \mu_h, e\rangle_n|^2 \le \frac{C}{n^2}\|\mu_h\|^2_{L^2(\Gamma)} \sum_{j=1}^{N_h}\left(\sum_{i=1}^{n} e_i\psi_j(x_i)\right)^2.$$

Since $e_i, i = 1, 2, \cdots, n$, are independent and identically random variables, we have

$$\mathbb{E}\left[\sup_{\mu_h\in Q_h\setminus\{0\}} \frac{|\langle \mu_h, e\rangle_n|^2}{\|\mu_h\|^2_{L^2(\Gamma)}}\right] \le C\frac{\sigma^2}{n^2}\sum_{j=1}^{N_h}\sum_{i=1}^{n}\psi_j(x_i)^2.$$

Since the number of measurement points in E, $\#\mathbb{T}_E \le Cnh_E$ and $N_h \le Ch^{-1}$, we obtain by using the inverse estimate that

$$\sum_{j=1}^{N_h}\sum_{i=1}^{n}\psi_j(x_i)^2 \le Cnh\sum_{j=1}^{N_h}\sum_{E\in\mathcal{E}_h}\|\psi_j\|^2_{L^\infty(E)} \le CN_hn \le Cnh^{-1}.$$

Therefore

$$\mathbb{E}\left[\sup_{\mu_h\in Q_h\setminus\{0\}} \frac{|\langle \mu_h, e\rangle_n|^2}{\|\mu_h\|^2_{L^2(\Gamma)}}\right] \le Ch^{-1}(\sigma^2 n^{-1}). \tag{2.11}$$

This, together with (2.10), yields

$$\mathbb{E}\left[\sup_{\mu_h\in Q_h\setminus\{0\}} \frac{|\langle \mu_h, e\rangle_n|^2}{\|\mu_h\|^2_{Q_h}}\right] \le Ch^{-2}(\sigma^2 n^{-1}),$$

which completes the proof. □

The following lemma will be useful in deriving the improved estimate for $\|u - u_h\circ\Phi_h^{-1}\|_{L^2(\Omega)}$.

Lemma 2.3.6. *We have*

$$\mathbb{E}\left[\sup_{\mu_h\in Q_h\setminus\{0\}} \frac{|\langle e, \mu_h\rangle_n|}{\|\mu_h\|_{H^{1/2}(\Gamma)}}\right] \le C|\ln h|(\sigma n^{-1/2}).$$

Proof. Let $h_0 = h \le 1$ and $h_i = h^{(p+1-i)/(p+1)}$ for $1 \le i \le p$, where $p \ge 1$ is an integer to be determined later. Obviously $h_i \le h_{i+1}$, $0 \le i \le p$. Let $\mathcal{E}_{h_i}$ be a uniform mesh over the boundary Γ and Q_{h_i} the finite element space defined in (2.3) over the mesh Q_{h_i}. Let $\{y^k_{h_i}\}_{k=1}^{N_{h_i}}$ be the nodes of the mesh $\mathcal{E}_{h_i}, i = 0, \cdots, p$. We introduce the following Clément-type interpolation operator $\pi_{h_i} : L^1(\Gamma) \to Q_{h_i}$ such that for any $v \in L^1(\Gamma)$,

$$(\pi_{h_i}v)(y^k_{h_i}) = \frac{1}{|S(y^k_{h_i})|}\int_{S(y^k_{h_i})} v(x)ds(x), \;\; 1 \le k \le N_{h_i},$$

where $S(y^k_{h_i})$ is the union of the two elements sharing the common node $y^k_{h_i}$. It is easy to show by scaling argument that

$$\|v - \pi_{h_i}v\|_{L^2(\Gamma)} \le h_i^m\|v\|_{H^m(\Gamma)}, \;\; \forall v \in H^1(\Gamma), m = 0, 1.$$

Thus by the theory of real interpolation of Sobolev spaces, e.g., [17, Proposition 12.1.5],

$$\|v - \pi_{h_i} v\|_{L^2(\Gamma)} \leq C h_i^{1/2} \|v\|_{H^{1/2}(\Gamma)}, \quad \forall v \in H^{1/2}(\Gamma). \tag{2.12}$$

Now we introduce the telescope sum

$$\mu_h = \sum_{i=0}^{p-1} (\mu_{h_i} - \mu_{h_{i+1}}) + \mu_{h_p}, \quad \forall \mu_h \in Q_h = Q_{h_0}, \tag{2.13}$$

where $\mu_{h_i} = \pi_{h_i} \mu_h \in Q_{h_i}$, $0 \leq i \leq p+1$. By (2.12)

$$\|\mu_{h_i} - \mu_{h_{i+1}}\|_{L^2(\Gamma)} \leq C h_{i+1}^{1/2} \|\mu_h\|_{H^{1/2}(\Gamma)}. \tag{2.14}$$

Then the same argument in proving (2.11) implies

$$\begin{aligned}
\mathbb{E}\left[\sup_{\mu_h \in Q_h \setminus \{0\}} \frac{|\langle e, \mu_{h_i} - \mu_{h_{i+1}} \rangle_n|}{\|\mu_h\|_{H^{1/2}(\Gamma)}} \right] &\leq C h_{i+1}^{1/2} h_i^{-1/2} (\sigma n^{-1/2}), \\
\mathbb{E}\left[\sup_{\mu_h \in Q_h \setminus \{0\}} \frac{|\langle e, \mu_{h_p} \rangle_n|}{\|\mu_h\|_{H^{1/2}(\Gamma)}} \right] &\leq C h_p^{-1/2} (\sigma n^{-1/2}).
\end{aligned}$$

By (2.13) we then obtain

$$\mathbb{E}\left[\sup_{\mu_h \in Q_h \setminus \{0\}} \frac{|\langle e, \mu_h \rangle_n|}{\|\mu_h\|_{H^{1/2}(\Gamma)}} \right] \leq C(p+1) h^{-\frac{1}{2(p+1)}} (\sigma n^{-1/2}).$$

This completes the proof by taking the integer p such that $p < |\ln h| \leq p+1$. □

Theorem 2.3.2. *We have*

$$\begin{aligned}
& \mathbb{E}\left[\|u - u_h \circ \Phi_h^{-1}\|_{L^2(\Omega)} \right] \\
\leq \;& C h^2 |\ln h| (\|u\|_{H^2(\Omega)} + \|f\|_{H^2(\Omega)} + \|g_0\|_{H^2(\Gamma)}) + C |\ln h| (\sigma n^{-1/2}).
\end{aligned}$$

Proof. Let $(w, p) \in H^1(\Omega) \times H^{-1/2}(\Gamma)$ be the solution of the following problem

$$\begin{aligned}
(\nabla w, \nabla v) + \langle p, v \rangle &= (u - \check{u}_h, v), \quad \forall v \in H^1(\Omega), && (2.15) \\
\langle \mu, w \rangle &= 0, \quad \forall \mu \in H^{-1/2}(\Gamma). && (2.16)
\end{aligned}$$

By the regularity theory of elliptic equations, $(w, p) \in H^2(\Omega) \times H^1(\Omega)$ and satisfies

$$\|w\|_{H^2(\Omega)} + \|p\|_{H^1(\Omega)} \leq C \|u - \check{u}_h\|_{L^2(\Omega)}. \tag{2.17}$$

Let $w_I = I_h(w \circ \Phi_h) \in V_h$ be the Lagrange interpolation of $w \in H^2(\Omega)$ and $p_I = r_h(p \circ \Phi_h) \in V_h$ be the Clément interpolation of $p \in H^1(\Omega)$. By (2.16) we know that $w = 0$ on Γ and consequently, $w_I = 0$ on Γ_h, $\check{w}_I = w_I \circ \Phi_h^{-1} = 0$

on Γ. Now by using (2.2)-(2.3), (2.9)-(2.10) we obtain

$$\begin{aligned} & \|u-\check{u}_h\|^2_{L^2(\Omega)} \\ = \;& (\nabla(w-\check{w}_I),\nabla(u-\check{u}_h)) + \langle p-\check{p}_I, u-\check{u}_h\rangle \\ + \;& [(f,\check{w}_I) - (I_h f,\check{w}_I)_{\Omega_h}] + [(\nabla w_I,\nabla u_h)_{\Omega_h} - (\nabla\check{w}_I,\nabla\check{u}_h)] \\ + \;& [\langle\check{p}_I, u-\check{u}_h\rangle - \langle\check{p}_I, u-\check{u}_h\rangle_n] - \langle\check{p}_I, e\rangle_n \\ := \;& \mathrm{I}+\cdots+\mathrm{VI}. \end{aligned} \tag{2.18}$$

By Lemma 2.3.1 and (2.17) we have

$$|\mathrm{I}| + |\mathrm{II}| \le Ch\|u-\check{u}_h\|_{H^1(\Omega)}\|u-\check{u}_h\|_{L^2(\Omega)}. \tag{2.19}$$

By (2.6) and (2.17)

$$\begin{aligned} |\mathrm{III}| &\le Ch^2|\ln h|^{1/2}\|f\|_{H^2(\Omega)}\|w_I\|_{V_h} \\ &\le Ch^2|\ln h|^{1/2}\|f\|_{H^2(\Omega)}\|u-\check{u}_h\|_{L^2(\Omega)}. \end{aligned} \tag{2.20}$$

Since $\Phi_h|_K = I$ for $K\in\mathcal{M}_h\backslash\mathcal{M}_h^*$, by (2.3), Lemma 2.3.1 and (2.17) we have

$$|\mathrm{IV}| \le Ch\|\check{w}_I\|_{H^1(\Omega^*)}\|\check{u}_h\|_{H^1(\Omega^*)}.$$

Now by using (2.5), Lemma 2.3.1, and (2.17), we have

$$\begin{aligned} &\|\check{u}_h\|_{H^1(\Omega^*)} \le \|u-\check{u}_h\|_{H^1(\Omega)} + Ch^{1/2}|\ln h|^{1/2}\|u\|_{H^2(\Omega)}, \\ &\|\check{w}_I\|_{H^1(\Omega^*)} \le Ch\|u-\check{u}_h\|_{L^2(\Omega)} + Ch^{1/2}|\ln h|^{1/2}\|u-\check{u}_h\|_{L^2(\Omega)}. \end{aligned}$$

This implies

$$|\mathrm{IV}| \le C\left[h\|u-\check{u}_h\|_{H^1(\Omega)} + h^2|\ln h|\|u\|_{H^2(\Omega)}\right]\|u-\check{u}\|_{L^2(\Omega)}. \tag{2.21}$$

To estimate the term V we first use the triangle inequality

$$|\mathrm{V}| \le |\langle\check{p}_I, u-\check{u}_I\rangle - \langle\check{p}_I, u-\check{u}_I\rangle_n| + |\langle\check{p}_I, u_I-\check{u}_h\rangle - \langle\check{p}_I,\check{u}_I-\check{u}_h\rangle_n|$$

By using Lemma 2.2.2 for $v=\check{p}_I(u-\check{u}_I)$ one obtains easily

$$\begin{aligned} |\langle\check{p}_I, u-\check{u}_I\rangle - \langle\check{p}_I, u-\check{u}_I\rangle_n| &\le Ch^2\|g_0\|_{H^2(\Gamma)}\|\check{p}_I\|_{L^2(\Gamma)} \\ &\le Ch^2\|g_0\|_{H^2(\Gamma)}\|u-\check{u}_h\|_{L^2(\Omega)}, \end{aligned}$$

where we have used the estimate $\|\check{p}_I\|_{L^2(\Gamma)} \le C\|p\|_{H^1(\Omega)} \le C\|u-\check{u}_h\|_{L^2(\Omega)}$. By (2.8) and (2.17) we have

$$\begin{aligned} |\langle\check{p}_I, u_I-\check{u}_h\rangle - \langle\check{p}_I,\check{u}_I-\check{u}_h\rangle_n| &\le Ch\|p_I\|_{H^1(\Omega_h)}\|u_I-u_h\|_{H^1(\Omega_h)} \\ &\le Ch\|u_I-u_h\|_{H^1(\Omega_h)}\|u-\check{u}_h\|_{L^2(\Omega)}. \end{aligned}$$

Thus

$$|\mathrm{V}| \le Ch^2(h^{-1}\|u-\check{u}_h\|_{H^1(\Omega)} + \|u\|_{H^2(\Omega)} + \|g_0\|_{H^2(\Gamma)})\|u-\check{u}_h\|_{L^2(\Omega)} \tag{2.22}$$

By inserting (2.19)-(2.22) into (2.18) we finally obtain that

$$\begin{aligned} \|u-\check{u}_h\|_{L^2(\Omega)} &\leq Ch^2|\ln h|\left[\|u\|_{H^2(\Omega)}+\|f\|_{H^2(\Omega)}+\|g_0\|_{H^2(\Gamma)}\right] \\ &+ Ch\|u-\check{u}_h\|_{H^1(\Omega)}+\sup_{\mu_h\in Q_h\setminus\{0\}}\frac{|\langle e,\mu_h\rangle_n|}{\|\mu_h\|_{H^{1/2}(\Gamma)}}. \end{aligned} \tag{2.23}$$

The lemma now follows from Theorem 2.3.1 and Lemma 2.3.6. □

2.4 Sub-Gaussian random errors

In this section, we will study the convergence of our finite element method when the random errors added to the boundary data are sub-Gaussian. We will use the theory of empirical processes [111, 112].

Definition 2.4.1. *A random variable X is called sub-Gaussian with parameter σ if*

$$\mathbb{E}[e^{\lambda(X-\mathbb{E}[X])}]\leq e^{\sigma^2\lambda^2/2},\ \ \forall\ \lambda\in\mathbf{R}.$$

The following definition on the Orilicz ψ_2-norm will be used in our analysis.

Definition 2.4.2. *Let $\psi_2=e^{x^2}-1$ and X be a random variable. The ψ_2 norm of X is defined as*

$$\|X\|_{\psi_2}=\inf\left\{C>0:\ \mathbb{E}\left[\psi_2\left(\frac{|X|}{C}\right)\right]\leq 1\right\}.$$

It is known that [112, Lemma 2.2.1] if $\|X\|_{\psi_2}\leq K$, then

$$\mathbb{P}(|X|>z)\leq 2\exp\left(-\frac{z^2}{K^2}\right),\ \forall\ z>0. \tag{2.1}$$

Inversely, if

$$\mathbb{P}(|X|>z)\leq C\exp\left(-\frac{z^2}{K^2}\right),\ \forall\ z>0, \tag{2.2}$$

then $\|X\|_{\psi_2}\leq\sqrt{1+C}K$.

Definition 2.4.3. *Let (T,d) be a semi-metric space, a stochastic process $\{X_t:\ t\in T\}$ is called a sub-Gaussian process with respect to the semi-metric d, if*

$$\mathbb{P}(|X_t-X_s|>z)\leq 2\exp\left(-\frac{1}{2}\frac{z^2}{d^2(t,s)}\right),\ \forall\ s,t\in T,\ z>0.$$

For a semi-metric space (T,d), the covering number $N(\varepsilon,T,d)$ is the minimum number of ε-balls that covers T. A set is called ε-separated if the distance of any two points in the set is strictly greater than ε. The packing number $D(\varepsilon,T,d)$ is the maximum number of ε-separated points in T. It is easy to check that [112, P.98]

$$N(\varepsilon,T,d)\leq D(\varepsilon,T,d)\leq N(\frac{\varepsilon}{2},T,d). \tag{2.3}$$

The following maximum inequality can be found in [112, Section 2.2.1].

Lemma 2.4.1. *If $\{X_t : t \in T\}$ is a separable sub-Gaussian process with respect to the semi-metric d, then*

$$\| \sup_{s,t\in T} |X_t - X_s| \|_{\psi_2} \le K \int_0^{\operatorname{diam} T} \sqrt{D(\varepsilon, T, d)}\, d\varepsilon.$$

Here $K > 0$ is some constant.

The following lemma provides the estimate of the covering number for finite dimensional subsets [32, Corollary 2.6].

Lemma 2.4.2. *Let G be a finite dimensional subspace of $L^2(D)$ of dimension $N > 0$ and $G_R = \{f \in G : \|f\|_{L^2(D)} \le R\}$. Then*

$$N(\varepsilon, G_R, \|\cdot\|_{L^2(D)}) \le (1 + 4R/\varepsilon)^N, \ \ \forall \varepsilon > 0.$$

Theorem 2.4.1. *We have*

$$\begin{aligned} & \| \|u - u_h \circ \Phi_h^{-1}\|_{H^1(\Omega)} \|_{\psi_2} + h^{1/2} \| \|\lambda - \lambda_h\|_{L^2(\Gamma)} \|_{\psi_2} \\ \le \ & Ch |\ln h|^{1/2} (\|u\|_{H^2(\Omega)} + \|f\|_{H^2(\Omega)}) + Ch^{-1}(\sigma h^{-1/2}). \end{aligned}$$

Proof. By Lemmas 2.3.3-2.3.5 we are left to estimate $\|M_{3h}\|_{\psi_2}$. Let $F_h = \{\mu_h \in Q_h : \ \|\mu_h\|_{L^2(\Gamma)} \le 1\}$, then

$$\|M_{3h}\|_{\psi_2} \le h^{-1/2} \| \sup_{\mu_h \in F_h} \langle |\mu_h, e\rangle_n| \|_{\psi_2}. \tag{2.4}$$

For any $\mu_h \in F_h$, denote by $E_n(\mu_h) = \langle \mu_h, e\rangle_n$. Then $E_n(\mu_h) - E_n(\mu_h') = \sum_{i=1}^n c_i e_i$, where $c_i = \alpha_i(\mu_h - \mu_h')(x_i)$, $i = 1, 2, \cdots, n$. For any $\lambda > 0$, since $\alpha_i \le B_4 n^{-1}$ by (2.8),

$$\mathbb{E}\left[e^{\lambda \sum_{i=1}^n c_i e_i}\right] \ \le \ e^{\frac{1}{2}\lambda^2\sigma^2 \sum_{i=1}^n c_i^2} \le e^{\frac{1}{2} B_4 \lambda^2 \sigma^2 n^{-1} \|\mu_h - \mu_h'\|_n^2} = e^{\frac{1}{2}\sigma_1^2 \lambda^2},$$

where $\sigma_1 = B_4 \sigma n^{-1/2} \|\mu_h - \mu_h'\|_n$. Thus $E_n(\mu_h) - E_n(\mu_h')$ is a sub-Gaussian process with the parameter σ_1. This implies by (2.1) that

$$\mathbb{P}(|E_n(\mu_h - \mu_h')| > z) \le 2e^{-z^2/2\sigma_1^2}, \ \ \forall z > 0.$$

Thus $E_n(\mu_h)$ is a sub-Gaussian random process with respect to the semi-distance $d(\mu_h, \mu_h') = \|\mu_h - \mu_h'\|_n^*$, where $\|\mu_h\|_n^* = B_4 \sigma n^{-1/2} \|\mu_h\|_n$.

By Lemma 2.2.4 we know that the diameter of F_h in terms of the semi-distance d is bounded by $2C_2 B_4 (\sigma n^{-1})$. By maximal inequality in Lemma 2.4.1 and (2.3) we have

$$\begin{aligned} \| \sup_{\mu_h \in F_h} |\langle \mu_h, e\rangle_n| \|_{\psi_2} \ &\le \ K \int_0^{2C_2 B_4 \sigma n^{-1/2}} \sqrt{\log N(\frac{\varepsilon}{2}, F_h, \|\cdot\|_n^*)}\, d\varepsilon \\ &= \ K \int_0^{2C_2 B_4 \sigma n^{-1/2}} \sqrt{\log N(\frac{\varepsilon}{2B_4 \sigma n^{-1/2}}, F_h, \|\cdot\|_n)}\, d\varepsilon. \end{aligned}$$

By Lemma 2.2.4 and Lemma 2.4.2 we know that for any $\delta > 0$,

$$N(\delta, F_h, \|\cdot\|_n) \le N(C_1^{-1}\delta, F_h, \|\cdot\|_{L^2(\Gamma)}) \le (1 + 4C_1/\delta)^{N_h},$$

where N_h is the dimension of Q_h which is bounded by Ch^{-1}. Therefore,

$$\begin{aligned} \| \sup_{\mu_h \in F_h} |\langle \mu_h, e\rangle_n| \|_{\psi_2} &\leq Ch^{-1/2} \int_0^{2C_2 B_4 \sigma n^{-1/2}} \sqrt{\log\left(1 + \frac{C\sigma n^{-1/2}}{\varepsilon}\right)}\, d\varepsilon \\ &\leq Ch^{-1/2}(\sigma n^{-1/2}). \end{aligned} \tag{2.5}$$

This shows $\|M_{3h}\|_{\psi_2} \leq Ch^{-1}(\sigma n^{-1/2})$ by (2.4). □

By (2.2), Theorem 2.4.1 implies that the probability of the H^1-finite element error violating the convergence order $O(h|\ln h|^{1/2}(\|u\|_{H^2(\Omega)}+\|f\|_{H^2(\Omega)})+ h^{-1}(\sigma n^{-1/2}))$ decays exponentially.

Theorem 2.4.2. *We have*

$$\begin{aligned} &\| \|u - u_h \circ \Phi_h^{-1}\|_{L^2(\Omega)} \|_{\psi_2} \\ \leq\ & Ch^2 |\ln h|(\|u\|_{H^2(\Omega)} + \|f\|_{H^2(\Omega)} + \|g_0\|_{H^2(\Gamma)}) + C|\ln h|(\sigma n^{-1/2}). \end{aligned}$$

Proof. Let $G_h = \{\mu_h \in Q_h : \|\mu_h\|_{H^{1/2}(\Gamma)} \leq 1\}$. By (2.23) we are left to show

$$\| \sup_{\mu_h \in G_h} |\langle \mu_h, e\rangle_n| \|_{\psi_2} \leq C|\ln h|(\sigma n^{-1/2}). \tag{2.6}$$

Again we use the telescope sum in (2.13) and obtain

$$\begin{aligned} &\| \sup_{\mu_h \in G_h} |\langle \mu_h, e\rangle_n| \|_{\psi_2} \\ \leq\ & \sum_{i=0}^{p-1} \| \sup_{\mu_h \in G_h} |\langle \mu_{h_i} - \mu_{h_{i+1}}, e\rangle_n| \|_{\psi_2} + \| \sup_{\mu_h \in G_h} |\langle \mu_{h_p}, e\rangle_n| \|_{\psi_2}. \end{aligned} \tag{2.7}$$

By the same argument as the one in the proof of (2.5) and using (2.14) we have

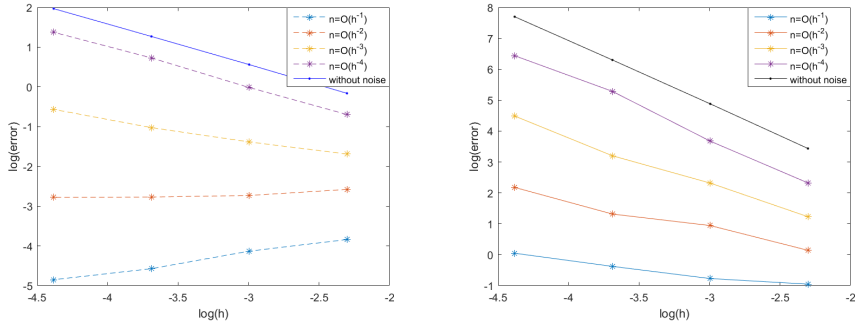
$$\begin{aligned} &\| \sup_{\mu_h \in G_h} |\langle \mu_{h_i} - \mu_{h_{i+1}}, e\rangle_n| \|_{\psi_2} \leq Ch_{i+1}^{1/2}(h_i^{-1/2} + h_{i+1}^{-1/2})(\sigma n^{-1/2}), \\ &\| \sup_{\mu_h \in G_{h_p}} |\langle \mu_h, e\rangle_n| \|_{\psi_2} \leq Ch_p^{-1/2}(\sigma n^{-1/2}). \end{aligned}$$

Inserting the above estimates into (2.7) shows (2.6) by taking p such that $|\ln h| < p \leq |\ln h| + 1$. □

By (2.2), Theorem 2.4.2 implies that the probability of the L^2-finite element error violating the convergence order $O(h^2|\ln h|(\|u\|_{H^2(\Omega)}+\|f\|_{H^2(\Omega)}+ \|g_0\|_{H^2(\Gamma)}) + |\ln h|(\sigma n^{-1/2}))$ decays exponentially.

2.5 Numerical examples

In this section, we show several numerical experiments to verify the theoretical analysis in this chapter. The analyses in Section 2.3 and Section 2.4 suggest that the optimal convergence rate can be achieved by taking $n = O(h^{-4})$. For the examples below, we take the exact solution $u_0 = \sin(5x+1)\sin(5y+1)$.



(a) H^1 convergence (b) L^2 convergence

FIGURE 2.1: The log-log plot of the convergence rate on the unit square.

n	h	H^1 error	α	L^2 error	β
$n = h^{-1}$	0.1000	8.8686		0.3978	
	0.0125	24.3951	0.4866	0.1394	-0.5043
$n = h^{-2}$	0.1000	2.8101		0.1348	
	0.0125	2.7125	-0.0170	0.0167	-1.0037
$n = h^{-3}$	0.1000	0.9637		0.0537	
	0.0125	0.3094	-0.5464	0.0017	-1.6649
$n = h^{-4}$	0.1000	0.6325		0.0380	
	0.0125	0.0838	-0.9721	6.3816e-4	-1.9656

TABLE 2.1: The convergence rate α in the H^1 norm and β in the L^2 norm on the unit square.

Example 2.5.1. *We take $\Omega = (0,1) \times (0,1)$. We construct the finite element mesh by first dividing the domain into $h^{-1} \times h^{-1}$ uniform rectangles and then connecting the lower left and upper right angle. We set $\{x_i\}_{i=1}^n$ being uniformly distributed on Γ, and $e_i,\ i = 1, 2, \cdots, n$, being independent normal random variables with variance $\sigma = 2$. We take different $n = h^{-i}$, $i = 1, 2, 3, 4$. Figure 2.1 shows the convergence rate of the error in the H^1 and L^2 norm for each choice of n. Table 2.1 show the convergence rate α in the H^1 norm and the convergence rate β in the L^2 norm.*

We observe that the numerical results confirm our theoretical analysis. The optimal convergence rate is achieved when choosing $n = h^{-4}$ while the other choices do not achieve optimal convergence. For example, when $n = h^{-2}$, the L^2 error is approximately $O(h^1)$ and no convergence for the H^1 error.

Example 2.5.2. *We take Ω to be the unit circle. The mesh is depicted in Figure 2.2. We set $\{x_i\}_{i=1}^n$ being uniformly distributed on Γ, and let $e_i = \eta_i + \alpha_i,\ i = 1, 2, \cdots, n$, where η_i and α_i are independent normal random variables with variance $\sigma_1 = 1$ and $\sigma_2 = 10e_i,\ i = 1, 2, \cdots, n$. We take different $n = h^{-i}$, $i = 1, 2, 3, 4$. Figure 2.3 shows the convergence rate of the error in the H^1 and L^2 norm for each choice of n. Table 2.2 shows the convergence rate α in the H^1*

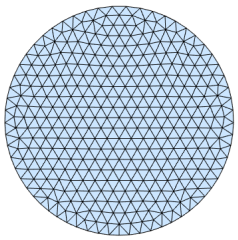
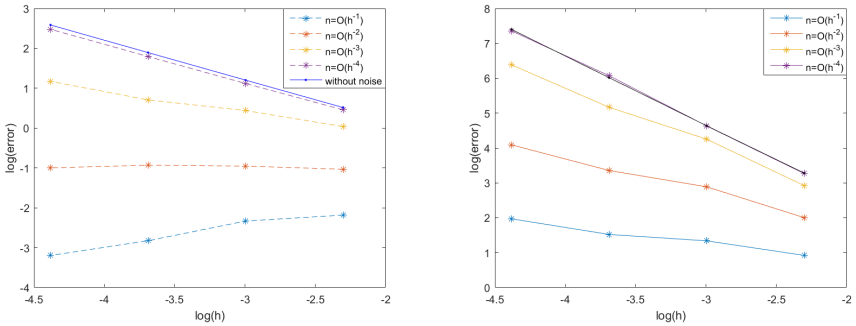


FIGURE 2.2: The uniform mesh for the unit circle with mesh size $h = 0.1$.



(a) H^1 convergence (b) L^2 convergence

FIGURE 2.3: The log-log plot of the convergence rate on the unit circle.

norm and the convergence rate β in the L^2 norm. Here again we observe the numerical results confirm our theoretical analysis.

n	h	H^1 error	α	L^2 error	β
$n = h^{-1}$	0.1000	46.5037		2.5872	
	0.0125	127.832	0.4863	0.9527	-0.4804
$n = h^{-2}$	0.1000	13.1775		0.8668	
	0.0125	16.1040	0.0964	0.1133	-0.9787
$n = h^{-3}$	0.1000	5.4157		0.2924	
	0.0125	1.7581	-0.5410	0.0113	-1.5665
$n = h^{-4}$	0.1000	2.0009		0.0980	
	0.0125	0.2527	-0.9950	0.0016	-1.9790

TABLE 2.2: The convergence rate α in the H^1 norm and β in the L^2 norm on the unit circle.

Part II

Two imaging methods and analysis of cell model for electropermeabilization

Chapter 3

The Linearized inverse problem in multifrequency EIT

3.1 Introduction

In this chapter, we propose a linearized method to solve the multifrequency EIT problem.

Electrical impedance tomography (EIT) is a diffusive imaging modality in which the conductivity distribution of the concerned object is recovered from the electrode voltage measurements on the boundary, induced by (multiple) known injected currents. The modality is safe, cheap and portable, and has the potential to be an established clinical imaging method in a multitude of applications [47]. However, the EIT inverse problem is severely ill-posed, and up to now, the resulting image quality is rather modest when compared with other modalities. This has motivated numerous studies on EIT imaging techniques.

The static imaging, which aims at recovering the absolute conductivity values, has so far achieved only limited success in practice. This is attributed to the fact that the electrode voltages are insensitive to localized conductivity changes, and thus the reconstructions are very sensitive to inevitable data noise as well as modelling errors, e.g., the boundary shape and electrode positions (and contact impedances). Hence, apart from accurate data, a very accurate forward model is also required for the success of static imaging, which is often difficult to obtain in practice. One prominent idea is to use difference imaging, in the hope of cancelling out the errors due to boundary shape, electrode and other systematic errors. One traditional approach to overcome these issues is the time difference imaging, where an image of the resulting conductivity change is produced by inverting a linearized sensitivity model. A second approach is multifrequency electrical impedance tomography (mfEIT), also known as electrical impedance tomography spectroscopy, and it has attracted some attention in recent years.

In mfEIT, one exploits the frequency dependence of the conductivity distribution. Experimentally, researchers have observed that many biological tissues exhibit strong frequency dependence within certain frequency ranges [31, 30, 69]. In mfEIT, the boundary voltage measurements are recorded simultaneously, whilst varying the modulation frequency of the injected current. The multifrequency approach is expected to be especially useful for diagnostic imaging of physiological conditions such as acute stroke, brain injury, and breast cancer, since patients are admitted into care after

the onset of the pathology and a baseline recording of healthy tissue is not available.

There have been several studies on frequency-difference imaging. Several earlier pieces of works include [38, 94, 119]. For example, a multifrequency experimental design which provides up to 64 electrodes for imaging the head was described in [119]. In these useful works, the simple frequency difference (between two neighboring frequencies) was often employed. Seo et al. [96] proposed a weighted frequency difference imaging technique, which is based on a suitable weighted voltage difference between any *two* sets of data. It was numerically shown that the approach can accommodate geometrical errors, including imperfectly known boundary. This approach can improve the reconstruction quality when the background is frequency dependent. Recently, Malone et al. [76] proposed a nonlinear reconstruction scheme, which uses all multifrequency data directly to reconstruct the volume fraction distribution of the tissues, and validated the approach on phantom experimental data; see also [75] for a recent probabilistic reconstruction-classification based technique. We also refer to [60] for a related mathematical study.

In this work, we shall analyze mfEIT in the linearized regime, by linearizing the forward model around the homogeneous background conductivity, as customarily adopted in practice [47, 2]. We shall discuss both the mathematically convenient continuum model and the practically popular complete electrode model. Our main contributions are as follows. First, we systematically discuss mfEIT reconstruction in the following three different scenarios, i.e. known spectral profiles, partially known spectral profiles and unknown spectral profiles. This analysis generalizes the existing studies, especially [96]. Second, we provide a rigorous justification of mfEIT for handling geometrical errors. Third, we present a novel group sparse reconstruction algorithm of iterative shrinkage type, which is easy to implement and converges fast. The extensive numerical experiments fully confirm our discussions.

The rest of the chapter is organized as follows. In Section 3.2, we mathematically formulate mfEIT using a continuum model, and analyze three important scenarios, depending on the knowledge of the spectral profiles. Then, in Section 3.3, we illustrate the potential of mfEIT in handling the modelling error due to an imperfectly known boundary shape. These analyses are then extended to the more realistic complete electrode model in Section 3.4. In Section 3.5, we present a novel group sparse reconstruction algorithm. Last, in Section 3.6, extensive numerical experiments are presented to illustrate the approach both with a known and with an imperfectly known boundary.

3.2 The Continuum Model

In this section, we mathematically formulate multifrequency electrical impedance tomography (mfEIT) in the continuum model. The extensions to an imperfectly known boundary and the complete electrode model will be described in Sections 3.3 and 3.4 below. Let Ω be an open bounded domain in $\mathbb{R}^d$ ($d = 2, 3$), occupied by the object, with a smooth boundary $\partial\Omega$. Then the mfEIT forward problem reads: given any input flux $f \in L^2(\partial\Omega)$ with

$\int_{\partial\Omega} f ds = 0$ and the frequency-dependent conductivity distribution $\sigma(x,\omega)$, find $u(x,\omega)$ such that

$$\begin{cases} -\nabla\cdot(\sigma(x,\omega)\nabla u(x,\omega)) = 0 \quad \text{in } \Omega, \\ \sigma(x,\omega)\dfrac{\partial u}{\partial \nu} = f(x) \quad \text{on } \partial\Omega, \\ \displaystyle\int_{\partial\Omega} u(x,\omega) ds = 0, \end{cases} \tag{3.1}$$

where ω is the frequency, ∇ denotes the gradient with respect to the spatial variable x, and ν is the unit outward normal direction to the boundary $\partial\Omega$. The weak formulation of problem (3.1) is given by: find $u = u(x,\omega) \in H^1(\Omega)$, with the grounding condition $\int_{\partial\Omega} u(x,\omega) ds = 0$, such that

$$\int_\Omega \sigma\nabla u\cdot\nabla v dx = \int_{\partial\Omega} f v ds, \qquad v \in H^1(\Omega).$$

Throughout, we assume that the frequency-dependent conductivity distribution $\sigma(x,\omega)$ takes the following separable form:

$$\sigma(x,\omega) = \sum_{k=0}^{K} \sigma_k(x) s_k(\omega), \tag{3.2}$$

where $K+1$ is the number of spectral profiles, and $\{s_k(\omega)\}_{k=0}^K$ are a collection of (possibly only partially known) material spectra, often referred to as endmembers, and $\{\delta\sigma_k(x)\}_{k=0}^K$ are scalar functions representing the corresponding proportions, also known as abundances, following the terminology in the hyperspectral unmixing literature [59]. Further, we shall assume

$$\begin{aligned} \sigma_0(x) &= 1 + \delta\sigma_0(x), \\ \sigma_k(x) &= \delta\sigma_k(x), \quad k = 1,\dots,K, \end{aligned}$$

where the $\delta\sigma_k$s are small (in suitable $L^p(\Omega)$ norms) so that a linearized model is valid. The $\delta\sigma_k$s, including the background $\delta\sigma_0$, are all unknown and represent the small inclusions/anomalies in the object Ω. We assume that they have compact spatial supports in the domain Ω, and are disjoint from each other. We also assume that the background spectral profile $s_0(\omega)$ is known.

In order to gain sufficient information about the conductivity distribution $\sigma(x,\omega)$, we apply M linearly independent input currents $\{f_n\}_{n=1}^M \subset L^2(\partial\Omega)$, with the grounding condition $\int_{\partial\Omega} f_n ds = 0$, to the boundary $\partial\Omega$. Let $\{u_n \equiv u_n(x,\omega)\}_{n=1}^M \subset H^1(\Omega)$ be the corresponding solutions to (3.1), i.e.,

$$\int_\Omega \sigma\nabla u_n\cdot\nabla v dx = \int_{\partial\Omega} f_n v ds, \qquad v \in H^1(\Omega). \tag{3.3}$$

Then the mfEIT inverse problem is to recover the abundances $\{\delta\sigma_k\}_{k=0}^K$ and/or other quantities of interest from noisy measurements of the electrode voltage $\{u_n(x,\omega)\}_{n=1}^M$ on the boundary $\partial\Omega$ at a number of modulating frequencies $\{\omega_q\}_{q=1}^Q$.

Next we derive the linearized model for the inverse problem. The linearized model is customarily employed in EIT applications due to its computational convenience [47, 2]. In this work, we adopt a linearized model

based on an integral representation. Let $v_m \in H^1(\Omega)$ be the potential corresponding to the unperturbed conductivity distribution $\sigma_0(x,\omega) \equiv s_0(\omega)$ with the input current f_m, namely

$$\int_\Omega \sigma_0 \nabla v_m \cdot \nabla v dx = \int_{\partial\Omega} f_m v ds, \qquad v \in H^1(\Omega). \tag{3.4}$$

We observe that $v_m = v_m^*/s_0(\omega)$, where v_m^* is the solution of (3.4) corresponding to the case $s_0 \equiv 1$. In other words, the dependence of v_m on the frequency ω is explicit. By taking $v = v_m$ in (3.3) and $v = u_n$ in (3.4), and subtracting the two identities, we obtain

$$\sum_{k=0}^{K} s_k(\omega) \int_\Omega \delta\sigma_k \nabla u_n \cdot \nabla v_m dx = \int_{\partial\Omega} (f_n v_m - f_m u_n) ds.$$

Hence, under the approximation $\nabla u_n(x,\omega) \approx \nabla v_n(x,\omega)$ in the domain Ω (valid in the linear regime), and using the identity $v_m = v_m^*/s_0(\omega)$, we arrive at the following linearized model

$$\sum_{k=0}^{K} s_k(\omega) \int_\Omega \delta\sigma_k \nabla v_n^* \cdot \nabla v_m^* dx = s_0(\omega)^2 \int_{\partial\Omega} (f_n v_m - f_m u_n) ds. \tag{3.5}$$

Note that the right hand side of this identity is completely known: u_n is the measured electrode voltage data (and thus depends on the frequency ω), and v_m can be computed (using the reference conductivity $\sigma_0(x,\omega) \equiv s_0(\omega)$), and thus it is known upon simple computation. Next, we triangulate the domain Ω into a shape regular quasi-uniform mesh of simplicial elements $\{\Omega_l\}_{l=1}^L$ such that $\overline{\Omega} = \cup_{l=1}^L \Omega_l$, and consider a piecewise constant approximation of the inclusions $\delta\sigma_k$s: for $k = 0, \ldots, K$,

$$\delta\sigma_k(x) \approx \sum_{l=1}^{L} (\delta\sigma_k)_l \chi_{\Omega_l}(x),$$

where χ_{Ω_l} is the characteristic function of the lth element Ω_l, and $(\delta\sigma_k)_l$ denotes the value of the kth abundance $\delta\sigma_k$ in the lth element Ω_l. Upon substituting the approximation into (3.5), we have the following finite-dimensional linear inverse problem

$$\sum_{k=0}^{K} s_k(\omega) \sum_{l=1}^{L} (\delta\sigma_k)_l \int_{\Omega_l} \nabla v_n^* \cdot \nabla v_m^* dx = s_0(\omega)^2 \int_{\partial\Omega} (f_n v_m - f_m u_n) ds.$$

Last, we introduce the sensitivity matrix M and the data vector X. We use a single index $j = 1, \ldots, J$ with $J = M^2$ for the index pair (m,n) with $j = M(m-1)+n$, and introduce the sensitivity matrix $M = [M_{jl}] \in \mathbb{R}^{J\times L}$ with its entries M_{jl} given by

$$M_{jl} = \int_{\Omega_l} \nabla v_n^* \cdot \nabla v_m^* dx \quad (j \leftrightarrow (m,n)),$$

which is independent of the frequency ω. Likewise, we introduce a data vector $X(\omega) \in \mathbb{R}^J$ with its jth entry $X_j(\omega)$ given by

$$X_j(\omega) = s_0(\omega)^2 \int_{\partial\Omega} (f_n v_m(\omega) - f_m u_n(\omega)) ds \quad (j \leftrightarrow (m,n)).$$

Upon writing the vectors $A_k = (\delta\sigma_k)_l \in \mathbb{R}^L$, $k = 0, \ldots, K$, we obtain the following linear system (parameterized by the frequency ω)

$$M \sum_{k=0}^{K} s_k(\omega) A_k = X(\omega). \tag{3.6}$$

In the mfEIT inverse problem, the abundance vectors $\{A_k\}_{k=0}^K$ are of primary interest and have to be estimated from the frequency dependent data $X(\omega)$, and occasionally the spectral profiles $\{s_k(\omega)\}_{k=\gamma 1}^K$ are also of interest. (Recall that the background spectral profile $s_0(\omega)$ is always supposed to be known.) Depending on the further a priori knowledge available about the spectral profiles $\{s_k(\omega)\}_{k=1}^K$, we distinguish the following three cases:

(a) All the spectral profiles $\{s_k(\omega)\}_{k=0}^K$ are known.

(b) The spectral profiles $\{s_k(\omega)\}_{k=1}^K$ may not be fully known, but their frequency dependence differs substantially.

(c) The spectral profiles are only partially known, and we aim at a partial recovery of the abundances.

These three cases are of different degree of challenge, and we shall discuss them separately below.

3.2.1 Case (a): Known Spectral Profiles

First we consider the case when the spectral profiles $\{s_k(\omega)\}_{k=0}^K$ are all known. In some applications, this does not represent a restriction, since the spectral profiles of many materials can actually be measured accurately (see e.g. [29] for the electrical conductivity of tissues at frequencies below 1 MHz). Suppose that we can measure the boundary voltage $u(x,\omega)$ at Q distinct frequencies $\{\omega_q\}_{q=1}^Q$. Then by writing $S = (S_{kq}) \in \mathbb{R}^{(K+1)\times Q}$, with $S_{kq} = s_k(\omega_q)$, we get from (3.6) the following matrix equation

$$MAS = X, \tag{3.7}$$

where the matrix $X = [X(\omega_1) \ \ldots \ X(\omega_Q)] \in \mathbb{R}^{J\times Q}$. In equation (3.7), the sensitivity matrix M can be precomputed, and the spectral profile matrix S and data X are known: Only the abundance matrix $A = [A_0 \ \ldots \ A_K] \in \mathbb{R}^{L\times(K+1)}$ is unknown. It is natural to assume that a sufficiently large number of frequencies are taken, so that the corresponding spectral profile matrix S is incoherent in the sense that $Q \geq K+1$ and $\operatorname{rank}(S) = K+1$ (and presumably S is also well-conditioned). Then the matrix S admits a right inverse S^{-1}. By letting $Y = XS^{-1}$ we obtain

$$MA = Y.$$

These are $K+1$ decoupled (and usually undetermined) linear system. By letting $Y = [Y_0 \ \dots \ Y_K] \in \mathbb{R}^{J\times(K+1)}$, we have $K+1$ independent (finite-dimensional) linear inverse problems

$$MA_k = Y_k, \quad k = 0, \dots, K, \tag{3.8}$$

where A_k represents the kth abundance. Here each linear system determines one and only one abundance A_k. The stable and accurate numerical solution of these ill-conditioned linear systems by suitable regularization techniques will be discussed in detail in Section 3.5 below.

The condition $\operatorname{rank}(S) = K+1$ on the matrix S is necessary and sufficient for the full decoupling of the abundances, and the well-conditioning of S ensures a stable decoupling procedure. It specifies the condition under which the abundance unmixing is practically feasible, and also the proper selection of the frequencies $\{\omega_q\}_{q=1}^{Q}$ such that $\operatorname{rank}(S) = K+1$. Note that the condition $\operatorname{rank}(S) = K+1$ depends essentially on the independence/incoherence of the continuous spectral profiles $\{s_k(\omega)\}_{k=0}^{K}$ (or frequency contrast). In the absence of this spectral incoherence, a fully decoupling is impossible. For example, consider the simple case of two endmembers, with $s_0(\omega) = 1+\omega$, $s_1(\omega) = 2+2\omega$. Then no matter how many frequencies one chooses, the spectral matrix S is always of rank one. Thus it is impossible to separate the two abundances, and instead only a linear combination can be obtained.

The right inverse $Y = XS^{-1}$ can alternatively be viewed as a least-squares procedure

$$\min_{Y\in\mathbb{R}^{J\times(K+1)}} \|X - YS\|_F.$$

This formulation exhibits clearly that for a rank-deficient spectral matrix S, the proposed approach yields the minimum-norm matrix Y compatible with the data, and for an inconsistent S, it yields a best approximation via projection (even though the physical interpretation is less clear). In addition, by the perturbation theory for least-squares problems [37], the well-conditioning of the spectral matrix S implies that the procedure is also stable with respect to small perturbations in the spectral profiles.

It is worth noting that this approach represents a natural generalization of the weighted frequency difference EIT (fdEIT) method proposed in [96], where only two abundances and two frequencies are considered ($K = 1$ and $Q = 2$).

Example 1. *Consider the case with $K = 1$ and $Q = 2$, namely, two inclusions and two frequencies. We write*

$$X = [X(\omega_1)\ X(\omega_2)] \quad \textit{and} \quad S = \begin{bmatrix} s_0(\omega_1) & s_0(\omega_2) \\ s_1(\omega_1) & s_1(\omega_2) \end{bmatrix}.$$

Therefore, if S is invertible, we obtain

$$Y = XS^{-1} = \frac{s_0(\omega_1)}{\det S}\left[\frac{s_1(\omega_2)}{s_0(\omega_1)}X(\omega_1) - \frac{s_1(\omega_1)}{s_0(\omega_1)}X(\omega_2) \quad X(\omega_2) - \frac{s_0(\omega_2)}{s_0(\omega_1)}X(\omega_1)\right].$$

The second column of Y recovers exactly the weighted fdEIT method discussed in [96]. Thus our method generalizes [96], as the recent work [56]. Our approach is slightly more general since it directly incorporates the multifrequency data. The

use of multiple frequencies is expected to improve the numerical stability, especially when high correlation may occur between neighboring frequencies and possibly the spectral profiles themselves are not precisely known. Further, the use of multiple frequencies enables completely decoupling multiple inclusions, instead of only one inclusion, which can be very useful in practice. It is also worth noting that in the special case of constant background spectral profile $s_0(\omega_1) = s_0(\omega_2)$, the approach reduces to the usual frequency difference. This delineates the region of validity of the usual frequency difference imaging for processing multifrequency data. We also note that the invertibility of spectral file matrix S is equivalent to the fact that the two rows of S are not collinear, i.e., the two frequency profiles are incoherent, and there is frequency dependent conductivity contrast.

3.2.2 Case (b): Spectral Profiles with Substantially Different Frequency Dependence

Next we consider the case when some of (or, possibly all) the spectral profiles $\{s_k(\omega)\}_{k=1}^K$ are not known, but do not change rapidly with the frequency ω, if compared to the remaining ones. Thus, instead of using the data $X(\omega)$ directly, it is natural to differentiate the relation (3.6) with respect to ω to eliminate the contributions from the abundances whose endmembers do not vary much with the frequency ω. This discriminating effect is useful in practice. For example, it is known that the conductivity of the malign tissues is more sensitive with respect to the frequency variation in certain frequency ranges [106, 69], even though the conductivity of healthy tissues in the background may exhibit fairly complex structure. The differentiation procedure provides a valuable tool in such scenarios.

More precisely, let $\mathcal{P} \subseteq \{0, 1, \dots, K\}$ be such that

$$\frac{d}{d\omega} s_p(\omega_q) \gg \frac{d}{d\omega} s_k(\omega_q), \qquad p \in \mathcal{P}, k \in \{0, 1, \dots, K\} \setminus \mathcal{P}. \tag{3.9}$$

By differentiating (3.6) with respect to the frequency ω we obtain

$$M \sum_{k=0}^{K} A_k \frac{d}{d\omega} s_k(\omega) = \frac{d}{d\omega} X(\omega).$$

In view of the assumption (3.9), this equality may be rewritten as

$$M \sum_{p \in \mathcal{P}} A_p \frac{d}{d\omega} s_p(\omega) \approx \frac{d}{d\omega} X(\omega). \tag{3.10}$$

In other words, the contributions from the remaining profiles are negligible, provided that the abundances A_k are comparable in magnitude. Different reconstruction schemes should be used depending on whether the spectral profiles $\{s_p(\omega)\}_{p \in \mathcal{P}}$ are known.

Case (b1): The Spectral Profiles $\{s_p(\omega)\}_{p \in \mathcal{P}}$ are not Known

In the case when the spectral profiles $\{s_p(\omega)\}_{p \in \mathcal{P}}$ are not known, the linear system (3.10) cannot be simplified further. By solving this underdetermined system, we can recover at most $\sum_{p \in \mathcal{P}} s'_p(\omega) A_p$, namely a linear combination

of the inclusions. Since the weights $\{s'_p(\omega)\}_{p\in\mathcal{P}}$ are unknown, it is impossible to separate the abundances. However, in the particular case, when $\mathcal{P} = \{p\}$ (i.e., $|\mathcal{P}| = 1$), the abundance $\delta\sigma_p$ may be recovered up to an unknown multiplicative constant, which gives the support information.

We illustrate the technique with an example.

Example 2. *Consider the case $K = 1$, and two linear frequency profiles, i.e.,*

$$s_0(\omega) = \alpha_0 + \beta_0\omega \quad \textit{and} \quad s_1(\omega) = \alpha_1 + \beta_1\omega,$$

with $\beta_0 \ll \beta_1$. Then the differentiation imaging amounts to

$$\beta_0 M A_0 + \beta_1 M A_1 = X'(\omega).$$

If MA_0 and MA_1 are comparable, then $\beta_0 \ll \beta_1$ implies that the contribution of $\beta_0 MA_0$ to the data is negligible (in comparison with $\beta_1 MA_1$). Hence upon differentiation, the technique allows to recover the dominant component $\beta_1 MA_1$, which upon linear inversion yields $\beta_1 A_1$, which in particular contains the support information about the abundance A_1, and also its magnitude up to a multiplicative constant. Further, for known β_1, it allows full recovery of the abundance A_1.

Case (b2): The Spectral Profiles $\{s_p(\omega)\}_{p\in\mathcal{P}}$ are Known

If the spectral profiles $\{s_p(\omega)\}_{p\in\mathcal{P}}$ are known, it is possible to perform the same analysis of Case (a) (in Section 3.2.1) to system (3.10). Taking measurements at Q distinct frequencies $\omega_1, \ldots, \omega_Q$, we have

$$M \sum_{p\in\mathcal{P}} A_p s'_p(\omega_q) \approx X'(\omega_q), \qquad q = 1, \ldots, Q.$$

Then, by writing $\widetilde{S} = (\widetilde{S}_{pq}) \in \mathbb{R}^{|\mathcal{P}|\times Q}$, with $\widetilde{S}_{pq} = s'_p(\omega_q)$, we get

$$MA\widetilde{S} = X'$$

where the matrix $X' = [X'(\omega_1) \ \ldots \ X'(\omega_Q)] \in \mathbb{R}^{J\times Q}$. Then the inversion step is completely analogous to that discussed in Section 3.2.1, provided that the incoherence condition $\operatorname{rank}\widetilde{S} = |\mathcal{P}|$ (as well as well-conditioning) holds. All the inclusions A_p, $p \in \mathcal{P}$, can be recovered.

Numerical Implementation

In the implementation, we take the forward difference between neighboring frequencies

$$M \sum_{k=0}^{K} A_k \frac{s_k(\omega_{q+1}) - s_k(\omega_q)}{\omega_{q+1} - \omega_q} = \frac{X(\omega_{q+1}) - X(\omega_q)}{\omega_{q+1} - \omega_q}. \tag{3.11}$$

It approximates the first order derivative $s'_k(\omega_q)$ with the forward difference

$$\frac{d}{d\omega} s_k(\omega_q) \approx \frac{s_k(\omega_{q+1}) - s_k(\omega_q)}{\omega_{q+1} - \omega_q}.$$

If more than one neighboring frequencies are available, it is also possible to use higher order difference formulas to get more accurate approximations of the derivative $s_k'(\omega)$, e.g., central difference scheme (on a nonuniform frequency grid)

$$\frac{d}{d\omega}s_k(\omega_q) \approx \frac{\delta\omega_{q-1}^2(s_k(\omega_{q+1}) - s_k(\omega_q)) + \delta\omega_q^2(s_k(\omega_q) - s_k(\omega_{q-1}))}{\delta\omega_q\delta\omega_{q-1}(\delta\omega_q + \delta\omega_{q-1})},$$

where $\delta\omega_q = \omega_{q+1} - \omega_q$. In practice, these represent different ways to perform differentiation imaging. However, their robustness with respect to noise in the data might differ due to the well-known ill-posed nature of numerical differentiation [43]. In this work, we shall use the forward difference scheme (3.11).

3.2.3 Case (c): Partially Known Spectral Profiles, Partial Recovery of the Abundances

In practice, it is also of interest to recover some information about the abundances when the spectral profiles $\{s_k(\omega)\}$ are only partially known. Unfortunately, in general, this is infeasible. But, one can still obtain some information under certain a priori knowledge. To discuss the situation, recall the notation $Y_k = MA_k$, cf. (3.8). Then

$$Y_0 s_0(\omega_q) + \ldots + Y_K s_K(\omega_q) = X(\omega_q), \qquad q = 1, \ldots, Q. \tag{3.12}$$

Now suppose the frequency dependence of the spectral profiles $\{s_k(\omega)\}_{k=0}^K$ are polynomial type, namely

$$s_k(\omega) = \sum_{n=0}^{N} \alpha_k^n \omega^n.$$

Inserting this expression in the identity (3.12) yields

$$\sum_{n=0}^{N}\sum_{k=0}^{K}(\alpha_k^n Y_k)\omega^n = X(\omega).$$

By taking a sufficiently large number of modulating frequencies $\{\omega_q\}_{q=1}^Q$, and using the identity principle for polynomials, we can compute the quantities

$$B_n := \sum_{k=0}^{K} \alpha_k^n Y_k, \qquad n = 0, \ldots, N.$$

Note that adding more frequencies would not add more information about Y_k and α_k^j than $\{B_n\}_{n=0}^N$. In other words, the quantities $\{B_n\}_{n=0}^N$ represent the essential information content in the data $\{X(\omega_q)\}_{q=1}^Q$ about the unknowns $\{Y_k\}_{k=0}^K$ and $\{\alpha_k^n : k = 0, \ldots, K,\ n = 0, \ldots, N\}$. Depending on K, N and the number of unknowns among the weights α_k^n of the spectral profiles, some inclusions Y_k can be reconstructed without knowing the corresponding spectral profiles. In other situations, there may be more unknowns than the number of equations, and it may be infeasible to determine all of them.

Instead of providing a general analysis of all possible cases, we present two examples that explain the different situations that may appear.

Example 3. *Consider the case $K = 1$. For every n we have $B_0 = \alpha_0^0 Y_0 + \alpha_1^0 Y_1$ and $B_n = \alpha_0^n Y_0 + \alpha_1^n Y_1$, whence*

$$Y_1 = (\alpha_0^0 \alpha_1^n - \alpha_1^0 \alpha_0^n)^{-1} (\alpha_0^0 B_n - \alpha_0^n B_0).$$

Since the spectral profile s_0 is always assumed to be known, so are the quantities α_0^0 and α_0^n. Therefore, Y_1 may be reconstructed up to a multiplicative constant c, provided that $\alpha_0^0 \alpha_1^n - \alpha_1^0 \alpha_0^n \neq 0$, without assuming any knowledge of the corresponding spectral profile s_1. Note that this nonzero condition simply represents the incoherence of the spectral profiles s_0 and s_1. Finally, by solving the underdetermined system $MA_1 = cY_1$, the inclusion $\delta\sigma_1$ can be reconstructed up to the multiplicative constant c.

In addition, assuming a unique recovery of the linearized EIT inverse problem, the knowledge of the quantity B_0 allows to recover an unknown linear combination of the abundances A_0 and A_1, and in particular the union of their supports. Since the supports of A_0 and A_1 are assumed to be disjoint from each other, this piece of information allows the support of the abundance A_0 to be recovered, given that the support of the abundance A_1 has already been reconstructed.

Example 4. *Note that if $K = 2$ and $N = 1$, we get only*

$$\alpha_0^0 Y_0 + \alpha_1^0 Y_1 + \alpha_2^0 Y_2 = B_0 \quad \text{and} \quad \alpha_0^1 Y_0 + \alpha_1^1 Y_1 + \alpha_2^1 Y_2 = B_1$$

which is vastly insufficient to determine all the unknowns. However, a calculation similar to the one presented in the previous example shows that Y_2 can be determined up to a multiplicative constant if $K = N = 2$ and s_1 is known, provided that a certain nonzero condition is satisfied. Like before, by solving the underdetermined system $MA_2 = cY_2$, the inclusion $\delta\sigma_2$ can be reconstructed up to a multiplicative constant, in particular its support. Like before, assuming a unique recovery with the linearized inverse problem, the union of the supports of $\delta\sigma_0$ and $\delta\sigma_1$ may be determined.

With obvious modifications, the preceding discussion is also valid for more general basis functions $\phi_n(\omega)$ which form a unisolvent system on the set of measured frequencies $\{\omega_q\}_{q=1}^Q$ [25, pp. 31–32].

3.3 Imperfectly Known Boundary

In this part, we illustrate the significant potentials of mfEIT for handling modelling errors, as in the case of an imperfectly known boundary. The analysis may be extended to other interesting scenarios, including imperfectly known contact impedances or injected currents. Here we shall discuss only the specific case of an imperfectly known boundary. This has long been one of the main obstacles in some practical applications of the EIT imaging [1, 63, 62]. It is known that the use of a slightly incorrect boundary can lead to significant errors in the reconstruction. mfEIT was proposed as one promising strategy to partially overcome the challenge in [96], where its potential was also numerically demonstrated. Here we shall present an analysis of the approach in the linearized regime which justifies these numerical findings.

We denote the true but unknown physical domain by $\widetilde{\Omega}$, and the computational domain by Ω, which is an approximation to $\widetilde{\Omega}$. Next we introduce a forward map $F : \widetilde{\Omega} \to \Omega$, $\widetilde{x} \to x$, which is assumed to be a smooth orientation preserving map with a sufficiently smooth inverse map $F^{-1} : \Omega \to \widetilde{\Omega}$. We denote the Jacobian of the map F by J_F, and the Jacobian of F with respect to the surface integral by J_F^S.

Suppose now that the function $\widetilde{u}_n(\widetilde{x}, \omega)$ satisfies problem (3.1) in the true domain $\widetilde{\Omega}$ with a conductivity $\widetilde{\sigma}(\widetilde{x}, \omega)$ and input current $\widetilde{f}_n$ on $\partial\widetilde{\Omega}$ with $\int_{\partial\widetilde{\Omega}} \widetilde{f}_n ds = 0$, namely

$$\begin{cases} -\nabla_{\widetilde{x}} \cdot (\widetilde{\sigma}(\widetilde{x}, \omega)\nabla_{\widetilde{x}}\widetilde{u}_n(\widetilde{x}, \omega)) = 0 \quad \text{in } \widetilde{\Omega}, \\ \widetilde{\sigma}(\widetilde{x}, \omega)\dfrac{\partial \widetilde{u}_n(\widetilde{x}, \omega)}{\partial \widetilde{\nu}} = \widetilde{f}_n \quad \text{on } \partial\widetilde{\Omega}, \\ \displaystyle\int_{\partial\widetilde{\Omega}} \widetilde{u}_n(\widetilde{x}, \omega) d\widetilde{s} = 0. \end{cases} \tag{3.13}$$

Here the frequency-dependent conductivity $\widetilde{\sigma}(\widetilde{x}, \omega)$ takes a separable form (cf. (3.2))

$$\widetilde{\sigma}(\widetilde{x}, \omega) = \sum_{k=0}^{K} s_k(\omega)\widetilde{\sigma}_k(\widetilde{x}), \tag{3.14}$$

with $\widetilde{\sigma}_0(\widetilde{x}) = 1 + \delta\widetilde{\sigma}_0(\widetilde{x})$, and $\widetilde{\sigma}_k(\widetilde{x}) = \delta\widetilde{\sigma}_k(\widetilde{x})$, $k = 1, \ldots, K$, where $\delta\widetilde{\sigma}_k$ are small and their supports are disjoint and stay away from the boundary $\partial\widetilde{\Omega}$. The weak formulation of problem (3.13) (by suppressing the dependence on the frequency ω) is given by: find $\widetilde{u}_n(\cdot, \omega) \in H^1(\widetilde{\Omega})$ with $\int_{\partial\widetilde{\Omega}} \widetilde{u}_n(\widetilde{x}, \omega) d\widetilde{s} = 0$ such that

$$\int_{\widetilde{\Omega}} \widetilde{\sigma}(\widetilde{x})\nabla_{\widetilde{x}}\widetilde{u}_n(\widetilde{x}) \cdot \nabla_{\widetilde{x}}\widetilde{v}(\widetilde{x}) d\widetilde{x} = \int_{\partial\widetilde{\Omega}} \widetilde{f}_n \widetilde{v} d\widetilde{s}, \qquad \widetilde{v} \in H^1(\widetilde{\Omega}). \tag{3.15}$$

Let us now discuss the experimental setup. The practitioner chooses a current density defined on the computational domain, namely a function f_n on $\partial\Omega$ such that $\int_{\partial\Omega} f_n \, ds = 0$. This current is then applied to the unknown boundary $\partial\widetilde{\Omega}$ of the real domain $\widetilde{\Omega}$. The deformation of the boundary has to be taken into account: the applied current $\widetilde{f}_n$ on $\partial\widetilde{\Omega}$ results to be

$$\widetilde{f}_n = (f_n \circ F)|\det J_F^S|. \tag{3.16}$$

Note that this implies directly $\int_{\partial\widetilde{\Omega}} \widetilde{f}_n d\widetilde{s} = 0$, as desired. This induces the electric potential $\widetilde{u}_n \in H^1(\widetilde{\Omega})$ given by (3.13) or, equivalently, by (3.15). Like in Section 3.2, the electric potential $\widetilde{u}_n$ is assumed to be measured on $\partial\widetilde{\Omega}$. However, because of the incorrect knowledge of the boundary, the measured quantity is in fact $u_n := \widetilde{u}_n \circ F^{-1}$ restricted to the computational boundary $\partial\Omega$.

Remark 3.3.1. *The factor $|\det J_F^S|$ has a certain physical interpretation. The current density on $\partial\widetilde{\Omega}$ is locally defined by $\widetilde{J} = I/\mathrm{area}(\widetilde{A})$, where I is the current injected through a small surface $\widetilde{A} \subseteq \partial\widetilde{\Omega}$. Thus*

$$\widetilde{J} = \frac{I}{\mathrm{area}(\widetilde{A})} = \frac{I}{\mathrm{area}(A)} \frac{\mathrm{area}(A)}{\mathrm{area}(\widetilde{A})} = J \frac{\mathrm{area}(A)}{\mathrm{area}(\widetilde{A})},$$

where J is the corresponding current density on $A := F(\widetilde{A}) \subseteq \partial\Omega$. Therefore, the factor $|\det J_F^S|$ is nothing other than the infinitesimal version of $\frac{\text{area}(A)}{\text{area}(\widetilde{A})}$ as $\text{area}(\widetilde{A}) \to 0$. It is worth pointing out that the following relation holds trivially true:

$$\int_{\partial\widetilde{\Omega}} \widetilde{f}_n \widetilde{u}_n d\widetilde{s} = \int_{\partial\Omega} f_n u_n ds.$$

In physical terms, the integral $\int_{\partial\Omega} f_n u_n ds$ represents the power needed to maintain the potential u_n on the boundary $\partial\Omega$. In other words, the choice (3.16) *preserves the needed power for the measured data, and it agrees with the one adopted in [63].*

We shall consider only the case that the computational domain Ω is a small variation of the true physical domain $\widetilde{\Omega}$ (but comparable with the inclusions $\delta\sigma_k$), so that the linearized regime is valid. Specifically, we write the map $F : \widetilde{\Omega} \to \Omega$ by $F(\widetilde{x}) = \widetilde{x} + \epsilon\widetilde{\phi}(\widetilde{x})$, where ϵ is a small scalar parameter and the smooth function $\widetilde{\phi}(\widetilde{x})$ characterizes the domain deformation. Further, let $F^{-1}(x) = x + \epsilon\phi(x)$ be the inverse map, which is also smooth.

To examine the influences of the domain deformation on the linearized inverse problem discussed in Section 3.2, we introduce the solution $v_m \in H^1(\Omega)$ relative to the reference conductivity distribution $\sigma_0(x,\omega) = s_0(\omega)$ in the computational domain Ω corresponding to the flux f_m, i.e.,

$$\int_\Omega \sigma_0 \nabla v_m \cdot \nabla v dx = \int_{\partial\Omega} f_m v ds, \qquad v \in H^1(\Omega), \tag{3.17}$$

which is computable over the computational domain Ω (and solvable by the assumption $\int_{\partial\Omega} f_m ds = 0$).

We can now state the corresponding linearized inverse problem. As a byproduct of the proof, we have that, even for an isotropic conductivity $\widetilde{\sigma}$ in the true domain $\widetilde{\Omega}$, cf. (3.14), in the computational domain Ω the equivalent conductivity $\sigma(x,\omega)$ is generally anisotropic (or matrix valued).

Proposition 1. *Set $\delta\sigma_k = \delta\widetilde{\sigma}_k \circ F^{-1}$ for $k = 0, 1, \ldots, K$ and $v_m^* = s_0(\omega) v_m$ for $m = 1, \ldots, M$. The linearized inverse problem on the computational domain Ω is given by*

$$s_0(\omega)\epsilon \int_\Omega \Psi \nabla v_n^* \cdot \nabla v_m^* dx + \sum_{k=0}^{K} s_k(\omega) \int_\Omega \delta\sigma_k \nabla v_n^* \cdot \nabla v_m^* dx = s_0(\omega)^2 \int_{\partial\Omega} (f_n v_m - f_m u_n) ds, \tag{3.18}$$

for some smooth function $\Psi : \Omega \to \mathbb{R}^{d\times d}$, which is independent of the frequency ω.

Proof. First, we derive the governing equation for the variable $u_n = \widetilde{u}_n \circ F^{-1}$ in the domain Ω from (3.15). Denote by $v = \widetilde{v} \circ F^{-1} \in H^1(\Omega)$. By the chain rule we have $\nabla_{\widetilde{x}} \widetilde{u}_n \circ F^{-1} = (J_F^t \circ F^{-1}) \nabla_x u_n$, where the superscript t denotes

the matrix transpose. Thus, we deduce

$$\begin{aligned}
&\int_{\widetilde{\Omega}} \widetilde{\sigma}(\widetilde{x})\nabla_{\widetilde{x}}\widetilde{u}_n(\widetilde{x})\cdot\nabla_{\widetilde{x}}\widetilde{v}(\widetilde{x})d\widetilde{x}\\
&= \int_{\Omega} (\widetilde{\sigma}\circ F^{-1})(x)(J_F^t\circ F^{-1})(x)\nabla u_n(x)\cdot(J_F^t\circ F^{-1})(x)\nabla v(x)|\det J_{F^{-1}}(x)|dx,\\
&= \int_{\Omega} (J_F\circ F^{-1})(x)(\widetilde{\sigma}\circ F^{-1})(x)(J_F^t\circ F^{-1})(x)\nabla u_n(x)\cdot\nabla v(x)|\det J_{F^{-1}}(x)|dx\\
&= \int_{\Omega} \sigma(x,\omega)\nabla u_n(x)\cdot\nabla v(x)dx,
\end{aligned}$$

where the transformed conductivity $\sigma(x,\omega)$ is given by [107, 63, 62]

$$\sigma(x,\omega) = \left(\frac{J_F(\cdot)\widetilde{\sigma}(\cdot,\omega)J_F^t(\cdot)}{|\det J_F(\cdot)|}\circ F^{-1}\right)(x). \tag{3.19}$$

Moreover, by (3.16) we have $\int_{\partial\widetilde{\Omega}} \widetilde{f}_n\widetilde{v}d\widetilde{s} = \int_{\partial\Omega} f_n v ds$. Therefore, in view of (3.15) the potential u_n satisfies

$$\int_{\Omega} \sigma(x,\omega)\nabla u_n(x,\omega)\cdot\nabla v(x)dx = \int_{\partial\Omega} f_n v ds, \qquad v\in H^1(\Omega). \tag{3.20}$$

Then by choosing $v = v_m$ in (3.20) and $v = u_n$ in (3.17), we arrive at

$$\int_{\Omega} (\sigma-\sigma_0)\nabla u_n\cdot\nabla v_m dx = \int_{\partial\Omega} (f_n v_m - f_m u_n)ds. \tag{3.21}$$

Note that $J_F = I + \epsilon J_{\widetilde{\phi}}$, and $J_{F^{-1}} = I + \epsilon J_\phi = I - \epsilon J_{\widetilde{\phi}}\circ F^{-1} + O(\epsilon^2)$, since ϵ is small. It is known that $|\det J_F| = 1 + \epsilon\mathrm{div}\widetilde{\phi} + O(\epsilon^2)$ [45, equation (2.10)]. Then the transformed conductivity $\sigma(x,\omega)$ can be explicitly written as

$$\begin{aligned}
\sigma(x,\omega) &= \widetilde{\sigma}(\cdot,\omega)(1+\epsilon\mathrm{div}\widetilde{\phi}(\cdot))^{-1}(I+\epsilon(J_{\widetilde{\phi}}(\cdot)+J_{\widetilde{\phi}}^t(\cdot)))\circ F^{-1}(x) + O(\epsilon^2)\\
&= \widetilde{\sigma}(\cdot,\omega)((1-\epsilon\mathrm{div}\widetilde{\phi}(\cdot))I+\epsilon(J_{\widetilde{\phi}}(\cdot)+J_{\widetilde{\phi}}^t(\cdot)))\circ F^{-1}(x) + O(\epsilon^2)\\
&= \widetilde{\sigma}(\cdot,\omega)\circ F^{-1}(x) + \Psi(x)\epsilon + O(\epsilon^2).
\end{aligned}$$

where $\Psi = (J_{\widetilde{\phi}} + J_{\widetilde{\phi}}^t - \mathrm{div}\widetilde{\phi}I)\circ F^{-1}$ is smooth and independent of the frequency ω. Upon collecting terms, this together with the separable form of $\widetilde{\sigma}(\widetilde{x},\omega)$ in (3.14) yields

$$\sigma(x,\omega) \approx s_0(\omega)I + \epsilon s_0(\omega)\Psi(x) + \sum_{k=0}^{K}\delta\sigma_k(x)s_k(\omega)I. \tag{3.22}$$

Upon substituting it into (3.21) and invoking the approximation $\nabla u_n \approx \nabla v_n$ in the domain, we obtain the desired expression. □

By Proposition 1, in the presence of an imperfectly known boundary with the deformation magnitude ϵ comparable with the inclusions $\{\delta\sigma_k\}_{k=0}^K$, there is one additional dominant source of errors in the linearized inverse problem: the perturbed sensitivity system contains an additional anisotropic component $\epsilon\Psi$, resulting from the domain deformation. As a consequence,

a direct inversion of the linearized model (3.18) is unsuitable. This is consistent with the empirical observation that a slightly incorrect boundary can lead to completely erroneous reconstructions [1, 34].

This issue can be resolved by using the proposed multifrequency approach. Indeed, by rearranging the terms in (3.18) we obtain

$$s_0(\omega)\int_\Omega(\epsilon\Psi+\delta\sigma_0)\nabla v_n^*\cdot\nabla v_m^* dx+\sum_{k=1}^{K} s_k(\omega)\int_\Omega \delta\sigma_k\nabla v_n^*\cdot\nabla v_m^* dx = s_0(\omega)^2\int_{\partial\Omega}(f_n v_m - f_m u_n)ds. \tag{3.23}$$

This equation is completely analogous to (3.5), with the only difference lying in the additional term $\epsilon\Psi$. Therefore, all the methods discussed in Section 3.2 may be applied straightforwardly, since the right hand side is known. The background perturbation $\delta\sigma_0$ will never be properly reconstructed, due to the pollution of the error term $\epsilon\Psi$. However, the inclusions corresponding to the other frequency profiles may be reconstructed, since they are affected by the deformation only through $\delta\sigma_k = \delta\widetilde{\sigma}_k \circ F^{-1}$. In other words, the location and shape can be slightly deformed. Thus we have shown that mfEIT is a very effective method to eliminate the modelling errors caused by the boundary uncertainty. Only the background anomaly (i.e., the inclusion $\delta\sigma_0$) is affected, and so cannot be reconstructed. All the other inclusions may be imaged successfully.

3.4 The Complete Electrode Model

In this section we adapt the approach discussed in Sections 3.2 and 3.3 to the more realistic complete electrode model, which has been shown to reproduce the experimental data within measurement precision [20] and is currently regarded as the most accurate model in a number of applications. We discuss the cases of a perfectly known and an imperfectly known boundary separately.

3.4.1 Perfectly Known Boundary

First we consider the case of a perfectly known boundary. Let Ω be an open bounded domain in $\mathbb{R}^d$ ($d = 2, 3$), with a smooth boundary $\partial\Omega$. We denote the set of electrodes by $\{e_j\}_{j=1}^E \subset \partial\Omega$, which are disjoint from each other, i.e., $\bar{e}_i \cap \bar{e}_k = \emptyset$ if $i \neq k$. The applied current on the jth electrode e_j is denoted by I_j, and the current vector $I = (I_1, \ldots, I_E)^t$ satisfies $\sum_{j=1}^E I_j = 0$ by the law of charge conservation. Let the space $\mathbb{R}^E_\diamond$ be the subspace of the vector space $\mathbb{R}^E$ with zero mean, i.e., $I \in \mathbb{R}^E_\diamond$. The electrode voltages $U = (U_1, \ldots, U_E)^t$ are also normalized so that $U \in \mathbb{R}^E_\diamond$. Then the mathematical formulation of the complete electrode model (CEM) reads [20, 102]: given the frequency-dependent conductivity distribution $\sigma(x, \omega)$, frequency-independent positive contact impedances $\{z_j\}_{j=1}^E$ and an input current pattern $I \in \mathbb{R}^E_\diamond$, find the potential $u(\cdot, \omega) \in H^1(\Omega)$ and the electrode

voltages $U \in \mathbb{R}^E_\diamond$ such that

$$\begin{cases} -\nabla \cdot (\sigma(x,\omega)\nabla u(x,\omega)) = 0 \text{ in } \Omega, \\ u + z_j \dfrac{\partial u}{\partial \nu_\sigma} = U_j \text{ on } e_j,\ j = 1,2,\dots,E, \\ \displaystyle\int_{e_j} \frac{\partial u}{\partial \nu_\sigma} ds = I_j \text{ for } j = 1,2,\dots,E, \\ \dfrac{\partial u}{\partial \nu_\sigma} = 0 \text{ on } \partial\Omega \backslash \cup_{j=1}^E e_j, \end{cases} \tag{3.24}$$

where $\frac{\partial u}{\partial \nu_\sigma}$ denotes the co-normal derivative $(\sigma\nabla u)\cdot\nu$. The second line describes the contact impedance effect: When injecting electrical currents into the object, a highly resistive thin layer forms at the electrode-electrolyte interface (due to certain electrochemical processes), which causes potential drops across this interface. The potential drop is described by Ohm's law, with the positive constants $\{z_j\}_{j=1}^E$ being contact impedances. In practice, it was observed that the contact impedances $\{z_j\}_{j=1}^E$ are inversely proportional to the conductivity of the object [48, 52], and thus we can write

$$z_j = s_0(\omega)^{-1} c_j, \tag{3.25}$$

for some constants $c_j > 0$ independently of the frequency, since by assumption, near the boundary $\partial\Omega$ we have $\sigma(x,\omega) = s_0(\omega)$. The metallic electrodes are perfect conductors, and hence the voltage U_j on the jth electrode e_j is a constant. The weak formulation of model (3.24) is given by: find $(u,U) \in \mathbb{H} := H^1(\Omega)\times\mathbb{R}^E_\diamond$ (equipped with the product norm) such that [102]

$$\int_\Omega \sigma(x,\omega)\nabla u(x,\omega)\cdot\nabla v(x)dx + \sum_{j=1}^E z_j^{-1}\int_{e_j}(u-U_j)(v-V_j)ds = \sum_{j=1}^E I_j U_j, \qquad (v,V)\in\mathbb{H}.$$

The bilinear form defined on the left hand side is coercive and continuous on the space $\mathbb{H}$, and thus by Lax-Milgram theorem there exists a unique solution $(u(\cdot,\omega), U(\omega)) \in \mathbb{H}$.

Consider M input currents $\{I_n\}_{n=1}^M \subset \mathbb{R}^E_\diamond$, and let $\{(u_n,U_n)\}_{n=1}^M \subset \mathbb{H}$ be the corresponding solutions to the complete electrode model (3.24), i.e.,

$$\int_\Omega \sigma(x,\omega)\nabla u_n(x,\omega)\cdot\nabla v(x,\omega)dx + \sum_{j=1}^E z_j^{-1}\int_{e_j}(u_n-U_{n,j})(v-V_j)ds = \sum_{j=1}^E I_{n,j}V_j, \qquad (v,V)\in\mathbb{H}. \tag{3.26}$$

The electrode voltages $U_n \in \mathbb{R}^E_\diamond$ can be measured in practice, and are used to recover the conductivity distribution $\sigma(x,\omega)$. To derive a linearized model, like before, let $(v_m, V_m) \in \mathbb{H}$ be the solution corresponding to the unperturbed conductivity field with $\sigma_0(x,\omega) = s_0(\omega)$:

$$\int_\Omega \sigma_0(x,\omega)\nabla v_m(x,\omega)\cdot\nabla v(x,\omega)dx + \sum_{j=1}^E z_j^{-1}\int_{e_j}(v_m-V_{m,j})(v-V_j)ds = \sum_{j=1}^E I_{m,j}V_j, \qquad (v,V)\in\mathbb{H}. \tag{3.27}$$

Like in the continuum model, in view of the relation (3.25), we can write

$(v_m^*, V_m^*) = s_0(\omega)(v_m, V_m)$ for the solution corresponding to the case $\sigma_0 \equiv 1$. Now we assume that the conductivity $\sigma(x,\omega)$ follows the separable form (3.2). Using the weak formulations for (u_n, U_n) and (v_m, V_m), we deduce immediately

$$\sum_{k=0}^{K} s_k(\omega) \int_\Omega \delta\sigma_k(x) \nabla u_n(x,\omega) \cdot \nabla v_m(x,\omega) dx = \sum_{j=1}^{E} (I_{n,j} V_{m,j} - I_{m,j} U_{n,j}).$$

Then, under the approximation $\nabla u_n \approx \nabla v_n$ in the domain Ω, and the piecewise constant approximation on the inclusions $\delta\sigma_k$s on the quasi-uniform triangulation $\{\Omega_l\}_{l=1}^L$ of the domain Ω, we have

$$\sum_{k=0}^{K} s_k(\omega) \sum_{l=1}^{L} (\delta\sigma_k)_l \int_{\Omega_l} \nabla v_n^* \cdot \nabla v_m^* dx = s_0(\omega)^2 \sum_{j=1}^{E} (I_{n,j} V_{m,j} - I_{m,j} U_{n,j}).$$

This formula is almost identical with that for the continuum model, cf. (3.5), and formally their only difference lies in the computation of the data vector $X(\omega)$. Hence, all the discussions in Section 3.2 can be adapted straightforwardly to the complete electrode model. In particular, all inversion methods discussed there can be directly applied to this case.

3.4.2 Imperfectly Known Boundary

Now we consider the case of an imperfectly known boundary. As in Section 3.3, let $\widetilde{\Omega}$ be the unknown true domain with a smooth boundary $\partial\widetilde{\Omega}$, and Ω be the computational domain with a smooth boundary $\partial\Omega$. Accordingly, let $\{\widetilde{e}_j\}_{j=1}^E \subset \partial\widetilde{\Omega}$ and $\{e_j\}_{j=1}^E \subset \partial\Omega$ be the real and computational electrodes, respectively and assume they satisfy the usual conditions discussed above. Then we introduce a sufficiently smooth orientation preserving forward map $F : \widetilde{\Omega} \to \Omega$, with a sufficiently smooth inverse $F^{-1} : \Omega \to \widetilde{\Omega}$, and we denote the restriction of F to the boundary $\partial\widetilde{\Omega}$ by $f : \partial\widetilde{\Omega} \to \partial\Omega$. We write $F^{-1}(x) = x + \epsilon\phi(x)$, where $\epsilon > 0$ denotes the magnitude of the domain deformation. For simplicity, further, it is assumed that there is no further electrode movement apart from domain deformation, i.e., $e_j = f(\widetilde{e}_j)$, $j = 1, \ldots, E$. With the frequency-dependent conductivity $\widetilde{\sigma}(\widetilde{x},\omega)$ of the separable form (3.14) and input current $I_n \in \mathbb{R}^E_\diamond$, by (3.24), the quantity $(\widetilde{u}_n(\widetilde{x},\omega), \widetilde{U}_n(\omega)) \in \widetilde{\mathbb{H}} \equiv H^1(\widetilde{\Omega}) \times \mathbb{R}^E_\diamond$ satisfies

$$\begin{cases} -\nabla_{\widetilde{x}} \cdot (\widetilde{\sigma}(\widetilde{x},\omega) \nabla_{\widetilde{x}} \widetilde{u}_n(\widetilde{x},\omega)) = 0 \quad \text{in } \widetilde{\Omega}, \\ \displaystyle\int_{\widetilde{e}_j} \frac{\partial \widetilde{u}_n}{\partial \widetilde{\nu}_{\widetilde{\sigma}}} d\widetilde{s} = I_{n,j} \quad \text{on } \widetilde{e}_j, j = 1, 2, \ldots, E, \\ \displaystyle z_j \frac{\partial \widetilde{u}_n}{\partial \widetilde{\nu}_{\widetilde{\sigma}}} + \widetilde{u}_n = \widetilde{U}_{n,j} \quad \text{on } \widetilde{e}_j, j = 1, 2, \ldots, E \\ \displaystyle\frac{\partial \widetilde{u}_n}{\partial \widetilde{\nu}_{\widetilde{\sigma}}} = 0 \quad \text{on } \partial\widetilde{\Omega} \setminus \cup_{j=1}^E \widetilde{e}_j, \end{cases} \tag{3.28}$$

The weak formulation of the problem is given by: find $(\widetilde{u}_n(\widetilde{x},\omega), \widetilde{U}_n(\omega)) \in \widetilde{\mathbb{H}}$ such that

$$\int_{\widetilde{\Omega}} \widetilde{\sigma}(\widetilde{x},\omega)\nabla_{\widetilde{x}}\widetilde{u}_n(\widetilde{x},\omega)\cdot\nabla_{\widetilde{x}}\widetilde{v}(\widetilde{x})d\widetilde{x} + \sum_{j=1}^{E} z_j^{-1}\int_{\widetilde{e}_j}(\widetilde{u}_n(\widetilde{x},\omega)-\widetilde{U}_{n,j}(\omega))(\widetilde{v}(\widetilde{x})-\widetilde{V}_j)d\widetilde{s} = \sum_{j=1}^{E} I_{n,j}\widetilde{V}_j, \qquad (\widetilde{v},\widetilde{V})\in\widetilde{\mathbb{H}}.$$

In the experimental setting, on the computational domain Ω, the injected current $I_n \in \mathbb{R}^E$ on the electrodes $\{e_j\}_{j=1}^E$ is known, and the corresponding voltage $\widetilde{U}_n(\omega) \in \mathbb{R}^E$ can be measured. The inverse problem is to recover the inclusion profiles $\{\delta\widetilde{\sigma}_k\}_{k=0}^K$ from the measured electrode voltages $\{\widetilde{U}_n(\omega)\}_{n=1}^M$ at a number of frequencies $\{\omega_q\}_{q=1}^Q$.

Now we can state the corresponding linearized inverse problem for the complete electrode model with an imperfectly known boundary. Consider the potential $u_n(\cdot,\omega) = \widetilde{u}_n(\cdot,\omega)\circ F^{-1}$, and the associated electrode voltages $U_n = \widetilde{U}_n$.

Proposition 2. *Let the reference solutions $(v_m, V_m) \in \mathbb{H}$ be defined by* (3.27) *and the conductivity $\widetilde{\sigma}$ be of the form* (3.14)*. Set $z = |\det J_{F^{-1}}^S|$, $\delta\sigma_k = \delta\widetilde{\sigma}_k \circ F^{-1}$ for $k = 0,1,\dots,K$ and $(v_m^*, V_m^*) = s_0(\omega)(v_m, V_m)$ for $m = 1,\dots,M$. The linearized inverse problem on the computational domain Ω is given by*

$$\begin{aligned} &s_0(\omega)\epsilon\int_\Omega \Psi\nabla v_n^*\cdot\nabla v_m^* dx + \sum_{k=0}^{K} s_k(\omega)\int_\Omega \delta\sigma_k\nabla v_n^*\cdot\nabla v_m^* dx \\ &= s_0(\omega)^2\sum_{j=1}^{E}(I_{n,j}V_{m,j} - I_{m,j}U_{n,j}) - s_0(\omega)\sum_{j=1}^{E} c_j\int_{e_j}(z-1)\left(\frac{\partial v_m^*}{\partial\nu}\right)^2 ds. \end{aligned} \tag{3.29}$$

for some smooth function $\Psi : \Omega \to \mathbb{R}^{d\times d}$, which is independent of the frequency ω.

Proof. Proceeding as in the proof of Proposition 1, by a change of variables (and suppressing the variable ω), since $e_j = f(\widetilde{e}_j)$ we deduce

$$\begin{aligned} &\int_{\widetilde{\Omega}} \widetilde{\sigma}(\widetilde{x})\nabla_{\widetilde{x}}\widetilde{u}_n(\widetilde{x})\cdot\nabla_{\widetilde{x}}\widetilde{v}(\widetilde{x})d\widetilde{x} \\ &= \int_\Omega (\widetilde{\sigma}\circ F^{-1})(x)(J_F^t\circ F^{-1})(x)\nabla u_n\cdot(J_F^t\circ F^{-1})(x)\nabla v(x)|\det J_{F^{-1}}(x)|dx. \\ &\int_{\widetilde{e}_j}(\widetilde{u}_n - \widetilde{U}_{n,j})(\widetilde{v}-\widetilde{V}_j)d\widetilde{s} = \int_{e_j}(u_n - U_{n,j})(v-V_j)|\det J_{F^{-1}}^S|ds, \end{aligned}$$

where $v = \widetilde{v}\circ F^{-1} \in H^1(\Omega)$ and $V_j = \widetilde{V}_j$. Hence, the pair $(u_n(\cdot,\omega), U_n(\omega))$ satisfies

$$\int_\Omega \sigma(x,\omega)\nabla u_n(x,\omega)\cdot\nabla v(x)dx + \sum_{j=1}^{E} z_j^{-1}\int_{e_j}(u_n - U_{n,j})(v-V_j)z\,ds = \sum_{j=1}^{E} I_{n,j}V_j, \qquad (v,V)\in\mathbb{H},$$

where the transformed conductivity $\sigma(x,\omega)$ is given by (3.19). By combining this identity with (3.27), we obtain

$$\int_\Omega (\sigma-\sigma_0)\nabla u_n\cdot\nabla v_m dx = \sum_{j=1}^{E}(I_{n,j}V_{m,j}-I_{m,j}U_{n,j})+\sum_{j=1}^{E}\int_{e_j}(z-1)(u_n-U_{n,j})\frac{\partial v_m}{\partial \nu_{\sigma_0}}ds.$$

In view of [45, 44] we can expand z as

$$z = 1+\epsilon(\mathrm{Div}\phi_t-(d-1)H\phi_\nu)+O(\epsilon^2), \tag{3.30}$$

where Div denotes the surface divergence, ϕ_t and ϕ_ν denote the tangential and normal components of the vectorial function ϕ on the boundary $\partial\Omega$, respectively, and H is the mean curvature of $\partial\Omega$. In particular, $z-1=O(\epsilon)$. Therefore, by linearization we can write

$$\int_{e_j}(z-1)(u_n-U_{n,j})\frac{\partial v_m}{\partial\nu_{\sigma_0}}ds \approx \int_{e_j}(z-1)(v_n-V_{n,j})\frac{\partial v_m}{\partial\nu_{\sigma_0}}ds = -z_j\int_{e_j}(z-1)\left(\frac{\partial v_m}{\partial\nu_{\sigma_0}}\right)^2 ds.$$

Inserting this approximation in the above identity we obtain

$$\int_\Omega (\sigma-\sigma_0)\nabla u_n\cdot\nabla v_m dx = \sum_{j=1}^{E}(I_{n,j}V_{m,j}-I_{m,j}U_{n,j})-\sum_{j=1}^{E}z_j\int_{e_j}(z-1)\left(\frac{\partial v_m}{\partial\nu_{\sigma_0}}\right)^2 ds.$$

The rest of the proof follows as in Proposition 1, and thus it is omitted. □

By proceeding as in the continuum model, we can rewrite (3.29) as

$$\begin{aligned} &s_0(\omega)\int_\Omega(\epsilon\Psi+\delta\sigma_0)\nabla v_n^*\cdot\nabla v_m^* dx+\sum_{k=1}^{K}s_k(\omega)\int_\Omega\delta\sigma_k\nabla v_n^*\cdot\nabla v_m^* dx \\ &= s_0(\omega)^2\sum_{j=1}^{E}(I_{n,j}V_{m,j}-I_{m,j}U_{n,j})-s_0(\omega)\sum_{j=1}^{E}c_j\int_{e_j}(z-1)\left(\frac{\partial v_m^*}{\partial\nu}\right)^2 ds. \end{aligned} \tag{3.31}$$

When compared with the linearized model in the continuum case, cf. (3.23), we observe the presence of the additional error term $s_0(\omega)C_m$, where

$$C_m(\omega) := -\sum_{j=1}^{E}c_j\int_{e_j}(z-1)\left(\frac{\partial v_m^*}{\partial\nu}\right)^2 ds,$$

that comes from the boundary deformation. The formula (3.31) is perfectly consistent with (3.23): in the continuum case, the contact impedance effect is not taken into account, and $u_n = U_n$ on the electrodes, namely $c_j = 0$, whence $C_m = 0$.

All the preceding analysis easily carries forward to the case $c_j > 0$, as we now discuss. Before treating the general case, let us consider the simple scenario where $z\equiv 1$ on the electrodes $\cup_j e_j$.

Example 5. *Recall that $z(x) = |\det J^S_{F^{-1}}(x)|$ for x on the boundary $\partial\Omega$. Physically, the factor z represents the length deformation relative to the map $F^{-1}\colon \partial\Omega\to$*

$\partial\widetilde{\Omega}$, as we remarked earlier. Thus, it may be reasonable to assume that the parametrization of the electrodes $\{e_j\}_{j=1}^E$ is known, which implies $z \equiv 1$ on the electrodes $\cup_j e_j$.

In such a case, we immediately obtain $C_m \equiv 0$. whence

$$s_0(\omega)\int_\Omega(\epsilon\Psi+\delta\sigma_0)\nabla v_n^*\cdot\nabla v_m^* dx+\sum_{k=1}^K s_k(\omega)\int_\Omega \delta\sigma_k\nabla v_n^*\cdot\nabla v_m^* dx = s_0(\omega)^2\sum_{j=1}^E(I_{n,j}V_{m,j}-I_{m,j}U_{n,j}).$$

This identity is completely analogous to (3.23)*, and the same comments on the reconstruction procedure are valid here, since the right hand side is known. In particular, by applying any of the techniques discussed in Section 3.2 to multifrequency measurements, it is possible to eliminate the error coming from the domain perturbation $\epsilon\Psi$, as this affects only the inclusion $\delta\sigma_0$ (corresponding to s_0). All the other inclusions $\delta\sigma_k$, $k = 1, \ldots, K$, may be successfully reconstructed.*

Now we consider the general case when $z \not\equiv 1$ on the electrodes $\cup_j e_j$. This corresponds to a situation where the length (or the area) of the electrodes is not precisely known. Thus, the additional error term $s_0(\omega)C_m$ in the linearized model (3.31) has to be taken into account. The key observation is that C_m is independent of the frequency ω. The difference imaging method discussed in Section 3.2.2 may be directly applied here, provided that $0 \notin \mathcal{P}$, i.e., if the frequency profile $s_0(\omega)$ does not vary substantially with respect to the frequency ω. Indeed, in this case the error term $s_0(\omega)C_m$ disappears upon differentiating the relation (3.31), and the inversion step may be performed as in Section 3.2.2.

The method of Section 3.2.1 (and, thus, the particular case discussed in Section 3.2.3) may be also directly applied, since the dependence of the error term $s_0(\omega)C_m$ with respect to the frequency ω follows exactly the spectral profile $s_0(\omega)$. Namely, its influence on the reconstruction step can be essentially lumped into the component $\delta\sigma_0$, like the conductivity perturbation $\epsilon\Psi$ discussed earlier. Thus, all the inclusions $\delta\sigma_k$, $k = 1, \ldots, K$, corresponding to the remaining frequency profiles $s_1, \ldots, s_K$ may be completely reconstructed. Alternatively, one may see this from the linear system as follows. When multiplying the right hand side of the system of equations associated to (3.31) by S^{-1}, the error term $s_0(\omega)C_m$ cancels out in all the systems $MA_k = Y_k$, for $k = 1, \ldots, K$. This follows by elementary linear algebra, since

$$\begin{aligned} S^{-1} &= C\,[s_0(\omega_1)\,,\,\ldots\,,\,s_0(\omega_Q)]\begin{bmatrix} s_0(\omega_1) & \cdots & s_0(\omega_Q) \\ \vdots & \vdots & \vdots \\ s_K(\omega_1) & \cdots & s_K(\omega_Q)\end{bmatrix}^{-1} \\ &= [C\,,\,0\,,\,\ldots\,,\,0], \end{aligned}$$

where the notation C denotes the column vector corresponding to the error terms C_m (see Example 1 for the simple case when $Q = 2$ and $K = 1$).

3.5 Group Sparse Reconstruction Algorithm

For all the scenarios discussed in the previous sections, one arrives at a number of (decoupled) linear systems

$$MA_k = Y_k \quad k = 0, \ldots, K, \tag{3.32}$$

where $M \in \mathbb{R}^{J\times L}$ is the sensitivity matrix, $A_k \in \mathbb{R}^L$ are the unknown abundances, and $Y_k \in \mathbb{R}^J$ is a known piece of data. These linear systems are often under-determined, and severely ill-conditioned, due to the inherent ill-posed nature of the EIT inverse problem. Below we describe one strategy for the stable and accuracy solution of the linear system (3.32) based on the idea of regularization; we refer to [93, 95, 53] for general discussions on regularization methods.

There are several natural aspects to take into consideration for the regularization term, especially sparsity, grouping, disjoint sparsity and bound constraints.

(1) For $k = 0, 1, 2, \ldots, K$, we can assume that the abundances $A_k = (\delta\sigma_k)_l \in \mathbb{R}^L$ are sparse with respect to the pixel basis (piecewise constant approximation). This suggests minimizing

$$\min_{A_k\in\Lambda} \|A_k\|_1 \quad \text{subject to } \|MA_k - Y_k\| \le \epsilon_k$$

for each $k = 0, 1, \ldots, K$. Here $\|\cdot\|_1$ denotes the ℓ^1 norm of a vector. The set Λ represents a box constraint on the unknown vector A_k, since the conductivity σ remains bounded from below and above by positive constants, due to physical constraint, and $\epsilon_k > 0$ is an estimate of the noise level in the data Y_k. This ℓ^1 optimization problem can be solved efficiently by many algorithms, e.g., iterative soft thresholding.

(2) In EIT applications, it is also reasonable to assume that each abundance A_k is clustered, and this refers to the concept of group sparsity. The grouping effect is useful to remove the undesirable spikes typically observed for the ℓ^1 penalty alone. There are several different approaches to this task. The elastic net [57] is one simple way to realize grouping. In this work, we shall develop an approach easy to implement, inspired by the dynamic group sparsity proposed in [51]. It allows to dynamically realize group sparsity without knowing the supports of the A_ks nor their sizes.

(3) The support of the inclusions A_k are assumed to be disjoint from each other. The disjoint supports of A_ks can be simply promoted by adding a term that penalizes the scalar product of the absolute values of the A_ks, as was done in [113].

Remark 3.5.1. *Note that (1) and (2) refer to methods in which the abundances A_ks are recovered separately, while (3) to methods in which all the A_ks are reconstructed simultaneously.*

Next we construct an efficient iterative algorithm, termed as group iterative soft thresholding, for achieving the goals outlined above. It combines the strengths of the classical iterative soft thresholding algorithm [24] and the grouping effect in the dynamical group sparse recovery [51]: the former is easy to implement and has a built-in regularizing effect, whereas the latter encourages the group sparsity pattern. In implementation, it is a simple modification of the classical iterative soft thresholding algorithm for ℓ^1 optimization (by omitting the subscript i in the abundance): given an initial guess A^0, construct an approximation iteratively by

$$A^{j+1} = S_{s^j\alpha}(g^j),$$

where the proxy g^j is defined by

$$g^j = A^j - s^j M^t(MA^j - Y). \tag{3.33}$$

Observe that $M^t(MA^j - Y)$ is the gradient of the fitting term $\frac{1}{2}\|MA - Y\|^2$ at the current iterate A^j, and hence g^j is essentially a gradient descent update of the current reconstruction A^j. The scalar $\alpha > 0$ is a regularization parameter and $s^j > 0$ is the step length. One simple choice of the step size s^j is the constant one $s^j = 1/\|M\|^2$, which ensures the convergence of the algorithm [24]. The soft thresholding operator S_λ for $\lambda > 0$ is defined by

$$S_\lambda(t) = \max(|t| - \lambda, 0)\,\mathrm{sign}(t),$$

and it is applied componentwise when the argument is a vector.

In the proposed group iterative soft thresholding algorithm, instead of performing the thresholding on the proxy g^j directly, we take into account the neighboring influence. Following [51], this is easily achieved by computing a generalized proxy d_l^j of the lth element by

$$d_l^j = |g_l^j|^2 + \sum_{k\in\mathcal{N}_l} w_{lk}|g_k^j|^2, \tag{3.34}$$

where w_{lk} are nonnegative weights, and $\mathcal{N}_l$ denotes the neighborhood of the lth element. The weights w_{lk} determine the strength of correlation between the components: The smaller the magnitude of w_{lk} is, the weaker the correlation strength between the lth and the kth components is, and if $w_{lk} = 0$ for all $k \in \mathcal{N}_l$, it does not encourage grouping at all. In our implementation, we take $w_{lk} = \beta$, for some constant $\beta > 0$, for all neighboring elements. Physically, in EIT, the neighborhood $\mathcal{N}_l$ of the lth element consists of all elements in the triangulation that share one edge with the lth element, and may be expanded to include also elements sharing one node. Then the vector d^j is used to reweigh the thresholding step by

$$\bar{d}^j = \max(d^j)^{-1} d^j, \tag{3.35}$$

The quantity $\bar{d}^j$ indicates a normalized grouping effect: the larger is $\bar{d}_l^j$, the more likely the lth element belongs to the group, and thus the less thresholding should be applied to it. This can be easily achieved by rescaling the regularization parameter α to be inversely proportional to $\bar{d}_l^j$, with

$$\bar{\alpha}_l^j = \alpha/\bar{d}_l^j, \quad l = 1, \ldots, L, \tag{3.36}$$

and last perform the projected thresholding with a spatially variable regularization parameter $\bar{\alpha}^j$

$$A^{j+1} = P_\Lambda(S_{s^j\bar{\alpha}^j}(g^j)). \tag{3.37}$$

where P_Λ denotes the pointwise projection onto the constraint set Λ. The complete procedure is listed in Algorithm 3.5.1. Here N is the maximum number of iterations. Since the solution A is expected to be sparse, a natural choice of the initial guess A^0 is the zero vector. The regularization parameter α plays a crucial role in the performance of the reconstruction quality: the larger the value α is, the sparser the reconstructed abundance is. There

are several strategies available for the classical iterative soft thresholding, e.g., discrepancy principle and balancing principle. In this work, we shall test the feasibility of the algorithm only with the regularization parameter α determined in a trial-and-error manner. One can terminate the algorithm by monitoring the relative change of the iterates.

Algorithm 3.5.1. GROUP ITERATIVE SOFT THRESHOLDING.

1: *Input* $M, Y, W, \mathcal{N}, \alpha, N$ *and* A^0.
2: ***for*** $j = 1, \ldots, N$ ***do***
3: *Compute the proxy* g^j *by* (3.33).
4: *Compute the generalized proxy* d^j *by* (3.34).
5: *Compute the normalized proxy* $\bar{d}^j$ *by* (3.35).
6: *Adapt the regularization parameter* $\bar{\alpha}^j$ *by* (3.36).
7: *Update the abundance* A^{j+1} *by the group thresholding* (3.37).
8: *Check the stopping criterion.*
9: ***end for***

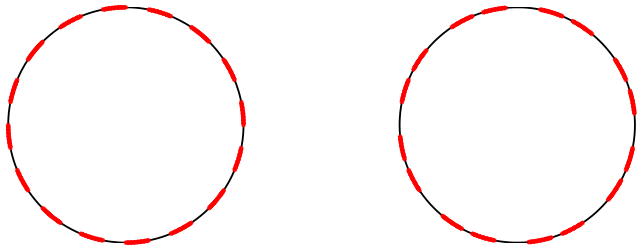
Last, we note that the disjoint sparsity can be enforced conveniently in Algorithm 3.5.1. Specifically, we first compute the normalized proxy $\bar{d}^{k,j}$ for the abundance A_k separately according to (3.35), and then at each component $l = 1, \ldots, L$, we update them by

$$\bar{d}_l^{k,j} = \begin{cases} \bar{d}_l^{k,j} & \text{if } k = k_l^*, \\ \varepsilon & \text{otherwise,} \end{cases} \qquad k_l^* = \underset{k=0,\ldots,K}{\arg\max}\ \bar{d}_l^{k,j},$$

where $\varepsilon > 0$ is a small number to avoid numerical overflow. This step effectively only retains the most likely abundance component (with the likelihood for the kth abundance given by $\bar{d}^{k,j}$), and sets to zero all the remaining ones. Hence, it enforces the disjoint sparsity as desired.

3.6 Numerical Experiments and Discussions

In this section we present some numerical results to illustrate the analytic study in Sections 3.2-3.4. We present results only for the complete electrode model (3.24), since the results for the simpler continuum model (3.1) are similar. The general setting for the numerical experiments below is as follows. The computational domain is taken to be the unit circle $\Omega = \{(x_1, x_2) : x_1^2 + x_2^2 < 1\}$. There are sixteen electrodes $\{e_j\}_{j=1}^E$ (with $E = 16$) evenly distributed along the boundary $\partial\Omega$, each of length $\pi/16$, thus occupying one half of the boundary $\partial\Omega$; see Fig. 3.1(a) for a schematic illustration of the electrode placement. The contact impedances $\{z_j\}_{j=1}^E$ on the electrodes $\{e_j\}_{j=1}^E$ are all set to unit, and the background conductivity σ_0 is taken to be $\sigma_0 \equiv 1$. Further, we shall always assume that the spectral profile $s_0(\omega)$ for the background is the constant one $s_0 \equiv 1$. This is not a restriction, since $s_0(\omega)$ is always known, and one can rescale the spectral profiles so that $s_0 \equiv 1$. We measure the electrode voltages U for all 15 sinusoidal input currents. The complete electrode model (3.24) is discretized using a piecewise linear finite element method on a shape regular quasi-uniform triangulation of the domain Ω [33]. The conductivity is represented on a coarser finite element mesh using a piecewise constant finite element basis. The electrode voltages are generated on a much finer mesh in order to avoid



(a) computational domain Ω (b) imperfectly known domain $\widetilde{\Omega}$

FIGURE 3.1: The electrode arrangement for the computational domain Ω and for an imperfectly known domain $\widetilde{\Omega}$ (used in Example 9). The curved segments in red denote the electrodes. The electrodes in (a) are equally spaced, but those in (b) are not.

the inverse crime. Then the noisy data U^δ is generated by adding componentwise Gaussian noise to the exact data $U^\dagger := U(\sigma^\dagger)$ corresponding to the true conductivity $\sigma^\dagger(x,\omega)$ as follows

$$U_j^\delta = U_j^\dagger + \epsilon \max_l |U_l^\dagger - U_l(\sigma_0)|\varepsilon_j, \quad j = 1,\dots,E,$$

where ϵ is the noise level, and ε_j follow the standard normal distribution.

We shall present the numerical results for the cases of known boundary and of imperfectly known boundary separately, and discuss only cases a) and b) with spectral profiles that are either fully known or have substantially different frequency dependence. Case c), corresponding to the case of partially known spectral profiles, will not be discussed, since the inversion is totally analogous to that of case a), except for simple algebraic manipulations. For the solutions of the underdetermined linear systems (3.32), we use the group iterative soft thresholding algorithm listed in Algorithm 3.5.1 with a constant step size. The regularization parameter α used in each separate reconstruction was determined by a trial-and-error manner, and it was set to 10^{-2} for all examples presented below. We did not implement the disjoint sparsity, since in all the examples under consideration the reconstruction are already very satisfactory. However, it is expected that with higher noise levels or in the case of almost touching inclusions, enforcing disjoint sparsity might give enhanced reconstructions. The algorithm is always initialized with a zero vector. Numerically, we observe that it converges steady and fast. All the computations were performed using MATLAB 2013a on a 2.5G Hz and 6G RAM personal laptop.

3.6.1 Perfectly Known Boundary

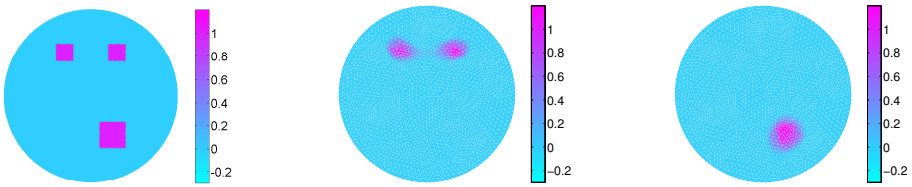
First, we illustrate the approach in the case of a perfectly known boundary.

Example 6. *Consider three square inclusions; the two inclusions on the top share the same spectral profile $s_1(\omega)$, and the one on the bottom has a second spectral profile $s_2(\omega)$, cf. Fig. 3.2(a) for an illustration. In the experiments, we consider the following two cases:*

(i) The spectral profiles $s_1(\omega)$ *and* $s_2(\omega)$ *are given by* $s_1(\omega) = 0.1\omega + 0.1$ *and* $s_2(\omega) = 0.2\omega$.

(ii) The spectral profiles $s_1(\omega)$ *and* $s_2(\omega)$ *are given by* $s_1(\omega) = 0.1\omega + 0.1$ *and* $s_2(\omega) = 0.02\omega$.

In either case, we take measurements at $Q = 3$ *frequencies,* $\omega_1 = 0$, $\omega_2 = 0.5$ *and* $\omega_3 = 1$.

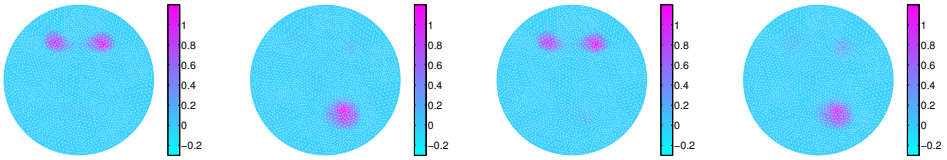


(a) true $\delta\sigma_k$s (b) recovered $\delta\sigma_1$ (c) recovered $\delta\sigma_2$

FIGURE 3.2: Numerical results for Example 6(i) with 1% noise in the data, and fully known spectral profiles. The reconstructions were obtained using the direct approach discussed in Section 3.2.1.

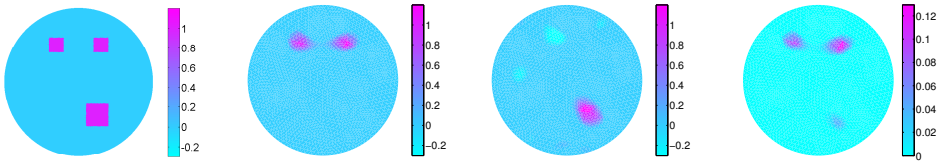
The numerical results for Example 6 with $\epsilon = 1\%$ noise in the data are shown in Figs. 3.2 and 3.4 for cases (i) and (ii), respectively. In case (i), the two frequencies have about the same magnitude, and the (rectangular) spectral matrix S is nonsingular. Assuming the knowledge of the spectral profiles, the direct approach discussed in Section 3.2.1 separates well the two sets of inclusions thanks to their incoherent spectral profiles. The recovery is very localized within a clean background, the supports match closely the true supports (and are clearly disjoint from each other) and the magnitude of the inclusions are well retrieved. The latter observation is a distinct feature of the proposed group sparse recovery algorithm discussed in Section 3.5. Hence, if both profiles are known exactly and incoherent, then the two sets of inclusions can be fairly recovered. Case (ii) is similar, except that the variation of the spectral profile $s_2(\omega)$ is now much smaller. The preceding observations remain largely valid, except that the inclusion corresponding to $s_2(\omega)$ involves minor spurious oscillations. This is attributed to the presence of noise in the data: the noise level is comparable with effective contributions from the inclusion. Hence, for the accurate recovery of the inclusions separately, the data should be reasonably accurate, as expected.

The well-conditioning of the spectral profile matrix S implies the robustness of the direct approach with respect to perturbations of the spectral profiles, as mentioned in Section 3.2.1. To confirm this, we present in Fig. 3.3 the reconstructions using imprecisely known spectral profiles, where the spectral matrix is perturbed by additive Gaussian noise with a zero mean and standard deviation proportional to the entry magnitude. Even only with three modulating frequencies, the reconstructions remain fairly stable, up to 20% perturbation of the spectral profiles, indicating the robustness of the approach. This is consistent with the experimental findings in [75], where the feasibility of the abundance separation with imprecise spectral profiles was numerically demonstrated.



(a) recovered $\delta\sigma_1$ (b) recovered $\delta\sigma_2$ (c) recovered $\delta\sigma_1$ (d) recovered $\delta\sigma_2$

FIGURE 3.3: Numerical results for Example 6(i) with 1% noise in the data, and imprecisely known spectral profiles. The reconstructions in (a) and (b) are obtained with the spectral matrix S perturbed by additive Gaussian noise with mean zero and standard deviation 10% of the entry magnitude, and those in (c) and (d) with 20% noise, both by the direct approach of Section 3.2.1.



(a) true $\delta\sigma_k$s (b) recovered $\delta\sigma_1$ (c) recovered $\delta\sigma_2$ (d) recovered $\delta\sigma_1$

FIGURE 3.4: Numerical results for Example 6(ii) with 1% noise in the data. The reconstructions in panels (b) and (c) are obtained with fully known spectral profiles using the direct approach in Section 3.2.1, and that in panel (d) is obtained without knowing the spectral profiles, using difference imaging in Section 3.2.2.

With Example 6(ii), we also illustrate difference imaging discussed in Section 3.2.2. Since the variation of the frequency profile $s_2(\omega)$ is small, this technique is also applicable, and the reconstruction of the first set of inclusions, in the absence of the knowledge of the spectral profiles, is shown in Fig. 3.4(d). The reconstruction recovers the inclusions up to a multiplicative constant very well, and it is almost completely free from spurious oscillations. This clearly shows the capability of difference imaging in Section 3.2.2 for spectral profiles with substantially different frequency dependence.

Example 7. *Consider three rectangular inclusions on the top left, top right and bottom with spectral profiles $s_1(\omega)$, $s_2(\omega)$ and $s_3(\omega)$, respectively, cf. Fig. 3.5a for an illustration. In the experiments, we consider the following two cases:*

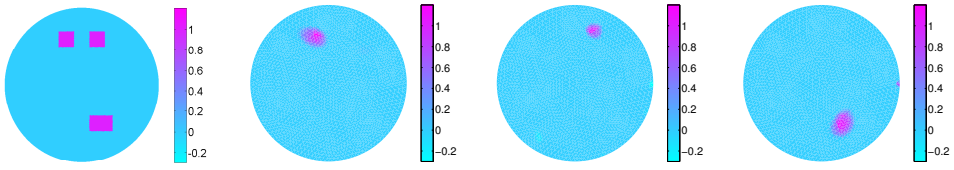
(i) The spectral profiles $s_1(\omega)$, $s_2(\omega)$ and $s_3(\omega)$ are given by $s_1(\omega) = 0.2\omega + 0.2$, $s_2(\omega) = 0.1\omega^2$, and $s_3(\omega) = 0.2\omega + 0.1$, respectively.

(ii) The spectral profiles $s_1(\omega)$, $s_2(\omega)$ and $s_3(\omega)$ are given by $s_1(\omega) = 0.02\omega + 0.02$ and $s_2(\omega) = 0.1\omega^2$, and $s_3(\omega) = 0.2\omega + 0.1$, respectively.

In either case, we take measurements at three frequencies, $\omega_1 = 0$, $\omega_2 = 0.5$ and $\omega_3 = 1$.

The numerical results for Example 7(i) and 7(ii) are shown in Figs. 3.5 and 3.6, respectively. If all three spectral profiles are known, the use of three frequencies yields almost perfect separation of the three inclusions by using the method of Section 3.2.1: the recovered inclusions are well clustered in a clean background, and their supports and magnitudes are correctly identified. Note that in the case of Example 7(ii), the spectral profile $s_1(\omega)$ is much smaller, and thus the recovery of the inclusion $\delta\sigma_1$ is more susceptible to noise, whereas the recovery of the remaining two are far more stable.

The results in Fig. 3.6 indicate that with known spectral profiles $s_2(\omega)$ and $s_3(\omega)$ and unknown $s_1(\omega)$, since $s_1(\omega)$ varies little with respect to ω, the difference imaging technique proposed in Section 3.2.2 can recover accurately both the magnitude and support of the inclusions $\delta\sigma_2$ and $\delta\sigma_3$. These observations fully confirm the discussions in Section 3.2.2.

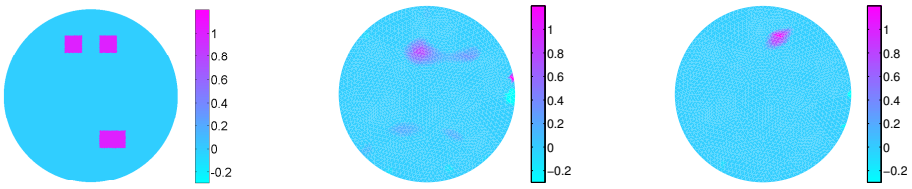


(a) true $\delta\sigma_k$s (b) recovered $\delta\sigma_1$ (c) recovered $\delta\sigma_2$ (d) recovered $\delta\sigma_3$

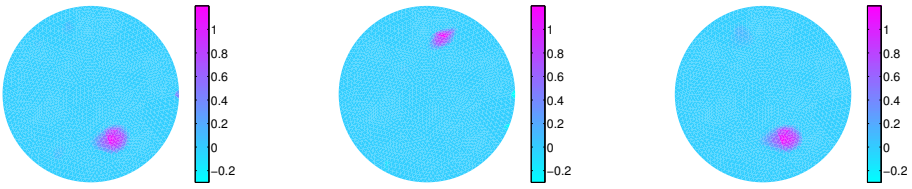
FIGURE 3.5: Numerical results for Example 7(i) with 1% noise in the data, with fully known spectral profiles. The reconstructions are obtained by the direct approach in Section 3.2.1.

3.6.2 Imperfectly Known Boundary

Now we illustrate the approach in the case of an imperfectly known boundary. In the first example, the unknown true domain $\widetilde{\Omega}$ is an ellipse centered



(a) true $\delta\sigma_k$s (b) recovered $\delta\sigma_1$ (c) recovered $\delta\sigma_2$



(d) recovered $\delta\sigma_3$ (e) recovered $\delta\sigma_2$ (f) recovered $\delta\sigma_3$

FIGURE 3.6: Numerical results for Example 7(ii) with 1% noise in the data. Here (b)-(d) are the reconstructions with fully known spectral profiles, while for (e) and (f) only the spectral profiles $s_2(\omega)$ and $s_3(\omega)$ are known, and the reconstructions are obtained by difference imaging in Section 3.2.2.

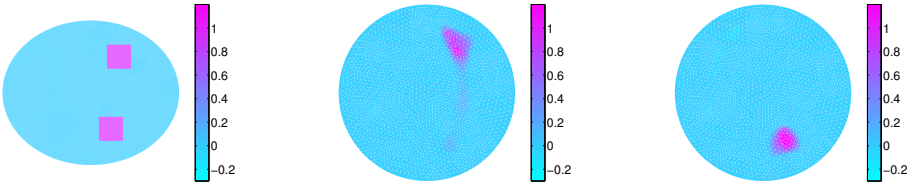
at the origin with semi-axes a and b, $\mathcal{E}_{a,b} = \{(x_1, x_2) : x_1^2/a^2 + x_2^2/b^2 < 1\}$, and the computational domain Ω is taken to be the unit circle.

Example 8. *Consider two square inclusions, on the top and on the bottom of the ellipse, with spectral profiles $s_1(\omega) = 0.2\omega + 0.2$ and $s_2(\omega) = 0.1\omega^2$, respectively. We consider the following two cases:*

(i) The true domain $\widetilde{\Omega}$ is $\mathcal{E}_{a,b}$ with $a = 1.1$ and $b = 0.9$;

(ii) The true domain $\widetilde{\Omega}$ is $\mathcal{E}_{a,b}$ with $a = 1.2$ and $b = 0.8$.

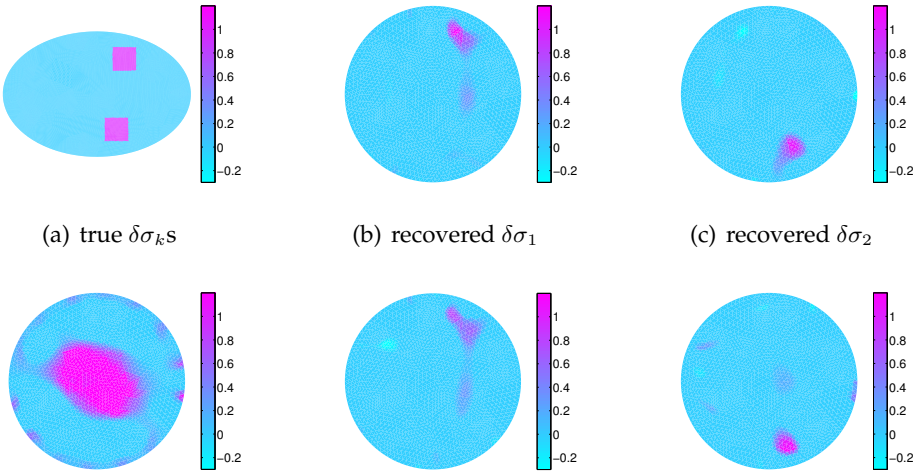
In either case, we take three frequencies, $\omega_1 = 0$, $\omega_2 = 0.5$ and $\omega_3 = 1$.



(a) true $\delta\sigma_k$s (b) recovered $\delta\sigma_1$ (c) recovered $\delta\sigma_2$

FIGURE 3.7: Numerical results for Example 8(i) with 0.1% noise in the data, fully known spectral profiles. The reconstructions are obtained using difference imaging technique in Section 3.2.2.

The numerical results are given in Figs. 3.7 and 3.8 with 0.1% noise in the data, for cases (i) and (ii), respectively. Even though not presented, we note that a direct application of the classical EIT imaging technique can only



(a) true $\delta\sigma_k$s (b) recovered $\delta\sigma_1$ (c) recovered $\delta\sigma_2$

(d) recovered $\delta\sigma_0$ (e) recovered $\delta\sigma_1$ (f) recovered $\delta\sigma_2$

FIGURE 3.8: Numerical results for Example 8(ii) with 0.1% noise in the data, fully known spectral profiles. The reconstructions in panels (b)-(c) are based on difference imaging technique in Section 3.2.2, and those in panels (d)-(f) are based on the direct approach in Section 3.2.1.

produce useless reconstruction, due to the presence of significant modelling errors. Numerically one can verify that for both cases, the contribution from domain deformation is actually one order of magnitude larger than that due to the genuine inclusions, which justify the much smaller noise level 0.1%. By exploiting the frequency incoherence, the multi-frequency EIT allows the separation of contributions from different abundances, and hence recovering each inclusion accurately.

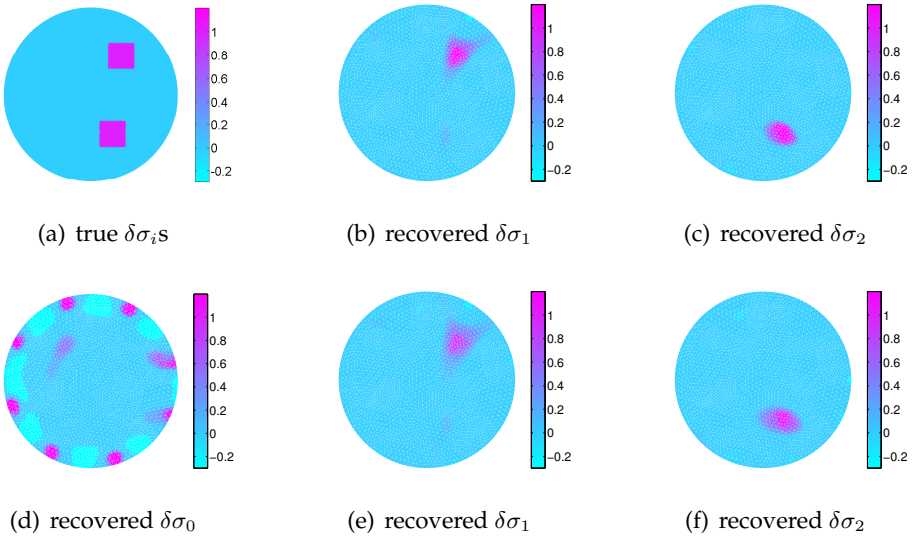
From Fig. 3.7, we observe that the difference imaging from Section 3.2.2 can recover the two inclusions accurately, and further, the two inclusions can be separated, due to their incoherent spectral profiles. However, the shape of the recovered inclusion tends to be slightly deformed and location slightly shifted. This is consistent with the discussions in Section 3.4: the unknown boundary induces a slightly deformed conductivity of the inclusions, in addition to the anisotropic component.

In Fig. 3.8 we present the results related to Example 8(ii). The preceding observations on difference imaging still hold, cf. Fig. 3.8(a) and (b). The direct approach of Section 3.2.1 works equally well: the recovered $\delta\sigma_1$ and $\delta\sigma_2$ are fairly accurate in terms of the location and magnitude; and the results are comparable with those obtained by difference imaging. Surely, the recovered $\delta\sigma_0$ contains only the spurious conductivity induced by the domain deformation. Should there be any true inclusion $\delta\sigma_0$ corresponding to the spectral profile $s_0(\omega)$, it will be washed away by the deformation error $\epsilon\Psi$ in (3.31). The preceding discussions fully confirm the analysis in Section 3.4: the multifrequency approach is capable of discriminating the perturbation due to domain deformation from the genuine inclusions by either the direct reconstruction in Section 3.2.1 or the difference imaging in Section 3.2.2.

Last we present one example where the electrodes are misplaced, but

the length of the electrodes do not change, i.e., the factor z in the boundary integral can be set to the unit (see Example 5). This is a special case of the imperfectly known boundary case, where the forward map F maps the domain Ω onto itself. However, the forward map is not the identity or a rotation operator, and thus it will induce an anisotropic conductivity, especially in the regions near the boundary.

Example 9. *The true domain $\widetilde{\Omega}$ is identical with the computational domain Ω, the unit circle, but every other electrode is shifted by an angle of $\pi/32$, while the length of each electrode remains unchanged; see Fig. 3.1(b) for a schematic illustration. There are two rectangular inclusions, on the top and on the bottom of the ellipse, with spectral profiles $s_1(\omega) = 0.2\omega + 0.2$ and $s_2(\omega) = 0.1\omega^2$, respectively. We take the measurement at three frequencies $\omega_1 = 0$, $\omega_2 = 0.5$ and $\omega_3 = 1$.*



(a) true $\delta\sigma_i$s (b) recovered $\delta\sigma_1$ (c) recovered $\delta\sigma_2$

(d) recovered $\delta\sigma_0$ (e) recovered $\delta\sigma_1$ (f) recovered $\delta\sigma_2$

FIGURE 3.9: Numerical results for Example 9 with 0.1% noise in the data, fully known spectral profiles. The reconstructions shown in panels (b)-(c) are based on difference imaging in Section 3.2.2, whereas those in panels (d)–(f) are based on the direct approach of the method discussed in Section 3.2.1.

The numerical results for Example 9 are summarized in Fig. 3.9. The analysis in Section 3.3 and in §3.4.2 indicates that the conductivity perturbation due to the domain deformation can be limited to the background component $\delta\sigma_0$, as above. The numerical results confirm the analysis: when using the direct approach discussed in §3.2.1, there are many pronounced perturbations around the boundary in the reconstructed $\delta\sigma_0$, due to the domain deformation. However, the reconstructed $\delta\sigma_1$ and $\delta\sigma_2$ are fairly reasonable in location and size, albeit slightly deformed. The differentiation imaging can also remove the contributions due to unknown electrode locations and separate the contributions from the two inclusions. This is feasible since the frequency profiles are incoherent both before and after the differentiation.

In summary, as expected from the analysis of Sections 3.3 and 3.4.2, the mfEIT technique has significant potential in handling modelling errors. The

inclusion $\delta\sigma_0$ corresponding to the background frequency profile s_0 may not be reconstructed. However, by using the multifrequency method, the remaining inclusions $\{\delta\sigma_k\}_{k=1}^K$ can be correctly recovered by either the direct approach in Section 3.2.1 or the difference imaging approach in Section 3.2.2, provided that the corresponding spectral profiles (or their derivatives) are sufficiently incoherent.

Chapter 4

A method to image anisotropy conductivity

4.1 Introduction

In this chapter, we describe a novel method of reconstructing images of an anisotropic conductivity tensor distribution inside an electrically conducting subject in Magneto-acoustic Tomography with Magnetic Induction (MAT-MI).

MAT-MI is a new noninvasive modality for imaging electrical conductivity distributions of biological tissue [118, 72, 77, 122, 71, 117, 74]. In the experiments, the biological tissue is placed in a static magnetic field. A pulsed magnetic field is applied to induce an eddy current inside the conductive tissue. In the process, the tissue emits ultrasound waves which can be measured around the tissue. The first step in the MAT-MI imaging problem is to recover the acoustic source in the scalar wave equation from measurements at a set of locations around the object. This problem has been studied in many works, including [28, 41, 50, 65, 103]. The second step in the MAT-MI is to reconstruct the electrical conductivity distribution from knowledge of the acoustic source.

Considering the fact that most biological tissues are known to have anisotropic conductivity values [10, 78, 110], the primary goal of MAT-MI should be the imaging of an anisotropic conductivity tensor distribution. However, up to now, all techniques have assumed an isotropic conductivity distribution in the image reconstruction problem to simplify the underlying mathematical theory [5, 90]. In this chapter, we firstly formulate a new image reconstruction method of an anisotropic conductivity tensor distribution. We combine MAT-MI with Diffusion Tensor Imaging (DTI).

DTI is a non-invasive technique for characterizing the diffusion properties of water molecules in tissues (see e.g. [14] and the references therein). Imaging conductivity tensors in the tissue with DTI is based on the correlation property between diffusion and conductivity tensors [110]. This linear relationship can be used to characterize the conductivity tensor. Once the conductivity directions of anisotropy are determined, one needs only to reconstruct a cross-property factor which is a scalar function. In [46, 66], it is shown how to recover this factor in Current Density Impedance Imaging. In [9], a multifrequency electrical impedance approach is developed for estimating the ratio between the largest and the lowest eigenvalue of the electrical conductivity tensor. An iterative procedure for reconstructing anisotropic conductivities from internal current densities has been proposed in [98].

In the process of the MAT-MI experiment, the tissue is placed in a constant static magnetic background field $B_0 = (0, 0, 1)$. A pulsed magnetic stimulation of the form $B_1 u(t)$ is applied, where the vector field B_1 is constant and $u(t)$ is the time variation. Let γ denote the conductivity, E_γ denote electric field, and Ω be the domain to be imaged. Then the electric field satisfies the following Maxwell equations

$$\begin{cases} \nabla \times E_\gamma &= B_1 \quad \text{in } \Omega, \\ \nabla \cdot (\gamma E_\gamma) &= 0 \quad \text{on } \Omega, \\ \gamma E_\gamma \cdot \nu &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{4.1.1}$$

The second step of MAT-MI is to reconstruct γ from the known internal data $\nabla \cdot (\gamma E_\gamma \times B_0)$ on Ω.

In this chapter, we will consider the anisotropic conductivity case with γ being a tensor. With DTI, one can measure the water self-diffusion tensor, which will characterize the electrical conductivity tensor [110]. Then we can assume that the tensor γ is of the form

$$\gamma(x) = \sigma(x) D(x), \tag{4.1.2}$$

with the tensor $D(x)$ being measured from DTI and the cross-property factor σ being a scalar function to be reconstructed. We will focus on the second step of MAT-MI combined with DTI, i.e., on reconstructing the cross-property factor σ from the internal data given by $\nabla \cdot (\gamma E_\gamma \times B_0)$ with known conductivity tensor $D(x)$.

In the following, we assume that $D(x)$ is a positive definite symmetric matrix everywhere and write it as $D = T^{'} \Sigma T$, where $D = \mathrm{diag}(e_1, e_2, e_3)$, $e_1 \geq e_2 \geq e_3$ are the eigenvalues of $D(x)$. The columns of $T^{'}$ are the corresponding eigenvectors. As we can always write $\sigma = \sigma_0 e_1 T^{'} \mathrm{diag}(1, e_2/e_1, e_3/e_1) T$, we assume that $e_1 = 1$ hereinafter.

4.2 Notation and preliminaries

In this section, we introduce the notation for the mathematical analysis. Let Ω be a bounded Lipschitz domain in $\mathbb{R}^3$. A typical point $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ denotes the spatial variable. Throughout this chapter, the standard notation for Hölder and Sobolev spaces and their norms is used. If there is no confusion, we omit the dependence on the domain.

Assumption 1. *Let σ and D be positive functions belonging to $C^{1,\beta}$, $\beta > 0$ and assume that*

$$c_1 \leq \sigma(x) \leq c_2, \quad \forall x \in \Omega, \tag{4.2.1}$$

and

$$\lambda \|\xi\|_2^2 \leq \xi^{'} D \xi \leq \|\xi\|_2^2, \quad \forall \xi \in R^3, \tag{4.2.2}$$

for some constants $\lambda, c_1, c_2 > 0$.

Here we state several useful results on elliptic partial differential equations with Neumann boundary conditions.

We say that $u \in H^1$ is a weak solution of the Neumann boundary value problem

$$\begin{cases} \nabla \cdot (\sigma D \nabla u) &= -\nabla \cdot E, \quad \text{in } \Omega, \\ (\sigma D \nabla u + E) \cdot \nu &= 0, \quad \text{on } \partial\Omega, \end{cases} \tag{4.2.3}$$

if

$$\int_\Omega \sigma D\nabla u \cdot \nabla \varphi \,\mathrm{d}x = -\int_\Omega E \cdot \nabla \varphi \,\mathrm{d}x, \quad \forall \varphi \in H^1. \tag{4.2.4}$$

We give a brief proof of the following regularity result and standard energy estimate.

Proposition 3. *Suppose that σ and D satisfy Assumption 1. For field $E \in L^2$, the Neumann problem* (4.2.3) *has a solution $u \in H^1$. The solution u is unique up to an additive constant and satisfies the estimate*

$$\|\nabla u\|_{L^2} \leq c_1^{-1}\lambda^{-1}\|E\|_{L^2}. \tag{4.2.5}$$

Proof. The proof of the existence and uniqueness up to an additive constant is a standard result by the Lax-Milgram Theorem. We refer the readers to [109]. In the following, we prove the gradient estimate (4.2.5).

It follows from the ellipticity condition (4.2.2) that

$$c_1\lambda\|\nabla u\|_{L^2}^2 \leq \int_\Omega \sigma\nabla u \cdot D\nabla u \,\mathrm{d}x.$$

Taking the test function φ in Definition 4.2.4 to be the solution u, we have that

$$\int_\Omega \sigma D\nabla u \cdot \nabla u \,\mathrm{d}x = -\int_\Omega E \cdot \nabla u \,\mathrm{d}x.$$

Consequently, applying the Cauchy-Schwarz inequality, we obtain that

$$c_1\lambda\|\nabla u\|_{L^2}^2 \leq \left|-\int_\Omega E \cdot \nabla u \,\mathrm{d}x\right| \leq \|\nabla u\|_{L^2}\|E\|_{L^2},$$

and (4.2.5) follows. □

Proposition 4. *Let σ and D satisfy Assumption 1. Then the system* (4.1.1) *is uniquely solvable and there exists a constant C and C_i $(1 \leq i \leq 3)$ depending on λ, c_1, c_2 and Ω, such that*

$$\|E_{\sigma D}\|_{L^2} \leq C_1, \tag{4.2.6}$$

$$\|E_{\sigma D}\|_{L^\infty(\Omega)} \leq C_2, \tag{4.2.7}$$

$$\|E_{\sigma D}\|_{C^{1,\gamma}(\Omega)} \leq C_3. \tag{4.2.8}$$

Moreover, if σ_1 and σ_2 satisfy Assumption 1, we have the following bound for the electric field difference,

$$\|E_{\sigma_1 D} - E_{\sigma_2 D}\|_{L^2(\Omega)} \leq C\|\sigma_1 - \sigma_2\|_{L^2(\Omega)}. \tag{4.2.9}$$

$$\|E_{\sigma_1 D} - E_{\sigma_2 D}\|_{H^1(\Omega)} \leq C\|\sigma_1 - \sigma_2\|_{H^1(\Omega)}. \tag{4.2.10}$$

Proof. Let us first reduce the system (4.1.1) to a Neumann boundary value problem. Let $\tilde{E} = \frac{1}{2}(-y, x, 0)$. We can readily check that $\nabla \times \tilde{E} = B_1$. Hence $\nabla \times (E_{\sigma D} - \tilde{E}) = 0$ and we can write $E_{\sigma D} = \tilde{E} + \nabla u$. Substituting this into

(4.1.1), we have that u solves the Neumann boundary value problem

$$\begin{cases} \nabla \cdot (\sigma D \nabla u) & = -\nabla \cdot (\sigma D \tilde{E}), & \text{in } \Omega, \\ (\sigma D \nabla u + \sigma D \tilde{E}) \cdot \nu & = 0, & \text{on } \partial\Omega. \end{cases} \tag{4.2.11}$$

With the help of proposition 3, we get

$$\|E_{\sigma D}\|_{L^2} \leq C_1.$$

From the standard L^p estimate for elliptic equations [35, Chapter 9] and the Sobolev Embedding Theorem, we know that $E_{\sigma D}$ is a bounded function in $W^{2,p}(\Omega)$ for any $p > 2$. Hence, $E_{\sigma D}$ is uniformly bounded by a constant C, which depends only on r_0, λ, c_1, c_2, and Ω. Then (4.2.7) is proved.

With the assumption of $C^{1,\gamma}$ property, we would obtain the $C^{2,\gamma}$ Hölder continuity[39] of u, i.e., the $C^{1,\gamma}$ continuity of $E_{\sigma D}$. Estimate (4.2.8) has been proven.

Next, we estimate the electric field difference. We denote $E_i = E_{\sigma_i D}$, for $i = 1, 2$. Note that $E_1 - E_2$ is curl-free. We set

$$\nabla u = E_1 - E_2.$$

Then, u satisfies the equation

$$\begin{cases} \nabla \cdot (\sigma_1 D \nabla u) & = -\nabla \cdot ((\sigma_1 - \sigma_2) D E_2), & \text{in } \Omega, \\ D \nabla u \cdot \nu & = 0, & \text{on } \partial\Omega. \end{cases} \tag{4.2.12}$$

With the same argument for proving (4.2.6), we obtain that

$$\|\nabla u\|_{L^2} \leq c_1^{-1} \lambda^{-1} \|(\sigma_1 - \sigma_2) D E_2\|_{L^2}.$$

Thus, we conclude from (4.2.7) that

$$\|E_1 - E_2\|_{L^2} \leq C \|(\sigma_1 - \sigma_2)\|_{L^2}.$$

From the standard theory of elliptic equations

$$\|u\|_{H^2(\Omega)} \leq C \|\nabla \cdot ((\sigma_1 - \sigma_2) D E_2)\|_{L^2(\Omega)}, \tag{4.2.13}$$

which implies (4.2.10). □

4.3 Uniqueness and stability

With the notation of the previous section, we will show the well-posedness of the inverse problem in a certain functional space.

We prove the following theorem on the stability of the inverse problem.

Theorem 4.3.1. *Let $F(\sigma) = \nabla \cdot (\sigma D E_{\sigma D} \times B_0)$. Suppose Assumption 1 is satisfied and $\sigma_1 - \sigma_2 \in W_0^{1,\infty}$. If there exist constants K, L and η such that*

$$\|\nabla \sigma_i\|_{L^\infty} \leq K \tag{4.3.1}$$

$$|1 - \lambda| \leq \eta \tag{4.3.2}$$

$$\|\nabla(\sigma_1 - \sigma_2)\|_{L^2(\Omega)} \leq L\|\sigma_1 - \sigma_2\|_{L^2(\Omega)}, \tag{4.3.3}$$

then

$$c\|\sigma_1 - \sigma_2\|_{L^2(\Omega)} \leq \|F(\sigma_1) - F(\sigma_2)\|_{L^2(\Omega)} \tag{4.3.4}$$

holds for some positive constant c.

Proof. We denote $E_i = E_{\sigma_i D}$, $i = 1, 2$ and write the data difference as

$$\begin{aligned} F(\sigma_1) - F(\sigma_2) &= \nabla \cdot (\sigma_1 D E_1 \times B_0) - \nabla \cdot (\sigma_2 D E_2 \times B_0) \\ &= \nabla \cdot ((\sigma_1 - \sigma_2) D E_1 \times B_0) + \nabla \cdot (\sigma_2 D (E_1 - E_2) \times B_0). \end{aligned}$$

Then, we can rewrite $F(\sigma_1) - F(\sigma_2)$ as

$$F(\sigma_1) - F(\sigma_2) = I_1 + I_2 + I_3 + I_4,$$

where

$$\begin{aligned} I_1 &= \nabla \cdot ((\sigma_1 - \sigma_2) E_1 \times B_0), \\ I_2 &= \nabla \cdot (\sigma_2 (E_1 - E_2) \times B_0), \\ I_3 &= \nabla \cdot ((\sigma_1 - \sigma_2)(D - I) E_1 \times B_0), \\ I_4 &= \nabla \cdot (\sigma_2 (D - I)(E_1 - E_2) \times B_0), \end{aligned}$$

where I is the identity matrix.

Next we multiply $F(\sigma_1) - F(\sigma_2)$ by $\sigma_1 - \sigma_2$ and integrate over Ω. For I_i, $i = 1, 2, 3, 4$, we can estimate the integrals $\int_\Omega (\sigma_1 - \sigma_2) I_i$ separately. We have

$$\begin{aligned} \int_\Omega (\sigma_1 - \sigma_2) I_1 dx &= \int_\Omega \Big((\sigma_1 - \sigma_2)(\sigma_1 - \sigma_2) \nabla \cdot (E_1 \times B_0) \\ &\quad + (\sigma_1 - \sigma_2) \nabla (\sigma_1 - \sigma_2) \cdot (E_1 \times B_0) \Big) dx \\ &= \frac{1}{2} \int_\Omega (\sigma_1 - \sigma_2)(\sigma_1 - \sigma_2) \nabla \cdot (E_1 \times B_0) dx \\ &= \frac{1}{2} \|\sigma_1 - \sigma_2\|^2_{L^2(\Omega)}. \end{aligned}$$

Here we use the equality $\nabla \cdot (E_1 \times B_0) = 1$ which can be easily checked from the identity $\nabla \cdot (E_1 \times B_0) = B_0 \cdot (\nabla \times E_1) - E_1 \cdot (\nabla \times B_0) = 1$. On the other hand,

$$\begin{aligned} \left| \int_\Omega (\sigma_1 - \sigma_2) I_2 dx \right| &= \left| \int_\Omega (\sigma_1 - \sigma_2) \nabla \sigma_2 \cdot ((E_1 - E_2) \times B_0)) dx \right| \\ &\leq KC \|\sigma_1 - \sigma_2\|^2_{L^2(\Omega)}. \end{aligned}$$

Here the assumption (4.3.1) and inequality (4.2.9) have been used.

Now we turn to the terms I_3 and I_4. We have

$$\begin{aligned}\left|\int_\Omega(\sigma_1-\sigma_2)I_3dx\right| &= \Big|\int_\Omega(\sigma_1-\sigma_2)(\sigma_1-\sigma_2)\nabla\cdot((D-I)E_1\times B_0)\\ &\quad +(\sigma_1-\sigma_2)\nabla(\sigma_1-\sigma_2)\cdot((D-I)E_1\times B_0))dx\Big|\\ &= \Big|-\int_\Omega(\sigma_1-\sigma_2)\nabla(\sigma_1-\sigma_2)\cdot((D-I)E_1\times B_0)dx\Big|\\ &\leq \eta LC\|\sigma_1-\sigma_2\|^2_{L^2(\Omega)}.\end{aligned}$$

In the last inequality we have used estimate (4.2.7) together with the assumptions (4.3.2) and (4.3.3). Finally, we have

$$\begin{aligned}\left|\int_\Omega(\sigma_1-\sigma_2)I_4dx\right| &= \Big|\int_\Omega(\sigma_1-\sigma_2)\sigma_2\nabla\cdot((D-I)(E_1-E_2)\times B_0)\\ &\quad +(\sigma_1-\sigma_2)\nabla\sigma_2\cdot((D-I)(E_1-E_2)\times B_0))dx\Big|\\ &= \Big|-\int_\Omega\sigma_2\nabla(\sigma_1-\sigma_2)\cdot((D-I)(E_1-E_2)\times B_0)dx\Big|\\ &\leq \eta LC\|\sigma_1-\sigma_2\|^2_{L^2(\Omega)}.\end{aligned}$$

Here we have used the assumptions (4.3.2), (4.3.3) and inequality (4.2.9).

Let K and η be such that $KC+2\eta LC<\frac{1}{2}$. We obtain

$$\int_\Omega(\sigma_1-\sigma_2)(F(\sigma_1)-F(\sigma_2))\geq c\|\sigma_1-\sigma_2\|^2_{L^2(\Omega)},$$

for some constant c, which proves the theorem. □

Now we are ready to introduce a functional framework for which the inverse problem is well defined. We assume that σ is known on the boundary of Ω. In what follows, we let σ_*, the true cross-property factor of Ω, belong to a bounded convex subset of $C^{1,\beta}(\Omega)$ given by

$$\widetilde{\mathcal{S}}=\{\sigma:=\sigma_0+\alpha|\,\alpha\in\mathcal{S}\},$$

where σ_0 is some positive function satisfying Assumption 1 and

$$\begin{aligned}\mathcal{S}= \quad \{\alpha\in C_0^{1,\beta}(\Omega)|\; & c_1\leq\alpha+\sigma_0\leq c_2,\ |\nabla(\alpha+\sigma_0)|_{L^\infty}\leq K,\\ & \|\nabla\alpha\|_{L^2(\Omega)}\leq L\|\alpha\|_{L^2(\Omega)},\ \|\alpha\|_{L^2(\Omega)}\leq c_3\}\end{aligned}\tag{4.3.5}$$

with c_1, c_2, c_3 and c_3 being positive constants. In other words, we can write $\widetilde{\mathcal{S}}=\sigma_0+\mathcal{S}$.

It is clear that the distribution of the electric field $E_{\sigma D}$ depends nonlinearly on the factor σ and $\nabla\cdot(\sigma DE_{\sigma D}\times B_0)$ is nonlinear with respect to σ. We first examine the Fréchet differentiability of the forward operator F. Then, some useful properties of the Fréchet derivative at σ, $DF[\sigma]$, are presented.

To introduce the Fréchet derivative, we consider the following Neumann boundary value problem

$$\begin{cases} \nabla \cdot (\sigma D \nabla \varphi_h) & = -\nabla \cdot (hDE_{\sigma D}), & \text{in } \Omega, \\ (\sigma D \nabla \varphi_h + hDE_{\sigma D}) \cdot \nu & = 0, & \text{on } \partial\Omega, \end{cases} \tag{4.3.6}$$

and

$$\begin{cases} \nabla \cdot (\sigma D \nabla \Phi_g) & = -\nabla \cdot (\sigma D(B_0 \times \nabla g)), & \text{in } \Omega, \\ (\sigma D \nabla \Phi_g + \sigma D(B_0 \times \nabla g)) \cdot \nu & = 0, & \text{on } \partial\Omega, \end{cases} \tag{4.3.7}$$

where $h \in \mathcal{S}$ is the increment to the factor σ.

By the same arguments as those in [90], together with Theorem 4.3.1, it is natural to conclude the following result that insures the well-posedness of the inverse problem.

Theorem 4.3.2. *For σ and D satisfying Assumption 1 and (4.3.2), the operator F is bounded and Fréchet differentiable at $\sigma \in \widetilde{\mathcal{S}}$. Its Fréchet derivative at σ, $DF[\sigma]$, is given by*

$$DF[\sigma](h) = \nabla \cdot ((\sigma D \nabla \varphi_h + hDE_{\sigma D}) \times B_0), \tag{4.3.8}$$

where φ_h solves (4.3.6)*. Meanwhile, $DF[\sigma]^*$, i.e., the adjoint of $DF[\sigma]$ is defined as below,*

$$DF[\sigma]^*(g) = -DE_{\sigma D} \cdot \nabla \Phi_g - \nabla g \cdot (DE_{\sigma D} \times B_0), \tag{4.3.9}$$

where Φ_g solves (4.3.7). Furthermore, we have the following stability result,

$$c\|h\|_{L^2(\Omega)} \leq \|DF[\sigma](h)\|_{L^2} \leq C\|h\|_{L^2(\Omega)}, \quad \forall h \in \mathcal{S}, \tag{4.3.10}$$

for some constant C which depends on λ, c_1, c_2 and Ω and the constant c is defined in Theorem 4.3.1.

Proof. The definition of $DF[\sigma](h)$ and (4.3.10) follow from [90] and Theorem 4.3.1. Here we only give a brief proof of the formulation of $DF[\sigma]^*$.

First by multiplying (4.3.6) by Φ_g and (4.3.7) by φ_h, we get after an integration by parts,

$$\int_\Omega \sigma D(B_0 \times \nabla g) \cdot \nabla \varphi_h dx = -\int_\Omega \sigma D \nabla \varphi_h \cdot \nabla \Phi_g dx = \int_\Omega hDE_{\sigma D} \cdot \nabla \Phi_g dx. \tag{4.3.11}$$

Then we are ready to compute $DF[\sigma]^*(g)$. We have

$$\begin{aligned} \int_\Omega DF[\sigma](h) g dx &= \int_\Omega \nabla \cdot ((\sigma D \nabla \varphi_h + hDE_{\sigma D}) \times B_0) g dx \\ &= -\int_\Omega ((\sigma D \nabla \varphi_h + hDE_{\sigma D}) \times B_0) \cdot \nabla g dx \\ &= -\int_\Omega -\sigma D \nabla \varphi_h \cdot \nabla \Phi_g + h(DE_{\sigma D} \times B_0) \cdot \nabla g dx \\ &= -\int_\Omega hDE_{\sigma D} \cdot \nabla \Phi_g + h(DE_{\sigma D} \times B_0) \cdot \nabla g dx. \end{aligned}$$

This proves (4.3.9). □

4.4 The reconstruction method

4.4.1 Optimization scheme

It is natural to formulate the reconstruction problem for σ_* as a least-square problem. To find σ_* we minimize the functional

$$J(\sigma) = \frac{1}{2}\|F(\sigma) - F(\sigma_*)\|^2_{L^2(\Omega)}$$

over $\sigma \in \widetilde{\mathcal{S}}$.

We can now apply the gradient descent method to minimize the discrepancy functional J. Define the iterates

$$\sigma_{n+1} = T[\sigma_n] - \mu DJ[T[\sigma_n]], \tag{4.4.1}$$

where $\mu > 0$ is the step size and $T[f]$ is any approximation of the Hilbert projection from $L^2(\Omega)$ onto $\overline{\widetilde{\mathcal{S}}}$ with $\overline{\widetilde{\mathcal{S}}}$ being the closure of $\widetilde{\mathcal{S}}$ (in the L^2-norm).

The presence of the projection T is necessary because σ_n might not be in $\widetilde{\mathcal{S}}$.

Using the definition of J we can show that the optimal control algorithm (4.4.1) is nothing else than the following projected Landweber iteration [68, 42, 49] given by

$$\sigma_{n+1} = T[\sigma_n] - \mu DF^*[T[\sigma_n]](F(T[\sigma_n]) - F(\sigma_*)). \tag{4.4.2}$$

For completeness, we state the convergence result of Landweber scheme here without proof. We refer to [7, 49] for details.

Theorem 4.4.1. *The sequence defined in* (4.4.2) *converges to the true cross-property factor σ_* of Ω in the following sense: there exists $\epsilon > 0$ such that if $\|T[\sigma_1] - \sigma_*\|_{L^2(\Omega)} < \epsilon$, then*

$$\lim_{n\to+\infty} \|\sigma_n - \sigma_*\|_{L^2(\Omega)} = 0.$$

4.4.2 A quasi-Newton method

It has been observed in [90] that the challenge of the Landweber iteration lies in the difficulty of evaluating the adjoint operator of the Fréchet derivative. To avoid taking too many derivatives, we introduce a more efficient way to reconstruct the conductivity. This is a generalization of the quasi-Newton method proposed in [90] for the anisotropic case with known conformal class. In the following, we describe this algorithm and prove its convergence in $\widetilde{\mathcal{S}}$.

Let σ be the scalar conductivity distribution function and let D be the known conformal class matrix-valued function. The forward operator is given by

$$F(\sigma) = \nabla \cdot (\sigma D E_{\sigma D} \times B_0),$$

where E_σ satisfies the system

$$\begin{cases} \nabla \cdot (\sigma D E_{\sigma D}) & = 0, & \text{in } \Omega, \\ \nabla \times E_{\sigma D} & = B_1, & \text{in } \Omega, \\ \sigma D E_{\sigma D} \cdot \nu & = 0, & \text{on } \partial\Omega. \end{cases}$$

Algorithm 4.4.1.

Step 0. *Select an initial conductivity $\sigma_1 \in \widetilde{\mathcal{S}}$ and set $k = 1$.*

Step 1. *Calculate the associated electric field E_k by solving the boundary value problem*

$$\begin{cases} \nabla \times E_k & = B_1, & \text{in } \Omega, \\ \nabla \cdot (\sigma_k D E_k) & = 0, & \text{in } \Omega, \\ \sigma_k D E_k \cdot \nu & = 0, & \text{on } \partial\Omega. \end{cases} \tag{4.4.3}$$

Step 2. *Calculate the updated conductivity by solving the stationary advection-diffusion equation with the inflow boundary condition:*

$$\begin{cases} \nabla \cdot (\sigma_{k+1/2} D E_k \times B_0) & = g, & \text{in } \Omega, \\ \sigma_{k+1/2} & = \sigma_*, & \text{on } \partial\Omega^-, \end{cases} \tag{4.4.4}$$

where $\partial\Omega^- = \{x \in \partial\Omega \mid DE_k(x) \times B_0 \cdot \nu(x) < 0\}$.

Step 3. *Let $\sigma_{k+1} = T[\sigma_{k+1/2}]$, where T is the Hilbert projection operator onto $\tilde{S}$. Set $k = k + 1$ and go to (4.4.3).*

4.4.3 Convergence analysis

In the algorithm above, one updates the electric field E and then updates the cross-factor σ later. Using the same argument as for proving the well-posedness, we could get the following convergence results.

Theorem 4.4.2. *Suppose that the cross-factor $\sigma_* \in \widetilde{\mathcal{S}}$ and D satisfies Assumption 1 and* (4.3.2)*. Let $\{\sigma_k\}$ be determined by the Algorithm 4.4.1. Then for proper constants K, L and η in* (4.3.5) *and* (4.3.2)*, there exists a constant $c < 1$ such that*

$$\|\sigma_{k+1} - \sigma_*\|_{L^2(\Omega)} \leq c\|\sigma_k - \sigma_*\|_{L^2(\Omega)}. \tag{4.4.5}$$

Proof. Note that T is a projection and $\widetilde{\mathcal{S}}$ is convex. Then, we have that T is nonexpansive and $\|\sigma_{k+1} - \sigma_*\| \leq \|\sigma_{k+1/2} - \sigma_*\|$ follows. It is left to estimate $\|\sigma_{k+1/2} - \sigma_*\|$.

First we subtract $\nabla \cdot (\sigma_* D E_k \times B_0)$ from both sides of (4.4.4)to get

$$\nabla \cdot ((\sigma_{k+1/2} - \sigma_*) D E_k \times B_0) = \nabla \cdot (\sigma_* D(E_* - E_k) \times B_0). \tag{4.4.6}$$

Multiplying by $\sigma_{k+1/2} - \sigma_*$ and integrating over Ω yields

$$\int_\Omega (\sigma_{k+1/2} - \sigma_*) \nabla \cdot ((\sigma_{k+1/2} - \sigma_*) D E_k \times B_0) = \int_\Omega (\sigma_{k+1/2} - \sigma_*) \nabla \cdot (\sigma_* D(E_* - E_k) \times B_0). \tag{4.4.7}$$

We split the terms into

$$\begin{aligned} & \int_\Omega (\sigma_{k+1/2} - \sigma_*) \nabla \cdot ((\sigma_{k+1/2} - \sigma_*) E_k \times B_0) \\ & + (\sigma_{k+1/2} - \sigma_*) \nabla \cdot ((\sigma_{k+1/2} - \sigma_*)(D - I) E_k \times B_0) \\ = & \int_\Omega (\sigma_{k+1/2} - \sigma_*) \nabla \sigma_* \cdot (D(E_* - E_k) \times B_0) + (\sigma_{k+1/2} - \sigma_*) \sigma_* \nabla \cdot ((D - I)(E_* - E_k) \times B_0). \end{aligned}$$

Integrating by parts gives $\int_\Omega (\sigma_{k+1/2} - \sigma_*) \nabla \cdot ((\sigma_{k+1/2} - \sigma_*) E_k \times B_0) = \frac{1}{2} \|\sigma_{k+1/2} - \sigma_*\|^2_{L^2(\Omega)}$.

The second term in the left hand side can be estimated as follows:

$$\begin{aligned} & \left| \int_\Omega (\sigma_{k+1/2} - \sigma_*) \nabla \cdot ((\sigma_{k+1/2} - \sigma_*)(D - I)E_k \times B_0) \right| \\ = & \left| \int_\Omega (\sigma_{k+1/2} - \sigma_*)^2 \nabla \cdot ((D - I)E_k \times B_0) \right| \\ \leq & C\eta \|\sigma_{k+1/2} - \sigma_*\|^2_{L^2(\Omega)}. \end{aligned}$$

Here the smallness of $D - I$ and the C^1 property of E_k (4.2.8) have been used.

For the right hand side, we have

$$\left| \int_\Omega (\sigma_{k+1/2} - \sigma_*) \nabla \sigma_* \cdot (D(E_* - E_k) \times B_0) \right| \leq CK \|\sigma_{k+1/2} - \sigma_*\|_{L^2(\Omega)} \|\sigma_k - \sigma_*\|_{L^2(\Omega)},$$

and

$$\begin{aligned} & \left| \int_\Omega (\sigma_{k+1/2} - \sigma_*) \sigma_* \nabla \cdot ((D - I)(E_* - E_k) \times B_0) \right| \\ \leq & C\eta \|\sigma_{k+1/2} - \sigma_*\|_{L^2(\Omega)} \|\sigma_k - \sigma_*\|_{H^1(\Omega)} \\ \leq & C(L + 1)\eta \|\sigma_{k+1/2} - \sigma_*\|_{L^2(\Omega)} \|\sigma_k - \sigma_*\|_{L^2(\Omega)}. \end{aligned}$$

Here we have used property (4.2.10) and the fact that $\sigma_k \in \tilde{S}$.

With the above estimates, as we did in Theorem 4.3.1, let $KC + \eta(L + 2)C < \frac{1}{2}$. We derive

$$\|\sigma_{k+1/2} - \sigma_*\|_{L^2(\Omega)} \leq c \|\sigma_k - \sigma_*\|_{L^2(\Omega)}, \tag{4.4.8}$$

where c is a constant smaller than 1. Hence,

$$\|\sigma_{k+1} - \sigma_*\|_{L^2(\Omega)} \leq c \|\sigma_k - \sigma_*\|_{L^2(\Omega)}, \tag{4.4.9}$$

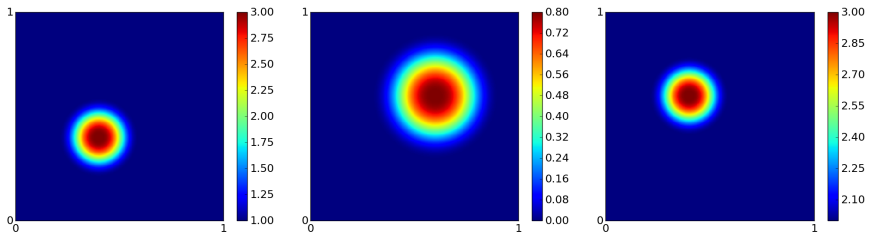
which proves the theorem. □

4.4.4 Numerical experiments

In this section, we present some numerical experiments to validate the reconstruction method proposed in Algorithm 4.4.1 and evaluate its robustness to measurement noise. To simplify the computation, we convert this three-dimensional problem into an equivalent two-dimensional problem assuming that the domain of interest is the cube $[0, 1]^3$ and the conductivity and the diffusion tensors are invariant along the third dimension. Moreover, we assume that the diffusion tensor D is of form

$$D = \begin{pmatrix} d_{11} & d_{12} & 0 \\ d_{21} & d_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.4.10}$$

where d_{ij}'s are constant plus some perturbations as shown in Figure 4.1. The non-zero part of the perturbation functions are used to characterize the anisotropy.



(a) d_{11} (b) d_{12} (c) d_{22}

FIGURE 4.1: Components of the diffusion tensor.

We use a uniform finite element triangular mesh over the two-dimensional unit square. The number of cells is 256 in each direction. The total number of triangles and vertices are 2^{17} and 257^2, respectively. Both the elliptic equation with a Neumann boundary condition and the stationary advection-diffusion equation are solved using the finite element method of first order implemented with FEniCS [73]. The internal data $F(\sigma)$ used for the reconstruction are synthetic data that are generated using the same solver. These data are commonly used to refer to the "noise-free" data, although they may contain some numerical errors.

For all examples, we use the same initial guess, constant function 0.2, and the same true cross-property factor (Figure 4.2(a)) given by

$$\sigma(x_1, x_2) = \begin{cases} 0.6, & r \leq 0.12, \\ 0.4s^3(6s^2 - 15s + 10) + 0.2, & 0.12 < r < 0.46, \\ 0.2, & \text{others}, \end{cases}$$

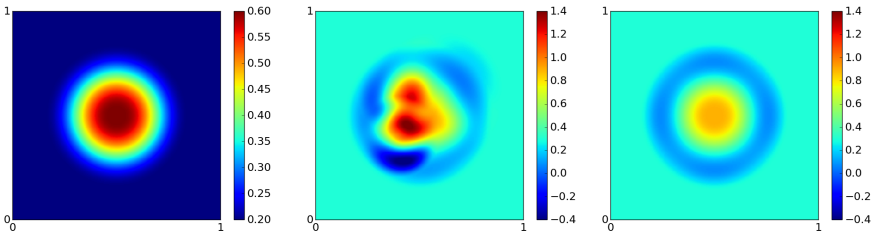
where $r(x_1, x_2) = \sqrt{(x_1 - 0.5)^2 + (x_2 - 0.5)^2}$ and $s = \frac{0.46 - r}{0.12}$. The internal data generated with the diffusion tensor as in (4.4.10) is shown in Figure 4.2(b). We also produce the data in the isotropic case (Figure 4.2(c)). The effect of the anisotropy can be observed clearly. The error-decay of the reconstruction with the noise-free data is shown in Figure 4.4(a). The final error is smaller than 2×10^{-3}. We only display the last iterate here.

This inverse problem bears a Lipschitz type stability and we avoid lowering the regularity of the cross-property factor using Algorithm 4.4.1. Therefore, the robustness of the reconstruction scheme to noisy data is expected. We perform the numerical tests with noisy data by perturbing the internal functional g in the following way:

$$g_\delta = g + \delta \|g\| \frac{w}{\|w\|},$$

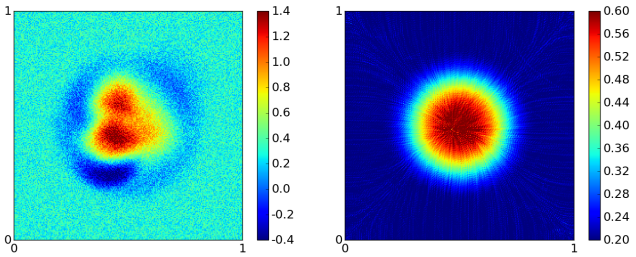
where w is a function taking values uniformly distributed in $[-1, 1]$ and δ is the noise level.

Figure 4.3 shows the noisy data with noise level $\delta = 24\%$ and the reconstructed cross-property factor. We do not use further regularization techniques since the regularization method may depend on the type of the noise in practical cases. But the projection onto the feasible space acts as a regularization scheme.



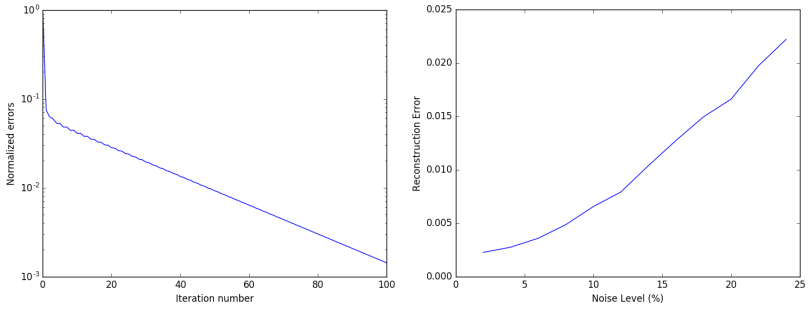
(a) True conductivity (b) Internal data (c) Internal data with isotropic diffusion tensor

FIGURE 4.2: Conductivity distribution and the internal data.



(a) Noisy data (b) Reconstructed σ

FIGURE 4.3: Reconstruction with noisy data ($\delta = 24\%$).



(a) Error decay with noise-free data. (b) Reconstruction error versus noise level.

FIGURE 4.4: Reconstruction error.

4.5 Concluding remarks

In this chapter, we have considered the reconstruction of an anisotropic conductivity from MAT-MI data which is conformal to a known diffusion tensor measured from DTI. The data is the internal functional $\nabla \cdot (\sigma E_\sigma \times B_0)$ throughout the domain. We have analyzed the linearization of the problem and the stability of the inversion. A local Lipschitz type stability estimate has been established for a certain class of anisotropic conductivities. A quasi-Newton type reconstruction method with projection has been introduced and its convergence has been proved. Numerical experiments demonstrate the effectiveness of the proposed approach and its robustness to noise.

In light of the numerical experiments, we have the following observations.

1. The effect of the electrical anisotropy is remarkably significant and can not be neglected in the reconstruction of electrical conductivity in MAT-MI.

2. There is still a room for improvement of the admissible class of conductivities. The convergence of the proposed algorithm has been observed for more general cases.

3. For the inversion with noisy data, oscillation in the reconstructed conductivity is observed. Regularization methods prompting sparsity, such as total variation regularization may be employed for a more stable reconstruction.

Chapter 5

Analysis of cell model for electropermeabilization

5.1 Introduction

In this chapter, we give the analysis of the cell model for electropermeabilization and use a homogenization method to show the effects of cell parameter to the effective conductivity.

The technique of electropermeabilization (formerly referred to as electroporation) is employed to make the chemotherapeutical treatment of cancer more efficient and avoid side-effects. Instead of spreading out drugs over the whole body, electropermeabilization makes it possible to focus drug application on special areas. The mechanism of electropermeabilization relies on careful exposition of biological tissue to electrical fields: this changes the membrane properties of the cells such that treatment can enter more easily just at precisely defined areas of the tissue [54, 81].

The local change in microscopic tissue properties, which electropermeabilization effects, occurs only with field strengths above a certain threshold. On the other hand, too strong fields result in cell death. One therefore thinks of electropermeabilization occurring within a certain threshold of intensity of the local electric field [26].

For treatment planning in electropermeabilization, one is interested in the percentage of electroporated cells over the whole tissue to form decisions in the short term how to gear treatment [64, 26, 81].

One would like supervise the electropermeabilization using measurements of the electric field distribution with image modalities like in [64]. In that work, measurements of magnetic resonance electrical impedance tomography [97] have been employed to find the electrical field distribution. A threshold is then applied to find the electroporated cells.

Yet this approach is only the first step in a larger program:

- the electrical field distribution reconstructed by an imaging modality is a macroscopic quantity;
- the thresholding hypothesis is a simplification and should be refined [26];
- the minimum transmembrane voltage governing electropermeabilization is determined by specific cell characteristics like the curvature of the cell membrane [67].

One solution to find about microscopic parameters from measurements is to take general models and do a specific parameter fitting with preselected

cells like in [26]. In clinical practice, though, a preselected cell population may be unavailable for the analysis.

In this chapter, we tackle the next step in electropermeabilization monitoring and investigate the question to determine microscopic parameters from macroscopic measurements. The modelling used stems from general physiological tissue models for cells, asymptotically simplified by Neu and Krassowska [83]. Whereas the mathematical well-posedness of the model of that model is not available in the literature, there exists an investigation of well-posedness for a similar model in [58]. In this chapter, demonstrate the local well-posedness of the asymptotic cell model of [83], as well as the absence of a blow up. A variant of the model is shown to be globally well-posed.

In order to describe the relation between macroscopic and microscopic quantities, we apply the homogenization scheme in [4] to the cell model of Neu and Krassowska [83]. This not only describes isotropic effective parameters such as classical theory [85], but includes also anisotropy. We provide a convergence analysis for the homogenized solution.

Then we study numerically the sensitivity of the effective parameters to:

- the conductivities of the extra- and intracellular media;
- the shape of the cell membrane;
- the volume fraction of the cells;
- the lattice structure of the cells.

We refer to research in [67, 55, 80, 84, 89], where these critical parameters for electropermeabilization have been investigated, partly from an empirical or computer simulation point of view.

The structure of the chapter is as follows. In Section 5.2, we introduce the model of [83] on the cellular scale. In Section 5.3, we investigate its well-posedness properties. In Section 5.4, we perform the homogenization and show the convergence of the homogenized solution. In Section 5.5, we provide a sensitivity analysis of the effective parameters, showing dependence on microscopic properties, summarized in Table 5.2. A discussion and final remarks in Section 5.6 conclude the article.

5.2 Modelling electropermeabilization on the cellular scale

5.2.1 Membrane model

Let $Y \Subset \mathbb{R}^d$ be a bounded domain representing the cell, and let $\Gamma \subset Y$ be the membrane of the cell. Let

$$Y \setminus \Gamma = Y_i \cup Y_e,$$

where Y_i (resp. Y_e) is be the inner (resp. the outer) domain. Let $\sigma_i(x)$ be the conductivity of the cell domain Y_i, and $\sigma_e(x)$ be the conductivity outside the cells on Y_e.

Let u_0 be an imposed voltage on the boundary of Y. An electrostatic model for the electrical field $u(x,t)$ on Y in the inner and outer domain is

$$\nabla \cdot (\sigma(x)\, \nabla u(x,t)) = 0 \qquad \text{on } Y \setminus \Gamma = Y_i \cup Y_e, \tag{5.2.1}$$

$$u = u_0 \qquad \text{on } \partial Y, \text{ with } \Delta u_0 = 0 \text{ in } Y, \tag{5.2.2}$$

$$\sigma_e \underline{n} \cdot \nabla u^+ = \sigma_i \underline{n} \cdot \nabla u^- = \sigma \underline{n} \cdot \nabla u = \sigma \partial_n u \qquad \text{on } \Gamma. \tag{5.2.3}$$

Here and throughout this chapter, ∂_n denotes the normal derivative.

5.2.2 Electropermeabilization models

In addition to the membrane model, a time-varying conductivity $\sigma_m(x,t)$ for $x \in \Gamma$ is taken account of. The general effect of electropermeabilization is described by relating σ_m and the membrane thickness δ to the transmembrane potential (TMP) jump $u^+(x,t) - u^-(x,t) := [u](x,t)$ in an ordinary differential equation (ODE) on Γ:

$$\sigma(x) \underline{n} \cdot \nabla u(x,t) = \frac{c_m}{\delta} \partial_t [u](x,t) + \frac{\sigma_m\left([u](x,t),t\right)}{\delta} [u](x,t) \qquad \text{on } \Gamma. \tag{5.2.4}$$

Here, the vector $\underline{n}$ is the outward normal to Γ, ∂_n is the normal derivative, the superscripts $\pm$ denote the limits for outside and inside Y_i, and c_m is a positive constant.

The membrane conductivity σ_m in (5.2.4) is described by different models. In [55], Mir et al. propose a static model based on

$$\sigma_m([u]) = \sigma_{m0} + K\ (e^{\beta[u]} - 1), \tag{5.2.5}$$

for some constants σ_{m0}, K, and β, and used the model (5.2.1)-(5.2.4) and (5.2.5) as a boundary-value problem for an elliptic equation with nonlinear transmission conditions at the membrane.

The classical and more involved model for σ_m due to Neu and Krassowska [83] is explained in the following. It assumes that σ_m is the sum of σ_{m0} and an electropermeabilization current. The latter is proportional to the pore density N, which in turn is governed by an ordinary differential equation:

$$\sigma_m([u],t) = \sigma_{m0} + \beta N([u],t) \qquad \text{on } \Gamma \times]0,T[, \tag{5.2.6}$$

$$N([u],0) = N_0(x) \qquad \text{on } \Gamma, \tag{5.2.7}$$

$$\partial_t N([u],t) = \alpha\, e^{\left(\frac{[u](x,t)}{V_{ep}}\right)^2} \left(1 - \frac{N([u],t)}{N_0} e^{-q\left(\frac{[u](x,t)}{V_{\text{ep}}}\right)^2}\right) \qquad \text{on } \Gamma \times]0,T[, \tag{5.2.8}$$

where α, β, q, and N_0 are constants, V_{ep} is the minimum transmembrane voltage for electropermeabilization, and T is the final time.

Given the condition

$$u(x,t) = u_{\text{ref}} \qquad \text{on } Y \text{ for } t < 0, \tag{5.2.9}$$

the initial value problem (5.2.1)-(5.2.4) and (5.2.6)-(5.2.9) is then solved on $Y\times]0,T[$.

Another model for σ_m has been developed in [58]. Together with (5.2.6), (5.2.7), one uses the dynamics

$$\partial_t N([u],t) \quad = \max\left(\tfrac{\beta([u])-N([u],t)}{\tau_{\text{ep}}}, \tfrac{\beta([u])-N([u],t)}{\tau_{\text{res}}}\right)$$

with

$$\beta(\lambda) = (1+\tanh(k_{\text{ep}}(|\lambda|-V_{\text{ep}})))/2,$$

and given constants τ_{ep}, τ_{res}, and k_{ep}.

5.3 Wellposedness of the electropermeabilization model

In this section, we treat the classical electropermeabilization model model (5.2.1)-(5.2.4) and (5.2.6)-(5.2.9) and study it in the form of an ODE on the membrane Γ.

As a preliminary step, let us prove the following representation of the pore density N.

Lemma 5.3.1. *(i) For $[u] = v$, the solution of the initial value problem in (5.2.7), (5.2.8) is*

$$N(x,t) = e^{-\int_0^t \frac{\alpha}{N_0}\, e^{(1-q)\left(\frac{v(x,\tau)}{V_{\text{ep}}}\right)^2} d\tau}\, N_0 + \left(\int_0^t \alpha\, e^{\left(\frac{v(x,s)}{u_0}\right)^2} e^{-\int_s^t \frac{\alpha}{N_0}\, e^{(1-q)\left(\frac{v(x,\tau)}{V_{\text{ep}}}\right)^2} d\tau}\, ds\right). \tag{5.3.1}$$

(ii) The pore density N, considered as a mapping $v(x,t) \mapsto N(v(x,t),t)$

$$C([0,T],C(\Gamma))\times[0,T] \to C([0,T],C(\Gamma)), \tag{5.3.2}$$

maps bounded sets to bounded sets.

Proof. Note that the solution to a linear inhomogeneous ordinary differential equation

$$\frac{\partial}{\partial t} N(t) = A(t)\, N(t) + b(t) \tag{5.3.3}$$

is given by [3, Thm. 5.14]

$$N(t) = U(t,0)\, N_0 + \int_s^t U(t,s)\, b(s)\, ds, \tag{5.3.4}$$

where

$$U(t,s) = \int_s^t A(\tau)\, d\tau.$$

Equation (5.2.8) is a special form of (5.3.3), and the coefficients A and b are

$$A(t) = -\frac{\alpha}{N_0}\, e^{(1-q)\left(\frac{[u](t)}{V_{\text{ep}}}\right)^2},$$

and

$$b(t) = \alpha\, e^{\left(\frac{[u](t)}{V_{\rm ep}}\right)^2}.$$

Inserting A and b into the general solution (5.3.4), we directly obtain the representation (5.3.1) in (i).

Using the norm $\|v\|_{C(\Gamma)} = \sup_{x\in\Gamma} |v(x)|$, the boundedness property in (ii) is then immediate. □

Remark 5.3.1. *In practice, it is clear that the potential v stays finite. One may therefore choose a real number $M > 0$ and work instead of $N(v,t)$ with the function*

$$N_M(v,t) := N(v_M,t) \qquad \text{with } v_M := \begin{cases} |v| & |v| \le M \\ M & |v| > M \\ -M & |v| < M \end{cases}. \tag{5.3.5}$$

For $\|v\|_{L^\infty(\Gamma)} < M$, this cutoff preserves the pore density: $N_M(v,t) = N(v,t)$. In Lemma 5.3.3, it is shown that the function $v \mapsto N_M(v,t)v_M$, considered in $C((0,T);L^2(\Gamma))$, has a global Lipschitz property.

5.3.1 Reduction to an ordinary differential equation

Definition 1 (Stekhlov-Poincaré operators)**.** *Let $H^s(\Gamma)$ be the standard Sobolev space on Γ of order s. Let $f \in H^{\frac{1}{2}}(\Gamma)$ be given. Define solutions of Dirichlet boundary value problems and assign the Neumann data via the* Stekhlov-Poincaré operators *Λ_c, Λ_e: $H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$ and Λ_0: $H^{1/2}(\partial\Omega) \to H^{-1/2}(\Gamma)$,*

$$\Lambda_c f := \partial_n P_1, \quad \Lambda_e f := \partial_n P_2, \quad \Lambda_0 f := \partial_n P_3,$$

where $P_i, i = 1,2,3$ are solutions to

$$\begin{cases} \Delta P_1 &= 0 & \text{in } Y_i, \\ P_1 &= f & \text{on } \Gamma, \end{cases}$$

and

$$\begin{cases} \Delta P_2 &= 0 & \text{in } Y_e, \\ P_2 &= 0 & \text{on } \partial Y, \\ P_2 &= f & \text{on } \Gamma, \end{cases} \qquad \begin{cases} \Delta P_3 &= 0 & \text{in } Y_e, \\ P_3 &= f & \text{on } \partial Y, \\ P_3 &= 0 & \text{on } \Gamma. \end{cases}$$

The following results hold.

Lemma 5.3.2. *(i) Solving the problem* (5.2.1)-(5.2.4) *and* (5.2.6)-(5.2.9) *for $u = (u_i, v, u_e)$ on $Y_i \cup \Gamma \cup Y_e$ is equivalent to solving the initial value problem*

$$\begin{gathered} \frac{c_m}{\delta}\partial_t v + \frac{\sigma_m}{\delta}(v,t)v + \Lambda_c B^{-1} v = G, \\ v(0) = \varphi, \end{gathered} \tag{5.3.6}$$

for v on Γ, with the correspondence

$$\begin{aligned} u_i &= -B^{-1}(v + \Lambda_e^{-1}\Lambda_0 g), \\ u_e &= u_i + v. \end{aligned}$$

Here, $B = Id + \Lambda_e^{-1}\Lambda_0, G = -\Lambda_c B^{-1}\Lambda_e^{-1}\Lambda_0 g$, and

$$\sigma_m(v,t) = \sigma_{m0}(x) + \beta N(v,t). \tag{5.3.7}$$

(ii) The linear operator $\Lambda_c B^{-1} : H^1(\Gamma) \to L^2(\Gamma)$ *is m-accretive. In particular, one has*

$$\forall v : \qquad \langle \Lambda_c B^{-1} v, v \rangle_{L^2} \geq 0, \tag{5.3.8}$$

where $\langle\ ,\ \rangle_{L^2}$ *is the scalar product on* $L^2(\Gamma)$.

Proof. The reduction of the time-dependent model on Ω to the initial value problem on Γ in (5.3.6), using the Steklov-Poincaré operators, is the same as in [58, Lemma 9]. The property in (ii) is shown in [58, Lemma 8]. □

For establishing existence and uniqueness results (in Theorem 5.3.1), we use the following lemma on the Lipschitz property of the function N_M introduced in Remark 5.3.1.

Lemma 5.3.3. *Let* $M > 0$, *and let* $N_M(v,t) = N(v_M, t)$ *with* $v_M = \operatorname{sgn}(v)\min(|v|, M)$ *be the modified pore density defined by* (5.3.5). *Then*

$$v \longmapsto N_M(v,t)v_M$$

is global Lipschitz in $C((0,T); L^2(\Gamma))$.

Proof. Let v_1, $v_2 \in C((0,T); L^2(\Gamma))$. One has the algebraic identity

$$\begin{aligned} &N(v_{1M},t)v_{1M} - N(v_{2M},t)v_{2M} = \\ &\quad \big(N(v_{1M},t)v_{1M} - N(v_{1M},t)v_{2M}\big) + \big(N(v_{1M},t)v_{2M} - N(v_{2M},t)v_{2M}\big). \end{aligned} \tag{5.3.9}$$

Using the boundedness of v_M, (5.3.9) shows that it suffices to prove that $N(v_M,t)$ is global Lipschitz in $C((0,T); L^2(\Omega))$.

Consider the explicit form of $N(v,t)$ in (5.3.1). As $\|v_M\|_{L^\infty} \leq M$, there exists a constant $L(M)$ such that

$$|N(v_{1M},t) - N(v_{2M},t)|^2 \leq L(M) \int_0^t |v_1(x,s) - v_2(x,s)|^2 ds. \tag{5.3.10}$$

Therefore, we have

$$\|N(v_{1M},t)v_{1M} - N(v_{2M},t)v_{2M}\|_{C((0,T);L^2(\Gamma))} \leq C(M)\|v_1 - v_2\|_{C((0,T);L^2(\Gamma))}, \tag{5.3.11}$$

and the global Lipschitz property of N_M in $C((0,T); L^2(\Omega))$ holds. □

Using Lemma 5.3.3, we now come to the well-posedness results. For this end, we introduce the following auxiliary problem. As a variant to (5.2.4), we consider

$$\sigma(x)\underline{\mathrm{n}} \cdot \nabla u(x,t) = \frac{c_m}{\delta}\partial_t[u](x,t) + \frac{\sigma_m\left([u]_M(x,t),t\right)}{\delta}[u]_M(x,t) \quad \text{on } \Gamma. \tag{4'}$$

Using the same procedure as in Lemma 5.3.2, we find that the model (5.2.1)-(5.2.3),(4') and (5.2.6)-(5.2.9) is equivalent to solving

$$\begin{aligned} \frac{c_m}{\delta}\partial_t\tilde{v} + \frac{\sigma_m(\tilde{v}_M,t)}{\delta}\tilde{v}_M + \Lambda_c B^{-1}\tilde{v} &= G, \\ \tilde{v}(0) &= \varphi. \end{aligned} \tag{5.3.12}$$

Let us now state the well-posedness properties of our initial value problems on Γ.

Theorem 5.3.1. *Let $G \in C^1((0,T);H^1(\Gamma))$ and $\varphi \in H^2(\Gamma)$.*

(i) The initial value problem in (5.3.12) *has a unique global solution $\tilde{v} \in C([0,T];H^2(\Gamma))$.*

(ii) For the initial value problem (5.3.6)*, there is a $t_0 > 0$ such that there exists a solution $v \in C([0,t_0[;H^2(\Gamma))$.*

(iii) The solution in (ii) is unique on $C([0,t_1],H^2(\Gamma))$ for any closed interval $[0,t_1] \subset [0,t_0[$.

Proof. (i): Let $M > \|\varphi\|_{L^\infty}$ be a constant and consider the initial value problem (5.3.12). Fix a number $T > 0$.

Due to the global Lipschitz property of $N_M v_M$ shown in Lemma 5.3.3, one can apply the fixed point argument in [58, Thm.10]) to conclude that there exists a unique solution $\tilde{v} \in C([0,T];L^2(\Gamma))$ solving (5.3.12).

If one additionally assumes that $G \in C^1([0,T];H^1(\Gamma))$ and $\varphi \in H^2(\Gamma)$, then one can likewise conclude $\tilde{v} \in C^1([0,T];H^2(\Gamma))$. Then we have that $\partial_n u_i \in L^2(\Gamma)$. With such boundary regularity, we infer $\tilde{u}_i \in H^{3/2}(Y_i)$, similarly $\tilde{u}_e \in H^{3/2}(Y_e)$. Then $\tilde{v} = \tilde{u}_e - \tilde{u}_i \in C([0,T];H^1(\Gamma))$. Using this argument once again, we have that $\tilde{v} = \tilde{u}_e - \tilde{u}_i \in C([0,T];H^2(\Gamma))$.

(ii): We will now show that the solution $\tilde{v}$ to (5.3.12) found in point (i) solves locally the original problem (5.3.6). – Using the Sobolev embedding theorem one has that

$$\Lambda_c B^{-1}\tilde{v} \in C([0,T];H^1(\Gamma)) \hookrightarrow C([0,T];C(\Gamma)).$$

Take a constant C_M such that, for any $t \leq T$, one has

$$\left\| \frac{\sigma_m(\tilde{v}_M,t)}{\delta}\tilde{v}_M + \Lambda_c B^{-1}\tilde{v} + G \right\|_{C(\Gamma)} \leq C_M.$$

Define

$$t_0 := \frac{c_m}{\delta}\frac{M - \|\varphi\|_{L^\infty}}{C_M}.$$

Then, for $t \leq t_0$, one gets

$$\begin{aligned} \|\tilde{v}(x,t)\|_{L^\infty(\Gamma)} &\leq \|\varphi\|_{L^\infty} + tC_M, \\ &\leq M. \end{aligned}$$

But for $\|\tilde{v}\|_\infty < M$, one has that $\tilde{v}_M = \tilde{v}$ and $N_M(\tilde{v},t) = N(\tilde{v},t)$. Therefore, the expressions in (5.3.6) and (5.3.12) are the same, which implies that, locally, $\tilde{v}$ solves as well the original initial value problem (5.3.6).

(iii): Take two solutions v, w to (5.3.6) in $C^1([0,t_1],H^2(\Gamma))$. Due to closedness of $[0,t_1]$ and continuity of the norm $\|.\|_{H^2} \to \mathbb{R}$, there exists a $M > 0$ such that for every $t \in [0,t_1]$, one has

$$\|v(t)\|_{H^2} < M \qquad \text{and} \qquad \|w(t)\|_{H^2} < M.$$

But then the cutoff with respect to M does not change the functions: $v_M = v$ and $w_M = w$. Therefore, v and w also solve (5.3.12). But for that ODE, one has a global uniqueness property. Therefore $v = w$ on $[0,t_1]$. □

We now give a more detailed analysis of the terms in equation (5.3.6) to show that a solution cannot blow up in finite time (see Theorem 5.3.2).

Note that for σ_m given by (5.2.6), there exists a $C \in \mathbb{R}$ such that one has for all v that

$$\langle \sigma_m(v,t)v, v\rangle_{L^2} \geq C\|v\|^2_{L^2}. \tag{5.3.13}$$

This immediately follows from the expression of the membrane conductivity in (5.2.6) and the fact that both the pore density N as well as N_M in (5.3.5) are positive.

Theorem 5.3.2. *For a function $v \in C^1([0,t_0[, L^2(\Gamma))$ which solves* (5.3.6)*, it is impossible that*

$$\|v(t_k)\|_{L^2(\Gamma)} \longrightarrow_{t\to b} \infty \qquad \textit{for } b \in [0,t_0[.$$

Proof. Take as an indirect assumption a blow up-sequence $\|v(t_k)\|_X \to \infty$ with $t_k \to b$. Without loss of generalization, we may choose $t_k \in [0,t_0[\cap W$, where W is a neighborhood of b such that v is nonzero on $[0,t_0[\cap W$. Due to the C^1-regularity property of $v(t)$ and $v \neq 0$, the function

$$[0,t_0[\cap W \to \mathbb{R}: \qquad t \longmapsto \|v(t)\|_{L^2}$$

is then continuously differentiable.

The sequence $t_k \to b(x)$ having the Cauchy property, the slope of the secants satisfies

$$\frac{|\|v(t_{k+1})\|_- \|v(t_k)\||}{t_{k+1}-t_k} \longrightarrow \infty,$$

as well. We then will work with a sequence τ_k such that

$$\partial_t \|v(\tau_k)\|_{L^2} \longrightarrow \infty, \tag{5.3.14}$$

chosen by the mean-value theorem.

Consider equivalently to (5.3.6) the equation

$$\sigma(v)v = G - C_m\partial_t v - \Lambda_c B^{-1}v.$$

Take the L^2-scalar product with v and take account of $\langle \partial_t v, v\rangle = \|v\|\ \partial_t\|v\|$. Then estimate the right-hand side with the Cauchy-Schwarz inequality and the accretivity property (5.3.8):

$$\begin{aligned}\langle \sigma(v)v, v\rangle_{L^2} &= \langle G, v\rangle_{L^2} - C_m\langle \partial_t v, v\rangle_{L^2} - \langle \Lambda_c B^{-1}v, v\rangle_{L^2},\\ &\leq \|G\|_{L^2}\|v\|_{L^2} - C_m\|v\|_{L^2}\ \partial_t\|v\|_{L^2}.\end{aligned}$$

Divide by $\|v\|_{L^2}$ to find

$$\frac{\langle \sigma(v)v, v\rangle_2}{\|v\|_{L^2}} \leq \|G\|_{L^2} - C_m\partial_t\|v\|_{L^2}. \tag{5.3.15}$$

From (5.3.13), we already know that the left-hand side stays positive.

Evaluate then expressions in inequality (5.3.15) for the sequence τ_k in (5.3.14). The result is that the right-hand side would tend to $-\infty$, which is impossible. This shows that no blow up of v in L^2 can occur. □

5.4 Homogenization

Let Ω be a bounded domain in $\mathbb{R}^2$, which carries a periodic structure made up by periodic open sets εY. The reference domain $Y = Y_i \cup Y_e \cup \Gamma$ contains a cell inside with membrane Γ, where Y_i is the intracellular domain and Y_e is the extracellular domain. The whole domain Ω is thus composed of

$$\Omega = \Omega^+ \cup \Omega^- \cup \Gamma_\varepsilon,$$

where Ω^+ is the collection of extracellular domains, Ω^- is the collection of intracellular domains and Γ_ε is the collection of membranes.

We write the thickness of the membrane of the cells εY in the form

$$\delta = \varepsilon\, \delta_0,$$

where ϵ is the scale of the cell and δ_0 is the reference cell membrane thickness for Y.

As in [10], we want to study behavior of the electrical field on this cell cluster and recover features of the microscopic cell model from tissue measurements. Considering the cell model in (5.2.1)-(5.2.4) and (5.2.6)-(5.2.9) for a domain Y, we first give the model equation for u_ε in Ω:

$$\begin{aligned}
\nabla \cdot (\sigma(x)\nabla u_\varepsilon(x,t)) &= 0 && \text{in } \Omega^+,\\
\nabla \cdot (\sigma(x)\nabla u_\varepsilon(x,t)) &= 0 && \text{in } \Omega^-,\\
[\sigma \nabla u_\varepsilon \cdot \mathbf{n}] &= 0 && \text{on } \Gamma_\varepsilon,\\
\frac{c_m}{\delta}\frac{\partial}{\partial t}[u_\varepsilon] + \frac{1}{\delta}\sigma_m([u_\varepsilon]_M, t)[u_\varepsilon]_M &= \sigma \partial_n u_\varepsilon^- && \text{on } \Gamma_\varepsilon,\\
[u_\varepsilon](x,0) &= S_\varepsilon && \text{on } \Gamma_\varepsilon,\\
u_\varepsilon(x,t) &= 0 && \text{on } \partial\Omega,
\end{aligned} \tag{5.4.1}$$

where $S_\varepsilon(x) = \varepsilon S_1(x, \frac{x}{\varepsilon}) + R(\varepsilon)$ and $\sigma_m = \sigma_{m0} + \beta N([u_\varepsilon], t)$. The pore density $N([u_\varepsilon], t)$ is governed by (5.2.8).

Here, in the second equation on Γ_ε, the quantity $[u_\varepsilon]_M$ is understood in the sense of the definition in (5.3.5), i.e., $[u_\varepsilon]_M = \operatorname{sgn}([u_\varepsilon]) \min(|[u_\varepsilon]|, M)$ for a constant $M > 0$.

Given the physical observation that the voltage v stays bounded, it is reasonable that for proper $M > 0$, the system (5.4.1) is an accurate model for the real potential. Given Lemma 5.3.2 and Theorem 5.3.1, it is also well-posed.

We want to explore the limit of the solution u_ε as $\varepsilon \to 0$. For this end, we start with an energy estimate on the solution u_ε which will be needed later when investigating the limit.

Proposition 5. *(i) We have for u_ε in* (5.4.1) *the energy estimate*

$$\int_0^t \int_\Omega \sigma |\nabla u_\varepsilon|^2 dx\, dt + \frac{1}{\varepsilon}\int_{\Gamma_\varepsilon} [u_\varepsilon]^2(x,t) dS \le C. \tag{5.4.2}$$

(ii) In particular, the estimate

$$\int_{\Gamma_\varepsilon} [u_\varepsilon]^2 dS \le C\varepsilon \tag{5.4.3}$$

holds.

Proof. Multiply (5.4.1) by u_ε, then integrate by parts to find

$$\int_0^t \int_\Omega \sigma |\nabla u_\varepsilon|^2 dx\, dt + \frac{\alpha}{2\varepsilon} \int_{\Gamma_\varepsilon} [u_\varepsilon]^2(x,t) dS + \frac{1}{\varepsilon} \int_0^t \int_{\Gamma_\varepsilon} \sigma_m([u_\varepsilon]_M, \tau)[u_\varepsilon][u_\varepsilon]_M(x,t) dS\, dt = \frac{\alpha}{2\varepsilon} \int_{\Gamma_\varepsilon} [S_\varepsilon]^2(x) dS. \quad (5.4.4)$$

The statement is then derived from the fact that

$$\sigma_m [u_\varepsilon][u_\varepsilon]_M \geq 0$$

and $S_\varepsilon(x) = \varepsilon S_1(x, \frac{x}{\varepsilon}) + o(\varepsilon)$. □

For now, let us formally assume that the solution u_ε of (5.4.1) has the form

$$u_\varepsilon(x,t) = u_0(x,t) + \varepsilon u_1(x, \frac{x}{\varepsilon}, t) + o(\varepsilon). \quad (5.4.5)$$

We will calculate the equation for u_0 in Subsection 5.4.1 and then prove rigorously that u_ε converges in an appropriate sense to u_0 in Subsection 5.4.2.

5.4.1 Formal calculation of the homogenization limit

To find the precise form of the terms in the ansatz (5.4.5), we can apply the arguments developed in [4]. For this end, it is required that for the membrane conductivity one has that

$$\sigma_m(0,t) = \text{constant}.$$

(see [4, Secs. 3.2 and 3.3]). This condition can be ensured for the model (5.2.6), together with (5.2.8): From (5.3.1), one can prove that $N(0,t) = N_0$, and therefore $\sigma_m(0,t) = \text{constant}$.

Before calculating the limit, we first give some definitions. Introduce the transform

$$T : H^{1/2}(\Gamma) \to C([0,T], H^1_p(Y)),$$

where

$$H^1_p(Y) = \left\{ u \text{ is periodic in } Y : u|_{Y_i} \in H^1(Y_i) \text{ and } u|_{Y_e} \in H^1(Y_e),\ \int_Y u = 0 \right\},$$

by

$$T(s)(y,t) := v(y,t)$$

with v being the solution to the following system with boundary data s:

$$\begin{aligned}
\nabla \cdot (\sigma(x) \nabla v) &= 0 && \text{in } Y_i, \\
\nabla \cdot (\sigma(x) \nabla v) &= 0 && \text{in } Y_e, \\
[\sigma \nabla v \cdot \mathbf{n}] &= 0 && \text{on } \Gamma, \\
\frac{c_m}{\delta_0} \frac{\partial}{\partial t}[v] + \frac{1}{\delta_0} \sigma_m(0,t)[v] &= \sigma \partial_n v^- && \text{on } \Gamma, \\
[v](x,0) &= s && \text{on } \Gamma.
\end{aligned}$$

We define next the cell problems $\chi^0 : \Omega \to \mathbb{R}^d$ and $\chi^1 : \Omega \times (0,T) \to \mathbb{R}^d$. For this, let $\underline{e}_h$ be the h-th unit vector in $\mathbb{R}^d$. Then the component $\chi^0_h \in H^1_p(Y)$ satisfies

$$\begin{aligned} \nabla \cdot (\sigma(x)\nabla\chi^0_h) &= 0 && \text{in } Y_i, \\ \nabla \cdot (\sigma(x)\nabla\chi^0_h) &= 0 && \text{in } Y_e, \\ [\sigma(\nabla_y\chi^0_h - \underline{e}_h)\cdot \underline{n}] &= 0 && \text{on } \Gamma, \\ [\chi^0_h](x,0) &= 0 && \text{on } \Gamma. \end{aligned}$$

The component χ^1_h is defined by

$$\chi^1_h = T(\sigma(\nabla_y\chi^0_h - \underline{e}_h)\cdot \underline{n}). \tag{5.4.6}$$

By a calculation analogous to [4, Sec.3], one finds that the candidate u_0 in equation (5.4.5) satisfies

$$\operatorname{div}\left[-\sigma_0\nabla_x u_0 - A^0\nabla_x u_0 - \int_0^t A^1(t-\tau)\nabla_x u_0(x,\tau)d\tau + \underline{F}(x,t)\right] = 0. \tag{5.4.7}$$

Here, the matrices A^0, A^1, and $\underline{F}(x,t)$ are defined by

$$\begin{cases} \sigma_0 = \sigma_1|Y_i| + \sigma_2|Y_e|, \\ (A^0)_{jh} = \displaystyle\int_\Gamma [\sigma]\chi^0_h n_j dS, \\ (A^1)_{jh} = \displaystyle\int_\Gamma [\sigma\chi^1_h] n_j dS, \\ \underline{F} = \displaystyle\int_\Gamma [\sigma T(S_1(x,\cdot))](y,t)\ \underline{n}\ dS, \end{cases} \tag{5.4.8}$$

where $\sigma_i = \sigma|_{Y_i}$ and $\sigma_e = \sigma|_{Y_e}$, with χ^0_h, χ^1_h and T given above.

5.4.2 Convergence

While in Subsection 5.4.1, we derived the formal limit (5.4.7) for the ansatz of the asymptotic expansion (5.4.5), we now state its convergence properties.

Theorem 5.4.1. *For the periodic solution u_ε in* (5.4.1) *and the homogenized solution u_0 in* (5.4.7)*, we have the convergence*

$$u_\varepsilon \to u_0$$

weakly in $L^2([0,T]\times\Omega)$ and strongly in $L^1_{\text{loc}}([0,T],\Omega)$.

The proof relies on arguments developed in [4]. For the sake of a readability, we outline them in the appendix, and only prove here the crucial lemma needed for their adaption to our case.

Lemma 5.4.1. *For $M > 1$, there exists a constant $C(M)$ such that*

$$\int_0^T \int_{\Gamma_\varepsilon} |\sigma_m(0,t)[u_\varepsilon] - \sigma_m([u_\varepsilon]_M, t)\ [u_\varepsilon]_M\ dS\ |dt \le C\varepsilon. \tag{5.4.9}$$

Symbol	Value	Definition
σ_i	0.455	intracellular conductivity
σ_e	5	extracellular conductivity
L	2×10^{-4}	computation domain size
r	0.5×10^{-4}	cell radius
δ	5×10^{-9}	membrane thickness
r_p	0.76	pore radius
σ_p	0.0746	pore conductivity
V_{ep}	0.258	characteristic voltage of electropermeabilization
α	10^9	electropermeabilization parameter
N_0	1.5×10^9	equilibrium pore density
c_m	9.5×10^{-12}	membrane capacitance

TABLE 5.1: Model parameters used for the numerical computations.

Proof. We have

$$\begin{aligned} & \sigma_m(0,t)[u_\varepsilon] - \sigma_m([u_\varepsilon]_M,t)\ [u_\varepsilon]_M \\ = \ & \sigma_m(0,t)[u_\varepsilon] - \sigma_m([u_\varepsilon]_M,t)\ [u_\varepsilon] + \sigma_m([u_\varepsilon]_M,t)\ ([u_\varepsilon] - [u_\varepsilon]_M). \end{aligned}$$

By the explicit form of $N(v,t)$ in (5.3.1) and $|v_M|_{L^\infty} \leq M$, there exists a constant $L(M)$ such that

$$|N([u_\varepsilon]_M,t) - N(0,t)|^2 \leq L(M) \int_0^t [u_\varepsilon]_M^2 ds, \tag{5.4.10}$$

and $\sigma_m([u_\varepsilon]_M,t) \leq C(M)$.

Together with the fact that $\left| \int_0^T [u_\varepsilon] - [u_\varepsilon]_M \ ds \right| \leq \int_0^T [u_\varepsilon]^2 ds$, we can thus conclude that

$$\int_0^T \int_{\Gamma_\varepsilon} \left| \sigma_m(0,t)[u_\varepsilon] - \sigma_m([u_\varepsilon]_M,t)\ [u_\varepsilon]_M \right| \ dS\ dt \leq C(M)\varepsilon.$$

The lemma then follows by the energy estimate (5.4.3). □

5.5 Numerical experiments

In the preceding section, we have modeled macroscopic processes as homogenized quantities with specific effective material parameters. In this section we show the sensitivity of the effective parameters to microscopic properties relevant in electropermeabilization.

We use FEM with mesh generator [86] to implement all the numerical simulations. We present the numerical experiments from two aspects: First we will simulate the single cell model (5.3.6) and show the electropermeabilization at cell level. Next we show how the microscopic parameters affect effective parameters and anisotropy properties in the homogenized model (5.4.7).

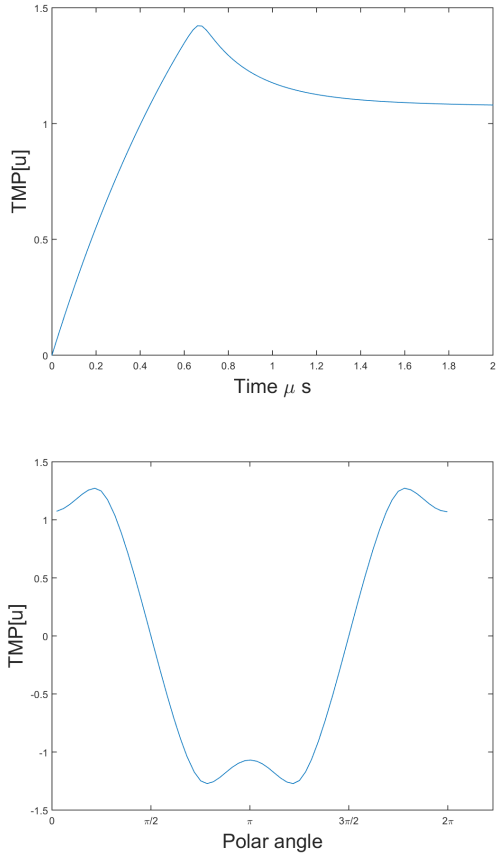


FIGURE 5.1: **(a)** Evolution of the transmembrane potential (TMP) v at the pole of the cell. **(b)** TMP along the cell membrane after 2 $\mu s'$.

5.5.1 Electropermeabilization simulation for a single cell

We simulate the single cell model (5.3.6) in a square domain $[0, L]\times[0, L]$, the cell is a circular in the center of the square with cell radius r. The parameter β in (5.2.8) is given by

$$\beta = \frac{2\pi r_p^2 \sigma_p \delta}{\pi r_p + 2\delta}. \tag{5.5.1}$$

All the parameters are given in Table 5.1. Figure 5.1 shows the results for the time evolution and the voltage after 2 $\mu s'$.

5.5.2 Homogenization for electropermeabilization model

In this section, we show the sensitivity of the effective parameters σ_0, A^0, and A^1 in (5.4.7) to

- the conductivities σ_o and σ_i;
- the shape of the cell with membrane Γ;
- the volume fraction $f = \frac{\text{vol}(Y_i)}{\text{vol}(Y)}$;
- the lattice of the cells in the domain Ω.

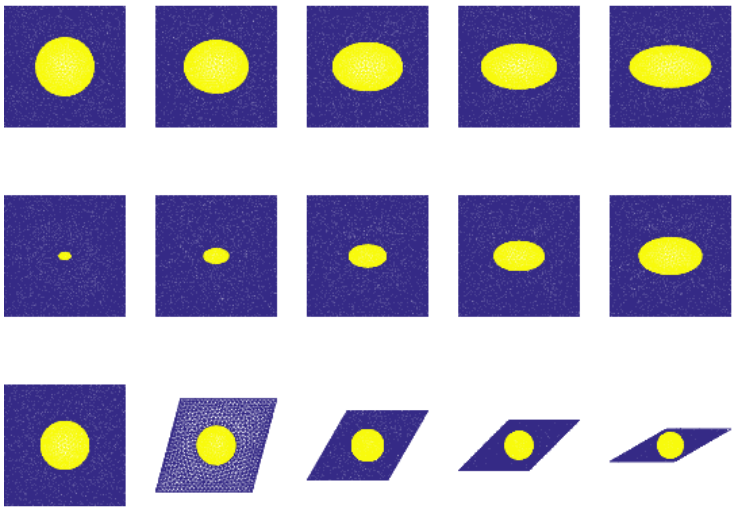


FIGURE 5.2: Cell shapes used in numerical examples (see text and Table 5.2). Example 1 uses the first mesh. Example 2 uses the cells in the first row. Example 3 uses the cells in the second row. Example 4 uses the cells in the last row.

We perform four experiments, the results of which are found in Table 5.2.

Example 1. We fix the shape and size of the cell and change the ratio of the interior and exterior conductivities σ_i and σ_e.

Example 2. In this example, we show how the shape of the cell membrane produces different effective anisotropy properties. We fix conductivities and the volume fraction of the cell, but take as cell shapes ellipses with different excentricity a/b.

Example 3. We investigate the effect of different volume fractions of a cell with the same shape.

Example 4. In this example, we show how the angle of the lattice in which the cells are arranged affects the effective parameters.

For all these experiments, Table 5.2 presents the reactions of the effective conductivity σ_0 and the effective anisotropy properties A^0 and $A^1(0)$ to the microscopical change. One sees clearly that σ_0, as well as A^0 and A^1 react to a change of cell and conductivity parameters. Most of the sensitivity functions are in fact monotonic.

The best contrast is seen in:

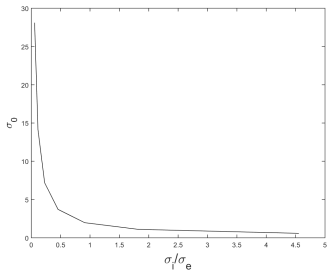
- the reaction of σ_0 to the change in conductivity σ_i/σ_e and to a change in the lattice angle ϕ;
- the reaction of both A^0 and A^1 to the cell shape.

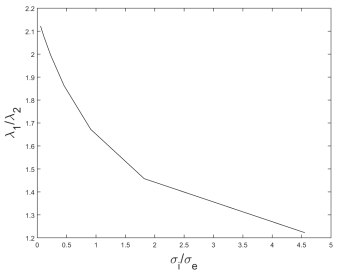
The volume fraction alone does not show so much contrast in the anisotropy of the effective parameters.

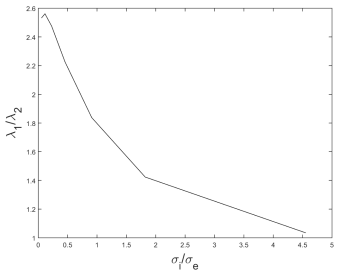
Given the results of the sensitivity analysis, it is promising to infer shape parameters from macroscopic effective properties in electropermeabilization, as it was done in [10] from multifrequency admittivity measurements.

effective conductivity σ_0	eigenvalues λ_1/λ_2 of A^0	eigenvalues λ_1/λ_2 of $A^1(0)$.

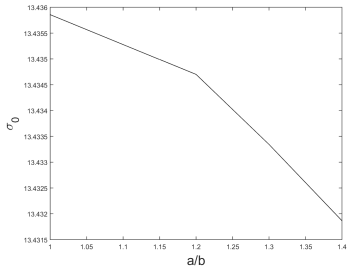
Example 1: Difference in conductivity (ratio σ_i/σ_e of interior and exterior conductivity).

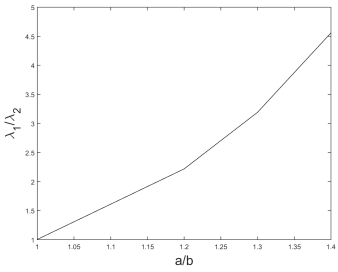


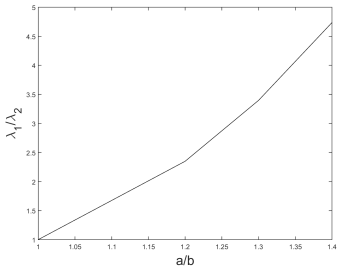




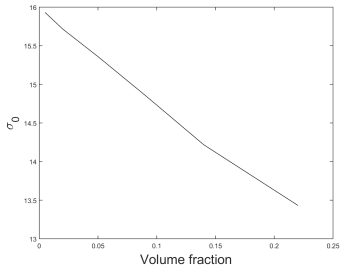
Example 2: Difference in cell shape: change of the excentricity a/b (see Fig. 5.2, 1st row).

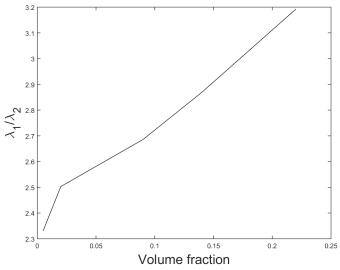


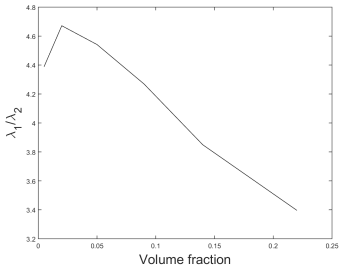




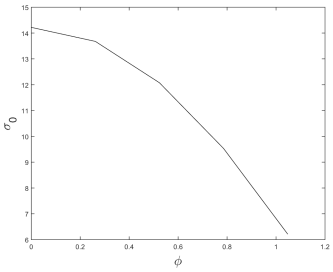
Example 3: Difference in volume fraction of the cells (see Fig. 5.2, 2nd row).

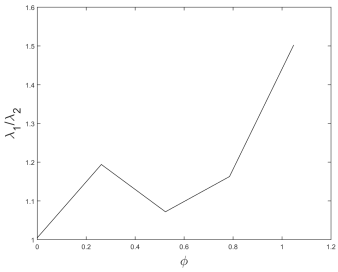






Example 4: Difference in angle ϕ of the lattice arrangement (see Fig. 5.2, 3rd row).





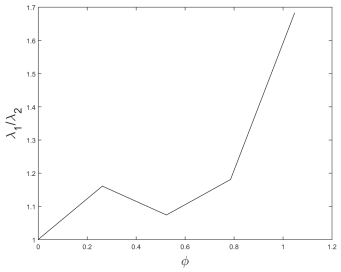


TABLE 5.2: Changes in microscopic parameters and the reaction of the effective parameters in (5.4.7).

5.6 Concluding remarks

We introduced a homogenization scheme relating critical microscopic and macroscopic quantities in electropermeabilization. The sensitivity analysis of the effective parameters showed this dependence and opens the door to solve the inverse problem to monitor those critical microscopic quantities in practice.

While setup optimization for electropermeabilization has been studied using computer simulations, for instance, in [81, 105, 36, 79, 80], from our approach comes an additional constraint: for mapping of the effective parameters A^1 and A^0, two currents have to be applied which are nowhere parallel. An electrode configuration providing this allows for unique reconstruction [61].

Concluding remarks

In this thesis, two different parts have been discussed.

In many scientific and engineering applications involving partial differential equations, the input data such as sources or boundary conditions are usually given through the measurements which may be subject to random noises. We have considered two cases in the finite element method with observational data in Part I: the thin plate spline model and the elliptic boundary equations with uncertain boundary data. We have analyzed the stochastic convergence based on the empirical process theory and investigated the stochastic convergence of the FEM which characterizes the tail property of the probability distribution function of the finite element error.

The methods used in these two cases can be extended to other forward problems with random input. This gives a way to choose the optimal discrete size (mesh size or system's degree of freedoms) to balance the random error. Also in the case of the thin plate spline model it has been shown that the self-consistent iterative Algorithm 1.5.1 is quiet efficient for determining the smoothing parameter. This kind of algorithm might be applied to other similar problems in practice.

Another promising aspect of the thesis is inverse problems. With uncertain input data, inverse problems can be very unstable because of their inherent ill-posedness. The methods introduced in the first parts pave a way to improve these issues. Yet more work is needed in this direction.

In Part II, we have discussed two imaging methods: the linearized model in multi-frequency EIT and the imaging of anisotropic conductivity using Diffusion Tensor, for isotropic and anisotropic conductivity reconstruction respectively. We have also analyzed the well-posedness of the cell model for electropermeabilization and a dynamical homogenization scheme. The mathematical models in this part help us to understand the dependence of the conductivity of the tissue on the frequency and the microoaganisation of the cells.

For the first imaging method, we have systematically discussed mfEIT reconstruction in the following three different scenarios: known spectral profiles, partially known spectral profiles and unknown spectral profiles. This analysis generalizes the existing studies. It has very promising applications to other multi-frequency imaging problems and multi-wavelength problems.

The second imaging method gives a way to reconstruct anisotropic conductivity distributions with DIT. We have firstly formulated a new image reconstruction method of an anisotropic conductivity tensor distribution by combining the MAT-MI and the DTI techniques. But a more difficult and important problem is to reconstruct full anisotropy conductivity and to develop a stable imaging technique for such problems.

We have also analyzed the well-posedness of the cell model for electropermeabilization and proposed a dynamical homogenization scheme.

We have studied the effective parameters in a homogenization model as the next step to monitor the microscopic properties in clinical practice. We have numerically demonstrated the sensitivity of these effective parameters to critical microscopic parameters governing electropermeabilization. This opens the door to solve the inverse problem of reconstructing these parameters.

With these three chapters in Part II, we could better understand the effective conductivity and have better ways to image the isotropic or anisotropic conductivity distributions. Also, they show the challenging directions for improving our understanding of the corresponding fundamental problems.

Appendix A

Proof of Lemma 1.3.3 when $d = 3$

The proof is very similar to the proof for 2D case in Section 1.4. We will construct $\Pi_h v_h$ by using the three dimensional C^1 element of Zhang constructed in [121] which simplifies an earlier construction of Zenisek [120]. For any tetrahedron $K \in \mathcal{M}_h$, the $C^1 - P_9$ element in [121] is a triple (K, P_K, Λ_K), where $P_K = P_9(K)$ and the set of degrees of freedom Λ_K consists of the following 220 functionals: for any $p \in C^2(K)$,

1° The nodal values of $p(a_i), Dp(a_i)(a_j - a_i), D^2p(a_i)(a_j - a_i, a_k - a_i), D^3(a_i)(a_j - a_i, a_k - a_i, a_l - a_i), D^4p(a_i)(a_j - a_i, a_k - a_i, a_l - a_i, a_n - a_i), 1 \le i \le 4, 1 \le j \le k \le l \le n \le 4, i \notin \{j, k, l, n\}$, where $\{a_i\}_{i=1}^4$ are the vertices of K; (120 functionals)

2° The 2 first order normal derivatives $\partial_{\nu_k} p(a_{ij})$ and 3 second order normal derivatives $\partial^2_{\nu_k\nu_l} p(b_{ij}), \partial^2_{\nu_k\nu_l} p(c_{ij})$ on the edge with vertices a_i, a_j, $1 \le i \ne j \le 4$, where $\nu_k, k = 1, 2$, are unit vectors perpendicular to the edge, and $a_{ij} = (a_i + a_j)/2$, $b_{ij} = (2a_i + a_j)/3$, $c_{ij} = (a_i + 2a_j)/3$; (48 functionals)

3° The nodal value $p(a_{ijk})$ and 6 normal derivatives $\partial_\nu p(a^n_{ijk})$ on the face with vertices a_i, a_j, a_k, $1 \le i, j, k \le 4, i \ne j, j \ne k, k \ne i$, $n = 1, 2, \cdots, 6$, where a_{ijk} is the barycenter of the face and $a^1_{ijk} = (2a_i + a_j + a_k)/4, a^2_{ijk} = (a_i + 2a_j + a_k)/4, a^3_{ijk} = (a_i + a_j + 2a_k)/4, a^4_{ijk} = (4a_i + a_j + a_k)/6, a^5_{ijk} = (a_i + 4a_j + a_k)/6, a^6_{ijk} = (a_i + a_j + 4a_k)/6$; (24 functionals)

4° The nodal values $p(d_i)$, $1 \le i \le 4$, at internal points $d_1 = (2a_1 + a_2 + a_3 + a_4)/5, d_2 = (a_1 + 2a_2 + a_3 + a_4)/5, d_3 = (a_1 + a_2 + 2a_3 + a_4)/5, d_4 = (a_1 + a_2 + a_3 + 2a_4)/5$. (4 functionals)

Let X_h be the finite element space

$$X_h = \{v_h : v_h|_K \in P_9(K), \forall K \in \mathcal{M}_h, f(v_h|_{K_1}) = f(v_h|_{K_2}), \forall f \in \Lambda_{K_1} \cap \Lambda_{K_2}\}.$$

It is known that $X_h \subset H^2(\Omega)$. We define the operator Π_h as follows. For any $v_h \in V_h$, $w_h := \Pi_h v_h \in X_h$ such that for any $K \in \mathcal{M}_h$, $w_h|_K \in P_9(K)$, for the degrees of freedom at vertices a_i, $1 \le i \le 4$,

$$\partial^\alpha (w_h|_K)(a_i) = \frac{1}{N(a_i)} \sum_{K' \in \mathcal{M}_h(a_i)} \partial^\alpha (v_h|_{K'})(a_i), \quad |\alpha| \le 4, \tag{A.0.1}$$

for the degrees of freedom on the edge with vertices a_i, a_j, $1 \le i \ne j \le 4$,

$$\partial_{\nu_k}(w_h|_K)(a_{ij}) = \frac{1}{N(a_{ij})} \sum_{K' \in \mathcal{M}_h(a_{ij})} \partial_{\nu_k}(v_h|_{K'})(a_{ij}), \quad k = 1, 2, \tag{A.0.2}$$

$$\partial_{\nu_k\nu_l}(w_h|_K)(b_{ij}) = \frac{1}{N(b_{ij})} \sum_{K' \in \mathcal{M}_h(b_{ij})} \partial_{\nu_k\nu_l}(v_h|_{K'})(b_{ij}), \quad k, l = 1, 2, \tag{A.0.3}$$

$$\partial_{\nu_k\nu_l}(w_h|_K)(c_{ij}) = \frac{1}{N(c_{ij})} \sum_{K' \in \mathcal{M}_h(c_{ij})} \partial_{\nu_k\nu_l}(v_h|_{K'})(c_{ij}), \quad k, l = 1, 2, \tag{A.0.4}$$

for the degrees of freedom on the faces with vertices a_i, a_j, a_k, $1 \le i, j, k \le 4, i \ne j, j \ne k, k \ne i$,

$$(w_h|_K)(a_{ijk}) = \frac{1}{N(a_{ijk})} \sum_{K' \in \mathcal{M}_h(a_{ijk})} (v_h|_{K'})(a_{ijk}), \tag{A.0.5}$$

$$\partial_\nu(w_h|_K)(a^n_{ijk}) = \frac{1}{N(a^n_{ijk})} \sum_{K' \in \mathcal{M}_h(a^n_{ijk})} \partial_\nu(v_h|_{K'})(a^n_{ijk}), \quad n = 1, 2 \cdots, 6, \tag{A.0.6}$$

and finally for the degrees of freedom at the interior points $d_i, 1 \le i \le 4$,

$$(w_h|_K)(d_i) = (v_h|_K)(d_i). \tag{A.0.7}$$

To show the desired estimate (1.18) in 3D we use the C^0-P_9 element in [121] which is a triple (K, P_K, Θ_K), where $P_K = P_9(K)$ and the set of degrees of freedom Θ_K is defined by replacing some of the degrees of freedom of the $C^1 - P_9$ element Λ_K as follows:

1° For the edge with vertices a_i, a_j, $1 \le i \ne j \le 4$, replace the 2 edge first order normal derivatives by $Dp(a_{ij})(a_k - a_{ij}), Dp(a_{ij})(a_l - a_{ij})$ and denote the corresponding nodal basis functions $p^k_{ij}(x), p^l_{ij}(x)$, where a_k, a_l are the other 2 vertices of K other than a_i, a_j;

2° For the edge with vertices a_i, a_j, $1 \le i \ne j \le 4$, replace the 3 edge second order normal derivatives by $D^2p(b_{ij})(a_k - b_{ij}, a_l - b_{ij}), D^2p(c_{ij})(a_k - b_{ij}, a_l - b_{ij})$ and denote the corresponding nodal basis functions $p^{kl}_{ij}(x), q^{kl}_{ij}(x)$, where a_k, a_l are the other 2 vertices of K other than a_i, a_j;

3° For the face with vertices a_i, a_j, a_k, $1 \le i, j, k \le 4, i \ne j, j \ne k, k \ne i$, replace the face normal derivatives by $Dp(a^n_{ijk})(a_l - a^n_{ijk})$ and denote the corresponding nodal basis functions $p^n_{ijk}(x)$, where a_l is the vertex of K other than a_i, a_j, a_k, $n = 1, 2 \cdots, 6$.

A regular family of this $C^0 - P_9$ element is affine-equivalent. For any $v_h \in V_h$, we also define an operator $q_h := \Lambda_h v_h$ in a similar way as the definition of Π_h by replacing the average normal derivatives in (A.0.2)-(A.0.4) and (A.0.6) by the corresponding directional derivatives in the definition of degrees of freedom for the $C^0 - P_9$ element. By the same argument as that

in the proof of 2D case in Section 1.4 we have

$$|v_h - q_h|_{H^m(K)} \leq Ch^{2-m} \left(\sum_{K' \in \mathcal{M}_h(K)} |v_h|^2_{H^2(K')} \right)^{1/2}, \quad m = 0, 1, 2. \quad \text{(A.0.8)}$$

Next we expend $q_h - w_h \in P_9(K)$ in terms of the nodal basis functions of the $C^0 - P_9$ element. From the definition of the $C^1 - P_9$ and $C^0 - P_9$ elements, we have $q_h - w_h = \phi_e + \phi_f$ in K, where the edge part of the function $q_h - w_h$ is

$$\begin{aligned}
\phi_e(x) \quad = \quad & \sum_{\substack{1 \leq i \neq j \leq 4 \\ \{k,l\} \in \{1,2,3,4\} \setminus \{i,j\}, k \neq l}} \Big[D(q_h|_K - w_h|_K)(a_{ij})(a_k - a_{ij}) p^k_{ij}(x) \\
& \qquad + D(q_h|_K - w_h|_K)(a_{ij})(a_l - a_{ij}) p^l_{ij}(x) \Big] \\
& + \sum_{\substack{1 \leq i \neq j \leq 4 \\ \{k,l\} \in \{1,2,3,4\} \setminus \{i,j\}, k \leq l}} \Big[D^2(q_h|_K - w_h|_K)(b_{ij})(a_k - b_{ij}, a_l - b_{ij}) p^{kl}_{ij}(x) \\
& \qquad + D^2(q_h|_K - w_h|_K)(c_{ij})(a_k - c_{ij}, a_l - c_{ij}) q^{kl}_{ij}(x) \Big],
\end{aligned}$$

and the face part of the function $q_h - w_h$ is

$$\phi_f(x) = \sum_{\substack{1 \leq i,j,k \leq 4, i \neq j, j \neq k, k \neq i \\ \{l\} \in \{1,2,3,4\} \setminus \{i,j,k\}}} \sum_{n=1}^{6} D(q_h|_K - w_h|_K)(a^n_{ijk})(a_l - a^n_{ijk}) p^n_{ijk}(x).$$

Since the tangential derivatives of $q_h - w_h$ along the edges vanish, we obtain by the same argument as that in the proof of 2D case in Section 1.4 that

$$|\phi_e|_{H^m(K)} \leq Ch^{2-m} \left(\sum_{K' \in \mathcal{M}_h(K)} |v_h|^2_{H^2(K')} \right)^{1/2}, \quad m = 0, 1, 2. \quad \text{(A.0.9)}$$

On any face F of K, $q_h - w_h - \phi_e \in P_9(F)$ and its nodal values at 3 vertices up to 4th order derivatives vanish, its first order normal derivative at the midpoint and two second order normal derivatives at two internal trisection points on 3 edges vanish, and the nodal value at the barycenter also vanishes. This implies $q_h - w_h - \phi_e = 0$ on any face of the element K. Let τ^n_{ijk} be the tangential unit vector on the face of vertices a_i, a_j, a_k such that

$$a_l - a^n_{ijk} = [(a_l - a^n_{ijk}) \cdot \tau^n_{ijk}] \tau^n_{ijk} + [(a_l - a^n_{ijk}) \cdot \nu] \nu.$$

Now by (A.0.4), (A.0.8)-(A.0.9), and the inverse estimate we have

$$\begin{aligned}
& |D(q_h|_K - w_h|_K)(a^n_{ijk})(a_l - a^n_{ijk})| \\
\leq \quad & |[(a_l - a^n_{ijk}) \cdot \tau^n_{ijk}] D\phi_e(a^n_{ijk}) \tau^n_{ijk}| + |[(a_l - a^n_{ijk}) \cdot \nu] D(q_h|_K - w_h|_K)(a^n_{ijk}) \nu| \\
\leq \quad & Ch^{1/2} \left(\sum_{K' \in \mathcal{M}_h(K)} |v_h|^2_{h^2(K')} \right)^{1/2}. \qquad \text{(A.0.10)}
\end{aligned}$$

Since a regular family of $C^0 - P_9$ element is affine-equivalent, we have $|p^n_{ijk}|_{H^m(K)} \leq Ch^{3/2-m}$, $m = 0, 1, 2$. Therefore, by (A.0.10) we obtain

$$|\phi_f|_{H^m(K)} \leq Ch^{2-m} \left(\sum_{K' \in \mathcal{M}_h(K)} |v_h|^2_{H^2(K')} \right)^{1/2}, \quad m = 0, 1, 2. \tag{A.0.11}$$

Combining (A.0.8), (A.0.9), (A.0.11) yields the desired estimate (1.18) in 3D since $v_h - w_h = (v_h - q_h) + \phi_e + \phi_f$ in K. The estimate (1.19) can be proved in the same way as the proof for the 2D case in Section 1.4. This completes the proof. □

Appendix B

Convergence for homogenization

B.1 Convergence for homogenization

We give here the outline of the method used in [4]. It shows how Lemma 5.4.1 is used to prove Theorem 5.4.1 for our application.

Theorem B.1.1. *For the solution u_ε in* (5.4.1) *and the homogenized solution u_0 in* (5.4.7)*, we have the convergence*

$$u_\varepsilon \to u_0$$

weakly in $L^2([0,T]\times\Omega)$ and strongly in $L^1_{\text{loc}}([0,T],\Omega)$.

Proof. From the estimate (5.4.2) we get, extracting subsequences if needed

$$\begin{aligned} &u_\varepsilon \to u_0, \qquad \sigma\nabla u_\varepsilon \to \xi \qquad \text{weakly in } L^2([0,T]\times\Omega),\\ &u_\varepsilon \to u_0 \qquad\qquad\qquad\quad \text{strongly in } L^1_{\text{loc}}([0,T],\Omega). \end{aligned} \tag{B.1.1}$$

Next, consider the weak formulation of system (5.4.1):

$$\begin{aligned} &\int_0^T\int_\Omega \sigma\nabla u_\varepsilon\cdot\nabla\psi\ dx\ dt + \frac{1}{\varepsilon}\int_0^T\int_{\Gamma_\varepsilon}\sigma_m([u_\varepsilon]_M)[u_\varepsilon]_M[\psi]\ dS\ dt\\ &\quad -\frac{c_m}{\delta}\int_0^T\int_{\Gamma_\varepsilon}[u_\varepsilon]\frac{\partial}{\partial t}[\psi]\ dS\ dt - \frac{c_m}{\delta}\int_{\Gamma\varepsilon}[u_\varepsilon](0)[\psi](0)\ dS = 0. \end{aligned} \tag{B.1.2}$$

The general idea is to pass to the limit $\varepsilon\to 0$ in this equation, and therefore to obtain the equation for u_0. This is possible for special test functions ψ.

Choose for ψ the functions $\varphi\ w_h^\varepsilon(x,t)$ for $h=1,\dots,d$, where φ is a smooth with compact support on Ω, and w_h^ε is built by the cell functions χ^1 and χ^2:

$$w_h^\varepsilon(x,t) := x_h - \varepsilon\chi_h^0\left(\frac{x}{\varepsilon}-\varepsilon\right)\int_t^T \chi_h^1\left(\frac{x}{\varepsilon},\tau-t\right)\ d\tau.$$

For this definition, given in [4, (5.1)] one has the weak formulation in [4, (5.2)-(5.4)].

By subtracting the weak equation (B.1.2) for $\psi = w_h^\varepsilon(x,t)$ and the equations [4, (5.2)-(5.4)], one can isolate the term $\int\int\sigma\nabla u_\varepsilon\nabla\varphi w_h^\varepsilon\ dx\ dt$:

$$\int_0^T\int_{\Gamma_\varepsilon}\sigma\nabla u_\varepsilon\nabla\varphi w_h^\varepsilon\ dx\ dt = K_{1\varepsilon}+K_{2\varepsilon}+K_{3\varepsilon}, \tag{B.1.3}$$

with

$$
\begin{aligned}
K_{1\varepsilon} &= \int_0^T \int_{\Gamma_\varepsilon} \sigma \nabla w_h^\varepsilon \nabla \varphi u_\varepsilon \, dx \, dt, \\
K_{2\varepsilon} &= -c_m \varepsilon \int_{\Gamma_\varepsilon} (S_1(x, \frac{x}{\varepsilon}) + R_\varepsilon)\varphi \int_0^T [\chi_h^2](\frac{x}{\varepsilon}, \tau) d\tau dS, \\
K_{3\varepsilon} &= \frac{1}{\varepsilon} \int_0^T \int_{\Gamma_\varepsilon} \Big(\sigma_m(0,t)[u_\varepsilon] - \sigma_m([u_\varepsilon], t)[u_\varepsilon] \Big) [w_h^\varepsilon] \varphi \, dS \, dt.
\end{aligned}
\tag{B.1.4}
$$

The limits of $K_{1\varepsilon}$ and $K_{2\varepsilon}$ are the same as in [4, p.18], whereas for the limit $K_{3\varepsilon}$, one can show that $K_{3\varepsilon} \to 0$ by Lemma 5.4.1. One can take then the limit $\varepsilon \to 0$ in (B.1.3) in order to obtain information on the specific form of the limit u_0 in (B.1.1). We get

$$
\begin{aligned}
-\int_0^T \int_\Omega \xi \cdot \nabla \varphi x_h \, dx \, dt = \int_0^T \int_\Omega \varphi(x) F_h(x,\tau); dx \, d\tau \\
+ \int_0^T \int_\Omega u_0(x,t)(\sigma_0 I + A^0)\underline{\mathrm{e}}_h + \int_0^t u_0(x,\tau) A^1(t-\tau)\underline{\mathrm{e}}_h \, d\tau \cdot \nabla \varphi(x) \, dx \, dt
\end{aligned}
\tag{B.1.5}
$$

with A^0, A^1, $\underline{\mathrm{F}}$ defined as in (5.4.8). Choosing $\psi = \varphi \, x_h$ in (B.1.2), combining with (B.1.5), and differentiating in T gives then expressions which show that $u_0 \in L^2([0,T], H^1(\Omega))$ and that actually (5.4.7) is the correct equation of the limit u_0.

□

Bibliography

[1] A. Adler, R. Guardo, and Y. Berthiaume. "Impedance imaging of lung ventilation: do we need to account for chest expansion". In: *IEEE Trans. Biomed. Eng.* 43.4 (1996), pp. 414–420.

[2] Andy Adler, Romina Gaburro, and William Lionheart. "Electrical Impedance Tomography". English. In: *Handbook of Mathematical Methods in Imaging*. Ed. by O. Scherzer. Springer New York, 2011, pp. 599–654. ISBN: 978-0-387-92919-4. DOI: 10.1007/978-0-387-92920-0_14. URL: http://dx.doi.org/10.1007/978-0-387-92920-0_14.

[3] H. Amann. *Ordinary differential equations. An introduction to nonlinear analysis*. de Gruyter Studies in Mathematics. Walter de Gruyter, Berlin, New York, 1990.

[4] M. Amar, D. Andreucci, P. Bisegna, and R. Gianni. "Evolution and momory effects in the homogenization limit for electrical conduction in biological tissues". In: *Math. Models Methods Appl. Sci.* 14 (2004), pp. 1261–1295.

[5] H. Ammari, S. Boulmier, and P. Millien. "A mathematical and numerical framework for magnetoacoustic tomography with magnetic induction". In: *J. Differential Equations* 259 (2015), pp. 5379–5405.

[6] H. Ammari, T. Widlak, and W. Zhang. "Towards monitoring critical microscopic parameters for electropermeabilization". In: *Quarterly of Applied Mathematics* 75 (2017), pp. 1–17.

[7] H. Ammari, L. Giovangigli, L. H. Nguyen, and J. K. Seo. *Admittivity imaging from multi-frequency micro-electrical impedance tomography*. arXiv:1403.5708.

[8] H. Ammari, L. Qiu, F. Santosa, and W. Zhang. "Determining anisotropic conductivity using diffusion tensor imaging data in magneto-acoustic tomography with magnetic induction". In: (), arXiv:1702.05187.

[9] H. Ammari, L. Giovangigli, H. Kwon, J.K. Seo, and T. Wintz. "Spectroscopic conductivity imaging of a cell culture". In: *Asympt. Anal.* 100 (2016), pp. 87–109.

[10] H. Ammari, J. Garnier, L. Giovangigli, W. Jing, and J.K. Seo. "Spectroscopic imaging of a dilute cell suspension". In: *J. Math. Pures Appl.* 105 (2016), pp. 603–661.

[11] H. Ammari, G. S. Alberti, B. Jin, J. K. Seo, and W. Zhang. "The Linearized inverse problem in multifrequency electrical impedance tomography". In: *SIAM Journal on Imaging Sciences* 9 (2016), pp. 1525–1551.

[12] R. Arcangéli, R. Manzanilla, and J.J. Torrens. "Approximation spline de surfaces de type explicite comportant des failles". In: *Math. Model. Numer. Anal.* 31 (1997), pp. 643–676.

[13] R. Arcangéli, R. Manzanilla, and J.J. Torrens. "Approximation spline de surfaces de type explicite comportant des failles". In: *Math. Model. Numer. Anal.* 31 (1997), pp. 643–676.

[14] D. L. Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet. "Imaging ofintravoxel Incoherent Motions - Application to Diffusion and Perfusion in Neurologic Disorders". In: *Radiology* 161 (1986), pp. 401–407.

[15] M.S. Birman and M.Z. Solomyak. "Piecewise polynomial approximations of functions of the classes W_α^k". In: *Mat. Sb.* 73 (1997), pp. 331–355.

[16] S.C. Brenner and L.R. Scott. *Mixed Finite Element Methods and Applications*. Springer-Verlag, Berlin, 2013.

[17] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, New York, 1994.

[18] Z. Chen, T. Rui, and W. Zhang. "A proper sampling finite element method for elliptic problems with observational boundary data". In: *arXiv:1702.05188* ().

[19] Z. Chen, T. Rui, and W. Zhang. "Stochastic Convergence of A Nonconforming Finite Element Method for the Thin Plate Spline Smoother for Observational Data". In: *arXiv:1701.08626* ().

[20] K.S. Cheng, D. Isaacson, J. C. Newell, and D. G. Gisser. "Electrode models for electric current computed tomography". In: *IEEE Trans. Biomed. Eng.* 36.9 (1989), pp. 918–924.

[21] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.

[22] P. Clément. "Approximation by finite element functions using local regularization". In: *RAIRO Anal. Numer.* 9 (1975), pp. 77–84.

[23] A. Cohen and R. DeVore. "Approximation of high-dimensional parametric PDEs". In: *Acta Numerica* 24 (2015), pp. 1–159.

[24] I. Daubechies, M. Defrise, and C. De Mol. "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint". In: *Comm. Pure Appl. Math.* 57.11 (2004), pp. 1413–1457. ISSN: 0010-3640. DOI: 10.1002/cpa.20042. URL: http://dx.doi.org/10.1002/cpa.20042.

[25] Philip J. Davis. *Interpolation and approximation*. Blaisdell Publishing Co. Ginn and Co. New York-Toronto-London, 1963, pp. xiv+393.

[26] J. Dermol and D. Miklavčič. "Predicting electroporation of cells in an inhomogeneous electric field based on mathematical modeling and experimental CHO-cell permeabilization to propidium iodide determination". In: *Bioelectrochemistry* 100 (2014), pp. 52–61.

[27] J. Duchon. "Splines minimizing rotation-invariant semi-norms in Sobolev spaces". In: *Constructive Theory of Functions of Several Variables, Lecture Notes in Mathematics* 571 (1997), pp. 85–100.

[28] D. Finch and Rakesh. *Recovering a function from its spherical mean values in two and three dimensions*. Photoacoustic Imaging and Spectroscopy, 2009.

[29] C. Gabriel, A. Peyman, and E. H. Grant. "Electrical conductivity of tissue at frequencies below 1MHz". In: *Phys. Med. Biol.* 54.16 (2009), pp. 4863–4878.

[30] S. Gabriely, R. W. Lau, and C. Gabriel. "The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz". In: *Phys. Med. Biol.* 41.11 (1996), pp. 2251–2269.

[31] L. A. Geddes and L. E. Baker. "The specific resistance of biological material – A compendium of data for the biomedical engineeer and physiologist". In: *Med. & Biol. Engng* 3 (1967), pp. 271–293.

[32] S.A. van de Geer. *Empirical process in M-estimation*. Cambridge University Press, Cambridge, 1978.

[33] M. Gehre, B. Jin, and X. Lu. "An analysis of finite element approximation in electrical impedance tomography". In: *Inverse Problems* 30.4 (2014), pp. 045013, 24. ISSN: 0266-5611. DOI: 10.1088/0266-5611/30/4/045013. URL: http://dx.doi.org/10.1088/0266-5611/30/4/045013.

[34] E. Gersing, B. Hofmann, and M. Osypka. "Influence of changing peripheral geometry on electrical impedance tomography measurements". In: *Med. Biol. Eng. Comput.* 34.5 (1996), pp. 359–361.

[35] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*. Vol. 224. Springer Science & Business Media, 2001.

[36] A. Golberg and B. Rubinsky. "Towards electroporation based treatment planning considering electric field induced muscle contractions". In: *Technol. Cancer. Res. Treat.* 11 (2012), pp. 189–201.

[37] Joseph F. Grcar. "Spectral condition numbers of orthogonal projections and full rank linear least squares residuals". In: *SIAM J. Matrix Anal. Appl.* 31.5 (2010), pp. 2934–2949. ISSN: 0895-4798. DOI: 10.1137/090777773. URL: http://dx.doi.org/10.1137/090777773.

[38] H. Griffiths and A. Ahmed. "A dual-frequency applied potential tomography technique: computer simulations". In: *Clin. Phys. Physiol. Meas.* 8.4A (1987), pp. 103–107.

[39] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman Advanced Publishing, 1985.

[40] M.D. Gunzburger, C.G. Webster, and G. Zhang. "Stochastic finite element methods for partial differential equations with random input data". In: *Acta Numerica* 23 (2014), pp. 521–650.

[41] M. Haltmeier, T. Schuster, and O. Scherzer. "Filtered backprojection for thermoacoustic computed tomography in spherical geometry". In: *Math. Methods Appl. Sci.* 28 (2005), pp. 1919–1937.

[42] M. Hanke, A. Neubauer, and O. Scherzer. "A convergence analysis of the Landweber iteration for nonlinear ill-posed problems". In: *Numer. Math.* 72 (1995), pp. 21–37.

[43] M. Hanke and O. Scherzer. "Inverse problems light: numerical differentiation". In: *The Amer. Math. Monthly* 108.6 (2001), pp. 512–521.

[44] F. Hettlich. "Erratum: Frechet derivatives in inverse scattering". In: *Inverse Problems* 14.1 (1998), pp. 209–210.

[45] Frank Hettlich. "Fréchet derivatives in inverse obstacle scattering". In: *Inverse Problems* 11.2 (1995), pp. 371–382. ISSN: 0266-5611. URL: http://stacks.iop.org/0266-5611/11/371.

[46] N. Hoell, A. Moradifam, and A. Nachman. "Current Density Impedance Imaging of an Anisotropic Conductivity in a Known Conformal Class". In: *SIAM J. Math Anal.* 46 (2014), pp. 1820–1842.

[47] D. Holder. *Electrical Impedance Tomography: Methods, History and Applications*. Institute of Physics, Bristol, 2005.

[48] R. Holm. *Electric Contacts: Theory and Applications*. Berlin Heidelberg: Springer-Verlag, 1967.

[49] M. V. de Hoop, L. Qiu, and O. Scherzer. "An analysis of a multi-level projected steepest descent iteration for nonlinear inverse problems in Banach spaces subject to stability constraints". In: *Numerische Mathematik* 129 (2015), pp. 127–148.

[50] Y. Hristova, P. Kuchment, and L. Nguyen. "Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media". In: *Inverse Problems* 24 (2008), p. 055006.

[51] J. Huang, X. Huang, and D. Metaxas. "Learning with dynamic group sparsity". In: *2009 IEEE 12th Int. Conf. Computer Vision*. 2009, pp. 64–71.

[52] J.H. Hwang, K. Kirkpatrick, T. Mason, and E. Garboczi. "Experimental limitations in impedance spectroscopy: Part iv. electrode contact effects". In: *Solid State Ionics* 98.12 (1997), pp. 93–104.

[53] K. Ito and B. Jin. *Inverse Problems: Tikhonov Theory and Algorithms*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015, pp. x+318. ISBN: 978-981-4596-19-0.

[54] A. Ivorra. "Tissue electroporation as a bioelectric phenomenon: Basic concepts". In: *B. Rubinsky, editor, Irreversible Electroporation, Series in Biomedical Engineering* 11 (2010), pp. 23–61.

[55] A. Ivorra, J. Villemejane, and L. M. Mir. "Electrical modeling of the influence of medium conductivity on electroporation". In: *Phys. Chem. Chem. Phys.* 12 (2010), pp. 10055–10064.

[56] J. Jang and J. K. Seo. "Detection of admittivity anomaly on high-contrast heterogeneous backgrounds using frequency difference EIT". In: *Phys. Meas.* 36.6 (2015), pp. 1179–1192.

[57] B. Jin, Dirk A. Lorenz, and S. Schiffler. "Elastic-net regularization: error estimates and active set methods". In: *Inverse Problems* 25.11 (2009), pp. 115022, 26. ISSN: 0266-5611. DOI: 10.1088/0266-5611/25/11/115022. URL: http://dx.doi.org/10.1088/0266-5611/25/11/115022.

[58] O. Kavian, M. Leguèbe, C. Poignard, and L. Weynans. "electropermeabilization modeling at the cell scale". In: *J. Math. Biol.* 68 (2014), pp. 235–265.

[59] N. Keshava and J. F. Mustard. "Spectral unmixing". In: *IEEE Signal Proc. Mag.* 19.1 (2002), pp. 44–57.

[60] S. Kim and A. Tamasan. "Reconstructing small perturbations in electrical admittivity at low frequencies". In: *Inverse Problems* 30.3 (2014), pp. 035006, 18. ISSN: 0266-5611. DOI: 10.1088/0266-5611/30/3/035006. URL: http://dx.doi.org/10.1088/0266-5611/30/3/035006.

[61] Y.J. Kim, O. Kwon, J.K. Seo, and E.J. Woo. "Uniqueness and convergence of conductivity imge reconstruction in magnetic resonance electrical impedance tomography". In: *Inverse Problems* 19 (2003), pp. 1213–1225.

[62] V. Kolehmainen, M. Lassas, and P. Ola. "Electrical impedance tomography problem with inaccurately known boundary and contact impedances". In: *IEEE Trans. Med. Imag.* 27.10 (2008), pp. 1404–1414.

[63] V. Kolehmainen, M. Lassas, and P. Ola. "The inverse conductivity problem with an imperfectly known boundary". In: *SIAM J. Appl. Math.* 66.2 (2005), pp. 365–383. ISSN: 0036-1399. DOI: 10.1137/040612737. URL: http://dx.doi.org/10.1137/040612737.

[64] M. Kranjc, B. Markelc, F. Bajd, M. Čemažar, T. Blagus I. Serša, and D. Miklavčič. "In situ monitoring of electric field distribution in mouse tumor during electroporation". In: *Radiology* 274 (2015), pp. 115–123.

[65] P. Kuchment and L. Kunyansky. "Mathematics of thermoacoustic tomography". In: *European J. Appl. Math.* 19 (2008), pp. 191–224.

[66] O. I. Kwon, W. C. Jeong, S. Z. K Sajib, H. J. Kim, and E. J. Woo. "Anisotropic conducticity tensor imaging in MREIT using directional diffusion rate of water molecules". In: *Phys. Medicine Biol.* 59 (2014), pp. 2955–2974.

[67] T. Kotnik L. Towhidi, G. Pucihar, S.M.P. Firoozabadi, H. Mozdarani, and D. Miklavčič. "Variability of the minimal transmembrane voltage resulting in detectable membrane electroporation". In: *Electromagn. Biol. Med.* 27 (2008), pp. 372–385.

[68] L. Landweber. "An iteration formula for Fredholm integral equations of the first kind". In: *Amer. J. Math.* 73 (1951), pp. 615–624.

[69] S. Laufer, A. Ivorra, V. E. Reuter, B. Rubinsky, and S. B. Solomon. "Electrical impedance characterization of normal and cancerous human hepatic tissue". In: *Physiol. Meas.* 31.7 (2010), pp. 995–1009.

[70] M. Lenoir. "Optimal isoparametric finite elements and error estimates for domains involving curved boundaries". In: *SIAM J. Numer. Anal.* 23 (1986), pp. 562–580.

[71] X. Li and B. He. "Multi-excitation magnetoacoustic tomography with magnetic induction for bioimpedance imaging". In: *IEEE Trans Med Imaging* 29 (2010), pp. 1759–1767.

[72] X. Li, Y. Xu, and B. He. "Magnetoacoustic tomography with magnetic induction for imaging electrical impedance of biological tissue". In: *Journal of applied physics* 99 (2006), p. 066112.

[73] A. Logg, K.A. Mardal, G. N. Wells, and et al. *Automated Solution of Differential Equations by the Finite Element Method*. Springer, 2012.

[74] Q. Ma and B. He. "Magnetoacoustic tomography with magnetic induction: A rigorous theory". In: *IEEE Trans Biomed Eng.* 55 (2008), pp. 813–816.

[75] E. Malone, G. Sato dos Santos, D. Holder, and S. Arridge. "A reconstruction-classification method for multifrequency electrical impedance tomography". In: *IEEE Trans. Med. Imag.* 34.7 (2015), pp. 1486–1497.

[76] E. Malone, G. Sato dos Santos, D. Holder, and S. Arridge. "Multifrequency electrical impedance tomography using spectral constraints". In: *IEEE Trans. Med. Imag.* 33.2 (2014), pp. 340–350.

[77] L. Mariappan, G. Hu, and B. He. "Magnetoacoustic tomography with magnetic induction for high-resolution bioimepedance imaging through vector source reconstruction under the static field of mri magnet". In: *Medical physics* 41 (2014), p. 022902.

[78] O.G. Martinsen, S. Grimnes, and H.P. Schwan. "Interface phenomena and dielectric properties of biological tissue". In: *Encylopedia Surface Colloid Sci.* (2002), pp. 2643–2652.

[79] D. Miklavčič, D. Šemrov, H. Mekid, and L. M. Mir. "A validated model of in vivo electric field distribution in tissues for electrochemotherapy and for DNA electrotransfer for gene therapy". In: *Biochim. Biophys. Acta* 1523 (2000), pp. 73–83.

[80] D. Miklavčič, K. Beravs, D. Šemrov, M. Čemačar, F. Demsar, and G. Serša. "The importance of electric field distribution for effective in vivo electroporation of tissues". In: *Biophys. J.* 74 (1998), pp. 2152–5158.

[81] D. Miklavčič et al. "Towards treatment planning and treatment of deep-seated solid tumors by electrochemotherapy". In: *Biomed. Eng. Online* 9 (2010).

[82] L.S.D. Morley. "The triangular equilibrium element in the solution of plate bending problems". In: *Aero. Quart.* 19 (1968), pp. 208–234.

[83] J. C. Neu and W. Krassowska. "Asymptotic model of electroporation". In: *Phys. Rev. E* 59 (1999), pp. 3471–3482.

[84] M. Pavlin, N. Pavšelj, and D. Miklavčič. "Dependence of induced transmembrane potential on cell density, arrangement and cell position inside a cell system". In: *IEEE Trans. Biomed. Eng.* 49 (2002), pp. 605–612.

[85] M. Pavlin, T. Slivnik, and D. Miklavčič. "Effective conductivitiy of cell suspensions". In: *IEEE Trans. Biomed. Eng.* 49 (2002), pp. 77–80.

[86] P.O. Persson and G. Strang. "A simple mesh generator in MATLAB". In: *SIAM Rev.* 46 (2004), pp. 329–345.

[87] J. Pitkäranta. "Boundary subspaees for the finite element method with Lagrange multipliers". In: *Numer. Math.* 33 (1979), pp. 272–289.

[88] J. Pitkäranta. "Local stability conditions for the Babuška method of Lagrange multiplier". In: *Math. Comp.* 35 (1980), pp. 1113–1129.

[89] G. Pucihar, T. Kotnik, B. Valič, and D. Miklavčič. "Numerical determination of transmembrane voltage induced on irregularly shaped cells". In: *Ann. Biomed. Eng.* 34 (2006), pp. 642–652.

[90] L. Qiu and F. Santosa. "Analysis of the Magnetoacoustic Tomography with Magnetic Induction". In: *SIAM Journal on Imaging Sciences* 8 (2015), pp. 2070–2086.

[91] X. Ren and J. Wei. "On a two-dimensional elliptic problem with large exponent in nonlinearity". In: *Trans. Amer. Math. Soc.* 343 (1994), pp. 749–763.

[92] S. Roberts, M. Hegland, and I. Altas. "Approximation of a thin plate spline smoother using continuous piecewise polynomial functions". In: *SIAM J. Numer. Anal.* 41 (2003), pp. 208–234.

[93] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. Springer, New York, 2009, pp. xiv+320. ISBN: 978-0-387-30931-6.

[94] J. Schlappa, E. Annese, and H. Griffiths. "Systematic errors in multi-frequency EIT". In: *Physiol. Meas.* 21.1 (2000), pp. 111–118.

[95] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. *Regularization Methods in Banach Spaces*. Walter de Gruyter GmbH & Co. KG, Berlin, 2012, pp. xii+283. ISBN: 978-3-11-025524-9. DOI: 10.1515/9783110255720. URL: http://dx.doi.org/10.1515/9783110255720.

[96] J. K. Seo, J. Lee, S. W. Kim, H. Zribi, and E. J. Woo. "Frequency-difference electrical impedance tomography (fdEIT): algorithm development and feasibility study". In: *Phys. Meas.* 29.8 (2008), pp. 929–944.

[97] J.K. Seo and E.J. Woo. "Magnetic resonance electrical impedance tomography (MREIT)". In: *SIAM Rev.* 53 (2011), pp. 40–68.

[98] J.K. Seo, F.C. Pyo, C. Park, O. Kwon, and E.J. Woo. "Image reconstruction of anisotropic conductivity tensor distribution in MREIT: computer simulation study". In: *Phys. Medicine Biol.* 49 (2004), pp. 4371–4382.

[99] Z.C. Shi. "On the error estimates of Morley element". In: *Numerica Mathematica Sinica* 12 (1990), pp. 113–118.

[100] I. Babuˇ ska. "The finite element method with Lagrangian multipliers". In: *Numer. Math.* 20 (1973), pp. 179–192.

[101] I. Babuˇ ska, R. Tempone, and G.E. Zouraris. "Galerkin finite element approximations of Stochastic elliptic partial differential equations". In: *SIAM J. Numer. Anal.* 42 (2004), pp. 800–825.

[102] E. Somersalo, M. Cheney, and D. Isaacson. "Existence and uniqueness for electrode models for electric current computed tomography". In: *SIAM J. Appl. Math.* 52.4 (1992), pp. 1023–1040. ISSN: 0036-1399. DOI: 10.1137/0152060. URL: http://dx.doi.org/10.1137/0152060.

[103] P. Stefanov and G. Uhlmann. "Thermoacoustic tomography with variable sound speed". In: *Inverse Problems* 25 (2009), p. 075011.

[104] R. Stenberg. "On some techniques for approximating boundary conditions in the finite element method". In: *J. Comp. Appl. Math.* 63 (1995), pp. 139–148.

[105] K. Sugibayashi, M. Yoshida, K. Mori, T. Watanabe, and T. Hasegawa. "Electric field analysis on the improved skin concentration of benzoate by electroporation". In: *Int. J. Pharm.* 219 (2001), pp. 107–112.

[106] A.J. Surowiec, S.S. Stuchly, J.R. Barr, and A. Swarup. "Dielectric properties of breast carcinoma and the surrounding tissues". In: *Biomedical Engineering, IEEE Transactions on* 35.4 (1988), pp. 257–263. ISSN: 0018-9294. DOI: 10.1109/10.1374.

[107] J. Sylvester. "An anisotropic inverse boundary value problem". In: *Comm. Pure Appl. Math.* 43.2 (1990), pp. 201–232. ISSN: 0010-3640. DOI: 10.1002/cpa.3160430203. URL: http://dx.doi.org/10.1002/cpa.3160430203.

[108] T. Tang and T. Zhou. "Recent developments in high order numerical methods for uncertainty quantification". In: *Sci. Sin. China* 45 (2015), pp. 891–928.

[109] M. E. Taylor. *Partial differential equations I. Basic theory, vol. 115 of Applied Mathematical Sciences*. Springer, New York, second ed., 2011.

[110] D. S. Tuch, V. J. Wedeen, A. M. Dale, J. S. George, and J. W. Belliveau. "Conductivity tensor mapping of the human brain using diffusion tensor MRI". In: *Proc. Natl. Acad. Sci.* 20 (1998), pp. 11697–11701.

[111] F.I. Utreras. "Convergence rates for multivariate smoothing spline functions". In: *J. Approx.Theory* 52 (1988), pp. 1–27.

[112] A.W. van der Vaart and J.A. Wellner. *Weak Convergence and Empirical Processes: with Applications to Statistics*. Springer, New York, 1996.

[113] K. Vervier, P. Mahé, D'Aspremont, J.B. Veyrieras, and J.P. Vert. "On learning matrices with orthogonal columns or disjoint supports". In: *LNCS 8726*. Springer-Verlag, 2014, pp. 274–289.

[114] G. Wahba. *Spline Models for Observational Data*. SIAM, Philadelphia, 1990.

[115] G. Wahba. *Spline Models for Observational Data*. SIAM, Philadelphia, 1990.

[116] M. Wang and J. Xu. "The Morley element for fourth order elliptic equations in any dimensions". In: *Numer. Math.* 103 (2006), pp. 155–169.

[117] R. Xia, X. Li, and B. He. "Magnetoacoustic tomographic imaging of electrical impedance with magnetic induction". In: *Applied physics letters* 91 (2007), p. 083903.

[118] Y. Xu and B. He. "Magnetoacoustic tomography with magnetic induction (MAT-MI)". In: *Physics in Medicine and Biology* 50 (2005), pp. 5175–5187.

[119] R. J. Yerworth, R. H. Bayford, B. Brown, P. Milnes, M. Conway, and D. S. Holder. "Electrical impedance tomography spectroscopy (EITS) for human head imaging". In: *Physiol. Meas.* 24.2 (2003), pp. 477–489.

[120] A. Zenisek. "Alexander polynomial approximation on tetrahedrons in the finite element method". In: *J. Approx. Theory* 7 (1973), pp. 334–351.

[121] S. Zhang. "A family of 3D continuously differentiable finite elements on tetrahedral grids". In: *Appl. Numer. Math.* 59 (2009), pp. 219–233.

[122] L. Zhou, S. Zhu, and B. He. "A reconstruction algorithm of magnetoacoustic tomography with magnetic induction for an acoustically inhomogeneous tissue". In: *IEEE transactions on bio-medical engineering* 61 (2014), pp. 1739–1746.

Résumé

Cette thèse contient deux sujets différents. Dans la première partie, nous avons considéré deux cas. L'un est le modèle plus lisse de la plaque mince et l'autre est les équations des limites elliptiques avec des données limites incertaines. Dans cette partie, les convergences stochastiques des méthodes des éléments finis sont prouvées pour chaque problème.

Dans la deuxième partie, nous fournissons une analyse mathématique du problème inverse linéarisé dans la tomographie d'impédance électrique multifréquence. Nous présentons un cadre mathématique et numérique pour une procédure d'imagerie du tenseur de conductivité électrique anisotrope en utilisant une nouvelle technique appelée Tentomètre de diffusion Magnéto-acoustographie et proposons une approche de contrôle optimale pour reconstruire le cross-property facteur reliant le tenseur de diffusion au tenseur de conductivité électrique anisotrope. Nous démontrons la convergence et la stabilité du type Lipschitz de l'algorithme et présentons des exemples numériques pour illustrer sa précision. Le modèle cellulaire pour Electropermécanisme est démontré. Nous étudions les paramètres efficace dans un modèle d'homogénéisation. Nous démontrons numériquement la sensibilité de ces paramètres efficaces aux paramètres microscopiques critiques régissant l'électropermécanisme.

Mots Clés

Méthode des éléments finis, données d'observation, imagerie de conductivité, multi fréquences, homogénéisation, électropermécanisme

Abstract

This thesis contains two different subjects. In first part, two cases are considered. One is the thin plate spline smoother model and the other one is the elliptic boundary equations with uncertain boundary data. In this part, stochastic convergences of the finite element methods are proved for each problem.

In second part, we provide a mathematical analysis of the linearized inverse problem in multifrequency electrical impedance tomography. We present a mathematical and numerical framework for a procedure of imaging anisotropic electrical conductivity tensor using a novel technique called Diffusion Tensor Magneto-acoustography and propose an optimal control approach for reconstructing the cross-property factor relating the diffusion tensor to the anisotropic electrical conductivity tensor. We prove convergence and Lipschitz type stability of the algorithm and present numerical examples to illustrate its accuracy. The cell model for Electropermeabilization is demonstrated. We study effective parameters in a homogenization model. We demonstrate numerically the sensitivity of these effective parameters to critical microscopic parameters governing electropermeabilization.

Keywords

Finite element method, observational data, conductivity imaging, multi frequency, homogenization, electropermeabilization