Stefan Köthe, BSc BSc

Deterministic and stochastic shape and topology optimization of an electrical machine

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Mathematics

submitted to

Graz University of Technology

Supervisor

Dr. P. Gangl

Institute of Applied Mathematics

Graz, May 2020

Abstract

This thesis deals with a design optimization problem of an electrical machine coming from a real world application. The problem is motivated from an electrical drive used in an X-Ray tube. Here, a synchronous reluctance machine is considered as drive for a rotary anode X-Ray tube because of its beneficial properties. The performance of the machine depends strongly on the electromagnetic fields in its interior which among others depend on the geometry of the rotating part of the machine (rotor). The task is to find a rotor design of the synchronous reluctance machine that maximizes the torque for a given impressed current density and a fixed rotor position. We focus on a two dimensional geometrical model of the electrical machine. The mathematical formulation of this problem results in an optimization problem which is constrained by a partial differential equation (PDE) and where the unknown is the design of the rotor geometry. In general the PDE-constraint turns out to be the two dimensional boundary value problem of nonlinear magnetostatics, however we will also consider a simplification to a linear PDE and an extension to a stochastic nonlinear PDE. The latter is used to model uncertainty in the material parameters due to e.g. measurement errors and yields a so called stochastic optimization problem.

In a first step we consider a shape optimization method based on sensitivity information in order to solve all three PDE-constrained design optimization problems. Here, the crucial ingredient is the computation of the so called shape derivative. We derive the shape derivative not only for the deterministic optimization problems but also for the stochastic problem at hand and employ it in suitable algorithms. Finally, we compare our results to the results of a parametric shape optimization approach which is commonly used for design optimization of electrical machines.

In the next step we concentrate on topology optimization methods based on topological sensitivities in order to solve the design optimization problems. In contrast to shape optimization methods which can only vary the boundary or interface of a domain, this class allows for topological changes of the design domain, that is, introduce holes or components. We derive the so called topological derivative of our optimization problem with the help of a recently presented derivation technique based on a Lagrangian framework. We use a well established level set algorithm which employs the topological derivative to solve the deterministic problems. Finally, we present a novel extension of the level set algorithm based on the stochastic gradient method to solve the stochastic optimization problem.

Contents

Introduction

Electrical machines are the principal components of many appliances, industrial equipment and systems. They are the foundation of the power industry and the core of any electrical drive. Electrical machines consume about 46% of the worldwide generated electricity resulting in about 6040 Mega-tonnes of CO_2 emission [30]. Therefore, constructing motors with high energy efficiency is indispensable for energy conservation and environmental sustainability. Besides energy efficiency there are many other performance criteria/requirements imposed on electrical machines such as high torque and power density, high reliability or low noise. So, finding designs such that these machines fulfill these criteria as well as possible has top priority and to achieve this goal design optimization is necessary [30].

In general, a design optimization problem consists of a cost/shape function $J(\Omega, u(\Omega))$ which is a measure for the performance criterion. Here, the shape function depends on $\Omega \subset \mathbb{R}^d$ and on $u(\Omega)$, where the state u satisfies the constraint $E(\Omega, u(\Omega)) = 0$. In the context of electrical machines the constraint E is often a partial differential equation or a system of PDEs describing electromagnetic phenomena, thermal fields, vibration noise or the coupling of these fields. Then, the objective is to minimize (or maximize) the cost function J over some admissible set $\Xi \subset \{\Omega : \Omega \subset \mathbb{R}^d\}$, that is,

$$\text{minimize } J(\Omega, u) \text{ over } \Omega \in \Xi \tag{0.1a}$$
$$\text{subject to } u \in \mathcal{X} \text{ solves } E(\Omega, u) = 0, \tag{0.1b}$$

where $\mathcal{X}$ is a function space. We also refer to this class as deterministic optimization problems. On the contrary in stochastic design optimization problems the constraint E is additionally subject to uncertainty due to e.g. uncertain coefficients, that is $E(\Omega, u(\Omega, \omega), \xi(\omega))$ where ξ is a finite dimensional random vector and $\omega \in X$ an element of the sample space X. The shape function then depends on the uncertainty as well, i.e. $J(\Omega, u(\Omega, \omega), \xi(\omega))$. It is then desirable to determine optimal designs that account for and in some sense are resilient to this uncertainty. A general problem formulation reads

$$\text{minimize } \mathcal{R}(J(\Omega, u, \xi(\omega)) \text{ over } \Omega \in \Xi \tag{0.2a}$$
$$\text{subject to } u = u_\omega \in \mathcal{X} \text{ solves } E(\Omega, u, \xi(\omega)) = 0 \text{ for } \mathbb{P} \text{ a.e. } \omega \in X, \tag{0.2b}$$

where $\mathcal{R}$ is a functional that maps random variables into real numbers and where $\mathbb{P}$ is a probability measure. For example $\mathcal{R}$ could be the expectation operator or any

risk measure. In the context of electrical machines we are faced with both types of problems depending on the modeling and the objectives.

There are several approaches to look on design optimization problems and they differ on the way on how the geometry is allowed to vary. We make the following classification of methods:

1. Parametric optimization: In that case the design Ω is described by a finite number of different parameters m such as lengths, thicknesses, distances, orientations etc. Therefore, the design optimization problem becomes a parametric (shape) optimization problem over the finite dimensional vector space $\mathbb{R}^m$. These problems can be solved by standard gradient based algorithms or by the class of evolutionary algorithms. The latter are less prone to getting stuck in local minima as they do not need any sensitivity information and are widely used in design optimization of electrical machines not only for single-objective problems but also for multi-objective ones. However, they are known to be computational expensive.

2. Shape optimization: In that point of view, we start with an initial design Ω_0 and want to find the optimal design Ω^* by considering smooth perturbations of the boundary of Ω_0. However, we consider the topology of the design domain to be fixed. In order to study the behavior of shape functions under these perturbations shape sensitivity analysis was introduced. This allows to set up gradient based algorithms to solve the optimization problem numerically.

3. Topology optimization: This is the most general class of design optimization methods. Here, we consider not only smooth boundary variations but also topological changes of the design domain. The objective is to find the optimal distribution of different materials within a given design domain.

In this thesis we focus on the latter two approaches. Therefore, we discuss in the following some aspects about shape and topology optimization.

Shape Optimization

One principle of shape sensitivity analysis is to identify shapes with functions to handle the lack of vector space structure of the set of admissible shapes. An important notion is the so called shape derivative, which is for a shape function J and a smooth vector field V defined by

$$dJ(\Omega; V) = \lim_{t \searrow 0} \frac{J(\Omega_t) - J(\Omega)}{t},$$

when the limit exists and $V \mapsto dJ(\Omega; V)$ is continuous and linear. Here, $\Omega_t = T_t(\Omega)$ describes the perturbation of the domain Ω under the flow T_t which is generated by V, see Definition 2.2. The idea of shape sensitivity was first used by Hadamard in his study of elastic plates [20]. The structure theorem says that the shape derivative can

be represented for domains with sufficiently smooth boundary as

$$dJ(\Omega; V) = \int_{\partial\Omega} gV \cdot n \, \mathrm{d}s,$$

where $g \in L^1(\partial\Omega)$ and V a sufficiently smooth vector field, see [41, 11]. Besides this boundary form, the shape derivative can alternatively be written as volume integral, i.e.

$$dJ(\Omega; V) = \int_{\Omega} F(V) \, \mathrm{d}x,$$

where F is an operator acting on V and its derivatives. This formulation requires less regularity of the domain and therefore is more general than the boundary form. In [27] it was pointed out that the volume form has several advantages in terms of numerical accuracy and numerical implementation as well. For this reason we use the volume form in this thesis. Typically, the shape derivative is used to set up gradient based algorithms to solve the shape optimization problems. Here, the basic procedure consists of the following steps:

- Compute the shape derivative.

- Extract a descent direction.

- Move the boundary of the shape along this direction.

Additional challenges arise when considering stochastic (shape) optimization problems. A widely used technique to solve these problems is the stochastic approximation approach which dates back to the paper of Robbins and Monro [36]. The crucial ingredient is the use of a so called stochastic gradient instead of a gradient to iteratively minimize the expected value of a random objective function. Recently, this approach has been combined with the gradient descent method based on the shape derivative to solve stochastic shape optimization problems on shape spaces [17, 18].

Topology optimization

The literature of methods that allow for topological changes of the design domain is very rich and we mention for example the homogenization method [1], On/Off-type methods or density based methods [7]. Perhaps, the latter are the most used in commercial software and a popular representative of them is the SIMP (Solid Isotropic Material with Penalization) method. For an overview of these methods we refer the reader to [39] and the references therein.

Besides the above mentioned approaches, topology optimization methods based on the topological derivative have gained an increasing attention not only from a theoretical but also application point of view. The topological derivative measures the sensitivity of a shape function with respect to a topological perturbation of the shape.

More precisely, the topological derivative of a shape function $J = J(\Omega)$ in a spatial point z is defined as

$$dJ(\Omega)(z) = \lim_{\varepsilon \searrow 0} \frac{J(\Omega_\varepsilon) - J(\Omega)}{|\omega_\varepsilon|},$$

if the limit exists. Here, Ω_ε denotes the topologically perturbed domain resulting from an insertion of a hole ω_ε. This concept was first introduced in [40] in a mathematically rigorous way. Several papers deal with the derivation of the topological derivative for optimization problems constrained by PDEs, see for example [2]. There are different ways to use the topological derivative in a numerical procedure. One way is to combine this derivative with shape optimization algorithms. Here, the typical procedure is to perform repeatedly several iterations of shape optimization and then make a topological perturbation where the topological derivative is most negative. Another approach which solely employs this sensitivity information is the level set algorithm introduced in [4].

In the stochastic case there are several articles which deal with stochastic topology optimization methods, for example see [34, 10] and the references therein. However, none of these methods employ the topological derivative to our knowledge.

Structure of this thesis

In this thesis we focus on solving a design optimization problem of an electrical machine by means of shape and topology optimization methods based on the shape and topological derivative respectively. The thesis is organized as follows:
In Chapter 1 we formulate two deterministic design optimization problems, a nonlinear and a linear one, which are motivated by the physical model of an electrical machine. Further, we investigate its mathematical properties. Chapter 2 deals with shape optimization. In the first part we derive the shape derivative for both deterministic problems and present numerical results. Subsequently, we extend the deterministic model and concentrate on a stochastic design optimization problem. We apply the stochastic shape gradient method [17] to solve the problem and present numerical results. In the final part we compare our results to the results of a parametric optimization approach applied to the nonlinear problem. In Chapter 3 we focus on topology optimization. The first part is concerned with the derivation of the topological derivative for the deterministic problem(s) and the numerical implementation. In particular, we discuss the level set algorithm and a finite element method for unfitted interfaces. We present numerical results for two different initial designs. In the final part a novel approach to solve the stochastic design optimization problem is presented. More precisely, we propose an extension of the level set algorithm based on the stochastic gradient method. We apply this algorithm to the stochastic problem at hand and show numerical results.

1 Problem formulation and analysis

In this chapter we introduce the deterministic model problems used throughout this thesis and investigate its mathematical properties. Therefore we briefly describe the physical background of electrical machines and especially have a closer look on a synchronous reluctance motor (SyRM) since it serves as our model problem. The following sections are mainly based on [13].

1.1 A brief introduction to electrical machines and the physical model

Electrical machines convert electrical and mechanical energy into each other. An electric motor converts electric energy into mechanical energy while a generator does the opposite. Electric motors can be classified into DC motors, which operate on direct current and AC motors, where alternating current is induced. The latter can be distinguished into induction motors and synchronous motors. In this thesis we will focus on a special type of synchronous motors namely synchronous reluctance motors. This motor consists of a fixed part called stator and a rotating part called the rotor. Here, the stator contains coil areas where electric current is induced which generates a magnetic field. Stator and rotor are separated by an air gap. The key component of this motor is that the rotor has no permanent magnet or coil areas such that a torque producing force will only be generated through magnetic reluctance. Compared to other synchronous motors the advantage of these motors lie in the simple construction and the robustness whereas the disadvantage lies in worse performance. For a thorough introduction to electrical machines we refer the reader to [8].

As starting point to obtain a physical model for electric machines serve Maxwell's equations which read

$$\operatorname{curl} \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}, \tag{1.1a}$$

$$\operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{1.1b}$$

$$\operatorname{div} \boldsymbol{B} = 0, \tag{1.1c}$$

$$\operatorname{div} \boldsymbol{D} = \rho, \tag{1.1d}$$

where $\boldsymbol{H}$ denotes the magnetic field intensity, $\boldsymbol{B}$ the magnetic flux density, $\boldsymbol{E}$ the electric field density, $\boldsymbol{D}$ the electric flux density, $\boldsymbol{J}$ the current density and finally ρ the charge density. The boldface letters denote vector valued quantities which depend in general on space and time. This set of equations is complemented by the constitutive equations

$$\boldsymbol{B} = \mu\boldsymbol{H} + \boldsymbol{M}, \tag{1.2a}$$

$$\boldsymbol{D} = \varepsilon\boldsymbol{E} + \boldsymbol{P}, \tag{1.2b}$$

$$\boldsymbol{J} = \boldsymbol{J}_i + \sigma\boldsymbol{E}, \tag{1.2c}$$

where $\boldsymbol{M}$ denotes the magnetization, $\boldsymbol{P}$ an electric polarization and $\boldsymbol{J}_i$ an impressed current density. Furthermore these equations involve the magnetic permeability μ, the electric permittivity ε and the electric conductivity σ. In general these three quantities are second order tensors depending on space, time and the electromagnetic fields. However, in this thesis we assume isotropic material behavior such that these quantities can be considered as scalar functions. Moreover we neglect possible effects of hysteresis. For the mathematical problem we will use a quantity which is closely related to μ, namely the magnetic reluctivity ν which is defined as the reciprocal of μ. It can be seen from (1.2a) that the magnetic reluctivity satisfies the relation

$$\boldsymbol{H} = \nu(\boldsymbol{B} - \boldsymbol{M}). \tag{1.3}$$

In the context of electrical machines we have to deal with ferromagnetic materials. In this case the magnetic reluctivity is described as a nonlinear function depending on the magnitude of the magnetic flux density, $\nu = \nu(|\boldsymbol{B}|)$.

For the simulation of electrical machines it is not necessary to work with the full set of equations in (1.1). A usual simplification is to neglect displacement currents $\frac{\partial \boldsymbol{D}}{\partial t}$ since for low frequency applications hold that (cf. [32, 25])

$$\left|\frac{\partial \boldsymbol{D}}{\partial t}\right| \ll |\boldsymbol{J}|. \tag{1.4}$$

Thus, we arrive at the so called *magnetoquasistatic* problem or *eddy current* problem formulation which reads in vector potential formulation: (cf.[13, 26])

$$\sigma\frac{\partial \boldsymbol{A}}{\partial t} + \operatorname{curl}(\nu(|\operatorname{curl}\boldsymbol{A}|)\operatorname{curl}\boldsymbol{A}) = \boldsymbol{J}_i + \operatorname{curl}(\nu\boldsymbol{M}), \tag{1.5}$$

where $\boldsymbol{A}$ denotes a vector potential which fulfills

$$\boldsymbol{B} = \operatorname{curl}\boldsymbol{A}. \tag{1.6}$$

This model can be used to describe the behavior of an electric motor in its starting phase, this means when the motor is accelerated from a resting position. We are

interested in the phase where the rotor has reached a constant speed. Therefore we can assume that all electromagnetic field quantities are independent of time and conclude that the derivative with respect to the time variable vanishes. As a result we obtain the *magnetostatic* model in 3D for determining the magnetic flux density $\boldsymbol{B} = \operatorname{curl} \boldsymbol{A}$,

$$\operatorname{curl}(\nu(|\operatorname{curl} \boldsymbol{A}|) \operatorname{curl} \boldsymbol{A}) = \boldsymbol{J}_i + \operatorname{curl}(\nu \boldsymbol{M}). \tag{1.7a}$$

Note that the above system of partial differential equations is not yet complete. In order to be able to solve (1.7a) numerically we have to impose suitable boundary and interface conditions. Therefore let us introduce a bounded domain $\hat{D}$ and further split the boundary of $\partial \hat{D}$ into two parts Γ_B and Γ_H such that $\hat{D} = \overline{\Gamma}_B \cup \overline{\Gamma}_H$ and $\Gamma_B \cap \Gamma_H = \emptyset$. One possible choice is to consider the boundary conditions

$$\boldsymbol{B} \cdot n = 0 \text{ on } \Gamma_B, \tag{1.7b}$$

$$\boldsymbol{H} \times n = 0 \text{ on } \Gamma_H, \tag{1.7c}$$

where n denotes the outer unit normal vector to $\hat{D}$. The condition (1.7b) is called induction boundary condition and means that no magnetic flux can leave the computational domain. The Perfect Magnetic Conductor (PMC) boundary condition (1.7c) models materials with very high permeability, cf.[44]. We now focus on suitable interface conditions on an interface Γ_I where the magnetic reluctivity ν jumps. We denote by $[\![v]\!]$ the jump of a function v on the interface Γ_I, i.e.

$$[\![v]\!] = v^+\big|_{\Gamma_I} - v^-\big|_{\Gamma_I},$$

where v^+ and v^- denote the restrictions of v to the corresponding subdomains. Then the interface conditions read

$$[\![\boldsymbol{B} \cdot n]\!] = 0 \text{ on } \Gamma_I, \tag{1.7d}$$

$$[\![\boldsymbol{H} \times n]\!] = 0 \text{ on } \Gamma_I. \tag{1.7e}$$

Remark 1.1. Note that a solution to (1.5) as well as (1.7) is only unique up to a gradient field $\nabla \phi$. This issue can be fixed by imposing additionally e.g. that div $\boldsymbol{A} = 0$. This choice is referred to as Coloumb gauge.

The model (1.7) can be reduced into a 2D setting if certain assumptions on the geometry of the computational domain $\hat{D}$ and the fields are satisfied. The assumptions are that [13]

- one space dimension is much larger compared to the other, i.e.

$$\hat{D} = D \times (-\ell, \ell) \text{ with } \ell \gg \operatorname{diam}(D),$$

- the fields $\boldsymbol{J}$, $\boldsymbol{H}$, $\boldsymbol{M}$ are of the form

$$\boldsymbol{J}_i = \begin{pmatrix} 0 \\ 0 \\ J_3(x_1, x_2) \end{pmatrix}, \boldsymbol{H} = \begin{pmatrix} H_1(x_1, x_2) \\ H_2(x_1, x_2) \\ 0 \end{pmatrix}, \boldsymbol{M} = \begin{pmatrix} M_1(x_1, x_2) \\ M_2(x_1, x_2) \\ 0 \end{pmatrix}$$

for $(x_1, x_2) \in D$.

Together with the ansatz

$$\boldsymbol{A} = \begin{pmatrix} 0 \\ 0 \\ u(x_1, x_2) \end{pmatrix}, (x_1, x_2) \in D, \tag{1.8}$$

we obtain the boundary value problem for two dimensional magnetostatics which reads [13]

$$-\operatorname{div}(\nu \nabla u) = J_3 - \nu \operatorname{div} M^\perp, \ \text{in } D \subset \mathbb{R}^2, \tag{1.9a}$$

$$u = 0, \qquad \text{on } \Gamma_B, \tag{1.9b}$$

$$\nu(x, |\nabla u|) \nabla u \cdot n = 0, \qquad \text{on } \Gamma_H, \tag{1.9c}$$

$$[\![u]\!] = 0, \qquad \text{on } \Gamma_I, \tag{1.9d}$$

$$[\![\nu(x, |\nabla u|) \nabla u \cdot n]\!] = 0, \qquad \text{on } \Gamma_I \tag{1.9e}$$

where $M^\perp = (-M_2, M_1)^T$ and Γ_B and Γ_H denote the part of the boundary of ∂D where we impose $\boldsymbol{B} \cdot n = 0$ and $\boldsymbol{H} \times n = 0$ respectively. Note that with the ansatz in (1.8) the Coloumb gauge condition $\operatorname{div} \boldsymbol{A} = 0$ is satisfied as well. More details on the derivation for the reduction into a 2D model can be found in [32]. For the simulation of the magnetic flux density in a synchronous reluctance motor we can use the model (1.9) in the following regime: Firstly, since the rotor does not contain any permanent magnets the magnetization term $M^\perp$ can be omitted. Secondly, it is common practice to assume homogenous Dirichlet boundary conditions on the entire boundary ∂D. This implies $\boldsymbol{B} \cdot n\big|_{\partial D} = 0$ which means that no magnetic flux leaves the domain. Summarizing we obtain the following model:

$$-\operatorname{div}(\nu \nabla u) = J_3, \ \text{in } D \subset \mathbb{R}^2, \tag{1.10a}$$

$$u = 0, \quad \text{on } \partial D, \tag{1.10b}$$

$$[\![u]\!] = 0, \quad \text{on } \Gamma_I, \tag{1.10c}$$

$$[\![\nu(x, |\nabla u|) \nabla u \cdot n]\!] = 0, \quad \text{on } \Gamma_I \tag{1.10d}$$

1.2 Model problem

In this section we describe the mathematical problem used throughout this thesis in detail. We follow the lines in [13]. As already mentioned we consider a synchronous reluctance motor as a model problem. We are interested in the magnetic flux density at a fixed rotor position when the rotor has reached a constant rotational speed. Furthermore the geometry of the motor fulfills the assumptions for the reduction into the 2D setting very well so that we can work with the model in (1.10). In the following we will mathematically formulate the problem of maximizing the torque of the motor
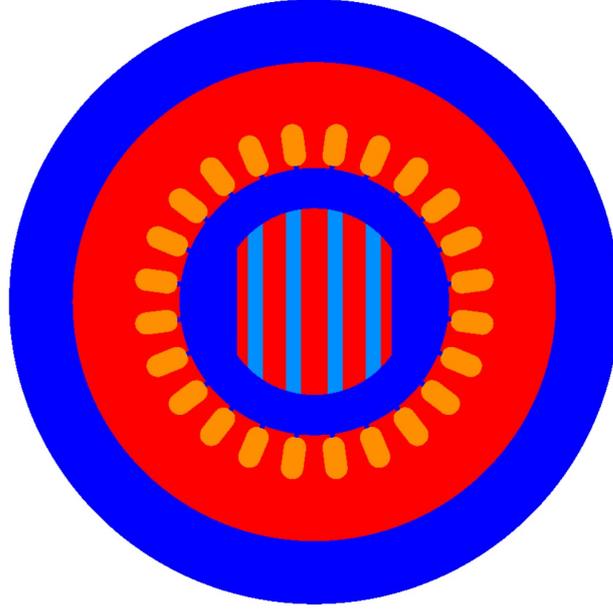
Figure 1.1: Computational domain indicating different subdomains. Red: ferromagnetic subdomain Ω_f^{ref}. Orange: coil area Ω_c. Blue: air. Light blue: non-ferromagnetic layers of the rotor.

with respect to the design of the rotor when the rotor reached a constant rotational speed.

Let the hold all domain $D \subset \mathbb{R}^2$ denote the two dimensional computational domain of the motor which is depicted in Figure 1.1. The domain contains all parts of the motor including an axially-layered rotor, stator, coil areas and an air gap between rotor and stator. We denote by Ω_f^{ref} the ferromagnetic reference subdomain of the motor and by $\Omega_{air}^{ref} = D \setminus \overline{\Omega_f^{ref}}$ its complement, see Figure 1.1. The subdomain Ω_{air}^{ref} consists of all areas which are not ferromagnetic. These include the air gap between rotor and stator, some layers of the rotor, the coil areas which we denote by Ω_c as well as a thin layer of air outside the stator. Furthermore we denote by Ω^d the design domain which is the whole rotor. We want to find the optimal distribution of ferromagnetic material and air regions inside Ω^d such that the torque gets maximized. For this reason we denote the subdomain of Ω^d which is currently occupied with ferromagnetic material with Ω [13]. Hence, the distribution of ferromagnetic material in the motor for any Ω is given by

$$\Omega_f := (\Omega_f^{ref} \setminus \Omega^d) \cup \Omega. \tag{1.11}$$

As before we define $\Omega_{air} = D \setminus \overline{\Omega_f}$. The reluctivity function ν in (1.10) attains different

values in the different subdomains of the motor. To be more specific, we have

$$\nu(x, |\nabla u|) = \begin{cases} \hat{\nu}(|\nabla u|) & x \in \Omega_f, \\ \nu_0 & x \in \Omega_{air}, \end{cases} \tag{1.12}$$

where $\hat{\nu}$ is a nonlinear function given by the B-H relation which will be discussed in the next section. The constant

$$\nu_0 = 10^7/(4\pi) \tag{1.13}$$

describes the magnetic reluctivity of vacuum which is practically the same as that of air. Note that also the coil areas consisting of copper and the non ferromagnetic layers of the rotor have the same magnetic reluctivity as of air. Since the reluctivity in the electric motor depends on the current design Ω as well, we have $\nu = \nu_\Omega$ and write

$$\nu_\Omega(x, |\nabla u|) = \hat{\nu}(|\nabla u|)\chi_{\Omega_f}(x) + \nu_0\chi_{\Omega_{air}}(x) \tag{1.14}$$

where χ_U denotes the characteristic function on the set U. In order to get a variational formulation of (1.10) we define the nonlinear operator $A_\Omega : H_0^1(D) \to H^{-1}(D)$ as

$$\langle A_\Omega(u), \eta \rangle = \int_D \nu_\Omega(x, |\nabla u|)\nabla u \cdot \nabla \eta \, \mathrm{d}x, \tag{1.15}$$

for all $u, \eta \in H_0^1(D)$. The weak formulation of the right hand side of (1.10a) reads

$$\langle F, \eta \rangle = \int_D J_3 \eta \, \mathrm{d}x, \tag{1.16}$$

for all $\eta \in H_0^1(D)$, where $\langle \cdot, \cdot \rangle$ denotes the duality product between $H^{-1}(D)$ and $H_0^1(D)$. Note that we consider a fixed rotor position and therefore the impressed current density J_3 is a piecewise constant function in the coil areas Ω_c and vanishes outside Ω_c. Further note that the interface conditions (1.10c) and (1.10d) are already included in the weak formulation of the problem. Finally, the variational formulation of model (1.10) is stated as follows: Find $u \in H_0^1(D)$ such that

$$\langle A_\Omega(u), \eta \rangle = \langle F, \eta \rangle \text{ for all } \eta \in H_0^1(D). \tag{1.17}$$

In our optimization problem we want to maximize a domain dependent functional representing the current torque for a fixed rotor position of the electric motor. There are several methods used for the calculation of the torque in the literature, for instance the eggshell method [23], the Maxwell stress tensor, Coenergy derivation or Arkkio's method. For an overview and a comparison between the latter three methods we refer the reader to [37]. We focus on Arkkio's method which is a variant of the Maxwell stress tensor. Here, the expression for the torque reads

$$T = \frac{L}{\mu_0(r_s - r_r)} \int_{\Omega_g} r B_r B_\phi \mathrm{d}S \tag{1.18}$$

where Ω_g denotes the whole volume in the air gap comprised between the layers of radii r_r and r_s, B_r and B_ϕ the radial and tangential component of the magnetic flux density in the air gap, respectively, and L the length of the rotor. Introducing the unit vectors pointing in radial and tangential direction, that is

$$n_g(x) := \frac{1}{|x|} \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \quad \text{and} \quad \tau_g(x) := \frac{1}{|x|} \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix} \quad \text{for } x \in \Omega_g, \qquad (1.19)$$

we can define B_r and B_ϕ as

$$B_r(x) := \boldsymbol{B}(x) \cdot n_g(x) \quad \text{and} \quad B_\phi(x) := \boldsymbol{B}(x) \cdot \tau_g(x) \quad \text{for } x \in \Omega_g. \qquad (1.20)$$

Noting that $\boldsymbol{B} = \operatorname{curl}((0,0,u)^T) = (\partial_2 u, -\partial_1 u, 0)^T$ and $r = |x|$ the equation for the torque (1.18) can be written as

$$T(u) = \frac{L}{\mu_0(r_s - r_r)} \int_{\Omega_g} \nabla u^T Q(x) \nabla u \, \mathrm{d}x \qquad (1.21)$$

with the symmetric matrix

$$Q(x) := \frac{1}{\sqrt{x_1^2 + x_2^2}} \begin{pmatrix} x_1 x_2 & \frac{x_2^2 - x_1^2}{2} \\ \frac{x_2^2 - x_1^2}{2} & -x_1 x_2 \end{pmatrix}. \qquad (1.22)$$

Observe that the functional in (1.21) only depends on the current shape $\Omega \subset \Omega^d$ via the solution u of the state equation, i.e. $T(u) = T(u(\Omega))$. Since we are interested in a high torque, we minimize the objective function

$$\mathcal{J}(u) := -T(u) \qquad (1.23)$$

with T defined by (1.21).

Summarizing we are faced with the PDE-constrained optimization problem

$$\inf_{\Omega \in \mathcal{O}} \mathcal{J}(u) \qquad (1.24\text{a})$$

$$\text{subject to } u \in H_0^1(D) : \langle A_\Omega(u), \eta \rangle = \langle F, \eta \rangle \text{ for all } \eta \in H_0^1(D), \qquad (1.24\text{b})$$

where $\mathcal{J}$ is given by (1.23), A_Ω and F are defined by (1.15) and (1.16) respectively. Here, $\mathcal{O}$ denotes the set of admissible shapes which we define as

$$\mathcal{O} = \{\Omega \subset \Omega^d : \Omega \text{ open, Lipschitz with uniform Lipschitz constant } L_\mathcal{O}\}. \qquad (1.24\text{c})$$

In our simulations we will also investigate a linear version of problem (1.24), that is, we want to minimize $\mathcal{J}(u)$ subject to a linear version of the magnetostatics boundary value problem. In this case we assume only linear behavior of the materials. Therefore

the magnetic reluctivity simplifies to a function which only depends on the spatial variable, more precisely

$$\nu_\Omega(x) = \nu_1 \chi_{\Omega_f}(x) + \nu_0 \chi_{\Omega_{air}}(x),$$

where $\nu_0 > \nu_1 > 0$ denotes a constant magnetic reluctivity for the ferromagnetic material. Introducing the bilinearform $a : H_0^1(D) \times H_0^1(D) \to \mathbb{R}$ defined by

$$a_\Omega(u, \eta) = \int_D \nu_\Omega(x) \nabla u \cdot \nabla \eta \, \mathrm{d}x \tag{1.25}$$

we can state the following optimization problem:

$$\inf_{\Omega \in \mathcal{O}} \mathcal{J}(u) \tag{1.26a}$$

$$\text{subject to } u \in H_0^1(D) : a_\Omega(u, \eta) = \langle F, \eta \rangle \text{ for all } \eta \in H_0^1(D), \tag{1.26b}$$

where the quantities $\mathcal{J}$, F and $\mathcal{O}$ are given as in (1.24). Although the problem above is less realistic than problem (1.24), a comparison in the context of design optimization might be interesting. In Chapter 2 we will extend problem (1.24) and tackle a stochastic PDE-constrained optimization problem.

1.3 Physical properties of B-H-curves

The material influence in our models appear in form of the magnetic reluctivity which is determined through a B-H-curve obtained from physical measurements for a specific material. In this section we will focus on the relation between the magnetic field intensity $\boldsymbol{H}$ and the magnetic flux density $\boldsymbol{B}$. As a result we obtain some regularity for $\hat{\nu}$ which will be the key ingredient for the analysis. Recall that we only consider isotropic materials such that these two fields are parallel and neglect possible effects of hysteresis.

The magnitude $B := |\boldsymbol{B}|$ of the magnetic flux density $\boldsymbol{B}$ depends on the magnitude $H := |\boldsymbol{H}|$ of the magnetic field intensity and on the properties of the material [33]. For materials like vacuum or air this relation is linear and as a result we have that $B = \mu H$ with a constant magnetic permeability μ. However, ferromagnetic materials respond strongly to a magnetic field and the relation between B and H is a nonlinear one described by the so called B-H curve

$$f : \mathbb{R}_0^+ \to \mathbb{R}_0^+, \ H \mapsto B := f(H),$$

where $\mathbb{R}_0^+$ represents the set of non-negative real numbers. Based on this notion one defines the magnetic permeability μ and the magnetic reluctivity introduced in (1.2a),(1.3) as

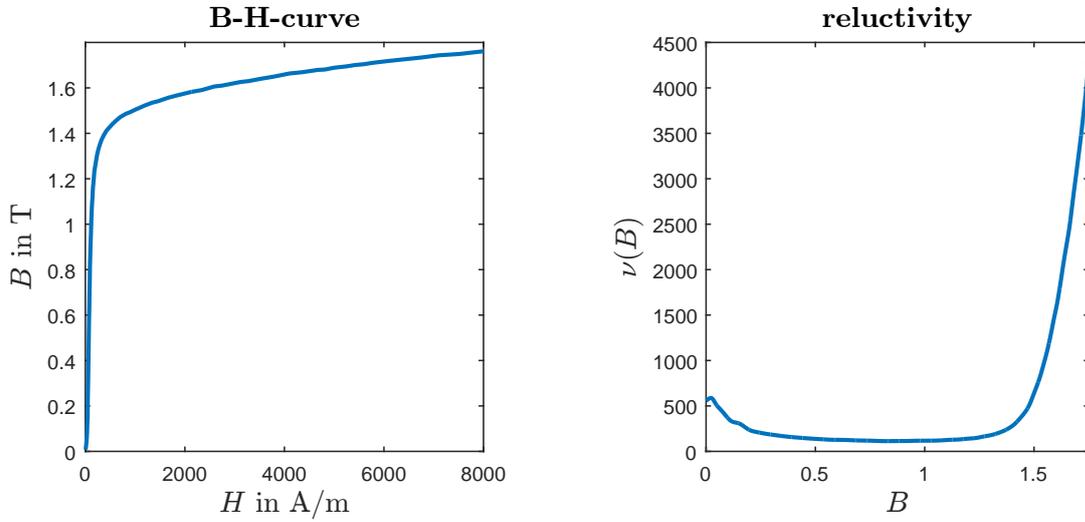$$\mu(s) := f(s)/s \quad \text{and} \quad \nu(s) := f^{-1}(s)/s \tag{1.27}$$

Figure 1.2: B-H-curve f and magnetic reluctivity ν of the ferromagnetic material used in our simulations of the SyRM.

such that we obtain

$$\boldsymbol{B} = \mu(|\boldsymbol{H}|)\boldsymbol{H} \quad \text{and} \quad \boldsymbol{H} = \nu(|\boldsymbol{B}|)\boldsymbol{B} \tag{1.28}$$

The B-H-curve used in our simulations for the model problem (1.24) is shown in Figure 1.2. It shows a typical B-H-curve of a ferromagnetic material which is characterized by a high amplification of B for small values of H and an amplification of B close to that of vacuum, for high values of H. These natural physical properties lead to the following Assumptions on a B-H-curve f (cf. [32]).

Assumption 1.1. Let $f : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ be a B-H-curve. Then f fulfills the following conditions :

1. f is continuously differentiable on $\mathbb{R}_0^+$,

2. $f(0) = 0$,

3. $f'(s) \geq \mu_0$ for all $s \geq 0$,

4. $\lim_{s \to \infty} f'(s) = \mu_0$, where μ_0 is the permeability of vacuum.

Note that μ_0 is the reciprocal of the magnetic reluctivity in vacuum introduced in (1.13), i.e. $\mu_0 = 1/\nu_0$. The following Lemma is an immediate consequence of Assumption 1.1.

Lemma 1.1 ([13])**.** *Let Assumption 1.1 be satisfied. Then the following statements hold:*

1. $\hat{\nu}$ is continuously differentiable on $(0, \infty)$

2. There exists $\underline{\nu} > 0$ such that for all $s \in \mathbb{R}_0^+$ we have

$$\underline{\nu} \leq \hat{\nu}(s) \leq \nu_0 \tag{1.29}$$

$$\underline{\nu} \leq (\hat{\nu}(s)s)' \leq \nu_0 \tag{1.30}$$

3. The mapping $s \mapsto \hat{\nu}(s)s$ is strongly monotone with monotonicity constant $\underline{\nu}$, i.e.,

$$(\hat{\nu}(s)s - \hat{\nu}(t)t)(s - t) \geq \underline{\nu}(s - t)^2, \quad \forall s, t \in \mathbb{R}_0^+, \tag{1.31}$$

and Lipschitz continuous with Lipschitz constant ν_0, i.e.,

$$|\hat{\nu}(s)s - \hat{\nu}(t)t| \leq \nu_0 |s - t| \quad \forall s, t \in \mathbb{R}_0^+. \tag{1.32}$$

Proof. A proof of statement 1. and 2. can be found in [32]. The last property easily follows from (1.30) together with the mean value theorem. $\qquad \square$

The properties (1.31) and (1.32) play an important role for the well-posedness of problem (1.17) as we will see in the next section.

Remark 1.2. Note that the derivative in (1.30) is strictly positive. This is an important condition in order to be able to apply Newton's method to solve the nonlinear state equation (1.17) numerically. More precisely, it guarantees the well-posedness of an boundary value problem that represents a linearization to the state equation, see Section 2.4.1.

1.4 Analysis of the state equation

We are going to show that the problem stated in (1.17) has a unique solution. In order to do so we will use the following theorem.

Theorem 1.2 (Zarantonello). *Let X be a real Hilbert space and $A : X \to X^\star$ be a nonlinear operator satisfying the following conditions:*

1. *A is strongly monotone, i.e there exists a constant $c_1 > 0$ such that for all $u, v \in X$,*

$$\langle A(u) - A(v), u - v \rangle \geq c \|u - v\|^2 \tag{1.33}$$

2. *A is Lipschitz continous, i.e there exists a constant c_2 such that for all $u, v \in X$,*

$$\|A(u) - A(v)\| \leq c_2 \|u - v\|. \tag{1.34}$$

Then, for each $b \in X^\star$ the operator equation $A(u) = b$, $u \in X$ has a unique solution which depends continuously on b.

In order to get well-posedness of the boundary value problem (1.17), we have to show that the operator A_Ω defined in (1.15) is strongly monotone and Lipschitz continuous. The following Lemma guarantees these two conditions under additional requirements.

Lemma 1.3 ([13, 32])**.** *Let the mapping $s \mapsto \nu_\Omega(s)s$ be strongly monotone and Lipschitz continuous from $\mathbb{R}_0^+ \to \mathbb{R}_0^+$ with monotonicity constant $\underline{\nu}$ and Lipschitz constant ν_0. Then the nonlinear operator A_Ω defined in (1.15) is strongly monotone with monotonicity constant $\underline{\nu}$ and Lipschitz continuous with Lipschitz constant $3\nu_0$.*

Now, we formulate the main result of this section.

Theorem 1.4. *Let $\hat{\nu}$ be a magnetic reluctivity according to a B-H curve f satisfying Assumption 1.1. Then the boundary value problem defined in (1.17), (1.15), (1.16) has a unique solution $u \in H_0^1(D)$ and there exists a constant $c > 0$ such that*

$$\|u\|_{H_0^1(D)} \leq c\|F\|_{H^{-1}(D)} \tag{1.35}$$

Proof. Thanks to the physical properties of a B-H curve, that is Assumption 1.1, we have that the mapping $s \mapsto \hat{\nu}(s)s$ is strongly monotone and Lipschitz continuous on the ferromagnetic subdomain Ω_f according to Lemma 1.1. Clearly, these conditions are satisfied when $\hat{\nu}$ is replaced by a constant reluctivity function, $\tilde{\nu}(s) = \nu_0$, as well. This means that the mapping $s \mapsto \nu_\Omega(x, s)s$ with the global magnetic reluctivity ν_Ω defined in (1.14) is strongly monotone and Lipschitz continuous. Applying Lemma 1.3 we get the strong monotonicity and the Lipschitz continuity of A_Ω. It remains to check that $F \in H^{-1}(D) = H_0^1(D)^\star$. Using Cauchy's inequality we get

$$|\langle F, \eta \rangle| \leq \|J_3\|_{L^2(D)}\|\eta\|_{L^2(D)} \leq \|J_3\|_{L^2(D)}\|\eta\|_{H^1(D)}.$$

Applying Theorem 1.2 yields to the above statement. $\qquad\square$

Finally, we want to mention that also the linear boundary value problem in (1.26b) has a unique solution $u \in H_0^1(D)$ according to the Lemma of Lax-Milgram.

2 Shape optimization

In this chapter we apply shape optimization methods based on the shape derivative to solve the design optimization problems stated in the previous chapter. The shape derivative represents the sensitivity of a shape function $J(\Omega)$ with respect to a variation of the domain. There are several methods to derive shape differentiability when the cost function is constrained by a PDE, for instance the material derivative method [41], the minimax formulation [11] or Céa's Lagrange method [9]. We will make use of an other approach, namely the so-called averaged adjoint method [42, 43] which is advantageous when dealing with a nonlinear constraint. With the help of the shape derivative we can compute deformation fields which we will use in a gradient based optimization algorithm.

This chapter is organized as follows: In Section 2.1 we introduce some basic definitions and results in shape optimization. We show existence of a solution to both model problems in Section 2.2. Based on the results in [13], we derive in Section 2.3 the shape derivative via the averaged adjoint method for our problem. Furthermore, we use this derivative to implement a numerical optimization algorithm and present the obtained results in Section 2.4. In the final section (Section 2.5) we extend model problem (1.24) to account for uncertain material parameters. Therefore we assume that the magnetic reluctivity is subject to uncertainty due to measurement errors. The resulting stochastic model is solved via a novel stochastic shape gradient method proposed in [17] and the optimization results are presented.

2.1 Basic definitions and results

In this section we recall some definitions together with an appropriate notion of differentiability in shape optimization. Unless stated otherwise we consider an open bounded and fixed subset $D \subset \mathbb{R}^d$ with Lipschitz boundary ∂D.

Definition 2.1 ([11]). Let $D \subset \mathbb{R}^d$ be a set and denote by $\mathcal{P}(D) = \{\Omega : \Omega \subset D\}$ the set of subsets of D. A shape function is a map

$$J : \Xi \to \mathbb{R}, \quad \Omega \mapsto J(\Omega)$$

for some family of admissible sets $\Xi \subset \mathcal{P}(D)$.

A typical shape optimization problem has the form

$$\min J(\Omega) \quad \text{over } \Omega \subset \Xi.$$

Here, the shape function J might implicitly depend on the solution of a PDE. In that case we call the shape optimization problem PDE-constrained.

Remark 2.1. The model problem (1.24) can be considered in this framework too. Introducing the reduced functional $J(\Omega) := \mathcal{J}(u_\Omega)$ where u_Ω is the unique solution to the underlying PDE yields to the above setting.

The shape derivative should give us information about the sensitivity of a shape function with respect to a perturbation of the domain Ω. However it is a priori not clear how a perturbation of the domain has to be understood since the set of shapes Ξ does not have a vector space structure. Throughout this thesis we focus on the so called velocity (or speed) method to describe domain perturbations. In the velocity method a reference domain $\Omega \subset D$ is transformed under the action of a velocity field V on D into a new domain. This transformation is defined in the following way:

Definition 2.2. Assume that the velocity field $V : [0, \tau] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies

$$\forall x \in \mathbb{R}^d, V(\cdot, x) \in C([0, \tau]; \mathbb{R}^d), \tag{2.1}$$

$$\exists c > 0, \forall x, y \in \mathbb{R}^d, \|V(\cdot, y) - V(\cdot, x)\|_{C([0,\tau];\mathbb{R}^d)} \leq c \, |y - x| . \tag{2.2}$$

The transformation (also called flow) $T_t : \mathbb{R}^d \to \mathbb{R}^d$ associated to the velocity field V is defined for each $x_0 \in \mathbb{R}^d$ as $T_t(x_0) := x(t)$ where $x : [0, \tau] \to \mathbb{R}^d$ solves the differential equation

$$\dot{x}(t) = V(t, x(t)), \quad \text{in } (0, \tau), \quad x(0) = x_0. \tag{2.3}$$

The flow T_t transforms the domain Ω into $\Omega_t := T_t(\Omega) = \{T_t(x_0) : x_0 \in \Omega\}$. Note that the conditions (2.1)-(2.2) guarantee that the ordinary differential equation (2.3) has a unique solution. In the case that the velocity field V is autonomous, i.e $V = V(x) \neq V(t, x)$ the assumptions reduce to $V \in C^{0,1}(\mathbb{R}^d, \mathbb{R}^d)$. It is worth to mention that the flow T_t generates a homeomorphism from $\overline{D}$ to itself if the vector field is tangential to the boundary ∂D [43]. Moreover, we have that interior (boundary) points of D are mapped onto interior (boundary) points. These facts are due to Nagumo's theorem and are summarized in the following Lemma for autonomous velocity fields but they also hold in the time-dependent case (cf. [43]).

Lemma 2.1. *Let $D \subset \mathbb{R}^d$ be a bounded C^k domain, $k \geq 1$ and $\tau > 0$. Suppose that $V \in C^{0,1}(\mathbb{R}^d, \mathbb{R}^d)$ is a vector field satisfying*

$$V(x) \cdot n(x) = 0 \quad \forall x \in \partial D.$$

Then for the flow T_t generated by V according to Definition 2.2 follows

$$T_t(\Omega) = \Omega, \quad \text{and} \quad T_t(\partial\Omega) = \partial\Omega \quad \forall t \geq 0.$$

and

$$\forall t \in [0, \tau] : x \mapsto T_t(x) \in Hom(\overline{D}).$$

From now on we consider only autonomous vector fields. In the following let $J : \Xi(D) \to \mathbb{R}$ be a shape function defined on an admissible set $\Xi(D) \subset \mathcal{P}(D)$.

Definition 2.3 (Eulerian semiderivative). Let $V \in C_c^k(D, \mathbb{R}^d)$ for $k \geq 1$ and $\Omega_t := T_t(\Omega), \Omega \in \Xi(D)$ according to Definition 2.2. The Eulerian semiderivative of the shape function J at Ω in the direction V is defined as

$$dJ(\Omega; V) := \lim_{t \searrow 0} \frac{J(\Omega_t) - J(\Omega)}{t} \tag{2.4}$$

when the limit exist and is finite.

Note that here we suppose that $V \in C_c^k(D, \mathbb{R}^d)$ which ensures additionally to the results in Lemma 2.1 that the flow T_t is a diffeomorphism. This guarantees that $\Omega_t \in \Xi$ is again an element of the admissible set as diffeomorphisms preserve the topology of the underlying set. Also note that V is extended by zero to $\mathbb{R}^d$ to be in the setting of Definition 2.2.

Definition 2.4. The shape function J is shape differentiable at Ω if the Eulerian semiderivative at Ω exists for all $V \in C_c^k(D, \mathbb{R}^d), k \geq 1$ and the mapping

$$V \mapsto dJ(\Omega; V)$$

is linear and continuous from $C_c^k(D, \mathbb{R}^d)$ to $\mathbb{R}$.

Finally we have a closer look on the structure of the shape derivative. The following fundamental result of shape optimization states that when $\partial\Omega$ and V are smooth enough the shape derivative can be represented as an integral over $\partial\Omega$ and only depends on the normal component of V.

Theorem 2.2 (Structure Theorem of Hadamard-Zolésio). *Let J be a shape functional. Assume that J is shape differentiable at $\Omega \in \Xi(D)$ and that $\partial\Omega$ is compact and of class $C^{k+1}, k \geq 0$. Then there exists a unique outward unit normal vector field $n \in C^k(\partial\Omega, \mathbb{R}^d)$ and a scalar distribution $g \in C^k(\partial\Omega)^*$ such that for all $V \in C_c^k(D, \mathbb{R}^d)$*

$$dJ(\Omega; V) = \langle g, \gamma_{\partial\Omega}(V) \cdot n \rangle_{C^k(\partial\Omega)^* \times C^k(\partial\Omega)} \tag{2.5}$$

where $\gamma_{\partial\Omega} : C_c^k(D, \mathbb{R}^d) \to C^k(\partial\Omega, \mathbb{R}^d)$. When $g \in L^1(\partial\Omega)$ we may write

$$dJ(\Omega; V) = \int_{\partial\Omega} gV \cdot n \, \mathrm{d}s. \tag{2.6}$$

Proof. A proof can be found in [11]. $\qquad\square$

2.2 Existence of optimal shapes

In this section we show that our model problems (1.24) and (1.26) have a solution $\Omega^\star \in \mathcal{O}$. Existence to shape optimization problems can often be shown by the direct method of calculus of variations. To apply this method it is necessary to have a suitable topology on the set of admissible shapes which also provide a form of compactness. So let us recall the notion of convergence in the sense of characteristic functions and a compactness result.

Definition 2.5. Let D be an open and bounded set and denote by $\Xi(D)$ the set of measurable subsets of D. We say a sequence Ω_n in $\Xi(D)$ converges to $\Omega \in \Xi(D)$ in the sense of characteristic functions if χ_{Ω_n} converges to χ_Ω in $L^1(D)$.

We make use of the following compactness result.

Theorem 2.3 ([22, Theorem 2.4.10]). *Let Ω_n be a sequence in $\mathcal{O}$. Then there exists $\Omega^\star \in \mathcal{O}$ and a subsequence Ω_{n_k} that converges to $\Omega^\star$ in the sense of characteristic functions.*

2.2.1 The linear case

We use the direct method of calculus of variations to show existence of a solution to problem (1.26). For this reason we need the following continuity result:

Theorem 2.4. *Let Ω_n be a sequence in $\mathcal{O}$ converging in the sense of characteristic functions to $\Omega \in \mathcal{O}$. Denote by u_n and u the solution to (1.26b) with Ω_n and Ω respectively. Then the sequence $u_n \in H_0^1(D)$ converges to u strongly in $H_0^1(D)$*

Proof. Let $\chi_n \to \chi$ in $L^1(D)$ and denote by $\chi_{f,n} := \chi_{\Omega_f^{ref}\setminus\Omega^d} + \chi_n$ the ferromagnetic subdomain for every n. From the Lax-Milgram theorem we have that

$$\|u_n\|_{H_0^1(D)} \le c_1^{-1}\|F\|_{H^{-1}(D)},$$

where c_1 denotes the ellipticity constant. Thus, the sequence u_n is bounded in $H_0^1(D)$ and we can extract a subsequence u_{n_k} which converges to some $\tilde{u} \in H_0^1(D)$. Note that $\|\nu_1\chi_{f,n_k} + \nu_0(1 - \chi_{f,n_k})\|_{L^\infty(D)} \le \max(\nu_1,\nu_0)$ and $(\nu_1\chi_{f,n_k} + \nu_0(1 - \chi_{f,n_k})) \to \nu_1\chi_f + \nu_0(1 - \chi_f)$ in $L^2(D)$ since $\chi_{f,n_k} \to \chi_f$ in $L^2(D)$. Furthermore $\nabla u_{n_k} \rightharpoonup \nabla\tilde{u}$ in $L^2(D)$. Applying [46, Lemma 4.21] we obtain for all $\eta \in H_0^1(D)$

$$\int_D (\nu_1\chi_{f,n_k} + \nu_0(1 - \chi_{f,n_k}))\nabla u_{n_k} \cdot \nabla\eta \mathrm{d}x \to \int_D (\nu_1\chi_f + \nu_0(1 - \chi_f))\nabla\tilde{u} \cdot \nabla\eta \mathrm{d}x.$$

Thus, $\tilde{u} \in H_0^1(D)$ is the solution to

$$\int_D (\nu_1\chi_f + \nu_0(1 - \chi_f))\nabla\tilde{u} \cdot \nabla\eta \mathrm{d}x = \langle F, \eta\rangle, \quad \forall\eta \in H_0^1(D).$$

Since u is defined as the solution to that same problem and since this solution is unique due to Lax-Milgram, it follows that $\tilde{u} = u$. By the above argument we have that every weakly converging subsequence of u_n converges to the same limit u. Together with the boundedness of u_n in $H_0^1(D)$ we get that the whole sequence u_n converges weakly to u. Now, we are in position to show strong convergence:

$$
\begin{aligned}
\min(\nu_1, \nu_0)\|\nabla(u_n - u)\|_{L^2(D)}^2 &\leq \int_D (\nu_1 \chi_{f,n} + \nu_0(1 - \chi_{f,n}))\nabla(u_n - u) \cdot \nabla(u_n - u)\mathrm{d}x \\
&= \int_D (\nu_1 \chi_{f,n} + \nu_0(1 - \chi_{f,n}))\nabla u_n \cdot \nabla u_n \mathrm{d}x \\
&\quad - 2 \int_D (\nu_1 \chi_{f,n} + \nu_0(1 - \chi_{f,n}))\nabla u_n \cdot \nabla u \mathrm{d}x \\
&\quad + \int_D (\nu_1 \chi_{f,n} + \nu_0(1 - \chi_{f,n}))\nabla u \cdot \nabla u \mathrm{d}x \\
&= \langle F, u_n \rangle - 2\langle F, u \rangle + \int_D (\nu_1 \chi_{f,n} + \nu_0(1 - \chi_{f,n}))\nabla u \cdot \nabla u \mathrm{d}x \\
&\to -\langle F, u \rangle + \int_D (\nu_1 \chi_f + \nu_0(1 - \chi_f))\nabla u \cdot \nabla u \mathrm{d}x = 0.
\end{aligned}
$$

In the last step above we used the weak convergence of u_n in $H_0^1(D)$. Hence, we have shown that $u_n \to u$ in $H_0^1(D)$. □

Now, take a minimizing sequence $\Omega_n \in \mathcal{O}$ for problem (1.26), which exists since $m := \inf_\Omega \mathcal{J}(u_\Omega) \geq 0$. According to Theorem 2.3 there exists a subsequence which we still denote by Ω_n which converges to some $\Omega^\star \in \mathcal{O}$ in the sense of characteristic functions. We denote by u_n and $u^\star$ the solutions to (1.26b) with Ω_n and $\Omega^\star$ respectively. Then, we have

$$
m = \inf_\Omega J(\Omega) = \lim_{n \to \infty} J(\Omega_n) = \lim_{n \to \infty} \mathcal{J}(u_n) = \mathcal{J}(u^\star) = J(\Omega^\star),
$$

where we used the continuity result of Theorem 2.4 and the continuity of $\mathcal{J}$.

2.2.2 The nonlinear case

We show existence of a solution to problem (1.24). Set $m := \inf_\Omega \mathcal{J}(u_\Omega)$. Since $\mathcal{J} \geq 0$ we have that $m \geq 0$. Now, let $\Omega_n \in \mathcal{O}$ be a minimizing sequence. Then, according to Theorem 2.3 we can extract a subsequence which we still denote by Ω_n, which converges to some $\Omega^\star \in \mathcal{O}$ in the sense of characteristic functions. Let us denote by u_n and $u^\star$ the solutions to (1.17) with Ω_n and $\Omega^\star$ respectively. It can be shown that $u_n \to u^\star$ in $H_0^1(D)$.

Proposition 2.5 ([14]). *Let $\Omega_n \in \mathcal{O}$ be a minimizing sequence for problem* (1.24) *and $\Omega^\star$ be an accumulation point of this sequence as in Theorem 2.3. Assume there exists*

a $\varepsilon > 0$ such that the solution u of (1.17) *satisfies*

$$\|u\|_{H^{1+\varepsilon}(D) \cap H_0^1(D)} \leq c$$

where c depends on F and on D. Then the sequence $u_n \in H_0^1(D)$ corresponding to Ω_n converges to $u^\star$ strongly in $H_0^1(D)$, where $u^\star$ is the solution to (1.17) *in $\Omega^\star$.*

Note that the mapping $u \mapsto \mathcal{J}(u)$ is continuous from $H_0^1(D)$ to $\mathbb{R}$ using the theorem of Lebesgue. Summarizing, the existence of a minimizer can be seen as follows:

$$m = \inf_{\Omega} J(\Omega) = \lim_{n \to \infty} J(\Omega_n) = \lim_{n \to \infty} \mathcal{J}(u_n) = \mathcal{J}(u^\star) = J(\Omega^\star),$$

where we used that Ω_n is a minimizing sequence, the definition of the reduced functional together with the notation $u_{\Omega_n} = u_n$, i.e. $J(\Omega_n) := \mathcal{J}(u_{\Omega_n}) = \mathcal{J}(u_n)$ and the continuity of $\mathcal{J}$.

Remark 2.2. Note that the existence of a minimizer for the linear case is also covered from the proof above. However, for completeness we treated the linear case separately.

2.3 Shape derivative via the averaged adjoint method

In this section we derive the shape derivative of model problem (1.24) and (1.26) using the averaged adjoint method. The averaged adjoint method is a quite novel Lagrangian approach introduced in [42] and is particularly well-suited for problems involving nonlinear PDE constraints. One of the main features of this method is that it allows to compute the shape derivative without the knowledge of the material derivative.

Let us briefly describe the procedure to calculate the shape derivative when using the averaged adjoint method [27]. Denote by $E = E(\Omega)$, $F = F(\Omega)$ some vector spaces and consider for $\varphi \in E$, $\psi \in F$ the Lagrangian $\mathcal{L}(\Omega, \varphi, \psi)$ of a PDE constrained shape optimization problem. Then the shape function on a perturbed domain $\Omega_t = T_t(\Omega)$ can be written as

$$J(\Omega_t) = \mathcal{L}(\Omega_t, u_t, \hat{\psi}),$$

where $u_t \in E(\Omega_t)$ is the solution to the perturbed state equation and $\hat{\psi} \in F(\Omega_t)$. Now, the shape derivative can be obtained by differentiating the Lagrangian with respect to t, however this step is crucial. In order to differentiate $\mathcal{L}(\Omega_t, \hat{\varphi}, \hat{\psi})$ with respect to t, the integrals in $\mathcal{L}(\Omega_t, \hat{\varphi}, \hat{\psi})$ need to be brought back to the fixed domain Ω by the transformation T_t. Unfortunately this leads to integrands including functions of the form $\hat{\varphi} \circ T_t$ and $\hat{\psi} \circ T_t$ which might be non-differentiable [14]. To overcome this issue we can reparameterize the spaces $E(\Omega_t)$, $F(\Omega_t)$ by composing the elements of $E(\Omega)$, $F(\Omega)$

with a suitable bijection. If E and F are H^1-spaces an appropriate transformation is given by T_t^{-1} and we may introduce

$$G(t, \varphi, \psi) := \mathcal{L}(\Omega_t, \varphi \circ T_t^{-1}, \psi \circ T_t^{-1}).$$

The main result in the following section states that the shape derivative of J can be obtained only by computing the partial derivative of G with respect to t and assigning proper values for φ and ψ.

2.3.1 An abstract result

Let E and F be Banach spaces. Consider for $\tau > 0$ a function

$$G : [0, \tau] \times E \times F \to \mathbb{R}, \quad (t, \varphi, \psi) \mapsto G(t, \varphi, \psi) \tag{2.7}$$

such that $\psi \mapsto G(t, \varphi, \psi)$ is affine for all $(t, \varphi) \in [0, \tau] \times E$. Define for $t \in [0, \tau]$ the set

$$E(t) := \{u \in E \mid d_\psi G(t, u, 0; \hat{\psi}) = 0 \text{ for all } \hat{\psi} \in F\},$$

called the set of solutions to the state equation. Let us introduce the following hypothesis.

Assumption 2.1 (H0). For every $(t, \psi) \in [0, \tau] \times F$ we assume that

(i) the set $E(t)$ is single-valued and we write $E(t) = \{u^t\}$,

(ii) the mapping $[0, 1] \ni s \mapsto G(t, su^t + (1 - s)u^0, \psi)$ is absolutely continuous,

(iii) the mapping $[0, 1] \ni s \mapsto d_\varphi G(t, su^t + (1 - s)u^0, \psi; \eta)$ belongs to $L^1(0, 1)$ for all $\eta \in E$.

For $t \in [0, \tau]$ and $u^t \in E(t)$ let us introduce the set

$$Y(t, u^t, u^0) := \left\{ q \in F \mid \forall \eta \in E : \int_0^1 d_\varphi G(t, su^t + (1 - s)u^0, q; \eta) \mathrm{d}s = 0 \right\} \tag{2.8}$$

which is called solution set of the averaged adjoint equation with respect to t, u^t and u^0. Setting $t = 0$ we obtain the solution set of the usual adjoint state equation:

$$Y(0, u^0, u^0) = \left\{ q \in F \mid d_\varphi G(0, u^0, q; \eta) = 0 \text{ for all } \eta \in E \right\}.$$

The following result proven in [42] allows us to calculate the shape derivative according to Definition 2.3 without the need of the material derivative $\dot{u}$. The key ingredient is the introduction of the set (2.8).

Theorem 2.6. *Let G be a function as in* (2.7) *satisfying Assumption (H0). Additionally let the following conditions hold:*

(H1) *For all $t \in [0, \tau]$ and $\psi \in F$ the derivative $\partial_t G(t, u^0, \psi)$ exists,*

(H2) *For all $t \in [0, \tau]$ the set $Y(t, u^t, u^0)$ is single-valued and we write $Y(t, u^t, u^0) = \{p^t\}$,*

(H3) *For any sequence of nonnegative real numbers $(t_n)_{n \in \mathbb{N}}$ converging to zero there exists a subsequence $(t_{n_k})_{k \in \mathbb{N}}$ such that*

$$\lim_{\substack{k \to \infty \\ s \searrow 0}} \partial_t G(s, u^0, p^{t_{n_k}}) = \partial_t G(0, u^0, p^0).$$

Then, for any $\psi \in F$ we obtain

$$\frac{d}{dt}(G(t, u^t, \psi))\,|_{t=0} = \partial_t G(0, u^0, p^0).$$

Remark 2.3 ([42]). In concrete applications the conditions (H0)-(H3) have the following meaning:

  (i) The condition (H0) ensures that we can apply the fundamental theorem of calculus with respect to u. By (H1) we are allowed to apply the mean value theorem with respect to t.

 (ii) (H2) guarantees the unique solvability of some averaged adjoint equation.

(iii) The condition (H3) can be shown by showing that p^t converges weakly to p^0 in F and that $(t, \psi) \mapsto G(t, \varphi, \psi)$ is weakly continuous for all $\varphi \in E$.

 (iv) The set $E(t)$ corresponds to the solution of the state equation on the perturbed domain $\Omega_t = T_t(\Omega)$ brought back to the fixed domain Ω.

2.3.2 Preliminaries: Newton operator, adjoint equation, properties of flows

This section serves as preparation for Section 2.3.3, where the derivation of the shape derivative for the nonlinear shape optimization problem is performed in detail. Here, we collect some necessary results for this computation which include an analysis of the adjoint equation and properties of flows.

Frechét derivative of the state operator

Recall the definition of the operator A_Ω defining the state equation (1.17)

$$\langle A_\Omega(u), \eta \rangle = \int_D \nu_\Omega(x, |\nabla u|) \nabla u \cdot \nabla \eta \, \mathrm{d}x,$$

for all $u, \eta \in H_0^1(D)$. We need to calculate the Frechét derivative of A_Ω (also called the Newton operator) because it will show up not only in the equation defining the adjoint

state but also in Newton's method when computing a solution to the nonlinear sate equation. For a more compact notation let us introduce the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T(W) := \hat{\nu}(|W|)W. \tag{2.9}$$

The Jacobian of T is given for $W \in \mathbb{R}^2$ by

$$DT(W) = \begin{cases} \hat{\nu}(|W|)I + \frac{\hat{\nu}'(|W|)}{|W|}W \otimes W, & W \neq (0,0)^T \\ \hat{\nu}(0)I, & W = (0,0)^T \end{cases} \tag{2.10}$$

where I denotes the two dimensional identity matrix and $\otimes$ the outer product of two column vectors, i.e.

$$a \otimes b := ab^T.$$

Note that DT is continuous also in $W = (0,0)^T$. In order to show the existence of a directional derivative it is sufficient to show that for any $u, \varphi, \eta \in H_0^1(D)$

$$t \mapsto \int_D \nu_\Omega(x, |\nabla(u + t\varphi)|)\nabla(u + t\varphi) \cdot \nabla\eta \, \mathrm{d}x,$$

is continuously differentiable on $\mathbb{R}$. Moreover it is sufficient to show that the above mapping is differentiable on the respective subdomains Ω_f and $\Omega_{air} = D \setminus \overline{\Omega_f}$. Here the only difficult part is the nonlinear term, so we show that

$$t \mapsto \int_{\Omega_f} \hat{\nu}(|\nabla(u + t\varphi)|)\nabla(u + t\varphi) \cdot \nabla\eta \, \mathrm{d}x = \int_{\Omega_f} T(\nabla(u + t\varphi)) \cdot \nabla\eta \, \mathrm{d}x \tag{2.11}$$

is differentiable on $\mathbb{R}$. Since T is continuously differentiable we get that the function $t \mapsto \gamma_t(x) := T(\nabla(u + t\varphi)(x)) \cdot \nabla\eta(x)$ is differentiable for almost every $x \in \Omega_f$. The derivative reads

$$\frac{d}{dt}\gamma_t(x) = DT(\nabla(u + t\varphi)(x))\nabla\varphi(x) \cdot \nabla\eta(x).$$

Using property (1.29) and (1.30) we get the estimate $|\hat{\nu}(s)| + |\hat{\nu}'(s)s| \leq 2\nu_0$ which we can use to conclude that

$$\left|\frac{d}{dt}\gamma_t(x)\right| \leq 2\nu_0 |\nabla\varphi(x)| |\nabla\eta(x)|, \text{ for almost every } x \in \Omega_f, \text{ for all } t \in \mathbb{R}. \tag{2.12}$$

Since $t \mapsto \frac{d}{dt}\gamma_t(x)$ is also continuous on $\mathbb{R}$ we may apply the fundamental theorem of calculus to get

$$\frac{\gamma_{t+h}(x) - \gamma_t(x)}{h} = \frac{1}{h}\int_t^{t+h} \frac{d}{ds}\gamma_s(x) \, \mathrm{d}s \leq 2\nu_0 |\nabla\varphi(x)| |\nabla\eta(x)|$$

for almost all $x \in \Omega_f$ where we used (2.12). Thus we can apply the theorem of Lebesgue to obtain

$$\frac{d}{dt} \int_{\Omega_f} \gamma_t(x) \, dx = \lim_{h \to 0} \int_{\Omega_f} \frac{\gamma_{t+h}(x) - \gamma_t(x)}{h} \, dx$$

$$= \int_{\Omega_f} \lim_{h \to 0} \frac{\gamma_{t+h}(x) - \gamma_t(x)}{h} \, dx = \int_{\Omega_f} \frac{d}{dt} \gamma_t(x) \, dx$$

which shows that (2.11) is indeed differentiable. The derivative reads

$$\frac{d}{dt} \int_{\Omega_f} \hat{\nu}(|\nabla(u + t\varphi)|) \nabla(u + t\varphi) \cdot \nabla\eta \, dx = \int_{\Omega_f} DT(\nabla(u + t\varphi)) \nabla\varphi \cdot \nabla\eta \, dx$$

and is obviously continuous. Put $t = 0$ we conclude that the directional derivative exists for all $u, \varphi \in H_0^1(D)$. Introduce the function $\mathcal{A}_\Omega : D \times \mathbb{R}^2 \to \mathbb{R}^{2,2}$ defined by

$$\mathcal{A}_\Omega(x, W) = DT(W)\chi_{\Omega_f}(x) + \nu_0 I \chi_{D \setminus \Omega_f}(x). \tag{2.13}$$

Then, the Gateaux derivative of $A_\Omega : H_0^1(D) \to H^{-1}(D)$ in $u \in H_0^1(D)$ is given by the linear and continuous mapping $A_\Omega'(u) : H_0^1(D) \to H^{-1}(D)$ defined by

$$\langle A_\Omega'(u)\varphi, \eta \rangle = \int_D \mathcal{A}_\Omega(x, \nabla u)\nabla\varphi \cdot \nabla\eta \text{ for all } \varphi, \eta \in H_0^1(D). \tag{2.14}$$

Moreover we get with the help of the mean value theorem that

$$|\langle A_\Omega(u + \varphi) - A_\Omega(u) - A_\Omega'(u)\varphi, \eta \rangle|$$

$$= \left| \int_{\Omega_f} \int_0^1 (DT(\nabla u + s\nabla\varphi) - DT(\nabla u)) ds \nabla\varphi \cdot \nabla\eta \, dx \right|$$

$$\leq \sup_{s \in [0,1]} \|DT(\nabla u + s\nabla\varphi) - DT(\nabla u)\| \|\varphi\|_{H^1(D)} \|\eta\|_{H^1(D)}$$

which we can use to see that

$$\frac{\|A_\Omega(u + \varphi) - A_\Omega(u) - A_\Omega'(u)\varphi\|_{H^{-1}(D)}}{\|\varphi\|_{H^1(D)}} \leq \sup_{s \in [0,1]} \|DT(\nabla u + s\nabla\varphi) - DT(\nabla u)\|$$

The right hand side of this goes to zero as $\varphi \to 0$ in $H_0^1(D)$ since DT is continuous. Therefore we conclude that $A_\Omega'(u)$ is the Fréchet derivative of A_Ω in u.

Let us have a closer look on the matrix $DT(W)$ for $W = (w_1, w_2)^T \in \mathbb{R}^2$. The eigenvalues and corresponding eigenvectors are [13]

$$\begin{array}{ll} \lambda_1 = \hat{\nu}(|W|), & v_1 = W^\perp = (-w_2, w_1)^T, \\ \lambda_2 = \hat{\nu}(|W|) + \hat{\nu}'(|W|)\,|W|, & v_2 = W. \end{array} \tag{2.15}$$

Note that $DT(W)$ is symmetric and due to (1.29) and (1.30) positive definite for all $W \in \mathbb{R}^2$. Furthermore it holds that

$$\lambda |z|^2 \leq z^T DT(W)z \leq \Lambda |z|^2, \quad \forall z, W \in \mathbb{R}^2 \tag{2.16}$$

with $\lambda = \min\{\lambda_1, \lambda_2\} \geq \underline{\nu}$ and $\Lambda = \max\{\lambda_1, \lambda_2\} \leq \nu_0$.

Adjoint equation

Define the Lagrangian associated to model problem (1.24) for $\varphi, \psi \in H_0^1(D)$ as

$$G(\Omega, \varphi, \psi) = -C_J \int_{\Omega_g} \nabla\varphi^T Q(x) \nabla\varphi \, \mathrm{d}x + \int_D \nu_\Omega(x, |\nabla\varphi|) \nabla\varphi \cdot \nabla\psi \, \mathrm{d}x - \langle F, \psi \rangle$$

Here, C_J denotes the constant prefactor of the cost function, i.e. $C_J = L/(\mu_0(r_s - r_r))$. The adjoint equation can be obtained by differentiating G with respect to φ at $\varphi = u$ and $\psi = p$ satisfying

$$d_\varphi G(\Omega, u, p; \eta) = 0 \quad \text{for all } \eta \in H_0^1(D).$$

A formal computation gives

$$\int_D \nu_\Omega(x, |\nabla u|) \nabla\eta \cdot \nabla p \, \mathrm{d}x + \int_{\Omega_f} \frac{\hat{\nu}'(|\nabla u|)}{|\nabla u|} (\nabla u \cdot \nabla\eta)(\nabla u \cdot \nabla p) \, \mathrm{d}x$$
$$= 2C_J \int_{\Omega_g} Q(x) \nabla u \cdot \nabla\eta \, \mathrm{d}x \text{ for all } \eta \in H_0^1(D).$$

Using the identity

$$(a \cdot b)(a \cdot c) = (a \otimes a)c \cdot b, \quad \text{for all } a, b, c \in \mathbb{R}^2$$

the previous equation can be written in more compact form as

$$\int_D \mathcal{A}(x, \nabla u)\nabla p \cdot \nabla\eta \mathrm{d}x = 2C_J \int_{\Omega_g} Q(x)\nabla u \cdot \nabla\eta \, \mathrm{d}x \text{ for all } \eta \in H_0^1(D)$$

where $\mathcal{A} : D \times \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$\mathcal{A}(x, y) = DT(y)\chi_{\Omega_f}(x) + \nu_0 I \chi_{D \setminus \Omega_f}(x), \tag{2.17}$$

with DT as in (2.10).

Lemma 2.7. *Let $\hat{\nu}$ be a reluctivity function coming from a B-H-curve satisfying Assumption 1.1. For $u \in H_0^1(D)$ the equation*

$$\int_D \mathcal{A}(x, \nabla u)\nabla p \cdot \nabla\eta \mathrm{d}x = 2C_J \int_{\Omega_g} Q(x)\nabla u \cdot \nabla\eta \, \mathrm{d}x \tag{2.18}$$

has a unique solution $p \in H_0^1(D)$.

Proof. Consider for $u \in H_0^1(D)$ the bilinear form

$$a'(u; \cdot, \cdot) : H_0^1(D) \times H_0^1(D) \to \mathbb{R} \tag{2.19}$$

$$(v, \eta) \mapsto \int_D \mathcal{A}(x, \nabla u)\nabla v \cdot \nabla\eta \, \mathrm{d}x \tag{2.20}$$

We want to apply the lemma of Lax-Milgram, therefore we check that the bilinear form $a'(u; \cdot, \cdot)$ is elliptic

$$
\begin{aligned}
a'(u; v, v) &= \int_D \mathcal{A}(x, \nabla u) \nabla v \cdot \nabla v \; \mathrm{d}x \\
&= \int_D \chi_{\Omega_f}(x) DT(\nabla u) \nabla v \cdot \nabla v + \chi_{D \setminus \Omega_f}(x) \nu_0 \nabla v \cdot \nabla v \; \mathrm{d}x \\
&\geq \underline{\nu} \int_D |\nabla u|^2 \; \mathrm{d}x \geq C \underline{\nu} \|v\|_{H^1(D)},
\end{aligned}
$$

Here we have used property (2.16) and the norm equivalence between the L^2 norm of the gradient and the H^1 norm since $v \in H_0^1(D)$. In order to show boundedness, recall that for fixed $\rho \in \mathbb{R}^2$ the matrix $DT(\rho)$ is symmetric (and positive definite) and therefore the spectral norm of $DT(\rho)$ coincides with its greatest eigenvalue, i.e $|DT(\rho)| := |DT(\rho)|_2 = \max\{\lambda_1(|\rho|), \lambda_2(|\rho|)\}$. Following this argument we have

$$
\begin{aligned}
|a'(u; v, \eta)| &\leq \int_D |\mathcal{A}(x, \nabla u) \nabla v \cdot \nabla \eta| \; \mathrm{d}x \\
&\leq \int_D \chi_{\Omega_f}(x) |DT(\nabla u) \nabla v \cdot \nabla \eta| + \chi_{D \setminus \Omega_f}(x) \nu_0 |\nabla v \cdot \nabla \eta| \; \mathrm{d}x \\
&\leq \int_D \max\{\lambda_1(|\nabla u|), \lambda_2(|\nabla u|), \nu_0\} |\nabla v| |\nabla \eta| \; \mathrm{d}x \\
&\leq \nu_0 \|\nabla v\|_{L^2(D)} \|\nabla \eta\|_{L^2(D)} \leq C \nu_0 \|v\|_{H^1(D)} \|\eta\|_{H^1(D)},
\end{aligned}
$$

where in the last step we exploited again the norm equivalence for functions in $H_0^1(D)$. Next we have to check that the linear functional

$$
\langle F_u, \eta \rangle := 2 C_J \int_{\Omega_g} Q(x) \nabla u \cdot \nabla \eta \; \mathrm{d}x
$$

is bounded. This can be seen as follows

$$
|\langle F_u, \eta \rangle| \leq 2 C_J \|Q\|_{C(\Omega_g, \mathbb{R}^{2,2})} \|\nabla u\|_{L^2(D)} \|\eta\|_{H^1(D)}.
$$

Thus, the lemma of Lax-Milgram yields a unique solution $p \in H_0^1(D)$ to problem (2.18). $\qquad \square$

Properties of flows and perturbed state equation

Let $V \in C_c^1(D, \mathbb{R}^d)$ be a vector field and T_t its associated flow. The important observation of the following result is that the flow T_t creates an isomorphism between $H^1(\Omega)$ and $H^1(\Omega_t)$ where $\Omega_t := T_t(\Omega)$.

Lemma 2.8. *Let $p \geq 1$ and suppose that $T : \mathbb{R}^d \to \mathbb{R}^d$ is a bi-Lipschitz mapping. Let U be an open subset of $\mathbb{R}^d$ and set $W := T^{-1}(U)$. Then we have*

$$u \in W^{1,p}(U) \Leftrightarrow u \circ T \in W^{1,p}(W)$$

Proof. A proof can be found in [45, p.52 Theorem 2.2.2] $\qquad\square$

In order to derive the shape derivative we have to consider the perturbed state equation, that is u_t is the weak solution to (1.17) on Ω_t, more precisely $u_t \in H^1_0(D)$ solves

$$\int_D \nu_{\Omega_t}(|\nabla u_t|)\nabla u_t \cdot \nabla \hat{\eta} \, \mathrm{d}x = \int_D J_3 \hat{\eta} \, \mathrm{d}x \quad \text{for all } \hat{\eta} \in H^1_0(D). \tag{2.21}$$

Notice that an application of the chain rule yields $(\nabla f) \circ T_t = \partial T_t^{-T} \nabla (f \circ T_t)$ for $f \in H^1(D)$. Hence, using the change of variables $y = T_t(x)$ and Lemma 2.8 in (2.21) shows that $u^t := u_t \circ T_t$ satisfies

$$\int_D A(t)\nu_{\Omega}(x, |B(t)\nabla u^t|)\nabla u^t \cdot \nabla \psi \, \mathrm{d}x = \int_D \xi(t)(J_3 \circ T_t)\psi \, \mathrm{d}x \text{ for all } \psi \in H^1_0(D), \tag{2.22}$$

where we used the abbreviations

$$\xi(t) := \det(\partial T_t), \tag{2.23a}$$

$$B(t) := \partial T_t^{-T}, \tag{2.23b}$$

$$A(t) := \det(\partial T_t)\partial T_t^{-1}\partial T_t^{-T}. \tag{2.23c}$$

Remark 2.4. Note that the function $\xi(t)$ is positive for small values of t. Therefore we omit the absolute value of $\xi(t)$ in (2.22).

One can show that the functions ξ, A, B are differentiable.

Lemma 2.9 ([43]). *Let $V \in C^1_c(D, \mathbb{R}^d)$ and T_t the flow associated with V. The mappings $t \mapsto \xi(t)$, $t \mapsto B(t)$ and $t \mapsto A(t)$ according to (2.23) are differentiable on $[0, \tau]$ and the derivatives read*

$$\frac{d}{dt}\xi(t) = \operatorname{tr}(\partial V^t B^T(t))\xi(t), \tag{2.24}$$

$$\frac{d}{dt}B(t) = -B(t)(\partial V^t)^T B(t), \tag{2.25}$$

$$\frac{d}{dt}A(t) = \operatorname{tr}(\partial V^t B^T(t))A(t) - B(t)^T \partial V^t A(t) - (B(t)\partial V^t A(t))^T \tag{2.26}$$

where $V^t(x) := V(T_t(x))$.

Here, differentiable in $t = 0$ respectively in $t = \tau$ means the existence of the right sided respectively left sided derivative.

Corollary 2.10. *In particular we get from Lemma 2.9 that*

$$\xi'(0) = \operatorname{div} V \tag{2.27}$$

$$B'(0) = -\partial V^T \tag{2.28}$$

$$A'(0) = \operatorname{div} V - \partial V - \partial V^T \tag{2.29}$$

Furthermore it holds $\xi(0) = 1$ and $A(0) = B(0) = I$.

In addition to the previous result we will make use of another property of A and ξ.

Lemma 2.11 ([43]). *Let $A \in C([0,\tau]; C(\overline{D}, \mathbb{R}^{d,d}))$ and $\xi \in C([0,\tau]; C(\overline{D}))$ be given and assume that $A(0) = I$ and $\xi(0) = 1$. Then there exist constants $\alpha_1, \alpha_2, \beta_1, \beta_2 > 0$ and $\tilde{\tau} > 0$ such that for all $\zeta \in \mathbb{R}^d$ and for all $t \in [0, \tilde{\tau}]$ it holds*

$$\alpha_1 |\zeta|^2 \le A(t)\zeta \cdot \zeta \le \alpha_2 |\zeta|^2, \tag{2.30}$$

$$\beta_1 \le \xi(t) \le \beta_2. \tag{2.31}$$

2.3.3 Shape derivative of the cost function

In this section we derive the shape derivative of the cost function $\mathcal{J}$ of problem (1.24) by applying Theorem 2.6. From this we can easily conclude the shape derivative of the linear problem (1.26).
We assume that $V \in C_c^1(D, \mathbb{R}^2)$ and $\operatorname{supp}(V) \cap \Omega_g = \emptyset$. Further let $V = 0$ on the boundary of the design domain, that is $\partial \Omega^d$. Let T_t be the flow associated to V. The conditions on V ensures that $\Omega_t = T_t(\Omega) \subset \Omega^d$ and that $T_t(\Omega^d) = \Omega^d$ for t small enough. First, we address the nonlinear problem.

Theorem 2.12. *Let $\hat{\nu}$ be a reluctivity function according to B-H-curve satisfying Assumption 1.1. Then the reduced functional J defined by (1.24) is shape differentiable and its shape derivative in direction V reads*

$$dJ(\Omega; V) = \int_D (\operatorname{div}(V)I - \partial V - \partial V^T)\nu_\Omega(x, |\nabla u|)\nabla u \cdot \nabla p \,\mathrm{d}x$$
$$- \int_{\Omega_f} \frac{\hat{\nu}'(|\nabla u|)}{|\nabla u|}(\partial V^T \nabla u \cdot \nabla u)(\nabla u \cdot \nabla p) \,\mathrm{d}x \tag{2.32}$$

Proof. Consider the Lagrangian on the perturbed domain $\Omega_t := T_t(\Omega)$ for $\varphi, \psi \in H_0^1(D)$, i.e.

$$\mathcal{L}(\Omega_t, \varphi, \psi) = -C_J \int_{\Omega_g} Q(x)\nabla\varphi \cdot \nabla\varphi \,\mathrm{d}x + \int_D \nu_{\Omega_t}(x, |\nabla\varphi|)\nabla\varphi \cdot \nabla\psi \,\mathrm{d}x - \langle F, \psi \rangle$$

with F as in (1.16) and where C_J is given by $C_J = L/(\mu_0(r_s - r_r))$. Following the methodology described at the beginning of Section 2.3 we introduce

$$G(t, \varphi, \psi) := \mathcal{L}(\Omega_t, \varphi \circ T_t^{-1}, \psi \circ T_t^{-1})$$

which after the change of variables $y = T_t(x)$ reads

$$
\begin{aligned}
G(t, \varphi, \psi) = &-C_J \int_{\Omega_g} \xi(t) B(t)^T Q^t B(t) \nabla \varphi \cdot \nabla \varphi \, \mathrm{d}x \\
&+ \int_D A(t) \nu_\Omega(x, |B(t)\nabla\varphi|) \nabla \varphi \cdot \nabla \psi \, \mathrm{d}x - \int_D \xi(t)(J_3 \circ T_t)\psi \, \mathrm{d}x,
\end{aligned}
\tag{2.33}
$$

where $Q^t = Q \circ T_t$ and ξ, A and B are defined as in Lemma 2.9. We apply Theorem 2.6 to this G with $E = F = H_0^1(D)$. Note that $J(\Omega_t) = G(t, u^t, \psi)$, where $u^t \in H_0^1(D)$ solves

$$
\int_D A(t)\nu_\Omega(x, |B(t)\nabla u^t|)\nabla u^t \cdot \nabla \psi \, \mathrm{d}x = \int_D \xi(t)(J_3 \circ T_t)\psi \, \mathrm{d}x \text{ for all } \psi \in H_0^1(D). \tag{2.34}
$$

Let us now verify Hypotheses (H0)-(H3).

(H0) Condition (i) is satisfied due to the Theorem of Zarantonello 1.2 and the properties of $t \mapsto A(t)$ for all $t \in [0, \tau]$. Let $u_s^t := su^t + (1-s)u^0$ and $\psi \in H_0^1(D)$. In order to show condition (ii) it is sufficient that the mapping $s \mapsto G(t, u_s^t, \psi)$ is continuous differentiable for all $(t, \psi) \in [0, \tau] \times H_0^1(D)$. Here the only difficult part is the nonlinear term coming from the state equation. However, we have seen that the nonlinear operator A_Ω is Fréchet-differentiable and the derivative is continuous. Therefore condition (ii) is satisfied. Condition (iii) is satisfied by construction.

(H1) We have that the operator A_Ω is differentiable and that the functions $\xi(t)$, $A(t)$ and $B(t)$ are differentiable according to Lemma 2.9. Therefore the mapping $t \mapsto G(t, \varphi, \psi)$ is differentiable for all $t \in [0, \tau]$ and all $\varphi, \psi \in H_0^1(D)$.

(H2) We have already proved that $E(t) = \{u^t\}$, where u^t is the solution to (2.34). We show that $Y(t, u^t, u^0) = \{p^t\}$ where $p^t \in H_0^1(D)$ is the unique solution to

$$
\begin{aligned}
\int_0^1 \int_D &\xi(t)\mathcal{A}(x, B(t)\nabla u_s^t)B(t)\nabla p^t \cdot B(t)\nabla\eta \, \mathrm{d}x\mathrm{d}s \\
&= 2C_J \int_0^1 \int_{\Omega_g} \xi(t)Q^t B(t)\nabla u_s^t \cdot B(t)\nabla\eta \, \mathrm{d}x\mathrm{d}s \text{ for all } \psi \in H_0^1(D),
\end{aligned}
\tag{2.35}
$$

where $u_s^t := su_t + (1-s)u^0$ and $\mathcal{A}(\cdot, \cdot)$ is given by (2.17). Due to (H0) this equation is well defined. The existence of a unique solution follows from the Lemma of Lax-Milgram exploiting the properties of Lemma 2.11. In particular $p^0 = p \in Y(0, u^0, u^0)$ is the unique solution of the adjoint equation (2.18).

(H3) We show that for any sequence $(t_n)_{n \in \mathbb{N}} \searrow 0$ there exists a subsequence $(t_{n_k})_{k \in \mathbb{N}}$ such that the sequence $(p^{t_{n_k}})_{k \in \mathbb{N}}$ where $p^{t_{n_k}} \in Y(t_k, u^{t_{n_k}}, u^0)$ converges weakly in $H_0^1(D)$ to the solution of the adjoint equation and that $(t, \psi) \mapsto \partial_t G(t, \varphi, \psi)$ is weakly continuous. We will use the following two lemmas.

Lemma 2.13. *The mapping $t \mapsto u^t$ is continuous from the right in 0, i.e.*

$$
\lim_{t \searrow 0} \|u^t - u\|_{H^1(D)} = 0.
$$

Proof. A proof can be found in [14]. □

Lemma 2.14. *There are constants $c > 0$ and $\tau > 0$ such that*

$$\left\|u^t\right\|_{H^1(D} \leq c \text{ for all } t \in [0, \tau].$$

Proof. Inserting $\psi = u^t$ in (2.34) we get

$$c\left\|u^t\right\|^2_{H^1(D)} \leq \alpha_1 \underline{\nu} \int_D \left|\nabla u^t\right|^2 \, \mathrm{d}x \leq \underline{\nu} \int_D A(t) \nabla u^t \cdot \nabla u^t \, \mathrm{d}x$$

$$\leq \int_D A(t)\nu_\Omega(x, \left|B(t)\nabla u^t\right|)\nabla u^t \cdot \nabla \psi \, \mathrm{d}x$$

$$= \int_D \xi(t) J_3 \circ T_t \psi \, \mathrm{d}x \leq \beta_2 \|J_3 \circ T_t\|_{L^2(D)} \left\|u^t\right\|_{H^1(D)}$$

where we used Poincaré's inequality, Lemma 2.11 and the lower bound for the function ν_Ω. Dividing by $\|u^t\|_{H^1(D)}$ yields to the claimed result. □

Now, we are able to show the following.

Lemma 2.15. *For any sequence of $(t_n)_{n\in\mathbb{N}}$ of non negative real numbers converging to zero there exists a subsequence $(t_{n_k})_{k\in\mathbb{N}}$ such that $(p^{t_{n_k}})_{k\in\mathbb{N}}$ where $p^{t_{n_k}}$ solves (2.35) for $t = t_{n_k}$ converges weakly in $H^1_0(D)$ to the solution p of the adjoint equation (2.18).*

Proof. The existence of a solution to (2.35) follows from the Lemma of Lax-Milgram. Inserting $\psi = p^t$ as test function in (2.35) we obtain with the help of Lemma 2.14 and the estimate

$$\underline{\nu} z \cdot z \leq \mathcal{A}(x, \rho) z \cdot z \leq \nu_0 z \cdot z \text{ for all } z, \rho \in \mathbb{R}^2,$$

that $\|p^t\| \leq \tilde{C}$ for all sufficiently small t, where $\tilde{C} > 0$ is a constant. Now, let $(t_n)_{n\in\mathbb{N}}$ be a sequence of non negative real numbers converging to zero. Since $(p^{t_n})_{n\in\mathbb{N}}$ is bounded we may extract a subsequence $(p^{t_{n_k}})_{k\in\mathbb{N}}$ which converges weakly in $H^1_0(D)$ to some $q \in H^1_0(D)$. Consider (2.35) with $t = t_{n_k}$:

$$\int_0^1 \int_D \xi(t_{n_k})\mathcal{A}(x, B(t_{n_k})\nabla u_s^{t_{n_k}})B(t_{n_k})\nabla p^{t_{n_k}} \cdot B(t_{n_k})\nabla \eta \, \mathrm{d}x\mathrm{d}s$$

$$= 2C_J \int_0^1 \int_{\Omega_g} \xi(t_{n_k})Q^{t_{n_k}}B(t_{n_k})\nabla u_s^{t_{n_k}} \cdot B(t_{n_k})\nabla \eta \, \mathrm{d}x\mathrm{d}s \text{ for all } \psi \in H^1_0(D),$$

$$(2.36)$$

Thanks to Lemma 2.13 we have that $u^t \to u$ in $H^1_0(D)$ and using the continuity of ξ and B we can pass to the limit to see that $q \in H^1_0(D)$ is the solution to the adjoint equation (2.18). By uniqueness of a solution of the adjoint equation we conclude $q = p$. □

Finally, note that the derivative of (2.33) for $t > 0$ is given by

$$
\begin{aligned}
\partial_t G(t, \varphi, \psi) = & - C_J \int_{\Omega_g} \xi'(t) Q^t B(t) \nabla\varphi \cdot B(t) \nabla\varphi \, \mathrm{d}x \\
& - C_J \int_{\Omega_g} \xi(t) (\partial Q \odot V) B(t) \nabla\varphi \cdot B(t) \nabla\varphi \, \mathrm{d}x \\
& - C_J \int_{\Omega_g} \xi(t) Q^t B'(t) \nabla\varphi \cdot B(t) \nabla\varphi \, \mathrm{d}x \\
& - C_J \int_{\Omega_g} \xi(t) Q^t B(t) \nabla\varphi \cdot B'(t) \nabla\varphi \, \mathrm{d}x \\
& + \int_D A'(t) \nu_\Omega(x, |B(t)\nabla\varphi|) \nabla\varphi \cdot \nabla\psi \, \mathrm{d}x \\
& + \int_{\Omega_f} A(t) \frac{\hat{\nu}'(|B(t)\nabla\varphi|)}{|B(t)\nabla\varphi|} (B(t)\nabla\varphi \cdot B'(t)\nabla\varphi)(\nabla\varphi \cdot \nabla\psi) \, \mathrm{d}x \\
& - \int_D \xi'(t) J_3 \circ T_t \psi \mathrm{d}x
\end{aligned}
\tag{2.37}
$$

where $\partial Q \odot V := (\nabla q_{ij} \cdot V)_{i,j=1,2}$ and where q_{ij} are the components of the matrix Q. We see that for fixed $\varphi \in H_0^1(D)$ the mapping $(t, \psi) \mapsto \partial_t G(t, \varphi, \psi)$ is weakly continuous. This finishes the proof of (H3) and so we can apply Theorem 2.6 to obtain $dJ(\Omega, V) = \partial_t G(0, u, p)$ where $u \in H_0^1(D)$ is the solution to the state equation (1.17) and $p \in H_0^1(D)$ solves the adjoint equation (2.18). In order to calculate $\partial_t G(0, u, p)$ note that the integrals on Ω_g and the integrals involving J_3 vanish because J_3 is supported only on Ω_c and $V = 0$ on Ω_g and on Ω_c respectively. Taking into account Corollary 2.10 we arrive at formula (2.32) which finishes the proof of Theorem 2.12. ∎

Finally, we deduce the shape derivative of the linear model problem.

Corollary 2.16. *The reduced functional J defined by (1.26) is shape differentiable and its shape derivative in direction V is given by*

$$
dJ(\Omega; V) = \int_D (\mathrm{div}(V) I_2 - \partial V - \partial V^T) \nu_\Omega(x) \nabla u \cdot \nabla p \, \mathrm{d}x
\tag{2.38}
$$

where $u \in H_0^1(D)$ satisfies (1.26b) and $p \in H_0^1(D)$ is the solution to the problem

$$
\int_D \nu_\Omega(x) \nabla\eta \cdot \nabla p \, \mathrm{d}x = 2C_J \int_{\Omega_g} Q(x) \nabla u \cdot \nabla\eta \, \mathrm{d}x, \text{ for all } \eta \in H_0^1(D).
\tag{2.39}
$$

Proof. All the steps from the proof of Theorem 2.12 remain valid when considering a piecewise constant reluctivity $\nu_\Omega(x)$. Thus, the formula in (2.32) reduces to (2.38). □

2.4 Simulation of the deterministic model problems

In this section we have a closer look on the numerical procedure to solve the model problems and present the obtained results. So far, we have calculated the shape derivative for both model problems. We will use this sensitivity information in a gradient-based optimization algorithm to compute in every iteration a deformation field V which represents a descent direction for J, i.e. $dJ(\Omega; V) < 0$ and move the interface between the ferromagnetic material and air a certain step into the direction of V. In the following section we address the question how to compute a decent direction and discuss the implemented algorithm in detail.

2.4.1 Numerical method

Descent direction

Recall the definition of a descent direction.

Definition 2.6 (descent direction [27])**.** The vector field $V \in C_c^{0,1}(D, \mathbb{R}^d)$ is called a descent direction for J at Ω if there exists $\tau > 0$ such that

$$J(\Omega_t) < J(\Omega) \text{ for all } t \in (0, \tau).$$

A descent direction leads to a decrease of the shape function J and can be used in an iterative procedure to find a (local) minimizer of J. We are interested in finding a vector field V such that

$$dJ(\Omega; V) < 0.$$

Note that this V is indeed a descent direction since by definition of the shape derivative we can find $\tau > 0$ such that

$$\frac{J(\Omega_t) - J(\Omega)}{t} < 0 \quad \text{for all } t \in (0, \tau),$$

and therefore $J(\Omega_t) < J(\Omega)$ for all $t \in (0, \tau)$. In order to be able to compute a descent direction we make the following assumption:

Assumption 2.2. We assume that $dJ(\Omega; \cdot)$ is a linear and bounded functional on $H_0^1(\Omega, \mathbb{R}^d)$.

Note that this assumption is satisfied for the shape derivatives in (2.32) and (2.38) by density of $C_c^\infty(D, \mathbb{R}^2)$ in $H_0^1(D, \mathbb{R}^2)$. Now, we choose a symmetric and positive definite bilinear form

$$b : H_0^1(\Omega^d, \mathbb{R}^2) \times H_0^1(\Omega^d, \mathbb{R}^2) \to \mathbb{R}$$

which is defined on the rotor subdomain Ω^d of the hold all domain D and solve the auxiliary boundary value problem: Find $V \in P_h$ such that

$$b(V, W) = -dJ(\Omega, W) \text{ for all } W \in P_h. \tag{2.40}$$

Here, $P_h \subset H_0^1(\Omega^d, \mathbb{R}^2)$ is a suitable finite dimensional space. Outside of Ω^d we extend V by zero. This ensures that $V = 0$ on Ω_g which is assumed in Section 2.3.3. Note that this V is a descent direction since

$$dJ(\Omega; V) = -b(V, V) < 0.$$

Moreover V will be in $W^{1,\infty}$ and thus the flow associated to this V will be well defined [43]. The choice of the bilinear form is problem depending and we only mention here that some forms may lead to better mesh quality than others [12, 24]. In our experiments we choose the bilinear form associated to the linear elasticity problem, i.e.

$$b(V, W) = \int_{\Omega^d} (\lambda \operatorname{tr}(\varepsilon(V))I + 2\mu\varepsilon(V)) : \varepsilon(W)\mathrm{d}x, \qquad (2.41)$$

where $\varepsilon(U) := 1/2(\partial U + \partial U^T)$, $A : B$ the Frobenius inner product between two matrices A, B and λ, μ denote the Lamé parameters. For our computations we set $\lambda = 0$ and $\mu = 10$.

Remark 2.5. In the course of optimization we have to solve (2.40) on the transformed domain $\Omega_t = T_t(\Omega_0)$ where Ω_0 denotes the initial domain. In order to solve this problem efficiently we reformulate it to a problem on the fixed domain Ω_0 using a change of variables $y = T_t(x)$, $x \in \Omega_0$ and a reparametrization of the space H^1 according to Lemma 2.8.

Optimization step and algorithm

Let Ω_0 denote the initial domain. After having computed a descent direction V we choose a stepsize $\tau > 0$ according to a so called backtracking (line search) procedure, that is: decrease τ until the value of the cost function on the updated geometry has decreased. In particular choose $\tau = \max\{1, 1/2, 1/4, ...\}$ such that $J((id + \tau V)(\Omega_0)) < J(\Omega_0)$. When the step size becomes too small, that means if no decrease could be achieved the algorithm is stopped. Otherwise we set $\Omega_1 := (id + \tau V)(\Omega_0)$. This procedure is used to set up the following optimization algorithm.

Algorithm 2.1. Initialization: Set $k = 0$, choose initial design Ω_0, compute $J(\Omega_0)$, set $T_0 = id$, choose $\bar{\tau} > 0$;

(1) compute descent direction V_k by solving (2.40);

(2) find $\tilde{\tau}_k > 0$ such that $J((T_k + \tau_k V_k)(\Omega_0)) < J(T_k(\Omega_0))$ via backtracking, where $\tau_k = \tilde{\tau}_k \bar{\tau}/\|V_k\|_\infty$;

(3) if no decrease could be achieved: stop;

(4) set $T_{k+1} = T_k + \tau_k V_k$ and $k \leftarrow k + 1$ and go to (1);

Note that in order to evaluate the cost function J the state equation has to be solved. Moreover for each computation of a descent direction V_k via (2.40) the state u

and the adjoint state p on the current design Ω_k are required. The parameter $\bar{\tau}$ allows for an additional scaling of the step size and is chosen by the user. In our experiments we set $\bar{\tau}$ to be the maximal mesh size. The calculation of V_k in step 1 has to be understood as in Remark 2.5. Observe that in the previous algorithm we only choose once the initial domain Ω_0 on which all computations are performed. That means that we do not move any nodes of the mesh but instead we work with the transformation maps T_k to store deformations which avoids computational complexity. The deformed shapes are given by $\Omega_k = T_k(\Omega_0) = (id + \tau_1 V_1) \circ \dots \circ (id + \tau_{k-1} V_{k-1})(\Omega_0)$.

Newtons'method

In order to solve the nonlinear state equation (1.17) of problem (1.24) we use Newton's method to the operator equation $R_\Omega(u) := A_\Omega(u) - F$ in $H^{-1}(D)$. Therefore we start with an initial guess $u^0 \in H_0^1(D)$ and compute for $k = 0, 1, 2, \dots$ the next iterate as $u^{k+1} = u^k + w^k$, where the update $w^k \in H_0^1(D)$ is the solution to

$$\left\langle A'_\Omega(u^k)w^k, \eta \right\rangle = -\left\langle R_\Omega(u^k), \eta \right\rangle \text{ for all } \eta \in H_0^1(D).$$

Note that this problem is well defined due to Lemma 2.7. It can be shown that the discretized operator to A'_Ω is Lipschitz continuous and thus the method converges locally quadratically [32]. Global convergence can be achieved by using the damped version of Newton's method, i.e. setting $u^{k+1} = u^k + \tau^k w^k$ with $\tau^k \in (0, 1]$ sufficiently small [13],[25].

2.4.2 Numerical results

First, we apply Algorithm 2.1 to solve model problem (1.26) which is constrained by the equation of linear magnetostatics. The shape derivative, the state and the adjoint equation are given by (2.38),(1.26b) and (2.39) respectively. In all our computations, which were done with the Finite Element software `Netgen/NGSolve` [38], we use piecewise linear finite elements on a triangular grid to solve the boundary value problems numerically. The algorithm terminates after 30 iterations and the results are shown in Figure 2.1. We observe a decrease of the cost function from -1.026442 Nm to -1.317712 Nm. Moreover Figure 2.1 shows the final design together with the deformed mesh. Observe that nodes on the boundary of the rotor do not move which is due to the homogenous Dirichlet boundary condition for the deformation field V. Note that in this model we do not consider any saturation effects of the material which may lead to a very thin distribution of ferromagnetic material in the rotor.

Now, we turn to the simulation of the more realistic model problem (1.24) which is constrained by the nonlinear equation of magnetostatics. Here the shape derivative, the state and the adjoint equation are given by (2.32),(1.17) and (2.18) respectively. The final design of the rotor after 25 iterations of Algorithm 2.1 is depicted in Figure 2.2. The torque of the motor increases from 1.029028 Nm to 1.259251 Nm. Figure

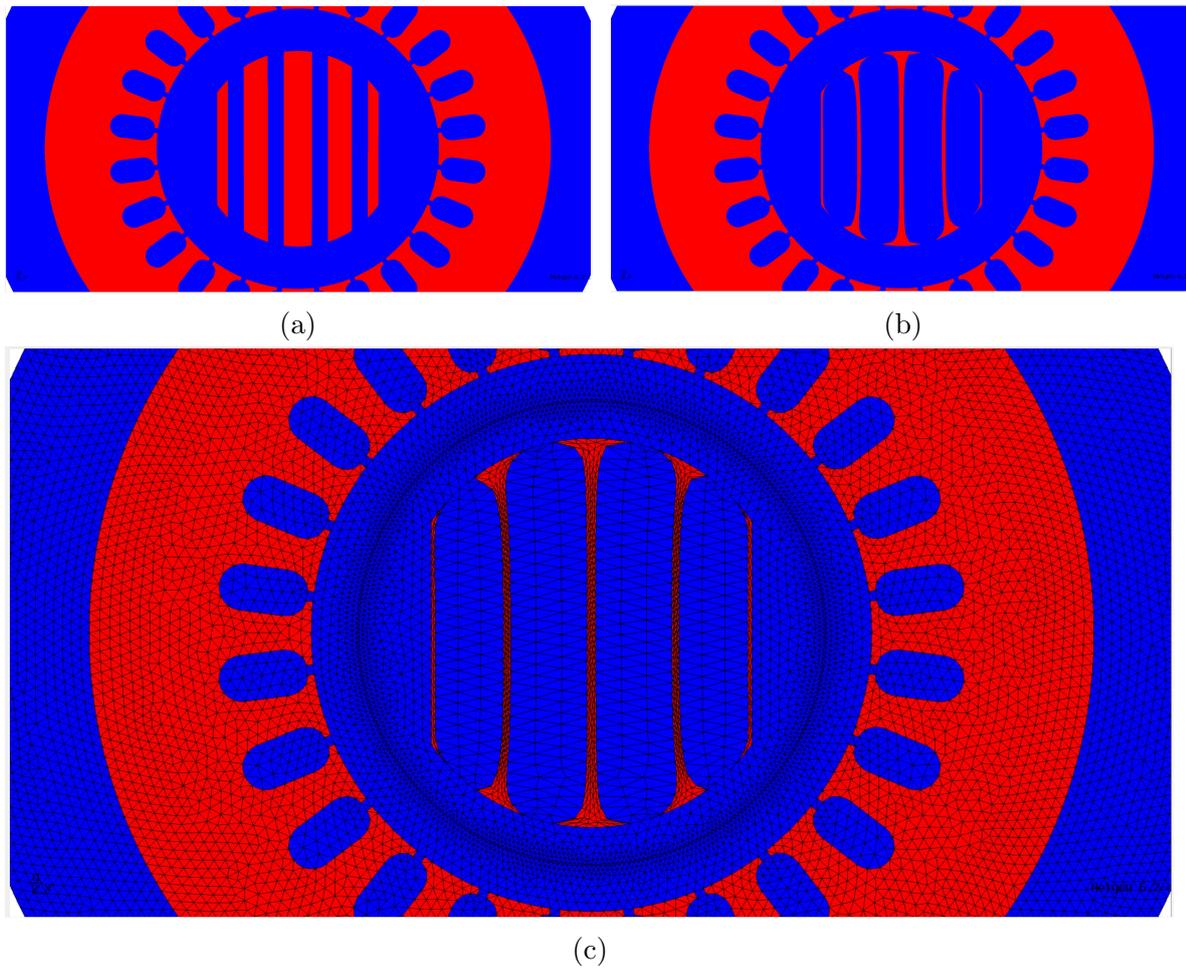(a)   (b)

(c)

Figure 2.1: Application of Algorithm 2.1 to linear problem (1.26). (a) Initial design: $T = 1.026442$ Nm.(b) Final design: $T = 1.317712$ Nm. (c) Final design including deformed mesh.

2.3 shows the magnetic flux density of the final design. Recalling the B-H curve of the ferromagnetic material (Figure 1.2) one can see that some layers of the rotor are close to saturation. Finally, since there is a significant difference in the final torque between the linear and nonlinear problem we conclude that the linear problem is not very realistic here.

(a)   (b)
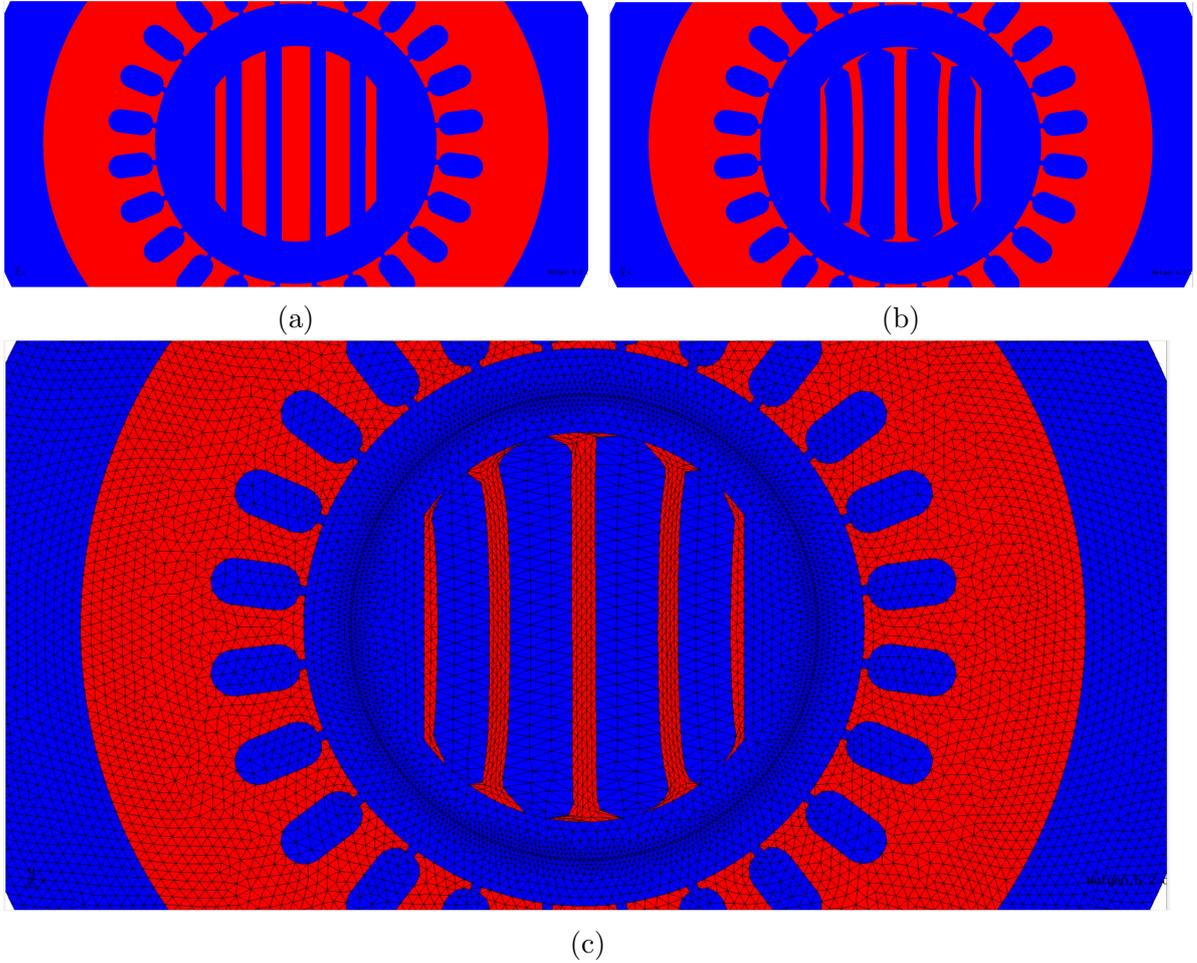
(c)

Figure 2.2: Application of Algorithm 2.1 to nonlinear problem (1.24). (a) Initial design: $T = 1.029028$ Nm. (b) Final design: $T = 1.259251$ Nm. (c) Final design including deformed mesh

2.5 Shape optimization of a synchronous reluctance motor under uncertainties

So far we dealt with the nonlinear behavior of the ferromagnetic material. In practice the magnetic reluctivity is obtained from measurements and is subject to measurement errors. Therefore the reluctivity curve might not be known exactly but rather be randomly distributed according to a probability distribution obtained empirically [17]. In this section we want to incorporate uncertainties in the ferromagnetic reluctivity of the motor leading to a stochastic problem formulation. Further we apply a novel stochastic shape optimization approach [17] to tackle this problem and hence to obtain a more robust design of the rotor.
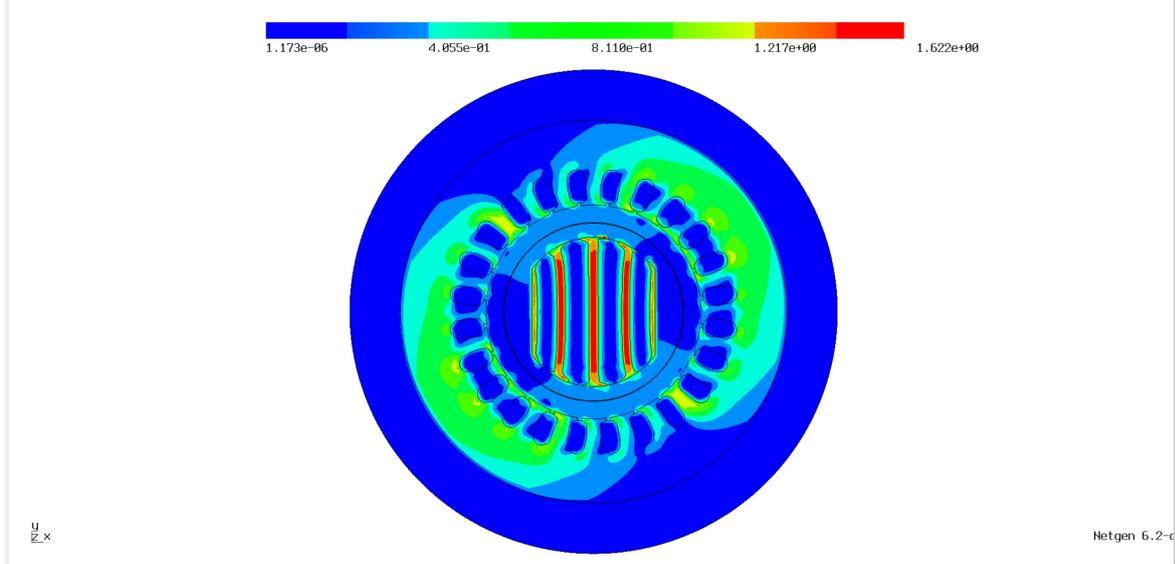
Figure 2.3: Distribution of the magnitude of the magnetic flux density $|\boldsymbol{B}|$ for the final design. Unit of the Color bar is Tesla.

2.5.1 A stochastic model problem

We consider the nonlinear model problem (1.24) and extend it to allow for stochastic data in the material term. Let $(X, \mathcal{F}, \mathbb{P})$ be a complete probability space, where $\mathcal{F} \subset 2^X$ is the σ-algebra of events and $\mathbb{P} : X \to [0,1]$ is a probability measure. We assume that the uncertain magnetic reluctivity is of the form

$$\nu_\Omega(x, |\nabla u|, \omega) = \begin{cases} \hat{\nu}(|\nabla u|)(1 + \delta\xi(\omega)) & x \in \Omega_f, \omega \in X \\ \nu_0 & x \in \Omega_{air}, \end{cases} \tag{2.42}$$

where $\omega \in X$ and ξ is a uniformly distributed random variable in the interval $[-1, 1]$. The parameter $\delta > 0$ describes the relative magnitude of the perturbation in the ferromagnetic material. Unless stated otherwise we set $\delta = 0.1$ which corresponds to a perturbation of 10%. Then, the weak formulation of the nonlinear state equation (1.17) in the stochastic setting for a fixed realization $\omega \in X$ reads: find $u = u(\cdot, \omega) \in H_0^1(D)$ such that

$$a_\omega(u, \eta) = \langle F, \eta \rangle \text{ for all } \eta \in H_0^1(D) \tag{2.43}$$

with

$$a_\omega(u, \eta) = \int_D \nu_\Omega(x, |\nabla u|, \omega) \nabla u(x, \omega) \cdot \nabla \eta(x) \, \mathrm{d}x \text{ and} \tag{2.44}$$

$$\langle F, \eta \rangle = \int_D J_3(x)\eta(x) \, \mathrm{d}x. \tag{2.45}$$

Now, the objective function depends on the random data as well and so for a single realization $\omega \in X$ we define

$$\mathcal{J}_\omega(u) := \mathcal{J}(u(\cdot, \omega)) \tag{2.46}$$

where $\mathcal{J}$ is defined as in (1.23). Since we are interested in a high torque over all possible outcomes of ω we introduce the expectation as

$$\mathbb{E}[\mathcal{J}_\omega(u)] := \int_X \mathcal{J}_\omega(u) \mathrm{d}\mathbb{P}(\omega)$$

and consider the following stochastic shape optimization problem:

$$\inf_{\Omega \in \mathcal{O}} \mathbb{E}[\mathcal{J}_\omega(u)] \tag{2.47a}$$

$$\text{subject to } u \in H_0^1(D) : a_\omega(u, \eta) = \langle F, \eta \rangle \text{ for all } \eta \in H_0^1(D), \ \omega \in X \tag{2.47b}$$

where the admissible set of shapes $\mathcal{O}$ is given as in (1.24c), a_ω and F are defined above respectively. Note that the solution of the state equation depends on the current shape, i.e. $u = u(\Omega)$. Hence, we also introduce the stochastic reduced functional $J_\omega(\Omega) := \mathcal{J}_\omega(u_\Omega)$. In order to solve the above model problem we need to calculate its shape derivative. For fixed $\omega \in X$ this can be done as in the deterministic case and so the shape derivative is given by

$$\begin{aligned}
dJ_\omega(\Omega; V) = &\int_D (\mathrm{div}(V)I - \partial V - \partial V^T)\nu_\Omega(x, |\nabla u|, \omega)\nabla u \cdot \nabla p \mathrm{d}x \\
&- \int_{\Omega_f} \frac{\hat{\nu}'(|\nabla u|)(1 + \delta\xi(w))}{|\nabla u|}(\partial V^T \nabla u \cdot \nabla u)(\nabla u \cdot \nabla p)\mathrm{d}x,
\end{aligned} \tag{2.48}$$

where $p = p(\cdot, \omega) \in H_0^1(D)$ is the solution to the adjoint equation

$$\int_D \mathcal{A}_\omega(x, \nabla u)\nabla p \cdot \nabla \eta \ \mathrm{d}x = 2C_J \int_{\Omega_g} Q(x)\nabla u \cdot \nabla \eta \ \mathrm{d}x, \text{ for all } \eta \in H_0^1(D). \tag{2.49}$$

Here, the function $\mathcal{A}_\omega : D \times \mathbb{R}^2 \to \mathbb{R}^2$ represents the stochastic version of (2.17) and is given by

$$\mathcal{A}_\omega(x, y) = (1 + \delta\xi(\omega))DT(y)\chi_{\Omega_f}(x) + \nu_0 I \chi_{D \setminus \Omega_f}(x). \tag{2.50}$$

2.5.2 Numerical method

In this section we discuss the numerical procedure to solve problem (2.47). We use a generalization of the so called stochastic gradient method which is proposed for shape optimization in [17] to tackle our model problem. The stochastic gradient method is an algorithm for computing solutions to problems of the form

$$\min_x \mathbb{E}[J_\omega(x)].$$

It uses iterates of the form $x_{n+1} = x_n - s_n \nabla J_{\omega_n}(x_n)$ where $\nabla J_{\omega_n}(x_n) \approx \nabla \mathbb{E}[J_\omega(x_n)]$ is called stochastic gradient at x_n for a random sample $\omega_n \in X$ and represents an approximation to $\nabla \mathbb{E}[J_\omega(x_n)]$. A challenging task is the choice of a proper step size s_n that leads to convergence of the method. One possible choice is the so called Robbins-Monro step size given by

$$s_n \geq 0, \quad \sum_{n=0}^{\infty} s_n = \infty, \quad \sum_{n=0}^{\infty} (s_n)^2 < \infty, \tag{2.51}$$

which dates back to [36]. Note that the Armijo rule may not converge without additional variance reduction techniques and therefore should be used with caution for stochastic problems [17].

As seen above, the crucial ingredient of the stochastic gradient method is that we can choose descent directions as the gradient of J_ω for a randomly chosen $\omega \in X$. Therefore it is sufficient to compute the shape derivative for J_ω instead for $\mathbb{E}[J_\omega(x_n)]$. For our problem a descent direction can be computed in the same way as in (2.40), that is for fixed $\omega \in X$ solve the auxiliary problem: find $V \in H_0^1(\Omega^d, \mathbb{R}^2)$ such that

$$b(V, W) = -dJ_\omega(\Omega; W) \text{ for all } W \in H_0^1(\Omega^d, \mathbb{R}^2), \tag{2.52}$$

where $b(\cdot, \cdot)$ is a symmetric and coercive bilinear form. Here, $b(\cdot, \cdot)$ is again chosen to be the linear elasticity bilinear form (2.41).

The procedure to solve the stochastic problem (2.47) is summarized in the following optimization algorithm:

Algorithm 2.2. Initialization: Set $k = 0$, choose initial design Ω_0, set $T_0 = id$;

(1) generate $\omega_k \in X$

(2) solve state equation (2.43) for $\omega = \omega_k$

(3) solve adjoint equation (2.49) for $\omega = \omega_k$

(4) compute descent direction V_k by solving (2.52) for $\omega = \omega_k$;

(5) choose step size $\tau_k > 0$;

(6) set $T_{k+1} = T_k + \tau_k V_k$ and $k \leftarrow k + 1$ and go to (1);

Note that in step (5) we choose a step size according to the Robbins-Monro rule (2.51). The convergence of this method has been proven recently in [18].

2.5.3 Numerical results

The results of applying Algorithm 2.2 to problem (2.47) with $\delta = 0.3$ are shown in Figure 2.4. Note that we use a single sample to generate the stochastic shape gradient at each iteration k. However, in order to calculate objective function values for the

initial and final design we use additional sampling. To be more precise, we approximate $\mathbb{E}[J_\omega(\Omega_k)]$ by the estimate

$$\hat{j}_k := \frac{1}{m} \sum_{\ell=1}^m J_{\omega_{k,\ell}}(\Omega_k) \approx \mathbb{E}[J_\omega(\Omega_k)],$$

with m identically and independently distributed samples $\{\omega_{k,1}, ..., \omega_{k,m}\}$ generated at iteration k. Note that we choose $m = 100$ to calculate $\hat{j}_k$ at the first and last iteration. We let the algorithm iterate 50 times and we could achieve an decrease of the objective function from $\hat{j}_0 = -1.02900047$ to $\hat{j}_{50} = -1.26053252$. Having a closer look on the final design in Figure 2.4 reveals that there are only minor changes in the final rotor design compared to the deterministic case in Figure 2.2. It seems that the randomness in the ferromagnetic material of the motor has only minimal influence on the final rotor design. A deeper investigation shows that the expected torque $\mathbb{E}[J_\omega(\Omega_{final})]$ for the final design Ω_{final} of the deterministic case is $|\mathbb{E}[J_\omega(\Omega_{final})]| \approx 1.25969783$ Nm which is a little lower than $|\hat{j}_{50}|$. This comparison confirms the suspicion that the final design of the stochastic model is slightly more robust. Finally, we mention that the expectation is a risk-neutral measure and other measures may lead to more conservative designs.

2.6 Comparison to parametric shape optimization methods

In this section we want to compare our results to the results of a parametric shape optimization approach applied to model problem (1.24). Using parametric shape optimization one has to define the parameters and dependencies of the geometry which are allowed to vary. In general, without any a priori knowledge of what could be a good design, it is very difficult to decide which geometrical parameters should change. However, thanks to our results of the previous section the parameter setting shown in Figure 2.5 is defined. Here, the parameters $d1$, $d2$ and $d3$ describe the thickness of the ferromagnetic layers whereas $d4$, $d5$ and $d6$ are used to describe the thickness of the non-ferromagnetic layers. Further, the parameters $l1$ to $l4$ and s_1 to $s4$ are used to define the length of and the distance between the ferromagnetic layers respectively. As additional design variable the angle $\beta \in [35°, 55°]$ representing the position of the rotor is chosen. The objectives are to maximize the torque and to minimize the area of ferromagnetic material inside the design domain. Note that this is different to the model problem (1.24) formulated in Chapter 1, where we considered only a fixed rotor position, that is $\beta = 45°$ and did not formulate any constraint to the ferromagnetic area of the design domain. The simulation of the parametric shape optimization problem has been done with the software `JMAG` and the results have been kindly provided by C. Mellak. In order to solve the parametric problem the Multi-objective Genetic

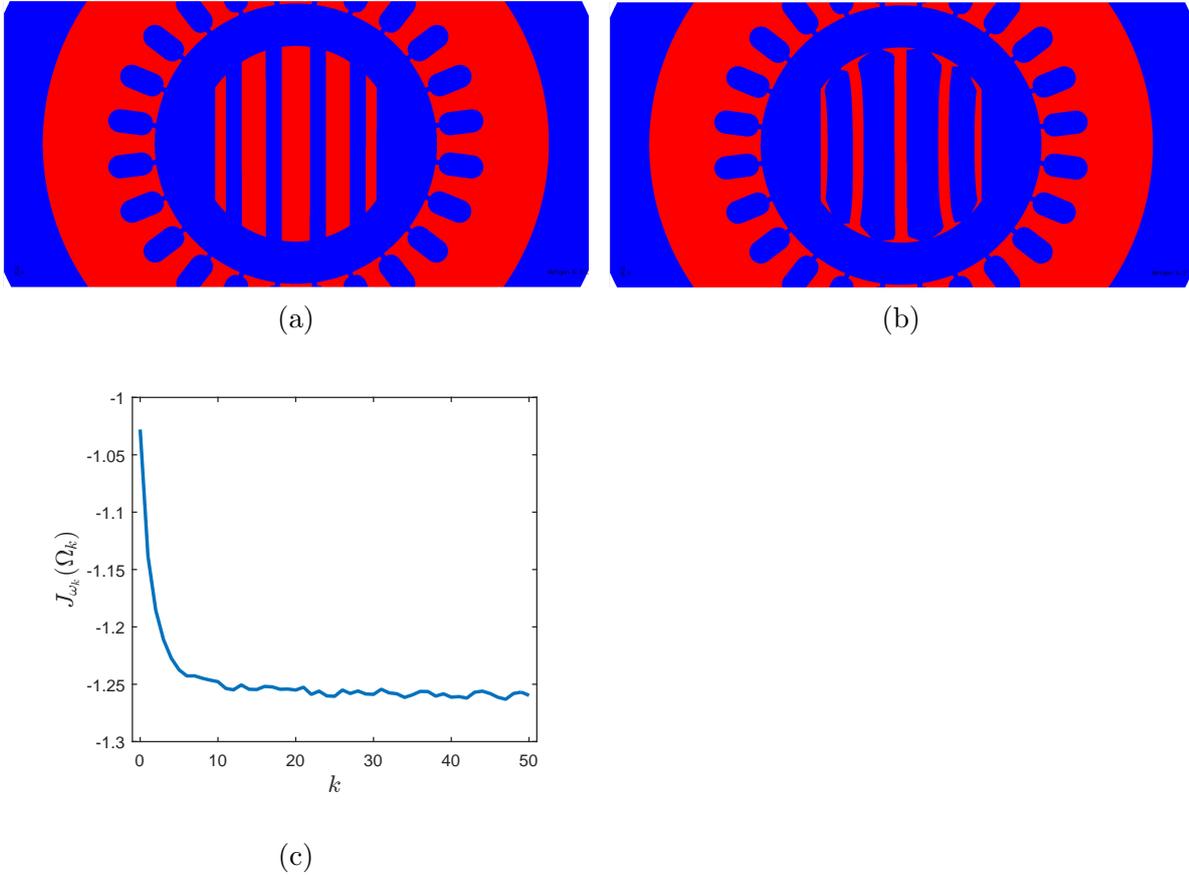(a)   (b)

(c)

Figure 2.4: Application of algorithm 2.2 to problem (2.47). (a) Initial design: $|\hat{j}_0| = 1.02900047$ Nm. (b) Final design: $|\hat{j}_{50}| = 1.26053252$ Nm. (c) Decrease of the objective function.

Algorithm (MOGA) with a population size of 200 and a number of generations of 30 is used. In contrast to our approach this algorithm does not need any sensitivity information but instead creates individuals and selects the fittest ones along every generation. The results are shown in Figure 2.6. The algorithm tested thousands of different designs with respect to the objectives and the final design obtained in case 5761 has a torque of 1.221 Nm where $\beta = 49.15°$. The total computational time was around 16 hours. Compared to the results of the sensitivity based approach in Section 2.4.2, where the total computational time of solving problem (1.24) was around 5 minutes and the final design has a torque of 1.259 Nm, the parametric approach solved by the MOGA performed significantly worse. The difference in the results rely on two aspects: Firstly, due to parametric formulation of the shape optimization problem one reduces the design space from an infinite dimensional to a finite dimensional one resulting in a loss of flexibility in the designs. Further, one has to a priori decide which geometrical parameters should change. Secondly, solving the parametric problem with

a genetic algorithm is responsible for the high computational effort.

Summarizing, even though the parametric problem is posed slightly different here, we can see the large potential of a gradient based shape optimization approach compared to a genetic algorithm.
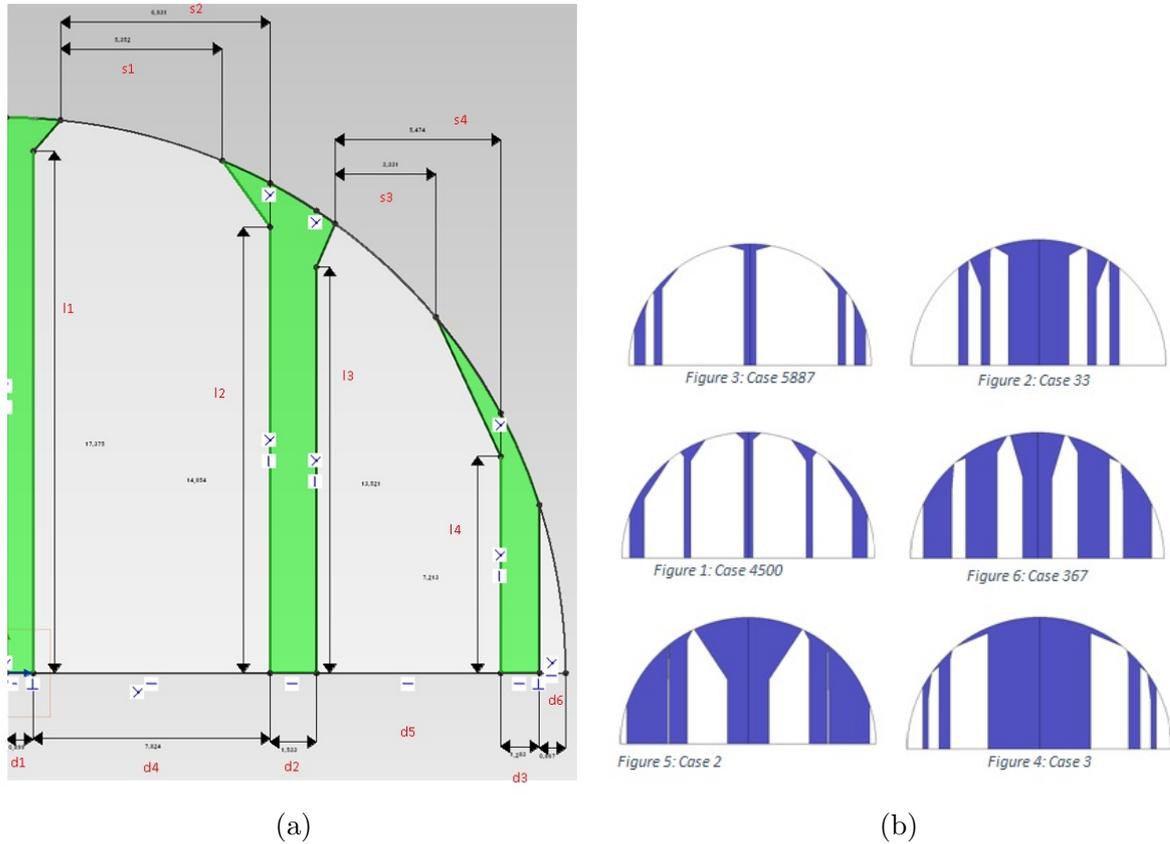

(a)   (b)

Figure 2.5: Parametric setting of the rotor. (a) Chosen geometrical parameters. (b) Some possible rotor designs.

(a)

(b)

(c)
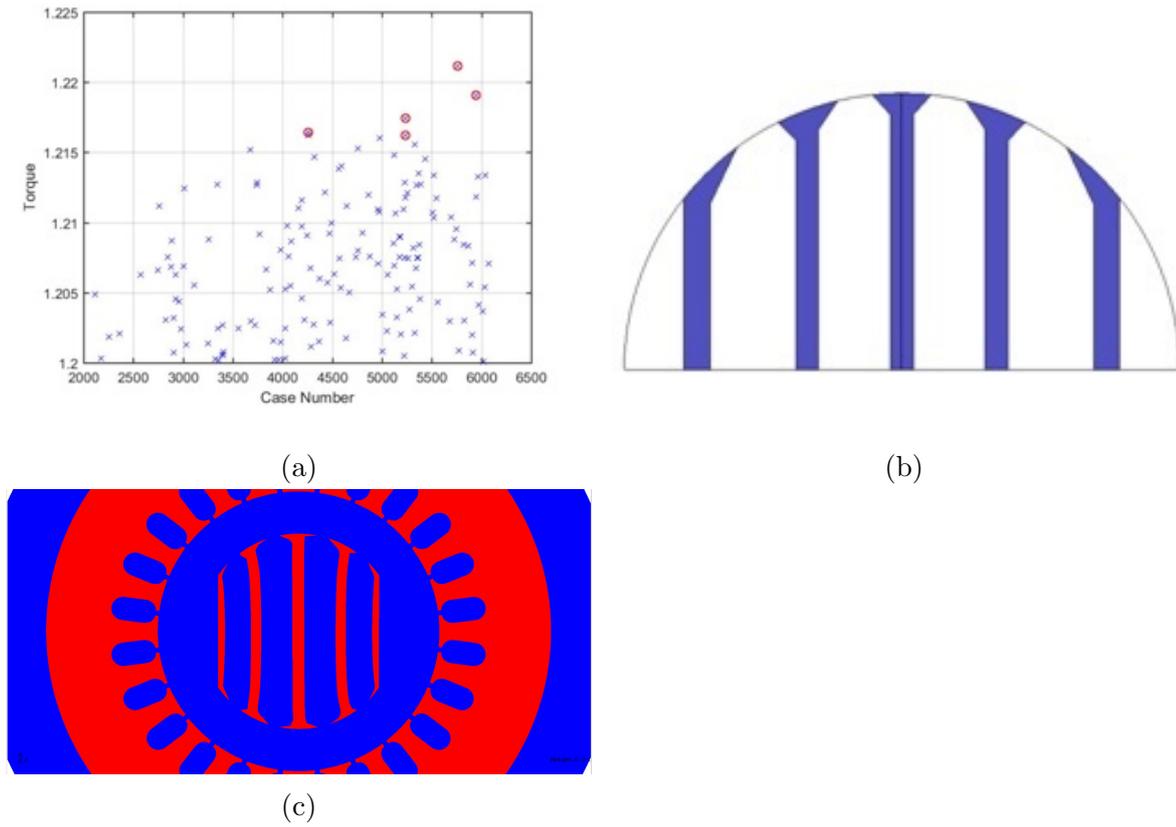
Figure 2.6: (a) Results (Cases) of the genetic algorithm . (b) Final design of the best case. Torque: 1.221 Nm at $\beta = 49.15°$. Computational time: $\sim$ 16h. (c) Final design of the sensitivity based approach applied to (1.24). Torque: 1.259 Nm. Computational time: $\sim$ 5min

3 Topology optimization

The aim of this chapter is to apply a sensitivity based topology optimization method to solve the model problems introduced in Chapter 1. The main ingredient of this approach is the use of the so called topological derivative. The topological derivative measures the sensitivity of a shape functional $J = J(\Omega), \Omega \subset \mathbb{R}^d$ with respect to topological perturbations of the shape Ω. That means it indicates whether an introduction of a hole around a spatial point x_0 would lead to an increase or decrease of the shape functional. Further, this information is used in a level set algorithm which was introduced in [4] to solve the design optimization problems numerically. The strength of this method in comparison to the shape optimization method considered in the previous chapter is that it allows for topological changes in the design domain which may lead to more flexible designs. This chapter is organized as follows: In Section 3.1 we derive the topological derivative for a class of quasi-linear model problems using the technique described in [16] and apply these results to the special case of problem (1.24). Subsequently, in Section 3.2 we focus on the numerical implementation of the design optimization problem. Here, we present the used level set algorithm and briefly discuss the Nitsche-Extended Finite Element Method (Nitsche-XFEM) as discretization method for boundary value problems with unfitted interface. The results for both optimization problems (1.26) and (1.24) are shown in Section 3.3 for different initial designs. Finally, in Section 3.4 we present an extended level set algorithm which is used to solve the stochastic optimization problem (2.47).

3.1 Topological derivative using a Lagrangian framework

In this section we derive the topological derivative for the objective function

$$J(\Omega) = -C_J \int_{\Omega_g} Q(x) \nabla u \cdot \nabla u \; \mathrm{d}x, \tag{3.1a}$$

where $C_J > 0$, $Q \in C(D, \mathbb{R}^{2,2})$, which is subject to the constraint that $u \in H_0^1(D)$ solves

$$\int_D \mathscr{A}_\Omega(x, \nabla u) \cdot \nabla \varphi \; \mathrm{d}x = \int_{\Omega_c} f\varphi \; \mathrm{d}x \text{ for all } \varphi \in H_0^1(D). \tag{3.1b}$$

Here, Ω_c is an open set, $f \in L^2(D)$ and $\mathscr{A}_\Omega : D \times \mathbb{R}^2 \to \mathbb{R}^2$ is a piecewise nonlinear function defined by

$$\mathscr{A}_\Omega(x, y) := \begin{cases} a_1(y) & \text{for } x \in \Omega \\ a_2(y) & \text{for } x \in D \setminus \Omega, \end{cases} \tag{3.2}$$

with functions $a_1, a_2 : \mathbb{R}^2 \to \mathbb{R}^2$ satisfying the following assumptions:

Assumption 3.1 (A). The functions $a_i : \mathbb{R}^2 \to \mathbb{R}^2$, $i = 1, 2$ are differentiable and there are constants $c_1, c_2, c_3 > 0$ such that they satisfy:

(i) $(a_i(x) - a_i(y)) \cdot (x - y) \geq c_1 \|x - y\|^2$, for all $x, y \in \mathbb{R}^2$.

(ii) $\|a_i(x) - a_i(y)\| \leq c_2 \|x - y\|$ for all $x, y \in \mathbb{R}^2$.

(iii) $\|\partial a_i(x) - \partial a_i(y)\| \leq c_3 \|x - y\|$ for all $x, y \in \mathbb{R}^2$.

Note that the design optimization problem of the electric motor (1.24) fits in the framework of quasi-linear problems above. We mention that the boundary value problem (3.1b) models a nonlinear potential equation defined on a domain consisting of two different materials which may exhibit nonlinear behavior. In the special case that $a_1(y) = \hat{\nu}(y)y$ and $a_2(y) = \nu_0 y$ this equation coincides with the magnetostatic boundary value problem in (1.17).
The rest of this section is organized as follows: We start by fixing notations and definitions for the topological setting. In the following part we present a Lagrangian framework which is used to calculate the topological derivative of problem (3.1). Finally, we apply these results to the electric motor design optimization problem (1.24) and derive the topological sensitivities in this special context.

3.1.1 Preliminaries

We fix the setting for the topological derivative. We use the following definition of the topological derivative:

Definition 3.1 ([16]). [Topological derivative] Let $D \subset \mathbb{R}^3$ be an open set and $\Omega \subset D$ an open subset. Let $\omega \subset \mathbb{R}^3$ be open with $0 \in \omega$ and define for $z \in \mathbb{R}^3, \omega_\varepsilon(z) := z + \varepsilon\omega$. The topological derivative of a shape function J at Ω at the point $z \in D \setminus \partial\Omega$ is defined by

$$dJ(\Omega)(z) = \begin{cases} \lim_{\varepsilon \searrow 0} \frac{J(\Omega \setminus \omega_\varepsilon(z)) - J(\Omega)}{|\omega_\varepsilon(z)|} & \text{if } z \in \Omega, \\ \lim_{\varepsilon \searrow 0} \frac{J(\Omega \cup \omega_\varepsilon(z)) - J(\Omega)}{|\omega_\varepsilon(z)|} & \text{if } z \in D \setminus \overline{\Omega}. \end{cases} \tag{3.3}$$

As it can be seen from the definition above the topological derivative for problem (3.1) is in general different for introducing material a_1 in regions where material a_2 is present (case $z \in D \setminus \overline{\Omega}$) and vice versa (cf. Figure 3.1). We will focus on the derivation for the case that $z \in D \setminus \overline{\Omega}$ and mention that the derivation for the case $z \in \Omega$ is

analogous. In the context of magnetostatics we will have to use both sensitivities, i.e. that for introducing air into ferromagnetic material and that for introducing ferromagnetic material into a domain of air to set up a bidirectional optimization algorithm which is capable of both, introducing and removing material at the most favorable positions [13].
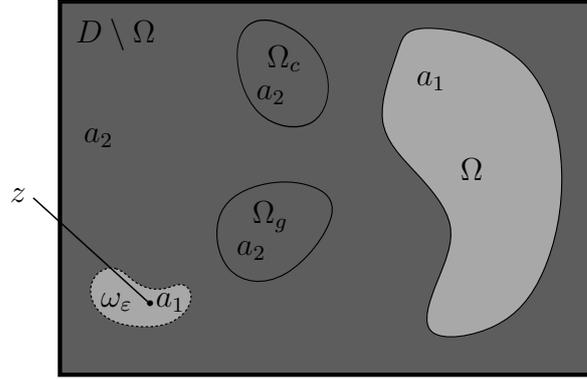

Figure 3.1: Setting for the topological perturbation: Inclusion ω_ε of material a_1 around point $z \in D \setminus \overline{\Omega}$ where material a_2 is present.

We use the following setting for the topological perturbation for model problem (3.1), see Figure 3.1: We fix

- an open and bounded set $\omega \subset \mathbb{R}^2$ with $0 \in \omega$,

- an open set $\Omega \Subset D$ and the inclusion point $z := 0 \in D \setminus \overline{\Omega}$,

- the perturbation $\omega_\varepsilon(z) := z + \varepsilon\omega = \varepsilon\omega$ and $\varepsilon \in [0, \tau]$, where $\tau > 0$ is such that $\omega_\varepsilon(z) \Subset D \setminus \overline{\Omega}$ for all $\varepsilon \in [0, \tau]$,

- the perturbed shape $\Omega_\varepsilon(z) := \Omega \cup \omega_\varepsilon(z)$,

- the transformation $T_\varepsilon(x) := \varepsilon x$, $x \in \mathbb{R}^2$, $\varepsilon \geq 0$.

In order to simplify notation we will use the abbreviations ω_ε for $\omega_\varepsilon(z)$, Ω_ε for $\Omega_\varepsilon(z)$ and x_ε for $T_\varepsilon(x)$.

For the subsequent sensitivity analysis we define for $d \geq 1$ the set $BL(\mathbb{R}^d) := \{u \in H^1_{loc}(\mathbb{R}^d) : \nabla u \in L^2(\mathbb{R}^d)^d\}$ and introduce the Beppo-Levi space as the quotient space $\dot{BL}(\mathbb{R}^d) := BL(\mathbb{R}^d)/\mathbb{R}$, where $/\mathbb{R}$ means that we factor out the constant functions. We denote by $[u]$ the equivalence classes of $\dot{BL}(\mathbb{R}^d)$ and equipped with the norm

$$\|[u]\|_{\dot{BL}(\mathbb{R}^d)} := \|\nabla u\|_{L^2(\mathbb{R}^d)^d}, \quad u \in [u], \tag{3.4}$$

the Beppo-Levi space is a Hilbert space [31]. Moreover $C_c^\infty(\mathbb{R}^d)/\mathbb{R}$ is dense in $\dot{BL}(\mathbb{R}^d)$.

3.1.2 Lagrangian framework

In this section we present an abstract result which will be used to calculate the topological derivative of the previous model problem. This section is based on [16].

Definition 3.2 (parametrised Lagrangian). Let X and Y be vector spaces and $\tau > 0$. A parametrised Lagrangian (or short Lagrangian) is a function

$$(\varepsilon, u, p) \mapsto G(\varepsilon, u, p) : [0, \tau] \times X \times Y \to \mathbb{R},$$

such that $p \mapsto G(\varepsilon, u, p)$ is affine on Y.

Definition 3.3. Let $\varepsilon \in [0, \tau]$. The state equation is defined by: find $u_\varepsilon \in X$, such that

$$\partial_p G(\varepsilon, u_\varepsilon, 0)(\varphi) = 0, \quad \text{for all } \varphi \in Y. \tag{3.5}$$

The set of states is denoted by $E(\varepsilon)$. The adjoint state is defined by: find $p_\varepsilon \in Y$, such that

$$\partial_u G(\varepsilon, u_\epsilon, p_\epsilon)(\varphi) = 0 \quad \text{for all } \varphi \in X. \tag{3.6}$$

The set of adjoint states associated with $(\varepsilon, u_\varepsilon)$ is denoted by $Y(\varepsilon, u_\varepsilon)$.

Definition 3.4 (ℓ-differentiable Lagrangian). Let X and Y be vector spaces and $\tau > 0$. Let $\ell : [0, \tau] \to \mathbb{R}$ be a given function satisfying $\ell(0) = 0$ and $\ell(\varepsilon) > 0$ for $\varepsilon \in (0, \tau]$. A ℓ-differentiable parametrised Lagrangian is a parametrised Lagrangian $G : [0, \tau] \times X \times Y \to \mathbb{R}$, satisfying,

(a) for all $v, w \in X$ and $p \in Y$,

$$s \mapsto G(\varepsilon, v + sw, p) \text{ is continuously differentiable on } [0, 1]. \tag{3.7}$$

(b) for all $u_0 \in E(0)$ and $p_0 \in Y(0, u_0)$ the limit

$$\partial_\ell G(0, u_0, p_0) := \lim_{\varepsilon \searrow 0} \frac{G(\varepsilon, u_0, p_0) - G(0, u_0, p_0)}{\ell(\varepsilon)} \text{ exists.} \tag{3.8}$$

We need the following assumptions.

Assumption 3.2 (A0). We assume that

(i) the set $E(\varepsilon) = \{u_\varepsilon\}$ is a singleton for all $\varepsilon \in [0, \tau]$,

(ii) the adjoint equation for $\varepsilon = 0$, $\partial_u G(0, u_0, p_0)(\varphi) = 0$ for all $\varphi \in X$, admits a unique solution.

Now, let us introduce the function

$$\begin{aligned} g : [0, \tau] &\to \mathbb{R} \\ \varepsilon &\mapsto g(\varepsilon) := G(\varepsilon, u_\varepsilon, 0) \end{aligned} \tag{3.9}$$

The following theorem gives sufficient conditions when the function g is one-sided ℓ-differentiable, that means when the limit

$$d_\ell g(0) := \lim_{\varepsilon \searrow 0} \frac{g(\varepsilon) - g(0)}{\ell(\varepsilon)} \tag{3.10}$$

exists, where $\ell : [0, \tau] \to \mathbb{R}$ is a given function satisfying $\ell(0) = 0$ and $\ell(\varepsilon) > 0$ for $\varepsilon \in (0, \tau]$.

Theorem 3.1. *Let $G : [0, \tau] \times X \times Y \to \mathbb{R}$ be a ℓ-differentiable parametrised Lagrangian satisfying Assumption (A0). Define for $\varepsilon > 0$,*

$$R_1^\varepsilon(u_0, p_0) := \frac{1}{\ell(\varepsilon)} \int_0^1 (\partial_u G(\varepsilon, su_\varepsilon + (1-s)u_0, p_0) - \partial_u G(\varepsilon, u_0, p_0))(u_\varepsilon - u_0) \, \mathrm{d}s \tag{3.11}$$

and

$$R_2^\varepsilon(u_0, p_0) := \frac{1}{\ell(\varepsilon)} (\partial_u G(\varepsilon, u_0, p_0) - \partial_u G(0, u_0, p_0))(u_\varepsilon - u_0). \tag{3.12}$$

If the limits $R_1(u_0, p_0) := \lim_{\varepsilon \searrow 0} R_1^\varepsilon(u_0, p_0)$ and $R_2(u_0, p_0) := \lim_{\varepsilon \searrow 0} R_2^\varepsilon(u_0, p_0)$ exist, then

$$d_\ell g(0) = \partial_\ell G(0, u_0, p_0) + R_1(u_0, p_0) + R_2(u_0, p_0).$$

Proof. By inserting a productive 0 and then using the fundamental theorem of calculus, which is possible due to assumption (3.7), we obtain

$$\begin{aligned}
g(\varepsilon) - g(0) &= G(\varepsilon, u_\varepsilon, p_0) - G(0, u_0, p_0) \\
&= G(\varepsilon, u_\varepsilon, p_0) - G(\varepsilon, u_0, p_0) + G(\varepsilon, u_0, p_0) - G(0, u_0, p_0) \\
&= \int_0^1 \partial_u G(\varepsilon, su_\varepsilon + (1-s)u_0, p_0)(u_\varepsilon - u_0) \, \mathrm{d}s + G(\varepsilon, u_0, p_0) - G(0, u_0, p_0) \\
&= \int_0^1 (\partial_u G(\varepsilon, su_\varepsilon + (1-s)u_0, p_0) - \partial_u G(\varepsilon, u_0, p_0))(u_\varepsilon - u_0) \, \mathrm{d}s \\
&\quad + (\partial_u G(\varepsilon, u_0, p_0) - \partial_u G(0, u_0, p_0))(u_\varepsilon - u_0) \\
&\quad + G(\varepsilon, u_0, p_0) - G(0, u_0, p_0).
\end{aligned}$$

In the last step we used that $\partial_u G(0, u_0, p_0)(\varphi) = 0$ for the special choice $\varphi = u_\varepsilon - u_0 \in X$ due to (A0). Now, dividing by $\ell(\varepsilon)$, using (3.8) and that $R_1(u_0, p_0)$ and $R_2(u_0, p_0)$ exist, we can pass to the limit $\varepsilon \searrow 0$ and arrive at the claimed result. $\qquad\square$

3.1.3 The topological derivative

In this section we apply Theorem 3.1 to the Lagrangian $G : [0, \tau] \times H_0^1(D) \times H_0^1(D) \to \mathbb{R}$ associated to model problem (3.1) which is defined as

$$G(\varepsilon, u, p) := -C_J \int_{\Omega_g} Q(x)\nabla u \cdot \nabla u \, \mathrm{d}x + \int_D \mathscr{A}_{\Omega_\varepsilon}(x, \nabla u) \cdot \nabla p \, \mathrm{d}x - \int_{\Omega_c} fp \, \mathrm{d}x. \tag{3.13}$$

Here, the operator $\mathscr{A}_{\Omega_\varepsilon}$ is defined according to (3.2) with $\Omega_\varepsilon = \Omega \cup \omega_\varepsilon$, that is $\mathscr{A}_{\Omega_\varepsilon}(x,y) := a_1(y)\chi_{\Omega_\varepsilon}(x) + a_2(y)\chi_{D\setminus\Omega_\varepsilon}(x)$. Note that the Lagrangian in (3.13) is considered on the topologically perturbed domain. Moreover with this Lagrangian it holds for the function g defined in (3.9) that $g(\varepsilon) = J(\Omega_\varepsilon)$ and when using $\ell(\varepsilon) = |\omega_\varepsilon|$ we obtain that the derivative in (3.10) corresponds to the topological derivative defined in (3.3). For this reason we always consider $\ell(\varepsilon) = |\omega_\varepsilon|$ from now on.

Analysis of the perturbed state equation

First, we are going to investigate the perturbed state equation, that is find $u_\varepsilon \in H_0^1(D)$ such that

$$\partial_p G(\varepsilon, u_\varepsilon, 0)(\varphi) = 0 \quad \text{for all } \varphi \in H_0^1(D) \tag{3.14}$$

or equivalently $u_\varepsilon \in H_0^1(D)$ satisfies

$$\int_D \mathscr{A}_\varepsilon(x, \nabla u_\varepsilon) \cdot \nabla\varphi \, \mathrm{d}x = \int_{\Omega_c} f\varphi \, \mathrm{d}x \quad \text{for all } \varphi \in H_0^1(D). \tag{3.15}$$

Here, we used the abbreviation $\mathscr{A}_\varepsilon(x,y) := \mathscr{A}_{\Omega_\varepsilon}(x,y)$ for $x, y \in \mathbb{R}^d$.

Lemma 3.2. *Let Assumption A(i) and A(ii) be satisfied. Then, for all $\varepsilon \in [0, \tau]$ there exists a unique solution $u_\varepsilon \in H_0^1(D)$ to the perturbed state equation* (3.15).

Proof. Let $\varepsilon \in [0, \tau]$ and introduce the operator $B_\varepsilon : H_0^1(D) \to (H_0^1(D))^*$ defined by

$$\langle B_\varepsilon\varphi, \psi \rangle := \int_D \mathscr{A}_\varepsilon(x, \nabla\varphi) \cdot \psi \, \mathrm{d}x.$$

Now, properties (i) and (ii) of Assumption A imply that the operator B_ε is Lipschitz continuous and strongly monotone for all measurable $\Omega_\varepsilon \subset D$ by Lemma 1.3. Hence, the perturbed state equation (3.15) has a unique solution by Theorem 1.2. $\qquad\square$

In order to study the behavior of $u_\varepsilon - u_0$ we make the following definition.

Definition 3.5 ([16])**.** We define the variation of the state by

$$K_\varepsilon := \frac{(u_\varepsilon - u_0) \circ T_\varepsilon}{\varepsilon} \in H_0^1(\varepsilon^{-1}D), \quad \varepsilon > 0. \tag{3.16}$$

Note that by extending u_ε and u_0 by zero outside of D we can view K_ε as an element of $\dot{B}L(\mathbb{R}^d)$. We have the following estimate:

Lemma 3.3 ([16])**.** *Let Assumption A(i), (ii) be satisfied. Then, there exists a constant $C > 0$, such that for all small $\varepsilon > 0$,*

$$\|u_\varepsilon - u_0\|_{H^1(D)} \leq C\varepsilon. \tag{3.17}$$

A direct consequence of Lemma 3.3 is the following estimate for ∇K_ε.

Corollary 3.4. *Let Assumption A(i), (ii) be satisfied. There exists a constant $C > 0$, such that for all small $\varepsilon > 0$*

$$\|\nabla K_\varepsilon\|_{L^2(\mathbb{R}^2)^2} \leq C. \tag{3.18}$$

Proof. From Lemma 3.3 we have that

$$\int_D |\nabla(u_\varepsilon - u_0)|^2 \, \mathrm{d}y + \int_D (u_\varepsilon - u_0)^2 \, \mathrm{d}y \leq C^2 \varepsilon^2. \tag{3.19}$$

Using the change of variables $y = T_\varepsilon(x)$ and then noting that $(\nabla f) \circ T_\varepsilon = \varepsilon^{-1}\nabla(f \circ T_\varepsilon)$ for $f \in H^1(D)$ we get

$$\int_{\varepsilon^{-1}D} |(\nabla(u_\varepsilon - u_0)) \circ T_\varepsilon|^2 \, \varepsilon^2 \, \mathrm{d}x + \int_{\varepsilon^{-1}D} ((u_\varepsilon - u_0) \circ T_\varepsilon)^2 \, \varepsilon^2 \, \mathrm{d}x \leq C^2\varepsilon^2$$

$$\Longleftrightarrow \int_{\varepsilon^{-1}D} \left|\frac{1}{\varepsilon}\nabla\left((u_\varepsilon - u_0) \circ T_\varepsilon\right)\right|^2 \varepsilon^2 \, \mathrm{d}x + \int_{\varepsilon^{-1}D} \left(\varepsilon\frac{(u_\varepsilon - u_0) \circ T_\varepsilon}{\varepsilon}\right)^2 \varepsilon^2 \, \mathrm{d}x \leq C^2\varepsilon^2$$

$$\Longleftrightarrow \int_{\varepsilon^{-1}D} |\nabla K_\varepsilon|^2 \, \mathrm{d}x + \int_{\varepsilon^{-1}D} (\varepsilon K_\varepsilon)^2 \, \mathrm{d}x \leq C^2$$

Now, extending K_ε by zero outside of $\varepsilon^{-1}D$ we may conclude

$$\|\nabla K_\varepsilon\|_{L^2(\mathbb{R}^2)^2} \leq C.$$

$\square$

The crucial result in our analysis is the following theorem:

Theorem 3.5 ([16])**.** *Let Assumption A(i) and A(ii) be satisfied. Then,*

(a) there exists a unique solution $K \in \dot{B}L(\mathbb{R}^d)$ to

$$\int_{\mathbb{R}^d} (\mathscr{A}_\omega(x, \nabla K + U_0) - \mathscr{A}_\omega(x, U_0)) \cdot \nabla\varphi \, \mathrm{d}x$$
$$= -\int_\omega (a_1(U_0) - a_2(U_0)) \cdot \nabla\varphi \, \mathrm{d}x \quad \text{for all } \varphi \in BL(\mathbb{R}^d), \tag{3.20}$$

where $U_0 := \nabla u_0(z)$ and $\mathscr{A}_\omega(x, y) := a_1(y)\chi_\omega(x) + a_2(y)\chi_{\mathbb{R}^d \setminus \omega}(x)$,

(b) we have $\nabla K_\varepsilon \to \nabla K$ strongly in $L^2(\mathbb{R}^d)^d$ as $\varepsilon \searrow 0$.

Adjoint equation

The adjoint equation associated to the Lagrangian in (3.13) is given as the solution $p \in H_0^1(D)$ to

$$\int_D \partial_u \mathscr{A}_\Omega(x, \nabla u)(\nabla \varphi) \cdot \nabla p \, \mathrm{d}x = 2C_J \int_{\Omega_g} Q(x) \nabla u \cdot \nabla \varphi \, \mathrm{d}x \quad \text{for all } \varphi \in H_0^1(D). \quad (3.21)$$

Lemma 3.6. *Let Assumption A hold and let $u \in H_0^1(D)$ be the solution to the unperturbed state equation* (3.1b)*. Then there exists a unique solution $p \in H_0^1(D)$ to the adjoint equation* (3.21)*.*

Proof. The proof is similar to the proof of Lemma 2.7 so we only point out the main differences. Introduce for fixed $u \in H_0^1(D)$ the bilinear form

$$a'(u; \cdot, \cdot) : H_0^1(D) \times H_0^1(D) \to \mathbb{R} \quad (3.22)$$

$$(\varphi, \psi) \mapsto \int_D \partial_u \mathscr{A}_\Omega(x, \nabla u)(\nabla \varphi) \cdot \nabla \psi \, \mathrm{d}x \quad (3.23)$$

We show that a' is bounded and coercive. Note that from Assumption A(ii) we get

$$\|\partial a_i(x)v\| = \lim_{t \searrow 0} \|(a_i(x + tv) - a_i(x))\|/t \le c_2\|v\|, \quad (3.24)$$

for $i = 1, 2$ and all $x, v \in \mathbb{R}^d$. Using this and Hölder's inequality we get the boundedness of a':

$$\left| \int_D \partial_u \mathscr{A}_\Omega(x, \nabla u)(\nabla \varphi) \cdot \nabla \psi \, \mathrm{d}x \right| \le c_2 \int_D |\nabla \varphi| \, |\nabla \psi| \, \mathrm{d}x \le c_2 \|\varphi\|_{H^1(D)} \|\psi\|_{H^1(D)}.$$

In order to show that a' is coercive we use the following estimate due to Assumption A(i):

$$\partial a_i(x)v \cdot v = \lim_{t \searrow 0} \frac{1}{t^2}(a_i(x + tv) - a_i(x)) \cdot tv \ge c_1 \|v\|^2 \quad (3.25)$$

for $i = 1, 2$ and all $x, v \in \mathbb{R}^d$. Thus, the equation (3.21) has a unique solution according to the lemma of Lax-Milgram. $\qquad \square$

Main result of section 3.1.3

Theorem 3.7 (Main result)**.** *Let Assumption A be statisfied. Let $\Omega \subset D$ open and u_0 be the solution of the state equation* (3.1b) *and p_0 the solution to the adjoint equation* (3.21)*. Let $z \in D \setminus \overline{\Omega}$ such that $z \notin (\Omega_g \cup \Omega_c)$ and suppose that $u_0 \in C^{1,\alpha}(\overline{B_\delta(z)})$ and $p_0 \in C^1(\overline{B_\delta(z)})$ for some $\delta > 0$ and $0 < \alpha < 1$. Moreover assume that $\nabla p_0 \in L^\infty(D)^2$. Then the Lagrangian G given by* (3.13) *satisfies the assumptions of Theorem 3.1 and hence the topological derivative at $z \in D \setminus \overline{\Omega}$ is given by*

$$dJ(\Omega)(z) = \partial_\ell G(0, u_0, p_0) + R_1(u_0, p_0) + R_2(u_0, p_0). \quad (3.26)$$

Furthermore, we have

$$\partial_\ell G(0, u_0, p_0) = (a_1(U_0) - a_2(U_0)) \cdot P_0 \tag{3.27}$$

and

$$R_1(u_0, p_0) = \frac{1}{|\omega|} \left(\int_{\mathbb{R}^2} [\mathscr{A}_\omega(x, \nabla K + U_0) - \mathscr{A}_\omega(x, U_0) - \partial_u \mathscr{A}_\omega(x, U_0)(\nabla K)] \cdot P_0 \right) \tag{3.28}$$

and

$$R_2(u_0, p_0) = \frac{1}{|\omega|} \int_\omega [\partial_u a_1(U_0) - \partial_u a_2(U_0)](\nabla K) \cdot P_0 \; \mathrm{d}x, \tag{3.29}$$

where $U_0 := \nabla u_0(z)$, $P_0 := \nabla p_0(z)$ and $\mathscr{A}_\omega(x, y) := a_1(y)\chi_\omega(x) + a_2(y)\chi_{\mathbb{R}^d \setminus \omega}(x)$. Here, $K \in \dot{BL}(\mathbb{R}^d)$ is the unique solution to (3.20)

Proof. Without loss of generality we assume that $z = 0$. We apply Theorem 3.1 to the Lagrangian (3.13). In view of Lemma 3.2 and Lemma 3.6 we already have that hypothesis (A0) is satisfied. It remains to check that the limits $R_1(u_0, p_0)$ and $R_2(u_0, p_0)$ exist. First, a computation shows that

$$\partial_u G(\varepsilon, su_\varepsilon + (1-s)u_0, p_0)(\varphi) = -2C_J \int_{\Omega_g} Q(x)\nabla(su_\varepsilon + (1-s)u_0) \cdot \nabla\varphi \; \mathrm{d}x$$

$$+ \int_D \partial_u \mathscr{A}_\varepsilon(x, \nabla(su_\varepsilon + (1-s)u_0))(\nabla\varphi) \cdot \nabla p_0 \; \mathrm{d}x$$

and

$$\partial_u G(\varepsilon, u_0, p_0) = -2C_J \int_{\Omega_g} Q(x)\nabla u_0 \cdot \nabla\varphi \; \mathrm{d}x + \int_D \partial_u \mathscr{A}_\varepsilon(x, \nabla u_0)(\nabla\varphi) \cdot \nabla p_0 \; \mathrm{d}x$$

Plugging this into the definition of R_1, setting $\varphi := u_\varepsilon - u_0$ and then using the change of variables $T_\varepsilon(x) = \varepsilon x$ and extending u_ε, u_0, p_0 by zero outside of $\varepsilon^{-1}D$ yields:

$$R_1^\varepsilon(u_0, p_0)$$

$$= \frac{1}{\ell(\varepsilon)} \int_0^1 \int_D (\partial_u \mathscr{A}_\varepsilon(x, \nabla(su_\varepsilon + (1-s)u_0)) - \partial_u \mathscr{A}_\varepsilon(x, \nabla u_0))(\nabla(u_\varepsilon - u_0)) \cdot \nabla p_0 \mathrm{d}x \mathrm{d}s$$

$$- \frac{1}{\ell(\varepsilon)} C_J \int_{\Omega_g} Q(x)\nabla(u_\varepsilon - u_0) \cdot \nabla(u_\varepsilon - u_0) \; \mathrm{d}x$$

$$= \frac{1}{|\omega|} \underbrace{\int_0^1 \int_{\mathbb{R}^2} (\partial_u \mathscr{A}_\omega(x, s\nabla K_\varepsilon + \nabla u_0(x_\varepsilon)) - \partial_u \mathscr{A}_\omega(x, \nabla u_0(x_\varepsilon)))(\nabla K_\varepsilon) \cdot \nabla p_0(x_\varepsilon) \; \mathrm{d}x \mathrm{d}s}_{=:I_\varepsilon}$$

$$- \frac{1}{|\omega|} C_J \underbrace{\int_{\varepsilon^{-1}\Omega_g} Q(x_\varepsilon)\nabla K_\varepsilon \cdot \nabla K_\varepsilon \mathrm{d}x}_{=:II_\varepsilon}$$

$$\to \frac{1}{|\omega|} \int_0^1 \int_{\mathbb{R}^2} (\partial_u \mathscr{A}_\omega(x, s\nabla K + U_0) - \partial_u \mathscr{A}_\omega(x, U_0))(\nabla K) \cdot \nabla P_0 \; \mathrm{d}x \mathrm{d}s \tag{3.30}$$

Thanks to Theorem 3.5 we have that $\nabla K_\varepsilon \to \nabla K$ strongly in $L^2(\mathbb{R}^2)^2$ as $\varepsilon \searrow 0$ and since $\varepsilon^{-1}\Omega_g$ goes to 'infinity' because $z \notin \Omega_g$ it follows $II_\varepsilon \to 0$ as $\varepsilon \searrow 0$. In order to see convergence of the first term I_ε, we may write

$$
I_\varepsilon = \int_0^1 \int_{\mathbb{R}^2} (\partial_u \mathscr{A}_\omega(x, s\nabla K_\varepsilon + \nabla u_0(x_\varepsilon)) - \partial_u \mathscr{A}_\omega(x, s\nabla K + \nabla u_0(x_\varepsilon)))(\nabla K_\varepsilon) \cdot \nabla p_0(x_\varepsilon) \, \mathrm{d}x\mathrm{d}s
$$
$$
+ \int_0^1 \int_{\mathbb{R}^2} (\partial_u \mathscr{A}_\omega(x, s\nabla K + \nabla u_0(x_\varepsilon)) - \partial_u \mathscr{A}_\omega(x, \nabla u_0(x_\varepsilon)))(\nabla (K_\varepsilon - K)) \cdot \nabla p_0(x_\varepsilon) \, \mathrm{d}x\mathrm{d}s
$$
$$
+ \int_0^1 \int_{\mathbb{R}^2} (\partial_u \mathscr{A}_\omega(x, s\nabla K + \nabla u_0(x_\varepsilon)) - \partial_u \mathscr{A}_\omega(x, \nabla u_0(x_\varepsilon)))(\nabla K) \cdot \nabla p_0(x_\varepsilon) \, \mathrm{d}x\mathrm{d}s.
$$

Using Assumption A(iii) and that $\nabla p_0 \in L^\infty(D)^2$ we see that the absolute value of the first and second term can be bounded by

$$
c_3 \|\nabla p_0\|_{L^\infty(D)^2} \|\nabla (K_\varepsilon - K)\|_{L^2(\mathbb{R}^2)^2} \|\nabla K_\varepsilon\|_{L^2(\mathbb{R}^2)^2}
$$

and

$$
c_3 \|\nabla p_0\|_{L^\infty(D)^2} \|\nabla (K_\varepsilon - K)\|_{L^2(\mathbb{R}^2)^2} \|\nabla K\|_{L^2(\mathbb{R}^2)^2}
$$

respectively. Hence, using the boundedness of ∇K_ε due to Corollary 3.4 and that $\nabla K_\varepsilon \to \nabla K$ in $L^2(\mathbb{R}^2)^2$ as $\varepsilon \searrow 0$ both terms disappear in the limit. It remains to show convergence of the third term. Note that using the continuity of $\partial_u a_1, \partial_u a_2, \nabla u_0, \nabla p_0$ we get pointwise convergence of the integrand of the third term. Further, using Assumption A(iii) the integrand is bounded by an L^1 function, i.e.

$$
|(\partial_u \mathscr{A}_\omega(x, s\nabla K(x) + \nabla u_0(x_\varepsilon)) - \partial_u \mathscr{A}_\omega(x, \nabla u_0(x_\varepsilon)))(\nabla K(x)) \cdot \nabla p_0(x_\varepsilon)|
$$
$$
\leq s \|\nabla p_0\|_{L^\infty(D)^2} \|\nabla K(x)\|^2 \text{ for a.e. } x \in \mathbb{R}^2.
$$

Hence, we can apply the theorem of Lebesgue to see that I_ε converges to the limit in (3.30). Finally, using the fundamental theorem we obtain the expression in (3.28). Now, we address the computation of $R_2(u_0, p_0)$. Using the definition $\ell(\varepsilon) = |\omega_\varepsilon| = \varepsilon^2 |\omega|$ and the change of variables T_ε we may write

$$
R_2^\varepsilon(u_0, p_0) = \frac{1}{|\omega_\varepsilon|} (\partial_u G(\varepsilon, u_0, p_0) - \partial_u G(0, u_0, p_0))(u_\varepsilon - u_0)
$$
$$
= \frac{1}{|\omega_\varepsilon|} \int_D (\partial_u \mathscr{A}_\varepsilon(x, \nabla u_0) - \partial_u \mathscr{A}_\Omega(x, \nabla u_0))(\nabla (u_\varepsilon - u_0)) \cdot \nabla p_0 \, \mathrm{d}x
$$
$$
= \frac{1}{|\omega_\varepsilon|} \int_{\omega_\varepsilon} (\partial_u a_1(\nabla u_0) - \partial_u a_2(\nabla u_0))(\nabla (u_\varepsilon - u_0)) \cdot \nabla p_0 \, \mathrm{d}x
$$
$$
= \frac{1}{|\omega|} \int_\omega (\partial_u a_1(\nabla u_0(x_\varepsilon)) - \partial_u a_2(\nabla u_0(x_\varepsilon)))(\nabla K_\varepsilon) \cdot \nabla p_0(x_\varepsilon) \, \mathrm{d}x
$$
$$
\to \frac{1}{|\omega|} \int_\omega (\partial_u a_1(U_0) - \partial_u a_2(U_0))(\nabla K) \cdot P_0 \, \mathrm{d}x
$$

The convergence in the last step as $\varepsilon \searrow 0$ can be seen as follows: By the continuity of ∇u_0, ∇p_0 in z, the continuity of $\partial_u a_1, \partial_u a_2$ and $\nabla K_\varepsilon \to \nabla K$ in $L^2(\mathbb{R}^2)^2$ we get pointwise convergence of the above integrand. Furthermore, using (3.24) and the boundedness of ∇K_ε due to Corollary 3.4, that is $\|\nabla K_\varepsilon\|_{L^2(\mathbb{R}^2)} \leq C$ for all $\varepsilon > 0$, the integrand can be bounded by an L^1 function, i.e.

$$|(\partial_u a_1(\nabla u_0(x_\varepsilon)) - \partial_u a_2(\nabla u_0(x_\varepsilon)))(\nabla K_\varepsilon) \cdot \nabla p_0(x_\varepsilon)|$$
$$\leq \tilde{C}\|\nabla p_0\|_{L^\infty(D)^2} \text{ for a.e. } x \in \mathbb{R}^2.$$

Hence, we may apply again the theorem of Lebesgue and obtain the desired limit. $\qquad\square$

Remark 3.1 ([16, 15]). Let us make a few remarks.

- In the proof we assume without loss of generality that $z = 0$. In the general case this situation can be achieved by a simple change of coordinate system.

- The topological derivative is computed for the case $z \in D \setminus \overline{\Omega}$. However, the same proof can be applied in the case that $z \in \Omega$ and $z \notin (\Omega_g \cup \Omega_c)$. In this situation the formula for the topological derivative is obtained by switching the roles of a_1 and a_2 in the theorem above and in particular in the definition of $\mathscr{A}_\omega$.

- The case $z \in \Omega_g$ can be dealt with in a similar manner. More precisely, the derivation of [16] shows that if $z \in \Omega_g$ then an additional term $-C_J \int_{\mathbb{R}^2} Q\nabla K \cdot \nabla K \, \mathrm{d}x$ in the topological derivative $dJ(\Omega)(z)$ appears. The case $z \in \Omega_c$ has to be treated separately since in this case the right hand side of (3.1b) becomes domain dependent.

Remark 3.2. As it was shown in [16] the formula of the topological derivative coincides with the formula obtained in [5, Thm. 2 and Thm. 3] for the respective special cases. More precisely, introducing the problem defining the variation of the adjoint state $Q \in \dot{B}L(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2} \partial_u \mathscr{A}_\omega(x, U_0)(\nabla \varphi) \cdot \nabla Q \, \mathrm{d}x = -\int_\omega (\partial_u a_1(U_0) - \partial_u a_2(U_0))(\nabla \varphi) \cdot P_0 \mathrm{d}x \qquad (3.31)$$

for all $\varphi \in BL(\mathbb{R}^2)$ the topological derivative can be written as

$$dJ(\Omega)(z) = \frac{1}{|\omega|}\left[(a_1(U_0) - a_2(U_0)) \cdot \int_\omega P_0 + \nabla Q \, \mathrm{d}x \right.$$
$$\left. + \int_{\mathbb{R}^2} (\mathscr{A}_\omega(x, \nabla K + U_0) - \mathscr{A}_\omega(x, U_0) - \partial_u \mathscr{A}_\omega(x, U_0)(\nabla K)) \cdot (P_0 + \nabla Q) \, \mathrm{d}x \right],$$
$$(3.32)$$

which is up to a scaling by $1/|\omega|$ the same formula as obtained in [5]. The different scaling comes from a different definition of topological derivative used in this work.

Remark 3.3 ([16]). Having a closer look on equation (3.31) it can be seen that ∇Q depends linearly on P_0. Thus, there exists a matrix $\mathscr{M} = \mathscr{M}(\omega, \partial a_1(U_0), \partial a_2(U_0))$ which is related to the concept of polarization matrices such that

$$\int_\omega \nabla Q \mathrm{d}x = \mathscr{M} P_0.$$

In the next section we will exploit this fact to simplify the topological derivative for the setting of 2D magnetostatics of our electric motor design problem.

3.1.4 Application to synchronous reluctance machine

In this section we want to show that the electric motor design optimization problem in (1.24) applies to the setting considered in the previous section. Moreover we are going to specify the respective topological sensitivities, i.e. that for introducing an inclusion of ferromagnetic material inside an area that is occupied with air which we denote by $\mathcal{T}^{air \to f}$ and the other way round which we denote by $\mathcal{T}^{f \to air}$. Note that it is important the have access to both sensitivities in order to employ suitable optimization algorithms. Recall the geometric model of the electric motor which is depicted in Figure 1.1. We are interested in the topological derivative on the design domain $\Omega^d \subset D$, where D denotes the whole computational domain. The ferromagnetic part of the design domain is denoted by Ω_f^d and corresponds to the ferromagnetic part of the motor restricted to the design domain, i.e. $\Omega_f^d = \Omega_f \cap \Omega^d$. The area of the design domain which is occupied with air is given by $\Omega_{air}^d = (D \setminus \Omega_f) \cap \Omega^d$. We consider the cost function in (1.23), i.e.

$$J(\Omega) = -C_J \int_{\Omega_g} Q(x) \nabla u_\Omega \cdot \nabla u_\Omega \ \mathrm{d}x,$$

where C_J is an abbreviation for the constant prefactor, that is $C_J = L/(\mu_0(r_s - r_r))$. In the notation of this section we define

$$\begin{aligned} a_1(y) &:= \hat{\nu}(|y|)y \\ a_2(y) &:= \nu_0 y \\ f &:= J_3 \end{aligned} \tag{3.33}$$

where $y \in \mathbb{R}^2$, such that boundary value problem (3.1b) becomes the 2D magnetostatic problem (1.17). Note that the impressed current density J_3 is supported only in the coil areas Ω_c and we have that $\Omega^d \cap \Omega_c = \emptyset$. In order to be able to apply Theorem 3.7 we have to show that Assumption 3.1 is fulfilled for a_1, a_2 as chosen in (3.33). Therefore, we require the magnetic reluctivity $\hat{\nu}$ to satisfy an additional smoothness assumption.

Assumption 3.3. We assume that $\hat{\nu} \in C^2(\mathbb{R}_0^+)$, $\hat{\nu}'(0) = 0$ and that there is a constant $c > 0$ such that for all $s \in \mathbb{R}_0^+$, $\hat{\nu}'(s) \leq c$ and $\hat{\nu}''(s) \leq c$.

Note that in practice the magnetic reluctivity is obtained by interpolation from measured data. For this reason the smoothness assumption above is justified [33, 15]. We are now in position to show the following:

Lemma 3.8. *Let $\hat{\nu}$ be a reluctivity according to a B-H-curve satisfying Assumption 1.1 and Assumption 3.3. Define $a_1(y) := \hat{\nu}(|y|)y$ and $a_2(y) := \nu_0 y$ for $y \in \mathbb{R}^2$. Then Assumption 3.1 is satisfied.*

Proof. All properties of Assumption 3.1 are fulfilled for the linear function a_2. For a_1, items (i) and (ii) of Assumption 3.1 follow from Lemma 1.1. Furthermore in [13, Lemma 4.7] is shown that the smoothness assumptions on $\hat{\nu}$ imply that a_1 is twice continuously differentiable which is sufficient for Assumption 3.1 (iii). $\qquad\square$

Hence, we can apply Theorem 3.7 with $\Omega = \Omega_f$ and the topological derivative in the case $z \in \Omega_{air}^d$ reads

$$
\mathcal{T}^{air \to f}(z) := dJ(\Omega)(z) = \frac{1}{|\omega|} \left[(\hat{\nu}(|U_0|)U_0 - \nu_0 U_0) \cdot \int_\omega (P_0 + \nabla Q)\mathrm{d}x \right.
$$
$$
\left. + \int_{\mathbb{R}^2} (T(\nabla K + U_0) - T(U_0) - DT(U_0)\nabla K)\chi_\omega \cdot (P_0 + \nabla Q)\mathrm{d}x \right],
$$
$$(3.34)$$

where T is defined in (2.9) and DT is its Jacobian given in (2.10). In view of Remark 3.3, we have that the mapping

$$
P_0 \mapsto (\hat{\nu}(|U_0|) - \nu_0) \int_\omega (P_0 + \nabla Q)\mathrm{d}x
$$

is linear from $\mathbb{R}^2$ to $\mathbb{R}^2$. Thus, there exists a matrix $\mathcal{M} = \mathcal{M}(\omega, DT(U_0), \nu_0 I)$ such that

$$
(\hat{\nu}(|U_0|) - \nu_0) \int_\omega (P_0 + \nabla Q)\mathrm{d}x = \mathcal{M}P_0.
$$

In [5, Theorem 6.9] an explicit formula for the polarization matrix $\mathcal{M}$ has been computed for the case that $\omega = B(0,1)$ is the unit disk in $\mathbb{R}^2$. The matrix reads

$$
\mathcal{M} = 2|\omega|\nu_0 R \begin{pmatrix} \frac{\lambda_1 - \nu_0}{\lambda_2 + \nu_0} & 0 \\ 0 & \frac{\lambda_1 - \nu_0}{\lambda_1 + \nu_0} \end{pmatrix} R^T,
$$
$$(3.35)$$

where $\lambda_1 = \hat{\nu}(|U_0|)$, $\lambda_2 = \hat{\nu}(|U_0|) + \hat{\nu}'(|U_0|)|U_0|$ and R denotes the rotation matrix around the angle between U_0 and the x-axis such that

$$
U_0 = R \begin{pmatrix} |U_0| \\ 0 \end{pmatrix}.
$$

Hence, the topological derivative in (3.34) simplifies to

$$
\mathcal{T}^{air \to f}(z) = \frac{1}{|\omega|} \left[U_0^T \mathcal{M} P_0 \right.
$$
$$
\left. + \int_{\mathbb{R}^2} (T(\nabla K + U_0) - T(U_0) - DT(U_0)\nabla K)\chi_\omega \cdot (P_0 + \nabla Q)\mathrm{d}x \right]. \tag{3.36}
$$

The sensitivity for the case $z \in \Omega_f^d$ (i.e. that for introducing an inclusion of air inside an area of the design domain that is occupied with ferromagnetic material) can be obtained by switching the roles of a_1 and a_2 in Theorem 3.7 (cf. Remark 3.1) and thus reads

$$
\mathcal{T}^{f \to air}(z) = \frac{1}{|\omega|} \left[(\nu_0 U_0 - \hat{\nu}(|U_0|)U_0) \cdot \int_\omega (P_0 + \nabla Q^{(2)})\mathrm{d}x \right.
$$
$$
\left. + \int_{\mathbb{R}^2} (T(\nabla K + U_0) - T(U_0) - DT(U_0)\nabla K)\chi_{\mathbb{R}^2 \setminus \omega} \cdot (P_0 + \nabla Q^{(2)})\mathrm{d}x \right]. \tag{3.37}
$$

Here, T, DT are given according to (2.9),(2.10) and $Q^{(2)}$ is the solution to (3.31) with respect to switched roles of a_1 and a_2. Similar as before, the mapping

$$
P_0 \mapsto (\nu_0 - \hat{\nu}(|U_0|)) \int_\omega (P_0 + \nabla Q^{(2)})\mathrm{d}x
$$

is linear from $\mathbb{R}^2$ to $\mathbb{R}^2$ by virtue of Remark 3.3 and thus there exists a matrix $\mathcal{M}^{(2)} = \mathcal{M}^{(2)}(\omega, \nu_0 I, DT(U_0))$ such that

$$
(\nu_0 - \hat{\nu}(|U_0|)) \int_\omega (P_0 + \nabla Q^{(2)})\mathrm{d}x = \mathcal{M}^{(2)} P_0.
$$

Assuming again that $\omega = B(0,1)$ is the unit disk in $\mathbb{R}^2$ the topological derivative in (3.37) can be written as

$$
\mathcal{T}^{f \to air}(z) = \frac{1}{|\omega|} \left[U_0^T \mathcal{M}^{(2)} P_0 \right.
$$
$$
\left. + \int_{\mathbb{R}^2} (T(\nabla K + U_0) - T(U_0) - DT(U_0)\nabla K)\chi_{\mathbb{R}^2 \setminus \omega} \cdot (P_0 + \nabla Q^{(2)})\mathrm{d}x \right], \tag{3.38}
$$

where the polarization matrix $\mathcal{M}^{(2)}$ is given by (see [5, Theorem 6.6]):

$$
\mathcal{M}^{(2)} = (\nu_0 - \lambda_1)\,|\omega|\,R \begin{pmatrix} \frac{\lambda_2 + \sqrt{\lambda_1 \lambda_2}}{\nu_0 + \sqrt{\lambda_1 \lambda_2}} & 0 \\ 0 & \frac{\lambda_1 + \sqrt{\lambda_1 \lambda_2}}{\nu_0 + \sqrt{\lambda_1 \lambda_2}} \end{pmatrix} R^T. \tag{3.39}
$$

Here, λ_1, λ_2 and R are given as in (3.35).

3.2 Computational aspects

In this section we are going to describe the level set method introduced in [4] which is based on the topological derivative. Here, the design domain is represented implicitly by means of a level set function which allows to deal with topological changes easily. However, we have to think about appropriate discretization methods for the boundary value problems involved in the computation of the topological derivative as we will be concerned with unfitted interface problems. A suitable method to handle this issue is the Extended Finite Element Method (XFEM) which we will briefly discuss in the subsequent section.

3.2.1 A level set algorithm

Recall from the notations of the computational domain in Section 1.2 that Ω denotes the part of the design domain Ω^d which is currently occupied with ferromagnetic material. The current design of the rotor is represented by means of a level set function $\psi : \Omega^d \to \mathbb{R}$ which is positive in the ferromagnetic subdomain and negative in the air subdomain. The interface between these subdomains is represented by the zero level set of ψ. So, we have

$$\begin{aligned}
\psi(x) > 0 &\iff x \in \Omega, \\
\psi(x) < 0 &\iff x \in \Omega^d \setminus \overline{\Omega}, \\
\psi(x) = 0 &\iff x \in \partial\Omega.
\end{aligned} \tag{3.40}$$

The evolution of this level set function is guided by the generalized topological derivative, which, for a given level set function ψ, is defined as

$$G_\psi(x) := \begin{cases} \mathcal{T}^{f \to air}(x), & x \in \Omega, \\ -\mathcal{T}^{air \to f}(x), & x \in \Omega^d \setminus \overline{\Omega}. \end{cases} \tag{3.41}$$

The algorithm is based on the observation, that a sufficient local optimality condition for a given design represented by ψ reads $\psi = cG_\psi$ for some $c > 0$. This is shown in the following lemma:

Lemma 3.9 ([13]). *Let $\psi : \Omega^d \to \mathbb{R}$ be a level set function representing the domain Ω via (3.40) and G_ψ the generalized topological derivative according to (3.41). Assume that ψ is such that for all $x \in \Omega \cup (\Omega^d \setminus \overline{\Omega})$*

$$\psi(x) = cG_\psi(x), \tag{3.42}$$

for some constant $c > 0$. Then we are in presence of a local minimum.

Proof. Let $z \in \Omega$. Then, by (3.40),(3.42) and (3.41) we have

$$0 < \psi(z) = cG_\psi(z) = c\mathcal{T}^{f \to air}(z),$$

and hence $\mathcal{T}^{f \to air}(z) > 0$. Thus, by Definition 3.1 introducing a small inclusion of air around z will always increase the objective function. An analogous argument holds for the case $z \in \Omega^d \setminus \overline{\Omega}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Now, the idea of the algorithm is to approximate the topological optimality condition (3.42) by a fixed point iteration on the unit sphere $\mathcal{S}$ of $L^2(D)$. Note that for $\psi \in \mathcal{S}$ condition (3.42) is equivalent to

$$\theta := \arccos\left(\psi, \frac{G_\psi}{\|G_\psi\|}\right) = 0, \tag{3.43}$$

where θ is the angle between functions ψ and G_ψ in L^2 sense. This fact will be used in the algorithm described below.

Algorithm 3.1. Initialization: Set $k = 0$, choose initial design ψ_0 such that $\|\psi_0\| = 1$, compute $J(\psi_0)$, choose $\varepsilon_\theta > 0$

(1) compute G_{ψ_k} according to (3.41);

(2) set $\theta_k = \arccos\left(\psi_k, \frac{G_{\psi_k}}{\|G_{\psi_k}\|}\right)$

(3) if $\theta_k < \varepsilon_\theta$ then stop, else set

$$\psi_{k+1} = \frac{1}{\sin\theta_k}\left(\sin((1-\kappa_k)\theta_k)\psi_k + \sin(\kappa_k\theta_k)\frac{G_{\psi_k}}{\|G_{\psi_k}\|}\right),$$

where $\kappa_k = \max\{1, 1/2, 1/4, ...\}$ such that $J(\psi_{k+1}) < J(\psi_k)$;

(4) Set $k \leftarrow k + 1$ and go to (1);

Note that the evaluation of the generalized topological derivative involves computing the solution u to the state equation and the solution p to the adjoint equation. Algorithm 3.1 aims to generate a sequence of level set functions $\{\psi_k\}$ that minimizes the angle between ψ_k and G_{ψ_k} until for some n this angle becomes small enough, that is, the optimality condition (3.43) is satisfied up to a small tolerance ε_θ. Note that in step (3) the function ψ_{k+1} is a combination between ψ_k and G_{ψ_k} up to a positive multiplicative constant and that by construction of the iteration formula we have $\psi_{k+1} \in \mathcal{S}$ [6]. Here, we also choose a step size κ according to a line search procedure in order to get a decrease of the objective function. The reason why the iteration is performed on the unit sphere is that a level set function ψ represents the same design domain when multiplied by a positive constant. Therefore, by choosing unitary functions we get rid of this useless degree of freedom which may be beneficial for stability purposes of the algorithm [4]. A more detailed description of the algorithm can be found in [4, 3].

FE-discretization of level set function and generalized topological derivative

Let us make a few comments about discretization of the continuous level set function ψ and the generalized topological derivative. Let $\mathscr{T}_h$ be a simplicial triangulation of the hold all domain D which we assume to be shape regular. We approximate ψ by a continuous and piecewise linear finite element function. However, to evaluate the generalized topological derivative element wise it is beneficial to approximate it by discontinuous finite elements. More precisely,

$$\psi_h \in V_h := \{v_h \in H^1(\Omega^d) : v_h\big|_T \in P_1(T), \text{ for all } T \in \mathscr{T}_h\}$$
$$G_{\psi_h} \in W_h := \{v_h \in L^2(\Omega^d) : v_h\big|_T \in P_0(T), \text{ for all } T \in \mathscr{T}_h\} \tag{3.44}$$

In order to update the discrete level set function ψ_h according to step (iii) of Algorithm 3.1 we need a nodal representation of G_{ψ_h}. Therefore we introduce a suitable interpolation between the discontinuous finite element space W_h and the continuous finite element space V_h. This interpolation relies on a nodal representation of the finite element space V_h and employs a simple averaging. Let $\mathcal{K}$ denote the set of all vertices associated to the mesh $\mathcal{T}_h$. Then we define for any $z_i \in \mathcal{K}$ the set of elements that contain z_i by

$$\omega(z_i) = \{T \in \mathcal{T}_h : z_i \in \overline{T}\}.$$

For any finite element node z_i and $w_h \in W_h$ we define the local average as

$$A_{z_i}(w_h) := \frac{1}{|\omega(z_i)|} \sum_{T \in \omega(z_i)} w_h\big|_T(z_i),$$

where $|\cdot|$ denotes the cardinality of the set $\omega(z_i)$. Now, the interpolation operator $\mathcal{I}_h^{av} : W_h \to V_h$ is defined as

$$\mathcal{I}_h^{av}(w_h) := \sum_{z_i \in \mathcal{K}} A_{z_i}(w_h)\phi_i, \tag{3.45}$$

where ϕ_i denotes the nodal basis function corresponding to z_i. Applying this interpolation to G_{ψ_h} yields a function $\hat{G}_h \in V_h$ with nodal values $\hat{G}_h(z_i) = A_{z_i}(G_{\psi_h})$, $z_i \in \mathcal{K}$. Thus, $\hat{G}_h$ is continuous across the material interface which can be seen as smoothing of the topological derivative. Moreover, this smoothing is necessary as otherwise the level set function may be discontinuous after convergence.

3.2.2 Discretization of the state and adjoint equation via Nitsche-XFEM

In the proposed algorithm the interface is represented by a level set function on a fixed mesh and moves after every optimization step. This means that the interface is not

aligned to the mesh elements and so we are faced with an unfitted interface problem when solving the state and adjoint equation. Note that this is in contrast to Chapter 2 where the interface is aligned with the mesh and so standard finite element methods could be used to take the discontinuity of the reluctivity along the interface into account. However, applying standard finite elements to unfitted interface problems leads to poor convergence behavior and to maintain optimal approximation orders the standard finite element spaces need to be extended. In this section we present the Nitsche-XFEM as discretization method for unfitted interface problems. We briefly discuss the construction of a suitable approximation space allowing for discontinuities along the interface. Subsequently, we use the Nitsche method to obtain a discrete variational formulation of the state and adjoint equation.

FE-space for unfitted interface problems (XFEM/CutFEM)

Let us consider solutions u to unfitted interface problems which are domain-wise smooth but discontinuous along the interface Γ. Further, consider the standard finite element space V_h of piecewise linear and globally continuous functions. It can be shown that if we choose V_h as approximation space for u then there holds only an approximation estimate of the form

$$\inf_{v_h \in V_h} \|u - v_h\|_{L^2(D)} \le c\sqrt{h}\|u\|_{H^k(\Omega_1 \cup \Omega_2)}, \ k \ge 1,$$

see [19, Section 7.9.1]. When considering problems like (1.10), where the solution u does not have a jump discontinuity but exhibits a kink along the interface, that is a jump discontinuity in the derivative due to different material values, the approximation is better but still sub-optimal [28]:

$$\inf_{v_h \in V_h} \|u - v_h\|_{L^2(D)} \le ch^{3/2}\|u\|_{H^k(\Omega_1 \cup \Omega_2)}, \ k \ge 1.$$

Both estimates are sharp and do not improve when higher order finite elements are used. This motivates the construction of approximation spaces which provide a remedy to this problem.

In order to achieve more accuracy when approximating jumps and kinks that are not fitted to the mesh, the idea is to add more degrees of freedom in regions near the interface. We need the restriction operator $\mathcal{R}_i : L^2(D) \to L^2(\Omega_i)$ on domain Ω_i, that is $\mathcal{R}_i v = v\big|_{\Omega_i}$. Then, we define the unfitted finite element space by

$$V_h^\Gamma := \mathcal{R}_1 V_h \oplus \mathcal{R}_2 V_h. \tag{3.46}$$

Note that in literature a finite element method based on (3.46) is referred to as Cut-FEM. For the space V_h^Γ we can expect optimal approximation orders again.

Lemma 3.10 ([19, Theorem 7.9.3]). *Let $u \in H^2(\Omega_1 \cup \Omega_2)$. Then, we have*

$$\inf_{v_h \in V_h^\Gamma} \|u - v_h\|_{L^2(D)} \le ch^2 \|u\|_{H^2(\Omega_1 \cup \Omega_2)} \tag{3.47}$$

A different characterization of this space which is better suited for implementation purposes can be made by an enrichment of the standard space V_h. The construction is as follows [19]:

Define the index set $\mathcal{J} = \{1, 2, ..., n\}$ where $n = \dim V_h$. Let $(\phi_i)_{i \in \mathcal{J}}$ be the nodal basis of V_h. The index set of nodes associated to basis functions whose support intersects with the interface is given by

$$\mathcal{J}_\Gamma = \{j \in \mathcal{J} : \Gamma \cap \operatorname{supp}(\phi_j) \ne \emptyset\}.$$

Let H_Γ be the Heaviside function with $H_\Gamma(x) = 0$ for $x \in \Omega_1$ and $H_\Gamma(x) = 1$ for $x \in \Omega_2$. Now, we introduce an enrichment function

$$\Phi_j(x) = |H_\Gamma(x) - H_\Gamma(x_j)| , \, j \in \mathcal{J}.$$

The function Φ_j is used to define the new basis functions $\phi_j^\Gamma := \phi_j \Phi_j, j \in \mathcal{J}_\Gamma$. Then, the unfitted finite element space in (3.46) can be represented as

$$V_h^\Gamma = V_h \oplus V_h^x, \text{ with } V_h^x := \operatorname{span}\{\phi_j^\Gamma : \ j \in \mathcal{J}_\Gamma\}. \tag{3.48}$$

Note that a finite element discretization which uses the space V_h^Γ in (3.48) is often called extended finite element method (XFEM). By construction the new basis functions ϕ_j^Γ vanish at all nodes, that is $\phi_j^\Gamma(x_i) = 0$ for all $x_i \in \mathcal{J}$ and are supported only on the cut elements, i.e., $\operatorname{supp} \phi_j^\Gamma \subset \Omega^\Gamma$.

Nitsche-XFEM formulation of the state and adjoint equation

We consider the interface problem associated to the linear model problem (1.26) which reads

$$-\operatorname{div}(\nu \nabla u) = J_3, \ \text{in } D \subset \mathbb{R}^2, \tag{3.49a}$$
$$u = 0, \quad \text{on } \partial D, \tag{3.49b}$$
$$[\![u]\!] = 0, \quad \text{on } \Gamma, \tag{3.49c}$$
$$[\![\nu \nabla u \cdot n]\!] = 0, \quad \text{on } \Gamma. \tag{3.49d}$$

Recall that the computational domain D is decomposed into a ferromagnetic and a non ferromagnetic part, i.e., $D = \Omega_f \cup \Omega_{air}$, separated by the interface $\Gamma := \overline{\Omega}_f \cap \overline{\Omega}_{air}$. Also recall that the function ν takes the constant values ν_0 on Ω_{air} and ν_1 on Ω_f respectively. We assume a simplicial triangulation of D which is not fitted to Γ and that the interface is represented as the zero level of a given level set function ψ.

The weak formulation of this problem is stated in (1.26b). Let $V_h^0 \subset H_0^1(D)$ be the standard finite element space of continuous piecewise linear functions which vanish on the boundary ∂D. In order to discretize (3.49) we choose the space

$$V_h^\Gamma = V_h^0\big|_{\Omega_{air}} \oplus V_h^0\big|_{\Omega_f}.$$

Remark 3.4. Note that in our numerical simulations we only described the design domain Ω^d by means of a level set function. Outside Ω^d we used a mesh that is fitted to the interface. This choice allows for a sharp resolution of the interface outside Ω^d and avoids geometrical approximation which occur when discretizing a level set function. However, to explain the idea of the Nitsche method we assume here for simplicity that the whole domain D is characterized by a level set function exactly even on a discrete level. In [29] is mentioned that the geometrical error due to a $P1$ approximation of the level set function does not deteriorate any error estimates.

To simplify notation we set from now on $\Omega_0 := \Omega_{air}$ and $\Omega_1 := \Omega_f$. The Nitsche method is a method to enforce boundary or interface conditions on unfitted meshes in a weak sense by using an adapted discrete variational formulation. More precisely, it uses a consistent penalization to enforce the interface conditions [28]. The Nitsche approach for interface problems like (3.49) has been introduced in [21]. The discrete formulation of problem (3.49) using the Nitsche technique reads: find $u_h \in V_h^\Gamma$ such that

$$A_h(u_h, v_h) := a_h(u_h, v_h) + N_h^c(u_h, v_h) + N_h^c(v_h, u_h) + N_h^s(u_h, v_h) = F_h(v_h), \ \forall v_h \in V_h^\Gamma \tag{3.50}$$

with the bilinear forms

$$a_h(u, v) = \sum_{i=0,1} \int_{\Omega_i} \nu_i \nabla u \cdot \nabla v \ \mathrm{d}x, \tag{3.51a}$$

$$N_h^c(u, v) = -\int_\Gamma \{\!\!\{ \nu \nabla u \cdot n \}\!\!\} [\![v]\!] \ \mathrm{d}s, \tag{3.51b}$$

$$N_h^s(u, v) = \frac{\lambda}{h} \bar{\nu} \int_\Gamma [\![u]\!] [\![v]\!] \ \mathrm{d}s \tag{3.51c}$$

for $u, v \in V_h^\Gamma + V_{reg}$ with $V_{reg} := H_0^1(D) \cap H^2(\Omega_1 \cup \Omega_2)$.
Here, n is the outer normal vector corresponding to Ω_0, $\lambda > 0$ a (sufficiently large) stabilization parameter and $\bar{\nu} = (\nu_0 + \nu_1)/2$ is the mean reluctivity. Furthermore, $\{\!\!\{ \cdot \}\!\!\}$ denotes the averaging operator, which we define for $\kappa_0, \kappa_1 > 0$ with $\kappa_0 + \kappa_1 = 1$ as

$$\{\!\!\{ w \}\!\!\} = \kappa_0 w_0\big|_\Gamma + \kappa_1 w_1\big|_\Gamma,$$

where $w_i = w\big|_{\Omega_i}$ is the restriction of the function w to Ω_i. The linear form $F_h : V_h^\Gamma \to \mathbb{R}$ on the right hand side reads

$$F_h(v_h) = \int_D J_3 v_h \ \mathrm{d}x.$$

Note that the bilinear form N_h^c is not symmetric however A_h is. The bilinear form N_h^s accounts for the stability of the method and guarantees that A_h is coercive provided that λ is chosen sufficiently large. Another crucial point for stability is the choice of the weights κ_i, $i = 0, 1$. We use the typical choice in case of piecewise linear functions, that is, define for every $T \in \mathscr{T}_h$

$$\kappa_i\big|_T = \frac{|T_i|}{|T|}, \quad i = 0, 1$$

where $T_i = T \cap \Omega_i$ denotes the part of T in Ω_i. This particular choice has been introduced in [21]. In order to summarize the properties of $A_h(\cdot, \cdot)$ we introduce the norm

$$\|v\|_h^2 := |v|_1^2 + \|[\![v]\!]\|_{\frac{1}{2},h,\Gamma}^2 + \|\{\!\{\nu \nabla v \cdot n\}\!\}\|_{-\frac{1}{2},h,\Gamma}^2$$

with

$$\|v\|_{\pm\frac{1}{2},h,\Gamma}^2 := \sum_{T \in \mathscr{T}^\Gamma} (\bar{\nu}/h_T)^{\pm 1} \|v\|_{L^2(\Gamma_T)}^2 \quad \text{and} \quad |v|_1^2 := \sum_{i=0,1} \nu_i \|\nabla v\|_{L^2(\Omega_i)}^2.$$

Now, we can state the following result:

Lemma 3.11. *Let $A_h(\cdot, \cdot)$ be the bilinear form according to (3.50)-(3.51). Then,*

(i) $A_h(\cdot, \cdot)$ is continuous,

$$A_h(u, v) \leq C\|u\|_h \|v\|_h \quad \forall u, v \in V_{reg} + V_h^\Gamma, \tag{3.52}$$

(ii) $A_h(\cdot, \cdot)$ is coercive for $\lambda > 0$ sufficiently large,

$$A_h(v_h, v_h) \geq C\|v_h\|_h^2 \quad \forall v_h \in V_h^\Gamma. \tag{3.53}$$

Proof. A proof can be found in [21, Lemma 5]. □

This implies that the discrete variational problem (3.50) has a unique solution $u_h \in V_h^\Gamma$. Moreover, we have the following a priori estimates with respect to the norm $\|\cdot\|_h$:

Theorem 3.12. *Let $u \in V_{reg}$ be the solution to (3.49) and $u_h \in V_h^\Gamma$ be the solution to (3.50). Then, we have*

(i) $\|u - u_h\|_h \leq Ch\|u\|_{H^2(\Omega_1 \cup \Omega_2)}$,

(ii) $\|u - u_h\|_{L^2(D)} \leq Ch^2\|u\|_{H^2(\Omega_1 \cup \Omega_2)}$.

Proof. A proof can be found in [21, Theorem 6]. □

In order to discretize the adjoint problem (2.39) associated to the linear model problem (1.26) using the Nitsche technique we proceed similar as before. The discrete variational problem reads: Find $p_h \in V_h^\Gamma$ such that

$$A_h(v_h, p_h) = 2C_J \int_{\Omega_g} Q(x) \nabla u_h \cdot \nabla v_h \, \mathrm{d}x \quad \text{for all } v_h \in V_h^\Gamma. \tag{3.54}$$

Note that this problem has a unique solution $p_h \in V_h^\Gamma$ due to Lemma 3.11. Furthermore, it can be shown that p_h satisfies the same a priori error estimates as u_h (cf. [35]).

Remark 3.5 (Nonlinear model problem). In order to discretize the state and adjoint equation associated to the nonlinear model problem (1.24) we decided for each element of the triangulation whether it should belong to the ferromagnetic or non-ferromagnetic domain. Using this approach yields to a jagged interface, however we can use a standard finite element method to solve the boundary value problems.

3.3 Numerical results

Here, we present the obtained numerical results. We mention that all computations were done with the finite element library `NGSolve` [38] using the add-on library `ngsxfem`.

3.3.1 The linear model problem

In this section we apply Algorithm 3.1 in order to solve model problem (1.26) for two different initial designs Ω^d of the electric motor. The topological derivative for the linear problem can be obtained as a special case from the derivation in Section 3.1.4. Note that in the linear case we have $\lambda_1 = \lambda_2 = \nu_1$ in (3.35) and therefore the topological derivative in (3.36) simplifies to

$$\mathcal{T}^{air \to f}(z) = \frac{1}{|\omega|} U_0^T \mathcal{M} P_0, \quad z \in \Omega_{air}^d \tag{3.55}$$

with the matrix

$$\mathcal{M} = 2 |\omega| \nu_0 \frac{\nu_1 - \nu_0}{\nu_1 + \nu_0} I. \tag{3.56}$$

Similarly, the second sensitivity in (3.38) reduces to

$$\mathcal{T}^{f \to air}(z) = \frac{1}{|\omega|} U_0^T \mathcal{M}^{(2)} P_0, \quad z \in \Omega_f^d \tag{3.57}$$

with the matrix

$$\mathcal{M}^{(2)} = 2 |\omega| \nu_1 \frac{\nu_0 - \nu_1}{\nu_0 + \nu_1} I. \tag{3.58}$$

As described in the previous section the state and the adjoint equation will be solved numerically according to (3.50) and (3.54). Figures 3.2 and 3.3 show the results obtained when applying Algorithm 3.1 to the linear model problem. In Figure 3.2 we considered an axially layered rotor as initial design. Here, we could achieve an increase of the torque from 0.531 Nm to 1.23941 Nm. Observe that the topology has changed when comparing initial and final design. In Figure 3.3 we chose an initial design where we assumed to have no a prior knowledge of what could be a good starting point. Therefore we considered a massive rotor form. Note that from a physical point of view choosing this design means that there will be no torque generated. We could achieve an increase of the torque from $2.112 * 10^{-5}$ Nm to 0.9661 Nm. This demonstrates the powerfulness of the method.
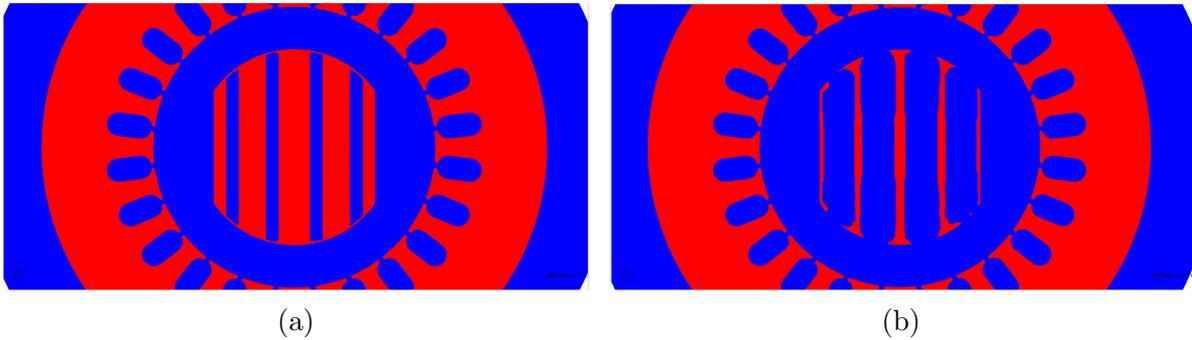

(a) (b)

Figure 3.2: Application of Algorithm 3.1 to linear problem (1.26). (a) Initial design: $T = 0.531$ Nm. (b) Final design: $T = 1.23941$ Nm.

(a) (b)

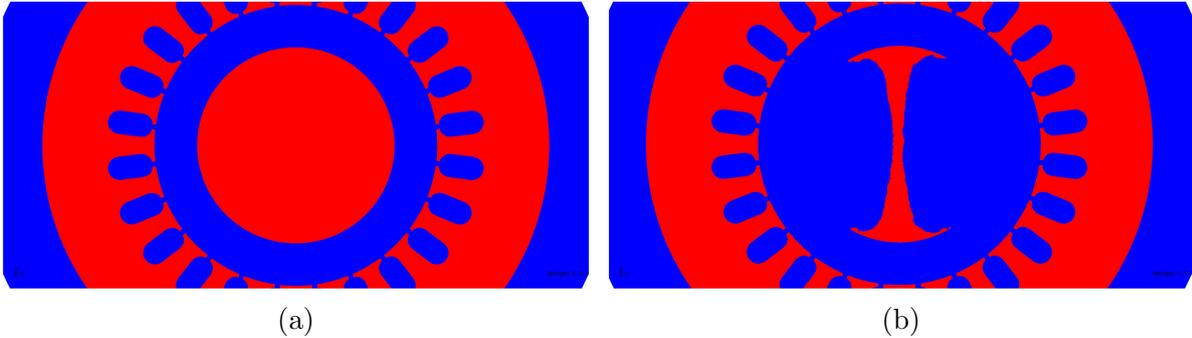Figure 3.3: Application of Algorithm 3.1 to linear problem (1.26). (a) Initial design: $T = 2.112 * 10^{-5}$ Nm. (b) Final design: $T = 0.9661$ Nm.

3.3.2 The nonlinear model problem

We apply Algortihm 3.1 to the nonlinear model problem (1.24) using the topological sensitivities obtained in (3.36) and (3.38) respectively. However, we neglect the ex-

pensive computation of the second term involving K in (3.36) and (3.38). In [5] is pointed out that this second term is comparably small to the first one and so evaluating only the first term might be a good approximation of the topological derivative. For a discussion on the efficient numerical evaluation of this second term we refer to [5]. The state and adjoint equation (1.17) and (2.18) will be solved numerically with the approach mentioned in Remark 3.5. Figures 3.4 and 3.5 show the obtained results of the level set algorithm with respect to two different initial designs. At first, we considered an axially layered initial design of the rotor shown in Figure 3.4. Here, we could achieve an increase of the objective value from 1.045806 Nm to 1.20176796 Nm. Similarly as before, the second chosen initial design is a massive rotor design shown in Figure 3.5. Here, we could achieve an increase of the torque from $2.106 * 10^{-5}$ Nm to 0.932723 Nm. Both examples show that there is a significant difference in the final designs compared to the final designs associated to the linear model problem.
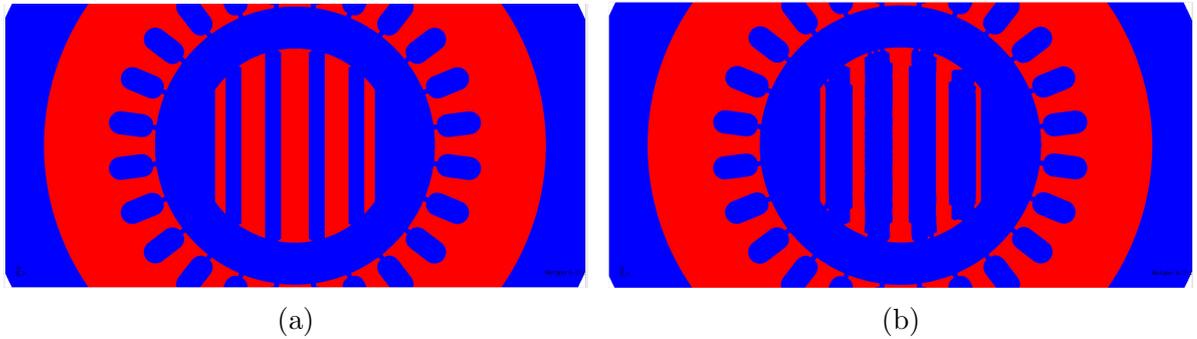

(a)   (b)

Figure 3.4: Application of Algorithm 3.1 to nonlinear problem (1.24). (a) Initial design: $T = 1.045806$ Nm. (b) Final design: $T = 1.20176796$ Nm.

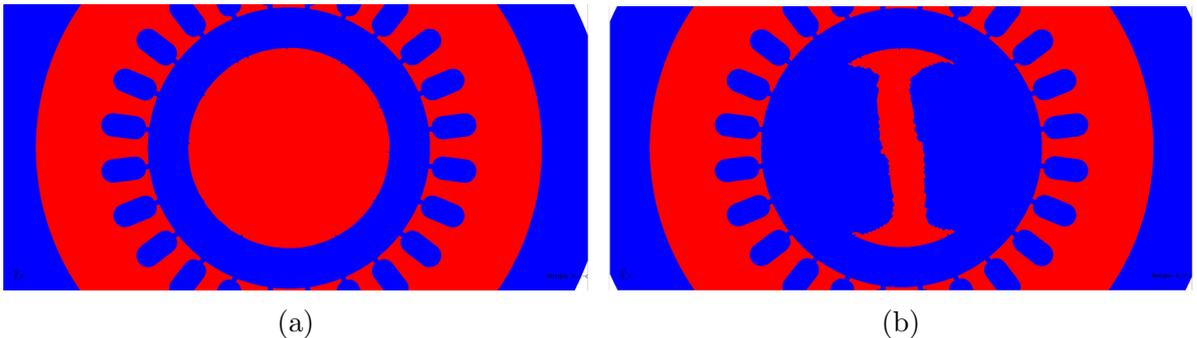(a)   (b)

Figure 3.5: Application of Algorithm 3.1 to nonlinear problem (1.24). (a) Initial design: $T = 2.106 * 10^{-5}$ Nm. (b) Final design: $T = 0.932723$ Nm.

3.4 Topology optimization of a synchronous reluctance motor under uncertainties

In this section we want to solve the stochastic model problem introduced in Section 2.5 via a level set method. Therefore we extend the algorithm in Section 3.2.1 to be capable of solving stochastic design optimization problems. This new algorithm is motivated by the stochastic gradient method which was also the basic idea of the stochastic shape gradient method introduced in Section 2.5.2. We present the algorithm in detail in Section 3.4.1. Subsequently, we apply this novel approach to solve the stochastic model problem (2.47) and present the obtained results.

3.4.1 A level set method for stochastic problems

Recall (2.47) and in particular recall the objective function considered therein, that is,

$$J_{\hat{\omega}}(\Omega) = -C_J \int_{\Omega_g} Q(x) \nabla u(x, \hat{\omega}) \cdot \nabla u(x, \hat{\omega}) \mathrm{d}x \qquad (3.59)$$

for a fixed $\hat{\omega} \in X$. In order to solve model problem (2.47) we combine the level set algorithm (Algorithm 3.1) with the idea of the stochastic gradient method. For this reason we need to compute the topological derivative of the (deterministic) shape function $J_{\hat{\omega}}$ which can be done similarly as in Section 3.1.4 by setting $a_1(y) = (1 + \delta\xi(\hat{\omega}))\hat{\nu}(|y|)y$ and $a_2(y) = \nu_0 y$, $y \in \mathbb{R}^2$ and applying Theorem 3.7 for fixed $\hat{\omega} \in X$. The derivatives for a fixed realization $\hat{\omega} \in X$ read

$$\mathcal{T}_{\hat{\omega}}^{air \to f}(z) := dJ_{\hat{\omega}}(\Omega)(z) = \frac{1}{|\omega|} \Big[U_0^T \mathcal{M} P_0$$
$$+ \int_{\mathbb{R}^2} (T_{\hat{\omega}}(\nabla K + U_0) - T_{\hat{\omega}}(U_0) - DT_{\hat{\omega}}(U_0)\nabla K)\chi_\omega \cdot (P_0 + \nabla Q) \mathrm{d}x \Big] \qquad (3.60)$$

for $z \in \Omega_{air}^d$ and

$$\mathcal{T}_{\hat{\omega}}^{f \to air}(z) := dJ_{\hat{\omega}}(\Omega)(z) = \frac{1}{|\omega|} \Big[U_0^T \mathcal{M}^{(2)} P_0$$
$$+ \int_{\mathbb{R}^2} (T_{\hat{\omega}}(\nabla K + U_0) - T_{\hat{\omega}}(U_0) - DT_{\hat{\omega}}(U_0)\nabla K)\chi_{\mathbb{R}^2 \setminus \omega} \cdot (P_0 + \nabla Q^{(2)}) \mathrm{d}x \Big] \qquad (3.61)$$

for $z \in \Omega_f^d$. Here, $T_{\hat{\omega}}$ and $DT_{\hat{\omega}}$ represent the stochastic versions of (2.9) and (2.10), that is,

$$T_{\hat{\omega}}(y) := (1 + \delta\xi(\hat{\omega}))T(y) \quad \text{and} \quad DT_{\hat{\omega}}(y) := (1 + \delta\xi(\hat{\omega}))DT(y), \ y \in \mathbb{R}^2.$$

Furthermore, $\mathcal{M}$ and $\mathcal{M}^{(2)}$ are defined according to (3.35) and (3.39) respectively. Note that the quantities U_0, P_0, K, Q and $Q^{(2)}$ depend on $\hat{\omega}$ as well. In the stochastic level set method we use a generalized topological derivative which for a given level set function ψ according to (3.40) and a fixed realization $\hat{\omega} \in X$ is defined as

$$G_\psi^{\hat{\omega}}(x) := \begin{cases} \mathcal{T}_{\hat{\omega}}^{f \to air}(x), & x \in \Omega, \\ -\mathcal{T}_{\hat{\omega}}^{air \to f}(x), & x \in \Omega^d \setminus \overline{\Omega}. \end{cases} \tag{3.62}$$

Now, we propose the following algorithm:

Algorithm 3.2. Initialization: Set $k = 0$, choose initial design ψ_0 such that $\|\psi_0\| = 1$, choose $\varepsilon_\theta > 0$

(1) generate $\hat{\omega}_k$

(2) compute $G_{\psi_k}^{\hat{\omega}}$ according to (3.62) for $\hat{\omega} = \hat{\omega}_k$;

(3) set $\theta_k = \arccos\left(\psi_k, \frac{G_{\psi_k}^{\hat{\omega}}}{\|G_{\psi_k}^{\hat{\omega}}\|}\right)$ for $\hat{\omega} = \hat{\omega}_k$;

(4) if $\theta_k < \varepsilon_\theta$ then stop, else set

$$\psi_{k+1} = \frac{1}{\sin\theta_k}\left(\sin((1 - \kappa_k)\theta_k)\psi_k + \sin(\kappa_k\theta_k)\frac{G_{\psi_k}^{\hat{\omega}}}{\|G_{\psi_k}^{\hat{\omega}}\|}\right) \quad \text{for } \hat{\omega} = \hat{\omega}_k,$$

where $\kappa_k = \max\{1, 1/2, 1/4, ...\}$ such that $J_{\hat{\omega}_k}(\psi_{k+1}) < J_{\hat{\omega}_k}(\psi_k)$;

(5) Set $k \leftarrow k + 1$ and go to (1);

Note that in order to evaluate the generalized topological derivative in step (2) one has to compute the solution $u = u(\cdot, \hat{\omega})$ to the state equation (2.43) and the solution $p = p(\cdot, \hat{\omega})$ to the adjoint equation (2.49) for $\hat{\omega} = \hat{\omega}_k$ respectively. In step (4) we choose the step size κ_k according to a linesearch procedure, which showed a satisfying convergence behavior in our experiments. However, from the shape gradient method is known that this choice might be critical and other step size rules like Robbins-Monro might be better suited.

3.4.2 Numerical results

We apply Algorithm 3.2 in order to solve model problem (2.47) numerically. To simplify the computation of the topological sensitivities we proceed similarly as for the nonlinear model problem and neglect the calculation of the second term involving K in (3.60) and (3.61). The discretization of the state and adjoint equation has been done as described in Remark 3.5 using piecewise linear finite elements. In our simulations we use a single sample $\hat{\omega}_k$ at each iteration k to compute the generalized topological

derivative. However, in order to calculate objective function values at the first and last iteration we use additional sampling. More precisely, we approximate $\mathbb{E}[J_{\hat{\omega}}(\psi_k)]$ by the estimate

$$\hat{j}_k := \frac{1}{m}\sum_{\ell=1}^{m}J_{\hat{\omega}_{k,\ell}}(\psi_k) \approx \mathbb{E}[J_\omega(\psi_k)],$$

with m identically and independently distributed samples $\{\hat{\omega}_{k,1},...,\hat{\omega}_{k,m}\}$ generated at iteration k. Observe that $-\hat{j}_k$ represents the torque. Note that we set $m = 100$ to compute $\hat{j}_k$ for the first and last iteration. As before we consider two different initial designs of the rotor, that are, an axially layered rotor design and a massive rotor design. The results for an axially layered initial design are shown in Figure 3.6. We could achieve an decrease of the objective function from $\hat{j}_0 = -1.04576424$ to $\hat{j}_6 = -1.20181864$. Compared to the deterministic case in Figure 3.4 there are hardly any changes in the design. A computation of $\mathbb{E}[J_{\hat{\omega}}(\psi_{final})]$ for the final design ψ_{final} of the deterministic case gives $|\mathbb{E}[J_{\hat{\omega}}(\psi_{final})]| \approx 1.20171399$ Nm which is a little bit lower than $|\hat{j}_6|$. The results for a massive initial design are shown in Figure 3.7. Here, the objective function decreases from $\hat{j}_0 = -2.10639442*10^{-5}$ to $\hat{j}_{19} = -0.93796927$. Compared to the results in Figure 3.5 we observe slight changes in the final design. Again, a deeper investigation shows that $|\mathbb{E}[J_{\hat{\omega}}(\psi_{final}^{m})]| \approx 0.93272016$ Nm where ψ_{final}^{m} denotes the final design of the massive rotor in the deterministic case is lower than $|\hat{j}_{19}|$. Thus, the final design in the stochastic case may be more robust.

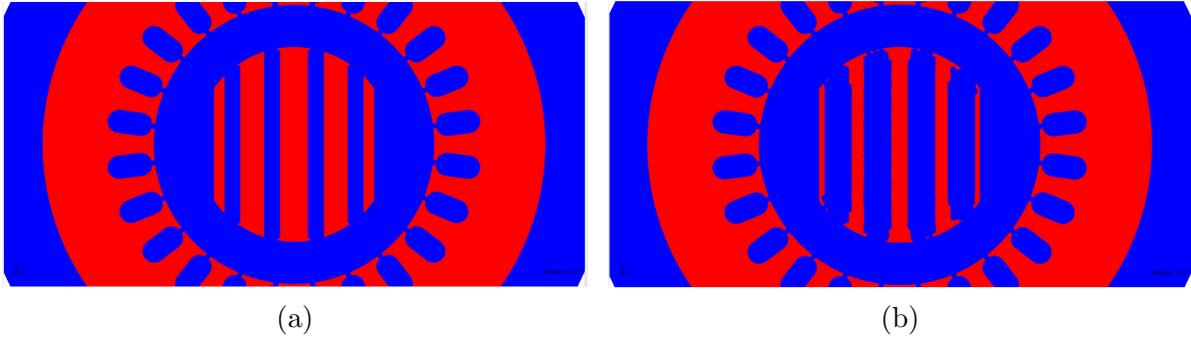| (a) | (b) |

Figure 3.6: Application of Algorithm 3.2 to nonlinear stochastic problem (2.47). (a) Initial design: $|\hat{j}_0| = 1.04576424$ Nm. (b) Final design: $|\hat{j}_6| = 1.20181864$ Nm.

<div align="center">(a)   (b)</div>
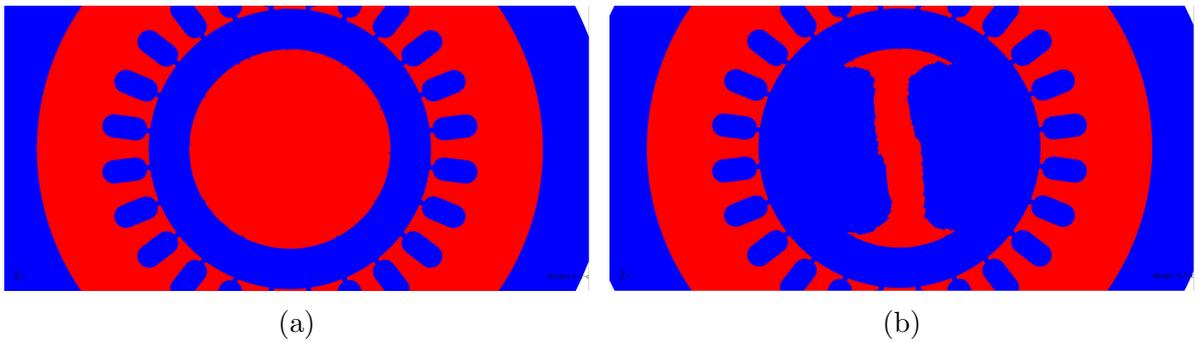
Figure 3.7: Application of Algorithm 3.2 to nonlinear stochastic problem (2.47). (a) Initial design: $|\hat{j}_0| = 2.10639442 * 10^{-5}$ Nm. (b) Final design: $|\hat{j}_{19}| = 0.93796927$ Nm.

Conclusion

In this thesis we solved a design optimization problem of an electrical machine motivated from an application in medical engineering by means of shape and topology optimization methods based on sensitivity information. We considered two deterministic PDE-constrained optimization problems, i.e., (i) one with a quasi linear PDE-constraint and (ii) one with a linear PDE-constraint as well as (iii) a stochastic model formulation, where we assumed the reluctivity subject to uncertainty.

To begin with we solved the problems from a shape optimization point of view. In the first part we paid attention to the two deterministic problems. Here, we derived the shape derivative for problem (i) rigorously and described the used numerical procedure in detail. The resulting rotor designs of both deterministic problems could achieve a significant increase of the torque (22% for (i) and 28% for (ii)) compared to the preexisting axially layered rotor design. A comparison between these results revealed that for design optimization it is beneficial to consider the more realistic model (i). Next, we extended problem (i) and considered a stochastic optimization problem. We applied a stochastic gradient method for shape spaces to solve it. The results showed that the final rotor design may be more robust than the rotor design of problem (i) meaning that it accounts for the uncertainty in the material. However, the sensitivity of the rotor design with respect to this uncertainty seems to be small. In the final part we demonstrated the efficiency of the used shape optimization method compared to a parametric optimization approach applied to problem (i).

As shape optimization methods do not allow for topological changes of the design domain, we decided to solve the same problems from the perspective of topology optimization. We derived the topological derivative for problem (i) in detail via a Lagrangian framework. We used a level set algorithm which solely employs the topological derivative together with a finite element method for unfitted interfaces to solve the deterministic problems. We showed numerical results for two different initial designs which demonstrated the flexibility of the method. Finally, we presented a novel algorithm which combines the level set method and the stochastic gradient method to solve the stochastic problem (iii). Numerical results showed that the final rotor designs may be more robust than the final rotor designs associated to problem (i).

Open questions and possible future work

The work presented in this thesis can be extended towards several directions:

While we neglect the computation of the second term of the topological derivative of (i) as mentioned in [5], it could be beneficial to use the full sensitivity information. Further, using a Nitsche-XFEM formulation for the boundary value problems associated to (i) in Section 3.3.2 might lead to a better resolution of the interface in the final designs.

Convergence analysis and the role of the step size for the stochastic level set algorithm is an open question.

We considered a fixed rotor position when maximizing the torque. It would be interesting to obtain optimal designs which account for several rotor positions and for a varying impressed current density.

While we used a two dimensional geometrical model of the motor, it would be interesting to investigate the effect of considering a three dimensional setting on the rotor designs. Therefore, the shape derivative has to be computed for the curl-curl setting which should be possible under mild modifications. The derivation of the topological derivative for 3D quasi-linear magnetostatics was recently done in [15].

Bibliography

[1] G. Allaire. *Shape optimization by the homogenization method.* Applied Mathematical Sciences. Springer, New York, 2002.

[2] S. Amstutz. Sensitivity analysis with respect to a local perturbation of the material property. *Asymptotic analysis*, 49(1, 2):87–108, 2006.

[3] S. Amstutz. Analysis of a level set method for topology optimization. *Optimization Methods and Software*, 26(4-5):555–573, 2011.

[4] S. Amstutz and H. Andrä. A new algorithm for topology optimization using a level-set method. *Journal of Computational Physics*, 216(2):573–588, 2006.

[5] S. Amstutz and P. Gangl. Topological derivative for the nonlinear magnetostatic problem. *Electron. Trans. Numer. Anal.*, 51:169–218, 2019.

[6] S. Amstutz, A. Novotny, and E. de Souza Neto. Topological derivative-based topology structural optimization under drucker-prager-type stress constraints. *Blucher Mechanical Engineering Proceedings*, 1(1):4683–4702, 2012.

[7] M. P. Bendsøe and O. Sigmund. *Topology Optimization: Theory, Methods, and Applications.* Springer, Berlin, 2003.

[8] A. Binder. *Elektrische Maschinen und Antriebe: Grundlagen, Betriebsverhalten.* Springer, 2012.

[9] J. Céa. Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction coût. *ESAIM: Mathematical Modelling and Numerical Analysis*, 20(3):371–402, 1986.

[10] S. De, J. Hampton, K. Maute, and A. Doostan. Topology optimization under uncertainty using a stochastic gradient-based approach. *arXiv preprint arXiv:1902.04562*, 2019.

[11] M. C. Delfour and J.-P. Zolésio. *Shapes and geometries*, volume 22. SIAM, second edition, 2011. Metrics, analysis, differential calculus, and optimization.

[12] R. P. Dwight. Robust mesh deformation using the linear elasticity equations. In *Computational fluid dynamics 2006*, pages 401–406. Springer, 2009.

[13] P. Gangl. *Sensitivity-Based Topology and Shape Optimization with Application to Electrical Machines.* PhD thesis, Johannes Kepler University Linz, 2016.

[14] P. Gangl, U. Langer, A. Laurain, H. Meftahi, and K. Sturm. Shape optimization of an electric motor subject to nonlinear magnetostatics. *SIAM Journal on Scientific Computing*, 37(6):B1002–B1025, 2015.

[15] P. Gangl and K. Sturm. Asymptotic analysis and topological derivative for 3d quasi-linear magnetostatics. *arXiv preprint arXiv:1908.10775*, 2019.

[16] P. Gangl and K. Sturm. A simplified derivation technique of topological derivatives for quasi-linear transmission problems. *arXiv preprint arXiv:1907.13420*, 2019.

[17] C. Geiersbach, E. Loayza, and K. Welker. Computational aspects for interface identification problems with stochastic modelling. *arXiv preprint arXiv:1902.01160*, 2019.

[18] C. Geiersbach, E. Loayza-Romero, and K. Welker. Stochastic approximation for optimization in shape spaces. *arXiv preprint arXiv:2001.10786*, 2020.

[19] S. Gross and A. Reusken. *Numerical methods for two-phase incompressible flows.* Springer, 2011.

[20] J. Hadamard. *Mémoire sur le problème d'analyse relatif à l'équilibre des plaques élastiques encastrées*, volume 33. Imprimerie nationale, 1908.

[21] A. Hansbo and P. Hansbo. An unfitted finite element method, based on Nitsche's method, for elliptic interface problems. *Computer methods in applied mechanics and engineering*, 191(47-48):5537–5552, 2002.

[22] A. Henrot and M. Pierre. *Shape variation and optimization: a geometrical analysis.* European Mathematical Society Publishing House, 2018.

[23] F. Henrotte, G. Deliége, and K. Hameyer. The eggshell approach for the computation of electromagnetic forces in 2d and 3d. *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering*, 23(4):996–1005, 2004.

[24] J. A. Iglesias, K. Sturm, and F. Wechsung. Two-dimensional shape optimization with nearly conformal transformations. *SIAM Journal on Scientific Computing*, 40(6):A3807–A3830, 2018.

[25] M. Jung and U. Langer. *Methode der finiten Elemente für Ingenieure: Eine Einführung in die numerischen Grundlagen und Computersimulation.* Springer-Vieweg, 2013. 2., überarbeitete und erweiterte Auflage.

[26] A. Kost. *Numerische Methoden in der Berechnung elektromagnetischer Felder.* Springer-Verlag Berlin Heidelberg, 1994.

[27] A. Laurain and K. Sturm. Distributed shape derivative via averaged adjoint method and applications. *ESAIM: Mathematical Modelling and Numerical Analysis*, 50(4):1241–1267, 2016.

[28] C. Lehrenfeld. *On a space-time extended finite element method for the solution of a class of two-phase mass transport problems.* PhD thesis, RWTH Aachen University, 2015.

[29] C. Lehrenfeld and A. Reusken. Analysis of a high-order unfitted finite element method for elliptic interface problems. *IMA Journal of Numerical Analysis*, 38(3):1351–1387, 2018.

[30] G. Lei, J. Zhu, Y. Guo, C. Liu, and B. Ma. A review of design optimization methods for electrical machines. *Energies*, 10(12):1962, 2017.

[31] C. Ortner and E. Suli. A note on linear elliptic systems on $\mathbb{R}^d$. *arXiv preprint arXiv:1202.3970*, 2012.

[32] C. Pechstein. Multigrid-Newton-methods for nonlinear magnetostatic problems. Master's thesis, Johannes Kepler University Linz, 2004.

[33] C. Pechstein and B. Jüttler. Monotonicity-preserving interproximation of B–H-curves. *Journal of Computational and Applied Mathematics*, 196(1):45–57, 2006.

[34] P. Putek, R. Pulch, A. Bartel, E. J. W. ter Maten, M. Günther, and K. M. Gawrylczyk. Shape and topology optimization of a permanent-magnet machine under uncertainties. *Journal of Mathematics in Industry*, 6(1):11, 2016.

[35] H.-G. Raumer. Shape optimization for interface problems using unfitted finite elements. Master's thesis, University of Göttingen, 2018.

[36] H. Robbins and S. Monro. A stochastic approximation method. *The annals of mathematical statistics*, 22(3):400–407, 1951.

[37] N. Sadowski, Y. Lefèvre, M. Lajoie-Mazenc, and J. Cros. Finite element torque calculation in electrical machines while considering the movement. *IEEE Transactions on Magnetics*, 28(2):1410–1413, 1992.

[38] J. Schöberl. C++ 11 implementation of finite elements in NGSolve. Technical Report 30, Institute for analysis and scientific computing, Vienna University of Technology, 2014.

[39] O. Sigmund and K. Maute. Topology optimization approaches: A comparative review. *Structural and Multidisciplinary Optimization*, 48(6):1031–1055, 2013.

[40] J. Sokolowski and A. Zochowski. On the topological derivative in shape optimization. *SIAM Journal on Control and Optimization*, 37(4):1251–1272, 1999.

[41] J. Sokolowski and J.-P. Zolésio. *Introduction to shape optimization*, volume 16 of *Springer Series in Computational Mathematics*. Springer, Berlin, 1992.

[42] K. Sturm. Minimax Lagrangian approach to the differentiability of nonlinear PDE constrained shape functions without saddle point assumption. *SIAM Journal on Control and Optimization*, 53(4):2017–2039, 2015.

[43] K. Sturm. *On shape optimization with non-linear partial differential equations*. PhD thesis, Technische Universität Berlin, 2015. `http://dx.doi.org/10.14279/depositonce-4339`.

[44] S. Zaglmayr. *High Order Finite Element Methods for Electromagnetic Field Computation*. PhD thesis, Johannes Kepler University Linz, 2006.

[45] W. P. Ziemer. *Weakly differentiable functions: Sobolev spaces and functions of bounded variation*, volume 120 of *Graduate Texts in Mathematics*. Springer, New York, 1989.

[46] J.-P. Zolésio. Introduction to Shape Optimization Problems and Free Boundary Problems. In *Shape Optimization and Free Boundaries*, pages 397–457. Springer, Dordrecht, 1992.

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

_____   _____
             Date   Signature