

## SYMPLECTIC PONTRYAGIN APPROXIMATIONS FOR OPTIMAL DESIGN \*

JESPER CARLSSON<sup>1</sup>, MATTIAS SANDBERG<sup>2</sup> AND ANDERS SZEPESSY<sup>1,3</sup>

**Abstract.** The powerful Hamilton-Jacobi theory is used for constructing regularizations and error estimates for optimal design problems. The constructed Pontryagin method is a simple and general method for optimal design and reconstruction: the first, analytical, step is to regularize the Hamiltonian; next the solution to its stationary Hamiltonian system, a nonlinear partial differential equation, is computed with the Newton method. The method is efficient for designs where the Hamiltonian function can be explicitly formulated and when the Jacobian is sparse, but becomes impractical otherwise (*e.g.* for non local control constraints). An error estimate for the difference between exact and approximate objective functions is derived, depending only on the difference of the Hamiltonian and its finite dimensional regularization along the solution path and its  $L^2$  projection, *i.e.* not on the difference of the exact and approximate solutions to the Hamiltonian systems.

**Mathematics Subject Classification.** 65N21, 49L25.

Received April 4, 2007. Revised April 8, 2008.

Published online October 16, 2008.

### 1. INTRODUCTION TO OPTIMAL DESIGN

As the computational capacity increases it becomes possible to solve more demanding construction problems. For instance, instead of only computing the deformation of a given construction, it is possible to computationally design an optimal construction with minimal deformation for a given load. In a mathematical setting optimal design is a particular inverse problem where the goal is to determine input functions in a partial differential equation that meet the design criteria in an optimal way; for instance, to distribute a fixed amount of material in space to construct a bridge with minimal deformation, for a given load, means to solve the elasticity equations and determine a material configuration. The start of this computational work is presented in [30] and has now become an activity with large international optimal design conferences and several books [1,4,30]. Inverse problems are often ill posed, *e.g.* small perturbations of data lead to large changes in the solution. To computationally solve inverse problems therefore requires some regularization, *cf.* [1,4,19,30,36]. The standard Tikhonov method [35] requires to choose a penalty, usually a norm and a parameter, with the purpose to regularize the computational method. Although there is good understanding how to choose the penalty for some

---

*Keywords and phrases.* Topology optimization, inverse problems, Hamilton-Jacobi, regularization, error estimates, impedance tomography, convexification, homogenization.

\* Support by the Swedish Research Council grants 2002-6285 and 2002-4961, and the European network HYKE, funded by the EC as contract HPRN-CT-2002-00282, is acknowledged.

<sup>1</sup> Department of Numerical Analysis, Kungl. Tekniska Högskolan, 100 44 Stockholm, Sweden. [jesperc@kth.se](mailto:jesperc@kth.se)

<sup>2</sup> CMA, University of Oslo, P.O. Box 1053 Blindern, 0316 Oslo, Norway. [mattias.sandberg@cma.uio.no](mailto:mattias.sandberg@cma.uio.no)

<sup>3</sup> Department of Mathematics, Kungl. Tekniska Högskolan, 100 44 Stockholm, Sweden. [szepessy@kth.se](mailto:szepessy@kth.se)

problems, *e.g.* in tomography [29], with norms related to the required regularity of solutions and parameters related to the error in data, there is no complete theory for how to regularize general nonlinear problems.

The objective of this work is to show how the powerful theory of viscosity solutions, for time dependent optimal control problems in the dynamic programming setting, can be used as a theoretical tool to find a regularization criterion and estimate its approximation error; based on this regularization we then construct a simple and general computational method also for some highly nonlinear time independent optimal design problems, extending the work [32] on time dependent problems. Our method [33] reduces to solve a Hamiltonian system, where the Hamiltonian is a  $C^2$ -regularized version of the original Hamiltonian. This Hamiltonian system is a nonlinear partial differential equation, where the Newton method with a sparse Jacobian becomes efficient and simple to use, *e.g.* in standard PDE software. The idea is to start by computing solutions to the Hamiltonian system with a highly regularized Hamiltonian. When this is done a less regularized Hamiltonian is chosen, and the previously computed solutions serve as a starting position for the Newton method with the new Hamiltonian. This procedure of successively decreasing the regularization is repeated as far as possible. As it is sufficient to use a one-parameter family of regularized Hamiltonians, it is only needed to tune this one parameter. In the ideal situation, however, this parameter may successively be decreased to zero. Therefore only one parameter needs to be set at the start of the first Newton iteration. This is an advantage compared to methods based on regularizations using penalization of terms involving the control variable, and (sometimes many) differentials of it; in such methods there are often several parameters.

A clear limitation of our method is the requirement to obtain an explicit formula for the regularized Hamiltonian. The optimal design and reconstruction problems presented in this paper need local constraints on the control and have explicit formulas. Non local constraints on the control or Lagrangians depending on derivatives of the control typically do not admit explicit formulas, and then our method may become impractical; in other words, we use an advantage of those particular problems having explicit formulas of the Hamiltonian. When penalty terms with derivatives are present, these derivatives must be considered as control variables. Another limitation is that the solution to the Hamiltonian system gives the primal and dual variables, but not the control function directly. As will be shown in the examples considered in this paper, it is however often possible to obtain an approximate control using the approximate primal and dual solutions.

We present a natural regularization to meet the (minimal) requirement that there is a regularized discrete solution, which is accurately approximating the optimal objective function (but not necessary the optimal control function). This regularization is derived from viscosity solutions theory for time dependent problems. Future work may use better insight also on the more demanding task to be able to iteratively compute solutions (instead of as above only approximate them) – *e.g.* in the time independent setting – to find further improved regularizations.

We derive an error estimate for the difference between exact and approximate objective functions, depending only on the difference of the Hamiltonian and its finite dimensional regularization along the solution path and its  $L^2$  projection, *i.e.* not on the difference of the exact and approximate solutions to the Hamiltonian systems. This error estimate is of the right form to be a good starting point for new studies on adaptive methods for optimal control problems: the difference of the Hamiltonians yields an error density which measures errors both from discretization and the regularization. Our experiments show that the estimate is relatively sharp.

One way to characterize so called symplectic time discretizations for Hamiltonian systems is that these approximations are exact solutions to another Hamiltonian system, *cf.* [23]. In this sense, our approximation in space and regularization shares this symplectic property. An equivalent definition of symplectic time discretization methods for Hamiltonian systems derived from optimal control problems is that the first variation of the discrete value function agrees with the discretization of the Lagrange multiplier, see [33] where symplectic time discretizations are analyzed for optimal control problems with the similar use of viscosity solution theory as here. This property that the first variation of the discrete value function agrees with the discretization of the Lagrange multiplier only makes sense for time dependent problems, which is one reason our analysis starts by extending the original time independent optimal control problem to an artificial time dependent dynamic programming formulation.

We study three different examples where the nonlinear PDEs are of different character: scalar concave maximization, scalar non-concave maximization and an elliptic system for reconstruction. The homogenization method is a theoretically powerful way to regularize some optimal design problems [1], based on a change of control variables related to composite micro structure. This regularization changes the differential operator part of the Lagrangian. Section 3.2 suggests a simple alternative regularization based on the material constraint part of the Lagrangian: the constraint on material volume  $\int_{\Omega} \sigma dx = C$ , which usually is included in the Lagrangian by a term  $\eta \int_{\Omega} \sigma dx$ , is now instead represented by  $\eta' \int_{\Omega} \sigma^{-1} dx$  which turns out to give the same regularized Hamiltonian as the homogenization method for a scalar problem.

## 2. SYMPLECTIC PONTRYAGIN APPROXIMATIONS IN OPTIMAL CONTROL

Consider a differential equation constrained minimization problem with solution  $\varphi$  in some Hilbert space  $V$  on a domain  $\Omega$  and control  $\sigma \in \mathcal{A} := \{\sigma : \Omega \rightarrow B\}$ :

$$\begin{aligned} f(\varphi, \sigma) &= 0 \quad \text{in distribution,} \\ \min_{\sigma \in \mathcal{A}} h(\varphi, \sigma), \end{aligned} \tag{2.1}$$

and its approximation with solution  $\bar{\varphi} \in \bar{V} \subset V$  and control  $\bar{\sigma} \in \bar{\mathcal{A}}$ :

$$\begin{aligned} \bar{f}(\bar{\varphi}, \bar{\sigma}) &= 0 \\ \min_{\bar{\sigma} \in \bar{\mathcal{A}}} \bar{h}(\bar{\varphi}, \bar{\sigma}). \end{aligned} \tag{2.2}$$

**Example 2.1** (optimal conductor). Section 3.1 presents minimization of the power loss in an electric conductor, by distributing a given amount of conducting material in a domain  $\Omega \subset \mathbb{R}^d$ , with a surface current  $q$ . Let  $\eta \in \mathbb{R}$  be a constant Lagrange multiplier, associated to the given amount of material, and find an optimal conduction distribution  $\sigma : \Omega \rightarrow \{\sigma_-, \sigma_+\} =: B$ , where  $\sigma_{\pm} > 0$ , such that

$$\min_{\sigma} \left\{ \int_{\partial\Omega} q\varphi \, ds + \eta \int_{\Omega} \sigma \, dx \mid f(\varphi, \sigma) := \operatorname{div}(\sigma \nabla \varphi) \Big|_{\Omega} = 0, \sigma \frac{\partial \varphi}{\partial n} \Big|_{\partial\Omega} = q \right\}. \tag{2.3}$$

Here  $\partial/\partial n$  denotes the normal derivative,  $ds$  is the surface measure on  $\partial\Omega$  and  $\varphi \in V := \{v \in H^1(\Omega) : \int_{\Omega} v \, dx = 0\}$  is the electric potential.

The corresponding Lagrangians

$$\begin{aligned} \mathcal{L}(\lambda, \varphi, \sigma) &:= \langle \lambda, f(\varphi, \sigma) \rangle + h(\varphi, \sigma), \\ \bar{\mathcal{L}}(\bar{\lambda}, \bar{\varphi}, \bar{\sigma}) &:= \langle \bar{\lambda}, \bar{f}(\bar{\varphi}, \bar{\sigma}) \rangle + \bar{h}(\bar{\varphi}, \bar{\sigma}), \end{aligned}$$

can be used to formulate the conditions

$$\begin{aligned} \partial_1 \mathcal{L} &= f(\varphi^*, \sigma^*) = 0, \\ \partial_2 \mathcal{L} &= \langle \lambda^*, \partial f(\varphi^*, \sigma^*) \rangle + \partial h(\varphi^*, \sigma^*) = 0, \\ \sigma^* &\in \operatorname{argmin}_{\sigma \in \mathcal{A}} \{ \langle \lambda^*, f(\varphi^*, \sigma) \rangle + h(\varphi^*, \sigma) \}, \end{aligned} \tag{2.4}$$

inspired by the corresponding time dependent condition (2.10). Here,  $\partial_1 =: \partial$  and  $\partial_2$  are the Gateaux derivatives with respect to the first and second arguments respectively, and  $\langle v, w \rangle$  is the duality pairing on  $V$ , which reduces to the  $L^2(\Omega)$  inner product if  $v, w \in L^2(\Omega)$ .

Optimal control problems are inverse problems. It is well known that inverse problems often are ill-posed; therefore they need to be regularized. We will use a formulation of (2.4) based on the Hamiltonian to regularize our optimal design problems. The Hamiltonians  $H : V \times V \rightarrow \mathbb{R}$  and  $\bar{H} : V \times \bar{V} \rightarrow \mathbb{R}$  defined as

$$\begin{aligned} H(\lambda, \varphi) &:= \min_{\sigma \in \mathcal{A}} \{ \langle \lambda, f(\varphi, \sigma) \rangle + h(\varphi, \sigma) \} \\ \bar{H}(\lambda, \bar{\varphi}) &:= \min_{\bar{\sigma} \in \bar{\mathcal{A}}} \{ \langle \lambda, \bar{f}(\bar{\varphi}, \bar{\sigma}) \rangle + \bar{h}(\bar{\varphi}, \bar{\sigma}) \} \end{aligned} \quad (2.5)$$

eliminate the control variables in the local equilibrium conditions

$$\begin{aligned} \partial_1 H(\lambda^*, \varphi^*) &= 0 \\ -\partial_2 H(\lambda^*, \varphi^*) &= 0, \end{aligned} \quad (2.6)$$

and similarly for  $(\bar{\lambda}^*, \bar{\varphi}^*)$

$$\begin{aligned} \partial_1 \bar{H}(\bar{\lambda}^*, \bar{\varphi}^*) &= 0 \\ -\partial_2 \bar{H}(\bar{\lambda}^*, \bar{\varphi}^*) &= 0. \end{aligned}$$

It turns out that for our optimal design problems it is easy to find an explicit expression for the Hamiltonian; for some other constrained minimization problems the Hamiltonian can be too complicated to use computationally. The Hamiltonian is in general only Lipschitz continuous even if  $f, g$  and  $h$  are smooth and we shall see that in particular for optimal design the Hamiltonians indeed are not differentiable everywhere. We also see that the stationary Hamiltonian system (2.6) becomes undefined where the Hamiltonian is not differentiable. At a point where the Hamiltonian is not differentiable the optimal control depends discontinuously on  $(\lambda, \varphi)$ . The Hamiltonian form has the advantage that the Newton method can be used to iteratively solve the nonlinear constrained optimization problem (2.1) when the Hamiltonian can be written using a simple formula so that the Hessian of  $H$  is explicitly known, while Hessian information is in general not available for direct constrained minimization based on the control variable  $\sigma$ . We want to understand how to regularize and to estimate errors introduced by approximation and regularization of the optimal control problem. In particular we seek an estimate of the error

$$E := \min_{\bar{\sigma} \in \bar{\mathcal{A}}} \{ \bar{h}(\bar{\varphi}, \bar{\sigma}) \mid \bar{f}(\bar{\varphi}, \bar{\sigma}) = 0 \} - \min_{\sigma \in \mathcal{A}} \{ h(\varphi, \sigma) \mid f(\varphi, \sigma) = 0 \}.$$

The definition of  $\lambda^*, \varphi^*$  and  $H$  imply

$$\min_{\sigma \in \mathcal{A}} \{ h(\varphi, \sigma) \mid f(\varphi, \sigma) = 0 \} = H(\lambda^*, \varphi^*) \quad (2.7)$$

$$\min_{\bar{\sigma} \in \bar{\mathcal{A}}} \{ \bar{h}(\bar{\varphi}, \bar{\sigma}) \mid \bar{f}(\bar{\varphi}, \bar{\sigma}) = 0 \} = \bar{H}(\bar{\lambda}^*, \bar{\varphi}^*) \quad (2.8)$$

which seems to require estimates of  $(\bar{\varphi}^* - \varphi^*, \bar{\lambda}^* - \lambda^*)$  to yield a bound on  $E$ . To obtain bounds on  $\bar{\varphi}^* - \varphi^*$  is in fact a harder problem than to estimate  $E$ . The situation is similar to minimization of a non strictly convex function where convergence of the minimum value may hold without having convergence of the minimizing points. In our case the error in the objective function can be small although the difference of the controls is large, *e.g.* near a point where  $H$  is not differentiable.

We shall see that the corresponding time dependent optimal control problem is useful for understanding regularizations and error estimates; in particular we use the time dependent formulation to derive bounds on  $E$  depending only on the difference of the two Hamiltonians along the *same path*, *i.e.* depending on  $H(\lambda, \varphi) - \bar{H}(\lambda, \bar{\varphi})$ , so that no estimate of  $\varphi^* - \bar{\varphi}^*$  or  $\lambda^* - \bar{\lambda}^*$  is needed. Let us now state and compare computational

methods for time dependent optimal control problems. Consider two controlled differential equations

$$\begin{aligned}\partial_t \varphi_t &= f(\varphi_t, \sigma_t), \\ \partial_t \bar{\varphi}_t &= \bar{f}(\bar{\varphi}_t, \bar{\sigma}_t),\end{aligned}$$

with solutions  $\varphi : [0, T] \rightarrow V$ , and  $\bar{\varphi} : [0, T] \rightarrow \bar{V}$ , and given initial values  $\varphi_0, \bar{\varphi}_0$ . Here,  $\partial_t$  denotes the partial derivative with respect to time and  $\varphi_t := \varphi(t)$ ,  $\sigma_t := \sigma(t)$ . The objective is to minimize

$$\begin{aligned}\min_{\sigma \in \mathcal{B}} \left\{ \int_0^T h(\varphi_t, \sigma_t) dt + g(\varphi_T) \mid \partial_t \varphi_t = f(\varphi_t, \sigma_t) \right\}, \quad \mathcal{B} &:= \{\sigma : [0, T] \rightarrow \mathcal{A}\}, \\ \min_{\bar{\sigma} \in \bar{\mathcal{B}}} \left\{ \int_0^T \bar{h}(\bar{\varphi}_t, \bar{\sigma}_t) dt + \bar{g}(\bar{\varphi}_T) \mid \partial_t \bar{\varphi}_t = \bar{f}(\bar{\varphi}_t, \bar{\sigma}_t) \right\}, \quad \bar{\mathcal{B}} &:= \{\bar{\sigma} : [0, T] \rightarrow \bar{\mathcal{A}}\}.\end{aligned}\tag{2.9}$$

These optimal control problems can be solved either directly using constrained minimization or by dynamic programming. The Lagrangian becomes

$$\mathcal{L} := \int_0^T \langle \lambda_t, f(\varphi_t, \sigma_t) - \partial_t \varphi_t \rangle + h(\varphi_t, \sigma_t) dt + g(\varphi_T)$$

and the constrained minimization method is based on the Pontryagin method

$$\begin{aligned}\partial_t \varphi_t &= f(\varphi_t, \sigma_t) \\ \partial_t \lambda_t &= -\langle \lambda_t, \partial f(\varphi_t, \sigma_t) \rangle + \partial h(\varphi_t, \sigma_t) \\ \sigma_t &\in \operatorname{argmin}_{\sigma \in \mathcal{A}} \{ \langle \lambda_t, f(\varphi_t, \sigma) \rangle + h(\varphi_t, \sigma) \}.\end{aligned}\tag{2.10}$$

This can be written as a Hamiltonian system

$$\begin{aligned}\partial_t \varphi_t &= \partial_1 H(\lambda_t, \varphi_t) \\ \partial_t \lambda_t &= -\partial_2 H(\lambda_t, \varphi_t)\end{aligned}\tag{2.11}$$

with  $\varphi_0$  given, and  $\lambda_T = \partial g(\varphi_T)$ .

The alternative dynamic programming method is based on the value functions  $u : V \times [0, T] \rightarrow \mathbb{R}$  and  $\bar{u} : \bar{V} \times [0, T] \rightarrow \mathbb{R}$ ,

$$\begin{aligned}u(\phi, \tau) &:= \inf_{\sigma \in \mathcal{B}} \left\{ \int_\tau^T h(\varphi_t, \sigma_t) dt + g(\varphi_T) \mid \partial_t \varphi_t = f(\varphi_t, \sigma_t), \varphi_\tau = \phi \in V \right\} \\ \bar{u}(\phi, \tau) &:= \inf_{\bar{\sigma} \in \bar{\mathcal{B}}} \left\{ \int_\tau^T \bar{h}(\bar{\varphi}_t, \bar{\sigma}_t) dt + \bar{g}(\bar{\varphi}_T) \mid \partial_t \bar{\varphi}_t = \bar{f}(\bar{\varphi}_t, \bar{\sigma}_t), \bar{\varphi}_\tau = \phi \in \bar{V} \right\},\end{aligned}\tag{2.12}$$

which solve the nonlinear Hamilton-Jacobi-Bellman equations

$$\begin{aligned}\partial_t u(\phi, t) + H(\partial u(\phi, t), \phi) &= 0, \quad u(\cdot, T) = g, \\ \partial_t \bar{u}(\phi, t) + \bar{H}(\partial \bar{u}(\phi, t), \phi) &= 0, \quad \bar{u}(\cdot, T) = \bar{g},\end{aligned}\tag{2.13}$$

with Hamiltonians defined as in (2.5).

The Hamilton-Jacobi formulation has two advantages and a severe disadvantage:

- + there is complete well posedness theory for Hamilton-Jacobi equations, based on viscosity solutions, see [15], although, in general, the value function is not everywhere differentiable, corresponding in the constrained optimization method to optimal backward paths  $\varphi^*$  that collide and hence the Lagrange multiplier  $\lambda^*$  becomes ill defined in a standard sense;
- + the Hamilton-Jacobi formulation finds a global minimum, while constrained minimization focuses on local minima;
- the drawback with dynamic programming is that the method is only computationally feasible for problems in low dimension,  $\bar{\varphi}_t \in \mathbb{R}^n$ , while constrained minimization is computable also for high dimensional problems where  $\bar{\varphi}$  is an approximation of a solution to a partial differential equation with  $n \gg 1$ .

Therefore the computational option is to use constrained minimization for problems in high dimension and we will choose a discretization of the stationary Hamiltonian system (2.6) to solve optimal design problems. However, we shall use the Hamilton-Jacobi equation in infinite dimension to understand regularizations and to derive error estimates. The additional structure extending our optimal design problems to dynamic programming problems is hence a useful theoretical tool. Note however that not all constrained optimal control problems have such extensions. Note also that solving the Hamiltonian system (2.11) is the method of characteristics for the Hamilton-Jacobi equation (2.13), with  $\lambda_t = \partial u(\varphi_t, t)$ .

**Example 2.2** (artificial time dependent optimal conductor). The time dependent extension of Example 2.1 is to find an optimal time dependent conductivity  $\sigma : \Omega \times [0, T] \rightarrow \{\sigma_-, \sigma_+\}$  to minimize the power loss under constraint of the parabolic equation

$$\partial_t \varphi = \operatorname{div}(\sigma \nabla \varphi),$$

where  $\varphi = \varphi(x, t)$ . The Lagrangian takes the form

$$\mathcal{L}(\sigma, \lambda, \varphi) := \int_0^T \int_{\partial\Omega} q(\varphi + \lambda) \, ds \, dt + \int_0^T \int_{\Omega} \underbrace{\sigma(\eta - \nabla \varphi \cdot \nabla \lambda)}_v - \partial_t \varphi \lambda \, dx \, dt,$$

with  $\lambda = \lambda(x, t)$  and the Hamiltonian

$$\begin{aligned} H(\lambda, \varphi) &= \min_{\sigma: \Omega \rightarrow \{\sigma_{\pm}\}} \left\{ \int_{\Omega} \sigma v \, dx + \int_{\partial\Omega} q(\varphi + \lambda) \, ds \right\} \\ &= \int_{\Omega} \underbrace{\min_{\sigma \in \sigma_{\pm}} \{\sigma v\}}_{\mathfrak{h}(v)} \, dx + \int_{\partial\Omega} q(\varphi + \lambda) \, ds, \end{aligned}$$

where we have introduced the function  $\mathfrak{h}(v) = \min_{\sigma \in \sigma_{\pm}} \{\sigma v\}$ . The value function

$$u(\phi, \tau) = \inf_{\sigma} \left\{ \int_{\tau}^T \int_{\partial\Omega} q\varphi \, ds + \eta \int_{\Omega} \sigma \, dx \, dt \mid \partial_t \varphi = \operatorname{div}(\sigma \nabla \varphi), \varphi_{\tau} = \phi \right\}$$

yields the infinite dimensional Hamilton-Jacobi equation

$$\partial_t u(\phi, t) + H(\partial u(\phi, t), \phi) = 0 \quad t < T, \quad u(\cdot, T) = 0,$$

using the Gateaux derivative  $\partial u(\phi, t)$  of the functional  $u(\phi, t)$  in  $L^2(\Omega)$ . The corresponding Hamiltonian system is the parabolic system

$$\begin{aligned} \int_{\Omega} \partial_t \varphi w + \mathfrak{h}'(\eta - \nabla \varphi \cdot \nabla \lambda) \nabla \varphi \cdot \nabla w \, dx &= \int_{\partial \Omega} qw \, ds, \quad \varphi(\cdot, 0) \text{ given,} \\ \int_{\Omega} -\partial_t \lambda v + \mathfrak{h}'(\eta - \nabla \varphi \cdot \nabla \lambda) \nabla \lambda \cdot \nabla v \, dx &= \int_{\partial \Omega} qv \, ds, \quad \lambda(\cdot, T) = 0, \end{aligned}$$

for all test functions  $v, w \in V \equiv \{v \in H^1(\Omega) : \int_{\Omega} v \, dx = 0\}$ .

## 2.1. Derivation of approximation error

We simplify by considering the case when  $\bar{g} = g$ . A similar approximation study is in [32]. Let us first derive the approximation error,  $\bar{u}(\bar{\varphi}_0, 0) - u(\varphi_0, 0) =: \tilde{E}$ , of the value functions (2.12) given solutions  $\sigma, \bar{\sigma}$  and  $\varphi, \bar{\varphi}$  to the time dependent optimal control problems (2.9)

$$\begin{aligned} \tilde{E} &:= \int_0^T \bar{h}(\bar{\varphi}_t, \bar{\sigma}_t) \, dt + g(\bar{\varphi}_T) - \left( \int_0^T h(\varphi_t, \sigma_t) \, dt + g(\varphi_T) \right) \\ &= \int_0^T \bar{h}(\bar{\varphi}_t, \bar{\sigma}_t) \, dt + u(\bar{\varphi}_T, T) - u(\varphi_0, 0) \\ &= \int_0^T \bar{h}(\bar{\varphi}_t, \bar{\sigma}_t) \, dt + u(\bar{\varphi}_T, T) - u(\bar{\varphi}_0, 0) + u(\bar{\varphi}_0, 0) - u(\varphi_0, 0). \end{aligned} \tag{2.14}$$

To simplify the analysis choose the initial data for the exact path to coincide with the initial data for the approximate path, *i.e.*  $\varphi_0 = \bar{\varphi}_0$ . Also assume that  $u$  is Gateaux differentiable; the general case with only sub differentiable  $u$  is in Section 2.2. Then the right hand side simplifies to

$$\begin{aligned} \int_0^T du(\bar{\varphi}_t, t) + \int_0^T \bar{h}(\bar{\varphi}, \bar{\sigma}) \, dt &= \int_0^T \partial_t u(\bar{\varphi}_t, t) + \langle \partial u(\bar{\varphi}_t, t), \bar{f}(\bar{\varphi}_t, \bar{\sigma}_t) \rangle + \bar{h}(\bar{\varphi}_t, \bar{\sigma}_t) \, dt \\ &\geq \int_0^T \underbrace{-H(\partial u(\bar{\varphi}_t, t), \bar{\varphi}_t)}_{= \partial_t u(\bar{\varphi}_t, t)} + \bar{H}(\partial u(\bar{\varphi}_t, t), \bar{\varphi}_t) \, dt, \end{aligned} \tag{2.15}$$

where the inequality follows from the definition (2.5) of  $\bar{H}$ . Note that the Pontryagin principle

$$\langle \partial \bar{u}(\bar{\varphi}_t, t), \bar{f}(\bar{\varphi}_t, \bar{\sigma}_t) \rangle + \bar{h}(\bar{\varphi}_t, \bar{\sigma}_t) = \bar{H}(\partial \bar{u}(\bar{\varphi}_t, t), \bar{\varphi}_t)$$

is not applicable here since we have  $\partial u$  instead of  $\partial \bar{u}$ . The more general case with  $\bar{g} \neq g$  yields the additional error term

$$(g - \bar{g})(\bar{\varphi}_T)$$

to the right hand side in (2.15).

Similarly, exchange the role of the exact value function along the approximate path,  $(u, \bar{\varphi})$ , with the approximate value function along the exact path,  $(\bar{u}, \varphi)$ , to obtain an upper bound on  $\tilde{E}$ . This requires a new step; to give meaning to  $\bar{u}$  along the exact path  $\varphi_t$ . For this purpose we introduce the projection  $P : V \rightarrow \bar{V}$ . We have,

using  $\bar{\varphi}_0 = \varphi_0 = P\varphi_0$ ,

$$\begin{aligned}
-\tilde{E} &= \int_0^T h(\varphi_t, \sigma_t) dt + g(\varphi_T) - \left( \int_0^T \bar{h}(\bar{\varphi}_t, \bar{\sigma}_t) dt + g(\bar{\varphi}_T) \right) \\
&= \int_0^T h(\varphi_t, \sigma_t) dt + g(\varphi_T) + \bar{u}(P\varphi_T, T) - g(P\varphi_T) - \bar{u}(\bar{\varphi}_0, 0) \\
&= \int_0^T h(\varphi_t, \sigma_t) dt + \bar{u}(P\varphi_T, T) - \bar{u}(P\varphi_0, 0) + g(\varphi_T) - g(P\varphi_T).
\end{aligned} \tag{2.16}$$

The first three terms in the right hand side become

$$\begin{aligned}
\int_0^T d\bar{u}(P\varphi_t, t) + \int_0^T h(\varphi_t, \sigma_t) dt &= \int_0^T \partial_t \bar{u}(P\varphi_t, t) + \langle \partial \bar{u}(P\varphi_t, t), Pf(\varphi_t, \sigma_t) \rangle + h(\varphi_t, \sigma_t) dt \\
&= \int_0^T \partial_t \bar{u}(P\varphi_t, t) + \langle P\partial \bar{u}(P\varphi_t, t), f(\varphi_t, \sigma_t) \rangle + h(\varphi_t, \sigma_t) dt \\
&\geq \int_0^T \partial_t \bar{u}(P\varphi_t, t) + H(\partial \bar{u}(P\varphi_t, t), \varphi_t) dt \\
&= \int_0^T -\bar{H}(\partial \bar{u}(P\varphi_t, t), P\varphi_t) + H(\partial \bar{u}(P\varphi_t, t), \varphi_t) dt.
\end{aligned} \tag{2.17}$$

Combining (2.16) and (2.17) we now have

$$\begin{aligned}
\tilde{E} &\leq \int_0^T \bar{H}(\partial \bar{u}(P\varphi_t, t), P\varphi_t) - H(\partial \bar{u}(P\varphi_t, t), \varphi_t) dt - g(\varphi_T) + g(P\varphi_T) \\
&= \int_0^T (\bar{H} - H)(\partial \bar{u}(P\varphi_t, t), P\varphi_t) dt \\
&\quad + \int_0^T H(\partial \bar{u}(P\varphi_t, t), P\varphi_t) - H(\partial \bar{u}(P\varphi_t, t), \varphi_t) dt + g(P\varphi_T) - g(\varphi_T).
\end{aligned} \tag{2.18}$$

Assume now that  $h, g$  and  $H$  are Lipschitz continuous in  $V$ , with respect to the variable  $\varphi$ . Then the projection error terms in the right hand side of (2.18) are

$$\begin{aligned}
\tilde{E}_P &:= \left| \int_0^T H(\partial \bar{u}(P\varphi_t, t), P\varphi_t) - H(\partial \bar{u}(P\varphi_t, t), \varphi_t) dt + g(P\varphi_T) - g(\varphi_T) \right| \\
&= \sup_{t \in [0, T]} \|P\varphi_t - \varphi_t\|_V \mathcal{O}(T).
\end{aligned}$$

Combine (2.15) and (2.18) to obtain

$$\int_0^T (\bar{H} - H)(\partial u(\bar{\varphi}_t, t), \bar{\varphi}_t) dt \leq \tilde{E} \leq \int_0^T (\bar{H} - H)(\partial \bar{u}(P\varphi_t, t), P\varphi_t) dt + \tilde{E}_P. \tag{2.19}$$

**Remark 2.3** (no minimizers). If there are no minimizers to (2.9), then for every  $\varepsilon > 0$ , we can choose controls  $\sigma, \bar{\sigma}$  with corresponding states  $\varphi, \bar{\varphi}$  such that

$$E_{lhs} - \varepsilon \leq \tilde{E} \leq E_{rhs} + \varepsilon$$

with  $E_{lhs}, E_{rhs}$  being the left and right hand sides of (2.19).



To estimate the error in the case of time independent optimal control problems with solutions  $\varphi$  and  $\bar{\varphi}$ , we assume that the time dependent control problems with initial data  $\varphi_0 = \bar{\varphi}_0$  for some given  $\bar{\varphi}_0 \in \bar{V}$  (close to some approximation of  $\bar{\varphi}$ ) asymptotically have the same solutions as the time independent versions, *i.e.*

$$\begin{aligned} \lim_{T \rightarrow \infty} \inf_{\sigma \in \mathcal{A}} \left\{ \frac{1}{T} \int_0^T h(\varphi_t, \sigma_t) dt \mid \partial_t \varphi_t = f(\varphi_t, \sigma_t), \varphi_0 = \bar{\varphi}_0 \right\} &= \inf_{\sigma \in B} \{h(\varphi, \sigma) \mid f(\varphi, \sigma) = 0\} \\ \lim_{T \rightarrow \infty} \inf_{\bar{\sigma} \in \bar{\mathcal{A}}} \left\{ \frac{1}{T} \int_0^T \bar{h}(\bar{\varphi}_t, \bar{\sigma}_t) dt \mid \partial_t \bar{\varphi}_t = f(\bar{\varphi}_t, \bar{\sigma}_t), \bar{\varphi}_0 = \varphi_0 \right\} &= \inf_{\bar{\sigma} \in \bar{B}} \{\bar{h}(\bar{\varphi}, \bar{\sigma}) \mid \bar{f}(\bar{\varphi}, \bar{\sigma}) = 0\}, \end{aligned} \quad (2.20)$$

which implies:

**Theorem 2.4.** *Assume that (2.20) holds, and that  $g$  is a bounded function. Then the error  $E$  satisfies*

$$E = \lim_{T \rightarrow \infty} \frac{\tilde{E}}{T},$$

where  $\tilde{E}$ , given in (2.14), (2.16) and Section 2.2, only depends on the difference between the Hamiltonians  $H$  and  $\bar{H}$  along a solution path and on the projection error  $\|\varphi_t - P\varphi_t\|_V$ , but not on the error between the paths  $(\varphi - \bar{\varphi}, \lambda - \bar{\lambda})$ .

## 2.2. Non differentiable solution to Hamilton-Jacobi equations

Solutions to Hamilton-Jacobi equations are in general not differentiable. Let us extend the derivation (2.15), (2.17) to a case when  $u$  is not differentiable. The theory of viscosity solutions to Hamilton-Jacobi equations gives well posedness for solutions, which are continuous but not necessarily differentiable, *cf.* [2,3,20]. This theory is now rather complete in the finite dimensional setting, *cf.* [16]. Let us therefore consider a case when  $V$  and  $\bar{V}$  are two finite element spaces, with  $\bar{V} \subset V$  so that the corresponding Hamilton-Jacobi equations are defined on finite dimensional spaces. By theory in *e.g.* [10], which covers finite dimensional optimal control systems, it follows that the value functions,  $u$  and  $\bar{u}$ , in all problems treated in this report, are semiconcave on, respectively,  $V \times [0, T]$  and  $\bar{V} \times [0, T]$ . One of the requirements for semiconcavity in [10] is that the flux (here  $f(\varphi, \sigma)$ ) and its spatial derivative ( $\partial_\varphi f(\varphi, \sigma)$ ) must both be Lipschitz in the state variable,  $\varphi$ , with a constant independent of  $\sigma$ . This can be verified for the fluxes in the present problems using a discrete  $H^2$  norm which is equivalent to the Euclidean norm since the spaces are finite dimensional. The other requirements for semiconcavity are easily verified. Therefore the error estimate (2.19) extends to the general case when  $u, \bar{u}$  are viscosity solutions that are not differentiable functions as follows. If  $u$  is a non differentiable semiconcave solution to a Hamilton-Jacobi equation the definition of viscosity solution reduces to

$$\begin{aligned} q + H(p, \varphi) &\geq 0 \quad \text{for all } (p, q) \in D^+u(\varphi, t) \\ u(\cdot, T) &= g, \end{aligned} \quad (2.21)$$

where  $D^+u(x) := \{z \in V \times \mathbb{R} : u(y+x) - u(x) - \langle z, y \rangle \leq o(\|y\|)\}$  is the super differential of  $u$  at the point  $x = (\varphi, t)$ . This means that in (2.15) we can for each  $t$  choose a point  $(p, q) \in D^+u(\bar{\varphi}_t, t)$  so that

$$\int_0^T du(\bar{\varphi}_t, t) + \int_0^T \bar{h}(\bar{\varphi}_t, t) dt = \int_0^T q + \langle p, \bar{f}(\bar{\varphi}_t, \bar{\sigma}_t) \rangle + \bar{h}(\bar{\varphi}_t, t) dt,$$

and by the definition (2.8) of  $\bar{H}$  we have

$$\int_0^T q + \langle p, \bar{f}(\bar{\varphi}_t, \bar{\sigma}_t) \rangle + \bar{h}(\bar{\varphi}_t, t) dt \geq \int_0^T (q + \bar{H}(p, \bar{\varphi}_t)) dt.$$

Finally, by (2.21),

$$\int_0^T (q + \bar{H}(p, \bar{\varphi}_t)) dt \geq \int_0^T (-H + \bar{H})(p, \bar{\varphi}_t) dt.$$

The analogous formulation holds for  $\bar{u}$ . Consequently (2.19) holds for some  $(p, q) \in D^+u(\bar{\varphi}_t, t)$  replacing  $(\partial u(\bar{\varphi}_t, t), \partial_t u(\bar{\varphi}_t, t))$  and some  $(\bar{p}, \bar{q}) \in D^+\bar{u}(P\varphi_t, t)$  replacing  $(\partial \bar{u}(P\varphi_t, t), \partial_t \bar{u}(P\varphi_t, t))$ .

The present analysis is, however, in principle valid even when we let  $V$  be an infinite dimensional Hilbert space, although existence and semiconcavity of solutions is not derived in full generality. For instance parabolic problems with fluxes  $f$  where the terms including second order derivatives depend on the control (as here) seem to not have been studied. In [8,9] the case of semilinear control problems is treated. This theory is used in [32] to perform analysis similar to the one in this section when  $V$  is infinite dimensional. For theory involving more nonlinearly operators, see *e.g.* [34].

### 2.3. Derivation of regularization error

In the examples treated in this report the Hamiltonian,  $H$ , is nondifferentiable, as the function  $\mathfrak{h}$  is nondifferentiable. Therefore it can not be expected that using the Hamiltonian system (2.11), even in the discretized case, would give an optimal path which could be used to determine the value of  $\bar{u}$ . For this reason we will consider solutions to the regularized Hamiltonian system

$$\begin{aligned} \partial_t \bar{\varphi}_t &= \partial_1 \bar{H}_\delta(\bar{\lambda}_t, \bar{\varphi}_t), \\ \partial_t \bar{\lambda}_t &= -\partial_2 \bar{H}_\delta(\bar{\lambda}_t, \bar{\varphi}_t), \end{aligned} \tag{2.22}$$

where  $\bar{H} \equiv \bar{H}_\delta$  is a smooth regularization of  $H$ , which is also concave in the  $\bar{\lambda}$  variable, for  $\delta > 0$  and  $\bar{H}_0 = H$ . To find an optimal control problem corresponding to (2.22), we may relate to the Hamiltonian,  $\bar{H}_\delta$ , the Legendre transform in the  $\bar{\lambda}$  variable:

$$L(\bar{\varphi}, \bar{l}) \equiv \sup_{\bar{\lambda} \in \bar{V}} \{ -\langle \bar{l}, \bar{\lambda} \rangle + \bar{H}_\delta(\bar{\lambda}, \bar{\varphi}) \}. \tag{2.23}$$

The function  $L$  is a running cost for the following variational problem:

$$\bar{u}_\delta(\phi, t_0) = \inf \left\{ \int_{t_0}^T L(\bar{\varphi}_t, \partial_t \bar{\varphi}_t) dt + g(\bar{\varphi}_T) \mid \bar{\varphi}_{t_0} = \phi \right\}, \tag{2.24}$$

where the infimum is taken over all absolutely continuous functions  $\varphi : [t_0, T] \rightarrow \bar{V}$ . This can be formulated as the optimal control problem

$$\inf_{\bar{\sigma} \in L^1([t_0, T]; \bar{V})} \left\{ \int_{t_0}^T L(\bar{\varphi}_t, \bar{\sigma}) dt + g(\bar{\varphi}_T) \mid \partial_t \bar{\varphi} = \bar{\sigma}, \bar{\varphi}_{t_0} = \phi \right\}, \tag{2.25}$$

and its associated Hamiltonian system is (2.22), since the Legendre transform of  $L$  gives

$$\bar{H}_\delta(\bar{\lambda}, \bar{\varphi}) = -\sup_{\bar{\sigma} \in \bar{V}} \{ -\langle \bar{\lambda}, \bar{\sigma} \rangle - L(\bar{\varphi}, \bar{\sigma}) \} = \inf_{\bar{\sigma} \in \bar{V}} \{ \langle \bar{\lambda}, \bar{\sigma} \rangle + L(\bar{\varphi}, \bar{\sigma}) \}. \tag{2.26}$$

Note that the equivalent problem with time reversed,  $s = T - t$ , corresponds to

$$\inf \left\{ \int_{t_0}^T \tilde{L}(\bar{\varphi}_s, \partial_s \bar{\varphi}_s) ds + g(\bar{\varphi}_0) \mid \bar{\varphi}_T = \phi \right\}$$

where

$$\tilde{L}(\bar{\varphi}, \bar{l}) := \sup_{\bar{\lambda} \in \bar{V}} \{ \langle \bar{l}, \bar{\lambda} \rangle + \bar{H}_\delta(\bar{\lambda}, \bar{\varphi}) \} = L(\bar{\varphi}, -\bar{l})$$

is the usual Legendre transformation of the convex function  $-\bar{H}_\delta(\cdot, \bar{\varphi})$ .

The problems described in Sections 3.1 and 3.3 have concave Hamiltonians which are not coercive, which implies that their corresponding running costs,  $L$ , takes the value  $+\infty$  for some arguments. Such running costs are treated in [14], where it is shown that the problem (2.24), with the aforementioned running cost, has a minimizer  $\bar{\varphi} : [t_0, T] \rightarrow \bar{V}$ . Furthermore, such a minimizer solves the Hamiltonian system (2.22) together with a function  $\bar{\lambda}$ , granting existence of a solution to this Hamiltonian system.

The value function  $\bar{u}_\delta$  is a viscosity solution to the Hamilton-Jacobi equation

$$\partial_t \bar{u}_\delta(\phi, t) + \bar{H}_\delta(\partial \bar{u}_\delta(\phi, t), \phi) = 0, \quad \bar{u}_\delta(\cdot, T) = \bar{g}(\cdot).$$

This result can easily be obtained from Theorem 6.4.5 in [10]. This theorem treats only running costs,  $L$ , with finite values, but the proof is basically unchanged by allowing the running costs of interest here. The error estimate of Theorem 2.4 is applicable both to estimate  $u - \bar{u}_\delta$ , with approximation of both  $V$  and  $H$ , and to  $\bar{u}_0 - \bar{u}_\delta$ , with approximation only of  $H$ . We may alternatively estimate the difference between  $\bar{u}$  and  $\bar{u}_\delta$  by using known results of the Hamilton-Jacobi equations

$$\begin{aligned} \partial_t \bar{u}_\delta(\phi, t) + \bar{H}_\delta(\partial \bar{u}_\delta(\phi, t), \phi) &= 0, \\ \partial_t \bar{u}_0(\phi, t) + \bar{H}_0(\partial \bar{u}_0(\phi, t), \phi) &= 0, \end{aligned}$$

and the fact that  $\bar{u}_0(\cdot, T) = \bar{u}_\delta(\cdot, T) = \bar{g}(\cdot)$ ; the comparison principle for viscosity solutions gives that

$$\|\bar{u} - \bar{u}_\delta\|_{\mathcal{C}(\bar{V} \times [0, T])} \leq T \|\bar{H}_0 - \bar{H}_\delta\|_{\mathcal{C}(\bar{V} \times \bar{V})},$$

see [33].

The value of  $\bar{u}_\delta$  for a case with constant solutions  $\bar{\varphi}^*$  and  $\bar{\lambda}^*$  to (2.22) is approximately  $T \cdot L(\bar{\varphi}^*, 0)$  when  $T$  is large (so that we can neglect  $\bar{g}(\bar{\varphi}_T)$ ). The definition of  $L$  gives that

$$L(\bar{\varphi}^*, 0) = \bar{H}_\delta(\hat{\lambda}, \bar{\varphi}^*), \tag{2.27}$$

where  $\hat{\lambda}$  is the maximizer of  $\bar{H}_\delta(\cdot, \bar{\varphi}^*)$ . As the Hamiltonian system for constant solutions is

$$\partial_1 \bar{H}_\delta = \partial_2 \bar{H}_\delta = 0, \tag{2.28}$$

and  $\bar{H}_\delta$  is concave in the  $\lambda$  argument we have that  $\hat{\lambda} = \bar{\lambda}^*$ . Hence the candidate for a value approximating (2.8) is  $\bar{H}_\delta(\bar{\lambda}^*, \bar{\varphi}^*)$ , where  $\bar{\varphi}^*$  and  $\bar{\lambda}^*$  are solutions to (2.28).

#### 2.4. Smoothed Hamiltonian as Tikhonov regularization

Tikhonov regularization in optimal control consists of penalization functions containing the control variable, and sometimes derivatives of it. The Legendre transform makes it possible to relate the regularized Hamiltonian with such penalization functions. At every point  $(\bar{\lambda}, \bar{\varphi})$  where the derivative  $\partial_1 \bar{H}$  exists, there is an element  $\bar{\sigma} \in \bar{\mathcal{A}}$  such that  $\partial_1 \bar{H}(\bar{\lambda}, \bar{\varphi}) = \bar{f}(\bar{\varphi}, \bar{\sigma})$ , see [7, 21]. We now make the following assumption:

$$\partial_1 \bar{H}_\delta(\bar{\lambda}, \bar{\varphi}) \in \text{co } \bar{f}(\bar{\varphi}, \bar{\mathcal{A}}) \text{ for all } (\bar{\lambda}, \bar{\varphi}) \in \bar{V} \times \bar{V}, \tag{2.29}$$

where  $\text{co}$  denotes the convex hull. This assumption holds for all regularizations we will consider in this paper.

It also holds for regularizations by convolution

$$\bar{H}_\delta(\bar{\lambda}, \bar{\varphi}) = \int_V \bar{H}(\bar{\lambda} - \bar{y}, \bar{\varphi}) \eta(\bar{y}) d\bar{y},$$

where  $\eta$  is a mollifier with support in the closed ball with radius  $\delta$  centered at the origin.

By the definition of  $L$  in (2.23), it follows that  $L(\bar{\varphi}, \bar{l}) = +\infty$ , if

$$\bar{l} \notin \overline{\{\partial_1 \bar{H}_\delta(\bar{\lambda}, \bar{\varphi}) \mid \bar{\lambda} \in \bar{V}\}}, \quad (2.30)$$

where the big bar denotes set closure. Since the sets  $\bar{f}(\bar{\varphi}, \bar{\mathcal{A}})$  are supposed to be closed, the assumption in (2.29) implies that

$$\overline{\{\partial_1 \bar{H}_\delta(\bar{\lambda}, \bar{\varphi}) \mid \bar{\lambda} \in \bar{V}\}} \subset \text{co } \bar{f}(\bar{\varphi}, \bar{\mathcal{A}}).$$

The infimum in (2.26) may therefore be reduced to an infimum over the set  $\text{co } \bar{f}(\bar{\varphi}, \bar{\mathcal{A}})$ . Furthermore, the function  $L(\bar{\varphi}, \cdot)$  is lower semicontinuous, see [31], and therefore the infimum is attained:

$$\bar{H}_\delta(\bar{\lambda}, \bar{\varphi}) = \min \{ \langle \bar{\lambda}, \bar{l} \rangle + L(\bar{\varphi}, \bar{l}) \mid \bar{l} \in \text{co } \bar{f}(\bar{\varphi}, \bar{\mathcal{A}}) \}. \quad (2.31)$$

In order to see the relation with ordinary Tikhonov regularization, let us now assume that the sets  $\bar{f}(\bar{\varphi}, \bar{\mathcal{A}})$  are convex. Then the regularized Hamiltonian satisfies

$$\bar{H}_\delta(\bar{\lambda}, \bar{\varphi}) = \min_{\bar{\sigma} \in \bar{\mathcal{A}}} \{ \langle \bar{\lambda}, \bar{f}(\bar{\varphi}, \bar{\sigma}) \rangle + \bar{h}_\delta(\bar{\varphi}, \bar{\sigma}) \},$$

where

$$\bar{h}_\delta(\bar{\varphi}, \bar{\sigma}) = L(\bar{\varphi}, \bar{f}(\bar{\varphi}, \bar{\sigma})).$$

Comparing with the expression for the Hamiltonian without regularization,  $\bar{H}$ , in (2.5), we see that the difference  $\bar{h}_\delta(\bar{\varphi}, \bar{\sigma}) - \bar{h}(\bar{\varphi}, \bar{\sigma})$  is a penalization function of Tikhonov-type for the optimal control problem. Relation (2.31) is inspired by such a comparison for an example in [11].

### 3. THREE DIFFERENT CONDUCTION DESIGNS

In the following sections we will study numerical approximation of three optimal control problems related to optimal design, using the Hamiltonian system (2.11) with a regularized Hamiltonian. By optimal design we mean that we seek to optimize some physical property, such as energy loss, by distributing a discrete valued control, such as an electric conductor, on a fixed domain. The problems considered are: to optimally design an electric conductor, to design an elastic domain and to reconstruct the interior of an object from measured electrical surface currents.

All three problems produce non-smooth controls due to lack of regularity in the Hamiltonian, which for the success of a discrete Pontryagin Principle needs to be regularized. However, in the time-independent setting, even a smooth Hamiltonian may result in an ill-posed minimization problem in the sense that one cannot find a minimizer as the limit of a minimizing sequence. The existence of such a minimizer essentially depends on the weak lower semicontinuity of the Hamiltonian, which in the standard theory of variational calculus is a necessary condition closely connected to quasi-convexity [18].

Ill-posed problems related to optimal design, as the one described in Section 3.2, has been studied extensively in the context of relaxation by quasi-convexification and homogenization in [1, 22, 24, 26–28].

In Section 3.3, we study the now classical problem of impedance tomography, reviewed in [5]. Since there seems to be no algorithm to directly compute the quasi-convexification of a general problem we will here show that a simple regularization, which in Sections 3.1 and 3.2 much resembles a Tikhonov regularization, can produce good approximations in the value functions, with the advantage that, by the Pontryagin approach, the Newton method with a sparse Hessian can be used.

### 3.1. Concave maximization

A concave problem of electric conduction is to distribute a given amount of conducting material in a domain  $\Omega \subset \mathbb{R}^d$  in order to minimize the power production for a surface current  $q$ , satisfying  $\int_{\partial\Omega} q \, ds = 0$ : let  $C$  be the given amount of material and find an optimal conduction distribution  $\sigma : \Omega \rightarrow \{\sigma_-, \sigma_+\}$ , where  $\sigma_{\pm} > 0$ , such that

$$\min_{\sigma} \left\{ \int_{\partial\Omega} q\varphi \, ds \mid \operatorname{div}(\sigma\nabla\varphi) = 0 \text{ in } \Omega, \sigma \frac{\partial\varphi}{\partial n} \Big|_{\partial\Omega} = q, \int_{\Omega} \sigma \, dx = C \right\}. \quad (3.1)$$

Here  $\partial/\partial n$  denotes the normal derivative and  $ds$  is the surface measure on  $\partial\Omega$  and  $\varphi \in V \equiv \{v \in H^1(\Omega) : \int_{\Omega} v \, dx = 0\}$  is the electric potential. Note that (3.1) implies that the power loss satisfies

$$\int_{\partial\Omega} q\varphi \, ds = - \int_{\Omega} \operatorname{div}(\sigma\nabla\varphi)\varphi \, dx + \int_{\partial\Omega} \sigma \frac{\partial\varphi}{\partial n} \varphi \, ds = \int_{\Omega} \sigma |\nabla\varphi|^2 \, dx.$$

For simplicity, let  $\eta > 0$  be a constant, associated to the given amount of material, and replace (3.1) with the easier problem to find an optimal conduction distribution such that

$$\min_{\sigma} \left\{ \int_{\partial\Omega} q\varphi \, ds + \eta \int_{\Omega} \sigma \, dx \mid \operatorname{div}(\sigma\nabla\varphi) = 0 \text{ in } \Omega, \sigma \frac{\partial\varphi}{\partial n} \Big|_{\partial\Omega} = q \right\}. \quad (3.2)$$

Observe, that although there exists a corresponding multiplier  $\eta$  for each volume constraint  $C$ , the converse may not be true.

The Lagrangian takes the form

$$\int_{\Omega} \sigma \underbrace{(\eta - \nabla\varphi \cdot \nabla\lambda)}_v \, dx + \int_{\partial\Omega} q(\varphi + \lambda) \, ds$$

and the Hamiltonian becomes

$$H(\lambda, \varphi) = \min_{\sigma} \left\{ \int_{\Omega} \sigma v \, dx + \int_{\partial\Omega} q(\varphi + \lambda) \, ds \right\} = \int_{\Omega} \underbrace{\min_{\sigma} \{\sigma v\}}_{\mathfrak{h}(v)} \, dx + \int_{\partial\Omega} q(\varphi + \lambda) \, ds \quad (3.3)$$

with the concave regularization

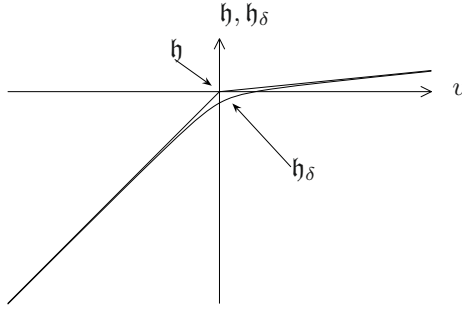
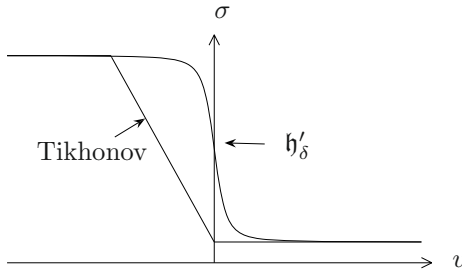
$$\bar{H}_{\delta}(\lambda, \varphi) = \int_{\Omega} \mathfrak{h}_{\delta}(\eta - \nabla\varphi \cdot \nabla\lambda) \, dx + \int_{\partial\Omega} q(\varphi + \lambda) \, ds, \quad (3.4)$$

depending on a smooth approximation,  $\mathfrak{h}_{\delta} \in C^2(\mathbb{R})$ , for  $\delta > 0$ , of the Lipschitz continuous and monotonically increasing function  $\mathfrak{h}$  (with discontinuous derivative  $\mathfrak{h}'$  at the origin, see Fig. 1). Note that  $\mathfrak{h}_0 = \mathfrak{h}$ . In this case the regularization  $\bar{H}_{\delta}$  is therefore similar to a Tikhonov regularization with penalty  $\delta \int_{\Omega} \sigma^2 dx$ , see Figure 2.

Note that  $\sigma$  need not to be restricted to discrete values in (3.2), since  $\sigma : \Omega \rightarrow [\sigma_-, \sigma_+]$  will lead to the same Hamiltonian.

By symmetry in the stationary Hamiltonian system (2.6), the primal and dual variables are equal,  $\lambda = \varphi$ , and the Hamiltonian system for the electric potential reduces to finite element discretizations of the nonlinear elliptic partial differential equation

$$\operatorname{div}(\mathfrak{h}'_{\delta}(\eta - |\nabla\varphi|^2)\nabla\varphi) = 0 \text{ in } \Omega, \quad \mathfrak{h}'_{\delta} \frac{\partial\varphi}{\partial n} \Big|_{\partial\Omega} = q, \quad (3.5)$$

FIGURE 1. The function  $\mathfrak{h}$  and its regularization  $\mathfrak{h}_\delta$  with respect to  $v$ .FIGURE 2. A Tikhonov type penalty  $\delta \int_{\Omega} \sigma^2 dx$  compared to the approximation  $\mathfrak{h}'_\delta$ .

which can be formulated as the concave maximization problem:  $\bar{\varphi} \in \bar{V}$  is the unique maximizer of

$$\bar{H}_\delta(\bar{\varphi}) = \int_{\Omega} \mathfrak{h}_\delta(\eta - |\nabla \bar{\varphi}|^2) dx + 2 \int_{\partial\Omega} q \bar{\varphi} ds, \quad (3.6)$$

where  $\bar{H}_\delta(\bar{\varphi})$  means  $\bar{H}_\delta(\bar{\varphi}, \bar{\varphi})$  and  $\bar{V} \subset V$  denotes a finite element subspace, while  $\varphi \in V$  is the unique maximizer of

$$H(\varphi) = \int_{\Omega} \mathfrak{h}_0(\eta - |\nabla \varphi|^2) dx + 2 \int_{\partial\Omega} q \varphi ds, \quad (3.7)$$

where  $H(\varphi)$  means  $H(\varphi, \varphi)$ .

An advantage with the Pontryagin approach (3.5) is that the Hessian  $D^2 \bar{H}_\delta$  can be determined explicitly and is sparse, so that the Newton method can be used for iterative solution of (3.5). In fact, the Newton method works well to solve the finite element version of (3.5) by successively decreasing  $\delta$ , see Section 4.1.

Since  $\sigma$  can be determined explicitly by the Pontryagin method an alternative approach would be to solve (3.5) according to the scheme

$$\sigma_{i+1} = \mathfrak{h}'_\delta(\eta - |\nabla \bar{\varphi}_i|^2) \quad (3.8)$$

where  $\bar{\varphi}_i$  solves

$$\operatorname{div}(\sigma_i \nabla \bar{\varphi}_i(x)) = 0, \quad x \in \Omega, \quad \sigma_i \frac{\partial \bar{\varphi}_i}{\partial n} \Big|_{\partial\Omega} = q, \quad (3.9)$$

given an initial guess  $\sigma_0$ . This type of scheme, which essentially is the Jacobi method, is highly unstable with respect to the initial guess since information from the Hessian is lost. In Section 3.2, we will however use this method, with  $\delta = 0$ , as a post-processing method to eliminate areas of intermediate density generated by the Newton method. These iterations are allowed as long as the value of the Hamiltonian stays relatively unchanged. As pointed out in [22] convergence to a global maximum of (3.6) by iterating in  $\sigma \equiv \mathfrak{h}'_0$  and  $\varphi$  separately, can not be guaranteed since it can be compared to minimizing  $f(x, y) = |x - y| + \frac{1}{2}|x + y|$  by iterating in  $x$  and  $y$  separately; such iterations would terminate at  $x = y$  although  $f$  is convex.

### 3.2. Non-concave maximization

Consider the conduction problem (3.2) where the objective now is changed to maximize the power production

$$\max_{\sigma} \left\{ \int_{\partial\Omega} q\varphi \, ds + \eta \int_{\Omega} \sigma \, dx \mid \operatorname{div}(\sigma \nabla \varphi) = 0 \text{ in } \Omega, \quad \sigma \frac{\partial \varphi}{\partial n} \Big|_{\partial\Omega} = q \right\}.$$

A problem with the same qualitative property of nonconcave maximization is to maximize the torsional rigidity of the cross section  $\Omega$  of an infinitely long elastic bar

$$\max_{\sigma} \left\{ \int_{\Omega} \varphi \, dx + \eta \int_{\Omega} \sigma \, dx \mid -\operatorname{div}(\sigma \nabla \varphi) = 1 \text{ in } \Omega, \quad \varphi|_{\partial\Omega} = 0 \right\}, \quad (3.10)$$

with shear moduli  $\sigma^{-1}$ , see [1,22,24].

The maximization problem (3.10) has the Lagrangian

$$\int_{\Omega} (\lambda + \varphi) \, dx + \int_{\Omega} \sigma \underbrace{(\eta - \nabla \varphi \cdot \nabla \lambda)}_v \, dx$$

and the Hamiltonian

$$H(\lambda, \varphi) = \int_{\Omega} (\lambda + \varphi) \, dx + \int_{\Omega} \underbrace{\max_{\sigma} \{\sigma v\}}_{\mathfrak{h}(v)} \, dx$$

which, as in Section 3.1, is regularized by  $\bar{H}_{\delta}$  with the  $C^2$ -approximation  $\mathfrak{h}_{\delta}$  of the Lipschitz continuous function  $\mathfrak{h}$ . Similarly to (3.6) we have  $\varphi = \lambda$  by symmetry and from the Hamiltonian system we arrive at finite element discretizations of the nonlinear elliptic partial differential equation

$$-\operatorname{div}(\mathfrak{h}'_{\delta}(\eta - |\nabla \varphi|^2) \nabla \varphi) = 1 \text{ in } \Omega, \quad \varphi|_{\partial\Omega} = 0, \quad (3.11)$$

which is the Euler-Lagrange equation of the problem to find an extremal point  $\varphi \in H_0^1(\Omega)$  of

$$\bar{H}_{\delta}(\varphi) = \int_{\Omega} 2\varphi \, dx + \int_{\Omega} \mathfrak{h}_{\delta}(\eta - |\nabla \varphi|^2) \, dx. \quad (3.12)$$

In contrast to (3.6), the existence of an extremal point can not be guaranteed for (3.12) since it lacks weak lower or upper semicontinuity as  $\delta$  becomes small.

Note that existence of minimizers to general functionals

$$F(\varphi) = \int_{\Omega} f(x, \varphi(x), \nabla \varphi(x)) \, dx, \quad (3.13)$$

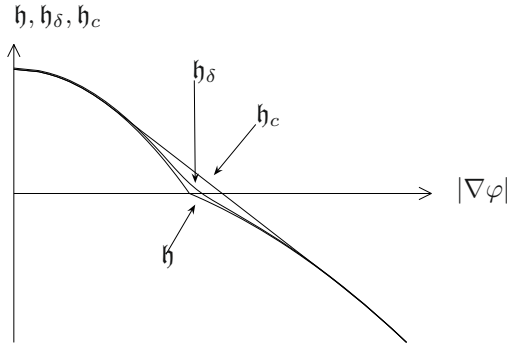


FIGURE 3. The function  $\mathfrak{h}$ , its regularization  $\mathfrak{h}_\delta$  and its concavification  $\mathfrak{h}_c$  with respect to  $|\nabla\varphi|$ , for the non-concave case.

where  $\Omega \subset \mathbb{R}^d$  is a bounded open set,  $\varphi : \Omega \rightarrow \mathbb{R}^n$ , can be guaranteed if  $F$  fulfills appropriate growth conditions (coercivity) and is weakly lower semi-continuous on  $H_0^1(\Omega)$ . Weak lower semi-continuity is generally hard to verify, but for the scalar case  $n = 1$ , or  $d = 1$ ,  $F$  is weakly lower semicontinuous if and only if  $f(x, \varphi, \cdot)$  is convex [18]. For the vector case convexity is a sufficient but far from necessary condition and can be replaced by quasi-convexity which is both a necessary and sufficient condition, but almost as hard to verify as weak lower semi-continuity.

To achieve, in this case, a weakly upper semicontinuous functional one can replace the function  $\mathfrak{h}_\delta$ , for  $\delta = 0$ , in (3.12) with its concavification

$$\mathfrak{h}_c = \begin{cases} \sigma_+(\eta - |\nabla\varphi|^2), & |\nabla\varphi|^2 < \eta \frac{\sigma_-}{\sigma_+}, \\ \eta(\sigma_+ + \sigma_-) - 2\sqrt{\eta\sigma_+\sigma_-}|\nabla\varphi|, & \eta \frac{\sigma_-}{\sigma_+} \leq |\nabla\varphi|^2 \leq \eta \frac{\sigma_+}{\sigma_-}, \\ \sigma_-(\eta - |\nabla\varphi|^2), & |\nabla\varphi|^2 > \eta \frac{\sigma_+}{\sigma_-}, \end{cases} \quad (3.14)$$

as in [22,24], see Figure 3. This gives a concave functional in (3.12) which not only has a maximizer but achieves the same supremum as the Hamiltonian  $H_0$ , and has maximizers which are exactly the weak limits of maximizing sequences for  $H_0$ . If  $d > 1$  and  $n > 1$ , maximizers with equivalent properties are given by quasi-concavification, see [18].

Numerical experiments using a finite element discretization of (3.11) shows that, although existence of solutions cannot be guaranteed for small  $\delta$ , the Pontryagin approach generates approximations close to the true concavified solutions in the sense that the error in the value functions is small, see Section 4.2. Of course, the sensitivity of the controls with respect to the value function may still be large.

An alternative to the above concavification (3.14) is to simply replace the original maximization problem in (3.10) by

$$\max_{\sigma} \left\{ \int_{\Omega} \varphi \, dx - \gamma \int_{\Omega} \frac{1}{\sigma} \, dx \mid -\operatorname{div}(\sigma \nabla \varphi) = 1 \text{ in } \Omega, \varphi|_{\partial\Omega} = 0 \right\}, \quad (3.15)$$

with a given multiplier  $\gamma \geq 0$  and  $\sigma : \Omega \rightarrow \{\sigma_-, \sigma_+\}$ . This formulation only differs in the choice of the given constant  $\gamma$ , associated to the amount of material.

From the new Hamiltonian

$$H(\varphi) = \int_{\Omega} 2\varphi \, dx + \int_{\Omega} \underbrace{\max_{\sigma} \left\{ -\frac{\gamma}{\sigma} - \sigma |\nabla\varphi|^2 \right\}}_{\mathfrak{h}} \, dx. \quad (3.16)$$



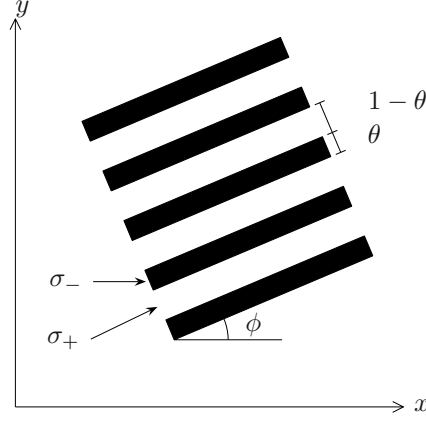


FIGURE 4. Laminate.

We then see that, allowing intermediate controls  $\sigma : \Omega \rightarrow [\sigma_-, \sigma_+]$ , the explicit maximization in  $\sigma$  gives the concave function

$$\mathfrak{h} = \begin{cases} -\frac{\gamma}{\sigma_+} - \sigma_+ |\nabla\varphi|^2, & |\nabla\varphi| < \frac{\sqrt{\gamma}}{\sigma_+}, \\ -2\sqrt{\gamma}|\nabla\varphi|, & \frac{\sqrt{\gamma}}{\sigma_+} < |\nabla\varphi| < \frac{\sqrt{\gamma}}{\sigma_-}, \\ -\frac{\gamma}{\sigma_-} - \sigma_- |\nabla\varphi|^2, & \frac{\sqrt{\gamma}}{\sigma_-} < |\nabla\varphi|, \end{cases}$$

and we thus have the concave maximization problem: to find the unique maximizer  $\varphi \in H_0^1(\Omega)$  to (3.16). In fact, the formulation (3.15) is related to relaxation by the homogenization method [1]. Instead of just expanding the set of admissible controls for the original problem (3.10) to  $\sigma : \Omega \rightarrow [\sigma_-, \sigma_+]$ , we look at the problem

$$\max_{\theta, \phi} \left\{ \int_{\Omega} \varphi \, dx - \eta \int_{\Omega} \theta \, dx \mid -\operatorname{div}(\sigma^*(\theta, \phi)\nabla\varphi) = 1 \text{ in } \Omega, \varphi|_{\partial\Omega} = 0 \right\}, \quad (3.17)$$

with  $\theta : \Omega \rightarrow [0, 1]$ ,  $\phi : \Omega \rightarrow [0, \pi]$  and the rank-1 laminate tensor

$$\sigma^* = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \lambda_{\theta}^+ & 0 \\ 0 & \lambda_{\theta}^- \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix},$$

with

$$\lambda_{\theta}^- = \left( \frac{\theta}{\sigma_-} + \frac{1-\theta}{\sigma_+} \right)^{-1}, \quad \lambda_{\theta}^+ = \theta\sigma_- + (1-\theta)\sigma_+.$$

The tensor  $\sigma^*$  is obtained from rotation and mixing of the two tensor valued controls  $\sigma_-I$  and  $\sigma_+I$  in proportions  $\theta$  and  $1-\theta$  and direction  $\phi$ , see Figure 4. We have thus enlarged the set of admissible controls by introducing two new parameters  $\theta, \phi$  describing a laminated material, where the effective conductivities in the principal directions of the material is  $\lambda_{\theta}^+$  and  $\lambda_{\theta}^-$ . It is easy to see that  $(\lambda_{\theta}^+)^{-1}$  and  $(\lambda_{\theta}^-)^{-1}$  corresponds to the total resistances for resistors connected in parallel and in series, respectively.

Using symmetry,  $\varphi = \lambda$ , the Hamiltonian derived from (3.17) is

$$H(\varphi) = \int_{\Omega} 2\varphi \, dx + \int_{\Omega} \max_{\theta, \phi} \left\{ -\eta\theta - (\sigma^*(\theta, \phi)\nabla\varphi) \cdot \nabla\varphi \right\} \, dx. \quad (3.18)$$

Rewriting the maximization in (3.18) as

$$\max_{\theta, \phi} \left\{ -\eta\theta - (\text{diag}(\lambda_\theta^+, \lambda_\theta^-)q_\phi) \cdot q_\phi \right\} \quad (3.19)$$

where  $q_\phi$  is the rotation of  $\nabla\varphi$ , it is evident that since  $\lambda_\theta^- \leq \lambda_\theta^+$ , aligning  $q_\phi$  in the  $\lambda_\theta^-$ -direction or equivalently aligning the material perpendicular to  $\nabla\varphi$ , maximizes (3.19) with respect to  $\phi$ . The maximization over  $\phi$  thus gives

$$H(\varphi) = \int_{\Omega} 2\varphi \, dx + \int_{\Omega} \max_{\sigma} \left\{ -\frac{\eta\sigma_+\sigma_-}{\sigma_+ - \sigma_-} \left( \frac{1}{\sigma} - \frac{1}{\sigma_+} \right) - \sigma|\nabla\varphi|^2 \right\} dx, \quad (3.20)$$

with the change of variables  $\sigma = \lambda_\theta^-$ . Defining  $\gamma = \eta\sigma_+\sigma_-(\sigma_+ - \sigma_-)^{-1}$ , (3.16) and (3.20) have the same minimizer  $\varphi$ .

### 3.3. Interior reconstruction

In the previous sections we discussed problems with symmetry, *i.e.*  $\varphi = \pm\lambda$ , for which convexification is a straightforward and simple approach. Although symmetry is present in many optimization problems connected to minimization of energy, there are other important problems, such as inverse problems related to reconstruction from measurements, where both  $\varphi$  and  $\lambda$  need to be determined. Even the simplest reconstruction problems are known to be highly ill-posed [19].

We will here focus on the problem to reconstruct the interior of an object from measured electrical surface currents, *i.e.* electric impedance tomography [5]: let  $\sigma^* : \Omega \rightarrow \{\sigma_-, \sigma_+\}$  denote a real valued unknown conductivity distribution, with  $\sigma_{\pm} > 0$ , in a given domain  $\Omega \subset \mathbb{R}^d$ . Using given surface currents  $q_i$ ,  $i = 1, \dots, N$  on  $\partial\Omega$ , satisfying  $\int_{\partial\Omega} q_i \, ds = 0$ , and the resulting measured surface potentials  $\varphi_i^*$  on  $\partial\Omega$ , the goal in this inverse problem is to find the optimal conductivity distribution  $\sigma : \Omega \rightarrow \{\sigma_-, \sigma_+\}$  such that

$$\min_{\sigma} \left\{ \sum_{i=1}^N \int_{\partial\Omega} (\varphi_i - \varphi_i^*)^2 \, ds \mid \text{div}(\sigma\nabla\varphi_i) = 0 \text{ in } \Omega, \sigma \frac{\partial\varphi_i}{\partial n} \Big|_{\partial\Omega} = q_i \right\}, \quad (3.21)$$

with  $\varphi_i \in V \equiv \{v \in H^1(\Omega) : \int_{\Omega} v \, dx = 0\}$ . Note, that we have here chosen the simpler case with measurements on the whole boundary; in reality often only a discrete number of contacts are allowed.

The Lagrangian becomes

$$\sum_{i=1}^N \int_{\partial\Omega} (\varphi_i - \varphi_i^*)^2 + \lambda_i q_i \, ds + \int_{\Omega} \sigma \underbrace{\sum_{i=1}^N -\nabla\varphi_i \cdot \nabla\lambda_i}_{v} \, dx$$

and the Hamiltonian

$$H(\lambda_1, \dots, \lambda_N, \varphi_1, \dots, \varphi_N) = \sum_{i=1}^N \int_{\partial\Omega} (\varphi_i - \varphi_i^*)^2 + \lambda_i q_i \, ds + \int_{\Omega} \underbrace{\min_{\sigma} \{\sigma v\}}_{\mathfrak{h}(v)} \, dx.$$

As in previous sections the Hamiltonian needs to be regularized such that

$$\bar{H}_{\delta}(\lambda_1, \dots, \lambda_N, \varphi_1, \dots, \varphi_N) = \sum_{i=1}^N \int_{\partial\Omega} (\varphi_i - \varphi_i^*)^2 + \lambda_i q_i \, ds + \int_{\Omega} \mathfrak{h}_{\delta}(v) \, dx, \quad (3.22)$$

which generates the coupled non-linear elliptic partial differential equations

$$\begin{aligned} \operatorname{div}\left(\mathfrak{h}'_{\delta}\left(\sum_{k=1}^N -\nabla\varphi_k \cdot \nabla\lambda_k\right)\nabla\varphi_i\right) &= 0, \text{ in } \Omega, & \mathfrak{h}'_{\delta}\frac{\partial\varphi_i}{\partial n}\Big|_{\partial\Omega} &= q_i \\ \operatorname{div}\left(\mathfrak{h}'_{\delta}\left(\sum_{k=1}^N -\nabla\varphi_k \cdot \nabla\lambda_k\right)\nabla\lambda_i\right) &= 0, \text{ in } \Omega, & \mathfrak{h}'_{\delta}\frac{\partial\lambda_i}{\partial n}\Big|_{\partial\Omega} &= 2(\varphi_i - \varphi_i^*) \end{aligned} \quad (3.23)$$

for  $i = 1, \dots, N$ . Even though the lack of symmetry prohibits any simplification, this system is only locally coupled, and finite element discretizations can be solved by the Newton method with a sparse Hessian, see Section 4.3.

It is clear that the minimization problem (3.21) attains its minimum for  $\sigma = \sigma^*$  and  $\varphi_i = \varphi_i^*$ , but it has not necessarily a unique solution. To determine uniqueness of solutions would require knowledge of the Neumann-Dirichlet map

$$\Lambda_{\sigma} : \sigma \frac{\partial\varphi_i}{\partial n}\Big|_{\partial\Omega} \rightarrow \varphi_i\Big|_{\partial\Omega}$$

associating boundary voltages with currents at the boundary for a fixed  $\sigma$ . Perfect knowledge of the map  $\Lambda_{\sigma}$  can in general only be gained by measuring the resulting potentials for all possible input currents. However, if  $\sigma \in \{\sigma_-, \sigma_+\}$  inside or outside a possibly multiple-connected domain  $D \subset \Omega$ , it is possible to uniquely determine  $\sigma$  with only partial knowledge of  $\Lambda_{\sigma}$ , *i.e.* by using only a finite number of experiments, see [5] for references.

If the reconstruction problem (3.21) has a unique solution it is still ill-posed in the sense that the slightest disturbance in measurements  $\varphi_i$  or having a true conductivity which allows intermediate values  $\sigma^* \in [\sigma_-, \sigma_+]$  would destroy all results on existence and uniqueness of solutions. This is also the case for a discrete number of contacts or for measurements on only parts of the boundary.

Alternative formulations of the impedance tomography problem related to relaxation of functionals as in Section 3.2 is found in [25,37]. In the latter the reconstruction problem is formulated as

$$\min_{\sigma, \varphi_i, J_i} \left\{ I := \sum_{i=1}^N \int_{\Omega} \left| \sigma^{\frac{1}{2}} \nabla\varphi_i + \sigma^{-\frac{1}{2}} J_i \right|^2 dx \mid \operatorname{div}(J_i) = 0 \text{ in } \Omega, J_i \cdot n|_{\partial\Omega} = q_i, \varphi_i|_{\partial\Omega} = \varphi_i^* \right\}.$$

Here the need for regularization to guarantee existence of solutions comes from the lack of lower semicontinuity of the functional  $I$ . This formulation also allows an explicit expression for  $\sigma$  since expanding the squares gives

$$\min_{\sigma > 0, \varphi, J} \left\{ \sigma \sum_{i=1}^N \int_{\Omega} |\nabla\varphi_i|^2 dx + \sigma^{-1} \sum_{i=1}^N \int_{\Omega} |J_i|^2 dx \mid \operatorname{div}(J_i) = 0 \text{ in } \Omega, J_i \cdot n|_{\partial\Omega} = q_i, \varphi_i|_{\partial\Omega} = \varphi_i^* \right\}$$

from the constraint  $\operatorname{div}(J) = 0$ , which allows pointwise minimization in  $\sigma$ , as in our case, such that

$$\sigma = \left( \sum_{i=1}^N |J_i| \right)^{\frac{1}{2}} \left( \sum_{i=1}^N |\nabla\varphi_i| \right)^{-\frac{1}{2}}.$$

## 4. NUMERICAL EXAMPLES

## 4.1. Electric conduction

In all numerical tests for the electric conduction problem (3.2) we let  $\Omega$  be the unit square, sometimes with holes cut out, and apply currents on contacts at the boundary. We also let  $\bar{V} \subset V \equiv \{v \in H^1(\Omega) : \int_{\Omega} v \, dx = 0\}$  be the linear finite element subspace with  $\Omega$  partitioned into a quasi uniform mesh with triangles of maximum diameter  $h_{\max} = 0.01$ . The bounds on the conductivity are  $\sigma_- = 10^{-3}$  and  $\sigma_+ = 1$ , and the regularized function  $\mathfrak{h}_{\delta}$  is chosen to be a hyperbola with asymptotes coinciding with  $\mathfrak{h}$  and with a closest distance  $\delta$  from the origin, see Figure 1.

For solving the non-linear partial differential equation (3.5), or equivalently maximizing (3.6), we use the Newton method  $\bar{\varphi}_{\text{new}} = \bar{\varphi}_{\text{old}} + \phi$ , where the update  $\phi = \sum_{i=1}^N \phi_i v_i$  comes from solving the system

$$\sum_{j=1}^N D_{v_i, v_j}^2 \bar{H}_{\delta}(\bar{\varphi}_{\text{old}}) \phi_j = -D_{v_i} \bar{H}_{\delta}(\bar{\varphi}_{\text{old}}), \quad i = 1, \dots, N, \quad (4.1)$$

with the sparse positive definite Hessian

$$\begin{aligned} D_{v_i, v_j}^2 \bar{H}_{\delta}(\bar{\varphi}) &= -2 \int_{\Omega} \mathfrak{h}'_{\delta}(\eta - |\nabla \bar{\varphi}|^2) \nabla v_i \cdot \nabla v_j \, dx \\ &\quad + 4 \int_{\Omega} \mathfrak{h}''_{\delta}(\eta - |\nabla \bar{\varphi}|^2) (\nabla \bar{\varphi} \cdot \nabla v_j) (\nabla \bar{\varphi} \cdot \nabla v_i) \, dx, \end{aligned}$$

and

$$D_{v_i} \bar{H}_{\delta}(\bar{\varphi}) = -2 \int_{\Omega} \mathfrak{h}'_{\delta}(\eta - |\nabla \bar{\varphi}|^2) \nabla \bar{\varphi} \cdot \nabla v_i \, dx + 2 \int_{\partial\Omega} q v_i \, ds.$$

Here  $v_i \in \bar{V}$  denotes the nodal Lagrange element basis functions of  $\bar{V}$ , with  $v_i(x_j) = \delta_{ij}$  and  $x_j$  denoting the corners of the triangles and  $\delta_{ij}$  the Kronecker delta.

To decrease the regularization  $\delta$  we use the following scheme, with the idea to decrease  $\delta$  to  $\alpha\delta$ ,  $0 < \alpha < 1$ , if the Newton method converged and increase  $\delta$  to the average of  $\delta$  and the last successful regularization if the iterations did not converge:

- (1) Choose an initial stepsize  $\alpha_{\text{old}} = 0.5$  and an initial regularization  $\delta_{\text{old}}$ .
- (2) If the Newton method for  $\delta_{\text{old}}$  converged choose

$$\delta_{\text{new}} = \alpha_{\text{old}} \delta_{\text{old}}, \quad \alpha_{\text{new}} = \alpha_{\text{old}}$$

otherwise let

$$\delta_{\text{new}} = \frac{1}{2} \delta_{\text{old}} \left(1 + \frac{1}{\alpha_{\text{old}}}\right), \quad \alpha_{\text{new}} = \frac{2\alpha_{\text{old}}}{\alpha_{\text{old}} + 1}.$$

- (3) Set  $\delta_{\text{old}} = \delta_{\text{new}}$ ,  $\alpha_{\text{old}} = \alpha_{\text{new}}$  and go to (2).

Results for different regularizations can be seen in Figure 5. Figure 6 shows solutions for different multipliers  $\eta$ , corresponding to different volume constraints. Solutions for other geometries and boundary conditions are presented in Figure 7. The Newton method works well but requires some additional iteration steps for smaller regularizations or finer grids since the Hessian becomes ill-conditioned. It is possible that other methods using or approximating Hessian information, such as quasi-Newton methods, may be used. However, from our experience we conclude that good approximation of the Hessian is vital for convergence. Some experiments using the non-linear multigrid method with a modification preventing large course-level corrections close to the jump in  $\mathfrak{h}_{\delta}$  has also showed good results.

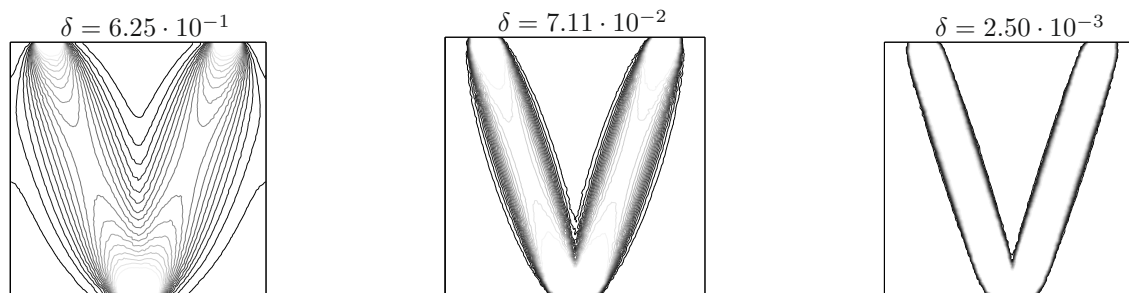


FIGURE 5. Electric conduction: contour plots of  $h'_\delta$  for different regularizations and with  $\eta = 0.5$ . The contours are equally spaced in the range  $[\sigma_-, \sigma_+]$ , with  $\sigma_+$  at the contacts. Current enters the top contacts ( $q = -1$ ) and leaves on the bottom contact ( $q = 2$ ). All contacts are of width 0.1. The regularization was initialized to  $\delta = 10$  and the pictures show the result after 5, 10 and 30 reductions of  $\delta$ . For each value of  $\delta$  no more than 5 Newton steps was necessary to reach a residual error of  $10^{-6}$  in the maximum norm. For two values of  $\delta$  the Newton method fails to converge, leading to a change in the update parameter  $\alpha$ .

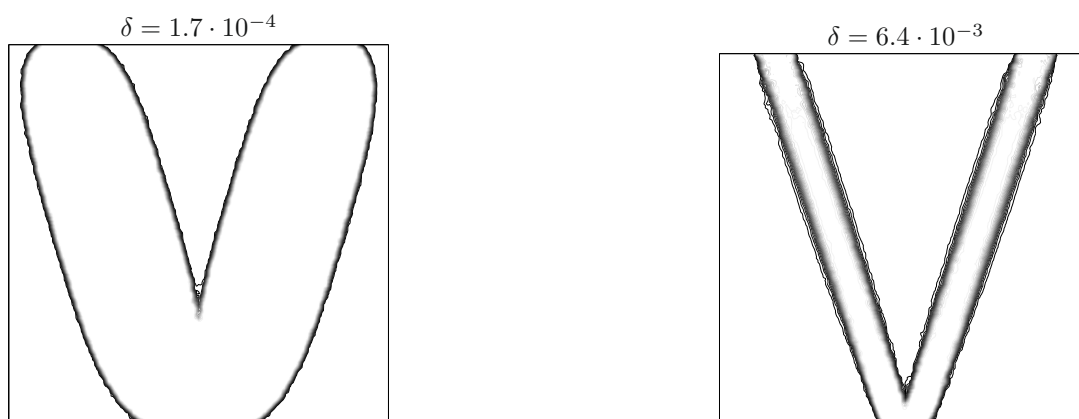


FIGURE 6. Electric conduction: contour plots of  $h'_\delta$  for different values of  $\eta$ . Left: data as in Figure 5 but with  $\eta = 0.1$ . Right: data as in Figure 5 but with  $\eta = 1.5$ . The left example has reached  $\delta \approx 10^{-4}$  in 23 reductions of  $\delta$  starting with  $\delta = 10$ , while the right example was stopped after 30 reductions of  $\delta$ .

To verify the error estimates in Section 2.1 we see that for given solutions  $\varphi \in V$  to (3.7) and  $\bar{\varphi} \in \bar{V}$  to (3.6) the error in the value functions is by (2.24)–(2.27)

$$E = \bar{H}_\delta(\bar{\varphi}, \bar{\varphi}) - H(\varphi, \varphi). \quad (4.2)$$

Section 2.1 and Theorem 2.4 estimate such errors in terms of the  $L^2$  projection error  $|P\varphi - \varphi|$  and the difference between the Hamiltonians  $H$  and  $\bar{H}$  along the same path. Now we apply this using the Hamiltonian  $\bar{H}_\delta$  defined in (3.4). The Legendre transform can be used as in Section 2.3 to obtain from  $\bar{H}_\delta$  a flux  $\bar{f}$  and a running cost  $\bar{h}$  defined on  $\bar{V}$ . By (2.5) this defines a Hamiltonian  $\bar{H} : V \times \bar{V} \rightarrow \mathbb{R}$ , which is used in the analysis in Section 2.1.

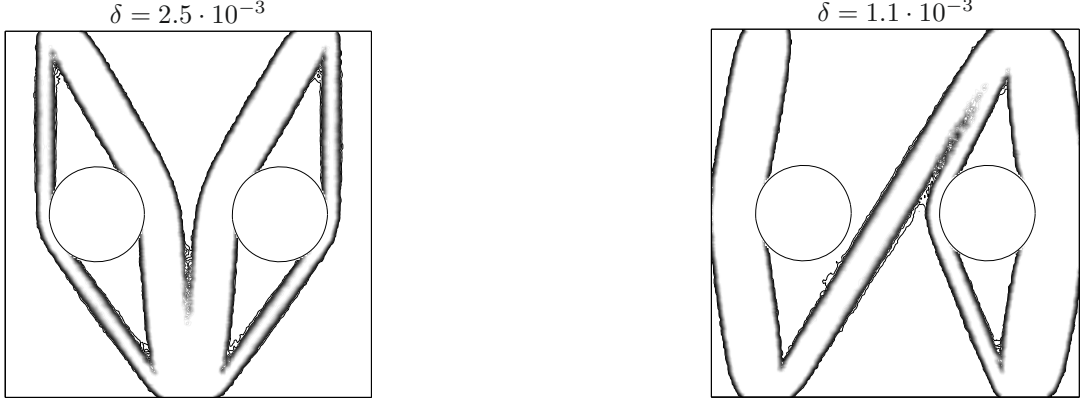


FIGURE 7. Electric conduction: contour plots of  $\mathfrak{h}'_\delta$  for different geometries and boundary currents  $q$ . Left: data as in Figure 5 but with two holes cut out. Right: data as in Figure 5 but with four contacts of width 0.1 and currents  $q = -1$  and  $q = -2$  on the top contacts and  $q = 1.5$  on the bottom contacts.

Definition (2.5) and the fact that  $\bar{H}$  and  $\bar{H}_\delta$  coincide on  $\bar{V} \times \bar{V}$  show that

$$\bar{H}(\lambda, \bar{\varphi}) = \bar{H}(P\lambda, \bar{\varphi}) = \bar{H}_\delta(P\lambda, \bar{\varphi}). \quad (4.3)$$

We expect  $\lambda = \partial u$  and  $\bar{\lambda} = \partial \bar{u}$  in (2.15) and (2.17). Replacing  $\partial u$  by  $\lambda$  and  $\partial \bar{u}$  by  $\bar{\lambda}$  would give an upper and a lower bound

$$\hat{E}_+ := \bar{H}_\delta(\bar{\lambda}, P\varphi) - H(\bar{\lambda}, \varphi), \quad \hat{E}_- := (\bar{H}_\delta - H)(\lambda, \bar{\varphi}),$$

of the error. The symmetry  $\lambda = \varphi$ ,  $\bar{\lambda} = \bar{\varphi}$  and (4.3) imply  $\hat{E}_+ = \hat{E}_- = \hat{E}$ , where

$$\begin{aligned} \hat{E} &:= \bar{H}_\delta(\bar{\varphi}, P\varphi) - H(\bar{\varphi}, \varphi) \\ &= \int_{\Omega} \mathfrak{h}_\delta(\underbrace{\eta - \nabla \bar{\varphi} \cdot \nabla P\varphi}_{\bar{v}}) - \mathfrak{h}(\underbrace{\eta - \nabla \bar{\varphi} \cdot \nabla \varphi}_{v}) \, dx + \int_{\partial\Omega} q(P\varphi - \varphi) \, ds, \end{aligned} \quad (4.4)$$

and from (4.4) we get the error bound

$$\begin{aligned} |\hat{E}| &= \left| \int_{\Omega} \mathfrak{h}_\delta(\bar{v}) - \mathfrak{h}(v) \, dx + \int_{\partial\Omega} q(P\varphi - \varphi) \, ds \right| \\ &= \left| \int_{\Omega} \mathfrak{h}_\delta(\bar{v}) - \mathfrak{h}_\delta(v) + \mathfrak{h}_\delta(v) - \mathfrak{h}(v) \, dx + \int_{\partial\Omega} q(P\varphi - \varphi) \, ds \right| \\ &\leq C_0\delta + \left| \int_{\Omega} \mathfrak{h}_\delta(\bar{v}) - \mathfrak{h}_\delta(v) \, dx + \int_{\partial\Omega} q(P\varphi - \varphi) \, ds \right| \\ &= C_0\delta + \left| \int_{\Omega} \int_0^1 \mathfrak{h}'_\delta(t\bar{v} + (1-t)v)(\bar{v} - v) \, dt \, dx + \int_{\partial\Omega} q(P\varphi - \varphi) \, ds \right| \\ &\leq C_0\delta + \underbrace{\int_{\Omega} \sigma_+ |\nabla \bar{\varphi} \cdot \nabla(\varphi - P\varphi)| \, dx + \left| \int_{\partial\Omega} q(P\varphi - \varphi) \, ds \right|}_{E_3}, \end{aligned} \quad (4.5)$$

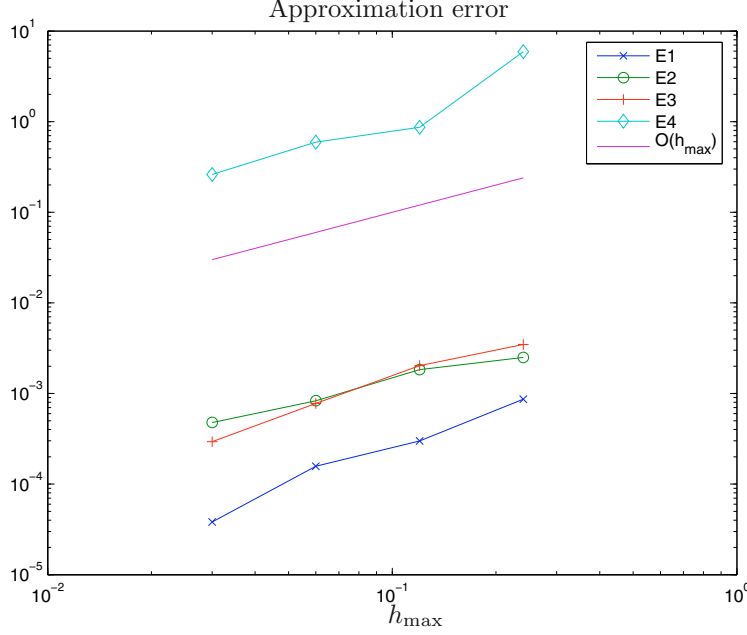


FIGURE 8. Electric conduction: error estimates for different meshes. Both  $\bar{\varphi}$  and  $\varphi$  are solutions to the regularized problem with  $\delta \approx 10^{-5}$ . The solution  $\varphi$  is computed on a mesh with  $h_{\max} = 0.015$  and the mesh for  $\bar{\varphi}$  is varying.

which can be estimated by

$$|\hat{E}| \leq C_0 \delta + \underbrace{C_1 h_{\max} \|\bar{\varphi}\|_{W^{1,\infty}(\Omega)} \|\varphi\|_{W^{2,1}(\Omega)} + C_2 h_{\max} \|q\|_{L^\infty(\partial\Omega)} \|\varphi\|_{W^{2,1}(\Omega)}}_{E_4}. \quad (4.6)$$

This estimate follows from stability of the  $L^2$  projection onto  $V$ , *cf.* [17], combined with a standard interpolation error estimate [6]. The regularized function  $\mathfrak{h}_\delta$  in (4.5) is chosen such that  $C_0 \approx 0.05$  independently of  $\delta$ .

To numerically verify the different estimates (4.2), (4.4), (4.5) and (4.6), we let  $\bar{\varphi} \in \bar{V}$  and  $\varphi \in V$  be finite element solutions to (3.6) and (3.7) with Hamiltonians  $\bar{H}_\delta$  and  $H = \bar{H}_{\delta_0}$  where  $\delta_0 \approx 0$ . Estimate (4.2) then becomes

$$E_1 := \bar{H}_\delta(\bar{\varphi}, \bar{\varphi}) - H_{\delta_0}(\varphi, \varphi) = \int_{\Omega} \mathfrak{h}_\delta(\eta - |\nabla \bar{\varphi}|^2) - \mathfrak{h}_{\delta_0}(\eta - |\nabla \varphi|^2) \, dx + 2 \int_{\partial\Omega} q(\bar{\varphi} - \varphi) \, ds, \quad (4.7)$$

and (4.4) becomes

$$E_2 := \bar{H}_\delta(\bar{\varphi}, P\varphi) - H_{\delta_0}(\bar{\varphi}, \varphi) = \int_{\Omega} \mathfrak{h}_\delta(\eta - \nabla \bar{\varphi} \cdot \nabla P\varphi) - \mathfrak{h}_{\delta_0}(\eta - \nabla \bar{\varphi} \cdot \nabla \varphi) \, dx + \int_{\partial\Omega} q(P\varphi - \varphi) \, ds. \quad (4.8)$$

Figure 8 shows the approximation error by comparing the different estimates  $E_1, E_2, E_3$  and  $E_4$  when  $\varphi$  is computed on a finite element mesh considerably finer than the one for  $\bar{\varphi}$ , and  $\delta = \delta_0 \approx 0$ . We see that  $E_2$  and  $E_3$  are accurate approximations of the true error  $E_1 \approx E$ , while  $E_4$  overestimates the error although it has the correct rate. Note that the interpolation constants satisfy  $C_1 \approx C_2 \approx 1$ . In Figure 9 the regularization error is shown by using the same mesh to compute  $\varphi$  and  $\bar{\varphi}$  for different regularizations  $\delta$  and a fixed  $\delta_0 \approx 0$ . We see that  $E_2$  again is an accurate approximation of  $E_1$  while  $E_4$  overestimates the error although it has

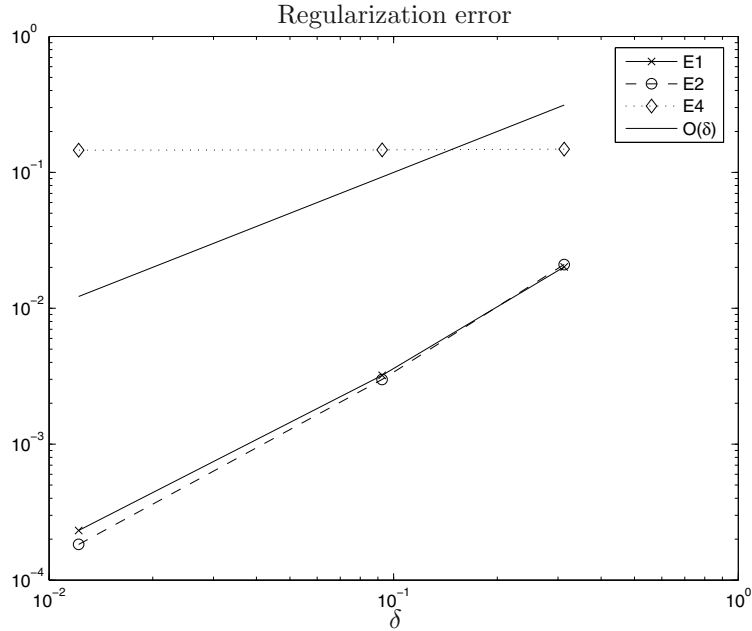


FIGURE 9. Electric conduction: error estimates for different regularizations. Both  $\bar{\varphi}$  and  $\varphi$  are solutions to regularized problems with  $h_{\max} = 0.015$ . The regularization for  $\varphi$  is  $\delta_0 \approx 10^{-3}$  and the regularization for  $\bar{\varphi}$  is varying.  $E_3$  disappears and  $E_4$  remains constant since no approximation error is present in this case.

the correct rate; it does not vary with  $\delta$ . Note that  $E_3$  is not presented in Figure 9 since it has the order of machine precision. We conclude that the approximation of the error estimate (4.4) in Theorem 2.4 is accurate.

## 4.2. Elastic domain

For the problem of maximizing the torsional rigidity of an elastic bar (3.10), we let the cross section  $\Omega$  of the bar be the unit square. The inverse shear moduli and the multiplier is, as in [1], chosen to be  $\sigma_- = 1$ ,  $\sigma_+ = 2$  and  $\eta = 0.0165$ , respectively. We also let  $\bar{V}$  be the linear finite element subspace of  $V \equiv H_0^1(\Omega)$  with maximum triangle diameter  $h_{\max} = 0.01$  and choose a regularization as in Section 4.1.

The maximization problem (3.10) has a Hamiltonian (3.12) which for small regularizations lacks upper semicontinuity, and thus there may not exist any solutions to the corresponding non linear partial differential equation (3.11). We solve (3.11) using the Newton method and a simple scheme to reduce the regularization, as in the previous section. As expected, the Newton method works well for sufficiently large regularizations, but does not converge for small regularizations, see Figure 10.

In Section 3.2 we concavify the unregularized Hamiltonian, which not only guarantees existence of at least one solution, but also gives a maximum that coincides with the supremum of the unregularized Hamiltonian. Figure 10 compares the result from solving the regularized problem (3.11) with the solution of the concavified problem. The regularized problem was solved by successively reducing the regularization 30 times, starting with  $\delta = 2$ .

In Figure 11 we see how the value of the regularized Hamiltonian approaches the value of the concavified Hamiltonian as the regularization decreases. We can also see that the Newton iterations fail when  $\delta$  becomes too small.

Since the goal is to find an optimal design with a discrete control  $\sigma \in \{\sigma_-, \sigma_+\}$  a few additional elementwise iterations with the Discrete Pontryagin method in (3.8)–(3.9) is done for postprocessing. These iterations are allowed as long as the value function does not increase substantially. In general, the discrete iterations does



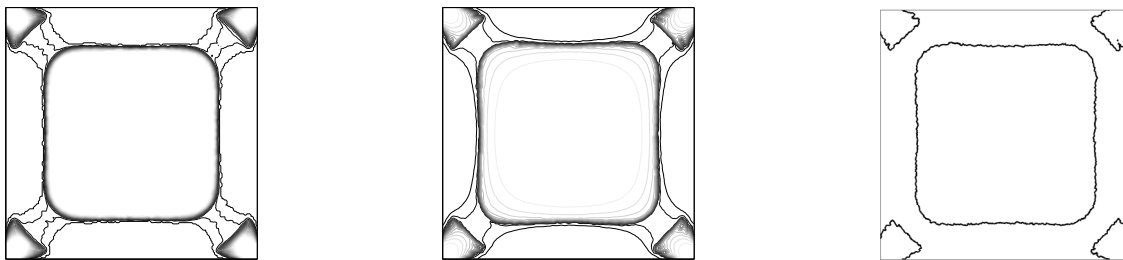


FIGURE 10. Elastic domain. Left: contour plot of  $h'_c$  for the concavified solution with  $\eta = 0.0165$ . The value of the Hamiltonian is 0.0555. Middle: contour plot of  $h'_\delta$  with  $\delta = 0.08$  and  $\eta = 0.0165$ . The value of the Hamiltonian is 0.0570. Right: five discrete iterations with (3.8)–(3.9) using initial data from the middle figure. The value of the Hamiltonian has converged to 0.0554.

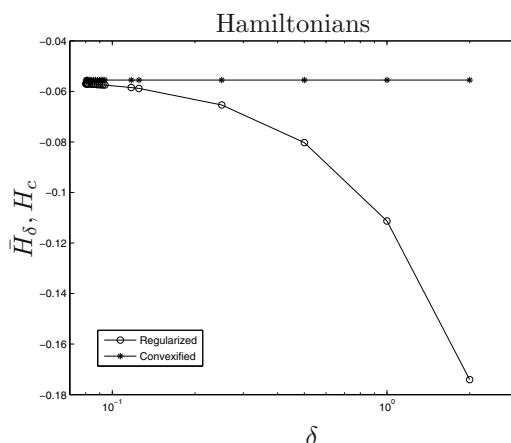


FIGURE 11. Elastic domain: plot of the regularized and concavified Hamiltonians for the solutions in Figure 10 with respect to regularization. Only regularizations for which the Newton method has converged are plotted. The regularized Hamiltonian approaches the concavified Hamiltonian as  $\delta \rightarrow 0$ , and the relative error for the smallest regularization,  $\delta = 0.08$ , is 2.7%.

not converge and we may need to control the amount of material allowed to change in each iteration; for the non-concave problem this appears however not necessary. The right plot in Figure 10 shows the solutions after a few discrete iterations with initial data from the middle figure.

The Discrete Pontryagin method much resembles the method in [12], which uses topological shape derivatives and starts from a domain with  $\sigma = \sigma_+$  and successively replaces volume fractions with  $\sigma_-$ . This method is appealing since it is simple and gives interesting designs, but it may not converge to the true optimal design if it is possible to remove too much material, which never can be added again.

Finally, in Figure 12, we show the results from solving the concavified and regularized problem with a different multiplier  $\eta = 0.025$ .

### 4.3. Impedance tomography

When solving the impedance tomography problem (3.21) one major issue affecting the reconstruction of the interior conductivity is the choice of input currents  $q_1, \dots, q_N$ . Consider applying a surface current  $q$  on two

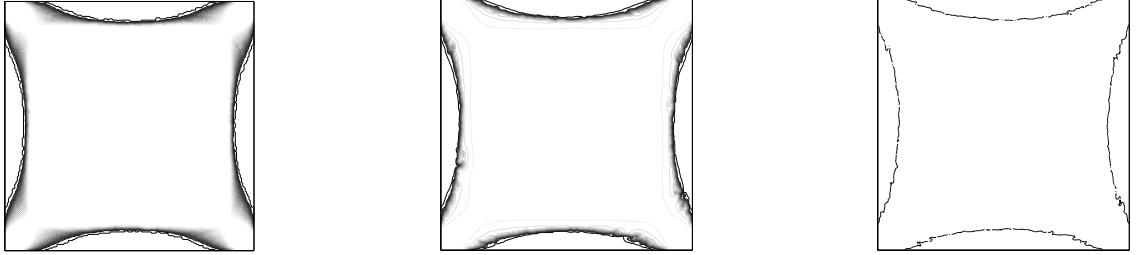


FIGURE 12. Elastic domain. Left: contour plot of  $h'_c$  for the concavified solution with  $\eta = 0.025$ . The value of the Hamiltonian is 0.0695. Middle: contour plot of  $h'_\delta$  with  $\delta = 0.1$  and  $\eta = 0.025$ . The value of the Hamiltonian is 0.0715. Right: five discrete iterations with (3.8)–(3.9) using initial data from the middle figure. The value of the Hamiltonian has converged to 0.0695, with a relative error of 0.9%.

different conductivity distributions  $\sigma$  and  $\sigma^*$  and measuring the resulting potentials  $\varphi = \Lambda_\sigma q$  and  $\varphi^* = \Lambda_{\sigma^*} q$ . Due to the nonlinearity of the inverse map  $\Lambda_\sigma \rightarrow \sigma$ , the different conductivities  $\sigma$  and  $\sigma^*$  may produce similar surface potentials  $\varphi$  and  $\varphi^*$  when subjected to a certain input current  $q$ , thus causing redundancy in the coupled equations (3.23). To prevent this we choose, following [13], the input current  $q$  to be optimal in the sense that it best distinguishes one conductivity distribution from another, *i.e.*

$$\max_q \left\{ \|\Lambda_\sigma q - \Lambda_{\sigma^*} q\|_{L^2(\partial\Omega)} \mid \|q\|_{L^2(\partial\Omega)} = 1, \int_{\partial\Omega} q \, ds = 0 \right\}, \quad (4.9)$$

which from self-adjointness of the Neumann-to-Dirichlet map  $\Lambda_\sigma$  is maximized by the eigenfunction corresponding to the largest eigenvalue of  $\Lambda_\sigma - \Lambda_{\sigma^*}$ . For multiple experiments we choose  $q_1, \dots, q_N$  to be the eigenfunctions corresponding to the  $N$  largest eigenvalues.

In the numerical experiments we have calculated the input currents using  $\sigma = 1.5$  and a true conductivity profile with  $\sigma^* = 1$  inside the region marked by a dashed line in Figure 14, and  $\sigma^* = 2$  outside. Note, that to calculate an optimal discretized boundary current  $q$  iteratively, *e.g.* by the power method, only measurements from the corresponding boundary potentials  $\Lambda_\sigma q$  and  $\Lambda_{\sigma^*} q$  are needed in each iteration. In Figure 13 the currents corresponding to the eigenfunctions of the four largest eigenvalues are shown. The boundary currents and potentials were calculated using a maximum element diameter of  $h_{\max} = 0.01$ .

The results from solving the coupled non-linear partial differential equations (3.23) using piecewise linear elements with  $h_{\max} = 0.03$  can be seen in Figures 14 and 15. Here, the regularization parameter  $\delta$  must not be too small since different grids are used for measurements and solutions in Figure 14, and 10% white noise is added to the measurements in Figure 15. This means that the discrete solution  $\varphi$  is disturbed componentwise by independent standard normal distributed stochastic variables  $\xi_i$  according to  $\varphi(x_i)(1 + 0.1\xi_i)$ , with nodal points  $x_i$ .

How to choose a good value of the regularization parameter,  $\delta$ , in a real problem may be difficult, at least a separate problem. Here we test how our regularization works by choosing the best possible regularization parameter value, using the “unknown” solution  $\sigma^*$  to find the parameter value, but nothing else. In the calculations we thus start with an initial regularization  $\delta = 1$ , reduce  $\delta$  at most 30 times and stop if  $\delta \leq 10^{-4}$  or if there is no decrease in the  $L^2(\Omega)$ -norm of  $h'_\delta - \sigma^*$ .

From Figure 14 it is evident that increasing the number of experiments has a significant effect on the reconstruction. The impedance tomography problem seems here to behave as the concave problem in Section 4.1. However in our experience there are cases, such as large differences between  $\sigma^-$  and  $\sigma^+$ , where (3.23) is harder to solve. In Figure 16 the approximation  $h'_\delta$  of the conductivity  $\sigma^*$  is shown.

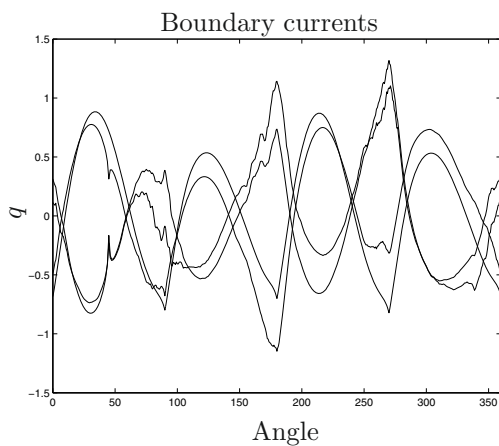


FIGURE 13. Boundary currents used for the impedance tomography problem plotted as a function of the angle variable in polar coordinates starting at the lower left corner of  $\Omega$ .

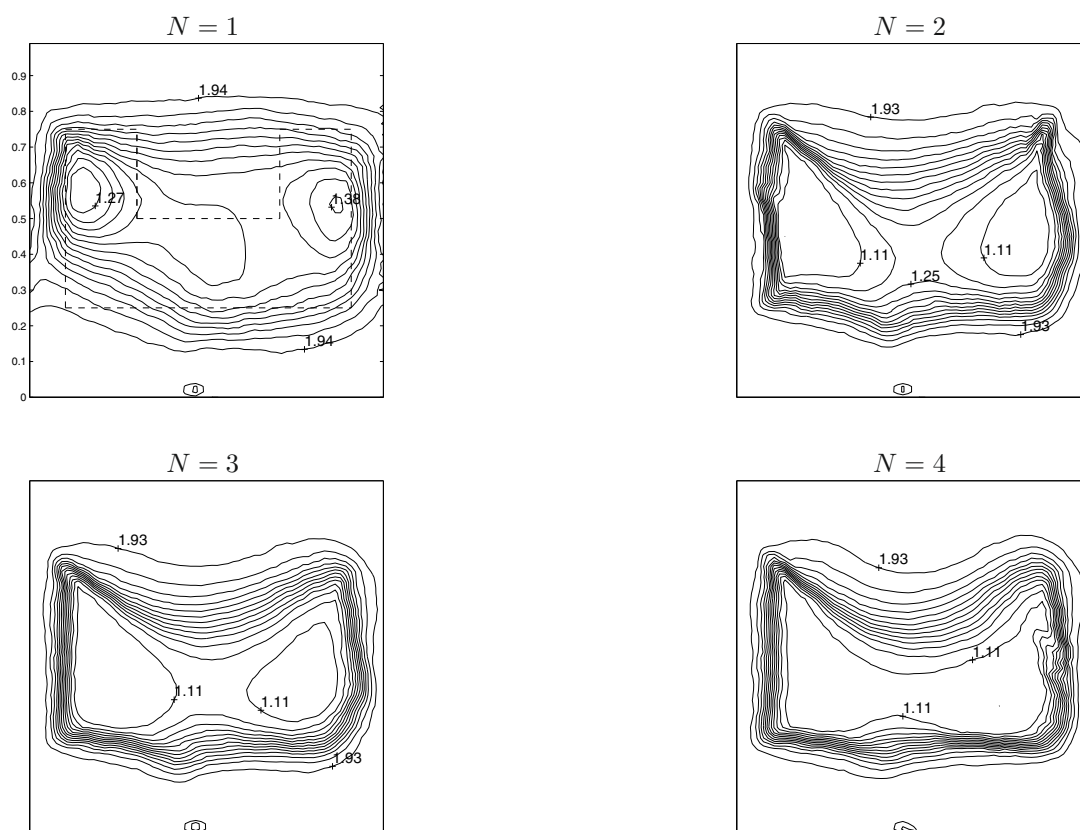


FIGURE 14. Impedance tomography: plots of  $h'_\delta$  for different number of experiments with  $h_{\max} = 0.03$ . The measured data was generated from solving the forward problem with  $h_{\max} = 0.01$  using the conductivity profile marked with a dashed line in the upper left figure. The true conductivity is  $\sigma^* = 1$  inside and  $\sigma^* = 2$  outside the marked region.

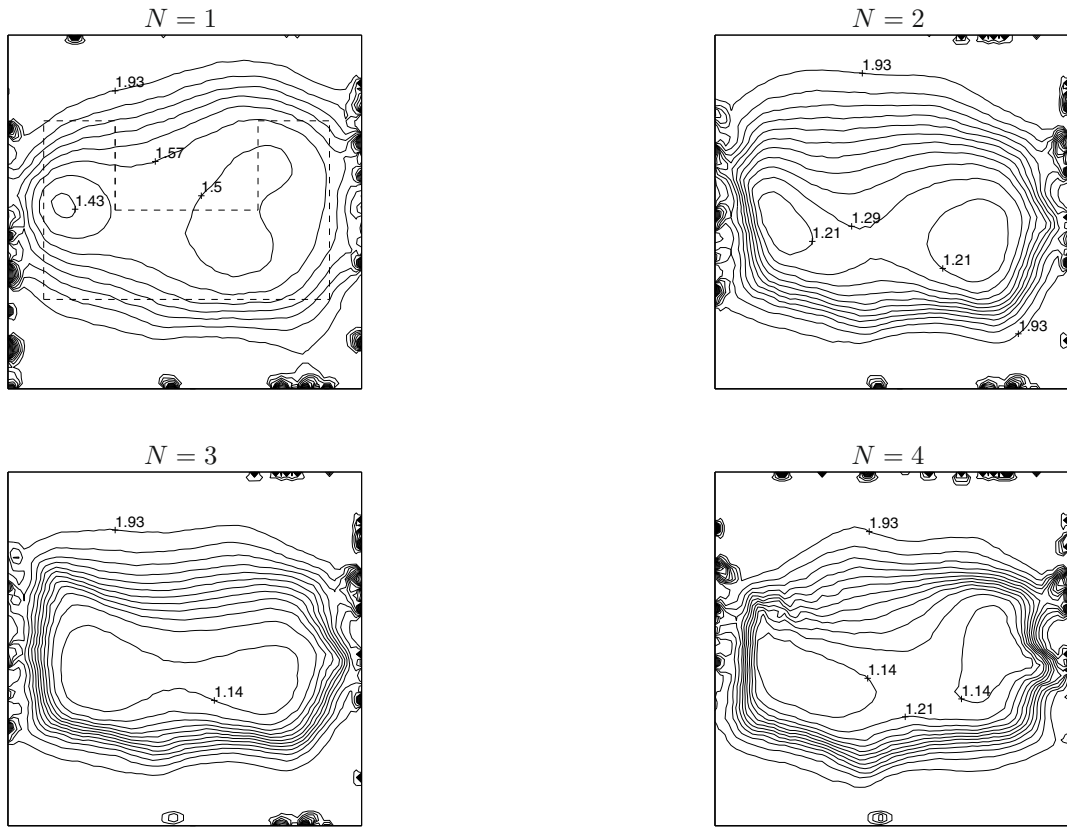


FIGURE 15. Impedance tomography: plots of  $h'_0$  for different number of experiments. Data as in Figure 14 but with 10% noise added to the measurements.

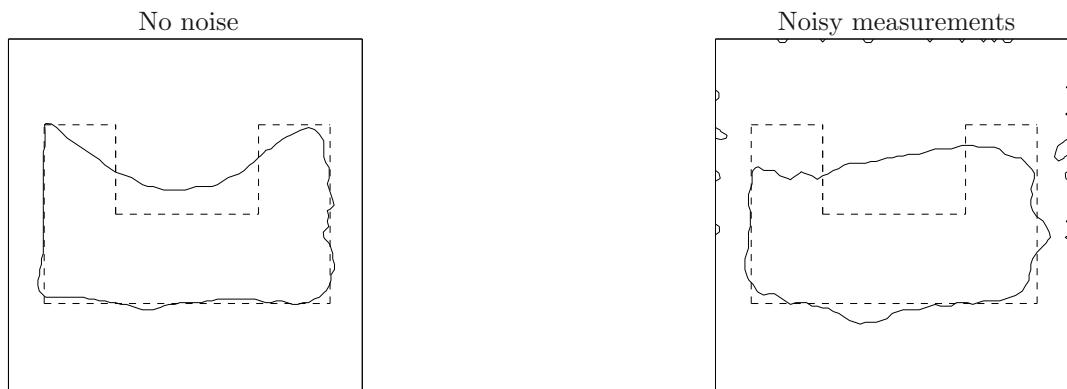


FIGURE 16. Impedance tomography: plots of  $h'_0$  after one discrete iteration ( $N = 4$ ). Left: initial data taken from the lower right plot in Figure 14. Right: initial data taken from the lower right plot in Figure 15.

## REFERENCES

- [1] G. Allaire, *Shape optimization by the homogenization method*, *Applied Mathematical Sciences* **146**. Springer-Verlag, New York (2002).
- [2] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, *Systems & Control: Foundations & Applications*. Birkhäuser Boston Inc., Boston, USA (1997). With appendices by M. Falcone and P. Soravia.
- [3] G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*, *Mathématiques & Applications* **17** (Berlin) [Mathematics & Applications]. Springer-Verlag, Paris (1994).
- [4] M.P. Bendsøe and O. Sigmund, *Topology optimization, Theory, methods and applications*. Springer-Verlag, Berlin (2003).
- [5] L. Borcea, Electrical impedance tomography. *Inverse Problems* **18** (2002) R99–R136.
- [6] S.C. Brenner and L. Ridgway Scott, *The mathematical theory of finite element methods*, *Texts in Applied Mathematics* **15**. Springer-Verlag, New York (1994).
- [7] P. Cannarsa and H. Frankowska, Some characterizations of optimal trajectories in control theory. *SIAM J. Control Optim.* **29** (1991) 1322–1347.
- [8] P. Cannarsa and H. Frankowska, Value function and optimality conditions for semilinear control problems. *Appl. Math. Optim.* **26** (1992) 139–169.
- [9] P. Cannarsa and H. Frankowska, Value function and optimality condition for semilinear co problems. II. Parabolic case. *Appl. Math. Optim.* **33** (1996) 1–33.
- [10] P. Cannarsa and C. Sinestrari, *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, *Progress in Nonlinear Differential Equations and their Applications* **58**. Birkhäuser Boston Inc., Boston, USA (2004).
- [11] J. Carlsson, *Symplectic reconstruction of data for heat and wave equations*. Preprint (2008) <http://arxiv.org/abs/0809.3621>.
- [12] J. Céa, S. Garreau, P. Guillaume and M. Masmoudi, The shape and topological optimizations connection. *Comput. Methods Appl. Mech. Engrg.* **188** (2000) 713–726. IV WCCM, Part II (Buenos Aires, 1998).
- [13] M. Cheney and D. Isaacson, Distinguishability in impedance imaging. *IEEE Trans. Biomed. Eng.* **39** (1992) 852–860.
- [14] F.H. Clarke, *Optimization and Nonsmooth Analysis*, *Canadian Mathematical Society Series in Mathematics*. John Wiley and Sons, Inc. (1983).
- [15] M.G. Crandall, L.C. Evans and P.-L. Lions, Some properties of viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.* **282** (1984) 487–502.
- [16] M.G. Crandall, H. Ishii and P.-L. Lions, User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)* **27** (1992) 1–67.
- [17] M. Crouzeix and V. Thomée, The stability in  $L_p$  and  $W_p^1$  of the  $L_2$ -projection onto finite element function spaces. *Math. Comp.* **48(178)** (1987) 521–532.
- [18] B. Dacorogna, *Direct methods in the calculus of variations*, *Appl. Math. Sci.* **78**. Springer-Verlag, Berlin (1989).
- [19] H.W. Engl, M. Hanke and A. Neubauer, *Regularization of inverse problems*, *Mathematics and its Applications* **375**. Kluwer Academic Publishers Group, Dordrecht (1996).
- [20] L.C. Evans, *Partial differential equations*, *Graduate Studies in Mathematics* **19**. American Mathematical Society, Providence, USA (1998).
- [21] H. Frankowska, Contingent cones to reachable sets of control systems. *SIAM J. Control Optim.* **27** (1989) 170–198.
- [22] J. Goodman, R.V. Kohn and L. Reyna, Numerical study of a relaxed variational problem from optimal design. *Comput. Methods Appl. Mech. Engrg.* **57** (1986) 107–127.
- [23] E. Hairer, C. Lubich and G. Wanner, *Geometric numerical integration, Structure-preserving algorithms for ordinary differential equations*, *Springer Series in Computational Mathematics* **31**. Springer-Verlag, Berlin (2002).
- [24] B. Kawohl, J. Stará and G. Wittum, Analysis and numerical studies of a problem of shape design. *Arch. Rational Mech. Anal.* **114** (1991) 349–363.
- [25] R.V. Kohn and A. McKenney, Numerical implementation of a variational method for electrical impedance tomography. *Inverse Problems* **6** (1990) 389–414.
- [26] R.V. Kohn and G. Strang, Optimal design and relaxation of variational problems. I. *Comm. Pure Appl. Math.* **39** (1986) 113–137.
- [27] R.V. Kohn and G. Strang, Optimal design and relaxation of variational problems. II. *Comm. Pure Appl. Math.* **39** (1986) 139–182.
- [28] R.V. Kohn and G. Strang, Optimal design and relaxation of variational problems. III. *Comm. Pure Appl. Math.* **39** (1986) 353–377.
- [29] F. Natterer, *The mathematics of computerized tomography*, *Classics in Applied Mathematics* **32**. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, USA (2001). Reprint of the 1986 original.
- [30] O. Pironneau, *Optimal shape design for elliptic systems*, *Springer Series in Computational Physics*. Springer-Verlag, New York (1984).
- [31] R.T. Rockafellar, *Convex analysis*, *Princeton Mathematical Series* **28**. Princeton University Press, Princeton, USA (1970).

- [32] M. Sandberg, *Convergence rates for numerical approximations of an optimally controlled Ginzburg-Landau equation*. Preprint (2008) <http://arxiv.org/abs/0809.1834>.
- [33] M. Sandberg and A. Szepessy, Convergence rates of symplectic Pontryagin approximations in optimal control theory. *ESAIM: M2AN* **40** (2006) 149–173.
- [34] D. Tataru, Viscosity solutions for Hamilton-Jacobi equations with unbounded nonlinear term: a simplified approach. *J. Diff. Eq.* **111** (1994) 123–146.
- [35] A.N. Tikhonov and V.Y. Arsenin, Solutions of ill-posed problems, in *Scripta Series in Mathematics*, V.H. Winston & Sons, Washington, D.C., John Wiley & Sons, New York (1977). Translated from the Russian, Preface by translation editor F. John.
- [36] C.R. Vogel, *Computational methods for inverse problems*, *Frontiers in Applied Mathematics* **23**. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2002). With a foreword by H.T. Banks.
- [37] A. Wexler, B. Fry and M.R. Neuman, Impedance-computed tomography algorithm and system. *Appl. Opt.* **24** (1985) 3985–3992.