On a continuation approach in Tikhonov regularization and its application in piecewise-constant parameter identification

V Melicher and V Vrábel

Research Group for Numerical Analysis and Mathematical Modelling, Department of Mathematical Analysis, Galglaan 2, B-9000 Gent, Belgium

E-mail: Valdemar.Melicher@UGent.be and Vladimir.Vrabel@UGent.be

Received 1 July 2013, in final form 16 August 2013
Published 1 October 2013
Online at stacks.iop.org/IP/29/115008

Abstract
We present a new approach to the convexification of the Tikhonov regularization using a continuation method strategy. We embed the original minimization problem into a one-parameter family of minimization problems. Both the penalty term and the minimizer of the Tikhonov functional become dependent on a continuation parameter. In this way we can independently treat two main roles of the regularization term, which are the stabilization of the ill-posed problem and introduction of the a priori knowledge. For zero continuation parameter we solve a relaxed regularization problem, which stabilizes the ill-posed problem in a weaker sense. The problem is recast to the original minimization by the continuation method and so the a priori knowledge is enforced. We apply this approach in the context of topology-to-shape geometry identification, where it allows us to avoid the convergence of gradient-based methods to a local minima. We present illustrative results for magnetic induction tomography which is an example of PDE-constrained inverse problem.

(Some figures may appear in colour only in the online journal)

1. Introduction

In this paper, we propose and study a continuation-based approach for the Tikhonov regularization of ill-posed problems that can be written in the form of an operator equation

\[ Fu = v, \]

where \( F : \mathcal{D}(F) \subseteq U \rightarrow V \) is a (in general nonlinear) forward operator, mapping between Banach spaces \( U \) and \( V \). By \( v \) we understand certain exact measurements projected on \( V \). We assume that only noisy data \( v^\delta \) are available, such that \( \| v - v^\delta \|_V \leq \delta \), where \( \delta \) is the level of noise.

Let us introduce a suitable regularization \( \mathcal{R} : U \rightarrow [0, +\infty] \) with the domain \( \mathcal{D}(\mathcal{R}) := \{ u \in U \mid \mathcal{R}(u) \neq +\infty \} \). It is a proper and convex functional. The general
convention is to consider only those solutions $u$ to ill-posed operator equation (1), where $\mathcal{R}(u)$ is sufficiently small. An element $u^\dagger$ is called an $\mathcal{R}$-minimizing solution (e.g. [1]) if
\[
\mathcal{R}(u^\dagger) = \min\{\mathcal{R}(u) \mid Fu = v\} < \infty.
\] (2)

We follow the classical Tikhonov idea [2, 3] and consider minimizers of functional
\[
\mathcal{T}_\alpha(u) := \|F(u) - v^\delta\|^2 + \alpha \mathcal{R}(u)
\] (3)
for a suitable regularization parameter $\alpha > 0$, which depends on both the noise level and data, i.e. $\alpha(\delta, v^\delta)$. The first term in (3) is called the fidelity functional (term). It ensures that the minima of the Tikhonov functional are approximate solutions of the operator equation (1), i.e. the problem which we want to solve in the first place. As usual, we denote a minimizer of (3) as
\[
u_\delta^\alpha \in \arg\min_{u \in U} \mathcal{T}_\alpha(u).
\] (4)

It is a well-known fact that under certain reasonable assumptions $\nu_\delta^\alpha$ are stable approximations of an $\mathcal{R}$-minimizing solution to (1), also in a rather general Banach space setting [1]. The resulting problem of regularization is to find a suitable $\alpha$ and the corresponding minimizer $\nu_\delta^\alpha$ of the Tikhonov functional (3), such that $\nu_\delta^\alpha$ approximates $u^\dagger$ as close as possible.

In this paper, we attempt to construct a sequence converging to the global minimizer $\nu_\delta^\alpha$. The biggest challenge is how to avoid the convergence of a numerical minimization method to a local minimum of (3), which is a common problem for standard gradient-based minimization methods (GBMM).

The possible reasons for the existence of local minima of (3) are triadic: the forward operator $F$ itself, the noise in the measurements and the penalty term $\mathcal{R}(u)$. The forward operator is case-specific and the noise is inherent to ill-posed problems. We have however full freedom of choice of regularization.

When one applies a GBMM to (3), the regularization $\mathcal{R}(u)$ enforces that the whole resulting minimizing sequence belongs to $U$. However, the underlying direct problem (1) generally requires a far less regularity of a solution than it is asked by $\mathcal{R}(u)$. Even if we expect our final solution to belong to $U$, it is not necessary to consider only minimizing sequences from $U$. This restriction is often the reason that a GBMM converges to a local minimum.

Let us recall that the purpose of adding the regularization is to stabilize the ill-posed problem and to ensure the desired properties of the solution. The main idea of the paper is to provide these two roles of the regularization term gradually.

1.1. Continuation immersion approach

Let us consider a Banach space $W$, such that $U$ is a proper subset of $W$ and the problem (1) is well defined in $W$, i.e. $U \subseteq W$ and $\mathcal{D}(F) \cap W \neq \emptyset$. We can introduce a new Tikhonov functional analogical to (3)
\[
\tilde{\mathcal{T}}_\alpha(w) := \|F(w) - v^\delta\|^2 + \alpha \mathcal{R}_W(w),
\] (5)
with a regularization term $\mathcal{R}_W : W \to [0, +\infty]$. It is again a convex and proper functional with the domain $\mathcal{D}(\mathcal{R}_W) := \{w \in W \mid \mathcal{R}_W(w) \neq +\infty\}$.

The strategy is to continuously transform the relaxed functional $\tilde{\mathcal{T}}_\alpha(w)$ to the original $\mathcal{T}_\alpha(u)$ together with the corresponding minimization problems by making use of the continuation method [4]. We will stabilize the problem (1) using $W$-based regularization, i.e. in a ‘broader’ sense. The ‘less’ constrained problem will be easier to solve and it will provide a very good starting point for minimization in $U$. The extra desired properties will be progressively imposed on the solution via continuation-based projection a posteriori.
We consider a one-parameter family of the Tikhonov functionals
\[ T^H_u(u, w, \lambda) = \| F(z) - v^H \|^2 + \alpha \{ \lambda R_U(u) + (1 - \lambda)R_W(w) \}, \]
where \( \lambda \in [0, 1] \) and
\[ z = \lambda u + (1 - \lambda)w. \]
The regularization term \( R_U \) stands for the original regularization in (3). For simplicity we assume that the regularization parameter \( \alpha \) does not depend on \( \lambda \). The forward problem \( F \) incident on (6) can be understood as acting on the parametrized family \( z \in W \). The regularization part
\[ R_{U,W}(u, w, \lambda) := \lambda R_U(u) + (1 - \lambda)R_W(w) \]
is better to be understood as a function on \( U \times W \).

We consequently deal with a one-parameter family of minimization problems1. For \( \lambda \in (0, 1) \) we look for a couple from \( U \times W \), which minimizes the functional (6), that is,
\[ (u^\lambda, w^\lambda) \in \arg\min_{(u, w) \in U \times W} T^H_u(u, w, \lambda). \]

For \( \lambda = 1 \), we get the original minimization problem of \( T_u \) and for \( \lambda = 0 \) the problem reduces to the minimization of (5). By the abuse of notation we sometimes write that \( (u, w) \) is a minimizer of \( T^H_u \) for \( \lambda \in [0, 1] \) to denote a minimizing couple \( (u, w) \in U \times W \) if \( \lambda \in (0, 1) \) and also to denote a minimizing element \( w \in W \) if \( \lambda = 0 \) or \( u \in U \) if \( \lambda = 1 \).

Analogically to the notion of the \( R \)-minimizing solution (2), let us define for each given \( \lambda \in [0, 1] \) an \( R_{U,W} \)-minimizing solution as a couple \( (u^\lambda, w^\lambda) \in U \times W \), such that
\[ R_{U,W}(u^\lambda, w^\lambda, \lambda) = \min \{ R_{U,W}(u, w, \lambda) : F(z) = v \} < \infty. \]
The paper is organized as follows. In section 2, we analyze the continuation approach in general. In section 3, we deal with piecewise-constant parameter identification problems (PCPIPs), which have been our motivation to study continuation methods in the context of Tikhonov regularization. We review the relevant state of the art in PCPIPs. Then, we introduce topology-to-shape continuation method (TSCM). In section 4, we apply the TSCM to magnetic induction tomography (MIT), which has many applications, e.g. in biomedical imaging and non-destructive testing of materials. Section 4.4 contains the implementation of the TSCM and several numerical experiments for MIT are presented in section 4.5.

2. Continuation approach for Tikhonov regularization

This section deals with theoretical aspects of the continuation approach for Tikhonov regularization. The functional \( T^H_u \) defined by (6) is always minimized with respect to the variables \( (u, w) \) and the variable \( \lambda \in [0, 1] \) is taken as a fixed parameter
\[ T^H_u(u, w, \lambda) \to \min, \quad \lambda u + (1 - \lambda)w = z \in D(F). \]
Throughout the section we make the following assumptions:

(A1) Let \( V \) be a Hilbert space with the scalar product \( (\cdot, \cdot) \) and the norm \( \| \cdot \| \). Let \( W \) be a reflexive Banach space and let \( U \) be its proper subspace, i.e. \( U \subset \subset W \), and \( \| \cdot \|_{U} \geq \tilde{C} \| \cdot \|_{W} \) where \( \tilde{C} \) is a positive constant. The space \( U \) is equipped with the topology of \( W \), i.e. a topology weaker than its norm topology.

1 The above formulas might bring the augmented Lagrangian method to mind. Among the differences between these two methods, we stress that we minimize here in the two independent variables \( u \) and \( w \). This turns out very convenient, mainly from the numerical point of view, as we will show later.
\textbf{(A2)} \( F : \mathcal{D}(F) \subseteq W \rightarrow V \), where \( \mathcal{D}(F) \) is closed and convex, and \( \mathcal{D} := \mathcal{D}(F) \cap U \neq \emptyset \). The map \( F \) is \textit{strongly continuous}, i.e.
\[ w_n \rightarrow w \text{ implies } F(w_n) \rightarrow F(w). \tag{10} \]

It is furthermore a \( C^1 \)-map.

\textbf{(A3)} \( \mathcal{R}_W : W \rightarrow [0, \infty) \) is a convex \( C^1 \)-map. There exists a constant \( p \geq 1 \) such that \( \mathcal{R}_W(w) \geq \|w\|_p^p \) for all \( w \in W \).

\textbf{(A4)} \( \mathcal{R}_U : U \rightarrow [0, \infty) \) is a convex \( C^1 \)-map. There exists a constant \( q \geq 1 \) such that \( \mathcal{R}_U(u) \geq \|u\|_q^q \) for all \( u \in U \). The functional \( \mathcal{R}_U \) is weakly lower semicontinuous with respect to the \( W \)-topology.

Under the assumptions (A1) and (A2), the strongly continuous operator \( F \) is moreover \textit{completely continuous}, i.e. compact and continuous. This makes problem (1) ill-posed (compare with \cite[Theorem 10.1]{2}). We remark that the reflexivity of \( U \) is not needed. It follows from (A1) that bounded sets in \( U \) are also bounded in \( W \) and hence they are weakly compact in \( W \). Given that \( U \) is reflexive, the results of this section are valid in the strong topology of \( U \) as well. The exponents \( p \) and \( q \) in the assumptions (A3) and (A4) naturally arise in practical applications when working with different Lebesgue and Sobolev spaces.

The assertion below provides a classical result about the existence of a minimizer of (6) and its characterization.

\textbf{Lemma 2.1} \textit{(well-posedness).} Assume (A1)–(A4). Let \( \lambda \in [0, 1] \) be arbitrary. Then there exists a minimizer \( (u^\lambda_\alpha, w^\lambda_\alpha) \) of \( T^H_\alpha \) for any \( \alpha > 0 \), which moreover satisfies the necessary condition
\[ D_\alpha T^H_\alpha(u^\lambda_\alpha, w^\lambda_\alpha, \lambda) = 0, \]
\[ D_\alpha T^H_\alpha(u^\lambda_\alpha, w^\lambda_\alpha, \lambda) = 0. \tag{11} \]

\textbf{Proof.} The proof is a straightforward application of the variational calculus. Let \( \lambda \in (0, 1) \). Since \( F \) is strongly continuous, the fidelity term is weakly lower semicontinuous. So is the regularization \( \mathcal{R}_W \) by the continuity and convexity argument and \( \mathcal{R}_U \) directly by the assumption (A4). The functional \( T^H_\alpha \) is their conical sum and hence it is weakly lower semicontinuous in \( W \) as well. It is also weakly coercive, i.e.
\[ T^H_\alpha(u, w, \lambda) > \alpha (\lambda \|u\|_U^p + (1 - \lambda) \|w\|_U^p) \rightarrow \infty \]
as \( \|u\|_U + \|w\|_W \rightarrow \infty \). Both properties of \( T^H_\alpha \) together imply that the functional \( T^H_\alpha \) attains its minimum (cf \cite{5, theorem 25.D}). As \( T^H_\alpha \) is differentiable, the minimizer solves equation (11). The case when \( \lambda = 0 \) and \( \lambda = 1 \) follows the same lines. \( \square \)

Expanding the condition (11) for \( \lambda \in (0, 1) \) reveals
\[ 2(F'(z) \cdot F(z) - v) + \alpha (\mathcal{R}_U'(u, \cdot)_{U'}) \lambda = 0, \]
and thus
\[ \langle \mathcal{R}_U'(u, \cdot)_{U'}, \cdot \rangle_{U'} = \langle \mathcal{R}_U'(w, \cdot)_{W'}, \cdot \rangle_{W'}. \]
where \( (\cdot, \cdot)_{U'} \) and \( (\cdot, \cdot)_{W'} \) are the dual pairing between \( U' \) and \( U \) and \( W' \) and \( W \), respectively. The above formula can be used to establish the so-called Ritz projection from the space \( W \) to its subspace \( U \).

\footnote{Note that \( F' : W \rightarrow L(W, V) \), and so \( F'(z) \in L(W, V) \) for \( z \in W \) and \( F'(z)h \in V \) for \( h \in W \).}
Lemma 2.2 (Ritz projection). Assume (A1), (A3) and (A4). Let \( U \) be a reflexive space and let the operator \( \mathcal{R}_U \) be strongly monotone. Then the map \( \mathcal{P} : W \to U \) such that \( w \mapsto \mathcal{P}(w) = u \), where \( u \) is the solution of the problem

\[
\langle \mathcal{R}_U'(u), h \rangle_U = \langle \mathcal{R}_W'(w), h \rangle_W, \quad \text{for all } h \in U,
\]

is well-defined and continuous.

Proof. It is sufficient to prove the unique solvability of problem (12). Since \( U \subset W \), it follows that \( W^* \subset U^* \), and hence \( \mathcal{R}_W'(w) \in U^* \). The assumption (A4) implies that \( \mathcal{R}_U' : U \to U^* \) is hemicontinuous, i.e. \( t \mapsto \langle \mathcal{R}_U'(u_1 + tu_2), h \rangle_U \) is continuous on \( [0, 1] \) for all \( u_1, u_2, h \in U \). Strong monotonicity implies coercivity. The theory of monotone operators (see [5, theorem 26.A]) then guarantees that for any \( w \in W \) there exists a unique \( u = \mathcal{P}(w) \) such that

\[
\mathcal{R}_U'(\mathcal{P}(w)) = \mathcal{R}_W'(w),
\]

and that \( [\mathcal{R}_U'(u)]^{-1} \) is Lipschitz continuous.

Remark 2.1. If the assumptions of the above lemma are satisfied, then for \( \lambda \in (0, 1) \) system (11) is equivalent to the system

\[
D_u T_u^H(\mathcal{P}(w), w, \lambda) = 0, \\
\mathcal{R}_U'(\mathcal{P}(w)) = \mathcal{R}_W'(w),
\]

and for \( \lambda = 0 \) we can still define ‘the minimizer’ \( u_0^\lambda(0) \) as the projection \( \mathcal{P}(w_0^\lambda(0)) \).

The following theorem provides the main result of this section. It establishes a continuous dependence of the minimizer of \( T_u^H \) on the parameter \( \lambda \) under a strong assumption of the unique minimizer of \( T_u^H \).

Theorem 2.1 (Continuous dependence on \( \lambda \)). Assume (A1)–(A4). Let \( \alpha > 0 \) and \( v^\delta \in V \). Assume that there exists a unique global minimizer \( (u_0^\lambda(\alpha), w_0^\lambda(\alpha)) \) of (6) for any \( \lambda \in [0, 1] \).

Then the mappings

\[
u_0^\lambda : [0, 1) \to W, \quad \lambda \mapsto u_0^\lambda(\lambda), \\
u_0^\lambda : (0, 1] \to U, \quad \lambda \mapsto u_0^\lambda(\lambda)
\]

are continuous.

Proof. We begin the proof with a simple estimate for the fidelity term in \( T_u^H \). By the mean value theorem we obtain

\[
\|F(z) - v^\delta\|^2_V = \|F(z) - F(w) + F(w) - v^\delta\|^2_V \\
\leq 2\|F(z) - F(w)\|^2_V + 2\|F(w) - v^\delta\|^2_V \\
\leq 2\|F(\xi)(\lambda u + (1 - \lambda)w - w)\|^2_V + 2\|F(w) - v^\delta\|^2_V \\
\leq 2\|F\|_{L(S,V)} \lambda \|u - w\|^2_V + 2\|F(w) - v^\delta\|^2_V,
\]

where the set \( S \) is the line segment \( u + t(w - u), t \in [0, 1] \). It is also evident that

\[
\mathcal{R}_{U,W}(u, v, \lambda) \leq \mathcal{R}_U(u) + \mathcal{R}_W(w)
\]

for any \( u \in U, \; v \in W \) and \( \lambda \in [0, 1] \).

This assumption is fulfilled for example for the case of a linear bounded forward operator mapping between Hilbert spaces and a strictly convex regularization term ([2, theorem 5.1]).

As we have mentioned, if \( \lambda = 0 \) and \( \lambda = 1 \), we consider just \( u_0^\lambda(0) \) and \( u_0^\lambda(1) \), respectively.
Let now $\lambda_k \to \lambda \in [0, 1]$ as $k \to \infty$. Denote by $(u_k, w_k)$ the corresponding global minimizer $(u^*_{\lambda_k}(\lambda_k), w^*_{\lambda_k}(\lambda_k))$ and set

$$z_k = \lambda_k u_k + (1 - \lambda_k) w_k.$$ 

By the definition of minimizer, it holds true that

$$T^H_\alpha (u_k, w_k, \lambda_k) \leq T^H_\alpha (u, w, \lambda_k)$$

for any $(u, w) \in \mathcal{D} \times \mathcal{D}(F)$. The minimum of $T^H_\alpha$ can be uniformly bounded for any $\lambda \in [0, 1]$ with the estimates (15) and (14):

$$\|F(z_k) - v^\delta\|^2_V + \alpha \|\lambda_k R_U(u_k) + (1 - \lambda_k) R_W (w_k)\|$$

$$\leq \|F(z) - v^\delta\|^2_V + \alpha \|R_U, W (u, w, \lambda_k)\|$$

$$\leq 2\|F'(u, w, \lambda)\|_V \|u - w\|^2 + 2\|F(w) - v^\delta\|^2_V + \alpha \|R_U + R_W (w)\|. \quad (16)$$

The assumptions (A3) and (A4) then show that

$$(1 - \lambda_k) \|w_k\|_W^p \leq (1 - \lambda_k) R_W (w_k) \leq C,$$

and

$$\lambda_k (\tilde{C} \|u_k\|_W)^q \leq \lambda_k \|u_k\|_U^q \leq \lambda_k R_U (u_k) \leq C,$$

where $C$ denotes a generic constant. Since $\lambda \in [0, 1]$ and $p, q \geq 1$, it easy to see that

$$\|\lambda_k u_k\|_W \leq \lambda_k \|u_k\|_W \leq C \quad \text{as well as} \quad \|\lambda_k u_k\|_W \leq C.$$

This on the other hand forces

$$\|z_k\|_W = \|\lambda_k u_k + (1 - \lambda_k) w_k\|_W$$

$$\leq \|\lambda_k u_k\|_W + \|\lambda_k u_k\|_W$$

$$\leq C^{1/q} + C^{1/p} \leq C.$$

Therefore, the sequence $[z_k]$ is uniformly bounded in $W$ for any $\lambda \in [0, 1]$. Both sequences $[u_k]$ and $[w_k]$ are bounded in $W$, unless $\lambda_k \to 0$ or $\lambda_k \to 1$, where the corresponding estimates are applicable only for $[u_k]$ or $[u_k]$, respectively.

Bounded sequences in reflexive spaces are weakly compact and so we can choose weakly convergent subsequences in $W$:

$$u_m \rightharpoonup \overline{u}, \quad w_m \rightharpoonup \overline{w} \quad \text{and} \quad z_m \rightharpoonup \overline{z} \quad \text{as} \quad m \to \infty. \quad (17)$$

The definition of $z$ guarantees that

$$\overline{z} = \lambda \overline{u} + (1 - \lambda) \overline{w} \quad \text{for any} \lambda \in [0, 1].$$

We then consecutively deduce by the weak lower semicontinuity of $T^H_\alpha$ and the definition of minimizer that

$$\liminf_{m \to \infty} \|F(z_m) - v^\delta\|^2_V + \alpha R_U, W (\overline{u}, \overline{w}, \lambda)$$

$$\leq \liminf_{m \to \infty} \|F(z_m) - v^\delta\|^2_V + \alpha R_U, W (u_m, w_m, \lambda_m)$$

$$\leq \limsup_{m \to \infty} \|F(\lambda_m u_m + (1 - \lambda_m) w_m) - v^\delta\|^2_V + \alpha R_U, W (u_m, w_m, \lambda_m)$$

$$\leq \lim_{m \to \infty} \|F(\lambda_m u + (1 - \lambda_m) w) - v^\delta\|^2_V + \alpha R_U, W (u, w, \lambda_m)$$

$$= \|F(z) - v^\delta\|^2_V + \alpha R_U, W (u, w, \lambda)$$

for all $(u, w) \in \mathcal{D} \times \mathcal{D}(F)$. This shows that $(\overline{u}, \overline{w})$ is a minimizer of (9) and that

$$\lim_{m \to \infty} T^H_\alpha (u_m, w_m, \lambda_m) = T^H_\alpha (\overline{u}, \overline{w}, \lambda). \quad (18)$$
Assume now that \((u_n, w_n) \not\to (\bar{u}, \bar{w})\). Then \(c := \limsup_{n \to \infty} T_a^H(u_n, w_n, \lambda) > T_a^H(\bar{u}, \bar{w}, \lambda)\) and there exists a subsequence \(\{(u_n, w_n)\}\) of \(\{(u_m, w_m)\}\) such that \((u_n, w_n) \to (\bar{u}, \bar{w})\) and \(T_a^H(u_n, w_n, \lambda) \to c\). As a consequence of \((18)\), we obtain
\[
\lim_{n \to \infty} T_a^H(u_n, w_n, \lambda) = \lim_{n \to \infty} T_a^H(u_n, w_n, \lambda) \pm T_a^H(u_n, w_n, \lambda) \\
= T_a^H(\bar{u}, \bar{w}, \lambda) + (T_a^H(\bar{u}, \bar{w}, \lambda) - c) < T_a^H(\bar{u}, \bar{w}, \lambda),
\]
which is in contradiction with weak lower semicontinuity of \(T_a^H\).

Since the global minimizer \((\bar{u}, \bar{w})\) is unique for any \(\lambda \in [0, 1]\), the above considerations demonstrate that every sequence \(\{(u_n, w_n)\}\) contains a subsequence strongly converging toward \((\bar{u}, \bar{w})\), and therefore, the functions \(u^\alpha_n\) and \(w^\alpha_n\) are continuous on the intervals \((0, 1)\) and \([0, 1]\), respectively.

The immediate consequence of the above theorem is that the continuation method arrives at a solution of the original problem, which is defined by \((4)\); it is \(u^\alpha_n(\lambda) \to u^\alpha(\lambda)\) as \(\lambda \to 1\).

The following theorems address the questions of stability and convergence of minimizers of \(T_a^H\). We omit their proofs, because they go along the same lines as e.g. in [2, theorems 10.2 and 10.3].

**Theorem 2.2** (stability). Assume \((A1)–(A4)\), \(\alpha > 0\) and \(v^\delta \in V\). Let \(\lambda \in [0, 1]\) be fixed and let \(\{v_k\}\) and \(\{(u_k, w_k)\}\) be sequences such that \(v_k \to v^\delta\) and \((u_k, w_k)\) is a minimizer of \((6)\) with \(v^\delta\) replaced by \(v_k\). Then there exists a convergent subsequence of \(\{(u_k, w_k)\}\) and the limit of every convergent subsequence is a minimizer of \((6)\).

**Theorem 2.3** (convergence). Assume \((A1)–(A4)\). Let \(v^\delta \in V\) with \(\|v - v^\delta\|_V \leq \delta\) and let \(\lambda \in [0, 1]\) be fixed. Let \(\alpha(\delta)\) be such that \(\alpha(\delta) \to 0\), and \(\delta^2/\alpha(\delta) \to 0\) as \(\delta \to 0\). Then every sequence \(\{(u_k^\delta(\lambda), w_k^\delta(\lambda))\}\), where \(\delta_k \to 0\), \(\alpha_k := \alpha(\delta_k)\) and \((u_k^\alpha(\lambda), w_k^\alpha(\lambda))\) is a solution of \((9)\), has a convergent subsequence. The limit of every convergent subsequence is an \(R_{U,W}\)-minimizing solution. If in addition, the \(R_{U,W}\)-minimizing solution \((u^\delta, w^\delta)\) is unique, then
\[
\lim_{\delta \to 0} (u^\delta(\alpha(\delta)), w^\delta(\alpha(\delta))) = (u^\lambda, w^\lambda).
\]

The last result about the existence of an \(R_{U,W}\)-minimizing solution is essentially due to [1].

**Lemma 2.3.** Assume \((A1)–(A4)\). If there exists a solution of \((1)\), then there exists an \(R_{U,W}\)-minimizing solution for any \(\lambda \in [0, 1]\).

### 3. Piecewise-constant parameter identification problems

Our motivation to study minimizers of \((6)\) comes from PCPIPs. We analyze partial differential equation (PDE)-constrained problems with the unknown parameter being a coefficient of the PDE-constraint.

For illustration purposes, we consider merely a binary piecewise-constant parameter
\[
\sigma_{PC} = \sigma_1 \chi_D + \sigma_2 \chi_{\Omega \setminus D}, \quad \sigma_1, \sigma_2 \in \mathbb{R},
\]
where the domain \(\Omega\) is an open-bounded set, on which the PDE-constrained problem is defined. The symbols \(\chi_D\) and \(\chi_{\Omega \setminus D}\) stand for the characteristic function of subset \(D \subset \Omega\) and its complement, respectively. The goal is to find the subdomain \(D\) and the unknown numbers \(\sigma_1\) and \(\sigma_2\) based on suitable observations of the state variable of the PDE-constraint. A classical example here is the problem of inverse electric impedance tomography (EIT).
We are primarily concerned with building a robust and efficient numerical algorithm to recover the unknown \( \sigma_{PC} \). In the case of EIT, the problem is extensively studied in the literature, see a comprehensive review [6].

Why do we look for the solution in the space of piecewise-constant functions? Such a choice is natural, given a problem like EIT. First, this class of functions is rich enough in order to be applicable. Second, as in the case of EIT, one usually has only a finite number of measurements on the boundary \( \Gamma \) corresponding to the Neumann-to-Dirichlet operator. For a two-dimensional domain \( \Omega \), these measurements are one dimensional. It is reasonable to assume, that we can successfully recover at most a one-dimensional unknown inside the domain. This is precisely what one does by considering (19). The goal is as a matter of fact to find the interface between the two regions of \( \Omega \). The choice of space plays a role of regularization.

\[
U = BV(\Omega) \text{: the most suitable type of regularization for PCPIPs is the } BV(\Omega)-\text{regularization [7]. The space } BV(\Omega) \text{ is the subspace of functions } u \in L^1(\Omega) \text{ such that the quantity } \\
J(u, \Omega) := \sup \left\{ \int_{\Omega} u(x) \nabla \cdot \xi(x) \, dx \mid \xi \in C^\infty_c(\Omega, \mathbb{R}^n), \|\xi\|_{L^\infty(\Omega, \mathbb{R}^n)} \leq 1 \right\}, \\
\text{is finite, where } C^\infty_c(\Omega, \mathbb{R}^n) \text{ is the set of smooth functions in } C^\infty(\mathbb{R}^n) \text{ with compact support in } \Omega. \end{equation}
Endowed with the norm

\[
\|u\|_{BV(\Omega)} := \|u\|_{L^1(\Omega)} + J(u, \Omega),
\]

it is a Banach space. For \( u \in W^{1,1}(\Omega) \), the seminorm \( J(u, \Omega) \) can be written as

\[
J(u, \Omega) = \int_{\Omega} |\nabla u| \, dx.
\]

A slightly more general functional is usually considered, namely

\[
J_\varepsilon(u, \Omega) = \int_{\Omega} \sqrt{|\nabla u|^2 + \varepsilon} \, dx,
\]

where \( 0 < \varepsilon \ll 1 \). Its variational definition for any \( u \in BV(\Omega) \) can be found in [7]. The functional \( J_\varepsilon \) efficiently approximates the \( BV \)-seminorm [7, theorem 2.2] and it is differentiable if \( |\nabla u| = 0 \). We remark that \( \varepsilon \) will be used subsequently in different situations where it always represents a small positive number.

Tikhonov regularization formulation for the PCPIP then reads as

\[
T_\alpha(\sigma_{PC}) := \|F(\sigma_{PC}) - v^\delta\|_V^2 + \alpha \|\sigma_{PC}\|_{BV(\Omega)}^2,
\]

where \( F \) is the operator associated with the forward problem. This functional is a particular case of the functional (3) from the introduction when we set \( U = BV(\Omega) \).

3.1. State of the art of geometry (shape) identification

In case the constants \( \sigma_1 \) and \( \sigma_2 \) in (19) are identified, the piecewise-constant parameter \( \sigma \) estimation is equivalent to the geometry identification of the subdomain \( D \).

The classical methods to identify the structural information are mostly based on a study of the sensitivity of certain cost functional to an infinitesimal change of the shape of the structure itself, see [8] and the references therein. This shape-sensitivity approach yields eventually to the notion of shape derivative [9].

The methods based on the shape-sensitivity approach, level set method (LSM) parameterizations including [10, 11], are updating the shape of domain first, not the topology.

---

5 We do not claim that certain two-dimensional recovery is impossible.
The topology is prescribed *a priori* by an initial guess. The choice of a good initial guess becomes very important for the method to converge to the optimal shape. Even if some proposed (and well-designed) algorithms are able to find the optimal shape [12], the convergence is usually very slow. The speed of the convergence is again strongly dependent on a good initial guess.

The second class of methods is based on the *homogenization* theory, see the pioneering work [13] or the monograph [14]. The optimal geometry is obtained in an enriched space of composite designs. The corresponding classical design can be retrieved via thresholding or penalization. This approach overcomes some restrictions of the classical shape-sensitivity approach. Both the topology and shape are optimized at once. The final acquired geometries are close to the optimal ones. Unfortunately, this approach is limited to certain types of problems and its rigorous application is a non-trivial task.

A method based on an iterative inclusion of new holes (the so-called bubbles) into the geometry was investigated in [15]. This idea is actually closely related to the one of the homogenization approach. In [16], a pointwise limit of such inclusions was used in linear elasticity to find an optimal design characterized by the so-called compliance functional. The importance of this contribution was recognized in [17–19], where the idea was extended to shape functionals and the notion of *topological derivative* was introduced and further developed. Since the introduction of the topological derivative, a great number of contributions were made using this concept both in science and engineering. We are interested particularly in those where topological and shape-sensitivity concepts are used in conjunction.

In [20], the authors first considered the shape-derivative-based LSM. The motion of the interface described by the LSM is governed by a nonlinear Hamilton–Jacobi equation, where speed is dependent on shape derivative of the cost functional, as usual. The idea was to introduce a new source term into the Hamilton–Jacobi equation, dependent on the topological derivative. This term allows for the nucleation of new holes in the domain. The approach was generalized in [21].

In [22], the authors study shape-derivative-based LSM for structural optimization. They do not use the topological derivative in the work itself, but, to our best knowledge, for the first time the topological derivative is suggested to be used for the initialization of the algorithms based on the shape-sensitivity approach. They study the idea in [23], where an alternating algorithm using both the shape and the topological derivatives is proposed.

In [24], the authors propose a variant of a binary level set approach for solving elliptic problems with piecewise-constant coefficients. The inverse problem is solved by a variational augmented Lagrangian approach with a total variation regularization. Their implementation was able to recover rather complicated geometries without assuming anything about *Da priori*, i.e. without any initial guess. As we will understand later on, it is due to the nature of the augmented Lagrangian approach which imposes the piecewise-constant constraint gradually. The results of [24] are applied to piecewise-constant LSM (PCLSM) parametrization in [25]. They are employed to study an optimization problem. The PCLSM methods for the identification of discontinuous parameters in ill-posed problems are considered in [26]. Both a Tikhonov regularization approach using operator splitting techniques and an augmented Lagrangian approach are introduced and analyzed.

In [27], topological-sensitivity-based initial guess is used as the starting point for shape-sensitivity LSM to solve an EIT problem.

Despite all the effort in combining topology and shape-sensitivity concepts and some very positive results as stated above, no clear idea, to our knowledge, has yet been presented how these concepts could be unified in one framework. We quote [28]: ‘it is still an open problem to
devise how the combination of boundary variations and singular perturbations of geometrical domains enters in a general framework of shape optimization.

3.2. Topology-to-shape continuation method

In this section, we introduce a continuation approach to shape identification which combines topology and shape sensitivities.

The main idea is based on the following reasoning. Roughly speaking, topological properties of a particular shape are those which stay invariant under various continuous transformations. A shape itself is a certain topology modified by those continuous boundary-like transformations, see the above section. Therefore, the topology is the ‘coarse’ information about a particular shape. In this line of reasoning, it is intuitive to first look for the topology itself and to consider continuation methods to transform it to the particular shape.

We will consider the relaxed parametrization of $\sigma_{pc}$,

$$\sigma = (1 - \lambda)\sigma_{L^2} + \lambda\sigma_{PC},$$  

(22)

analogously to (7). We assume that $\sigma_{L^2} \in L^2(\Omega)$, because the space $U = BV(\Omega)$ is included at most in $W = L^2(\Omega)$, in the case if the domain $\Omega \subset \mathbb{R}^2$.

The function $\sigma_{L^2}$ can be interpreted as topological derivative. It is almost everywhere locally defined and represents the distribution of the mass in $\Omega$. The optimization with respect to $\sigma_{L^2}$ means adding and removing mass locally at a given point in the domain. On the other hand, the optimization with respect to $\sigma_{PC}$ is driven by shape derivative flux and moves only the interface $\partial D$.

In analogy to (8), let us first consider the following functional

$$\lambda \|\sigma_{PC}\|^2_{BV(\Omega)} + (1 - \lambda) \|\sigma_{L^2}\|^2_{L^2(\Omega)}$$

as a candidate for $R_{U,W}(\sigma_{PC}, \sigma_{L^2}, \lambda)$ and let us verify the assumptions of the theoretical framework in section 2. It is easy to see that the choice of the spaces $W = L^2(\Omega)$ and $U = BV(\Omega)$ complies with the assumption (A1). Likewise, the $R_W = \|\cdot\|^2_{L^2(\Omega)}$ trivially fulfils the assumption (A3) for $p = 2$. The problematic assumption is (A4), as the $BV$-norm is not differentiable. We will work instead with the functional $R_{U,W}$ below:

$$R_{U,W}(\sigma_{PC}, \sigma_{L^2}, \lambda) = \lambda \int_{\Omega} \sqrt{\|\sigma_{PC}\|^2 + \varepsilon} \, dx + \lambda J_{\varepsilon}(\sigma_{PC}, \Omega) + (1 - \lambda) \int_{\Omega} |\sigma_{L^2}|^2 \, dx,$$  

(23)

where the first term approximates the $L^1$-norm and the second one has been introduced by (20). The functional (23) is now differentiable. It is convex and weakly lower semicontinuous [7, theorem 2.3] with respect to the $W$-topology as well. Consequently, the regularization functional above also fulfils the assumption (A4). We conclude that for an admissible forward operator $F$ (see (A2)), the TSCM with $R_{U,W}$ as defined in (23) lies within the proposed continuation framework (sections 1 and 2).

Let us remark that we deal with a constrained problem because of the piecewise-constant restriction for $\sigma_{PC}$. Using an LSM parametrization we obtain again an unconstrained problem with the corresponding level set function $\phi$ as the free variable (section 4.4). For a proper analysis, a regularization term penalizing $\phi$ has to be added to the Tikhonov functional. We refer the reader to [29] for a detailed theoretical treatment.

---

6 In our case, when the ‘shape’ of the piecewise-constant $\sigma$ is defined by (19), the topology is determined by the number of connected components of $D$ and their equivalent classes (ball, torus etc).

7 A possible remedy is to approximate $BV(\Omega)$ by its reflexive subspace $W^{1,1+r}(\Omega), 0 < \eta \ll 1$. 

---
4. Magnetic induction tomography

In this section, we apply the framework to an inverse problem in MIT.

MIT is a non-invasive visualization technique, which is a very promising member of the broader electromagnetic imaging family. It has many potential applications, for instance non-destructive testing, industrial and medical imaging [30]. We refer the reader to the paper [31] for a comprehensive review. MIT is a non-contact technique, in contrast to widely studied electrical impedance tomography [32]. Another advantage of MIT is its explicit frequency dependence, which allows for more accurate reconstruction of the body properties [33].

4.1. Mathematical formulation

We proceed to the mathematical description of MIT. Electromagnetic phenomena in general are governed by the Maxwell equations. Considering the linear isotropic case, the time-harmonic regime with the angular velocity \( \omega > 0 \) and making use of the magnetic vector potential \( \mathbf{A} \) \( (\mathbf{B} = \nabla \times \mathbf{A}) \), we can write them in the form:

\[
\nabla \times (\mu^{-1} \nabla \times \mathbf{A}) + i \omega (\sigma + i \omega \varepsilon) \mathbf{A} = \mathbf{J}, \\
\n\nabla \cdot (\varepsilon \mathbf{A}) = 0.
\]

The scalar potential \( V \) is eliminated by the temporal gauge; \( V = 0 \). The permeability \( \mu \) and the permittivity \( \varepsilon \) are the known strictly positive scalar functions of the space variable. The conductivity \( \sigma \) is assumed to be positive in the imaged body and it vanishes in the surrounding non-conducting region; \( \mathbf{J} \) stands for the applied current from the excitation coil. For more on various MIT models we refer to [31, 34].

We formulate a simplified MIT boundary value problem. Let \( \Omega \) be a bounded two-dimensional domain in the \( xy \)-plane with the sufficiently smooth boundary \( \partial \Omega =: \Gamma \). It represents a cross section of the imaged body. Assume that the applied current \( \mathbf{J} \) is perpendicular to \( xy \)-plane and does not depend on \( z \)-coordinate. The induced eddy currents can be then described by the \( z \)-component of the potential \( \mathbf{A} \) which we will simply denote by \( A \). We restrict ourselves to the imaged body region, where the conductivity is strictly positive, \( \sigma \geq \sigma_{\min} > 0 \). The domain source \( \mathbf{J} \) is modeled by a boundary source \( \mathbf{e} \), which is imposed via the Neumann boundary condition on \( \Gamma \). The corresponding experimental setup is depicted in figure 1.

For an experimental realization see [35].
We use the eddy current approximation of the Maxwell equations, where the displacement current term $iωεA$ in (24) is disregarded. The state variable $A$ then satisfies the forward problem

$$\nabla \cdot (μ^{-1}\nabla A) + iωσA = 0 \quad \text{in } Ω,$$

$$μ^{-1}\nabla A \cdot n = e \quad \text{on } Γ.$$  \hspace{1cm} (25)

Let us remark that under physiological conditions for higher excitation frequencies $ω$, the displacement current term can have a significant contribution and has to be taken into account.

4.2. Forward problem

We now show that the MIT forward problem satisfies the assumption (A2) of section 2.

Let us first introduce some notation. The standard scalar product of two complex-valued functions in the space $L^2(Ω)$ is denoted by $(u, v) = \int_Ω u(x)v(x) dx$. We write $∥u∥ = √(u, u)$ for the induced norm. The subscript $Γ$ indicates integration over the boundary in $L^2(Γ)$-sense. The symbol $H^1(Ω)$ stands for the Sobolev space of the complex-valued functions with first weak derivatives. It is compactly embedded in the all Lebesgue spaces but $L^∞(Ω)$ (e.g. [36, theorem 5.8.2]):

$$H^1(Ω) \hookrightarrow L^q(Ω) \quad \text{for any } q ∈ [1, ∞).$$  \hspace{1cm} (26)

The weak formulation of (25) reads as

$$(μ^{-1}\nabla A, \nabla ϕ) + (iωσA, ϕ) = (e, ϕ)_{Γ} \quad ∀ϕ ∈ H^1(Ω).$$  \hspace{1cm} (27)

This variational problem defines the impedance map $Λ$, the so-called Neumann-to-Dirichlet map

$$Λ : (σ, ω, e) ↦ A|Γ.$$  \hspace{1cm} (28)

Lemma 4.1. The impedance map

$$Λ : σ ↦ Λ(σ) = A|Γ,$$

where the function $A$ is the solution of problem (27) for any $e ∈ L^2(Γ)$ and $ω > 0$ fixed, is a well-defined and strongly continuous map from the set

$$M = \{σ ∈ L^q(Ω), q > 1 : σ ≥ σ_{min} > 0\}$$

to the space $L^2(Γ)$.

Proof. The Sobolev embedding (26) implies that term in (27) containing $σ$ makes sense for any $σ ∈ L^q(Ω), q > 1$. Given arbitrary $σ ∈ M$, the existence of a unique solution $A ∈ H^1(Ω)$ follows readily from the Lax–Milgram theorem for sesquilinear forms.

Let now $σ_n ↦ σ$ as $n → ∞$. It holds that $σ ∈ M$, because $M$ is closed and convex. Denote by $A_n$ and $A$ the corresponding solutions of (27) for $σ_n$ and the weak limit $σ$, respectively. The subtraction of the variational formulas from each other gives

$$(μ^{-1}\nabla (A - A_n), \nabla ϕ) + (iωσ(A - A_n), ϕ) = (iω(σ_n - σ)A, ϕ).$$

The sesquilinear form on the left-hand side is equivalent to the $H^1(Ω)$-scalar product which leads to a one-to-one correspondence between test functions $ϕ$ and linear functionals on $H^1(Ω)$. Since $A_ϕ ∈ L^{(q-1)}(Ω)$, the right-hand side tends to zero for any $ϕ ∈ H^1(Ω)$ as $n → ∞$. We hence see that

$$A_n → A \quad \text{in } H^1(Ω).$$
It follows from continuity of the trace mapping $H^1(\Omega) \to H^{1/2}(\Gamma)$ and the compact embedding $H^{1/2}(\Gamma) \hookrightarrow L^2(\Gamma)$ that
\[ A_n \to A \quad \text{in } L^2(\Gamma). \]

The differentiation of (27) at $\sigma$ in the direction $h$ yields
\[ \mu^{-1} \nabla \delta A, \nabla \varphi + (i \omega \varphi, \varphi) = - (i \omega h, \varphi) \quad \forall \varphi \in H^1(\Omega). \] (29)

The symbol $\delta A := \delta A(\sigma; h)$ stands for the variation (Gâteaux differential) of $A = A(\sigma)$ in the direction $h$. The variation $\delta A$ is sometimes called the sensitivity of $A$ and (29) the sensitivity equation, which is a well-posed problem with the unique solution $\delta A$ for any $h$ from $L^q(\Omega)$, $q > 1$. It is straightforward to verify that for given $\sigma$ the mapping $h \mapsto \delta A(\sigma; h)|_{\Gamma}$ is linear and bounded operator in $L(M, L^2(\Gamma))$. Recalling the relationship between the variation and Fréchet derivative, we see that $\Lambda$ is Fréchet differentiable at $\sigma$ and
\[ \Lambda'(\sigma) h = \delta \Lambda(\sigma; h)|_{\Gamma}. \]

The map $\Lambda' : M \to L(M, L^2(\Gamma))$ is continuous in $\sigma$ by the similar reasoning as in the proof of lemma 4.1 and so we have the following assertion.

**Lemma 4.2.** The impedance map $\Lambda : M \to L^2(\Gamma)$ is $C^1$-Fréchet differentiable.

### 4.3. Inverse problem

By the inverse problem in MIT we will understand the reconstruction of the piecewise-constant conductivity $\sigma$ in the imaged body based on a finite number of Dirichlet-to-Neumann data $(e, m)$ corresponding to the impedance map (28). The boundary data $m$ are essentially voltage measurements associated with excitations $e$. Lemma 4.1 implies that $\Lambda$ is a compact operator and so the recovery of $\sigma$ is inherently an ill-posed problem.

We employ the TSCM from section 3.2 to solve MIT. We look for the conductivity $\sigma$ in the form (22), i.e.
\[ \sigma = (1 - \lambda) \sigma_{L^2} + \lambda \sigma_{PC}, \]
where $\sigma_{PC}$ is a double-valued piecewise-constant function as it is considered in section 3 for the example of electrical impedance tomography. The associated continuation Tikhonov functional for MIT read as
\[ T^h_{\mu}(\sigma) = F(\sigma) + \alpha R_{U,W}(\sigma_{PC}, \sigma_{L^2}, \lambda), \] (30)
where the regularization part $R_{U,W}$ is given by (23) and $F$ is the fidelity term
\[ F(\sigma) = \int_{\Gamma} |\Lambda(\sigma, \omega, e) - m|^2 \, dS. \] (31)

The forward problem operator $\Lambda$ of MIT is an admissible operator fulfilling assumption (A2) of section 2 as it is shown in section 4.2. Altogether, the theory of section 2 is applicable to the inverse problem of MIT as stated in this section.

#### 4.3.1 Adjoint problem.

In section 4.5, we will use a gradient-based method (the steepest descent method) to find a minimizer of (30). Let us express the derivative of fidelity term (31) using an adjoint variable. The variation of $F$ in the direction $h$ reads as
\[ \delta F(\sigma; h) = \lim_{t \to 0} \frac{F(\sigma + th) - F(\sigma)}{t} \]
\[ = (\Lambda(\sigma) - m, \delta \Lambda(\sigma; h))_{\Gamma} + (\delta \Lambda(\sigma; h), \Lambda(\sigma) - m)_{\Gamma} \]
\[ = 2i \Re [(\delta \Lambda(\sigma; h), \Lambda(\sigma) - m)_{\Gamma}], \]
where the variation $\delta \Lambda (\sigma; h) \equiv \delta A$ solves the sensitivity equation (29). We now introduce the adjoint variable $Z$ which satisfies

$$
(\mu - 1 \nabla \phi, \nabla Z) + (i \omega \sigma \phi, Z) = - (\phi, \Lambda (\sigma) - m)_{\Gamma} \quad \forall \phi \text{ in } H^1(\Omega),
$$

(32)


to establish that

$$
\delta F(\sigma; h) = 2 \Re \left[ (\delta \Lambda (\sigma; h), \Lambda (\sigma) - m)_{\Gamma} \right] = 2 \Re \left[ - (\mu - 1 \nabla \delta A, \nabla Z) - (i \omega \sigma \delta A, Z) \right] = 2 \Re \left[ (i \omega h A, Z) \right].
$$

(33)

Let us note, that the variational problem (32) for $Z$ is uniquely solvable given the properties of the material parameters and of the impedance map $\Lambda$. We assume that $m \in L^2(\Gamma)$.

4.4. Implementation of the TSCM

In this section, we describe the implementation of the TSCM for the problem of the MIT.

The practical implementation of the TSCM algorithm presented in algorithm 1 closely follows the theoretical exposition. The outer loop successively increases the value of $\lambda$ by the increment $\Delta \lambda$ starting from $\lambda = 0$. It terminates when $\lambda = 1$ is reached. The number of steps is determined by $\Delta \lambda$. The inner loop constitute more or less a standard adjoint-variable-based steepest descent algorithm for the minimization of (30) for the fixed $\lambda$. The number $n$ stands for the total number of iterations through both loops in algorithm 1.

**Algorithm 1**: Topology-to-shape continuation algorithm.

**Data**: $n = 0$; $\lambda = 0$; $\sigma_n = \sigma_{l2,n} = \delta_1$; $\phi_n = -\delta_1$;

**do**

$s_n = 2$;

**do**

**Compute the derivatives:**

$\sigma_n$ $\rightarrow$ direct problem (27) $\rightarrow$ $A_n$;

$(\sigma_n, A_n)$ $\rightarrow$ adjoint problem (32) $\rightarrow$ $Z_n$;

$(A_n, Z_n)$ $\rightarrow$ cost functional derivative (33) $\rightarrow$ $\nabla_\sigma F_n$;

$\nabla_\sigma F_n + (40) + (37) \rightarrow \nabla_\phi T_{\theta,n}$;

$\nabla_\phi T_{\theta,n} + (39) + (38) \rightarrow \nabla_{\phi T_{\theta,n}} T_{\theta,n}$;

**Find the optimal step:**

$s_n = \text{Linesearch}(\sigma_n, \nabla_\sigma T_{\theta,n}, \nabla_\phi T_{\theta,n})$;

$\text{Update } \sigma_n$: $\sigma_{l2,n+1} = \sigma_{l2,n} - s_n \nabla_\sigma T_{\theta,n}$;

$\phi_{n+1} = \phi_n - s_n \nabla_\phi T_{\theta,n}$;

$\sigma_{n+1} = \lambda \sigma_{PC}(\phi_{n+1}) + (1 - \lambda) \sigma_{l2,n+1}$;

$n = n + 1$;

**while** $|\nabla_\sigma T_{\theta,n}|^2 + |\nabla_\phi T_{\theta,n}|^2 > \tau_1^2$ and $s_n > \tau_2$;

$\lambda = \lambda + \Delta \lambda$;

**while** $\lambda < 1$;

We use the LSM [37] to parametrize the conductivity $\sigma_{PC}$ introduced in (19). One first defines the level set function $\phi$ for the subset $D \subset \Omega$ with its boundary $\partial D$:

$$
\phi(x) = \begin{cases} 
\text{distance}(x, \partial D) & x \in D, \\
-\text{distance}(x, \partial D) & x \in \Omega/D.
\end{cases}
$$

14
The zero-level set of $\phi$ represents the boundary of $D$ (its ‘interface’). The piecewise-constant conductivity $\sigma_{PC}$ is then parametrized as
\[
\sigma_{PC}(\phi) = \sigma_1 H(\phi) + \sigma_2 (1 - H(\phi)),
\] (34)
where $H$ stands for the unit step Heaviside function. We use the following smooth approximations of $H$ and its derivative:
\[
H_k(\phi) = \frac{1}{\pi} \arctan \frac{\phi}{k} + \frac{1}{2}, \quad H_k'(\phi) = \delta_k(\phi) = \frac{\epsilon}{\pi (\phi^2 + \epsilon^2)}.
\] (35)

The gradient of (23) with respect to $\sigma_{PC}$ is evaluated as the solution of the variational problem
\[
(\nabla_{\sigma_{PC}} R_{U,W}, h) = \lambda \left( \frac{\nabla \sigma_{PC}}{\sqrt{|\nabla \sigma_{PC}|^2 + \epsilon}}, \nabla h \right) + \left( \frac{\sigma_{PC}}{\sqrt{|\sigma_{PC}|^2 + \epsilon}}, h \right)
\] (36)
for all $h \in H_1^0(\Omega)$. It is, in fact, a projection of $\delta_{\sigma_{PC}} R_{U,W}$ onto the nodes of the finite element mesh. We remark that all the variational problems ((27), (32) etc) are solved by the finite element method where $H^1(\Omega)$ is approximated by linear Lagrange basis functions. Using (34) together with (35), we have
\[
\nabla_{\phi} R_{U,W} = (\sigma_1 - \sigma_2) H_k'(\phi) \nabla_{\sigma_{PC}} R_{U,W}. \quad (37)
\]
The gradient of (23) with respect to $\sigma_{L_2}$ is simply
\[
\nabla_{\sigma_{L_2}} R_{U,W} = 2(1 - \lambda) \sigma_{L_2}. \quad (38)
\]
The gradient $\nabla_{\sigma} F$ of the fidelity term $F$ with respect to $\sigma$ is evaluated from (33) again by projection onto the nodes of the finite element mesh as in (36):
\[
\nabla_{\sigma} F = 2h[\text{ItoAZ}].
\]
This yields
\[
\nabla_{\sigma_{L_2}} F = (1 - \lambda) \nabla_{\sigma} F
\] (39)
and
\[
\nabla_{\phi} F = \lambda (\sigma_1 - \sigma_2) H_k'(\phi) \nabla_{\sigma} F. \quad (40)
\]

We do not optimize with respect to the constants $\sigma_1$ and $\sigma_2$, which we consider to be known. However, algorithm 1 is easily extendable to the case of unknown $\sigma_1$ and $\sigma_2$.

We emphasize that we do not assume any $a$ priori knowledge about the shape of $D$. The unknowns $\phi$ and $\sigma_{L_2}$ are initiated as $\phi = -\delta_1$ and $\sigma_{L_2} = \delta_2$ with $\delta_1$ and $\delta_2$ being some positive constants, $\delta_2 \approx \sigma_{\text{min}}$. It means that initially ($\lambda = 0$) the whole domain $\Omega$ is occupied by a weak phase. In addition, we have zero inclusion $D$ and thus the value of $\sigma_{PC}$ is $\sigma_2$ in the whole domain.

In algorithm 1, the search for an optimal step-size $s_n$ might be the most time-consuming part, since the Linesearch-algorithm detects the optimal $s_n$ by the evaluation of the cost functional for different intermediate values of $s_n$ and one such evaluation means to solve one forward problem (27). However, we do not need to find the optimal value of $s_n$ for which the drop of $T^H_u$ is maximal. It is enough to find one value for which $T^H_u$ drops sufficiently (the method is then no more steepest descent). We update $s$ according to the following simple rule [38]:
\[
s_{n+1} = 2s_n \text{ if } T^H_u(\sigma_n(s_{n-1})) < T^H_u(\sigma_{n-1}),
\]
i.e. when $s_{n-1} := s_{n-2}$ gave a reduction of cost functional value, we try double the step. If in the next step $s_n$ does not give a descent, we take the step with the smallest $k$ from the sequence $s_n^k = s_{n-1}^k/2, \quad k = 1, \ldots, \infty$ such that we have descent. The last part is the actual update process. The inner cycle of algorithm 1 stops when the norm of gradient is sufficiently small ($\leq \tau_1$) or the computed gradient is not a descent direction anymore, i.e. $s_n \leq \tau_2$, where $\tau_1$ and $\tau_2$ are suitable constants.
In all the experiments we use synthetic data. The number $N$ of the measurements for every experiment corresponds to the number of excitation coils $N(e)$ (see figure 1) multiplied with the number of excitation frequencies $N(\omega)$. The fidelity functional reads

$$
\mathcal{F}(\sigma) = \sum_{\omega} \sum_{e} \int_{\Gamma} |\Lambda(\sigma, \omega, e) - m|^2 \, dS. 
$$

We take $\sigma_1 = 20 \, \text{S} \, \text{m}^{-1}$ and $\sigma_2 = 2 \, \text{S} \, \text{m}^{-1}$ and $\mu = \mu_0$ which complies with physiological conditions. For comparison, in non-destructive testing of metallic pieces normal magnitudes of $\sigma$ are in millions of $\text{S} \cdot \text{m}^{-1}$ and $\mu \gg \mu_0$.

All the excitation currents $e_i = 1 \, A \, \text{m}^{-1}$, $i = 1, \ldots, N(e)$. The angular excitation frequencies $\omega_i = 2\pi f_i = 2\pi 2^{15+i}$, $i = 0, \ldots, N(\omega) - 1$. The basic frequency $f_0 = 2^{15}$ is set so that $\mu^{-1} > \omega_0 \max(\sigma_1, \sigma_2)$. For such a base frequency, the magnetic phenomena dominate the electric ones.

The parameters in algorithm 1 are $\tau_1 = 10^{-5}$, $\tau_2 = 10^{-6}$, $\delta_1 = 1$, $\delta_2 = 0.01$. We implemented the algorithm in FreeFem++ [39] and the source code is freely available at [40]. In all the experiments for both $\sigma_1$, $\mu$ we use identical fixed regular meshes with homogeneous division of the boundary $\Gamma$. We also always consider 28 excitation coils, i.e. $N(e) = 28$, and the regularization parameter $\alpha = 0.00001$. In (35) we take $\epsilon = h^2$, where $h$ is the diameter of the finite element mesh. If not stated otherwise, we take $\Delta \lambda = 0.1$.

We first compare the performance of the continuation algorithm (TSCM) and the standard LSM on an example with a non-trivial topology (figure 2). The blue dotted line represents in all the figures the exact phantom and the red line is the numerical approximation. The initial shape of $\sigma_{PC}$ for the standard LSM is depicted in figure 2(a). Figure 2 displays the results for the base angular frequency $\omega_0$. The LSM in figure 2(a) ended up in a local minimum after 37 iterations. The algorithm stopped because the computed gradient was not a descent direction anymore, i.e. $\nabla \mathcal{F} < \tau_2$. We see that without a proper initial guess, the standard LSM failed to recover the desired shape. On the other hand, the TSCM in figure 2(c) for zero noise provided a decent approximation. Both bigger phantoms are recovered quite successfully but they stay connected. The smallest phantom is not identified properly. Only certain allocation of its mass is identified along the proximal boundary. Even for 1% noise the TSCM method provided a decent approximation (figure 2(d)). The method seems to be rather stable with respect to noise. We recall, that the standard LSM is very sensitive when only boundary measurements are available, e.g. in [41, figure 7] only a noise level of 0.01% is considered in a case of a complicated phantom for the problem of EIT.

We next perform numerical experiments that use explicit dependence of MIT model on the frequency $\omega$. The results are presented in figure 3 for the phantom identical to the previous single-frequency experiment in figure 2. We consider the 4-frequency case $N(\omega) = 4$ and four levels of noise: 1%, 5%, 10% and 20%. The blue line is again the exact shape and the red line is its TSCM-identification. As expected, we got more accurate recovery of $\sigma_{PC}$. For the noise levels up to 10%, all the components of the phantom are quite accurately identified, accuracy gradually decreasing. Even for noise level of 20%, the identification is surprisingly accurate and all the components are identified; however, two bigger components stay connected by a bridge. This experiment confirms our conjecture that the method is very stable with respect to the non-systematic noise.

Noise causes non-convexity of the fidelity term $\mathcal{F}$ with respect to data regardless the properties of the forward operator $F$. Provided the data contain sufficient information to identify the phantom, the TSCM is able to eliminate this type of non-convexity. We are convinced
the reason lies within the nature of the method. The TSCM is essentially a convexification approach.

The convergences of the fidelity term $F(\sigma)$ and of the relative error between the computed conductivity $\sigma_{TSCM}$ and the exact conductivity $\sigma_{exact}$

$$e(\sigma) = \frac{\|\sigma_{TSCM} - \sigma_{exact}\|_{L^2(\Omega)}}{\|\sigma_{exact}\|_{L^2(\Omega)}}$$

with respect to the total number of iterations $n$ of algorithm 1 are depicted in figure 4(a). These graphs correspond to the experiment of figure 3(a). The distribution of the number of iterations for different $\lambda$-steps is depicted in figure 4(b). In general, the first iteration of the TSCM for $\lambda = 0$ is the most time consuming, which is natural, because it is nothing else than the minimization of $T^H_\alpha$ in the space $L^2(\Omega)$. It provides the information about ‘the optimal topology’ for $\sigma_{PC}$. Once this good initial guess is found, the continuation method rather quickly transforms this function to the desired piecewise-constant conductivity $\sigma_{PC}$.

Next, we consider a more complicated phantom with its two components touching and one of them being a torus. We again consider four excitation frequencies $N(\omega) = 4$. The results for two levels of noise, 1% and 10%, are depicted in figure 5. Again, we obtained a decent reconstruction even for 10% noise. Except the outside boundary, the hole of the torus

$\textbf{Figure 2.}$ Comparison between the standard LSM and TSCM. (a) LSM, no noise, initial $\phi_0$; (b) LSM, no noise, final $\phi_{37}$; (c) TSCM, no noise, final $\phi_{228}$; (d) TSCM, 1% noise, final $\phi_{138}$. 


is also well identified. The less resolved regions are those where the components are touching and the center of the domain.

Last, we examine the behavior of the TSCM regarding $\Delta \lambda$, i.e. regarding the number of $\lambda$-iterations $N(\lambda)$. We take the noise level of 1% and $N(\omega) = 2$. In figure 6(a), the total
number of iterations $n$ of algorithm 1 and in figure 6(b) the corresponding relative error of the conductivity $e(\sigma)$ are plotted against $\ln(N(\lambda))$. We see that $n$ shows a tendency to grow and $e(\sigma)$ a tendency to decrease. The results are obtained from a single-problem sample for each $N(\lambda)$. In figure 7, two particular examples are presented for $N(\lambda) = 2$ and for $N(\lambda) = 4$. 

Figure 5. TSCM: experiment 2; $N(\omega) = 4$. (a) 1% noise, final $\phi_{533}$; (b) 10% noise, final $\phi_{202}$.

Figure 6. TSCM: dependence on $\Delta \lambda$; $\rho = 1$%; $N(\omega) = 2$. (a) Number of iterations. (b) Relative error of sigma.

Figure 7. TSCM: dependence on $\Delta \lambda$; $\rho = 1$%; $N(\omega) = 2$. (a) $N(\lambda) = 2$, final $\phi_{175}$; (b) $N(\lambda) = 4$, final $\phi_{247}$.
We see that to correctly identify the shape and particularly its topology, it is necessary to consider at least $N(\lambda) = 4$. The continuation method has to be allowed to perform a sufficient number of steps to shift the information from $\sigma_L$ to $\sigma_{PC}$, i.e. the process has to be sufficiently continuous.

5. Conclusions

In this paper, we have presented a continuation approach for Tikhonov regularization and employed it to perform shape identification without any initial knowledge of topology. We have successfully applied the resulting TSCM to a MIT problem.

The TSCM does not explicitly perform any singular perturbations of the geometry (section 3.1). As a consequence, the usual difficulties in coupling the local and global (the shape and topology) sensitivities vanish and the method appears to be a very promising candidate for an ultimate framework unifying both sensitivity concepts. To establish such a claim more rigorously, it is necessary to provide a deeper analysis of the continuation approach with respect to the homotopy parameter $\lambda$, which is a possible future work. Any result in this direction will be dependent on a particular choice of the functional spaces $W$ and $U$ and their properties. Our understanding of the underlying concepts suggests that for TSCM-specific choice of the functional spaces such an analysis is attainable.

In this paper, we have provided more or less standard results on well-posedness, stability and convergence of the framework. Under a strong condition of uniqueness, we have provided a local convergence result of the continuation approach (theorem 2.1).

The numerical results of the TSCM for multiple-frequency MIT show that the method is, at least in certain settings, a globally convergent one. They demonstrate its decent accuracy and above all excellent stability of its reconstruction with respect to noise. It suggest that a generalization to multiple-valued piecewise-constant parameters scenario is feasible and should be fairly straightforward. As already known for MIT, simultaneous reconstruction of both conductivity and permittivity is possible. Altogether, the MIT with the TSCM as a solver could be used as a diagnostic method.

The regularization parameter $\alpha$ could be considered as a function of $\lambda$ as well. This should lead to $\lambda$-adaptive parameter choice rules and consequently a more efficient implementation of the TSCM algorithm. From the numerical point of view also conjugate gradient, quasi-Newton or Gauss–Newton algorithm extensions are possible.

Finally, let us quote from [42], where a penalty method is used to solve PCPIPs: ‘From our numerical experiences, we find that it is better to neglect the regularization term at the beginning stage of the iteration. At this stage, we should let the output-least-squares term to drag $\phi$ into the right direction without thinking about the regularity of $q$. In the context of continuation it is easy to explain this observation from [42]. The minimization without the total variation regularization term essentially behaves as Landweber type of regularization method, where the number of iterations plays the role of regularization [2], and the method converges to the least-squares solution in $L^\infty$-sense. The gradually increasing regularization parameter in the front of the total variation term functions as the continuation parameter $\lambda$. The same insight explains the global convergence of augmented Lagrangian methods [26].

Acknowledgments

VM would like to acknowledge the support of the BOF doctor-assistant research mandate 01P09209T of Ghent University, Ghent, Belgium. VV was supported by the BOF grant

$^8$ $\phi$—piecewise-constant level set, $q$—function coefficient to be recovered.
References

[1] Hofmann B, Kaltenbacher B, Pöschl C and Scherzer O 2007 A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators Inverse Problems 23 987–1010
[2] Engl H W, Hanke M and Neubauer A 1996 Regularization of Inverse Problems (Mathematics and its Applications vol 375) (Dordrecht: Kluwer)
[3] Morozov V A 1984 Methods for Solving Incorrectly Posed Problems (Berlin: Springer)
[4] Allgower E L and Georg K 2003 Introduction to Numerical Continuation Methods (Philadelphia, PA: SIAM)
[5] Zeidler E and Boron L F 1989 Nonlinear Functional Analysis and Its Applications: Part 2 B: Nonlinear Monotone Operators (Nonlinear Functional Analysis and Its Applications) (Berlin: Springer)
[6] Borcea L 2002 Electrical impedance tomography Inverse Problems 18 R99
[7] Acar R and Vogel C R 1994 Analysis of bounded variation penalty methods for ill-posed problems Inverse Problems 10 1217–29
[8] Pironneau O 1984 Optimal Shape Design for Elliptic Systems (Berlin: Springer) ISBN: 0-387-2069-6
[9] Sokolowski J and Zolésio J P 1992 Introduction to Shape Optimization. Shape Sensitivity Analysis vol 16 (Berlin: Springer)
[10] Santosa F 1996 A level-set approach for inverse problems involving obstacles ESAIM Contrôle Optim. Calc. Var. 1 17–33
[11] Fang W and Ito K 2003 Identification of contact regions in semiconductor transistors by level-set methods J. Comput. Appl. Math. 159 399–410
[12] Chan T, Esedoglu S, Park F and Yip A 2005 Total variation image restoration: overview and recent developments The Handbook of Mathematical Models in Computer Vision (Berlin: Springer) pp 17–32
[13] Bendste M P and Kikuchi N 1988 Generating optimal topologies in structural design using a homogenization method Comput. Methods Appl. Mech. Eng. 71 197–224
[14] Allaire G 2002 Shape Optimization by the Homogenization Method (Applied Mathematical Sciences vol 146) (Berlin: Springer)
[15] Eschenauer H A, Kobelev V V and Schumacher A 1994 Bubble method for topology and shape optimization of structures Struct. Multidiscip. Optim. 8 42–51
[16] Schumacher A 1996 Topologieoptimierung von Bauteilstrukturen unter Verwendung von Lochpositionierkriterien PhD Thesis Siegen University, Siegen, Germany
[17] Sokolowski J and Zochowski A 1997 On topological derivative in shape optimisation Technical Report 3170 INRIA-Lorraine
[18] Sokolowski J and Zochowski A 1999 On the topological derivative in shape optimization SIAM J. Control Optim. 37 1251–72
[19] Sokolowski J and Zochowski A 1999 Topological derivatives for elliptic problems Inverse Problems 15 123–34
[20] Burger M, Hackl B and Ring W 2004 Incorporating topological derivatives into level set methods J. Comput. Phys. 194 544–62
[21] He L, Kao C-Y and Osher S 2007 Incorporating topological derivatives into shape derivatives based level set methods J. Comput. Phys. 225 891–909
[22] Allaire G, Jouve F and Toader A-M 2004 Structural optimization using sensitivity analysis and a level-set method J. Comput. Phys. 194 363–93
[23] Allaire G, de Gourary F, Jouve F and Toader A-M 2005 Structural optimization using topological and shape sensitivity via a level set method Control Cybern. 34 59–80
[24] Nielsen L K, Tai X-C, Aanonsen Si I and Espedal M 2007 A binary level set model for elliptic inverse problems with discontinuous coefficients Int. J. Numer. Anal. Modeling 4 74–99
[25] Zhu S, Wu Q and Liu C 2011 Shape and topology optimization for elliptic boundary value problems using a piecewise constant level set method Appl. Numer. Math. 61 752–67
[26] De Cezaro A, Leitao A and Tai X-C 2013 On piecewise constant level-set (PCLS) methods for the identification of discontinuous parameters in ill-posed problems Inverse Problems 29 015003
[27] Hintermüller M and Laurain A 2008 Electrical impedance tomography: from topology to shape Control Cybern. 37 913–33
[28] Fulmanski P, Laurain A, Scheid J-F and Sokolowski J 2008 Level set method with topological derivatives in shape optimization Int. J. Comput. Math. 85 1491–514
[29] Frühauf F, Scherzer O and Leitão A 2005 Analysis of regularization methods for the solution of ill-posed problems involving discontinuous operators SIAM J. Numer. Anal. 43 767–86
[30] Griffiths H 2001 Magnetic induction tomography Meas. Sci. Technol. 12 1126–31
[31] Soleimani M 2008 Computational aspects of low frequency electrical and electromagnetic tomography: a review study Int. J. Numer. Anal. Modeling 5 407–40
[32] Cheney M, Isaacson D and Newell J C 1999 Electrical impedance tomography SIAM Rev. 41 85–101
[33] Brunner P, Merwal R, Missner A, Rosell J, Hollaus K and Scharfetter H 2006 Reconstruction of the shape of conductivity spectra using differential multi-frequency magnetic induction tomography Physiol. Meas. 27 237–48
[34] Zolgharni M, Ledger P D and Griffiths H 2009 Forward modelling of magnetic induction tomography: a sensitivity study for detecting haemorrhagic cerebral stroke Med. Biol. Eng. Comput. 47 1301–13
[35] Korjenevsky A, Cherepin V and Sapetisky S 2000 Magnetic induction tomography: experimental realization Physiol. Meas. 21 89–94
[36] Kufner A, John O and Fučík S 1977 Function spaces Monographs and Textbooks on Mechanics of Solids and Fluids Mechanics: Analysis (Leyden, MA: Noordhoff International Publishing)
[37] Osher S and Sethian J A 1988 Fronts propagating with curvature dependent speed: algorithms based on Hamilton–Jacobi formulations J. Comput. Phys. 79 12–49
[38] Cimrák I and Melicher V 2010 Determination of precession and dissipation parameters in micromagnetism J. Comput. Appl. Math. 234 2239–49
[39] Hecht F, Pironneau O, Morice J, Le Hyaric A and Ohitsuka K 2009 FreeFem++ Version 3.2 3rd edn (Paris: Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie) (www.freefem.org/ff++)
[40] Melicher V and Vrabel V 2013 TSCM implementation for magnetic induction tomography www.nam2.ugent.be/software/TSCM.tar.gz
[41] Chung E T, Chan T F and Tai X C 2005 Electrical impedance tomography using level set representation and total variational regularization J. Comput. Phys. 205 357–72
[42] Tai X-C and Li H 2007 A piecewise constant level set method for elliptic inverse problems ICSC ’05: Int. Conf. on Scientific Computing (Nanjing University, Nanjing, 4–8 June); Appl. Numer. Math. 57 686–96