Multiple level sets for piecewise constant surface reconstruction in highly ill-posed problems

K. van den Doel * U. M. Ascher † A. Leitão ‡

November 24, 2009

Abstract

This paper considers highly ill-posed surface recovery inverse problems, where the sought surface in 2D or 3D is piecewise constant with several possible level values. These levels may further be potentially unknown. Multiple level set functions are used when there are more than two such levels, and we extend the methods and theory of our previous works to handle such more complex situations. A rather efficient method is developed. Inverse potential problems in two and three space dimensions are solved numerically, demonstrating the method's capabilities for several both known and unknown level values.

Key words: Inverse problem, Level set, Inverse potential, Tikhonov functional, Dynamic regularization.

1 Introduction

Several important applications give rise to problems involving the recovery of distributed parameter functions from inverse problems with elliptic forward PDEs; see, e.g., [40, 28, 35, 12, 9, 33, 34, 23, 24, 15]. In most of these references the parameter function to be recovered, typically related to material conductivity, current, charge, or mass distribution, is assumed smooth. This enables a reasonably stable practical reconstruction of regularized solutions using, e.g., Tikhonov-type functionals [43, 20, 44]. But when discontinuities arise in the distributed parameter model, which commonly occurs in practical situations, the reconstruction problem becomes very difficult both theoretically and practically; see [29, 2] as well as the unsatisfactory results reported in Section 4 upon using the popular total variation (TV) regularization.

Fortunately, it is often reasonable to assume further that the solution may take on at each point only one of two values, thus yielding a shape recovery problem. Better results are then obtained in a much more stable fashion using algorithms that take this a priori information directly into account [31, 5, 7, 21, 36]. Level sets are a popular tool in this context, because they allow description of a discontinuous unknown function in terms of a differentiable one, thus enabling a more stable iterative process and the use of simpler and more versatile regularization

*Department of Computer Science, University of British Columbia, Vancouver, Canada (kvdoel@cs.ubc.ca), supported in part under NSERC Discovery Grant 84306.

†Department of Computer Science, University of British Columbia, Vancouver, Canada (ascher@cs.ubc.ca), supported in part under NSERC Discovery Grant 84306.

‡Department of Mathematics, Fed. University of St Catarina, Florianopolis, Brazil (acgleitao@gmail.com), supported in part under CNPq grants 306020/2006-8 and 474593/2007-0 and by the Alexander von Humboldt Foundation AvH.

functionals, and for their insensitivity to topological changes. Many authors have reported rather satisfactory practical experience using this approach; see, e.g., [36, 22, 6, 10, 21, 13, 14]. For reviews on the use of level set methods for inverse problems, see [18, 41]. Some of the algorithms described in these references can be rather inefficient, but this is not an inherent defect of level set methods [14].

We note that at present there is want of complete theory for practical level set methods in this context. A glimpse at the partial efforts reported in [21], where otherwise routine tasks, such as establishing existence results for corresponding Tikhonov regularizations, become rather delicate and mandate the addition of unlikely terms to the corresponding functionals, provides some insight into the issues involved. Of course, the practical solution of highly ill-posed problems with sparse data is never a completely routine task, but solution discontinuities and corresponding level sets describing them make the analysis even more complicated. In essence, whereas it is easy to devise a guaranteed descent method that improves a data fitting term at each iteration step (see [36, 6] and Section 3 below), the task of choosing a level set method that yields a reasonable shape solution in a stable way is more delicate. Nonetheless, the methods proposed in [13, 14] provide a highly efficient and often robust approach for the approximate solution of such problems.

Further, often situations arise where the sought model may take on at each point only one of *several* values. For instance, think of $l-1$ homogeneous bodies made of different materials all buried in a homogeneous sand that serves as background. One can use n level set functions, with n defined by

$$2^{n-1} < l \leq 2^n,$$

in order to describe a surface with l distinct constant values [41, 11]. If n gets large then the additional stability that the piecewise constant modeling assumption yields evaporates [1], but it is useful to consider cases with $n = 2$ or 3, say. For notational simplicity we assume until Section 4 that $l = 2^n$: if $l < 2^n$ then some of the 2^n values can be taken the same to arrive at l different values.

Several multiple level set approaches have been proposed in recent literature, especially in the area of image analysis. In [42] a piecewise constant level set approach (i.e., the level set function is assumed to be piecewise constant) using TV regularization is proposed to solve an EIT problem with a single interior measurement. In [41] the operator P defined in (2) below is introduced for the first time. A Tikhonov functional based on TV regularization is proposed and minimized using gradient type methods. In [8] a multiple level set approach based on TV regularization is proposed. Approximate solutions are obtained by minimizing a Tikhonov-TV type functional, which is defined on a set of characteristic functions. The results of [21] have been extended in [11] to the two level set case. It is worth noting that the multiple level set approach used in [7, 11] to represent the solution of the inverse problem is similar to the one considered here. However, the method proposed in this article differs from the ones introduced in these references, and is based on the efficient level set method proposed in [13, 14].

At this point, to be more specific, we introduce some notation. A given *forward operator*, $F(m)$, predicts the data for each *model* argument function $m = m(\mathbf{x})$, where m is defined over a given domain Ω in 2D or 3D. In the *inverse problem* a model m is sought such that $F(m)$ matches observed data b up to the noise level in the data measurements. The forward model is further given by

$$F(m) = Qu, \tag{1a}$$

$$u = G(m), \tag{1b}$$

where Q projects the field u to data locations (e.g., along the boundary $\partial\Omega$), and $G(m)$ is the inverse of an elliptic-type PDE system defined on the closure of Ω. Examples of G are given in Section 4 and elsewhere. Importantly, we assume in addition that there are 2^n values, $a_{i_1,\ldots,i_n}$, $i_k = 0, 1$, such that at each point $\mathbf{x} \in \Omega$, $m(\mathbf{x})$ takes on one of these values. Then we can describe the model m using n level set functions $\psi^1(\mathbf{x}), \ldots, \psi^n(\mathbf{x})$ as

$$m(\mathbf{x}) = \sum_{i_1,\ldots,i_n=0,1} a_{i_1,\ldots,i_n} \hat{H}_{i_1}(\psi^1(\mathbf{x})) \cdots \hat{H}_{i_n}(\psi^n(\mathbf{x})) \quad \text{at each } \mathbf{x}, \text{where} \tag{2a}$$

$$\hat{H}_i(s) = \begin{cases} H(s), & i = 0, \\ 1 - H(s), & i = 1, \end{cases} \tag{2b}$$

with H the Heaviside function. Let us write this as $m = P(\psi)$, where $\psi = (\psi^1, \ldots, \psi^n)$.

The so-called output least squares method would attempt to find m by fitting the data in the least squares sense, i.e., by attempting to bring the expression

$$\phi(\psi; 0) = \frac{1}{2}\|F(m) - b\|_2^2 \tag{3a}$$

to within the noise level. However, the Fréchet derivative (or Jacobian matrix on the discrete level), $J = F'(m)$, has a nontrivial null-space in our applications and a regularization is thus required.

A Tikhonov-type regularization then calls for the minimization with respect to ψ of

$$\phi(\psi; \alpha) = \frac{1}{2}\|F(P(\psi)) - b\|_2^2 + \alpha R(\psi)\,, \tag{3b}$$

where α is a parameter and we now consider the functional

$$R(\psi) = \sum_{i=1}^{n} \int_\Omega \frac{1}{2}|\boldsymbol{\nabla}\hat{\psi}^i|^2 + \frac{1}{2}\beta_1|\hat{\psi}^i|^2 + \beta_2|\boldsymbol{\nabla}H(\psi^i)|\,. \tag{3c}$$

Here $\hat{\psi} = \psi - \psi_0$, with ψ_0 a given reference function. This function as well as the parameters $\beta_j \geq 0$ are to be determined in the sequel. Notice that the last integral in (3c) corresponds to $|H(\psi^i)|_{BV}$, the BV-seminorm of $H(\psi^i)$, which is formally defined by

$$|v|_{BV(\Omega)} = \sup\{\int_\Omega v \operatorname{div} w \, dx;\ w \in C_c^1(\Omega)\}\,,$$

where $C_c^1(\Omega)$ is the set of C^1 functions with compact support in Ω. The various terms in (3c) are all familiar, having appeared and been justified in the literature before. However, in the sequel we will find occasions for dropping several of them (i.e., setting some of β_1, β_2 and ψ_0 to 0).

In Section 2 we recall key points from the theory developed in [11, 21] for such a regularization. For this we must choose $\beta_i > 0$, $i = 1, 2$, and we set $\beta_1 = 1$. We further extend this theory for the case of unknown contrasts, i.e., where some of the values $a_{i_1,\ldots,i_n}$ (for a given n) are unknown.

Of course, to practically find an approximate minimizer for (3) one must discretize the equations, and we consider a finite volume or finite element discretization of the PDE problem that defines the forward operator F. This involves a discrete mesh implying a certain resolution, or mesh width h (even for nonuniform meshes), so we also consider m as a mesh function, cf.

[25]. We further smooth the projector P, replacing the Heaviside function H in a standard way by a differentiable function H_h whose derivative is proportional to h^{-1} over a segment of length $O(h)$ [21, 13]. The theory of Section 2 assures us at least of some regular passage to the limit $h \to 0$. But for a finite, positive resolution h additional approximations can be usefully considered. In Section 3 we extend the fast methods of [13, 14] to the cases of multiple level sets with known and unknown levels $a_{i_1,\ldots,i_n}$, and we provide a framework that includes also the method of [21, 11]. Our method does not employ a Tikhonov regularization, though it is closely related to one. As was shown in [13] for the electrical impedance tomography (EIT) problem in two dimensions, the Tikhonov-type regularization method often performs poorly in practice, requiring in particular a delicate adjustment of the various parameters. On the other hand, the dynamic regularization method, further extended for large scale problems in [14], allows us to drop the Tikhonov term (i.e., set $\alpha = 0$ in (3b)), and rely on a finite number of outer and inner iterations in our scheme to regularize the problem. The only remnant of (3c) is then a preconditioner, or smoother, which is essential for its stability.

This paper can be seen as gradually developing a more efficient and more stable method for a family of difficult inverse problems. Curiously, the more efficient and more stable the multiple level set method gets, the further it seems to be from our theoretical base. In Section 4 we demonstrate the efficacy of our method for an inverse potential problem [28, 21, 13, 11, 29] in 2D and 3D. Conclusions are offered in Section 5.

2 Tikhonov-type functionals in continuous spaces

In this section we present a convergence analysis for the Tikhonov-type approach introduced in Section 1.

2.1 Results for known level values

Here we consider the case where the level values $a_{i_1,\ldots,i_n}$ in (2a) are known. Under this assumption, the operator P defined by (2) does not depend on the values $a_{i_1,\ldots,i_n}$, i.e., $P = P(\psi)$.

Let ϕ be the Tikhonov functional in (3b), where R is the H^1–TV penalization term defined in (3c) and $\beta_i > 0$ $i = 1, 2$, are chosen as above. Since the Heaviside function H used in (2) is discontinuous, the operator P is discontinuous as well. Thus, one cannot prove that the Tikhonov functional in (3b) attains a minimizer on the set in $(H^1(\Omega))^n$ of level set functions ψ (see [20]).

In order to guarantee existence of a minimizer of the Tikhonov functional ϕ, we introduced the concept of generalized minimizers in [11]. According to this concept, $\phi(\psi; \alpha)$ is no longer considered as a functional on $(H^1(\Omega))^n$, but as a functional defined on the w-closure of the graph of P, contained in $(H^1(\Omega))^n \times (L^\infty(\Omega))^n$. This requires the introduction of a smooth approximation to the Heaviside function H used in (2b). For $\varepsilon > 0$, let H_ε be the function defined by

$$H_\varepsilon(s) = \frac{1}{2}(\tanh(s/\varepsilon) + 1)\,. \tag{4}$$

Both operators H and H_ε are considered as mappings from H^1 into L^∞. More precisely, we have:

Definition 1. *A* generalized minimizer *of the Tikhonov functional* $\phi(\psi;\alpha)$ *in* (3b) *is defined as a minimizer of the functional*

$$\tilde{\phi}(z,\psi;\alpha) = \frac{1}{2}\|F(P(\psi)) - b\|_2^2 + \alpha\mathcal{R}(z,\psi)\,, \tag{5}$$

on the set of admissible parameters

$$\begin{aligned} Ad \;=\; & \{(z,\psi) = (z^1,\dots,z^n,\psi^1,\dots,\psi^n) \in (L^\infty(\Omega))^n \times (H^1(\Omega))^n\,; \\ & \exists\ \{\psi_k^1\},\dots,\{\psi_k^n\} \in H^1 \text{ and } \{\varepsilon_k\} \in \mathbb{R}^+ \text{ with } \lim_{k\to\infty}\varepsilon_k = 0,\ \text{s.t.} \\ & \lim_{k\to\infty}\|\psi_k^j - \psi^j\|_{L^2} = 0 \text{ and } \lim_{k\to\infty}\|H_{\varepsilon_k}(\psi_k^j) - z^j\|_{L^1} = 0 \text{ for } j = 1,2,\dots,n\}\,, \end{aligned}$$

where the penalization functional $\mathcal{R}$ *is defined by*

$$\mathcal{R}(z,\psi) = \inf_{\{\psi_k^j\},\{\varepsilon_k\}} \liminf_{k\to\infty} \sum_{j=1}^{n} \left\{\|\psi_k^j - \psi_0^j\|_{H^1}^2 + |H_{\varepsilon_k}(\psi_k)|_{BV}\right\}. \tag{6}$$

Notice that the infimum in (6) is taken over all sequences $\{\psi_k^1\},\dots,\{\psi_k^n\}$ and $\{\varepsilon_k\}$, characterizing (z,ψ) as an element of the set of admissible parameters Ad. In the above definition the functional $\mathcal{R} : Ad \to \mathbb{R}^+$ corresponds to a relaxation of the penalization term R in (3c).

In order to prove existence of a minimizer of $\tilde{\phi}(z,\phi;\alpha)$ in Ad, two properties of $\mathcal{R}$ are required, namely, coerciveness and weak lower semi-continuity. For the proof of these properties the assumption that F is a continuous operator w.r.t. the L^1–topology is crucial [11].

In the next lemma we recall relevant properties of the generalized minimizers of ϕ, as well as the above mentioned regularity properties of the functional $\mathcal{R}$ in (6):

Lemma 1 (Properties of generalized minimizers and relaxed penalization functional)**.**

(a) The set of admissible parameters Ad is sequentially closed with respect to the $(L^1(\Omega))^n \times (L^2(\Omega))^n$ *topology.*

(b) For each $(z,\psi) = (z^1,\dots,z^n,\psi^1,\dots,\psi^n) \in Ad$ *we have* $\sum_{j=1}^n \left(|z^j|_{BV} + \|\psi^j - \psi_0^j\|_{H^1(\Omega)}^2\right) \le \mathcal{R}(z,\psi)$*.*

(c) If $\{(z_k,\psi_k) = (z_k^1,\dots,z_k^n,\psi_k^1,\dots,\psi_k^n)\}_{k\in\mathbb{N}}$ *is a sequence in Ad with* $z_k^j \to z^j$ *in* $L^1(\Omega)$ *and* $\psi_k^j \rightharpoonup \psi^j$ *in* $H^1(\Omega)$ *for some* $(z^1,\dots,z^n,\psi^1,\dots,\psi^n) = (z,\psi) \in Ad$*, then* $\mathcal{R}(z,\psi) \le \liminf_{k\to\infty}\mathcal{R}(z_k,\psi_k)$*.*

Proof. These results are straightforward generalizations of proofs given in [11], Section 3. □

Once the regularity properties of $\mathcal{R}$ in Lemma 1 are verified, the classical analysis of Tikhonov-type regularization methods [20, Chapter 10] applies to the functional ϕ, as we see next:

Theorem 1 (Convergence analysis)**.**

(a) **(Well-Posedness)** *The functional* $\phi(\psi;\alpha)$ *in* (3b) *attains generalized minimizers on Ad.*

(b) **(Stability)** *Given* $\alpha > 0$*, the problem of minimizing* $\phi(\psi;\alpha)$ *over* ψ *is stable in the sense of continuous dependence of the solutions on the data b.*

In the sequel, for every $\alpha > 0$, denote by (z_α, ψ_α) a generalized minimizer of $\phi(\psi;\alpha)$ on Ad.

(c) **(Convergence for exact data)** *Suppose that there exists noiseless data b^0 such that $F(m^*) = b^0$ for some "ground truth", or "true solution", m^*. Let $b = b^0$ and $\beta_i > 0$. Then, for every sequence of positive numbers $\{\alpha_k\}_{k\in\mathbb{N}}$ converging to zero there exists a subsequence, denoted again by $\{\alpha_k\}_{l\in\mathbb{N}}$, such that $(z_{\alpha_k}, \psi_{\alpha_k})$ is strongly convergent in $(L^1(\Omega))^n \times (L^2(\Omega))^n$. Moreover, the limit (z,ψ) is an element of Ad satisfying $F(P(z)) = b^0$.*

(d) **(Convergence for noisy data)** *Let $\alpha = \alpha(\delta)$ be a function satisfying $\alpha(\delta) \overset{\delta\to 0}{\longrightarrow} 0$ and $\delta^2\alpha(\delta)^{-1} \overset{\delta\to 0}{\longrightarrow} 0$. Moreover, let $\{\delta_k\}_{k\in\mathbb{N}}$ be a sequence of positive numbers converging to zero, and let $b^{\delta_k} \in Y$ be corresponding noisy data satisfying $\|b^{\delta_k} - b^0\| \le \delta_k$. Then there exists a subsequence, denoted again by $\{\delta_k\}$, and a sequence $\{\alpha_k = \alpha(\delta_k)\}$ such that $(z_{\alpha_k}, \psi_{\alpha_k})$ converges in $(L^1(\Omega))^2 \times (L^2(\Omega))^2$ to some $(z,\psi) = (z^1, \dots, z^n, \psi^1, \dots, \psi^n)$ that satisfies $F(m) = b^0$ for $m = m(z) = P(z^1, \dots, z^n)$.*

Proof. These results are straightforward generalizations of proofs given in [11], Section 4. □

In [11] the generalized minimizers of $\phi(\psi;\alpha)$ are approximated by minimizers of smoothed functionals $\phi_\varepsilon(\psi;\alpha)$, $\varepsilon > 0$. Then the first order optimality conditions for a minimizer of these smoothed functionals are used as motivation for the derivation of a level set type method.

2.2 Results for unknown level values

Here we consider the case where not all level values $a_{i_1,\dots,i_n}$ in (2a) are known. Assume that $a_{i_1,\dots,i_n}$ are determined by $l_u \le l$ unknown model parameters. Let us parameterize the unknown levels in terms of the variables $m_1, \dots, m_{l_u}$ which are to be recovered, i.e., $a_{i_1,\dots,i_n} = a_{i_1,\dots,i_n}(m_1, \dots, m_{l_u})$ with prescribed (model) functions $a_{i_1,\dots,i_n}$. We now proceed as before, treating the additional degrees of freedom $m_1, \dots, m_{l_u}$ on par with the level set functions. Notice that the operator P defined by (2) can be written as $P = P(\psi, m_1, \dots, m_{l_u})$, or in shorthand notation, $P = P(\psi, m_u)$.

Analogous to (3), we define the Tikhonov functional

$$\Phi(\psi, m_u; \alpha) = \frac{1}{2}\|F(P(\psi, m_u)) - b\|_2^2 + \alpha S(\psi, m_u)\,, \tag{7}$$

where $S(\psi, m_u) = R(\psi) + \frac{\beta_3}{2}\sum_{i=1}^{l_u} |m_i - m_{i,0}|^2$, and $m_{i,0}$ are given initial guesses for the unknown values m_i, $i = 1, \dots, l_u$.

Notice that the functional Φ has l_u more degrees of freedom than the functional ϕ in (3b).

Definition 2. *A* generalized minimizer *of the Tikhonov functional $\Phi(\psi, m_u; \alpha)$ is defined as a minimizer of the functional*

$$\tilde{\Phi}(z, \psi, m_u; \alpha) = \frac{1}{2}\|F(P(\psi, m_u)) - b\|_2^2 + \alpha\mathcal{S}(z, \psi, m_u)\,, \tag{8}$$

on the set of admissible parameters $\widetilde{Ad} = Ad \times \mathbb{R}^{l_u}$, where the penalization term $\mathcal{S}$ is defined by $\mathcal{S}(z, \psi, m_u) = \mathcal{R}(z,\psi) + \frac{\beta_3}{2}\sum_{i=1}^{l_u} |m_i - m_{i,0}|^2$.

Analogous to Lemma 1 we have the following:

Corollary 1. *(a) The set $\widetilde{Ad}$ is sequentially closed with respect to the $(L^1(\Omega))^n \times (L^2(\Omega))^n \times \mathbb{R}^{l_u}$ topology.*

(b) The functional $\mathcal{S}(z, \psi, m_u)$ is coercive and weak lower semi-continuous on the set of admissible parameters $\widetilde{Ad}$.

Sketch of the proof. The proof of item (a) follows from Lemma 1(a) and the fact that $\mathbb{R}^{l_u}$ is a metric space. To prove item (b), notice that up to a multiplicative constant the functional $\mathcal{S}(z, \psi, m_u)$ is obtained by adding the term $\sum_{i=1}^{l_u} |m_i - m_{i,0}|^2$ to $\mathcal{R}(z, \psi)$. Then, using Lemma 1(b) and (c), together with the fact that the $\mathbb{R}^{l_u}$-norm is coercive and weak lower semi-continuous, the assertion in item (b) follows. □

Once the regularity properties of $\mathcal{S}$ in Corollary 1(b) are established, the classical analysis of Tikhonov type regularization methods applies to the functional $\Phi(\psi, m_u; \alpha)$:

Theorem 2 (Convergence Analysis)**.**

(a) The Tikhonov functional $\Phi(\psi, m_u; \alpha)$ in (7) attains generalized minimizers on $\widetilde{Ad}$.

(b) Given $\alpha > 0$, the problem of minimizing $\Phi(\psi, m_u; \alpha)$ is stable in the sense of continuous dependence of the solutions on the data b.

For every $\alpha > 0$, denote by $(z_\alpha, \psi_\alpha, m_{u,\alpha})$ a generalized minimizer of $\Phi(\psi, m_u; \alpha)$ on $\widetilde{Ad}$.

(c) Suppose that there exists noiseless data b^0 such that $F(m^*) = b^0$ for some ground truth m^*. Let $b = b^0$ and $\beta_i > 0$. Then, for every sequence of positive numbers $\{\alpha_k\}_{k\in\mathbb{N}}$ converging to zero there exists a subsequence, denoted again by $\{\alpha_k\}_{l\in\mathbb{N}}$, such that $(z_{\alpha_k}, \psi_{\alpha_k}, m_{u,\alpha_k})$ is strongly convergent in $(L^1(\Omega))^n \times (L^2(\Omega))^n \times \mathbb{R}^{l_u}$. Moreover, the limit (z, ψ, m_u) is an element of $\widetilde{Ad}$ satisfying $F(P(z, m_u)) = b^0$.

(d) Let $\alpha = \alpha(\delta)$ be a function satisfying $\alpha(\delta) \xrightarrow{\delta\to 0} 0$ and $\delta^2\alpha(\delta)^{-1} \xrightarrow{\delta\to 0} 0$. Moreover, let $\{\delta_k\}_{k\in\mathbb{N}}$ be a sequence of positive numbers converging to zero and $b^{\delta_k} \in Y$ be corresponding noisy data satisfying $\|b^{\delta_k} - b^0\| \le \delta_k$. Then, there exist a subsequence, denoted again by $\{\delta_k\}$, and a sequence $\{\alpha_k = \alpha(\delta_k)\}$ such that $(z_{\alpha_k}, \psi_{\alpha_k}, m_{u,\alpha_k})$ converges in $(L^1(\Omega))^2 \times (L^2(\Omega))^2 \times \mathbb{R}^{l_u}$ to some $(z, \psi, m_u) = (z^1, \dots, z^n, \psi^1, \dots, \psi^n, m_1, \dots, m_{l_u})$ such that $F(m) = b^0$ for $m = m(z, m_u) = P(z^1, \dots, z^n, m_1, \dots, m_{l_u})$.

Sketch of the proof. Since $(0, 0, 0) \in \widetilde{Ad}$, the set of admissible parameters is not empty. Then, given a minimizing sequence $\{(z_k, \psi_k, m_{u,k})\}$ for $\tilde{\Phi}$, it follows from the boundedness of the sequence $\{\tilde{\Phi}(z_k, \psi_k, m_{u,k}; \alpha)\}$ that $\{(z_k, \psi_k, m_{u,k})\}$ is bounded in $BV(\Omega)^n \times H^1(\Omega)^n \times \mathbb{R}^{l_u}$. Thus, the Sobolev compact embedding theorem together with the Bolzano-Weierstrass theorem guarantee the existence of a subsequence converging to some $(z, \psi, m_u) \in L^1(\Omega)^n \times H^1(\Omega)^n \times \mathbb{R}^{l_u}$. Now, Corollary 1(a) guarantees that $(z, \psi, m_u) \in \widetilde{Ad}$. Arguing aided by Corollary 1(b) and the weak continuity of F and P, we deduce that (z, ψ, m_u) is a generalized minimizer of Φ, proving item (a).

The proof of assertion (b) follows the lines of the proof of [20, Theorem 10.2] (notice that Corollary 1(b) is essential to this proof).

If the data is exact, then the sequence of real numbers $\{\tilde{\Phi}(z_{\alpha_k}, \psi_{\alpha_k}, m_{u,\alpha_k}; \alpha_k)\}$ is bounded. Thus, the sequences $\{z_{\alpha_k}\}$, $\{\psi_{\alpha_k}\}$ and $\{m_{u,\alpha_k}\}$ are bounded in $BV(\Omega)^n$, $H^1(\Omega)^n$ and $\mathbb{R}^{l_u}$, respectively. Arguing as in (a), one can extract subsequences (denoted by the same indexes)

converging to some limit (z, ψ, m_u) in the topology of $(L^1(\Omega))^n \times (L^2(\Omega))^n \times \mathbb{R}^{l_u}$. Moreover, $(z, \psi, m_u) \in \widetilde{Ad}$. Now, arguing aided by the continuity of F and the lower semi-continuity of $\mathcal{S}$, one concludes that $F(P(z, m_u)) = b^0$, proving (c).
The existence of $(z^*, \psi^*, m_u^*) \in \widetilde{Ad}$ with $F(P(z^*, m_u^*)) = b^0$ follows from the assumptions in item (d). Thus, $\tilde{\Phi}(z_{\alpha_k}, \psi_{\alpha_k}, m_{u,\alpha_k}; \alpha_k) \leq \delta_k^2 + \alpha_k \mathcal{S}(z_{\alpha_k}, \psi_{\alpha_k}, m_{u,\alpha_k})$. Taking the limit $k \to \infty$, it follows from the assumptions that $\|F(P(z_{\alpha_k}, m_{u,\alpha_k})) - b^{\delta_k}\| \to 0$ as $k \to \infty$. Therefore, $\lim_{k\to\infty} F(P(z_{\alpha_k}, m_{u,\alpha_k})) = b^0$. On the other hand, it follows from the definition of $\tilde{\Phi}$ that $\mathcal{S}(z_{\alpha_k}, \psi_{\alpha_k}, m_{u,\alpha_k}) \leq \delta_k^2 \alpha_k^{-1} + \mathcal{S}(z^*, \psi^*, m_u^*)$. Thus, we conclude from the assumptions that $\limsup_{k\to\infty} \mathcal{S}(z_{\alpha_k}, \psi_{\alpha_k}, m_{u,\alpha_k}) \leq \mathcal{S}(z^*, \psi^*, m_u^*)$. The proof of assertion (d) follows similarly to the proof of assertion (c). □

3 Numerical algorithms

As mentioned in Section 1, our first step for designing numerical algorithms is to discretize the PDE problem that gets inverted in (1b) using a finite volume or finite element method. This implies a mesh with finite resolution h and we correspondingly discretize also m, ψ, and H. There is also a finite number of data values in b. Abusing notation in an obvious way and reshaping mesh functions into vectors, we can write the discretized forward problem again as (1), where m and u are now vectors and Q becomes a matrix with typically many more columns than rows. Also define the Jacobian matrices

$$J = \frac{\partial F}{\partial m} \quad \text{and} \quad \hat{J} = \frac{\partial F}{\partial \psi}\,, \tag{9}$$

see, e.g., [24]. Finally, defining

$$P' = \frac{\partial P(\psi)}{\partial \psi} \quad \text{and} \quad P_i' = \frac{\partial P(\psi)}{\partial \psi^i}$$

we obtain

$$\hat{J} = JP'\,.$$

Note that P' is block diagonal, since P_i' is diagonal.

3.1 Minimizing discrete Tikhonov-type functionals

Let us assume at first that the levels $a_{i_1,\ldots,i_n}$ are known. This assumption is removed towards the end of the present section. The necessary conditions for a minimum of $\phi(\psi; \alpha)$ in (3b) are

$$\boldsymbol{\nabla}_\psi \phi(\psi; \alpha) \equiv \hat{J}^T(F(P(\psi)) - b) + \alpha R'(\psi) = 0. \tag{10}$$

Further, on the discrete level, for all the methods described below the perimeter of the reconstructed shapes is controlled also by other means (see the first paragraph of Section 3.2), so we conveniently set $\beta_2 = 0$ in (3c). Thus, consider

$$R'(\psi) = \begin{pmatrix} (\beta_1 I - \Delta)\hat{\psi}^1 \\ \vdots \\ (\beta_1 I - \Delta)\hat{\psi}^n \end{pmatrix}\,, \tag{11}$$

with the differential terms appropriately discretized.

A damped Gauss-Newton method for solving this minimization problem leads to the following update rule for ψ:

$$K(\psi;\alpha)\ \delta\psi \;=\; -\boldsymbol{\nabla}_\psi\phi(\psi;\alpha)\,, \quad \delta\psi = \begin{pmatrix}\delta\psi^1\\ \vdots\\ \delta\psi^n\end{pmatrix}\,, \tag{12a}$$

$$\psi \;\leftarrow\; \psi + \tau\delta\psi, \tag{12b}$$

$$K(\psi;\alpha) \;=\; \hat{J}^T\hat{J} + \alpha \begin{pmatrix}\beta_1 I - \Delta & & 0\\ & \ddots & \\ 0 & & \beta_1 I - \Delta\end{pmatrix} = \hat{J}^T\hat{J} + \alpha\ \mathrm{diag}(\beta_1 I - \Delta)\,. \tag{12c}$$

The default value for the step size is $\tau = 1$, and it gets decreased if necessary by a standard weak line search [13] in order to obtain a sufficient decrease in $\phi(\psi;\alpha)$.

In [11], following [21], the authors ignore the expensive first term in $K(\psi;\alpha)$ and set $\beta_1 = 1$. They also update the reference function ψ_0 at the beginning of each iteration (or "time step"), setting it to the current ψ and holding it fixed for the duration of the current iteration. This yields an iterated Tikhonov method [38, 39]. We have tested this method on the problem described in Sections 4.1–4.2. It produces qualitatively similar results to those obtained by the method described next, thus validating both programs, and has the advantage of providing a smooth, stable approach to the solution. However, dropping the data fitting contribution in the Gauss-Newton matrix also slows down convergence significantly, and the methods described below are often faster by as much as two orders of magnitude.

3.2 Fast dynamic regularization

In [13, 14] we have proposed a dynamic regularization method where we set $\alpha = 0$ in the right hand side of (12a). Note that the resulting iterative scheme can no longer be considered as optimizing a Tikhonov functional. Instead, terminating the iteration after a finite number of steps yields regularization [13]. Further, we observed there that using $\beta_1 = 0$ or $\beta_1 = 1$ produced indistinguishable results, so we simply set $\beta_1 = 0$ here. Unlike the fast Tikhonov-type methods studied in [13], this method does not require $\beta_2 > 0$, i.e., we can set $\beta_2 = 0$. The role of the BV term is to penalize rough boundaries, to arrive at a smooth reconstruction, and in practice we have found that the discretized Laplacian Δ used in the dynamic regularization scheme already takes care of this smoothing. All of this has the combined effect of replacing the iteration (12a) by

$$\Big(\hat{J}^T\hat{J} - \alpha\ \mathrm{diag}(\Delta)\Big)\,\delta\psi = -\hat{J}^T(F(P(\psi)) - b)\,. \tag{13}$$

Note that since the matrices in (13) are nonsingular for $\alpha > 0$ (we ensure this in a standard way) the obtained direction $\delta\psi$ is a descent direction with respect to the output least squares function $\phi(\psi;0)$ defined in (3a). With a careful line search we obtain a monotonic decrease in $\phi(\psi;0)$ as the iteration proceeds, and we use the discrepancy principle directly on this function to stop the iteration. This process is therefore more direct than the corresponding one using continuation with Tikhonov functionals. Indeed, the method is best considered as directly attempting to minimize $\phi(\psi;0)$ until noise takes over, with the α-term in (13) stabilizing the matrix to keep it away from singularity.

Note that due to the smoothing role of Δ one cannot simply replace it by the identity, as the Levenberg-Marquardt method would have it. In fact, for Tikhonov-type regularization, even

the discretized Laplacian operator with natural boundary conditions does not always suffice to produce a desirable level set function. We refer to [13] for another regularization term, called R_{4n} there, which occasionally does a better job in this respect.

For larger problems, including almost anything in 3D, the explicit calculation and storage of the sensitivity matrix J quickly become prohibitive. A preconditioned conjugate gradient (PCG) inner iteration using the Laplacian as a preconditioner is thus employed instead. The preconditioner acts as a smoothing operator on the iterative solution, rather than as a means to speed up convergence. Indeed, without the preconditioner we can also obtain convergence in not too many more iterations, but the solution is undesirably rough, with many fragmented regions. We thus view the preconditioner as a smoothing operator.

A single inner PCG iteration requires only two evaluations of F for a matrix-vector multiplication in (13). Furthermore, instead of solving the system (13) to high accuracy using PCG we can apply only a few inner iterations, say 3 or 5, which has a regularizing effect, cf. [26], and this allows not only a very cheap overall iteration but also the reduction of α to (almost) 0. A highly efficient algorithm is thus obtained, as demonstrated in Section 4 and in [14].

Unknown levels

With a small variation of the above method we can also solve problems where not all values $a_{i_1,\ldots,i_n}$ are known. We consider the most general case where they are determined by l_u unknown model parameters. As in Section 2.2 these are parameterized in terms of the variables $m_1, \ldots, m_{l_u}$.

We proceed as before, treating the additional degrees of freedom $m_1, \ldots, m_{l_u}$ on par with the level set functions. This leads again to the iteration (13), where the vector $\delta\psi$ is replaced by $(\delta\psi, \delta m_1, \ldots \delta m_{l_u})^T$ and the modified Jacobian $\hat{J}$ is now defined as

$$\hat{J} = \frac{\partial F}{\partial(\psi, m_1, \ldots, m_{l_u})} ,$$

which can be written as

$$\hat{J} = J \left(P', \sum_{i_1,\ldots,i_n=0,1} \hat{H}_{i_1} \cdots \hat{H}_{i_n} \frac{\partial a_{i_1,\ldots,i_n}}{\partial(m_1, \ldots, m_{l_u})} \right) .$$

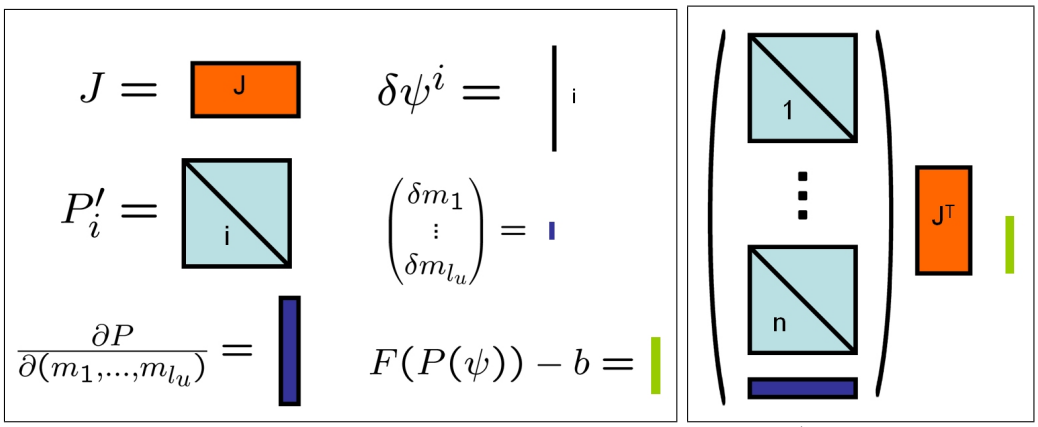
The matrix forms of these equations are clarified in Fig. 1.

4 Numerical experiments

In this section we consider several examples with $l_u < l$ specified and $n = \lceil \log_2(l) \rceil$.

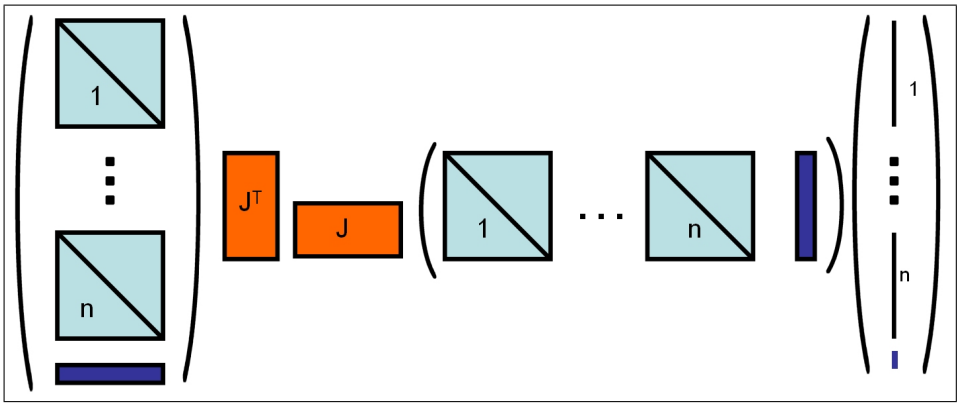
The cost of one iteration, or step, of our algorithm depends mainly on the cost of solving the PDEs involved. There are two types of such PDEs. The first is the system (1b) that gets solved for the forward problem. This is also the essential effort in carrying out one matrix-vector multiplication involving J_i or J_i^T [24, 14]. Let us denote this cost by C_F. The other PDE inversion involves the discrete Laplacian equipped with natural boundary conditions (and made nonsingular in a standard fashion), and we denote this cost by C_P. For the inverse potential problem discussed in Section 4.1, $C_F \approx C_P$, whereas for EIT or DC resistivity problems considered in [14] the forward solution cost C_F dominates C_P.

Using this, the cost of evaluating the right hand side in (13) is estimated by $(n+1)C_F$, where n is the number of level set functions. For the iteration of [11] the cost is therefore



(a) Pictorial representation of matrices and vectors

(b) $\hat{J}^T(F(P(\psi)) - b)$



(c) $\hat{J}^T\hat{J}(\delta\psi, \delta m_1, \ldots, \delta m_{l_u})^T$

Figure 1: The matrix structure of terms of (13) in a picture. For a grid with N points, boundary data on N_B points, l_u unknown values m_j and n level sets, the P_i' are $N \times N$ diagonal matrices, J is $N_B \times N$, and $\frac{\partial P}{\partial(m_1,\ldots,m_{l_u})}$ is $N \times l_u$.

$nC_P + (n+1)C_F$. This is not much more than the cost of a simple gradient descent iteration. The iteration in (13) can be more expensive, though. If a fixed number N_{in} of PCG inner iterations is used then the cost of one outer iteration is estimated by $n(2C_F + C_P)N_{in} + (n+1)C_F$.

4.1 Inverse potential problem

Our generic test problem in this paper is similar to the one considered in [11, 21, 13, 28]. The forward problem $F(m)$ consists of solving on a domain Ω, for a given source function $m(\mathbf{x})$, the Poisson problem

$$\begin{aligned} -\Delta u &= m, \quad \text{in } \Omega, \\ u &= 0, \quad \text{on } \partial\Omega, \end{aligned} \tag{14}$$

and obtaining predicted data Qu as the values of the derivative of u on the boundary $\partial\Omega$. The inverse problem then consists of the recovery of $m(\mathbf{x})$ from the boundary data. We thus have for $G(m)$ in (1b) the solution of a simple Poisson problem.

Note that this inverse problem has no unique solution even under idealized conditions [27, 29, 28], unlike the EIT problem. A generalization of this inverse problem, with the Laplacian in (14) replaced by a generalized Poisson operator with a complicated conductivity structure, has applications in inverse gravimetry [4, 30], EEG [32], ECG [3], and EMG [16, 15].

In those applications the solution studied is often the minimum norm solution. Another favourite regularization method is the "smooth solution", obtained from a Tikhonov penalty function discretizing $\int_\Omega |\nabla m|^2$. However, the minimum norm solution method has a tendency to place the sources as close as possible to the boundary, whereas the smooth solution method introduces excessive smearing of sharp features. In recent years much attention has been devoted to total variation (TV) regularization, which involves a Tikhonov penalty functional discretizing a standard modification of $\int_\Omega |\nabla m|$ in an attempt to preserve jumps in m; see, e.g., [44, 2, 30]. In the present context, however, this method does not perform well either, as we shall see.

In the sequel, m is assumed to be piecewise constant with three possible values m_1, m_2, and m_3, i.e., $l = 3$. Without loss of generality we can set the background m_3 value to 0, as it can always be reset by a redefinition of variables in (14). Even under these restrictions, though, the solution is not unique, and with our method the selection between possible solutions is made by the choice of initial guess.

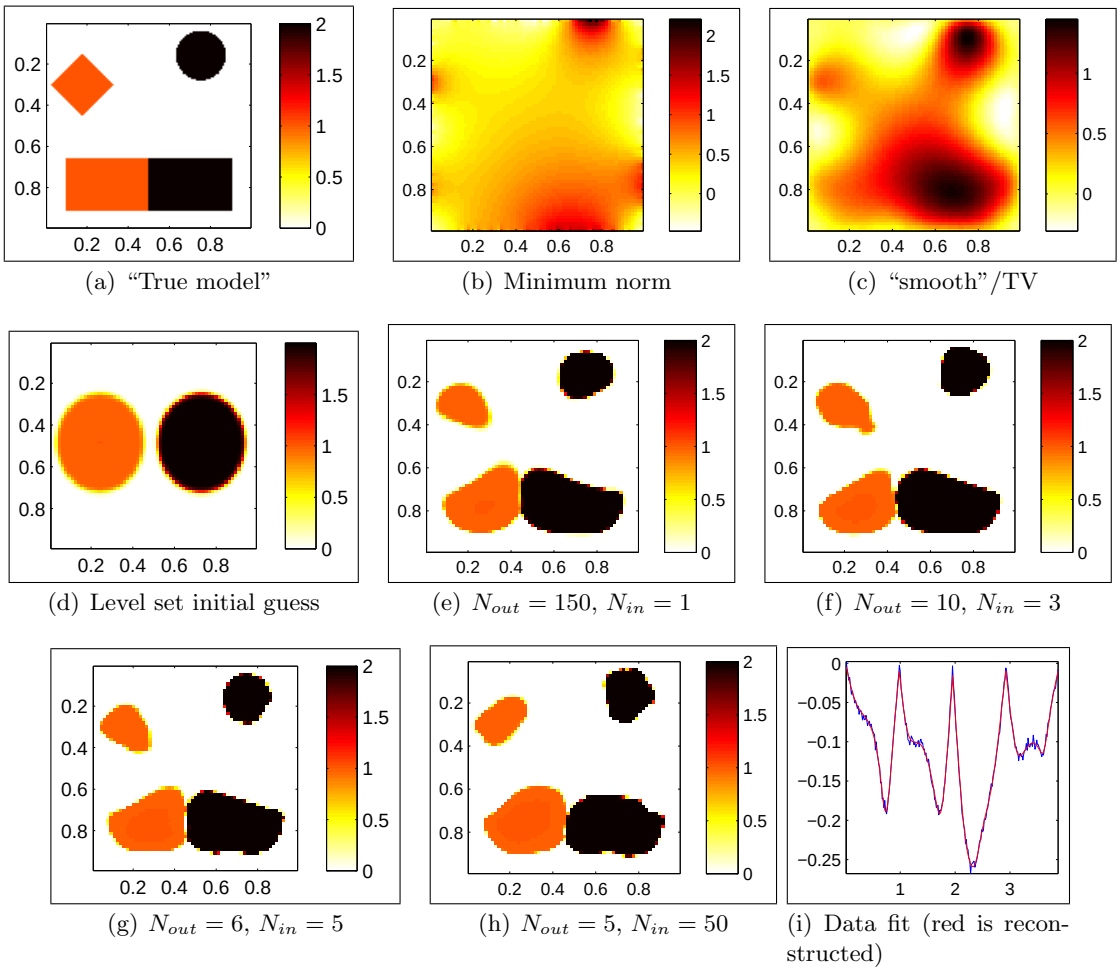
The resulting number of level set functions is $n = 2$, and in (2a) we set $a_{00} = a_{11} = 0$, $a_{01} = m_1$, and $a_{10} = m_2$. Unlike [11] where finite elements were used, we discretize (14) using the standard finite difference scheme on a uniform mesh with spacing h. The model m is discretized as a piecewise constant function on the dual mesh. In approximating (2) we use (4) with $\varepsilon = h$.

4.2 Numerical experiments in 2D

In this section we follow [11] in the experimental setup, selecting the unit square for Ω and obtaining predicted data Qu as the outward normal derivative of u on the mesh points discretizing the entire boundary $\partial\Omega$.

We have selected a 64×64 discretization mesh, i.e., $h = 2^{-6}$. (Similar results were obtained upon using $h = 2^{-7}$.) At first we experiment with known values m_i. The "true solution" depicted in Fig. 2(a) was used on a mesh with width $h/2$ to generate artificial "exact data" on the entire boundary mesh, to which we added 3% Gaussian noise to create the data used in the following experiment. For comparison with our level set based reconstructions we computed the minimum norm solution, the "smooth solution" and the TV solution, depicted in Figs. 2(b-c). The latter two reconstructions were computed by minimizing the standard Tikhonov regularized least-squares data fitting error function, and selecting the regularization parameter α using the discrepancy principle [20], to obtain a misfit (defined in (15) below) of about 5%.

The minimum norm solution is concentrated near the boundary, as predicted, and as such is the worst. The smooth solution and the TV regularized solution are visually indistinguishable! The reason why the TV regularization fails to sharpen the edges, as it is reputed to do in image processing applications such as denoising and deblurring, is that there is no information present in the surface data to indicate sharp edges in the reconstructed solution. Note that even though the TV regularization allows for sharp edges, it still discourages them. The penalty on jumps in m can be further reduced by other functionals such as Tukey's [37], but this introduces nonconvexity and thus potentially spurious solutions into the nonlinear Tikhonov minimization process. Note that our results contrast the findings in [4], where TV regularization was claimed to lead to reconstructions with sharp edges. Knowledge about edge existence and location was



(a) "True model" (b) Minimum norm (c) "smooth"/TV

(d) Level set initial guess (e) $N_{out} = 150$, $N_{in} = 1$ (f) $N_{out} = 10$, $N_{in} = 3$

(g) $N_{out} = 6$, $N_{in} = 5$ (h) $N_{out} = 5$, $N_{in} = 50$ (i) Data fit (red is reconstructed)

Figure 2: Model reconstruction for the inverse potential problem in 2D with known m_i-levels. The minimum norm, smooth solution and TV reconstructions all yield poor results. The level set reconstructions are much more pleasing. The choices $N_{in} = 3$ or 5 yield particularly efficient computations.

introduced there through their initial guess, constituting a restriction which we avoid here.

Turning to the level set based method, we initialized the level set functions as depicted in Fig. 2(d). Resulting reconstructions are shown in Figs. 2(e–h) based on various choices for the number of inner iterations. The process is terminated in each case after the misfit value falls below 5%. The misfit is defined as

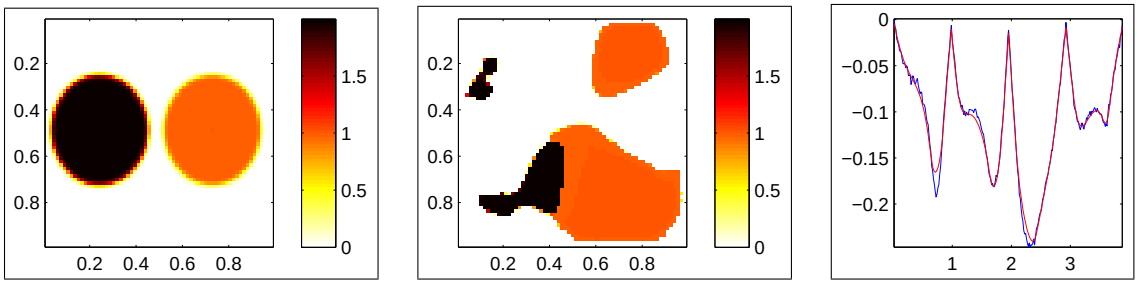
$$||F(P(\psi)) - b||_2/||b||_2\,. \tag{15}$$

Employing $N_{in} = 3$ or 5 for (13) with $\alpha = 0$ yields a very stable and fast process, despite the topological change from initial guess to final shapes. With this, the iteration terminates after $N_{out} = 10$ or 6 approximate Gauss-Newton iterations, respectively, with misfit $\approx 3\%$. If we take $N_{in} = 50$, corresponding to an attempt to solve the outer iteration (13) exactly, then of course we must select $\alpha > 0$ or else we would be attempting to solve a singular system. The choice $\alpha = 10^{-3}$ yields results of a similar quality at about 10 times the price, compare Figs. 2(g) and 2(h). If we set $N_{in} = 1$ then we are doing preconditioned steepest descent

[14, 17], and the iteration terminates at $N_{out} = 150$. Note the excellent data fitting depicted in Fig. 2(i), which is common to all level set method variants.

Turning to the method of [11, 21], we estimate costs as in the beginning of Section 4. Note that using $N_{in} = 5$ and $C_F = C_P$ the cost of one iteration of (13) is equivalent to that of $33/5 \approx 7$ artificial time steps of [11]. Thus, $N_{out} = 6$ outer iterations of (13) are equivalent in computational cost to less than 50 simpler artificial time steps. Since simple artificial time methods typically require thousands of steps to converge (cf. [44, 19, 7, 21]), the demonstrated superiority of (13) is by a significant margin.

If we start the iteration with poor initial conditions, for example with m_1 and m_2 reversed, the iteration can get stuck in a local minimum, as depicted in Fig. 3. In order to arrive at the correct solution the reconstructed regions would have to change places, and in the process of doing so would temporarily have to increase the data fitting measure. In this run, after 100 additional outer iterations the misfit did not decrease any further.



(a) Initial guess

(b) $N_{out} = 50, N_{in} = 5,$ misfit $= 5\%$

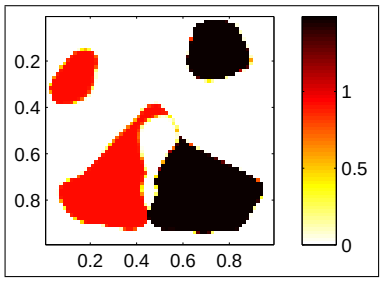
(c) Data fit (red is reconstructed)

Figure 3: When the iteration is started with an unfortunate configuration the solution may not converge and get stuck in a local optimum. The misfit in this case did not reduce below 5%.

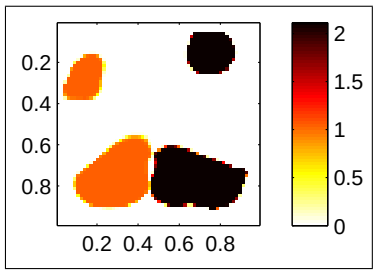
Next we assume that the exact levels of m_1 and m_2 are unknown in the same setup as above, i.e., $l_u = 2$, $l = 3$. The reconstruction results generally depend on the quality of the initial guesses m_1^0 and m_2^0, so really our procedure is effective for sharpening ball-park estimates, not for determining the values of m_i from thin air. Fig. 4 displays some relatively successful reconstructions obtained using $N_{in} = 5$ and $\alpha = 0$. Interestingly, the unfortunate initial condition as in Fig. 3 is now easily recovered from, see Fig. 4(a). This is because the additional degrees of freedom allow the shapes to "change identity", by changing their m_i values.

The added flexibility with respect to the initial guess configuration afforded when the m_i levels are assumed unknown suggests a method for recovery from unfortunate initial guesses for the problem with known m_i levels. We start with a few iterations as in Fig. 4(a) and then switch to iterating with the known m_i values. If the initial choice was such that the shapes would have had to change place or pass though each other, the effect can be approximately achieved within the first couple of iterations by refining the parameters m_i.

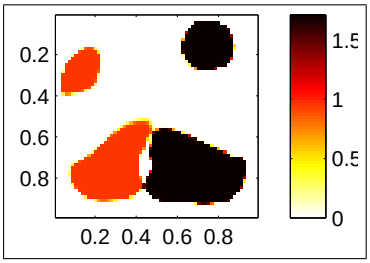
This is demonstrated in Fig. 5. Of course, the above is not a foolproof recipe for solving all hard, large, nonconvex problems of the type considered here, but it is a simple idea that is easy to program and apply, and which occasionally works well.



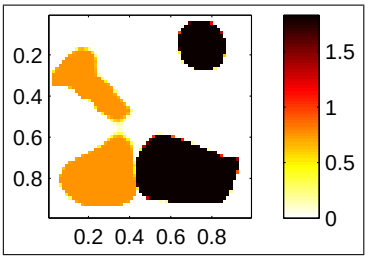
(a) $m_1^0 = 2$, $m_2^0 = 1$, i.e., same as in Fig. 3



(b) $m_1^0 = 1.5$, $m_2^0 = 1.5$

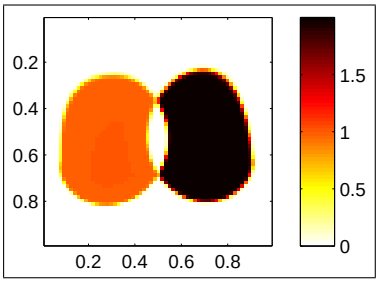


(c) $m_1^0 = 1.0$, $m_2^0 = 1.0$

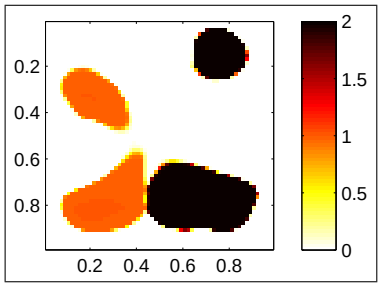


(d) $m_1^0 = 0.5$, $m_2^0 = 1.0$

Figure 4: Model reconstruction for the inverse potential problem with unknown m_1 and m_2 levels. Results: (a) $m_1 = 0.91$, $m_2 = 1.48$, $misfit = .025$, $N_{out} = 10$; (b) $m_1 = 1.04$, $m_2 = 2.11$, $misfit = .026$, $N_{out} = 10$; (c) $m_1 = .96$, $m_2 = 1.71$, $misfit = .026$, $N_{out} = 9$; (c) $m_1 = .75$, $m_2 = 1.83$, $misfit = .024$, $N_{out} = 9$.



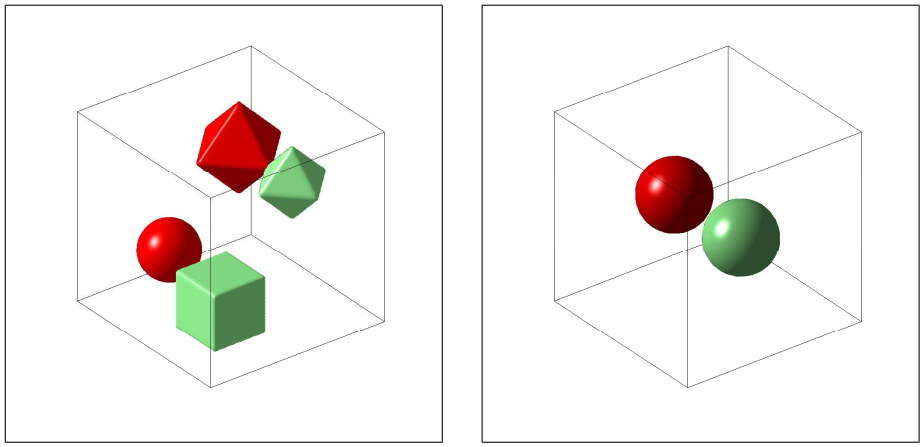
(a) Starting with $m_1^0 = 2$, $m_2^0 = 1$, i.e., same as in Fig. 3; result after 3 iterations



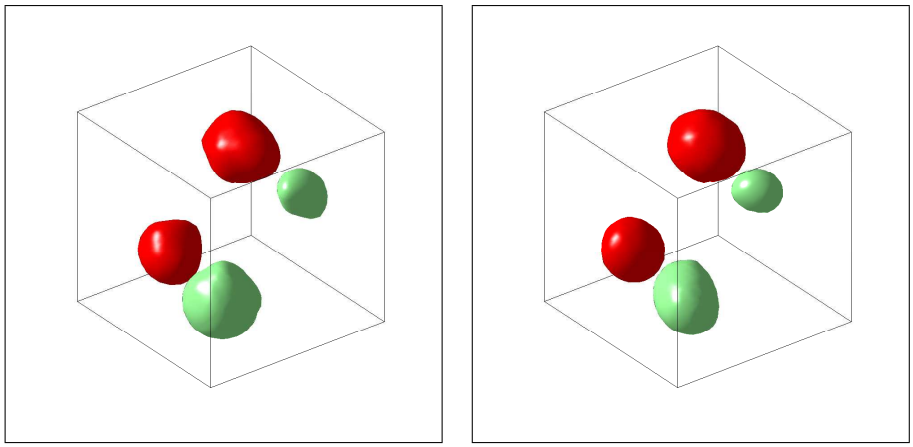
(b) Starting with the result of (a) for known levels $m_1 = 1$, $m_2 = 2$; result after 6 iterations

Figure 5: Obtaining good results for the problem of Fig. 3: (a) a warm start configuration is obtained first by pretending that the m_i are unknown and applying 3 iterations; (b) using the results of (a) as the initial configuration, good results are obtained for the known m_i.

4.3 Numerical experiments in 3D



(a) True model (b) Initial guess



(c) $N_{out} = 6, N_{in} = 5$, known levels (d) $N_{out} = 8$, $N_{in} = 5$, unknown levels

Figure 6: Model reconstruction in 3D for the inverse potential problem with known and unknown m_i-levels, from data on the boundary. The green (light) shapes in the true model have $m_1 = 1$, the red (dark) shapes have $m_2 = 2$. The shapes defined by level set functions ψ^1 and ψ^2 have the corresponding colors. The recovered m values for the experiment with unknown levels starting at $m_i^0 = 3$ were $m_1 = 1.42$ and $m_2 = 1.78$.

We have tested the fast method extending [14] and described in Section 3.2 for the same problem setup as above but in 3D, with the unit cube replacing the unit square for the domain Ω. A uniform 32^3 grid was used for the inversion and a uniform 64^3 grid was employed for the artificial data generation with the same amount of noise (3%) as before added. The discretization of the forward problem is the obvious extension of the 2D method. The forward problem is solved iteratively using PCG with an incomplete Cholesky factorization as a preconditioner.

Fig. 6 depicts typical reconstructions with known levels $m_1 = 1$ and $m_2 = 2$ and with unknown levels starting at $m_1^0 = m_2^0 = 3$. As before we have chosen initial guesses such that an iterative process converging to something near the "true model", or "ground truth", would have to go through a topological change. This latter aspect of the setup often proves to be a major stumbling block for methods that do not use level sets; see Section 4.2 as well as, e.g.,

[14].

The resulting performance of our method is essentially similar to that recorded in the 2D case. As such, the efficiency is more impressive in 3D because there are many more grid points and hence each matrix-vector multiplication is more expensive. A single outer iteration on our 3Ghz desktop computer takes about 15 seconds.

4.4 Partial surface data in 3D

In an attempt at a somewhat more realistic simulation resembling gravimetry, we restrict the data to the upper surface, and test how well two distinct objects with densities 1 and 2 can be recovered when they are placed deeper and deeper. We place two objects at eight different depths, and reconstruct their positions (thus, there are a total of eight experiments). Again the background density is 0. The two objects are depicted in Fig. 7(a,b), and the last two rows of subfigures in Fig. 7 show a side view of their positions at various depths from almost touching the upper surface (where the data is taken) to all the way at the bottom. The shapes and locations are then reconstructed using three methods: (i) our present fast method with two level set functions with known values, i.e., $l = 3,\ l_u = 0,\ n = 2,\ m_1 = 1,\ m_2 = 2,\ m_3 = 0$, see Figs. 8 and 11; (ii) using TV regularization with the results thresholded, see Figs. 9 and 12; and (iii) our method from [14] using one level set function with assumed density $m_1 = 1$ and background $m_2 = 0$, see Figs. 10 and 13. The latter illustrates the robustness of the method against modeling uncertainties. As before we injected 3% artificial noise, set $N_{in} = 5$, and terminated the iteration at a misfit of 5%. The number of outer iterations varied between 5 for shallow objects and up to 50 for the deepest objects.

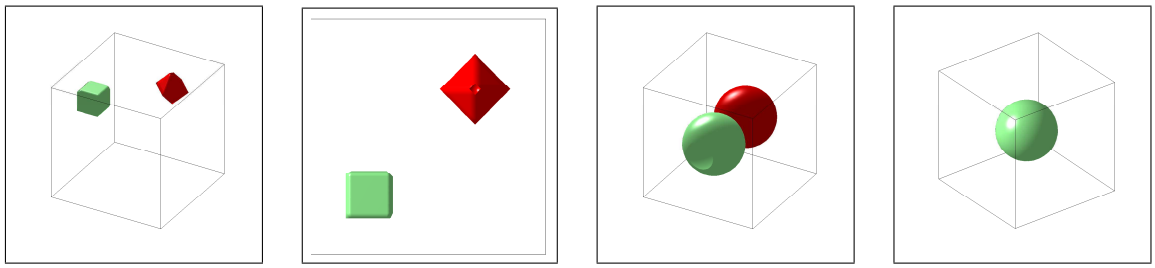
We see that beyond a certain depth the two level set method can no longer distinguish between the two densities, and both objects are identified as having mass 2, yet the localization of the objects remains fairly accurate. By contrast the TV regularization method cannot resolve the depth of the objects at all, and it finds a solution with an equivalent source distribution near the surface. The reconstruction with one level set shows significant shape distortions for objects near the surface, but for deeper objects it tracks the depth only slightly less accurately than the two level set method.

We have also experimented with the levels m_1 and m_2 taken to be unknown. However, the results are not good enough to be reproduced here: the reconstruction favors large objects with low density near the surface. This is not surprising as the potential field of a ball of radius ρ and fixed total mass (or charge) μ is independent of ρ outside the ball, so the actual density cannot be recovered for distant sources.

5 Conclusions

In this paper we have extended the methods and theory of our previous works [21, 11, 13, 14] to handle highly ill-posed surface recovery inverse problems, where the forward problem involves inverting elliptic PDEs, and where the sought surface in 2D or 3D is piecewise constant with several (not just two) levels that are potentially unknown. Multiple level set functions are used.

Currently, there is no adequate theory to support practical methods for this class of problems. The theory developed in Section 2, which extends that developed in [21, 11], helps to highlight some of the difficulties involved: although in practice many technical difficulties are avoided by using the smoothed H_h rather than H, the limit case as $h \to 0$ remains fragile.

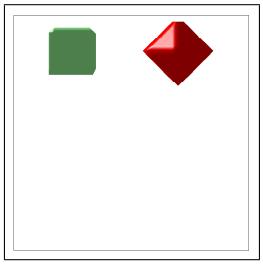


(a) True model (objects near top)

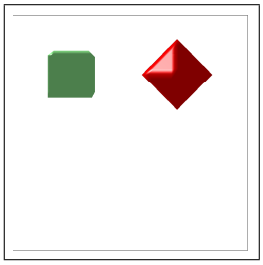
(b) Top view of true model

(c) Initial configuration for 2 level set functions

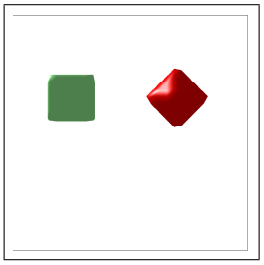
(d) Initial configuration for 1 level set function



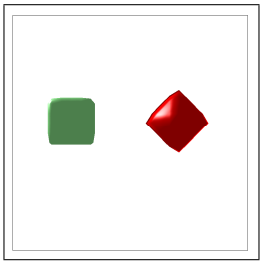
(e) True model at 1/8 depth, side view



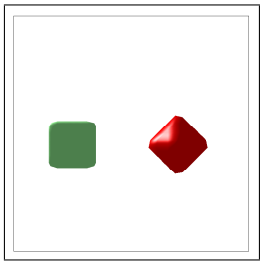
(f) True model at 1/4 depth, side view



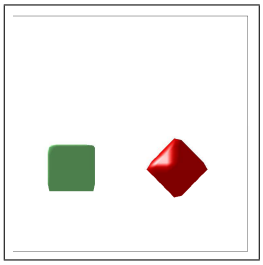
(g) True model at 3/8 depth, side view



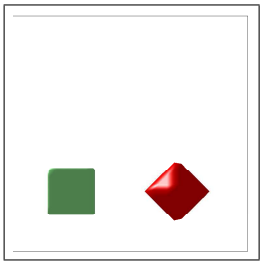
(h) True model at 1/2 depth, side view



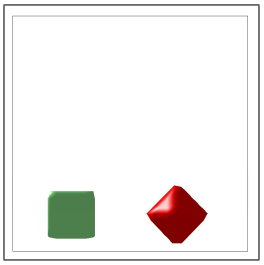
(i) True model at 5/8 depth, side view



(j) True model at 3/4 depth, side view



(k) True model at 7/8 depth, side view



(l) True model at bottom, side view

Figure 7: Recovery of two objects of density 1 (green/light) and 2 (red/dark) from data on the top surface only, as described in Section 4.4. The initial configurations for the reconstruction based on two level sets and on one level set are depicted in (c) and (d), respectively. The last two rows depict a side view of the true model with the objects placed at various depths.

We have further placed the methods developed in the above references on a common platform that allows understanding their relative logic and merit. The methods presented and developed in Section 3 get progressively more efficient and, frustratingly, also seemingly further away from supporting theory. The method that we end up using in practice, presented in Section 3.2 for both cases of known and unknown levels, no longer uses a Tikhonov-type regularization and relies for smoothing effects on a limited number of PCG inner iterations. This yields efficiency that in many instances improves by an astounding amount on other methods that have been proposed in the literature in recent years, see for example [19, 7, 18, 11], reducing typical CPU run times on a laptop from several hours to a minute or two. Correctness of this method has been verified by comparing against another method from Section 3.1.

Our fast method has been tested here on several inverse potential problems in two and three space variables. It also applies to other problems such as EIT, DC resistivity and electromagnetic data inversion [14, 17], although we have not tested cases of unknown levels for

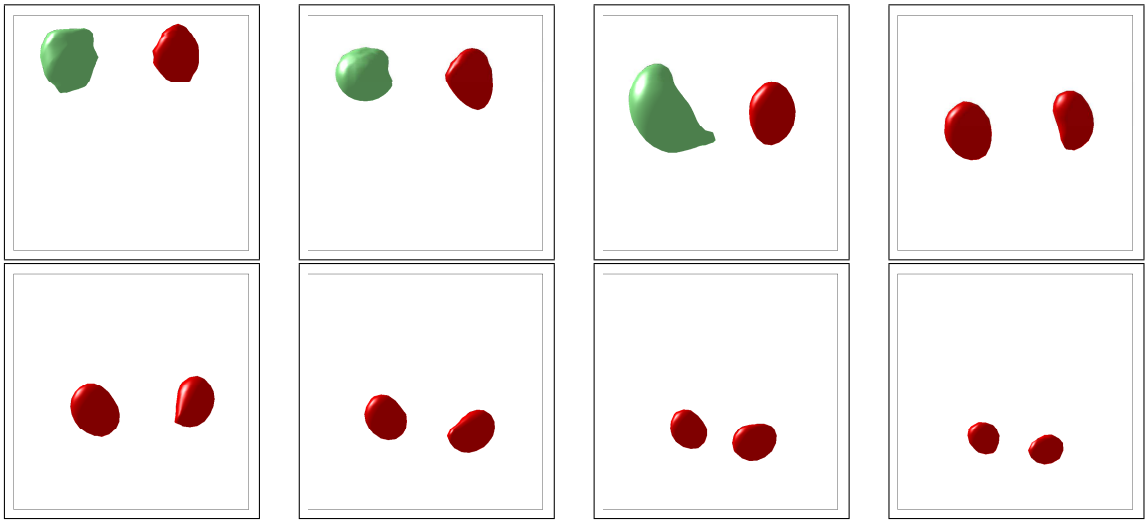


Figure 8: Side views of reconstructions of two objects of density 1 (green/light) and 2 (red/dark) at different depths from data on the top surface only, using two level set functions. Compare respectively with the true models in Fig. 7(e-l). The noise level was 3% and the iteration was terminated when the misfit fell below 5%. Observe the good localization achieved even in depths where the two nonzero densities can no longer be distinguished from each other.

those instances. In addition to being fast, an important advantage of our method is that there is no need to choose the regularization parameter α.

For the inverse potential problem, where the data are given only on the boundary, we have shown that such information that is not extractable with a TV regularization can yield successful reconstructions with a method that uses a few level set functions. A clear case has been made for the use of our fast multiple level set method in practical applications.

For cases with unknown levels m_i, reasonable recovery can often be realized from ball-park estimates of the levels. Moreover, we have shown an example where pretending not to know the levels has in effect improved the initial guess to a point where an impasse created by a bad initial guess got resolved. On the other hand, the practical use of the method for gravimetry problems in the presence of unknown m_i is less clear.

Our level set functions ψ^i are in $H^1(\Omega)$. One may argue that using instead level set functions that are required to be only in $L^2(\Omega)$ (as, e.g., in [36]) could simplify calculations. However, we have found that using the smoother level set functions produces a more stable iterative procedure with a more logical shape evolution. Throughout many calculations we have not encountered any need to re-initialize, as is common in other level set implementations. The term $|\boldsymbol{\nabla}\hat{\psi}^i|^2$ in (3c), which yields the Laplacian operator in (12c) or, for the case $\alpha = 0$, corresponds to the Laplacian preconditioner, is at the heart of our algorithms. With it we obtain extremely efficient and stable methods, while replacing this Laplacian by the identity often produces poor results.

References

[1] A. Alessandrini and S. Vessella. Lipschitz stability for the inverse conductivity problem. *Adv. Appl. Math.*, 35:207–241, 2005.

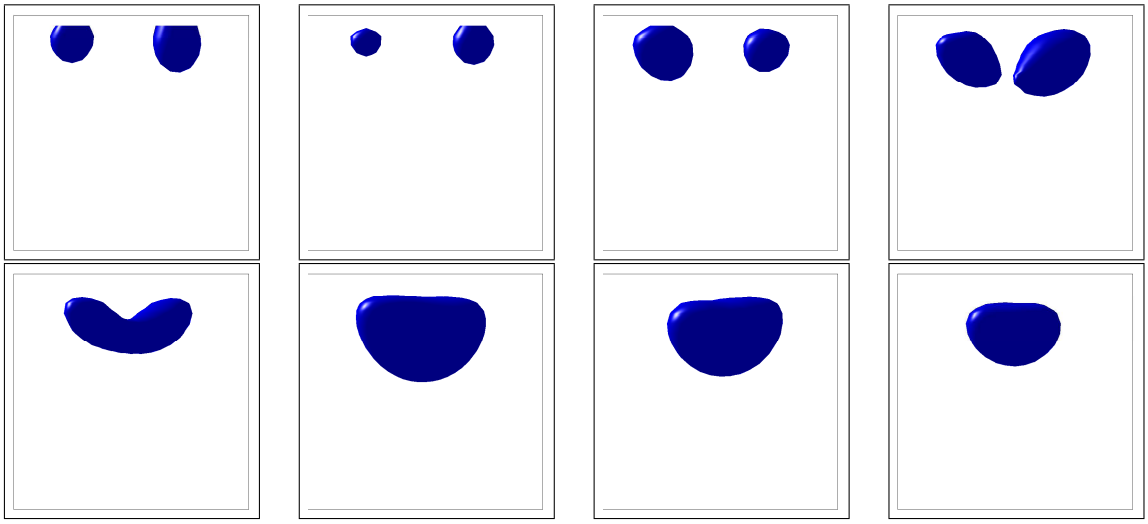


Figure 9: Side views of reconstructions of two objects of density 1 and 2 at different depths from data on the top surface only, using TV regularization. Compare respectively with the true models in Fig. 7(e-l). To allow visual comparison we manually defined and depicted a level set to indicate the region where most mass was located. The noise level was 3% and the iteration was terminated when the misfit fell below 5%. The results are significantly poorer than in Fig. 8.

[2] U. Ascher, E. Haber, and H. Huang. On effective methods for implicit piecewise smooth surface recovery. *SIAM J. Sci. Comput.*, 28:339–358, 2006.

[3] T. Berger, F. Hintringer, and G. Fischer. Noninvasive Imaging of Cardiac Electrophysiology. *Indian Pacing Electrophysiol J.*, 7(3):160–165, 2007.

[4] H. Bertete-Aquirre, E. Cherkaev, and M. Oristaglio. Non-smooth gravity problem with total variation penalization functional. *Geophys. J. Int.*, 149:499–507, 2002.

[5] M. Burger. A level set method for inverse problems. *Inverse Problems*, 17:1327–1355, 2001.

[6] M. Burger. Levenberg-marquardt level set methods for inverse obstacle problems. *Inverse problems*, 20:259–282, 2004.

[7] T. Chan and X. Tai. Level set and total variation regularization for elliptic inverse problems with discontinuous coefficients. *J. Comp. Phys.*, 193:40–66, 2003.

[8] T. Chan and L. Vese. A multiphase level set framework for image segmentation using the mumford and shah model. *Int. J. Comput. Vis.*, 50:271–293, 2002.

[9] M. Cheney, D. Isaacson, and J. C. Newell. Electrical impedance tomography. *SIAM Review*, 41:85–101, 1999.

[10] E. Chung, T. Chan, and X. Tai. Electrical impedance tomography using level set representations and total variation regularization. *J. Comp. Phys.*, 205:357–372, 2005.

[11] A. DeCezaro, A. Leitão, and X.-C. Tai. On multiple level-set regularization methods for inverse problems. *Inverse Problems*, 25:035004, 2009.

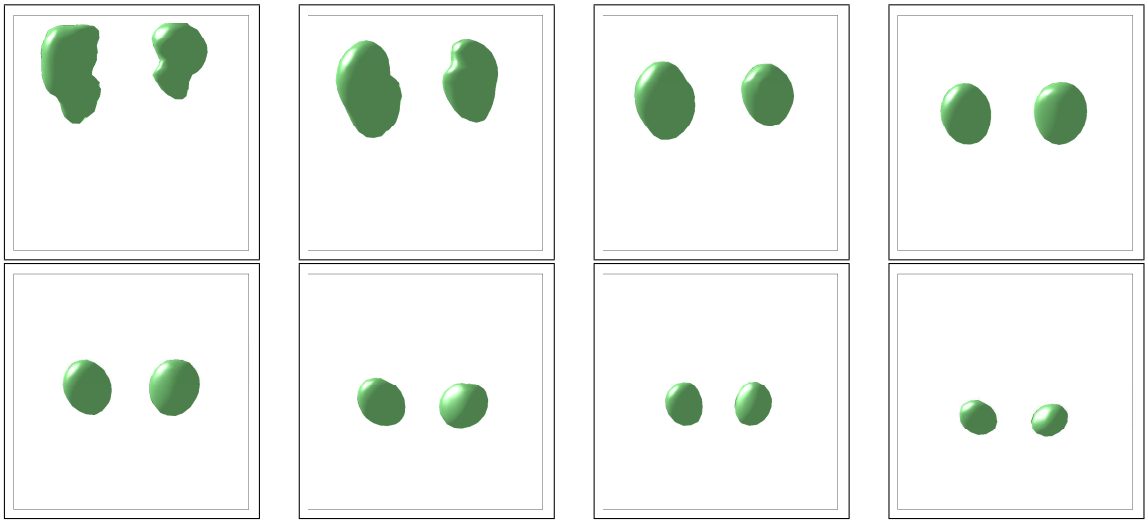


Figure 10: Side views of reconstructions of two objects of density 1 and 2 at different depths from data on the top surface only, using one level set function. Compare respectively with the true models in Fig. 7(e-l). The noise level was 3% and the iteration was terminated when the misfit fell below 5%. Note the relatively good localization achieved despite the inevitable error in the density levels.

[12] A. J. Devaney. The limited-view problem in diffraction tomography. *Inverse Problems*, 5:510–523, 1989.

[13] K. van den Doel and U. Ascher. On level set regularization for highly ill-posed distributed parameter estimation problems. *J. Comp. Phys.*, 216:707–723, 2006.

[14] K. van den Doel and U. Ascher. Dynamic level set regularization for large distributed parameter estimation problems. *Inverse Problems*, 23:1271–1288, 2007.

[15] K. van den Doel, U. Ascher, and D. Pai. Computed myography: three dimensional reconstruction of motor functions from surface EMG data. *Inverse Problems*, 24:065010, 2008.

[16] K. van den Doel, U. M. Ascher, A. Curt, J. Steeves, and D. K. Pai. Computed myography: three dimensional reconstruction of motor functions from surface EMG data. In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, Canada, August 20-24*, pages 550–554, 2008.

[17] O. Dorn and U. Ascher. Shape reconstruction in 3D electromagnetic induction tomography using a level set technique. In *Proc. 23rd International Review of Progress in Applied Computational Electromagnetics ACES*, pages 1–6, 2007.

[18] O. Dorn and D. Lesselier. Level set methods for inverse scattering. *Inverse Problems*, 22:R67–R131, 2006. Topical Review.

[19] O. Dorn, E. L. Miller, and C. M. Rappaport. A shape reconstruction method for electromagnetic tomography using adjoint fields and level sets. *Inverse Problems*, 16, 2000. 1119-1156.

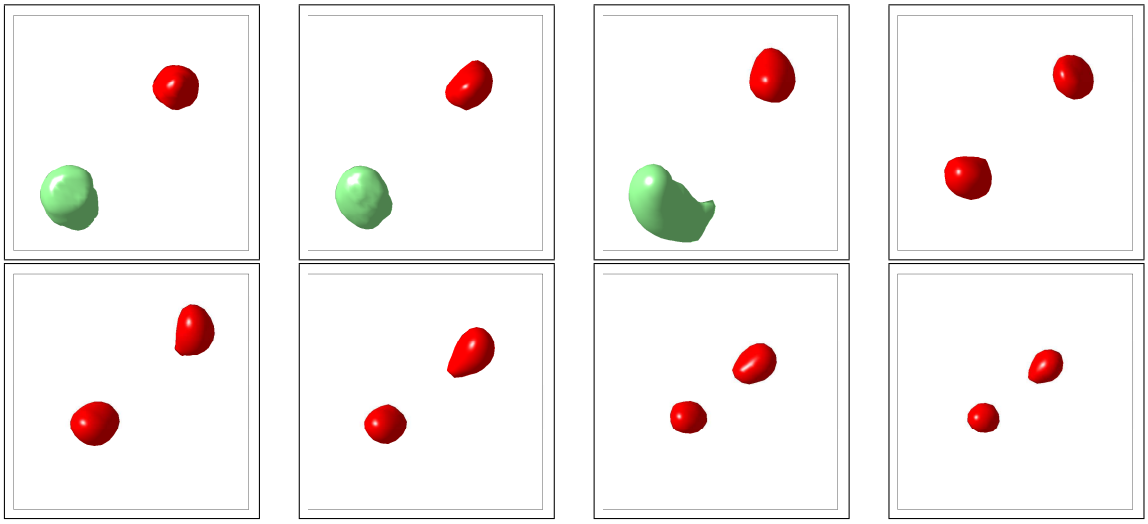


Figure 11: Top views corresponding to Fig. 8.

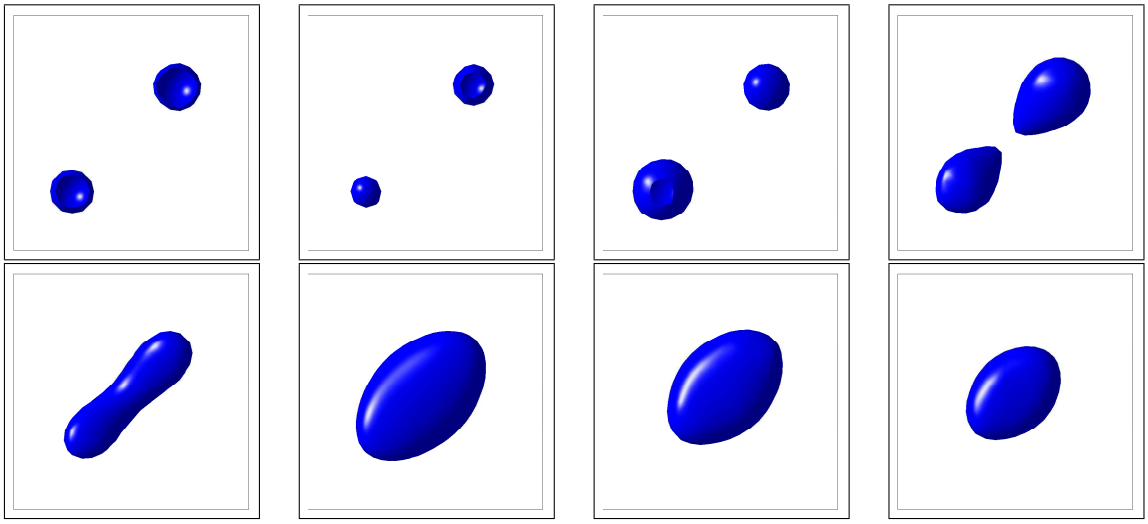


Figure 12: Top views corresponding to Fig. 9.

[20] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer, Dordrecht, 1996.

[21] F Fruhauf, O. Scherzer, and A. Leitão. Analysis of regularization methods for the solution of ill-posed problems involving unbounded operators and a relation to constraint optimization. *SIAM J. Numer. Anal.*, 43:767–786, 2005.

[22] E. Haber. A multilevel, level-set method for optimizing eigenvalues in shape design problems. *J. Comp. Phys.*, 198:518–534, 2004.

[23] E. Haber, U. Ascher, D. Aruliah, and D. Oldenburg. Fast simulation of 3D electromagnetic using potentials. *J. Comput. Phys.*, 163:150–171, 2000.

[24] E. Haber, U. Ascher, and D. Oldenburg. Inversion of 3D electromagnetic data in frequency and time domain using an inexact all-at-once approach. *Geophysics*, 69:1216–1228, 2004.

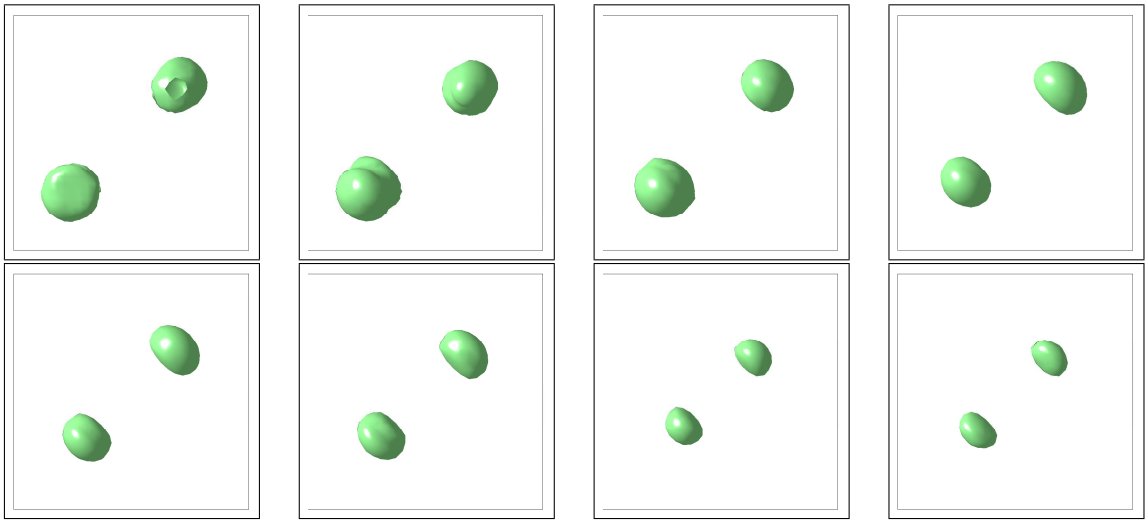


Figure 13: Tops views corresponding to Fig. 10.

[25] E. Haber, S. Heldmann, and U. Ascher. Adaptive finite volume method for distributed non-smooth parameter identification. *Inverse Problems*, 23:1659–1676, 2007.

[26] M. Hanke and P. C. Hansen. Regularization methods for large scale problems. *Survey on Mathematics for Industry*, 3:253–315, 1993.

[27] H. L. F. Helmholtz. Ueber einige gesetze der verheilung elektrischer ströme in körperlicher leitern mit anwendung auf die thierisch-elektrischen versuche. *Ann. Physik und Chemie*, 9:211–233, 1853.

[28] F. Hettlich and W. Rundell. Iterative methods for the reconstruction of an inverse potential problem. *Inverse Problems*, 12:251–266, 1996.

[29] V. Isakov. *Inverse Problems for Partial Differential Equations*. Springer, 2006.

[30] B. Kirkendall, Y. Li, and D. Oldenburg. Imaging cargo containers using gravity gradiometry. *IEEE Trans. Geoscience and Remote Sensing*, 45:1786–1797, 2007.

[31] A. Leitão and O. Scherzer. On the relation between constraint regularization, level sets, and shape optimization. *Inverse Problems*, 19:L1–L11, 2003.

[32] C. M. Michel, M. M. Murray, G. Lantz, S Gonzalez, L. Spinelli, and R. G. de Peralta. EEG source imaging. *Clinical neurophysiology*, 115:2195–2222, 2004.

[33] G. Newman and D. Alumbaugh. Three-dimensional massively parallel electromagnetic inversion–I. theory. *Geophysical journal international*, 128:345–354, 1997.

[34] G. Newman and D. Alumbaugh. Three-dimensional massively parallel electromagnetic inversion–II, analysis of a crosswell electromagnetic experiment. *Geophysical journal international*, 128:355–367, 1997.

[35] R. L. Parker. *Geophysical Inverse Theory*. Princeton University Press, Princeton NJ, 1994.

[36] F. Santosa. A level-set approach for inverse problems involving obstacles. *ESAIM Controle Optim. Calc. Var.*, 1:17–33, 1996.

[37] G. Sapiro. *Geometric Partial Differential Equations and Image Analysis*. Cambridge, 2001.

[38] O. Scherzer. Convergence rates of iterated tikhonov regularized solutions of nonlinear ill-posed problems. *Numer. Math.*, 66:259–279, 1993.

[39] O. Scherzer. Scale-space methods and regularization for denoising and inverse problems. *Advances in Image and Electron Physics*, 128:445–530, 2003.

[40] N. C. Smith and K. Vozoff. Two dimensional DC resistivity inversion for dipole dipole data. *IEEE Trans. on geoscience and remote sensing*, GE 22:21–28, 1984.

[41] X.-C. Tai and T. Chan. A survey on multiple level set methods with applications to identifying piecewise constant functions. *Int. J. Num. Anal. Model.*, 1:25–47, 2004.

[42] X.-C. Tai and H. Li. A piecewise constant level set method for elliptic inverse problems. *Appl. Numer. Math.*, 57:686–696, 2007.

[43] A. N. Tikhonov and V. Ya. Arsenin. *Methods for Solving Ill-posed Problems*. John Wiley and Sons, Inc., 1977.

[44] C. Vogel. *Computational methods for inverse problem*. SIAM, Philadelphia, 2002.