*Journal of Global Optimization* **25:** 407–423, 2003.



# Solving an Inverse Problem for an Elliptic Equation by d.c. Programming

LE THI HOAI AN[1], PHAM DINH TAO[1] and DINH NHO HÀO[2,⋆⋆‡]
[1]*Institut National des Sciences Appliquées de Rouen, Laboratoire de Modélisation et Optimisation-Recherche Opérationnelle, BP 8, 76131 Mont Saint Aignan, France; E-mail: lethi@lmi.insa-rouen.fr, pham@insa-rouen.fr;* [2]*Hanoi Institute of Mathematics, P.O. Box 631 Bo Ho, 10 000 Hanoi, Vietnam; E-mail: hao@thevinh.ncst.ac.vn*

**Abstract.** An inverse problem of determination of a coefficient in an elliptic equation is considered. This problem is ill-posed in the sense of Hadamard and Tikhonov's regularization method is used for solving it in a stable way. This method requires globally solving nonconvex optimization problems, the solution methods for which have been very little studied in the inverse problems community. It is proved that the objective function of the corresponding optimization problem for our inverse problem can be represented as the difference of two convex functions (d.c. functions), and the difference of convex functions algorithm (DCA) in combination with a branch-and-bound technique can be used to globally solve it. Numerical examples are presented which show the efficiency of the method.

**Key words:** branch-and-bound technique, DCA, d.c. programming, ill-posed problem, inverse problem, Tikhonov regularization

## 1. Introduction

Inverse problems for differential equations are frequently encounted in science, engineering, geophysics, medicine, etc. Following J.B. Keller [12], "we call two problems **inverse** of one another if the formulation of each involves all or part of the solution of the other. Often, for historical reasons, one of the two problems has been extensively studied for some time, while the other is newer and not so well understood. In such cases, the former is called the **direct problem**, while the latter is called the **inverse problem**". For references on inverse problems we refer the reader to Bui [1], Groetsch [9], Isakov [11], Engl et al [5], Tikhonov et al [28] and the references therein.

Let $\Omega$ be a bounded domain in $\mathbb{R}^n$ $(n = 1, 2, 3)$. Consider the boundary value problem

$$\begin{aligned} -\Delta u + cu = f \quad & \text{in } \Omega \\ u = g \quad & \text{on } \partial\Omega. \end{aligned} \tag{1}$$

‡ Present address: Vrije Universiteit Brussel, Department of Electronics and Information Processing, Pleinlaan 2, 1050 Brussel, Belgium, E-mail: hao@etro.vub.ac.be

⋆⋆ Partially supported by the PAST visiting programme and Alexander von Humboldt Foundation.

Traditionally, in this problem the functions $c, f, g$ are given and $u$ is to be found; this is the *direct problem*. With appropriate conditions on $\Omega, c, f$ and $g$ there is a unique solution of (1) (see Section 2) and it is well-posed in the sense of Hadamard. In this paper we consider *the inverse problem* of determination of the coefficient $c(x)$ from $u$ given in the whole domain $\Omega$.

Inverse problems are sometimes non-linear (the above inverse problem is non-linear) and mainly ill-posed. The last means that (1) inverse problems might not have a solution in some sense, (2) if there is a solution, it might not be unique, (3) and/or might not continuously depend on the data. The instability of the solution with respect to the data makes severe difficulties, as a small error in the data may cause dramatically large errors in the solution. The classical numerical methods for well-posed problems are no more applicable to ill-posed problems, the new appropriate ones (*regularization methods*) are required. For references on such methods, again we refer the reader to Isakov [11], Engl et al. [5], Tikhonov et al. [28] and the references therein.

One of the most popular methods for solving ill-posed problems is Tikhonov's regularization method. As the inverse problem of determination of $c$ from $u$ is ill-posed (see Section 2 below), we have to use a regularization method for it. Let $z \in L^2(\Omega)$ be an observation for $u$. One of the variants of Tikhonov's regularization method for our inverse problem has the form

$$\min \left\{ J_\alpha(c) := \frac{1}{2}\|u(c) - z\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|L(c - c^*)\|^2_Y, c \in \mathcal{C}. \right\} \tag{2}$$

Here $\alpha > 0$ is the Tikhonov regularization parameter which has to be determined from the noisy data $z$ and noise level $\epsilon$, $\mathcal{C}$ is the domain in which $c$ lives, $c^*$ plays the role of a selection criterion or a guess for the unknown $c$, $L$ and $Y$ are some appropriate operator and space, respectively. (For more details, see Section 2). The solution of the problem (2) is stable with respect to the noise level in $z$, and if $\alpha$ is chosen properly with respect to the noise level $\epsilon$, it can be proved that the solution of the problem (2) converges to the exact solution of our inverse problem as the noise level tends to zero (see Engl et al. [5, 8]). However, one of the main problems of Tikhonov regularization method, how to solve the non-linear optimization problem (2), has been very little studied in the inverse problems community. If one cannot find a global solution of (2), the regularization method of Tikhonov does not seem to work. The main purpose of this paper is to fill this gap in the literature.

Our motivation comes from the fact that the functional $J_\alpha(x)$ is a *d.c.* one. It means that it can be represented as the **d***ifference of two* **c***onvex functions*. Therefore, we can consider (2) as a d.c. program. One of the most efficient and robust algorithms for large scale d.c. programming is the *difference of convex functions algorithm* (DCA) suggested by Pham Dinh Tao [22, 23] (see also [15, 18, 24, 25]). DCA is an iterative method which solves, at each iteration, a convex subproblem. The main idea of this algorithm is to work with convex programs which approximate the primal and dual problems, via the d.c. duality. Since DCA is a local method for nonconvex optimization problems, in general, we should combine it with some other techniques to find global solutions of the problem (2). We shall use the branch-and-bound technique for this purpose in this paper.

In the next section we shall study the inverse problem of determining the coefficient $c$ from the observation $z$ of $u$. We shall prove that this problem is ill-posed, and the functional (2) is d.c. In Section 3 we shall outline DCA and branch-and-bound techniques for (2). Finally, in Section 4 a numerical example is presented.

This paper is the second one of our work on applying the techniques of d.c. programming to non-linear ill-posed problems (see [20, 21]).

## 2. Inverse problem

Let $\Omega$ be a bounded domain in $\mathbb{R}^n (n = 1, 2, 3)$. For $n = 1$ we set $\Omega = (0, 1)$. For $n = 2, 3$ we suppose that $\Omega$ is a sphere, a sphere shell, a parallelepiped, or a domain that can be transformed into one of these domains by a regular mapping $y = y(x) \in C^2(\bar{\Omega})$ (see [14, p. 74]). Consider the boundary value problem

$$\begin{aligned} -\Delta u + cu &= f \quad && \text{in } \Omega \\ u &= g && \text{on } \partial\Omega, \end{aligned} \tag{3}$$

where $f$ and $g$ are functions in $L^2(\Omega)$ and $H^{3/2}(\partial\Omega)$, respectively. We suppose that $c \in \mathcal{C}$, where

$$\mathcal{C} = \{c \in L^2(\Omega), 0 < c_1 \leqslant c(x) \leqslant c_2 < \infty \text{ a.e.}\},$$

$c_1, c_2$ are given positive numbers.

We note that with these assumptions there exists a unique solution $u \in H^2(\Omega)$ of the problem (3). Further, there exists a constant $k$ that depends only on $\Omega$ and $c_2$ such that

$$\|u\|_{H^2(\Omega)} \leqslant k(\|f\|_{L^2(\Omega)} + \|g\|_{H^{3/2}(\partial\Omega)}). \tag{4}$$

In what follows for short we set

$$\mathcal{R} := \|f\|_{L^2(\Omega)} + \|g\|_{H^{3/2}(\partial\Omega)}.$$

Now we consider the inverse problem (we call it (IP)) of determining the coefficient $c(x)$ from a noisy observation $z \in L^2(\Omega)$ of $u$:

$$\|z - u\|_{L^2(\Omega)} \leqslant \epsilon. \tag{5}$$

(The number $\epsilon$ is called the noise level.) This problem is ill-posed as the following example shows.

EXAMPLE . Consider the direct problem

$$\begin{aligned} & -u_{xx}(x) + c(x)u(x) = f, \quad 0 < x < \pi, \\ & u(0) = g_1 \quad u(\pi) = g_2, \end{aligned} \tag{6}$$

Suppose that $0 < c_1 \leqslant c(x) \leqslant c_2$, $f \in L^2(0, \pi)$. Then there is a unique solution $u$ to this problem. We choose $g_1, g_2$ and $f$ in such a way that $u > 1$, or simple prescribe $u > 1$ and then choose $g_1, g_2$ and $f$. Let now $u$ be inexactly given by $u^\epsilon$ and $c$ be found from $u^\epsilon$. For example, let

$$u^\epsilon(x) = u(x) + \epsilon \sin(nx), \quad 0 < \epsilon < 1.$$

Then we can verify that

$$c^\epsilon(x) = \frac{f(x) + u_{xx}(x) - n^2\epsilon \sin(nx)}{u(x) + \epsilon \sin(nx)}$$

is a solution of the inverse problem with respect to the data $u^\epsilon$.

We observe that

$$c - c^\epsilon = c\frac{\epsilon \sin(nx)}{u + \epsilon \sin(nx)} + n^2 \frac{\epsilon \sin(nx)}{u + \epsilon \sin(nx)}.$$

Take, for example, $x = \pi/2$, $n = 2m + 1$. Since $u > 1$, $0 < \epsilon < 1$,

$$c - c^\epsilon = c\frac{\epsilon}{u + \epsilon} + (2m+1)^2 \frac{\epsilon}{u + \epsilon} \to \infty, \text{ as } m \to \infty.$$

Thus, although the noise level $\epsilon$ in the data is small, the errors in the solution of the inverse problem are arbitrary large. The problem is ill-posed!

A natural way to solving the inverse problem (IP) is to minimize the output least squares functional

$$J(c) = \frac{1}{2}\|u(c) - z\|^2_{L^2(\Omega)}$$

over $\mathcal{C}$. However, since (IP) is ill-posed, this variational problem has the same nature. Tikhonov's regularization method should therefore be used. Namely, we minimize the regularized functional

$$J_\alpha(c) := J(c) + \frac{\alpha}{2}\|c - c^*\|^2_{L^2(\Omega)} \tag{7}$$

over $\mathcal{C}$. Here $c^* \in L^\infty(\Omega)$ is a guess for $c$ or a selection for $c$. As it has been proved in Chavent and Kunisch [2], Engl et al. [8], if we choose $\alpha = \sqrt{\epsilon}$, then as $\epsilon$ tends to zero the solution of (7) tends to the solution of (IP) which is nearest $c^*$. There are many other methods of choosing the regularization parameter $\alpha$ to guarantee this convergence property, see, e.g., Engl et al. [5], Kunisch and Ring [13], Tikhonov et al. [28], however in this paper we use this method only, as our main aim is to suggest a method for globally solving (7).

Now we study some properties of the functional $J$. The nonlinear mapping $F : D(F) \subseteq L^2(\Omega) \to H^2(\Omega)$ is defined as the parameter-to-solution mapping

$$F(c) = u(c),$$

with $u(c)$ being the solution of (3) and, with some $\hat{\epsilon} > 0$,

$$D(F) := \{c \in L^2(\Omega) : \|c - \hat{c}\| \leqslant \hat{\epsilon} \text{ for some } \hat{c} \in \mathcal{U}\}.$$

Following [3, 4] it can be proved that $F : L^2(\Omega) \to H^2(\Omega)$ is twice continuously Fréchet differentiable, its first differential $F'(c)h = u'(c)h := \eta$ in direction $h \in L^2(\Omega)$ is the solution of the problem

$$\begin{aligned} & -\Delta\eta + c\eta = -hu(c) \quad \text{in } \Omega \\ & \eta = 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{8}$$

and its second derivative in direction $(h, h_1) \in L^2(\Omega) \times L^2(\Omega)$

$$(u''(c)h, h_1) := \xi(h, h_1) := \xi$$

is the unique solution of the problem

$$\begin{aligned} & -\Delta\xi + c\xi = -h(u'(c)h_1) - h_1(u'(c)h) \quad \text{in } \Omega \\ & \xi = 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{9}$$

Since $u \in H^2(\Omega)$, $n = 1, 2, 3$, $u \in L^\infty(\Omega)$. It follows that $hu \in L^2(\Omega)$. Hence, $\eta \in H^2(\Omega) \cap H_0^1(\Omega)$. Analogously, $\xi \in H^2(\Omega) \cap H_0^1(\Omega)$.

It is clear now that $J$ is twice continuously Fréchet differentiable and

$$J'(c)h = (u'(c)h, u(c) - z) \tag{10}$$

and

$$(J''(c)h, h) = \|u'(c)h\|^2 + (u(c) - z, (u''(c)h, h)). \tag{11}$$

Here and hereafter $(\cdot, \cdot)$ and $\|\cdot\|$ are the inner product and norm in $L^2(\Omega)$, respectively.

Now we estimate $(J''(c)h, h)$. In doing so, we note that, since $n \in \{1, 2, 3\}$, for any $\psi \in H^2(\Omega)$, there is a constant $k_1$ depends only on $\Omega$ such that (see, e.g. [14, p. 38–39])

$$\|\psi\|_{L^\infty(\Omega)} \leqslant k_1 \|\psi\|_{H^2(\Omega)}. \tag{12}$$

Hence, from (8), (4) and (12),

$$\begin{aligned} \|u'(c)h\| & \leqslant k\|hu(c)\| \\ & \leqslant k\|u(c)\|_{L^\infty(\Omega)}\|h\| \\ & \leqslant kk_1\|u(c)\|_{H^2(\Omega)}\|h\| \\ & \leqslant k^2 k_1 \mathcal{R}\|h\|. \end{aligned} \tag{13}$$

Further, the last inequality and (9) yield

$$\begin{aligned}\|(u''(c)h,h)\| &\leqslant 2k\|h(u'(c)h)\|\\ &\leqslant 2k\|u'(c)h\|_{L^\infty(\Omega)}\|h\|\\ &\leqslant 2kk_1\|u'(c)h\|_{H^2(\Omega)}\|h\|\\ &\leqslant 2k^3k_1^2\mathcal{R}\|h\|^2.\end{aligned}$$

Thus,

$$|(J''(c)h,h)| \leqslant k^4k_1^2\mathcal{R}^2\|h\|^2 + 2k^3k_1^2\mathcal{R}(k\mathcal{R}+\|z\|)\|h\|^2.$$

Hence, with

$$\rho = k^4k_1^2\mathcal{R}^2 + 2k^2k_1^2\mathcal{R}(k\mathcal{R}+\|z\|) \tag{14}$$

the functional

$$J(c) + \frac{\rho}{2}\|c\|^2$$

is convex. Thus, the functional $J_\alpha(c)$ is d.c., since we can represent it as the difference of two convex functions, for example,

$$J_\alpha(c) = \frac{\rho}{2}\|c\|^2 - \left(\frac{\rho}{2}\|c\|^2 - J(c) - \frac{\alpha}{2}\|c\|^2\right). \tag{15}$$

REMARK 2.1. In many cases we can estimate the sup-norm of the solution of a boundary value problem for elliptic equations by the maximum principle so that the upper bound for $\|J''\|$ can be easier obtained.

To apply DCA (see Section 3)) to our problem we need also a convenient representation for $J'(c)$. It can be seen from (10) that $J'(c)h = ((u'(c)^*)(u(c)-z),h)$. We prove that in fact

$$J'(c)h = -\int_\Omega u(c)\varphi h\,\mathrm{d}x, \tag{16}$$

where $\varphi$ be the solution of the adjoint problem

$$\begin{aligned}&-\Delta\varphi + c\varphi = u(c) - z \quad &&\text{in } \Omega\\ &\varphi = 0 &&\text{on } \partial\Omega.\end{aligned} \tag{17}$$

This may have already been proved somewhere else, however, for completeness we outline a proof here.

First, we note that $\varphi \in H^2(\Omega)\cap H_0^1(\Omega)$, and

$$\|\varphi\|_{H^2(\Omega)} \leqslant k\|u(c)-z\| \leqslant k(k\mathcal{R}+\|z\|). \tag{18}$$

For $h \in L^2(\Omega) \cap L^\infty(\Omega)$, set

$$v = u(c+h) - u(c).$$

We have

$$v = u(c+h) - u(c) = u'(c)h + r,$$

where $r := r(c,h)$ and

$$\lim_{\|h\|\to 0} \frac{\|r\|}{\|h\|} = 0.$$

Besides, $v \in H^2(\Omega) \cap H_0^2(\Omega)$ is the unique solution of the problem

$$\begin{array}{ll} -\Delta v + cv = -hu(c+h) & \text{in } \Omega \\ v = 0 & \text{on } \partial\Omega. \end{array} \tag{19}$$

It easily follows

$$\begin{aligned} J(c+h) - J(c) &= (u(c) - z, v) + \frac{1}{2}\|v\|^2 \\ &= -\int_\Omega u(c+h)\varphi h \,\mathrm{d}x + \frac{1}{2}\|v\|^2 \\ &= -\int_\Omega (u(c) + u'(c)h + r)\varphi h \,\mathrm{d}x + \frac{1}{2}\|v\|^2 \\ &= -\int_\Omega u(c)\varphi h dx - \int_\Omega ((u'(c)h) + r)\varphi h \,\mathrm{d}x + \frac{1}{2}\|v\|^2. \end{aligned}$$

Routine estimates yield

$$J(c+h) - J(c) = -\int_\Omega u(c)\varphi h \,\mathrm{d}x + o(\|h\|).$$

On the other hand, from (18), (4) and (17) we have

$$\begin{aligned} \left| -\int_\Omega u(c)\varphi h \,\mathrm{d}x \right| &\leqslant \|\varphi\|_{L^\infty(\Omega)} \|u(c)\| \cdot \|h\| \\ &\leqslant k_1 \|\varphi\|_{H^2(\Omega)} \|u(c)\| \cdot \|h\| \\ &\leqslant k^2 k_1 (k\mathcal{R} + \|z\|)\mathcal{R}\|h\|. \end{aligned}$$

Thus, (16) is proved.

## 3. DCA and branch-and-bound method

To solve (7) numerically we have to work with its finite-dimensional setting. For example, suppose that for any given $c \in \mathcal{C}$ with some discretization method we can approximate the solution of the direct problem (3) via $u_1, u_2, \dots, u_n$ and the solution of the adjoint problem (17) via $\varphi_1, \varphi_2, \dots, \varphi_n$. (We note that (7) and (17) of the same type, so to solve them numerically, we can use the same code). Suppose further that this discretization method approximates $c$ with the aid of $[c] = (c_1, c_2, \dots, c_n)$ and the $c_i$ belong to $\tilde{\mathcal{C}} := \{c : \tilde{c}_1 \leqslant c \leqslant \tilde{c}_2\}$. The functional $J_\alpha$ has now the form

$$I_\alpha([c]) = \frac{1}{2}\sum_{i=1}^{n} \beta_i |u_i([c]) - z_i|^2 + \frac{\alpha}{2}\beta_i \sum_{i=1}^{n}(c_i - c_i^*)^2.$$

Here we use the symbol $I$ for finite-dimensional functionals and $\beta_i$, $i = 1, 2, \dots, n$ are some defined numbers, which depend on the discretization method.

The gradient of $I_\alpha([c])$ has the form

$$\frac{\partial I_\alpha([c])}{\partial c_i} = -\beta_i u_i([c])\varphi_i + \alpha\beta_i(c_i - c_i^*).$$

Thus, we have a finite-dimensional (non-convex) optimization problem:

$$\min\{I_\alpha([c]) : c \in \tilde{\mathcal{C}}\}. \tag{20}$$

In this section we describe two approaches to solve a more general problem of the form

$$\alpha := \min\{f(x) : x \in \Pi_1^n[a_i, b_i], -\infty < a_i \leqslant b_i < \infty, i = 1, 2, \dots, n\}, \tag{21}$$

where $f$ is a real valued function that is not given in explicit form but a gradient of $f$ can be computed. In the next section we will show how our approaches work for (20). Before going further we outline some notions in convex analysis for the reader's convenience.

We follow [27] for definitions of usual tools of convex analysis where functions could take infinite values $\pm\infty$. A function $\theta : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be proper if it takes nowhere the value $-\infty$ and is not identically equal to $+\infty$. The effective domain of $\theta$, denoted by dom $\theta$, is

$$\operatorname{dom}\theta = \{x \in \mathbb{R}^n : \theta(x) < \infty\}.$$

The "convex cone" of all lower semicontinuous proper convex functions on $\mathbb{R}^n$ is denoted $\Gamma_0(\mathbb{R}^n)$. For $g \in \Gamma_0(\mathbb{R}^n)$, the conjugate function $g^*$ of g is a function belonging to $\Gamma_0(\mathbb{R}^n)$ and defined by

$$g^*(y) = \sup\{\langle x, y\rangle - g(x) : x \in \mathbb{R}^n\}.$$

and we have $g^{**} = g$.

Let $g \in \Gamma_0(\mathbb{R}^n)$ and let $x^0 \in \operatorname{dom} g$, then $\partial g(x^0)$ stands for the *subdifferential* of $g$ at $x^0$ and is given by

$$\partial_\epsilon g(x^0) = \{y^0 \in \mathbb{R}^n : g(x) \geqslant g(x^0) + \langle x - x^0, y^0\rangle, \forall x \in \mathbb{R}^n\}.$$

Further, by $\rho(g)$ we denote the modulus of strong convexity of $g$,

$$\rho(g) = \sup\{\rho \geqslant 0 : g - (\rho/2)\|.\|^2 \text{ be convex on } \mathbb{R}\}.$$

### 3.1. A DCA SCHEME

In the first approach we follow the general scheme DCA introduced by Pham Dinh Tao in [22] (see also [15, 18, 19, 25]) to solve the d.c. program

$$\beta = \inf\{f(x) := g(x) - h(x) : x \in \mathbb{R}^n\} \tag{22}$$

Here $f$ and $g$ belong to $\Gamma_0(\mathbb{R}^n)$.

The dual program of the last is also d.c. ([24]):

$$\beta = \inf\{h^*(y) - g^*(y) : y \in \mathbb{R}^n\}. \tag{23}$$

Here $h^*$ and $g^*$ are dual functions of $h$ and $g$, respectively.

The DCA is a primal-dual subdifferential method based on the d.c. duality and the local optimality. It consists in the construction of two sequences $\{x^k\}$ and $\{y^k\}$ such that $x^{k+1}$ (resp. $y^k$) is a solution to the convex program (P$_k$) (resp. (D$_k$)) defined by

$$(\mathrm{P}_k) \qquad \begin{cases} \text{Minimize} & g(x) - [h(x^k) + \langle x - x^k, y^k\rangle] \\ \text{s.t.} & x \in \mathbb{R}^n \end{cases}$$

$$(\mathrm{D}_k) \qquad \begin{cases} \text{Minimize} & h^*(y) - [g^*(y^{k-1}) + \langle x^k, y - y^{k-1}\rangle] \\ \text{s.t.} & y \in \mathbb{R}^n \end{cases} .$$

In view of the relation: (P$_k$) (resp. (D$_k$)) is obtained from (22) (resp. (23)) by replacing $h$ (resp. $g^*$) with its affine minorization defined by $y^k \in \partial h(x^k)$ (resp. $x^k \in \partial g^*(y^{k-1})$), the DCA yields the next scheme:

$$y^k \in \partial h(x^k); \quad x^{k+1} \in \partial g^*(y^k). \tag{24}$$

By this way there hold ([15], [18], [24], [25])

(i) The sequences $\{g(x^k) - h(x^k)\}$ and $\{h^*(y^k) - g^*(y^k)\}$ are decreasing and

- $g(x^{k+1}) - h(x^{k+1}) = g(x^k) - h(x^k)$ if and only if $y^k \in \partial g(x^k) \cap \partial h(x^k)$, $y^k \in \partial g(x^{k+1}) \cap \partial h(x^{k+1})$ and $[\rho(g) + \rho(h)]\|x^{k+1} - x^k\| = 0$.
- $h^*(y^{k+1}) - g^*(y^{k+1}) = h^*(y^k) - g^*(y^k)$ if and only if $x^{k+1} \in \partial g^*(y^k) \cap \partial h^*(y^k)$, $x^{k+1} \in \partial g^*(y^{k+1}) \cap \partial h^*(y^{k+1})$ and $[\rho(g^*) + \rho(h^*)]\|y^{k+1} - y^k\| = 0$. DCA terminates at the $k^{th}$ iteration if either of the above equalities holds.

(ii) If $\rho(g) + \rho(h) > 0$ (resp. $\rho(g^*) + \rho(h^*) > 0$), then the series $\{\|x^{k+1} - x^k\|^2\}$ (resp. $\{\|y^{k+1} - y^k\|^2\}$) converges.

(iii) If the optimal value $\beta$ of problem ($P_{d.c.}$) is finite and the sequences $\{x^k\}$ and $\{y^k\}$ are bounded, then every limit point $x^\infty$ (resp. $y^\infty$) of the sequence $\{x^k\}$ (resp. $\{y^k\}$) is a critical point of $g - h$ (resp. $h^* - g^*$). (A point $x$ is called a *critical point* of $g - h$ if $\partial g(x) \cap \partial h(x) \neq \emptyset$).

(iv) If DCA converges to a point that admits a neighbourhood in which the objective is finite and convex, and if the second d.c. component of the objective function is differentiable at this point, then it is a local minimizer for ($P_{d.c.}$).

(v) DCA converges to a global minimizer for ($P_{d.c.}$) if $f = g - h$ is actually convex ($f$ then is called "false" d.c. function).

The DCA was first introduced by Pham Dinh [22] and later developed in his joint works with Le Thi, see, e.g. Le Thi [15], Le Thi–Pham Dinh [17, 18], Pham Dinh–Le Thi [24, 25] and references therein. It is actually one of a few algorithms in the local approach which has been successfully applied to many large-scale d.c. optimization problems and proved to be more robust and efficient than related standard methods. For all details of the DCA we refer the reader to the just mentioned references.

It is well known that Tikhonov's functional (2) is strongly convex in a neighborhood of a solution $c_0$, so if we start our DCA in an appropriate subset of this neighborhood, we can get a global solution. However, it is hard to find this neighborhood, and since DCA is a local method for global optimization, we shall combine it with the branch-and-bound technique to find global solutions of the problem.

To solve (21) by DCA we first take a $\rho > 0$ such that $\rho/2\|x\|^2 - f(x)$ is convex. Such a $\rho$ always exists when $f$ is twice continuously differentiable (see (14), (15)). We then decompose the function $f$ in the form

$$f(x) = \frac{\rho}{2}\|x\|^2 - \left(\frac{\rho}{2}\|x\|^2 - f(x)\right).$$

It is easily seen that with this d.c. decomposition DCA applied to (21) yields the sequence

$$x^{k+1} = \mathrm{Proj}_{\Pi_1^n[a_i,b_i]}\left(x^k - \frac{1}{\rho}\xi_k\right), \quad \text{with } \xi_k \in \partial f(x^k).$$

Here $\mathrm{Proj}_{\Pi_1^n[a_i,b_i]}$ is the orthogonal projection into the box $\Pi_1^n[a_i, b_i]$. Immediate advantages of this algorithm are: (1) it works when the function $f$ is not given in an explicit form, and (2) the projection of a point onto a box is explicitly and inexpensively computed. The algorithm can be described as follows:

**Algorithm 1** *(DCA for (21))*:

Let $x^0$ be a point in $\mathbb{R}^n$ and $\varepsilon$ be a sufficiently small positive number. Set $er \leftarrow \varepsilon + 1, k \leftarrow 0$

**Do while** $er \geqslant \varepsilon$

Compute $\xi_k \in \partial f(x^k)$. Set

$$y^k = \rho x^k - \xi_k, \text{ with } \xi_k \in \partial f(x^k)$$

and

$$x^{k+1} = \mathrm{Proj}_{\Pi_1^n[a_i,b_i]}(y^k/\rho).$$

$er \leftarrow \|x^{k+1} - x^k\|, k+1 \leftarrow k$
**end do**
We note that if $f$ is differentiable, then $\xi_k$ is nothing other than $\nabla f(x^k)$.

## 3.2. A BRANCH-AND-BOUND ALGORITHM

### 3.2.1. *Lower bound*

In the bounding procedure we relax the objective function, knowing that it is not given in an explicit form. As above we suppose that $f$ is twice continuously differentiable.

To compute a lower bound of $f$ over a subrectangle $B = \{x : \underline{r}_i \leqslant x_i \leqslant \overline{r_i}, : i = 1, \dots, n\}$ we find a convex minorization of $f$ over $B$ as follows. We express $f$ by another d.c. decomposition:

$$f(x) = \left(\frac{\rho}{2}\|x\|^2 + f(x)\right) - \frac{\rho}{2}\|x\|^2 := g(x) - h(x),$$

and then approximate the concave function $-h$ in $B$ by its convex envelope over $B$ which is actually affine. The last, in fact, is $-l(x)$, where

$$l(x) := \sum_{i=1}^{n} l_i(x_i) = \sum_{i=1}^{n} \frac{\rho}{2}[(\underline{r}_i + \overline{r_i})x_i - \underline{r}_i\overline{r_i}] \geqslant h(x).$$

The function

$$\tilde{f}(x) = \left(\frac{\rho}{2}\|x\|^2 + f(x)\right) - \sum_{i=1}^{n} l_i(x_i)$$

is convex and $\tilde{f}(x) \leqslant f(x)$ in $B$. A lower bound $\beta_B$ of $f$ on $B$ can be determined by solving the convex program

$$\beta_B := \min\{\tilde{f}(x) : x \in B\}. \tag{25}$$

Since $f$ is not explicitly defined, we can use the projection gradient method ([26]) to solve (25):

$$x^{k+1} = \mathrm{Proj}_B\left(x^k - \delta_k \eta^k\right), \quad \delta_k \to 0, \sum_{k=0}^{\infty} \delta_k = +\infty \tag{26}$$

where $\eta^k$ is the gradient of $\tilde{f}$ at $x^k$. However, this method converges rather slow. We shall use DCA instead. To this end, we represent $\tilde{f}$ in the form

$$\tilde{f}(x) = \frac{2\rho}{2}\|x\|^2 - \left(\frac{\rho}{2}\|x\|^2 - f(x) + \sum_{i=1}^{n} l_i(x_i)\right).$$

DCA applied to the last has the form

$$x^{k+1} = \mathrm{Proj}_B\left(x^k - \frac{1}{2\rho}(\rho x^k - \nabla f(x^k) + \sum_{i=1}^{n}\frac{\rho}{2}(\underline{r}_i + \overline{r_i})e_i\right),$$

($e_i$ being the $i^{\text{th}}$ vector of the canonical basic of $\mathbb{R}^n$).

### 3.2.2. *Upper bound*

At each iteration one may update the upper bound by computing $g(x^B) - h(x^B)$, where $x^B$ is an optimal solution of (25). However, one can obtain a better upper bound by using the DCA. In fact, if $x^B$ is the best current feasible point, then the DCA starting with $x^B$ can provide a better feasible point. This is guaranteed by the fact that the DCA generates the sequence $\{x^k\}$ such that $\{(g-h)(x^k)\}$ is decreasing (see, e.g., [23, 25]).

### 3.2.3. *Branching operation*

Exploiting the special structure of the feasible set we use the rectangular subdivision for branching. Rectangular subdivision procedures play an important role in branch-and-bound methods. The approach of Phillips-Rosen (see also Kalantari-Rosen [7]) uses exhaustive subdivisions, i.e., any nested sequence of rectangles generated by the algorithm will tend to a single point. In Horst-Tuy [10] a concept of “normal rectangular subdivision” was introduced for the class of separable concave minimization problems that includes the Kalantari-Rosen approach and a subdivision proposed earlier in Falk-Soland [6]. In this paper we use the adaptive bisection introduced in [17]. More precisely, let $B_k = \{y : l_i^k \leqslant y_i \leqslant L_i^k\}$ be the rectangle selected for subdivision. Choose an index $i_k$ satisfying

$$i_k \in \arg\max_i\{l_{ki}(x_i^{B_k}) - h_i(x_i^{B_k})\}$$

and subdivide $B_k$ into two subrectangles

$$B_{k1} = \{y \in B_k : x_{i_k} \leqslant x_{i_k}^{B_k}\}, \quad B_{k2} = \{y \in B_k : x_{i_k} \geqslant x_{i_k}^{B_k}\}.$$

#### 3.2.4. **The combined DCA-branch-and-bound algorithm (BBDCA)**

**Initialization.**
Set $B_0 := B$
Solve the convex problem (by Projection subgradient method (26))

$$\beta_0 := \min\{\tilde{f}(x) : x \in B_0\}.$$

to obtain an optimal solution $x^{B_0}$ and a lower bound of the optimal value $\alpha$.
Solve Problem (21) by the DCA from the starting point $x^{B_0}$ to obtain an upper bound $\gamma_0$ and a feasible point $x^0$.
**If** $\gamma_0 - \beta_0 \leqslant \varepsilon|\gamma_0|$, **then** $stop \leftarrow true$, $x^0$ is an $\varepsilon$- optimal solution of (21)
**else** $stop \leftarrow false$
**endif**
Set $\mathcal{R} \leftarrow \{B_0\}, \quad k \leftarrow 0.$
**While** $stop = false$ **do**
Select a rectangle $B_k \in \mathcal{R}$ such that $\beta_k := \beta_{B_k} = \min\{\beta_B : B \in \mathcal{R}\}$.
Bisect $B_k$ into $B_{k1}$ and $B_{k2}$ by the adaptive rectangular subdivision.
Solve Problems (CP$^{ki}$) (i=1,2)

$$(\mathrm{CP}^{ki}) \qquad \beta_{B_{ki}} := \min\{\tilde{f}(x) : x \in B_{ki}\}$$

to obtain $\beta(B_{ki})$ and $x^{B_{ki}}$.
If $x^{B_{ki}}$ is the best point, i.e., $f(x^{B_{ki}}) < \gamma_{k-1}$, then update the upper bound $\gamma_k$ by applying the DCA to problem (21) from the starting point $x^{B_{ki}}$.
Let $x^k$ be a point in $K$ such that $f(x^k) = \gamma_k$.
Set $\mathcal{R} \leftarrow \mathcal{R} \cup \{B_{ki} : \beta(R_{ki}) < \gamma_k - \varepsilon|\gamma_k|, \ i = 1, 2\} \setminus \{B_k\}$.
**If** $\mathcal{R} = \emptyset$, then $stop \leftarrow true$, $x^k$ is an $\varepsilon$-optimal solution.
**else** Set $k \leftarrow k + 1$.
**endif**
**endwhile**

THEOREM 3.1. (Convergence of BBDCA). *(i) If the algorithm terminates at the iteration $k$, then $x^k$ is a global optimal solution to problem (21). (ii) If the algorithm is infinite, then it generates a bounded sequence $\{x^k\}$ every accumulation point of which is a global optimal solution of (21), and*

$$\gamma_k \searrow \alpha, \quad \beta_k \nearrow \alpha.$$

## 4. Numerical test

Consider the system

$$\begin{aligned} &-u_{xx} + c(x)u = f(x) \quad \text{in } (0, 1), \\ &u(0) = u(1) = 0, \end{aligned} \tag{27}$$

where $f$ is $L^2(\Omega)$. The coefficient $c(x)$ is to be found from a noisy observation $z = u^\epsilon \in L^2(\Omega)$ of $u$:

$$\|z - u\|_{L^2(0,1)} \leqslant \epsilon.$$

Suppose that $\mathcal{C} := \{0 < c_1 \leqslant c \leqslant c_2\}$. We shall minimize

$$J_\alpha(c) = \frac{1}{2}\|u(c) - z\|^2_{L^2(0,1)} + \frac{\alpha}{2}\|c - c^*\|^2_{L^2(0,1)}$$

over $\mathcal{C}$ with $u(c)$ being the solution of (27). As it has been proved in section 2 that (see (16))

$$J'_\alpha(c) = -u(c)\varphi + \alpha(c - c^*),$$

where $\varphi$ is the solution of the adjoint problem

$$\begin{aligned} &-\varphi_{xx} + c(x)\varphi = u(c) - z \quad \text{in } (0,1),\\ &\varphi(0) = \varphi(1) = 0. \end{aligned} \tag{28}$$

To solve the above optimization problem we simply use the finite difference method. We divide the interval $[0, 1]$ into $n$ equal subintervals: $x_0 = 0, x_1 = h, \dots, x_i = ih, \dots, x_n = 1$, where $h = 1/n$. The finite difference scheme for (27) takes the form

$$\begin{aligned} &-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + c_i u_i = f_i, \quad i = 1, 2, \dots, n-1\\ &u_0 = u_n = 0. \end{aligned} \tag{29}$$

Here we denote $g_i = g(x_i)$ for any function $g$ defined on $[0, 1]$. And the discrete version of $J_\alpha$ is

$$I_\alpha([c]) := \frac{1}{2}\sum_{i=1}^{n-1} h|u_i([c]) - z_i|^2 + \frac{\alpha}{2}\sum_{i=1}^{n-1} h(c_i - c_i^*)^2.$$

Here, we denote $[c] = (c_1, c_2, \dots, c_{n-1})$.

The discrete adjoint problem is

$$\begin{aligned} &-\frac{\varphi_{i+1} - 2\varphi_i + \varphi_{i-1}}{h^2} + c_i\varphi_i = u_i([c]) - z_i, \quad i = 1, 2, \dots, n-1\\ &\varphi_0 = \varphi_n = 0. \end{aligned} \tag{30}$$

The gradient of $I_\alpha([c])$ can be represented by

$$\frac{\partial I_\alpha([c])}{\partial c_i} = -hu_i\varphi_i + \alpha h(c_i - c_i^*), i = 1, 2, \dots, n-1.$$

Thus, to find the gradient of $I_\alpha([c])$, in every iteration we have to solve two direct problems (29) and (30).

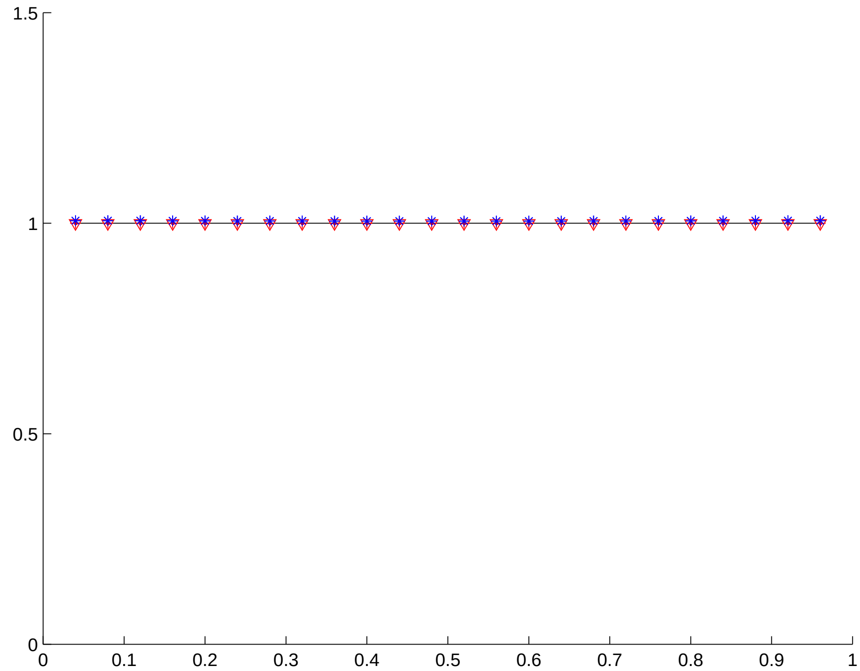


*Figure 1.* First example: exact solution $c = 1$, "$* * *$" for $c^* = x$, "$- - -$" for $c^* = 1.1$, noise level $\epsilon = 10^{-3}$, CPU time $< 1$ min.

In the first test we take $c(x) = 1$, $u(x) = \sin(\pi x)$ and $f = (\pi^2 + 1)\sin(\pi x)$. In the second test we take $c(x) = 1 + x^2$, $u(x) = \sin(\pi x)$, and $f = \sin(\pi x)(1 + x^2 + \pi^2)$. The data $z$ is $u + \epsilon \sin(n\pi x)$. The noise level $\epsilon$ in the both tests is $10^{-3}$, and the dimension of discretization is $n = 25$.

The numerical results for these two examples are presented below. Our code was written in FORTRAN 77 and run on a SUN-SPARC 20. We took $\varepsilon = 10^{-7}$ and $\varepsilon = 10^{-3}$ in Algorithm 1 (DCA) and Algorithm BBDCA, respectively. DCA gives in the most cases an optimal solution: the optimale value given by DCA is *very near* zero and so we need not branching operation anymore. And the algorithm is very fast, the CPU time of our algorithm in all cases is not greater than one minute. Although numerical results are very good, we observed that the choice of the guess function $c^*$ is crucial in the quality of the numerical results. A bad guess may lead to bad numerical results. Thus, suggestions from the real world practice are quite important in solving inverse problems (Figures 1 and 2).

## 5. Conclusions

We have suggested a useful method for solving an inverse problem for elliptic equations which is well known to be non-linear and ill-posed. From our numerical tests it seems to be quite efficient. The method is based on the standard Tikhonov

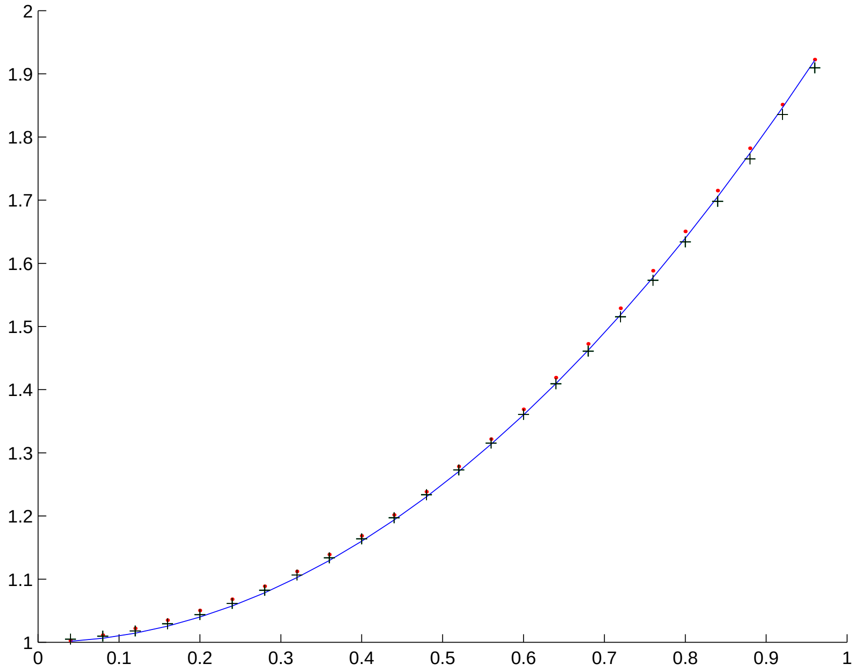


*Figure 2.* Second example: exact solution $c = 1 + x^2$, "..." for $c^* = 1.2 + x^2$, "$+ + +$" for $c^* = 1 + x^2$, "$- - -$" for $c^* = 1.1 + x^2$, noise level $\epsilon = 10^{-3}$, CPU time $< 1$ min.

regularization method but in a further step: globally solve non-convex optimization problems in Tikhonov's regularization method which is a d.c. program. Such kind of non-convex optimization problems has been very little studied in the inverse problems community and there only only local methods are applied. Our aim is to fill this gap in the literature. The DCA seems to fit well to many other non-linear ill-posed problems, say non-convex quadratic problems ([20]), Hammerstein equations, autoconvolution equations, bilinear inverse problems, etc ([21]).

## References

1. Bui, H. D. (1994), *Inverse Problems in the Mechanics of Materials*, CRC Press, Boca Raton, FL.
2. Chavent, G. and Kunisch, K. (1994) Convergence of Tikhonov regularization for constrained ill-posed inverse problems, *Inverse Problems* 10: 63–76.
3. Colonius, F. and Kunisch, K. (1986), Stability for parameter estimation in two point boundary value problem, *J. Reine Angewandte Math.* 370: 1–29.
4. Colonius, F. and Kunisch, K. (1989), Output least squares stability in elliptic systems, *Appl. Math. Opt.* 19: 33-63.
5. Engl, H. W., Hanke, M. and Neubauer, A. (1996), *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht.
6. Falk, J. E. and Soland, R. M. (1969), An algorithm for separable nonconvex programming problems, *Management Science* 15: 550–569.

7. Kalantari, B. and Rosen, J. B. (1987), Algorithm for global minimization of linearly constrained concave quadratic functions, *Mathematics of Operations Research* 12: 544–561.
8. Engl, H. W., Kunisch, K. and Neubauer, A. (1989), Convergence rates for Tikhonov regularisation of non-linear ill-posed problems, *Inverse Problems* 5: 523–540,
9. Groetsch, Ch. W. (1993), *Inverse Problems in the Mathematical Sciences.* Friedr. Vieweg & Sohn, Braunschweig.
10. Horst, R. and Tuy, H. (1993), *Global Optimization: Deterministic Approaches*, 2 edition, Springer, Berlin, New York.
11. Isakov, V. (1998) *Inverse Problems for Partial Differential Equations*, Springer, New York.
12. Keller, J. B. (1976), Inverse problems, *Amer. Math. Monthly* 83: 107–118.
13. Kunisch, K. and Ring, W. (1993), Regularization of nonlinear ill-possed problems with closed operators, *Numer. Funct. Anal. Optimiz.* 14: 389–404.
14. Ladyzhenskaya, O. A. (1985), *The Boundary Value Problems of Mathematical Physics*, Springer, New York, Berlin, Heidelberg, Tokyo.
15. Le Thi Hoai An (1997), *Contribution à l'optimisation non convexe et l'optimisation globale: Théorie, Algorithmes et Applications* , Habilitation à Diriger des Recherches, Université de Rouen.
16. Le Thi Hoai An and Pham Dinh Tao (1997), Solving a class of linearly constrained indefinite quadratic problems by d.c. algorithms, *Journal of Global Optimization* 11: 253–285.
17. Le Thi Hoai An and Pham Dinh Tao (1998), A branch-and-bound method via d.c. optimization algorithm and ellipsoidal technique for box constrained nonconvex quadratic programming problems, *Journal of Global Optimization*, 13: 171–206.
18. Le Thi Hoai An and Pham Dinh Tao (2001), The D.C. programming and DCA revisited with D.C. models of real world nonconvex optimization problems, Proceedings of The 5$^{\text{th}}$ International Conference on Optimization: Techniques and Applications, Hong Kong December 15-17, 2001 (ICOTA 2001) Editor D. Li, Volume 3, pp. 1324–1333.
19. Le Thi Hoai An and Pham Dinh Tao (2002), Large Scale Global Molecular Optimization from exact distance matrices by a d.c. optimization approach, Revised version in *SIAM Journal on Optimization*.
20. Le Thi Hoai An, Pham Dinh Tao, and Dinh Nho Hào (2001), On the ill-posedness of the trust region subproblem. *Journal of Inverse and Ill-Posed Problems* (to appear)
21. Le Thi Hoai An, Pham Dinh Tao, and Dinh Nho Hào (2001), D.c. programming approach to Tikhonov regularization for non-linear ill-posed problems, (in preparation).
22. Pham Dinh Tao and E. B. Souad (1986), Algorithms for solving a class of non convex optimization problems. Methods of subgradients, *Fermat days 85. Mathematics for Optimization*, Elsevier, North-Holland, Amsterdam: 249–270.
23. Pham Dinh Tao and E. B. Souad (1988) Duality in d. c. (difference of convex functions) optimization. Subgradient methods, *Trends in Mathematical Optimization, International Series of Numer Math.* 84 (Birkhäuser), 277–293.
24. Pham Dinh Tao and Le Thi Hoai An (1998), D.c. optimization algorithms for solving the trust region subproblem, *SIAM J. Optimization* 8: 476–505.
25. Pham Dinh Tao and Le Thi Hoai An (1997), Convex analysis approach to d.c. programming: Theory, Algorithms and Applications. *Acta Mathematica Vietnamica* 22(1): 289–355.
26. Polyak, B. (1987), *Introduction to Optimization.* (Optimization Software Inc., Publication Division, New York.
27. Rockafellar, R. T. (1970), *Convex Analysis*, Princeton University Press, Princeton.
28. Tikhonov, A. N., Leonov, A. S. and Yagola, A. G. (1998), *Non-linear Ill-Posed Problems*, Vol. 1, 2, Chapman and Hall, London.