Inverse Problems **14** (1998) 83–100. Printed in the UK PII: S0266-5611(98)86112-1

Numerical identifications of parameters in parabolic systems

Yee Lo Keung† and Jun Zou‡
Department of Mathematics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Received 21 July 1997, in final form 30 October 1997

Abstract. In this paper, we investigate the numerical identifications of physical parameters in parabolic initial-boundary value problems. The identifying problem is first formulated as a constrained minimization one using the output least squares approach with the H^1-regularization or BV-regularization. Then a simple finite element method is used to approximate the constrained minimization problem and the convergence of the approximation is shown for both regularizations. The discrete constrained problem can be reduced to a sequence of unconstrained minimization problems. Numerical experiments are presented to show the efficiency of the proposed method, even for identifying highly discontinuous and oscillating parameters.

1. Introduction

In this paper, we consider a finite element approach, combined with the output least squares method, for identifying the parameter $q(x)$ in the following parabolic problem

$$\frac{\partial u}{\partial t} - \nabla \cdot (q(x)\nabla u) = f(x,t) \qquad \text{in } \Omega \times (0,T) \tag{1.1}$$

with the initial condition

$$u(x,0) = u_0(x) \qquad \text{in } \Omega \tag{1.2}$$

and the Dirichlet boundary condition

$$u(x,t) = 0 \qquad \text{on } \partial\Omega \times (0,T). \tag{1.3}$$

In practical applications, we are often given the terminal status observation

$$u(x,T) = z(x)$$

(possibly through the interpolation of the point observation values) and asked to recover the physical parameter $q(x)$. We shall carry out the recovery process in such a way that the solution u (e.g. the absolute temperature) matches its terminal status observation data z optimally in the energy norm or the L^2-norm. The physical domain Ω can be any bounded domain in R^d ($d \geqslant 1$), with a piecewise smooth boundary Γ, and $f \in H^{-1}(\Omega)$ is a given source term. The problem outlines the heat conduction of some material occupying the domain Ω. We refer to the works by Bank and Kunisch [1] and Engl *et al* [7] for a more applied background of the problem. Also, there are many existing analytical and numerical methods for solving the inverse problem, see [1, 2, 5, 7–10, 14, 16] and the references therein.

† E-mail address: ylkeung@math.cuhk.edu.hk
‡ Author to whom correspondence should be addressed. E-mail address: zou@math.cuhk.edu.hk

0266-5611/98/010083+18$19.50 © 1998 IOP Publishing Ltd

The major novelty of the paper is to investigate the finite element method for solving the nonlinear minimization problem which is formulated using the output least squares method with H^1-regularization or BV-regularization, and to show the finite element convergence for both regularizations. We then reduce the nonlinear constrained minimization to a sequence of unconstrained minimizations. Finally, the Armijo algorithm is suggested for solving the unconstrained finite element minimization problems. The numerical experiments are presented to indicate the stable and global convergence of the proposed numerical algorithms and their effectiveness for the identification of highly discontinuous and oscillating parameters.

Later, we will need the following space of functions with bounded variation

$$BV(\Omega) = \{q \in L^1(\Omega);\ \|q\|_{BV(\Omega)} < \infty\}$$

where $\|q\|_{BV(\Omega)} = \|q\|_{L^1(\Omega)} + \int_\Omega |Dq|$. The notation $\int_\Omega |Dq|$ is not for an integral but for a quantity defined by

$$\int_\Omega |Dq| = \sup\left\{ \int_\Omega q \operatorname{div} g \, \mathrm{d}x;\ g \in (C_0^1(\Omega))^d \text{ and } |g(x)| \leqslant 1 \text{ in } \Omega \right\}.$$

We now consider the parameter identifying problem as the following constrained minimizing process

$$\text{minimize} \quad J(q) = \frac{1}{2}\int_\Omega q(x)|\nabla(v(q;T) - z)|^2 \, \mathrm{d}x + \gamma N(q) \tag{1.4}$$

$$\text{subject to} \quad q \in K \quad \text{and} \quad v \equiv v(q;t) \in H_0^1(\Omega) \quad \text{satisfying} \quad v(x,0) = u_0(x) \text{ in } \Omega \tag{1.5}$$

$$\int_\Omega v_t \phi \, \mathrm{d}x + \int_\Omega q(x)\nabla v \cdot \nabla\phi \, \mathrm{d}x = \int_\Omega f(x,t)\phi \, \mathrm{d}x \qquad \forall \phi \in H_0^1(\Omega) \tag{1.6}$$

for a.e. $t \in (0,T)$. Note that the system (1.5) and (1.6) is the variational formulation associated with the parabolic problem (1.1)–(1.3). Subsequently, we may denote the solution of this variational problem as $v(q;t)$ or $v(q)$ or $v(q)(x,t)$, or simply v in the case where there is no confusion. The function $z \in H_0^1(\Omega)$ appearing in (1.4) is the measured data, and $N(q)$ is a regularization term with a weight coefficient $\gamma > 0$. Throughout the paper, $N(q)$ is taken to be

$$N(q) = \int_\Omega |\nabla q|^2 \, \mathrm{d}x \qquad \text{or} \qquad N(q) = \int_\Omega |Dq| \tag{1.7}$$

namely the semi-norm in $H^1(\Omega)$ or the semi-norm in the BV-space. The constrained set K above is a subset of $H^1(\Omega)$ or $BV(\Omega)$ defined by

$$K = \{q \in L^1(\Omega);\ |||q||| < \infty \text{ and } \alpha_1 \leqslant q(x) \leqslant \alpha_2 \text{ a.e. in } \Omega\}.$$

Here the norm $|||q||| = \|q\|_{H^1(\Omega)}$ or $|||q||| = \|q\|_{BV(\Omega)}$ corresponds to the forms of $N(q)$, α_1 and α_2 are two positive constants.

Note that the evaluation of the cost functional $J(q)$ requires the availability of the terminal status value of the solution $v(q;t)$ to the system (1.5) and (1.6) at $t = T$, this assumes the regularity $v \in C(0,T;H_0^1(\Omega))$. This may not be true in many real applications, for example, with a discontinuous coefficient $q(x)$ or source term $f(x,t)$.

To cover general cases, we will reformulate the problem (1.4)–(1.6) in a weaker and more practical sense in section 2. The remaining sections of the paper are arranged as follows: in section 3 we will discuss the discretization of the minimization problem of section 2 by using a simple finite element method, then reduce the constrained finite element problem to a sequence of unconstrained minimizations. Finally, in section 4 we will derive the Armijo algorithm for solving discrete unconstrained minimizations and present some numerical experiments in section 5.

2. An averaging-terminal status formulation and existence of its solutions

Throughout our analysis, we will make the following assumptions on the given source term and initial data for the parabolic problem (1.1)–(1.3)

$$f \in L^2(Q_T) \qquad \text{and} \qquad u_0 \in H^1(\Omega) \tag{2.1}$$

where $Q_T = \Omega \times (0, T)$. Assuming (2.1), we know from standard parabolic theory that for each $q \in K$ there exists a unique solution $v(q;t)$ to the parabolic problem or equivalently to the variational problem (1.5) and (1.6) and that it has the following regularities

$$v(q) \in L^2(0, T; H_0^1(\Omega)) \qquad v(q) \in H^1(0, T; L^2(\Omega)) \qquad v(q) \in C(0, T; L^2(\Omega)).$$

Instead of the system (1.4)–(1.6), we will use the following weaker and more practical formulation

$$\text{minimize} \quad J(q) = \frac{1}{2}\int_{T-\sigma}^{T}\int_\Omega q(x)|\nabla(v(q;t) - z)|^2\,\mathrm{d}x\,\mathrm{d}t + \gamma N(q) \tag{2.2}$$

$$\text{subject to} \quad q \in K \quad \text{and} \quad v \equiv v(q;t) \in H_0^1(\Omega) \quad \text{satisfying} \quad v(x,0) = u_0(x) \text{ in } \Omega \tag{2.3}$$

$$\int_\Omega v_t\phi\,\mathrm{d}x + \int_\Omega q(x)\nabla v\cdot\nabla\phi\,\mathrm{d}x = \int_\Omega f(x,t)\phi\,\mathrm{d}x \qquad \forall\phi \in H_0^1(\Omega) \tag{2.4}$$

for a.e. $t \in (0, T)$. In the above, σ is a small constant number. In our numerical implementation, we often take σ to be one or two discrete time-step sizes.

In our later analysis, we will make no difference between the semi-norm $\|\nabla\cdot\|_{L^2(\Omega)}$ and the full-norm $\|\cdot\|_{H^1(\Omega)}$ in $H_0^1(\Omega)$ as they are equivalent by Poincaré's inequality.

We are now going to show the existence of minimizers to the problem (2.2)–(2.4). To do so, we need the following lemma.

Lemma 2.1. *For any sequence $\{q_n\}$ in K which converges to some $q \in K$ in $L^1(\Omega)$ as $n \to \infty$, we have*

$$\lim_{n\to\infty}\int_{T-\sigma}^{T}\int_\Omega q_n(x)|\nabla(v(q_n) - z)|^2\,\mathrm{d}x\,\mathrm{d}t = \int_{T-\sigma}^{T}\int_\Omega q(x)|\nabla(v(q) - z)|^2\,\mathrm{d}x\,\mathrm{d}t.$$

Proof. First, taking any $q(x) \in K$ and $\phi = v(q;t)$ in (2.4) and then integrating with respect to t, we derive that

$$\|v(q;t)\|^2_{L^2(\Omega)} + \alpha_1\int_0^t\int_\Omega|\nabla v(q;t)|^2\,\mathrm{d}x\,\mathrm{d}t \leqslant \|u_0\|^2_{L^2(\Omega)} + \frac{1}{\alpha_1}\|f\|^2_{L^2(0,T;H^{-1}(\Omega))} \tag{2.5}$$

for any $t \in (0, T]$. This implies the sequence $\{v(q_n)\}$ is bounded in the space $L^2(0, T; H_0^1(\Omega))$; hence we may extract a subsequence, still denoted by $\{v(q_n)\}$, such that

$$v(q_n) \to v^* \qquad \text{weakly in } L^2(0, T; H_0^1(\Omega)). \tag{2.6}$$

We next show that $v^* = v(q)$. To do this we multiply both sides of the equation

$$\int_\Omega v(q_n)_t\phi\,\mathrm{d}x + \int_\Omega q_n(x)\nabla v(q_n)\cdot\nabla\phi\,\mathrm{d}x = \int_\Omega f\phi\,\mathrm{d}x \qquad \forall\phi \in H_0^1(\Omega) \tag{2.7}$$

by a function $\eta(t) \in C^1[0, T]$ with $\eta(T) = 0$. Then by integrating with respect to t, we get

$$\begin{aligned} -\int_\Omega u_0\eta(0)\phi\,\mathrm{d}x &+ \int_0^T\int_\Omega \eta(t)f\phi\,\mathrm{d}x\,\mathrm{d}t \\ &= -\int_0^T\int_\Omega v(q_n)\phi\eta_t(t)\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_\Omega \eta(t)q(x)\nabla v(q_n)\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t \\ &\quad + \int_0^T\int_\Omega \eta(t)(q_n(x) - q(x))\nabla\phi\cdot\nabla v(q_n)\,\mathrm{d}x\,\mathrm{d}t. \end{aligned} \tag{2.8}$$

The last term in (2.8) converges to zero by (2.5) for $v(q_n)$ and the Lebesgue dominant convergence theorem. Thus, letting $n \to \infty$ in (2.8) and using (2.6), we obtain

$$-\int_\Omega u_0\eta(0)\phi \, \mathrm{d}x + \int_0^T\int_\Omega \eta(t) f\phi \, \mathrm{d}x \, \mathrm{d}t = -\int_0^T\int_\Omega v^*\phi\eta_t(t) \, \mathrm{d}x \, \mathrm{d}t + \int_0^T\int_\Omega \eta(t)q(x)\nabla v^* \cdot \nabla\phi \, \mathrm{d}x \, \mathrm{d}t \tag{2.9}$$

which is valid for any $\eta(t) \in C^1[0, T]$ with $\eta(T) = 0$. Hence, we have

$$\int_\Omega v_t^*\phi \, \mathrm{d}x + \int_\Omega q(x)\nabla v^* \cdot \nabla\phi \, \mathrm{d}x = \int_\Omega f\phi \, \mathrm{d}x \qquad \forall\phi \in H_0^1(\Omega)$$

and $v^*(0) = u_0$. Therefore, $v^* = v(q)$ by the definition of $v(q)$.

Finally, we are ready to prove the desired result of the lemma. We rewrite (2.7) in the form

$$\int_\Omega (v(q_n) - z)_t\phi \, \mathrm{d}x + \int_\Omega q_n(x)\nabla(v(q_n) - z) \cdot \nabla\phi \, \mathrm{d}x = \int_\Omega f\phi \, \mathrm{d}x - \int_\Omega q_n(x)\nabla z \cdot \nabla\phi \, \mathrm{d}x$$

and then take $\phi = v(q_n) - z$ giving

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|v(q_n) - z\|^2_{L^2(\Omega)} + \int_\Omega q_n(x)|\nabla(v(q_n) - z)|^2 \, \mathrm{d}x = \int_\Omega f(v(q_n) - z) \, \mathrm{d}x - \int_\Omega q_n(x)\nabla z \cdot \nabla(v(q_n) - z) \, \mathrm{d}x. \tag{2.10}$$

Similar relations hold for $v(q)$, namely

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|v(q) - z\|^2_{L^2(\Omega)} + \int_\Omega q(x)|\nabla(v(q) - z)|^2 \, \mathrm{d}x = \int_\Omega f(v(q) - z) \, \mathrm{d}x - \int_\Omega q(x)\nabla z \cdot \nabla(v(q) - z) \, \mathrm{d}x. \tag{2.11}$$

Subtracting (2.11) from (2.10) and after some simple manipulations we derive

$$\left\{\int_\Omega q_n(x)|\nabla(v(q_n) - z)|^2 \, \mathrm{d}x - \int_\Omega q(x)|\nabla(v(q) - z)|^2 \, \mathrm{d}x\right\} + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|v(q_n) - v(q)\|^2_{L^2(\Omega)}$$
$$= \int_\Omega q(x)\nabla v(q) \cdot \nabla(v(q_n) - v(q)) \, \mathrm{d}x + \int_\Omega (q_n(x) - q(x))|\nabla z|^2 \, \mathrm{d}x + \int_\Omega \{q_n(x)\nabla v(q_n) - q(x)\nabla v(q)\} \cdot \nabla(v(q) - 2z) \, \mathrm{d}x \equiv: R_n^1 \tag{2.12}$$

where we have used equation (2.4) for both $v(q_n)$ and $v(q)$.

Then by rewriting the first term on the left-hand side of (2.12), we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|v(q_n) - v(q)\|^2_{L^2(\Omega)} + \int_\Omega q_n(x)|\nabla(v(q_n) - v(q))|^2 \, \mathrm{d}x$$
$$= R_n^1 + \left\{\int_\Omega (q(x) - q_n(x))|\nabla(v(q) - z)|^2 \, \mathrm{d}x - 2\int_\Omega q_n\nabla(v(q_n) - v(q)) \cdot \nabla(v(q) - z) \, \mathrm{d}x\right\} \equiv: R_n^1 + R_n^2$$

and integrating over the interval $(0, t)$ for any $t \leqslant T$, we get

$$\frac{1}{2}\|v(q_n; t) - v(q; t)\|^2_{L^2(\Omega)} \leqslant \int_0^T |R_n^1 + R_n^2| \, \mathrm{d}t.$$

By the weak convergence of $v(q_n)$ and the assumed convergence on q_n, it is easy to show

$$\int_0^T |R_n^1 + R_n^2| \, \mathrm{d}t \to 0 \qquad \text{as } n \to \infty.$$

Therefore, we have proved

$$\max_{t\in[0,T]} \|v(q_n; t) - v(q; t)\|_{L^2(\Omega)} \to 0 \qquad \text{as } n \to \infty. \tag{2.13}$$

Now the desired convergence of the lemma follows immediately by integrating (2.12) over $[T - \sigma, T]$ and using (2.13). □

Using lemma 2.1, one can prove (cf [12, 13]):

Theorem 2.1. *There exists at least a minimizer to the optimization problem (2.2)–(2.4).*

Remark 2.1. *All the results of the paper are easily generalized to the L^2-norm case in the cost functional $J(q)$, i.e. its first term is replaced by*

$$\frac{1}{2}\int_{T-\sigma}^T \int_\Omega |v(q; t) - z|^2 \, \mathrm{d}x \, \mathrm{d}t.$$

3. Finite element method and its convergence

We now propose a finite element method for solving the continuous minimization problem (2.2)–(2.4). We first triangulate the polyhedral domain Ω with a regular triangulation $\mathcal{T}^h$ of simplicial elements, namely intervals in one dimension, triangles in two dimensions and tetrahedra in three dimensions (cf Ciarlet [6]). Then we define the finite element space V_h to be the continuous and piecewise linear space over the triangulation $\mathcal{T}^h$, and $\mathring{V}_h$ a subspace of V_h with all functions vanishing on the boundary $\partial\Omega$. Let $\{x_i\}_{i=1}^N$ be the set of all the nodal points of the triangulation $\mathcal{T}^h$, then the constrained subset K is approximated by

$$K_h = \{v_h \in V_h; \alpha_1 \leqslant v_h(x_i) \leqslant \alpha_2 \text{ for } i = 1, 2, \ldots, N\}.$$

To fully discretize the parabolic system (2.3) and (2.4), we also need the time discretization. To do so, we divide the time interval $(0, T)$ into M equally-spaced subintervals by using nodal points

$$0 = t^0 < t^1 < \cdots < t^M = T$$

with $t^n = n\tau$, $\tau = T/M$. For a continuous mapping $u : [0, T] \to L^2(\Omega)$, we define $u^n = u(\cdot, n\tau)$ for $0 \leqslant n \leqslant M$. For a given sequence $\{u^n\}_{n=0}^M \subset L^2(\Omega)$ we define the difference quotient and the averaging function

$$\partial_\tau u^n = \frac{u^n - u^{n-1}}{\tau} \qquad \bar{u}^n = \frac{1}{\tau}\int_{t^{n-1}}^{t^n} u(t) \, \mathrm{d}t.$$

With the above notations, we can formulate the finite element problem corresponding to (2.2)–(2.4) as follows

$$\text{minimize} \qquad J_h^M(q_h) = \frac{\tau}{2}\sum_{n=M-n_0}^{M} \int_\Omega q_h(x)|\nabla(v_h^n - z)|^2 \, \mathrm{d}x + \gamma N_h(q_h) \tag{3.1}$$

$$\text{subject to} \quad q_h \in K_h \quad \text{and} \quad v_h^n \equiv v_h^n(q_h) \in \mathring{V}_h \quad \text{satisfying} \quad v_h^0 = Q_h u_0(x) \text{ in } \Omega \tag{3.2}$$

$$\int_\Omega \partial_\tau v_h^n \phi_h \, \mathrm{d}x + \int_\Omega q_h(x)\nabla v_h^n \cdot \nabla\phi_h \, \mathrm{d}x = \int_\Omega \bar{f}^n \phi_h \, \mathrm{d}x \qquad \forall \phi_h \in \mathring{V}_h \tag{3.3}$$

for $n = 1, 2, \ldots, M$. The integer $n_0 \geqslant 0$ and the parameter σ are assumed to satisfy $\sigma = (n_0 + 1)\tau$ for simplicity. The term $N_h(q_h)$ is the discrete regularization defined by

$$N_h(q_h) = \int_\Omega |\nabla q_h|^2 \, \mathrm{d}x \qquad \text{or} \qquad N_h(q_h) = \int_\Omega \sqrt{|\nabla q_h|^2 + \delta(h)} \, \mathrm{d}x \quad (3.4)$$

corresponding to the continuous forms of $N(q)$. Here $\delta(h)$ is any positive function satisfying $\lim_{h\to 0} \delta(h) = \delta(0) = 0$, and its role is to smooth the non-differentiable function $|\cdot|$. The operator Q_h used in (3.2) is the L^2-projection from $L^2(\Omega)$ onto $\mathring{V}_h$, which is defined by

$$\int_\Omega Q_h v \phi \, \mathrm{d}x = \int_\Omega v\phi \, \mathrm{d}x \qquad \forall v \in L^2(\Omega), \quad \phi \in \mathring{V}_h. \quad (3.5)$$

The operator Q_h can be replaced by some other computationally less expensive operators with similar approximation properties to the L^2-projection (cf Chan *et al* [3, 4]) including the finite element interpolant (if the initial value u_0 is continuous). This does not affect any of our later convergence results.

Let I_h: $C(\bar{\Omega}) \to V_h$ be the standard nodal value interpolant associated with V_h. Then for any $p > d = \dim(\Omega)$, we have (cf Ciarlet [6] and Xu [17])

$$\lim_{h\to 0} \|v - I_h v\|_{W^{1,p}(\Omega)} = 0 \qquad \forall v \in W^{1,p}(\Omega) \quad (3.6)$$

$$\lim_{h\to 0} \|v - Q_h v\|_{H_0^1(\Omega)} = 0 \qquad \forall v \in H_0^1(\Omega) \quad (3.7)$$

and for any $v \in H_0^1(\Omega)$ we have

$$\|Q_h v\|_{L^2(\Omega)} \leqslant C\|v\|_{L^2(\Omega)} \qquad \|\nabla Q_h v\|_{L^2(\Omega)} \leqslant C\|\nabla v\|_{L^2(\Omega)}. \quad (3.8)$$

Concerning the existence of the minimizers to the finite element problem (3.1)–(3.3), we can show (cf [9, 12]):

Theorem 3.1. *There exists at least a minimizer to the finite element problem (3.1)–(3.3).*

In our later convergence analysis of the finite element approximation, we will need the following two lemmas.

Lemma 3.1. *Let $v_h^n(q_h)$ be the solutions of the finite element system (3.2) and (3.3) corresponding to $q_h \in K_h$, then we have the following stability estimates*

$$\max_{1\leqslant n\leqslant M} \|v_h^n(q_h)\|_{L^2(\Omega)}^2 + \tau \sum_{n=1}^{M} \|\nabla v_h^n(q_h)\|_{L^2(\Omega)}^2 \leqslant C(\|u_0\|_{L^2(\Omega)}^2 + \|f\|_{L^2(Q_T)}^2) \quad (3.9)$$

$$\max_{1\leqslant n\leqslant M} \|\nabla v_h^n(q_h)\|_{L^2(\Omega)}^2 + \tau \sum_{n=1}^{M} \|\partial_\tau v_h^n(q_h)\|_{L^2(\Omega)}^2 \leqslant C(\|\nabla u_0\|_{L^2(\Omega)}^2 + \|f\|_{L^2(Q_T)}^2) \quad (3.10)$$

with C independent of q_h, h and τ.

Lemma 3.2. *For any sequence $\{q_h\}$ in K_h and some $q \in K$, if q_h converges to q in $L^1(\Omega)$ as h tends to 0, then*

$$\sum_{n=M-n_0}^{M} \tau \int_\Omega q_h(x) |\nabla(v_h^n(q_h) - z)|^2 \, \mathrm{d}x \to \int_{T-\sigma}^{T} \int_\Omega q(x) |\nabla(v(q) - z)|^2 \, \mathrm{d}x \, \mathrm{d}t$$

$$\text{as } \tau, h \to 0.$$

Proof. We shall use the notation v_h^n and v to denote $v_h^n(q_h)$ and $v(q)$, respectively, and

$$v^n = v(q; t_n) = v(q; n\tau) \qquad \text{for } 0 \leqslant n \leqslant M$$

$$\bar{v}^n = \bar{v}^n(q) = \frac{1}{\tau}\int_{t^{n-1}}^{t^n} v(q; t)\,\mathrm{d}t \qquad \text{for } 1 \leqslant n \leqslant M \qquad \bar{v}^0 = \bar{v}^0(q) = u_0.$$

Taking $\phi = \tau^{-1}\phi_h$ in (2.4), then integrating over $[t^{n-1}, t^n]$ and subtracting it from (3.3) yields

$$\int_\Omega \partial_\tau(v_h^n - v^n)\phi_h\,\mathrm{d}x + \frac{1}{\tau}\int_{t^{n-1}}^{t^n}\int_\Omega q_h\nabla(v_h^n - v)\cdot\nabla\phi_h\,\mathrm{d}x\,\mathrm{d}t$$
$$= \frac{1}{\tau}\int_{t^{n-1}}^{t^n}\int_\Omega (q - q_h)\nabla v\cdot\nabla\phi_h\,\mathrm{d}x\,\mathrm{d}t.$$

Letting $\eta_h^n = v_h^n - Q_h\bar{v}^n$, and taking $\phi_h = \tau\eta_h^n$ in the above equation gives

$$\begin{aligned}
&\frac{1}{2}\|\eta_h^n\|_{L^2(\Omega)}^2 - \frac{1}{2}\|\eta_h^{n-1}\|_{L^2(\Omega)}^2 + \alpha_1\tau\|\nabla\eta_h^n\|_{L^2(\Omega)}^2 \\
&\qquad \leqslant \tau\int_\Omega \partial_\tau(v^n - Q_h\bar{v}^n)\eta_h^n\,\mathrm{d}x + \int_{t^{n-1}}^{t^n}\int_\Omega (q - q_h)\nabla v\cdot\nabla\eta_h^n\,\mathrm{d}x\,\mathrm{d}t \\
&\qquad\quad + \int_{t^{n-1}}^{t^n}\int_\Omega q_h\nabla(v - Q_h\bar{v}^n)\cdot\nabla\eta_h^n\,\mathrm{d}x\,\mathrm{d}t \\
&\qquad \equiv\, : (\mathrm{I})_1 + (\mathrm{I})_2 + (\mathrm{I})_3.
\end{aligned} \tag{3.11}$$

Summing the above equation over $n = 1, 2, \ldots, k \leqslant M$, we obtain

$$\frac{1}{2}\|\eta_h^k\|_{L^2(\Omega)}^2 - \frac{1}{2}\|\eta_h^0\|_{L^2(\Omega)}^2 + \alpha_1\tau\sum_{n=1}^k\|\nabla\eta_h^n\|_{L^2(\Omega)}^2 \leqslant \sum_{n=1}^k(\mathrm{I})_1 + \sum_{n=1}^k(\mathrm{I})_2 + \sum_{n=1}^k(\mathrm{I})_3. \tag{3.12}$$

We next estimate $(\mathrm{I})_1$, $(\mathrm{I})_2$ and $(\mathrm{I})_3$. First for $(\mathrm{I})_1$, from the definition of Q_h and the following formula, which holds for any sequences $\{a_n\}$ and $\{b_n\}$

$$\sum_{n=1}^k(a_n - a_{n-1})b_n = a_kb_k - a_0b_0 - \sum_{n=1}^k a_{n-1}(b_n - b_{n-1}) \tag{3.13}$$

we have

$$\begin{aligned}
\sum_{n=1}^k(\mathrm{I})_1 &= \int_\Omega (v^k - \bar{v}^k)\eta_h^k\,\mathrm{d}x - \tau\sum_{n=1}^k\int_\Omega (v^{n-1} - \bar{v}^{n-1})\partial_\tau\eta_h^n\,\mathrm{d}x \\
&\leqslant \sqrt{\tau}\left\{\int_{t^{k-1}}^{t^k}\|v_t\|_{L^2(\Omega)}^2\,\mathrm{d}t\right\}^{1/2}\|\eta_h^k\|_{L^2(\Omega)} \\
&\quad + \tau\left\{\int_0^T\|v_t\|_{L^2(\Omega)}^2\,\mathrm{d}t\right\}^{1/2}\left\{\tau\sum_{n=1}^k\|\partial_\tau\eta_h^n\|_{L^2(\Omega)}^2\right\}^{1/2} \leqslant C\sqrt{\tau}
\end{aligned}$$

where we have used the stability estimates (3.9) and (3.10) and the property of Q_h.

The estimate of $(\mathrm{I})_2$ can be carried out easily using Young's inequality

$$\sum_{n=1}^k(\mathrm{I})_2 \leqslant \frac{\alpha_1}{2}\tau\sum_{n=1}^k\|\nabla\eta_h^n\|_{L^2(\Omega)}^2 + \frac{\alpha_2}{\alpha_1}\int_0^T\int_\Omega |q - q_h||\nabla v|^2\,\mathrm{d}x\,\mathrm{d}t.$$

Finally, we can decompose $(\mathrm{I})_3$ into

$$(\mathrm{I})_3 = \int_{t^{n-1}}^{t^n}\int_\Omega q_h\nabla(v - Q_hv)\cdot\nabla\eta_h^n\,\mathrm{d}x\,\mathrm{d}t + \int_{t^{n-1}}^{t^n}\int_\Omega q_h\nabla Q_h(v - \bar{v}^n)\cdot\nabla\eta_h^n\,\mathrm{d}x\,\mathrm{d}t$$

and then by applying Young's inequality and the property of Q_h, this yields

$$\sum_{n=1}^{k}(\mathrm{I})_3 \leqslant \frac{\alpha_1}{4}\tau\sum_{n=1}^{k}\|\nabla\eta_h^n\|_{L^2(\Omega)}^2 + 2\frac{\alpha_2}{\alpha_1}\int_0^T\int_\Omega q_h|\nabla(v-Q_hv)|^2\,\mathrm{d}x\,\mathrm{d}t$$
$$+2\frac{\alpha_2^2}{\alpha_1}\int_\Omega\sum_{n=1}^{k}\int_{t^{n-1}}^{t^n}|\nabla(v-\bar{v}^n)|^2\,\mathrm{d}t\,\mathrm{d}x.$$

Now using the above estimates on $(\mathrm{I})_1$, $(\mathrm{I})_2$ and $(\mathrm{I})_3$ and the properties of the averaging function (cf [15, ch 6]) we derive that as $\tau, h\to 0$

$$\max_{1\leqslant n\leqslant M}\|\eta_h^n\|_{L^2(\Omega)}^2\to 0 \qquad\text{and}\qquad \tau\sum_{n=1}^{M}\|\nabla\eta_h^n\|_{L^2(\Omega)}^2\to 0. \tag{3.14}$$

By means of the results in (3.14) and the relation

$$v_h^n-\bar{v}^n=(v_h^n-Q_h\bar{v}^n)+(Q_h\bar{v}^n-\bar{v}^n)$$

we obtain immediately the convergence

$$\max_{1\leqslant n\leqslant M}\|v_h^n-\bar{v}^n\|_{L^2(\Omega)}^2\to 0 \qquad\text{and}\qquad \tau\sum_{n=1}^{M}\|\nabla(v_h^n-\bar{v}^n)\|_{L^2(\Omega)}^2\to 0 \tag{3.15}$$

as $\tau, h\to 0$. Finally, we are ready to show the desired convergence in lemma 3.2. By the convergence of Q_h and the boundedness of v_h^n and q_h, it suffices to prove

$$I_h^M - I_h^\sigma \equiv \tau\sum_{n=M-n_0}^{M}\int_\Omega q_h|\nabla(v_h^n-z_h)|^2\,\mathrm{d}x - \int_{T-\sigma}^{T}\int_\Omega q|\nabla(v-z_h)|^2\,\mathrm{d}x\,\mathrm{d}t\to 0$$

as $h,\tau\to 0$, where $z_h = Q_h z$. We can then rewrite it in the following form

$$I_h^M - I_h^\sigma = \sum_{n=M-n_0}^{M}\int_{t^{n-1}}^{t^n}\int_\Omega q_h(|\nabla(v_h^n-z_h)|^2-|\nabla(v-z_h)|^2)\,\mathrm{d}x\,\mathrm{d}t$$
$$+\int_{T-\sigma}^{T}\int_\Omega(q_h-q)|\nabla(v-z_h)|^2\,\mathrm{d}x\,\mathrm{d}t \equiv: (\mathrm{II})_1+(\mathrm{II})_2.$$

For $(\mathrm{II})_2$, we know by the Lebesgue dominant convergence theorem and (3.7) that

$$|(\mathrm{II})_2|\leqslant 2\int_{T-\sigma}^{T}\int_\Omega|q_h-q||\nabla(v-z)|^2\,\mathrm{d}x\,\mathrm{d}t + 2\int_{T-\sigma}^{T}\int_\Omega|q_h-q||\nabla(z-z_h)|^2\,\mathrm{d}x\,\mathrm{d}t\to 0.$$

For $(\mathrm{II})_1$, we obtain by the Cauchy–Schwarz inequality and (3.9) that

$$|(\mathrm{II})_1|\leqslant\alpha_2\left(\sum_{n=M-n_0}^{M}\int_{t^{n-1}}^{t^n}\|\nabla(v_h^n-v)\|_{L^2(\Omega)}^2\,\mathrm{d}t\right)^{1/2}$$
$$\cdot\left(\sum_{n=M-n_0}^{M}\int_{t^{n-1}}^{t^n}\|\nabla(v_h^n+v-2z_h))\|_{L^2(\Omega)}^2\,\mathrm{d}t\right)^{1/2}$$
$$\leqslant C\left(\sum_{n=M-n_0}^{M}\int_{t^{n-1}}^{t^n}\|\nabla(v_h^n-v)\|_{L^2(\Omega)}^2\,\mathrm{d}t\right)^{1/2}$$

which converges to zero by using

$$\|\nabla(v_h^n-v)\|_{L^2(\Omega)}\leqslant\|\nabla(v_h^n-\bar{v}^n)\|_{L^2(\Omega)}+\|\nabla(v-\bar{v}^n)\|_{L^2(\Omega)}$$

and the convergence in (3.15). Thus we have proved that $I_h^M - I_h^\sigma\to 0$ as $\tau, h\to 0$. □

Finally, we can state the following convergence theorem about the finite element problem (3.1)–(3.3) (see [12] for details).

Theorem 3.2. *Let $\{q_h^*\}_{h>0}$ be a sequence of minimizers to the discrete minimization problem (3.1)–(3.3). Then each subsequence of $\{q_h^*\}_{h>0}$ has a subsequence converging to a minimizer of the continuous problem (2.2)–(2.4).*

4. Armijo algorithm

To solve the discretized constrained minimization of $J_h^M(\cdot)$ over K_h in (3.1)–(3.3), we reduce it into a sequence of unconstrained minimizations of the following functional $J_h^M(\varepsilon;\cdot)$ over the entire space V_h with $\varepsilon > 0$

$$\text{minimize } J_h^M(\varepsilon; q_h) = J_h^M(q_h) + \frac{1}{\varepsilon}\int_\Omega P(q_h)(x)\,\mathrm{d}x \tag{4.1}$$

subject to $q_h \in V_h$ and $v_h^n \equiv v_h^n(q_h) \in \mathring{V}_h$ satisfying

$$\int_\Omega \partial_\tau v_h^n \phi_h\,\mathrm{d}x + \int_\Omega q_h \nabla v_h^n \cdot \nabla\phi_h\,\mathrm{d}x = \int_\Omega \bar{f}^n \phi_h\,\mathrm{d}x \qquad \forall \phi_h \in \mathring{V}_h \tag{4.2}$$

for $n = 1, 2, \ldots, M$, with $v_h^0 = Q_h u_0$ and $P(q_h)$ defined by

$$P(q_h)(x) = \tfrac{1}{2}(q_h(x) - \alpha_2)_+^2 + \tfrac{1}{2}(\alpha_1 - q_h(x))_+^2.$$

We now formulate the Armijo algorithm for solving the problem (4.1)–(4.2). To do so, we need to evaluate the Gateaux derivatives of the following cost functional

$$J_h^M(\varepsilon; q_h) = \frac{\tau}{2}\sum_{n=M-n_0}^{M}\int_\Omega q_h|\nabla(v_h^n(q_h) - z_h)|^2\,\mathrm{d}x + \gamma N_h(q_h) + \frac{1}{\varepsilon}\int_\Omega P(q_h)\,\mathrm{d}x$$

where $z_h = I_h z$ is the finite element interpolant of the measured data z. First, for each $q_h \in K_h$, we can easily get the Gateaux derivative $v_h^n(q_h)' : V_h \to \mathring{V}_h$ of the function $v_h^n(\cdot)$, which satisfies $v_h^0(q_h)' p_h = 0$ and $v_h^n(q_h)' p_h \in \mathring{V}_h$ $(n \geqslant 1)$ for any $p_h \in V_h$ and

$$\int_\Omega \partial_\tau(v_h^n(q_h)' p_h)\phi_h\,\mathrm{d}x + \int_\Omega q_h \nabla(v_h^n(q_h)' p_h)\cdot\nabla\phi_h\,\mathrm{d}x$$
$$= -\int_\Omega p_h \nabla v_h^n(q_h)\cdot\nabla\phi_h\,\mathrm{d}x \qquad \forall\phi_h \in \mathring{V}_h. \tag{4.3}$$

To compute the derivative of the first term in the cost function, i.e.

$$J_1(q_h) = \frac{\tau}{2}\sum_{n=M-n_0}^{M}\int_\Omega q_h|\nabla(v_h^n(q_h) - z_h)|^2\,\mathrm{d}x$$

we introduce the discrete adjoint state system to (4.2) (cf Lions [11]).

Thus, we need to find w_h^n, $n = M, M-1, \ldots, 1, 0$, such that $w_h^M = 0$ and w_h^n $(n < M)$ satisfy

$$-\int_\Omega \partial_\tau w_h^n \phi_h\,\mathrm{d}x + \int_\Omega q_h \nabla w_h^{n-1}\cdot\nabla\phi_h\,\mathrm{d}x = \mu_n\int_\Omega q_h\nabla(v_h^n(q_h) - z_h)\cdot\nabla\phi_h\,\mathrm{d}x$$

where $\mu_n = 1$ for $M - n_0 \leqslant n \leqslant M$ and $\mu_n = 0$ otherwise. Then we can derive (see [12] for details)

$$J_1(q_h)' p_h = -\tau\sum_{n=1}^{M}\int_\Omega p_h\nabla v_h^n(q_h)\cdot\nabla w_h^{n-1}\,\mathrm{d}x + \frac{\tau}{2}\sum_{n=M-n_0}^{M}\int_\Omega p_h|\nabla(v_h^n(q_h) - z_h)|^2)\,\mathrm{d}x.$$

For the functionals $J_2(q_h) = N_h(q_h)$ and $J_3(q_h) = \int_\Omega P(q_h)\,\mathrm{d}x$, we can easily obtain their derivatives as follows

$$J_2(q_h)' p_h = \begin{cases} \displaystyle\int_\Omega \frac{\nabla q_h \cdot \nabla p_h}{\sqrt{|\nabla q_h|^2 + \delta(h)}}\,\mathrm{d}x & \text{for } N_h(q_h) = \displaystyle\int_\Omega \sqrt{|\nabla q_h|^2 + \delta(h)}\,\mathrm{d}x \\ 2\displaystyle\int_\Omega \nabla q_h \cdot \nabla p_h\,\mathrm{d}x & \text{for } N_h(q_h) = \displaystyle\int_\Omega |\nabla q_h|^2\,\mathrm{d}x \end{cases}$$

and

$$J_3(q_h)' p_h = \int_\Omega P'(q_h) p_h\,\mathrm{d}x.$$

With these derivatives, we can represent the derivative of $J_h^M(\varepsilon; q_h)$ by

$$\begin{aligned}(J_h^M(\varepsilon; q_h))' p_h = {} & -\tau \sum_{n=1}^{M} \int_\Omega p_h \nabla v_h^n(q_h) \cdot \nabla w_h^{n-1}\,\mathrm{d}x \\ & + \frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_\Omega p_h |\nabla(v_h^n(q_h) - z_h)|^2\,\mathrm{d}x + \gamma J_2(q_h)' p_h + \frac{1}{\varepsilon}\int_\Omega P'(q_h) p_h\,\mathrm{d}x.\end{aligned}$$

Now we are able to present the Armijo algorithm.

Armijo algorithm. Given a step size control fraction $\beta \in (0.5, 1)$, a penalty constant $\varepsilon \in (0, 1)$, and an initial guess $q_h^0 \in K_h$. Set $j = 0$.

1. Compute $v_h^n \equiv v_h^n(q_h^j) \in \mathring{V}_h$ by solving

$$\begin{aligned} & v_h^0 = Q_h u_0(x) \qquad \text{in } \Omega \\ & \int_\Omega \partial_\tau v_h^n \phi_h\,\mathrm{d}x + \int_\Omega q_h^j \nabla v_h^n \cdot \nabla\phi_h\,\mathrm{d}x = \int_\Omega f\phi_h\,\mathrm{d}x \qquad \forall \phi_h \in \mathring{V}_h.\end{aligned}$$

Compute $w_h^M = 0$ and $w_h^n \in \mathring{V}_h$, $n = M-1, \ldots, 1, 0$ by solving

$$-\int_\Omega \partial_\tau w_h^n \phi_h\,\mathrm{d}x + \int_\Omega q_h^j \nabla w_h^{n-1} \cdot \nabla\phi_h\,\mathrm{d}x = \mu_n \int_\Omega q_h^j \nabla(v_h^n(q_h^j) - z_h) \cdot \nabla\phi_h\,\mathrm{d}x.$$

2. Compute the components of $(J_h^M(\varepsilon; q_h^j))'$ corresponding to the lth basis ϕ_l by

$$\begin{aligned} g_l = {} & -\tau \sum_{n=1}^{M} \int_\Omega \phi_l \nabla v_h^n(q_h^j) \cdot \nabla w_h^{n-1}\,\mathrm{d}x + \frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_\Omega \phi_l |\nabla(v_h^n(q_h^j) - z_h)|^2\,\mathrm{d}x + \gamma J_2(q_h^j)'\phi_l \\ & + \frac{1}{\varepsilon}\int_\Omega P'(q_h^j)\phi_l\,\mathrm{d}x.\end{aligned}$$

Set $g_h^j = \sum_{l=1}^N g_l \phi_l$.

3. Compute the gradient norm $e_j = (h^d \sum_{l=1}^N g_l^2)^{1/2}$.
4. Set $\lambda = 1$.
   (i) Compute $\Delta = J_h^M(\varepsilon; q_h^j + \lambda g_h^j) - J_h^M(\varepsilon; q_h^j) + \frac{1}{2}\lambda\,\mathrm{e}_j^2$.
   (ii) If $\Delta \leqslant 0$, go to (iii); otherwise set $\lambda = \beta\lambda$, go to (i).
   (iii) Compute $q_h^{j+1} = q_h^j + \lambda g_h^j$. If $\|q_h^{j+1} - q_h^j\| <$ tolerance, stop; otherwise set $j = j + 1$, go to step 1.

5. Numerical experiments

We now show some numerical experiments on the proposed methods for parameter identifications. The test problem is

$$\begin{aligned}
&\frac{\partial u}{\partial t} - \nabla\cdot(q(x)\nabla u) = f(x,t) \qquad (x,t)\in\Omega\times(0,T)\\
&u(x,0) = u_0(x) \qquad x\in\Omega\\
&u(x,t) = 0 \qquad (x,t)\in\partial\Omega\times(0,T)
\end{aligned} \tag{5.1}$$

where $\Omega = (0,1)$ and $T = 1$. Unless otherwise specified, all the numerical experiments use the Armijo algorithm with H^1-regularization. The numerical results using BV-regularization work equally as well as the H^1-regularization for both the smooth and discontinuous parameters.

Most parameters related in the algorithm are attached in each figure. The error shown is the relative L^2-norm error between the exact parameter $q(x)$ to be identified and the computed parameter q_h. The penalty parameter ε and the step size control parameter β in the Armijo algorithm are taken to be 10^{-5} and $3/4$, respectively. The finite element mesh size h and the time-step size τ are both taken to be $1/100$. The lower and upper bounds α_1 and α_2 in the constrained set K are taken to be 0.5 and 20.0. The constant σ is chosen to be one time-step size τ, so $n_0 = 0$.

Example 1. We take the observed data z as

$$z = u(x,1) = \sin(2\pi x)$$

and the exact solution as

$$u(x,t) = \mathrm{e}^{\sin(\pi t)}\sin(2\pi x)$$

but the identifying coefficient $q(x)$ as

$$q(x) = 3 + 2x^2 - 2\sin(2\pi x).$$

The function $f(x,t)$ is then computed through equation (5.1) using $u(x,t)$ and $q(x)$. Figure 1 shows the exact solution $q(x)$ (the full curve) and the numerically identified solution $q_h(x)$ (open circles). Note that the exact coefficient function $q(x)$ is very smooth in the example and the finite element identified solution $q_h(x)$ is nearly indistinguishable from the true solution $q(x)$. The initial guess q_h^0 is taken to be the constant 10.0 everywhere, which is not a good initial guess at all, but the numerical method still converges steadily and the approximation appears to be quite accurate.

When the above observed data z has the following noised form

$$z^\delta(x) = z(x) + \delta\sin(3\pi x)$$

the numerical result is shown in figure 2 with the noise parameter $\delta = 1\%$, and the output least square norm in (2.2) is taken to be the L^2-norm instead of the original energy-norm. The numerical recovered coefficient matches the true parameter $q(x)$ well, except for some oscillations around two singular points $x = 0.25$ and 0.75 where $u_x = 0$.

Example 2. We take the observed data z as

$$z = u(x,1) = \sin(\pi x)$$

and the exact solution as

$$u(x,t) = \mathrm{e}^{\sin(\pi t)}\sin(\pi x)$$

and the identifying coefficient $q(x)$ as the highly discontinuous function

$$q(x) = \begin{cases} 2 - x & x \in [0, 0.3] \\ 1 - x + 4x^2 & x \in (0.3, 0.7) \\ 3 & x \in [0.7, 1]. \end{cases}$$

The function $f(x, t)$ is computed through equation (5.1) using $u(x, t)$ and $q(x)$, and is also discontinuous. Figure 3 shows the exact solution $q(x)$ (the broken curve) and the numerically identified solution $q_h(x)$ (the full curve). We note that although the exact coefficient function $q(x)$ is highly discontinuous, the finite element identified solution $q_h(x)$ matches very well with $q(x)$ except for small oscillations around two discontinuous points $x = 0.3$ and $x = 0.7$. The initial guess q_h^0 can be taken to be much worse than $q_h^0 = 4.0$ here, say $q_h^0 = 10.0$ or 20.0, the method converges still very stably and gives the same accurate result as $q_h^0 = 4.0$.

When the above observed data z has the following noised form

$$z^\delta(x) = z(x) + \delta \sin(3\pi x)$$

the numerical identified result is shown in figure 4 with the output least squares norm in (2.2) taken to be the L^2-norm instead of the original energy-norm. The numerical identified parameter appears to be reasonably good considering that the identifying parameter is highly discontinuous. When the noise parameter δ goes over 1%, the error exceeds 0.05 but the numerical solution still keeps a very good shape of the true parameter function.

We also plot the numerical experiment with BV-regularization for comparison in figure 5. Here we take the smoothing parameter $\delta(h)$ in $N_h(q_h)$ to be 0.2. The numerical solution is close to that with H^1-regularization plotted in figure 3.

Example 3. We take the observed data z as

$$z = u(x, 1) = \sin(\pi x)$$

and the exact solution as

$$u(x, t) = \mathrm{e}^{\sin(\pi t)} \sin(\pi x)$$

but the identifying coefficient $q(x)$ as the highly discontinuous and oscillating function

$$q(x) = \begin{cases} 2 & x \in [0, 0.3] \\ 4 & x \in (0.3, 0.6) \\ 2 + \sin(10\pi x) & x \in [0.6, 1]. \end{cases}$$

The function $f(x, t)$ is computed through equation (5.1) using $u(x, t)$ and $q(x)$, and is also discontinuous and oscillating. Figure 6 shows the exact coefficient $q(x)$ (the broken curve) to be recovered and the numerically identified solution $q_h(x)$ (the full curve). Note that the exact parameter function $q(x)$ is highly discontinuous and oscillating in this example, but the finite element identified solution matches well with $q(x)$ except for small perturbations around two discontinuous points $x = 0.3$ and $x = 0.6$.

Example 4. We take the observed data z to be the hat function

$$z = u(x, 1) = 0.5 - |x - 0.5|$$

and the exact solution as

$$u(x, t) = \mathrm{e}^{\sin(\pi t)}(0.5 - |x - 0.5|)$$

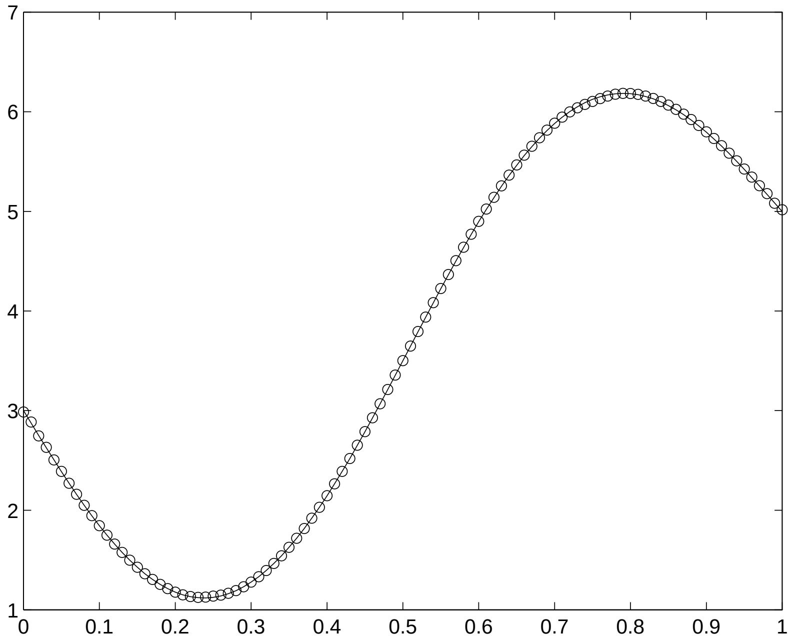

Figure 1. $q_h^{(0)} = 10.0$, $\gamma = 10^{-7}$, error $= 0.0013$.

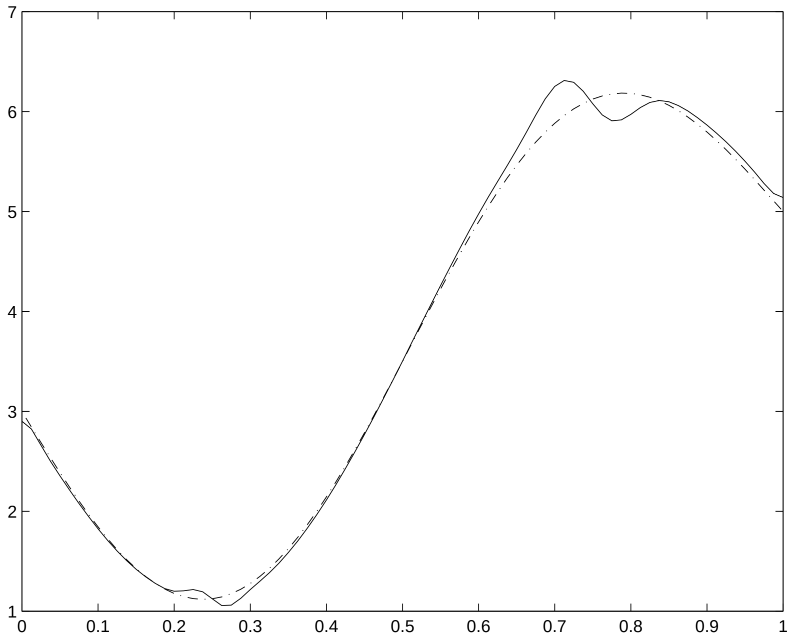

Figure 2. $q_h^{(0)} = 4.0$, $\gamma = 10^{-10}$, $\delta = 1\%$, error $= 0.0276$.

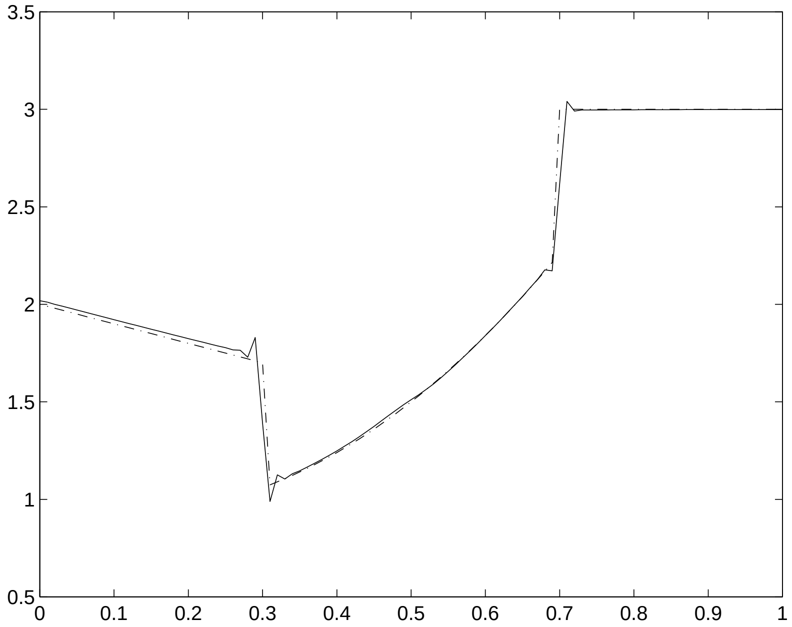

Figure 3. $q_h^{(0)} = 4.0$, $\gamma = 10^{-7}$, error $= 0.024$.

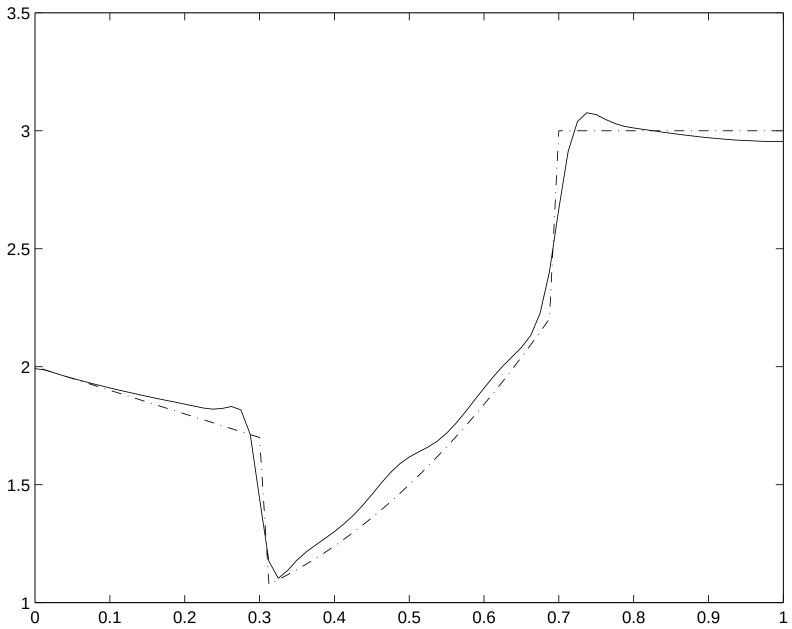

Figure 4. $q_h^{(0)} = 4.0$, $\gamma = 10^{-10}$, $\delta = 0.5\%$, error $= 0.034$.

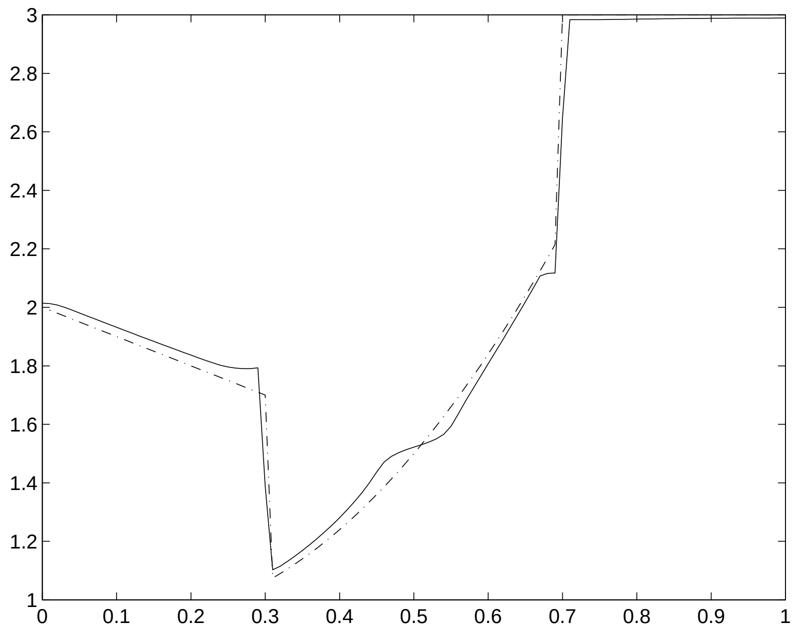

Figure 5. BV-norm, $q_h^{(0)} = 4.0$, $\gamma = 10^{-5}$, error $= 0.027$.

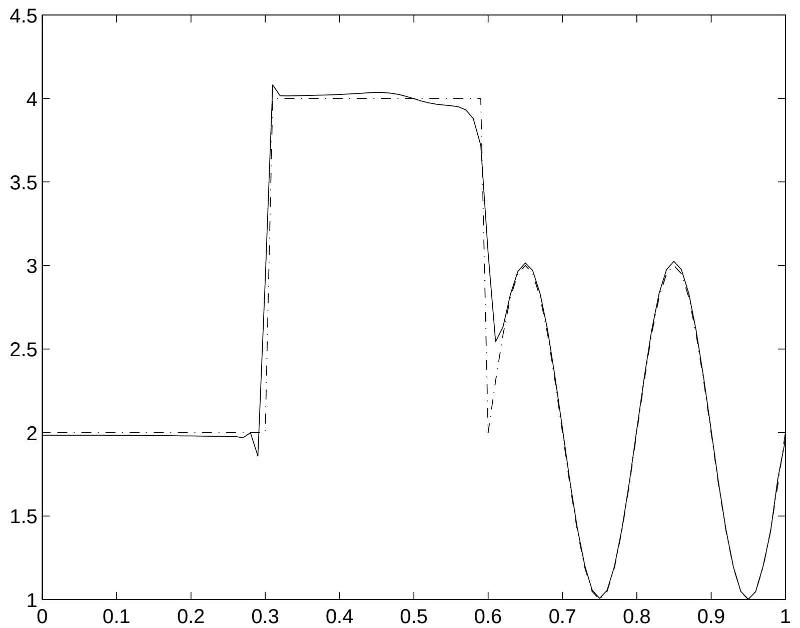

Figure 6. $q_h^{(0)} = 3.0$, $\gamma = 10^{-7}$, error $= 0.054$.

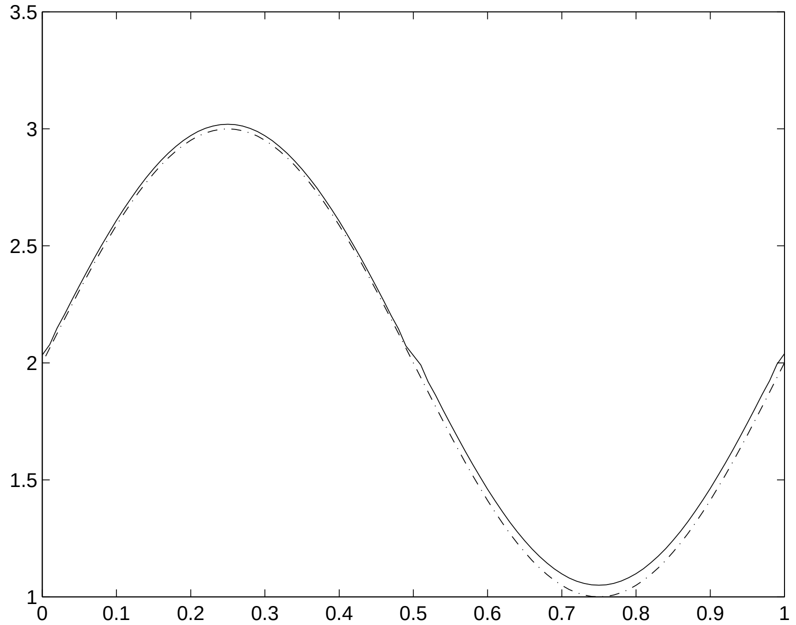

Figure 7. $q_h^{(0)} = 3.0$, $\gamma = 10^{-8}$, error $= 0.018$.

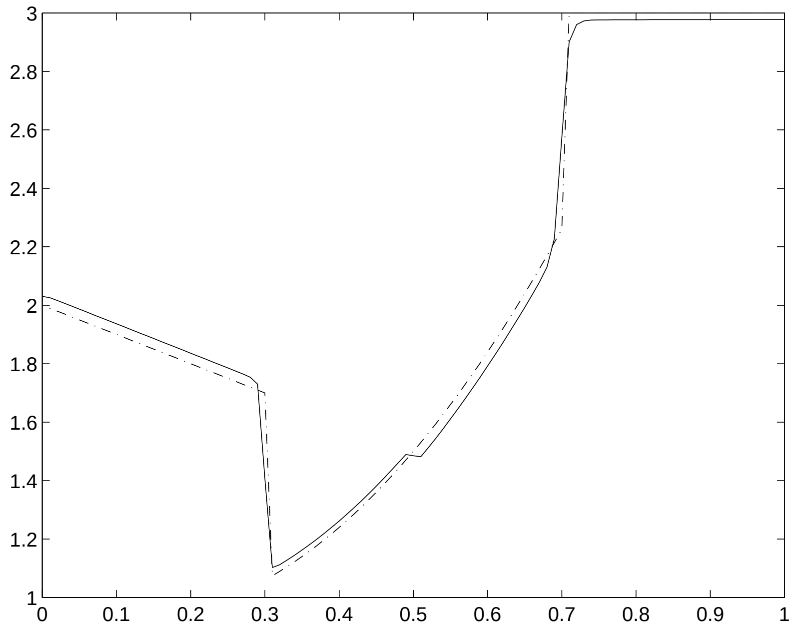

Figure 8. $q_h^{(0)} = 4.0$, $\gamma = 10^{-7}$, error $= 0.026$.

but a smooth coefficient $q(x) = 2 + \sin(2\pi x)$. Note that the observed data have lowest regularity in this case, i.e. $z \in H^1(\Omega)$, but the exact parameter $q(x)$ is still smooth. Figure 7 shows the exact solution $q(x)$ (the broken curve) and the numerically identified solution $q_h(x)$ (the full curve). We see that the lack of regularity on the observed data does not affect our numerical method too much, which gives a very satisfactory approximation.

Example 5. We take the observed data z to be the hat function

$$z = u(x, 1) = 0.5 - |x - 0.5|$$

and the exact solution as

$$u(x, t) = \mathrm{e}^{\sin(\pi t)}(0.5 - |x - 0.5|)$$

but a highly discontinuous coefficient $q(x)$ as

$$q(x) = \begin{cases} 2 - x & x \in [0, 0.3] \\ 1 - x + 4x^2 & \in (0.3, 0.7) \\ 3 & x \in [0.7, 1]. \end{cases}$$

Note that in this case not only does the observed data have a lowest regularity, i.e. $z \in H^1(\Omega)$, but the identifying parameter $q(x)$ is also highly discontinuous. Figure 8 shows the exact solution $q(x)$ (the broken curve) and the numerically identified solution $q_h(x)$ (the full curve). We can see that the lack of regularity on both the observed data and the parameter to be identified does not affect our numerical method too much and the numerical location of the discontinuity and singularity points ($x = 0.3, 0.5, 0.7$) is very accurate.

Acknowledgment

The work of one of the authors (JZ) was partially supported by Hong Kong RGC grant No CUHK 338/96E.

References

[1] Bank H T and Kunisch K 1989 *Estimation Techniques for Distributed Parameter System* (Boston, MA: Birkhäuser)

[2] Cannon J 1984 *The One-dimensional Heat Equation (Encyclopedia of Mathematics 23)* (London: Addison-Wesley)

[3] Chan T, Smith B and Zou J 1996 Overlapping Schwarz methods on unstructured meshes using non-matching coarse grids *Num. Math.* **73** 149–67

[4] Chan T and Zou J 1994 Additive Schwarz domain decomposition methods for elliptic problems on unstructured meshes *Num. Algorithms* **8** 329–46

[5] Chen Z and Zou J 1997 An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems *Technical Report* 97-06 (116), Department of Mathematics, The Chinese University of Hong Kong

[6] Ciarlet P 1991 Basic error estimates for elliptic problems *Handbook of Numerical Analysis* vol II, ed P G Ciarlet and J-L Lions (Amsterdam: North-Holland) pp 17–352

[7] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)

[8] Guenther R, Hudspeth R, McDougal W and Gerlach J 1985 Remarks on parameter identification I *Num. Math.* **47** 355–61

[9] Gutman S 1990 Identification of discontinuous parameters in flow equations *SIAM J. Control Optim.* **28** 1049–60

[10] Ito K and Kunisch K 1996 Augmented Lagrangian-SQP-methods in Hilbert spaces and applications to control in the coefficients problems *SIAM J. Control Optim.* **6** 96–125

[11] Lions J L 1971 *Optimal Control of Systems Governed by Partial Differential Equations* (New York: Springer)
[12] Keung Y L and Zou J 1997 Identifying parameters in parabolic systems by finite element methods *Technical Report* 97-16 (126), Department of Mathematics, The Chinese University of Hong Kong
[13] Kunisch K and Ring W 1993 Regularization of nonlinear illposed problems with closed operators *Num. Funct. Anal. Optimiz.* **14** 389–404
[14] Kunisch K and White L 1985 The parameter estimation problem for parabolic equations and discontinuous observation operators *SIAM J. Control Optim.* **23** 900–27
[15] Royden H 1988 *Real Analysis* 3rd edn (New York: Macmillan)
[16] Seid Y F and Zou J 1997 Finite element simulations in parameter identifications *Technical Report* 97-10 (120), Department of Mathematics, The Chinese University of Hong Kong
[17] Xu J 1989 Theory of multilevel methods *PhD Thesis* Cornell University