Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Computational and Applied Mathematics 193 (2006) 183–203
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Solving an inverse parabolic problem by optimization from final measurement data

Qun Chen[1], Jijun Liu*

Department of Mathematics, Southeast University, Nanjing 210096, PR China

Received 21 March 2005; received in revised form 11 June 2005

Abstract

We consider an inverse problem of reconstructing the coefficient q in the parabolic equation $u_t - \Delta u + q(x)u = 0$ from the final measurement $u(x, T)$, where q is in some subset of $L^1(\Omega)$. The optimization method, combined with the finite element method, is applied to get the numerical solution under some assumption on q. The existence of minimizer, as well as the convergence of approximate solution in finite-dimensional space, is proven. The new ingredient in this paper is that we do not need uniformly a priori bounds of H^1-norm on q. Numerical implementations are also presented.

MSC: 35R30; 35J05; 76Q05

Keywords: Inverse problem; Parabolic equation; Optimization; Finite element method; Convergence; Numerics

1. Introduction

Let $\Omega \subset \mathscr{R}^m$ $(m \geqslant 1)$ be a given bounded domain. Consider the heat conduction problem governed by

$$\begin{aligned} &u_t - \Delta u + q(x)u = 0, && (x,t) \in Q_T := \Omega \times (0,T), \\ &u(x,t) = 0, && (x,t) \in \partial\Omega \times (0,T), \\ &u(x,0) = u_0(x), && x \in \Omega, \end{aligned} \tag{1.1}$$

* Corresponding author. Tel.: +86 258 379 4934; fax: +86 258 379 2880.
E-mail addresses: qunchen@163.com (Q. Chen), jjliu@seu.edu.cn (J. Liu).
[1] This author now works at Nanjing University of Information Science and Technology.

0377-0427/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2005.06.003

with some initial temperature $u_0 \in H_0^1(\Omega)$. Then it follows from the standard theory on the parabolic equation [10] that, if q satisfies $c_0 \leqslant q(x) \leqslant \beta_0 < \infty$ a.e in Ω for two constants c_0, β_0, there exists a unique solution in $L^2((0,T), W_0^{1,2}(\Omega))$. We also denote by $u(q)(x,t)$ to show the dependence of solution on $q(x)$.

Physically speaking, this model describes the heat conduction procedure in a given homogeneous medium Ω. If the medium is inhomogeneous with some input source $f(x,t)$, then the equation should be written as

$$u_t - \nabla(p(x)\nabla u)u + q(x)u = f(x,t), \quad (x,t) \in \Omega.$$

The coefficient $p(x)$ represents the heat conduction property such as heat capacity, while $q(x)u(x,t)$ is also the heat source which depends both on location x and on temperature, except for the heat source $f(x,t)$. That is, $q(x)$ in fact describes the medium property of generating heat source or heat sink.

The tasks of inverse heat problems are detection of heat conduction properties of the medium from some information of the solution, i.e., the determination of the unknown coefficient(s) in the heat equation from some additional information about $u(x,t)$. For example, if $u(x,t)$ is given on $\partial\Omega$ for some time interval $[0,t_0]$, which is the so-called lateral overdetermination problem, then the inverse problems of determining the unknown coefficients have been studied thoroughly, see [5, Chapter 9] for single and many measurements.

However, there is another possibility to give the additional temperature for inverse heat problems. That is, the additional data is $u(x,t)$ given at some final time $t=T$ for all $x \in \Omega$, rather than $u(x,t)$ for $x \in \partial\Omega$ and all time. Due to this practical requirement, the inverse heat conduction problems using final temperature as inversion input data have been considered carefully, see [3,8] for determining $p(x)$ in the above equation with $q(x)=0$ from the measurement given $u(x,T)$. In this model, the heat conduction procedure is considered only of the input linear source $f(x,t)$, ignoring the nonlinear source $q(x)u(x,t)$ within the medium. In [6,7], the authors considered the determination of principal coefficient $a(x)$ in 1D equation

$$u_t - a(x)u_{xx} + b(x)u_x + c(x)u = f(x,t)$$

from final overdetermination data $u(x,t_0)$ under an optimization control framework. The existence of $a(x)$ and a well-posed algorithm are obtained. The uniqueness and stability of determining $a(x)$ in the parabolic equation

$$\partial_t u + \mathscr{A}u + a(x)u = 0$$

from the final measurement data is considered in [4], where zero initial condition and nonzero boundary condition $\mathscr{B}u = g$ in $\partial\Omega \times (0,T)$ are assumed. The determination of $q(x)$ by Hölder space method from additional information given at $t=T$ can also be found in [2]. Finally, the theoretical issues such as existence and uniqueness of coefficients inversion for parabolic equation are also studied in [5,11].

This paper considers the inversion scheme of determining $q(x)$ in system (1.1) from the final measurement

$$u(x,T) = z(x). \tag{1.2}$$

The particular difficulty of this inverse problem is that the information of $q(x)$ contained in $u(x,T)$ is very weak due to the exponentially decay of $u(x,T)$ with respect to T. On the other hand, since there is no uniqueness for this inverse problem with general input data u_0 and single measurement (1.2), the optimization technique should be applied to get some general solution.

The optimization technique is a classical tool to yield "general solution" for inverse problems without unique solution [9]. The basic idea is to restrict the solution under consideration to some compact set, and then take the minimizer of some cost functional as the general solution. However, the a priori compactness of admissible set means some smooth assumption on the exact solution, which is not suitable for many practical problems [3].

In this paper, by adjusting the penalty term for the cost functional, we prove the existence of minimizer in a relaxed admissible set for this functional corresponding to our inverse problem. Then the finite element method is applied to solve the optimization problem by the adjoint space technique. The convergence of approximate minimizer in a finite-dimensional space to the exact minimizer is proven. The motivation of this inversion scheme comes from the basic idea used in [8]. Finally, we present some numerical examples to show the validity of the inversion method.

2. Existence of minimizer for cost functional

In this section, we state the optimization version for the inverse problem of recovering $q(x)$ described by (1.1) and (1.2).

Introduce the admissible set

$$\mathscr{K} := \{q(x) \in L^1(\Omega) : \|q\|_{H^1} < \infty,\ 0 \leqslant q(x) \leqslant \beta_0 \text{ a.e. in } \Omega\}$$

for known constant $\beta_0 > 0$ and the functionals

$$J_1(q) := \frac{1}{2}\int_{T-\sigma}^{T} \mathrm{d}t \int_\Omega q(x)|u(q)(x,t) - z(x)|^2\,\mathrm{d}x,$$

$$J(q) := J_1(q) + \alpha \int_\Omega |\nabla q|^2\,\mathrm{d}x \tag{2.1}$$

for given small constant $\sigma > 0$ and regularizing parameter $\alpha > 0$, where $u(q)(x,t)$ solves (1.1) for $q \in \mathscr{K}$. The constant σ introduced here is used to restrict the property of $u(\cdot, t)$ nearing the final time T, which will be useful in the numerical scheme.

We give an explanation to the cost functional constructed here. In contrast to the obvious least-square formulation $\frac{1}{2}\int_{T-\sigma}^{T} \mathrm{d}t \int_\Omega |u(q)(x,t) - z(x)|^2\,\mathrm{d}x$, here we use $J_1(q)$. The motivation to this kind of functional comes from the consideration of nonlinearity of original inverse problem. It is obvious that the determination of q from $u(q)(x,T)$ is nonlinear due to the term $q(x)u(x,t)$ in the equation. So we hope to fully take into consideration for this nonlinearity in the optimization functional. For this purpose, the "weight function" $q(x)$ is introduced to the functional to match the nonlinear term $q(x)u(x,t)$ to some extent.

To make $J_1(q)$ a "true" cost functional, we should assume $q(x) \geqslant 0$ to keep $J_1(q)$ nonnegative. However, this requirement on $q(x)$ is not essential. For $q(x)$ with lower bound $c_0 < 0$, we can introduce the transform $v(x,t) = u(x,t)e^{c_0 t}$, which satisfies

$$v_t - \Delta v + (q(x) - c_0)v = 0.$$

So the same kind of inverse problem for function $v(x,t)$ is constituted with $q(x) - c_0 \geqslant 0$. We can use the optimization technique proposed in this paper to recover $q(x) - c_0$. That is, our optimization inversion scheme is essentially suitable for $q(x)$ with (negative) lower bound. In other words, the non-negative assumption on q in the admissible set $\mathscr{K}$ is not essential for our inversion scheme as well as the convergence property.

The uniqueness for the inverse problem constituted by (1.1) and (1.2) is an important issue, which means whether the information $u(x,T)$ is enough for determining $q(x)$. With the uniqueness result, we can expect the minimizing sequence from optimization indeed converges to the unique classical solution at least in some general sense. To the authors' knowledge, however, these kinds of results seem to be open for general $u_0(x)$. However, we can prove the following uniqueness for non-negative initial temperature.

Theorem 2.1. *Let the initial temperature $u_0(x) \geqslant 0$, $u_0(x) \not\equiv 0$. Assume that $q_1, q_2 \in \mathscr{K}$. Then we get $q_1 = q_2$ if $u_1(x,T) = u_2(x,T)$.*

The proof is lengthy but completely using the same arguments as those in [11]. We omit the details here.

Now we begin to consider the optimization problem.

Firstly, we assert that the functional $J_1(q)$ is of some continuous property in $\mathscr{K}$ in the following sense.

Lemma 2.2. *If a sequence $\{q_n\} \subset \mathscr{K}$ with the convergence $q_n \to q^* \in \mathscr{K}$ in $L^1(\Omega)$, then there exists a subsequence $\{q_{n_k}\} \subset \{q_n\}$ such that*

$$\lim_{n_k \to \infty} J_1(q_{n_k}) = J_1(q^*). \tag{2.2}$$

Proof. The proof is standard and similar to that used in [8] for the reconstruction of $p(x)$. We only give the outlines here.

Step 1: By taking the weak form of equation in (1.1) for $q = q^n$ and choosing the test function as $u(q_n)(\cdot,t)$, we can prove that there exists a subsequence of $\{u(q_n) : n = 1, \ldots\} \subset L^2((0,T); H_0^1(\Omega))$ such that

$$u(q_{n_k})(x,t) \rightharpoonup u^*(x,t) \in L^2((0,T); H_0^1(\Omega)) \tag{2.3}$$

for some function $u^*(x,t)$ as $n_k \to \infty$.

Step 2: Prove $u^*(x,t) = u(q^*)(x,t)$.

Firstly, we take $q = q_{n_k}$ in the weak form of equation in (1.1) and multiply both sides by $\eta(t) \in C^1[0,T]$ with $\eta(T) = 0$ for any $\phi(x) \in H_0^1(\Omega)$. By integrating the resulted equation in $[0,T]$ for t, we are led to

$$\int_0^T \left[\int_\Omega (u^*)_t \phi \, \mathrm{d}x + \int_\Omega (\nabla u^* \cdot \nabla \phi) \, \mathrm{d}x + \int_\Omega q^* u^* \phi \, \mathrm{d}x \right] \eta(t) \, \mathrm{d}t = 0 \tag{2.4}$$

for all $\phi(x) \in H_0^1(\Omega)$ and $\eta(t) \in \mathscr{B} := \{h(t) \in C^1[0,T] : h(T) = 0\}$ due to $q_{n_k} \to q^*$ in $L^1(\Omega)$. Since $C_0^\infty(0,T) \subset \mathscr{B}$, we know that (2.4) holds for all $\eta(t) \in C_0^\infty(0,T)$. So (2.4) is essentially a weak form of (1.1) corresponding to $q = q^*$. That is, $u^*(x,t) = u(q^*)(x,t)$ in $L^2((0,T); H_0^1(\Omega))$.

Step 3: Prove $J_1(q_{n_k}) \to J_1(q^*)$ as $n_k \to \infty$.
It follows from (2.3) and Step 2 that

$$u(q_{n_k})(x,t) \rightharpoonup u(q^*)(x,t) \text{ in } L^2((0,T); H_0^1(\Omega)), \quad n_k \to \infty. \tag{2.5}$$

By rewriting the weak form of equation for $q = q^*, q^{n_k}$ and taking test function $\phi(x) = u(q^*)(x,t) - z(x), u(q^{n_k})(x,t) - z(x)$, respectively, we get

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(q_{n_k}) - u(q^*)\|_{L^2}^2 + \int_\Omega q_{n_k}(u(q_{n_k}) - z)^2\,\mathrm{d}x - \int_\Omega q^*(u(q^*) - z)^2\,\mathrm{d}x \\
&\quad = \int_\Omega (q_{n_k} - q^*)[z^2 + u(q_{n_k})(u(q^*) - 2z)]\,\mathrm{d}x - \int_\Omega |\nabla(u(q_{n_k}) - u(q^*))|^2\,\mathrm{d}x \\
&\qquad + 2\int_\Omega q^*[u(q_{n_k}) - u(q^*)]((u(q^*) - z))\,\mathrm{d}x.
\end{aligned} \tag{2.6}$$

This relation can be rewritten as

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(q_{n_k}) - u(q^*)\|_{L^2}^2 + \int_\Omega q_{n_k}(u(q_{n_k}) - u(q^*))^2\,\mathrm{d}x \\
&\quad = \int_\Omega (q_{n_k} - q^*)u(q^*)[u(q^*) - u(q_{n_k})]\,\mathrm{d}x - \int_\Omega |\nabla(u(q_{n_k}) - u(q^*))|^2\,\mathrm{d}x
\end{aligned} \tag{2.7}$$

by simple computations, which generates

$$\begin{aligned}
&\|u(q_{n_k})(\cdot,t) - u(q^*)(\cdot,t)\|_{L^2}^2 \\
&\quad \leqslant 2T\sqrt{2\beta_0\|q_{n_k} - q^*\|_{L^1}}\left(\int_{Q_T} |u(q^*)|^2|u(q^*) - u(q_{n_k})|^2\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}
\end{aligned} \tag{2.8}$$

uniformly in $t \in [0,T]$. Therefore we get

$$\max_{[0,T]} \|u(q_{n_k})(\cdot,t) - u(q^*)(\cdot,t)\|_{L^2}^2 \to 0 \text{ as } n_k \to \infty, \tag{2.9}$$

since $q_{n_k} \to q^*$ in $L^1(\Omega)$ and $u(q^*), u(q_{n_k})$ is uniformly bounded for $q_{n_k}, q^* \in \mathscr{K}$. On the other hand, it follows from (2.7) that

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(q_{n_k}) - u(q^*)\|_{L^2}^2 + \int_\Omega |\nabla(u(q_{n_k}) - u(q^*))|^2\,\mathrm{d}x \\
&\quad \leqslant \int_\Omega (q_{n_k} - q^*)u(q^*)[u(q^*) - u(q_{n_k})]\,\mathrm{d}x.
\end{aligned}$$

By integrating this inequality for $t \in [T-\sigma, T]$, we get

$$\int_{T-\sigma}^T \int_\Omega |\nabla(u(q_{n_k}) - u(q^*))|^2\,\mathrm{d}x\,\mathrm{d}t \to 0 \tag{2.10}$$

as $n_k \to \infty$ from (2.8), (2.9). Now we integrate (2.6) for $t \in [T-\sigma, T]$ again; the convergence (2.5), (2.9) and (2.10) completes the proof. □

With this continuous property for $J_1(q)$, now we can prove the existence of minimizer.

Theorem 2.3. *There exists at least one* $q^*(x) \in \mathscr{K}$ *for* $\alpha > 0$ *such that*

$$J(q^*) = \inf_{\mathscr{K}} J(q). \tag{2.11}$$

Proof. Firstly, there exists a sequence $\{q_n\} \subset \mathscr{K}$ such that

$$\lim_{n\to\infty} J(q_n) = \inf_{\mathscr{K}} J(q), \tag{2.12}$$

which means $\int_\Omega |\nabla q_n|^2 \, dx$ is uniformly bounded with given $\alpha > 0$ and therefore $\|q_n\|^2_{H^1(\Omega)} \leqslant C$. So there exists a subsequence of $\{q_n\}$ such that

$$q_{n_k}(x) \rightharpoonup q^*(x) \in H^1(\Omega) \text{ as } n_k \to \infty. \tag{2.13}$$

Now the Sobolev embedding theorem yields $q_{n_k}(x) \to q^*(x)$ in $L^1(\Omega)$ as $n_k \to \infty$. Obviously $0 \leqslant q^*(x) \leqslant \beta_0$ since $q_{n_k}(x) \in \mathscr{K}$. So we get as $n_k \to \infty$ that

$$q_{n_k}(x) \to q^*(x) \in \mathscr{K} \tag{2.14}$$

in $L^1(\Omega)$. On the other hand, (2.13) generates

$$\int_\Omega |\nabla q^*|^2 = \lim_{n_k\to\infty} \int_\Omega \nabla q_{n_k} \cdot \nabla q^* \leqslant \lim_{n_k\to\infty} \sqrt{\int_\Omega |\nabla q_{n_k}|^2 \, dx \int_\Omega |\nabla q^*|^2 \, dx}. \tag{2.15}$$

Now we apply Lemma 2.2 to the sequence $\{q_{n_k}\}$. That is, there exists a subsequence of $\{q_{n_k}\}$, denoted by $\{q_{m_j}\}$ such that $\lim_{m_j\to\infty} J_1(q_{m_j}) = J_1(q^*)$. So we get from (2.14), Lemma 2.2, (2.15) and (2.12) that

$$J(q^*) = \lim_{m_j\to\infty} J_1(q_{m_j}) + \int_\Omega |\nabla q^*|^2 \, dx \leqslant \lim_{m_j\to\infty} J(q_{m_j}) = \inf_{\mathscr{K}} J(q).$$

The proof is complete. □

Remark 2.4. We have proven the existence of a minimizer, which depends indeed on the regularizing parameter α, σ as well as the error level δ if the noisy data $z^\delta(x)$ is used in the functional. So the other interesting topic to be considered is the convergence property of minimizer as $\alpha \to 0$ and $\delta \to 0$. Although both α and σ play the roles of regularization, the introduction of parameter α in the functional leads us to an approximate optimization problem to original inverse problem. Therefore, it is also necessary to consider the relation between α and σ. All these problems should be studied furthermore.

Since q^* can be found only in a finite space, we should consider the approximation relation $q_h - q^*$, where q_h is the approximation with the discrete parameter h for finite-dimensional space. This is the topic in the next section.

3. Find minimizer in finite-dimensional space

The determination of q^* can be stated as: minimize the functional

$$J(q) := \frac{1}{2}\int_{T-\sigma}^{T} \mathrm{d}t \int_\Omega q(x)|u(q)(x,t) - z(x)|^2 \,\mathrm{d}x + \alpha \int_\Omega |\nabla q|^2 \,\mathrm{d}x \tag{3.1}$$

subject to $q \in \mathscr{K}$, where $u(q)(\cdot, t) \in H_0^1(\Omega)$ meets the constraint

$$\int_\Omega u_t(q)\phi \,\mathrm{d}x + \int_\Omega \nabla u(q) \cdot \nabla\phi \,\mathrm{d}x + \int_\Omega q u(q)\phi \,\mathrm{d}x = 0, \quad \forall \phi(x) \in H_0^1(\Omega),$$
$$u(q)(x,0) = u_0(x), \quad x \in \Omega \tag{3.2}$$

for all $t \in (0, T]$.

The basic idea of solving (3.1)–(3.2) numerically is to find the minimizer by some iteration process and solve the variational problem (3.2) by the finite element method (FEM) for any fixed time $t \in (0, T]$ at each iteration step, where the t derivative is approximated by forward difference scheme.

First of all, both $u(q)(\cdot, t)$ for any fixed $t > 0$ and $q(\cdot)$ are functions defined in Ω. Triangulate the bounded domain Ω by the regular triangulation $\mathscr{T}^h$, where $h > 0$ describes the size of each pixel. Introduce

$$V_h := \{f(x) : \text{continuous and piecewise linear function in } \mathscr{T}^h\},$$

$$\dot{V}_h := \{f(x) \in V_h,\ f(x) = 0 \text{ for } x \in \partial\Omega\}.$$

Denote by $\{x_i\}_{i=1}^N$ the set of all nodal points of $\mathscr{T}^h$ and approximate $\mathscr{K}$ by

$$\mathscr{K}_h = \{q_h(x) \in V_h, 0 \leqslant q_h(x_i) \leqslant \beta_0, i = 1, \ldots, N\}.$$

As for time discretization, we divide $[0, T]$ by nodal $t^j = \tau j$ for $j = 0, 1, \ldots, M$ with step $\tau = T/M$. Denote by $u^j(\cdot) = u(\cdot, j\tau)$, while $u_h^j(x)$ is the projection of $u^j(\cdot)$ onto $\dot{V}_h$ for $u^j(x)$ with zero value in $\partial\Omega$. Moreover, we approximate u_t by backward difference, that is,

$$\partial_\tau u^j(x) := \partial_t u(x,t)|_{t=j\tau} = \frac{u^j(x) - u^{j-1}(x)}{\tau}.$$

For simplicity, we take the small parameter $\sigma > 0$ such that $\sigma = (n_0 + 1)\tau$ for some non-negative integer n_0. Under the above notation, the optimization problem (3.1)–(3.2) in finite space V_h has the following form:

Minimize the functional

$$J_h^{n_0}(q_h) =: \frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_\Omega q_h |u_h^n(q_h)(x) - z(x)|^2 \,\mathrm{d}x + \alpha \int_\Omega |\nabla q_h|^2 \,\mathrm{d}x \tag{3.3}$$

subject to $q_h \in \mathscr{K}_h \subset V_h$, where $u_h^n(q_h)(x) = u_h(q_h)(x, n\tau) \in \dot{V}_h$ meets

$$\int_\Omega \partial_\tau u_h^j(q_h)\phi_h \,\mathrm{d}x + \int_\Omega \nabla u_h^j(q_h) \cdot \nabla\phi_h \,\mathrm{d}x + \int_\Omega q_h u_h^j(q_h)\phi_h \,\mathrm{d}x = 0,$$
$$u_h^0(q_h)(x) = Q_h u_0(x) \tag{3.4}$$

for all $j=0,1,\ldots,M$ and the test function $\phi_h \in \dot{V}_h$, where Q_h is the L^2 projection from $L^2(\Omega)$ onto $\dot{V}_h$ defined by

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

Denote by $I_h : C(\Omega) \to V_h$ the interpolation operator. Then the following property holds for operators I_h, Q_h (see [1,12]).

Lemma 3.1. *Assume $p > dimension(\Omega)$. Then it follows that*

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

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

$$\|Q_h u\|_{L^2} \leqslant C\|u\|_{L^2}, \|Q_h \nabla u\|_{L^2} \leqslant C\|\nabla u\|_{L^2}, \quad \forall u \in H_0^1(\Omega).$$

Similar to the continuous version of optimization problem discussed in Section 2, we should discuss the existence of minimizer in $\mathscr{K}_h$ contained in a finite-dimensional space. However, the proof is somewhat similar and simple due to the equivalence of norms in finite-dimensional space.

The key to the existence proof of minimizer is the following continuous property of $u_h^j(q_h)$ with respect to q_h in $\mathscr{K}_h$ for any fixed $j=0,1,\ldots,M$ and h.

Lemma 3.2. *Fix $\tau, h > 0$. Let a sequence $\{q_h^k\} \subset \mathscr{K}_h$ which tends to q_h as $k \to \infty$ in $L^1(\Omega)$. Then we get for the solution $u_h(q_h^k)(x,t)$ that*

$$u_h^j(q_h^k) \to u_h^j(q_h) \tag{3.5}$$

as $k \to \infty$ in $H_0^1(\Omega)$ for any fixed $j = 1,2,\ldots,M$.

Proof. For q_h^k, q_h in (3.4), the corresponding form is

$$\int_\Omega \partial_t u_h^j(q_h^k)\phi_h \,\mathrm{d}x + \int_\Omega \nabla u_h^j(q_h^k)\cdot\nabla\phi_h \,\mathrm{d}x + \int_\Omega q_h^k u_h^j(q_h^k)\phi_h \,\mathrm{d}x = 0, \tag{3.6}$$

$$\int_\Omega \partial_t u_h^j(q_h)\phi_h \,\mathrm{d}x + \int_\Omega \nabla u_h^j(q_h)\cdot\nabla\phi_h \,\mathrm{d}x + \int_\Omega q_h u_h^j(q_h)\phi_h \,\mathrm{d}x = 0. \tag{3.7}$$

By taking $\phi_h = \tau u_h^j(q_h^k)(x)$ in (3.6), we get that

$$\frac{1}{2}\|u_h^j(q_h^k)\|^2 - \frac{1}{2}\|u_h^{j-1}(q_h^k)\|^2 + \tau\|\nabla u_h^j(q_h^k)\|^2 \leqslant 0. \tag{3.8}$$

Taking summation for $j = 1,2,\ldots,n \leqslant M$ yields

$$\|u_h^n(q_h^k)\|^2 \leqslant C\|u_0\|^2, \quad \tau\sum_{j=1}^n \|\nabla u_h^j(q_h^k)\|^2 \leqslant C\|u_0\|^2 \tag{3.9}$$

uniformly for all $n = 1, 2, \ldots, M$ and $k = 1, 2, \ldots$ with constant C independent of h, τ, k, noticing Lemma 3.1.

Now we can estimate the error $w_h^n(k) := u_h^n(q_h^k) - u_h^n(q_h)$. The subtraction (3.7) from (3.6) yields

$$\int_\Omega \partial_t w_h^j(k)\phi_h \,\mathrm{d}x + \int_\Omega \nabla w_h^j(k) \cdot \nabla \phi_h \,\mathrm{d}x + \int_\Omega q_h^k w_h^j(k)\phi_h \,\mathrm{d}x = \int_\Omega (q_h - q_h^k) u_h^j(q_h)\phi_h \,\mathrm{d}x$$

with $w_h^0(k) = 0$. Choosing $\phi_h = \tau w_h^j(k)$ in this relation and taking summation for $j = 1, 2, \ldots, n \leqslant M$ again yields from (3.9) and $w_h^0(k) = 0$ that

$$\begin{aligned} \frac{1}{2}\|w_h^n(k)\|_{L^2}^2 &\leqslant \tau \|q_h - q_h^k\|_\infty \sum_{j=1}^n \|u_h^j(q_h^k)\| \, \|w_h^j(k)\| \\ &\leqslant 2Cn\tau \|q_h - q_h^k\|_\infty \leqslant 2CT\|q_h - q_h^k\|_\infty \end{aligned} \tag{3.10}$$

for all $n = 1, \ldots, M$. Inserting this estimate to the above inequality yields

$$\begin{aligned} \|\nabla w_h^j(k)\|^2 &\leqslant \|q_h - q_h^k\|_\infty \|u_h^j(q_h)\| \, \|w_h^j(k)\| + \frac{1}{\tau}\frac{1}{2}\|w_h^{j-1}(k)\|^2 \\ &\leqslant 2C\|q_h - q_h^k\|_\infty + \frac{2CT}{\tau}\|q_h - q_h^k\|_\infty. \end{aligned}$$

So we get $\|w_h^j(k)\|_{H^1}^2 \leqslant C\|q_h - q_h^k\|_\infty \to 0$ as $k \to \infty$ for all $j = 1, \ldots, M$ from (3.10), with constant C independent of h but dependent on τ. □

With this continuous property, we can prove the existence of minimizer for the optimization problem (3.3)–(3.4) in set $\mathscr{K}_h \subset V_h$.

Theorem 3.3. *There exists at least one minimizer $q_h^* \in \mathscr{K}_h$ for the optimization problem* (3.3)–(3.4).

The proof is standard and similar to the continuous case (Theorem 2.3), so we omit it.

Remark 3.4. The continuous property applied here only needs the L^2 convergence of $u_h^n(q_h^{n_k})$ as $n_k \to \infty$ in Lemma 3.2.

4. Convergence of approximate solution

To take the minimizer q_h^* in $\mathscr{K}_h$ as the approximate solution to original optimization problem (2.1), we should prove that the minimizer q_h^* of functional $J_h^{n_0}(q_h)$ in finite-dimensional space indeed converges to q^*, the minimizer of optimization problem (2.1) in infinite-dimensional space.

The first step to this aim is to prove the approximate solution to direct problem (1.1) from its variational form (3.2) converges to the exact solution. That is, when the discrete size $\tau, h \to 0$, the discrete functional should tend to the continuous one. For given $\sigma > 0$ in (2.1), we still take $\tau \to 0$ in such a way that $\sigma = (n_0 + 1)\tau$ for some integer n_0.

We need the following lemma in the proof of convergence property.

Lemma 4.1. *Let* $\{u_h^j(q_h)\}$ *be the solution to* (3.4) *for* $q_h \in \mathscr{K}_h$. *Then*

$$\max_{n=1,\ldots,M} \|u_h^n(q_h)\|^2 + \tau \sum_{n=1}^{M} \|u_h^n(q_h)\|^2 + \tau \sum_{n=1}^{M} \|\nabla u_h^n(q_h)\|^2 \leqslant C\|u_0\|^2,$$

$$\max_{n=1,\ldots,M} \|\nabla u_h^n(q_h)\|^2 + \tau \sum_{n=1}^{M} \|\partial_\tau u_h^n(q_h)\|^2 \leqslant C(\|\nabla u_0\|^2 + \|u_0\|^2)$$

hold with constant C *independent of* q_h, h, τ.

Proof. The first result comes from the analogy of (3.9), while the second one is obtained by taking test function $\phi_h = \tau\partial_\tau u_h^j$ in (3.4) and Lemma 3.1. □

For simplicity, we denote $u_h^n(q_h), u(q), u(q)(x, n\tau)$ by u_h^n, u, u^n, respectively. Moreover, introduce the average value of $u(q)$ for $t \in [t_{n-1}, t_n]$ by

$$\overline{u}^n := \frac{1}{\tau}\int_{t_{n-1}}^{t_n} u(q)(x,t)\,\mathrm{d}t, \quad n = 1, \ldots, M, \;\; \overline{u}^0 := u_0(q).$$

The continuous property of the functional (3.3) can be stated as:

Lemma 4.2. *If* $\{q_h(x)\} \subset K_h$ *converges to* $q(x) \in \mathscr{K}$ *in* $L^1(\Omega)$ *as* $h \to 0$, *then*

$$\lim_{h,\tau\to 0} \sum_{n=M-n_0}^{M} \tau \int_\Omega q_h |u_h^n(q_h) - z_h|^2\,\mathrm{d}x = \int_{T-\sigma}^{T} \mathrm{d}t \int_\Omega q|u(q) - z|^2\,\mathrm{d}x,$$

where $u(q) = u(q)(x,t), u_h^n(q_h) = u_h(q_h)(x, n\tau)$ *are the solutions to infinite-dimensional variational form* (3.2) *and its approximation* (3.4), *respectively.*

Proof. Taking $\phi = \tau^{-1}\phi_h(x)$ in (3.2) and integrating for $t \in [t_{j-1}, t_j]$ yield

$$\int_\Omega \partial_\tau u^j \phi_h\,\mathrm{d}x + \frac{1}{\tau}\int_{t_{j-1}}^{t_j}\int_\Omega \nabla u \cdot \nabla\phi_h\,\mathrm{d}x\,\mathrm{d}t + \frac{1}{\tau}\int_{t_{j-1}}^{t_j}\int_\Omega q u \phi_h\,\mathrm{d}x\,\mathrm{d}t = 0$$

from the definition of $\partial_\tau u^n$. Subtracting the above equality from (3.4) and taking $\phi_h = \tau\eta_h^j$ we get

$$\begin{aligned}
&\tau\int_\Omega \partial_\tau\eta_h^j\,\eta_h^j\,\mathrm{d}x + \tau\int_\Omega \partial_\tau(Q_h\overline{u}^j - u^j)\eta_h^j\,\mathrm{d}x + \int_{t_{j-1}}^{t_j}\int_\Omega |\nabla\eta_h^j|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \int_{t_{j-1}}^{t_j}\int_\Omega \nabla(Q_h\overline{u}^j - u)\cdot\nabla\eta_h^j\,\mathrm{d}x\,\mathrm{d}t + \int_{t_{j-1}}^{t_j}\int_\Omega q_h|\eta_h^j|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \int_{t_{j-1}}^{t_j}\int_\Omega q_h(Q_h\overline{u}^j - u)\eta_h^j\,\mathrm{d}x\,\mathrm{d}t = \int_{t_{j-1}}^{t_j}\int_\Omega (q - q_h)u\eta_h^j\,\mathrm{d}x\,\mathrm{d}t,
\end{aligned}$$

where we introduce $\eta_h^j := u_h^j - Q_h\overline{u}^j$. Noticing $q_h \geqslant 0$ and $\tau\partial_\tau\eta_h^j\ \eta_h^j = (\eta_h^j)^2 - \eta_h^j\eta_h^{j-1} \geqslant [(\eta_h^j)^2 - (\eta_h^{j-1})^2]/2$, the above equality tells us

$$
\begin{aligned}
&\frac{1}{2}\|\eta_h^j\|^2 - \frac{1}{2}\|\eta_h^{j-1}\|^2 + \tau\|\nabla\eta_h^j\|^2 \\
&\quad\leqslant \tau\int_\Omega \partial_\tau(u^j - Q_h\overline{u}^j)\eta_h^j\,\mathrm{d}x + \int_{t_{j-1}}^{t_j}\int_\Omega (q - q_h)u\eta_h^j\,\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \int_{t_{j-1}}^{t_j}\int_\Omega \nabla(u - Q_h\overline{u}^j)\cdot\nabla\eta_h^j\,\mathrm{d}x\,\mathrm{d}t + \int_{t_{j-1}}^{t_j}\int_\Omega q_h(u - Q_h\overline{u}^j)\eta_h^j\,\mathrm{d}x\,\mathrm{d}t \\
&\quad =: I_1(j) + I_2(j) + I_3(j) + I_4(j).
\end{aligned} \tag{4.1}
$$

Taking summation for $j = 1, \ldots, n \leqslant M$ and noticing $\eta_h^0 = 0$ generate

$$
\frac{1}{2}\|\eta_h^n\|^2 + \tau\sum_{j=1}^n\|\nabla\eta_h^j\|^2 \leqslant \sum_{j=1}^n[I_1(j) + I_2(j) + I_3(j) + I_4(j)]. \tag{4.2}
$$

Noticing the identity $\sum_{j=1}^n(a_j - a_{j-1})b_j = a_nb_n - a_0b_0 - \sum_{j=1}^n a_{j-1}(b_j - b_{j-1})$ and the definitions of Q_h, ∂_τ^n, we get

$$
\begin{aligned}
\sum_{j=1}^n I_1(j) &= \sum_{j=1}^n\int_\Omega[(u^j - \overline{u}^j) - (u^{j-1} - \overline{u}^{j-1})]\eta_h^j\,\mathrm{d}x \\
&= \int_\Omega(u^n - \overline{u}^n)\eta_h^n\,\mathrm{d}x - \tau\sum_{j=1}^n\int_\Omega(u^{j-1} - \overline{u}^{j-1})\partial_\tau\eta_h^j\,\mathrm{d}x.
\end{aligned} \tag{4.3}
$$

Noticing the fact that

$$
|u^n(x) - \overline{u}^n(x)| = \left|\frac{1}{\tau}\int_{t_{n-1}}^{t_n}\int_y^{t_n} u_t(x,t)\,\mathrm{d}t\mathrm{d}y\right| \leqslant \sqrt{\tau}\sqrt{\int_{t_{n-1}}^{t_n}|u_t(x,t)|^2\,\mathrm{d}t}, \tag{4.4}
$$

the first term in the right-hand side of (4.3) can be estimated by

$$
\left|\int_\Omega(u^n - \overline{u}^n)\eta_h^n\,\mathrm{d}x\right| \leqslant \|u^n - \overline{u}^n\|\|\eta_h^n\| \leqslant \sqrt{\tau}\left(\int_{t_{n-1}}^{t_n}\|u_t\|^2\,\mathrm{d}t\right)^{1/2}\|\eta_h^n\|, \tag{4.5}
$$

while the second one can be estimated from

$$
\begin{aligned}
&\sum_{j=1}^n\int_\Omega|u^{j-1} - \overline{u}^{j-1}||\partial_\tau\eta_h^j|\,\mathrm{d}x \leqslant \sqrt{\sum_{j=1}^n\|u^{j-1} - \overline{u}^{j-1}\|^2\sum_{j=1}^n\|\partial_\tau\eta_h^j\|^2} \\
&\quad\leqslant \sqrt{\sum_{j=2}^n\int_\Omega\tau\int_{t_{j-2}}^{t_{j-1}}|u_t|^2\,\mathrm{d}t\,\mathrm{d}x\sum_{j=1}^n\|\partial_\tau\eta_h^j\|^2} \leqslant \|u_t\|_{L^2(Q_T)}\sqrt{\tau\sum_{j=1}^n\|\partial_\tau\eta_h^j\|^2}
\end{aligned} \tag{4.6}
$$

due to (4.4). Now inserting (4.5) and (4.6) into (4.3) we get

$$\sum_{j=1}^{n} I_1(j) \leqslant C\sqrt{\tau} \tag{4.7}$$

with constant C independent of h, τ from Lemma 4.1 and the boundedness of $\|u_t\|_{L^2(Q_T)}$. For $I_2(j)$, we can see that

$$\begin{aligned}\sum_{j=1}^{n} I_2(j) &\leqslant \sum_{j=1}^{n}\left(\int_{t_{j-1}}^{t_j}\int_\Omega |q-q_h|^2u^2\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}\left(\int_{t_{j-1}}^{t_j}\int_\Omega |\eta_h^j|^2\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}\\ &\leqslant \left(\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}\int_\Omega |q-q_h|^2u^2\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}\left(\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}\int_\Omega |\eta_h^j|^2\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}\\ &\leqslant C\left(\int_0^T\int_\Omega |q-q_h|u^2\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}\end{aligned} \tag{4.8}$$

from $\|q-q_h\|_{L^\infty} \leqslant C$, noticing that $\tau\sum_{j=1}^{n}\|\eta_h^j\|^2 \leqslant Cn\tau \leqslant CT$ again from Lemma 4.1. We estimate $I_3(j)$ by $|ab| \leqslant \frac{1}{4}|a|^2+|b|^2$ to obtain that

$$\begin{aligned}I_3(j) &= \int_{t_{j-1}}^{t_j}\int_\Omega \nabla(u-Q_hu)\cdot\nabla\eta_h^j\,\mathrm{d}x\,\mathrm{d}t + \int_{t_{j-1}}^{t_j}\int_\Omega \nabla(Q_hu-Q_h\overline{u}^j)\cdot\nabla\eta_h^j\,\mathrm{d}x\,\mathrm{d}t\\ &\leqslant \frac{\tau}{2}\|\nabla\eta_h^j\|^2 + \int_{t_{j-1}}^{t_j}\int_\Omega [|\nabla(u-Q_hu)|^2 + |\nabla(Q_hu-Q_h\overline{u}^j)|^2]\,\mathrm{d}x\,\mathrm{d}t,\end{aligned}$$

which yields

$$\begin{aligned}\sum_{j=1}^{n} I_3(j) &\leqslant \frac{\tau}{2}\sum_{j=1}^{n}\|\nabla\eta_h^j\|^2 + \int_0^T \|\nabla(u-Q_hu)\|^2\,\mathrm{d}t\\ &\quad + C\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}\|\nabla(Q_h(u-\overline{u}^j))\|^2\,\mathrm{d}t.\end{aligned} \tag{4.9}$$

Finally we consider $I_4(j)$. Noticing η_h^j is independent of t, we get that

$$I_4(j \leqslant C\|\eta_h^j\|\int_{t_{j-1}}^{t_j}[\|u-Q_hu\|_{L^2(\Omega)} + \|Q_h(u-\overline{u}^j)\|_{L^2(\Omega)}]\,\mathrm{d}t$$

from the property of Q_h, which yields

$$\sum_{j=1}^{n} I_4(j) \leqslant C\int_0^T \|u-Q_hu\|_{L^2(\Omega)}\,\mathrm{d}t + C\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}\|Q_h(u-\overline{u}^j)\|_{L^2(\Omega)}\,\mathrm{d}t. \tag{4.10}$$

Inserting (4.7)–(4.10) into (4.2) generates

$$
\begin{aligned}
\frac{1}{2}\|\eta_h^j\|^2 + \frac{\tau}{2}\sum_{j=1}^{n}\|\nabla\eta_h^j\|^2 \leqslant C\sqrt{\tau} + C\left(\int_0^T\int_\Omega |q-q_h|u^2\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2} \\
+C\int_0^T \|u-Q_hu\|_{H^1(\Omega)}\,\mathrm{d}t + C\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}\|Q_h(u-\overline{u}^j)\|_{H^1}\,\mathrm{d}t \qquad (4.11)
\end{aligned}
$$

from the property of Q_h. On the other hand,

$$
\|Q_h(u-\overline{u}^j)\|_{H^1} \leqslant C\|u-\overline{u}^j\|_{H^1} \leqslant \frac{C}{\tau}\int_{t_{j-1}}^{t_j}\|u(\cdot,t)-u(\cdot,\xi)\|_{H^1}\,\mathrm{d}\xi.
$$

Since $t, \xi \in [t_{j-1}, t_j]$ means $|t-\xi| \leqslant \tau$, it follows for $\forall\varepsilon > 0$ that $\|u(\cdot,t)-u(\cdot,\xi)\|_{H^1} \leqslant \varepsilon$ for $\tau > 0$ small enough. So $\|Q_h(u-\overline{u}^j)\| \leqslant C\varepsilon$, which means

$$
\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}\|Q_h(u-\overline{u}^j)\|_{H^1}\,\mathrm{d}t \leqslant C\varepsilon n\tau \leqslant CT\varepsilon.
$$

Now by letting $\tau, h \to 0$ in (4.11), we get $\|\eta_h^j\|^2$, $\frac{\tau}{2}\sum_{j=1}^{n}\|\nabla\eta_h^j\|^2 \leqslant C\varepsilon$ uniformly for all $j = 1, \ldots, n \leqslant M$ for $h, \tau > 0$ small enough. The arbitrariness of ε means

$$
\lim_{h,\tau\to 0}\max_{n=1,\ldots,M}\|\eta_h^n\|^2 = \lim_{h,\tau\to 0}\frac{\tau}{2}\sum_{j=1}^{M}\|\nabla\eta_h^j\|^2 = 0. \qquad (4.12)
$$

That is,

$$
\tau\sum_{j=1}^{M}\|\eta_h^j\|^2 \leqslant M\tau \max_{j=1,\ldots,M}\|\eta_h^j\|^2 = T \max_{j=1,\ldots,M}\|\eta_h^j\|^2 \to 0 \qquad (4.13)
$$

as $h, \tau \to 0$. By this estimate, we get from $u_h^n - \overline{u}^n = \eta_h^n + (Q_h\overline{u}^n - \overline{u}^n)$ that

$$
\begin{aligned}
\tau\sum_{n=1}^{M}\|u_h^n-\overline{u}^n\|_{L^2(\Omega)}^2 &\leqslant 2\tau\sum_{n=1}^{M}\|\eta_h^n\|^2 + 2\tau\sum_{n=1}^{M}\|Q_h\overline{u}^n-\overline{u}^n\|^2 \\
&\leqslant 2\tau\sum_{n=1}^{M}\|\eta_h^n\|^2 + 2T\max_{n=1,\ldots,M}\|Q_h\overline{u}^n-\overline{u}^n\|^2 \to 0 \qquad (4.14)
\end{aligned}
$$

as $h, \tau \to 0$ from (4.13) and the property of Q_h.

Now we can complete the proof of Theorem 4.2. By defining $z_h := Q_h z$, it is enough to prove that

$$
I(h,\tau) - I(h) := \tau\sum_{n=M-n_0}^{M}\int_\Omega q_h|u_h^n - z_h|^2\,\mathrm{d}x - \int_{T-\sigma}^{T}\int_\Omega q|u-z_h|^2\,\mathrm{d}x\,\mathrm{d}t \to 0
$$

as $h, \tau \to 0$, noticing M depending on τ. We rewrite $I(h,\tau) - I(h)$ as

$$I(h,\tau) - I(h) = \sum_{n=M-n_0}^{M} \int_{t_{n-1}}^{t_n} \int_\Omega q_h(|u_h^n - z_h|^2 - |u - z_h|^2)\,\mathrm{d}x\,\mathrm{d}t + \int_{T-\sigma}^{T} \int_\Omega (q_h - q)|u - z_h|^2\,\mathrm{d}x\,\mathrm{d}t =: S_1(h,\tau) + S_2(h). \tag{4.15}$$

For the first term, it follows from Hölder inequality and Lemma 4.1 that

$$\begin{aligned} S_1(h,\tau) &= \sum_{n=M-n_0}^{M} \int_{t_{n-1}}^{t_n} \int_\Omega q_h(u_h^n - u)[u_h^n + u - 2z_h]\,\mathrm{d}x\,\mathrm{d}t \\ &\leqslant C\left(\sum_{n=M-n_0}^{M} \int_{t_{n-1}}^{t_n} \|u_h^n - u\|^2\right)^{1/2} \left(\sum_{n=M-n_0}^{M} \int_{t_{n-1}}^{t_n} \|u_h^n + u - 2z_h\|^2\right)^{1/2} \\ &\leqslant C\left(\sum_{n=M-n_0}^{M} \int_{t_{n-1}}^{t_n} \|u_h^n - u\|^2\,\mathrm{d}t\right)^{1/2} \end{aligned} \tag{4.16}$$

due to the boundedness of q_h. On the other hand, the triangle inequality means

$$\sum_{n=M-n_0}^{M} \int_{t_{n-1}}^{t_n} \|u_h^n - u\|^2\,\mathrm{d}t \leqslant 2\tau \sum_{n=1}^{M} \|u_h^n - \overline{u}^n\|^2 + 2\sum_{n=1}^{M} \int_{t_{n-1}}^{t_n} \|\overline{u}^n - u\|^2\,\mathrm{d}t. \tag{4.17}$$

The first term tends to 0 as $h, \tau \to 0$ from (4.14), while the second term also tends to 0 from the property of average function. So we get $S_1(h,\tau) \to 0$ as $h, \tau \to 0$. That $S_2(h) \to 0$ as $h \to 0$ is obvious since $q_h \to q$ in L^1. Therefore we conclude that $I(h,\tau) - I(h) \to 0$ as $h, \tau \to 0$. The proof is complete. □

Now we can prove the convergence property of solution sequence $\{q_h\}_{h>0}$.

Theorem 4.3. *Assume that* $\{q_h^*\}_{h>0}$ *is the minimizer to discrete optimization problem* (3.3)–(3.4). *Then there exists a subsequence of* $\{q_h^*\}_{h>0}$ *such that this subsequence converges to the minimizer* (3.1)–(3.2).

Proof. Noticing the constant $\beta_0 \in \mathscr{K}_h$, we get $J_h^{n_0}(q_h^*) \leqslant J_h^{n_0}(\beta_0) \leqslant C$ with some constant C independent of h, τ from (3.4) and (3.9), which generates $\|q_h^*\|_{H^1} \leqslant C$ for any fixed $\alpha > 0$ due to $q_h^* \in \mathscr{K}_h$ and (3.3). Noticing the uniformly boundedness of $\{q_h^*\}_{h>0}$ in L^2, there exists a subsequence of $\{q_h^*\}_{h>0}$, which is still denoted by $\{q_h^*\}_{h>0}$ such that

$$q_h^* \to q^* \in \mathscr{K} \tag{4.18}$$

in $L^2(\Omega) \subset L^1(\Omega)$ as $h \to 0$. On the other hand, the denseness of $C_0^\infty(\overline{\Omega})$ in $W^{1,2}(\Omega)$ means that any $q \in \mathscr{K}$ can be approximated by a function in $C_0^\infty(\overline{\Omega})$. That is, for arbitrary $\varepsilon > 0$, there exists $q_\varepsilon \in C_0^\infty(\overline{\Omega})$ such that

$$\|q_\varepsilon - q\|_{L^2} \leqslant \varepsilon, \quad \|\nabla(q_\varepsilon - q)\|_{L^2} \leqslant \varepsilon \tag{4.19}$$

for any $q \in H^1(\Omega)$. Restrict $q_\varepsilon \in \mathscr{K}$ by defining

$$\hat{q}_\varepsilon(x) := \begin{cases} q_\varepsilon(x), & 0 \leqslant q_\varepsilon(x) \leqslant \beta_0, \\ 0, & q_\varepsilon(x) \leqslant 0, \\ \beta_0, & q_\varepsilon(x) \geqslant \beta_0, \end{cases} \tag{4.20}$$

then $\hat{q}_\varepsilon \in W^{1,\infty}(\Omega) \bigcap \mathscr{K}$ and

$$\|\hat{q}_\varepsilon - q\|_{L^1(\Omega)} \leqslant \|q_\varepsilon - q\|_{L^1(\Omega)} \leqslant \varepsilon, \quad \|\nabla q_\varepsilon\|_{L^2(\Omega)} \leqslant \|\nabla q\|_{L^2(\Omega)} + \varepsilon \tag{4.21}$$

from this definition and $q \in \mathscr{K}$ as well as (4.19), which means

$$\begin{aligned} \int_\Omega |\nabla \hat{q}_\varepsilon|^2 \, \mathrm{d}x &= \int_{x:\hat{q}_\varepsilon = q} |\nabla \hat{q}_\varepsilon|^2 \, \mathrm{d}x = \int_{x:\hat{q}_\varepsilon = q} |\nabla q_\varepsilon|^2 \, \mathrm{d}x \\ &\leqslant \int_\Omega |\nabla q_\varepsilon|^2 \, \mathrm{d}x \leqslant \int_\Omega |\nabla q|^2 \, \mathrm{d}x + 2\|\nabla q\|\varepsilon + \varepsilon^2. \end{aligned} \tag{4.22}$$

Now for any $q \in \mathscr{K}$, define $\hat{q}_\varepsilon$ in the previous way and then define $q_h^\varepsilon = I_h \hat{q}_\varepsilon \in \mathscr{K}_h$. Since q_h^* is the minimizer in $\mathscr{K}_h$, it follows that

$$J_h^{n_0}(q_h^*) \leqslant J_h^{n_0}(q_h^\varepsilon) = \frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_\Omega q_h^\varepsilon |u_h^n(q_h^\varepsilon) - z|^2 \, \mathrm{d}x + \alpha \int_\Omega |\nabla q_h^\varepsilon|^2 \, \mathrm{d}x. \tag{4.23}$$

Now Lemma 4.2 and inequality (2.15) generate

$$\begin{aligned} J(q^*) &\leqslant \lim_{h,\tau \to 0} \frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_\Omega q_h^* |u_h^n(q_h^*) - z|^2 \, \mathrm{d}x + \alpha \lim_{h \to 0} \int_\Omega |\nabla q_h^*|^2 \, \mathrm{d}x \\ &\leqslant \lim_{h,\tau \to 0} \left[\frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_\Omega q_h^\varepsilon |u_h^n(q_h^\varepsilon) - z|^2 \, \mathrm{d}x + \alpha \int_\Omega |\nabla q_h^\varepsilon|^2 \, \mathrm{d}x \right] \end{aligned} \tag{4.24}$$

from (4.23). For any fixed $\varepsilon > 0$, it follows from $q_h^\varepsilon = I_h \hat{q}_\varepsilon \to \hat{q}_\varepsilon$ in $H^1(\Omega)$ and $\|u - I_h u\|_{L^2(\Omega)} \to 0$ as $h \to 0$ and (4.24) that

$$\begin{aligned} J(q^*) &\leqslant \frac{1}{2} \int_{T-\sigma}^{T} \int_\Omega q^\varepsilon |u(q^\varepsilon) - z|^2 \, \mathrm{d}x + \alpha \int_\Omega |\nabla q^\varepsilon|^2 \, \mathrm{d}x \\ &\leqslant \frac{1}{2} \int_{T-\sigma}^{T} \int_\Omega q^\varepsilon |u(q^\varepsilon) - z|^2 \, \mathrm{d}x + \alpha \left(\int_\Omega |\nabla q|^2 \, \mathrm{d}x + 2\|\nabla q\|\varepsilon + \varepsilon^2 \right) \end{aligned}$$

from (4.22). Finally, we take $\varepsilon \to 0$ in the estimate, we get $J(q^*) \leqslant J(q)$ for any $q \in \mathscr{K}$ from Lemma 2.2 and (4.21). That is, q^* is the minimizer. □

This theorem guarantees theoretically that the minimizer q_h^* for finite-dimensional problem can be used to approximate the exact minimizer q for small $h > 0$. However, the optimization problem (3.3)–(3.4) in the set $\mathscr{K}_h$ is constrained by the requirement $0 \leqslant q(x) \leqslant \beta_0$. This condition will cause some difficulty when we find $q_h^*(x)$ numerically. As usual, we convert this problem into one without a priori quantity assumption on $q_h(x)$. For this purpose, construct a new functional from $J_h^{n_0}(q)$ for $q_h \in \mathscr{K}_h$ by

$$\tilde{J}_h^{n_0}(\varepsilon, q_h) = J_h^{n_0}(q_h) + \frac{1}{\varepsilon}\int_\Omega P(q_h)(x)\,dx \tag{4.25}$$

with the admissible set V_h and small $\varepsilon > 0$, where the penalty term is

$$P(q_h)(x) := \tfrac{1}{2}[(q_h(x) - \beta_0)_+^2 + (q_h(x))_+^2]. \tag{4.26}$$

The notation a_+ is defined as $a_+ = 0$ for $a \leqslant 0$ and $a_+ = a$ for $a \geqslant 0$. We consider the minimizer of functional $\tilde{J}_h^{n_0}(\varepsilon, q_h)$ in the entire space V_h. By the same argument as that in proving Theorem 3.3, we know that there exists a minimizer $q_h^\varepsilon(x) \in V_h$ to this problem. The following theorem states the relation between $q_h^\varepsilon(x)$ and q_h^*, the minimizer of (3.3)–(3.4).

Theorem 4.4. *Assume that $\{\varepsilon_i\}$ is a positive sequence satisfying $\varepsilon_1 > \varepsilon_2 > \cdots \varepsilon_n \to 0$ as $n \to \infty$ and denote by $q_h^{\varepsilon_n}$ the minimizer of functional* (4.25) *in V_h. Then there exists a subsequence of $\{q_h^{\varepsilon_n}\}$, which tends to q_h^* as $n \to \infty$.*

Proof. Since $q_h^{\varepsilon_i}$ is a minimizer of $\tilde{J}_h^{n_0}(\varepsilon_i, \cdot)$ in V_h and $\varepsilon_{i+1} < \varepsilon_i$, we get

$$\tilde{J}_h^{n_0}(\varepsilon_i, q_h^{\varepsilon_i}) \leqslant \tilde{J}_h^{n_0}(\varepsilon_i, q_h^{\varepsilon_{i+1}}) \leqslant \tilde{J}_h^{n_0}(\varepsilon_{i+1}, q_h^{\varepsilon_{i+1}}),$$

that is, the sequence $\tilde{J}_h^{n_0}(\varepsilon_i, q_h^{\varepsilon_i})$ increases with respect to i. On the other hand, it follows for $q_h \in \mathscr{K}_h \subset V_h$ that

$$\frac{1}{\varepsilon_i}\int_\Omega P(q_h^{\varepsilon_i})(x)\,dx \leqslant \tilde{J}_h^{n_0}(\varepsilon_i, q_h^{\varepsilon_i}) \leqslant \tilde{J}_h^{n_0}(\varepsilon_i, q_h) = J_h^{n_0}(q_h)$$

due to $P(q_h) = 0$. This fact means

$$\lim_{\varepsilon_i \to 0}\int_\Omega P(q_h^{\varepsilon_i})(x)\,dx \to 0, \quad \varepsilon_i \to 0. \tag{4.27}$$

So expression (4.26) generates

$$0 \leqslant \lim_{\varepsilon_i \to 0} q_h^{\varepsilon_i}(x) \leqslant \beta_0, \tag{4.28}$$

since $q_h^{\varepsilon_i}(x)$ is a continuous positive function. Moreover, it follows from (4.26) and (4.27) that $\{\|q_h^{\varepsilon_i}\|_{L^2(\Omega)}\}$ is uniformly bounded for all i. Noticing $q_h^{\varepsilon_i}$ lies in a finite-dimensional space for fixed $h > 0$, all the norms are equivalent. Therefore there exists a subsequence, denoted still by $q_h^{\varepsilon_i}$, which converges to some q_h^* in any norm. So (4.28) means $q_h^* \in \mathscr{K}_h$.

Finally, the definition of $q_h^{\varepsilon_i}$ and (4.26) means for any $q_h \in \mathscr{K}_h \subset V_h$ that

$$J_h^{n_0}(q_h^{\varepsilon_i}) \leqslant \tilde{J}_h^{n_0}(\varepsilon_i, q_h^{\varepsilon_i}) \leqslant \tilde{J}_h^{n_0}(\varepsilon_i, q_h) = J_h^{n_0}(q_h).$$

Now the application of $q_h^{\varepsilon_i} \to q_h^*$ in any norm and Lemma 4.2 to the above inequality means $J_h^{n_0}(q_h^*) \leqslant J_h^{n_0}(q_h)$ for any $q_h \in \mathscr{K}_h$. That is, q_h^* is the minimizer in $\mathscr{K}_h$. The proof is complete. □

Our next work is to minimize the function $\tilde{J}_h^{n_0}(\varepsilon, q_h)$ given by (4.25) in the whole space V_h for given $\sigma, h, \varepsilon > 0$. As usual, the iterative scheme will be used to solve this optimization problem.

5. Numerical implementations for optimization problem

Having obtained the theoretical results for our inverse problems, we consider the numerical inversion schemes for the optimization problem in finite-dimensional space. We apply Armijo algorithm to minimize the functional

$$\tilde{J}_h^{n_0}(\varepsilon, q_h) = \frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_\Omega q_h |u_h^n(q_h) - z|^2 \, dx + \alpha \int_\Omega |\nabla q_h|^2 \, dx + \frac{1}{\varepsilon} \int_\Omega P(q_h) \, dx$$

$$=: \tilde{J}_1(q_h) + \alpha \tilde{J}_2(q_h) + \frac{1}{\varepsilon} \tilde{J}_3(q_h) \tag{5.1}$$

over the finite space V_h.

A key ingredient in Armijo algorithm is the Gateaux derivative of $\tilde{J}_h^{n_0}(\varepsilon, q_h)$. Since $\varepsilon > 0$ is given in this algorithm, we abbreviate $\tilde{J}_h^{n_0}(\varepsilon, q_h)$ as $\tilde{J}_h^{n_0}(q_h)$ in this section, without any confusion.

By lengthy but trivial computation, we can obtain the Gateaux derivatives of functional $\tilde{J}_h^{n_0}(q_h)$ at q_h along direction $p(x) \in V_h$. This result is stated as:

Theorem 5.1. *The Gateaux difference of $\tilde{J}_h^{n_0}(q_h)$ at point $q_h \in V_h$ along direction $p(x)$ is determined from the following expressions*:

$$\tilde{J}_2'(q_h)p = 2 \int_\Omega \nabla q_h(x) \cdot \nabla p(x) \, dx, \quad \tilde{J}_3'(q_h)p = \int_\Omega P'(q_h) p \, dx, \tag{5.2}$$

with the function

$$P'(q_h)p = \begin{cases} (q_h(x) - \beta_0)p(x), & q_h(x) > \beta_0, \\ 0, & 0 \leqslant q_h(x) \leqslant \beta_0, \\ q_h(x)p(x), & q_h(x) \leqslant 0. \end{cases} \tag{5.3}$$

For derivative $\tilde{J}_1'(q_h)p$, it can be computed by the adjoint operator method with the expression

$$\tilde{J}_1'(q_h)p = \frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_\Omega p(x) |u_h^n(q_h) - z|^2 \, dx + \tau \sum_{n=1}^{M} \int_\Omega p(x) w_h^{n-1}(x) u_h^n(q_h) \, dx, \tag{5.4}$$

where the adjoint function $w_h^n(x)$ for $n = 1, 2, \ldots, M$ satisfies that

$$- \int_\Omega \partial_\tau w_h^n \phi_h \, dx + \int_\Omega \nabla w_h^{n-1} \cdot \nabla \phi_h \, dx + \int_\Omega q_h w_h^{n-1} \phi_h \, dx$$

$$+ \mu_n \int_\Omega q_h^j(x)(u_h^n - z)\phi_h \, dx = 0,$$

$$w_h^M(x) = 0, \ \partial_\tau w_h^n := \frac{w_h^n(x) - w_h^{n-1}(x)}{\tau} \tag{5.5}$$

for $n = M, \ldots, 1$, $\phi_h(x) \in \dot{V}_h$ is the test function and $\mu_n = 1$ for $n = M, \ldots, M - n_0$ and $\mu_n = 0$ for $n = 1, 2, \ldots, M - n_0 - 1$.

Now we state the steps of Armijo algorithm as follows:

Step 1: Give iteration step adjusting parameter η_0 and iteration stop tolerance parameter ε_0. Set iteration initial guess $q_h^0(x) \in V_h$.

Step 2: For given $q_h^j(x) \in V_h$, solve the direct problem for $u(x,t)$ in $(x,t) \in \Omega \times (0,T)$ by its variational form

$$\int_\Omega \partial_\tau u_h^n \phi_h \,\mathrm{d}x + \int_\Omega \nabla u_h^n \cdot \nabla \phi_h \,\mathrm{d}x + \int_\Omega q_h^j u_h^n \phi_h \,\mathrm{d}x = 0, \ \forall \phi_h \in \dot{V}_h,$$

$$u_h^0(x) = I_h u_0(x), \ \partial_\tau u_h^n := \frac{u_h^n(x) - u_h^{n-1}(x)}{\tau} \tag{5.6}$$

for $n = 1, \ldots, M$ along time direction, where ϕ_h is the base function of space $\dot{V}_h$. Then we solve the adjoint problem for $w(x,t)$ by the variational form (5.5).

Step 3: Compute the Gateaux derivative $\tilde{J}_h^{n_0}(q_h^j)\psi_l := g_l$ for $l = 1, 2, \ldots, N$ along direction ψ_l, where ψ_l is the base function of V_h:

$$g_l = \frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_\Omega \psi_l(x) |u_h^n(q_h^j) - z|^2 \,\mathrm{d}x + \tau \sum_{n=1}^{M} \int_\Omega \psi_l(x) w_h^{n-1}(x) u_h^n(q_h^j) \,\mathrm{d}x$$
$$+ 2\alpha \int_\Omega \nabla q_h^j(x) \cdot \nabla \psi_l(x) \,\mathrm{d}x + \frac{1}{\varepsilon} \tilde{J}_3'(q_h^j)\psi_l, \tag{5.7}$$

where $J_3'(q_h^j)\psi_l$ is given by (5.2) and (5.3). Then we get the iteration direction from jth step to $(j+1)$th step $g_h^j(x) = \sum_{l=1}^N g_l \psi_l(x)$.

Step 4: Compute the norm of $g_h^j(x)$ at jth step: $e_j = (h\sum_{l=1}^N g_l^2)^{1/2}$.

Step 5: Determine the iteration step length from jth step to $(j+1)$th step or stop the iteration from the following step:

(5a) Set $\lambda = 1$.

(5b) Compute err $:= \tilde{J}_h^{n_0}(q_h^j + \lambda g_h^j) - \tilde{J}_h^{n_0}(q_h^j) + \frac{1}{2}\lambda e_j^2$.

(5c) If err $\leqslant 0$, then come to (5d); otherwise set $\lambda = \eta_0 \lambda$ and come to (5b).

(5d) Set $q_h^{j+1} = q_h^j + \lambda g_h^j$. If $\|\lambda g_h^j\| \leqslant \varepsilon_0$, exit and stop the iteration scheme. Otherwise set $j = j + 1$ and come to Step 2.

Next we apply this algorithm to some examples to show the validity of our inversion scheme. In our numerics, the step sizes h, τ and regularizing parameter α as well as penalty parameter ε are given in advance. Moreover, the condition err $\leqslant 0$ in (5c) is replaced by err $\leqslant \delta_0$ for some small parameter $\delta_0 > 0$.

Example 1: We take $a = 0.01$, $T = 1$ with $q(x) = e^x$. The initial function for iteration is $q^0(x) \equiv 3$ with the regularizing parameter $\alpha = 10^{-5}$. The reconstruction results for different iteration times are shown in Fig. 1.

We can see from this figure that the main shape can be recovered with satisfactory accuracy, except the boundary points. However, we also find in our numerics that the boundary status can also be revealed well with different regularizing parameter $\alpha = 10^{-8}$.

Example 2. This example aims to reconstruct a oscillatory function $q(x) = 3 + 2x^2 - 2\sin(2\pi x)$. In numerics, we take $a = 0.01$, $T = 1$.

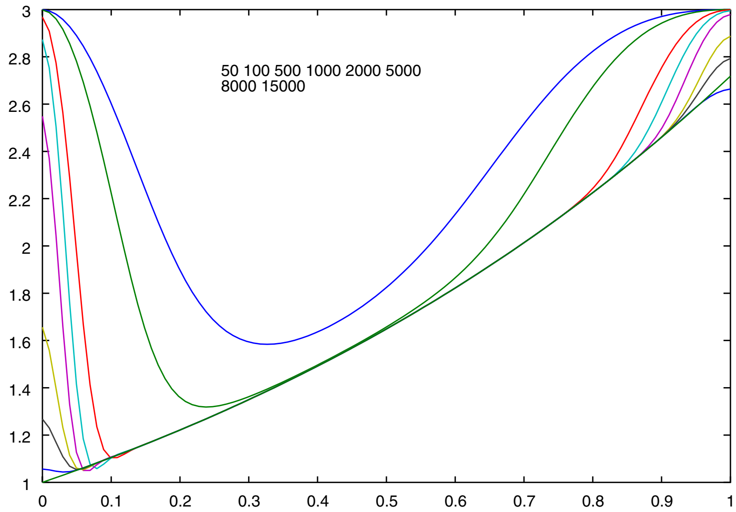

Fig. 1. Recovery at different iteration step.

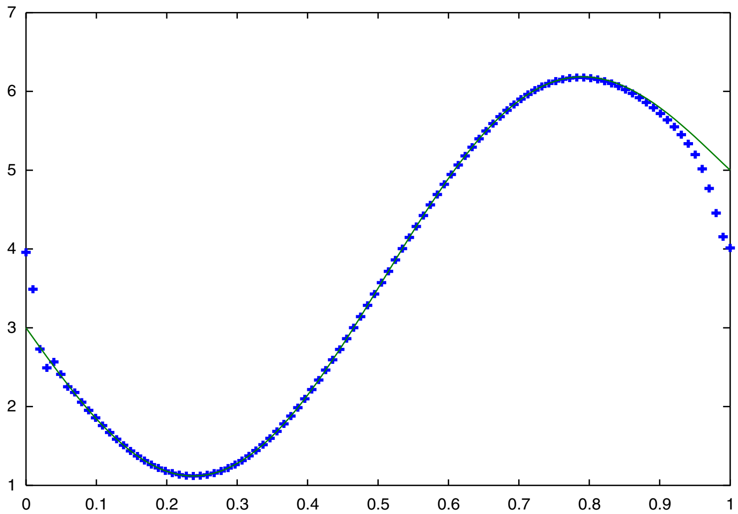

Fig. 2. Recovery of $q(x) = 3 + 2x^2 - 2\sin(2\pi x)$.

The same situation arises as that in Example 1. If we apply the noisy data generated in the form

$$u^\delta(x, T) = u(x, T)[1 + \delta \times \text{random}(x)]$$

with $\delta = 0.05$, that is, with 5% relative error, the reconstruction is also satisfactory, see Fig. 2.

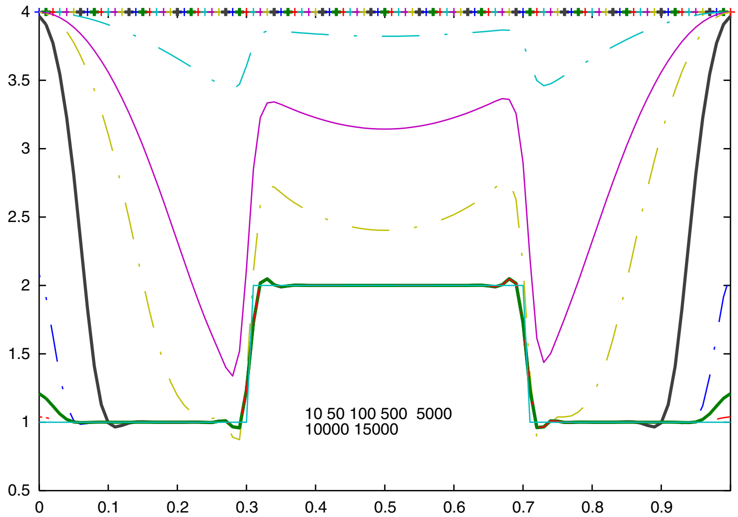

Fig. 3. Recovery of discontinuous $q(x)$ from constant initial value.

Example 3: We take $a = 0.01$, $T = 1$ with discontinuous $q(x)$ given by

$$q(x) = \begin{cases} 1, & x \in [1, 0.3], \\ 2, & x \in (0.3, 0.7], \\ 1, & x \in (0.7, 1]. \end{cases} \tag{5.8}$$

The initial function for iteration is $q^0(x) \equiv 4$. For different iteration times n, the results are shown in Fig. 3.

We can see that the discontinuous property of $q(x)$ is recovered very well after 15 000 iterations. Also, the rough shape of $q(x)$ has been obtained after few iterations.

Acknowledgements

This work is supported by NSFC (no. 10371018). The authors thank the anonymous referees for their helpful suggestions, which led to the present version of this paper.

References

[1] P. Ciarlet, Basic error estimates for elleptic problems, in: P.G. Ciarlet, J.-L. Lions (Eds.), Handbook of Numerical Analysis, vol. II, North-Holland, Amsterdam, 1991, pp. 17–352.
[2] M. Chouli, M. Yamamoto, Generic well-posedness of an inverse parabolic problem—the Holder space approach, Inverse Problems 12 (3) (1996) 195–206.
[3] S. Gutman, Identification of discontinuous parameters in flow equations, SIAM J. Control and Optimization 28 (5) (1990) 1049–1060.
[4] V. Isakov, Inverse parabolic problems with the final overdetermination, Comm. Pure Appl. Maths 44 (2) (1991) 185–209.
[5] V. Isakov, Inverse Problems for Partial Differential Equations, Springer, New York, 1998.

[6] L.S. Jiang, Q. Chen, L. Wang, J.E. Zhang, A new well-posed algorithm to recover implied local volatility, Quantitative Finance 3 (2003) 451–457.
[7] L.S. Jiang, Y. Tao, Identifying the volatility of underlying assets from option prices, Inverse Problems 17 (2001) 137–155.
[8] Y.L. Keung, J. Zou, Numerical identifications of parameters in parabolic equations, Inverse Problems 14 (1998) 83–100.
[9] C. Kravaris, J.H. Seinfeld, Identification of parameters in distributed parameter systems by regularization, SIAM J. Control and Optimization 23 (2) (1985) 217–241.
[10] O.A. Ladyzhenskaya, The Boundary Value Problems of Mathematical Physics, Springer, New York, 1985.
[11] A.I. Prilepko, D.G. Orlovsky, I.A. Vasin, Methods for Solving Inverse Problems in Mathematical Physics, vol. 1, Marcel Dekker, New York, 2000.
[12] J. Xu, Theory of multilevel methods, Ph.D Thesis, Cornell University, May 1989.