INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **17** (2001) 1181–1202 www.iop.org/Journals/ip PII: S0266-5611(01)17705-1

Simultaneous reconstruction of the initial temperature and heat radiative coefficient

Masahiro Yamamoto[1] **and Jun Zou**[2]

[1] Department of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo 153, Japan
[2] Department of Mathematics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, People's Republic of China

E-mail: myama@ms.u-tokyo.ac.jp and zou@math.cuhk.edu.hk

Received 4 October 2000

Abstract

Given the measurement of temperature at a fixed time $\theta > 0$ and the measurement of temperature in a subregion of the physical domain, we investigate the simultaneous reconstruction of the initial temperature and heat radiative coefficient in a heat conductive system. The stability of the inverse problem is first established, and then the numerical reconstruction is mainly studied. The reconstruction process is done by Tikhonov regularization with the regularizing terms being the L^2-norms of gradients, and is carried out in such a way that the temperature solution of the heat equation matches its fixed time observation and its subregion observation optimally in the L^2-norm sense. The continuous nonlinear optimization system will be discretized by the piecewise linear finite element method, and the existence of discrete minimizers and convergence of the finite element approximation are shown. The discrete finite element problem is solved by a nonlinear gradient method with an efficient nonlinear multigrid technique for accelerating the reconstruction process. Numerical experiments are given to demonstrate the efficiency of the proposed nonlinear multigrid gradient method for solving the inverse parabolic problem.

1. Introduction

Consider the following heat conduction problem:

$$u_t(x,t) = \Delta u(x,t) + p(x)\,u(x,t) \qquad \text{in} \quad \Omega \times (0,T) \tag{1.1}$$

with the initial condition

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

and the Dirichlet boundary condition

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

0266-5611/01/041181+22$30.00

where the physical domain Ω is an open bounded domain in $\mathbb{R}^d$ ($d = 1, 2, 3$), with a piecewise smooth boundary Γ.

The goal of this paper is to investigate the possibility of the simultaneous numerical reconstruction of the initial temperature distribution μ and the heat radiative coefficient p in (1.1), (1.2). Such reconstructions are extremely important in many practical applications. But unfortunately, they are highly ill-posed in most situations. In order to achieve some reasonable numerical reconstructions, one has to take into consideration the ill-posedness of the problem and restoration of stability under suitable *a priori* information. However, only few theoretical results seem available for the investigation.

First of all, it is well known that in the case where p is given, the reconstruction of the initial temperature μ from $u(x, \theta)$ for $x \in \Omega$ and fixed $\theta > 0$ is highly ill-posed and impossible in most cases. On the other hand, in the case where μ is given, the inverse problem of determining p from the observation $u(x, \theta)$, $x \in \Omega$, with $\theta > 0$, can be transformed to a Fredholm equation of the second kind, where there might exist a non-trivial solution which implies the non-uniqueness for such an inverse problem (e.g. [16]).

Thus we are motivated to adopt a different formulation which should be conditionally stable and also physically acceptable. Then, as observation of the temperature, we take $u(x, \theta)$, $x \in \Omega$ and $u_{|\omega\times(0,T)}$, where $\theta > 0$ is fixed and $\omega \subset \Omega$ is an arbitrarily prescribed subregion. In general, ω can be assumed to be sufficiently small. We can expect a Lipschitz stability for determining the coefficient p but a very weak stability for determining an initial status μ (theorem 2.1). For our numerical reconstruction, we will take some inexact data $z_\theta(x)$, $x \in \Omega$ and $z_{|\omega\times(0,T)}$, instead of the exact data $u(\cdot, \theta)_{|\Omega}$ and $u_{|\omega\times(0,T)}$:

$$z_\theta(x) \approx u(x, \theta), \qquad x \in \Omega \qquad \text{and} \qquad z(x, t) \approx u(x, t), \qquad (x, t) \in \omega \times (0, T) \tag{1.4}$$

where ω is a subregion of Ω. The distributional observation data z_θ and z are possibly obtained through interpolations of the point observation values in practice.

The essence of our formulation for the inverse problem is to simultaneously reconstruct the initial value μ and the radiative coefficient p in the heat conduction system (1.1)–(1.3) from the measurements of temperatures over $\omega \times (0, T)$ and at a fixed moment $\theta > 0$. In the case of $\theta = 0$, our formulation corresponds to the determination of the coefficient p only, and the stability and uniqueness are still open problems for an arbitrary subregion ω in Ω (e.g. [16]). For our simultaneous reconstruction, we choose a fixed $\theta \in (0, T)$. Note that the measurement of $u(\cdot, \theta)$ for $\theta > 0$ is often easier to achieve than the one of $u(\cdot, 0)$, so our formulation also seems reasonable from the practical point of view.

The main purpose of this paper is to propose some numerical methods for solving the considered inverse problem, based on our stability analysis. Our numerical experiments demonstrate that we are indeed able to achieve satisfactory numerical reconstructions for both the initial data and the heat radiative coefficient (section 6), where the adopted norm is much weaker and more practical than the norm adopted for our theoretical stability analysis. The gap here between both norms is a finite order of differentiation, so we conjecture that we should have the Hölder stability for determining the radiative coefficient $p(x)$ in the case where we choose the same norm (cf (3.1)–(3.3)). This may be done using the Sobolev interpolation theory, and will be studied elsewhere. Furthermore, in a succeeding paper, we will give an *a priori* strategy for choosing regularizing parameters based on our stability results. For some existing strategies, see, for example, [9]. As we know, the realization of the *a priori* strategy is often a hard problem for actual inverse problems.

To the authors' knowledge, there exist few publications which discuss numerical reconstruction grounded on the theoretical stability analysis for the inverse problem with

restricted measurements such as $u(x,t)$ for $x \in \omega \subset \Omega$ and $0 < t < T$. For other types of inverse parabolic problems, we refer to [3, 13, 18, 20]. For references on analytical and numerical methods for solving the inverse elliptic problems, see [1, 5, 9, 12, 17] and the references therein.

Our simultaneous reconstruction is carried out in such a way that the temperature solution of the heat conduction equation matches its subregion observation data z and the fixed time observation z_θ optimally in the L^2-norm sense, with the help of a Tikhonov regularization of L^2-norms of the gradients for both radiative coefficients and initial temperatures (section 3). We will discretize the continuous nonlinear optimization system and the parabolic equation by the piecewise linear finite element method, and prove the existence of discrete minimizers of the finite element system and their convergence to the global minimizers of the continuous optimization problem (section 4). Then the nonlinear finite element minimization problem is solved using a gradient method, which usually converges rather slowly, though very stably in most instances. For accelerating the entire reconstruction process, we will formulate a nonlinear multigrid gradient method (MGM) (section 5). Numerical experiments will be given to demonstrate the efficiency of the proposed nonlinear MGM for the simultaneous reconstruction of the thermal radiative coefficient and the initial temperature status (section 6). In particular, we would like to stress that our numerical reconstructions are also quite satisfactory in two dimensions.

2. Stability of the inverse problem

In this section, we present some conditional stability results for the inverse problem formulated in section 1. Such stability is fundamental for our subsequent numerical reconstruction of the initial temperature and the heat radiative coefficient in the heat conduction system. The stability was discussed in Choulli and Yamamoto [7] and Choulli *et al* [6] for a heat conductive problem with homogeneous Neumann boundary conditions. Here we will prove the conditional stability for the system (1.1), (1.2) with the Dirichlet boundary condition (1.3), which appears to be much more technical from the numerical analysis point of view. However, the stability established here is still not perfectly adjusted to the same case as adopted in our later numerical reconstruction. The major difference lies in the considered solution classes: we adopt H^1-solutions in the numerical reconstruction while the stability of this section will mainly consider the sufficiently smooth solutions for simplicity. The stability analysis for the H^1-solution case is much more complicated and will be considered elsewhere.

Throughout this section, for any $\gamma \in (0,1)$ and non-negative integer m, we use

$$C^{\gamma+m}(\bar{\Omega}), \qquad C^{\gamma+m,\frac{\gamma+m}{2}}(\bar{\Omega}\times[0,T]), \qquad C^{\gamma+m,\frac{\gamma+m}{2}}(\partial\Omega\times[0,T])$$

to denote the usual Hölder spaces (cf [10,21]). We assume that $\partial\Omega$ is of class $C^{\gamma+3}$ with some $\gamma \in (0,1)$ and the boundary function $\eta = \eta(x,t)$ in (1.3) satisfies

$$\eta \in C^{\gamma+3,\frac{\gamma+3}{2}}(\partial\Omega\times[0,T]), \qquad \eta(x,t) > \eta_0 \qquad \text{on} \quad \partial\Omega\times[0,T] \tag{2.1}$$

with some constant $\eta_0 > 0$. For any fixed $M > 0$ and $r_0 > 0$, we set

$$\begin{aligned}\mathcal{U} = \{&(p,\mu) \in C^{\gamma+1}(\bar{\Omega})\times C^{\gamma+3}(\bar{\Omega});\ \mu_{|\partial\Omega} = \eta(\cdot,0),\ (\Delta\mu + p\mu)_{|\partial\Omega}\\ &= \eta_t(\cdot,0),\ \|p\|_{C^{\gamma+1}(\bar{\Omega})} \leqslant M,\ \|\mu\|_{C^{\gamma+3}(\bar{\Omega})} \leqslant M,\ \mu(x) \geqslant r_0 > 0 \text{ in } \bar{\Omega}\}.\end{aligned} \tag{2.2}$$

Then, by the classical parabolic theory (cf [21]), for any $(p,\mu) \in \mathcal{U}$, there exists a unique solution $u(p,\mu) \in C^{\gamma+3,\frac{\gamma+3}{2}}(\bar{\Omega}\times[0,T])$ to the problem (1.1)–(1.3). Let $\omega \subset \Omega$ be any subdomain, and $\theta \in (0,T)$ be fixed. The inverse problem considered here is to determine

$p = p(x)$ and $\mu = \mu(x)$ in Ω from the measurement given in (1.4). This section aims to establish a stability estimate for the inverse parabolic problem.

The main result of this section will answer the stability of the inverse problem with the Dirichlet observation data $u_{|\omega\times(0,T)}$ within the class of sufficiently smooth solutions for the case of $\theta > 0$. For any $(p, \mu), (q, \nu) \in \mathcal{U}$, let

$$E((p, \mu), (q, \nu)) = \|u(p, \mu) - u(q, \nu)\|_{H^1(0,T;L^2(\omega))} + \|(u(p, \mu) - u(q, \nu))(\cdot, \theta)\|_{H^2(\Omega)}.$$

Our main result of this section is stated in the following theorem.

Theorem 2.1. *There exists a constant $C = C(M, \Omega, T, \theta, \eta, \gamma, \omega) > 0$ such that*

$$\|p - q\|_{L^2(\Omega)} \leqslant C\, E((p, \mu), (q, \nu)), \tag{2.3}$$

$$\|\mu - \nu\|_{L^2(\Omega)} \leqslant C\, \{|\log E((p, \mu), (q, \nu))|\}^{-1} \tag{2.4}$$

for all $(p, \mu), (q, \nu) \in \mathcal{U}$.

Remark 2.1. The determination of initial values involves the backward parabolic equation, and so one cannot expect good stability in general. However, the determination of heat radiative coefficients seems much more stable. In the case where $u_{|\omega\times(0,T)}$ is replaced by the Neumann data on any part of $\partial\Omega$, one can prove similar stability.

Remark 2.2. One can also take the Neumann data $\frac{\partial u}{\partial n}|_{\Gamma\times(0,T)}$ as the observation data in place of $u_{|\omega\times(0,T)}$, with $\Gamma \subset \partial\Omega$ being an arbitrary relatively open subset. Bukhgeim [2] and Klibanov [19] studied this case and proved the uniqueness of the inverse problem for $\theta = 0$, provided that Γ is a sufficiently large part of $\partial\Omega$. When Γ does not occupy a sufficient large part of $\partial\Omega$, the uniqueness is still open. On the other hand, for $\theta > 0$, Isakov [16] proved the uniqueness of the inverse problem with the above-mentioned Neumann data on an arbitrarily small part Γ over any time interval $(0, T)$.

Proof of theorem 2.1. We divide the proof into four steps.

First step. We first establish a Carleman estimate. Set

$$(Pu)(x, t) = u_t(x, t) - \Delta u(x, t) + \sum_{i=1}^{d} a_i(x, t)\frac{\partial u}{\partial x_i} + b(x, t)u, \qquad (x, t) \in Q$$

where $Q = \Omega \times (0, T)$ and the coefficients $\{a_i\}_{i=1}^d$ and b satisfy

$$\|b\|_{L^\infty(Q)} \leqslant M_1, \qquad \|a_i\|_{L^\infty(Q)} \leqslant M_1, \qquad i = 1, 2, \ldots, d. \tag{2.5}$$

The next lemma can be found in Fursikov and Imanuvilov [11] and Imanuvilov [15].

Lemma 2.1. *Let ω_0 be an arbitrary subdomain of Ω such that $\bar{\omega}_0 \subset \omega$. Then there exists a function $\psi \in C^2(\overline{\Omega})$ such that $\psi(x) = 0$ for $x \in \partial\Omega$ and*

$$\psi(x) > 0, \qquad x \in \Omega; \qquad |\nabla\psi(x)| > 0, \qquad x \in \overline{\Omega \setminus \omega_0}. \tag{2.6}$$

The following lemma is a Carleman estimate due to [11, 15].

Lemma 2.2. *There exists a constant $\hat{\lambda} = \hat{\lambda}(M_1) > 0$ such that for any $\lambda \geqslant \hat{\lambda}$ we can choose $s_0(\lambda) > 0$ and a constant $C = C(\lambda, M_1) > 0$, independent of each choice $\{a_i\}_{i=1}^d$ and b, such that for all $s \geqslant s_0(\lambda)$, the following inequality holds:*

$$\int_Q \left\{\frac{1}{s\varphi}\left(|z_t|^2 + \sum_{i,j=1}^{d}\left|\frac{\partial^2 z}{\partial x_i \partial x_j}\right|^2\right) + s\varphi|\nabla z|^2 + s^3\varphi^3 z^2\right\} e^{2s\alpha}\, dx\, dt$$
$$\leqslant C\int_Q |Pz|^2 e^{2s\alpha}\, dx\, dt + C\int_{Q_\omega} s^3\varphi^3 z^2 e^{2s\alpha}\, dx\, dt$$

for all $z \in W_2^{1,2}(Q)$ *with* $z(x) = 0$ *for* $x \in \partial\Omega$. *Here*

$$\varphi(t,x) = \frac{e^{\lambda\psi(x)}}{t(T-t)}, \qquad \alpha(t,x) = \frac{e^{\lambda\psi(x)} - e^{2\lambda\|\psi\|_{C(\overline{\Omega})}}}{t(T-t)}.$$

Remark 2.3. The Carleman estimate is true for a general parabolic operator of second order with coefficients depending on x and t (cf [15]). For a similar Carleman estimate, we refer to Tataru [26] where only a sketch of the proof was given.

Second step. First by the maximum principle (cf [21]), we have for any $x \in \overline{\Omega}$ and $t \in (0,T)$,

$$u(p,\mu)(x,t) \geqslant e^{-MT} \min\{\min_{x\in\overline{\Omega}} \mu(x), \min_{x\in\partial\Omega, t\geqslant 0} \eta(x,t)\} \geqslant e^{-MT} \min\{\eta_0, r_0\}. \tag{2.7}$$

So $u(p,\mu)$ is always positive. For simplicity, we will translate the time variable and consider the problem in the time interval $(-\delta, T)$ with $\delta > 0$:

$$\begin{aligned} &u_t(x,t) = \Delta u(x,t) + p(x)u(x,t), && x \in \Omega, \quad -\delta < t < T \\ &u(x,-\delta) = \mu(x), && x \in \Omega \\ &u(x,t) = \eta(x,t), && x \in \partial\Omega, \quad -\delta < t < T. \end{aligned} \tag{2.8}$$

Moreover, we can set $\theta = \frac{T}{2}$ after a suitable translation of time variable. Then using the compatibility condition of the first order following from $(p,\mu) \in \mathcal{U}$ and the classical parabolic theory (cf [10, 21]), we have the following lemma.

Lemma 2.3. *There exists a constant* $M_2 = M_2(\mathcal{U}, \delta, T, \Omega) > 0$ *such that*

$$\|u(p,\mu)\|_{C^{\gamma+3,\frac{\gamma+3}{2}}(\bar{\Omega}\times[-\delta,T])} \leqslant M_2, \qquad \forall\,(p,\mu) \in \mathcal{U}.$$

Third step. We now prove (2.3). Let

$$\begin{aligned} &w = u(p,\mu) - u(q,\nu), && f = p - q, \quad R = u(q,\nu), \\ &a = u(p,\mu)(\cdot,\theta) - u(q,\nu)(\cdot,\theta). \end{aligned} \tag{2.9}$$

Then w satisfies

$$\begin{aligned} &w_t = \Delta w + pw + fR, && (x,t) \in \Omega \times (-\delta, T) \\ &w(x,\theta) = a(x), && x \in \Omega \\ &w = 0 && (x,t) \in \partial\Omega \times (-\delta, T). \end{aligned} \tag{2.10}$$

Consider $z = w/R$: it is well defined by (2.7). Using the fact that $\frac{\partial R}{\partial t} = \Delta R + qR$, we have

$$\begin{aligned} &z_t = \Delta z + \frac{2\nabla R}{R}\cdot\nabla z + (p-q)z + f && \text{in} \quad \Omega \times (-\delta, T) \\ &z(x,\theta) = \frac{a(x)}{R(x,\theta)}, && x \in \Omega \\ &z = 0 && \text{on} \quad \partial\Omega \times (-\delta, T) \end{aligned} \tag{2.11}$$

and $y = z_t$ satisfies

$$\begin{aligned} &y_t = \Delta y + \frac{2\nabla R}{R}\cdot\nabla y + (p-q)y + \frac{\partial}{\partial t}\left(\frac{2\nabla R}{R}\right)\cdot\nabla z && \text{in} \quad \Omega \times (-\delta, T) \\ &y = 0 && \text{on} \quad \partial\Omega \times (-\delta, T). \end{aligned} \tag{2.12}$$

By lemma 2.3, we know that $\frac{2\nabla R}{R}$ is uniformly bounded for $(q, \mu) \in \mathcal{U}$. Then we obtain by applying lemma 2.2 to (2.12) that, for large $s > 0$,

$$\int_Q \left(\frac{1}{s\varphi}(|y_t|^2 + |\Delta y|^2) + s\varphi|\nabla y|^2 + s^3\varphi^3 y^2 \right) \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t$$
$$\leqslant C \int_Q \left| \frac{\partial}{\partial t} \left(\frac{2\nabla R}{R} \right) \cdot \nabla z \right|^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t + C \int_{Q_\omega} s^3\varphi^3 y^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t. \tag{2.13}$$

For the first term of the right-hand side, we have using $z(x, \theta) = a(x)/R(x, \theta)$ that

$$z(x, t) = \int_\theta^t y(x, \xi) \, \mathrm{d}\xi + \frac{a(x)}{R(x, \theta)}. \tag{2.14}$$

Then by lemma 2.3,

$$\int_Q \left| \frac{\partial}{\partial t} \left(\frac{2\nabla R}{R} \right) \cdot \nabla z \right|^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t$$
$$\leqslant C_1 \int_Q \left| \int_\theta^t |\nabla y(x, \xi)| \, \mathrm{d}\xi \right|^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t + C_1 \int_Q \left| \nabla \left(\frac{a(x)}{R(x, \theta)} \right) \right|^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t$$
$$\leqslant C_1 \int_Q \left| \int_\theta^t |\nabla y(x, \xi)| \, \mathrm{d}\xi \right|^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t + C_1(\|a\mathrm{e}^{s\alpha}\|^2_{L^2(Q)} + \|(\nabla a)\mathrm{e}^{s\alpha}\|^2_{L^2(Q)}).$$

By a direct manipulation, we can show that for any $\alpha \in C^1(\overline{Q})$ satisfying

$$\left(t - \frac{T}{2} \right) \frac{\partial \alpha}{\partial t} \leqslant 0, \qquad x \in \overline{\Omega}, \quad 0 \leqslant t \leqslant T,$$

the following inequality holds for all $u \in L^2(Q)$:

$$\int_Q \left| \int_\theta^t |u(x, \xi)| \, \mathrm{d}\xi \right|^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t \leqslant C_2(T) \int_Q |u(x, t)|^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t.$$

(This type of inequality is essential for applications of Carleman estimates to inverse problems: see [2, 16, 19], for example.) This implies

$$\int_Q \left| \frac{\partial}{\partial t} \left(\frac{2\nabla R}{R} \right) \cdot \nabla z \right|^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t$$
$$\leqslant C_2 \int_Q |\nabla y(x, t)|^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t + C_2(\|a\mathrm{e}^{s\alpha}\|^2_{L^2(Q)} + \|(\nabla a)\mathrm{e}^{s\alpha}\|^2_{L^2(Q)})$$

with $C_2 > 0$ independent of $s > 0$. Thus, taking large $s > 0$, we obtain from (2.13) that

$$\int_Q \left(\frac{1}{s\varphi}(|y_t|^2 + |\Delta y|^2) + s\varphi|\nabla y|^2 + s^3\varphi^3 y^2 \right) \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t$$
$$\leqslant C \int_{Q_\omega} s^3\varphi^3 y^2 \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t + C(\|a\mathrm{e}^{s\alpha}\|^2_{L^2(Q)} + \|(\nabla a)\mathrm{e}^{s\alpha}\|^2_{L^2(Q)})$$

with $Q_\omega = \omega \times (0, T)$. For such fixed $s > 0$, we have $\mathrm{e}^{2s\alpha}, s^3\varphi^3\mathrm{e}^{2s\alpha} \leqslant C_3$ in Q and so

$$\int_{\frac{T}{4}}^{\frac{3T}{4}} \int_\Omega \left(\frac{1}{s\varphi}(|y_t|^2 + |\Delta y|^2) + s\varphi|\nabla y|^2 + s^3\varphi^3 y^2 \right) \mathrm{e}^{2s\alpha} \, \mathrm{d}x \, \mathrm{d}t$$
$$\leqslant C \int_{Q_\omega} |z_t|^2 \, \mathrm{d}x \, \mathrm{d}t + C\|a\|^2_{H^1(\Omega)}.$$

Moreover, noting that

$$\frac{1}{s\varphi}e^{2s\alpha}, s\varphi e^{2s\alpha}, s^3\varphi^3 e^{2s\alpha} \geqslant C_4 > 0 \qquad \text{in} \quad \Omega \times \left(\frac{T}{4}, \frac{3T}{4}\right)$$

for the fixed $s > 0$, we have

$$\int_{\frac{T}{4}}^{\frac{3T}{4}} \int_\Omega (|y_t|^2 + |\Delta y|^2 + |\nabla y|^2 + y^2)\,dx\,dt \leqslant C \int_{Q_\omega} \Big\{ |u(p,\mu) - u(q,\nu)|^2 + \left| \frac{\partial}{\partial t}(u(p,\mu) - u(q,\nu)) \right|^2 \Big\} dx\,dt + C\|a\|^2_{H^1(\Omega)}. \tag{2.15}$$

Furthermore, by (2.11), (2.14) and $y = z_t$ we have

$$\begin{aligned}
\frac{T}{2} f &= \int_{\frac{T}{4}}^{\frac{3T}{4}} \left(z_t - \Delta z - \frac{2\nabla R}{R} \cdot \nabla z - (p-q)z \right) dt \\
&= \int_{\frac{T}{4}}^{\frac{3T}{4}} \left(y - \int_\theta^t \Delta y(x,\xi)\,d\xi - \frac{2\nabla R}{R} \cdot \int_\theta^t \nabla y(x,\xi)\,d\xi \right. \\
&\quad \left. -(p-q)\int_\theta^t y(x,\xi)\,d\xi \right) dt - \frac{T}{2}\left(\Delta\left(\frac{a(x)}{R(x,\theta)} \right) + (p-q)\frac{a(x)}{R(x,\theta)} \right) \\
&\quad - \int_{\frac{T}{4}}^{\frac{3T}{4}} \frac{2\nabla R(x,t)}{R(x,t)} \cdot \nabla \left(\frac{a(x)}{R(x,\theta)} \right) dt.
\end{aligned}$$

Then by lemma 2.3 and the fact that $(p,\mu), (q,\nu) \in \mathcal{U}$, we further deduce

$$\begin{aligned}
\frac{T}{2}\|f\|^2_{L^2(\Omega)} &\leqslant \int_\Omega \left| \int_{\frac{T}{4}}^{\frac{3T}{4}} \left(y - \int_\theta^t \Delta y(x,\xi)\,d\xi - \frac{2\nabla R}{R} \cdot \int_\theta^t \nabla y(x,\xi)\,d\xi \right. \right. \\
&\quad \left. \left. -(p-q)\int_\theta^t y(x,\xi)\,d\xi \right) dt \right|^2 dx + C_5(\|a\|^2_{H^2(\Omega)} + \|p-q\|^2_{L^\infty(\Omega)}\|a\|^2_{L^2(\Omega)}) \\
&\leqslant C_5 \int_{\frac{T}{4}}^{\frac{3T}{4}} \int_\Omega (y^2 + |\Delta y|^2 + |\nabla y|^2 + \|p-q\|^2_{L^\infty(\Omega)} y^2)\,dx\,dt \\
&\quad + C_5(\|a\|^2_{H^2(\Omega)} + \|p-q\|^2_{L^\infty(\Omega)}\|a\|^2_{L^2(\Omega)}).
\end{aligned}$$

Now (2.3) follows from this and (2.15).

Fourth step. Finally, we prove (2.4). Setting $w_1 = w_t$ in (2.10), we have

$$\begin{aligned}
&(w_1)_t = \Delta w_1 + p(x)w_1 + f(x)R_t(x,t), && x \in \Omega, \quad -\delta < t < T \\
&w_1(x,\theta) = \Delta a(x) + p(x)a(x) + f(x)R(x,\theta) \equiv b(x), && x \in \Omega \\
&w_1(x,t) = 0, && x \in \partial\Omega, \quad -\delta < t < T.
\end{aligned} \tag{2.16}$$

We then decompose (2.16) as follows:

$$\begin{aligned}
&v_t = \Delta v + p(x)v + f(x)R_t(x,t), && x \in \Omega, \quad -\delta < t < T \\
&v(x,-\delta) = 0, && x \in \Omega \\
&v(x,t) = 0, && x \in \partial\Omega, \quad -\delta < t < T
\end{aligned} \tag{2.17}$$

and

$$\begin{aligned}
&u_t = \Delta u + p(x)u(x,t), && x \in \Omega, \quad -\delta < t < T \\
&u(x,\theta) = b(x) - v(x,\theta), && x \in \Omega \\
&u(x,t) = 0, && x \in \partial\Omega, \quad -\delta < t < T.
\end{aligned} \tag{2.18}$$

We can verify that

$$w_1 = u + v; \qquad w_1(x, -\delta) = u(x, -\delta), \qquad x \in \Omega. \tag{2.19}$$

Noting that $w_1 = w_t$, thus $w_1(\cdot, -\delta) = w_t(\cdot, -\delta)$, and we have by lemma 2.3 that

$$\|w_1(\cdot, -\delta)\|_{L^\infty(\Omega)} \leqslant C_6 \|u(p, \mu) - u(q, \nu)\|_{C^{\gamma+3, \frac{\gamma+3}{2}}(\overline{Q})} \leqslant C_6 M_2,$$

which, with (2.19), implies

$$\|u(\cdot, -\delta)\|_{L^\infty(\Omega)} \leqslant C_6 M_2.$$

Thus we can apply the method of the logarithmic convexity (cf Payne [23]) to obtain

$$\|u(\cdot, t)\|_{L^2(\Omega)} \leqslant (C_6 M_2)^{1-\frac{t+\delta}{\theta+\delta}} \|u(\cdot, \theta)\|_{L^2(\Omega)}^{\frac{t+\delta}{\theta+\delta}}, \qquad -\delta \leqslant t \leqslant \theta. \tag{2.20}$$

Furthermore, by the semigroup theory (e.g. Pazy [24]), we can write

$$v(t) = v(\cdot, t) = \int_{-\delta}^{t} \mathrm{e}^{-(t-\xi)A} p v(\xi)\, \mathrm{d}\xi + \int_{-\delta}^{t} \mathrm{e}^{-(t-\xi)A} f R_\xi(\xi)\, \mathrm{d}\xi, \qquad -\delta < t < \theta.$$

Here $A = -\Delta$ with $\mathcal{D}(A) = H^2(\Omega) \cap H_0^1(\Omega)$. By $\|p\|_{L^\infty(\Omega)} \leqslant M$ and lemma 2.3, we have

$$\|v(t)\|_{L^2(\Omega)} \leqslant C_7 \int_{-\delta}^{t} \|v(\xi)\|_{L^2(\Omega)}\, \mathrm{d}\xi + C_7 \|f\|_{L^2(\Omega)}, \qquad -\delta < t < \theta.$$

The Gronwall inequality then yields

$$\|v(t)\|_{L^2(\Omega)} \leqslant C_8 \|f\|_{L^2(\Omega)}, \qquad -\delta \leqslant t \leqslant \theta. \tag{2.21}$$

In view of (2.16), (2.18) and (2.19), the inequalities (2.20) and (2.21) imply

$$\begin{aligned} \|w_t(\cdot, t)\|_{L^2(\Omega)} &\leqslant \|u(\cdot, t)\|_{L^2(\Omega)} + \|v(\cdot, t)\|_{L^2(\Omega)} \\ &\leqslant C_9 \|u(\cdot, \theta)\|_{L^2(\Omega)}^{\frac{t+\delta}{\theta+\delta}} + C_9 \|f\|_{L^2(\Omega)} \\ &\leqslant C_9 (\|a\|_{H^2(\Omega)} + \|f\|_{L^2(\Omega)})^{\frac{t+\delta}{\theta+\delta}} + C_9 \|f\|_{L^2(\Omega)}, \end{aligned} \tag{2.22}$$

which yields

$$\begin{aligned} \|\mu - \nu\|_{L^2(\Omega)} &= \|u(p, \mu)(\cdot, -\delta) - u(q, \nu)(\cdot, -\delta)\|_{L^2(\Omega)} = \|w(\cdot, -\delta)\|_{L^2(\Omega)} \\ &= \left\| - \int_{-\delta}^{\theta} w_t(\cdot, \xi)\, \mathrm{d}\xi + w(\cdot, \theta) \right\|_{L^2(\Omega)} \\ &\leqslant C_9 \int_{-\delta}^{\theta} (\|a\|_{H^2(\Omega)} + \|f\|_{L^2(\Omega)})^{\frac{\xi+\delta}{\theta+\delta}}\, \mathrm{d}\xi + C_9 \|f\|_{L^2(\Omega)} (\theta + \delta) + \|a\|_{L^2(\Omega)} \\ &= C_9 (\theta + \delta) \frac{|1 - (\|a\|_{H^2(\Omega)} + \|f\|_{L^2(\Omega)})|}{|\log(\|a\|_{H^2(\Omega)} + \|f\|_{L^2(\Omega)})|} + C_9 \|f\|_{L^2(\Omega)} (\theta + \delta) + \|a\|_{L^2(\Omega)} \\ &\leqslant \frac{C_{10}}{|\log(\|a\|_{H^2(\Omega)} + \|f\|_{L^2(\Omega)})|}. \end{aligned}$$

This with (2.3) leads to (2.4), and so completes the proof of theorem 2.1. □

3. The continuous formulation

From the stability established in section 2 we know that the inverse problem of reconstructing the initial temperature and the heat radiative coefficient is conditionally stable under a smooth class of solutions to the corresponding parabolic system and the observation data $u \in H^1(0, T; L^2(\omega))$ and $u(\cdot, \theta) \in H^2(\Omega)$. Clearly, such a requirement on the excessive amount of observation data is not very practical, but these stability results give us some important guidance for the possible numerical reconstruction. Moreover, taking into consideration the finiteness of the gap between the order of the norms for the measurement data adopted for theorem 2.1 and the one taken in the subsequent numerical analysis, it is very likely to derive some stability of Hölder type in determining the heat radiative coefficient for the more practical requirement on the measurement data to be discussed below. But the analysis for this case is much more technical and complicated and will be left for our forthcoming work.

For the simultaneous numerical reconstruction, we assume that only the perturbed observation data $z_\theta(x)$ and $z(x, t)$ given in (1.4) are available. Then we formulate the reconstruction of the initial temperature μ and the thermal radiative coefficient p in (1.1), (1.2) as the following constrained minimizing process: minimize

$$J(p, \mu) = \tfrac{1}{2} \int_0^T \int_\omega (u(p, \mu) - z)^2 \, \mathrm{d}x \, \mathrm{d}t + \tfrac{1}{2} \int_\Omega (u(p, \mu)(x, \theta) - z_\theta(x))^2 \, \mathrm{d}x + \beta \int_\Omega |\nabla p|^2 \, \mathrm{d}x + \gamma \int_\Omega |\nabla \mu|^2 \, \mathrm{d}x \tag{3.1}$$

subject to $p \in K_1$, $\mu \in K_2$ and $u(\cdot, t) \equiv u(p, \mu)(\cdot, t) \in H^1(\Omega)$ satisfying

$$u(x, 0) = \mu(x) \qquad \text{in} \quad \Omega; \qquad u(x, t) = \eta(x, t) \qquad \text{on} \quad \partial\Omega \times (0, T) \tag{3.2}$$

$$\int_\Omega u_t \, \phi \, \mathrm{d}x + \int_\Omega \nabla u \cdot \nabla \phi \, \mathrm{d}x = \int_\Omega p(x) \, u \, \phi \, \mathrm{d}x \qquad \forall \, \phi \in H_0^1(\Omega) \tag{3.3}$$

for a.e. $t \in (0, T)$. Note that the system (3.2), (3.3) is the variational formulation associated with the parabolic problem (1.1)–(1.3). In what follows, we denote the solution of this variational problem as $u(p, \mu)$ or $u(p, \mu)(x, t)$. The measurement functions z and z_θ satisfy that $z \in L^2(0, T; L^2(\omega))$ and $z_\theta \in L^2(\Omega)$, the positive constants β and γ are regularization parameters. The constraint sets K_1 and K_2 above are chosen to be as follows:

$$K_1 = \{p \in H^1(\Omega);\ |p(x)| \leqslant \alpha_1 \text{ a.e. in } \Omega\},$$
$$K_2 = \{\mu \in H^1(\Omega);\ 0 < \mu(x) \leqslant \alpha_2 \text{ a.e. in } \Omega\}.$$

Here α_1 and α_2 are two positive constants.

For the minimization problem (3.1)–(3.3), we can show the following theorem.

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

Proof. It is easy to verify that $\min J(p, \mu)$ is finite over $K_1 \times K_2$, and thus there exists a minimizing sequence $\{p_n, \mu_n\}$ from $K_1 \times K_2$ such that

$$\lim_{n \to \infty} J(p_n, \mu_n) = \inf_{\substack{p \in K_1 \\ \mu \in K_2}} J(p, \mu).$$

This implies the boundedness of $\{p_n, \mu_n\}$ in $H^1(\Omega) \times H^1(\Omega)$ and therefore there is a subsequence (still denoted as $\{p_n, \mu_n\}$) such that both p_n and μ_n converge weakly to p^* and μ^* respectively in $H^1(\Omega)$. Similarly to the proof of lemma 2.1 in Keung and Zou [18],

we can show that $u(p_n, \mu_n)$ converges weakly to $u(p^*, \mu^*)$ in $L^2(0, T; H^1(\Omega))$ and

$$\lim_{n\to\infty} \int_0^T \int_\omega (u(p_n, \mu_n) - z)^2 \, \mathrm{d}x \, \mathrm{d}t = \int_0^T \int_\omega (u(p^*, \mu^*) - z)^2 \, \mathrm{d}x \, \mathrm{d}t,$$
$$\lim_{n\to\infty} \int_\Omega (u(p_n, \mu_n)(x, \theta) - z_\theta(x))^2 \, \mathrm{d}x = \int_\Omega (u(p^*, \mu^*)(x, \theta) - z_\theta(x))^2 \, \mathrm{d}x.$$

Now the proof of (p^*, μ^*) being a minimizer of J over $K_1 \times K_2$ is routine (cf [18]). □

4. Discretization and its convergence

We now propose a finite element method for solving the continuous minimization problem (3.1)–(3.3). For the purpose, we triangulate the polyhedral domain Ω with a regular triangulation $\mathcal{T}^h$ of simplicial elements (cf Ciarlet [8]), then define the finite element space V_h to be the space of all continuous piecewise linear functions over $\mathcal{T}^h$, and $\overset{\circ}{V}_h$ a subspace of V_h with all functions vanishing on the boundary $\partial\Omega$. Let $\mathcal{N}_h = \{x_i\}_{i=1}^{N_h}$ be the set of all nodal points of $\mathcal{T}^h$. Then we approximate the constraint subsets K_1 and K_2 by

$$K_{1h} = \{p_h \in V_h; |p_h(x_i)| \leqslant \alpha_1 \text{ for } x_i \in \mathcal{N}_h\},$$
$$K_{2h} = \{\mu_h \in V_h; 0 < \mu_h(x_i) \leqslant \alpha_2 \text{ for } x_i \in \mathcal{N}_h\}.$$

For the time discretization, we partition the interval $[0, T]$ using the following 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 its difference quotient and the averaging $\bar{u}^n$ of a given function $u(\cdot, t)$ as follows:

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

where for $n = 0$ we let $\bar{u}^0 = u(\cdot, 0)$.

For our later analysis, we now introduce two discrete projection operators. The first is a quasi-L^2 projection operator π_h: $L^2(\Omega) \to V_h$ (see [28, 29], for example), which possesses the following properties (cf Xu [28]):

$$\|\pi_h w\|_{L^2(\Omega)} \leqslant C\|w\|_{L^2(\Omega)}, \qquad \lim_{h\to 0} \|w - \pi_h w\|_{L^2(\Omega)} = 0, \qquad \forall\, w \in L^2(\Omega); \tag{4.1}$$

$$\|\pi_h w\|_{H^1(\Omega)} \leqslant C\|w\|_{H^1(\Omega)}, \qquad \lim_{h\to 0} \|w - \pi_h w\|_{H^1(\Omega)} = 0, \qquad \forall\, w \in H^1(\Omega), \tag{4.2}$$

and we have $\pi_h w \in K_{1h}$ (resp. K_{2h}) for any $w \in K_1$ (resp. K_2). This property is crucial to our later convergence analysis. It is important to note that the standard L^2 projection does not have such a nice property [4].

The second operator we need is a discrete operator $Q_h : H^1(\Omega) \to V_h$ which preserves the boundary values, namely $Q_h w = Q_h v$ on $\partial\Omega$ for any two functions $w, v \in H^1(\Omega)$ satisfying $w = v$ on $\partial\Omega$. Clearly, the quasi-L^2 projection π_h does not have such a property. The standard nodal value interpolation associated with V_h has the aforementioned two properties but it is defined only for smoother functions and it does not have the stability estimates shown in (4.1), (4.2) either. Next we introduce such a boundary value preserving operator Q_h, which was first proposed by Scott–Zhang [25]. For any $x_i \in \mathcal{N}_h$ (the set of nodal points in $\mathcal{T}^h$), choose τ_i to be a $(d-1)$-simplex from the triangulation $\mathcal{T}^h$ with vertices z_l $(l = 1, \ldots, d)$

such that $z_1 = x_i$. The choice of τ_i is not unique in general, but if $x_i \in \partial\Omega$, we take $\tau_i \subset \partial\Omega$. Let $\theta_i \in \mathcal{P}_1(\tau_i)$, the set of all linear polynomials on τ_i, be the unique function satisfying

$$\int_{\tau_i} \theta_i \lambda_l \, \mathrm{d}x = \delta_{l1}, \qquad l = 1, \ldots, d,$$

where λ_l is the barycentric coordinate of τ_i (see Ciarlet [8]) with respect to z_l. Obviously,

$$\int_{\tau_i} \theta_i \, v \, \mathrm{d}x = v(x_i) \qquad \text{for any} \quad v \in \mathcal{P}_1(\tau_i).$$

The average nodal value interpolant Q_h is then defined by

$$Q_h v(x) = \sum_{i=1}^{N_h} \phi_i(x) \int_{\tau_i} \theta_i \, v \, \mathrm{d}x. \tag{4.3}$$

For any $v \in H^{1/2}(\partial\Omega)$, we define

$$Q_h^b v(x) = \sum_{x_i \in \partial\Omega} \phi_i(x) \int_{\tau_i} \theta_i \, v \, \mathrm{d}x \equiv Q_h v(x) \qquad \forall\, x \in \partial\Omega. \tag{4.4}$$

Note that the above definitions for Q_h and Q_h^b do not make sense for $n = 1$. In this case, we let Q_h and Q_h^b be the standard nodal value interpolant I_h on V_h, which is well defined since $H^1(\Omega) \subset C^0(\bar{\Omega})$ for $n = 1$.

Lemma 4.1 (cf [25]). *The operator Q_h defined by (4.3) satisfies:*

(a) $Q_h w = Q_h v$ *on* $\partial\Omega$ *if* $w, v \in H^1(\Omega)$ *and* $w = v$ *on* $\partial\Omega$,
(b) $|v - Q_h v|_{H^t(\Omega)} \leqslant h^{s-t} |v|_{H^s(\Omega)} \forall\, v \in H^s(\Omega)$ $(s = 1, 2;\ t = 0, 1)$,
(c) $|Q_h v|_{H^1(\Omega)} \leqslant |v|_{H^1(\Omega)}$ *and* $\|Q_h v\|_{H^1(\Omega)} \leqslant \|v\|_{H^1(\Omega)} \forall\, v \in H^1(\Omega)$.

With the above notation and the discretization of the first term of (3.1) by the composite trapezoidal rule in time, we formulate the finite element approximation of the problem (3.1)–(3.3) as follows: minimize

$$J_h(p_h, \mu_h) = \frac{\tau}{2} \sum_{n=0}^{M} \alpha_n \int_\omega (u_h^n(p_h, \mu_h) - z^n)^2 \, \mathrm{d}x + \frac{1}{2} \int_\Omega (u_h^{n_0}(p_h, \mu_h) - z_\theta)^2 \, \mathrm{d}x$$
$$+ \beta \int_\Omega |\nabla p_h|^2 \, \mathrm{d}x + \gamma \int_\Omega |\nabla \mu_h|^2 \, \mathrm{d}x \tag{4.5}$$

over all $p_h \in K_{1h}$ and $\mu_h \in K_{2h}$ with $u_h^n \equiv u_h^n(p_h, \mu_h) \in V_h$ satisfying

$$u_h^0 = \mu_h \qquad \text{and} \qquad u_h^n = Q_h^b \bar{\eta}^n + \hat{u}_h^n, \tag{4.6}$$

$$\int_\Omega \partial_\tau u_h^n \, \phi_h \, \mathrm{d}x + \int_\Omega \nabla u_h^n \cdot \nabla \phi_h \, \mathrm{d}x = \int_\Omega p_h \, u_h^n \, \phi_h \qquad \forall\, \phi_h \in \mathring{V}_h \tag{4.7}$$

for $n = 1, 2, \ldots, M$. Here $\alpha_0 = \alpha_M = \frac{1}{2}$ and $\alpha_n = 1$ for all $n \neq 0, M$ while $\hat{u}_h^n \in \mathring{V}_h$ and $n_0 > 0$ is an integer such that $n_0 \tau = \theta$. Note that $\bar{\eta}^n$ is the average of $\eta(x, t)$ on $[t^{n-1}, t^n]$, so $Q_h^b \bar{\eta}^n$ is easy to calculate on the boundary $\partial\Omega$ using (4.4).

As for the existence of the minimizers to the finite element problem (4.5)–(4.7), we have the following theorem, which can be proved basically along the same line as given in Keung and Zou [18].

Theorem 4.1. *There exists at least a minimizer to the finite element problem (4.5)–(4.7).*

In the rest of this section we are going to prove the convergence of the discrete minimizers of (4.5) to the global minimizers of (3.1). For this purpose, we need three auxiliary lemmas, and the detailed proofs are given in [29]. The first one is about the error estimate of the composite trapezoidal rule, as presented in the following lemma.

Lemma 4.2. *For any given function $f \in H^1(0, T)$, we have*

$$E(f^2) \leqslant 2\tau \left(\int_0^T f'(t)^2 \, \mathrm{d}t \right)^{1/2} \left(\int_0^T f^2(t) \, \mathrm{d}t + \max_{t\in[0,T]} f^2(t) \right)^{1/2} \tag{4.8}$$

where $E(f^2)$ is the approximation error given by

$$E(f^2) = \int_0^T f^2(t) \, \mathrm{d}t - \frac{\tau}{2} \sum_{n=1}^{M} \{ f^2(t^{n-1}) + f^2(t^n) \}.$$

From now on, we assume that the given boundary function in (1.3) satisfies

$$\eta(x, t) \in H^1(0, T; H^{1/2}(\partial\Omega)),$$

and then extend η over the domain Ω by solving the Dirichlet problem at each time t:

$$\Delta\eta_*(x, t) = 0 \qquad \text{in} \quad \Omega; \qquad \eta_*(x, t) = \eta(x, t) \qquad \text{on} \quad \partial\Omega.$$

With this extension we have $\eta_*(x, t) \in H^1(0, T; H^1(\Omega))$. We remark that this extension is only used in the convergence analysis and will never be needed in our numerical algorithm.

Remark 4.1. Note that $\eta_*(x, t) = \eta(x, t)$ on $\partial\Omega$, so their averages over $[t^{n-1}, t^n]$ are the same, i.e. $\bar{\eta}_*^n(x) = \bar{\eta}^n(x)$ on $\partial\Omega$. Using this, we can write $Q_h\bar{\eta}_*^n = Q_h^b\bar{\eta}^n + \tilde{u}_h^n$ with $\tilde{u}_h^n \in \overset{\circ}{V}_h$. Then the solution u_h^n in (4.6) can be decomposed as

$$u_h^n = Q_h^b\bar{\eta}^n + \hat{u}_h^n = Q_h\bar{\eta}_*^n + (\hat{u}_h^n - \tilde{u}_h^n). \tag{4.9}$$

The relation (4.9) will be very important for the convergence analysis. Using (4.9), (4.7) and some careful manipulations, we can derive the following two lemmas.

Lemma 4.3. *Let $u_h^n(p_h, \mu_h)$ be the solution of the finite element problem (4.6), (4.7) corresponding to $p_h \in K_{1h}$ and $\mu_h \in K_{2h}$. Then the following stability estimates hold:*

$$\max_{1\leqslant n\leqslant M} \|u_h^n\|_{H^1(\Omega)}^2 + \tau \sum_{n=1}^{M} (\|\nabla u_h^n\|_{L^2(\Omega)}^2 + \|\partial_\tau u_h^n\|^2) \leqslant C \left(\|\mu_h\|_{H^1(\Omega)}^2 + \|\eta_*\|_{H^1(0,T;H^1(\Omega))}^2 \right)$$

with C independent of p_h, μ_h, h and τ.

Lemma 4.4. *For any sequence $\{p_h, \mu_h\}$ in $V_h \times V_h$ converging to $\{p, \mu\}$ weakly in $H^1(\Omega) \times H^1(\Omega)$ as $h \to 0$, we have*

$$\int_\Omega (u_h^{n_0}(p_h, \mu_h)(x) - z_\theta(x))^2 \, \mathrm{d}x \to \int_\Omega (u(p, \mu)(x, \theta) - z_\theta(x))^2 \, \mathrm{d}x,$$

$$\tau \sum_{n=0}^{M} \alpha_n \int_\omega (u_h^n(p_h, \mu_h) - z^n)^2 \, \mathrm{d}x \to \int_0^T \int_\omega (u(p, \mu) - z)^2 \, \mathrm{d}x \, \mathrm{d}t$$

when $h \to 0$ and $\tau \to 0$.

Finally, we have the following convergence theorem about the finite element problem (4.5)–(4.7).

Theorem 4.2. *Let $\{p_h^*, \mu_h^*\}_{h>0}$ be a sequence of minimizers to the finite element problem (4.5)–(4.7). Then each subsequence of $\{p_h^*, \mu_h^*\}_{h>0}$ has a subsequence converging in $L^2(\Omega)$ to a minimizer of the continuous problem (3.1)–(3.3). If the minimizers of the continuous problem are unique, then the whole sequence $\{p_h^*, \mu_h^*\}_{h>0}$ converges to the unique minimizer of (3.1)–(3.3).*

Proof. It is not difficult to see that $J_h(p_h^*, \mu_h^*) \leqslant C$ for some constant C independent of τ and h. This implies that $\{p_h^*\}$ and $\{\mu_h^*\}$ are both bounded in $H^1(\Omega)$ and so there exists a subsequence of $\{p_h^*\}$ and $\{\mu_h^*\}$ each, still denoted as $\{p_h^*\}$ and $\{\mu_h^*\}$, such that

$$p_h^* \to p^* \text{ weakly in } H^1(\Omega), \qquad \mu_h^* \to \mu^* \text{ weakly in } H^1(\Omega).$$

Then for any $p \in K_1$ and $\mu \in K_2$, we have by lemma 4.4 and the properties of π_h in (4.1), (4.2) that

$$\begin{aligned} J(p^*, \mu^*) &\leqslant \lim_{\substack{h\to 0\\ \tau\to 0}} \frac{\tau}{2} \sum_{n=0}^{M} \alpha_n \int_\omega (u_h^n(p_h^*, \mu_h^*) - z^n)^2 \,\mathrm{d}x + \lim_{\substack{h\to 0\\ \tau\to 0}} \frac{1}{2} \int_\Omega (u_h^{n_0}(p_h^*, \mu_h^*) - z_\theta)^2 \,\mathrm{d}x \\ &\quad + \beta \liminf_{\substack{h\to 0\\ \tau\to 0}} \|\nabla p_h^*\|^2 + \gamma \liminf_{\substack{h\to 0\\ \tau\to 0}} \|\nabla \mu_h^*\|^2 \\ &\leqslant \liminf_{\substack{h\to 0\\ \tau\to 0}} J_h(p_h^*, \mu_h^*) \leqslant \liminf_{\substack{h\to 0\\ \tau\to 0}} J_h(\pi_h p, \pi_h \mu) \\ &= \tfrac{1}{2} \int_0^T \int_\omega (u(p, \mu) - z)^2 \,\mathrm{d}x \,\mathrm{d}t + \tfrac{1}{2} \int_\Omega (u(x, \theta) - z_\theta(x))^2 \,\mathrm{d}x \\ &\quad + \beta \|\nabla p\|^2 + \gamma \|\nabla \mu\|^2 = J(p, \mu), \end{aligned}$$

that is, $\{p_h^*, \mu_h^*\}$ is a minimizer of the problem (3.1)–(3.3). The proof of the last statement is standard. □

5. Numerical algorithms

This section is devoted to some iterative methods for solving the discretized finite element minimization of $J_h(p_h, \mu_h)$ over K_{1h} and K_{2h} in (4.5)–(4.7). For the purpose we need to calculate the Gateaux derivative of $J_h(p_h, \mu_h)$ at any given direction $(q_h, \lambda_h) \in V_h \times \overset{\circ}{V}_h$, and the derivative will be written as $J_h(p_h, \mu_h)'(q_h, \lambda_h)$. Note that the directions for μ_h are searched only in $\overset{\circ}{V}_h$ instead of V_h because of the considered Dirichlet boundary condition. This can be easily adapted for Neumann boundary conditions.

First of all, we note that the Gateaux derivative for the discrete parabolic solution $u_h^n(p_h, \mu_h)$ at any given direction $(q_h, \lambda_h) \in V_h \times \overset{\circ}{V}_h$, denoted as $u_h^n(p_h, \mu_h)'(q_h, \lambda_h)$, solves the following discrete system: $\mathcal{U}_h^0 = \lambda_h$ and $\mathcal{U}_h^n \equiv u_h^n(p_h, \mu_h)'(q_h, \lambda_h) \in \overset{\circ}{V}_h$ for $n = 1, 2, \ldots$, satisfies

$$\int_\Omega \partial_\tau \mathcal{U}_h^n \,\phi_h \,\mathrm{d}x + \int_\Omega \nabla \mathcal{U}_h^n \cdot \nabla \phi_h \,\mathrm{d}x = \int_\Omega p_h \,\mathcal{U}_h^n \,\phi_h \,\mathrm{d}x + \int_\Omega q_h \,u_h^n(p_h, \mu_h)\phi_h \,\mathrm{d}x \quad \forall \phi_h \in \overset{\circ}{V}_h \,. \tag{5.1}$$

To derive an easy-to-implement formula for computing the derivative $J_h(p_h, \mu_h)'(q_h, \lambda_h)$, we use the adjoint equation technique (cf [18, 22]). For this, we define a discrete sequence $\{w_h^n\}_{n=0}^M \subset \overset{\circ}{V}_h$ such that $w_h^M = 0$ and $w_h^n \in \overset{\circ}{V}_h$ for $n < M$ solves the discrete backward

parabolic equation

$$-\int_\Omega \partial_\tau w_h^n\,\phi_h\,\mathrm{d}x+\int_\Omega \nabla w_h^{n-1}\cdot\nabla\phi_h\,\mathrm{d}x-\int_\Omega p_h\,w_h^{n-1}\,\phi_h\,\mathrm{d}x$$
$$=\tau\alpha_n\int_\omega (u_h^n(p_h,\mu_h)-z^n)\phi_h\,\mathrm{d}x+k_n\int_\Omega (u_h^n(p_h,\mu_h)-z_\theta)\phi_h\,\mathrm{d}x\quad\forall\,\phi_h\in\overset{\circ}{V}_h \tag{5.2}$$

where $k_{n_0}=1$ and $k_n=0$ for $n\neq n_0$.

Using (5.1) and the sequence $\{w_h^n\}_{n=0}^M$, we come to the following simple formula for the evaluation of the derivative of $J_h(p_h,\mu_h)$ (cf [29]):

$$J_h(p_h,\mu_h)'(q_h,\lambda_h)$$
$$=\tau\alpha_0\int_\omega(\mu_h-z^0)\lambda_h\,\mathrm{d}x+\tau^{-1}\int_\Omega w_h^0\,\lambda_h\,\mathrm{d}x+\sum_{n=1}^M\int_\Omega q_h\,w_h^{n-1}\,u_h^n(p_h,\mu_h)\,\mathrm{d}x$$
$$+2\beta\int_\Omega\nabla p_h\cdot\nabla q_h\,\mathrm{d}x+2\gamma\int_\Omega\nabla\mu_h\cdot\nabla\lambda_h\,\mathrm{d}x. \tag{5.3}$$

With the formula (5.3) we are now ready to present the following gradient method for solving the discrete minimization problem (4.5)–(4.7).

Gradient method I. Let an initial guess $(p_h^{(0)},\mu_h^{(0)})\in K_{1h}\times K_{2h}$ be given.

(a) Compute $u_h^0=\mu_h^{(0)}$ and $u_h^n\equiv u_h^n(p_h^{(0)},\mu_h^{(0)})=Q_h^b\bar\eta^n+\hat u_h^n$ from (4.6), (4.7).
Compute $w_h^M=0$ and $w_h^{n-1}\in\overset{\circ}{V}_h$ for $n=M,M-1,\ldots,1$ from (5.2).

(b) Compute the components of $(J_h(p_h^{(0)},\mu_h^{(0)}))'$ corresponding to all the basis functions $\{\phi_m\}$ from V_h and $\{\phi_l\}$ from $\overset{\circ}{V}_h$ using (5.3) respectively:

$$\bar g_m=\sum_{n=1}^M\int_\Omega u_h^n(p_h^{(0)},\mu_h^{(0)})\,w_h^{n-1}\,\phi_m\,\mathrm{d}x+2\beta\int_\Omega\nabla p_h^{(0)}\cdot\nabla\phi_m\,\mathrm{d}x;$$
$$g_l=\tau\alpha_0\int_\omega(\mu_h^{(0)}-z^0)\phi_l\,\mathrm{d}x+\tau^{-1}\int_\omega w_h^0\,\phi_l\,\mathrm{d}x+2\gamma\int_\Omega\nabla\mu_h^{(0)}\cdot\nabla\phi_l\,\mathrm{d}x.$$

Set $\bar g_h=\sum_m\bar g_m\phi_m$ and $g_h=\sum_l g_l\phi_l$.

(c) Find $\lambda^*>0$ such that $J_h(p_h^{(0)}-\lambda\,\bar g_h,\mu_h^{(0)}-\lambda\,g_h)$ is minimized over all $\lambda>0$.

(d) Project the outputs $\tilde p_h=p_h^{(0)}-\lambda^*\,\bar g_h$ and $\tilde\mu_h=\mu_h^{(0)}-\lambda^*\,g_h$ onto the constraint sets K_{1h} and K_{2h} respectively:

$$p_h^*=\min\{\max\{-\alpha_1,\tilde p_h\},\alpha_1\},\qquad \mu_h^*=\min\{\max\{0,\tilde\mu_h\},\alpha_2\}.$$

If $\|(p_h^*,\mu_h^*)-(p_h^{(0)},\mu_h^{(0)})\|\leqslant$ tolerance, stop;
otherwise set $(p_h^{(0)},\mu_h^{(0)})=(p_h^*,\mu_h^*)$, $j=j+1$, goto (a).

For later convenience, we denote the output of the above gradient method I as

$$(p_h^*,\mu_h^*)=\text{Gradient I}(J_h,p_h^{(0)},\mu_h^{(0)},m_h)$$

i.e., (p_h^*,μ_h^*) is the approximate solution obtained using m_h iterations of the gradient method I with the cost functional J_h and initial guess $(p_h^{(0)},\mu_h^{(0)})$.

Remark 5.1. There are many existing approaches for finding the step size λ in step (c). We usually use the parabolic curve search method and it takes about five searches on the average in our numerical experiments.

Through numerous numerical experiments, we find that the above gradient method converges in most cases globally (even with very bad initial guesses) and stably. In particular, we have observed that the first few iterations of the method often converge very fast and then the convergence slows down significantly. If one still continues with the gradient method after the first few iterations, then it will take a great many more iterations to reach the tolerance; a few thousand iterations are often needed even for one-dimensional inverse parabolic problems. This is very much like the performance of the classical iterative methods for solving second-order boundary value problems. For the latter, there exists a well developed MGM which can deal with such a slow-down very efficiently by making full use of the fast convergence of the first few iterations of the classical iterative methods. The MGM starts with a fine grid and iterates a few times using a classical iterative method (called a smoothing step) and then goes to a coarser grid to solve the residual equation to achieve some coarse correction for the approximate solution obtained on the fine grid, again applying the same iterative method a few iterations for the residual equation. The MGMs have been proved to be very effective for solving various direct problems for partial differential equations: see [14, 27, 30] and references therein. However, to our knowledge, there seem still to be no applications of the MGM for solving the highly nonlinear optimization systems arising from ill-posed inverse problems. For such applications, one first has to find a feasible way of formulating the MGM for the highly nonlinear minimization problem with constraints involving some initial-boundary value problems. Clearly this is not straightforward.

In the rest of this section we propose a nonlinear MGM for solving the nonlinear minimization system (4.5)–(4.7), and the numerical results will demonstrate its effectiveness in solving the inverse parabolic problem considered in this paper.

Assume that we are given a nested set of shape regular triangulations $\{\mathcal{T}^{h_k}\}_{k=0}^N$, with $\mathcal{T}^{h_{k+1}}$ being a refinement of $\mathcal{T}^{h_k}$. $\{V_{h_k}\}_{k=0}^N$ are the continuous piecewise linear finite element spaces defined on $\{\mathcal{T}^{h_k}\}_{k=0}^N$ such that

$$V_{h_0} \subset V_{h_1} \subset \cdots \subset V_{h_N} \equiv V_h.$$

Our goal is to solve the discrete minimization problem (4.5)–(4.7), which is defined on the finest space V^h, by making use of the auxiliary coarser spaces V_{h_k} for $0 \leqslant k < N$.

To do so, we need to introduce some more notation. Corresponding to each coarse triangulation $\mathcal{T}^{h_k}$, we divide the time interval $[0, T]$ into M_k subintervals using the points

$$0 = t_k^0 < t_k^1 < \cdots < t_k^{M_k} = T$$

with $t_k^n = n\tau_k$ and $\tau_k = T/M_k$. Similarly to K_{1h} and K_{2h}, we define two constrained subsets K_{1h_k} and K_{2h_k}. Furthermore, for the initialization step of the nonlinear MGM to be introduced below, we have to solve a coarse minimization problem on each coarse space V_{h_k}:

$$\begin{aligned}\text{minimize } J_k^0(p_{h_k}, \mu_{h_k}) = {} & \frac{\tau_k}{2}\sum_{n=0}^{M_k}\alpha_n \int_\omega (u_{h_k}^n - z^n)^2 \,\mathrm{d}x + \frac{1}{2}\int_\Omega (u_{h_k}^{n_k} - z_\theta)^2 \,\mathrm{d}x \\ & + \beta \int_\Omega |\nabla p_{h_k}|^2 \,\mathrm{d}x + \gamma \int_\Omega |\nabla \mu_{h_k}|^2 \,\mathrm{d}x \end{aligned} \tag{5.4}$$

over all $p_{h_k} \in K_{1h_k}$ and $\mu_{h_k} \in K_{2h_k}$, where $u_{h_k}^n \equiv u_h^n(p_{h_k}, \mu_{h_k}) \in V_{h_k}$ solves (4.6), (4.7) on $V_{h_k}^0$. Here $\alpha_0 = \alpha_{M_k} = \frac{1}{2}$ and $\alpha_n = 1$ for all $n \neq 0, M_k$ while $n_k > 0$ is an integer such that $n_k\tau_k = \theta$.

Another important step of the MGM is the coarse grid correction, for which we define the following nonlinear functional:

$$J_k(p_h + p_{h_k}, \mu_h + \mu_{h_k}) = \frac{\tau_k}{2} \sum_{n=0}^{M_k} \alpha_n \int_\omega (u_{h_k}^n - z^n)^2 \, \mathrm{d}x + \frac{1}{2} \int_\Omega (u_{h_k}^{n_k} - z_\theta)^2 \, \mathrm{d}x$$
$$+\beta \int_\Omega |\nabla(p_h + p_{h_k})|^2 \, \mathrm{d}x + \gamma \int_\Omega |\nabla(\mu_h + \mu_{h_k})|^2 \, \mathrm{d}x$$

where $p_{h_k} \in V_{h_k}$ and $\mu_{h_k} \in V_{h_k}^0$, while $u_{h_k}^0 = \mu_h + \mu_{h_k}$ and for $n = 1, 2, \ldots, M_k$, $u_{h_k}^n \equiv u_{h_k}^n(p_h + p_{h_k}, \mu_h + \mu_{h_k}) = Q_{h_k}^b \bar{\eta}^n + \hat{u}_{h_k}^n \in V_{h_k}$ solves

$$(\partial_\tau u_{h_k}^n, \phi_{h_k}) + (\nabla u_{h_k}^n, \nabla \phi_{h_k}) = (p_h + p_{h_k}, u_{h_k}^n \, \phi_{h_k}) \qquad \forall \, \phi_{h_k} \in V_{h_k}^0. \tag{5.5}$$

Finally, corresponding to the adjoint parabolic solution w_h^n given by (5.2), we define an adjoint solution on each coarse space V_{h_k} such that $w_{h_k}^{M_k} = 0$ and $w_{h_k}^{n-1}$ for $n = M, \ldots, 1$ solves

$$-\int_\Omega \partial_\tau w_{h_k}^n \, \phi_{h_k} \, \mathrm{d}x + \int_\Omega \nabla w_{h_k}^{n-1} \cdot \nabla \phi_{h_k} \, \mathrm{d}x - \int_\Omega (p_h + p_{h_k}) \, w_{h_k}^{n-1} \, \phi_{h_k} \, \mathrm{d}x$$
$$= \tau_k \alpha_n \int_\omega (u_{h_k}^n - z^n) \phi_{h_k} \, \mathrm{d}x + k_n \int_\Omega (u_{h_k}^n - z_\theta) \, \phi_{h_k} \, \mathrm{d}x \qquad \forall \, \phi_{h_k} \in V_{h_k}^0 \tag{5.6}$$

where $u_{h_k}^n$ is the solution of (5.5).

Before we state the nonlinear MGM, we first formulate the gradient method for solving the following minimization problem:

$$\min \; J_k(p_h + p_{h_k}, \mu_h + \mu_{h_k})$$

over all $p_{h_k} \in V_{h_k}$ and $\mu_{h_k} \in V_{h_k}^0$ such that $p_h + p_{h_k} \in K_{1h}$ and $\mu_h + \mu_{h_k} \in K_{2h}$. This minimization will constitute the coarse grid correction step in the MGM.

Gradient method C. Let an initial guess $p_{h_k}^{(0)} \in K_{1h_k}$ and $\mu_{h_k}^{(0)} \in K_{2h_k}$ be given.

(a) Solve the parabolic and backward parabolic equations (5.5), (5.6) for $\{u_{h_k}^n\}_{n=0}^{M_k}$ and $\{w_{h_k}^n\}_{n=0}^{M_k}$.

(b) Compute the components of $(J_k(p_h + p_{h_k}^{(0)}, \mu_h + \mu_{h_k}^{(0)}))'$ corresponding to all the basis functions $\{\phi_m\}$ from V_{h_k} and $\{\phi_l\}$ from $V_{h_k}^0$ respectively:

$$\bar{g}_{h_k,m} = \sum_{n=1}^{M_k} (u_{h_k}^n \, w_{h_k}^{n-1}, \phi_m) + 2\beta \, (\nabla(p_h + p_{h_k}^{(0)}), \nabla \phi_m);$$

$$g_{h_k,l} = \tau_k \alpha_0 \int_\omega (\mu_h + \mu_{h_k}^{(0)} - z^0) \phi_l \, \mathrm{d}x + \tau_k^{-1} \int_\omega w_{h_k}^0 \, \phi_l \, \mathrm{d}x + 2\gamma \, (\nabla(\mu_h + \mu_{h_k}^{(0)}), \nabla \phi_l).$$

Set $\bar{g}_{h_k} = \sum_m \bar{g}_{h_k,m} \, \phi_m$ and $g_{h_k} = \sum_l g_{h_k,l} \, \phi_l$.

(c) Find $\lambda^* > 0$ such that $J_k(p_h + p_{h_k}^{(0)} - \lambda \, \bar{g}_{h_k}, \; \mu_h + \mu_{h_k}^{(0)} - \lambda \, g_{h_k})$ is minimized over all $\lambda > 0$.

(d) Project $\tilde{p}_h = p_h + p_{h_k}^{(0)} - \lambda^* \bar{g}_{h_k}$ and $\tilde{\mu}_h = \mu_h + \mu_{h_k}^{(0)} - \lambda^* \, g_{h_k}$ onto the constraint sets K_{1h} and K_{2h}:

$$p_h^* = \min\{\max\{-\alpha_1, \tilde{p}_h\}, \alpha_1\}, \qquad \mu_h^* = \min\{\max\{0, \tilde{\mu}_h\}, \alpha_2\}.$$

Set $(p_{h_k}^{(0)}, \mu_{h_k}^{(0)}) := (p_{h_k}^{(0)} - \lambda \, \bar{g}_{h_k}, \; \mu_{h_k}^{(0)} - \lambda \, g_{h_k})$, $j := j + 1$, goto (a).

In what follows, we denote the output of the gradient method C above as

$$(p_h^*, \mu_h^*) = \text{Gradient C } (J_k, p_h, \mu_h, p_{h_k}^{(0)}, \mu_{h_k}^{(0)}, m_k)$$

i.e., (p_h^*, μ_h^*) is the coarse correction of (p_h, μ_h), obtained using m_k iterations of the gradient method C with the cost functional J_k and coarse initial guess $(p_{h_k}^{(0)}, \mu_{h_k}^{(0)})$.

With the above preparation, we are now ready to formulate the nonlinear MGM for solving the finite element minimization problem (4.5)–(4.7).

Nonlinear MGM. Let $(p_{h_0}^{(0)}, \mu_{h_0}^{(0)}) \in K_{1h_0} \times K_{2h_0}$ be a given initial guess on the coarsest finite element space V_{h_0}.

I. Coarse grid initialization.

For $k = 0, 1, \ldots, N-1$, do:
  If $k \neq 0$, calculate $p_{h_k}^{(0)} = \Pi_{k-1}^k p_{h_{k-1}}^*$ and $\mu_{h_k}^{(0)} = \Pi_{k-1}^k \mu_{h_{k-1}}^*$.
  Compute $(p_{h_k}^*, \mu_{h_k}^*) = \text{Gradient I } (J_k^0, p_{h_k}^{(0)}, \mu_{h_k}^{(0)}, m_k)$.
end;
Compute $p_h^{(0)} = \Pi_{N-1}^N p_{h_{N-1}}^*$ and $\mu_h^{(0)} = \Pi_{N-1}^N \mu_{h_{N-1}}^*$.

II. Smoothing and coarse grid correction. Set the iteration number $j = 0$.

(a) Set $\bar{p}_h^{(0)} = p_h^{(0)}$ and $\bar{\mu}_h^{(0)} = \mu_h^{(0)}$.
  For $k = N, N-1, \ldots, 1, 0$, do:
    If $k \neq N$, compute $p_h^{(0)} = p_h^*$ and $\mu_h^{(0)} = \mu_h^*$.
    Compute $(p_h^*, \mu_h^*) = \text{Gradient C } (J_k, p_h^{(0)}, \mu_h^{(0)}, 0, 0, n_k)$.
(b) If $\|(p_h^*, \mu_h^*) - (\bar{p}_h^{(0)}, \bar{\mu}_h^{(0)})\| \leqslant$ tolerance, stop;
  otherwise set $p_h^{(0)} = p_h^*$ and $\mu_h^{(0)} = \mu_h^*$, $j := j + 1$, goto (a).

Remark 5.2. The operators Π_{k-1}^k in step I can be any interpolation from $V_{h_{k-1}}$ onto V_{h_k}. We use the natural finite element injections since $V_{h_{k-1}} \subset V_{h_k}$. For the starting values $p_{h_k}^{(0)}$ and $\mu_{h_k}^{(0)}$ in step II, we take the most natural zero.

Remark 5.3. The major computational costs of the nonlinear MGM are from the iterations on the finest grid, while all the costs of iterations on the coarse grids are very small compared to the ones on the finest grid since the unknowns on the coarse spaces are considerably scaled down compared with those on the fine space. This is one of the essential ingredients of MGMs.

6. Numerical experiments

In this section we show some numerical experiments on the nonlinear MGM proposed in the previous section for the simultaneous reconstruction for the initial value $\mu(x)$ and the coefficient $p(x)$ in the following test problem:

$$\begin{aligned} &\frac{\partial u}{\partial t}(x,t) = \Delta u(x,t) + p(x)u(x,t), \qquad (x,t) \in \Omega \times (0,T) \\ &u(x,0) = \mu(x), \qquad x \in \Omega \\ &u(x,t) = \eta(x), \qquad (x,t) \in \partial\Omega \times (0,T). \end{aligned} \tag{6.1}$$

Here Ω is a one- or two-dimensional domain. The observed data will be taken to be

$$z_\theta(x) = u(x,\theta), \qquad x \in \Omega; \qquad z(x,t) = u(x,t), \qquad (x,t) \in \omega \times (0,T).$$

For the one-dimensional case, we take $\Omega = (0, 1)$, $T = 1$, $\theta = 1/2$ and $\omega = (1/4, 3/4)$. And we use the four-level nested finite element grids with mesh sizes $h_3 = 1/40$, $h_2 = 1/20$, $h_1 = 1/10$, $h_0 = 1/5$, and choose the parameters $m_k = 5$ and $n_k = 5$ in the nonlinear MGM. For the two-dimensional case, we take $\Omega = (0, 1) \times (0, 1)$, $T = 1$, $\theta = 1/2$ and $\omega = (1/4, 3/4) \times (1/4, 3/4)$, but we use the three-level nested finite element grids with mesh sizes $h_2 = 1/40$, $h_1 = 1/20$, $h_0 = 1/10$, and choose the parameters $m_k = 3$ and $n_k = 3$ in the nonlinear MGM. The regularization parameters are listed in our figure captions, and $\mathcal{E}$ denotes the relative L^2-norm error between the exact parameter $p(x)$ (or $\mu(x)$) to be identified and the numerically reconstructed parameter $p_h(x)$ (or $\mu_h(x)$). The upper bounds α_1 and α_2 in the two constraint sets K_1 and K_2 are both taken to be 100.

In our numerical implementations, we always assume the observation data have some observation errors. That is, instead of using the exact data $z_\theta(x)$ and $z(x, t)$, we take the noisy data of the following form:

$$z_\theta^\delta(x) = z_\theta(x) + \delta \sin(3\pi x), \qquad z^\delta(x, t) = z(x, t) + \delta \sin(3\pi x)$$

for the one-dimensional examples, and

$$z_\theta^\delta(x, y) = z_\theta(x, y) + \delta \sin(3\pi(x + y)), \qquad z^\delta(x, y, t) = z(x, y, t) + \delta \sin(3\pi(x + y))$$

for the two-dimensional examples. Here δ is a noise level parameter, and we use the function $\sin(3\pi x)$ instead of the random function simply for the convenience of the reader's numerical verification.

Example 1. We take the exact solution $u(x, t)$ as

$$u(x, t) = \exp(-3t - x^2 + x + 1/2) + \exp(t + x^2 - x - 1/2)$$

and the coefficient $p(x)$ and the initial data $\mu(x)$ as

$$p(x) = -4x^2 + 4x - 2, \qquad \mu(x) = \exp(-x^2 + x + 1/2) + \exp(x^2 - x - 1/2).$$

Figures 1 and 2 show the exact coefficient and initial data $\{p(x), \mu(x)\}$ (the dashed curves) and the numerically reconstructed $\{p_h(x), \mu_h(x)\}$ (the solid curves) when the noise level is $\delta = 1\%$. The initial guess $\{p_h^0(x), \mu_h^0(x)\}$ is taken to be the constant pair $\{-2.5, 2.0\}$ at all six grid points on the coarsest grid with $h_0 = 1/5$. That is, we start with only these six nodal point values and then iterate with the nonlinear MGM. Clearly, this initial guess is not good at all, but the MGM converges very stably and fast (the result shown is obtained from the tenth iteration) and the reconstruction of both p and μ appears to be rather satisfactory with 1% noise present.

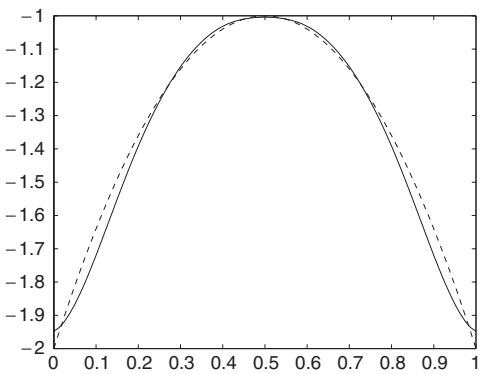

Figure 1. $p_h^{(0)} = -2.5$, $\beta = 5\text{E}{-6}$, $\mathcal{E} = 2.8\text{E}{-2}$.

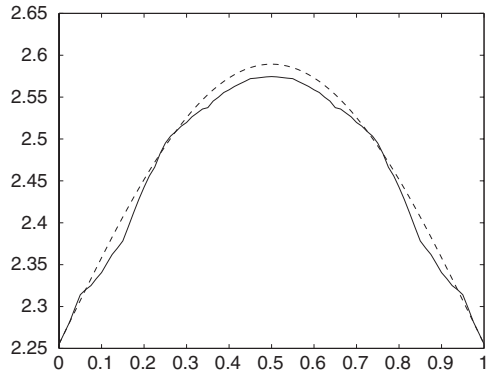

Figure 2. $\mu_h^{(0)} = 2.0$, $\gamma = 5\text{E}{-6}$, $\mathcal{E} = 5.3\text{E}{-3}$.

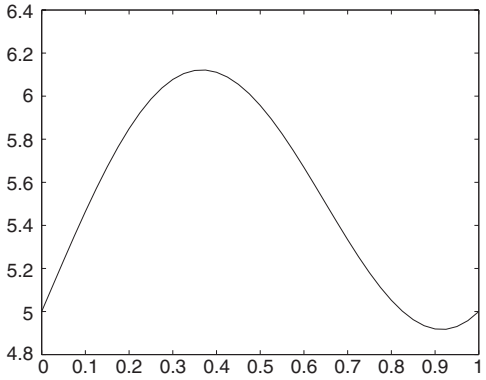

Figure 3. The exact observation $z_\theta = u(x, 1/2)$.

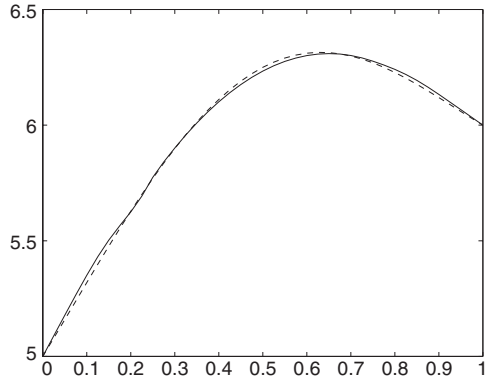

Figure 4. $\mu_h^0 = 5.0$, $\gamma = 1.0\text{E}{-7}$, $\mathcal{E} = 2.3\text{E}{-3}$, $\delta = 1\%$.

Example 2. This example tests the effectiveness of the reconstruction of only the initial temperature $\mu(x)$, with the radiative coefficient $p(x)$ known. To construct more general testing functions p and μ, we add a source function $f(x, t)$ to the right-hand side of equation (6.1). Then we take the exact solution $u(x, t) = \sin \pi x(t + 1) + x^2 + 5$ and the identifying initial data $\mu(x) = x^2 + 5 + \sin \pi x$ with $p(x) = 3 + \cos(2\pi x)$ given. The function $f(x, t)$ is then calculated through (6.1) using $u(x, t)$ and $p(x)$.

Figure 3 is the exact observation $z_\theta = u(x, \theta)$ given at $\theta = 1/2$, very different from the initial temperature $\mu(x)$ to be reconstructed. Figure 4 shows the exact initial data $\mu(x)$ (the dashed curve) and the numerically reconstructed $\mu_h(x)$ (the solid curve) obtained at the tenth iteration of the MGM with noise level $\delta = 1\%$. The initial guess $\mu_h^0(x)$ was taken to be the constant 5.0 on the coarsest grid everywhere. It can be seen that the reconstructed μ_h matches the exact μ very satisfactorily. We reduced the size of the subregion ω from $(1/4, 3/4)$ to $(3/8, 5/8)$ and then to $(7/16, 9/16)$; the reconstructions are nearly as accurate as the one shown in figure 4, although needing a few more iterations: 20 and 25 respectively. If ω is shifted to be in the left half or the right half of $\Omega = (0, 1)$, the reconstructions are still very satisfactory but not as accurate as in the previous cases. Finally, we tried to switch off the observation on the subregion $\omega \times (0, 1)$ with only the measurements given by $z_\theta = u(\cdot, \theta)$. The reconstructions are then always extremely bad even without any noise present. This is consistent with the well known instability of the backward parabolic problem.

Example 3. We take the exact solution $u(x, t) = \exp(t)(2 - \sin 2\pi x)$, the initial data $\mu(x) = u(x, 0)$ and the radiative coefficient $p(x)$ as

$$p(x) = \frac{2 - \sin 2\pi x - 4\pi^2 \sin 2\pi x}{2 - \sin 2\pi x}.$$

We see that the function p has a very large fluctuation between its maximum and minimum (difference $\geqslant 50$). Reconstruction of such a function is very hard from the numerical point of view. Figures 5 and 6 show the exact coefficient $p(x)$ with the initial data $\mu(x)$ (the dashed curves) and the numerically reconstructed $\{p_h(x), \mu_h(x)\}$ (the solid curve) and with noise level $\delta = 1\%$. The initial guess $\{p_h^0(x), \mu_h^0(x)\}$ is taken to be the constant pair $\{0.5, 2.0\}$ everywhere, which is certainly a very bad initial guess for the reconstruction, but the nonlinear MGM converges very stably and the reconstruction obtained using only 20 iterations (figure 5) appears to be satisfactory, considering the large fluctuation of function values of p.

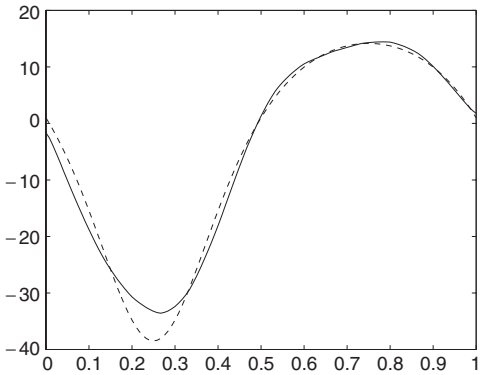

Figure 5. $p_h^{(0)} = 0.5$, $\beta = 5\text{E}{-9}$, $\mathcal{E} = 1.3\text{E}{-1}$.

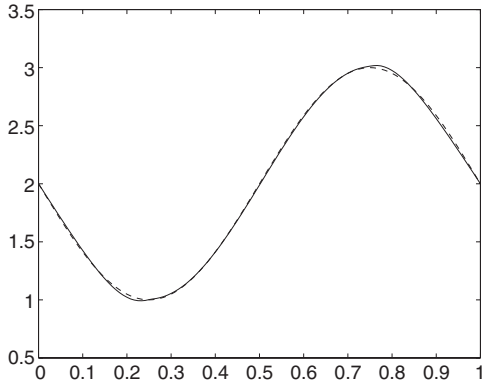

Figure 6. $\mu_h^{(0)} = 2.0$, $\gamma = 1.0\text{E}{-7}$, $\mathcal{E} = 8.6\text{E}{-3}$.

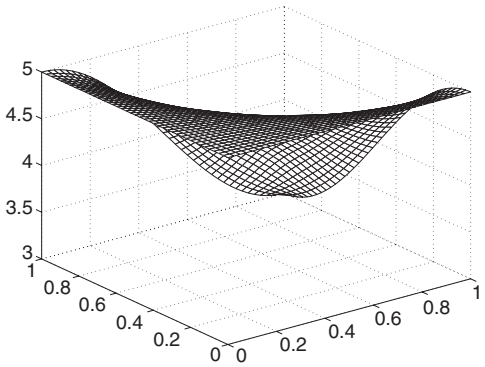

Figure 7. $\mu_h^{(0)} = 5.0$, $\gamma = 1.0\text{E}{-3}$, two iterations, $\mathcal{E} = 1.26\text{E}{-2}$.

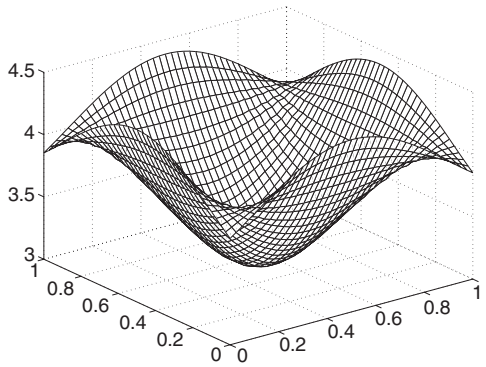

Figure 8. $\mu_h^{(0)} = 5.0$, $\gamma = 5.0\text{E}{-5}$, three iterations, $\mathcal{E} = 1.02\text{E}{-2}$.

Example 4. This example contains two two-dimensional numerical tests: one takes the exact solution $u(x, y, t) = 4 + \sin \pi xy \sin \pi t + \cos \pi xy \cos \pi t + t^2$, and the other takes $u(x, y, t) = 4 - \sin(5\pi x/4 - \pi/8) \sin(5\pi y/4 - \pi/8) \cos 2\pi t + t^2$. The radiative coefficient $p(x, y) = 2+\cos \pi xy$ and the initial data $\mu(x, y) = u(x, y, 0)$ are to be identified. These initial profiles are interesting and challenging. Figures 7 and 8 show the numerically reconstructed initial temperature $\mu_h(x, y)$ by the nonlinear MGM with the initial guesses both being the constant 5.0 at the grid points on the coarsest mesh $h_0 = 1/10$. These are certainly very bad initial guesses given the complicated shapes of the initial profiles. However, in spite of the instability of the inverse problem, surprisingly, the reconstructions are very satisfactory (see figures 7 and 8) with only two and three iterations respectively and the noise being present ($\delta = 1\%$).

Example 5. This example tries to reconstruct both $p(x, y)$ and $\mu(x, y)$. The exact solution $u(x, y, t)$ is chosen to be

$$u(x, y, t) = t^2 + (2 - \sin(\pi x)(1 - x) \sin(\pi y)(1 - y)) \cosh(t),$$

the radiative coefficient is $p(x, y) = 4 - xy(1 - x)(1 - y)$ and the identifying initial data $\mu(x, y) = u(x, y, 0)$. This example is a little challenging to our numerical method. The main difficulty lies in the fact that both functions change their values in a very small range (0.35 for

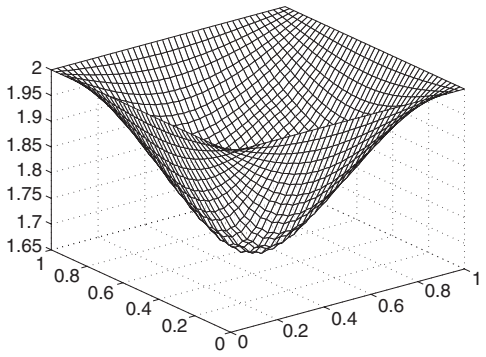

Figure 9. $\mu_h^{(0)} = 2.0$, $\gamma = 1.0\text{E}{-4}$, iter $= 2$, $\mathcal{E} = 2.48\text{E}{-2}$.

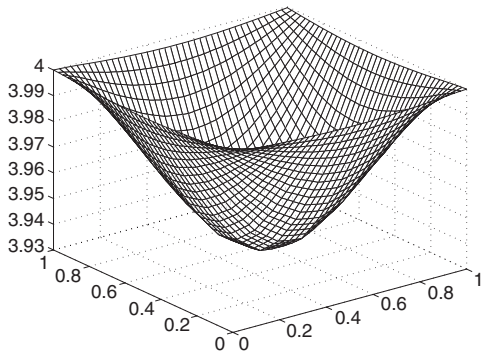

Figure 10. $p_h^{(0)} = 4.0$, $\beta = 1.0\text{E}{-7}$, iter $= 2$, $\mathcal{E} = 1.16\text{E}{-3}$.

$\mu(x, y)$ and 0.07 for $p(x, y)$: the former is five times the latter). Our nonlinear MGM starts with a constant initial guess 2.0 for $\mu(x, y)$ and 4.0 for $p(x, y)$ at the grid points on the coarsest mesh $h_0 = 1/10$ and achieves satisfactory outputs for both μ and p at the second iteration: see figures 9 and 10. We then tried to reduce the size of the observation domain ω to a much smaller subsquare centred at the point $(0.5, 0.5)$ with an area of $1/16$: the reconstruction is almost the same as in figures 9 and 10.

Acknowledgments

The authors thank Mr Kin Wing Lau for his great help in conducting all the numerical results in this paper. The work of JZ was partially supported by Hong Kong RGC grant no CUHK 4004/98P, CUHK Direct grant 98/99 and Visiting Scholar Foundation of Key Lab. in University (China).

References

[1] Banks H T and Kunisch K 1989 *Estimation Techniques for Distributed Parameter Systems* (Cambridge, MA: Birkhäuser)
[2] Bukhgeim A L 2000 *Introduction to the Theory of Inverse Problems* (The Netherlands: VSP)
[3] Cannon J R 1984 *The One-dimensional Heat Equation. Encyclopedia of Mathematics* vol 23 (Reading, MA: Addison-Wesley)
[4] Chan T, Smith B and Zou J 1996 Overlapping Schwarz methods on unstructured meshes using non-matching coarse grids *Numer. Math.* **73** 149–67
[5] Chen Z and Zou J 1999 An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems *SIAM J. Control Optim.* **37** 892–910
[6] Choulli M, Imanuvilov O Yu and Yamamoto M Stability in an inverse parabolic problem with a single measurement *Preprint*
[7] Choulli M and Yamamoto M On inverse parabolic problems *Preprint*
[8] Ciarlet P 1978 *The Finite Element Method for Elliptic Problems* (Amsterdam: North-Holland)
[9] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)
[10] Friedman A 1983 *Partial Differential Equations of Parabolic Type* (Florida: Krieger)
[11] Fursikov A V and Imanuvilov O Yu 1996 *Controllability of Evolution Equations (Lecture Notes Series No 34)* Seoul National University, Korea
[12] Guenther R, Hudspeth R, McDougal W and Gerlach J 1985 Remarks on parameter identification I *Numer. Math.* **47** 355–61
[13] Gutman S 1990 Identification of discontinuous parameters in flow equations *SIAM J. Control Optim.* **28** 1049–60
[14] Hackbusch W 1985 *Multigrid Methods and Applications* (Berlin: Springer)

[15] Imanuvilov O Yu 1995 Controllability of parabolic equations *Sbornik Math.* **186** 879–900
[16] Isakov V 1990 *Inverse Source Problems* (Providence, RI: American Mathematical Society)
[17] 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
[18] Keung Y L and Zou J 1998 Numerical identifications of parameters in parabolic systems *Inverse Problems* **14** 83–100
[19] Klibanov M V 1992 Inverse problems and Carleman estimates *Inverse Problems* **8** 575–96
[20] Kunisch K and White L 1985 The parameter estimation problem for parabolic equations and discontinuous observation operators *SIAM J. Control Optim.* **23** 900–27
[21] Ladyzenskaja O, Solonnikov V and Ural'ceva N N 1968 *Linear and Quasi-Linear Equations of Parabolic Type* (Providence, RI: American Mathematical Society)
[22] Lions J L 1971 *Optimal Control of Systems Governed by Partial Differential Equations* (New York: Springer)
[23] Payne L E 1975 *Improperly Posed Problems in Partial Differential Equations* (Philadelphia: SIAM)
[24] Pazy A 1983 *Semigroups of Linear Operators and Applications to Partial Differential Equations* (Berlin: Springer)
[25] Scott L and Zhang S 1990 Finite element interpolation of nonsmooth function satisfying boundary conditions *Math. Comput.* **54** 483–93
[26] Tataru D 1997 Carleman estimates, unique continuation and controllability for anisotropic PDE's *Contemp. Math.* **209** 267–79
[27] Wesseling P 1992 *An Introduction to Multigrid Methods* (New York: Wiley)
[28] Xu J 1989 Theory of multilevel methods *PhD Thesis* Cornell University
[29] Yamamoto M and Zou J 2000 Simultaneous reconstruction of the initial temperature and heat radiative coefficient *Technical Report CUHK-2000-02 (201)* Department of Mathematics, The Chinese University of Hong Kong webpage (http://www.math.cuhk.edu.hk)
[30] Zou J 1987 A new fast solver–monotone MG method (MMG) *J. Comput. Math.* **5** 325–35