Numerical Identifications of Parameters in Partial Differential Equations

Jingzhi Li\textsuperscript{1,3} and Jun Zou\textsuperscript{2,4}

\textsuperscript{1,2}Department of Mathematics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong.

E-mail: jzli@math.cuhk.edu.hk, zou@math.cuhk.edu.hk

Abstract. In this paper, we will review some recent theoretical and algorithmic developments in parameter identifications in partial differential equations by our research group, focusing on such aspects as variational formulations, convergence analysis, choice strategies of regularization parameters and algorithmic implementations.

Keywords: inverse problem, parameter identification, partial differential equation.

MSC2000: 35R25, 65M30, 65M60, 65N30

1. Introduction
Inverse problems in partial differential equations and integral equations arise from a wide variety of problems in science and engineering. In this paper, we will review some recent theoretical and algorithmic developments in parameter identifications in partial differential equations by our research group. We will mainly focus on such aspects as variational formulations, convergence analysis, choice strategies of regularization parameters and algorithmic implementations. Due to the space limitation, numerical simulations for each numerical method to be discussed in this review will not be shown, and interested readers are referred to all the original articles for results of numerical experiments with these methods.

The rest of the paper is arranged as follows. In the next section, preliminaries are introduced. In section 3, 4, 5 and 7, the identification of heat conductivity, the simultaneous reconstruction of initial temperature and radiative coefficient, the reconstruction of space and time dependent heat fluxes as well as the identification of singular source density are explored respectively, using output least-squares with Tikhonov regularizations and finite element discretizations. In section 6, an efficient solver, nonlinear multigrid gradient method (NMGG), is presented for solving the nonlinear optimization systems resulting from the output least-squares formulation with Tikhonov regularizations. A new convergence rate analysis of Tikhonov regularizations for parameter identifications in heat conduction is done in section 8. Model function choice strategies for regularization parameters are introduced in section 9.

\footnote{The work of this author was partially supported by a Direct Grant of CUHK. (Project CUHK4244/01P)}\footnote{The work of this author was fully supported by Hong Kong RGC Grants (Project 403403)}
2. Preliminaries

In this section, we shall describe some basic notations, concepts and terminologies which will be adopted in the subsequent sections.

2.1. Discretizations

The physical domain \( \Omega \) will be always considered to be a general polygon (2D) or a polyhedron (3D). We shall employ finite element methods for solving the parameter identification problems. Thus we will always assume to triangulate the domain \( \Omega \) with a regular triangulation \( T^h \) of simplicial elements (cf.[5]). Then we define the finite element space \( \tilde{V}_h \) to be the continuous and piecewise linear finite element space over the triangulation \( T^h \), \( \tilde{V}_h \) the subspace of \( V_h \) with all functions vanishing on the boundary \( \partial \Omega \), and \( N_h = \{ x_i \}_{i=1}^N \) the set of all the nodal points of the triangulation \( T^h \). We shall often use \( Q_h \) to denote some interpolation or projection operator onto \( V_h \), e.g., the standard nodal-value interpolation operator or the \( L^2 \)-projection operator.

To fully discretize the time dependent system, we shall also need the time discretization. We divide the time interval \( [0, T] \) into \( M \) equally-spaced subintervals using nodal points

\[
0 = t_0 < t_1 < \cdots < t_M = T
\]

with \( t_n = n \tau, \tau = T/M \). For a continuous mapping \( u : [0, T] \to L^2(\Omega) \), we define \( u^n = u(\cdot, t_n) \) for \( 0 \leq n \leq M \). For a given sequence \( \{ u^n \}_{n=0}^M \subset L^2(\Omega) \), we define the difference quotient \( \partial_t u^n \) and the average \( \bar{u}^n \) of a function \( u(\cdot, t) \) as follows:

\[
\partial_t u^n = \frac{u^n - u^{n-1}}{\tau}, \quad \bar{u}^n = \frac{1}{\tau} \int_{t_{n-1}}^{t_n} u(\cdot, t) \, dt,
\]

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

2.2. Regularizations

For the latter use, we introduce the space of functions with bounded variation

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

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

\[
\int_\Omega |Dq| = \sup \left\{ \int_\Omega q \nabla \cdot g \, dx : g \in (C^0_0(\Omega))^d \text{ and } |g(x)| \leq 1 \text{ in } \Omega \right\}.
\]

Throughout the paper, \( N(q) \) will represent a regularization term which is taken to be

\[
N(q) = \int_\Omega |\nabla q|^2 \, dx \quad \text{or} \quad N(q) = \int_\Omega |Dq|,
\]

namely the square of the \( H^1 \)-seminorm or the \( BV \)-seminorm.

The constrained set \( K \) is a subset of \( H^1(\Omega) \) or \( BV(\Omega) \) defined by

\[
K = \left\{ q \in L^1(\Omega); \| q \| < \infty \text{ and } \alpha_1 \leq q(x) \leq \alpha_2 \text{ a.e. in } \Omega \right\}.
\]

Here the norm \( \| q \| = \| q \|_{H^1(\Omega)} \) or \( \| q \| = \| q \|_{BV(\Omega)} \) corresponds to the forms of \( N(q) \) and \( \alpha_1 \) and \( \alpha_2 \) are two positive constants.

The corresponding discrete terms of \( K \) and \( N(q) \) are defined by

\[
K_h = \{ v_h \in V_h; \quad \alpha_1 \leq v_h(x) \leq \alpha_2, \quad \forall x \in \Omega \},
\]

\[
N_h(q_h) = \int_\Omega |\nabla q_h|^2 \, dx \quad \text{or} \quad N_h(q_h) = \int_\Omega \sqrt{|\nabla q_h|^2 + \delta(h)} \, dx,
\]

depending on the form of \( N(q) \). Here \( \delta(h) \) is any positive function satisfying \( \lim_{h \to 0} \delta(h) = 0 \).
3. Identification of heat conductivity

In this section, we review the numerical identification of the heat conductivity $q(x)$ in the following heat conductive system based on the study in [14]:

$$\frac{\partial u}{\partial t} - \nabla \cdot (q(x) \nabla u) = f(x, t) \quad \text{in} \quad \Omega \times (0, T) \quad (3.1)$$

with the initial condition

$$u(x, 0) = u_0(x) \quad \text{in} \quad \Omega \quad (3.2)$$

and the Dirichlet boundary condition

$$u(x, t) = 0 \quad \text{on} \quad \partial \Omega \times (0, T), \quad (3.3)$$

when some observation data $z(x)$ of $u(x, t)$ at the terminal time $t = T$ are available.

Throughout the review, we shall be mainly interested in the identification process by the output least-squares formulations with appropriate Tikhonov regularizations. The formulations are widely used as one of the most stable and reliable methodology, though they may not always be most efficient in terms of computational time.

We now formulate the parameter identifying problem in the parabolic initial-boundary value problem (3.1)-(3.3) as the following constrained minimizing one (cf.[14]):

$$\min J(q) = \frac{1}{2} \int_{\Omega} q \| \nabla v(q; t) - \nabla z \|^2 dx + \gamma N(q) \quad (3.4)$$

subject to $q \in K$ and $v \equiv v(q; t) \in H^1_0(\Omega)$ satisfying $v(x, 0) = u_0(x)$ in $\Omega$ and

$$\int_{\Omega} v_\ell \phi \, dx + \int_{\Omega} q(x) \nabla v \cdot \nabla \phi \, dx = \int_{\Omega} f(x, t) \phi \, dx, \quad \forall \ \phi \in H^1_0(\Omega) \quad (3.5)$$

for a.e. $t \in (0, T)$. Here $\nabla z$ in (3.4) represents the observation data of $\nabla u(x, t)$ at the terminal time $t = T$.

To formulate the finite element method for solving the continuous minimization problem (3.4)-(3.5), we do the same triangulation procedure and define similar finite element spaces as in Sect. 2.1. Also, we assume the following for the given source term and initial data

$$f \in L^2(\Omega \times (0, T)) \quad \text{and} \quad u_0 \in H^1(\Omega), \quad (3.6)$$

where $\Omega \times (0, T)$. Instead of the system (3.4)-(3.5), the following weaker and more practical formulation is recommended:

$$\min J(q) = \frac{1}{2} \int_{T-\sigma}^{T} \int_{\Omega} q \| \nabla v(q; t) - \nabla z \|^2 dxdt + \gamma N(q) \quad (3.7)$$

subject to $q \in K$ and $v \equiv v(q; t) \in H^1_0(\Omega)$ satisfying $v(x, 0) = u_0(x)$ in $\Omega$ and

$$\int_{\Omega} v_\ell \phi \, dx + \int_{\Omega} q(x) \nabla v \cdot \nabla \phi \, dx = \int_{\Omega} f(x, t) \phi \, dx, \quad \forall \ \phi \in H^1_0(\Omega) \quad (3.8)$$

for a.e. $t \in (0, T)$. Here $\sigma$ is a small constant number. In the numerical implementation, $\sigma$ is suggested to be taken as one or two discrete time-step sizes. One can show (cf.[14]) that for any sequence $\{q_n\}$ in $K$ and some $q$ in $K$ if $\{q_n\}$ converges to $q \in L^1(\Omega)$ as $n$ tends to $\infty$, then

$$\lim_{n \to \infty} \int_{T-\sigma}^{T} \int_{\Omega} q_n(x) \| \nabla (v(q_n) - z) \|^2 dxdt \to \int_{T-\sigma}^{T} \int_{\Omega} |\nabla (v(q) - z) |^2 dxdt. \quad (3.9)$$
Using this, one can show the existence of minimizers to the continuous constraint optimization problem (3.7)-(3.8).

We formulate the finite element approximation problem corresponding to (3.7)-(3.8) as follows:

$$
\min J_h^M(q_h) = \frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_{\Omega} q_h |\nabla v^n_h - \nabla z|^2 dx + \gamma N_h(q_h) \tag{3.10}
$$

subject to \(q_h \in K_h\) and \(v^n_h \equiv v^n_h(q_h) \in \overset{\circ}{V}_h\) satisfying \(v^0_h = Q_h u_0(x)\) in \(\Omega\) and

$$
\int_{\Omega} \partial_T v^n_h \phi_h dx + \int_{\Omega} q_h \nabla v_h \cdot \nabla \phi_h dx = \int_{\Omega} f \phi_h dx, \quad \forall \phi_h \in \overset{\circ}{V}_h \tag{3.11}
$$

where \(K_h\) is the constraint subset of \(K\) as defined in (2.5) and \(N_h(q_h)\) is the discrete regularization term as defined in (2.6).

Let \(P(q_h)\) be the functional defined by

$$
P(q_h) = \frac{1}{2} (q_h(x) - \alpha_2) + \frac{1}{2} (\alpha_1 - q_h(x))^2, \tag{3.12}
$$

then the discretized constrained minimization of \(J_h^M(\cdot)\) over \(K_h\) stated in (3.10)-(3.11) can be further reduced to a sequence of unconstrained minimizations of the following functional \(J_h^M(\varepsilon; \cdot)\) over the entire space \(V_h\) with \(\varepsilon > 0\) as follows:

$$
\min J_h^M(\varepsilon; q_h) = \frac{\tau}{2} \sum_{n=M-n_0}^{M} \int_{\Omega} q_h |\nabla v^n_h - \nabla z|^2 dx + \gamma N_h(q_h) + \frac{1}{\varepsilon} \int_{\Omega} P(q_h(x)) dx \tag{3.13}
$$

subject to \(q_h \in K_h\) and \(v^n_h \equiv v^n_h(q_h) \in \overset{\circ}{V}_h\) satisfying \(v^0_h = Q_h u_0(x)\) in \(\Omega\) and

$$
\int_{\Omega} \partial_T v^n_h \phi_h dx + \int_{\Omega} q_h \nabla v_h \cdot \nabla \phi_h dx = \int_{\Omega} f \phi_h dx, \quad \forall \phi_h \in \overset{\circ}{V}_h. \tag{3.14}
$$

It was proved in [14] that there exist minimizers to the finite element problem (3.10)-(3.11) and that for any sequence \(\{q_h\}_{h>0}\) in \(K_h\) and some \(q\) in \(K\), if \(\{q_h\}_{h>0}\) converges to \(q \in L^1(\Omega)\) as \(h\) tends to 0, then one has

$$
\sum_{n=M-n_0}^{M} \tau \int_{\Omega} q_h(x) |\nabla (v^n_h(q_h) - z)|^2 dx \rightarrow \int_{T-\sigma}^{T} \int_{\Omega} |\nabla (v(q) - z)|^2 dx dt, \tag{3.15}
$$

as \(\tau, h \rightarrow 0\). Also, the convergence of the sequence of minimizers to the discrete minimization problem (3.10)-(3.11) can be shown.

The formulations and convergence results of this section can be equally extended to the cases with observation data available only in a subdomain of \(\Omega\), or with pointwise observation data. Accordingly, when the observation data of \(u\) instead of \(\nabla u\) are available, the energy norm in (3.4) and (3.10) should be replaced by the \(L^2\)-norm, where the coefficient \(q\) or \(q_h\) in the \(L^2\)-norm can be kept or dropped. This remark holds true for most methods presented in this review.
4. Reconstruction of initial temperature and radiative coefficient

Given the measurement of temperature at a fixed time and the measurement of temperature in a subregion of the physical domain, the simultaneous reconstruction of the initial temperature and heat radiative coefficient in a heat conductive system is a rather challenging inverse problem. In [25], a simultaneous reconstruction process was initiated by Tikhonov regularizations with the regularization term being the $L^2$-norm of gradients, and was 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.

Consider the following heat conduction problem:

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

with the initial condition

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

and the Dirichlet boundary condition

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

where the physical domain $\Omega$ can be any polyhedral domain in $R^d(d \geq 1)$ with a piecewise smooth boundary $\partial \Omega$. The reconstruction of the initial temperature distribution $\mu$ and the heat radiative coefficients $p$ simultaneously was studied in [25]. Such reconstructions are extremely important in many practical applications, but they are highly ill-posed in most situations.

It is well known that in the case where $p$ is given, the reconstruction of the initial temperature $\mu$ from $u(x, \theta)$ for $x \in \Omega$ and fixed $\theta > 0$ is highly ill-posed, and this is called the backward parabolic problem. On the other hand, in the case where $\mu$ 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 of such an inverse problem (cf.[10]). A different formulation in [25] which are simultaneously was studied in [25]. Such reconstructions are extremely important in many practical applications, but they are highly ill-posed in most situations.

Now the identifying problem of parameters $p$ and $\mu$ in (4.1)-(4.3) can be formulated as the following constrained minimizing one:

$$\min J(p, \mu) = \frac{1}{2} \int_0^T \int_\omega (u(p, \mu) - z)^2 dx dt$$

$$+ \frac{1}{2} \int_\Omega (u(p, \mu)(x, \theta) - z_\theta(x))^2 dx + \beta \int_\Omega |\nabla p|^2 dx + \gamma \int_\Omega |\nabla \mu|^2 dx$$

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

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

where $\Omega$ is an arbitrarily prescribed subregion. As shown in [25], one can expect a Lipschitz stability for determining the coefficient $p$ but a very weak stability for determining an initial status $\mu$. For the numerical reconstruction, we take some inexact data $z_\theta(x)$ and $z_{|\omega \times (0, T)}$, instead of the exact data $u(\cdot, \theta)|_{\Omega}$ and $u_{|\omega \times (0, T)}$: $z_\theta(x) \approx u(x, \theta), \quad x \in \Omega; \quad z(x, t) \approx u(x, t), \quad (x, t) \in \omega \times (0, T)$. The distributional observation data $z_\theta$ and $z$ are possibly obtained through interpolations of the point observation values in practice.

Now the identifying problem of parameters $p$ and $\mu$ in (4.1)-(4.3) can be formulated as the following constrained minimizing one:
and

\[ \int_{\Omega} u(t) \phi \, dx + \int_{\Omega} \nabla u \cdot \nabla \phi \, dx = \int_{\Omega} p(x) \phi \, dx, \quad \forall \, \phi \in H_0^1(\Omega) \] (4.6)

for a.e. \( t \in (0, T) \), where the positive constants \( \beta \) and \( \gamma \) are regularization parameters. The constraint sets \( K_1 \) and \( K_2 \) above are chosen to be as follows:

\[ K_1 = \{ p \in H^1(\Omega); \ |p(x)| \leq \alpha_1 \ \text{a.e. in} \ \Omega \}, \]

\[ K_2 = \{ \mu \in H^1(\Omega); \ 0 < \mu(x) \leq \alpha_2 \ \text{a.e. in} \ \Omega \}. \]

Here \( \alpha_1 \) and \( \alpha_2 \) are two a priori constants. The existence of minimizers to the optimization problem (4.4)-(4.6) was proved in [25].

To formulate the finite element method for solving the continuous minimization problem (4.4)-(4.6), we do the same triangulation procedure and define similar finite element spaces as in Sect. 2.1. Then we approximate the constraint sets \( K_1 \) and \( K_2 \) by

\[ K_{1h} = \{ p_h \in V_h; \ |p_h(x_i)| \leq \alpha_1 \ \text{for} \ x_i \in \mathcal{N}_h \}, \]

\[ K_{2h} = \{ \mu_h \in V_h; \ 0 < \mu_h(x_i) \leq \alpha_2 \ \text{for} \ x_i \in \mathcal{N}_h \}. \]

By the composite trapezoidal rule in time, we formulate the finite element approximation of the problem (4.4)-(4.6) as follows (cf.[25]):

\[ \min J_h(p_h, \mu_h) = \frac{T}{2} \sum_{n=0}^{M} \alpha_0 \int_{\omega} (u_h^n(p_h, \mu_h) - z^n)^2 \, dx \] (4.7)

\[ + \frac{1}{2} \int_{\Omega} (u_h^n(p_h, \mu_h) - z^n)^2 \, dx + \beta \int_{\Omega} |\nabla p_h|^2 \, dx + \gamma \int_{\Omega} |\nabla \mu_h|^2 \, dx \]

subject to \( p_h \in K_{1h} \), \( \mu_h \in K_{2h} \) and \( u_h^n \equiv u_h^n(p_h, \mu_h) \in V_h \) satisfying

\[ u_h^0(x, 0) = \mu_h, \quad u_h^n = Q_h \tilde{\eta}^n + \bar{u}_h^n, \] (4.8)

\[ \int_{\Omega} \partial_t u_h^n \phi_h \, dx + \int_{\Omega} \nabla u_h^n \cdot \nabla \phi_h \, dx = \int_{\Omega} p_h u_h^n \phi_h \, dx, \quad \forall \, \phi_h \in V_h \] (4.9)

for \( n = 1, 2, \ldots, M \), where \( \alpha_0 = \alpha_M = \frac{1}{2} \) and \( \alpha_n = 1 \) for all \( n \neq 0, M \). Here \( \tilde{\eta} \in H^1(\Omega) \) can be any extension of \( \eta(x, t) \) into \( \Omega \), and \( \bar{u}_h^n \in V_h \).

As in [25], one can demonstrate the existence of minimizers to the finite element problem (4.7)-(4.9), and the convergence of the minimizers \( \{ q_h^n, \mu_h^n \} \) to the finite element system (4.7)-(4.9).

5. Numerical reconstruction of heat fluxes

When only the observation data of temperature in a small subregion near the outer boundary of the physical domain is available, how to reconstruct heat fluxes on an inner boundary of a heat conductive system is quite an interesting task, which arises from the steel industrial background (cf.[22] [23]).

Suppose a heat conductive system occupies an open bounded domain \( \Omega \) with an outer boundary \( \Gamma_o \) and an inner boundary \( \Gamma_i \). Consider the model problem as follows:

\[ \frac{\partial u}{\partial t} = \nabla \cdot (a(x, t) \nabla u) \quad \text{in} \quad \Omega \times (0, T) \] (5.1)
with the initial condition
\[ u(x, 0) = u_0(x) \quad \text{in} \quad \Omega, \] (5.2)
and the heat flux exchanges through the outer and inner boundary \( \Gamma_o \) and \( \Gamma_i \) as follows:
\[ -a(x, t) \frac{\partial u}{\partial n} = c(x, t)(u(x, t) - u_a(x, t)) \quad \text{on} \quad \Gamma_o \times (0, T), \] (5.3)
\[ -a(x, t) \frac{\partial u}{\partial n} = q(x, t) \quad \text{on} \quad \Gamma_i \times (0, T). \] (5.4)
Here \( a(x, t) \) is the heat conductivity, \( c(x, t) \) and \( u_a(x, t) \) are specified functions, and \( q(x, t) \) is the heat flux on the inner boundary \( \Gamma_i \).

Two different regularization formulations for this severely ill-posed inverse problem were proposed and the well-posedness was rigorously justified in [22] [23].

The first approach is to formulate the identification into the following constrained minimizing problem (5.5)-(5.6) in numerical implementations.

Using the composite trapezoidal rule for the time discretization of the first integral in (5.5) and the exact time integration for the second term, the fully discrete finite element approximation to problem (5.5)-(5.6) can be formulated as follows:

\[ \min J(q) = \frac{1}{2} \int_0^T \left( \int_\omega (u(q) - z)^2 \, dx \, dt + \frac{\beta}{2} \int_0^T \int_{\Gamma_i} q^2 \, ds \, dt \right) \] (5.5)

subject to \( q \in L^2(0, T; L^2(\Gamma_i)) \) and \( u(q) \equiv u(q)(\cdot, t) \in H^1(\Omega) \) satisfying \( u(x, 0) = u_0(x) \) in \( \Omega \) and
\[ \int_\Omega \frac{\partial u}{\partial t} \, v \, dx + \int_\Omega a \nabla u \cdot \nabla v \, dx + \int_{\Gamma_o} c \, u \, vds = \int_{\Gamma_o} c \, u_a \, vds - \int_{\Gamma_i} q \, vds \] (5.6)
for all \( v \in H^1(\Omega) \) and for a.e. \( t \in (0, T) \). Here \( z \) is the observation data of \( u \) in \( \omega \times (0, T) \), and \( \omega \) is a small subregion of \( \Omega \) near the outer boundary.

Another formulation for reconstruction of heat fluxes in the heat conductive system (5.1)-(5.4) is to use an \( L^2 \)-regularization in space and \( H^1 \)-regularization in time:

\[ \min J(q) = \frac{1}{2} \int_0^T \left( \int_\omega (u(q) - z)^2 \, dx \, dt + \frac{\beta}{2} \int_0^T \int_{\Gamma_i} |q_t(x, t)|^2 \, ds \, dt \right) \] (5.7)

subject to \( q \in H^1(0, T; L^2(\Gamma_i)) \) and \( u(q) \equiv u(q)(\cdot, t) \in H^1(\Omega) \) satisfying \( u(x, 0) = u_0(x) \) in \( \Omega \) and
\[ \int_\Omega \frac{\partial u}{\partial t} \, v \, dx + \int_\Omega a \nabla u \cdot \nabla v \, dx + \int_{\Gamma_o} c \, u \, vds = \int_{\Gamma_o} c \, u_a \, vds - \int_{\Gamma_i} q \, vds \] (5.8)
for all \( v \in H^1(\Omega) \) and a.e. \( t \in (0, T) \).

It was shown in [22] [23] that for both formulations above, the unique existence of the minimizer and the convergence of the minimization sequence hold true respectively. But the second formulation (5.7)-(5.8) seems to be able to generate more satisfactory results than the first one (5.5)-(5.6) in numerical implementations.

Two fully discrete finite element methods to approximate the resultant nonlinear minimization problems are introduced correspondingly in [22] [23]. Note that the triangulation of the domain \( \Omega \) is not as usual. The triangulation \( T^h \) of \( \Omega \) is formed by the sectorial elements. The arc segments on \( \Gamma_o \) and \( \Gamma_i \) generate naturally two triangulations of \( \Gamma_o \) and \( \Gamma_i \), respectively. To fully discretize the system (5.5)-(5.6), the time interval is divided as in Sect. 2.1.

Using the composite trapezoidal rule for the time discretization of the first integral in (5.5) and the exact time integration for the second term, the fully discrete finite element approximation problem (5.5)-(5.6) can be formulated as follows:

\[ \min J_{h,\tau}(q_{h,\tau}) = \frac{\tau}{2} \sum_{n=0}^{M} a_n \int_\omega (u_h^n - z^n)^2 \, dx + \frac{\beta \tau}{2} \sum_{n=1}^{M} \int_{\Gamma_i} |q_h^n|^2 \, ds \] (5.9)
over all \( q^n_\ell \in V^\ell_h \) with \( u^n_\ell \equiv u^n_\ell (q_{h,\tau}) \in V^h \) satisfying \( u^0_\ell = Q_h u_0(x) \) and

\[
\int_\Omega \partial_n u^n_\ell \phi_h dx + \int_\Omega \bar{a}^n \nabla u^n_\ell \cdot \nabla \phi_h dx + \int_{\Gamma_i} \bar{c}^n u^n_\ell \phi_h ds = 0 \quad (5.10)
\]

for \( n = 1, 2, \ldots, M \). Here \( \{a_n\} \) are the coefficients of the composite trapezoidal rule, i.e., \( a_0 = \frac{1}{2} \) and \( a_n = 1 \) for all \( n \neq 0, M \).

Instead of the trapezoidal rule, by adopting the Crank-Nicolson scheme, the following finite element approximation of (5.7)-(5.8) is proposed in [22] [23]:

\[
\min J_{h,\tau}(q_{h,\tau}) = \frac{\tau}{2} \sum_{n=0}^{M} a_n \int_\omega (u^n - z^n)^2 dx + \frac{\beta}{2} \left( \int_{\Gamma_i} |q^n_{h,\tau}|^2 ds + \sum_{n=1}^{M} \int_{\Gamma_i} |\partial_n q^n_{h,\tau}|^2 ds \right) \quad (5.11)
\]

over all \( q^n_\ell \in V^\ell_h \) with \( u^n_\ell \equiv u^n_\ell (q_{h,\tau}) \in V^h \) satisfying \( u^0_\ell = Q_h u_0 \) in \( \Omega \) and

\[
\int_\Omega \partial_n u^n_\ell \phi_h dx + \int_\Omega \bar{a}^n \nabla u^n_\ell + u^{n-1}_h \bar{a}^n \nabla \phi_h dx + \int_{\Gamma_i} \bar{c}^n u^n_\ell + u^{n-1}_h \bar{c}^n \phi_h ds = 0 \quad (5.12)
\]

for \( n = 1, 2, \ldots, M \). Here \( \{a_n\} \) are the coefficients of the composite trapezoidal rule: \( a_0 = \frac{1}{2} \) and \( a_n = 1 \) for all \( n \neq 0, M \). The heat flux \( q \) is approximated by \( q_{h,\tau} \), a piecewise linear interpolation based on \( \{q^\ell_h\} \) over the time partition in (2.1):

\[
q_{h,\tau}(x,t) = \frac{t - t_{n-1}}{\tau} q^n_h + \frac{t_n - t}{\tau} q^{n-1}_h, \quad t \in (t_{n-1}, t_n) \quad \text{for} \ n = 1, 2, \ldots, M - 1 \quad (5.13)
\]

The existence of the discrete minimizers of the finite element solutions for these two finite element formulations and the convergence of the finite element solutions were rigorously justified in [22] [23]. Moreover, an efficient conjugate gradient method was formulated to solve the nonlinear optimization problems (5.9)-(5.10) and (5.11)-(5.12). We refer interested readers to [22] [23] for more algorithmic details. It is very important and interesting to note that if the Crank-Nicolson scheme is replaced by the backward Euler scheme in (5.11)-(5.12), no convergence can be expected.

6. Nonlinear Multigrid Gradient Method (NMGG)

One of the most important issues in numerical parameter identifications is to solve the nonlinear optimization systems resulting from the output least-squares formulation with Tikhonov regularizations, e.g., the systems (3.10)-(3.11) and (4.7)-(4.9), etc.

Obviously, it is natural for us to resort to the classical gradient method to solve the nonlinear optimization problems. The classical gradient-type methods are among the most stable ones for inverse problems and relatively less sensitive with respect to the noise in observation data, but their convergences are rather slow in general, and the computational cost may become intolerably expensive even for small-scale inverse problems.

Yamamoto and Zou proposed a nonlinear multigrid gradient method (NMGG) in [25] for solving these nonlinear optimization systems. Some initial numerical experiments have demonstrated its advantages over the existing methods: it can greatly speed up the numerical identification process, and converge globally.
As pointed out in [25], the NMGG algorithm was motivated by the following observations. Through numerical experiments, the gradient method converges in most cases stably and globally, even with very bad initial guesses. Especially the first few iterations of the gradient 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 may take a great number of more iterations to reach the desired tolerance. 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 multigrid method (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 some classical iterative method a few iterations. The MGMs have been proved to be very effective for solving various direct problems for partial differential equations. For the latter, of more iterations to reach the desired tolerance. This is very much like the performance of the gradient 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 may take a great number of more iterations to reach the desired tolerance. 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 multigrid method (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 some classical iterative method a few iterations. The MGMs have been proved to be very effective for solving various direct problems for partial differential equations. If one still continues with the gradient method after the first few iterations, then it may take a great number of more iterations to reach the desired tolerance. 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 multigrid method (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 some classical iterative method a few iterations. The MGMs have been proved to be very effective for solving various direct problems for partial differential equations. However there are no applications of the MGM for solving the highly nonlinear optimization systems arising from ill-posed inverse problems. The key issue one has to take care of is to find a feasible way of formulating the MGM for the highly nonlinear minimization problem with constraints involving some initial-boundary value conditions. We now follow [25] to introduce the nonlinear multigrid gradient method, taking the discrete system (4.7)-(4.9) as an example.

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

\[
V^0 \subset V^1 \cdots \subset V^N \equiv V^h
\]

and \( V^h_k \) is the corresponding subspace of \( V^h \) with functions vanishing on the boundary. The goal of NMGG is to solve the discrete minimization problem (4.7)-(4.9), which is defined on the finest space \( V^h \), by making use of the auxiliary coarser spaces \( V^h_k \) for \( 0 \leq k < N \).

Corresponding to each coarse triangulation \( T^h_k \), we divide the time interval \([0,T]\) into \( M_k \) subintervals using the points

\[
0 = \tau^0_k < \tau^1_k < \cdots < \tau^M_k = T
\]

with \( \tau_k = T/M_k \). Corresponding to \( K^h_{1k} \) and \( K^h_{2k} \) on \( T^h \), we define two constrained subsets \( K^h_{1k} \) and \( K^h_{2k} \) on \( T^h_k \). Furthermore, for the initialization step of the NMGG to be introduced below, we have to solve a coarse minimization problem on each coarse space \( V^h_k \):

\[
\min \ J^0_k(p^h_k, \mu^h_k) = \frac{\tau_k}{2} \sum_{n=0}^{M_k} \alpha_n \int_{\omega} (u^n_{h_k} - z^n)^2 dx
\]

\[
+ \int \left( u^n_{h_k} - z^n \right) dx + \beta \int |\nabla p^n_{h_k}|^2 dx + \gamma \int |\nabla \mu^n_{h_k}|^2 dx
\]

over all \( p^h_k \in K^h_{1k} \) and \( \mu^h_k \in K^h_{2k} \), where \( u^n_{h_k} \equiv u^n_{h_k}(p^h_k, \mu^h_k) \in V^h \) solves (4.8)-(4.9) on \( V^h_k \). Here \( \alpha_0 = \alpha_{M_k} = \frac{1}{2} \) and \( \alpha_n = 1 \) for all \( n \neq 0, M_k \) while \( M_k > 0 \) is an integer such that \( n_k \tau_k = \theta \).

**NMGG Algorithm.** Let \((p^0_{h_0}, \mu^0_{h_0}) \in K^h_{10} \times K^h_{20} \) be a given initial guess on the coarsest finite element space \( V^h_0 \).

**Step 1. Coarse grid initialization.**

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

(b) Compute $p_{hk}^0 = \Pi_{N-1}^N p_{hN-1}^*$ and $\mu_{hk}^0 = \Pi_{N-1}^N \mu_{hN-1}^*$.

Step 2. Smoothing and coarse grid correction.
(a) Set the iteration number $j = 0$, $\tilde{p}_h^0 = p_h^0$ and $\tilde{\mu}_h^0 = \mu_h^0$.
For $k = N, N-1, \ldots, 1, 0$, do :
   If $k \neq N$, calculate $\tilde{p}_h^k = p_h^k$ and $\tilde{\mu}_h^k = \mu_h^k$.
   Compute $(\tilde{p}_h^k, \tilde{\mu}_h^k) = \text{Gradient C}(J_k, \tilde{p}_h^k, \tilde{\mu}_h^k, 0, 0, n_k)$.
end;
For $k = 1, \ldots, N$, do :
   Calculate $p_{hk}^0 = \Pi_{k-1}^k p_{hk-1}^*$ and $\mu_{hk}^0 = \Pi_{k-1}^k \mu_{hk-1}^*$.
   Compute $(p_{hk}^*, \mu_{hk}^*) = \text{Gradient I}(J_k, p_{hk}^0, \mu_{hk}^0, m_k)$.
end;
(b) If $\| (\tilde{p}_h^k, \tilde{\mu}_h^k) - (\tilde{p}_h^k, \tilde{\mu}_h^k) \| \leq \text{tolerance}$, stop;
otherwise, set $p_h^j = p_h^k$ and $\mu_h^j = \mu_h^k$, $j := j + 1$, go to (2a).

Readers are referred to [25] for detailed description of Gradient I and Gradient C used in steps 1 and 2.

The NMGG has demonstrated its good performance in different inverse problems.

7. Identification of singular source density
The identification of the boundary shape of some defects of a dielectric material or the interface between different materials is of great use in material sciences. In [12], a special identification problem was considered and transformed into a Fredholm integral equation of the first kind, whose solvability was investigated and allows the identification of some source densities of the distribution type.

We first follow [12] to review the formulation of the corresponding inverse problem. Consider the Maxwell equations:

$$\varepsilon(x)E_t + J(x, t) = \nabla \times H \quad \text{in } \Omega \times (0, T), \tag{7.1}$$

$$\mu(x)H_t = -\nabla \times E \quad \text{in } \Omega \times (0, T), \tag{7.2}$$

where $\Omega \subset \mathbb{R}^3$ is occupied by a dielectric material. $E$ and $H$ represent the electric and magnetic fields, and $J$ is the current density. The coefficients $\varepsilon(x)$ and $\mu(x)$ are the permittivity and permeability of the material. Eliminating $E$ gives

$$\mu(x)H_t + \nabla \times \left( \frac{1}{\varepsilon} \nabla \times H \right) = \nabla \times \left( \frac{1}{\varepsilon} J \right). \tag{7.3}$$

Then the steady-state case becomes

$$\nabla \times \left( \frac{1}{\varepsilon} \nabla \times H \right) = \nabla \times \left( \frac{1}{\varepsilon} J \right). \tag{7.4}$$

If the considered domain consists of two materials with different dielectric coefficients $\varepsilon(x)$, equation (7.4) can be regarded as the following differential equation with a continuous coefficient but a singular source density:

$$\nabla \times (\nabla \times H) = J(x) + g_\varepsilon(x). \tag{7.5}$$
Then the location of the interface can be determined once the singular source density \( g_e \) is available. 

Now consider a thin plate \( V = (0, d_0) \times \Omega \) with thickness \( d_0 \) and a two-dimensional planar domain \( \Omega \). Given the measurement of the eddy current \( F(x) \) at a position \( x \) of a vertical distance \( d \) from the thin plate \( V \), the magnetic field is generated so that it takes the form \( H = (0, H(x_1, x_2))^T \), then the component of \( F(x) \) along the vertical direction is

\[
F_3(x) = \int_V \frac{(x_1 - x'_1)H_{x_1} + (x_2 - x'_2)H_{x_2}}{|x - x'|^3} dx
\]

\[
= \int_V \nabla_{x'} \frac{1}{|x - x'|} \cdot \nabla_x H(x') dx
\]

\[
= -\int_V \nabla_{x'} H(x') \cdot dx' + \int_{\partial V} \frac{\partial H(x')}{\partial n} \frac{1}{|x - x'|} dx'.
\]

This with (7.5) and the fact that \( \nabla \times (\nabla \times H) = (0, 0, \nabla \cdot H_{x'}(x')) \) leads to the following inverse problem: Find the distribution \( \rho(x) \) such that

\[
(K_d \rho)(x) := \int_{\Omega} k_d(x, x')\rho(x')dx' = f(x), \quad x \in \Omega, \tag{7.6}
\]

where \( f(x) \) is available only in a noisy form due to the measurement error, and \( k_d(x, x') \) is given by

\[
k_d(x, x') = \frac{1}{\sqrt{|x - x'|^2 + d^2}}. \tag{7.7}
\]

The following output least-squares method with a regularization of bounded variations was proposed in [12] to solve (7.6):

\[
\min_{\rho \in H^1(\Omega)} J(\rho) = \frac{1}{2} ||K_d \rho - f||_{L_2(\Omega)}^2 + \beta ||\rho||_{BV} + \frac{\mu}{2} ||\nabla \rho||_{L_2(\Omega)}^2; \tag{7.8}
\]

where \( ||\rho||_{BV} \) is defined by:

\[
||\rho||_{BV} = \int_{\Omega} \sqrt{\rho_{x_1}^2 + \rho_{x_2}^2} dx \quad \text{or} \quad ||\rho||_{BV} = \int_{\Omega} \rho_{x_1} dx + \int_{\Omega} \rho_{x_2} dx.
\]

We remark that the parameter \( \beta \) in (7.8) plays the role of regularization in order to properly handle the noise in the data and the nonsmoothness of the function \( \rho(x) \). \( \mu \) is always taken to be smaller than \( \beta \), whose role is to further stabilize the formulation. Without the \( \mu \)-term, the well-posedness of the system (7.6) is not ensured. If the solution \( \rho(x) \) is smooth, one may drop the \( \beta \)-term. But the term is crucial to recover the nonsmooth function \( \rho(x) \).

Let \( \rho^* \) be the solution of (7.8) and \( A_d \)-inner product be defined by

\[
A_d(\sigma, \rho) = (K_d \sigma, K_d \rho) + \mu (\nabla \sigma, \nabla \rho) \quad \forall \ \sigma, \rho \in H^1(\Omega). \tag{7.9}
\]

In [12], Ito and Zou proved the well-posedness of the system (7.8); the minimizer \( \rho^* \) of the problem (7.8) is \( A_d \)-Lipschitz continuous with respect to the regularization parameter \( \beta \) and is continuous with respect to the distance parameter \( d \).

The following iterative algorithm was proposed in [12]: given an initial guess \( \rho^0 \in H^1(\Omega) \), one may generate the sequence \( \{\rho^n\}_{n=1}^{\infty} \subset H^1(\Omega) \) by solving

\[
\frac{\rho^{n+1} - \rho^n}{\Delta t} + (\alpha I + \mu \nabla \nabla)\rho^{n+1} + \beta \partial j(\rho^{n+1}) = K_d f + (\alpha I - K_d K_d) \rho^n, \tag{7.10}
\]
where $\alpha > 0$ is any constant and $\partial j$ the subdifferential of $j(\rho) = |\rho|_{BV}$.

The global convergence of this algorithm (7.10) in $H^1(\Omega)$ was demonstrated in [12]. An obvious advantages of the algorithm is that it reduces the evaluation of the action of the global integral operator $K_d$ to a minimum level.

One of the applications of the identification technique above is to locate the defects of the materials as well as the junction between different materials. Other approaches for similar identifications can be found in [1] [2] [3].

8. New convergence rate analysis of Tikhonov regularizations

In this section, we review a new convergence rate analysis of Tikhonov regularizations that was studied in [8] for parabolic inverse problems (3.1)-(3.3), where the terminal status observation data of $u(x, t)$ is available in a gradient form: $\nabla z(x, t) \approx \nabla u(x, t)$. The pointwise observation data can be treated as well (cf.[8]).

There exists a rich literature on the stability and convergence rate estimates for the Tikhonov regularization methods for nonlinear ill-posed problems, see [7] [16] [20] [21] and the references therein. But the assumptions and conditions in the existing theory are too restrictive and complicated.

Engl and Zou introduced in [8] some new tricks to get rid of restrictive conditions in the existing general convergence theory, e.g., the Fréchet differentiability of the parameter-to-solution map $F(q)$ and the uniform Lipschitz continuity of the Fréchet derivative $F'(q)$. The new techniques enables us to formulate a much simpler source condition and avoid the smallness condition for the source function. Moreover, the new source condition uses the map $F(q)$ itself, instead of its derivative $F'(q)$ and the adjoint $F'(q)^*$, and this makes the condition be interpreted much more easily. The key to this is the use of a modified adjoint which is closely related to the weak form of (3.1).

The identification of the heat conductivity in the system (3.1)-(3.3) are formulated as follows:

$$\min_{q \in K} J(q) = \int_{T-\sigma}^T \int_\omega q(x)|\nabla (u(q) - z)|^2 dx dt + \beta \int_\Omega |\nabla (q - q^*)|^2 dx ,$$

(8.1)

where $u(q)$ solves (3.1)-(3.3) in its weak form, and $\sigma > 0$ is a small number and $K$ defined as in (2.4).

Introduce a linear operator $F(q) : L^2(0,T; L^2(\Omega)^d) \to L^2(\Omega)$ as follows: for any $q \in K$, $F(q)$ is defined by

$$F(q) \phi = - \int_{T-\sigma}^T \nabla u(q) \cdot \phi dt \quad \forall \ \phi \in L^2(0,T; L^2(\Omega)^d).$$

Then under the condition that there exists a function

$$\phi \in H^1_0(T - \sigma, T; L^2(\Omega)) \cap L^2(T - \sigma, T; H^1_0(\Omega))$$

(8.2)

such that the following source condition holds

$$F(q^\dagger) \nabla \phi = \nabla^* \nabla (q^\dagger - q^*),$$

(8.3)

we have

$$\|\nabla q^\dagger - \nabla q^\dagger\| = O(\sqrt{\delta}), \quad \int_{T-\sigma}^T \|\nabla u(q^\dagger) - \nabla u(q^\dagger)\|^2 dt = O(\delta^2)$$

as long as $\beta$ is chosen to be proportional to $\delta$.

The convergence rate in the $L^2$-norm formulation and the general convergence results of the Tikhonov regularization method for heat conduction problem were also established in [8]. Moreover, the new simple source condition can be verified.
This new methodology can be applied to many other inverse problems including elliptic and parabolic problems with nonlinear source terms or nonlinear heat conductivity parameters, and to the case with measurements on a subdomain of Ω, see [6] [15].

9. Choice strategies for regularization parameter

Choosing reasonable regularization parameters is a very crucial step for the output least-squares formulation of inverse problems, which affect the final numerical results significantly. It is an extremely tedious job when the usual trial and error experiments are used. Clearly, it is an essential topic in numerical solution of inverse problems about how to find some general methods and principles for searching certain reasonable regularization parameters for the use in practical implementations. In [17] and [24], based on the Morozov and damped Morozov discrepancy principles, iterative methods like Newton’s method, quasi-Newton’s method, two-parameter model function method and its updated version were proposed for finding some reasonable regularization parameters in an efficient manner. We shall review the main results from [17] [24].

Consider inverse problems of the form

\[ Tf = z, \]  

(9.1)

where \( T \) is a linear bounded operator mapping the parameter space \( X \) into the observation space \( Y \), and \( z \in Y \) is the observation data which may be corrupted by error. The noisy data with noise level \( \delta \) are denoted by \( z^\delta \).

To transform the ill-posed problem into a well-posed problem and make a numerical solution feasible, we still consider the most reliable methodology to formulate the inverse problem as the output least-squares system:

\[
\min_{f \in X} J(f, \beta) = \frac{1}{2} \| Tf - z^\delta \|^2_Y + \frac{\beta}{2} \| f \|^2_X. 
\]  

(9.2)

For any \( \beta > 0 \) there exists a unique solution \( f(\beta) \) to the minimization problem (9.2), which can be characterized as the solution to the variational problem:

\[
(Tf, Tg)_Y + \beta(f, g)_X = (z^\delta, Tg)_Y, \quad \forall \ g \in X. 
\]  

(9.3)

Introduce the minimal cost functional of (9.2):

\[
F(\beta) = \frac{1}{2} \| Tf(\beta) - z^\delta \|^2_Y + \frac{\beta}{2} \| f(\beta) \|^2_X. 
\]  

(9.4)

Then one can show (cf.[17]) that both \( f(\beta) \) and \( F(\beta) \) are infinitely differentiable with respect to \( \beta \), and the derivative \( f^{(n)}(\beta) \in X \), for each \( n \geq 1 \), is the unique solution to the equation:

\[
(Tw, Tg)_Y + \beta(w, g)_X = -n(f^{(n-1)}(\beta), Tg)_X, \quad \forall \ g \in X. 
\]  

(9.5)

Moreover, the first and second derivatives of \( F(\beta) \) are given by the simple relations:

\[
F'(\beta) = \frac{1}{2} \| f(\beta) \|^2_X, \quad F''(\beta) = (f(\beta), f'(\beta))_X. 
\]  

(9.6)

It is important to observe that under the condition \( z^\delta \not\in ker T^* \), the non-negative function \( F(\beta) \) is strictly monotonically increasing and strictly concave.

The well-known Morozov principle has received a considerable amount of attention in linear inverse problems (cf.[4] [9] [18] [19]). It claims that the regularization parameter \( \beta \) should be
chosen such that the error due to the regularization is equal to the error due to the observation data. That is, \( \beta \) is chosen according to

\[
\| Tf(\beta) - z^\delta \|_Y^2 = \delta^2 \equiv \| z - z^\delta \|_Y^2. \tag{9.7}
\]

In terms of \( F(\beta) \), equation (9.7) can be written as

\[
F(\beta) - \beta F'(\beta) = \frac{1}{2} \delta^2. \tag{9.8}
\]

In many applications, the Morozov principle may not perform very satisfactorily. One can therefore consider a more general class of damped Morozov principles (cf.[17] [24]):

\[
\| Tf(\beta) - z^\delta \|_Y^2 + \beta \| f(\beta) \|_X^2 = \delta^2, \tag{9.9}
\]

where \( \gamma \in [1, \infty] \), or equivalently,

\[
F(\beta) + (\beta \gamma - \beta) F'(\beta) = \frac{1}{2} \delta^2. \tag{9.10}
\]

Different methods may be used to solve the damped nonlinear Morozov equation, see [17] [24].

**Newton’s method.** Let \( G(\beta) = F(\beta) + (\beta \gamma - \beta) F'(\beta) - \frac{1}{2} \delta^2 \), then

\[
G'(\beta) = \frac{1}{2} \beta \gamma^{-1} f(\beta), f(\beta) = (\beta \gamma - \beta) f(\beta), f'(\beta) \).
\]

Thus computing \( G'(\beta) \) involves the evaluation of \( f'(\beta) \) that solves the equation (9.5).

**Newton’s algorithm.** Given an initial guess \( \beta_0 \), generate the sequence \( \{\beta_k\}_{k\geq 1} \) by

\[
\beta_{k+1} = \beta_k - \frac{2G(\beta_k)}{\gamma \beta_k^{-1} \| f(\beta_k) \|_X^2 + 2(\beta_k^\gamma - \beta_k)(f(\beta_k), f'(\beta_k) \})}. \tag{9.11}
\]

**Quasi-Newton’s method.** The Newton’s method converges quadratically, but one has to evaluate both \( f(\beta) \) and \( f'(\beta) \) at each iteration. To avoid solving equation (9.5) for \( f'(\beta) \), it was proposed in [17] to replace \( f'(\beta_k) \) in (9.11) by \( f_k(\beta_k, \beta_{k-1}) = \frac{f(\beta_k) - f(\beta_{k-1})}{\beta_k - \beta_{k-1}} \).

**Quasi-Newton’s algorithm.** Given initial guesses \( \beta_0 \) and \( \beta_1 \), generate the quasi-Newton’s sequence \( \{\beta_k\}_{k\geq 1} \) according to

\[
\beta_{k+1} = \beta_k - \frac{2G(\beta_k)}{\gamma \beta_k^{-1} \| f(\beta_k) \|_X^2 + 2(\beta_k^\gamma - \beta_k)(f(\beta_k), f_k(\beta_k, \beta_{k-1}))}. \tag{9.12}
\]

One can show (cf.[17]) that under the condition that \( F(0) < \frac{1}{2} \delta^2 \leq F(1) \), there exists a positive constant \( \varepsilon \) such that whenever the initial guesses \( \beta_0 \) and \( \beta_1 \) belong to the interval \( I = [\beta_* - \varepsilon, \beta_* + \varepsilon] \), the whole sequence \( \{\beta_k\}_{k=0}^\infty \) generated by the quasi-Newton’s method is contained in \( I \) and converges to the unique solution of the Morozov equation (9.10) superlinearly.

**Model function method.** By a model function we mean a parametrized function \( m(\beta) \) which approximates or interpolates \( F(\beta) \) in some way and preserves the properties of the non-negative function \( F(\beta) \). Some results of a four-parameter model function approach for nonlinear inverse problems was investigated in [11]. We are now going to introduce two two-parameter model
function approaches which perform well for linear inverse problems, based on the studies in [17] and [24].

**Model function 1.** When the linear operator \( T \) has dense range in \( L^2(\Omega) \), then \( F(0) = 0 \). And in this sense, \( F(\beta) \) can be modeled by the simple model function

\[
m(\beta) = C(1 - \frac{T}{T + \beta}).
\]

(9.13)

To update the two parameters \( C \) and \( T \) in this model function, the following algorithm was suggested to approximately solve the general Morozov equation (9.10) in [17]:

**Algorithm 1.** Set \( k = 0 \) and choose \( \beta_0 > 0 \) and \( \epsilon > 0 \).

1. Solve \( f(\beta) \) from (9.3) and compute \( F(\beta) \) and \( F'(\beta_k) \) from (9.4) and (9.6). Compute \( T_k \) and \( C_k \) from

\[
m_k(\beta_k) = C_k(1 - \frac{T_k}{T_k + \beta_k}) = F(\beta_k), \quad m'_k(\beta_k) = \frac{C_k T_k}{(T_k + \beta_k)^2} = F'(\beta_k).
\]

(9.14)

2. Set \( m_k(\beta) = C_k(1 - \frac{T_k}{T_k + \beta}) \).

3. Solve for \( \beta_{k+1} \) the approximate Morozov’s equation

\[
m_k(\beta) + (\beta^\gamma - \beta)m'_k(\beta) = \frac{1}{2}\delta^2.
\]

(9.15)

4. If \( |\beta_{k+1} - \beta_k| \leq \epsilon \), stop; otherwise set \( k := k + 1 \), goto step 1.

The following formulae can be easily found from (9.14) for computing \( T_k \) and \( C_k \):

\[
T_k = \frac{\beta_k^2 F'(\beta_k)}{F(\beta_k) - \beta_k F'(\beta_k)}, \quad C_k = \frac{F^2(\beta_k)}{F(\beta_k) - \beta_k F'(\beta_k)}.
\]

(9.16)

Note that the denominators in (9.16) can be written as

\[
F(\beta_k) - \beta_k F'(\beta_k) = \frac{1}{2} \| Ff(\beta_k) - z_\delta \|^2_Y > 0.
\]

(9.17)

In some applications, this may approach zero when \( \beta_k \) becomes close to its convergent limit, so may cause some computational instability. This leads to a new model by Xie and Zou (cf.[24]):

**Model function 2.** Without assuming \( F(0) = 0 \), one can model \( F(\beta) \) by

\[
m(\beta) = \frac{1}{2} \| z_\delta \|^2_Y + \frac{\tilde{C}}{T + \beta}.
\]

(9.18)

This model function still keeps only two parameters \( \tilde{C} \) and \( \tilde{T} \).

Replacing the model function \( m(\beta) \) in the two-parameter algorithm I by the new model (9.18), a new two-parameter algorithm is suggested as follows:

**Algorithm 2.** Set \( k = 0 \) and choose \( \beta_0 > 0 \) and \( \epsilon > 0 \).

1. Solve \( f(\beta_k) \) from (9.3) and compute \( F(\beta_k) \) and \( F'(\beta_k) \) from (9.4) and (9.6). Compute \( T_k \) and \( C_k \) from

\[
m_k(\beta_k) = \frac{1}{2} \| z_\delta \|^2_Y + \frac{C_k}{T_k + \beta_k} = F(\beta_k), \quad m'_k(\beta_k) = -\frac{C_k}{(T_k + \beta_k)^2} = F'(\beta_k).
\]

(9.19)
Step 2. Set $m_k(\beta) = \frac{1}{2}\|z\|^2 + \frac{C_k}{T_k + \beta}$.

Step 3. Solve for $\beta_{k+1}$ the approximate Morozov’s equation

$$m_k(\beta) + (\beta^\gamma - \beta)m_k'(\beta) = \frac{1}{2}\delta^2.$$  \hfill (9.20)

Step 4. If $|\beta_{k+1} - \beta_k| \leq \varepsilon$, stop; otherwise set $k := k + 1$, goto step 1.

This new algorithm has some advantages over the two-parameter algorithm I. From the equalities (9.19) we can easily derive the formulae for updating the two parameters $C_k$ and $T_k$:

$$T_k = \frac{||Tf(\beta_k)||^2}{||T(\beta_k)||^2_X}, \quad C_k = -\frac{(||Tf(\beta_k)||^2 + \beta_k\|T(\beta_k)||^2_X)^2}{2\|T(\beta_k)||^2_X}. \hfill (9.21)$$

The new model function appears more consistent with the original approximation principle. In addition, the new formulae in (9.21) are more computationally stable than those in (9.16). For any $\beta > 0$, the new model function $m(\beta)$ still preserves the monotonicity and concavity of $F(\beta)$ (cf.[24]).

Further in [24], Xie and Zou proposed an improved version of the two-parameter algorithm 2, which not only can preserve the nice properties of the new model function (9.18) but also achieve a global convergence. For this, Let $G_k(\beta) = m_k(\beta) + (\beta^\gamma - \beta)m_k'(\beta)$, then we define a relaxation form:

$$\hat{G}_k(\beta) = G_k(\beta) + \alpha_k(G_k(\beta) - G_k(\beta_k)) \hfill (9.22)$$

where $\alpha_k$ can be chosen by some nice formulae. Note that $\hat{G}_k(\beta)$ still preserves the monotonicity, and this gives an improved algorithm:

**Algorithm 3.** Set $k = 0$ and choose $\beta_0 > 0$ and $\varepsilon > 0$.

Step 1. Solve $f(\beta_k)$ from (9.3) and compute $F(\beta_k)$ and $F'(\beta_k)$ from (9.4) and (9.6). Compute $T_k$ and $C_k$ from

$$m_k(\beta_k) = \frac{1}{2}\|z\|^2 + \frac{C_k}{T_k + \beta_k} = F(\beta_k), \quad m_k'(\beta_k) = -\frac{C_k}{(T_k + \beta_k)^2} = F'(\beta_k). \hfill (9.23)$$

Step 2. Set $m_k(\beta) = \frac{1}{2}\|z\|^2 + \frac{C_k}{T_k + \beta}$.

Step 3. Solve for $\beta_{k+1}$ the approximate Morozov’s equation

$$\hat{G}_k(\beta) = \frac{1}{2}\delta^2. \hfill (9.24)$$

Step 4. If $|\beta_{k+1} - \beta_k| \leq \varepsilon$, stop; otherwise set $k := k + 1$, goto step 1.

Under the condition that $\hat{G}_0(\beta_0) > \frac{1}{2}\delta^2$, the sequence $\{\beta_k\}$ generated by the two-parameter algorithm 3 is well defined. Moreover, the sequence $\{\beta_k\}$ either is finite and terminates at some $\beta_k$ satisfying $G(\beta_k) \leq \frac{1}{2}\delta^2$, or it is infinite and converges to the unique solution $\beta^*$ of the Morozov equation (9.10) strictly monotone decreasingly with any initial value $\beta_0$ lying in $(\beta^*, 1)$ (cf.[24]).

**References**

[1] H. Banks and F. Kojima, *Boundary shape identification problems in two-dimensional domains related to thermal testing of materials*, Quart. Appl. Math. 47 (1989), 273-293.

[2] H. Banks and F. Kojima, *Boundary shape identification in two-dimensional electrostatic problems using squids*, Technical report crsc-tr98-15, Center for Research in Scientific Computing, North Carolina State University, 1998.
[3] H. Banks, F. Kojima and W. Winfree, *Boundary estimation problems arising in thermal tomography*, Inverse Problems 6 (1990), 897-921.
[4] J. Baumeister, *Stable solutions of inverse problems*, Braunschweig, Vieweg, 1986.
[5] P. Ciarlet, *The finite element method for elliptic problems*, Amsterdam, North-Holland, 1978.
[6] H. W. Engl and P. Kugler, *The influence of the equation type on iterative parameter identification problems which are elliptic or hyperbolic in the parameter*, European J. of Appl. Math. 14 (2003), 129–163.
[7] H. W. Engl, K. Kunisch and A. Neubauer, *Convergence rates for Tikhonov regularization of nonlinear ill-posed problems*, Inverse Problems 5 (1989), 523-540.
[8] H. W. Engl and J. Zou, *A new approach to convergence rate analysis of Tikhonov regularization for parameter identification in heat conduction*, Inverse Problems 16 (2000), 1907-1923.
[9] C. W. Groetsch, *The theory of Tikhonov regularization for Fredholm equations of the first kind*, Boston, MA: Pitman, 1983.
[10] V. Isakov, *Inverse Source Problems*, Providence, RI, Amer. Math. Soc., 1990.
[11] K. Ito and K. Kunisch, *On the choice of the regularization parameters in nonlinear inverse problems*, SIAM J. Optim. 2 (1992), 376-404.
[12] K. Ito and J. Zou, *Identification of some source densities of the distribution type*, J. Comput. Appl. Math. 132 (2001), 295–308.
[13] Y. L. Keung and J. Zou, *An efficient linear solver for nonlinear parameter identification problems*, SIAM J. Sci. Comput. 22 (2000), 1511–1526.
[14] Y. L. Keung and J. Zou, *Numerical identifications of parameters in parabolic systems*, Inverse Problems 14 (1998), 83-100.
[15] P. Kugler, *Identification of a temperature dependent heat conductivity from single boundary measurements*, SIAM J. Numer. Anal. 41 (2003), 1543–1563.
[16] K. Kunisch and W. Ring, *Regularization of nonlinear ill-posed problems with closed operators*, Numer. Funct. Anal. Optim. 14 (1993), 389-404.
[17] K. Kunisch and J. Zou, *Iterative choices of regularization parameters in linear inverse problems*, Inverse Problems 14 (1998), 1247-1264.
[18] A. K. Louis, *Inverse und schlechtgestellte problems*, Stuttgart, Teubner, 1989.
[19] V. A. Morozov, *Methods for solving incorrectly posed problems*, N.Y., Springer, 1984.
[20] A. Neubauer, *Tikhonov regularization for nonlinear ill-posed problems: optimal convergence and finite dimensional approximation*, Inverse Problems 5 (1989), 541-557.
[21] T. I. Seidman and C. R. Vogel, *Well-posedness and convergence of some regularization methods for nonlinear ill-posed problems*, Inverse Problems 5 (1989), 227-238.
[22] J. Xie, *Numerical reconstruction of heat fluxes*, SIAM J. Numer. Anal., Accepted for publication, 2005.
[23] J. Xie and J. Zou, *Numerical reconstruction of heat fluxes*, Technical Report, Department of Mathematics, The Chinese University of Hong Kong, Hong Kong, 2003.
[24] J. Xie and J. Zou, *An improved model function method for choosing regularization parameters in linear inverse problems*, Inverse Problems 18 (2002), 631–643.
[25] M. Yamamoto and J. Zou, *Simultaneous reconstruction of the initial temperature and heat radiative coefficient*, Inverse Problems 17 (2001), 1181–1202.