Journal of Computational Mathematics, Vol.16, No.4, 1998, 305–318.

A QUASI-NEWTON METHOD IN INFINITE-DIMENSIONAL SPACES AND ITS APPLICATION FOR SOLVING A PARABOLIC INVERSE PROBLEM*[1)]

Wen-huan Yu
(*Department of Mathematics, Tianjin University, Tianjin 300072, P.R. China.*)

Abstract

A Quasi-Newton method in Infinite-dimensional Spaces (QNIS) for solving operator equations is presented and the convergence of a sequence generated by QNIS is also proved in the paper. Next, we suggest a finite-dimensional implementation of QNIS and prove that the sequence defined by the finite-dimensional algorithm converges to the root of the original operator equation providing that the later exists and that the Fréchet derivative of the governing operator is invertible. Finally, we apply QNIS to an inverse problem for a parabolic differential equation to illustrate the efficiency of the finite-dimensional algorithm.

Key words: Quasi-Newton method, parabolic differential equation, inverse problems in partial differential equations, linear and Q-superlinear rates of convergence

1. Introduction

Quasi-Newton methods play an important role in numerically solving non-linear systems of equations on the Euclidean spaces. But it seems that the quasi-Newton methods have not been applied directly to solving inverse problems in partial differential equations (PDE) up to now if we exclude those methods, by which inverse problems in PDEs are formulated as optimization problems with equality constraints.

We, first, suggest a Quasi-Newton method in Infinite-dimensional Spaces (QNIS) in §2, which can be used to solve an operator equation that is governed by a non-linear operator mapping sets in a Hilbert space into another Hilbert space.

Next, we prove in §3 that the sequence $\{q_n\}$ generated by the QNIS procedure converges to the root of the operator equation if the later exists and the Fréchet derivative of the governing operator is invertible. In §4 we, first, give a proof to show that a finite-dimensional, approximate equation has a root if the original equation does, and then prove that the roots of finite-dimensional approximate equations converge to the root of the original operator equation under proper conditions. Finally, apply the above-mentioned algorithm to an inverse problem for parabolic differential equation, which shows that QNIS is efficient.

There are a lot of papers dealing with computation of inverse problems. We only list a few of them according the methods used as follows:

* Received Feburary 17, 1995.
1) The work has partially been supported by the National Natural Science Foundation of China.

1. The gradient or conjugate gradient methods[6,20,19,21], which need computing the derivative maps of the operators described by partial differential equations;
2. The generalized pulse spectrum technique (GPST)[7−8,25];
3. The finite-dimensional approximate modal methods[1−3];
4. The regularization methods[22−24];
5. The sequential quadratic programming (SQP) methods[16].

Finally, it should be pointed out that a superlinear rate of convergence in infinite-dimensional spaces is not trivial as it does in finite-dimensional spaces. [10] showed out that Q-superlinear convergence for a Lipchitzian operator $F(q)$ can be achieved if an initial operator A_0 is close to $F'(q^*)$ up to an arbitrary compact perturbation.

By the way, QNIS presented in the paper can also be applied to inverse problems in other PDEs.

2. A Quasi-Newton Method in Infinite-dimensional Spaces

We consider an operator equation

$$\Phi(q, u) = 0, \tag{1}$$

where $\Phi \in C(Q \times U, \mathcal{F})$, $u \in U$ is a state of the system, $q \in Q$ is a parameter, U is a state space. Q is a topological space, U and $\mathcal{F}$ are Banach spaces, $C(Q \times U, \mathcal{F})$ is the set of all continuous maps on $Q \times U$ to $\mathcal{F}$.

We assume that (1) is well-posed, that is, $\forall q \in Q$ there is a unique $u \in U$ satisfying (1), and u depends continuously on q, then denote $u = u(q)$.

The inverse problem we address is to determine the pair (q, u) satisfying (1) and

$$\mathcal{M}u = z, \tag{2}$$

where $z \in Y$ is given, $\mathcal{M} : U \to Y$ is a given measurement operator.

The operator equations studied in the paper consist of partial differential equations and additional initial and/or boundary-value conditions.

For example, (1) is described by the following initial-boundary value problem for a parabolic equation:

$$\begin{aligned} &u_t = (q(x,y)u_x)_x + (q(x,y)u_y)_y + f(x,y,t), \quad (x,y) \in \Omega, \quad t \in (0,T) \\ &\partial_\nu u \mid_{\partial\Omega} = 0, \quad u(x,0) = u_0(x), \end{aligned} \tag{3}$$

which governs the temperature distribution in a nonhomogeneous isotropic solid or the pressure distribution in a fluid-containing porous medium. It is well-known that $\forall\, q \in L^\infty(\Omega)$ with $q(x) \geq c_0 > 0$, a.e. Ω, the problem (3) is well-posed and $U = H^1(\Omega \times (0,T))$.

The inverse problem considered is to determine $(q, u) \in Q \times U$ that satisfy (3) and

$$u \mid_{t=T} = z, \tag{4}$$

where $z \in H^1(\Omega)$ is given.

Next, we set

$$F : Q \to Y, \quad F(q) = \mathcal{M}u(q) - z. \tag{5}$$

Therefore, solving the inverse problem (1) + (2) will be reduced to solving the following operator equation: search $q \in Q$ such that

$$F(q) = 0. \tag{6}$$

We assume that q^* is a root of (6). By the Taylor theorem for vector-valued functions one has

$$F(q) = F(q) - F(q^*) = \int_0^1 F'(q^* + t(q - q^*))(q - q^*)\, dt \approx F'(q)(q - q^*) \tag{7}$$

if q approximates to q^*, and then gets the Newton iteration method:

$$s_k = -[F'(q_k)]^{-1} F(q_k), \quad q_{k+1} = q_k + s_k. \tag{8}$$

But, this method needs calculating the derivative operator and its inverse. On the contrary, the quasi-Newton methods do not need computing any derivative operator. We recall that the Broyden method, which belongs to [4] and which is one of the most successful quasi-Newton methods, for solving non-linear systems of equations:

$$P : \mathbf{R}^n \to \mathbf{R}^n, \quad P(x) = 0, \tag{9}$$

reads as follows:

$$\begin{aligned} &A_k s_k = -P(x_k), \quad x_{k+1} = x_k + s_k, \quad y_k = P(x_{k+1}) - P(x_k), \\ &A_{k+1} = A_k + (y_k - A_k s_k) s_k^T / (s_k^T s_k), \end{aligned} \tag{10}$$

where A_k is an $n \times n$ matrix. By the Sherman-Morrison-Woodbury formula it follows that

$$B_{k+1} = B_k + (s_k - B_k y_k) s_k^T B_k / (s_k^T B_k y_k), \tag{11}$$

where $B_k = A_k^{-1}$.

Obviously, in order to extend the above method one should overcome the following difficulties:

- the operator F is a map from a space to another space;
- Q and Y both are infinite-dimensional.

From now on, we assume that Q and Y are Hilbert spaces. To begin with, change the algorithm (10) as follows:

$$\begin{cases} A_k s_k = -F(q_k), \quad q_{k+1} = q_k + s_k, \quad y_k = F(q_{k+1}) - F(q_k), \\ A_{k+1} = A_k + (y_k - A_k s_k)(s_k, \cdot)/(s_k, s_k), \end{cases} \tag{12}$$

where $(\cdot, \cdot)$ is the inner product in Q and the operator $(s_k, \cdot) : Q \to \mathbf{R}$, i.e. $\forall q \in Q$, $(s_k, \cdot) q \equiv (s_k, q)$.

Obviously, $A_k \in \mathcal{L}(Q,Y)$ if $A_0 \in \mathcal{L}(Q,Y)$, where $\mathcal{L}(Q,Y)$ is the space of all linear continuous operators on Q to Y. It will be proved in the next section that A_k is invertible provided A_0 does. Therefore, by the generalized Sherman-Morrison-Woodbury formula it follows that

$$B_{k+1} = B_k + (s_k - B_k y_k)(s_k, B_k\cdot)/(s_k, B_k y_k), \tag{13}$$

where $B_k = A_k^{-1}$.

Thus, the formulas (12) form the Quasi-Newton algorithm in Infinite-dimensional Spaces (QNIS), which is a generalization of the Broyden-like method.

3. Convergence of QNIS

We, first, state the following lemma from [12]:

Lemma 3.1. *Suppose that $S \in \mathcal{L}(X,Y)$, where X and Y are Banach spaces, and that S has an inverse $S^{-1} \in \mathcal{L}(Y,X)$. Then $\forall T \in \mathcal{L}(X,Y)$ satisfying $\|T\| < 1/\|S^{-1}\|$ the operator $\tilde{T} = S + T$ is invertible and $\|\tilde{T}^{-1}\| \le \|S^{-1}\|/(1 - \|S^{-1}T\|) \le \|S^{-1}\|/(1 - \|S^{-1}\|\|T\|)$.*

If we are given $q_0 \in Q$ and $A_0 \in \mathcal{L}(Q,Y)$, then one gets $\{q_n\}$ and $\{A_n\}$ from the algorithm (12), their convergence can be obtained from the following lemma, which is an extension of Broyden et al.'s results [5].

Lemma 3.2. *We assume that $F : D \subset Q \to Y$ is continuously Fréchet differentiable in $D_o \subset D$, where D_o is a convex, open set, that $q^* \in D$ is a zero point of F, and that $\Lambda \equiv F'(q^*) \in \mathcal{L}(Q,Y) \setminus \{0\}$ is invertible, $\|\Lambda^{-1}\| \le \beta$, and satisfies the following inequality*

$$\|F'(q) - \Lambda\| \le L\|q - q^*\|, \quad \forall\, q \in D_o. \tag{14}$$

Furthermore, we assume that the operator sequence $\{A_n\}$ defined by (12) satisfies

$$\|A_{n+1} - \Lambda\| \le [1 + \alpha_1\sigma(q_n, q_{n+1})]\|A_n - \Lambda\| + \alpha_2\sigma(q_n, q_{n+1}), \tag{15}$$

where α_1 and α_2 are constants and $\sigma(q_n, q_{n+1}) = \max\{\|q_n - q^*\|, \|q_{n+1} - q^*\|\} \equiv \sigma_n$.

Then the sequence $\{q_n\}$ defined by (12) is well-defined, converges to q^*, and satisfies

$$\|q_{n+1} - q^*\| \le \gamma\|q_n - q^*\|, \quad n = 0, 1, \ldots, \tag{16}$$

where $\gamma \in (0,1)$ is arbitrarily given providing that A_0 and q_0 satisfy

$$\|q_0 - q^*\| \le \eta, \quad \|A_0 - \Lambda\| \le \delta, \tag{17}$$

where $\eta = \eta(\gamma)$ and $\delta = \delta(\gamma)$ are constant dependent on γ and satisfy the following inequalities:

$$6\beta(1+\gamma)\delta < \gamma, \tag{18}$$

$$(2\alpha_1\delta + \alpha_2)\eta/(1-\gamma) \le \delta, \quad \eta < \eta_0 \tag{19}$$

$$\beta(1+\gamma)^2(L\eta/2 + 3\delta) \le \gamma, \tag{20}$$

Furthermore, A_n^{-1} exists and the sequences $\{\|A_n\|\}$ and $\{\|A_n^{-1}\|\}$ are uniformly bounded.

Proof. Because D_o is open and $q^* \in D_o$, there exists $\eta_0 > 0$ such that the ball $B(\eta_0, q^*) \equiv \{q \in Q;\ \|q - q^*\| < \eta_0\} \subset D_o$. By the assumptions for any $\gamma \in (0,1)$ one can definitely choose $\delta = \delta(\gamma) > 0$ and $\eta = \eta(\gamma) > 0$ such that the inequalities (17)–(20) are valid.

Because $A_0 = \Lambda + (A_0 - \Lambda)$ and $\|A_0 - \Lambda\| \le \delta < 1/\beta \le \|\Lambda^{-1}\|^{-1}$, by Lemma 3.1 we have that A_0 is invertible and that $\|A_0^{-1}\| \le \|\Lambda^{-1}\|/(1 - \|\Lambda^{-1}\|\|A_0 - \Lambda\|) \le \beta/(1-\beta\delta) < \beta/(1 - 6\beta\delta)$. But, by (18) $1 - 6\beta\delta > 1 - \gamma/(1+\gamma) = 1/(1+\gamma)$, so

$$\|A_0^{-1}\| < (1+\gamma)\beta. \tag{21}$$

It follows by the mean-value theorem that

$$\begin{aligned}\|q_1 - q^*\| =& \|(q_1 - q_0) + (q_0 - q^*)\| = \| - A_0^{-1}F(q_0) + (q_0 - q^*)\| \\ =& \|A_0^{-1}\{-[F(q_0) - F(q^*) - F'(q^*)(q_0 - q^*)] + [A_0 - F'(q^*)](q_0 - q^*)\}\| \\ \le& (1+\gamma)\beta\Big\{\Big\| \int_0^1 [F'(q^* + t(q_0 - q^*)) - F'(q^*)](q_0 - q^*)\, dt\Big\| + \delta\|q_0 - q^*\|\Big\} \\ \le& (1+\gamma)\beta(L\eta/2 + \delta)\|q_0 - q^*\| \le \gamma\|q_0 - q^*\|.\end{aligned} \tag{22}$$

Next, by (15), (19), and (21)

$$\|A_1 - \Lambda\| \le (1 + \alpha_1\eta)\|A_0 - \Lambda\| + \alpha_2\eta \le \delta + (\alpha_1\delta + \alpha_2)\eta < 2\delta. \tag{23}$$

Using the induction, one proves

$$\|A_k - \Lambda\| \le 2\delta \quad \text{and} \quad \|q_{k+1} - q^*\| \le \gamma\|q_k - q^*\|. \tag{24}$$

In fact, we suppose that (24) are true for $k \le m-1$. By (15) we have

$$\|A_{k+1} - \Lambda\| \le (1 + \alpha_1\eta\gamma^k)\|A_k - \Lambda\| + \alpha_2\eta\gamma^k,$$

i.e.

$$\|A_{k+1} - \Lambda\| - \|A_k - \Lambda\| \le 2\alpha_1\eta\gamma^k\delta + \alpha_2\eta\gamma^k = (2\alpha_1\delta + \alpha_2)\eta\gamma^k. \tag{25}$$

Adding (25) from $k = 0$ to $m-1$, one gets

$$\|A_m - \Lambda\| \le \|A_0 - \Lambda\| + (2\alpha_1\delta + \alpha_2)\eta\sum_{k=0}^{m-1}\gamma^k < \delta + (2\alpha_1\delta + \alpha_2)\eta/(1-\gamma) \le 2\delta. \tag{26}$$

Therefore,

$$\|A_m - A_0\| \le \|A_m - \Lambda\| + \|A_0 - \Lambda\| < 3\delta, \quad \forall\, m. \tag{27}$$

In addition, $\|I - A_0^{-1}A_m\| \le \|A_0^{-1}\|\|A_m - A_0\| < 3(1+\gamma)\beta\delta < \gamma < 1$, where $I \in \mathcal{L}(Q)$ is the unit operator, by the Banach theorem $(A_0^{-1}A_m)^{-1}$ exists, hence A_m is invertible. Furthermore,

$$\|A_m^{-1}\| = \|[A_0 + (A_m - A_0)]^{-1}\| \le \|A_0^{-1}\| \sum \|A_0^{-1}\|^k\|A_m - A_0\|^k$$

$$
\begin{aligned}
&\le (1+\gamma)\beta \sum (3(1+\gamma)\beta\delta)^k = (1+\gamma)\beta/(1-3(1+\gamma)\beta\delta) \\
&< (1+\gamma)\beta/(1-6\beta\delta) < (1+\gamma)^2\beta, \quad \forall\, m,
\end{aligned} \tag{28}
$$

i.e. for any $m \in \mathbf{N}$, A_m is invertible and $\{\|A_m^{-1}\|\}$ are uniformly bounded as well. Moreover, by the assumption of induction one has $\|q_m - q^*\| < \eta$ and then

$$
\begin{aligned}
\|q_{m+1} - q^*\| =& \|(q_{m+1} - q_m) + (q_m - q^*)\| = \| - A_m^{-1}F(q_m) + (q_m - q^*)\| \\
=& \|A_m^{-1}\{-[F(q_m) - F(q^*) - F'(q^*)(q_m - q^*)] + [A_m - F'(q^*)](q_m - q^*)\}\| \\
\le& \|A_m^{-1}\| \Big\{ \Big\| \int_0^1 [F'(q^* + t(q_m - q^*)) - \Lambda](q_m - q^*)\, dt \Big\| \\
&+ \|A_m - \Lambda\| \|q_m - q^*\| \Big\} \le (1+\gamma)^2\beta(L\eta/2 + 2\delta)\|q_m - q^*\| \\
\le& \gamma\|q_m - q^*\|,
\end{aligned} \tag{29}
$$

$$
\|q_{m+1} - q^*\| \le \gamma\|q_m - q^*\| \le \cdots \le \gamma^{m+1}\|q_0 - q^*\| < \gamma^{m+1}\eta < \eta_0. \tag{30}
$$

Thus, $\{q_m\} \subset D_o$, $q_m \to q^*$ in Q, and $\{\|A_m\|\}$ and $\{\|A_m^{-1}\|\}$ are uniformly bounded. □

Now, we give the following main results:

Theorem 3.3. *We make the following assumptions:*

• *H1. Suppose that D is open, bounded, and convex and that $\exists q^* \in D$ such that $F(q^*) = 0$.*

• *H2. For any $q \in D$, $F(q)$ and $u(q)$ have continuous Fréchet derivatives of order two which both bounded above by a constant M. Moreover, $\Lambda \equiv F'(q^*) \in \mathcal{L}(Q, Y) \setminus \{0\}$ is invertible and $\|\Lambda^{-1}\| \le \beta$.*

Then the sequence $\{q_n\}$ defined by QNIS (12) *is well-defined and $q_n \to q^*$ in Q only providing that q_0 and A_0 satisfy the requirements of Lemma* 3.2.

Proof. First, by the argument of Lemma 3.2 we know that under the assumptions of Theorem 3.3 there exists $A_0 \in \mathcal{L}(Q, Y)$ satisfying the requirements of Lemma 3.2.

Next, in order to prove Theorem 3.3 we only need by Lemma 3.2 verifying the condition (15), i.e. we verify if the operator A_{n+1} defined by (12) satisfies (15) for some α_1 and α_2.

As a matter of fact, one gets

$$
\begin{aligned}
A_{n+1} - \Lambda &= A_n - \Lambda + (s_n, \cdot)(y_n - A_n s_n)/(s_n, s_n) \\
&= (A_n - \Lambda)[I - (s_n, \cdot)s_n/(s_n, s_n)] + (s_n, \cdot)(y_n - \Lambda s_n)/(s_n, s_n),
\end{aligned} \tag{31}
$$

where $I \in \mathcal{L}(Q)$ is the unit operator, and $\mathcal{L}(Q)$ is the linear bounded operator space on Q to Q. It is easy to check that $C_n \equiv I - (s_n, \cdot)s_n/(s_n, s_n) \in \mathcal{L}(Q)$ is self-adjoint and by the Cauchy inequality $\forall q \in Q$ $(C_n q, q) = (q, q) - (s_n, q)^2/(s_n, s_n) \ge 0$, and then

$$
\|C_n\| = \sup_{\|q\|=1} (C_n q, q) \le 1. \tag{32}
$$

Moreover, by the assumption H2 and the mean-value theorem one gets

$$
\|y_n - \Lambda s_n\| = \Big\| \int_0^1 [F'(q_n + t s_n) - F'(q^*)] s_n dt \Big\|
$$

$$\leq M\|s_n\| \int_0^1 \|t(q_{n+1}-q^*)+(1-t)(q_n-q^*)\|dt \leq M\sigma_n\|s_n\|. \tag{33}$$

Considering (32) and (33) from (31), one has

$$\|A_{n+1}-\Lambda\| \leq \|A_n-\Lambda\| + M\sigma_n. \tag{34}$$

So, the condition (15) is satisfied with $\alpha_1 = 0$ and $\alpha_2 = M$, and then it follows by Lemma 3.2 that the conclusions of Theorem 3.3 are true. □

About the Q-superlinear convergence for $\{q_n\}$ from [10] one can get

Theorem 3.4. *We assume that the requirements in Theorem 3.3 are satisfied and that $A_0 - F'(q^*)$ is compact. Then the sequence $\{q_n\}$ generated by QNIS is Q-superlinear convergent, i.e.*

$$\lim_{k\to\infty} \|q_{k+1}-q^*\|/\|q_k-q^*\| = 0,$$

provided $\|q_0-q^*\|$ is sufficiently small.

4. A Finite-dimensional Approximation of QNIS

We assume, besides the assumptions H1-H2 being true, in this section that Q and Y both are Hilbert spaces and that Q^n and Y^n are their n-dimensional subspaces, respectively, which satisfy the following

$$H3\begin{cases} \lim\limits_{n\to\infty} q^n = q, & \forall\, q\in Q \quad \text{with} \quad q^n = P_q^n q \\ \lim\limits_{n\to\infty} y^n = y, & \forall\, y\in Y \quad \text{with} \quad y^n = P_y^n y \end{cases}$$

where P_q^n and P_y^n are project operators on Q to Q^n and Y to Y^n, respectively, i.e. P_q^n and P_y^n both are self-adjoint, $P_q^n = (P_q^n)^2$, $P_y^n = (P_y^n)^2$, moreover, $Q^n = P_q^n Q$ and $Y^n = P_y^n Y$.

In order to implement QNIS by use of a computer we consider the map

$$F^m:\ Q^n \to Y^m, \quad F^m(q) = P_y^m F(q), \quad \forall\, q\in Q^n, \tag{35}$$

instead of $F(q)$. Similarly, we define

$$\Lambda^m:\ Q^n \to Y^m, \quad \Lambda^m q = P_y^m \Lambda q, \quad \forall\, q \in Q^n. \tag{36}$$

Since Λ is linear, one can choose m so large that $Y^m \supset \Lambda(Q^n)$, and then Λ^m is also invertible by the assumption H2. In addition, it follows by Lemma 3.1 that the operator $(F^m(\cdot))' : B \longrightarrow \mathcal{L}(Q^n, Y^m)$, $(F^m(q))'h = P_y^m F'(q)h$, $\forall h \in Q^n$, has a continuous inverse operator $[(F^m)'(q)]^{-1} \in \mathcal{L}(Y^m, Q^n)$, where $B \subset Q$ is the ball about q^* with radius δ.

Therefore, instead of (6) one considers the following approximation:

$$F^m(q) = 0. \tag{37}$$

Corresponding to F^m the Broyden-like procedure will be read as follows:

$$\begin{cases} \text{Given } q_0^n \in Q^n, \quad A_0^m \in \mathcal{L}(Q^n, Y^m), \\ A_k^m s_k^n = -F^m(q_k^n), \quad q_{k+1}^n = q_k^n + s_k^n, \\ y_k^m = F^m(q_{k+1}^n) - F^m(q_k^n), \quad A_{k+1}^m = A_k^m + (s_k^n, \cdot)_n (y_k^m - A_k^m s_k^n)/(s_k^n, s_k^n)_n, \end{cases} \tag{38}$$

where the inner product $(\cdot,\cdot)_n$ is induced by that of Q.

The differences between the algorithm (38) and the usual Broyden-like methods consist in: 1). the inner product $(\cdot,\cdot)_n$ is that of the n-dimensional vector space Q_n but, generally speaking, is not that of the Euclidean space, and 2). the operator $A_k^m \in \mathcal{L}(Q^n, Y^m)$ is not one mapping a space into itself.

Before proving the convergence of $\{q_k^n\}$ defined by (38) we give the following lemma by use of a secant method:

Lemma 4.1. *Let X and Z be Banach spaces and let $f : \mathcal{D}(f) \subset X \to Z$, where $\mathcal{D}(f)$ is the unit ball of X, satisfy*

- *f has continuous Fréchet derivatives of order two in $\mathcal{D}(f)$, both bounded above by a constant M, which is assumed to exceed 2.*
- *there exists a linear right inverse map $L(x)$ of $f'(x)$ with domain $\mathcal{D}(L) = \mathcal{D}(f)$ and range in the space $\mathcal{L}(Z, X)$, such that*

$$\|L(x)h\| \le M\|h\|, \quad \forall h \in Z, \quad x \in \mathcal{D}(L), \tag{39}$$

$$f'(x)L(x)h = h, \quad \forall h \in Z, \quad x \in \mathcal{D}(L). \tag{40}$$

Then, if $\|f(0)\| < M^{-4}$, it follows that $0 \in f(\mathcal{D}(f))$.

Proof. Take any $\theta \in (1/4, 7/8)$. Set $x_0 = 0$ and $\xi_n = -L(x_n)f(x_n)$, $x_{n+1} = x_n + \xi_n$, $\forall\, n \ge 0$.

We will prove inductively that

$$x_n \in \mathcal{D}(L), \quad \forall\, n \ge 1 \tag{41}$$

$$\|\xi_n\| \le \theta\|\xi_{n-1}\|, \quad \forall\, n \ge 1. \tag{42}$$

For $n = 0$,

$$\|x_1\| = \|\xi_0\| = \|L(0)f(0)\| \le M\|f(0)\| \le M^{-3} < 1, \tag{43}$$

so $x_1 \in \mathcal{D}(L)$. To prove that (42) is true when $n = 1$, we use the mean-value theorem with Lagrange remainder to $f(x+k)$:

$$f(x+k) = f(x) + f'(x)k + \int_0^1 (1-t)f''(x+tk)k^2\,dt.$$

Therefore,

$$\begin{aligned}
\|\xi_1\| =& \|L(x_1)f(x_1)\| \le M\|f(x_1)\| = M\Big\|f(0) + f'(0)\xi_0 + \int_0^1 (1-t)f''(t\xi_0)\xi_0^2\,dt\Big\| \\
=& M\Big\|[f(0) - f'(0)L(0)f(0)] + \int_0^1 (1-t)f''(t\xi_0)\xi_0^2\,dt\Big\| = M\Big\|\int_0^1 (1-t)f''(t\xi_0)\xi_0^2\,dt\Big\| \\
\le& M(M/2)\|\xi_0\|^2 \le M^2M^{-3}\|\xi_0\|/2 \le \|\xi_0\|/4 \le \theta\|\xi_0\|,
\end{aligned}$$

i.e. (42) is also true when $n = 1$.

Suppose that (41) and (42) are valid for $n \le m-1$, then

$$\begin{aligned}
\|x_m\| &= \|(x_m - x_{m-1}) + \cdots + (x_2 - x_1) + x_1\| \\
&\le \|\xi_{m-1}\| + \cdots + \|\xi_1\| + \|\xi_0\| \le (\theta^{m-1} + \cdots + \theta + 1)\|\xi_0\|
\end{aligned}$$

$$\leq \|\xi_0\|/(1-\theta) \leq M^{-3}/(1-\theta) \leq 1/[8(1-\theta)] < 1,$$

hence, $x_m \in \mathcal{D}(L)$. Next,

$$\begin{aligned}
\|\xi_m\| &= \|L(x_m)f(x_m)\| \\
&= \Big\|L(x_{m-1})\Big\{f(x_{m-1}) + f'(x_{m-1})\xi_{m-1} + \int_0^1 (1-t)f''(x_{m-1}+t\xi_{m-1})\xi_{m-1}^2\,dt\Big\}\Big\| \\
&\leq M\Big\|f(x_{m-1}) - f'(x_{m-1})L(x_{m-1})f(x_{m-1}) + \int_0^1 (1-t)f''(x_{m-1}+t\xi_{m-1})\xi_{m-1}^2 dt\Big\| \\
&\leq M\int_0^1 (1-t)\|f''(x_{m-1}+t\xi_{m-1})\|dt\|\xi_{m-1}\|^2 \\
&\leq (M^2/2)\|\xi_{m-1}\|\|\xi_{m-1}\| \leq M^2/2\,\theta^{m-1}\|\xi_0\|\|\xi_{m-1}\| \leq \theta\|\xi_{m-1}\|.
\end{aligned}$$

It follows by (42) that x_n converges to some $\tilde{x} \in \mathcal{D}(f)$. In addition, by means of $\|f(x_n)\| = \|f'(x_n)L(x_n)f(x_n)\| = \|f'(x_n)\xi_n\| \leq M\theta^n\|\xi_0\| \to 0$, we immediately get $f(\tilde{x}) = 0$. □

Theorem 4.2. *If the assumptions H1-H3 are valid. Then the sequence* $\{q_k^n\}$ *defined by the algorithm* (38) *is well-defined and* $q_k^n \to q^*$ *in* Q *as* $n \to +\infty$ *and* $k \to +\infty$*.*

Proof. Let $\epsilon > 0$ be arbitrary.

Since $F(q^*) = 0$ and the continuity of F, there is $\delta_2 \in (0, \epsilon)$ such that

$$\|F^m(q)\| \leq M^{-4}, \quad \forall\, q \in B(q^*, \delta_2) \cap Q^n \equiv B_2^n, \tag{44}$$

if ϵ is small enough and n is large enough, where $B(q^*, \delta_2) \equiv B_2$ is the ball of radius δ_2 about q^* in Q.

Next, it follows by the assumptions H1-H3 that $F^m : Q^n \to Y^m$ is twice continuously Fréchet differentiable, $[F^m(q)]' = P_y^m F'(q)$, and $[F^m(q)]'' = P_y^m F''(q)$. So, it is evident that $\|[F^m(q)]'\|$, $\|[F^m(q)]''\| \leq M$, $\forall q \in B_2^n$.

Moreover, the operator $f_m : \mathcal{D}(f) \to \mathcal{L}(Q^n, Y^m)$, $f_m(q) = [F^m(q)]' \in \mathcal{L}(Q^n, Y^m)$, and its inverse operator $L_m : \mathcal{D}(L) \to \mathcal{L}(Y^m, Q^n)$, $L_m(q) = ([F^m(q)]')^{-1} \in \mathcal{L}(Y^m, Q^n)$, where $\mathcal{D}(f) = \mathcal{D}(L) = B_2^n$, satisfy the assumptions of Lemma 4.1.

Thus, it follows by Lemma 4.1 that there exists $q^n \in B_2 \cap Q^n$, i.e. $\|q^n - q^*\| < \delta_2 < \epsilon$, such that

$$F^m(q^n) = 0. \tag{45}$$

To use Lemma 3.2, it is sufficient to verify if the condition (15) is satisfied for F^m. In fact,

$$\begin{aligned}
A_{k+1}^m - \Lambda_n^m &= A_k^m - \Lambda_n^m + (s_k^n, \cdot)_n(y_k^m - A_k^m s_k^n)/(s_k^n, s_k^n)_n \\
&= (A_k^m - \Lambda_n^m)[I - (s_k^n, \cdot)_n s_k^n/(s_k^n, s_k^n)_n] + (s_k^n, \cdot)_n(y_k^m - \Lambda_n^m s_k^n)/(s_k^n, s_k^n)_n,
\end{aligned}$$

where $\Lambda_n^m \equiv [F^m(q)]'\,|_{q=q^n}$. Following the proof of Lemma 3.2, one gets $\|A_{k+1}^m - \Lambda_n^m\| \leq \|A_k^m - \Lambda_n^m\| + M\sigma_k^n$, where $\sigma_k^n \equiv \max(\|q_{k+1}^n - q^n\|, \|q_k^n - q^n\|)$. Thus, by Lemma 3.2 one has $\|q_k^n - q^n\| < \gamma^k\eta$, where $\gamma \in (0, 1)$ is arbitrary and $\eta = \eta(\gamma)$ is determined by Lemma 3.2. Hence, we take k so large that $\|q_k^n - q^n\| < \epsilon$. Therefore, $\|q_k^n - q^*\| \leq \|q_k^n - q^n\| + \|q^n - q^*\| < 2\epsilon$. □

5. A Parabolic Inverse Problem

Various physical phenomena have led to a study of mixed initial-boundary value problems for parabolic equations. The parabolic inverse problem we address is to find the function pair (u, q) satisfying the following:

$$\begin{aligned} &u_t - u_{xx} + q(x)u = 0, \quad (x,t) \in \Sigma \equiv (0,1) \times (0,\pi), \\ &u(0,t) = g(t), \quad u(1,t) = 0, \quad t \in [0,\pi] \\ &u(x,0) = 0, \quad x \in [0,1], \quad q \in Q^+. \end{aligned} \tag{46}$$

and

$$u(x,\pi) = z(x), \quad x \in [0,1], \tag{47}$$

where $g(t) = 1 - \cos t$, $t \in [0,\pi]$, $z \in Y^+$ is given, Y^+ is defined by $Y^+ \equiv \{y \in Y;\ y \geq 0,\ y(0) = g(\pi),\ y(1) = 0,\ v_t - y_{xx} \leq 0\}$, $Y \equiv C^{2+\lambda}([0,1])$, and v is the solution of (46) with $q = 0$. We assume in the example that the true value of q is $q_{tr} = 2 + \sin \pi x$, so $z(\cdot) = u(\cdot, \pi; q_{tr})$. Moreover, assume that $Q \equiv C^{\lambda}([0,1])$ and that $Q^+ \equiv \{q \in Q;\ q \geq 0\}$.

It is well-known that $\forall q \in Q$ there exists a unique solution $u \in V \equiv C^{2+\lambda}(\bar{\Sigma})$ to (46) by [17], one denotes it by $u = u(q) = u(x,t;q)$ to show the dependence of u on q.

To begin with, we have

Theorem 5.1. *The function $u : Q \to V$ defined by* (46) *is infinitely differentiable, i.e. $u \in C^{(n)}(Q;V)$, $n \in \mathbf{N} \cup \{0\}$, where $C^{(n)}(Q;V)$ denotes the linear space of n-times Fréchet continuously differentiable functions on Q to V. Moreover, the first Fréchet derivative $u'(\cdot) : Q \to \mathcal{L}(Q;V)$ and the second Fréchet derivative $u''(\cdot) : Q \to \mathcal{L}(Q;\mathcal{L}(Q;V))$ of u at q are determined implicitly by $u'(q)h = \dot{u}$ and $u''(q)hk = \ddot{u}$, $\forall h,k \in Q$, respectively, where $\dot{u}$ and $\ddot{u}$ are defined by the following systems:*

$$\begin{aligned} &\dot{u}_t - \dot{u}_{xx} + q\dot{u} = -hu(q), \quad (x,t) \in \Sigma, \\ &\dot{u}(0,t) = \dot{u}(1,t) = 0, \quad t \in [0,\pi], \\ &\dot{u}(x,0) = 0, \quad x \in [0,1], \end{aligned} \tag{48}$$

and

$$\begin{aligned} &\ddot{u}_t - \ddot{u}_{xx} + q\ddot{u} = -k[u'(q)h] - h[u'(q)k], \quad (x,t) \in \Sigma, \\ &\ddot{u}(0,t) = \ddot{u}(1,t) = 0, \quad t \in [0,\pi], \\ &\ddot{u}(x,0) = 0, \quad x \in [0,1]. \end{aligned} \tag{49}$$

Proof. First, we prove $u \in C^{(n)}(Q;V)$, $n \in \mathbf{N} \cup \{0\}$.

Taking, for example, $n = 0$, we can deduce the proof similarly for any order n. Take $q, \tilde{q} \in Q$, and then by (46) get $u = u(q)$ and $\tilde{u} = u(\tilde{q})$. Set $h = \tilde{q} - q$ and $\delta u = \tilde{u} - u$, so δu satisfies

$$(\delta u)_t - (\delta u)_{xx} + q(\delta u) = -h\tilde{u}, \quad (x,t) \in \Sigma,$$

$$
\begin{aligned}
&(\delta u)(0,t) = (\delta u)(1,t) = 0, \quad t \in [0,\pi], \\
&(\delta u)(x,0) = 0, \quad x \in [0,1].
\end{aligned}
\tag{50}
$$

Clearly, $\|\delta u\|_V = O(\|h\|_Q)$ by [17]. Hence, $u(\cdot) \in C(Q; V)$.

Secondly, we prove (48). Taking $q, h \in Q$, there exists a unique solution $\dot{u} \in V$ to the problem (48) by [17]. Thus, set $\tilde{q} = q + h$, $u = u(q)$, $\tilde{u} = u(\tilde{q})$, and $\hat{u} = \tilde{u} - u - \dot{u}$. It is evident that $\hat{u}$ satisfies

$$
\begin{aligned}
&\hat{u}_t - \hat{u}_{xx} + q\hat{u} = -h(\tilde{u} - u), \quad (x,t) \in \Sigma, \\
&\hat{u}(0,t) = \hat{u}(1,t) = 0, \quad t \in [0,\pi], \\
&\hat{u}(x,0) = 0, \quad x \in [0,1].
\end{aligned}
\tag{51}
$$

It follows by [17] that $\|\tilde{u} - u\|_V = o(1)$, and then $\|\hat{u}\|_V = o(\|h\|_Q)$. Thus, $\dot{u} = u'(q)h$.

Similarly, one can get $u''(q)hk = \ddot{u}$, $\forall h, k \in Q$.

Next, we recognize the inverse problem (46) with (47) as solving the nonlinear operator equation:

$$
F(q) = 0, \tag{52}
$$

where $F : Q \to Y$ is defined by $F(q) \equiv u(\cdot, \pi; q) - z$ and $Y \equiv C^{2+\lambda}([0,1])$.

We have the following properties about the operator $F(q)$:

1. $F(\cdot) \in C^{(n)}(Q; Y)$, $n \in \mathbf{N} \cup \{0\}$ by Theorem 5.1;
2. the operator $F(q)$ has a unique zero point $q^* \in Q^+$ since the inverse problem $u(\cdot, \pi; q) = z$, $\forall z \in Y^+$, is well-posed by [11];
3. $\forall q \in Q^+$, $u(q) \geq 0$, $u_t(q) \geq 0$, and $u(\cdot, \pi; q) > 0$ by $g \geq 0$, $g_t \geq 0$, and the maximum principle for parabolic equations;
4. $\forall q \in Q^+$, $\forall y \in \tilde{Y}$, where $\tilde{Y} \equiv \{y \in Y;\, y(0) = g(\pi),\, y(1) = 0\}$, the inverse problem $F'(q)h = u'(q)h(\cdot, \pi) \equiv \dot{u}(\cdot, \pi) = y$ has a unique solution $h \in Q$ by [11] and h continuously depends on y. That is, the operator $F'(q) : Q \to \tilde{Y}$ is invertible and the inverse operator is bounded.
5. The operator $\Lambda \equiv F'(q^*) \in \mathcal{L}(Q; Y)$ is compact, which can be deduced from the following formula about $F'(q^*)$:

$$
F'(q^*)h\,(x) = -\int_0^\pi \int_\Omega G(x, \xi, \pi, \tau) u(\xi, \tau; q^*) d\xi d\tau, \quad x \in (0,1),
$$

and the properties of the kernel function $G(x, \xi, t, \tau) u(\xi, \tau; q^*)$[17], where $G(x, \xi, t, \tau)$ is the Green function of (48), which is determined by the following

$$
\begin{aligned}
&[\partial_t - \partial_{xx} + q^*(x)] G(x, \xi, t, \tau) = \delta(x - \xi)\delta(t - \tau) \\
&\qquad G(x, \xi, \tau, \tau) = 0, \quad G(x, \xi, t, \tau)\,|_{\xi=0,1} = 0.
\end{aligned}
$$

If one uses QNIS to solve the equation (52), then the sequence $\{q_n\}$ generated by QNIS will Q-superlinearly converge to q^* by Theorem 3.4 because the initial guess A_0 is of finite rank and $A_0 - F'(q^*)$ is compact.

Since the classical solution to a PDE is also a generalized solution in the Ladyzenskaja sense, from now on, one takes $Q = L^2([0,1])$ and $V = H^1(\Sigma)$.

Therefore, $Y = H^1([0,1])$. We choose $M, N, L \in \mathbf{N}$ with $M \geq N$, and then set $h = 1/N$, $k = 1/M$, $\tau = \pi/L$, and

$$Q^N = \Big\{q \in Q; \quad q(x) = \sum_{i=0}^{N-1} c_i \chi_i(x), \quad c_i \in \mathbf{R}\Big\},$$
$$Y^M = \{y \in Y; \quad y'(x) = \sum_{i=0}^{M-1} d_i \tilde{\chi}_i(x), \quad d_i \in \mathbf{R}\Big\},$$

where $\chi_i(x)$ and $\tilde{\chi}_i(x)$ are characteristic functions of $\Omega_i = \{x; ih < x < (i+1)h\}$, $i = 0, \ldots, N-1$, and $\omega_i = \{x; ik < x < (i+1)k\}$, $i = 0, \ldots, M-1$, respectively.

We use the Crank-Nicolson implicit finite difference method to discretize (46), that is, expand u at the point $(ih, (j+0.5)\tau)$:

$$u(ih,(j+0.5)\tau) \approx (u_{i,j+1} + u_{i,j})/2, \quad u_t(ih,(j+0.5)\tau) \approx (u_{i,j+1} - u_{i,j})/\tau,$$
$$u_x(ih,(j+0.5)\tau) \approx \{(u_{i+1,j+1} - u_{i,j+1}) + (u_{i+1,j} - u_{i,j})\}/2h,$$
$$u_{xx}(ih,(j+0.5)\tau) \approx \{(u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}) + (u_{i+1,j} - 2u_{i,j} + u_{i-1,j})\}/2h^2.$$

where $u_{i,j}$ is the value of u at the mesh point $(ih, j\tau)$, i.e. $u_{i,j} = u(ih, j\tau)$, $\forall i, j = 0, 1, 2, \ldots$, etc.

Substitute the above into (46), and then one could obtain the following two-level stable approximate equations:

$$\begin{aligned} &-u_{i-1,j+1} + (2\rho + 2 + q_i h^2)u_{i,j+1} - u_{i+1,j+1} = -u_{i-1,j} - (2 - 2\rho + q_i h^2)u_{i,j} + u_{i+1,j}, \\ &u_{0,j} = g_j, \quad u_{N,j} = 0, \quad u_{i,0} = 0, \quad i = 0, 1, \ldots, N, \quad j = 0, 1, \ldots, L, \end{aligned} \tag{53}$$

where $\rho = \tau/2h^2$, $g_j = g(j\tau)$, and $q_i = q(ih)$.

The initial operator A_0 is defined by

$$A_0 s = \sum_{j=0}^{M} \sum_{i=0}^{N} \tilde{c}_{ij}(g^i, s) F^j, \tag{54}$$

where $\{F^0, \ldots, F^M\}$ and $\{g^0, \ldots, g^N\}$ are bases of Y^M and Q^N, respectively, $g^i = \{g_0^i, \ldots, g_N^i\}$, $(g^i, s) = h \sum_j [g_j^i s_j + (g^i)'(jh) s'(jh)]$, $(g^i)'(jh) = (g_{j+1}^i - g_j^i)/h$, and $s'(jh) = (s_{j+1} - s_j)/h$.

By a simple computation the operator $A_n \in \mathcal{L}(Q, Y)$ will become the following form: $A_n s = -f$, $f_j = -\sum_{i=0}^{N}(c_{ij} s_i + \lambda \delta_{ij})$, where λ is a small number, δ_{ij} is the Kronecker function, $f = (f_0, \ldots, f_M)$, and $s = (s_0, \ldots, s_N)$.

In the paper we take $h = 0.1$, $k = 0.1$, and $\tau = 0.05$. Moreover, take A_0 with $c_{ij} = \alpha \delta_{ij}$, $\forall i = 0, \ldots, N; j = 0, \ldots, M$.

The computational results are summarized in Tables 1 and 2, where q_i^{tr} and q_i^{cal} are the true value and the calculated value of q at ih, respectively, and

$$\delta q_i = q_i^{cal} - q_i^{tr} \qquad f_j = u_{j,L} - z(jh) \qquad m = \sum_{i=0}^{N} \delta q_i/(N+1)$$
$$p = \sum_{j=0}^{M} f_j/(M+1) \quad \sigma_1^2 = \sum_{i=0}^{N} (\delta q_i - m)^2/N \quad \sigma_2^2 = \sum_{j=0}^{M} (f_j - p)^2/M$$

We stop computation if the inequalities $\|s_l\| < \epsilon_1$ and $\|f^l\| < \epsilon_2$ both are true, where $s_l \equiv q^{l+1} - q^l$, $q^l \equiv (q_0^l, \ldots, q_I^N)$, and q_i^l is the value of q_i at the lth iteration, similarly, $f^l \equiv (f_0^l, \ldots, f_M^l)$. In the paper we take $\epsilon_1 = 10^{-4}$ and $\epsilon_2 = 10^{-5}$.

From these results one can obtain the following conclusions:

1. The initial guesses of q_0 and A_0 are very important for computation. In our example if take $q_0 > 2.5$ or $\alpha < 1.5$ in A_0, then the computation will be divergent.
2. QNIS is effective and saves time provided that the initial guess is taken properly.
3. Because the Fréchet derivative $F'(q)$ arisen from inverse problems in PDEs usually is compact, the convergence of the sequence $\{q_n\}$ generated by QNIS is Q-superlinear if the other assumptions of Theorem 3.4 are satisfied. Thus, the quasi-Newton method is suitable for solving inverse problems in PDEs.
4. In the paper we only prove local convergence. If we use some kind of hybrid method, the global convergence will be obtained, which is not stated here.

Table 1 Effect of Initial Value on Estimation of q, case 1

Iteration Times	m	σ_1^2	p	σ_2^2	Initial Guess
1	-2.0909	7.9091E-1	1.8256	3.0748	$c_{ij} = \alpha\delta_{ij}$
2	-6.0679E-1	5.1754E-1	7.6253E-1	6.0021E-1	$\alpha = 0.9$
3	-5.0045E-1	4.1754E-2	9.8252E-2	8.1785E-3	$q_0 = 2.1$
4	-7.0483E-2	9.0751E-3	5.0621E-3	4.5713E-5	
5	-4.8580E-2	5.8417E-3	4.6159E-3	7.4731E-5	
6	-1.7007E-2	9.5712E-4	1.3265E-3	2.0831E-6	
7	-7.0023E-3	3.8707E-5	8.8419E-4	7.3246E-7	
8	-1.1025E-4	8.3923E-7	4.9761E-5	8.0389E-8	
9	7.7385E-5	5.8038E-7	8.8586E-6	1.0715E-8	
10	3.6024E-6	6.6409E-8	3.4673E-6	5.7302E-9	

Table 2 Effect of Initial Value on Estimation of q, case 2

Iteration Times	m	σ_1^2	p	σ_2^2	Initial Guess
1	-2.0909	7.9091E-1	1.8256	3.0748	$A_0 = (c_{i,j})$
2	-9.0679E-2	5.0162E-4	1.5326E-2	6.4815E-4	$c_{i,j} = \alpha\delta_{i,j}$
3	-1.5064E-2	5.0162E-4	2.5019E-3	7.4852E-6	$q_0 = 0.8$
4	-4.5496E-3	9.7901E-5	7.3288E-4	5.8941E-8	$\alpha = 1.0$
5	2.1989E-4	4.1855E-8	3.7318E-4	4.9137E-8	
6	7.6296E-5	8.9047E-10	4.0518E-5	7.4392E-10	

References

[1] H.T. Banks, K. Kunisch, An approximation theory for nonlinear partial differential equations with applications to identification and control, *SIAM J. Control and Optimization,* **20** : 6 (1982), 815–849.

[2] H.T. Banks, K. Kunisch, Estimation Techniques for Distributed Parameter Systems, Birkhäuser, Boston, Basel, and Berlin, 1989.

[3] H.T. Banks, S. Reich, I.G. Rosen, An approximation theory for the identification of nonlinear distributed parameter systems, *SIAM J. Control and Optim.,* **28** : 1 (1990), 523–569.

[4] C.G. Broyden, A class of methods for solving nonlinear simultaneous equations, *Math. Comp.,* **19**(1965), 577–593.

[5] C.G. Broyden, J.E. Dennis, J.J. Moré, On the local and superlinear convergence of quasi-Newton methods, *J. Inst. Math. Appl.,* **12** (1973), 223–246.

[6] G. Chavent, Identification of distributed parameter systems: About the output least square method, its implementation and identifiability, in Identification and System Parameter Estimation, ed. by R. Isermann, 85-97, Pergamon, New York, 1982.

[7] Y.M. Chen, J.Q. Liu, An iterative numerical algorithm for solving multi-parameter inverse problems of evolutional partial differential equations, *J. Comput. Phys.,* **53** (1984), 429–442.

[8] Y.M. Chen, G.Q. Xie, An iterative method for simultaneous determination of balk and shear moduli and density variations, *J. Comput. Phys.,* **62** (1986), 143–163.

[9] Y.M. Chen, F.G. Zhang, Hierarchical multigrid strategy for efficiency improvement of the GPST inversion algorithm, *Applied Numerical Mathematics,* **6** (1990), 431–446.

[10] A. Griewank, The local convergence of Broyden-like methods on Lipchitzian problems in Hilbert spaces, *SIAM J. Numer. Anal.,* **24** (1987), 684–705.

[11] V. Isakov, Inverse Problems with the final overdetermination, *Comm. Pure and Appl. Math.,* **V. XLIV** (1991), 185–209.

[12] L.V. Kantorovich, G.P. Akilov, Functional Analysis in Normed Spaces, MacMillan Company, New York, 1964.

[13] C.T. Kelley, E.W. Sachs, A quasi-Newton method for elliptic boundary value problems, *SIAM J. Numer. Anal.,* **24** : 3 (1987), 516–531.

[14] C.T. Kelley, E.W. Sachs, Quasi-Newton methods and unconstrained optimal control problems, *SIAM J. Control and Optim.,* **23** (1987), 1503–1516.

[15] C. Kravaris, J.H. Seinfeld, Identifiability of spatially-varying conductivity from point observation as an inverse Sturm-Liouville problem, *SIAM J. Control* & *Optim.,* **24** :3 (1986), 522–542.

[16] C.K. Kunisch, E.W. Sachs, Reduced SQP methods for parameter identification problems, *SIAM J. Numer. Anal.,* **29** (1992), 1793–1820.

[17] O.A. Ladyzenskaja et al., Linear and Quasi-linear Equations of Parabolic Type, *AMS, Providence* (*RI*)*,* 1968.

[18] L. Lapidus, Numerical Solution of Partial Differential Equations in Sciences and Engineering, John Wiley & Sons, New York, 1982.

[19] Tai-Yong Lee, John H. Seinfeld, Estimation of two-phase petroleum reservoir properties by regularization, *J. Comput. Phys.,* **69** (1987), 397–419.

[20] J.H. Seinfeld, W.H. Chen, Identification of petroleum reservoir properties, Distributed Parameter Systems: Identification, Estimation and Control, ed. by W. H. Ray and D. G. Lainiotis, Marcel Dekker, Inc., New York and Basel, 1978.

[21] W.H. Yu, A Rapid convergent iteration method for identifying distributed parameters in partial differential equations, Preprints 1988'IFAC Symp. Identification and System Parameter Estimation, Beijing (China), Oct.,1988.

[22] W.H. Yu, J.H. Seinfeld, Identification of parabolic distributed parameter systems by regularization with differential operator, *J. Math. Anal. Appl.,* **132** (1988), 365–398.

[23] W.H. Yu, J.H. Seinfeld, Identification of distributed parameter systems with pointwise constrains on the parameters, *J. Math. Anal. Appl.,* **136** (1988), 497–520.

[24] W.H. Yu, Solving inverse problems for hyperbolic equations via the regularization method, *J. Computational Mathematics,* **11** : 2 (1993), 142–153.

[25] Jian-ping Zhu, Yung-ming Chen, Multilevel grid method for history matching multi-dimensional multi-phase reservoir models, *Applied Numer. Math.,* **10** (1992), 159–174.