A convergent numerical method to recover the initial condition of nonlinear parabolic equations from lateral Cauchy data

Thuy T. Le and Loc H. Nguyen
Department of Mathematics and Statistics, University of North Carolina at Charlotte, Charlotte, NC 28223
E-mail: thuy.le@uncc.edu, loc.nguyen@uncc.edu

Keywords: Nonlinear parabolic equations, initial condition, convergent numerical method, Carleman estimate, iteration

1. Introduction

Let \( d \geq 1 \) be the spatial dimension and \( T > 0 \). Let \( q : \mathbb{R} \rightarrow \mathbb{R} \) and \( c : \mathbb{R}^d \rightarrow \mathbb{R} \) be smooth functions in the class \( C^1 \). Assume that \( c(x) \geq c_0 \) for some \( c_0 > 0 \). Consider the problem

\[
\begin{align*}
    c(x)u_t(x,t) &= \Delta u(x,t) + q(u(x,t)) & x \in \mathbb{R}^d, t \in (0,T) \\
    u(x,0) &= p(x) & x \in \mathbb{R}^d,
\end{align*}
\]

where \( p \) is a source function compactly supported in an open and bounded domain \( \Omega \) of \( \mathbb{R}^d \) with smooth boundary \( \partial \Omega \). In this paper, we assume \( [1] \) has a unique bounded solution. The smoothness property of this bounded solution can be found in \( [1] \) Chapter V, §6]. We are interested in the following problem.
Problem 1 (Inverse Source Problem) Assume that $|u(x, t)|$ does not blow up as $t \to T$ for all $x \in \Omega$. Given the lateral Cauchy data

$$f(x, t) = u(x, t) \quad \text{and} \quad g(x, t) = \partial_x u(x, t)$$

for $x \in \partial \Omega$, $t \in [0, T]$, determine the function $u(x, 0) = p(x), x \in \Omega$.

Problem 1 arises from the problem of recovering the initial condition of parabolic equations from the lateral Cauchy data. It has many real-world applications; for e.g., determine the spatially distributed temperature inside a solid from the boundary measurement of the heat and heat flux in the time domain [2]; identify the pollution on the surface of the rivers or lakes [3]; effectively monitor the heat conductive processes in steel industries, glass and polymer forming and nuclear power station [4]. When the nonlinear term $q(u)$ takes the form $u(1 - u)$ (or $q(u) = u(1 - |u|^\alpha)$) for some $\alpha > 0$, the parabolic equation in (1) is called the high dimensional version of the well-known Fisher (or Fisher-Kolmogorov) equation. It is worth mentioning that Fisher equation occurs in ecology, physiology, combustion, crystallization, plasma physics, and in general phase transition problems, see [5]. Due to its realistic applications, the problem of determining the initial condition of parabolic equations has been studied intensively; however, up to the knowledge of the authors, numerical solutions are computed only in the case when the nonlinearity is absent, see e.g., [11]. The uniqueness of Problem 1 is well-known assuming that the nonlinearity $q$ is in the class $C^1$, see [6]. On the other hand, the logarithmic stability results were rigorously proved in [2, 4]. For the completeness, we briefly recall the logarithmic stability of Problem 1 in this paper. The natural approach to solve this problem is the optimal control method; that means, minimizing some mismatch functionals. However, since the known stability is logarithmic [2, 4], the optimal control approach might not give good numerical results; especially, when the initial guess, if provided, is far away from the true solution. A more important reason for us to not use the optimal control method is that the cost functional is nonconvex and have multi-minima. We draw the reader’s attention to the convexification method, see [7, 8], which can overcome the difficulty about the lack of the initial guess. In those papers [7, 8], the authors introduce a convex functional, whose minimizer yields the solution of their problem, by combining the quasi-reversibility method and Carleman weight functions.

As mentioned, since a good initial guess of the true solution of Problem 1 is not always available, the optimal control method, which is widely used in the scientific community, might not be applicable. To overcome this difficulty, we propose to solve Problem 1 in the Fourier domain with respect to a special basis of $L^2(0, T)$. This basis was first introduced in [9] by Klibanov and then successfully used very often to solve a variety of inverse problems by our research group. More precisely, we derive a system of elliptic PDEs whose solution consists of a finite number of the Fourier coefficients of the solution to the parabolic equation [1]. Solution of this system directly yields the knowledge of the function $u(x, t)$, from which the solution to our inverse problem follows. We numerically solve this nonlinear system by an iterative process. The initial solution
Recovering the initial conditions of nonlinear parabolic equations can be computed by solving the system obtained by removing the nonlinear term. Then, we approximate the nonlinear system by replacing the nonlinearity by the one acting on the initial solution obtained in the previous step. Solving this approximation system, we find an updated solution. Continuing this process, we get a fast convergent sequence reaching to the desired function. The convergence is proved by using a new Carleman estimate.

Two papers closely related to the current one are [10] and [11]. In [10], a source term for a nonlinear parabolic equation is computed and in [11], the second author and his collaborator computed the initial condition of linear parabolic equation from the lateral Cauchy data. On the other hand, the coefficient inverse problem for parabolic equations is also very interesting and studied intensively. We draw the reader’s attention to [12, 13, 14, 15, 16, 17, 18] for important numerical methods and good numerical results. Besides, the problem of recovering the initial conditions for hyperbolic equation is very interesting since it arises in many real-world applications. For instance, the problems thermo and photo acoustic tomography play the key roles in bio-medical imaging. We refer the reader to some important works in this field [19, 20, 21]. Applying the Fourier transform, one can reduce the problem of reconstructing the initial conditions for hyperbolic equations to some inverse source problems for the Helmholtz equation, see [22, 23, 24, 25, 26] for some recent results.

The paper is organized as follows. In Section 2, we derive a nonlinear system of elliptic PDEs, which leads to a numerical method to solve Problem 1. In this section, we also briefly discuss the uniqueness of this inverse problem for the completeness’ sake. In Section 3, we establish and prove a Carleman estimate. This estimate plays an important role in the proof of our main theorem in Section 4. In Section 4, we propose an iterative scheme to solve the nonlinear system obtained in Section 2. Especially, in Section 4, we prove the convergence of the iterative sequence. In Section 5, we discuss the implementation of our method and show several numerical results. Section 6 is for concluding remarks.

2. A numerical method to find the desired initial condition

The main aim of this section is to introduce the arguments leading to our algorithm to solve Problem 1. We first discuss the stability of the inverse problem.

2.1. The uniqueness and logarithmic stability

In this section, we briefly discuss about the uniqueness and stability of the reconstruction of \( p(x) \), \( x \in \Omega \), from the lateral Cauchy data \( f(x, t) \) and \( g(x, t) \), \( (x, t) \in \partial \Omega \). These results can be deduced from [2, Theorem 3].
Let \( u^1(x,t) \) and \( u^2(x,t) \) be solutions to
\[
\begin{align*}
  c(x)u^i_t(x,t) &= \Delta u^i(x,t) + q(u^i(x,t)) \quad (x,t) \in \Omega \times [0,T], \\
  u^i(x,t) &= f^i(x,t) \quad (x,t) \in \partial\Omega \times [0,T], \\
  \partial_n u^i(x,t) &= g^i(x,t) \quad (x,t) \in \partial\Omega \times [0,T], \\
  u^i(x,0) &= p^i(x) \quad x \in \Omega,
\end{align*}
\]
i \in \{1,2\}. Assume that
\[
\max\{\|u^1\|_{L^\infty(\Omega \times [0,T])}, \|u^2\|_{L^\infty(\Omega \times [0,T])}\} < M
\]
for some constant \( M \). The following result holds true.

**Proposition 1 (The uniqueness and the logarithmic stability)** Let \( A \) be a number with \( A > \max_{(x,t) \in \partial\Omega \times [0,T]}(\|f^1(x,t)| + |f^2(x,t)| + |g^1(x,t)| + |g^2(x,t)|) \). Denote by \( E \) the mismatch of the data
\[
E = \|f^2(x,t) - f^1(x,t)\|_{H^1(\partial\Omega \times [0,T])} + \|\partial_n (g^2(x,t) - g^1(x,t))\|_{L^2(\partial\Omega \times [0,T])}.
\]
Then, for any \( \beta \in (0,2) \), we can find a constant \( \epsilon_0 \in (0,1) \) depending only on \( M, c, q, T, \Omega \) and \( \beta \) such that
\[
\|p^1 - p^2\|_{L^2(\Omega)} \leq \frac{C}{\beta \ln\left(\frac{M}{\epsilon_0}\right)} \|\nabla(p^1 - p^2)\|_{L^2(\Omega)} + C\left(\frac{A}{\epsilon_0}\right)^\beta E^{2-\beta}
\]
where \( C \) is a constant depending only on \( M, c, q, T \) and \( R \).

**Proof:** Denote by the function \( v(x,t) = u^2(x,t) - u^1(x,t) \), \( x \in \Omega, t > 0 \). We have for all \( x \in \Omega \) and \( t \in [0,T] \)
\[
|c(x)v_t(x,t) - \Delta v(x,t)| = q(u^2(x,t)) - q(u^1(x,t)) \\
\leq \max_{s \in [-2M,2M]} |q'(s)| |v(x,t)|.
\]
Applying Theorem 3 in [2], we obtain (4). \( \square \)

**Remark 1** A direct consequence of (4) is the uniqueness of Problem 1. Due to the presence of the first term in the right hand side, the stability Problem 1 is of the logarithmic rate. The optimal control method might not lead to good numerical results.

In the next subsection, we derive a system whose solutions directly yields numerical solutions to Problem 1.

### 2.2. A system of nonlinear elliptic equations

We will employ a special basis of \( L^2(0,T) \). For each \( n = 1,2,\ldots, \) set \( \phi_n(t) = (t-T/2)^{n-1} \exp(t-T/2) \). The set \( \{\phi_n\}_{n=1}^\infty \) is complete in \( L^2(0,T) \). Applying the Gram-Schmidt orthonormalization process to this set, we obtain a basis of \( L^2(0,T) \), named as \( \{\Psi_n\}_{n=1}^\infty \). We have the proposition

**Proposition 2 (see [9])** The basis \( \{\Psi_n\}_{n=1}^\infty \) satisfies the following properties:
Recovering the initial conditions of nonlinear parabolic equations

(i) $\Psi_n' \text{ is not identically zero for all } n \geq 1,$

(ii) For all $m, n \geq 1$

\[ s_{mn} = \int_0^T \Psi_n'(t)\Psi_m(t)dt = \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{if } n < m. \end{cases} \]

As a result, for all integer $N > 1$, the matrix $S = (s_{mn})_{m,n=1}^N$, is invertible.

This basis was originally introduced to solve the electrical impedance tomography problem with partial data in [9]. Since then, this basis is widely used to solve a variety kinds of inverse problems. The most related paper with the current one is [11], in which the second author and his collaborator employed this basis to recover the initial condition for linear parabolic equations. For each point $x \in \Omega$, we can approximate $u(x, t), t \in [0, T]$ as

\[ u(x, t) = \sum_{n=1}^{\infty} u_n(x) \Psi_n(t) \simeq \sum_{n=1}^{N} u_n(x) \Psi_n(t) \]  \hspace{1cm} (5)

where

\[ u_n(x) = \int_0^T u(x, t)\Psi_n(t)dt \quad n \geq 1. \]  \hspace{1cm} (6)

The “cut-off” number $N$ will be determined numerically later in Section 5 using the given data, see also Figure 1 for an illustration. Due to (5), the function $u_t(x, t)$ is approximated by

\[ u_t(x, t) \simeq \sum_{n=1}^{N} u_n(x) \Psi_n'(t) \quad x \in \Omega, t \in [0, T]. \]  \hspace{1cm} (7)

From now on, we replace the approximation “$\simeq$” by the equality. Plugging (5) and (7) into the governing equation in (1), we obtain

\[ c(x) \sum_{n=1}^{N} u_n(x) \Psi_n'(t) = \sum_{n=1}^{N} \Delta u_n(x) \Psi_n(t) + q \left( \sum_{n=1}^{N} u_n(x) \Psi_n(t) \right) \]  \hspace{1cm} (8)

for all $x \in \Omega$. For each $m = 1, \ldots, N$, multiply $\Psi_m(t)$ to both sides of (8) and then integrate the resulting equation with respect to $t$ on $[0, T]$. For all $x \in \Omega$, we have

\[ c(x) \sum_{n=1}^{N} u_n(x) \int_0^T \Psi_n'(t)\Psi_m(t)dt \]

\[ = \sum_{n=1}^{N} \Delta u_n(x) \int_0^T \Psi_n(t)\Psi_m(t)dt + \int_0^T q \left( \sum_{n=1}^{N} u_n(x) \Psi_n(t) \right) \Psi_m(t)dt. \]  \hspace{1cm} (9)

The system (9) with $m = 1, \ldots, N$ becomes

\[ c(x) \sum_{n=1}^{N} s_{mn} u_n(x) = \Delta u_m(x) + q_m(u_1(x), u_2(x), \ldots, u_N(x)) \]  \hspace{1cm} (10)
where
\[ s_{mn} = \int_0^T \Psi'_n(t) \Psi_m(t) dt \]
and
\[ q_m(u_1(x), u_2(x), \ldots, u_N(x)) = \int_0^T q \left( \sum_{n=1}^N u_n(x) \Psi_n(t) \right) \Psi_m(t) dt. \quad (11) \]

Due to [6], each function \( u_m, m = 1, \ldots, N \), satisfies the Cauchy boundary conditions
\[
\begin{cases}
  u_m(x) = f_m(x) = \int_0^T f(x, t) \Psi_m(t) dt \\
  \partial_n u_m(x) = g_m(x) = \int_0^T g(x, t) \Psi_m(t) dt
\end{cases}
\quad (12)
\]
for all \( x \in \partial \Omega, m = 1, \ldots, N \). Here, \( f(x, t) \) and \( g(x, t) \) are the given data.

**Remark 2** Problem [4] becomes the problem of finding all functions \( u_m(x), x \in \Omega, m = 1, \ldots, N \), satisfying (10) and the Cauchy boundary conditions (12). In fact, if all of those functions are known, we can compute the function \( u(x, t), x \in \Omega, t \in [0, T] \) via (9). Then, the initial condition \( p(x) \) is given by the function \( u(x, 0) \).

**Remark 3** From now on, we consider the values of \( f_m(x) \) and \( g_m(x) \) on \( \partial \Omega, m = 1, \ldots, N \), as the “indirect data”, see (12). Denote by \( f^\delta_m(x) \) and \( g^\delta_m(x) \) the noiseless data. In numerical study, we set the noisy data as
\[
\begin{align*}
  f^\delta_m &= f^*_m(1 + \delta(-1 + 2\text{rand})) \\
  g^\delta_m &= g^*_m(1 + \delta(-1 + 2\text{rand}))
\end{align*}
\]
on \( \partial \Omega, 1 \leq m \leq N \) where \( \delta > 0 \) is the noise level and \( \text{rand} \) is the function taking uniformly distributed random numbers in the range [0, 1]. In our numerical study, \( \delta = 20\% \).

### 3. A Carleman estimate

In this section, we establish a Carleman estimate. This estimate and its corollary play a crucial role in the proof of our main result, Theorem [3].

**Theorem 1 (Carleman estimate)** Let \( x_0 \) be a point in \( \mathbb{R}^d \setminus \Omega \) such that \( r(x) = |x - x_0| > 1 \) for all \( x \in \Omega \). Let \( b > \max_{x \in \Omega} r(x) \) be a fixed constant. There exist positive constants \( \beta_0 \) depending only on \( b, x_0, \Omega \) and \( d \) such that for all function \( v \in C^2(\overline{\Omega}) \) satisfying
\[
  v(x) = \partial_n v(x) = 0 \quad \text{for all } x \in \partial \Omega,
\]
the following estimate holds true
\[
\begin{align*}
  &\int_{\Omega} e^{2 \lambda b - \beta r^2(x)} |\Delta v(x)|^2 dx \\
  &\geq \frac{C}{\lambda \beta^4 b^{\beta}} \sum_{i,j=1}^d \int_{\Omega} e^{2 \lambda b - \beta r^2(x)} v(x) \partial^2 x_i x_j v(x) dx \\
  &+ C \lambda^3 b^{3 - \beta} \int_{\Omega} r^2(x) e^{2 \lambda b - \beta r^2(x)} |v(x)|^2 dx + C \lambda^{1/2} b^{-\beta} \int_{\Omega} e^{2 \lambda b - \beta r^2(x)} |\nabla v(x)|^2 dx.
\end{align*}
\]
for $\beta \geq \beta_0$ and $\lambda \geq \lambda_0$. Here, $\lambda_0$ is a positive number with $\lambda_0 b^{-\beta} \gg 1$ and $C$ is a constant depending only on $b$, $\Omega$, $d$ and $x_0$.

We split the proof of Theorem 1 into several lemmas.

**Lemma 1** Let $v$ be the function as in Theorem 1. There exists a positive constants $\beta_0$ depending only on $b$, $x_0$, $\Omega$ and $d$ such that

$$
\int_\Omega \frac{e^{2\lambda b^{-\beta r^\beta}(x)}}{4\lambda b^{-\beta r^\beta - 2}(x)} |\Delta v(x)|^2 \, dx \\
\geq C \lambda^2 b^{-2\beta} \int_\Omega r^{2\beta}(x) e^{2\lambda b^{-\beta r^\beta}(x)} |v(x)|^2 \, dx - C \int_\Omega e^{2\lambda b^{-\beta r^\beta}(x)} |\nabla v(x)|^2 \, dx
$$

(13)

for all $\beta \geq \beta_0$ and $\lambda \geq \lambda_0$. Here, $\lambda_0$ is a constant such that $\lambda_0 b^{-\beta} \gg 1$.

**Proof:** By changing variables, if necessary, we can assume that $x_0 = 0$. Define the function

$$
w(x) = e^{\lambda b^{-\beta r^\beta}(x)} v(x) \quad \text{or} \quad v(x) = e^{-\lambda b^{-\beta r^\beta}(x)} w(x)
$$

(14)

for all $x \in \Omega$. Since $v$ vanishes on $\partial \Omega$, so does $w$. On the other hand, by the product rule in differentiation, for all $x \in \Omega,$

$$
\nabla v(x) = e^{-\lambda b^{-\beta r^\beta}(x)} \nabla w(x) - \beta \lambda b^{-\beta r^\beta - 2}(x) e^{-\lambda b^{-\beta r^\beta}(x)} w(x) x
$$

(15)

It follows that

$$
e^{-\lambda b^{-\beta r^\beta}(x)} \nabla w(x) \cdot v = \nabla v(x) \cdot v + \beta \lambda b^{-\beta r^\beta - 2}(x) e^{-\lambda b^{-\beta r^\beta}(x)} w(x) x = 0.
$$

for all $x \in \partial \Omega$. We thus obtain $w(x) = \partial_x v(x) = 0$ for all $x \in \partial \Omega$. Hence, from now on, whenever we apply the integration by parts formula on $v$ and $w$, the integrals on $\partial \Omega$ vanishes. We next compute the Laplacian of $v$ in terms of $w$. For all $x \in \Omega,$

$$\Delta v(x) = e^{-\lambda b^{-\beta r^\beta}(x)} \Delta w(x) + 2 \nabla e^{-\lambda b^{-\beta r^\beta}(x)} \cdot \nabla w(x) + w(x) \Delta (e^{-\lambda b^{-\beta r^\beta}(x)})$$

$$= e^{-\lambda b^{-\beta r^\beta}(x)} \left[ \Delta w(x) - 2 \lambda \beta b^{-\beta r^\beta - 2}(x) \nabla w(x) \cdot x \\
+ e^{\lambda b^{-\beta r^\beta}(x)} \Delta (e^{-\lambda b^{-\beta r^\beta}(x)}) w(x) \right].
$$

Using the inequality $(a - b + c)^2 \geq -2ab - 2bc$, we have

$$|\Delta v(x)|^2 \geq -4\lambda \beta b^{-\beta r^\beta - 2}(x) e^{-2\lambda b^{-\beta r^\beta}(x)} \left[ \Delta w(x) \nabla w(x) \cdot x \\
+ e^{\lambda b^{-\beta r^\beta}(x)} \Delta (e^{-\lambda b^{-\beta r^\beta}(x)}) w(x) \nabla w(x) \cdot x \right]
$$

(16)

for all $x \in \Omega$. By a straight forward computation, for $x \in \Omega,$

$$\Delta (e^{-\lambda b^{-\beta r^\beta}(x)}) = -\lambda \beta b^{-\beta} e^{-\lambda b^{-\beta r^\beta}(x)} r^{-2}(x) [(\beta - 2 + d) - \lambda b^{-\beta r^\beta}(x)].
$$

Plugging this into (16) gives

$$|\Delta v(x)|^2 \geq -4\lambda \beta b^{-\beta r^\beta - 2}(x) e^{-2\lambda b^{-\beta r^\beta}(x)} \left[ \Delta w(x) \nabla w(x) \cdot x \\
- \lambda \beta b^{-\beta r^\beta - 2}(x) [(\beta - 2 + d) - \lambda b^{-\beta r^\beta}(x)] w(x) \nabla w(x) \cdot x \right]$$
for all \( x \in \Omega \). Hence,

\[
\int_{\Omega} \frac{e^{2\lambda b^{-\beta}r^\beta(x)}}{4\lambda b^{-\beta}r^{\beta-2}(x)} |\Delta v(x)|^2 \, dx \geq I_1 + I_2 + I_3
\]  

(17)

where

\[
I_1 = - \int_{\Omega} \Delta w(x) \nabla w(x) \cdot xd\mathbf{x},
\]

(18)

\[
I_2 = \lambda \beta b^{-\beta} (\beta - 2 + d) \int_{\Omega} r^{\beta-2}(x) w(x) \nabla w(x) \cdot xd\mathbf{x},
\]

(19)

\[
I_3 = - \lambda^2 b^{-2\beta} \beta^2 \int_{\Omega} r^{2\beta-2}(x) w(x) \nabla w(x) \cdot xd\mathbf{x}.
\]

(20)

Estimate \( I_1 \). Write \( x = (x_1, \ldots, x_d) \) and integrating \( I_1 \) by parts. It follows from (18) that \( I_1 \) is equal to

\[
\int_{\Omega} \nabla w(x) \cdot \nabla [\nabla w(x) \cdot x] \, d\mathbf{x} = \sum_{i,j=1}^{d} \int_{\Omega} \partial_{x_i} w(x) \partial_{x_i} x_j \partial_{x_j} w(x) \, d\mathbf{x}
\]

\[
= \sum_{i,j=1}^{d} \int_{\Omega} \partial_{x_i} w(x) [\partial_{x_j} w(x) \delta_{ij} + x_j \partial_{x_i} x_j w(x)] \, d\mathbf{x}
\]

\[
= \sum_{i=1}^{d} \int_{\Omega} |\partial_{x_i} w(x)|^2 \, d\mathbf{x} + \sum_{i,j=1}^{d} \int_{\Omega} x_j \partial_{x_i} w(x) \partial_{x_i} x_j w(x) \, d\mathbf{x}.
\]

Using the identity \( \phi(x) \partial_{x_j} \phi(x) = \frac{1}{2} \partial_{x_j} (\phi(x)^2) \) with \( \Phi(x) = \partial_{x_i} w(x) \) gives

\[
I_1 = \int_{\Omega} |\nabla w(x)|^2 \, d\mathbf{x} + \frac{1}{2} \sum_{i,j=1}^{d} \int_{\Omega} x_j \partial_{x_j} (\partial_{x_i} w(x))^2 \, d\mathbf{x}
\]

\[
= \int_{\Omega} |\nabla w(x)|^2 \, d\mathbf{x} - \frac{1}{2} \sum_{i,j=1}^{d} \int_{\Omega} (\partial_{x_i} w(x))^2 \partial_{x_j} x_j \, d\mathbf{x}.
\]

Hence,

\[
I_1 = \left(1 - \frac{d}{2}\right) \int_{\Omega} |\nabla w(x)|^2 \, d\mathbf{x}.
\]  

(21)

Estimate \( I_2 \). We apply the identity \( w \nabla w = \frac{1}{2} \nabla|w|^2 \) to get from (19)

\[
I_2 = \frac{\lambda \beta b^{-\beta} (\beta - 2 + d)}{2} \int_{\Omega} r^{\beta-2}(x) \nabla |w(x)|^2 \cdot xd\mathbf{x}
\]

\[
= - \frac{\lambda \beta b^{-\beta} (\beta - 2 + d)}{2} \int_{\Omega} |w(x)|^2 \text{div}(r^{\beta-2}(x)\mathbf{x}) \, d\mathbf{x}.
\]

Here, the integration by parts formula was used. We, therefore, obtain

\[
I_2 = - \frac{\lambda \beta b^{-\beta} (\beta - 2 + d)^2}{2} \int_{\Omega} |w(x)|^2 \, d\mathbf{x}.
\]  

(22)

Estimate \( I_3 \). Using integration by parts formula again, by (20),

\[
I_3 = - \frac{\lambda^2 b^{-2\beta} \beta^2}{2} \int_{\Omega} r^{2\beta-2}(x) \nabla |w(x)|^2 \cdot xd\mathbf{x}
\]

\[
= \frac{\lambda^2 b^{-2\beta} \beta^2}{2} \int_{\Omega} |w(x)|^2 \text{div}(r^{2\beta-2}(x)\mathbf{x}) \, d\mathbf{x}.
\]
Hence,
\[ I_3 = \frac{\lambda^2 \beta^2 (2\beta - 2 + d) b^{-2\beta}}{2} \int_\Omega |w(x)|^2 r^{2\beta - 2}(x) dx \]
\[ \geq C \lambda^2 \beta^3 b^{-2\beta} \int_\Omega r^{2\beta}(x) |w(x)|^2 dx. \] (23)

Combining (17), (21), (22) and (23) and using the fact that \( \lambda b^{-\beta} \gg 1 \) (which implies \( \lambda b^{-\beta} \gg 1 \)), we get
\[ \int_\Omega \frac{e^{2\lambda \beta - \beta r^2}(x) \Delta v(x)}{4\lambda \beta b^{-\beta r^2}(x)} dx \]
\[ \geq C \lambda^2 \beta^3 b^{-2\beta} \int_\Omega r^{2\beta}(x) e^{2\lambda \beta - \beta r^2}(x) |v(x)|^2 dx - C \int_\Omega |\nabla w(x)|^2 dx. \] (24)

Recall (14) that \( w = e^{\lambda \beta - \beta r^2} v \). We have for all \( x \in \Omega \),
\[ \nabla w(x) = e^{\lambda \beta - \beta r^2}(x) [\nabla v(x) + \lambda \beta \beta - \beta r^2(\lambda \beta \beta - \beta r^2 - 2)(\lambda \beta \beta - \beta r^2)(x)x]. \] (25)

It follows from (14), (24), (25), the triangle inequality and the fact \( \beta^3 \gg \beta^2 \) that
\[ \int_\Omega \frac{e^{2\lambda \beta - \beta r^2}(x) \Delta v(x)}{4\lambda \beta b^{-\beta r^2}(x)} dx \]
\[ \geq C \lambda^2 \beta^3 b^{-2\beta} \int_\Omega r^{2\beta}(x) e^{2\lambda \beta - \beta r^2}(x) |v(x)|^2 dx - C \int_\Omega e^{2\lambda \beta - \beta r^2}(x) |\nabla v(x)|^2 dx. \]

Recall that \( \rho = \max_{x \in \Omega} r(x) \). We have obtained the desired inequality (13). \( \square \)

**Lemma 2** Let \( v \) be the function satisfying all hypotheses of Theorem 1. There exist positive constants \( \beta_0 \) and \( \lambda_0 \) depending only on \( b, \ x_0, \ \Omega \) and \( d \) such that
\[ -\int_\Omega e^{2\lambda \beta - \beta r^2}(x) v(x) \Delta v(x) dx \]
\[ \geq C \int_\Omega e^{2\lambda \beta - \beta r^2}(x) |\nabla v(x)|^2 dx - C \lambda^2 \beta^2 b^{-2\beta} \int_\Omega e^{2\lambda \beta - \beta r^2}(x) r^{2\beta}(x) |v(x)|^2 dx \] (26)

for all \( \beta \geq \beta_0 \) and \( \lambda \geq \lambda_0 \).

**Proof:** By integrating by parts, we have
\[ -\int_\Omega e^{2\lambda \beta - \beta r^2}(x) v(x) \Delta v(x) dx = \int_\Omega \nabla v(x) \cdot \nabla (e^{2\lambda \beta - \beta r^2}(x) v(x)) dx \]
\[ = \int_\Omega e^{2\lambda \beta - \beta r^2}(x) |\nabla v(x)|^2 dx + \int_\Omega v(x) \nabla v(x) \cdot \nabla (e^{2\lambda \beta - \beta r^2}(x)) dx. \] (27)

The absolute value of second integral in the right hand side of (27) can be estimated as
\[ \left| \int_\Omega v(x) \nabla v(x) \cdot \nabla (e^{2\lambda \beta - \beta r^2}(x)) dx \right| \]
\[ \leq 2\lambda \beta b^{-\beta} \int_\Omega r^{\beta - 1}(x) e^{2\lambda \beta - \beta r^2}(x) |v(x)||\nabla v(x)| dx \]
\[ \leq C \lambda^2 \beta^2 b^{-2\beta} \int_\Omega e^{2\lambda \beta - \beta r^2}(x) r^{2\beta}(x) |v(x)|^2 dx \]
\[ + \frac{1}{2} \int_\Omega e^{2\lambda \beta - \beta r^2}(x) |\nabla v(x)|^2 dx. \] (28)
This, (27) and (28) imply
\[-\int_{\Omega} e^{2\lambda b^{-\beta} r(x)} v(x) \Delta v(x) \, dx \geq C \int_{\Omega} e^{2\lambda b^{-\beta} r(x)} |\nabla v(x)|^2 \, dx - C\lambda^2 \beta^2 b^{-2\beta} \int_{\Omega} e^{2\lambda b^{-\beta} r(x)} r^{2\beta}(x) |v(x)|^2 \, dx.\]

The lemma is proved. \(\square\)

**Lemma 3** Let \(v\) be the function satisfying all hypotheses of Theorem 1. There exist positive constants \(\beta_0\) depending only on \(b, x_0, \Omega\) and \(d\) such that
\[
\int_{\Omega} e^{2\lambda b^{-\beta} r(x)} |\Delta v(x)|^2 \, dx \geq C \lambda^3 \beta^4 b^{-3\beta} \int_{\Omega} r^{2\beta}(x)e^{2\lambda b^{-\beta} r(x)} |v(x)|^2 \, dx
+ C \lambda \beta^{1/2} b^{-\beta} \int_{\Omega} e^{2\lambda b^{-\beta} r(x)} |\nabla v(x)|^2 \, dx
\]
for all \(\beta \geq \beta_0\) and \(\lambda \geq \lambda_0\). Here \(\lambda_0\) is a constant satisfying \(\lambda_0 b^{-\beta} > 1\).

**Proof:** Multiplying \(\beta^{1/4}\) to (26) and then applying the inequality \(-ab \leq a^2/2 + b^2/2\), we have
\[
\int_{\Omega} \lambda \beta^{3/2} b^{-\beta} e^{2\lambda b^{-\beta} r(x)} r^{-2}(x) |v(x)|^2 \, dx + \int_{\Omega} \frac{e^{2\lambda b^{-\beta} r(x)}}{4 \lambda b^{-\beta} r^{-2}(x)} |\Delta v(x)|^2 \, dx
\geq C \beta^{1/2} \int_{\Omega} e^{2\lambda b^{-\beta} r(x)} |\nabla v(x)|^2 \, dx
- C \lambda^2 \beta^5 b^{-2\beta} \int_{\Omega} r^{2\beta}(x) e^{2\lambda b^{-\beta} r(x)} |v(x)|^2 \, dx.\]

Since \(r(x) > 1\), \(\beta^{3/2} r^{-2}(x) \ll r^{2\beta}(x)\), we have
\[
\int_{\Omega} \frac{e^{2\lambda b^{-\beta} r(x)}}{4 \lambda \beta b^{-\beta} r^{-2}(x)} |\Delta v(x)|^2 \, dx \geq C \beta^{1/2} \int_{\Omega} e^{2\lambda b^{-\beta} r(x)} |\nabla v(x)|^2 \, dx
- C \lambda^2 \beta^5 b^{-2\beta} \int_{\Omega} r^{2\beta}(x) e^{2\lambda b^{-\beta} r(x)} |v(x)|^2 \, dx.\]

Here, we have used the fact that \(\lambda b^{-\beta} \gg 1\). Adding (30) and (13) together, we obtain
\[
\int_{\Omega} e^{2\lambda b^{-\beta} r(x)} |\Delta v(x)|^2 \, dx \geq C \lambda^2 \beta^3 b^{-2\beta} \int_{\Omega} r^{2\beta}(x) e^{2\lambda b^{-\beta} r(x)} |v(x)|^2 \, dx
+ C \beta^{1/2} \int_{\Omega} e^{2\lambda b^{-\beta} r(x)} |\nabla v(x)|^2 \, dx,\]
which implies (29). \(\square\)

**Lemma 4** Let \(v\) be the function satisfying all hypotheses of Theorem 1. There exist positive constants \(\beta_0\) and \(\lambda_0\) depending only on \(b, x_0, \Omega\) and \(d\) such that
\[
\frac{1}{\lambda \beta^{3/4} b^{-\beta}} \int_{\Omega} e^{2\lambda b^{-\beta} r(x)} |\Delta v(x)|^2 \, dx \geq \frac{C}{\lambda \beta^{3/4} b^{-\beta}} \sum_{i,j=1}^{d} \int_{\Omega} e^{2\lambda b^{-\beta} r(x)} r^{2\beta}(x) |\partial_{x_i x_j} v(x)|^2 \, dx
- C \lambda \beta^{1/4} b^{-\beta} \int_{\Omega} e^{2\lambda b^{-\beta} r(x)} |\nabla v(x)|^2 \, dx\]
for all \(\beta \geq \beta_0\) and \(\lambda \geq \lambda_0\).
Proof: By the density arguments, we can assume that $v \in C^2(\Omega)$. Write $x = (x_1, \ldots, x_d)$. We have

$$\int_{\Omega} e^{2\lambda - \beta \rho(x)} |\Delta v(x)|^2 \, dx = \sum_{i,j=1}^{d} \int_{\Omega} e^{2\lambda - \beta \rho(x)} \partial_{x_i}^2 v(x) \partial_{x_j}^2 v(x) \, dx$$

$$= \sum_{i,j=1}^{d} \int_{\Omega} \partial_{x_i} \left[ e^{2\lambda - \beta \rho(x)} \partial_{x_i}^2 v(x) \partial_{x_j} v(x) \right] \, dx$$

$$- \sum_{i,j=1}^{d} \int_{\Omega} \partial_{x_j} v(x) \partial_{x_j} \left[ e^{2\lambda - \beta \rho(x)} \partial_{x_i}^2 v(x) \right] \, dx.$$

The first integral in the right hand side above vanishes due to the divergence theorem. Hence

$$\int_{\Omega} e^{2\lambda - \beta \rho(x)} |\Delta v(x)|^2 \, dx = - \sum_{i,j=1}^{d} \int_{\Omega} e^{2\lambda - \beta \rho(x)} \partial_{x_j} v(x) \partial_{x_i}^2 v(x) \, dx$$

$$- \sum_{i,j=1}^{d} \int_{\Omega} \partial_{x_j} \left( e^{2\lambda - \beta \rho(x)} \partial_{x_i} v(x) \right) \partial_{x_i}^2 v(x) \, dx. \quad (32)$$

The first term in the right hand side of (32) is rewritten as

$$- \sum_{i,j=1}^{d} \int_{\Omega} e^{2\lambda - \beta \rho(x)} \partial_{x_j} v(x) \partial_{x_i}^2 v(x) \, dx$$

$$= \sum_{i,j=1}^{d} \int_{\Omega} \partial_{x_i} \left( e^{2\lambda - \beta \rho(x)} \partial_{x_j} v(x) \right) \partial_{x_i}^2 v(x) \, dx$$

$$= \sum_{i,j=1}^{d} \int_{\Omega} e^{2\lambda - \beta \rho(x)} \partial_{x_i}^2 v(x) \, dx + \sum_{i,j=1}^{d} \int_{\Omega} \partial_{x_j} v(x) \partial_{x_i} \left( e^{2\lambda - \beta \rho(x)} \right) \partial_{x_i}^2 v(x) \, dx.$$

Combining this and (32), we have

$$\int_{\Omega} e^{2\lambda - \beta \rho(x)} |\Delta v(x)|^2 \, dx = \sum_{i,j=1}^{d} \int_{\Omega} e^{2\lambda - \beta \rho(x)} \partial_{x_i}^2 v(x) \, dx$$

$$+ \sum_{i,j=1}^{d} \int_{\Omega} \left[ \partial_{x_j} v(x) \partial_{x_i} \left( e^{2\lambda - \beta \rho(x)} \right) \partial_{x_i}^2 v(x) - \partial_{x_j} \left( e^{2\lambda - \beta \rho(x)} \right) \partial_{x_i} v(x) \partial_{x_i}^2 v(x) \right] \, dx.$$

Hence,

$$\int_{\Omega} e^{2\lambda - \beta \rho(x)} |\Delta v(x)|^2 \, dx \geq \sum_{i,j=1}^{d} \int_{\Omega} e^{2\lambda - \beta \rho(x)} \partial_{x_i}^2 v(x) \, dx$$

$$- 2 \sum_{i,j=1}^{d} \int_{\Omega} |\partial_{x_j} v(x)||\partial_{x_i} \left( e^{2\lambda - \beta \rho(x)} \right)| \partial_{x_i}^2 v(x) \, dx.$$

Note that for all $i = 1, \ldots, d$,

$$\partial_{x_i} \left( e^{2\lambda - \beta \rho(x)} \right) = 2\lambda - \beta \rho x_i e^{2\lambda - \beta \rho(x)} x_i$$

for all $x \in \Omega$. 


Recovering the initial conditions of nonlinear parabolic equations

Using the inequality \( ab \leq a^2/2 + b^2/2 \), we obtain (31).

We now prove Theorem 1.

Proof of Theorem 1:
Adding (29) and (31) together, we obtain

\[
(1 + \frac{1}{\lambda^2 \beta^2/4b - \beta}) \int_\Omega e^{2\lambda b - \beta r^2(x)} |\Delta v(x)|^2 dx \\
\geq \frac{C}{\lambda \beta^2/4b - \beta} - \sum_{i,j=1}^d \int_\Omega e^{2\lambda b - \beta r^2(x)} |\partial^2_{x_i x_j} v(x)|^2 dx \\
+ C \lambda^3 \beta^2 b - 3\beta - \int_\Omega e^{2\lambda b - \beta r^2(x)} |v(x)|^2 dx + C \lambda \beta^{1/2} - \int_\Omega e^{2\lambda b - \beta r^2(x)} |\nabla v(x)|^2 dx.
\]

Theorem 1 has been proved.

Corollary 1 Recall \( \beta_0 \) and \( \lambda_0 \) as in Theorem 1. Fix \( \beta = \beta_0 \) and let the constant \( C \) depend on \( x_0 \), \( \Omega \), \( d \) and \( \beta \). There exists a constant \( \lambda_0 \) depending only on \( x_0 \), \( \Omega \), \( d \) and \( \beta \) such that for all function \( v \in H^2(\Omega) \) with

\[
v(x) = \partial_x v(x) = 0 \quad \text{on} \quad \partial \Omega,
\]

we have

\[
\int_\Omega e^{2\lambda b - \beta r^2(x)} |\Delta v(x)|^2 dx \\
\geq C \lambda^{-1} - \sum_{i,j=1}^d \int_\Omega e^{2\lambda b - \beta r^2(x)} |\partial^2_{x_i x_j} v(x)|^2 dx + C \lambda^3 \int_\Omega e^{2\lambda b - \beta r^2(x)} |v(x)|^2 dx \\
+ C \lambda \int_\Omega e^{2\lambda b - \beta r^2(x)} |\nabla v(x)|^2 dx
\]

for all \( \lambda \geq \lambda_0 \).

Remark 4 Although there are many versions of the Carleman estimate available, those versions are either too complicated, not suitable for us to prove Theorem 2 and Theorem 3, or not work in computations. The main ideas of the proof follow from [27, 28, 29, 30, 31].

Remark 5 The presence of the second derivatives in the right hand side of (33) is a new feature of our Carleman estimate. This allows us to prove the existence and uniqueness of the minimizers of the cost functionals in Section 4.

4. An iterative procedure

In this section, we propose a procedure to compute \( u_1(x), \ldots, u_N(x) \). We first approximate (10)–(12) by solving the following over-determined problem

\[
\begin{cases}
    c(x) \sum_{n=1}^N s_{mn} u_n^{(0)}(x) = \Delta u_m^{(0)}(x) & x \in \Omega, \\
    u_m^{(0)}(x) = f_m(x) & x \in \partial \Omega, \\
    \partial_n u_m^{(0)}(x) = g_m(x) & x \in \partial \Omega
\end{cases}
\]
for a vector value function \((u_1^{(0)}, \ldots, u_N^{(0)})\). Then, assume by induction that we know \((u_1^{(k-1)}, \ldots, u_N^{(k-1)})\), \(k \geq 1\), we find \((u_1^{(k)}, \ldots, u_N^{(k)})\) by solving

\[
\begin{align*}
& \quad \begin{cases}
\sum_{n=1}^N s_{mn} u_n^{(k)}(x) = \Delta u_m^{(k)}(x) \\
\quad q_m[P(u_1^{(k-1)}(x)), \ldots, P(u_N^{(k-1)}(x))] \quad x \in \Omega, \\
\quad f_m(x) \quad x \in \partial \Omega, \\
\quad g_m(x) \quad x \in \partial \Omega
\end{cases} \\
& \quad \partial_\nu u_m^{(k)}(x) = g_m(x)
\end{align*}
\]

where \(q_m\) is defined in (11) for \(m = 1, 2, \ldots, N\). Here,

\[
P(s) = \begin{cases} M \sqrt{T} & s \in (M \sqrt{T}, \infty), \\
-\sqrt{T} & s \in [-\sqrt{T}, \sqrt{T}], \\
-\sqrt{T} & s \in (-\infty, -\sqrt{T}]
\end{cases}
\]

serves as a cut-off function. Where \(M > \|u^*\|_{L^\infty(\Omega \times [0,T])}\) is a fixed constant.

In practice since both Dirichlet and Neumann conditions imposed, problem (34) and problem (35) might have no solution. However, since these two problems are linear, we can use the linear least squares method to find the “best fit” solutions. It is remarkable mentioning that a Carleman weight function will be included in the linear least squares method. Define the set of admissible solution

\[
H = \{(u_m)_{m=1}^N \in H^2(\Omega)^N : u_m|_{\partial \Omega} = f_m \text{ and } \partial_\nu u_m|_{\partial \Omega} = g_m, 1 \leq m \leq N\}.
\]

Throughout the paper, we assume that the set \(H\) is nonempty. In the analysis, we will need the following subspace of \(H^2(\Omega)^N\)

\[
H_0 = \{(v_1, \ldots, v_N) \in H^2(\Omega) : v_m(x) = \partial_\nu v_m(x) = 0\}.
\]

Let \(x_0\) be a point in \(\mathbb{R}^d \setminus \Omega\) with \(\min\{r(x) : x \in \Omega\} > 1\) and \(b > \max\{r(x) : x \in \Omega\}\) where

\[
r(x) = |x - x_0| \quad \text{for all } x \in \mathbb{R}^d.
\]

We choose \(x_0\) such that \(\min\{r(x) : x \in \Omega\} > 1\). To find \(u^{(0)}\), we minimize the functional \(J^{(0)} : H \to \mathbb{R}\) with

\[
J^{(0)}(u_1, \ldots, u_N) = \sum_{m=1}^N \int_\Omega e^{2\lambda b - \beta r(x)} \left| \Delta u_m - c(x) \sum_{n=1}^N s_{mn} u_n \right|^2 dx
\]

where \(\lambda\) and \(\beta\) are the numbers as in Corollary 1. The obtained minimizer \((u_m^{(0)})_{m=1}^N \in H\) is called the regularized solution to (34). Next, assume, by induction, that we know \((u_m^{(k-1)})_{m=1}^N, k \geq 1\), we set \((u_m^{(k)})_{m=1}^N\) as the minimizer of \(J^{(k)} : H \to \mathbb{R}\) defined as

\[
J^{(k)}(u_1, \ldots, u_N) = \sum_{m=1}^N \int_\Omega e^{2\lambda b - \beta r(x)} \left| \Delta u_m - c(x) \sum_{n=1}^N s_{mn} u_n + q_m[P(u_1^{(k-1)}), \ldots, P(u_N^{(k-1)})] \right|^2 dx.
\]

The following result guarantees the existence and uniqueness of the minimizer of (34) and the one of (35), \(k \geq 1\).

**Theorem 2** Assume that \(f_m\) and \(g_m\) are in \(L^2(\partial \Omega)\), \(m = 1, 2, \ldots, N\) and assume that \(H\) is nonempty. Then, each functional \(J^{(k)}, k \geq 0\), has a unique minimizer provided that both \(\lambda\) and \(\beta\) are sufficiently large.
Proof: We only prove Theorem 2 when \( k \geq 1 \). Since \( H \) is nonempty, we can find a vector valued function \((\varphi_m)_{m=1}^N \in H \). Define
\[
v_m(x) = u_m(x) - \varphi_m(x) \quad x \in \Omega, m = 1, \ldots, N.
\]
We minimize
\[
I^{(k)}(v_1, \ldots, v_N) = J^{(k)}(u_1 - \varphi_1, \ldots, u_N - \varphi_N)
\]
where \((v_m)_{m=1}^N \) varies in \( H_0 \), defined in (37). If \((v_m)_{m=1}^N \) minimizes \( I^{(k)} \), then by the variational principle,
\[
\sum_{m=1}^N \left \langle e^{2\lambda b - \beta r \beta(x)} \left( \Delta v_m - c(x) \sum_{n=1}^N s_{mn} v_n + \Delta \varphi_m - c(x) \sum_{n=1}^N s_{mn} \varphi_n 
+ q_m(P(u_1^{(k-1)}), \ldots, P(u_N^{(k-1)})) \right), \Delta h_m - c(x) \sum_{n=1}^N s_{mn} h_n \right \rangle_{L^2(\Omega)} = 0
\]
for all \((h_m)_{m=1}^N \) in \( H_0 \). The identity (40) is equivalent to
\[
\sum_{m=1}^N \left \langle e^{2\lambda b - \beta r \beta(x)} \Delta v_m - c(x) \sum_{n=1}^N s_{mn} v_n, \Delta h_m - c(x) \sum_{n=1}^N s_{mn} h_n \right \rangle_{L^2(\Omega)} = - \sum_{m=1}^N \left \langle e^{2\lambda b - \beta r \beta(x)} \left( \Delta \varphi_m - c(x) \sum_{n=1}^N s_{mn} \varphi_n + q_m(P(u_1^{(k-1)}), \ldots, P(u_N^{(k-1)})) \right), \Delta h_m - c(x) \sum_{n=1}^N s_{mn} h_n \right \rangle_{L^2(\Omega)}.
\]
The left hand side of (41) defines a bilinear form \( \langle \cdot, \cdot \rangle \) of a pair \(((v_m)_{m=1}^N, (h_m)_{m=1}^N)\) in \( H_0 \).

We claim that \( \langle \cdot, \cdot \rangle \) is coercive; that means,
\[
\langle (v_m)_{m=1}^N, (v_m)_{m=1}^N \rangle \geq C \| (v_m)_{m=1}^N \|^2_{H^2(\Omega)^N}
\]
for some constant \( C \). In fact, using the inequality \( (x - y)^2 \geq x^2/2 - y^2 \), we have
\[
\sum_{m=1}^N \int_\Omega e^{2\lambda b - \beta r \beta(x)} \left| \Delta v_m - c(x) \sum_{n=1}^N s_{mn} v_n \right|^2 dx \geq \sum_{m=1}^N \int_\Omega e^{2\lambda b - \beta r \beta(x)} \left| \Delta v_m \right|^2 dx
\]
\[
- \sum_{m=1}^N \int_\Omega e^{2\lambda b - \beta r \beta(x)} \left| c(x) \sum_{n=1}^N s_{mn} v_n \right|^2 dx.
\]
Applying the Carleman estimate (33) for the function \( v_m \) for each \( m \in \{1, \ldots, N \} \), we have
\[
\sum_{m=1}^N \int_\Omega e^{2\lambda b - \beta r \beta(x)} \left| \Delta v_m - c(x) \sum_{n=1}^N s_{mn} v_n \right|^2 dx
\]
\[
\geq C \lambda \left[ \sum_{i,j=1}^d \left| \partial_{x_i x_j}^2 v_m \right|^2 + C \lambda \left| \nabla v_m \right|^2 + C \lambda^2 \left| v_m \right|^2 \right] dx
\]
\[
- \sum_{m=1}^N \int_\Omega e^{2\lambda b - \beta r \beta(x)} \left| c(x) \sum_{n=1}^N s_{mn} v_n \right|^2 dx.
\]
Recovering the initial conditions of nonlinear parabolic equations

Since \( c(x) \) and \( s_{mn} \) are finite, we can choose \( \lambda \) sufficiently large such that

\[
\sum_{m=1}^{N} \int_{\Omega} e^{2\lambda b_{-\beta}r_{\beta}(x)} \left| \Delta v_{m} - c(x) \sum_{n=1}^{N} s_{mn}v_{n} \right|^2 dx \\
\geq C \max_{x \in \Omega} \left\{ e^{2\lambda b_{-\beta}r_{\beta}(x)} \right\} \lambda^{-1} \sum_{m=1}^{N} \| (v_m)_t \|_{H^2(\Omega)}^2.
\]

Applying the Lax-Milgram theorem, we can find a unique vector valued function \((v_m)_m^N\) satisfying (41). The vector valued function \((u_m)_m^N\) can be found via (39).

**Theorem 3** Assume that problem (10)–(12) has a unique solution \((u_m)_m^N\) satisfying (41). Then, there is a constant \( \lambda \) depending only on \( \Omega, T, d \) and \( N \) such that

\[
\sum_{m=1}^{N} \left\| e^{\lambda b_{-\beta}r_{\beta}(x)} (u_m^{(k)} - u_m^*) \right\|_{L^2(\Omega)}^2 \leq \left[ \frac{C}{\lambda^2} \right] k^{1-N} \sum_{m=1}^{N} \left\| e^{\lambda b_{-\beta}r_{\beta}(x)} (u_m^{(1)} - u_m^*) \right\|_{L^2(\Omega)}^2
\]

for \( k = 1, 2, \ldots \) where \( C \) is a constant depending only on \( \Omega, T, M, d, N \) and \( \| q \|_{C^3(\overline{\Omega})} \).

**Proof of Theorem 3** In the proof, \( C \) is a generous constant that might change from estimate to estimate.\

**Step 1.** Establish a priori bound. Recall \( H_0 \) as in (37). Since \((u_1^{(k)}, \ldots, u_N^{(k)})\) is the minimizer of \( J^{(k)} \), by the variational principle, for all \( h \in H_0 \)

\[
\sum_{m=1}^{N} \left\langle e^{\lambda b_{-\beta}r_{\beta}(x)} \left[ \Delta u^{(k)} - c(x) \sum_{n=1}^{N} s_{mn}u^{(k)} + q_m(P(u_1^{(k-1)}), \ldots, P(u_N^{(k-1)})) \right], e^{\lambda b_{-\beta}r_{\beta}(x)} \left[ \Delta h_m - c(x) \sum_{n=1}^{N} s_{mn}h_m \right] \right\rangle_{L^2(\Omega)} = 0. \tag{43}
\]

On the other hand, since \((u_1^*, \ldots, u_N^*)\) solves (10)–(12),

\[
\sum_{m=1}^{N} \left\langle e^{\lambda b_{-\beta}r_{\beta}(x)} \left[ \Delta u^* - c(x) \sum_{n=1}^{N} s_{mn}u^* + q_m(u_1^*, \ldots, u_N^*) \right], e^{\lambda b_{-\beta}r_{\beta}(x)} \left[ \Delta h_m - c(x) \sum_{n=1}^{N} s_{mn}h_m \right] \right\rangle_{L^2(\Omega)} = 0. \tag{44}
\]

It follows from (43) and (44) that

\[
\sum_{m=1}^{N} \left\langle e^{\lambda b_{-\beta}r_{\beta}(x)} \left[ \Delta (u^{(k)} - u^*) - c(x) \sum_{n=1}^{N} s_{mn}(u^{(k)} - u^*) \\
+ q_m(P(u_1^{(k-1)}), \ldots, P(u_N^{(k-1)})) - q_m(u_1^*, \ldots, u_N^*) \right], e^{\lambda b_{-\beta}r_{\beta}(x)} \left[ \Delta h_m - c(x) \sum_{n=1}^{N} s_{mn}h_m \right] \right\rangle_{L^2(\Omega)} = 0. \tag{45}
\]

Using the test function \( h_m = u_m^{(k)} - u_m^*, m = 1, \ldots, N \) in (45) and using Hölder’s inequality, we have

\[
\sum_{m=1}^{N} \left\| e^{\lambda b_{-\beta}r_{\beta}(x)} \left[ \Delta (u^{(k)} - u^*) - c(x) \sum_{n=1}^{N} s_{mn}(u^{(k)} - u^*) \right] \right\|_{L^2(\Omega)}^2
\]
\[ \leq \sum_{m=1}^{N} \left\| e^{\lambda b^{-\beta}r^\beta}(x) \left[ q_m(P(u_1^{(k-1)}), \ldots, P(u_N^{(k-1)})) - q_m(u_1^*, \ldots, u_N^*) \right] \right\|_{L^2(\Omega)} \]

\[ \times \left\| e^{\lambda b^{-\beta}r^\beta}(x) \left[ \Delta (u_m^{(k)} - u_m^*) - c(x) \sum_{n=1}^{N} s_{mn}(u_m^{(k)} - u_m^*) \right] \right\|_{L^2(\Omega)}. \] (46)

Using the inequality \( \sum_{m=1}^{N} a_m b_m \leq (\sum_{m=1}^{N} a_m^2)^{1/2}(\sum_{m=1}^{N} b_m^2)^{1/2} \) for the right hand side of (46) and simplifying the resulting, we get

\[ \sum_{m=1}^{N} \left\| e^{\lambda b^{-\beta}r^\beta}(x) \left[ \Delta (u_m^{(k)} - u_m^*) - c(x) \sum_{n=1}^{N} s_{mn}(u_m^{(k)} - u_m^*) \right] \right\|_{L^2(\Omega)}^2 \]

\[ \leq \sum_{m=1}^{N} \left\| e^{\lambda b^{-\beta}r^\beta}(x) \left[ q_m(P(u_1^{(k-1)}), \ldots, P(u_N^{(k-1)})) - q_m(u_1^*, \ldots, u_N^*) \right] \right\|_{L^2(\Omega)}^2. \] (47)

**Step 2.** Estimate the right hand side of (47). Since \( \|u^*(x, t)\|_{L^\infty} \leq M \), we have

\[ |u_m^*(x)| = \left| \int_0^T u^*(x, t)\Psi_m(t) dt \right| \leq \|u^*(x, t)\|_{L^2(0,T)} \|\Psi_m(t)\|_{L^2(0,T)} \leq M \sqrt{T} \]

for \( m = 1, \ldots, N \). Therefore,

\[ \left| q_m(P(u_1^{(k-1)}), \ldots, P(u_N^{(k-1)})) - q_m(u_1^*, \ldots, u_N^*) \right| \leq A_m \sum_{n=1}^{N} |u_n^{(k-1)} - u_n^*| \]

where

\[ A_m = \max \left\{ \|\nabla q_m(s_1, \ldots, s_N)\| : |s_i| \leq M \sqrt{T}, i = 1, \ldots, N \right\} \quad m = 1, \ldots, N. \]

Set \( A = \sum_{m=1}^{N} A_m \). The right hand side of (47) is bounded from above by

\[ \sum_{m=1}^{N} \left\| e^{\lambda b^{-\beta}r^\beta}(x) \left[ q_m(P(u_1^{(k-1)}), \ldots, P(u_N^{(k-1)})) - q_m(u_1^*, \ldots, u_N^*) \right] \right\|_{L^2(\Omega)}^2 \]

\[ \leq A \sum_{m=1}^{N} \left\| e^{\lambda b^{-\beta}r^\beta}(x) |P(u_m^{(k-1)}) - u_m^*| \right\|_{L^2(\Omega)}^2 \leq A \sum_{m=1}^{N} \left\| e^{\lambda b^{-\beta}r^\beta}(x) |u_m^{(k-1)} - u_m^*| \right\|_{L^2(\Omega)}^2. \] (48)

Combining (47) and (48) gives

\[ \sum_{m=1}^{N} \left\| e^{\lambda b^{-\beta}r^\beta}(x) \left[ \Delta (u_m^{(k)} - u_m^*) - c(x) \sum_{n=1}^{N} s_{mn}(u_m^{(k)} - u_m^*) \right] \right\|_{L^2(\Omega)}^2 \]

\[ \leq A \sum_{m=1}^{N} \left\| e^{\lambda b^{-\beta}r^\beta}(x) |u_m^{(k-1)} - u_m^*| \right\|_{L^2(\Omega)}^2. \] (49)

**Step 3.** Estimate the left hand side of (47). Using the inequality \((a - b)^2 \geq a^2/2 - 2b^2\) we have

\[ \sum_{m=1}^{N} \left\| e^{\lambda b^{-\beta}r^\beta}(x) \left[ \Delta (u_m^{(k)} - u_m^*) - c(x) \sum_{n=1}^{N} s_{mn}(u_m^{(k)} - u_m^*) \right] \right\|_{L^2(\Omega)}^2 \]

\[ \geq \sum_{m=1}^{N} \frac{1}{2} \left\| e^{\lambda b^{-\beta}r^\beta}(x) \Delta (u_m^{(k)} - u_m^*) \right\|_{L^2(\Omega)}^2 \]
In our numerical tests, the function $c(x) = 1 + 1/30 \left[ 3(1 - 3x)^2 e^{-9x^2 - (3y+1)^2} -10(3x/5 - 27x^3 - 243y^5)e^{-9x^2 - 9y^2} - 1/3e^{-(3x+1)^2 - 9y^2} \right]$

Applying Carleman estimate in Corollary 1 for the function $u^{(k)}_m - u^*$, $m = 1, \ldots, N$, we estimate

$$\sum_{m=1}^{N} \frac{1}{2} \left\| e^{\lambda - r^2} (u^{(k)}_m - u^*) \right\|_{L^2(\Omega)}^2 \geq C\lambda^2 \sum_{m=1}^{N} \left\| e^{\lambda - r^2} (u^{(k)}_m - u^*) \right\|_{L^2(\Omega)}^2,$$

Fix $\lambda \geq \lambda_0$ where $\lambda_0$ is as in Corollary 1. It follows from (50) and (51) that

$$\sum_{m=1}^{N} \frac{1}{2} \left\| e^{\lambda - r^2} (u^{(k)}_m - u^*) \right\|_{L^2(\Omega)}^2 - 2 \sum_{m=1}^{N} \left\| e^{\lambda - r^2} \sum_{n=1}^{N} s_{mn} (u^{(n)}_m - u^*_n) \right\|_{L^2(\Omega)}^2
\geq C\lambda^3 \sum_{m=1}^{N} \left\| e^{\lambda - r^2} (u^{(k)}_m - u^*) \right\|_{L^2(\Omega)}^2.$$

Combining (47), (48) and (52) gives

$$\sum_{m=1}^{N} \left\| e^{\lambda - r^2} (u^{(k)}_m - u^*) \right\|_{L^2(\Omega)}^2 \leq \frac{A}{C\lambda^3} \sum_{m=1}^{N} \left\| e^{\lambda - r^2} (u^{(k-1)}_m - u^*_m) \right\|_{L^2(\Omega)}^2.$$

By induction, we have

$$\sum_{m=1}^{N} \left\| e^{\lambda - r^2} (u^{(k)}_m - u^*) \right\|_{L^2(\Omega)}^2 \leq \left[ \frac{A}{C\lambda^3} \right]^{k-1} \sum_{m=1}^{N} \left\| e^{\lambda - r^2} (u^{(1)}_m - u^*_m) \right\|_{L^2(\Omega)}^2.$$

Replacing $A/C$ by the generous constant $C$, we have proved Theorem 3. \hfill \Box

**Remark 6** The technique of using the Carleman estimate to prove Theorem 3 is similar to the one in [32] in which a coefficient inverse problem for hyperbolic equations was considered. We also find that this technique is applicable to solve an inverse source problem for nonlinear parabolic equations [17] from the boundary and an additional internal measurements.

5. Numerical implementation

We only consider the case $d = 2$. We solve the forward problem of Problem 1 as follows.

Let $R_1 > R > 0$ be two positive numbers. Define the domains

$$\Omega_1 = (-R_1, R_1)^2 \quad \text{and} \quad \Omega = (-R, R)^2.$$

We approximate (1) defined on $\mathbb{R}^d \times (0, T)$ by the following problem defined on $\Omega_1 \times (0, T)$

$$\begin{cases}
  c(x) \Delta u(x, t) = q(u(x, t)) & x \in \Omega_1, t \in (0, T), \\
  u(x, 0) = p(x) & x \in \Omega_1, \\
  u(x, t) = 0 & x \in \partial \Omega_1, t \in [0, T].
\end{cases}
$$

In our numerical tests, the function $c$ is given by

$$c(x, y) = 1 + 1/30 \left[ 3(1 - 3x)^2 e^{-9x^2 - (3y+1)^2} -10(3x/5 - 27x^3 - 243y^5)e^{-9x^2 - 9y^2} - 1/3e^{-(3x+1)^2 - 9y^2} \right]$$
Recovering the initial conditions of nonlinear parabolic equations

for $\mathbf{x} = (x, y) \in \Omega$. The range of $c$ is $[0.8, 1.25]$, which is not a perturbation of the constant function 1.

Regarding to the forward problem, we solve \((53)\) by the finite difference method using the explicit formula. Then the data $f(\mathbf{x}, t) = u(\mathbf{x}, t)$ and $g(\mathbf{x}, t) = \partial_n u(\mathbf{x}, t)$ on $\partial \Omega \times [0, T]$ can be extracted easily. We next present the implementation for the inverse problem.

Fix a positive integer $N_x$. On $\Omega = [-R, R]^2$, we arrange an $N_x \times N_x$ uniform grid

$$
\mathcal{G} = \{(x_i, y_j) : x_i = -R + (i - 1)h, y_j = -R + (j - 1)h, 1 \leq i, j \leq N_x\}
$$

where $h = 2R/(N_x - 1)$ is the step size. In our computations, we set $R_1 = 6$, $R = 1$, $T = 1.5$ and $N_x = 80$. To solve Problem 1, we need to compute the discrete values of the function $u$ on the grid $\mathcal{G}$.

![Figures](image1.png)

**Figure 1:** The comparison of $f(x, R, t)$ and its partial Fourier sum $\sum_{n=1}^{N} f_n(x, R, t)$ on $\{(x, y = R) \in \partial \Omega\}$. The first row displays the graphs of the absolute differences of $f(x, R, t)$ and $\sum_{n=1}^{N} f_n(x, R) \Psi_n(t)$. The horizontal axis indicates $x$ and the vertical axis indicates $t$. It is evident that the bigger $N$, the smaller difference is. The second row shows the true data $f(x, y = R, T)$ (solid line) and its approximation $\sum_{n=1}^{N} f_n(x, y = R) \Psi_n(T)$ (dash–dot line). We observe that when $N = 35$, the two curves coincide.

The first step in our method is to find an appropriate cut off number $N$. We do so as follows. Take the data on $\{(x, y = R) \in \partial \Omega\}$, which is the top part of $\partial \Omega$, $f(x, y = R, t) = u_{true}(x, y = R, t)$ in Test 1 in Section 5.2. Then, we compare the
function \( f(x, R, t) \) and the function \( \sum_{n=1}^{N} f_n(x, y = R)\Psi_n(t) \) where \( f_n(x, y = R) \) is computed by [12]. Choose \( N \) such that the function 
\[
e_N(x, t) = \left| f(x, y = R, t) - \sum_{n=1}^{N} f_n(x, y = R)\Psi_n(t) \right|
\]
is small enough. We use the same number \( N \) for all numerical tests. In this paper, \( N = 35 \), see Figure [1] for an illustration.

**Remark 7** In our computations, when the cut-off number \( N \) is 15 or 25, the quality of the numerical results are poor. When \( N = 35 \), we obtain good numerical results. Increasing \( N > 35 \) does not improve the computed quality.

**Remark 8** In this numerical section, we choose the Carleman weight function \( e^{\lambda b^{-\beta}|x-x_0|^\beta} \) when defining \( J^{(k)} \), \( k \geq 0 \), where \( \lambda = 40 \) and \( \beta = 10 \). The point \( x_0 \) is \((0,1.5)\) and \( b = 5 \). This and the condition \( \lambda b^{-\beta} \) large conflict. However, in practice, the Carleman weight function with these values of \( \lambda \) and \( \beta \) already help provide good numerical solutions to Problem [7]. We numerically observe that the weight function blow-up when \( \lambda b^{-\beta} \gg 1 \), causing some unnecessary numerical difficulties.

### 5.1. Computing the vector valued function \((u_m)^N_{m=1}\)

Recall that \((u_m^{(0)}(x, y))^{N}_{m=1}\) minimizes \( J^{(0)} \) on \( H \). Similarly to the argument in the first step of the proof of Theorem [3] for all \( h \in H_0 \), see the definition of \( H_0 \) in (37), by the variational principle, we have

\[
\sum_{m=1}^{N} \left\langle e^{\lambda b^{-\beta}|x-x_0|^\beta} \left[ \Delta u_m^{(0)} - c(x) \sum_{n=1}^{N} s_{mn} u_m^{(0)} \right], e^{\lambda b^{-\beta}|x-x_0|^\beta} \left[ \Delta h_m - c(x) \sum_{n=1}^{N} s_{mn} h_m \right] \right\rangle_{L^2(\Omega)} = 0. \tag{54}
\]

For any \( u \in H \), we next associate the values of \( u_m \) \( \{u_m(x_i, y_j) : 1 \leq m \leq N, 1 \leq i, j \leq N_x\} \) with a \( N_x^2N \) dimentional vector \( u_i \) with

\[
u_i = u_m(x_i, y_j) \tag{55}\]

where

\[
i = (i - 1)N_xN + (j - 1)N + m \quad \text{for all } 1 \leq i, j \leq N_x, 1 \leq m \leq N. \tag{56}\]

The range of the index \( i \) is \( \{1, \ldots, N_x^2N\} \). The “line up” finite difference form of (54) is

\[
\langle (\mathcal{L} - S)u^{(0)}, (\mathcal{L} - S)h \rangle = 0 \tag{57}\]

where \( \Omega_x \), \( u^{(0)} \) and \( h \) are the line up versions of \( (W_\xi^N)_{m=1}^N \), \( (u_m^{(0)})_{m=1}^N \) and \( (h_m)_{m=1}^N \) respectively. Here, \( \langle \cdot, \cdot \rangle \) is the classical Euclidian inner product. In (57)

(i) the \( N_x^2N \times N_x^2N \) matrix \( \mathcal{L} \) is defined as

(a) \( (\mathcal{L})_{ii} = -\frac{4e^{\lambda b^{-\beta}|x_i-x_j|^\beta}}{d_x^2} \) for \( i \) as in (56) for \( 2 \leq i, j \leq N_x - 1, 1 \leq m \leq N; \)
(b) \((\mathcal{L})_{ij} = \frac{e^{\lambda_{ij}}}{d_{ij}^2} \) for \( j = (i \pm 1 - 1)N_xN + (j - 1)N + m \) and \( j = (i - 1)N_xN + (j \pm 1 - 1)N + m \); for \( 2 \leq i, j \leq N_x - 1 \), \( 1 \leq m \leq N \);
(c) the other entries are 0.

(ii) the \( N_x^2 \times N_x^2 \) matrix \( \mathcal{S} \) is defined as \((\mathcal{S})_{ij} = e^{\lambda_{ij}}x_{ij}c(x_i, y_j)s_{mn} \) for \( i \) as in (56) and \( j = (i - 1)N_xN + (j \pm 1 - 1)N + m \), for \( 2 \leq i, j \leq N_x - 1 \), \( 1 \leq m, n \leq N \). The other entries are 0.

On the other hand, since \((u^{0}_{m})_{m=1}^N\) satisfies the boundary constraints (12), we have
\[
\mathcal{D}u^{(0)} = \mathbf{f} \quad \text{and} \quad \mathcal{N}u^{(0)} = \mathbf{g}
\]

where
\[
(58)
\]

(i) The \( N_x^2 \times N_x^2 \) matrix \( \mathcal{D} \) is defined as \( \mathcal{D}_{ii} = 1 \) for \( i \) as in (56), \( i \in \{1, N_x\} \), \( 1 \leq j \leq N_x \) or \( 2 \leq i \leq N_x - 1 \), \( j \in \{1, N_x\} \). The other entries are 0.

(ii) The \( N_x^2 \times N_x^2 \) matrix \( \mathcal{N} \) is defined as
\[
\mathcal{N}_{ii} = \frac{1}{a_{x_i}} \quad \text{for } i \text{ as in (56), } i \in \{1, N_x\}, \quad 1 \leq j \leq N_x \text{ or } 2 \leq i \leq N_x - 1, \quad j \in \{1, N_x\}, \quad 1 \leq m \leq N;
\]
\[
\mathcal{N}_{ii} = -\frac{1}{a_{x_i}} \quad \text{for } i \text{ as in (56) and } j = (i + 1 - 1)N_xN + (j - 1)N + m, \quad i = 1, \quad 1 \leq j \leq N_x, \quad 1 \leq m \leq N;
\]
\[
\mathcal{N}_{ii} = -\frac{1}{a_{x_i}} \quad \text{for } i \text{ as in (56) and } j = (i - 1)N_xN + (j - 1)N + m, \quad i = N_x, \quad 1 \leq j \leq N_x, \quad 1 \leq m \leq N;
\]
\[
\mathcal{N}_{ii} = -\frac{1}{a_{x_i}} \quad \text{for } i \text{ as in (56) and } j = (i - 1)N_xN + (j - 1)N + m, \quad 1 \leq i \leq N_x, \quad j = 1, \quad 1 \leq m \leq N;
\]
\[
\mathcal{N}_{ii} = -\frac{1}{a_{x_i}} \quad \text{for } i \text{ as in (56) and } j = (i - 1)N_xN + (j - 1)N + m, \quad 2 \leq i \leq N_x - 1, \quad j = N_x, \quad 1 \leq m \leq N;
\]
\[
\text{The other entries are 0.}
\]

(iii) The \( N_x^2 \) dimensional vector \( \mathbf{f} \) is defined as \( f_i = f_m(x_i, y_j) \) for \( i \) as in (56), \( i \in \{1, N_x\} \), \( 1 \leq j \leq N_x \), \( 1 \leq m \leq N \) or \( 2 \leq i \leq N_x - 1 \), \( j \in \{1, N_x\} \).

(iv) The \( N_x^2 \) dimensional vector \( \mathbf{g} \) is defined as \( g_i = g_m(x_i, y_j) \) for \( i \) as in (56), \( i \in \{1, N_x\} \), \( 1 \leq j \leq N_x \), \( 1 \leq m \leq N \) or \( 2 \leq i \leq N_x - 1 \), \( j \in \{1, N_x\} \).

Solving (57)–(58) by the least square method with the command “lsqlin” built in Matlab, we obtain the vector \( u^{(0)} \) and hence the initial solution \((u^{(0)}_m(x_i, y_j))_{m=1}^N\) for \( 1 \leq i, j \leq N_x \).

**Remark 9** In computation, defining the matrices above is ineffective due to their large size, \( N_x^2 \times N_x^2 \) where \( N_x = 80 \) and \( N = 35 \). We note that most of those matrices’ entries are 0. So, instead of defining dense matrices, we use the invention of sparse matrices. Moreover, using sparse matrices significantly reduces the computational time.

We next compute the vector valued function \((u^{(k)}_m)_{m=1}^N, k \geq 1, \) assuming by induction that \((u^{(k-1)}_m)_{m=1}^N\) is known. Applying a very similar argument when deriving (57)–(58), the vector \( u^{(k)} \) the line up version of \((u^{(k)}_m(x_i, y_j))_{m=1}^N\) with \( 1 \leq i, j \leq N_x \) satisfies the equations
\[
(\mathcal{L} - \mathcal{S})^T(\mathcal{L} - \mathcal{S})u^{(k)} = -(\mathcal{L} - \mathcal{S})^T q^{(k-1)}.
\]

(59)
We numerically observe that

\[ D u^{(k)} = f \quad \text{and} \quad N u^{(k)} = g \]  

(60)

where \( q^{(k-1)} \) is the line up version of \((q_m(u_1^{(k-1)}(x,y), \ldots, u_N^{(k-1)}(x,y)))_{m=1}^N \). To find \( u^{(k)} \), we solve (59)–(60) by the least square method with the command “lsqlin” of Matlab. The value of the function \((u_m(x_i, y_j))_{m=1}^N \) follows. We next find \( u(x,y,t) \) via (5). The desired solution to Problem 1 \( p(x,y) \) is set to be \( u(x,y,0) \).

**Remark 10** In theory, we need to apply the cut-off function \( P \), see (36). This is only for our convenience to prove Theorem 3. However, in computation, we can obtain good numerical results without applying the cut-off technique. This can be explained by setting \( M \) sufficiently large.

We summarize the procedure to find \( p \) in Algorithm 1.

**Algorithm 1** The procedure to solve Problem 1

1: Choose \( N = 35 \), see Remark 7. Find \( \{\Psi_n\}_{n=1}^N \) as in the beginning of Section 2.2.
2: Compute matrices \( \mathcal{L}, \mathcal{S}, \mathcal{D} \) and \( \mathcal{N} \). Find the line up versions \( f \) and \( g \) of the data \( f_m(x_i,y_j) \) and \( g_m(x_i,y_j) \) for \((x_i,y_j) \in \mathcal{G} \cap \partial \Omega, 1 \leq m \leq N \).
3: Solve (57)–(58) by the least square method. The solution is denoted by \( u^{(0)} \).
   Compute \( u_m^{(0)}(x_i,y_j), 1 \leq i,j \leq N_x, 1 \leq m \leq N \) using \( u_m^{(0)}(x_i,y_j) = (u^{(0)})_i \) with \( i \) as in (56).
4: Set the initial solution \( p^{(0)} = \sum_{n=1}^N u_n^{(0)}(x_i,y_j)\Psi_n(0) \).
5: for \( k = 1 \) to 5 do
   6: Find \( q^{(k-1)} \), the line up version of \( q(P(u_1^{(k-1)}(x_i,y_j)), \ldots, u_N^{(k-1)}(x_i,y_j))) \), \( 1 \leq i,j \leq N_x, 1 \leq m \leq N \) in the same manner of (55) and (56).
   7: Solve (59)–(60) by the least square method. The solution is denoted by \( u^{(k)} \).
   Compute \( u_m^{(k)}(x_i,y_j), 1 \leq i,j \leq N_x, 1 \leq m \leq N \) using \( u_m^{(k)}(x_i,y_j) = (u^{(k)})_i \) with \( i \) as in (56).
   8: Set the initial solution \( p^{(k)} = \sum_{n=1}^N u_n^{(k)}(x_i,y_j)\Psi_n(0) \).
   9: Define the recursive error at step \( k \) as \( \|p^{(k)} - p^{(k-1)}\|_{L^\infty(\Omega)} \).
10: end for

**Remark 11** We numerically observe that \( \|p^{(5)} - p^{(4)}\|_\infty \) is sufficiently small in all tests in Section 5.2, i.e., our iterative scheme converges fast. Iterating the loop in Algorithm 1 five (5) times is enough to obtain good numerical results. Therefore, we stop the iterative process when \( k = 5 \).

### 5.2. Numerical examples

In this section, we show four (4) numerical results.

**Test 1.** The true source function is given by

\[
p_{\text{true}} = \begin{cases} 
8 \quad x^2 + (y - 0.3)^2 < 0.45^2, \\
0 \quad \text{otherwise}.
\end{cases}
\]
The nonlinearity $q$ is given by

$$q(s) = s(1 - s) \quad s \in \mathbb{R}.$$ 

In this case, the parabolic equation in (1) is the Fisher equation. The true and computed source functions $p$ are displayed in Figure 2. It appears in the graph of this source function a big inclusion with contrast 8.

Figure 2: Test 1. The reconstruction of the source function. (a) The function $p_{\text{true}}$ (b) The initial solution $p^{(0)}$ obtained by Step 3 in Algorithm 1 (c) The function $p^{(5)}$ obtained by Step 8 in Algorithm 1 (d) The true (solid), the initial source function (dot) in (b) and computed source function (dash-dot) on the vertical line in (c). (e) The curve $\|p^{(k)} - p^{(k-1)}\|_{L^\infty(\Omega)}$, $k = 1, \ldots, 5$. The noise level of the data in this test is 20%.

Our method to find the initial solution works very well in this case when effectively detect that inclusion. One can see in Figure 2b that by solving the system (57)–(58), we obtain the initial solution that clearly indicate the position of the inclusion. The value of the reconstructed function inside the inclusion is somewhat acceptable and will improve after several iterations, see Figure 2d. The reconstructed function $p_{\text{comp}} = p^{(5)}$ is a good approximation of the true function $p_{\text{true}}$, see Figures 2c and 2d. It is evident from Figure 2e that our method converges fast. The reconstructed maximal value inside the inclusion is 7.202 (relative error 9.98%).

**Test 2.** We test the case of multiple inclusions, each of which has different value. The
true source function $p_{\text{true}}$ is given by

$$p_{\text{true}}(x, y) = \begin{cases} 
12 & (x - 0.5)^2 + (y - 0.5)^2 < 0.35^2, \\
10 & (x + 0.5)^2 + (y + 0.5)^2 < 0.35^2, \\
14 & (x - 0.5)^2 + (y + 0.5)^2 < 0.35^2, \\
9 & (x + 0.5)^2 + (y - 0.5)^2 < 0.35^2, \\
0 & \text{otherwise.}
\end{cases}$$

In this test, the nonlinearity $q$ is given by

$$q(s) = -s(1 - \sqrt{|s|}), \quad s \in \mathbb{R}.$$

The true and computed source functions $p$ are displayed in Figure 3.

![Figure 3: Test 2. The reconstruction of the source function. (a) The function $p_{\text{true}}$. (b) The initial solution $p^{(0)}$ obtained by Step 3 in Algorithm 1. (c) The function $p^{(5)}$ obtained by Step 8 in Algorithm 1. (d) The true (solid), the initial solution (dot) and computed source function (dash-dot) on the diagonal line in (c). (e) The curve $\|p^{(k)} - p^{(k-1)}\|_{L^\infty(\Omega)}$, $k = 1, \ldots, 5$. The noise level of the data in this test is 20%.](image)

In this test, we successfully recover all four inclusions. On the other hand, the value of $p$ in each inclusion is high, making the true solution far away from the constant background $p_0 = 0$. Hence, $p_0 = 0$ might not serve as the initial guess. Our method to find the initial solution in Step 3 in Algorithm 1 is somewhat effective, see Figure 3b. The computed images of the initial solution does not completely separate the inclusions. Both computed values and images of the inclusions improve with iterations. The computed source function $p_{\text{comp}} = p^{(5)}$ is acceptable, see Figure 3c. Figure 3d shows
that the constructed values in the inclusions are good. The procedure converges very fast, see Figure 3e.

The true maximal value of the upper left inclusion is 9 and the computed one is 8.992 (relative error 0.0%). The true maximal value of the upper right inclusion is 12 and the computed one is 13.4 (relative error 11.67%). The true maximal value of the lower left inclusion is 10 and the computed one is 10.13 (relative error 1.3%). The true maximal value of the lower right inclusion is 14 and the computed one is 14.86 (relative error 6.14%).

**Test 3.** The true source function is given by

\[ p_{\text{true}} = \begin{cases} 1 & 0.2^2 < x^2 + y^2 < 0.8^2, \\ 0 & \text{otherwise}. \end{cases} \]

The nonlinearity is given by

\[ q(s) = s^2 \quad s \in \mathbb{R}. \]

The support of the function \( p_{\text{true}} \) is ring-like. This test is interesting due to the presence of the void and the nonlinearity grows fast. The true and computed source functions \( p \) are displayed in Figure 4.

![Figure 4](image-url)

Figure 4: Test 3. The reconstruction of the source function. (a) The function \( p_{\text{true}} \) (b) The initial solution \( p^{(0)} \) obtained by Step 3 in Algorithm 1 (c) The function \( p^{(5)} \) obtained by Step 8 in Algorithm 1 (d) The true (solid), initial solution (dot) and computed source function (dash-dot) on horizontal line in (c). (e) The curve \( \| p^{(k)} - p^{(k-1)} \|_{L^\infty(\Omega)}, \)

\( k = 1, \ldots, 5 \). The noise level of the data in this test is 20%.
In this test, our method to find the initial solution in Step 3 in Algorithm 1 is somewhat acceptable. The void in the initial solution \( p^{(0)} \) cannot be seen very well, see Figure 4b. The contrast and the void are improved with iteration. The final reconstructed source function \( p^{(5)} \) is satisfactory, see Figures 4c and 4d. The computed maximal value inside the ring is 1.094 (relative error = 9.4%).

**Test 4.** In this test, we identify two high contrast “lines”. The true source function is given by

\[
p_{\text{true}} = \begin{cases} 
10 \max\{|x|/4, 4|y - 0.6| < 0.9\} \text{ and } |x| < 0.8, \\
8 \max\{|x|/4, 4|y + 0.6| < 0.9\} \text{ and } |x| < 0.8, \\
0 \text{ otherwise.}
\end{cases}
\]

The nonlinearity is given by

\[ q(s) = -s^2 \quad s \in \mathbb{R}. \]

The true and computed source functions \( p \) are displayed in Figure 5.

![Figure 5: Test 3. The reconstruction of the source function.](image)

It is evident that Algorithm 1 provides good computed source function. The initial solution by Step 3 in Algorithm 1 is quite good although there is a “negative” artifact between the two detected lines, see Figure 5b. This artifact is reduced significantly with
iteration. We observe that the shape and contrasts of two lines are reconstructed very well, see Figures 5c and 5d. Our method converges fast, see Figure 5e.

The true maximal value of the source function in the upper line is 10 and the computed one is 9.714 (relative error 2.8%). The true maximal value of the source function in the lower line is 8 and the computed one is 8.041 (relative error 0.51%).

6. Concluding remarks

In this paper, we analytically and numerically solve the problem of recovering the initial condition of nonlinear parabolic equations. The first step in our method is to derive a system of nonlinear elliptic PDEs whose solutions are the Fourier coefficients of the solution to the governing nonlinear parabolic equation. We propose an iterative scheme to solve the system above. Finding the initial solution for this iterative process is a part of our algorithm. The convergence of this iterative method was proved. We show several numerical results to confirm the theoretical part.

Acknowledgment: The authors sincerely appreciate Michael V. Klibanov for many fruitful discussions that strongly improve the mathematical results and the presentation of this paper. The work of the second author was supported by US Army Research Laboratory and US Army Research Office grant W911NF-19-1-0044.

References

[1] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Ural’tseva. Linear and quasilinear equations of Parabolic Type, volume 23. American Mathematical Society, Providence, RI, 1968.
[2] M. V. Klibanov. Estimates of initial conditions of parabolic equations and inequalities via lateral Cauchy data. Inverse Problems, 22:495–514, 2006.
[3] A. El Badia and T. Ha-Duong. On an inverse source problem for the heat equation. Application to a pollution detection problem. Journal of Inverse and Ill-posed Problems, 10:585–599, 2002.
[4] J. Li, M. Yamamoto, and J. Zou. Conditional stability and numerical reconstruction of initial temperature. Communications on Pure and Applied Analysis, 8:361–382, 2009.
[5] R. A. Fisher. The wave of advance of advantageous genes. Annals of Eugenics, 7(4):355–369, 1937.
[6] M. M. Lavrent’ev, V. G. Romanov, and Shishat-ski˘i. Ill-Posed Problems of Mathematical Physics and Analysis. Translations of Mathematical Monographs. AMS, Providence: RI, 1986.
[7] A. B. Bakushinskii, M. V. Klibanov, and N. A. Kosheev. Carleman weight functions for a globally convergent numerical method for ill-posed Cauchy problems for some quasilinear pdes. Nonlinear Anal. Real World Appl., 34:201–224, 2017.
[8] M. V. Klibanov. Carleman weight functions for solving ill-posed Cauchy problems for quasilinear PDEs. Inverse Problems, 31:125007, 2015.
[9] M. V. Klibanov. Convexification of restricted Dirichlet to Neumann map. J. Inverse and Ill-Posed Problems, 25(5):669–685, 2017.
[10] M. Boulakia, M. de Buhan, and E. Schwindt. Numerical reconstruction based on Carleman estimates of a source term in a reaction-diffusion equation. preprint, hal-02185889, 2019.
[11] Q. Li and L. H. Nguyen. Recovering the initial condition of parabolic equations from lateral Cauchy data via the quasi-reversibility method. Inverse Problems in Science and Engineering, DOI: 10.1080/17415977.2019.1643850, 2019.
Recovering the initial conditions of nonlinear parabolic equations

[12] L. Borcea, V. Druskin, A. V. Mamonov, and M. Zaslavsky. A model reduction approach to numerical inversion for a parabolic partial differential equation. *Inverse Problems*, 30:125011, 2014.

[13] K. Cao and D. Lesnic. Determination of space-dependent coefficients from temperature measurements using the conjugate gradient method. *Numer Methods Partial Differential Eq.*, 34:1370–1400, 2018.

[14] K. Cao and D. Lesnic. Simultaneous reconstruction of the perfusion coefficient and initial temperature from time-average integral temperature measurements. *Applied Mathematical Modelling*, 68:523–539, 2019.

[15] Y. L. Keung and J. Zou. Numerical identifications of parameters in parabolic systems. *Inverse Problems*, 14:83–100, 1998.

[16] L. H. Nguyen. A new algorithm to determine the creation or depletion term of parabolic equations from boundary measurements. *preprint, arXiv:1906.01931*, 2019.

[17] P. M. Nguyen and L. H. Nguyen. A numerical method for an inverse source problem for parabolic equations and its application to a coefficient inverse problem. *Journal of Inverse and Ill-posed Problems*, DOI: https://doi.org/10.1515/jiip-2019-0026, 2019.

[18] L. Yang, J-N. Yu, and Y-C. Deng. An inverse problem of identifying the coefficient of parabolic equation. *Applied Mathematical Modelling*, 32:1984–1995, 2008.

[19] H. Liu and G. Uhlmann. Determining both sound speed and internal source in thermo- and photo-acoustic tomography. *Inverse Problems*, 31:105005, 2015.

[20] V. Katsnelson and L. V. Nguyen. On the convergence of time reversal method for thermoacoustic tomography in elastic media. *Applied Mathematics Letters*, 77:79–86, 2018.

[21] M. Haltmeier and L. V. Nguyen. Analysis of iterative methods in photoacoustic tomography with variable sound speed. *SIAM J. Imaging Sci.*, 10:751–781, 2017.

[22] L. H. Nguyen, Q. Li, and M. V. Klibanov. A convergent numerical method for a multi-frequency inverse source problem in inhomogenous media. *Inverse Problems and Imaging*, 13:1067–1094, 2019.

[23] X. Wang, Y. Guo, D. Zhang, and H. Liu. Fourier method for recovering acoustic sources from multi-frequency far-field data. *Inverse Problems*, 33:035001, 2017.

[24] X. Wang, Y. Guo, J. Li, and H. Liu. Mathematical design of a novel input/instruction device using a moving acoustic emitter. *Inverse Problems*, 33:105009, 2017.

[25] J. Li, H. Liu, and H. Sun. On a gesture-computing technique using electromagnetic waves. *Inverse Probl. Imaging*, 12:677–696, 2018.

[26] D. Zhang, Y. Guo, J. Li, and H. Liu. Retrieval of acoustic sources from multi-frequency phaseless data. *Inverse Problems*, 34:094001, 2018.

[27] L. Beilina and Michael V. Klibanov. *Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems*. Springer, New York, 2012.

[28] A. L. Bukhgeim and Michael V. Klibanov. Uniqueness in the large of a class of multidimensional inverse problems. *Soviet Math. Doklady*, 17:244–247, 1981.

[29] M. V. Klibanov. Carleman estimates for global uniqueness, stability and numerical methods for coefficient inverse problems. *J. Inverse and Ill-Posed Problems*, 21:477–560, 2013.

[30] M. H. Protter. Unique continuation for elliptic equations. *Trans. Amer. Math. Soc.*, 95(1):81–91, 1960.

[31] H. M. Nguyen and L. H. Nguyen. Cloaking using complementary media for the Helmholtz equation and a three spheres inequality for second order elliptic equations. *Transaction of the American Mathematical Society*, 2:93–112, 2015.

[32] L. Baudouin, M. de Buhan, and S. Ervedoza. Convergent algorithm based on Carleman estimates for the recovery of a potential in the wave equation. *SIAM J. Numer. Anal.*, 55:1578–1613, 2017.