Boundary Switch on/off Control Approach to Simultaneous Identification of Diffusion Coefficient and Initial State for One-Dimensional Heat Equation

Zhi-Xue Zhao[a,b*], Mapundi K. Banda[b], Bao-Zhu Guo[c,d]

[a] *College of Mathematical Science, Tianjin Normal University, Tianjin 300387, China*

[b] *Department of Mathematics and Applied Mathematics, University of Pretoria, Pretoria 0002, South Africa*

[c] *School of Mathematics and Physics, North China Electric Power University, Beijing 102206, China*

[d] *School of Computer Science and Applied Mathematics, University of the Witwatersrand, Wits 2050, Johannesburg, South Africa*

Abstract

In this paper, we consider simultaneous reconstruction of diffusion coefficient and initial state for a one-dimensional heat equation through boundary control and measurement. The boundary measurement is proposed to make the system approximately observable, and both the coefficient and initial state are shown to be identifiable by this measurement under a boundary switch on/off control. By a Dirichlet series representation for the observation, we can transform the problem into an inverse process of reconstruction of the spectrum and coefficients for the Dirichlet series in terms of observation. This happens to be the reconstruction of spectral data for an exponential sequence with measurement error, and it enables us to develop an algorithm based on the matrix pencil method in signal analysis. A theoretical error analysis for the algorithm concerning the coefficient reconstruction is carried out for the proposed method. The numerical simulations are presented to verify the proposed algorithm.

Keywords: Boundary switch on/off control, identifiability, identification, matrix pencil method, error analysis.

AMS subject classifications: 35K05, 35R30, 65M32, 65N21.

1 Introduction

It is recognized that many industrial controls are temperature control. The inverse heat conduction problem (IHCP) is one of the important control problems in science and engineering. Such kinds of problems

*The corresponding author. Email: zxzhao@amss.ac.cn

usually arise in modeling and process control with heat propagation in thermophysics, chemical engineering, and many other industrial and engineering applications. In the last decades, there are various class of IHCPs being investigated ranging from recovery of boundary heat flux [27]; estimation of medium parameters such as thermal conductivity coefficient [10] and radiative coefficient [5, 30]; recovery of spatial distribution of heat sources [31]; and reconstruction of initial state distributions [30]. For many other aspects including numerical solutions of inverse problems for PDEs, we refer to the monograph [13]. Most of the existing works, however, are devoted to single parameter identification. The simultaneous reconstruction of more than one different coefficients has not been sufficiently investigated, for which, to the best of our knowledge, only a few studies are available (see, e.g., [4, 5, 27, 30] and the references therein). To cope with the ill-posed nature of inverse problems, optimization methods and regularization techniques together with many other numerical methods such as finite difference methods and finite element methods are generally applied in the literature. In addition to the numerical methods mentioned above, a number of classical identifiability results are based on the inverse spectral theory, see, for instance, [15, 18, 23, 24]. In [18], the unique determination of eigenvalues and coefficients under certain conditions for a parabolic equation is considered by the Gel'fand-Levitan theory. Some uniqueness results on simultaneous identification of coefficients and initial values for parabolic equations are presented in [15, 23, 24]. However, most of the identifiability results require that the initial value can not be orthogonal to any of the eigenvectors. This restrictive condition is actually unverifiable in practice because the initial value is also unknown. Some other uniqueness results on the determination of constant coefficients are discussed in [14, 16]. But no numerical identification algorithm is attempted in these theoretical papers.

In this paper, we are concerned with reconstruction of the diffusion coefficient and initial state for a one-dimensional heat conduction equation in a homogeneous bar of unit length, which is described by

$$\begin{cases} u_t(x,t) = \alpha u_{xx}(x,t),\ 0 < x < 1,\ t > 0, \\ \alpha u_x(0,t) = f(t),\ u_x(1,t) = 0,\ t \geq 0, \\ y(t) = u(0,t),\ t \geq 0, \\ u(x,0) = u_0(x),\ 0 \leq x \leq 1, \end{cases} \tag{1.1}$$

where x represents the position, t the time. $\alpha \geq \alpha_0 > 0$ is an unknown constant that represents the thermal diffusivity, and $u_0(x)$ is the unknown initial temperature distribution. Here we do not impose any restriction on the initial value other than boundedness. The function $f(t)$ is the Neumann boundary control (input) which represents the heat flux through the left end of the bar, and $y(t)$ is the boundary temperature measurement. Sometimes we write the solution of (1.1) as $u = u(x,t;f,u_0)$ to represent its dependence on $f(t)$ and $u_0(x)$.

The inverse problem that we consider in this paper can be described as follows:

Inverse Problem: *Consider the problem as posed in Equation* (1.1)*. Design the boundary control* $f \in L^2(0,T)$ *for some finite time interval* $t \in (0,T)$*, to reconstruct simultaneously the* α *and* $u_0(x)$ *from the boundary observation* $\{y(t),\ t \in (0,T)\}$*.*

Let us briefly explain the main idea of this paper, which is inspired by an idea of [9]. By the superposition principle of linear systems, the output $y(t)$ of system (1.1) can be separated into two parts

$u(0,t;0,u_0)$ and $u(0,t;f,0)$, where the former is determined by the initial state only and the latter by the control. The first part $u(0,t;0,u_0)$ admits a Dirichlet series representation, which means that it can be determined by its restriction on any finite interval. By choosing the control $f(t)$ appropriately, we can design an algorithm to estimate the unknown coefficients in the Dirichlet series. The part of the output that has been determined by the initial value, $u(0,t;0,u_0)$, can be canceled from the output $y(t)$ such that the coefficient identification of α is equivalently transformed into the case with zero initial state. After estimating the diffusion coefficient, the remaining problem is a single reconstruction of the initial state.

The rest of the paper is organized as follows. Section 2 is devoted to simultaneous identifiability of the coefficient and initial value based on the Dirichlet series theory. The identification algorithm based on the matrix pencil method is introduced in Section 3. In Section 4, an error analysis of the matrix pencil method to the infinite spectral estimation problem is obtained. A numerical simulation is presented in Section 5 to show the validity of the algorithm introduced in Section 3.

2 Identifiability

The identifiability is to show that the boundary control $f \in L^2(0,T)$ and boundary observation $\{y(t),\ t \in (0,T)\}$ are sufficient to determine α and $u_0(x)$ uniquely. To begin with, let $\mathcal{H} = L^2(0,1)$ with the usual inner product $\langle\cdot,\cdot\rangle$ and inner product induced norm $\|\cdot\|$. Define the operator $\mathcal{A}:\ D(\mathcal{A})(\subset \mathcal{H}) \mapsto \mathcal{H}$ by

$$\begin{cases} [\mathcal{A}\psi](x) = -\alpha\psi''(x), \\ D(\mathcal{A}) = \{\psi \in H^2(0,1) \,|\, \psi'(0) = \psi'(1) = 0\}. \end{cases} \tag{2.1}$$

It is easy to verify that $\mathcal{A}$ is positive semidefinite in $\mathcal{H}$, and the eigenvalues $\{\lambda_n\}$ of $\mathcal{A}$ are given by

$$\lambda_n = \alpha n^2\pi^2,\ n = 0,1,2,\ldots, \tag{2.2}$$

with the corresponding eigenfunctions $\{\phi_n(x)\}_{n=0}^{\infty}$ given by

$$\phi_0(x) = 1,\ \phi_n(x) = \sqrt{2}\cos n\pi x,\ n \in \mathbb{N}^*. \tag{2.3}$$

It is well known that $\{\phi_n(x)\}_{n\in\mathbb{N}}$ forms an orthonormal basis for $\mathcal{H}$. Set

$$G(t,x,y) = \sum_{n=1}^{\infty} e^{-\lambda_n t}\phi_n(x)\phi_n(y) + 1,\quad A_n(x) = \langle u_0,\phi_n\rangle\phi_n(x),\ n \in \mathbb{N}. \tag{2.4}$$

It is evident that $G(t,0,0) = 2\sum_{n=1}^{\infty} e^{-\lambda_n t} + 1$ is continuous in $t > 0$. Suppose that the boundary control satisfies $f \in L^2(0,T)$ for some $T > 0$. Standard analysis (see, e.g., [2, Chapter 3]) shows that the boundary observation $y(t)$ of system (1.1) takes the form

$$y(t) = \sum_{n=0}^{\infty} A_n(0)e^{-\lambda_n t} - \int_0^t G(t-s,0,0)f(s)\,ds,\ 0 < t < T. \tag{2.5}$$

It is well known that system (1.1) (with $f = 0$) is not exactly observable in $\mathcal{H}$ because the following observability inequality

$$\int_0^{\tau} u^2(0,t)dt \geq c_\tau\|u_0\|^2,\ \forall\ u_0 \in \mathcal{H},$$

is not valid for any constants $\tau,\ c_\tau > 0$. However, from the uniqueness of Dirichlet series expansion (see, for example, [10, Lemma 2.4], and also [20, Chapter 9] for more details about Dirichlet series), it is easy to verify, from (2.5), that system (1.1) is approximately observable (see, e.g., [26, Chapter 6]), which ensures that the initial state can be uniquely determined from the output.

Suppose that $T_2 > T_1 > 0$ are two arbitrary positive numbers, and the boundary control $f(t)$ is chosen to be zero in the time interval $[0, T_2]$. In this case, it is deduced from (2.5) that the boundary observation is

$$y(t) \triangleq u(0,t;0,u_0) = \sum_{n=0}^{\infty} C_n e^{-\lambda_n t},\ \forall\, t \in [T_1, T_2], \tag{2.6}$$

where $C_n = A_n(0),\ n \in \mathbb{N}$. Since $u_0(x)$ is unknown, it is not clear whether $C_n \neq 0$ for any $n \in \mathbb{N}$. Define the set $\mathbb{K} \subset \mathbb{N}$, which is unknown as well and satisfies

$$C_k \neq 0,\ k \in \mathbb{K};\ C_k = 0,\ k \notin \mathbb{K}. \tag{2.7}$$

Theorem 2.1 below establishes the identifiability of the set $\{(C_k, \lambda_k)\}_{k\in\mathbb{K}}$.

Theorem 2.1. *Let $0 < T_1 < T_2 < \infty$, $u_0 \in L^2(0,1)$. Then the set $\{(C_k, \lambda_k)\}_{k\in\mathbb{K}}$ in (2.6) can be uniquely determined by the observation data $\{y(t) \mid t \in [T_1, T_2]\}$.*

Proof. The proof is accomplished by two steps.

Step 1: $\{(C_k, \lambda_k)\}_{k\in\mathbb{K}}$ can be uniquely determined by infinite time observation $\{y(t)|\ t \in (0,\infty)\}$.

Actually, since

$$\sum_{n=0}^{\infty} |C_n|^2 \leq 2\|u_0\|^2_{L^2(0,1)} < +\infty,$$

it follows that $\sup_{n\geq 0} |C_n| < \infty$. Since $\lambda_n = \alpha n^2\pi^2$, the series (2.6) converges uniformly in t over $[t_0, +\infty)$ for any $t_0 > 0$. Apply the Laplace transform to (2.6) to obtain

$$\hat{y}(s) = \sum_{n=0}^{\infty} \frac{C_n}{s+\lambda_n} = \sum_{k\in\mathbb{K}} \frac{C_k}{s+\lambda_k}, \tag{2.8}$$

where $\hat{}$ denotes the Laplace transform. Since we are only interested in those $C_k \neq 0$, there is no zero/pole cancelations. In other words, $-\lambda_k$ is a pole of $\hat{y}(s)$ and C_k is the residue of $\hat{y}(s)$ at $-\lambda_k$ for any $k \in \mathbb{K}$. By the uniqueness of the Laplace transform, $\{(C_k, \lambda_k)\}_{k\in\mathbb{K}}$ is uniquely determined by $\{y(t)|\ t \in (0,\infty)\}$.

Step 2: $\{(C_k, \lambda_k)\}_{k\in\mathbb{K}}$ can be uniquely determined by finite-time observation $\{y(t)|\ t \in [T_1, T_2]\}$.

By *Step 1*, we only need to show that the observation $y(t)$ in (2.6) for all $t > 0$ can be uniquely determined by its restriction on $I_1 = [T_1, T_2]$, or in other words, $y(t) = 0$ for $t \in [T_1, T_2]$ in (2.6) implies that $y(t) = 0$ for all $t > 0$. But this is evident because $y(t)$ is an analytic function in $t > 0$. This completes the proof of the theorem. □

Theorem 2.2. *Let $0 < T_1 < T_2 < T_3 < \infty$, $u_0 \in L^2(0,1)$. The control function $f \in L^2(0, T_3)$ satisfies*

$$\begin{cases} f(t) = 0, & \text{for } t \in [0, T_2), \\ f(t) \neq 0, & \text{for almost all } t \in [T_2, T_3], \end{cases} \tag{2.9}$$

and the corresponding boundary observation is $\{y(t) = u(0,t;f,u_0) \mid t \in [T_1, T_3]\}$. Then the diffusion coefficient α and the initial state $u_0(x)$ in system (1.1) can be uniquely determined by $\{y(t) \mid t \in [T_1, T_3]\}$.

Proof. By (2.5), it follows that

$$\widetilde{y}(t) \triangleq \sum_{k\in\mathbb{K}} C_k e^{-\lambda_k(t+T_2)} - y(t+T_2) = \int_0^t G(t-s,0,0)\widetilde{f}(s)\,ds,\ t\in(0,T_3-T_2], \tag{2.10}$$

where

$$\widetilde{f}(t) = f(t+T_2),\ t\in(0,T_3-T_2],$$

and

$$G(t,0,0) = 2\sum_{n=1}^{\infty} e^{-\lambda_n t} + 1 \triangleq \sum_{n=0}^{\infty} G_n e^{-\lambda_n t},\ t\in(0,T_3-T_2], \tag{2.11}$$

with

$$G_0 = 1, \lambda_0 = 0;\ G_n = 2, n\in\mathbb{N}^*.$$

Since $\widetilde{f}\in L^2(0,T_3-T_2)$ is nonzero for almost all $t\in(0,T_3-T_2)$ and $G(t,0,0)$ is continuous in $t>0$, it follows from [25, Theorem 151] that $\{G(t,0,0)\mid t\in(0,T_3-T_2]\}$ can be uniquely determined by $\{\widetilde{y}(t)\mid t\in(0,T_3-T_2]\}$. By Theorem 2.1, $\{(C_k,\lambda_k)\}_{k\in\mathbb{K}}$ can be uniquely determined from $\{y(t)\mid t\in[T_1,T_2]\}$, which shows, from the first equality of (2.10), that $\{\widetilde{y}(t)\mid t\in(0,T_3-T_2]\}$ can be obtained from $\{y(t)\mid t\in[T_1,T_3]\}$. In addition, since all the coefficients $\{G_n\}_{n\in\mathbb{N}}$ in (2.11) are nonzero, by Theorem 2.1 again, $\{\lambda_n\}_{n\in\mathbb{N}^*}$ can be uniquely determined by $\{G(t,0,0)\mid t\in(0,T_3-T_2]\}$. The exponents $\{\lambda_n\}_{n\in\mathbb{N}^*}$ are therefore uniquely determined by $\{y(t)\mid t\in[T_1,T_3]\}$, and then the diffusion coefficient α can be obtained from (2.2).

We now turn to the identifiability of the initial value $u_0(x)$. To be specific, it follows from (2.5) that

$$y(t) + \int_0^t G(t-s,0,0)f(s)\,ds = \sum_{n=0}^{\infty} A_n(0)e^{-\lambda_n t},\ t\in[T_1,T_3]. \tag{2.12}$$

Since the diffusion coefficient α can be uniquely determined by $\{y(t)\mid t\in[T_1,T_3]\}$, the left-hand side of (2.12) can also be uniquely determined by $\{y(t)\mid t\in[T_1,T_3]\}$. On the other hand, the right-hand side of (2.12) is a Dirichlet series (see, e.g., [20, Chapter 9]), it follows from the uniqueness of Dirichlet series expansion (e.g., [10, Lemma 2.4]) that all the coefficients $\{A_n(0)\}_{n=0}^{\infty}$ can be uniquely determined by the left-hand side of (2.12) or the boundary observation $\{y(t)\mid t\in[T_1,T_3]\}$. Therefore, the initial value

$$u_0(x) = \sum_{n=0}^{\infty}\langle u_0,\phi_n\rangle\phi_n(x) = A_0(0) + \frac{1}{\sqrt{2}}\sum_{n=1}^{\infty} A_n(0)\phi_n(x),$$

can also be uniquely determined by $\{y(t)\mid t\in[T_1,T_3]\}$. This ends the proof of the theorem. □

Remark 2.1. Actually, there are many papers studying simultaneous identifiability of parameters and initial values for parabolic equations, see, for instance, [15, 23, 24]. However, most of the identifiability results require that the initial value should be a generating element (see [23]) with respect to the system operator $\mathcal{A}$, that is,

$$\langle u_0,\phi_n\rangle \neq 0,\ \text{for any } n\in\mathbb{N}. \tag{2.13}$$

But this condition is unverifiable because the initial value $u_0(x)$ is also unknown. Theorem 2.2 shows that this restrictive condition can be removed by designing a boundary switch on/off control (2.9), which

is somehow like the persistently exciting condition introduced in [17, 21] to stimulate the plant behavior. It should also be emphasized that the boundary switch on/off control signal (2.9) is crucial in Theorem 2.2, without which the identifiability may not be valid anymore. The Example 2.1 below indicates the necessity of the control signal (2.9).

Example 2.1. Let $f(t) \equiv 0$ in system (1.1). Consider the following two cases:
Case 1: The diffusion coefficient α and initial value $u_0(x)$ are

$$\alpha = 1,\ u_0(x) = \phi_4(x) = \sqrt{2}\cos 4\pi x, \tag{2.14}$$

respectively.
Case 2: The diffusion coefficient and initial value are

$$\widetilde{\alpha} = 4,\ \widetilde{u}_0(x) = \phi_1(x) = \sqrt{2}\cos \pi x, \tag{2.15}$$

respectively.

Simple calculations from (2.5) show that both cases produce the same boundary measured data

$$y(t) = u(0,t;0,u_0) = u(0,t;0,\widetilde{u}_0) = \sqrt{2}e^{-4\pi^2 t},\ t > 0. \tag{2.16}$$

Hence, we can not distinguish $\{\alpha, u_0(x)\}$ and $\{\widetilde{\alpha}, \widetilde{u}_0(x)\}$ from the observation data $\{y(t)\}$ alone when $f(t) = 0$, or in other words, the identifiability may not be valid without the boundary control $f(t)$.

The simplest practically implementable control that satisfies (2.9) is

$$f(t) = \begin{cases} 0, & t \in [0, T_2), \\ 1, & t \in [T_2, T_3], \end{cases} \tag{2.17}$$

which is used in the numerical identification algorithm in Section 3.2. To illustrate the identifiability analysis more clearly, we present a block diagram in Figure 1.

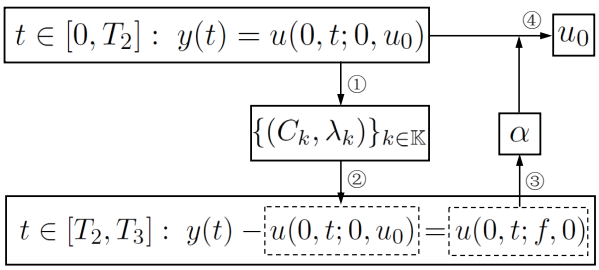

Figure 1: Block diagram of identifiability analysis

3 Numerical computation method

It is seen clearly from previous sections that the key point for identification is to recover the spectrum-coefficient data $\{(C_n, \lambda_n)\}_{n\in\mathbb{N}}$ from $\{y(t) \mid t \in [T_1, T_2]\}$ by the Dirichlet series representation (2.6). The difficulty is that there may exist infinitely many $C_n \neq 0$ in (2.6). In this section, we use the matrix pencil method to extract some of the $\{(C_n, \lambda_n)\}$ from the sum of the first M terms of the infinite series (2.6), and treat the remainder terms as a measurement error.

3.1 Finite dimensional approximation

Suppose $M \in \mathbb{N}^*$ and split the series in (2.6) into two parts,

$$y(t) = \sum_{n=0}^{M-1} C_n e^{-\lambda_n t} + e(M,t),\ \forall\, t \in [T_1, T_2], \tag{3.1}$$

where

$$e(M,t) = \sum_{n=M}^{\infty} C_n e^{-\lambda_n t},\ \forall\, t \in [T_1, T_2]. \tag{3.2}$$

Theorem 3.1 gives a bound of $e(M,t)$.

Theorem 3.1. *Suppose that the diffusion coefficient $\alpha \geq \alpha_0 > 0$ and the initial value $u_0(x)$ satisfies $\|u_0\|_{L^2(0,1)} \leq M_0$ for some $M_0 > 0$. Then for $t \geq T_1 > 0$,*

$$|e(M,t)| < \left(\sqrt{2} + \frac{1}{4M\pi^2\alpha_0 T_1}\right) M_0 e^{-\alpha_0 M^2\pi^2 T_1}. \tag{3.3}$$

Proof. For $t \geq T_1 > 0$, introduce

$$F(\alpha, M, t) = \int_M^{\infty} e^{-\alpha x^2\pi^2 t}\, dx, \tag{3.4}$$

to obtain

$$F^2(\alpha, M, t) = \iint_{D_1} e^{-\alpha(x^2+y^2)\pi^2 t}\, dxdy + \iint_{D_2} e^{-\alpha(x^2+y^2)\pi^2 t}\, dxdy \triangleq I_1 + I_2,$$

where

$$D_1 = \left\{(x,y) \,\middle|\, x \geq y \geq M\right\}, D_2 = \left\{(x,y) \,\middle|\, y > x \geq M\right\}.$$

By symmetry of the integration domains D_1 and D_2 with respect to x and y, $I_1 = I_2$. To compute I_1, we use a double integral in polar coordinates to convert it to iterated integrals

$$I_1 = \iint_{\widetilde{D}_1} e^{-\alpha\rho^2\pi^2 t}\rho\, d\rho d\theta = \int_0^{\frac{\pi}{4}} d\theta \int_{\frac{M}{\sin\theta}}^{\infty} e^{-\alpha\rho^2\pi^2 t}\rho\, d\rho = \frac{1}{2\alpha\pi^2 t}\int_0^{\frac{\pi}{4}} e^{-\frac{\alpha M^2\pi^2 t}{\sin^2\theta}}\, d\theta, \tag{3.5}$$

where

$$\widetilde{D}_1 = \left\{(\rho,\theta) \,\middle|\, \rho \geq \frac{M}{\sin\theta},\ 0 < \theta \leq \frac{\pi}{4}\right\}.$$

The variable substitution $u = \cot^2\theta$ in (3.5) yields

$$\begin{aligned} I_1 &= \frac{1}{4\alpha\pi^2 t}\int_1^{+\infty} e^{-\alpha M^2\pi^2 t(1+u)}\frac{1}{(1+u)\sqrt{u}}\, du \\ &< \frac{1}{4\alpha\pi^2 t}e^{-\alpha M^2\pi^2 t}\int_1^{+\infty} e^{-\alpha M^2\pi^2 t u}u^{-\frac{3}{2}}\, du \\ &= \frac{M}{4\pi\sqrt{\alpha t}}e^{-\alpha M^2\pi^2 t}\cdot\Gamma\left(-\frac{1}{2},\ \alpha M^2\pi^2 t\right), \end{aligned} \tag{3.6}$$

where

$$\Gamma(a,x) = \int_x^{\infty} t^{a-1}e^{-t}\, dt, \tag{3.7}$$

is the upper incomplete gamma function ([1, Section 6.5]). It is known that [1, p.263]

$$\Gamma(a,x) = e^{-x}x^a\left(\frac{1}{x+}\frac{1-a}{1+}\frac{1}{x+}\frac{2-a}{1+}\frac{2}{x+}\cdots\right),\ x > 0,\ |a| < \infty, \tag{3.8}$$

from which we have

$$\Gamma(a,x) < e^{-x}x^{a-1},\ x>0,\ a<0. \tag{3.9}$$

This together with (3.6) gives

$$F(\alpha,M,t) = \sqrt{2I_1} < \frac{1}{\sqrt{2}M\pi^2\alpha t}e^{-\alpha M^2\pi^2 t}. \tag{3.10}$$

We now turn to the estimation of $|e(M,t)|$. It can be computed that for $t \geq T_1 > 0$,

$$\begin{aligned} |e(M,t)| &\leq \left(\sum_{n=M}^{\infty}|C_n|^2\right)^{\frac{1}{2}}\left(\sum_{n=M}^{\infty}e^{-2\alpha n^2\pi^2 t}\right)^{\frac{1}{2}} \\ &< \sqrt{2}M_0\left[e^{-2\alpha M^2\pi^2 t} + \sum_{n=M+1}^{\infty}\int_{n-1}^{n}e^{-2\alpha x^2\pi^2 t}\,dx\right]^{\frac{1}{2}} \\ &= \sqrt{2}M_0\left[e^{-2\alpha M^2\pi^2 t} + F(2\alpha,M,t)\right]^{\frac{1}{2}} \\ &< \left(\sqrt{2}+\frac{1}{4M\pi^2\alpha_0 T_1}\right)M_0 e^{-\alpha_0 M^2\pi^2 T_1}. \end{aligned} \tag{3.11}$$

This completes the proof of the theorem. □

Remark 3.1. It can be seen from (3.3) that if $\alpha_0 M^2\pi^2 T_1$ is sufficiently large, then indeed

$$y(t) \approx \sum_{n=0}^{M-1} C_n e^{-\lambda_n t},\ \forall\, t\in[T_1,T_2], \tag{3.12}$$

with the truncation error $e(M,t)$ estimated in (3.3). It should also be noted that all eigenvalues are real and $-\lambda_n t \to -\infty$ as $n\to\infty$ with the order of n^2. This plays a crucial role in estimation of $e(M,t)$ and guarantees the feasibility of finite dimensional approximation for spectral estimation.

3.2 Matrix pencil method

The matrix pencil method was presented by Hua and Sarkar in [12] for estimation of signal parameters from a noisy exponential sequence. This method has been proved to be quite useful because of its computational efficiency and low sensitivity to the noise.

Suppose that the observed system response can be described by

$$y(t) = x(t) + n(t) = \sum_{i=1}^{M} R_i e^{s_i t} + n(t),\ \forall\, t\in[0,T], \tag{3.13}$$

where $n(t)$ is the noise, $x(t)$ is the system response, $y(t)$ is the noise contaminated observation. Let T_s be the sampling period, the discrete form of (3.13) can be expressed as follows,

$$y(kT_s) = \sum_{i=1}^{M} R_i z_i^k + n(kT_s),\ k=0,1,\ldots,N-1, \tag{3.14}$$

where $z_i = e^{s_i T_s}$, and N is the number of sample points which should be large enough. Generally, the number of exponential components M, the amplitudes R_i, and the poles z_i all can be unknown. In what

follows, we show how to estimate these numbers simultaneously from the observation $\{y(kT_s)\}_{k=0}^{N-1}$ by virtue of the matrix pencil method.

Let $x_k = x(kT_s)$ and $y_k = y(kT_s)$, and define

$$\mathbf{x_k} = [x_k, x_{k+1}, \ldots, x_{N-L+k-1}]^\top,\ \mathbf{y_k} = [y_k, y_{k+1}, \ldots, y_{N-L+k-1}]^\top, \quad k = 0, 1, 2, \ldots, L, \tag{3.15}$$

and

$$\begin{aligned} X_0 &= [\mathbf{x_{L-1}}, \mathbf{x_{L-2}}, \ldots, \mathbf{x_0}], & Y_0 &= [\mathbf{y_{L-1}}, \mathbf{y_{L-2}}, \ldots, \mathbf{y_0}], \\ X_1 &= [\mathbf{x_L}, \mathbf{x_{L-1}}, \ldots, \mathbf{x_1}], & Y_1 &= [\mathbf{y_L}, \mathbf{y_{L-1}}, \ldots, \mathbf{y_1}], \\ X &= [\mathbf{x_0}, \mathbf{x_1}, \ldots, \mathbf{x_L}], & Y &= [\mathbf{y_0}, \mathbf{y_1}, \ldots, \mathbf{y_L}], \end{aligned} \tag{3.16}$$

where the superscript "$\top$" denotes the transpose, and L is called the pencil parameter. It has been pointed out that the best choices for L are $N/3$ and $2N/3$, and all values satisfying $N/3 \le L \le 2N/3$ appear to be good choices in general [12]. In this paper, the pencil parameter L is always chosen to be $N/3$ or $\lfloor N/3 \rfloor + 1$ when $N/3$ is not an integer. Here and in the sequel, $\lfloor \cdot \rfloor$ is as usual the floor function and $\lfloor x \rfloor$ denotes the integer part of the number x.

3.2.1 Estimation of M

In case of noiseless observation, M is equal to the rank of X: $M = \mathrm{rank}(X)$. In case of the noise contaminated observation, suppose that $\{\sigma_i\}$ are the singular values of Y. Then M can be estimated by

$$M = \#\big\{\sigma_i \,|\, \sigma_i/\sigma_{max} \ge \varepsilon\big\}, \tag{3.17}$$

where σ_{max} is the maximal singular value of Y and ε is a threshold. Here and in the sequel, $\#S$ denotes the number of elements in the set S.

3.2.2 Estimation of $\{z_i\}_{i=1}^M$

In case of noiseless observation, it has been proved in [12, Theorem 2.1] that $\{z_i\}_{i=1}^M$ in (3.14) are the M nonzero eigenvalues of the matrix $X_0^\dagger X_1$ when $M \le L \le N - M$, here and in the sequel the superscript "$\dagger$" denotes the pseudoinverse. In case of the noise contaminated observation, suppose that the SVD of Y_0 is $Y_0 = U_0 \Sigma_0 V_0^\top$, and the rank-$M$ truncated pseudoinverse $Y_{0,M}^\dagger$ is defined as

$$Y_{0,M}^\dagger = \sum_{i=1}^{M} \frac{1}{\sigma_i} v_i u_i^* \triangleq V_{0,M} A^{-1} U_{0,M}^*, \tag{3.18}$$

where $\{\sigma_i\}_{i=1}^M$ are the M largest singular values of Y_0; v_i's and u_i's are the corresponding singular vectors. The superscript "$*$" in (3.18) denotes the conjugate transpose. Then the estimates of $\{z_i\}_{i=1}^M$ can be realized by computing the M nonzero eigenvalues of $Y_{0,M}^\dagger Y_1$, or equivalently, the eigenvalues of the $M \times M$ matrix

$$Z_E = A^{-1} U_{0,M}^* Y_1 V_{0,M}. \tag{3.19}$$

And the $\{s_i\}_{i=1}^M$ in (3.13) can be obtained by

$$s_i = \frac{\ln z_i}{T_s},\ i = 1, 2, \ldots, M. \tag{3.20}$$

3.2.3 Estimation of $\{R_i\}_{i=1}^M$

Having estimated the number M of the exponential components and all the poles $\{z_i\}_{i=1}^M$, the amplitudes R_i can be estimated by solving the following linear least squares problem

$$\{R_i\}_{i=1}^M = \operatorname{argmin} \sum_{k=0}^{N-1} \left[y_k - \sum_{i=1}^{M} R_i z_i^k \right]^2. \tag{3.21}$$

This subsection serves as a brief introduction of the matrix pencil method, and the interested readers can refer to [12] for more details. It is easily seen from (3.18) that only the first M largest singular values of Y_0 are retained and other smaller ones (may be dominated by the noise) are chopped off. This is actually the truncated singular value decomposition (TSVD) known as a regularization method, which is based on the perturbation theory for the singular value decomposition (see, e.g., [11, Chapter 4]).

3.3 Identification algorithm

Suppose that $0 < T_1 < T_2 < T_3$ are three arbitrary positive numbers, and the control function $f(t)$ is chosen as in (2.17) and the corresponding observation data is $\{y(t) = u(0,t;f,u_0)|\ t \in [T_1,T_3]\}$. In this section, we formulate the identification for the coefficient and initial value in several steps.

Step 1: Estimate several eigenvalues of system operator $\mathcal{A}$ from the observation without control by the matrix pencil method.

Specifically, let $T_1 = t_0 < \cdots < t_{N_1} = T_2$ be the uniform grids of $[T_1, T_2]$ with the sampling period $T_s = \frac{T_2 - T_1}{N_1}$, then the measured values at sample points are

$$y_i = y(t_i) = \sum_{k=0}^{K-1} \left(C_{n_k} e^{-\lambda_{n_k} T_1} \right) e^{-(\lambda_{n_k} T_s) i},\ i = 0, 1, \ldots, N_1 - 1, \tag{3.22}$$

where $K = \#\mathbb{K}$ with $\mathbb{K}$ defined by (2.7) and $\{C_{n_k}\}_{k=0}^{K-1}$ consists of all the nonzero elements in the series $\{C_n\}_{n\in\mathbb{N}}$. Then the number M of the estimable eigenvalues and the approximate eigenvalues $\{\widetilde{\lambda}_{n_k}\}_{k=0}^{M-1}$ can be obtained by virtue of the matrix pencil method following the process introduced in Sections 3.2.1 and 3.2.2.

Remark 3.2. As stated in Theorem 2.1, it is unknown whether $u_0(x)$ is orthogonal to some of the eigenvectors $\{\phi_n\}_{n\in\mathbb{N}}$. In case that $\langle u_0, \phi_n \rangle = 0$ for some $n \in \mathbb{N}$, then $C_n = 0$ and the observation has nothing to do with $C_n e^{-\lambda_n t_i}$. It is noteworthy that the $\{\widetilde{\lambda}_{n_k}\}_{k=0}^{M-1}$ recovered in *Step 1* are the approximates of some eigenvalues of $\mathcal{A}$, but may not be the first M eigenvalues, i.e., the relationships $\widetilde{\lambda}_{n_k} \approx \lambda_k\ (= \alpha k^2 \pi^2)$ are not always true. In fact, it is true only when $n_k = k$ or $\langle u_0, \phi_k \rangle \neq 0$ for $k = 0, 1, \ldots, M-1$, which is the case mentioned in [23] where such an initial value is said to be generic and in this case the *Steps 3* and *4* below are not necessary anymore. In other words, when $\langle u_0, \phi_k \rangle = 0$ for some k, we can not always recover α from $\{\widetilde{\lambda}_{n_k}\}$ directly.

Step 2: Estimate the coefficients $\{\widetilde{C}_{n_k}\}_{k=0}^{M-1}$ from (3.22) by solving the linear least squares problem

$$\min \sum_{i=0}^{N_1-1} \left[y_i - \sum_{k=0}^{M-1} \widetilde{C}_{n_k} e^{-\widetilde{\lambda}_{n_k} t_i} \right]^2. \tag{3.23}$$

Remark 3.3. After obtaining $\{(\widetilde{C}_{n_k}, \widetilde{\lambda}_{n_k})\}_{k=0}^{M-1}$, the control free part of the observation $u(0,t;0,u_0)$ can be estimated as

$$u(0,t;0,u_0) \approx \sum_{k=0}^{M-1} \widetilde{C}_{n_k} e^{-\widetilde{\lambda}_{n_k} t},\ t>0. \tag{3.24}$$

Step 3: Estimate an approximate value of the diffusion coefficient α by obtaining the first several eigenvalues of $\mathcal{A}$ through $\{y(t) \mid t \in [T_2,\ T_3]\}$ by virtue of the matrix pencil method.

Similar to *Step 1*, let $T_2 = t_0 < t_1 < \cdots < t_{N_2} = T_3$ be the uniform grids of $[T_2, T_3]$ with the sampling period $T_s' = \frac{T_3 - T_2}{N_2}$, and the control is chosen to be $f(t) = 1$ for $t \in [T_2, T_3]$. Then from (2.5) and (3.24) we obtain

$$\begin{aligned} y(t_i) &= u(0,t_i;0,u_0) + u(0,t_i;f,0) \\ &= \sum_{n=0}^{\infty} C_n e^{-\lambda_n t_i} - \int_0^{t_i} G(t_i - s, 0, 0) f(s) ds \\ &\approx \sum_{k=0}^{M-1} \widetilde{C}_{n_k} e^{-\widetilde{\lambda}_{n_k} t_i} - \frac{1}{3\alpha} - T_s' i + \sum_{n=1}^{\infty} \frac{2}{\lambda_n} e^{-\lambda_n' i}. \end{aligned} \tag{3.25}$$

Let

$$y_i' = y(t_i) - \sum_{k=0}^{M-1} \widetilde{C}_{n_k} e^{-\widetilde{\lambda}_{n_k} t_i} + T_s' i,\ i = 0, 1, \ldots, N_2 - 1, \tag{3.26}$$

and

$$C_0' = -\frac{1}{3\alpha}, \lambda_0' = 0;\ C_n' = \frac{2}{\lambda_n}, \lambda_n' = \lambda_n T_s',\ n \in \mathbb{N}^*. \tag{3.27}$$

Then (3.25) becomes

$$y_i' \approx \sum_{n=0}^{\infty} C_n' e^{-\lambda_n' i},\ i = 0, 1, \ldots, N_2 - 1. \tag{3.28}$$

Next, we estimate $\{\,(C_n', \lambda_n')\,\}_{n=0}^{M'-1}$ from (3.28) by repeating the processes in *Steps 1-2*, and then α can be obtained from (2.2) and (3.27).

Remark 3.4. The estimation for $\{\,(C_n', \lambda_n')\,\}_{n=0}^{M'-1}$ from (3.28) is slightly different from that in *Step 1* since none of the $\{C_n'\}_{n=0}^{M'-1}$ is zero although they are also unknown. Hence, we can recover α from the following relations

$$\lambda_n' = \alpha n^2 \pi^2 T_s',\ n = 1, 2, \ldots, M' - 1, \tag{3.29}$$

or

$$C_n' = \frac{2}{\alpha n^2 \pi^2},\ n = 1, 2, \ldots, M' - 1. \tag{3.30}$$

However, the α obtained from (3.29) may be different from that obtained from (3.30) since both $\{C_n'\}$ and $\{\lambda_n'\}$ are estimated values rather than exact ones. In simulations, the pairs (C_n', λ_n') that satisfy

$$C_n' \lambda_n' \approx 2T_s',\ n = 1, 2, \ldots, M' - 1, \tag{3.31}$$

seem to be more credible to estimate α. Actually, the estimated coefficient here is only for the identification of $\{n_k\}_{k=0}^{M-1}$, which is shown in succeeding *Step 4*.

Remark 3.5. It has been pointed out in [12] that the best choice for the pencil parameter L is $N/3$ or $2N/3$ and the relation $M \le L \le N - M$ is necessary in estimation of the poles $\{z_i\}_{i=1}^{M}$, from which it is easy to deduce that the number of sample points N should be no less than $3M$. Although M is also unknown, an estimate of $\overline{M}$, the maximum probable value of M, can be made according to Theorem 3.1 and the error level. Specifically, suppose that the error bound of the measurement is δ. Then the remainder term $e(\overline{M}, t)$ in (3.1) can be merged into the measurement error as long as $e(\overline{M}, t) \le \delta$, which can be solved based on the estimate (3.3). In this way, the number of sample points N can be chosen to be any number satisfying $N > 3\overline{M}$.

Step 4: Estimate α from $\left\{\widetilde{\lambda}_{n_k}\right\}_{k=0}^{M-1}$ and reconstruct the initial state $u_0(x)$.

To be specific, after estimating $\left\{\left(\widetilde{C}_{n_k}, \widetilde{\lambda}_{n_k}\right)\right\}_{k=0}^{M-1}$ in *Steps 1-2*, and recovering an approximate of α in *Step 3*, we can now determine the series $\mathbb{K}_M = \left\{n_k\right\}_{k=0}^{M-1}$ by

$$n_k = \left\lfloor \sqrt{\widetilde{\lambda}_{n_k}/(\alpha\pi^2)} \right\rceil, \ k = 0, 1, \ldots, M-1, \tag{3.32}$$

where $\lfloor x \rceil$ denotes the integer nearest to x. Then, α can be estimated by

$$\alpha_k = \widetilde{\lambda}_{n_k}/(n_k^2\pi^2) \text{ for } n_k \neq 0, \ k = 0, 1, \ldots, M-1. \tag{3.33}$$

Now we turn to the initial value. It is clear from (2.5) that

$$y(t) = \sum_{n=0}^{\widetilde{M}-1} A_n(0)e^{-\lambda_n t} + e(\widetilde{M}, t), \ \forall\, t \in [T_0, T_2), \tag{3.34}$$

where $T_0 \in (0, T_2)$. It follows from Theorem 3.1 that we can choose properly $\widetilde{M}$ and T_0 such that $\left|e(\widetilde{M}, t)\right|$ is sufficiently small. Suppose that only observations at the sample points $T_0 = t_0 < t_1 < \cdots < t_N = T_2$ are available. Then the coefficients $\{A_n(0)\}$ can be estimated by solving the following problem:

$$\min \sum_{i=0}^{N-1} \left[y(t_i) - \sum_{n=0}^{\widetilde{M}-1} A_n(0)e^{-\alpha n^2\pi^2 t_i} \right]^2, \tag{3.35}$$

or equivalently, finding the least squares solution of the matrix equation

$$CA = b, \tag{3.36}$$

where C is an $N \times \widetilde{M}$ matrix with the (i, j) elements

$$C(i, j) = e^{-\alpha(j-1)^2\pi^2 t_{i-1}}, \ i = 1, \ldots, N, \ j = 1, \ldots, \widetilde{M}, \tag{3.37}$$

and

$$A = [A_0(0), \cdots, A_{\widetilde{M}-1}(0)]^\top, \ b = [y(t_0), \cdots, y(t_{N-1})]^\top.$$

Since the reconstruction of the initial value is known to be ill-posed, it may result in the matrix equation (3.36) to be ill-posed as well. In order to obtain stable results, we use the TSVD regularization method to solve (3.36). Suppose that the SVD of matrix C is $C = U_C\Sigma_C V_C^\top$, where $U_C = [u_1', u_2', \cdots, u_N']$

and $V_C = [v_1', v_2', \cdots, v_{\widetilde{M}}']$ are orthonormal matrices with the left and right singular vectors, respectively. $\Sigma_C = \mathrm{diag}(\sigma_1, \sigma_2, \cdots)$ is a diagonal matrix with non-negative diagonal elements being the singular values of C. In the TSVD method, the matrix C is replaced by its rank-k approximation, and the regularized solution is given by

$$A_{reg} = \sum_{i=1}^{k} \frac{u_i'^{\top} b}{\sigma_i} v_i' \triangleq C^I b, \tag{3.38}$$

where $k \leq \mathrm{rank}(C)$ is the regularization parameter, which can be determined by the generalized cross-validation (GCV) criterion [8]. The GCV criterion determines the optimal regularization parameter k by minimizing the following GCV function:

$$G(k) = \frac{\|CA_{reg} - b\|^2}{(\mathrm{trace}(I_N - CC^I))^2}. \tag{3.39}$$

Having obtained the regularized solution A_{reg}, the initial value can be estimated by the asymptotic Fourier series expansion:

$$u_0(x) \approx \sum_{n=0}^{\widetilde{M}-1} A_n(0) \cos n\pi x. \tag{3.40}$$

Remark 3.6. It is evident that the reconstructed initial value, $\widetilde{u}_0(x)$, in (3.40) is an approximated Fourier series expansion of $u_0(x)$ with the first $\widetilde{M}$ terms. Actually, since α has been estimated, there are various methods for the initial state reconstruction, see, e.g., [19, 29] and the references therein. Compared with those methods, the method here is more direct and simpler.

Remark 3.7. Considering the applicability and stability of the numerical algorithm, we point out that the choices of the sampling parameters (i.e., N_1, N_2, T_1, T_2, T_3) are relatively flexible. For the numbers of sample points N_1, N_2, the only requirement is that they should be large enough. For the sampling times T_1, T_2, T_3, they should be properly small especially when the diffusion coefficient is large. Otherwise, the measured values at the sample points will indicate a constant trend due to the dissipative nature of heat equation (see, e.g., Theorem 3.1), and such measured values can not provide enough information for the inverse problems. Therefore, a feasible criterion for choosing them is that the measured values at the sample points should display a clear trend of decay.

To end this section, we indicate that the inverse problem concerned in this paper is ill-posed. In order to achieve good accuracy and stability, regularization methods are the most powerful and efficient methods in solving ill-posed problems. It can be seen that matrix pencil method plays a key role in the estimation of the diffusion coefficient, and as has been pointed out at the end of section 3.2, the TSVD as a regularization method is contained in the matrix pencil method, which enhances the stability of the numerical algorithm for coefficient estimation. In the last step of the identification algorithm, TSVD is used once again in solving the ill-posed matrix equation (3.36) in order to obtain stable results for the reconstruction of initial state.

4 Error analysis

Noise sensitivity of the matrix pencil method for estimation of finite signal parameters from a noisy exponential sequence is analyzed in [12]. But our case is different in two aspects. First, the number of unknown parameters in the infinite spectral estimation is not finite. Second, the perturbation, that is, the remainder term $e(M,t)$ in (3.2), is not random. Our error analysis is mainly based on the estimation of $|e(M,t)|$ for the finite dimensional approximation in Section 3.1 and the perturbation theory for matrix eigenvalue problems.

In this section, we establish an error analysis by applying the matrix pencil method to the infinite spectral estimation problem (3.1). We may suppose without loss of generality that $C_n \neq 0$ for any $n \in \mathbb{N}$. In fact, we consider only the first M nonzero terms in series (3.1), where M is defined as (3.17). Let $T_1 = t_0 < t_1 < \cdots < t_{N-1} = T_2$ be the points on a uniform grid of $[T_1, T_2]$ with the sampling period $T_s = \frac{T_2 - T_1}{N-1}$, and the observation data at sample points, t_i, are

$$y_i = y(t_i) \triangleq \sum_{n=0}^{M-1} \left(C_n e^{-\lambda_n T_1} \right) z_n^i + e(M, t_i), \tag{4.1}$$

where $z_n = e^{-\lambda_n T_s}$. By Theorem 3.1, it follows that

$$|y_i - x_i| = |e(M, t_i)| < M_\alpha e^{-\alpha_0 M^2 \pi^2 T_s i}, \tag{4.2}$$

where $x_i = \sum_{n=0}^{M-1} \left(C_n e^{-\lambda_n T_1} \right) z_n^i$ and

$$M_\alpha = \left(\sqrt{2} + \frac{1}{4M\pi^2 \alpha_0 T_1} \right) M_0 e^{-\alpha_0 M^2 \pi^2 T_1}. \tag{4.3}$$

Define the matrices X_0, Y_0, X_1, Y_1 as (3.15)-(3.16). Theorem 4.1 below gives the bounds of $\|Y_0 - X_0\|_F$ and $\|Y_1 - X_1\|_F$, where $\|\cdot\|_F$ denotes the matrix Frobenius norm.

Theorem 4.1. *Suppose that the number of sample points $N > 9$ and*

$$\theta = 2\alpha_0 M^2 \pi^2 T_s. \tag{4.4}$$

Then

$$\|Y_0 - X_0\|_F < M_\alpha \sqrt{M_{\theta,L} + (1 + \theta^{-1})^2}, \tag{4.5}$$

and

$$\|Y_1 - X_1\|_F < M_\alpha \sqrt{M_{\theta,L+1} + (\theta^{-1} + \theta^{-2})\, e^{-\theta}}, \tag{4.6}$$

where M_α is defined in (4.3) and

$$M_{\theta,L} = \begin{cases} e^{-\theta}, & \theta \geq 1, \\ 2\theta^{-1} e^{-1}, & (L-1)^{-1} < \theta < 1, \\ (L-1) e^{-(L-1)\theta}, & 0 < \theta \leq (L-1)^{-1}. \end{cases} \tag{4.7}$$

Proof. Notice that the matrix obtained from X_0 with its columns flipped in the left-right direction admits the Hankel structure ([7, p.339]) and so is Y_0, it is easy to deduce from the definition of Frobenius norm that

$$\begin{aligned}\|Y_0 - X_0\|_F^2 &= \sum_{i=0}^{L-1}(i+1)\,|y_i - x_i|^2 + \sum_{i=1}^{L} i\,|y_{N-1-i} - x_{N-1-i}|^2 + L\sum_{i=L}^{N-L-2}|y_i - x_i|^2 \\ &\le M_\alpha^2\Big[\sum_{i=0}^{L-1} ie^{-\theta i} + \sum_{i=0}^{L-1} e^{-\theta i} + \sum_{i=1}^{L} ie^{-\theta(N-1-i)} + L\sum_{i=L}^{N-L-2} e^{-\theta i}\Big] \\ &\triangleq M_\alpha^2\left[S_1 + S_2 + S_3 + S_4\right].\end{aligned}$$

To estimate S_1, we introduce

$$f(x) = xe^{-\theta x},\ x \ge 0, \tag{4.8}$$

which satisfies

$$f'(x) = (1-\theta x)e^{-\theta x},\ x \ge 0.$$

There are three different cases according to the values of θ.

Case 1: $\theta \ge 1$. In this case, $f'(x) \le 0$ for $x \ge 1$. Hence

$$ie^{-\theta i} \le \int_{i-1}^{i} xe^{-\theta x}dx,\ i = 2, 3, \ldots, L-1. \tag{4.9}$$

Therefore,

$$S_1 = e^{-\theta} + \sum_{i=2}^{L-1} ie^{-\theta i} \le e^{-\theta} + \sum_{i=2}^{L-1}\int_{i-1}^{i} xe^{-\theta x}dx = e^{-\theta} + \int_{1}^{L-1} xe^{-\theta x}dx.$$

Case 2: $\frac{1}{L-1} < \theta < 1$. In this case, $f'(x) \ge 0$ for $0 \le x \le \theta^{-1}$, and $f'(x) < 0$ for $x > \theta^{-1}$, which imply

$$f(x) \le f(\theta^{-1}) = \theta^{-1}e^{-1},\ x \ge 0,$$

and

$$ie^{-\theta i} \le \int_{i}^{i+1} xe^{-\theta x}dx,\ i = 1, \ldots, \lfloor\theta^{-1}\rfloor - 1, \tag{4.10}$$

$$ie^{-\theta i} \le \int_{i-1}^{i} xe^{-\theta x}dx,\ i = \lfloor\theta^{-1}\rfloor + 2, \ldots, L-1. \tag{4.11}$$

Thus, we have

$$\begin{aligned}S_1 &\le \sum_{i=1}^{\lfloor\frac{1}{\theta}\rfloor-1}\int_{i}^{i+1} xe^{-\theta x}dx + f\left(\lfloor\theta^{-1}\rfloor\right) + f\left(\lfloor\theta^{-1}\rfloor + 1\right) + \sum_{i=\lfloor\frac{1}{\theta}\rfloor+2}^{L-1}\int_{i-1}^{i} xe^{-\theta x}dx \\ &\le \int_{1}^{\lfloor\frac{1}{\theta}\rfloor} xe^{-\theta x}dx + 2f\left(\theta^{-1}\right) + \int_{\lfloor\frac{1}{\theta}\rfloor+1}^{L-1} xe^{-\theta x}dx \\ &\le \frac{2}{\theta}e^{-1} + \int_{1}^{L-1} xe^{-\theta x}dx.\end{aligned}$$

Case 3: $0 < \theta \le \dfrac{1}{L-1}$. In this case, $f'(x) \ge 0$ for $0 \le x \le L-1$. Hence

$$ie^{-\theta i} \le \int_{i}^{i+1} xe^{-\theta x}dx,\ i = 1, \ldots, L-2. \tag{4.12}$$

Therefore,

$$S_1 \le (L-1)e^{-(L-1)\theta} + \sum_{i=1}^{L-2} \int_i^{i+1} xe^{-\theta x} dx = (L-1)e^{-(L-1)\theta} + \int_1^{L-1} xe^{-\theta x} dx.$$

Combining the three cases discussed above gives

$$S_1 \le M_{\theta,L} + \int_1^{L-1} xe^{-\theta x} dx = M_{\theta,L} + \frac{1}{\theta}\left(1+\frac{1}{\theta}\right)e^{-\theta} - \frac{1}{\theta}\left(L-1+\frac{1}{\theta}\right)e^{-\theta(L-1)}, \tag{4.13}$$

where $M_{\theta,L}$ is defined in (4.7).

An analogous analysis of S_2, S_3, and S_4 gives

$$\begin{aligned} S_2 &\le 1 + \frac{1}{\theta} - \frac{1}{\theta}e^{-\theta(L-1)}, \\ S_3 &\le \frac{1}{\theta}e^{-\theta(N-2)}\left[\left(L+1-\frac{1}{\theta}\right)e^{\theta L} + \frac{1}{\theta} - 1\right] \triangleq M_{\theta,N}, \\ S_4 &\le \frac{L}{\theta}\left[e^{-\theta(L-1)} - e^{-\theta(N-L-2)}\right]. \end{aligned} \tag{4.14}$$

A simple calculation shows that when $N > 9$,

$$M_{\theta,N} < \frac{1}{\theta^2}e^{-\theta(L-1)} + \frac{L}{\theta}e^{-\theta(N-L-2)}. \tag{4.15}$$

Consequently,

$$\begin{aligned} S_1+S_2+S_3+S_4 &\le M_{\theta,L} + \frac{1}{\theta}\left(1+\frac{1}{\theta}\right)e^{-\theta} - \frac{1}{\theta}\left(L-1+\frac{1}{\theta}\right)e^{-\theta(L-1)} + 1 + \frac{1}{\theta} \\ &\quad -\frac{1}{\theta}e^{-\theta(L-1)} + M_{\theta,N} + \frac{L}{\theta}\left[e^{-\theta(L-1)} - e^{-\theta(N-L-2)}\right] \\ &= M_{\theta,L} + \frac{1}{\theta}\left(1+\frac{1}{\theta}\right)e^{-\theta} - \frac{1}{\theta^2}e^{-\theta(L-1)} + 1 + \frac{1}{\theta} + M_{\theta,N} - \frac{L}{\theta}e^{-\theta(N-L-2)} \\ &< M_{\theta,L} + \left(1+\frac{1}{\theta}\right)^2. \end{aligned}$$

Hence,

$$\|Y_0 - X_0\|_F < M_\alpha \sqrt{M_{\theta,L} + (1+\theta^{-1})^2}.$$

By applying similar analysis to $\|Y_1 - X_1\|_F$, we can achieve the estimation (4.6). The details are omitted. This completes the proof of the theorem. □

Suppose the singular values of Y_0 are $\sigma(Y_0) = \{\sigma_i\}$ and $Y_{0,M}$ is the rank-M truncated approximation of Y_0 defined by

$$Y_{0,M} = U_{0,M} A V_{0,M}^\top. \tag{4.16}$$

Now we are in a position to give an error analysis for the infinite spectral estimation problem (3.1) using the matrix pencil method.

Theorem 4.2. *Let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_M$ be the first M singular values of the matrix Y_0, and assume that $Y_{0,M}^{\dagger} Y_1$ is diagonalizable, i.e., $Y_{0,M}^{\dagger} Y_1 = X_M \widetilde{\Lambda}_M X_M^{-1}$, where*

$$\widetilde{\Lambda}_M = \mathrm{diag}(\widetilde{z}_1, \ldots, \widetilde{z}_M, 0, \ldots, 0),\ \widetilde{z}_1 \geq \widetilde{z}_2 \geq \cdots \geq \widetilde{z}_M.$$

The nonzero eigenvalues of the matrix $X_0^{\dagger} X_1$ are supposed to be

$$\Lambda_M = \{z_1, z_2, \ldots, z_M\},\ z_1 \geq z_2 \geq \cdots \geq z_M.$$

Let M_α, θ, $M_{\theta,L}$ be defined as in (4.3), (4.4), (4.7), respectively, and let

$$\rho = \left[\|Y_{0,M} - Y_0\|_2 + M_\alpha \sqrt{M_{\theta,L} + (1+\theta^{-1})^2} \right] \Big/ \sigma_M. \tag{4.17}$$

If $\rho < 1$, then

$$|\widetilde{z}_n - z_n| < \frac{\kappa(X_M)}{\sigma_M \cdot (1-\rho)} \left[\frac{1+\sqrt{5}}{2} \rho \|Y_1\|_2 + M_\alpha \sqrt{M_{\theta,L+1} + (\theta^{-1} + \theta^{-2})\, e^{-\theta}} \right]. \tag{4.18}$$

In particular, if $\theta > \dfrac{1}{L-1}$, then

$$|\widetilde{z}_n - z_n| < \frac{\kappa(X_M) \cdot \rho}{\sigma_M \cdot (1-\rho)} \left[\frac{1+\sqrt{5}}{2} \|Y_1\|_2 + \sigma_M \right]. \tag{4.19}$$

Proof. We need to estimate the matrix norm $\left\| Y_{0,M}^{\dagger} Y_1 - X_0^{\dagger} X_1 \right\|_2$ first, which is done as follows:

$$\begin{aligned} \left\| Y_{0,M}^{\dagger} Y_1 - X_0^{\dagger} X_1 \right\|_2 &= \left\| Y_{0,M}^{\dagger} Y_1 - X_0^{\dagger} Y_1 + X_0^{\dagger} Y_1 - X_0^{\dagger} X_1 \right\|_2 \\ &\leq \left\| Y_{0,M}^{\dagger} Y_1 - X_0^{\dagger} Y_1 \right\|_2 + \left\| X_0^{\dagger} Y_1 - X_0^{\dagger} X_1 \right\|_2 \\ &\leq \left\| Y_{0,M}^{\dagger} - X_0^{\dagger} \right\|_2 \cdot \|Y_1\|_2 + \left\| X_0^{\dagger} \right\|_2 \cdot \|Y_1 - X_1\|_2. \end{aligned}$$

Since $Y_{0,M}$ is the rank-M truncated matrix of Y_0, $\mathrm{rank}(X_0) = M = \mathrm{rank}(Y_{0,M})$. An application of [22, Theorem 3.4] yields

$$\left\| Y_{0,M}^{\dagger} - X_0^{\dagger} \right\|_2 \leq \frac{1+\sqrt{5}}{2} \left\| Y_{0,M}^{\dagger} \right\|_2 \cdot \left\| X_0^{\dagger} \right\|_2 \cdot \|Y_{0,M} - X_0\|_2 \,. \tag{4.20}$$

Hence,

$$\left\| Y_{0,M}^{\dagger} Y_1 - X_0^{\dagger} X_1 \right\|_2 \leq \frac{1+\sqrt{5}}{2} \left\| Y_{0,M}^{\dagger} \right\|_2 \cdot \left\| X_0^{\dagger} \right\|_2 \cdot \|Y_{0,M} - X_0\|_2 \cdot \|Y_1\|_2 + \left\| X_0^{\dagger} \right\|_2 \cdot \|Y_1 - X_1\|_2. \tag{4.21}$$

Since $\left\| Y_{0,M}^{\dagger} \right\|_2 = \sigma_M^{-1}$,

$$\|Y_{0,M} - X_0\|_2 \cdot \left\| Y_{0,M}^{\dagger} \right\|_2 \leq \left(\|Y_{0,M} - Y_0\|_2 + \|Y_0 - X_0\|_F \right) \cdot \left\| Y_{0,M}^{\dagger} \right\|_2 < \rho,$$

where the last inequality is based on the estimation of $\|Y_0 - X_0\|_F$ in Theorem 4.1. Since $\rho < 1$, it follows from [28, Lemma 3.1] that

$$\left\| X_0^{\dagger} \right\|_2 \leq \frac{\left\| Y_{0,M}^{\dagger} \right\|_2}{1 - \left\| Y_{0,M}^{\dagger} \right\|_2 \cdot \|Y_{0,M} - X_0\|_2} < \frac{1}{(1-\rho)\sigma_M}. \tag{4.22}$$

As a result,

$$\left\|Y_{0,M}^{\dagger}Y_1 - X_0^{\dagger}X_1\right\|_2 < \frac{1}{(1-\rho)\cdot\sigma_M}\left[\frac{1+\sqrt{5}}{2}\rho\|Y_1\|_2 + \|Y_1 - X_1||_F\right].$$

By [3, Theorem 3], we have

$$\begin{aligned}|\widetilde{z}_n - z_n| &\leq \kappa(X_M)\cdot\left\|Y_{0,M}^{\dagger}Y_1 - X_0^{\dagger}X_1\right\|_2 < \frac{\kappa(X_M)}{(1-\rho)\cdot\sigma_M}\left[\frac{1+\sqrt{5}}{2}\rho\|Y_1\|_2 + \|Y_1 - X_1||_F\right] \\ &< \frac{\kappa(X_M)}{\sigma_M\cdot(1-\rho)}\left[\frac{1+\sqrt{5}}{2}\rho\|Y_1\|_2 + M_\alpha\sqrt{M_{\theta,L+1} + (\theta^{-1}+\theta^{-2})\,e^{-\theta}}\right].\end{aligned}$$

In particular, if $\theta > \dfrac{1}{L-1}$, it follows from the definition of $M_{\theta,L}$ in (4.7) that

$$M_{\theta,L+1} = M_{\theta,L},\ \ \theta > \frac{1}{L-1}. \tag{4.23}$$

On the other hand, since

$$\frac{1}{\theta}\left(1+\frac{1}{\theta}\right)e^{-\theta} < \left(1+\frac{1}{\theta}\right)^2,\ \ \theta > 0,$$

it follows from (4.23) that

$$\sqrt{M_{\theta,L+1} + (\theta^{-1}+\theta^{-2})\,e^{-\theta}} < \sqrt{M_{\theta,L} + (1+\theta^{-1})^2},$$

and then

$$\begin{aligned}|\widetilde{z}_n - z_n| &< \frac{\kappa(X_M)}{\sigma_M\cdot(1-\rho)}\left[\frac{1+\sqrt{5}}{2}\rho\|Y_1\|_2 + M_\alpha\sqrt{M_{\theta,L} + (1+\theta^{-1})^2}\right] \\ &= \frac{\kappa(X_M)}{\sigma_M\cdot(1-\rho)}\left[\frac{1+\sqrt{5}}{2}\rho\|Y_1\|_2 + \rho\cdot\sigma_M - \|Y_{0,M} - Y_0\|_2\right] \\ &< \frac{\kappa(X_M)\cdot\rho}{\sigma_M\cdot(1-\rho)}\left[\frac{1+\sqrt{5}}{2}\|Y_1\|_2 + \sigma_M\right].\end{aligned}$$

This ends the proof of the theorem. □

Remark 4.1. By $z_n = e^{-\lambda_n T_s}$, we can also obtain an error estimation $|\widetilde{\lambda}_n - \lambda_n|$, between the estimated eigenvalues and the exact eigenvalues, that is (for $\theta > \frac{1}{L-1}$),

$$|\widetilde{\lambda}_n - \lambda_n| = \frac{|\ln\widetilde{z}_n - \ln z_n|}{T_s} = \frac{|\widetilde{z}_n - z_n|}{T_s\cdot\bar{z}_n} < \frac{\kappa(X_M)\cdot\rho}{\sigma_M(1-\rho)T_s\cdot\bar{z}_n}\left[\frac{1+\sqrt{5}}{2}\|Y_1\|_2 + \sigma_M\right], \tag{4.24}$$

where the mean value theorem has been applied in the second equality and $\bar{z}_n$ is between $\widetilde{z}_n$ and z_n. In addition, we can choose $\bar{z} = \widetilde{z}_n$ in case of $|\widetilde{z}_n - z_n| \ll |\widetilde{z}_n|$.

Remark 4.2. We point out that the estimation seems hard to improve further. It can be seen that the estimation of $\left\|Y_{0,M}^{\dagger}Y_1 - X_0^{\dagger}X_1\right\|_2$ plays a key role in the proof of Theorem 4.2. The condition $\rho < 1$ is mainly for the estimation of $\left\|X_0^{\dagger}\right\|$, which becomes extremely complicated for $\rho \geq 1$ due to the unknown nature of X_0. However, by (4.17), since the value of ρ is determined by M (determined by ε in (3.17)) and T_1, the parameters ε and T_1 can be chosen appropriately in applications to make ρ relatively small, and from (4.19), the error bound becomes smaller as ρ/σ_M becomes smaller.

5 Numerical simulation

In this section, two numerical examples are presented to illustrate the performance of the algorithm developed in Section 3, and all the numerical results are obtained in Matlab. It is known that the observation data are inevitably contaminated by some measurement errors. In the numerical computations, the noisy data are generated by adding a random perturbation, i.e.,

$$\widetilde{y}(t_i) = y(t_i)(1 + \delta \cdot \mathrm{rand}(i)), \tag{5.1}$$

where δ indicates the noise level and $\mathrm{rand}(i)$ generates the random number between $[-1, 1]$.

Example 5.1. Suppose the diffusion coefficient $\alpha^* = 0.1$ and initial value $u_0^*(x) = x - 9\cos\pi x + 5\cos 3\pi x$ in system (1.1). Since $\langle u_0^*, \phi_{2n}\rangle = 0,\ n \in \mathbb{N}^*$, this initial value is not generic ([23]). The observation $y(t) = u(0, t; f, u_0^*)$ is obtained by solving the direct problem (1.1) with the Crank-Nicolson method ([6]), where the step sizes for time $\Delta t = 0.001$ and for space $\Delta x = 0.01$.

For the case of this inverse problem, the time interval is chosen to be $[T_1, T_2, T_3] = [0.1, 0.6, 1.1]$ since the measured values indeed display a trend of decay (attributed to small α). The number of sample points $N_1 = N_2 = 50$, the pencil parameters $L = L' = 17$ and the sampling period $T_s = T_s' = 0.01$. For noise free measured data, i.e., $\delta = 0$ in (5.1), we can obtain the estimated $\left\{\widetilde{C}_{n_k}, \widetilde{\lambda}_{n_k}, \lambda_n', \alpha\right\}$ by following the Steps 1-3 in Section 3.3, and we can also determine the series $\mathbb{K}_M$ by (3.32) following Step 4. The results are partially presented in Table 1, and the coefficient estimation through (3.33) is $\alpha^* \approx 0.1000$. Then the initial value $u_0(x)$ can be estimated by the Fourier series expansion (3.40) by solving the matrix equation (3.36) with TSVD. The results are given in Figure 2, from which we can see that the numerical result is quite accurate.

Table 1: The estimated $\left\{\widetilde{C}_{n_k}, \widetilde{\lambda}_{n_k}, \lambda_n', \alpha, n_k, \alpha_k\right\}$ following Steps 1-4

n(or k)	$\widetilde{C}_{n_k}$	$\widetilde{\lambda}_{n_k}$	$100 * \lambda_n'$	α	n_k	α_k
0	0.5000	0.0000	-0.0105	–	0	–
1	-9.4053	0.9869	1.0630	0.1077	1	0.1000
2	4.9549	8.8761	4.3056	0.1091	3	0.0999
3	-0.0162	24.6246	9.9040	0.1115	5	0.0998
4	-0.0083	48.1762	18.2243	0.1154	7	0.0996

Similar to the identification process for the noise free case, numerical results for identification of α from noisy data with various relative noise levels $\delta = 0.5\%, 1\%$ are found to be 0.1087 and 0.1124, respectively, and the corresponding initial values are presented in Figure 2, which shows that the estimated initial values are reasonably in agreement with the real initial data.

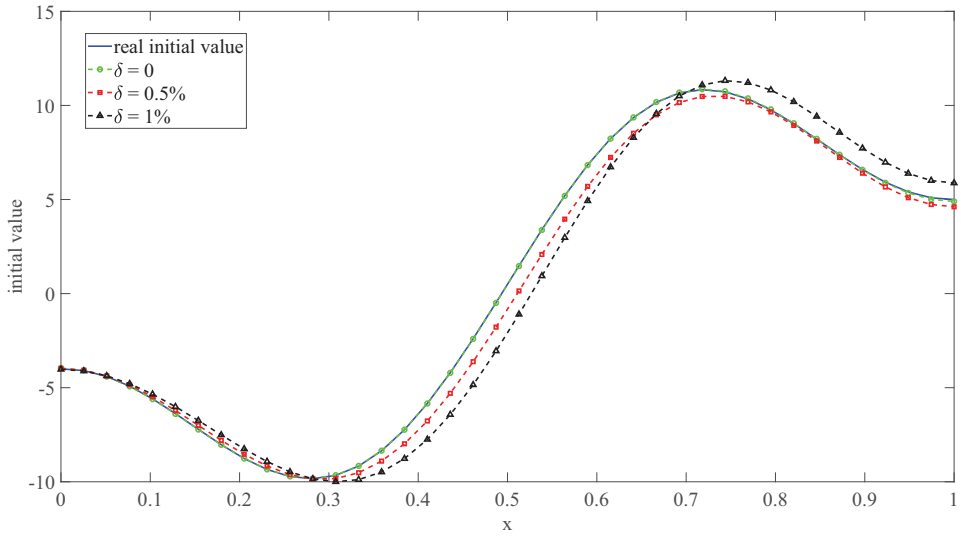

Figure 2: The initial values for Example 5.1 for various levels of noise

Example 5.2. Suppose the diffusion coefficient $\alpha^* = 10$ and initial value $u_0^*(x) = x^2 - x^4 + e^{-2x}\cos x$. The observation $y(t) = u(0, t; f, u_0^*)$ is obtained by solving the direct problem (1.1) with the Crank-Nicolson method. For the inverse problem, the time interval is chosen to be $[T_1, T_2, T_3] = [0.0001, 0.1001, 0.2001]$, which is much smaller than that in Example 5.1 since the measured values in larger time intervals are almost constants (attributed to large α) which cannot be used for identification. The number of sample points $N_1 = N_2 = 100$, the pencil parameters $L = L' = 34$ and the sampling period $T_s = T_s' = 0.001$. The numerical results of diffusion coefficient for various noise levels in the cases of $\delta = 0,\ 0.5\%,\ 1\%$ are 10.0000, 10.2810, 10.9125, respectively, and the corresponding estimated initial values are shown in Figure 3. Although the accuracy of the reconstruction does not seem to be as good as that in Example 5.1, we still obtain the overall shape of the initial value. The reason for the low accuracy lies in the larger diffusion coefficient which leads to more severe ill-posedness.

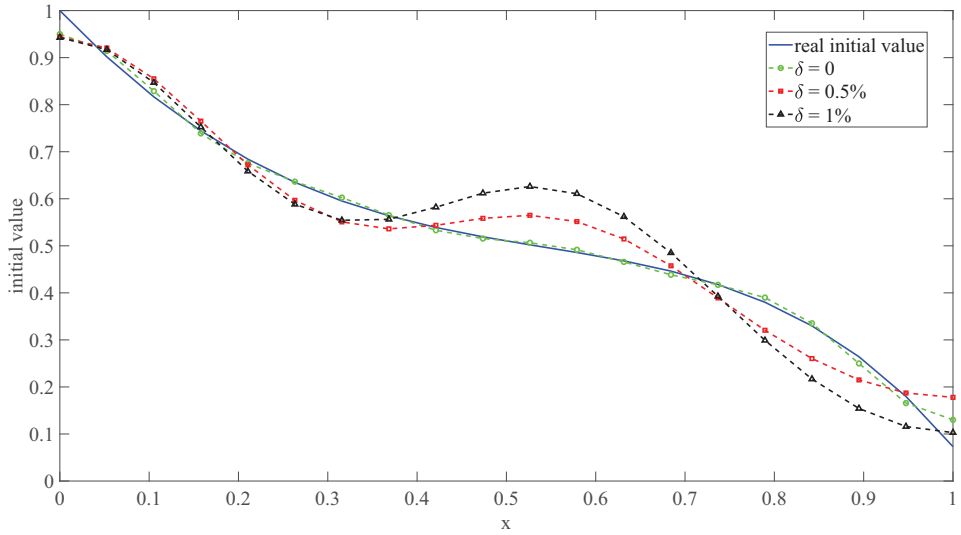

Figure 3: The initial values for Example 5.2 for various levels of noise

6 Concluding remarks

In this paper, we represent the boundary observation with boundary Neumann control for a one-dimensional heat equation into a Dirichlet series in terms of spectrum determined by the diffusivity and coefficients determined by the initial value. The identification of diffusion coefficient and initial value is therefore transformed into an inverse problem of reconstruction of spectrum-coefficient data from the observation. Taking the first finite terms of the series, the problem happens to be an inverse problem of finite exponential sequence with deterministic small perturbation. We are thus able to develop an algorithm to reconstruct simultaneously the diffusion coefficient and initial value by the matrix pencil method which is used in signal processing. An error analysis is presented and a numerical experiment is carried out to validate the efficiency and accuracy of the proposed algorithm. The method developed is promising and can be applied in identification of variable coefficients and other PDEs.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No. 11901433), the Science & Technology Development Fund of Tianjin Education Commission for Higher Education, China (No. 2018KJ157), Tianjin 131 Talents Project, Program for Innovative Research Team in Universities of Tianjin, China (No. TD13-5078), the National Research Foundation of South Africa (No. 93099, 93476), and the Postdoctoral Fellowship of the University of Pretoria.

References

[1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York, 1964.

[2] S.A. Avdonin and S.A. Ivanov, *Families of Exponentials: The Method of Moments in Controllability Problems for Distributed Parameter Systems*, Cambridge University Press, 1995.

[3] F.L. Bauer and C.T. Fike, Norms and exclusion theorems, *Numer. Math.*, 2(1960), 137–141.

[4] A. Benabdallah, P. Gaitan, and J.L. Rousseau, Stability of discontinuous diffusion coefficients and initial conditions in an inverse problem for the heat equation, *SIAM J. Control Optim.*, 46(2007), 1849-1881.

[5] M. Choulli and M. Yamamoto, Uniqueness and stability in determining the heat radiative coefficient, the initial temperature and a boundary coefficient in a parabolic equation, *Nonlinear Anal.*, 69(2008), 3983-3998.

[6] J. Crank and P. Nicolson, A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type, *Proc. Cambridge Phil. Soc.*, 43(1947), 50-67.

[7] F.R. Gantmacher, *The Theory of Matrices: Volume One*, Chelsea Publishing Company, New York, 1960.

[8] G.H. Golub, M. Heath and G. Wahba, Generalized cross-validation as a method for choosing a good ridge parameter, *Technometrics*, 21(1979), 215-223.

[9] B.Z. Guo and J.D. Chang, Simultaneous identifiability of coefficients, initial state and source for string and beam equations via boundary control and observation, *Proc. 8th Asian Control Conference,* Kaohsiung, 2011, 365–370.

[10] S. Gutman and J.H. Ha, Identifiability of piecewise constant conductivity in a heat conduction process, *SIAM J. Control Optim.*, 46(2007), 694–713.

[11] P.C. Hansen, *Discrete Inverse Problems: Insight and Algorithms*, SIAM, Philadelphia, 2010.

[12] Y.B. Hua and T.K. Sarkar, Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise, *IEEE Trans. Acoust. Speech Signal Process.,* 38(1990), 814–824.

[13] V. Isakov, *Inverse Problems for Partial Differential Equations*, Springer, New York, 1998.

[14] S. Kitamura and S. Nakagiri, Identifiability of spatially-varying and constant parameters in distributed systems of parabolic type, *SIAM J. Control Optim.*, 15(1977), 785–802.

[15] R. Murayama, The Gel'fand-Levitan theory and certain inverse problems for the parabolic equation, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 28(1981), 317–330.

[16] S. Nakagiri, Identifiability of linear systems in Hilbert spaces, *SIAM J. Control Optim.*, 21(1983), 501–530.

[17] Y. Orlov and J. Bentsman, Adaptive distributed parameter systems identification with enforceable identifiability conditions and reduced-order spatial differentiation, *IEEE Trans. Automat. Control*, 45(2000), 203–216.

[18] A. Pierce, Unique identification of eigenvalues and coefficients in a parabolic equation, *SIAM J. Contorl Optim.*, 17(1979), 494–499.

[19] K. Ramdani, M. Tucsnak, and G. Weiss, Recovering the initial state of an infinite-dimensional system using observers, *Automatica,* 46(2010), 1616–1625.

[20] S. Saks and A. Zygmund, *Analytic Functions*, Monografie Matematyczne, Warsaw, 1965.

[21] A. Smyshlyaev, Y. Orlov and M. Krstic, Adaptive identification of two unstable PDEs with boundary sensing and actuation, *Int. J. Adapt. Control Signal Process.*, 23(2009), 131–149.

[22] G.W. Stewart, On the perturbation of pseudo-inverses, projections and linear least squares problems, *SIAM Rev.*, 19(1977), 634–662.

[23] T. Suzuki and R. Murayama, A uniqueness theorem in an identification problem for coefficients of parabolic equations, *Proc. Japan Acad. Ser. A Math. Sci.*, 56(1980), 259–263.

[24] T. Suzuki, Uniqueness and nonuniqueness in an inverse problem for the parabolic equation, *J. Differential Equations*, 47(1983), 296–316.

[25] E.C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, 2nd Edtion, Clarendon Press, Oxford, 1948.

[26] M. Tucsnak and G. Weiss, *Observation and Control for Operator Semigroups*, Birkhäuser, Basel, 2009.

[27] Y.B. Wang, J. Cheng, J. Nakagawa, and M. Yamamoto, A numerical method for solving the inverse heat conduction problem without initial value, *Inverse Probl. Sci. Eng.*, 18(2010), 655–671.

[28] P.A. Wedin, Perturbation theory for pseudo-inverses, *BIT,* 13(1973), 217–232.

[29] G.Q. Xu, State reconstruction of a distributed parameter system with exact observability, *J. Math. Anal. Appl.*, 409(2014), 168–179.

[30] M. Yamamoto and J. Zou, Simultaneous reconstruction of the initial temperature and heat radiative coefficient, *Inverse Problems,* 17(2011), 1181–1202.

[31] G.H. Zheng and T. Wei, Recovering the source and initial value simultaneously in a parabolic equation, *Inverse Problems,* 30(2014), 065013 (35pp).