Graphical Abstract

Traceability of Water Pollution: An Inversion Scheme Via Dynamic Complex Geometrical Optics Solutions

Lingyun Qiu, Zhongjing Wang, Hui Yu, Shenwen Yu

Highlights

Traceability of Water Pollution: An Inversion Scheme Via Dynamic Complex Geometrical Optics Solutions

Lingyun Qiu, Zhongjing Wang, Hui Yu, Shenwen Yu

- Research highlight 1
- Research highlight 2

Traceability of Water Pollution: An Inversion Scheme Via Dynamic Complex Geometrical Optics Solutions

Lingyun Qiu[a,b], Zhongjing Wang[c,d], Hui Yu[a,b], Shenwen Yu[a]

[a] *Yau Mathematical Sciences Center, Tsinghua University, Beijing, 100084, China*
[b] *Yanqi Lake Beijing Institute of Mathematical Sciences and Applications, Beijing, 101408, China*
[c] *Department of Hydraulic Engineering, Tsinghua University, Beijing, 100084, China*
[d] *Breeding Base for State Key Laboratory of Land Degradation and Ecological Restoration in Northwest China, Ningxia University, Yinchuan, 750021, China*

Abstract

We investigate the identification of the time-dependent source term in the diffusion equation using boundary measurements. This facilitates tracing back the origins of environmental pollutants. Employing the concept of dynamic complex geometrical optics (CGO) solutions, a variational formulation of the inverse source problem is analyzed, leading to a proof of uniqueness result. Our proposed two-step reconstruction algorithm first determines the point source locations and subsequently reconstructs the Fourier components of the emission concentration functions. Numerical experiments on simulated data are conducted. The results demonstrate that the proposed two-step reconstruction algorithm can reliably reconstruct multiple point sources and accurately reconstruct the emission concentration functions. Additionally, by partitioning the algorithm into online and offline computations, and concentrating computational demand offline, real-time pollutant traceability becomes feasible. This method, applicable in various fields - especially those related to water pollution, can identify the source of a contaminant in the environment, thus serving as a valuable tool in environmental protection.

Keywords: inverse problem, parabolic equation, complex geometrical optics solutions, boundary control, point source, water pollution

Email addresses: `lyqiu@tsinghua.edu.cn` (Lingyun Qiu), `zj.wang@tsinghua.edu.cn` (Zhongjing Wang), `huiyu@tsinghua.edu.cn` (Hui Yu), `ysw22@mails.tsinghua.edu.cn` (Shenwen Yu)

1. Introduction

In the field of environmental science, managing water pollution is a significant concern. Water contamination disrupts ecosystems, compromises biodiversity, and poses health risks through fatal waterborne diseases. A pivotal step towards addressing this issue lies in accurately identifying pollution sources, which in turn aids in implementing timely preventative or mitigative actions.

In this paper, we contribute to this critical endeavor by focusing on an inverse source problem associated with water pollution dispersion. We apply the advection-diffusion-reaction equation to model the pollution dispersion as follows.

$$\partial_t u - \kappa \Delta u + v \cdot \nabla u + \gamma u = F(x,t), \quad (x,t) \in \Omega \times [0,T], \tag{1}$$

where Ω is the spatial domain, $[0,T]$ is the observation time interval, κ is the isotropic diffusion coefficient, v is the convective velocity vector and γ is the reaction coefficient. Here, $F(x,t)$ is the pollution source term that varies with space and time, and $u(x,t)$ represents the contaminant's density.

Our primary objective, which also represents the main challenge of our research, is to determine the source term F using measured boundary data $u|_{\partial\Omega\times[0,T]}$. These measurements span across the full boundary over the observation time interval. Successfully identifying the source term F not only advances the mathematical handling of inverse source problems for parabolic equations but also offers a practical tool for regulators in their fight against water pollution.

Recently, many progresses have been made in the research field on inverse source problems for parabolic equations. The problem involving Dirichlet boundary data is explored in [1]. The source term is assumed to take the form

$$F(x,t) = f(x,t)\phi(x,t) + g(x,t),$$

where ϕ and g are two known functions and f, belonging to $L^2(\Omega \times [0,T])$, is the term to be determined. They subsequently derive an optimal control problem with a Tikhonov regularization term to recover f. Where g and ϕ are provided and $f(x,t) = f(x)$ is a spatial function, the recovery of F is discussed in several works [2, 3, 4, 5, 6], each based on varying spatial assumptions and measured data. In [2], the stability of f is constructed by using Carleman estimate. Andrey B. Kostin and his cooperators consider the existence and the uniqueness of the problem in [3, 4], where the measured data is the integral observation $\int_0^T u(t,x)h(t)dt = \varphi(x)$. When $\phi(x,t) = e^{-\gamma t}$ with $\gamma > 0$ and $\gamma \neq n^2, n \in \mathbb{N}$, Cheng and liu prove the Hölder stability of f with measurement at two instants on

a subdomain of Ω in [5]. When Ω is on a hyperplane, the uniqueness of f, where the measured data is $u|_{\{\Gamma\subset\partial\Omega\}\times[0,T]}$, is proved in [6]. In [7], given the Neumann boundary data, the authors study the recovery of the semi-discrete source term

$$F(x,t) := \sum_{k=1}^{K} p_k(x)\chi_{[t_{k-1},t_k)}(t),$$

where χ is the characteristic function. Additionally, an inverse random source problem is explored in [8]. In general, the aforementioned papers primarily focus on scenarios demonstrating spatial continuity in the source term. Other novel contributions to parabolic inverse source problems also exist. The recovery of reaction coefficient function γ is studied in [9, 10, 11, 12]. Besides, the initial value problem, that is to recover $u(x,0)$, is discussed in [13, 14, 15].

In the context of addressing water pollution, it is more reasonable to assume that the source term consists of several isolated point sources, whose spatial coordinates are denoted by s_j, and the intensity functions depend on time only, characterized by $g_j(t)$. More precisely, we assume that

$$F(x,t) = \sum_{j=1}^{m} g_j(t)\delta(x-s_j). \tag{2}$$

Our objective is to determine the number of sources m, the intensity functions $g_j(t)$ and the source locations s_j. The well-posedness of the direct problem and uniqueness of this inverse problem has been investigated in [16, 17, 18]. To recover the source term, an algebraic identification technique is proposed in [19].

Inspired by the widely-utilized adjoint state method [20], we consider the weak form of the inverse problem, using some carefully chosen test functions to satisfy the adjoint state equation. These test functions eliminate some interior terms, and the boundary term is obtained from our measurement. The selection of the test functions is crucial to the reconstructive numerical method. In [19], the authors use the harmonic polynomials $v(x,t) = (x_1 + \mathrm{i}x_2)^n, n = 0, 1, 2, ...$ in $\mathbb{R}^2$, which are static and independent of time. This will lead to some imperfections in the recovery of intensity functions. Instead, we adopt the dynamic Complex Geometrical Optics (CGO) solutions. CGO solutions have been widely applied in inverse problems, as seen in [21, 22]. To the best of our knowledge, the introduction of dynamic CGO solutions appeared in [23]. These were used to extend the elliptic Carleman estimates in [24] to a hyperbolic Carleman inequality. To

recover the intensity functions, it is advantageous to use test functions in a time-variant form, thus making the dynamic CGO solutions our preferred choice. This method shows significant advantages over existing approaches.

After using the test functions, the weak form of the equation still remains an interior integration term. Boundary control theory is applied for its computation. Theoretically, the existence of boundary control is studied maturely, as seen in [25]. However, many computational difficulties are to be resolved, which are discussed in [26]. In this paper, we derive a new approximation way to compute a large number of boundary controls, which both saves computational time and guarantees accuracy.

The paper is organized in the following way. In Section 2, the variational form of the inverse source problem using dynamic CGO solutions is derived. The uniqueness and stability results are presented in Section 3. A numerical method, consisting of a novel approximation method of boundary control and the recovery of the source term, is proposed in Section 4. Section 5 presents the numerical results of the algorithm. The conclusion and the future work are discussed in Section 6.

2. The inverse source problem and its variational formulation

2.1. The direct problem

In this paper, we consider the basic situation where κ, γ and v in (1) are constants. In fact, by introducing the transformation

$$u(x,t) = \exp\left(\frac{v}{2\kappa}\cdot x - \left(\gamma + \frac{v\cdot v}{4\kappa}\right)t\right)u_1(x,t),$$

the function $u_1(x,t)$ satisfies the heat equation

$$\partial_t u_1 - \kappa\Delta u_1 = \exp\left(-\frac{v}{2\kappa}\cdot x + \left(\gamma + \frac{v\cdot v}{4\kappa}\right)t\right)F, \quad (x,t)\in\Omega\times[0,T],$$

For such a setting, the original problem boils down to the diffusion process of a pollutant. More precisely, the dynamics of the water pollution is governed by the heat equation in $\mathbb{R}^d$, where $d\in\mathbb{N}$, subject to certain initial and boundary conditions:

$$\begin{cases} \partial_t u - \kappa\Delta u &= F & \text{in } Q, \\ u(x,0) &= 0 & \text{in } \Omega, \\ \kappa\frac{\partial u}{\partial n} &= 0 & \text{on } \Sigma, \\ u(x,t) &= \varphi(x,t) & \text{on } \Sigma, \end{cases} \tag{3}$$

where κ is the chemical diffusivity, $u(x,t)$ is the contaminant concentration, and $F(x,t)$ is the contaminant source. Here $\Omega \subset \mathbb{R}^d$ stands for a bounded C^2-domain, with n representing the outward normal vector to its boundary, $\partial\Omega$. The observation spans the time interval $(0,T)$. The combined space-time domains are represented by $Q = \Omega \times (0,T)$ and $\Sigma = \partial\Omega \times (0,T)$. Specifically, the contaminant source is characterized by:

$$F(x,t) = F(\theta) = \sum_{j=1}^{m} g_j(t)\delta(x - s_j), \tag{4}$$

Here, m denotes the number of sources. The time-dependent intensity and position of the j-th point source are represented by $g_j(t)$ and $s_j \in \mathbb{R}^d$, respectively. The comprehensive information about the contaminant source is encapsulated in the set $\theta = \{m, (s_j, g_j)_{1\leq j\leq m}\}$. Without loss of generality, we assume that

H1 The point source positions s_j's are distinct.

H2 The source intensities $g_j \in L^2(0,T)$ are nonzero and real functions. Also, they are inactive after a given time $T^* \in (0,T)$. In other words,

$$\exists\, T^* \in (0,T) \text{ such that } g_j(t) = 0 \quad \forall t \in (T^*, T) \text{ and } \forall j \in \{1, 2, \ldots, m\}.$$

In this paper, we neglect the water depth and consider the inverse source problem in $\mathbb{R}^2$. As shown in [16], if u is the solution of (3) subject to the assumptions H1 and H2, then $\varphi = u|_\Sigma \in L^2(\Sigma)$.

2.2. The inverse problem

Due to the special form of the source term (4), we propose to employ the variational formulation of the problem. The following notations involving T^* will be used:

$$\begin{aligned} Q^- &= \Omega \times (0, T^*), \quad & \Sigma^- &= \partial\Omega \times (0, T^*), \\ Q^+ &= \Omega \times (T^*, T), \quad & \Sigma^+ &= \partial\Omega \times (T^*, T). \end{aligned}$$

The integration with the test function $v(x,t)$ over Q^- leads to

$$\int_{Q^-} (\partial_t u - \kappa\Delta u)\, v dx dt = \sum_{j=1}^{m} \int_0^{T^*} g_j(t) v\left(s_j, t\right) dt. \tag{5}$$

Define the test function space:

$$\mathcal{V} = \left\{ v \in L^2((0,T); H^1(\Omega)) \middle| \partial_t v + \kappa \Delta v = 0 \text{ in } Q \right\}.$$

Integrating by parts, we have:

Lemma 1. *For any $v \in \mathcal{V}$, we have*

$$\int_\Omega u(x,T^*)v(x,T^*)\,dx + \kappa \int_{\Sigma^-} \varphi \frac{\partial v}{\partial n}\,d\sigma = \sum_{j=1}^m \int_0^{T^*} g_j(t)v(s_j,t)\,dt.$$

Define a functional $\mathcal{R}(v)$ by

$$\mathcal{R}(v) := \int_\Omega u(x,T^*)v(x,T^*)\,dx + \kappa \int_{\Sigma^-} \varphi \frac{\partial v}{\partial n}\,d\sigma. \tag{6}$$

The second term comes directly from measured boundary data φ. We will obtain the first term using the boundary controllability theorem and the details are provided in Section 2.3. Lemma 1 transforms the inverse source problem into the weak form:

$$\mathcal{R}(v) = \sum_{j=1}^m \int_0^{T^*} g_j(t)v(s_j,t)\,dt \qquad \forall v \in \mathcal{V}. \tag{7}$$

We will demonstrate that a special subset of $\mathcal{V}$ is enough to recover the source. Given certain test function space $\mathcal{H} \subset \mathcal{V}$, we define the forward operator $\mathcal{P}$ as:

$$\mathcal{P}_\mathcal{H} : \theta = \{m, (s_j, g_j)_{1\le j\le m}\} \to \big(\mathcal{R}(v)\big)_{v\in\mathcal{H}}. \tag{8}$$

Then the variational inverse problem can be formulated as follows.

The Variational Inverse Problem:

Given $\big(\mathcal{R}(v)\big)_{v\in\mathcal{H}}$, find $\theta = \{m, (s_j, g_j(t))_{1\le j\le m}\}$, i.e., the number of sources m, their fixed positions s_j, and their time-dependent intensities $\{g_j\}_{j=1}^m$, such that

$$\mathcal{P}_\mathcal{H}(\theta) = \big(\mathcal{R}(v)\big)_{v\in\mathcal{H}}. \tag{9}$$

2.3. The computability of $\mathcal{R}(v)$

Note that $u(x,T^*)$ may not be available in applications. In order to evaluate $\mathcal{R}(v)$ through the measured boundary data φ, we need to transform the domain integral $\int_\Omega u(x,T^*)v(x,T^*)\,dx$ to a boundary integral. We employ the boundary

controllability theory to do this; see more details in [25]. For the sake of completeness, we state the main results in this section. In fact, we only need to evaluate

$$\int_\Omega u(x,T^*)v(x,T^*)\,dx \quad \text{for any given test function } v(x,t)\in\mathcal{H}. \tag{10}$$

We formulate a suitable boundary controllability problem to obtain the above integration. It is stated in two lemmas below.

Lemma 2. *For each given $p\in H^1(\Omega)$, there exists a pair of functions (ψ,ω) with $\psi\in L^2\big((T^*,T);H^1(\Omega)\big)$, $\omega\in L^\infty(\Sigma^+)$ which solves the following system*

$$\begin{cases} \partial_t\psi+\kappa\Delta\psi=0 & \textit{in } Q^+, & (11a)\\ \kappa\dfrac{\partial\psi}{\partial n}=\omega & \textit{on } \Sigma^+, & (11b)\\ \psi(x,T)=0 & \textit{in } \Omega, & (11c)\\ \psi(x,T^*)=p & \textit{in } \Omega. & (11d)\end{cases}$$

Proof. According to the Theorem 6.1 of Chapter 2 in [27], the following equations has a solution ω:

$$\begin{cases} \partial_t\psi_1+\kappa\Delta\psi_1=0 & \text{in } Q^+,\\ \kappa\frac{\partial\psi_1}{\partial n}=\omega & \text{on } \Sigma^+,\\ \psi_1(x,T)=-\psi_2(x,T) & \text{in } \Omega,\\ \psi_1(x,T^*)=0 & \text{in } \Omega,\end{cases}$$

where ψ_2 is the solution to

$$\begin{cases} \partial_t\psi_2+\kappa\Delta\psi_2=0 & \text{in } Q^+,\\ \kappa\frac{\partial\psi_2}{\partial n}=0 & \text{on } \Sigma^+,\\ \psi_2(x,T^*)=p & \text{in } \Omega.\end{cases}$$

Then $\psi=\psi_1+\psi_2$ and ω are the solution to (11). □

Using the boundary control theory, the computability of $\mathcal{R}(v)$ is given in [19]. We restate the lemma in our setting and provide a brief proof.

Lemma 3. *For a test function $v\in\mathcal{V}$, take $p=v(x,T^*)$ in* (11) *and ω is the corresponding control, and we have:*

$$\mathcal{R}(v)=\kappa\int_{\Sigma^-}\varphi\frac{\partial v}{\partial n}\,d\sigma+\int_{\Sigma^+}\varphi\omega\,d\sigma. \tag{12}$$

Proof. Multiplying the equation in the system (11) with the solution u, and integrating by parts in Q^+ leads to

$$\begin{aligned}
0 &= \int_{Q^+} \big(\partial_t \psi + \kappa \Delta \psi\big) u \, dxdt \\
&= \int_{Q^+} \Big[\partial_t(\psi u) - \psi \partial u\Big] \, dxdt + \int_{\Sigma^+} \kappa \left(\frac{\partial \psi}{\partial n} u - \psi \frac{\partial u}{\partial n} \right) d\sigma + \int_{Q^+} \kappa \psi \Delta u \, dxdt \\
&= \int_{\Omega} u\psi \, dx \Big|_{T^*}^{T} + \int_{\Sigma^+} \kappa \frac{\partial \psi}{\partial n} u \, d\sigma \\
&= - \int_{\Omega} u(x, T^*) v(x, T^*) \, dx + \int_{\Sigma^+} \omega \varphi \, d\sigma.
\end{aligned}$$

Recalling the definition of $\mathcal{R}(v)$ from (6), the proof is completed. □

Remark 1. *The boundary control ω that satisfies* (11) *is not necessarily unique. However, the value of $\mathcal{R}(v)$ remains invariant to the specific choice of ω. This invariance can be attributed to the fact that, given two boundary controls ω_1 and ω_2 associated with the test function v, the relation*

$$\int_{\Sigma^+} \varphi(\omega_1 - \omega_2) \, d\sigma = 0$$

holds true.

2.4. Choice of test functions

There exist various types of test functions that satisfy the adjoint problem, including harmonic polynomial functions of the form:

$$v(x) = (x \cdot a)^n \qquad \text{for } x = (x_1, \cdots, x_d) \in \mathbb{R}^d, a = (a_1, \cdots, a_d) \in \mathbb{C}^d, a \cdot a = 0.$$

In [19], it was investigated that harmonic polynomial functions are utilized to identify the information of sources. However, their algebraic method is prone to instability due to the need for calculating matrix eigenvalues. Another drawback is that as the power n increases, the function becomes numerically unstable. This instability directly affects the accuracy of the numerical computation of $R(v)$, thereby potentially disrupting the entire algorithm. Additionally, harmonic polynomials are independent of time. Consequently, from $R(v)$, we can only gather information regarding the total intensity $\int_0^{T^*} g_j(t)dt$. To recover the intensity functions, an alternative method must be employed, such as solving the

minimization problem described in [19]. This method will involve solving heat equations for each iterative step, thus leads to large computational time.

To enhance the numerical method, we incorporate the dynamic complex geometrical solutions (CGO) solutions $v(x,t) = e^{\alpha t + \rho \cdot x}$, where $\alpha \in \mathbb{C}$ and $\rho \in \mathbb{C}^d$. First, choosing different dynamic CGO solutions lead to the stability of the numerical method as affirmed by Theorem 9. Moreover, the time component of dynamic CGO solutions aids in recovering the Fourier coefficients of intensity functions, as indicated by the subsequent Theorem 4.

Define the function space $\mathcal{H} \subset \mathcal{V}$ and the subspaces $\mathcal{H}_k$ of $\mathcal{H}$ by

$$\mathcal{H} = \big\{ v \in L^2((0,T); H^1(\Omega)) \big| v(x,t) = e^{\alpha t + \rho \cdot x} \text{ with } \alpha \in \mathbb{C}, \rho \in \mathbb{C}^d \\ \text{satisfies } \alpha + \kappa \rho \cdot \rho = 0. \big\},$$

$$\mathcal{H}_k = \big\{ v \in L^2((0,T); H^1(\Omega)) \big| v(x,t) = e^{-\frac{2k\pi \mathrm{i}}{T^*} t + \rho \cdot x} \text{ with } \rho \in \mathbb{C}^d \\ \text{satisfies } -\frac{2k\pi \mathrm{i}}{T^*} + \kappa \rho \cdot \rho = 0. \big\}.$$

In the context of our problem, it's crucial to consider the frequency-dependent behavior of the system. By introducing subspaces $\mathcal{H}_k$, we aim to capture the essence of this frequency-dependent behavior. Each subspace $\mathcal{H}_k$ focuses on a specific frequency k, allowing us to isolate and analyze the system's response at that particular frequency. This decomposition into subspaces is not only mathematically elegant but also computationally efficient. The forthcoming algorithm, which solves the subproblem for each frequency k, can then leverage the information contained within the corresponding subspace $\mathcal{H}_k$, ensuring a more targeted and precise numerical solution.

For the integer k, let $\alpha_k = -\frac{2k\pi \mathrm{i}}{T^*}$, and define the corresponding Fourier coefficients of g_j by

$$\lambda_j(\alpha_k) = \int_0^{T^*} g_j(t) e^{\alpha_k t} \, dt. \tag{13}$$

Theorem 4. *Given $\big(m, (s_j)_{1 \le j \le m}\big)$, the k-th Fourier coefficients $\lambda_j(\alpha_k)$ can be calculated from $\{\mathcal{R}(v) | v \in \mathcal{H}_k\}$.*

Since the proof involves Theorem 8, we postpone it in Section 3.2.

Noticing that the intensity functions $g_j(t) \in L^2(0, T^*)$, the Fourier transform tells us:

$$g_j(t) \approx \frac{1}{T^*} \sum_k \lambda_j(\alpha_k) e^{-\alpha_k t} \qquad \text{for } j = 1, \ldots, m. \tag{14}$$

3. The well-posedness and stability of the inverse problem

In this section, we address the uniqueness of the solution to the inverse problem based on the new data type $\{\mathcal{R}(v)|v \in \mathcal{H}\}$. Additionally, we discuss the feasibility of our proposed numerical method. The analyses of these two main results rely on some certain constructive dynamic CGO solutions. Let $\rho = a + b\mathrm{i}$ with $a, b \in \mathbb{R}^d$. We will construct several dynamic CGO solutions, $v(x,t) = e^{\alpha t + (a+b\mathrm{i})\cdot x}$, based on the real vector a. This construction is detailed in the following lemma:

Lemma 5. *Fix $k \in \mathbb{R}$ and for real vector $a \in \mathbb{R}^d$ with $|a| \geq \sqrt{|k\pi|}$, there exists a real vector $b \in \mathbb{R}^d$ such that the complex vector $\rho = a + b\mathrm{i} \in \mathbb{C}^d$ satisfies $\rho \cdot \rho = 2k\pi\mathrm{i}$.*

Proof. Suppose $a = |a|(\cos\beta a_1, \sin\beta)$, where $a_1 \in \mathbb{R}^{d-1}$ with $|a_1| = 1$. If $|a| \neq 0$, we simply take $b = |a|(\cos\gamma a_1, \sin\gamma)$, where $\gamma = \beta + \arccos\left(\frac{k\pi}{|a|^2}\right)$. If $|a| = 0$, i.e., $k = 0$, then we take $b = \mathbf{0}$.

□

3.1. Uniqueness theorem

The uniqueness theorem, which maps the measured boundary data φ to the source set θ under assumptions H1 and H2, is established in [19, Theorem 1]. However, our numerical approach transforms the boundary data φ into a different data type, $\{\mathcal{R}(v)|v \in \mathcal{H}\}$. Given that we employ a smaller test function space $\mathcal{H}$, it becomes essential to establish the uniqueness of the mapping from $\{\mathcal{R}(v)|v \in \mathcal{H}\}$ to the source term θ. The following theorem addresses this:

Theorem 6. *The operator $\mathcal{P}_{\mathcal{H}}$ defined by* (8) *is an injection.*

Proof. We will show that there exists only a trivial solution θ such that

$$\mathcal{R}(v) = 0 \qquad \forall v \in \mathcal{H} \tag{15}$$

by contradiction. For $m = 1$, it is clear that θ must be a trivial solution and hence $\mathcal{P}_{\mathcal{H}}$ is an injection. Assume that we have more than one nonzero sources. Without loss of generality, order the sources in terms of locations such that

$$|s_1| \geq |s_j| \qquad \forall 2 \leq j \leq m.$$

Since all the locations are distinct, we have:

$$s_1 \cdot s_1 > s_1 \cdot s_j \qquad \forall 2 \leq j \leq m. \tag{16}$$

For each k, we employ Lemma 5 to construct the test functions

$$v_\xi(x,t) = e^{-\alpha_k t} w_\xi(x) = e^{-\alpha_k t} e^{\xi(s_1 + b_\xi \mathrm{i})\cdot x} \qquad \forall \xi > \xi_0 := \frac{1}{|s_1|}\sqrt{\frac{|k|\pi}{T^*}}.$$

Note that $|w_\xi(x)| = |w_{\xi_0}(x)|^{\frac{\xi}{\xi_0}}$, and the inequality (16) implies that

$$|w_{\xi_0}(s_1)| > |w_{\xi_0}(s_j)| \qquad \forall 2 \le j \le m. \tag{17}$$

Obviously $v_\xi \in \mathcal{H}$. Plugging the test function v_ξ into (15) and dividing both sides by $w_\xi(s_1)$ lead to

$$\begin{aligned} 0 &= \sum_{j=1}^{m} \int_0^{T^*} g_j(t) e^{-\alpha_k t}\, dt \frac{w_\xi(s_j)}{w_\xi(s_1)} \\ &= \int_0^{T^*} g_1(t) e^{-\alpha_k t}\, dt + \sum_{j=2}^{m} \int_0^{T^*} g_j(t) e^{-\alpha_k t}\, dt \frac{w_\xi(s_j)}{w_\xi(s_1)}. \end{aligned}$$

It follows that

$$\begin{aligned} \left| \int_0^{T^*} g_1(t) e^{-\alpha_k t}\, dt \right| &= \left| -\sum_{j=2}^{m} \int_0^{T^*} g_j(t) e^{-\alpha_k t}\, dt \frac{w_\xi(s_j)}{w_\xi(s_1)} \right| \\ &\le \sum_{j=2}^{m} \left| \int_0^{T^*} g_j(t) e^{-\alpha_k t}\, dt \right| \left| \frac{w_{\xi_0}(s_j)}{w_{\xi_0}(s_1)} \right|^{\frac{\xi}{\xi_0}}. \end{aligned}$$

Taking the limit as $\lambda \to +\infty$, together with (17), we conclude that

$$\int_0^{T^*} g_1(t) e^{-\alpha_k t}\, dt = 0 \qquad \forall k \in \mathbb{R}. \quad \Longrightarrow \quad g_1(t) \equiv 0 \qquad \forall t \in [0, T^*].$$

It contradicts the assumption H2, and the proof is complete. □

Remark 2. *Lemma 3 shows that the measured data φ uniquely determines the values of functional $\mathcal{R}(v)$. Consequently, the injectivity of $\mathcal{P}_{\mathcal{H}}$ supports [19, Theorem 1], which states that the source set θ is uniquely determined by the boundary measurement φ_θ.*

3.2. The feasibility of the numerical method and stability results

In the following discussion, we explore the reconstruction of the Fourier coefficients associated with the intensity functions. Building upon the forthcoming Section 4.2, which introduces an efficient numerical method to ascertain both the number of sources and their respective locations, we assume the availability of the values $\{m, (s_j)_{j=1}^m\}$ within the set θ. Our analysis will particularly emphasize the solvability of the terms $\lambda_j(\alpha_k)$.

Fix k, and suppose there exist m distinct complex vectors $\{\rho_l^{(k)}\}_{1\le l\le m}$ associated to α_k according to Lemma 5 such that

$$v_{k,l}(x,t) = e^{\alpha_k t + \rho_l^{(k)}\cdot x} \in \mathcal{H}.$$

Then equations (7) and (13) lead to

$$\mathcal{R}(v_{k,l}) = \sum_{j=1}^{m} \lambda_j(\alpha_k) e^{\rho_l^{(k)}\cdot s_j}. \tag{18}$$

We end up with a linear system of the Fourier coefficients for the k-th mode:

$$A_k \Lambda_k = \mathcal{R}_k, \tag{19}$$

where $A_k = \left[a_{l,j}^{(k)}\right]_{1\le l,j\le m}$ with $a_{l,j}^{(k)} = e^{\rho_l^{(k)}\cdot s_j}$, and the two vectors are

$$\Lambda_k = \Big(\lambda_1(\alpha_k), \cdots, \lambda_m(\alpha_k)\Big)^T, \quad \mathcal{R}_k = \Big(\mathcal{R}(v_{k,1}), \cdots, \mathcal{R}(v_{k,m})\Big)^T.$$

The solvability of Λ_k is equivalent to the invertibility of the matrix A_k. We will show that for properly chosen test functions, i.e., complex vectors $\{\rho_l^{(k)}\}_{1\le l\le m}$, all the eigenvalues of A_k are nonzero, and hence A_k is invertible. We start with a lemma to estimate the magnitude of the eigenvalues of a matrix.

Lemma 7. *Consider a general matrix $A = (a_{l,j})_{1\le l,j\le m}$ and its eigenvalue μ. For any given positive real numbers $p_1, \ldots, p_m$, we have*

$$\mu \in \bigcup_{j=1}^{m} \mathcal{B}\left(a_{j,j}, p_j \sum_{l\ne j} \frac{|a_{l,j}|}{p_l}\right)$$

where the open ball $\mathcal{B}(z_0, r) := \{z \in \mathbb{C} : |z - z_0| < r\}$.

Proof. Introduce a matrix

$$B = D^{-1}AD = [b_{l,j}]_{1\le l,j\le m} = \left[\frac{p_j}{p_l}a_{l,j}\right]_{1\le l,j\le m}, \tag{20}$$

where $D = \mathrm{diag}(p_1, \cdots, p_m)$. Since A and B are similar, they have the same eigenvalues. The result is obtained by applying the Gershgorin circle theorem in [28] to B^T. □

We now show that the open balls of A_k exclude the origin, implying that all eigenvalues are distinct from zero.

Theorem 8. *Given the set $(m, (s_j)_{j=1}^m)$ and an integer k, there exist test functions $v_{k,1}, \cdots, v_{k,m} \in \mathcal{H}_k$, such that the matrix A_k is invertible.*

Proof. By shifting the domain, we can guarantee that the source locations s_j are distinct from the origin. We then arrange them in the order:

$$|s_1| \ge |s_2| \ge \cdots \ge |s_m| > 0.$$

Given their distinctness, we can identify two positive constants, δ and r, satisfying:

$$|s_j - s_l| \ge \delta > \sqrt{\frac{2\ln(m-1)}{r}} \qquad \forall j \ne l. \tag{21}$$

Define r_k as

$$r_k := \max\left\{1, r, \frac{1}{|s_m|}\sqrt{\frac{|k|\pi}{T^*}}\right\}.$$

Referring to Lemma 5, a complex vector ρ_l exists such that:

$$\rho_l = r_k s_l + \mathrm{i}b_l, \quad \text{and} \quad \rho_l \cdot \rho_l = \frac{2k\pi\mathrm{i}}{T^*}.$$

The associated test function is $v_{k,l}(x,t) = e^{\alpha_k t + \rho_l \cdot x}$, and the matrix $A_k = [a_{l,j}]_{1\le l,j\le m}$ with $a_{l,j} = e^{\rho_l \cdot s_j}$. Here we omit the superindex k in ρ_l and $a_{l,j}$ for concise presentation.

Selecting $p_j = e^{\frac{r_k}{2}|s_j|^2}$ and for any z in the open ball $\mathcal{B}(a_{j,j}, p_j \sum_{l\neq j} \frac{|a_{l,j}|}{p_l})$, we deduce:

$$
\begin{aligned}
|z| &\geq |a_{j,j}| - p_j \sum_{l\neq j} \frac{|a_{l,j}|}{p_l} \\
&= \exp(r_k|s_j|^2) - \exp\left(\frac{r_k}{2}|s_j|^2\right) \sum_{l\neq j} \exp\left(\frac{r_k}{2}(2s_l \cdot s_j - |s_l|^2)\right) \\
&= \exp\left(r_k|s_j|^2\right) \left(1 - \exp\left(-\frac{r_k}{2}|s_j|^2\right) \sum_{l\neq j} \exp\left(\frac{r_k}{2}(2s_l \cdot s_j - |s_l|^2)\right)\right) \\
&= \exp\left(r_k|s_j|^2\right) \left(1 - \sum_{l\neq j} \exp\left(-\frac{r_k|s_j - s_l|^2}{2}\right)\right) \\
&\geq \exp\left(r_k|s_j|^2\right) \left(1 - (m-1)\exp\left(-\frac{r_k\delta^2}{2}\right)\right) \\
&\geq 1 - (m-1)e^{-\frac{r\delta^2}{2}}.
\end{aligned}
$$

To summarize, we have

$$
|z| \geq C(m;\delta,r) := 1 - (m-1)e^{-\frac{r\delta^2}{2}} > 0 \qquad \forall \gamma \in \mathcal{B}_j \text{ for } 1 \leq j \leq m. \quad (22)
$$

Given that any eigenvalue μ of A_k resides in one of the aforementioned open balls (as per Lemma 7), μ is distinct from zero. Consequently, A_k is invertible. □

Now we are ready to provide a proof of Theorem 4 as follows.

Proof of Theorem 4. For each integer k, choose the test functions $v_{k,l}$, for $l = 1, ..., m$ as described in Theorem 8. Then the resulting linear system $A_k\Lambda_k = \mathcal{R}_k$ possesses a unique solution Λ_k since A_k is invertible guaranteed by Theorem 8. □

For $k = 0$, the estimate on the eigenvalues in Theorem 8 leads to a stability result concerning the total intensities of the sources.

Theorem 9. *Given the set $(m, (s_j)_{j=1}^m)$, there exist test functions $v_{0,l}(x)$ for $1 \leq l \leq m$ such that the resulting linear system $A_0\Lambda_0 = \mathcal{R}_0$ is stable in the sense that*

$$
||\Lambda_0 - \widehat{\Lambda_0}||_2 \leq \frac{1}{C(m;\delta,r)} ||\mathcal{R}_0 - \widehat{\mathcal{R}_0}||_2, \quad (23)
$$

where the pairs $(\Lambda_0, \mathcal{R}_0)$ and $(\widehat{\Lambda_0}, \widehat{\mathcal{R}_0})$ are associated to two sets of source intensities $(g_j)_{j=1}^m$ and $(\widehat{g_j})_{j=1}^m$, respectively.

Proof. Construct the test functions $v_{0,l}$ analogously to Theorem 8, i.e., $\rho_l = r_0 s_l + \mathrm{i} b_l$. We aim to show that A_0 is a Hermitian matrix, i.e.

$$\rho_l \cdot s_j = \overline{\rho_j} \cdot s_l \quad \text{for } 1 \le l, j \le m.$$

Thanks to Lemma 5, we have for $1 \le l \le m$,

$$\rho_l \cdot \rho_l = 0. \quad \Longleftrightarrow \quad |b_l| = r_0 |s_l|, b_l \cdot s_l = 0.$$

In fact, b_l is a (clockwise) rotation of $r_0 s_l$ with an angle of $\pi/2$. It follows that

$$(b_l + b_j) \cdot (s_l + s_j) = 0 \qquad \text{for } 1 \le j, l \le m.$$

Thus, $b_l \cdot s_j + b_j \cdot s_l = 0$ for $1 \le j, l \le m$. Further, we can express

$$\begin{aligned} \rho_l \cdot s_j - \overline{\rho_j} \cdot s_l =& (r_0 s_l + \mathrm{i} b_l) \cdot s_j - (r_0 s_j - \mathrm{i} b_j) \cdot s_l \\ =& r_0 s_l \cdot s_j - r_0 s_j \cdot s_l + \mathrm{i}(b_l \cdot s_j + b_j \cdot s_l) = 0. \end{aligned}$$

Consequently, A_0 is Hermitian, and so is its inverse A_0^{-1}. Letting μ represent the eigenvalue of A_0, by (22), we have

$$\|A_0^{-1}\|_2 = \max_{\mu} \left| \frac{1}{\mu} \right| \le \frac{1}{C(m; \delta, r)}.$$

This directly yields

$$\|\Lambda_0 - \widehat{\Lambda_0}\|_2 = \|A_0^{-1}(\mathcal{R}_0 - \widehat{\mathcal{R}_0})\|_2 \le \frac{1}{C(m; \delta, r)} \|(\mathcal{R}_0 - \widehat{\mathcal{R}_0})\|_2.$$

□

Remark 3. *The assertions of Theorem 9 still hold if test functions are chosen such that the associated matrix A_0 is normal.*

4. Numerical algorithm

This section introduces the proposed reconstruction algorithm. Assumption H2 confirms that all intensity functions $g_j(t)$ are real, leading to the relation

$$\Lambda_k = \bar{\Lambda}_{-k}.$$

Consequently, it suffices to recover Λ_k for $k \geq 0$. Using the linear system for the Fourier coefficients of the k-th mode as described in (19), we formulate the inversion of the number of sources m, the Fourier coefficients of intensity functions Λ_k, and the source locations $S = (s_1, \ldots, s_m)^T$ as the following optimization problem:

$$\min_{m,\, S,\, \Lambda_k} \left(\sum_{0 \leq k \leq K} \|A_k(m, S)\Lambda_k - \mathcal{R}_k\|_2^2 \right). \tag{24}$$

To enhance efficiency and accuracy, we split this problem into two sub-problems. Drawing inspiration from the invertibility established in Theorem 8 for $k = 0$ and the stability result in Theorem 9, we first determine the number of sources, their locations, and total intensities by addressing:

$$(m, S, \Lambda_0) = \underset{m \leq M, S, \Lambda_0}{\arg\min} \|A_0(m, S)\Lambda_0 - \mathcal{R}_0\|_2^2. \tag{25}$$

Assuming the number of sources does not exceed M, we employ $3M$ test functions $v_{0,l}, 1 \leq l \leq 3M$, which are defined as:

$$v_{0,l}(x, t) = \exp(\rho_{0,l} \cdot x).$$

$$\rho_{0,l} = \frac{2\sqrt{2}}{\text{diag}(\Omega)} \left((\cos\beta_l, \sin\beta_l) + \mathrm{i}(\cos\gamma_l, \sin\gamma_l) \right), \beta_l = \frac{2j\pi}{M}, \gamma_l = \beta_l - \frac{\pi}{2},$$

Subsequently, we isolate the reconstruction process for the Fourier coefficients based on the value of k, reconstructing the coefficients for each frequency by resolving:

$$A_k(m, S)\Lambda_k = \mathcal{R}_k, \quad k = 1, 2, \ldots, K. \tag{26}$$

For $k > 0$, we only need no more than M test functions and we will give the definition of them in Section (5.3).

We can now outline the algorithm as follows:

(i) Given two integers K and M, construct the CGO-type test functions $v_{0,l}, 1 \leq l \leq 3M$ and $v_{k,l}, 1 \leq k \leq K, 1 \leq l \leq M$. Compute the approximation to the boundary control ω using the bases $\{\omega_\eta\}$ given in Sec. 4.1.

(ii) Compute the functional $\mathcal{R}_0 = \big(\mathcal{R}(v_{0,l})\big)_{1 \leq l \leq 3M}$ and $\mathcal{R}_k = \big(\mathcal{R}(v_{k,l})\big)_{1 \leq l \leq M}$ in Equation (12). Solve the minimization problem (25) to obtain the set (m, S, Λ_0).

(iii) With known (m, S), solve the linear systems (26) to obtain $\Lambda_k, 1 \leq k \leq K$, and then perform the inverse Fourier transform to retrieve $g_j, 1 \leq j \leq m$.

Next we will address the implementation of each step.

4.1. Efficient computation of boundary control ω and $\mathcal{R}(v)$

For our numerical method, a substantial number of calculations for $\mathcal{R}(v)$ are required. Solving the minimization problem of boundary control (11), as proposed in [25], can be computationally intensive as it involves solving the initial-boundary value problem of the heat equations repeatedly. To mitigate this computational demand, especially when computing a large number of boundary controls, we introduce a novel method that enhances efficiency in these scenarios. This method centers around the construction of a set of basis functions to approximate the space $L^2(\Sigma^+)$. Each basis function within this set is characterized by a triplet of indices, $\eta = (\eta_1, \eta_2, \eta_3)$ with $\eta_j = 1, 2, ..., W_j, j = 1, 2, 3$, where W_j's are given positive integers.

Specifically, for a spatial domain $\Omega = [0, L_x] \times [0, L_y]$, we define $\omega_\eta \in L^2(\Sigma^+)$ as

$$\omega_\eta(x, y, t) = \cos\left(\frac{2\eta_3 - 1}{2}\pi\frac{t - T^*}{T - T^*}\right)\begin{cases} -\kappa_2 \sin\left(\frac{\eta_1\pi}{L_x}x\right)\frac{\eta_2\pi}{L_y} & \text{for } y = 0, \\ \kappa_2 \sin\left(\frac{\eta_1\pi}{L_x}x\right)\frac{\eta_2\pi}{L_y}(-1)^{\eta_2} & \text{for } y = L_y, \\ -\kappa_1 \sin\left(\frac{\eta_2\pi}{L_y}y\right)\frac{\eta_1\pi}{L_x} & \text{for } x = 0, \\ \kappa_1 \sin\left(\frac{\eta_2\pi}{L_y}y\right)\frac{\eta_1\pi}{L_x}(-1)^{\eta_1} & \text{for } x = L_x. \end{cases}$$

For each basis function ω_η, there is an associated function ψ leading to $\psi_\eta(T^*)$, which can be determined using a standard PDE (Partial Differential Equation) solver for (11a-c). With this set of functions $\{\psi_\eta(T^*)\}$, we can solve the minimization function with respect to $v(T^*)$ as

$$\{a_\eta\} = \underset{\{\hat{a}_\eta\}}{\arg\min}\left\|\sum_\eta \hat{a}_\eta\psi_\eta(T^*) - v(T^*)\right\|^2_{L^2(\Omega)}. \tag{27}$$

The approximation of boundary control is then given by $\omega \approx \sum_\eta a_\eta\omega_\eta$.

Our methodology offers a significant reduction in computational time by requiring the PDE solver only a limited number of times to determine $\psi_\eta(T^*)$. The function approximation problem (27) is a typical least-squares problem and several methods such as QR decomposition are derived for it. We specifically employ the LSQR method for an efficient solution.

Crucially, the computation of the bases $\psi_\eta(T^*)$ is independent of both the boundary measurement φ and the test function v. This ensures that their precomputation, albeit initially time-intensive, becomes a one-time offline process.

As a result, not only do subsequent changes in the test functions lead to minor online computational adjustments, but modifications in the measured data also demand minimal online recalculations. This highlights a dual advantage: for a given spatial domain, after performing the offline computation once, we can conduct multiple measurements and run the inversion—which requires the boundary controls—rapidly. This level of efficiency is a marked improvement over the traditional methods described in [25], highlighting the advantages and versatility of our proposed approach.

After introducing our new methodology for efficient boundary control computation, we can outline the specific steps as follows:

(i-1) Given each boundary control basis ω_η, obtain $\psi_\eta(T^*)$ from the system (11a-c).

(i-2) Given the test functions $v_{0,l}, 1 \leq l \leq 3M$, solve (27) by the LSQR method.

(i-3) Obtain $\sum_\eta a_\eta \omega_\eta$ as the approximation to the boundary control ω.

For $1 \leq k \leq K$, the computation of the boundary control of $v_{k,l}$ can be performed analogously. Also, it is worth noting that step (i) is performed offline.

4.2. Identification of source number m and source location S

In this section, we discuss the process of recovering source number m, source location S and the zeroth mode Fourier coefficients Λ_0. We begin with substituting the boundary control obtained in step (i) into (12) to obtain $\mathcal{R}_0$. Our next objective is to solve (25). This task presents two major challenges. Firstly, the source number m remains unknown, demanding a strategy for its identification. Secondly, the gradient-based algorithm to solve (25) is prone to convergence towards undesirable local minima.

To mitigate these difficulties, we construct the following numerical algorithm.

(ii-1) Use the boundary control from step (i) to compute $\mathcal{R}_0$.

(ii-2) Initialize the source number as M in (25). Subsequently use the Sequential Least SQuares Programming(SLSQP) method to solve (25).

(ii-3) Use the random sampling method to bypass the local minimum. In practice, we examine several random initial guesses for the source location and the zeroth intensity. The solution with the minimum relative error from these trials is accepted as the genuine solution.

(ii-4) Assess the results. If two points are near each other or certain points exhibit significantly low intensities, some redundancy or non-existence of the points is implied. If N denotes the number of these extraneous points, then M is adjusted as $M = M - N$. We then revert to the previous step (ii-2). Otherwise, the algorithm finishes in the next step (ii-5).

(ii-5) Then the source number is defined as $m = M$, and the source position is determined as S.

In step (ii-2), fixing the source count aids in translating the minimization problem to one involving continuous variables, paving the way for gradient-based optimization techniques like SLSQP. The relatively modest complexity of equation (25) ensures a fast minimization, with the repeated minimizations in step (ii-3) causing only a negligible increase in computational time. Step (ii-4) is important, offering a systematic strategy for identifying the source number, seamlessly aligning with practical scenarios.

4.3. Computation of Fourier coefficients for $k > 0$

In Section 4.2, we obtain the source number m, the source location S and the zeroth mode coefficients Λ_0. Based on the inverse Fourier transform, we are ready to state the first intensity recovery method, referring to it by Inverse Fourier Method.

From a theoretical standpoint, for any $k > 0$, there exist m test functions $v_l(x,t) \in \mathcal{H}_k$ such that the matrix $A_k = (v_l(s_j, 0))_{1\le l,j\le m}$ is invertible. This is proved in Theorem 8. With (19), we can obtain the k-th mode Fourier coefficients.

In our numerical method, the intensity function's recovery hinges on the inverse Fourier transform, which might induce the Gibbs phenomenon near a jump discontinuity at T^*. To mitigate this, we extend g_j to be periodic over $[-T^*, T^*]$ and employ the test functions:

$$v_k(x,t) = \exp(\rho_k \cdot x) \exp\Big(-\frac{k\pi \mathrm{i}}{T^*} t \Big). \tag{28}$$

For $k \ge 1$, at most M test functions $v_{k,l}, 1 \le l \le M$ are required. The spatial vectors $\rho_{k,l}$ are defined as:

$$\rho_{k,l} = r_k \Big((\cos\beta_l, \sin\beta_l) + \mathrm{i}(\cos\gamma_l, \sin\gamma_l) \Big), \beta_l = \frac{2l\pi}{M}, \gamma_l = \beta_l - \arccos\frac{k\pi}{2r_k^2 T^*},$$

where r_k is a parameter depends on the size of space domain Ω and the index of the Fourier mode k. If r_k is excessively large, the minimization problem and

the computation of $\mathcal{R}(v)$ become unstable. Conversely, if r_k is too small, the definition of $\rho_{k,l}$ is not valid. Typically, we select:

$$r_k = \max\left\{\frac{2\sqrt{2}}{\text{diag}(\Omega)}, \sqrt{\frac{k\pi}{2T^*}}\right\}.$$

Analogous to (14), the inversion of intensity functions is based on:

$$g_j(t) \approx \frac{1}{2T^*}\sum_k \lambda_j\left(\frac{k\pi}{T^*}\right)e^{-\frac{k\pi \mathrm{i}}{T^*}t} \qquad \text{for } j = 1, \ldots, m.$$

Remark 4. *The test functions used in Section 4.2 and 4.3 are independent of the exact source locations S and the source number m. This independence allows for the advance preparation of boundary controls for these test functions.*

4.4. Approximation of intensity functions

To improve the accuracy of the recovery results, we introduce an alternative method for the reconstruction of the intensity functions, referring to it by Approximation Method. Similar to Section 4.3, the source number m and source location S are obtained in Step (ii). With this information, we aim to choose several bases of intensity functions, denoted as h_l for $0 \le l \le 2L$ with L being a positive integer. Suppose $f_{l,j}(x,t) = \delta(x - s_j)h_l(t)$ and we can define the function $\varphi_{l,j}$ that satisfies the direct problem

$$\begin{cases} \partial_t u - \kappa\Delta u &= f_{l,j} \quad \text{in } Q, \\ u(x,0) &= 0 \quad \text{in } \Omega, \\ \kappa\frac{\partial u}{\partial n} &= 0 \quad \text{on } \Sigma, \\ u(x,t) &= \varphi_{l,j} \quad \text{on } \Sigma. \end{cases} \tag{29}$$

From the above system we obtain the bases $\varphi_{l,j}$ for the boundary measurement φ. Consequently, the following minimization problem arises:

$$\{b_{l,j}\} = \arg\min_{\{\hat{b}_{l,j}\}} \left\| \sum_{0\le l\le 2L, 1\le j\le m} \hat{b}_{l,j}\varphi_{l,j} - \varphi \right\|^2_{L^2(\Sigma)}, \tag{30}$$

The recovery of intensity functions is then expressed as

$$g_j(t) = \sum_{l=0}^{2L} b_{l,j}h_l(t). \tag{31}$$

In real scenarios, the source intensity functions are non-negative functions. Therefore, post-reconstruction necessitates the truncation of the negative segment. Given that $\hat{g}(t)$ represents the intensity function procured via the aforementioned recovery technique in (31), the definitive result $g(t)$ is formulated as

$$g(t) = \begin{cases} \hat{g}(t) & \text{if } \hat{g}(t) \geq 0, \\ 0 & \text{if } \hat{g}(t) < 0, \end{cases} \qquad \text{for} \quad 0 \leq t \leq T^*.$$

In practice, the basis functions in (31) are chosen as

$$h_{2l} = \cos\left(\frac{l\pi}{T^*}t\right), \quad h_{2l+1} = \sin\left(\frac{l\pi}{T^*}t\right) \qquad \text{for } l = 0, \ldots, L.$$

To summarize, the algorithm presented in this subsection follows these steps:

(iv-1) Set L and determine the intensity function basis h_l for $0 \leq l \leq 2L$.

(iv-2) With m and S obtained in Step (ii), solve the heat equations (29) for $0 \leq l \leq 2L$.

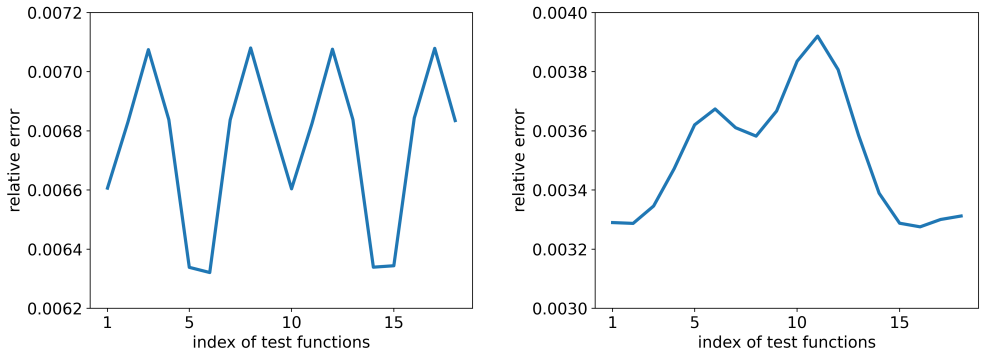
(iv-3) Use LSQR to solve the problem (30).

(iv-4) Reconstruct the nonnegative intensity functions by (31) and the post-reconstruction.

Remark 5. *The Approximation Method needs extra online computational time compared with method in Section 4.3. This is due to the necessity of solving the heat equations* (29) $(2L+1)m$ *times.*

5. Numerical simulations

For all numerical tests, we set $\Omega = [0, 1000] \times [0, 1000]$ m^2 with a constant diffusion coefficient $\kappa = 1$ m^2/s. The inactive moment T^* is set to 38.4 hours (equivalent to 138240 seconds) while the total observation duration T is 48 hours (equivalent to 172800 seconds). The intensity functions utilized in the numerical tests, when expressed in seconds, are as follows:



(a) Relative error of $v(T^*)$ (b) Relative error of $\mathcal{R}(v)$

Figure 1: Approximation Results of Boundary Control

$$q(t) = \frac{1}{2}\left[1 - \tanh\left(\frac{t - 0.9T^*}{21600}\right)\right];$$

$$g_1(t) = \left[3 + \frac{3}{2}\sin\left(\frac{2\pi t}{T^*}\right) + 3e^{-\left(\frac{10}{T^*}(t-36000)\right)^2} + \frac{5}{2}e^{-\left(\frac{15}{T^*}(t-100500)\right)^2}\right] q(t);$$

$$g_2(t) = \left[5 + \sin\left(\frac{4\pi t}{T^*}\right)\right] q(t); \qquad g_3(t) = 6\left[1 - \left(\frac{t}{T^*}\right)^2\right] q(t);$$

$$g_4(t) = 5q(t); \qquad g_5(t) = g_6(t) = 4q(t).$$

5.1. Boundary control reconstruction and calculation of $\mathcal{R}(v)$

The approximation method in Section 4.1 is used to calculate boundary control ω. We choose the basis function index $W_1 = W_2 = W_3 = 10$. Additionally, the maximal LSQR iteration step in Step (i-2) is defined to be 1000. The numerical results of the approximation of $v(T^*)$ and corresponding $\mathcal{R}(v)$ are shown in Figure 1. The relative error of $v(T^*)$ is given by $\frac{||v(T^*)-\tilde{v}||_{L^2(\Omega)}}{||v(T^*)||_{L^2(\Omega)}}$ and the relative error of $\mathcal{R}(v)$ is given by $\frac{|\mathcal{R}(v)-\tilde{\mathcal{R}}(v)|}{|\mathcal{R}(v)|}$, where $\tilde{v}$ and $\tilde{\mathcal{R}}(v)$ are the numerical approximations. The approximations of $v(T^*)$ and $\mathcal{R}(v)$ show a good agreement, which are both $\mathcal{O}(10^{-3})$. This demonstrates the effectiveness of our new method to compute the boundary control.

5.2. *The Identification of the Source Locations*

We perform numerical simulations for Step (ii) to demonstrate the accuracy of the approximation on m and S. For Step (ii-1), the initial guess of source number M is set to 7 and we choose 50 random initial guesses in Step (ii-3). The recovery results of source locations are presented in Figure 2. The position error is referred to as $\frac{||S-S_{\text{Iter}}||_2}{||S||_2}$, and the intensity error is referred to as $\frac{||\Lambda-\Lambda_{\text{Iter}}||_2}{||\Lambda||_2}$. The numerical results show that the algorithm correctly identifies the source number m. The recoveries of the source location S and total intensities are also highly precise. For the two-point case in Figure 2a, the reconstruction method handles the situation where the distance between two sources is 150 m. Generally, as the source number m increases, the source number is accurately determined. But the reconstruction becomes more unstable, particularly for the total intensities of the sources. Additionally, we observe that the sources near the boundary are better recovered than those in the middle of the whole domain.

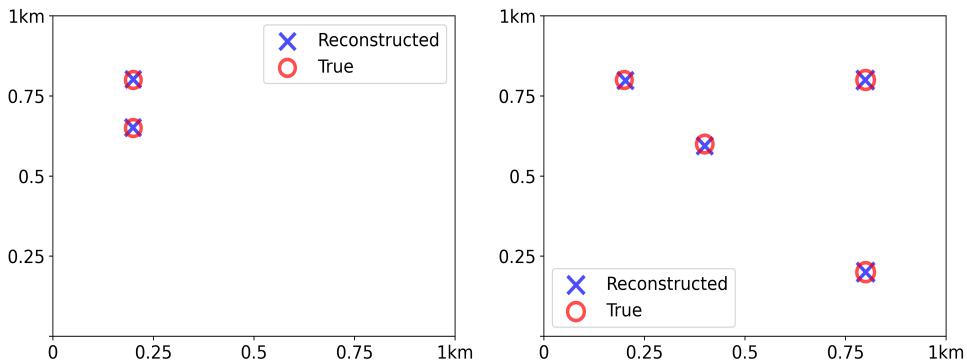
Table 1: Reconstruction of positions

		Source 1		Source 2		Source 3		Source 4		Source 5		Source 6	
2 points case	True position	200	800	200	650								
	Reconstructed position	200	802	199	651								
4 points case	True position	200	800	800	200	400	600	800	800				
	Reconstructed position	203	798	800	200	400	594	799	800				
5 points case	True position	200	800	800	200	400	600	800	800	500	100		
	Reconstructed position	202	799	801	200	394	597	798	799	501	102		
6 points case	True position	200	800	800	200	400	600	800	800	500	100	900	500
	Reconstructed position	176	819	790	241	350	663	807	798	501	109	1000	531

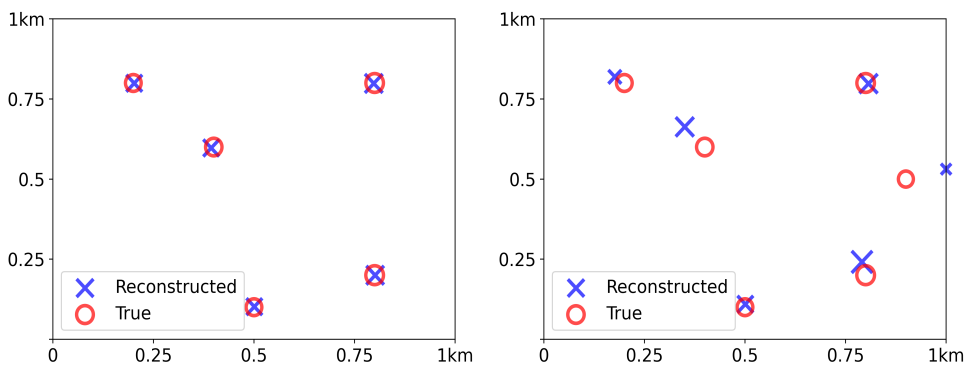
5.3. *Recovery of intensity functions' Fourier coefficients*

Leveraging the source locations of 5-points case, discussed in Section 5.2, where the intensity functions are given by $g_1(t), \cdots, g_5(t)$, we continue to present our algorithm. The recovery method, grounded in Fourier coefficients, specifies that $|k| \leq K = 8$ during Step (iii), followed by truncating any negative components. The results from this inversion process can be seen in Figure 3 under the title 'I-F Method' and the L^2 relative errors of the recovered intensity functions are tabulated in Table 2.

The outcomes demonstrate a satisfactory recovery of the intensity functions. The two peaks of both g_1 and g_2 are accurately determined. Apart from $g_3(t)$, the relative L^2 errors of the other four intensity functions remain less than 10%. Such observations underscore the notion that the accuracy in intensity function recovery



(a) Position error: 0.2%, intensity error: 1.7%. (b) Position error 0.3%, intensity error 2.7%.



(c) Position error 0.4%, intensity error 3.0%. (d) Position error 6.7%, intensity error 29.0%.

Figure 2: Reconstruction of positions

is closely linked to the precision of source location recovery. Specifically, given that the accuracy of the location s_3 is slightly poorer than the others, the resulting intensity function exhibits a notable disagreement. Additionally, it is noteworthy from Figure 3 that the recovery performance at the endpoint surpasses that at the starting point. This discrepancy can be associated with the extension phase of the intensity functions during Step (iii).

5.4. Refinement of intensity function recovery

Building upon the 5-point scenario described in Section 5.3, we employ an enhanced approximation method for refining the intensity function. In this method, denoted as Step (iv-1), the number of intensity basis functions is set to $L = 8$.

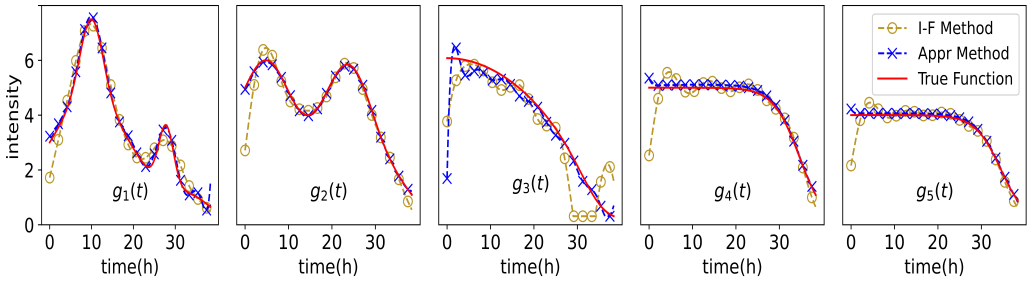


Figure 3: Reconstruction of intensity functions

Table 2: Reconstruction error of intensity functions

	$g_1(t)$	$g_2(t)$	$g_3(t)$	$g_4(t)$	$g_5(t)$
I-F Method	8.97%	8.36%	20.94%	8.97%	8.02%
Appr Method	3.30%	0.79%	10.28%	2.13%	1.66%

The results are presented in Figure 3, under the label "Appr Method", and the L^2 relative errors of the recovered intensity functions are tabulated in Table 2. Compared to the recovery results from Step (iii), the approximation method in Step (iv) show significantly better results, especially in the recovery of the function g_3. A noteworthy observation is the considerable reduction of the Gibbs phenomenon in this refined method, which leads to a substantial decrease in relative errors during the recovery process.

6. Conclusion

In this paper, we address a water pollution traceability problem based on the convection-diffusion-reaction equation. By employing dynamic CGO solutions, we transform the inverse problem into its weak form. We subsequently establish the uniqueness of the inverse problem and showcase the stability of the zeroth Fourier coefficient through an estimation of the coefficient matrix's norm.

To recover the source term, our proposed numerical method consists of three parts. It begins with the computation of boundary controls tied to the employed test functions. Following this, we identify both the count and positions of the pollutant sources. To round off the process, we introduce two methods for the recovery of intensity functions: one that deduces them through their Fourier coefficients and another that sets up a minimization framework for their reconstruction. The first one, referred to as the inverse Fourier approach, drastically shortens

online computational time, but offers a mediocre recovery of g_j's. In contrast, the approximation technique, while being more computationally intensive online, shines in its ability to reduce recovery errors in a great deal. These two methods can be flexibly used in practice based on specific scenarios. Compared to previous approaches, our numerical method is capable of effectively recovering cases with a greater number of sources, and the results for recovering the source positions and intensity functions are relatively successful.

Nevertheless, our method presents certain challenges. Persistent questions remain regarding the stability of all Fourier coefficients and the accuracy of the method in determining a larger number of sources. In the future, we aim to investigate the analytical questions that are crucial to inverse problems and employ our approach for more complex cases in the meantime. Further exploration of this approach and its applications pertaining to time-fractional equations is intriguing as well.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 11901339, 11971258, U2243601), State Key R&D Program (2021YFD1900600), and Tsinghua-Ningxia Yinchuan Joint Institute of Internet of Water & Digital Governance Key Project (SKL-IOW-2021TC2101).

References

[1] P. X. Thanh, Space-time finite element method for determination of a source in parabolic equations from boundary observations, Journal of Inverse and Ill-posed Problems 29 (5) (2021) 689–705. doi:10.1515/jiip-2019-0104.

[2] X. Huang, O. Y. Imanuvilov, M. Yamamoto, Stability for inverse source problems by Carleman estimates, Inverse Problems 36 (12) (2020) 125006. doi:10.1088/1361-6420/aba892.

[3] A. I. Prilepko, V. L. Kamynin, A. B. Kostin, Inverse source problem for parabolic equation with the condition of integral observation in time, Journal of Inverse and Ill-posed Problems 26 (4) (2018) 523–539. doi:10.1515/jiip-2017-0049.

[4] M. S. Hussein, D. Lesnic, V. L. Kamynin, A. B. Kostin, Direct and inverse source problems for degenerate parabolic equations, Journal of Inverse and Ill-posed Problems 28 (3) (2020) 425–448. doi:10.1515/jiip-2019-0046.

[5] J. Cheng, J. Liu, An inverse source problem for parabolic equations with local measurements, Applied Mathematics Letters 103 (2020) 106213. doi:10.1016/j.aml.2020.106213.

[6] J. Cheng, M. Yamamoto, Continuation of solutions to elliptic and parabolic equations on hyperplanes and application to inverse source problems, Inverse Problems 38 (8) (2022) 085005. doi:10.1088/1361-6420/ac7410.

[7] G. Lin, Z. Zhang, Z. Zhang, Theoretical and numerical studies of inverse source problem for the linear parabolic equation with sparse boundary measurements, Inverse Problems 38 (12) (2022) 125007. doi:10.1088/1361-6420/ac99f9.

[8] B. Wu, Q. Chen, Z. Wang, Carleman estimates for a stochastic degenerate parabolic equation and applications to null controllability and an inverse random source problem, Inverse Problems 36 (7) (2020) 075014. doi:10.1088/1361-6420/ab89c3.

[9] M. V. Klibanov, J. Li, W. Zhang, Convexification for an inverse parabolic problem, Inverse Problems 36 (8) (2020) 085008. doi:10.1088/1361-6420/ab9893.

[10] M. Bellassoued, O. B. Fraj, Stably determining time-dependent convection–diffusion coefficients from a partial Dirichlet-to-Neumann map, Inverse Problems 37 (4) (2021) 045011. doi:10.1088/1361-6420/abe10d.

[11] D.-H. Chen, D. Jiang, J. Zou, Convergence rates of Tikhonov regularizations for elliptic and parabolic inverse radiativity problems, Inverse Problems 36 (7) (2020) 075001. doi:10.1088/1361-6420/ab8449.

[12] J. Fan, Z. Duan, Determining a potential of the parabolic equation from partial boundary measurements*, Inverse Problems 37 (9) (2021) 095001. doi:10.1088/1361-6420/ac156d.

[13] L. Yang, Z.-C. Deng, Optimization method for a multi-parameters identification problem in degenerate parabolic equations, Journal of Inverse and Ill-posed Problems (Jan. 2023). doi:10.1515/jiip-2022-0038.

[14] Y.-H. Lin, H. Liu, X. Liu, S. Zhang, Simultaneous recoveries for semilinear parabolic systems, Inverse Problems 38 (11) (2022) 115006. doi:10.1088/1361-6420/ac91ee.

[15] E. Casas, K. Kunisch, Using sparse control methods to identify sources in linear diffusion-convection equations, Inverse Problems 35 (11) (2019) 114002. doi:10.1088/1361-6420/ab331c.

[16] M. Andrle, A. E. Badia, Identification of multiple moving pollution sources in surface waters or atmospheric media with boundary observations, Inverse Problems 28 (7) (2012) 075009. doi:10.1088/0266-5611/28/7/075009.

[17] A. E. Badia, 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 (6) (2002) 585–599. doi:doi:10.1515/jiip.2002.10.6.585.

[18] A. El-Badia, M. Farah, A stable recovering of dipole sources from partial boundary measurements, Inverse Problems 26 (11) (2010) 115006. doi:10.1088/0266-5611/26/11/115006.

[19] M. Andrle, A. E. Badia, On an inverse source problem for the heat equation. Application to a pollution detection problem, II, Inverse Problems in Science and Engineering 23 (3) (2015) 389–412. arXiv:https://doi.org/10.1080/17415977.2014.906415, doi:10.1080/17415977.2014.906415.

[20] R.-E. Plessix, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, Geophysical Journal International 167 (2) (2006) 495–503. doi:10.1111/j.1365-246X.2006.02978.x.

[21] G. Uhlmann, Inverse problems: Seeing the unseen, Bulletin of Mathematical Sciences 4 (2) (2014) 209–279. doi:10.1007/s13373-014-0051-9.

[22] I. Kocyigit, R.-Y. Lai, L. Qiu, Y. Yang, T. Zhou, Applications of CGO Solutions on Coupled-Physics Inverse Problems (Aug. 2016). arXiv:1512.06695, doi:10.48550/arXiv.1512.06695.

[23] M. Choulli, Y. Kian, Logarithmic stability in determining the time-dependent zero order coefficient in a parabolic equation from a partial Dirichlet-to-Neumann map. Application to the determination of a nonlinear term, Journal de Mathématiques Pures et Appliquées 114 (2018) 235–261. doi:10.1016/j.matpur.2017.12.003.

[24] A. L. Bukhgeim, G. Uhlmann, RECOVERING A POTENTIAL FROM PARTIAL CAUCHY DATA, Communications in Partial Differential Equations 27 (3-4) (2002) 653–668. doi:10.1081/PDE-120002868.

[25] C. Carthel, R. Glowinski, J. L. Lions, On exact and approximate boundary controllabilities for the heat equation: A numerical approach, Journal of Optimization Theory and Applications 82 (3) (1994) 429–484. doi:10.1007/BF02192213.

[26] A. Münch, E. Zuazua, Numerical approximation of null controls for the heat equation: Ill-posedness and remedies, Inverse Problems 26 (8) (2010) 085018. doi:10.1088/0266-5611/26/8/085018.

[27] A. Fursikov, Optimal Control of Distributed Systems. Theory and Applications, Vol. 187, American Mathematical Society, 2000.

[28] R. A. Brualdi, S. Mellendorf, Regions in the Complex Plane Containing the Eigenvalues of a Matrix, The American Mathematical Monthly 101 (10) (1994) 975–985. arXiv:2975164, doi:10.2307/2975164.